Principles and Recommendations for Population and Housing Censuses

Revision 4

United Nations

This publication is being provided in its current form without formal editing or copy preparation.

Preface

Since its earliest years, the United Nations has provided a series of international principles and recommendations on population and housing censuses. These guidelines are designed to assist national statistical offices and census officials worldwide in planning and carrying out effective and cost-efficient censuses. The first set of these principles and recommendations was issued in 1958 at the request of the United Nations Statistical Commission. This initiative addressed the need for developing international standards and formed a cornerstone of the first World Population and Housing Census Programme.

While the scope of these recommendations has evolved over time in response to current practices and national needs, they consistently offer guidance on the main characteristics of population and housing censuses, general aspects of census operations and methodologies, and detailed recommendations on census content.

Over the years, the United Nations Statistics Division has played a crucial role in coordinating the World Population and Housing Census Programme. This has involved issuing and revising international recommendations, providing technical assistance to countries in conducting censuses, and compiling and disseminating census results from countries or areas. The previous global census recommendations, titled *Principles and Recommendations for Population and Housing Censuses, Revision 3*[1]*,* were published in 2017.

Recognizing this publication as a primary methodological reference for countries planning and conducting their population and housing censuses, the Statistical Commission, at its fifty-fifth session,[2] welcomed the ongoing collaboration with an international expert group on the fourth revision of these principles and recommendations, in preparation for the 2030 World Population and Housing Census Programme.

The current revision of the principles and recommendations was overseen by an expert group comprising census experts representing all regions of the world. Their contributions were organized through the following task teams:

(i) Task Team 1 - Census planning, organizing and management
(ii) Task Team 2 - Census methodologies, use of administrative registers, and population concepts and definitions
(iii) Task Team 3 - Geospatial information in census operations
(iv) Task Team 4 - Quality assurance, indicators, and assessment
(v) Task Team 5 - Data collection and processing
(vi) Task Team 6 - Communication, dissemination, and utilization
(vii) Task Team 7 - Recommended topics on population and housing characteristics

[1] United Nations publication, Sales No. E.15.XVII.10.
[2] United Nations Statistical Commission. *Report on the fifty-fifth session (27 February - 1 March 2024)*. Economic and Social Council, Official Records, 2024, Supplement No. 4. E/2024/24-E/CN.3/2024/36. 2024.

As Secretariat of the World Population and Housing Census Programme, the United Nations Statistics Division coordinated the revision process for this fourth edition. The work of revision was undertaken by the seven task teams, each responsible for revising specific thematic components of the publication. The expert group, under which these task teams operated, held three meetings[3] to review the text of the draft of the fourth revision of the *Principles and Recommendations for Population and Housing Censuses,* taking into account good practices and contemporary approaches in census taking.

At its fifty-sixth session in March 2025, the United Nations Statistical Commission endorsed the *Principles and Recommendations for Population and Housing Censuses, Revision 4* and encouraged countries to begin its implementation.[4] The Commission further *adopted* the draft resolution on the 2030 World Population and Housing Census Programme and its submission to the Economic and Social Council for adoption. The resolution, in particular, recommends that Member States consider the fourth revision of the Principles and Recommendations as the international standard for conducting population and housing censuses and for concepts, definitions and classifications on population and related statistics, to promote integration across data sources and statistical domains and to enhance data comparability within and across countries.

The key areas of revision in the *Principles and Recommendations for Population and Housing Censuses, Revision 4* compared to Revision 3 include the use of administrative data, geospatial integration, quality assurance, and contingency planning. The revised publication is structured into five parts: Part One: Essential Features and Census Methodology; Part Two: Planning, Organization and Management; Part Three: Census Operation Activities; Part Four: Quality Assurance; and Part Five: Population and Housing Census Topics.
The revised census recommendations detail alternative methodologies to the full field enumeration for producing census statistics and methods for estimating population in unenumerated areas. Emphasising the crucial role of establishing a quality assurance system and setting quality standards from the outset, they incorporate the new Part Four on Quality assurance which closely aligns with the United Nations National Quality Assurance Framework (UNNQAF).[5]

Other notable changes involve the topics collected in censuses. Several new topics have been introduced, including reason for change of residence (for measuring internal migration), history of residing abroad, and reason for international migration. For countries with developing CRVS systems, data collection on vital event registration and legal identity assignment is recommended. Recognizing the significance of ICT in modern life, a new topic on individual ICT use has been added.

Changes have been made to the concept and measurement of status in employment, an economic characteristic, in accordance with the 2018 resolution on work relationships adopted by the Twentieth

[4] United Nations Statistical Commission. *Report on the fifty-sixth session (4-7 March 2025)*. Economic and Social Council, Official Records, 2025, Supplement No. 4. E/2025/24-E/CN.3/2025/37. 2025.
[5] United Nations, Department of Economic and Social Affairs. *National Quality Assurance Frameworks Manual for Official Statistics: Including recommendations, the framework and implementation guidance*. New York: United Nations, 2019.

International Conference of Labour Statisticians (20th ICLS)[6]. Furthermore, for countries aiming to create a frame of agricultural holdings for a future agricultural census, one agriculture topic has been updated to "Measure of farmland area and number of livestock," replacing "Characteristics of all agricultural activities during the last year" from Revision 3.

[6] International Labour Office, Department of Statistics. *Report of the Conference, 20th International Conference of Labour Statisticians (Geneva, 10-19 October 2018)* (Report III). Geneva: ILO, 2018.

Acknowledgements

The Principles and Recommendations for Population and Housing Censuses, Revision 4, was prepared by the United Nations Statistics Division (UNSD), a division of the Department of Economic and Social Affairs (DESA), under the direction of Mr. Stefan Schweinfest, Director of UNSD. The revision was led and substantively guided by Francesca Grum, Assistant Director and Chief of the Demographic and Social Statistics Branch of UNSD. Seiffe Tadesse, Statistician in the same Branch, coordinated and provided technical expertise and oversight throughout the revision process.

This publication reflects a collective effort, incorporating contributions from census experts across national statistical offices, census agencies, regional organizations, and UN and international agencies. The revision was carried out by an expert group comprising over 70 census specialists from all world regions and relevant organizations. The chairs of the seven task teams established within this expert group were instrumental in guiding the revision process, ensuring the publication aligns with current good practices and contemporary census approaches.

The United Nations Statistics Division expresses its appreciation to the following task team chairs:

- Geoffrey Bowlby, Assistant Chief Statistician, Census, Regional Services and Operations, Statistics Canada - Chair of Task Team 1: Census planning, organizing, and management
- Eric Schulte Nordholt, Project Leader, Census, Statistics Netherlands - Chair of Task Team 2: Census methodologies, use of administrative data, and population concepts and definitions
- Cayo de Oliveira Franco, Geography Manager, Brazilian Institute of Geography and Statistics (IBGE) - Chair of Task Team 3: Use of geospatial information in census operations
- Mauricio Rodríguez Abreu, Director General of Sociodemographic Statistics, National Institute of Statistics and Geography (INEGI) and Mitali Sen, Chief, Technical Assistance and Capacity Building Branch, International Programs Center (IPC), U.S. Census Bureau (former chair) - Chair of Task Team 4: Quality assurance, indicators, and assessments
- Luqmaan Omar, Chief Director for Data Operations, Statistics South Africa - Chair of Task Team 5: Data collection and processing
- Richard McMahon, Assistant Director General and Head of Social and Demographic Statistics Directorate, Central Statistics Office, Ireland - Chair of Task Team 6: Communication, dissemination, and utilization
- Thomas Spoorenberg, Senior Population Affairs Officer, UN Population Division, Department of Economic and Social Affairs (DESA) - Chair of Task Team 7: Recommended topics on population and housing characteristics

The United Nations Statistics Division also expresses its appreciation to the task team members for their review, comments, and written contributions to the revision. The revision benefited greatly from these census experts. UNSD especially thanks the following task team members:

Members of **Task Team 1**: Mark Harding (Australia), Cimar Azeredo (Brazil), Sergio Esteban Gordillo Alvarez (Colombia), Sanjay (India), Vijay Kumar (India), Sarpono (Indonesia), Job Nyandwaki Mose (Kenya), Salah Al Muzahmi (GCC-Stat), Carolina Cavada (UNECLAC), Tapiwa Jhamba (UNFPA-HQ), Willis Odek (UNFPA-ASRO), Margaret Mbogoni (UNSD)

Members of **Task Team 2**: Emily Walter (Australia), Ali Said (Indonesia), Gerardo Gallo and Donatella Zindato (Italy), Espen Andersen (Norway), Yousuf Al Riyami (Oman), Nasser Al Mohdi (Qatar), Mosidi Nhlapo (South Africa), Jorge Luis Vega Valle (Spain), Aisha Turki Al Ali and Ebtesam Al Shehhi (UAE), David Thorogood (Eurostat), Jeff Montgomery (Pacific Community (SPC)), Fiona Willis-Núñez (UNECE), Afsaneh Yazdani (UNESCAP), Mr. Romesh Silva (UNFPA-HQ), Christophe Lefranc (UNFPA-AP), Clare Menozzi (UN Population Division), Maria Isabel Cobos Hernandez and Haoyi Chen (UNSD)

Members of **Task Team 3**: Amal M. El Sharnoby (Egypt), Sarpono (Indonesia), Hellen Wanyoike (Kenya), Janusz Dygaszewicz, Tomasz Klimanek, Amelia Wardzińska-Sharif and Anna Sławińska (Poland), Josh Coutts and Julia Lopez (United States), Michele Melchiorri (European Commission), Claudia Baranzelli (OECD), Mathias Kuepie (UNFPA-WCA), Dennis Mwaniki (UNHABITAT), Brian Blankespoor (World Bank), Mark Iliffe, Adriana Skenderi and Elena De Jesus (UNSD)

Members of **Task Team 4**: Donatella Zindato (Italy), Assia Nijari (Morocco), Marcin Szymkowiak (Poland), Papa Ibrahima Silmang Sene (Senegal), Deborah Stempowski (United States), Ali bin Mahboob Al Raisi (UAE), Daniel Scheuregger (Eurostat), Tapiwa Jhamba (UNFPA-HQ), Fredrick Okwayo (UNFPA-ESA), Patrick Gerland and Thomas Spoorenberg (UN Population Division), Hiroaki Ogawa (UNSD)

Members of **Task Team 5**: Paul Waweru (Kenya), Norfariza Hanim Binti Kasim (Malaysia), Papa Ibrahima Silmang Sene (Senegal), Seif Kuchengo (Tanzania), Jennifer Reichert (United States), Oliver Fischer (IPC/US Census Bureau), Mehmet Dogu Karakaya (UNHABITAT), Sergiy Radyakin (World Bank), Margaret Mbogoni and Iliana Vaca Trigo (UNSD)

Members of **Task Team 6**: Aoife Connors (Australia), Larry Bartleet (United Kingdom), Fabian Bach (Eurostat), Fiona Willis-Núñez (UNECE), Fabiana Del Popolo (UNECLAC), Ipek Dedeoglu (UNESCWA), Alessio Cangiano (UNFPA-HQ), Willis Odek (UNFPA-ASRO), Olivier Dupriez (World Bank), Patrick Gerland (UN Population Division), Maria Isabel Cobos Hernandez and Hiroaki Ogawa (UNSD)

Members of **Task Team 7**: Sanjay (India), Vijay Kumar (India), Ali Said (Indonesia), Job Nyandwaki Mose and Paul Waweru (Kenya), Mauricio Rodríguez Abreu (Mexico), Yousuf AL Riyami (Oman), Mosidi Nhlapo and Angela Ngyende (South Africa), Ruth Davison Minja (Tanzania), Ali bin Mahboob Al Raisi (UAE), Lalita Sohai (CARICOM), Lewis Dijkstra (European Commission), David Thorogood (Eurostat), Elisa Benes and Kieran Walsh (ILO), Nobuko Mizoguchi (IPC/US Census Bureau), Fabiana Del Popolo and Xavier Mancero (UNECLAC), Ipek Dedeoglu (UNESCWA), Mehmet Dogu Karakaya (UNHABITAT), Romesh Silva (UNFPA-HQ), Sabrina Juran (UNFPA-LAC), Clare Menozzi (UN Population Division), Iliana Vaca Trigo and Elena De Jesus (UNSD)

The publication benefited from extensive global consultations with the broader statistical community and census experts in national statistical offices and regional and international organizations, who provided feedback on both its content and structure. The contribution of ECLAC and EGRISS in conducting a comprehensive review of the publication is noted.

The United Nations Statistics Division, as Secretariat of the decennial World Population and Housing Census Programme, planned, organized, and coordinated the revision process, including expert group meetings, text review, and publication preparation. The core team that contributed to the revision process comprised Francesca Grum, Maria Isabel Cobos Hernandez, Meryem Demirci (former staff), Seiffe Tadesse, Adriana Skenderi, Andrew Smith, Iliana Vaca Trigo, Hiroaki Ogawa, Elena De Jesus and

Predrag Savic (former staff). The team performed coordination roles and contributed to the substantive review of the publication.

The comprehensive review of the publication by Mr. Ian White, consultant to UNSD, is gratefully acknowledged.

The United Nations Statistics Division appreciates UNECE's close collaboration and contributions to the revision of the Principles and Recommendations, especially their efforts to maximize harmonization between the European and global census recommendations for the 2030 round of population and housing censuses.

Introduction

The population and housing census is a cornerstone within national statistical systems, offering a comprehensive snapshot of a country's human capital and its spatial distribution. It addresses fundamental questions about a nation, specifically "How many are we?" and "Who and where do we live?", thereby delivering crucial data that enables evidence-based decision-making. This information is of vital importance to governments, policymakers, researchers, and a wide range of stakeholders, empowering them to effectively address societal needs, allocate resources equitably, and monitor progress towards both national and international development goals.

The census provides detailed data encompassing demographic, social, and economic characteristics, including age, sex, education, employment, and housing conditions. This information is indispensable for a variety of critical purposes. It guides the formulation and planning of policies and programs across diverse sectors such as health, education, housing, and infrastructure. It also enables the equitable distribution of public funds and services based on the specific needs of the population. Furthermore, the census plays a key role in economic analysis, providing insights into labor force dynamics, market trends, and overall economic growth. It is essential for social development, facilitating efforts to address inequalities, promote social inclusion, and monitor the well-being of vulnerable populations. In addition, the census aids in disaster preparedness and response by identifying populations at risk and supporting effective relief efforts.

Beyond these practical applications, the census provides a rich dataset for in-depth research and analysis of population dynamics and societal trends. Finally, it is instrumental in monitoring progress towards national development plans and international frameworks, such as the Sustainable Development Goals (SDGs).

A critical function of the census is its role in providing the framework for subsequent sample surveys. The census provides a detailed and comprehensive listing of the population and housing units, which serves as the basis for selecting representative samples for more focused and in-depth surveys on specific topics. These sample surveys rely on the census frame to ensure that the selected sample accurately reflects the characteristics of the entire population, allowing for reliable and statistically valid inferences.

Moreover, the census plays an increasingly important role in integrating data from various sources. As countries modernize their statistical systems, they seek to combine census data with information from administrative registers, such as those related to vital events, social security, and taxation, as well as data from other surveys. This integration, often conceptualized as building a "population spine," allows for a more dynamic and nuanced understanding of the population, reducing the need for frequent large-scale enumerations and improving the timeliness and accuracy of statistics. The census serves as the core of this spine, providing the fundamental demographic data to which other data can be linked.

While the traditional approach of full field enumeration remains prevalent, modern censuses are increasingly incorporating new technologies and methodologies to enhance efficiency, accuracy, and timeliness. These advancements include the use of register-based censuses, which utilize existing administrative records, offering a more cost-effective and less burdensome approach for respondents. Combined censuses integrate data from administrative registers with data collected through traditional enumeration or sample surveys, to improve data quality and reduce the scope of fieldwork. Modern

censuses are also characterized by the integration of various technologies. Computer-assisted personal interviewing (CAPI) utilizes tablets or other mobile devices for data collection, while online self-enumeration allows individuals to complete census questionnaires electronically. Geographic information systems (GIS) are employed for improved census planning, enumeration, and dissemination. Finally, the efficiency of data processing, analysis, and dissemination is enhanced through the use of advanced computing power and statistical software.

These modern approaches, while offering significant advantages, necessitate careful planning, a robust technological infrastructure, and strong data protection mechanisms. The guidelines presented herein aim to provide a framework for countries to conduct censuses that are both methodologically sound and responsive to the evolving data needs of the 21st century.

Preface i
Acknowledgements iv
Introduction vii

PART ONE. ESSENTIAL FEATURES AND CENSUS METHODOLOGY 1
I. Essential roles of the census 1
II. Definitions and key features 2
A. Definitions 2
1. Population census 2
2. Housing census 3
B. Key features 3
1. Individual enumeration 3
2. Universality within a defined territory 3
3. Simultaneity 4
4. Defined periodicity 4
5. Capacity to produce small-area statistics 5
III. Uses of population and housing censuses 5
A. Uses of population censuses 6
1. Uses for policymaking, planning and administrative purposes 6
2. Uses for research purposes 7
3. Uses for business, industry and labour 7
4. Uses for boundary delimitation 8
5. Use for population estimates and projections 8
6. Use as a sampling frame for surveys 9
7. Use in humanitarian settings 10
8. Use for monitoring internationally agreed development goals and international comparisons 11
9. Use in small area estimation 12
10. Use in integrated population data systems 12
B. Uses of housing censuses 13
1. Uses for development of benchmark housing statistics 13
2. Uses for the formulation of housing policy and programmes 14
3. Assessment of the quality of housing 14
C. Relationships of population and housing censuses with other statistical activities 15
1. Relationship between the population census and the housing census 15
2. Relationship of population and housing censuses to intercensal sample surveys 16
3. Relationship of population and/or housing censuses to other types of censuses and other statistical investigations 16
4. Relationship between the population census and/or housing census and administrative data sources 20
D. Integration of census data with geospatial information and other data sources 25
1. Integration of census data with other data sources 25
2. Integration of census data with geospatial information 26
IV. Census methodology 29
A. Full field enumeration 32
1. Types of censuses conducted with full-field enumeration 34
2. Innovations in conducting field-based censuses 37
B. Register-based censuses 40
1. Necessary pre-conditions for the use of administrative registers for censuses 42
2. Types of registers 45
3. Establishing a statistical population register 48
4. Constructing integrated statistical registers 49

5. Transitioning from field-based censuses to register-based approaches 50
6. Types of register-based censuses 52
7. Advantages and challenges 53
C. Combined census methodologies 55
1. Types of combined census methodologies 57
2. Advantages and challenges 58
V. Methods for estimating population for areas not enumerated in the census 60
A. Requirements for estimating population 60
B. Methodological approaches 61
1. Use of administrative data sources 61
2. Use of survey data 62
3. Modelled methods 62
VI. Strategies for producing more frequent population and housing statistics 63
A. The need for more frequent data on population and housing characteristics 64
B. Use of administrative registers 64
C. Intercensal and continuous surveys 65
D. Use of mobile phone data 66

PART TWO. PLANNING, ORGANIZATION AND MANAGEMENT 68
I. Introduction 68
II. Census planning 68
A. Overall census planning 68
B. Planning for both a population and housing census 70
C. Special considerations when planning a combined or register-based census 70
III. Strategic objectives and management 71
A. Strategic objectives 71
B. Strategic management 74
C. Making censuses more inclusive 75
IV. Units, place and time of enumeration 77
A. Units of enumeration 77
1. Person 78
2. Household 78
3. Population in collective living quarters 80
4. Living quarters 80
5. Housing unit 80
6. Building 81
B. Place of enumeration 81
1. Concepts relating to the place of enumeration 81
2. Operational issues relating to the place of residence and the place of enumeration 83
C. Time of enumeration 85
D. Time reference period for data on the characteristics of the population and of living quarters 85
V. Legal basis 85
VI. Financial management 87
A. Financial basis for censuses 87
B. Budget and cost control 89
VII. Resource mobilization 91
VIII. Project management 92
A. Development of workplans 92
B. Emergency preparedness, risk management and contingency planning 93
1. Documentation of risks and mitigation plans 94
2. Emergency management, governance and decision-making 95
3. Disinformation and misinformation threats to the census 97

4. Crisis response training for census staff 97
5. Ensuring fast workflow during an emergency 98
6. Contingency planning for main plan failure 99
C. Change management 99
IX. Administrative organization 101
A. Overall overview 101
B. Roles and responsibilities of key stakeholders in the census process 102
C. Statistical leadership 104
X. User consultation, communication and publicity 105
XI. Census calendar 112
XII. Human resources management and training 114
XIII. Logistics management 117
A. Procurement management 117
B. Forward and reverse logistics 119
XIV. Contracting out 120
XV. Choice of technologies 124
XVI. Innovation to optimize census processes 128

PART THREE. CENSUS OPERATIONAL ACTIVITIES 130
I. Introduction 130
II. Census questionnaires: content and design 130
III. Building census infrastructure 137
IV. Geospatial information 139
A. Strategic planning 139
B. The role of geospatial information in the census 141
C. Geospatial information technology for census 144
D. Geographic information systems 146
E. Census geography 149
F. Implementation of a census geospatial information programme 155
G. Contracting out and partnerships for census geospatial operations 158
V. Census tests 160
VI. Establishing the census frame 162
VII. Field enumeration 164
A. Method of enumeration 164
B. Multi-mode data collection 170
C. Use of administrative data during field-based enumeration 172
D. Timing and length of the enumeration period 173
E. Management and supervision 175
F. Security during data collection 178
G. Technologies for field enumeration 178
VIII. Data processing 182
A. Method of processing 185
B. Preparation for data capture 186
C. Data capture 186
D. Coding 189
E. Data editing 191
F. Validation 192
G. Processing control 193
H. Master file 194
I. Methods of tabulation 194
J. Use of administrative data during data processing (for field-based censuses) 195
K. Security during data processing 196
IX. Census products, dissemination and utilization 198

A. Introduction 198
B. Plans for census products and data dissemination 198
1. Developing a dissemination strategy 198
2. Tabulation programme 203
3. Geospatial information for analysis and dissemination 203
4. Mode of dissemination of outputs 207
5. Confidentiality and privacy 209
6. Metadata 211
7. Promotion of, and training on, uses of census data 213
C. Census data dissemination: products and services 215
1. Provisional and final results 215
2. Census reports 216
3. Databases 220
4. Geographic products 224
5. Interactive electronic outputs 229
6. Microdata dissemination 230
7. Customized products 235
8. General interest products and special audience products 236
D. Census data utilization 238
1. General uses of population and housing censuses 238
2. Analysis of the results 242
3. Intercensal and post-census population estimates 243
4. Cross-cutting and emerging social issues 243
5. Development indicators 249
X. Overall evaluation of the census 250
A. Importance of evaluations 250
B. Planning for the evaluation 251
XI. Documentation of census experience 252
XII. Archiving individual records 253
A. Purpose of archiving individual records 253
B. Procedures for archiving 254
C. Archiving microdata 255

PART FOUR. QUALITY ASSURANCE 256
I. Quality assurance 256
A. Plans for quality assurance 256
B. Quality assurance components 256
1. Managing the census within the national statistical system 257
2. Managing the institutional environment 257
3. Managing census processes 257
4. Managing census outputs 258
C. Other international standards, tools and references 259
D. The role of managers 261
E. Quality improvement 263
1. Topic selection 264
2. Questionnaire design and development 265
3. Geospatial information 265
4. Development of the data capture system 267
5. Field operations and pilot testing 269
6. Processing 271
7. Analysis 274
8. Dissemination 275
9. Evaluation 275

10. Documentation 277
F. Ensuring and measuring the quality of data derived from administrative sources 277
II. Methods to evaluate the coverage and quality of census results 279
A. Purpose of evaluating the quality of census data 279
B. Methods of census evaluation 281
C. Post-enumeration survey 283
D. Demographic analysis for census evaluation 284
E. Comparison with administrative sources 286
F. Adjusting census data 287
1. Adjustments for full field enumeration censuses 288
2. Adjustments for register-based censuses 288
3. Adjustments for combined censuses 289
4. The decision to adjust, or not to adjust, census figures 290
G. Acceptance of results 290
H. User satisfaction 291

PART FIVE. POPULATION AND HOUSING CENSUS TOPICS 293
Chapter I: Population census topics 293
I. Factors determining the selection of topics 293
A. National priority 294
B. International comparability 294
C. Suitability 295
D. Resources 295
E. Alternative sources 295
II. List of topics 296
III. Population count 300
A. Population present count 301
B. Usual resident population count 302
C. Supplementary population counts 304
1. Service population count 304
2. Counts of population subgroups 305
D. Difficult-to-enumerate groups or difficult-to-reach groups 305
IV. Definitions and specifications of topics 309
A. Geographic and internal migration characteristics 309
1. Place of usual residence (core topic) 310
2. Place where present at time of census (core topic) 310
3. Place of birth (core topic) 311
4. Duration of residence (core topic) 312
5. Place of previous residence (core topic) 313
6. Place of residence at a specified date in the past (core topic) 313
7. Reason for change of residence 314
8. Total population (core topic) 314
9. Locality (core topic) 316
10. Urban and rural (core topic) 317
B. International migration characteristics 318
1. Country of birth (core topic) 319
2. Country of citizenship (core topic) 320
3. Acquisition of citizenship 321
4. Year or period of arrival in the country (core topic) 321
5. Ever resided abroad 322
6. Reason for international migration 323
C. Household and family characteristics 324
1. Relationship to the reference person of household (core topic) 326

2. Household and family composition (core topic) 328
3. Household and family status 330
D. Demographic and social characteristics 332
1. Sex (core topic) 332
2. Age (core topic) 333
3. Marital status (core topic) 334
4. Ethno-cultural characteristics 336
5. Religion 336
6. Language 337
7. Ethnicity 337
8. Indigenous peoples 339
9. Disability status (core topic) 340
a. Disability characteristics 341
b. Use of the census to measure disability at the aggregate level 341
c. Census question wording 344
d. Use of the census to screen for disability and follow-up with other surveys 345
E. Fertility and mortality 345
1. Children ever born alive (core topic) 349
2. Children living (core topic) 350
3. Date of birth and sex of last child born alive (core topic) 351
4. Births in the past 12 months 352
5. Deaths among children born in the past 12 months 352
6. Age, date or duration of first marriage or union 353
7. Age of mother at birth of (date or time since) first child born alive 353
8. Household deaths in the past 12 months (core topic) 354
9. Maternal or paternal orphanhood 356
F. Registration of vital events and assignment of legal identity 356
G. Educational characteristics 360
1. Literacy (core topic) 360
2. School attendance (core topic) 361
3. Educational attainment (core topic) 362
4. Field of education and educational qualifications 364
H. Individual use of information and communication technology 366
1. Individuals using the Internet 367
2. Individuals owning mobile cellular phones 367
I. Economic characteristics 367
1. Introduction 367
2. Conceptual framework for work statistics 370
3. Labour force status (core topic) 373
4. Characteristics of jobs and establishments 380
5. Status in employment (core topic) 381
6. Occupation (core topic) 386
7. Industry (core topic) 386
8. Place of work 387
9. Institutional sector of employment 389
10. Working time 389
11. Participation in own-use production of goods (core topic) 390
12. Income 391
J. Agriculture 393
1. Introduction 393
2. Own-account agriculture production 393
3. Measure of farmland area and number of livestock 393
Chapter II. Housing census topics 395

I. Factors determining the selection of topics 395
A. National priority 396
B. International comparability 397
C. Suitability 397
D. Resources 397
E. Alternative sources 397
II. List of topics 398
III. Definitions and specifications of topics 401
1. Living quarters – type of (core topic) 401
2. Location of living quarters (core topic) 410
3. Occupancy status (core topic) 411
4. Occupancy by one or more households (core topic) 412
5. Occupants – number of (core topic) 413
6. Ownership – type of (core topic) 413
7. Rooms – number of (core topic) 415
8. Bedrooms – number of 415
9. Useful floor space 415
10. Water supply system (core topic) 416
11. Drinking water – main source of (core topic) 417
12. Toilet – type of (core topic) 418
13. Sewage disposal (core topic) 419
14. Solid waste disposal – main type of (core topic) 419
15. Bathing facilities (core topic) 420
16. Kitchen – availability of (core topic) 421
17. Fuel used for cooking (core topic) 422
18. Energy used for lighting – type of (core topic) 422
19. Heating – type and energy used 422
20. Hot water – availability of 423
21. Piped gas – availability of 423
22. Use of the housing unit 423
23. Building – type of (core topic) 423
24. Year or period of construction 425
25. Dwellings in the building – number of 426
26. Position of dwelling in the building 426
27. Accessibility to dwelling 427
28. Construction material of outer walls (core topic) 427
29. Construction material of floor and roof 428
30. State of repair 428
31. Elevator – availability of 429
32. Farm building 429
33. Age and sex of the reference person of household (core topic) 429
34. Tenure (core topic) 430
35. Rental and housing costs 430
36. Furnished/unfurnished 431
37. Information and communication technology – availability of (at the household level) 431
38. Cars – number of available 432
39. Durable household appliances – availability of 432
40. Outdoor space – access to 432

Glossary of definitions and acronyms 434

PART ONE. ESSENTIAL FEATURES AND CENSUS METHODOLOGY

I. Essential roles of the census

1.1. Evidence-based decision-making is a universally recognized paradigm of efficient management of economic, social and environmental affairs and of overall effective governing of societies today. Making use of relevant, accurate and timely statistics is essential to this paradigm; producing detailed statistics for small areas and small population groups is its foundation. The role of the population and housing census is to produce and disseminate such detailed small-area statistics on population, its composition, characteristics, spatial distribution and organization into families, households and dwellings. Censuses have been promoted internationally since the end of the nineteenth century, when the International Statistical Congress recommended that all countries in the world conduct them.[7] Since 1958, the United Nations has been actively promoting the population and housing census by compiling principles and recommendations and launching regular decennial worldwide census programmes. In fact, the majority of the countries in the world conduct a population and housing census periodically. In order to increase the relevance of the data produced by censuses, several countries are now moving towards using alternative census methodologies to produce census-like statistics with a more frequent periodicity than the traditional decennial or quinquennial one.

1.2. While the roles of the population and housing census are many and will be elaborated in detail throughout the present revision of these *Principles and Recommendations for Population and Housing Censuses*, several of the essential roles are summarised here:

(a) **The population and housing census plays an essential role in public administration**. The results of a census are used as a critical reference to ensure equity in the distribution of resources, government services and representation nationwide by informing the distribution and allocation of government funds among various regions and districts for education, health services, delineating electoral districts at the national and local levels, and measuring the impact of industrial development, to name a few. Establishing a public consensus on priorities would be almost impossible without census counts. A wide range of other users, including the private sector, academia, civil society and individuals, make use of census outputs.

(b) **The census also plays an essential role in all elements of the national statistical system**, including the economic, social and environmental components. Census statistics are used as benchmarks for statistical compilation, definitions and classifications, and as a sampling frame for sample surveys. Today, the national statistical system of almost every country relies on sample surveys for efficient and reliable data collection, notwithstanding the increasing availability of non-traditional data sources. Without the sampling frame and population benchmarks derived from the

[7] *Report of the Proceedings of the Fourth Session of the International Statistics Congress, held in London July 16th, 1860, and the Five following Days*, printed by George Edward Eyre and William Spottiswoode, London, 1861.

population and housing census, the national statistical system would face difficulties in providing reliable official statistics for use by the government and the general public. For many countries with less developed statistical systems, a census is also an essential source providing data to integrate with survey data for producing model-based small-area estimations.

(c) **A basic function of a census is to generate statistics on small areas and small population groups**. While statistics on small areas are useful *per se*, they are important because they can be used to produce statistics on any geographic unit with arbitrary boundaries, subject to any statistical disclosure control restrictions for overlapping boundaries. For example, in planning the location of a school, it is necessary to have the data on the distribution of school-age children by school area, which may not necessarily correspond to the administrative area units. Similarly, small-area data from the census can be combined to approximate natural regions (for example, water catchments or vegetation zones) that do not follow administrative boundaries. Since census data can be tabulated for any geographic unit, it is possible to provide the required statistics in a remarkably flexible manner. This versatile feature of the census is also invaluable for use in the private sector for applications such as business planning and market analyses. Furthermore, the census is essential for generating accurate and reliable statistics for small population groups, such as ethnic minorities, indigenous peoples, migrants, and people with disabilities. These statistics are crucial for understanding the needs and experiences of these groups and for developing targeted policies and programmes.

(d) **Census data are used as a benchmark for research and analysis**. They are used to study population dynamics. Population projections, as an example, are one of the most important analytical outputs based on census data and are crucial for all segments of the public and private sectors.

1.3. It is critically important to produce detailed statistics for small areas and small population groups **as a building block for efficient governance at all levels**. For many countries the method for assembling this building block will be by conducting a population and housing census through universal and simultaneous individual enumeration of each set of living quarters and of each individual within the country's boundaries. Some countries will adopt alternative approaches; yet, all of these methods should produce detailed, timely statistics for small areas and small population groups at a well-defined point or period in time.

II. Definitions and key features

A. Definitions

1. Population census

1.4. A population census is the total process of planning, collecting, processing, analysing, disseminating and evaluating demographic, economic and social data at the smallest geographic level pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country.

1.5. In order to plan for, and implement, economic and social development, administrative activity or scientific research, it is necessary to have reliable and detailed data on the size, distribution and composition of population. The population census is a primary source of these basic benchmark statistics,

covering not only the settled population but also homeless persons[8] and nomadic groups. Data from population censuses should allow the presentation and analysis of statistics on persons and households and for a wide variety of geographic units, ranging from the country as a whole to individual small localities, villages or city blocks.

2. Housing census

1.6. A housing census is the total process of planning, collecting, processing, analysing, disseminating and evaluating statistical data relating to the number and condition of housing units and facilities as available to the households pertaining, at a specified time, to all living quarters[9] and occupants thereof in a country or in a well-delimited part of a country.

1.7. A housing census must provide information on the supply of housing units together with information on the structural characteristics and facilities that have a bearing upon the maintenance of privacy and health and living conditions. Sufficient demographic, social and economic data concerning the occupants must be collected to furnish a description of housing conditions and also to provide basic data for analysing the causes of housing deficiencies and for studying possibilities for remedial action. In this connection, data obtained as part of the population census, including data on homeless persons, are often used in the presentation and analysis of the results of the housing census, if both operations are conducted together or there is a link between them.

B. Key features

1.8. The key features of population and housing censuses are individual enumeration, universality within a defined territory, simultaneity, defined periodicity and the capacity to produce small-area statistics.

1. Individual enumeration

1.9. The term "census" implies that each individual and each set of living quarters is enumerated separately and that the characteristics thereof are separately recorded. Only by this procedure can the data on the various characteristics be cross-classified. The requirement of individual enumeration can be met by the collection of information in the field (face-to-face interviews or self-response), by the use of information contained in an appropriate administrative register or set of registers, or by a combination of these methods.

2. Universality within a defined territory

[8] Homelessness is a condition where individuals or households lack stable, safe, and adequate housing. Due to cultural differences in perceiving concepts such as "adequate housing" and "security of tenure," the definition of homelessness can vary across countries. See paragraph 2.46. for the definition of “homelessness”.
[9] For the definition of "living quarters", see paragraph 5.479.

1.10. The census should cover a precisely defined territory (for example, the entire country or a well-delimited part of it). The population census should include every person present and/or residing within its scope, depending upon the type of population count required, and should include them only once. The housing census should include every set of living quarters irrespective of type, and should include each one only once. This does not preclude the use of sampling techniques for obtaining data on specified characteristics, provided that the sample design is consistent with the size of the areas for which the data are to be tabulated and the degree of detail in the cross-tabulations to be made.

3. Simultaneity

1.11. Each person and each set of living quarters should be enumerated at the same well-defined point in time and the data collected should refer to a well-defined reference period. The key benefit of referring all data to a common date or reference period is to minimise over-and under-counting. The time reference period need not, however, be identical for all of the data collected. For most of the data, it will be the day of the census; in some instances, it may be a period prior to the census.[10] However, it is recognised that not all census methodologies satisfy this criterion.[11]

4. Defined periodicity

1.12. Censuses should, ideally, be taken at regular intervals so that comparable information is made available in a fixed sequence. A series of censuses makes it possible to appraise the past, accurately describe the present and estimate the future. It is recommended that a national census be taken at least every 10 years. Some countries may find it necessary to carry out censuses more frequently because of the rapidity of major changes in their population and/or its housing circumstances. The use of data derived from administrative sources may provide the opportunity for more frequently produced statistical outputs.

1.13. The census data of any country are of greater value nationally, regionally and internationally if they can be compared with the results of censuses of other countries that were taken at approximately the same time. Therefore, countries should make all efforts to undertake a census in years ending in “0” or at a time as near to those years as possible. It is obvious, however, that legal, administrative, financial and other considerations often make it impracticable for a country to strictly adhere to a standard international pattern in the timing of its censuses. Crises such as pandemics, wars, or other significant events can compel adjustments to the census schedule. Some regional groupings of countries, such as the European Union, have their own requirements regulating the dates of censuses among their members[12].

[10] For example, collecting information on the core topic of household deaths in the past 12 months (see paragraphs **Error! Reference source not found.**-5.280).

[11] Simultaneity is an essential feature of censuses that rely on full field enumeration (the so-called traditional approach) (see paragraphs 1.137. -1.158.) as well as of some types of register-based censuses (see paragraph 1.170. -1.240.).

[12] For EU countries it is mandatory to have censuses in years ending in "1".

In fixing a census date, therefore, such national and regional factors could be given greater weight than the desirability of international simultaneity.

5. Capacity to produce small-area statistics

1.14. The census should produce data on the number and characteristics of the population and housing units down to the lowest appropriate geographic level, compatible with national circumstances, and for small population groups, while protecting confidentiality of personal information on each individual.

III. Uses of population and housing censuses

1.15. Population and housing censuses are a principal means of collecting basic population and housing statistics. They form the core of an integrated programme of data collection and compilation aimed at providing a comprehensive source of statistical information for economic and social development planning, administration, assessing conditions in human settlements, research and commercial and other uses. Population and housing censuses are often a basis for sample surveys and provide insights and updates into buildings, household and population characteristics as well as increasingly being used as a source of environmental and sustainable development statistics and geospatial information – addresses, coordinates, population distribution, buildings and settlements, etc.

1.16. The value of population and housing censuses is increased if the results can be employed together with other data sources, as in the use of the census data as a basis or benchmark for current statistics, and if it can furnish the information needed for conducting other statistical data collection. The census can, for example, provide a statistical frame for other censuses or sample surveys. The population census is also important in developing the population estimates needed to calculate vital rates in combination with civil registration data (see paragraphs 1.84. –1.89.). In addition, these censuses are a major source of data used in official compilations of social indicators, particularly on topics that usually change slowly over time. It is important, therefore, to take into account the requirements of a continuous, coordinated programme of data collection and compilation when planning a census. When the relationships between the population census, the housing census and other statistical investigations are considered from the outset, their shared purposes can best be served and their benefits best harnessed. The use of consistent concepts and definitions throughout an integrated programme of data collection and compilation is essential if the advantages of these relationships are to be fully realized. Of course, census-type information can also be derived from population registers and can also be estimated from sample surveys without undertaking a complete enumeration. These alternative data sources are discussed under "Census methodology" in paragraphs 1.131. –1.265. .

1.17. Data to measure and monitor indicators are required by countries to track the progress towards internationally agreed development agendas, including the 2030 Agenda and its Sustainable Development Goals. As such, efforts must be made by census offices to produce the relevant data to meet these commitments.

1.18. Increasingly, population and housing censuses are a key source of data in the compilation of environment and sustainable development statistics. They provide environmental agencies with a range of information to analyze energy consumption, identify conservation opportunities, and forecast energy needs. Census data help national and local government and relief agencies in planning relief operations to assist populations and areas affected by natural disasters such as floods, hurricanes, tornadoes, and

earthquakes. The census also provides a rich source of data for public health authorities in identifying prevalence levels of disease, responses to pandemics and measurement of outcomes of public health initiatives. Moreover, census data are frequently integrated with other thematic data to derive essential indicators for comprehensive analyses.

1.19. A population and housing census also serves as the logical starting point for developing statistical products to meet ongoing national and local data needs in the intercensal period.[13] For example, census data can be used to create demographic projections, analyse social and economic trends, and inform policy decisions.

1.20. In addition to the statistical value obtained directly from the census results themselves, there are further, indirect benefits from taking a census, particularly to the organization responsible for the census, or the national statistical office (NSO). These benefits include:

(i) **Improved skills and experience:** Varied sets of skills are often required for administering a census that are not necessarily prominent in other parts of the organization, such as project management, procurement and commercial, communication, human resources, and geospatial (GIS) and information technology (IT) skills.

(ii) **Technological advancement:** Often a census requires new technology to support complex data collection and processing requirements. These developments may be reused for other exercises within the NSO or lead to new technological developments.

(iii) **New methods:** The development of methods for enumerating the whole population, or statistical methods (such as edit and imputation) developed for processing census results, can often be reused for other statistical exercises within the NSO such as stratification of the population, estimation and modelling population characteristics, or the development of indicators.

(iv) **Halo effect:** The extensive promotion and visibility of the census can have a positive ripple effect on public perception of other statistical activities and products, enhancing their image and increasing response rates in subsequent surveys. This phenomenon - "halo effect" - occurs when the success and trust generated by the census extends to other statistical investigations, bolstering their credibility and acceptance among respondents.[14]

A. Uses of population censuses

1. Uses for policymaking, planning and administrative purposes

1.21. The fundamental purpose of the population census is to provide the facts essential to national policymaking, planning and administration. Information on the size, distribution and characteristics of a country's population is essential for describing and assessing its economic, social and demographic circumstances and for developing sound policies and programmes aimed at fostering the welfare of a country and its population. The population census, by providing comparable basic statistics for a country

[13] See Chapter X, Part Three of this publication.

[14] The halo effect is a cognitive bias in which an observer's overall impression of a person, company, brand or product influences the observer's feelings and thoughts about that entity's character or properties. The halo effect is a specific type of confirmation bias, wherein positive feelings in one area cause ambiguous or neutral traits to be viewed positively.

as a whole and for each administrative unit, locality and small area therein, can make an important contribution to the overall planning process and the management of national affairs. Counts of the population overall, or of subgroups within the population, by geographic region are often used for the distribution of government funding, development programmes and services. Population censuses in many countries represent the very foundation of their national statistical systems, with the census providing important baseline data for policy development and planning, for managing and evaluating programme activities across a broad range of sectoral applications, and for monitoring overall development progress. Emerging uses of census data are the assessment of good governance by civil society groups and measuring the achievement of Sustainable Development Goals. The performance of a democratically elected government in improving the welfare of its citizens can be monitored from one census to another by ordinary citizens through the widespread and timely dissemination of census results.

1.22. Geo-referenced statistics, in particular at subnational and local levels, greatly increase the relevance of statistical information by providing the geographic context of the characteristics that censuses or surveys are portraying. This allows policymakers and researchers to more easily understand and analyse this geographic relationship, leading to the development of more targeted, locally relevant and actionable plans, such as access to physical infrastructure, urban-regional planning, and the planning to deal with climate change, disasters and pandemics.

2. Uses for research purposes

1.23. In addition to serving specific governmental policy purposes, the population census provides indispensable data for the scientific analysis and appraisal of the composition, distribution and past and prospective growth of the population. The changing patterns of urban–rural concentration, the development of urbanized areas, the geographic distribution of the population according to such variables as occupation and education, the changes in the sex and age structure of the population, household composition, and the mortality, fertility and migration differentials for various population groups, as well as the economic and social characteristics of the population and the labour force, are questions of scientific interest that are of importance both to research and for solving practical problems of industrial and commercial growth and understanding population dynamics.

3. Uses for business, industry and labour

1.24. In addition to those uses given above, the census has many important uses for individuals and institutions in business, industry and labour. Reliable estimates of consumer demand for an ever-expanding variety of goods and services depend on accurate information on the size and location of the population in subnational areas and its distribution at least by sex and age, since these characteristics heavily influence the demand for housing, transport, furnishings, food, clothing, recreational facilities, medical supplies and so forth. Furthermore, the census can be used to generate statistics on the size and characteristics of the supply and demand of labour needed for the production and distribution of such commodities and services in conformity with International Labour Organization statistical standards.[15]

[15] See Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (October 2013).

Such statistics on the local availability of labour may be important in determining the location and organization of enterprises.

4. Uses for boundary delimitation

1.25. One of the basic administrative uses of census data is to support political and administrative mapping. Detailed information on the geographic distribution of the population is indispensable for this purpose. Certain aspects of the legal or administrative status of territorial divisions may also depend on the size and characteristics of their populations, for example, whether a previously rural area is now to be declared as urban.

1.26. A compelling use of census data is in the redrawing of electoral constituency boundaries in most countries. This is often enshrined in the country's constitution and provides a legal basis for census taking. The current distribution of a country's population is thereby used to assign the number of elected officials who will represent people in the country's legislature.

5. Use for population estimates and projections

1.27. Census data plays a vital role in estimating and projecting future population trends. These estimates and projections are crucial for analytical and planning purposes.

1.28. Intercensal and post-censal population estimates are calculated using a combination of data sources and statistical methods, which can vary from country to country. These estimates typically include: the base population from the most recent census; estimates of births and deaths, both internal and international migration; and adjustments for specific subgroups, such as age groups. Results from the census evaluation are crucial for improving the accuracy of intercensal and post-census estimates. This evaluation helps identify and correct any deficiencies in the data collection phase, particularly regarding the base population figures.

1.29. It is important to acknowledge that population estimates and projections are not exact. They rely on the accuracy of the baseline population (including distributions by age and sex) as well as on assumptions about future trends, which can change and lead to deviations. Nevertheless, they remain the best tools for understanding the dynamics of future populations. Regularly re-basing these estimates with a population census helps minimize these deviations.

1.30. At the international level, the United Nations produces and issues population estimates and projections (through the World Population Prospects[16]) from 1950 to the present for 237 countries or areas. These projections are grounded in analyses of historical demographic trends, considering the results of national population censuses, civil registration systems, and nationally representative sample surveys.

[16] https://population.un.org/wpp/

6. Use as a sampling frame for surveys

1.31. For countries without comprehensive population registers, population censuses can serve as a foundation for the construction of a sampling frame for surveys during the intercensal years on many topics, such as the labour force, poverty, fertility, mortality, health, time-use, gender-based violence, disability and international migration.

1.32. Censuses provide an areal frame that contains hierarchical geographical areas from the largest (at the national level) to the smallest geographic division, usually called enumeration areas (EAs) and a list frame that contains the list of households located within each EA. For countries with appropriate resources, address-based sampling frames have been gaining popularity given their efficiency and quality. The addresses are updated regularly and important auxiliary variables are available on the frame to help improve sampling efficiency. For countries that do not have the resources or capacity to maintain a comprehensive list of addresses, a master sample frame is often used. With a master sample frame, the address list is updated only for selected EAs.

1.33. It is important, when the census is in the planning stage, to give careful consideration to the construction of a census frame for subsequent use as a survey sample frame. The census frame should be complete in coverage. This means, for example, that care must be taken to ensure that the entire country is divided into EAs without any omissions or overlaps, i.e. all land area belongs to one and only one EA. The coverage of population within the EAs should be complete, regardless whether they reside in private households or not.

1.34. Population and household counts for the EAs, taken from the census, are also a highly useful ingredient for post-census sample survey design planning. This information is often used to establish measures of size for the selection of first- or second-stage sampling units, or to help in various stratification schemes. Characteristics of living quarters, the households and individuals collected in the census are also useful as stratification variables to support efficient sampling design for surveys. For example, these may include the degree of urbanization, age, sex, education attainment, labour force status, nationality/statelessness status, forced displacement status, type of living quarter, as well as other characteristics that may be used to support design of specialized surveys targeting less numerous and/or difficult-to-reach populations such as international migrants, persons with disabilities, indigenous population, etc.

1.35. Efforts should be made to ensure census records are geospatially enabled, i.e., geocoded to a specific location. This facilitates various applications, such as selecting samples for household surveys and integrating census results with administrative data. Geospatially enabled census records allow for spatial analysis, integration with other geospatial information, and enhance the efficiency of data collection and analysis processes.

1.36. Many countries use telephone to follow up with respondents to address missing values and non-response.[17] For countries without a good telephone frame for phone surveys, phone numbers collected during censuses can be used for subsequent surveys. However, this requires the following of strict protocols and consent from respondents.

7. Use in humanitarian settings

1.37. In humanitarian settings, census data can be used to support the identification of crisis affected populations and specifically vulnerable groups.

1.38. Census data can be used in humanitarian settings in a number of ways, including:

- ***To assess the needs of the population***: Census data can be used to identify the number and location of people who have been affected by a humanitarian crisis. This information can be used to plan and deliver relief and recovery efforts. For example, census data can be used to identify the potential number of people who may need food, water, shelter and medical assistance.
- ***To identify, protect and assist vulnerable groups***: Census data can be used to identify population groups facing increased vulnerabilities. This information can be used to target assistance to those who need it most such as women from marginalized communities, women-headed households, children, older persons and persons with disabilities. By identifying vulnerable populations in locations such as refugee camps, census data contribute to policies and actions that help prevent violence, exploitation, and ensure their safety and well-being.

1.39. In addition to these general uses, census data can also be used to address specific humanitarian challenges. For example, census data can be used to:

- ***Track the spread of disease***: Census data plays a vital role in tracking the spread of disease in several ways, including by helping to identify population sub-groups that exhibit demographic characteristics that make them more susceptible to certain diseases based on age, socioeconomic status, or population and housing density. Researchers can use census data to create simulations of how a disease might spread throughout a population, helping to predict potential outbreaks and allowing for targeted interventions before a disease becomes widespread. By comparing data from different censuses, public health officials can track changes in disease prevalence over time. This helps them understand the effectiveness of existing interventions and identify emerging threats.
- ***Plan for long-term recovery***: Census data informs long-term recovery plans following a crisis. By identifying areas needing new schools or hospitals, it supports infrastructure rebuilding and economic development.

1.40. The use of census data is an essential tool for humanitarian organizations. By providing accurate and timely information on the population, such data can help humanitarian organizations to save lives and to build a better future for people affected by crisis. Given that the periodicity of census-taking can affect its usability, it is crucial to generate estimates in the intercensal period to ensure timely and

[17] United Nations Statistical Commission. 51st session, item 3(j), background document. Report on the Results of the UNSD Survey on 2020 round population and housing censuses. 2020. Available from: https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Survey-E.pdf

accurate information. To ensure a coordinated and efficient response during humanitarian crises, the UN System has developed the Common Operational Datasets (CODs).[18] These authoritative reference datasets, updated every two years, are essential for all preparedness and response activities of UN agencies. By utilizing CODs, the humanitarian community can work together to deliver timely and necessary assistance to those affected by crises.

8. Use for monitoring internationally agreed development goals and international comparisons

1.41. Census data are essential for monitoring international development goals. The Sustainable Development Goals (SDGs), adopted by all United Nations Member States in 2015, provide a blueprint for a better and more sustainable future for all. They address the global challenges that many countries face, including poverty, inequality, climate change, environmental degradation, peace and justice. Census data are crucial for calculating the denominators used in numerous SDG indicators, providing the foundational population figures necessary for accurate measurement and analysis.[19,20] By providing accurate and timely information on the population, census data can help to ensure that no one is left behind.

1.42. Census data can be used to track progress towards the SDGs in a number of ways, including:

- **Measuring progress towards SDG indicators**: Census data can be used to monitor progress towards achieving many of the SDG indicators. For example, census data can be used to measure SDG 1 on poverty, SDG 3 on health, SDG 4 on education, SDG 5 on gender equality, SDG 6 on water and sanitation, and SDG 7 on energy.
- **Identifying disparities**: Census data can be used to identify disparities in progress towards the SDGs. For example, such data can be used to identify disparities in access to education and healthcare between different groups of people, such as women and men, rural and urban populations, and among ethnic groups, migrants, forcibly displaced or stateless populations.
- **Monitoring the impact of interventions**: Census data can be used to monitor the impact of interventions designed to achieve the SDGs. For example, the data can be used to monitor the impact of a programme to reduce poverty or improve educational attainment.
- **Geographic disaggregation and geo-statistical integration**: The census is often the only source of information that allows analysis for small geographic areas and specific population subgroups. There are indicators that can only be generated with the integration of census information with geospatial information (e.g., SDG 9.1.1 on rural access to all-season roads).

1.43. As noted previously, data from the census are essential to demonstrate the changing patterns of urban–rural concentration, the development and expansion of urbanized areas and the population

[18] https://cod.unocha.org/

[19] UN Technical Report on Measuring Sustainable Development Goals Indicators through Population and Housing Censuses and Civil Registration and Vital Statistics Data. Link: https://unstats.un.org/unsd/demographic-social/census/documents/tr_on_sdg_in_phc_crvs.pdf

[20] The 2020 census round: challenges of the 2030 Agenda for Sustainable Development, the Sustainable Development Goals and the Montevideo Consensus on Population and Development. Link: https://www.cepal.org/en/publications/46727-2020-census-round-challenges-2030-agenda-sustainable-development-sustainable

distribution. This also applies to monitoring urbanization at a global level, allowing international comparison. The Degree of Urbanisation (DEGURBA)[21] combines population size and population density thresholds to classify the territory of a country on an urban-rural continuum and delineate cities, towns, suburbs, and rural areas.

1.44. Censuses are a key information source to harmonise the definition of urban and rural areas for international statistical comparison (Degree of Urbanisation) as several indicators require urban-rural disaggregation or are sensitive to how urban and rural areas are defined.

9. Use in small area estimation

1.45. Small Area Estimation (SAE) offers a means of estimating indicators when critical data are not available at lower levels of geography (such as district or municipality levels) or for domains of study in which the sample sizes are too small. Small area estimation methods let two sources borrow strength from each other. In producing estimates on poverty, for example, a household survey that collects detailed information on household expenditure or consumption is required. Given the sample size constraint, direct estimates of poverty typically do not reach a user-specified level of precision for smaller geographical areas. Population censuses, on the other hand, have comprehensive coverage of a country's population. The SAE method brings these two sources together to produce poverty estimates for smaller geographical areas. Furthermore, small area estimators can provide predictions for domains where no sample information is available. Other data sources that have been used to integrate with household surveys include administrative data sources and geospatial data.

1.46. SAE has several advantages especially when dealing with development needs. SAE allows better use of existing data within a national statistical system to identify smaller geographic areas or specific population groups with the greatest development needs[22]. With the ability to produce estimates at lower administrative levels or for population groups that are difficult to measure through household surveys, SAE also helps improve targeting of intervention and resource allocation.

10. Use in integrated population data systems

1.47. Population data systems generally refer to integrated frameworks for the collection, management, analysis, and dissemination of population-related data. They draw from diverse data sources, including but not limited to, censuses, surveys, administrative records, geospatial data, and alternative data sources. These systems play a pivotal role in guiding policy and decision-making processes, tracking population trends, highlighting disparities, assessing the effectiveness of programmes, and fostering research and innovation. With rapid advancements in technology and methodology, as well as increasing demands for frequent, granular data, there is a need to envision the future of these systems.

[21] The Degree of Urbanisation (DEGURBA) was adopted by the 51st Statistical Commission of the United Nations in March 2020.

[22] SAE4SDG Toolkit Wiki page (UNSD, April 2022); Introduction to Small Area Estimation Techniques: A Practical Guide for National Statistics Offices (ADB, May 2020).

1.48. Integrated population data systems rely heavily on the census as their foundational element. By presenting a detailed snapshot of the size of the population and its demographic and housing characteristics at a specific point in time, the census provides granular data that forms the backbone of such integrated systems.

1.49. Census data are crucial for generating intercensal estimates that provide a more continuous picture of population dynamics. However, traditional census methods have limitations. The 10-year periodicity can leave significant gaps in data for users who require more frequent updates. Additionally, the accuracy of estimates can deteriorate over time, especially in areas with rapid population shifts.

1.50. While these limitations exist, efforts should be made to make census data more interoperable by design. This means designing and collecting census data with the specific goal of integrating them smoothly with data from other sources. Some key considerations for achieving this include:

i) Thoroughly documenting the coverage and data collection methodologies used in the census makes this information readily available to data users. This transparency fosters trust and facilitates more accurate interpretation of the data.

ii) Ensuring that variables collected in both the census and other surveys use consistent concepts, definitions, and classifications is crucial. This allows for seamless comparisons and avoids misinterpretations when combining data from different sources.

iii) Assigning geographic codes (geocoding) to census data allows for spatial analysis and integration with other geospatial datasets. This enables users to visualize population trends across geographic areas and identify local variations.

iv) Collecting in the census information on unique ID numbers of people for countries that utilize them facilitates more effective linkage of data from different sources, provided that adequate measures are implemented to safeguard privacy and confidentiality.

1.51. Implementing these strategies can help unlock the full potential of the population and housing census within integrated population data systems. A more interoperable census becomes a stronger foundation, allowing countries to layer and analyse data from diverse sources to create a richer and more dynamic understanding of population trends.

B. Uses of housing censuses

1. Uses for development of benchmark housing statistics

1.52. The housing census produces benchmark statistics on the current housing situation and is vital for developing national housing and human settlements programmes and policies. The housing census is also valuable for providing the sampling frame for special housing and related surveys during the intercensal years.

1.53. Housing benchmark statistics are also critical for emergency planning for response to natural hazards (such as destructive storms, earthquakes, tsunami and fires), or post-conflict situations. Following such situations, these statistics can be used to estimate the numbers of people and structures affected, the need for emergency response and reconstruction requirements.

1.54. NSOs would need to develop, from housing censuses, the sort of benchmark statistics in housing that could be supplemented by current building and construction statistics that would provide continuous

up-to-date information of the housing position needed for the consideration of housing programmes. Since not all the basic information required to assess housing needs or to formulate housing programmes can be obtained through a housing census alone, additional requisite data should be obtained through the population census, special housing surveys and environmental surveys, and from vital statistics, economic statistics and so forth; but data obtained from the housing census will constitute the basic framework within which the estimates are made, indices computed and further statistical enquiries planned.

1.55. When population and housing censuses are carried out as a single operation or independently but in a well-coordinated fashion, the combined information provided is of much higher value than when conducted in isolation, since the essential features of both censuses are interrelated. The information on housing may be analysed in association with the spatial, demographic and socioeconomic conditions of the occupants and *vice versa*.

2. Uses for the formulation of housing policy and programmes

1.56. Housing census data play a critical role in formulating housing policies and programs. Objective data on the housing situation provide a foundation for policymakers to make informed decisions, considering social, economic, and political factors.

1.57. Additionally, census data help to identify the housing needs of various population groups by providing key information such as household size, housing conditions (e.g., overcrowding, access to basic amenities), and tenure status (e.g., owner-occupied, renter-occupied). Understanding these housing characteristics for different demographic groups allows policymakers to design targeted housing programmes at local and national levels. Furthermore, census data provide a baseline to monitor the impact of these programmes over time, enabling continuous improvement with geographically relevant and comparable data.

1.58. In most countries, housing programmes encompass both governmental and private activity. The data derived from a housing census are used by governmental authorities for making an analysis or diagnosis of the housing situation. Housing conditions are analysed in quantitative and qualitative terms and data from previous censuses are used to indicate the changes in the housing situation that have occurred during the intercensal periods; the housing stock and future housing requirements are estimated and compared with the rates of dwelling production being attained; and the characteristics of the households in need of housing are considered in relation to the availability and cost of housing. As part of overall development plans, such an analysis is necessary for the formulation of national housing programmes and for their execution.

1.59. Commercial users also study housing census data. Those engaged in the construction industry, financing institutions, and manufacturers of housing fixtures and equipment and household appliances use housing census data, in combination with population census data, to assess the potential demand for these goods and services.

3. Assessment of the quality of housing

1.60. The materials used for the construction of housing units (roof, walls, floors) are a significant indication of the quality of life experienced by the occupants and, viewed in aggregate across multiple dwellings in an area and, alongside other information such as weather conditions and date of

construction, can provide an indicator of disaster risk. Trends indicated by census data on the type of housing materials used can show improvements in the welfare of the population, such as where the percentage of poor-quality or slum-like housing facilities has decreased, or where there are differences in housing conditions between female- and male-headed households. Additionally, measures of water, sanitation and hygiene, as well as energy usage and telecommunication data are valuable for the planning and development of future infrastructure needs.

C. Relationships of population and housing censuses with other statistical activities

1. Relationship between the population census and the housing census

1.61. There is an especially close association between population censuses and housing censuses. The two censuses may constitute one statistical operation or they may be two separate but well-coordinated activities, but in either case they should never be considered completely independently of each other because essential elements of each census are common to both. For example, an essential feature of a population census is determining the location of each occupied set of living quarters and of the persons living therein, while an essential feature of a housing census is the collection of information on the characteristics of each set of living quarters in association with the number and characteristics of its occupants.

1.62. In many countries, the population and housing censuses are taken concurrently, often with the use of a single schedule. In this way, the information on population and living quarters can be more readily matched and processed. This also makes it possible to relate the information on demographic and economic characteristics of each household member that is routinely collected in the population census to the housing census data, thereby facilitating a wide range of multivariate analyses.

1.63. In censuses that are conducted with a field enumeration, the advantages of simultaneous investigation may be offset to some extent by the additional burden imposed on the respondent and the enumerator resulting from the increased amount of information that must be collected at one time. In countries where this is likely to be a serious problem, consideration might be given to collecting data for a limited number of topics based on a complete enumeration in the population and housing census, but with more complex data being collected on a sample basis only, either concurrently with or immediately following the full enumeration. Alternatively, consideration might be given to carrying out the housing census as part of the advance household listing operations of the population census.

1.64. The relationship between the population census and the housing census will affect the means by which data on homeless persons are obtained. In the case of simultaneous censuses of population and housing, data on homeless persons will be obtained as part of the population census. Where the housing census is carried out independently of the population census, it may be necessary to assign resources to enumerate homeless persons in the housing census also. Such information may reflect, among other things, the magnitude of the housing problem in a given locality.

1.65. Beyond the operational side, population and housing censuses are also fundamental statistical exercises that offer invaluable insights into a country's demographics, living conditions, and socioeconomic status. These intrinsically linked operations inform each other, painting a more detailed

picture of the social, infrastructural, and economic landscape. By exploring these synergies, policymakers can make informed decisions that address the multifaceted needs of the population.

1.66. For instance, high population density revealed by the population census can be combined with housing data to understand the availability and quality of housing in that area. This combined insight allows for targeted infrastructure development, such as building affordable housing or improving public transportation, to better serve the local population.

1.67. Additionally, spatial analysis of combined population and housing data helps identify patterns, hotspots, and disparities in circumstances such as housing affordability, access to basic services, and urban-rural divides. This information guides resource allocation and targeted policies to address inequalities and improve overall living standards.

1.68. Joint assessments of population distribution, housing characteristics, and homelessness inform disaster risk management and emergency response planning. These combined data identify vulnerable communities, allocates resources for mitigation, and shapes resilient housing strategies that protect lives and livelihoods.

2. Relationship of population and housing censuses to intercensal sample surveys

1.69. The rapidity of current changes in the size and other characteristics of populations, and the demand for additional detailed data on social and economic characteristics of population and housing characteristics that are not appropriate for collection in a full-scale census, have brought about the need for continuing programmes of intercensal household sampling surveys to collect current and detailed information on many topics.[23]

1.70. The population and housing census can provide the frame for scientific sample design in connection with such surveys (see paragraphs 1.31. -0); at the same time, it provides benchmark data for evaluating the validity of the overall survey results as well as a base against which changes in the characteristics investigated in both enquiries can be measured. To allow for the comparison of census and survey results, the definitions and classifications employed should be either common or harmonized, while remaining consistent with the aims of each investigation. Because of the relative permanence of living quarters, the lists available from the housing census (with suitable updating) may also provide a convenient frame for carrying out enquiries dealing with topics other than population and housing.

3. Relationship of population and/or housing censuses to other types of censuses and other statistical investigations

(i) Census of agriculture

[23] *Designing Household Survey Samples: Practical Guidelines No. 98* (United Nations publication, Sales No. E.06.XVII.13).

1.71. While the population and housing censuses have a close relationship, their relationship with the agricultural census is often limited to the coordination of data collection in households involved in agricultural production activities. As the result of increasing integration within programmes of data collection, the relationship between the population and housing census and the agricultural census is now far closer than in the past and countries are increasingly looking at new ways to strengthen this relationship. This relationship is particularly relevant in many developing countries where most agricultural production activities are in the household sector.

1.72. One conceptual issue in relating the two censuses is that they use different units of enumeration. The unit of enumeration in the agricultural census is the agricultural holding,[24] which is the economic unit of agricultural production, while the units of enumeration in the population census are the household and the individuals within the household. The unit common to the two censuses is therefore the household engaged in own-account agricultural production activities.[25] In many developing countries, there is usually a one-to-one relationship between households and agricultural holdings. In these cases, the same unit is enumerated in both types of censuses.

1.73. Another conceptual issue is the reference periods used by the two censuses. The population census normally collects information about a person's main job or work activity during a short reference period (usually one week prior to the census), while the census of agriculture collects information on work on agriculture during the year preceding the census because of the seasonality of many agricultural activities. In consequence, users may find some agricultural activity data from the agricultural census more comprehensive than from the population census because the latter may not identify persons connected with agricultural activity on a seasonal basis or as a secondary activity. On the other hand, the population census provides data on all persons working in agriculture, including as paid employees. The latter information (paid employees) is outside the scope of the agricultural census as its focus is on households with own-account agricultural production. To get a complete picture, agricultural data users will need both agricultural census data and population census data.

1.74. In planning the population and housing census, every opportunity for developing the relationship between this census and the agricultural census should be explored. This can take several forms. Definitions used in the population and housing censuses should be compatible with those used in the agricultural census so that meaningful comparisons can be made between the two data sets. The population and housing census can also be of use in the preparation of the agricultural census, such as in the demarcation of EAs, or the preparation of the frame for the agricultural census. One proviso, however, is that there must be a legislative framework in place to allow the NSO to use confidential unit record information collected in the population census for the purposes of the creating such a frame, particularly if the agricultural census is being carried out by an agency other than the NSO itself.

1.75. In planning the national census programme, serious consideration should be given to the possibility of collecting key agricultural information as part of the population and housing census exercise that would

[24] An agricultural holding is an economic unit of agricultural production under single management comprising all livestock kept and all land used wholly or partly for agricultural production purposes, without regard to title, legal form or size.

[25] That is, households with members engaged in agricultural production activities for own final use.

facilitate the preparation of an adequate frame of agricultural holdings in the household sector for a subsequent agricultural census. This integration can be achieved through pre-census cartographic work and/or listing exercise, or by adding two non-core agricultural topics to the census questionnaire (as detailed in paragraphs **Error! Reference source not found.**–5.456). Regardless of the chosen method (cartographic/listing or questionnaire), countries aiming to establish a frame for the agricultural census through the population and housing census should ensure synchronization. Conducting the agricultural census as soon as possible after the population and housing census leverages the most up-to-date and relevant frame data. This approach is detailed in the World Programme for the Census of Agriculture 2030[26] and *Guidelines for Linking Population and Housing Censuses with Agricultural Censuses*.[27]

1.76. Linking population and agricultural census data in countries where most agricultural production activities are in the household sector can bring many benefits. This could add considerable analytical value to data sets from both censuses and save on data collection costs. Many of the demographic and activity status data collected in the population census are also collected in the agricultural census. If data from the two censuses could be linked, it would no longer be necessary to collect these data again in the agricultural census, while still allowing for comprehensive cross-tabulations.

1.77. A few countries conduct the data collection for the population and agricultural censuses as a joint field operation. This is feasible only in countries with extensive methodological and field experience. Normally, each census retains its separate identity and uses its own questionnaire, but field operations are synchronized so that the two data collections can be done at the same time by the same enumerators. Occasionally, the two censuses are merged into one, particularly in countries where fieldwork costs are high and logistical challenges make conducting the two censuses separately more costly. This may have a number of advantages; however, as this is an increasingly complex operation, its impact on field operations and data quality needs to be carefully considered. Errors or shortcomings in the design and execution would impact equally on both censuses.

(ii) Census of establishments

1.78. While population censuses primarily focus on individuals and households, they may collect basic information about individuals' employment, such as their occupation and industry of employment. This information can be used to generate preliminary lists of businesses and other economic units, though these lists would lack the detail of a dedicated economic census or business register. However, this information can be valuable for various purposes, such as estimating the number and types of establishments in an area. This can help with regional planning and economic development initiatives. It can also be useful for identifying potential businesses for inclusion in future economic censuses or surveys. This can improve the coverage and efficiency of those data collection efforts. Furthermore, it can also supplement existing business registers. Census data can help identify new businesses or update information on existing ones. To maximize the utility of this information, it is important that the data is

[26] Food and Agriculture Organization of the United Nations, World Programme for the Census of Agriculture 2030, Statistical Development Series No. 20 (Rome, 2026).

[27] *Guidelines for Linking Population and Housing Censuses with Agricultural Censuses with Selected Country Practices*, Food and Agriculture Organization of the United Nations and the United Nations Population Fund, Rome, 2012.

geocoded or geospatially enabled (coordinates and/or addresses) to allow integration with other databases, including administrative records.

1.79. Many business registers cover only establishments with fixed visible premises in which more than some minimum of persons (usually 5 or 10) are employed. In these cases, the population census can be used to collect basic information (such as kind of activity and size) on those establishments with employment below this minimum number of persons (see paragraph 1.81.). However, special care should be taken in the choice of the unit of enumeration to ensure that there is no double counting of establishments.

1.80. When the information from the population census is to be used to construct a list-based sample frame, it is essential that the information from the population census be made available and used shortly after the enumeration is carried out because this information can quickly become outdated. This requirement is less imperative when the information is to be used to construct an area-based sample frame. Moreover, the legal authority of the NSO to use population census information for such purposes (equivalent to that noted in paragraph 1.74. above) should be established.

1.81. The population census information needed for these purposes is the status in employment, in order to identify employers and independent workers. For this subset of workers, information needs to be collected on the number of establishments operated, and for each of these, the kind of economic activity, the name and address of the establishment (if any), the number of workers engaged (including contributing family workers and employees) and whether the establishment is operated in partnership with other persons. If all of this information appears in the census questionnaire (or can be derived from other information collected), the number of small establishments can be extracted from the schedule or from the processing documents after the enumeration.

(iii) Census of buildings

1.82. In certain circumstances, it may be necessary, as part of the housing census operations, to enquire whether or not buildings (both residential and non-residential) are occupied. Thus, it may be convenient to record basic information for all buildings at the time of the housing census, even though detailed data may be collected only for those in which housing units or other sets of living quarters are located. The comprehensive list thus obtained sometimes provides the basis for a census of buildings, carried out concurrently with, or subsequent to, the housing census, or it may provide for the identification of special types of buildings significant for other enquiries, such as the census of establishments or the census of schools. In these cases, it is important that the information is geocoded or geospatially enabled (coordinates and/or addresses) to allow spatial analysis and integration with geospatial information and other databases, including administrative records. If a listing of households is to be carried out before the actual enumeration, this would be most ideal for collecting basic information about buildings. Building/dwelling registers with comprehensive coverage, when available, can be used to generate housing census data directly, reducing the need for additional data collection efforts. Even if not used directly, such registers can nevertheless serve as a valuable baseline for a housing census data collection. By using existing register data, NSOs can identify potential housing units, reducing the need for an extensive field enumeration.

(iv) System of current housing statistics

1.83. Current housing statistics refer to the latest data and trends in the housing market. They reflect the number of dwellings constructed and certain related information such as value, number of rooms, floor space, and so forth, as well as the number of dwellings destroyed or demolished. These data are usually obtained from a system of data collection based on the administrative procedures required in connection with activities in the housing market. For example, construction statistics may be derived from permits issued for the construction of dwellings, from records of dwelling starts or completions, or from certificates of occupancy. Statistics on dwellings destroyed may be obtained from the records maintained for the levying of rates and the collection of taxes. Compiled monthly or quarterly, current housing statistics reflect changes in the housing inventory and, although they may serve other purposes, they are also used to update the benchmark data obtained from housing censuses.

4. Relationship between the population census and/or housing census and administrative data sources

(i) Civil registration and vital statistics

1.84. Population censuses provide a fundamental baseline for the Civil Registration and Vital Statistics (CRVS) systems. By comparing census records with civil registration records, gaps in registration can be identified. In addition, census data can:

a. help identify population groups facing increased vulnerabilities such as refugees, internally displaced persons, stateless persons, irregular and undocumented migrants, or marginalized communities, who may face challenges in accessing civil registration services;
b. be used to estimate mortality rates (e.g., age-specific mortality rates) and fertility rates, which are essential components of vital statistics;
c. be used as a benchmark for evaluating the performance of CRVS systems.

1.85. The United Nations Statistics Division (UNSD) in its *Handbook on Civil Registration and Vital statistics Systems*[28] defines civil registration as "the continuous, permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population, as provided through decree or regulation in accordance with the legal requirements in each country". A well-functioning CRVS system helps ensure that every person has a legal identity and all vital events are registered, facilitating access to public benefits, social protections, and human rights. It also produces vital statistics (statistics on births, deaths, marriages and divorces) and serves as an optimal source of data on the prevalence and distribution of vital events for evidence-based decision policy, planning and programme implementation.

1.86. Additionally, the 2030 Agenda for Sustainable Development highlights the importance of CRVS system, as numerous targets and indicators rely on data from this system. Complete, reliable, and timely data from CRVS systems can provide cause-specific mortality rates to monitor many targets and indicators

[28] Handbook on Civil Registration and Vital Statistics Systems: Management, Operation and Maintenance, Revision 1. Sales No.: E.18.XVII.18. United Nations. New York, 2021.

listed in the SDGs. Furthermore, a complete CRVS system facilitates the calculation of birth rates, age-specific fertility rates, and general mortality rates.

1.87. Population census data serve as denominators for the computation of vital rates, especially rates specific for characteristics normally investigated only at the time of the census. Conversely, census results, time-adjusted by vital and migration statistics, can provide estimates of the future size, distribution and other characteristics of the population of the total country and subnational areas. Furthermore, census data on fertility can provide a benchmark check on the reliability of current birth statistics, and *vice versa*. It is consequently desirable that procedures for the collection of population census data, vital statistics and migration statistics be closely coordinated with regard to coverage, concepts, definitions, classifications and tabulations. It is important to stress the need to use the same definitions in vital statistics and in censuses. For example, for the place of birth, the same definition should be used in both sources (such as the place of residence of the mother at the time of birth or the place of registered birth – see paragraph 5.64).

1.88. Linking census and vital statistics data offers various possibilities to improve analysis. One of the benefits of linking geocoded census data with vital statistics is the ability to georeference births, deaths and other vital events, enabling valuable analysis of these phenomena. Some countries have linked individual census records for infants less than 1 year of age with birth registration reports for the year preceding the census date as a means of checking on the completeness of one or the other type of investigation. Linkage of death reports with census records has been used to compare the information on characteristics of the deceased as reported in the two sources.[29] While the many problems posed in the past by the one-to-one matching of two types of records have not been entirely solved, their severity has been mitigated by developments in computer technology. Before undertaking either of the procedures, however, countries should consider carefully the possible advantages of using household sample survey returns rather than census returns in the operation. Moreover, such operations have to be carried out in complete accord with national laws and policies governing the confidentiality of information obtained in the census if public confidence in the census is to be maintained.

1.89. In the establishment of a civil registration system, census results on the geographic distribution of the population can be useful for determining appropriate locations for registration offices.

(ii) Population register

1.90. The term "population register" was originally defined in 1969, in the publication entitled *Methodology and Evaluation of Population Registers and Similar Systems* (United Nations, 1969), as "an individualized data system, that is, a mechanism of continuous recording, and/or of coordinated linkage, of selected information pertaining to each member of the resident population of a country in such a way to provide the possibility of determining up-to-date information concerning the size and characteristics of that population at selected time intervals" (chap. I.A). According to the UN's *Handbook on Registers-Based Population and Housing Censuses* (2022), a population register is a system that "records all

[29] An elaboration of comparison between census and vital statistics is provided in *Principles and Recommendations for a Vital Statistics System, Revision 3*, United Nations publication, Sales Number E.13.XVII.10, United Nations, New York, 2013 (see paragraphs 595–597).

residents in the country, together with their basic demographic information, and updates for births, deaths, changes of marital status, international migration (both into and out of the country), and internal migration (by recording changes of address within the country)."

1.91. Thus, the population register is the product of a continuous process, in which notifications of certain events, which may have been recorded originally in different administrative systems, are automatically linked to it on a current basis. The method and sources of updating should cover all changes so that the characteristics of individuals in the register remain current. Because of the nature of a population register, its organization, as well as its operation, should have a legal basis.

1.92. Basic characteristics that may be included in a population register are date and place of birth, sex, date and place of death, date of arrival/departure, citizenship(s), marital status, and place of residence. Depending on the on the potential to link with data held in other registers, much additional information may be added to the single record, such as language(s), ethnicity, educational attainment, parity, activity status and occupation. In order to be useful, any additional information must be kept up to date. If complete and up-to-date, population registers can produce data on both internal and international migration through the recording of changes of residence as well as the recording of international arrivals and departures.

1.93. A key requirement is that each unit in the population register should be always uniquely identifiable. This is best achieved by using a system of identification codes (keys). Identification is, however, also possible without such a code being used, if sufficient information is available on each unit (for persons this could include name, sex, address and date of birth).

1.94. The population census and population register are essential data sources for understanding and managing a population. While serving distinct purposes, they can be mutually supportive. On the one hand, a well-maintained population register can improve the efficiency and accuracy of future censuses by providing a list of individuals to be enumerated, reducing the need for field operations and minimizing non-response rates. On the other hand, periodic censuses can validate the accuracy of the population register by identifying and correcting discrepancies, ensuring the register's reliability. Together, these data sources provide a comprehensive understanding of a population, with the census offering a static snapshot and the register tracking changes over time. By working together, they enable governments and organizations to make informed decisions and effectively manage their populations.

(iii) Building/dwelling or address register

1.95. Building and dwelling registers are primarily used for property management, taxation, and urban planning, containing detailed information about individual buildings or dwellings, such as construction type, size, occupancy status, ownership information, and valuation. In some cases, building and dwelling registers may include address information. In contrast, an address register is an up-to-date comprehensive list of all known physical addresses within a country. It is typically used for postal services, emergency response, and location-based services. It can serve as key piece of statistical infrastructure for a census, it should preferably contain essential information about individual addresses such as address text (including street name, house number and postal code; and purpose/use of the address) and geographic coordinates such as latitude and longitude (x- and y-coordinates). It could also include information on dwelling structure for residential addresses and non-private dwelling information.

1.96. An address register can enable the running of a census self-enumeration with Internet or paper questionnaire by providing a mail out frame which is a dataset of residential and non-private dwellings (such as hotels and caravan parks) and is used to support the mail out of census materials. The latitude and longitude and dwelling structure is particularly useful for helping to locate dwellings in remote areas.

1.97. An address register can also be used to facilitate a register-based or combined census (see paragraphs 1.170. -1.176. and paragraphs 1.241. -1.250. respectively) by using common variables (such as address text) to link to other government data sources in order to add further value to the current information, for example, the construction period or availability of electricity. Utilities data can also be linked to an address register to help determine whether a residential address is being regularly occupied or is only used at certain times during the year.

1.98. When data for certain census variables (particularly in a register-based or combined census) are obtained as inputs from administrative registers, it is important to ensure that the descriptions and concepts of these variables are harmonized to ensure consistency with required census outputs. For instance, non-private dwellings and dwelling structure both appear as an input from an address register but also as an output in a census. In order to ensure the outputs are accurate the definitions of the different non-private dwellings and dwelling structures need to be consistent across the address register and the census.

(iv) Business register

1.99. Statistical Business Registers (SBRs) provide the infrastructure to support the collection of high quality, reliable and consistent economic data and the production of economic statistics. SBRs are maintained by NSOs and used for statistical purposes.

1.100. To ensure effective and complete coverage, SBRs must be regularly maintained and updated. Administrative data sources such as legal or taxation business registers including unique identifiers are used to update less significant units with simpler structures. Large, complex and significant businesses are updated through profiling activity to transform these businesses from administrative data to statistical units. Other sources of updates may include economic censuses, survey feedback, big data, or web scraping (a process of using software to extract data from websites).

1.101. SBRs provide a coherent set of statistical units and classifications to collect and assemble data. SBRs include units represented in the economic units model, which describes the characteristics of businesses and the structural relationships between related businesses.

1.102. SBRs are used to create regular point-in-time frames (frozen or common frames) from which survey or sampling frames (subsets) for most economic surveys are drawn, ultimately feeding into the System of National Accounts. SBRs are also used to analyse and publish business demography statistics. Business demography includes events such as creation and cessation of units, survival information, and changes in business characteristics over time including type of activity, size measures (employment, turnover), location, institutional sector, employing versus non-employing units.

1.103. Business registers are valuable for population and housing censuses. They can provide a foundation for sampling, enhance the collection of economic data, and contribute to a more accurate and comprehensive picture of the population. Business registers can contribute to censuses in the following ways:

i) **Enhanced economic data collection**: Information within business registers, such as industry classification codes, can be directly linked to corresponding economic activity questions in the census. This allows for efficient coding and analysis of economic characteristics like industry of employment or type of business.
ii) **Accurate place of work data**: Business register data, including business locations, can be used to improve the accuracy of "place of work" information collected during the census. This fosters a better understanding of commuting patterns and workforce distribution.
iii) **Data integration and analysis**: Combining business register and census data can facilitate the analysis of intersections between business and issues of current interest, such as gender equality.

(v) Other administrative data sets

1.104. Other administrative data may include, for example, records of foreign citizens and foreign workers, people living in institutional places, enrolment and graduation records, social insurance, employment records, etc. The availability and use of these sources is very country specific.

1.105. There is an increasing availability of a range of administrative data sources (held either by governmental agencies or private sector bodies) containing information on the persons or the households within a country. The utility of these data sets for statistical analysis may be limited, in some countries, by their lack of population coverage, data accuracy or range of characteristics, especially in the case of newly-developed sources, and those designed for purposes in which completeness and accuracy are not key quality criteria. The linkage of these data sets with the census file, with its complete coverage of persons and households, can provide the ability to create new insights and new statistical products to leverage more value from the census.

1.106. Administrative data may replace (either wholly or partially) data obtained by field enumeration – for example, in some countries income data from the taxation or revenue department can replace the need to directly collect those data in the census. Administrative data can also extend census data – for example, census data can be linked with visa information or health information to expand the census data set into areas that may be too sensitive to collect on the census form, or with past education data to analyse longitudinally the impact of education on labour force outcomes. Moreover, administrative data can be used as a tool to verify the data collected through a field-based census; it can also replace missing data – for example, health records have been used in certain countries to estimate the count and characteristics of usual residents who did not respond during the census enumeration period (where a legal authority to do so has been established).

1.107. In the 2020 round of population and housing censuses, countries have successfully utilized data from public utilities (e.g., water, electricity) to complement or verify certain information collected through field operations. One example is the use of data from electricity departments. The unique electricity meter number assigned to each dwelling is linked to a unique identification number of the tenant. This linkage enables the extraction of information such as dwelling type, tenant nationality, and consumer type (household or business). This data can be helpful in determining dwelling occupancy type. Additionally, the amount of electricity consumed over a specific period can serve as a proxy to determine whether a housing unit is vacant or occupied and to estimate the number of persons living in the dwelling.

1.108. As described above in paragraph 1.88. , linkage operations should be undertaken with caution, ensuring not only that all national laws are met but also that the trust of the public in the census and the

statistical systems is maintained. See Chapter IV, Section B.1 for elaboration on necessary pre-conditions for the use data from administrative registers in censuses.

D. Integration of census data with geospatial information and other data sources

1. Integration of census data with other data sources

1.109. Data integration is the combination of technical and business processes used to combine data from disparate sources, either at the unit level (for example, persons, dwellings, businesses) or at the aggregated level. Data sources may include censuses, sample surveys, administrative data, or new data sources including geospatial information or big data. It should be noted that if some data sources are not controlled by the NSO, legislative authority maybe needed to facilitate data integration. This ensures that the census organization has the necessary access and permissions to combine data from various sources.

1.110. Data integration has the potential to improve the cost-effectiveness and relevance compared to single data collection activity. It also has the potential to enhance the coverage, granularity, timeliness and richness of available data, allowing more complex questions to be analysed and more insights to be drawn than would be possible from a single source. Many integrated data assets are also longitudinal in nature, meaning they allow changes and patterns in population, economy and environment to be better understood over time. Integrating data can also serve to reduce respondent burden by reducing the number of questions that need to be asked in censuses or surveys, as shown in some countries when quality administrative data sources are used with proper consent from respondents. Censuses are often used as a spine[30] for linking population data from multiple sources for statistical purposes.

1.111. To fully realise the potential of data integration, institutional and technical capacities must be in place. Institutional capacity mainly refers to:

(a) the availability and accessibility of relevant data sources;
(b) the existence of a strong legal basis supporting the use of and access to unit record data and metadata for statistical purposes and legislation ensuring privacy and confidentiality;
(c) appropriate mechanisms for collaboration with data holders (either from the public or private sector) and;
(d) human resources and ICT Infrastructure.

1.112. At the technical level, expertise and skill are required for:

(a) dealing with the lack of interoperability, in terms of lack of unique identifiers, differences in concepts, classifications, coverage, data formats, reference periods, etc.;
(b) handling data quality issues, such as missing data, erroneous values, and inconsistencies within and across data sources; as well as
(c) linking records across data sources.

[30] A **population spine** is a concept related to a reference population that is used to link together and manage different data sets about individuals. It is often a term used in relation to register-based censuses, which makes it easier to track changes in the population over time.

1.113. Facilitation of the integration of census with other data sources requires careful planning and designing of the census data collection, so that it is "interoperable" by design, while also addressing privacy and confidentiality concerns. This may include using standard concepts, definitions, and classification as well as collecting unique identifiers, such as national IDs, in the census. Georeferencing censuses is another important element to ensure that census data can be integrated with other sources through geospatial information.

1.114. Data integration at unit level is made possible with the use of key linking variables. In the case of people, this may be unique personal identification numbers (PIN), such as national IDs, social security numbers, etc. or a combination of identifying variables such as name, sex, date of birth, addresses, etc., if unique identifiers are not available across the data sources.

1.115. Data integration at the aggregated level required is carried out through modelling, for two different sources borrowing strength from each other. One example is the small area estimation methods that have been traditionally used for poverty mapping for lower geographical areas but have now been extended to cover a wide variety of outcome indicators including employment and social welfare. The method is also being used by countries to produce estimates for population groups facing increased vulnerabilities such as persons with disabilities.

1.116. For countries that do not have access to good administrative data, population censuses have been an important source to be integrated with household surveys, for two reasons. Firstly, population censuses are more comprehensive in coverage compared to household surveys. Secondly, population censuses collect a number of variables that can be used as auxiliary variables for the data integration. On the other hand, household surveys collect a lot of more in-depth information though with limitations on coverage. Combining the two sources, through methods such as small area estimation, can produce data that are both more granular and more timely.

1.117. Successful integration of census data with other data sources hinges on careful planning and design that prioritizes interoperability. This might involve using consistent concepts and definitions for key socio-demographic variables across all data sources. Additionally, georeferencing censuses plays a crucial role in enabling the integration of census data with other data sources. This linkage, ideally done through consistent, unique personal identifiers (PIDs), can also allow the dataset to use various additional geospatial information within the address register, geocode database and/or coordinates. In the absence of such a system, organisations may rely on other ways to reference location (such as geographical names), alternative sources (such as an address register from a utility provider) or higher-level geography (such as large geographical areas).

2. Integration of census data with geospatial information

1.118. Geospatial information (such as an address, x- and y-coordinates, geographical name, or unique identifiers related to geospatial features) can play an important role in bringing together information from various domains by enabling the integration of datasets from different sources using the common location information as a matching and universal key variable (for example, integrating administrative data with survey data using address or postal code that exists in both datasets).

1.119. Geospatially-enabled statistical information can provide critical knowledge by the integration with other data from a wide variety of sources to better understand multi-faceted issues that the society

currently faces, such as sustainable development, urbanization, and climate change. To support these efforts, the Global Statistical Geospatial Framework (GSGF) [31] details some of the very basic conceptual phases needed to create a flexible, interoperable national geospatial structure, fundamentally integrated with the statistical data collected during enumeration.

1.120. The GSGF enables a range of data to be integrated from both statistical and geospatial organizations and, through the application of its five principles (see paragraph 1.121.) and supporting key elements, permits the production of harmonised and standardised geospatially enabled statistical data. The resulting data can then be integrated with statistical, geospatial, and other information to inform and facilitate data-driven and evidence-based decision making to support local, sub-national, national, regional, and global development priorities and agendas, such as the census.

1.121. The five principles of the GSGF outline the broad processes by which a range of geospatial and statistical infrastructures and processes are applied to input data to enable integration. Firstly, the statistical data are geospatially enabled to the finest level possible. Then, geospatial tools and methods, such as common geographies and common standards of good practice, are used to ensure that the data are interoperable, accessible, and usable. The five principles are:

1. Use of fundamental geospatial infrastructure and geocoding;
2. Geocoded unit record data in a data management environment;
3. Common geographies for the dissemination of statistics;
4. Statistical and geospatial interoperability; and
5. Accessible and usable geospatially enabled statistics.

1.122. In addition to these five principles, four key elements play a fundamental role in enabling data to be obtained from different sources and the GSGF Principles to be applied. These elements comprise:

a. Standards and good practice;
b. National laws and policies;
c. Technical infrastructure; and
d. Institutional collaboration.

1.123. To ensure the quality of integration, standardising the geospatial information in the different datasets is critical. This standardisation should normally take place before the integration of datasets and can be done through, for example, matching location information in the datasets with a centralised standard system (such as address matching, coordinates, geocoding) which should be part of the National Spatial Data Infrastructure (NSDI).

1.124. The NSDI is a framework of policies, institutional arrangements, technologies, data, and people that enables the sharing and effective usage of geographic information by standardising formats and protocols for access and interoperability. The development of a NSDI is a multi-agency and multi-partner activity. All governmental authorities, private sector and public users can view NSDI data set and map layers based on user rules and permission to download, print, geoprocessing services and dynamic queries.

[31] *The Global Statistical Geospatial Framework*, United Nations Statistical Division, New York, 2019.

1.125. The benefits of using NSD are that it:

- Provides a unified standard base map with different scales;
- Reduces duplication of efforts among governmental authorities;
- Increases the benefits of using available spatial data;
- Lowers costs related to geographic information while making geographic data more accessible;
- Provides legal arrangements and governance; and
- Establishes key partnerships between Government authorities, states, cities, academic, and the private sector.

1.126. The advantages of NSDI are:

- Gains in time for searching and accessing the necessary spatial data;
- Access to the latest data and complete metadata;
- Increased flexibility;
- Contributes to openness and transparency;
- Continuously improved and updated; and
- Provides NSDI web mapping governmental portal.

1.127. The population and housing census provides an opportunity to bring the NSO and the national geospatial information agency (NGIA)[32] closer together in mapping initiatives and information/data exchange. This approach can optimize the use of resources both in the census operation and in the production of national mapping, also involving the contracting of services (such as the provision of satellite images, location collection equipment, software, etc.). It can be an input for the NSDI and for the NSO to integrate efforts to manage national geospatial information according to the Integrated Geospatial Information Framework (UN-IGIF).

1.128. The use of fundamental geospatial infrastructure and geocoding[33] should be used in censuses to obtain standardized, high-quality location references and cover precise coordinates or a small geographic area or standard grid reference for each population or household unit. Preferably, the location should be recorded through direct or indirect capture of x and y coordinates and/or address point. When this level of accuracy is not possible using a country's current geospatial and statistical infrastructure, adaptations based on more general locations or larger geographies will be necessary.

[32] National Geospatial Information Agency (NGIA) is used as an encompassing term to cover National Mapping, National Cartographic, National Geospatial Information Agencies and Authorities.

[33] Geocoding is generally defined as the process of geospatially enabling statistical unit records so that they can be used in geospatial analysis. More specifically, geocoding is the process of linking unreferenced location information (e.g. an address), that is associated with a statistical unit, to a geocode (i.e. a geospatially referenced object). Alternatively, the geocode can be directly incorporated into the statistical unit record. Geocodes are, preferably, fine scale geospatially referenced objects that are stored as a geometry data type, such as location coordinates (i.e. x, y, z coordinates), and/or small area geographies (e.g. mesh blocks, block faces or similar small building block geographies). Larger geographic units, such as enumeration geographies, can be used as geocodes where finer scale geospatial units are not available. The linkage of a geocode to a statistical unit record can occur through use of standard geographic coding systems, a Uniform Resource Identifier (URI) or through other computer-based linkage mechanisms.

1.129. The census should produce geospatially enabled data on the number and characteristics of the population and housing units that allow disaggregation to the lowest appropriate geographic level, compatible with national circumstances, and for small population groups [34] while protecting the confidentiality of personal information of each individual. This will also allow the definition of user-defined flexible areas or user-centred geographies as long as confidentiality issues are observed. Not all census methodologies satisfy this criterion.

1.130. The resulting geospatially enabled statistics and integrated data will support informed and evidence-based decisions, development plans at all levels within a country and provide valuable information on the underpinning pillars of a country's society, economy, and environment, as well as informing regional and global action.

IV. Census methodology

1.131. There are three basic approaches to conducting a census based on different methods of data collection; these are: the *full-field enumeration* in which data is collected directly from the public using census questionnaires; the use of data sources solely based on administrative records (commonly referred to as a *register-based census*); or combining data derived from administrative records with data collected from the field (the so-called *combined census* approach).

1.132. Taking account of the experiences of the previous rounds of population and housing censuses, it has become evident that an increasing number of countries have been exploring the use of alternative methodologies to the traditional field-enumeration based approach for collecting census information. The use of registers – primarily population registers – in combination with other sources is being considered in a number of countries for the purpose of producing detailed small-area statistics on population and housing, as well as the application of continuous survey methodology for the same purpose. Furthermore, these alternatives to the traditional method of conducting population and housing censuses are becoming more diverse in terms of developing combinations of various data collection methods, and it is thus a challenge to summarize and categorize them using generally accepted data source methodologies. Many countries have developed combined systems, making use of the information available in the registers or from administrative systems to complement information collected through field operations (in order, for example, to ensure the exhaustivity of the count) or, *vice versa*, making use of survey data, in order to complement information available in registers. Despite this complexity, a three-folded classification (field enumeration approach, register-based approaches, combined approaches) is still a useful basic classification to describe the main features of the different approaches and thus provide recommendations according to the main issues and challenges.

1.133. It should be noted that many countries are expected to continue using the full-field enumeration census approach – collecting information from each household – in the 2030 round of censuses,[35] while

[34] One example is the application 'Build a custom area profile' from the United Kingdom Office for National Statistics. The application a simple tool that allows the user to create a Census 2021 data profile covering any area in England and Wales, including one that you draw for yourself on a map.

[35] The 2030 round of censuses covers the decade 2025–2034.

at the same time it is anticipated that increasing numbers of countries will plan to use alternative methodologies. There are a number of reasons for exploring alternative approaches, principally:

(a) the need to produce more frequent and timely statistics;
(b) budgetary limitations for census taking;
(c) reluctance of the population to participate in the census;
(d) increased technical capacities to utilize data sources; and,
(e) the availability of data on more variables from multiple sources.

1.134. This section aims to briefly elaborate on possible methodologies for conducting censuses based on the recent experiences of countries. The section also describes the necessary conditions for using a specific methodology, its advantages and challenges, and its implications for the content and administration of the census. It should be kept in mind that countries using a specific census methodology might have significant differences in implementation of the methodology, arising from differing national conditions and expectations. Regardless of the approach, the key principle of providing detailed statistics at the lowest geographic level (and for small population groups) remains of paramount importance, reinforcing the importance of geocoded information.

1.135. The various census methodologies are represented in a matrix in Table 1. This presents only those options that either have been used, or are likely to be used, by countries and does not cover all possible variations or combinations.

1.136. The different approaches that are summarized in Table 1[36] are then described more fully in sections A to C. First, the full-field enumeration approaches (section A) and the register-based census (section B) are presented; then the combined methodologies are described (section C). Alternative approaches have been adopted in different ways by different countries, depending on national contexts, preferences and practices and the availability of appropriate data sources.

[36] For more details see *Guidelines on the use of registers and administrative data for population and housing census,* UNECE, 2018; *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices*, United Nations Economic Commission for Europe, ECE/CES/13, Geneva, 2011; and *Register-based statistics in the Nordic countries: Review of best practices with focus on population and social statistics*, United Nations publication, Sales Number E.07.II.E.11, United Nations, Geneva, 2007.

Table 1. Main features of census methodological approaches

Methodology	Approach	Description
Field enumeration (see Section A)	Full field enumeration (traditional census) or full field enumeration with yearly updates of characteristics (see A.1(i)-(iii))	Information on census topics concerning individuals and households is collected by census questionnaire directly from respondents using enumerators or other modes of data collection (such as telephone interview, mail out/mail back, Internet), or by applying a combination of different modes of data collection. In a traditional census, full field enumeration may include an in-built sample survey by use of the long form. It is common that a short form is used together with a long form. Both forms are utilized during the same time frame of the census. While the long form estimates are not based on full coverage, they are nevertheless regarded as census data. Through modelling, the full field enumeration with the continuous cumulative survey is used to generate yearly (or other interval) estimates of detailed characteristics for different geographic levels. With this approach, the sample can be cumulated over time to produce statistics at the lowest levels of geographic detail to provide more frequent and relevant data.
	Rolling census (see A.1(iv))	Information on individuals and households is collected through a continuous cumulative survey covering the whole country over a period of time (generally years), rather than on a particular day or short period of enumeration. The two main parameters of a rolling census are the length of the period of enumeration (which is linked to the frequency of updates required) and the sampling rate (which depends on the geographic levels required for dissemination purposes).
Register-based (see Section B)	Fully register based (see B.6(i))	Information on individuals and households is collected from existing administrative sources, namely different types of registers, of which the following are of primary importance: population, building/dwelling and address registers. These are linked at the individual level with information taken from other administrative or statistical sources, such as business, tax, education, employment and other relevant registers.
	Base register and existing sample surveys or integrated administrative	Information on individuals and households is collected from existing administrative sources, namely different kinds of registers, of which the following are of primary importance: population, building/dwelling and address registers. These are linked at the individual level with information from

	sources and existing sample surveys (see B.6(ii))	existing sample surveys. No field data collection will take place. Existing sample surveys include intercensal sample surveys on different topics, such as the labour force survey and the living standards survey.
Combined methodology (see Section C)		Information on individuals and households is collected by combining data collected from an *ad hoc* survey or full field enumeration with administrative data sources or statistical registers. Data from administrative sources or registers are employed not only as a frame or to support field operations, but directly as a data source for some census information. In some cases, register-based data are used to prefill the questionnaires to be verified or corrected during data collection. Full field enumeration or *ad hoc* sample survey is used to provide information on those census topics not available from administrative sources or to adjust data that are of poor quality in registers.

A. Full field enumeration

1.137. The full field enumeration (or traditional) approach comprises a complex operation of actively collecting information from individuals and households on a range of topics at a specified time, accompanied by the compilation, evaluation, analysis and dissemination of demographic, economic and social data pertaining to a country or a well-delimited part of the country. Members of the public respond to a census questionnaire, or enumerators are deployed to collect information from respondents. For interviewer-based censuses, enumerators assigned to different EAs cover all households and persons in their EA during a specified, and usually short, period of time in order to meet the requirements of universality and simultaneity.

1.138. Full field enumeration can be based on the concept of a usual resident population count (counting people at the place of their usual residence) or population present count (counting those where they are physically present at the reference date, regardless of the place of their usual residence) (see Part V, Chapter III for more elaboration). This could affect the length of enumeration periods, field operation planning, and census procedures. Various methods can be used for collecting the data, including: a face-to-face interview conducted by an enumerator (with paper or electronic questionnaire); via a telephone response facility; self-enumeration via the Internet or paper questionnaire (mailed out and mailed back or dropped off and picked up by an enumerator), or a combination of these methods. Countries employing the full field enumeration approach may utilize very different collection approaches in doing so.

1.139. The full field enumeration census has merit in providing a snapshot of the entire population at a specified period and data for small geographic areas and small population groups. In that sense, the 'traditional' census is perhaps unique in nature. This approach is particularly suitable for countries requiring population numbers by various social and economic characteristics simultaneously for all geographic levels to meet the needs of planning and the allocation of funds. The delimitation of electoral boundaries requires simultaneity, and for that reason also the traditional approach may be more appropriate. But at the same time, traditional censuses have been singled out as the most elaborate,

complex and costly data collection activity that national statistical authorities undertake. In addition to costs, this complex task requires full awareness and agreement of the public to participate in it.

Necessary conditions

1.140. It is essential to have national legislation for conducting the population and housing census to ensure confidentiality, transparency and the cooperation of the population. A permanent central census organization, which may or may not be part of the statistical office, needs to exist in the country, which can be expanded during the time of the census. Since a full field enumeration census requires substantial resources, sufficient funding for a field operation covering the entire country and subsequent data processing needs to be ensured. Other conditions necessary for this approach are the support of, and general acceptance by, the public to participate in the enumeration, and trust in the integrity of the statistical office or census agency.

Some considerations for census taking and content

1.141. Very precise planning is required for every stage of census taking in the traditional approach, due to the sheer volume of work and overlapping time frames. The recruitment and training of a large number of census takers adds to cost and complexity. The involvement of administrative machinery at the central, provincial and local levels is essential for successful field operations.

1.142. Since the data obtained in the full field enumeration method are respondent- and enumerator-based, there is risk of error in canvassing the questions and in the quality of response. This, however, can be minimized through a well-considered design and testing of the questionnaire, effective training and effective publicity.

1.143. Data can be provided for every administrative level subject to privacy and confidentiality considerations, which may not always be possible with other census methods if some parts of the data collection are based on sample surveys.

1.144. All the essential features of a population and housing census (described in Section II.B paragraphs 1.8. 1.14.) are fully satisfied with the full field (traditional) census method.

Advantages and challenges

1.145. The two most significant advantages of a full field enumeration are comprehensiveness of coverage and simultaneity. Another major advantage is the flexibility in deciding the topics to be covered and design of the questionnaire, and the ability to collect data on population groups for which there may be no information available from administrative or other sources. There is lesser need for complex data adjustment since processing of raw data provides all inputs. The census frame becomes the base for all subsequent sampling frames. Finally, the focused and time-bound nature of the field operation implies that the data collection is finished in a short period. However, it is important to note that this may not always be the case and the duration of data collection period may vary. Moreover, regardless of the duration, ongoing monitoring is necessary to ensure data quality and completeness throughout the process.

1.146. One of the biggest disadvantages of a full field enumeration is its cost and administrative complexity. Another disadvantage is that it has a very long processing time. Also, since it can be conducted

only at 5- or 10-year intervals, data tend to become quickly outdated. The burden on respondents can be considered a disadvantage, especially in countries where participation in the enumeration is declining. Finally, many countries experience increasing difficulties in enumerating specific population groups, such as persons with high mobility or with multiple residences, or who are difficult to reach for other reasons. Unexpected emergencies such as natural disasters and public health incidents could also significantly disrupt the traditional census.

1. Types of censuses conducted with full-field enumeration

1.147. This section outlines and discusses the four primary variants of the full field enumeration census: the census with a single questionnaire, the census with short and long forms, the census with yearly updates of characteristics, and the rolling census.

1.148. The two most widely used types of the full field enumeration census are the census with a single questionnaire and the census with both short and long questionnaires. In the latter case, the short form contains only questions intended for universal coverage, while the long form is used to collect information from only a sample of households and population. In addition to these two main types, there are also a number of alternative census approaches that combine elements of both types of questionnaire.

1.149. Overall, the decision of whether to use a full field enumeration census with both short and long questionnaires, a single-form census, or an alternative method depends on a number of factors, such as the budget, the resources available, and the specific needs of the country or region.

(i) Full field enumeration census with a single questionnaire

1.150. A full field enumeration census with a single questionnaire is a census in which all households receive the same questionnaire. This type of census is conducted using door-to-door enumeration or self-enumeration by mail or online or enumeration through a call center established at the census office.

Advantages and challenges

1.151. There are several advantages to using a full field enumeration census with a single questionnaire. Firstly, it is a proven method that has been used successfully for many years. Secondly, the use of a single form can reduce the complexity of the census (it is relatively simple to administer and process). Thirdly, it can help to ensure that all households are counted and weighted equally. Also, one of its advantages is its flexibility to include topics on which information from other sources is not available. However, this approach also presents challenges, such as the cost of the field operation and the burden placed on respondents in terms of the time and effort required to complete the questionnaire.

1.152. Full field enumeration censuses with single questionnaires are still the most common method of conducting a census today, but countries are increasingly looking for ways to improve the efficiency and effectiveness of such data collection operations, and exploring alternative sources of data, such as administrative records.

(ii) Full field enumeration census with both short and long questionnaires

1.153. A full field enumeration census with both short and long questionnaires is a census in which some households receive a short questionnaire, and other households receive a long questionnaire with more questions. The short questionnaire, meant for universal coverage, collects basic demographic information, such as name, age, sex, marital status and ethnicity. The long questionnaire, on the other hand, collects more detailed information on a variety of socioeconomic and housing topics, such as education, employment, income, and housing. Both forms are utilized during the same time frame of the census. While the data collected on the long form provide estimates that are not based on full coverage, they are, nevertheless, still regarded as census outputs (as the disseminated output covers the whole population).

Advantages and challenges

1.154. The use of two questionnaires allows the census to collect a broad range of data without burdening all households with the long questionnaire. However, this type of census can be more complex and expensive to administer than a census with just a single questionnaire. Some of the advantages of using a full field enumeration census with both short and long questionnaires are:

- collects a broad range of data;
- reduces the burden on households; and
- can be used to estimate population characteristics for small geographic areas with the expansion of the sample data (usually an aggregate of EAs).

But there are some challenges as well:

- more complex and expensive to administer than a census with a single questionnaire;
- may be subject to sampling error;
- may be difficult to reach certain populations.
- more difficult to explain the results to users;

1.155. Section II of Part Three, paragraphs 3.93.9-3.17 provide a more in-depth elaboration of the two-questionnaire approach.

(iii) Full field enumeration census with yearly updates of characteristics

1.156. A full field enumeration census with yearly updates of characteristics is a variation of the traditional census design combining a full enumeration with a yearly sample survey. This approach focuses on counting the population and collecting only basic data in the census year while detailed population and housing data are collected throughout the intercensal period. In a traditional census, typically conducted once every 10 years, all households are counted at a specific point in time. This means that the data collected in such a way can become outdated over time. In a continuous survey, a household sample is interviewed regularly, such as quarterly or annually. This means that the data collected from the continuous survey is more up-to-date than that collected from the decennial enumeration. However, the annual survey only collects data from a sample of the population, so its coverage is not as complete as a full field enumeration. But by combining the two methods, both timely and accurate data on the population can be collected. This approach generates yearly (or even more frequent) estimates of detailed characteristics of population and housing.

Advantages and challenges

1.157. The primary advantage of this approach is that it provides more frequent and relevant data on population and housing than would be available when a census is conducted only once a decade. However, such an approach might be costly and technically difficult, as it requires a continuous programme of comprehensive planning, development and testing.

1.158. Some of the advantages of adopting this approach are:

- produces more frequent and up-to-date population estimates;
- can be used to estimate population characteristics for small geographic areas with the expansion of the sample data (usually an aggregate of enumeration areas); and
- reduces the burden on households.

But (as with the short/long form method) there are some challenges as well:

- more complex and expensive to administer than a traditional census alone;
- may be subject to sampling error; and
- may be difficult to capture information on certain populations.
- more difficult to explain the results to users;

(iv) Rolling census

1.159. A rolling census is a variation of the traditional census where the entire population of a country is surveyed over an extended period. In a rolling census, information is collected on individuals, households, and dwellings through a continuous cumulative survey covering the whole country over several years, rather than on a specific day or short enumeration period.[37] This method involves conducting a sample survey of households each year and combining the data with previous years' data to produce annual population estimates. This approach allows for continuous data collection and provides more up-to-date information on population dynamics. However, it requires complex sampling and modelling techniques and may lack the snapshot view provided by a traditional census.

1.160. The two main parameters of a rolling census are the length of the period of enumeration and the sampling rate (which depends on the geographic levels required for dissemination purposes). For example, it is possible to build a sample framework in order to produce national results with one annual survey, regional results by cumulating three annual surveys, and small-area results by cumulating data over five years. Annual surveys may be conducted over the full course of the year or in a particular month or some other shorter time frame.

Necessary conditions

1.161. Implementation of such an approach requires: highly complex sampling and modelling techniques; a high-quality sampling frame in order to allow sampling at very low levels of geography; and successful consultation to gain acceptance of the approach with major stakeholders, including national and local governments and the user community.

[37] In the 2010 round of censuses France was the only country applying this concept.

1.162. The necessary conditions partly depend on the complexity of the sample framework. If the sampling units are addresses, a master address file is to be built first. But if the sampling units are larger, for example municipalities, it is only necessary to have enough information to spread the municipalities over the different years. It will be necessary to explain to statistics users the impact of the rolling sample on the use and interpretation of data, as many users are more used to snapshot data rather than period data.

Advantages and challenges

1.163. The main advantage of the rolling census approach is the potential for higher frequency for updating data: a full field enumeration census provides benchmarks every five or, more commonly, ten years. In contrast, the rolling census provides annual updates at the national level (with a few years of data collection needed for sub-national level updates). The high peak costs and labour requirements of a traditional census are instead spread over a longer period. Furthermore, it is possible to improve the census process over time, and to test methodological refinements and new technologies as they emerge.

1.164. The rolling census offers flexibility to introduce new topics as data requirements change, ensuring the census remains relevant. By adopting emerging technologies and establishing permanent teams, it enables continuous evaluation of data quality and training of field staff. However, introducing a new topic into a rolling census gives a comprehensive picture only after a complete data collection cycle for all regions, which may take several years. While results are generally unbiased and accurate from the first year in large geographical areas, they become accurate for all areas after a full cycle.

1.165. The rolling census approach does not provide a simultaneous snapshot of the whole population due to varying enumeration times, thus making comparisons between areas more challenging. The results represent an average over the cycle rather than a snapshot, which may not be fully relevant if trends are rapidly changing. For example, a sudden shift in migration patterns caused by a specific event (such as a natural disaster or conflict) could be obscured by the averaging effect of the rolling census.

Some considerations for census taking and content

1.166. It is better to initiate a rolling census just after a full traditional census, in order to exploit the recent census information in order to build the sample framework. As the operation is annual, the process must be very carefully prepared, since any delay can be problematic for the subsequent stages.

1.167. A rolling census is able to include all usual census topics. There is also the possibility of changing the questions more regularly than in a decennial cycle. This enables the census to be more reactive to changes in the needs of users, even if comparability over time should, in principle, be preserved. However, only if the questions are stable over a number of years can a rolling census produce statistics at the same level of detail as in the more traditional approach. Depending on the census organization and procedures, it may be possible to add some thematic surveys if required.

2. Innovations in conducting field-based censuses

1.168. There are a number of innovations that NSOs can pursue to improve the efficiency, accuracy, and accessibility of field-based censuses. Some of these innovations include:

a) **Technology**. Technology is playing an increasingly important role in conducting traditional censuses. For example, many countries are now using mobile devices to collect census and geographic coordinates data directly in the field, while the Internet is being utilized for self-enumeration. This can help to improve the efficiency and accuracy of data collection and the territorial coverage of the census operation. Technology can also facilitate the real time monitoring of field operations and response rates through digital field management systems (see paragraphs 3.209-3.213 for specific examples). Some countries are experimenting with using machine learning to develop new methods for estimating the population of small geographic areas, to improve the accuracy of annual population estimates, or to encode certain variables (transforming categorical or textual data into numerical representations so machine learning models can read them).

b) **Administrative data**. Administrative data is providing useful information to improve the preparations for, and the collection of information in, a census. Some examples of this include:
   - Creation of address register
   - Developing lists to replace or facilitate census listing process
   - Inform on planning and collection operations (for example, determining the size and boundaries of enumeration areas)
   - Dwelling occupancy verification
   - Informing on areas likely to respond using paper forms
   - Informing on likely areas of hard-to-reach populations
   - Geocoding special locations (such as places of education)
   - Determining non-response follow-up contacts
   - Filling in missing values and improving imputation procedures
   - Validating census data
   - Weighting/calibration for surveys related to the census

   Paragraphs 3.192-3.196 and paragraphs 3.302-3.306 describe, for example, how administrative data can be used, respectively, during the field operation and data processing.

c) **Geospatial information.** Geospatial information must be considered an integral part of the census methodology and operations. Geospatially relevant methods, tools and activities – such as the Geographic Information System (GIS) – should be integrated into the regular production processes of NSOs, so that the design and production of geospatially enabled statistics can be conducted in a systematic and consistent manner. Some examples of this include:
   - Satellite images and aerial photography
   - Global Navigation Satellite Systems (GNSS)[38]
   - Geospatial intelligence (for efficient placement of field offices)
   - Georeferenced address registry
   - Geocoding buildings and households
   - Map of enumeration areas

[38] Global Navigation Satellite System - GNSS is the standard generic term for satellite navigation systems that provide autonomous geo-spatial positioning with global coverage. This term includes e.g. the GPS, GLONASS, Galileo, Beidou and other regional systems.

- Geodatabase
- Digital census cartography
- Geocoded administrative data (for example from electricity agencies or distribution companies)
- Real-time monitoring of fieldwork with geoportals and dashboards
- Route planning and tracking enumerators
- Geopackages for handheld devices
- Geofencing[39]
- Heatmaps with concentration of refusals and vacant housing units
- Geographic population distribution (for example in urban, rural, remote areas)
- Spatially enabled topics/questions[40]
- Map services and geoportals to disseminate census results
- Degree of Urbanzation (DEGURBA)
- Grids
- Common Geographies and Functional Areas

Section IV in Part Three of these *Principles and Recommendations* discusses in more detail the use of geospatial information in the planning and execution of the population and housing census.

d) **Data integration.** Countries are also increasingly integrating data from different sources to produce more complete and accurate census data. For example, some countries are using administrative records, such as birth and death records, to update their population data, while others are using satellite imagery to estimate the population of remote areas. Others are using a combination of administrative records and sample surveys and geospatial data to produce more frequent population statistics. (See paragraphs 1.109. -1.130. on the integration of census data with geospatial information and other data sources).

e) **Using supplementary population bases to generate population statistics: Supplementary** population bases (such as service or daytime populations) may offer researchers and policymakers the most accurate and relevant information for specific analysis needs.

f) **Public engagement**: To ensure a successful and accurate census, innovative public engagement strategies are paramount. This includes utilizing interactive and accessible outreach through social media campaigns, user-friendly websites and mobile apps. Additionally, multilingual communication and engaging content like infographics and videos can further bridge the gap.

[39] Geofencing is a resource that uses technologies to create a virtual geographic perimeter, by obtaining data through Wi-Fi networks, GNSS, radio frequency identification (RFID) and other geolocation tools.

[40] Applicable for opening topics or questions for specific populations (e.g. native or indigenous population, informal settlement dwellers, quilombolas) in locations that are not defined in enumeration areas or where the population is dispersed. These can configure areas of operational interest (AOIs) that will define the application of the question through on-site geospatial verification via geolocation at the time of the interview.

Incentives for participation and gamification[41] elements can also increase engagement. Transparency and building trust through clear data privacy assurances, open data platforms, and community feedback mechanisms are equally important. Promoting civic engagement with school and intergenerational engagement activities programs can ensure everyone understands the importance of being counted. Through these innovative approaches, census operations can foster a more engaged and informed public, leading to higher participation rates and ultimately, more accurate and complete data. (See Section X in Part Three for a full discussion of use consultation, communication and publicity to promote the census.)

1.169. These are just a few examples of the many innovations that are being used to improve field-based censuses. By using these innovations, countries can produce more accurate, timely, and accessible census data. In addition to the above innovations, countries are also exploring the use of new technologies, such as artificial intelligence and machine learning, to produce population statistics. These technologies have the potential to revolutionize the way that population statistics are collected and produced. Such technologies must be evaluated carefully first before deployment to avoid biases and errors.

B. Register-based censuses

1.170. The concept of producing census-like results based on registers was developed in the 2000 round of censuses, although it has been debated and tested to various degrees since the 1970s, and several countries succeeded in using this approach to generate census data in the 1990 round of censuses. The philosophy underlying this concept is to take advantage of existing administrative sources, namely different kinds of registers on individuals, households and dwellings. These registers are linked at the individual record level with information held on business, tax, education, employment and other relevant registers. While it is theoretically possible to link records on the basis of the name and other unique details of the individuals, the existence of a unique identification numbers for each individual, household and dwelling allows a much more effective and reliable linkage of records from different registers.

1.171. Administrative data sources can, depending on their content and quality, be used in all phases of census taking. In principle, where greater amounts of information can be obtained from administrative sources, the production of census-type statistics will be faster and cheaper. The most complete use of registers will be where all core and, in some countries, non-core census topics can be based on register information. The quality of data collected through administrative registers can be enhanced by conducting sample surveys to measure and assess its accuracy. These surveys can also be used to gather additional information that may not be available in the administrative records. The sample survey(s) may either use the register as a sampling frame, or else be completely independent of the register.

1.172. To provide specific demographic data and socioeconomic characteristics, the register-based census relies on the collection of available administrative data and information from government and private entities, as well as other service providers. The adoption of a register-based approach will enable

[41] Gamification in the context of census data collection refers to the use of game-like elements or mechanics to engage and motivate the public to participate in the data collection process. By incorporating elements such as points, rewards, challenges, or competition, gamification can make the census process more enjoyable, interactive, and rewarding for participants, thus leading to increased participation rates and higher-quality data.

the provision of systematic, wide-ranging and timely statistical products that support informed decision-making.

1.173. The use of administrative data sources alone may reduce the flexibility of the census in terms of the variables that are available and their definitions. It may be difficult to change the variables as these are defined to fulfil the various administrative purposes for which the registers are created. A significant potential risk to the success of the census is that any administrative source will usually be beyond the control of the statistical authority. The influence of the NSO over the administrative source can often be very limited. The content and availability of such sources may change at relatively short notice and without reference to statistical needs. For example, a change in taxation legislation may mean that a key administrative register may no longer collect the specific information needed for the census. This risk can be minimized by establishing close and regular communication and cooperation between the statistical authority and the owners of the administrative sources.

1.174. Administrative registers can suffer from coverage issues, with certain populations or persons under-represented or over-represented in the data. Most of the time, only persons legally present in the country would get covered through this approach. There is a risk of certain population groups being absent in administrative data sources. For example, irregular or undocumented immigrants, recently arrived migrants not yet registered but fulfilling the usual residence criteria, homeless persons, nomadic or transient populations, persons involved in illegal employment or activities, etc. These groups are not likely to be recorded on certain occasions in any such administrative registers. On the other hand, registers may include persons who are no longer living in the country (for example persons who have emigrated but have not yet been removed from the registers), or, indeed, are no longer living at all (where, for example, their deaths have not yet been recorded). Some of these typical coverage problems could be corrected with the use of the "signs of life"[42] methods (see paragraph 1.206.).

1.175. Subject to the caveats mentioned in the paragraphs above, the essential feature of "individual enumeration" is satisfied in this approach as separate information is collected regarding the characteristics of each individual. As regards "universality within a defined territory", this criterion is satisfied as the enumeration base is taken from a population register. With regard to "simultaneity", the timing of the census extraction may require careful thought where register update cycles vary. With respect to "periodicity", this approach allows extraction at any desired frequency, including "at least once in 10 years", noting again the need to manage the updating cycles for the registers. Finally, in most cases, the requirement in terms of producing small-area statistics is largely met, as the information in the registers generally allows for such aggregates to be generated, provided that the records in the register have sufficient georeferenced information to create an output geography at the desired level.

1.176. It is possible that countries with larger populations may have greater difficulties when making the transition to a register-based census, since the volume of information that must be analyzed and

[42] The 'signs of life' method involves applying a series of criteria to identify active residents in a population register. These criteria, often referred to as 'activity rules,' help ensure that only individuals who are alive and meet specific residency requirements are included in the register-based census. By cross-checking data from various administrative sources, the method allows census organizations to enhance the coverage and accuracy of population statistics derived from the register.

integrated is higher. Moreover, it is also possible that larger countries, instead of having a single register at national level, have several registers at the regional and local levels, which would increase the number of data sources to be processed and further complicate the transition to a register-based census.

1. Necessary pre-conditions for the use of administrative registers for censuses

1.177. Before administrative data can be effectively utilized for censuses, several key conditions must be met. These prerequisites are crucial to ensure the successful integration of administrative data into the census process, ultimately leading to the cost-effective production of more accurate, efficient and timely population statistics. A well-established central population register, or systems of registers, must exist with the infrastructure to facilitate access to the data it houses. NSOs should, wherever possible, ensure that the registers are of high quality, accurate, reliable and relevant for statistical analysis, as well as have comprehensive geographic coverage, encompassing the entire population of the country. The registers should be updated continuously to ensure they reflect the most recent demographic changes. Furthermore, a consistent set of concepts and definitions should be adopted across all administrative registers to facilitate data integration and analysis. Having common identifiers across all registers also ensures an effective linkage mechanism. The availability of full metadata of administrative data (database structure, definitions, information related to data collection, etc.) is also important.

1.178. Some of the essential pre-conditions[43] that a country should have in order to be able to conduct a register-based census are:

a) Legal framework

1.179. If administrative data are to be used for census purposes, statistical authorities should have a clear legal mandate to use such data for statistical purposes. Individual historical, cultural and political factors of each country can lead to highly diverse legal frameworks, but ideally such legislation should encompass:

Data access

1.180. The legal framework should grant NSO the authority to access relevant administrative data held by other public authorities for the purpose of producing official statistics. This access should be clearly defined in the legislation, outlining any limitations related to data confidentiality or duration of access.

1.181. While specifying data sources in primary legislation might be inflexible, the legislation could acknowledge the NSO's right to access all relevant administrative data at the individual record level, with the understanding that specific sources may evolve based on user needs and administrative changes. Additionally, the legislation should explore avenues for the NSO to collaborate with data providers. This might involve some influence over the creation, revision, or deletion of administrative data elements crucial for statistical purposes, although achieving this might require further exploration and potentially separate agreements.

[43] For more elaboration of the preconditions, refer to the UN Handbook on Registers-based Population and Housing Censuses. Link: https://unstats.un.org/unsd/demographic-social/publication/handbook-registers-phc.pdf

Privacy, integrity and security

1.182. To ensure the safe and secure handling of data and foster public trust, a robust legal framework is essential. This framework should comprise several key acts, including a statistics act, a privacy act, and potentially a data protection act. These regulations should govern all aspects of data handling within the NSO and its interactions with other entities.

1.183. The legislation should emphasize the NSO's legal obligation to protect the confidentiality of administrative data. A "one-way traffic" principle should be established, whereby data flows from data providers to the NSO for statistical purposes but generally not the other way around. Specific circumstances permitting exceptions to this principle should be mandated in the legislation.

1.184. Regulations should clearly define how unit record data can be transferred to, and used within, the NSO. Additionally, the legislation should specify how the output microdata can be shared with other departments, organizations, and users, ensuring such access is strictly for statistical purposes and adheres to robust privacy-preserving and confidentiality control techniques.

Data use

1.185. The list of variables on which data is to be collected may vary from census to census. Thus, these should not be specifically prescribed in the 'framework' legislation; but can be stipulated in any secondary regulations relating to each particular census. Variables, with metadata, including identifiers of administrative data sources, should be listed completely and described clearly. Limitations to the data use (for example duration of use, deletion of microdata, recoding of national unique identifiers, etc.) should be described. Furthermore, it should be clearly defined that unit record data compiled for statistical purposes will not be retransmitted to the data-providing organization or other governmental authorities.

Transparency

1.186. Transparency serves as a foundational principle of a register-based census, fostering public trust, accountability, accuracy, understanding, privacy protection, and international comparability. By promoting openness and clarity throughout the census process, transparency contributes to the credibility and legitimacy of census results and supports evidence-based policymaking.

b) Political support and cooperation

1.187. A joint effort towards register-based statistics production requires close collaboration among relevant authorities (as well as a firm and explicit commitment at the highest possible political level to achieve the requisite legislation noted above). Cooperation between statistical and administrative authorities generates a mutual and deeper understanding of the primary purpose of the registers and the needs of the statistical authority. It is recommended that the collaboration between the administrative authorities and the NSO be formalized.

c) Confidentiality and public approval

1.188. In the context of a census, the most important principle for the population is the confidential use of the individual's personal information, as stated in the sixth Fundamental Principle of Official Statistics, which requires that the use by statistical authorities of individual data, whether they refer to natural or legal persons, be strictly confidential and used exclusively for statistical purposes.

1.189. The political decision concerning the use of administrative data in a census can be highly influenced by public approval or refusal. In the run-up to implementing a new or modified census methodology it is helpful to inform the public about the intentions. It can be expected that people will become increasingly sensitive towards the collection and analysis of personal data by governmental authorities. In addition to outlining the general benefits and risks of the use of administrative data, information to the public should focus on the confidentiality of personal microdata. Clear limits and rules regarding the use of administrative data provide a common understanding that individual data collected for statistical purposes will not be passed on to other governmental authorities.

1.190. User acceptance is crucial for leveraging administrative data sources in censuses. Transparency about data quality, limitations, and potential biases is essential to ensure that users will trust and utilize the resulting statistics effectively. Building trust in the use of administrative data requires clear communication with stakeholders and data users. Addressing concerns about privacy and data security is paramount for achieving widespread user acceptance.

d) Administrative routines

1.191. The decision about the use of register data for statistical purposes largely depends on the nature of the register itself, including qualities such as integrity, reliability and lifespan. Normally, data that are important for the administrative agencies responsible for the registers are recorded with a high degree of accuracy. There should be confidence in the administrative authority's capacity to be a reliable partner and data supplier. This implies the presence of administrative routines and safeguards. Does, for example, the administrative authority have extensive experience with the collection of the data that may be used for statistical purposes? Is the administrative authority well organized and is it anticipated that the necessary data collection will continue into the future? Are there existing quality guidelines for the administrative authority that guarantee long-term data quality? These are some of the questions that need to be answered when assessing the feasibility of exploiting administrative registers for statistical purposes in general, and for generating census statistics in particular.

e) Identifiers

1.192. Regardless of the census methodology adopted, it is extremely important that a unique unit identifier is used in all the data sources. The use of a unique identifier is essential in order to link unit record information from different sources successfully. This primary key may already exist in the country – for example, a national personal identification number (PIN). Where it does not exist, or exists but with poor quality (for example, where too many units are missing), it can be artificially created for statistical purposes. Such a statistical linkage key can be built from unchanging variables for persons, such as "family name at birth", "first name", "date of birth", "sex" and "place of birth". Care needs to be taken with alternative spellings, for example, incorrect or incomplete register entries, transcription errors, and the varying transcription of foreign alphabets, names or place names. In any case, linkage between different registers can be made both by deterministic criteria (complete equality of one or more variables) or by probabilistic criteria (high similarity of one or more variables).

1.193. Unique identifiers assist in the detection (and correction as necessary) of identical statistical units (duplicates). Duplicate records most often arise when collecting data from more than one decentralized register, but are also possible within one centralized register. The problem with duplicate data entries is the risk of multiple counting of identical statistical units.

1.194. In the case of fully register-based censuses, information from different registers should be matched using good-quality identifiers. It is important to define how often information from different sources is updated and the reference date of the information stored in the different registers. When two or more data deliveries with the same content from the same administrative authority are planned, a linkage key will enable validation of data quality with regard to the statistical reference period.

1.195. Successful data linkage may be compromised by poor quality of the source data. Information stored or provided by data owners may have errors resulting in non-linkage or multi-linkage of records. In these situations, probabilistic approaches that choose the closest candidate, or the use of geographic information (starting from the lowest detail level and gradually increasing), can help to improve the linkage process.

f) Institutional infrastructure and capacity

1.196. To transition to a register-based system, NSO staff must develop a wide range of skills, including relationship management, information management, data mining and analysis, data protection, communication, and specialized statistical skills for data linking processes. While NSOs typically have expertise in information management, a register-based system requires additional statistical knowledge, particularly for data linkage and matching. There is significant variation between countries in the availability and quality of administrative registers, necessitating strengthened capacity in NSOs, especially in developing countries, to improve the use of administrative registers for small area statistics. This includes identifying usable registers, assessing their quality, and designing census processes with appropriate methodologies.

1.197. A long-term strategy is essential for developing a register-based census, involving the improvement of key registers. NSOs need to develop infrastructure and training programs to equip staff with the necessary statistical and IT skills. Administrative registers should ideally align with NSO's statistical requirements, but systems must be prepared to transform administrative data into statistical records. This involves coding to standard classifications, editing records, linking data between registers, and creating new variables. Without these skills and a thorough understanding of existing variables and definitions, producing trusted statistical outputs may not be possible.

2. Types of registers

1.198. Administrative registers are produced on the basis of administrative processes to collect information on units (persons and dwellings) and variables that are defined by administrative rules and demands in a country. Although the content and process of registers may differ from one country to

another, the types of registers are usually very similar. The following provides definitions of the main concepts used in the system of administrative registers.[44]

a) A *register* is defined as systematic collection of unit-level data organized in such a way that *updating* is possible. Updating is the processing of identifiable information with the purpose of establishing, updating, correcting or extending the register.

b) *Administrative registers* are registers primarily used in an administrative information system. This means that the registers are used in the production of goods and services in public or private institutions or companies, or that the information is a result of such production.

c) *Base registers* are kept as a basic resource for public administration. The function is to keep stock of the population and to maintain identification information. They may also hold fundamental geospatial data such as addresses, geocoded buildings, establishments and dwellings, functional areas, geographical names etc. *Statistical base registers* are based on the corresponding administrative registers.

d) *Specialized registers* focus on a specific theme or domain, unlike base registers which serve a broad foundational purpose. Specialized registers often receive core demographic data (such as name, address, date of birth) from a base population register, but collect additional information relevant to their specific purpose. Examples include registers of vehicles, social security, education, business, health, etc.

e) *Statistical registers* are created by processing data from administrative registers for statistical purposes. A statistical register could be based on one or more administrative registers.

1.199. As noted above, administrative registers are maintained primarily for administrative purposes. Units and variables of administrative data are described according to administrative rules and demands. Before a register is used for census purposes, the suitability of its data in terms of definitions, concepts, content, reference date, accuracy and other criteria should be statistically examined by comparing them with previous census and survey results, and by conducting quality and compatibility surveys. A pilot census may be used for this purpose.

1.200. All persons in the register that are included within the definition of the census's population base are enumerated. The timing of the census extraction may require careful thought where register update cycles vary. Registration delays and administrative delays in updating between regional and national databases can otherwise have a serious impact on the quality of the output.

(i) Base registers

1.201. Base registers are those registers that hold the basic information relating to the stock of the entire set of 'population' units that are being counted in the census – typically persons, dwellings and buildings.

[44] *Register-based statistics in the Nordic countries: Review of best practices with focus on population and social statistics*, United Nations publication, Sales Number E.07.II.E.11, United Nations, Geneva, 2007.

Base registers serve as the foundation, and data from specialized registers are linked at the unit record level to create the statistical register. This linkage allows for the derivation of additional characteristics needed for the census. By combining the base register data with information from specialized registers, the requisite census attribute variables for the population and dwellings are created, ultimately enabling the production of census statistics.

1.202. The base registers that are most commonly used for the purpose of the census (and which are discussed in more detail in Chapter III, Section C.4 above) are:

(i) **Population register** – records all residents in the country, together with their basic demographic information (such as age, sex and marital status), and updates for births, deaths, changes of marital status, international migration (both into and out of the country), and internal migration (by recording changes of address within the country). See paragraphs 1.90. -1.94. .

(ii) **Building and/or dwelling register** - usually held by land and property valuation agencies or by local or municipal authorities responsible for the development of housing policies and urban planning. They can provide information on the location of all places where people might live (by distinguishing between wholly residential, partially residential and non-residential buildings) as well as some of the core census housing topics such as ownership, type of housing (for example whether a detached house, terrace house, or apartment), whether or not the housing is rented, living area, number of rooms, condition concerning availability of electricity, bathroom, kitchen, and toilet, and year of construction. With respect to a register-based population census such registers are of fundamental importance in the harmonisation with the population register, in that every person that is allocated a dwelling number can be associated with a particular housing unit, and (in some cases) a distinct household. Ideally, it is important that these records are geocoded with coordinates and/or addresses to facilitate their integration with other statistical and geospatial data. See paragraphs 1.95. -1.98. .

(iii) **Establishment (or business) register** - usually contains information about business establishments (including their industry, institutional sector, size of workforce, and location). Ideally, it is important that these records are geocoded with coordinates and/or addresses to facilitate their integration with other statistical and geospatial data. See paragraphs 1.99. -1.103. .

(iv) **Address register** – records all addresses, by location, and identifies whether the address is used for accommodation, commercial (business) or office purposes. The address register may also include links to the building (such as an apartment building) and/or the dwelling. Ideally, addresses should have associated coordinates to facilitate checking for duplicate records and integration with other databases. See paragraphs 1.95. -1.98. .

1.203. Without these base registers, it is very difficult to conduct a fully register-based census. However, even without a population register, it may still be possible to conduct a combined census, using data from other registers and information collected in a field enumeration (see paragraphs 1.241. -1.265. below).

(ii) Specialized registers

1.204. In addition to the base registers described above, there will be other specialised (or supplementary/subject-matter) administrative registers available to the NSO that contain records relating to those characteristics for which census information is required and from which the requisite variables

for core and non-core topics can be derived using the PIN for record linkage. Such registers may include the following:

a) Social security or pension registers;
b) Tax registers;
c) Employment, unemployment and jobseeker registers;
d) Education and student registers;
e) Health registers;
f) Border control; and
g) Other supplementary registers.

1.205. NSOs may also be able to access data from a range of other administrative sources either to provide topic-related census variables or to identify particular population sub-groups that are traditionally hard to count accurately (and for whom full information may thus not be recorded in the population register), such as the institutional population. For example (and this is by no means a complete list):

h) electoral registers;
i) registers of motor vehicles;
j) registers of foreign nationals or of residence permits;
k) lists of military service personnel;
l) registers of persons living in institutional places;
m) registers of persons with disabilities;
n) registers of recipients of social programmes;
o) registers of building permits; farm registers; and,
p) registers held by public facility service providers (such as for piped water supply, electricity and/or piped gas, and sewage and waste disposal).

1.206. Some specialized registers may help to identify issues of coverage by, in particular, identifying 'signs of life' activities. The basis of this methodology is that over time, a person living in a country inevitably leaves certain traces or markers of administrative activity in the form of records in different registers. This means it is possible to verify the person's presence in the country, as well as connections between persons and their locations, on an annual basis. Such verification is based on signs of life that are recorded in registers every year. For example: persons in the age 0-3 years old can be captured in health registers. School age population 4-17 years old can be captured in education enrolment registers. The remaining population may be captured in labour force and other registers.

3. Establishing a statistical population register

1.207. For the compilation of official statistics using registers, the development of a statistical population register is an important step. The statistical population register (SPR) is a systematized and indexed collection of individual records for every resident (including nationals and foreign citizens) of the country. It should be noted that a statistical population register is distinct from its administrative equivalent in that it is a comprehensive database designed specifically for statistical purposes, while the latter is primarily maintained for administrative or operational functions. Where developed and functioning, it represents the backbone of the compilation of official statistics and represents a master population frame as well for designing and running statistical surveys.

1.208. The statistical population register should be regularly updated with pertinent information, starting with the administrative population register. The legal framework for establishing, maintaining and exploiting the national statistical population register must ensure that it is used solely for compiling

aggregate statistics and that it cannot be accessed for any other purposes either within or outside of the national statistical authorities except for approved research, if lawfully authorized and strictly controlled and restricted only to anonymised records.

1.209. Creating an SPR is a key step towards achieving the use of administrative data for census purposes. It should be noted that it is important to distinguish between the registration of an individual unit (person) in the population register and the inclusion of that record in the computation of internationally comparable statistics on the stock of population and its basic characteristics, such as age and sex. While the listing of those persons in the population register will be in accordance with national legislation for the requisite administrative purpose, for census purposes only those persons meeting the requirements of being residents of the country should be included. For instance, persons who have left the country but are still holding a permanent residence permit should be still listed in the population register, but they should be excluded from the census population of usual residents. When using administrative registers, it is important to be able to identify those registered persons who have emigrated, especially in countries where there may be no incentive to report emigration to authorities. Some countries maintain administrative registers of their nationals living abroad, which can aid in this process.

1.210. An SPR is usually generated from the existing administrative population register by the NSO. In countries that have a reliable population register (for both nationals and foreign citizens), a statistical population register is routinely generated and updated from the existing population register. On the other hand, if there is no centralized population register, a statistical population register can be constructed by integrating local population registers and linking them with other existing registers. In either case the SPR should be updated regularly.

1.211. Thus, establishing a national SPR may involve some investment in collecting records from different existing administrative registers and/or administrative data sources, harmonizing them, linking information from different registers and running complex editing procedures to ensure the consistency and quality of individual records. Once established, and under the protocols spelled out in legal provisions, the national SPR is updated from administrative registers on a regular basis.

4. Constructing integrated statistical registers

1.212. A statistical register contains the data integrated from different sources including surveys and administrative data. The statistical registers are usually housed in the NSOs who usually can transform the data from administrative sources into the statistical data by: editing incorrect entries; removing inconsistencies and duplicates; harmonizing concepts, definitions, classifications and reference periods; and allocating new statistical identifications. Among the statistical registers, the population register is the backbone for any register-based statistical system.

1.213. An integrated statistical register typically plays the role of a data coordination tool, integrating data from several sources, both statistical and administrative. This may be done by linking records using common identifiers or by using matching techniques. It may sometimes be easier to use data from a single source, but in such cases it is often difficult to check the accuracy of that source. When several sources are used and integrated within a statistical register it is possible to have a much better view of the accuracy of the data. Unfortunately, the negative side of this is that it becomes necessary to have a strategy for dealing with conflicting data from different sources. However, if variables in statistical registers are stored with source codes and dates, automated algorithms can be used to prioritise sources and resolve most data conflicts.

1.214. An integrated statistical register may also provide the possibility of deriving new (census-related) variables not otherwise available from any one single register (such as industry of occupation, level of overcrowding or a housing quality index.)

1.215. Traditionally, statistical registers have been used as sampling frames for surveys, but as has been already noted, an SPR is at the very heart of a register-based census. Such a register may be established inside a statistical institute for statistical purposes. However, it is more common (and far simpler) for the NSO to transform the administrative register into a statistical register owned and managed by the agency itself, if permitted by law and the NSO has the resources to do so. This transformation process is usually a relatively simple task and may include editing, correcting, removal of duplicates, and translating the administrative ID into a statistical ID. The NSO may also then integrate data that the administrative register does not have with authenticated evidence of data from other sources (such as, for example, country of birth for people immigrating before the administrative register was established).

1.216. In the process of transforming administrative records into statistical records, NSOs should consider geocoding as a fundamental process in building integrated statistical records. Location, in many cases an identifier (coordinates and addresses) that will facilitate integration, should be incorporated into the NSO's regular production processes so that the planning and production of geospatially enabled statistical records can be conducted in a systematic and consistent manner. Although official geospatial information is mostly produced by the NGIA, all geo-located data is a constituent component of the data ecosystem, which NSOs often see themselves as custodians. In this sense, institutional integration and partnership must be fostered, taking advantage of the NSDI and global frameworks such as the GSGF and the UN-IGIF.

1.217. In addition to identifying the data sources to be used for any particular statistical operation, some of the key processes involved in the construction of a statistical register are:

- data quality management;
- linking the unit records from these sources;
- dealing with duplications;
- resolution of conflicting information relating to the same data items in different linked sources;
- updating; and
- editing and imputation.

5. Transitioning from field-based censuses to register-based approaches

1.218. Moving forward to a register-based census from the traditional field enumeration approach has, as already been noted, certain advantages such as reduced cost in the long run, improvement in the quality of data and ability to produce more frequent information at lower geographic levels. However, countries should consider their national context when deciding whether or not to utilize available administrative data for census purposes - and not all countries may wish or need to do so. For example, in some countries, privacy concerns discourage the use of administrative data for the census , or for other countries the data on variables necessary for particular policymaking (such as religion or ethnocultural characteristics) are not available in administrative data sources.

1.219. Country practices indicate that changing census methodology from a full field enumeration census to a fully register-based census usually takes several decades. There is no ideal method or duration for this transition and therefore each country should plan this process based on the initial assessment of the

availability of administrative registers as well as assessment of preconditions that have to be met before changing the data collection methodology. It should be noted that direct one step transition from a full field enumeration census to fully register-based census is, generally, not realistic, and that this transition should be planned gradually, introducing more variables each time, providing that the data from registers have been proven to be of good quality.

1.220. If a country decides to move to a combined or register-based census, the transition is usually done in several stages, often by first adopting one form of a combined census approach (discussed in Section C below), such as registers with full-field enumeration or registers with ad-hoc surveys, before moving to a fully register-based census. In this case, dependency on administrative data sources can be gradually increased and the cost of the change can then be spread over two or three census cycles. At the beginning of this process, key registers such as a population register might be incomplete or of insufficient quality to be used as source for the census. In such cases, it should be made clear from the beginning that the NSO should not attempt a move immediately and should continue conducting some form of field-based census. However, even when countries continue to do so, innovations making greater use of administrative data could help the NSO work more efficiently (as discussed in paragraph 1.168. above). It is beneficial if the relevant public authorities make administrative data sources available to the statistical institutes to produce proto-type register-based statistics. The government itself can help a great deal both by removing any legal barriers to data sharing and by subsidising the transition.

1.221. It should be emphasised that there is no unique way of moving to register-based approaches; in fact, previous national experiences show that there are many different ways of designing this transition, especially for countries that are planning to use administrative data sources for producing census information for the first time. Therefore, these recommendations provide only a general framework for moving to a combined or register-based censuses. Four key phases of the transition include:

(i) **Establishing/maintaining necessary pre-conditions**: For a successful transition from a full field enumeration to a combined or register-based census, it is necessary to establish legal framework for being able to use individual data for statistical purposes. Institutional capacity and establishing effective and continuous communication and cooperation with all stakeholders are also very critical.

(ii) **Identifying administrative data sources and other sources**: This phase needs to be carefully planned and managed in collaboration with register-holders and users. The NSO should build up a wide-ranging knowledge of the existing administrative data sources before making any decision on changing the census methodology.

(iii) **Transforming administrative data into statistical registers**. This phase includes a series of activities that are implemented for transforming administrative data into statistical data from which the census outputs can be produced. The design of this phase will differ from one country to another depending on the quality of existing population register and other administrative registers. Countries with good quality of population register can more easily construct an SPR. However, countries that do not have reliable population register, may have to carry out some form of field enumeration or conduct ad-hoc surveys in order to be able to construct the requisite SPR.

(iv) **Quality management and assessment**: Quality management and assessment is an overarching process implemented by NSOs throughout whole phases of this transition. For a successful design of this phase, it is important to develop a system for continuously monitoring the quality of activities and assessing the quality of each phase so as to be able to understand the effects of

the quality on the next phase and, finally, on the census outputs[45]. (Part Four of these *Principles and Recommendations* discusses quality issues in more detail.)

6. Types of register-based censuses

(i) Fully register-based census

1.222. A census is said to be fully register-based when information on all variables is obtained solely from different registers. This option allows the countries that have it implemented, to obtain information on the different census variables as frequently as they deem appropriate. Due to the advantages that this alternative provides, many countries have the medium or long-term objective of carrying out a census using this methodology.

1.223. For a successful fully register-based census, it must be verified that each and every one of the census characteristics can be obtained from registers with sufficient quality to meet users' needs. Therefore, a detailed study should be made of the information provided by these sources and contrasted with other results to verify its reliability. On the other hand, its level of coverage - which should encompass the entire population - must also be checked. Additionally, the periodic updating of information held in the registers must be carefully assessed.

(ii) Census with registers and existing sample surveys

1.224. A variant of the register-based census is when the data from population and/or integrated administrative registers are combined with those from existing sample survey(s). In this approach, no dedicated field-based data collection occurs as part of the census exercise, rather existing survey data are utilized. Different data sources are integrated as part of a "virtual census"[46] process. The data for the census exercise may be derived from many types of registers and surveys, covering different population groups throughout the country and its subregions. Compared to field-based data collection methods, this process is lower in cost and staff requirements, and uses more frequently updated data sources. More significantly, by combining data that are already available from other sources, the virtual census creates no additional respondent burden, thereby ensuring its public acceptability.

1.225. The option of a register-based census with existing sample surveys can be adopted only if all necessary census information is available from the various administrative or survey sources, and it is possible to link the information from the different sources at the unit record level. During the process of integrating individual records, care should be taken to check the accuracy of the data and remove inconsistencies prior to the production of statistical outputs.

[45] See UNECE Guidelines for assessing the quality of administrative sources for use in censuses, https://unece.org/statistics/publications/CensusAdminQuality

[46] The term "virtual census" originated in the Netherlands, where it was first developed, and here it refers to that specific methodology whereby data is derived from many types of registers and a few surveys.

1.226. However, a weakness of the use of survey data within the data collection operation is that for the variables derived from such surveys the sample size may not be sufficient to offer the geographically detailed outputs, or thematically disaggregated information with multiple variables, that are an essential feature of the census. Importantly though, the use of data from existing regular surveys will often mean that time series of data are available. By combining samples for several survey exercises, it may be possible to produce reliable estimates for small geographic areas (although there would be the perennial problem that the reference date would not be homogeneous for all the individuals).

1.227. In addition, the use of existing surveys does present a number of problems compared to the use of ad hoc surveys. The timing, content, statistical definitions and sampling approaches used in an existing survey may not be appropriate to allow the data to be readily combined with data from the administrative source(s). For example, many major household surveys are not designed to cover persons living in institutional households (such as student accommodation, hospitals, prisons and military establishments), meaning that an additional source of information is needed for these persons.

7. Advantages and challenges

1.228. The primary advantages of a register-based approach are: reduced long-term costs resulting from there being no need to conduct an extensive field operation; the opportunity for greater frequency of data, potentially on an annual basis rather than the decennial or quinquennial cycle; reducing non-response; and eliminating respondent burden. However, it should be recognised that there is always a considerable initial cost involved in setting up systems for moving from the full field enumeration to the registered-based census.

1.229. Another advantage is that the need for processing can be confined to those data items that have changed. In the long run costs are much reduced if information is collected just once and processed only if and when it changes, such as for example, when there is a change of address. Personal characteristics such as country of birth, citizenship, religion, levels of completed education and qualifications, and some housing characteristics such as period of construction, floor space and number of rooms change quite seldom if at all in some cases.

1.230. There are, however, a number of potential drawbacks and challenges with the use of administrative data sources that need to be taken into account. One limitation is that the scope of statistical topics, key definitions and, indeed, the population base depend on the information that can be compiled from the available registers. These, in turn, will be based on the underlying administrative purpose and procedures of the registers. This may impose restrictions with respect to the characteristics that are available for description and may also undermine comparability not only with data collected through full-field enumeration but also with international standards. When a registered data item is changed, new or updated information is not always registered immediately. In certain cases, new or updated information may not be registered at all. Where this occurs, the register information will not accurately reflect real circumstances. Also, the choice of new and emerging topics is limited in registered-based approach as compared to full-field enumeration which facilitates a lot more flexibility. This also applies to several census non-core topics that can be difficult to derive using administrative sources only, and to difficult-to-reach populations.

1.231. Furthermore, when using different data sources based on registers, not all of them have the same reference date; there are those that are updated continuously, weekly, monthly, etc. This can seriously affect the key element of “simultaneity”.

1.232. Depending on the amount of information available from different data sources and the quality and extent of data linkage, data processing can be more complex with these census methods than with a field-based enumeration, although good-quality results can, nevertheless, still be obtained.

1.233. The use of registers also imposes on the statistical agency a dependency on the authorities responsible for holding and maintaining the registers as well as on any changes in legislation and administrative policy and practices. The decision to use administrative sources in the statistical production process, therefore, requires close collaboration between the NSO and the administrative authorities responsible for the registers (as noted at paragraph 1.187.). During preparations for data delivery, all parties concerned must agree on date(s) of delivery and the content of the data. This implies a bilateral agreement at a high hierarchical level on a detailed data set description, scheduled delivery dates and the statistical reference period. Test data deliveries help to resolve or minimize problems with the subsequent data processing by the NSO. Validation techniques appropriate to administrative data should be applied, including checks on the plausibility, completeness and reference periods.

1.234. Another challenge of register-based censuses is that the visibility of the statistical agency may be reduced. Unlike traditional censuses, which involve actively canvassing individuals, a register-based approach that relies on existing administrative data, makes the process far less visible to the public. This can lead to a lack of awareness about the NSO's role and reputation in data collection and analysis.

1.235. Register-based information can suffer from data quality issues, such as errors, omissions, and implausible information (for example, records showing people as being implausibly old, invalid occupations, information about migration that is not consistent with other data). To address these problems, edit rules can be defined to highlight inconsistent or implausible information. If a data source has a significant number of errors, it may not be suitable for compiling census statistics. Correction or imputation of erroneous records can be attempted using other data sources or statistical techniques. Missing or implausible data can create challenges for data analysis. While deleting such cases may be tempting, it can lead to a loss of representativeness and completeness, introducing bias. Various imputation methods, including single and multiple imputation techniques, can be considered to address these issues.

1.236. Sometimes information about topics can be obtained through different data sources (registers and surveys). In this situation, it is very common that calibration techniques are used in order to reduce inconsistencies between data from such different sources. However, calibration only guarantees coherence to a certain geographic level, generally modifying or adjusting the sampling factors. It may be necessary to explain to users the reasons for any (remaining) differences. Furthermore, for the calibration results to be satisfactory, the survey must be sufficiently representative.

1.237. Overcoverage and undercoverage are common issues in administrative registers, particularly for certain population groups (see paragraph 1.174.). Overcoverage can occur when individuals are mistakenly included in the register, while undercoverage can occur when individuals are missing from the register. This can lead to inaccuracies in census data and affect the reliability of population estimates.

1.238. Households are an important unit in every census. However, identifying households from available administrative data sources in most cases is challenging, especially when the "housekeeping" concept rather than the "household dwelling" concept is used (see paragraphs 2.42-2.43). Household relationship information is not typically available in registers but could be derived by combining

information from different registers. For instance, for the purposes of identifying cohabiting couples, information such as if two persons living at the same address have a child in common, whether the persons are relatives or not, the mutual age difference, and the date of moving into the dwelling, can be used. This method makes it possible to identify cohabiting couples of opposite sex reliably, but there can still be challenges in some cases to detect cohabiting couples without common children.

1.239. Some ethnocultural characteristics such as religion and ethnicity are often difficult to gather by using administrative data sources. In many countries such information is regarded as sensitive and not suitable to hold in an administrative register.

1.240. Another challenge when conducting a register-based census can be the lack of data describing the situation before the register was established. In the census this is associated with topics such as place of birth and ever-resided abroad. There are also challenges in deriving topics related to human behaviour and attitudes, examples being mode of transport to work/school and type of energy used for heating.

C. Combined census methodologies

1.241. The combined census is a methodological approach that leverages both administrative registers and field operations to collect census data (the field operation could be either a full field enumeration or an *ad hoc* sample survey). This methodology can vary widely, with different countries adopting unique combinations of these two primary sources. Common approaches include using registers for the population count and conducting field operations for selected census topics or using a combination of registers and field operations for both the population counts and selected census variables.

1.242. The combined census is frequently used as a stepping stone strategy for NSOs aiming to move from the full field enumeration census to a fully register-based census. Compared to a register-based approach the combined census method is more expensive, and imposes more response burden on the public. On the other hand, it may help to increase the coverage and quality of statistical registers for variables collected in the field.

1.243. In recent years, it has been observed in a number of countries that:

(a) The availability and quality of the administrative registers is relatively good (at least for certain key census variables);
(b) Administrative registers, especially base registers, are being extensively geocoded with coordinates and/or addresses;
(c) Information for some census topics is not available in the administrative registers or the quality is not sufficiently high;
(d) The population generally, and certain population groups (in particular, those that are difficult to enumerate), are becoming more sensitive to the handling of personal information, and possibly more reluctant to cooperate with officialdom or more difficult to enumerate due to their high mobility or other reasons; and
(e) Administrative registers exist but the information on some topics cannot be shared with NSOs due to lack of coordination and/or legal constraints.

1.244. In these cases, a combined census that uses register(s) and questionnaire(s) could be an option. Essentially, the combined methodology makes use of registers holding data that is relevant to a census,

complemented by surveys or a full field enumeration. The use of survey and enumeration data is intended to:

(a) Improve the accuracy of the population counts;

(b) Provide information from population groups or regions where registration information may be more fragile or non-existent;

(c) Provide information for census variables that cannot be reliably based on administrative data;

(d) Check, update and improve the quality of census data derived from administrative sources;

(e) Add additional variables to the census;

(f) Provide a frame to link together unit record data from different sources;

(g) To pave a way forward towards a fully register-based census by analysing the accuracy of available multiple data sources.

1.245. The combined census methodology can thus be considered as a first-step towards a fully register-based census. It provides the opportunity to assess the accuracy, completeness of the available administrative data sources and hence, if the conditions are right, to plan for a fully register-based approach in subsequent censuses. In short, when a country is in a state of transition towards a register-based census, but may have only a limited set of records with partial information or geographic coverage, the combined census may be an appropriate option.

1.246. As noted at paragraph 1.241. , in the combined approach information on individuals, households and dwellings is collected by combining data from registers with data collected from a survey. Field data may be collected using full field enumeration, an ad hoc sample or rolling survey methods. Data from registers are employed not only as a sampling frame or to support field operations, but also directly as the data source for some census information. In a case where registers are used along with full field enumeration, data from registers may be prefilled in questionnaires, and respondents may be asked to check, update and confirm their details. Other questions relating to topics not available in the registers may also be canvassed during the field operation. When registers are used along with a sample survey *(ad hoc* or rolling survey), some census outputs may be produced entirely from the information available on the register(s), while for other outputs, information from the survey, duly weighted to the population totals, could be used. Some census tabulations may also be generated from model-based approaches where survey data and register data are used jointly (see paragraphs 1.279. -1.284. below on modelled methods).

1.247. It is worth emphasising that the results obtained from the registers will not always be consistent with the results obtained from the sample survey, even for data items relating to the same topic. To ensure coherence between the two sources, weighting techniques may be used. Additionally, sample survey can be used to evaluate the accuracy of register counts, including coverage.

1.248. The data sources should include verified and accurate personal information (name, ID number, date of birth, sex, marital status**,** family structure). In an ideal situation, a "base" register can be envisaged, to include unified identity codes for both people and address components in order to link more efficiently the related register and survey data. The link between persons and their dwellings is equally important, giving the household unit. Other administrative data sources can include tax files, social security files, public records of unemployed and registers of educational qualifications (as noted in paragraph 1.204. above). It is preferable to have a centralized base register. If this is not available, regional registers will need to be consolidated.

1.249. Data sharing protocols/procedures, agreed between the holders of administrative registers and the NSO, need to be in place. These procedures should be subject to the provisions prescribed in the legal framework and ensure the confidentiality of individual records. The census reference period should be aligned with the regular updating cycle of the registers that have to be used for producing census outputs.

1.250. The different data sources of a combined census may have overlaps in some variables and the information contained may be different in certain cases. If this happens, a decision will need to be made as to which data source is the most appropriate to use.

1. Types of combined census methodologies

1.251. There are many different approaches to the implementation of a combined census methodology. Two prominent approaches combine base registers (such as population or dwelling registers) and/or integrated administrative sources with either full enumeration or *ad hoc* sample survey.

(i) Base register[47] and/or integrated administrative sources[48] with full field enumeration

1.252. One approach can be to combine the full enumeration with a base register or registers.[49] The questionnaire used in the total enumeration may then contain fewer questions compared to a full field enumeration census questionnaire, but still covers the whole population of individuals, households and dwellings. Over time, countries may decide to adopt this model, increasing the use of integrated administrative registers and reducing the number of questions in the questionnaire in successive censuses. In any case, so as to carry out a census that is as efficient as possible, NSOs should consider that the information that can be obtained with sufficient quality from the administrative registers need then not be part of the census questionnaire. The registers can be used to prefill such information as name, address, education, occupation and dwelling characteristics on the questionnaire. The questionnaire can then be used to ask the respondents if the prefilled information is still valid or needs to be changed.

(ii) Base register and/or integrated administrative sources with ad hoc sample survey

[47] Examples of base registers could be population register and/or dwelling register; and address register and/or business register.

[48] Examples of integrated administrative sources could be administrative or statistical registers of education and/or occupation.

[49] In Europe in the 2010 round of censuses, four countries applied this approach: Estonia, Latvia, Liechtenstein and Lithuania. For more details, see *Measuring population and housing: practices of UNECE countries in the 2010 round of censuses,* United Nations Economic Commission for Europe, United Nations, New York and Geneva, 2014, page 12.

1.253. Another model involves the use of an *ad hoc* sample survey instead of a full enumeration.[50] The base register could then be used as a sample frame and also to prefill some information such as name and address on the questionnaire. The sample can be sized and stratified in such a way that data are available for small population groups and geographic areas.

1.254. This model requires far fewer enumerators than a full enumeration, so a more specific training operation with skilled and prepared professional interviewers can be carried out. The follow-up of the operation is also simpler. There is no need to obtain information from each member of the population, giving a clear reduction in response burden. Non-response can be corrected in the sample by the use of statistical techniques to ensure information is still representative of the total population. In any case, it is important to pay attention to ensuring that all traditionally difficult-to-count groups are also correctly represented in the survey responses.

1.255. By using a sample and a much smaller number of enumerators, a significant reduction in field costs is possible, compared with the full field enumeration approach.

1.256. The larger the size of the survey, the greater is the level of geographic and thematic detail of the information that can be produced, but the higher the costs of the operation will be. The extent to which this model can produce detailed statistics (in particular for municipalities or smaller geographic areas) will depend on the size of the sample survey. Even though a larger sample should allow more geographically detailed statistics to be produced, the level of such detailed information is likely to be lower than if a comprehensive approach were taken – such as with a full field enumeration census or fully register-based exercise.

2. Advantages and challenges

1.257. The main advantage with using the full enumeration with a base register or registers model is the reduction in response burden for respondents and the overall cost of the census compared to a full field enumeration. This model also allows the preparation of small-area statistics as all variables are collected as total counts. However, this model will still involve a large data collection exercise with the use of enumerators. While self-enumeration data collection may be utilized, many respondents may still require follow-up and assistance in completing the questionnaire. This model demands extensive planning for data collection, as it requires responses from the entire population.

1.258. The alternative model of register-based census combined with *ad hoc* sample survey offers several advantages:

(a) It is more cost-effective than a full field enumeration census collecting all census items from the whole population;

(b) It reduces the burden on both enumerators and respondents;

(c) It decreases non-response when information is obtained from registers;

[50] In the 2010 round of censuses, some of the countries that applied this approach included Germany, Israel, Poland, Spain, Switzerland and Turkey. For more details, see *Measuring population and housing: practices of UNECE countries in the 2010 round of censuses,* United Nations Economic Commission for Europe, United Nations, New York and Geneva, 2014, page 12.

(d) It allows for more effective imputation of missing survey data by leveraging information from administrative records to adjust for varying levels of non-response in different population groups;

(e) It provides greater flexibility to introduce questions on new and emerging topics.

1.259. Data integrated at the unit record level might be expected to provide very reliable results, because they are based on a maximum amount of information. The coverage of subpopulations may be more reliable because when data are missing in one source, another source may be used. Another advantage of micro-integration is that there will be less reason for confusion among statistics users. For example, there will be one figure on each socioeconomic phenomenon, instead of several different figures depending on which sources have been used.

1.260. One of the primary challenges of combined censuses that use sample surveys is the increased complexity involved in producing output tabulations. Weighting problems may arise, and statistical modelling may be necessary when integrating survey data with register data. Additionally, these censuses may lack the high public profile and publicity of traditional full field enumeration censuses, potentially leading to less interest in, and use of, the census results.

1.261. Finally, this type of census cannot have the same level of thematic and geographic detail as a full field enumeration census, and this may require giving explanations to the different census users. A challenge arises when the reference dates of the register-based data differ significantly from the census reference period, determined by the timing of the sample survey. Such discrepancies can make efficient data integration with administrative registers challenging.

1.262. Data validation, processing and dissemination may be more complicated, as this approach involves both total counts based on the register and sample data from survey. In addition, as some variables are based only on sample data, it may be impossible to meet the level of statistical and geographic detail required in some tables. It is advisable to be very didactic with census users: explain to them in a transparent way how the census results are obtained and the limitations they may have, which are similar to those involving the use of short and long census forms. On the other hand, the possibilities of reducing cost and response burden provide a very strong reason to adopt this approach.

1.263. Some of the required variables will need to be constructed from different sources. The census results obtained may differ to some extent from those that would be obtained from a full enumeration covering all census topics. This may have a negative impact on the comparability of results between countries and over time. An advantage of registers is that, in effect, they offer complete coverage subject to the quality of the data contained. It is preferable that statistical authorities make full use of the register data that are available.

1.264. For the combined census method (as with the full field enumeration approach), a number of different modes can be employed to collect information in the field, including paper and/or online questionnaires, handheld devices and telephone interviews. Electronic devices have important advantages that influence the quality of the information obtained: validation controls can be included in the different questions, time taken to respond to the questions is reduced and the analysis and dissemination of information is faster. Where data are collected via different routes (such as Internet collection in parallel with face-to-face interviews), controls are needed to avoid duplication of information. Employing different modes to collect information from different households may introduce some bias in the responses.

1.265. For the field data collection component of combined census methodologies, it is important to store control information and indicators at the lowest geographic level available in the central database, covering issues related to the census operation, such as progress with the fieldwork, response rates and comparisons with information in registers. This control information – normally based on web reports and analysed daily by project managers or regional offices – can be used to detect problems that appear during the fieldwork and to plan necessary actions to overcome these problems. If data are analysed in real-time or on a daily basis, close monitoring of the continuing field operation and data entry is possible. See Part III, Chapter VII, Section E for more elaboration on the management and supervision of census field operations.

V. Methods for estimating population for areas not enumerated in the census

1.266. Conducting a census can be disrupted by unforeseen circumstances such as natural disasters or security risks, hindering field operations in specific areas. In such scenarios, various approaches, including administrative data sources and existing sample surveys can be used to estimate population numbers and densities for these inaccessible regions. It has to be noted that the estimates generated through these approaches should not be considered as equivalent to census data. When incorporating population or housing estimates into census data, it is essential to clearly distinguish them from observed census counts, allowing users to differentiate between the two sources during analysis. It is also important to emphasize that though such estimates can provide valuable approximations and insights they are subject to inherent limitations and errors. Therefore, accuracy requirements should be carefully considered when selecting an estimation method.

1.267. Beyond traditional methods, innovative techniques are emerging for population estimation in inaccessible areas. One such example is the use of mobile positioning data to estimate daytime population distribution. As technology advances, NSOs can expect even more precise and efficient methods for handling these situations.

1.268. While complete coverage during a census is an ideal goal, achieving it can be challenging. Factors such as low response rates in self-enumeration methods, limited administrative data coverage, or security issues during field enumeration can all contribute to incomplete data in specific areas. In such cases, estimation approaches can be employed to fill these data gaps, but it is crucial to acknowledge them as estimates and distinct from actual counts. The inherent error associated with estimation methods needs to be considered when making this choice.

1.269. Estimation methods are often used to address data gaps at smaller geographical levels such as districts, municipalities, or villages. It is important to remember that these "small area estimates" are not replacements for actual data and may differ from reality. Therefore, accuracy requirements must be a key factor when selecting an appropriate estimation method.

A. Requirements for estimating population

1.270. In order to generate reasonably accurate population data for uncounted areas during a census, careful consideration is needed to select the appropriate estimation method. Understanding the acceptable level of error is crucial. The methodology used for estimation should be clearly documented

and accessible to users. The chosen method(s) should rely on data sources that are readily available, reliable, and of sufficient quality. Balancing accuracy needs with resource constraints is important.

1.271. Some key requirements for using estimation methods for parts of a country that are not possible to enumerate during a census include (based on the specific requirements by method):

i) Survey data: Existing survey data should be based on a high-quality sampling frame, ideally derived from a recent census or administrative data. The chosen sampling method (for example, simple random sampling, stratified sampling, or cluster sampling) should be appropriate for the target population and context. The survey questionnaire should be carefully designed to address the key data needs for population estimation and should have strived for a high response rate.
ii) Administrative data: The administrative data sources used should be accurate, complete, and up-to-date and should cover the target population in the inaccessible area as comprehensively as possible. The ability to link administrative data records to geographic locations is also essential for spatial analysis.
iii) Combination methods: The chosen data sources should be compatible and allow for meaningful integration, while robust statistical methods are required to combine data from different sources and produce reliable estimates.
iv) Geospatial information: A high-quality geographic framework, such as digital maps and boundaries, satellite imagery, and modeling techniques, is needed for spatial analysis to support population estimation, particularly in geographically remote areas.

B. Methodological approaches

1. Use of administrative data sources

1.272. Administrative data can be used to estimate the population not enumerated in a census in a number of ways, for example through data from birth records, death records, change of address, and school enrolment records. Migration is likely to account for a significant element of population change, therefore, the appropriate administrative sources, such as a register of foreigners, may be used to provide relevant data.

1.273. Administrative records can be used to estimate total number of housing units, and the proportion of housing units which are occupied, in areas where individual enumeration is not possible. Administrative records to consider for this approach may include electricity and water consumption, telecom services, property registrations and cadastral information.

1.274. Estimation of the total resident population in parts of the territory not covered by the main census can be achieved with the Housing-Unit Ratio Method[51] or one of its variants. This method requires separate estimates of the total number of housing units, the proportion of housing units which are occupied, and the average number of residents per occupied housing unit. Administrative data can be used to estimate the first two components, while estimates of the latter are likely to be sourced from

[51] Methods and Materials of Demography 2nd ed, Siegel and Swanson 2010, pp 550-551.

previous surveys and censuses. The average number of residents per occupied housing unit is generally equivalent to the average household size when using the 'household dwelling' concept and is true for most households when the 'housekeeping' concept is applied. However, specific circumstances may lead to minor discrepancies.

1.275. It is important to note that administrative data of insufficient coverage may not be a viable substitute for a census. Administrative data sets may not be complete or accurate, and they may not cover all members of the population. However, administrative data can be used to improve the accuracy of population estimates, and to identify populations that are at risk of being undercounted in the census. Despite these challenges, administrative data can be a valuable tool for estimating the population not enumerated in a census. By carefully considering the strengths and weaknesses of administrative data, it is possible to develop accurate and reliable estimates of the total population.

2. Use of survey data

1.276. When a census cannot reach certain areas due to disasters or security issues, existing sample surveys can help estimate the population and its characteristics. However, it's crucial to acknowledge the limitations and ensure the chosen survey data is suitable for this purpose.

1.277. Using existing survey data to estimate populations in inaccessible areas requires careful selection and analysis. The chosen survey must have collected information comparable to a census, including demographics, household characteristics, and socioeconomic factors. It would be important to critically evaluate the survey's sampling method to ensure it accurately represents the diversity of the inaccessible population. This is crucial for generating reliable estimates. Furthermore, it would be important to ensure the data has been analysed using appropriate statistical techniques, including proper weighting methods to "gross-up" the sample data and accurately represent the entire inaccessible population. This step ensures the estimates are as accurate and representative as possible.

1.278. It's important to note that estimates based on sample surveys aren't perfect. They have a margin of error, meaning they might not precisely match the true population characteristics. Smaller sample sizes generally lead to larger potential errors. The accuracy of these estimates depends heavily on how well the sampling method captures the diversity of the population being studied. By recognizing these limitations and using robust sampling techniques, statisticians can still produce reasonably accurate population data.

3. Modelled methods

1.279. In contexts where a population and housing census cannot be fully executed in all locations of a given territory, a modelled approach can be adopted to produce population estimates for small areas or for uniform grids with relatively small dimensions. Such an approach relies on the availability of data from previous censuses or surveys (complete population counts within small, defined areas, a partial census, a census-like population survey, or specifically designed micro-census surveys undertaken where it is possible to conduct fieldwork) and detailed spatial data.

1.280. Statistical models can then be used to link the population data to spatial data for the regions of interest to estimate population and housing characteristics in locations where census enumeration is not possible. In estimating population characteristics, the first step is to establish relationships between satellite image features (such as building density and land use patterns) and the population data collected

through censuses or surveys. The next step is to create population density maps by assigning population values to different land cover types or building clusters based on the established correlations. By aggregating these high-resolution predictions, population totals can also be produced for administrative units or for the national level if required.

1.281. Imageries from satellites or unmanned aerial vehicles (drones) offer a powerful tool for assessing population and housing characteristics. Such imageries are able to identify, for example, land cover types such as urban areas, agricultural lands, forests, water bodies, building heights, land slopes, and vegetation density. Using object-based image analysis, individual structures (such as buildings and roads) can be extracted from the imagery using segmentation algorithms. Satellite imageries are particularly useful for identifying housing characteristics and counting individual housing units. There are three main steps in using satellite imagery for estimating housing characteristics:

i) conducting building footprint analysis through measuring the size and shape of buildings to assess housing unit sizes and types;
ii) ii) determining the dominant roof materials (e.g., metal, tile, thatch) to infer socioeconomic status and construction practices; and
iii) iii) evaluating the presence and condition of infrastructure elements like roads, water supply, and sanitation facilities to assess the quality of housing.

1.282. The success of this approach relies on the identification of relevant geospatial covariates and the definition of functional relationships between geospatial datasets and population density to predict population numbers in locations where it is not possible to conduct census enumeration. It also requires robust geo-referencing of the geographical areas where population data are collected, and a design that captures as best as possible the range of population densities, demographics, and environmental dimensions that exist across the regions of interest. The covariates need to be strongly correlated to population density and available consistently across all areas where population estimation is required. Although access to high-quality and spatially comprehensive datasets has traditionally been difficult in resource-poor settings, the increasing availability of very high-resolution satellite imagery makes the production of high-quality covariates increasingly feasible for many settings.

1.283. Although the production of such modelled population estimates can never replace the rich data on individuals, families, households, or communities generated by a full-scale population and housing census, and should, therefore, never be disseminated as census data, their high level of geographical disaggregation makes them useful in contexts where census enumeration is not possible.

1.284. Small Area Estimation techniques can also be used to estimate population in areas not enumerated in the census (see also paragraphs 1.45. -1.46.). These techniques utilize statistical models to "borrow strength" from neighbouring areas or auxiliary data sources, such as administrative records or satellite imagery, to generate reliable population estimates for smaller geographic areas. By combining information from different sources, such techniques can provide valuable insights into population dynamics in areas where it is not possible to conduct census enumeration.

VI. Strategies for producing more frequent population and housing statistics

1.285. There are a number of alternative approaches that could be used for producing more frequent population statistics such as: intercensal or annual surveys for countries conducting a census with the full

field enumeration method, or use of other sample surveys; administrative data sources; geospatial information; or a combination of such approaches, for example, administrative records combined with sample surveys. The best approach for producing more frequent population statistics will vary depending on the specific needs of the country or region. However, all of these approaches have the potential to improve the timeliness and accuracy of population statistics. In addition to these approaches, countries are also increasingly exploring the use of new technologies, and alternative data sources, to produce population statistics. These sources can revolutionize the way population statistics are collected and produced, and make population statistics more frequent, more accurate, and more widely accessible.

A. The need for more frequent data on population and housing characteristics

1.286. There are a number of reasons why there is a need for more frequent data on population and housing characteristics. Some of these reasons include:

(i) **Planning government programs and services**: Timely and accurate data on population and housing characteristics allows governments to plan and deliver effective programmes, such as the building of new schools and hospitals based on population growth and distribution, or the provision of rental assistance based on housing affordability.

(ii) **Allocating resources**: Population and housing data guide fair and efficient resource allocation, such as directing funds to low-income communities based on poverty rates or to schools serving those communities based on educational attainment.

(iii) **Conducting research**: Population and housing data are crucial for researchers studying economic trends, social inequality, and public health; for instance, income and employment data are needed to analyze policy impacts, while ethno-cultural data can examine levels of discrimination.

1.287. In addition to these general reasons, there are also a number of specific reasons why more frequent data on population and housing characteristics are needed. For example, the COVID-19 pandemic has highlighted the need for more timely data on population health and housing conditions. The pandemic has also shown how quickly some population characteristics can change, making it more important than ever to collect data on a regular basis. More frequent data can be used to respond to unexpected events, such as natural disasters or economic crises. For example, data on population distribution can be used to target resources to areas that have been affected by a flooding or earthquakes.

1.288. More frequent data can also be used to identify emerging trends in population and housing characteristics. For example, data on housing affordability can be used to identify areas where housing prices are rising rapidly or where people are at risk of displacement and homelessness.

1.289. Overall, there is a strong need for more frequent data on population and housing characteristics that are essential tool for making informed decisions about the future. These data are essential for planning and delivering government programmes and services, allocating resources fairly and efficiently, and conducting research on a wide range of topics. By investing in more frequent data collection, governments and researchers can ensure that they have the information they need to make informed decisions that benefit everyone.

B. Use of administrative registers

1.290. As mentioned in previous sections, administrative data sources offer a promising avenue for producing more frequent population and housing statistics. However, there are preconditions that should

be met. First and foremost is the availability of quality administrative data source. There should be mechanisms in place that guarantee long term access to and use of administrative data sources by the NSO for statistical purposes. The use of administrative data can be as limited as the use of birth, death and migration data to update the total population count at geographical levels; or can be more extensive in the event that quality administrative data sources on other topics, such as education, employment and health, are also available. The necessary pre-conditions for the use of such data sources for the purpose of the census are discussed in more detail in paragraphs 1.177. -1.197. .

C. Intercensal and continuous surveys

1.291. Intercensal surveys can be used to produce more frequent population and housing statistics by collecting data from a sample of the population between censuses.

1.292. Intercensal surveys serve as valuable complements to full field enumeration censuses, offering several advantages: they can collect data on a wider range of topics, providing more in-depth information on specific aspects of the population (such as levels of disability for example); and, they can produce more timely data, allowing for more frequent updates and analysis. While intercensal surveys offer these benefits, it is important to note that they are not intended to replace full censuses but rather to provide additional insights and data points between decennial censuses.

1.293. However, intercensal surveys also present some challenges: they are subject to sampling error, which means that the estimates produced from the survey may not be as accurate as the estimates produced from a census; they may not be able to reach all segments of the population, which can lead to bias in the results; and they may not be able to provide accurate estimates at lower geographic levels.

1.294. Despite these disadvantages, intercensal surveys can be a valuable tool for producing more frequent population and housing statistics. By using such surveys, countries can produce more timely and accurate data on the population, which can be used to inform decision-making on a range of issues. The intercensal survey combined with administrative data on births, deaths and migration can be used to update the overall snapshot picture provided by the decennial census, even if this picture has a lower resolution than that of the census.

1.295. Continuous annual surveys involve collecting data from a representative sample of the population on a year-round basis, continuously collecting data on demographics and other characteristics. The same sample may be used throughout the continuous cycle, or a varying sample may be employed, as in the case of the rolling census. By collecting data throughout the year, such an approach to data collection allows better tracking of constantly evolving trends, identification of emerging issues, and measurement of the impact of policies in a timelier manner.

1.296. Continuous annual surveys offer further advantages over full field enumeration census methods. Firstly, they provide a significant enhancement in timeliness. By collecting data throughout the year, they reflect the most recent changes in the population, making them more suitable for tracking rapidly evolving social and economic trends. Secondly, continuous surveys enable a more granular analysis of trends. Data collected over shorter time periods allow users to not only see the overall direction of change, but also the pace at which things are evolving. This can be crucial for identifying emerging issues early on. Finally, since they collect data regularly throughout year they can help the NSO assess the effectiveness of recently implemented policies and guide any necessary adjustments. Annual continuous surveys also provide an opportunity to occasionally add specific modules to reflect emerging data needs.

1.297. However, continuous annual surveys also come with their own set of challenges. Maintaining the same representative sample throughout the year can be more complex and potentially more expensive than conducting a single large-scale census survey. Additionally, continuous surveys might face lower response rates compared to high-profile decennial censuses. Survey fatigue and a perception of less importance among respondents can contribute to this challenge. Finally, ensuring data comparability across different years is crucial for valid trend analysis. Even minor changes in how questions are asked, or data are collected can make it difficult to accurately compare results over time.

1.298. Despite the challenges, both intercensal and continuous annual surveys offer a valuable approach to data collection, providing a more dynamic and up-to-date understanding of populations dynamics. By combining these methods with traditional census data, users can gain a richer and more nuanced picture of an ever-changing world.

D. Use of mobile phone data

1.299. Mobile phone data has emerged as a promising source for generating more frequent and timely population statistics, particularly in countries with high mobile phone penetration. This data can help bridge the gaps between censuses, providing valuable insights into population dynamics. The advantages of these data are that often there is substantial coverage, potential for high accuracy, and minimal effort needed on the part of individuals to share the data.

1.300. Mobile phone data, particularly call detail records and location/positioning data, offer unique insights into population mobility and daytime population statistics. By analysing anonymized and aggregated data from mobile networks, NSOs can produce granular data on population movements, thereby identifying patterns in commuting, international temporary mobility, and other mobility trends. These data can be used for urban planning, emergency response, and transportation management, providing a dynamic picture of population distribution that static census data cannot offer.

1.301. The integration of mobile data with traditional data sources can improve the accuracy of resident population estimates. Mobile phone data can serve as a secondary data source in "signs of life" methods[52], which are used to validate and update register-based population data (see also footnote 36 and paragraph 1.206.). These methods help identify active users within a geographic area, providing an additional layer of verification and enhancing the reliability of population registers.

1.302. Mobile phone data could also be used to validate data from other sources and collection exercises. For instance, during household surveys or other field operations, discrepancies between collected data and mobile phone data can highlight areas that require further investigation. This triangulation of data sources can enhance the overall quality and robustness of population statistics.

[52] Bernardini, A., Chieppa, A., & Tamburrano, T. (2024). Discovering individual profiles from administrative signs of life useful for the estimation of census results. Rivista Italiana di Economia Demografia e Statistica, 78(1), 1-20. https://www.researchgate.net/publication/379079413

1.303. Despite its potential, the use of mobile phone data in official statistics is still at the experimental and innovative stages and comes with its own challenges. Access to mobile phone data is a significant barrier, often requiring negotiations with telecom operators and adherence to strict privacy regulations. Additionally, analysing mobile phone data requires specialized skills and technology that may not be readily available in many statistical offices. Investments in technological infrastructure and capacity building are essential to establish the necessary analytical capabilities.

1.304. Privacy and ethical considerations also play a critical role in the use of mobile phone data. Ensuring (and demonstrating) that data are anonymized and aggregated to protect individual privacy is paramount. NSOs must establish robust data governance frameworks to manage these concerns while maintaining public trust.

1.305. It is also important to keep in mind the limitations of mobile phone data with regard to representativity, particularly in regions with low mobile phone penetration or among demographic groups less likely to use mobile phones. Besides low penetration/under-coverage for selected groups, there could also be duplicated coverage issues for mobile data. Addressing these biases is crucial for ensuring the accuracy and reliability of the derived statistics.

PART TWO. PLANNING, ORGANIZATION AND MANAGEMENT

I. Introduction

2.1. Part Two of these *Principles and Recommendations* elaborates on the planning, organization and management of the population and housing census. This is a unique and most complex statistical exercise requiring multidimensional considerations during both the preparation and operation phases, regardless of the census methodology (e.g., full field enumeration, register-based or combined approaches) or the type of population base (e.g., usual resident or present population) adopted in the census.

II. Census planning

A. Overall census planning

2.2. A population and housing census (or a population census on its own) consists of a complex series of interrelated activities, often representing one of the largest, most intricate, and costly operations a country undertakes. Some activities, such as the mobilization of enumerators in the field and procurement of hand-held devices (tablets), may be massive in scale, while others, such as the training of supervisory staff, must be carried out uniformly in all parts of the country. For fully register-based or combined censuses, significant efforts are required to gather and merge administrative data sources, particularly during transitions from field-based to register-based approach. Additionally, since censuses generally take place at intervals of five to ten years, planning for each new census must consider evolving external conditions, new census methodologies and topics, technological advancement (such as web-based self-response mode), stakeholders' requirements, changes in personnel condition, as well as developing administrative frameworks.

2.3. To ensure that the various operations occur in the proper sequence and in a timely manner, the entire process, along with its individual components, must be carefully planned in advance. Even an apparently minor oversight in planning may lead to serious defects in census results and costly inefficiencies during operations. Careful planning is therefore critically important for a successful census, not only in countries with limited statistical capacity, but also in those with a well-developed statistical system. In addition, appropriate organizational and administrative arrangements are necessary. Such arrangements ensure that the extensive human and material resources that need to be mobilized for the census are used effectively and efficiently, and that the strict timelines and massive logistic requirements are met. For register-based censuses, which normally require fewer human resources, scaling up the census may not constitute a risk compared to a field-based census. However, other challenges must be addressed through effective and appropriate administration, such as preserving institutional memory to manage the complex processes involved.

2.4. It is important that at each stage of census planning and implementation, the various administrative arrangements should be developed with sound technical considerations.[53] The quality of census data – both in terms of accuracy and timeliness - hinges on comprehensive consideration of a diverse range of subject matters and statistical requirements throughout the process. This is especially valid for cross-cutting issues, such as information technology, which play a crucial role in many key phases of the census. For these reasons, the management of any large statistical operation, but particularly a population and housing census, should not be considered a routine administrative assignment.[54]

2.5. While not all censuses follow a uniform pattern, there are certain major elements that must be taken into account in every census operation. In general, census operations can be divided into the eight phases outlined in the Generalized Statistical Business Process Model (GSBPM),[55] as illustrated in Figure 2.1:

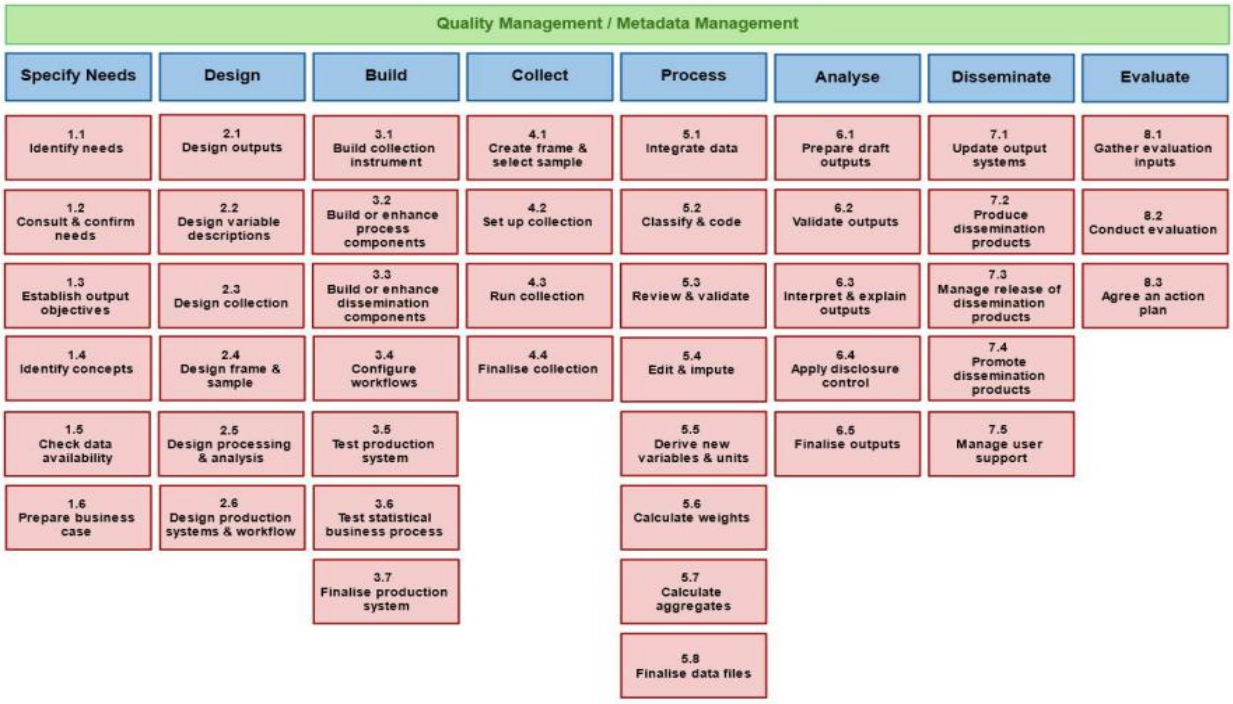
- **Specify needs**: Consultations take place, and objectives are set;
- **Design**: Outputs, data collection, processing and production systems are designed;
- **Build**: Census systems and processes are developed and tested;
- **Collect**: Field operations occur and/or administrative data are received;
- **Process**: Data are integrated, classified, edited, imputed and weighted for aggregates;
- **Analyse**: Data are validated, analysed, and finalized subject to disclosure control;
- **Disseminate**: Outputs are produced and published with user support; and finally
- **Evaluate**: Lessons learned are documented for consideration in the next census cycle.

Figure 2.1. Generalized Statistical Business Process Model

[53]For recommendations on administrative arrangements and management structures, see more details in the Handbook on the Management of Population and Housing Census: Revision 2, https://unstats.un.org/unsd/publication/seriesf/series_f83rev2en.pdf

[54] For general discussions on statistical management, see the Handbook on Management and Organization of National Statistical Systems, Fourth Edition, https://unstats.un.org/capacity-development/handbook/Handbook_20230417.pdf

[55] United Nations Economic Commission for Europe (UNECE), on behalf of the International Statistical Community, *Generic Statistical Business Process Model (GSBPM)*, Version 5.1 (January 2019).



2.6. Quality assurance measures should be applied throughout all phases to ensure that the outputs meet the expectation of census data users. Additionally, NSOs should continuously evaluate and document their experiences so that a thorough set of lessons learned can be discussed at the Evaluate phase. It is important to note that these phases may not always follow a strict sequence or be entirely separate. For example, some census results might be released before all data-processing activities are completed. Similarly, the analysis and dissemination phases often overlap quite extensively, and the systematic evaluation of the census may occur in stages, alongside other GSBPM phases. Furthermore, certain elements, such as the initial budget or staff requirements, may need to be adjusted as new circumstances arise during the later stages of the census operation. The details of each phase and their implication for effective census management are discussed in the following section.

B. Planning for both a population and housing census

2.7. As noted at paragraph 1.61. , a housing census may be conducted separately from a population census, or the two can be integrated into a single operation. While a joint population and housing census is more complex, it is more cost-effective than conducting both censuses independently, even when efforts are made to minimize additional costs. If conducted separately, there may be added expenses and challenges for training, resource mobilization, logistics, procurement, communication and data processing, among other operations. In addition to the financial and logistical advantages, amalgamating both a population and housing census will provide richer, more comprehensive data than when conducting each census separately. In assessing whether to integrate the two censuses, or conduct them separately, each country should carefully weigh the trade-offs based on its specific needs and circumstances (see also paragraphs 1.61. –1.68.).

C. Special considerations when planning a combined or register-based census

2.8. As described in Part One Chapter IV, some countries are transitioning from field-based to a combined or fully register-based approach. When considering the use of administrative registers in a census, a comprehensive evaluation of the benefits and challenges is essential. Such a transition is complex and may carry significant risks, so all aspects must be thoroughly studied. This assessment should be based on criteria established from international experience, such as those outlined in the *Handbook on Registers-based Population and Housing Censuses*[56]. (See Sections IV.B and IV.C in Part 1 for a full analysis of the merits of both these data collection methodologies.)

2.9. For any base register or administrative data source used in a census, it is advisable to select those that are established as a legal requirement in the country. NSOs planning for a combined or fully register-based census should assess their legal frameworks to ensure that they have the authority to use administrative information for statistical purposes.

2.10. The suitability of data must be evaluated, taking into account definitions, concepts, content, coverage, reference dates and other criteria. Birth and death records are the most important sources for the creation and maintenance of register-based census, and these, along with all other sources to update the base register, should be available in digital format. To link records from different registers, the existence or creation of a unique identification number for each individual, household and/or dwelling is essential. In addition, modelling techniques may be required to integrate and select the best available information from various administrative data sources. A well-defined methodology for linking different registers and maintaining them should be prescribed, including a plan for the frequency of updates. Transparency in these methodologies is crucial, and clear communication to census data users about data sources, linkage processes, and any modelling techniques employed is essential to build trust and understanding.

2.11. It is important to thoroughly evaluate public acceptance of the large-scale use of administrative data when assessing the feasibility of the transition. Some portion of the population may have concerns about privacy and the implications of using administrative data for population enumeration and characterization. Therefore, NSO should ensure that it has the legal authority and technical capacity to use large-scale administrative data securely, guaranteeing the safety and confidentiality of the data in its holdings.

III. Strategic objectives and management

A. Strategic objectives

2.12. The development of census plans should begin with setting strategic objectives. These can guide the implementation of the plans, establish standards and form a set of benchmarks for assessing the success of the census. Ideally, the process of developing these objectives would involve combining insights and lesson learned from previous census experiences, understanding stakeholders' information needs and requirements, and assessing societal and technological changes. However, in practice, some of this

[56] United Nations (2022). Handbook on Registers-based Population and Housing Censuses (Version: December 2022). https://unstats.un.org/unsd/demographic-social/publication/handbook-registers-phc.pdf

information may be difficult to obtain and often provides conflicting guidance. Nevertheless, these objectives can be used to assist in planning major elements of the census process. Although the strategic objectives will vary across countries, depending on local circumstances and priorities, they generally fall under the following categories:

(a) Census content;
(b) Impact on the public and census staff;
(c) Production of census results;
(d) Cost-effectiveness; and
(e) Cost-benefit.

The cost, benefits and ability to harness technology must align with each country's overall context.

2.13. *Census content.* The objective is to ensure that a census is relevant, with topics that meet the clear needs of users and stakeholders (see paragraph 2.170 for key stakeholders in the census process). This must be balanced against considerations of cost-effectiveness, available human resources, time constraints, and minimizing respondents' burden. The content of a census should be determined by:

(a) Stakeholders' needs through appropriate consultations at all stages;
(b) Measurable quality standards that reflect stakeholders' needs;
(c) Adequate testing of new topics to ensure reliable census outputs;
(d) Outputs' comparability at international, regional and subnational level; and
(e) Suitability of the census as the data source in comparison to other potential sources.

2.14. The census may serve as the backbone of a broader programme of socio-economic statistics. Therefore, the content objectives must consider similar data from other sources and determine where census content best fits within the overall framework of a nation's socio-economic statistics.

2.15. Furthermore, since administrative data often uses different definitions than those used for census purposes, these differences should be carefully considered when designing related household surveys. While full harmonization might not be feasible, NSOs should still strive to collaborate with administrative data providers to establish clear conceptual guidelines, including agreed-upon definitions, classifications, and units of measure, to the extent possible. This will help improve the consistency and comparability of data used for statistical purposes.

2.16. *Impact on the public and on census staff.* The objective is to ensure that all aspects of the field enumeration operations, administrative data intake and use, as well as the dissemination of results, are acceptable to the public and fully comply with legal and ethical standards for protecting the confidentiality of individual information. The public should be fully informed about the objectives, contents and methods of the census, and their rights and obligations with respect to the census. Similarly, all census staff must fully understand their responsibilities. Subsidiary objectives include:

(a) Keeping completed forms, accessed administrative records and other records containing personal information secure and confidential;
(b) Ensuring strong public support for all aspects of the census, regardless of census approaches; and
(c) Producing requested customized outputs while preventing the disclosure of personal information, adhering to established quality and privacy standards, and implementing policies that safeguard access to census products by various stakeholders.

2.17. *Production of good quality census results.* The objective is to deliver census products and services that meet legal obligations and stakeholders' needs, while adhering to defined quality standards and timelines. Subsidiary objectives include:

(a) Producing timely outputs with minimal error, ensuing that the data are fit for their intended uses;
(b) Providing standard outputs for the main results and services for customized output;
(c) Ensuring timely and universal access to outputs, while safeguarding statistical confidentiality;
(d) Using appropriate geographic bases for collecting and referencing data for output;
(e) Improving enumeration methods, whether field-based or register-based, is crucial for producing high-quality census results. Focusing on difficult-to-reach areas and minimizing undercoverage, overcoverage, and response errors are key. Employing a variety of data collection methods, such as personal interviews or self-response options (paper or web), allows for flexibility in addressing regional or other challenges during census collection, ultimately contributing to more accurate and reliable results;
(f) Improving methods for evaluating census outputs and effectively communicating them to users; and
(g) Developing measures for assessing coverage, quality and targets.

2.18. *Cost-effectiveness*. The objective is to minimize costs while achieving the main strategic objectives. Subsidiary objectives include:

(a) Adopting more efficient and technologically advanced approaches for data collection, capture and processing;
(b) Contracting out appropriate parts of the operation;
(c) Exploring alternative funding sources and, if appropriate, developing proposals for cost recovery and income generation;
(d) International, regional and in-country collaboration and coordination to co-purchase and share equipment (such as tablets) and expert services (such as mapping and GIS experts);
(e) Encouraging the public to self-respond online or on paper, where required;
(f) Using administrative data through linking different registers; and
(g) Exploring shorter periodic production of census results (for register-based censuses).

2.19. *Cost–benefit.* The objective is to increase the value or benefits of the census while managing its overall cost. A cost-benefit analysis should also highlight the costs and consequences of not having the necessary information and/or the cost of obtaining the same information from an alternative source. This requires clearly articulated benefits that arise from census activities and outputs, along with plans to realise these benefits. Some benefits can be quantified, such as those resulting from the use of census data, while others may be more difficult to measure but are still important and should be documented. Subsidiary objectives include illustrating the value of the census as:

(a) an educational tool and framework;
(b) a resource for comparative studies at national and international level; and
(c) a cornerstone of the national statistical system.

The results of this cost-benefit analysis are helpful for effectively advocating the benefits of the census, mobilizing resources, and engaging respondents. By demonstrating the clear value and return on investment of the census, it is possible to secure the necessary support and funding for future operations while also building public trust and participation.

2.20. These objectives can serve as benchmarks to assess user requirements. They may also be built into appraisal systems to compare and evaluate different options with appropriate weighting. In general,

the strategic objectives of the population and housing census must be clearly emphasized throughout the entire process – from preparation and execution to the utilization of census data.

B. Strategic management

2.21. The primary value of strategic management is to assist census organizations to operate successfully in a dynamic, complex and uncertain environment. Crafting a census management strategy is critical for a successful execution, and requires a wide range of knowledge and skills across different areas. It must be governed by a systematic process to ensure that all relevant issues are carefully examined and considered, influencing the future direction of the programme. NSOs may consider adopting the following strategic management process to guide operations in all phases of the census.

2.22. The strategic management process consists of four phases:

(a) Strategy analysis;
(b) Strategy formulation;
(c) Strategy implementation; and
(d) Strategy monitoring and review.

2.23. *Strategy analysis* involves identifying the issues that need to be addressed to take corrective action or to chart a new direction. It seeks to transform the organizational setup to one that is systemic, holistic, comprehensive, and coordinated. Strategy analysis includes:

(a) Setting the direction for the census operation by reaffirming its purpose and determine how it should conduct its business;
(b) Setting strategic goals by defining what the system aims to achieve in terms of its highest goals and outcomes. These strategic outcomes must be aligned to users' expectation and address their needs. A strategy driven by outcomes requires "planning backwards" – starting from the desired outcomes and determining the best way to achieve them; and,
(c) Conducting a strategic analysis of census operations by assessing the current and previous situation, and identifying key issues that need attention. This includes analysing the internal and external environments to provide the evidence base for developing the strategy. Census organizations may consider using a SWOT (strengths, weaknesses, opportunities, threats) analysis to do a strategic analysis of the internal and external environment.

2.24. *Strategy formulation* involves defining how and where the census organization must respond. Strategy formulation includes:

(a) Developing and reviewing the value chain;
(b) Developing strategic objectives;
(c) Compiling a strategy map;
(d) Defining the strategic intent;
(e) Identifying critical success factors;
(f) Identifying strategic risks; and
(g) Identifying and developing strategic interventions.

2.25. *Strategy implementation* is the process of turning strategies and plans into actions to accomplish strategic objectives. Strategy implementation includes:

(a) Compiling a census strategic plan, work programme and operational plan with the following elements:

i. Work planning: Defining inputs, processes and outputs in the operational plan;

ii. Human resource planning: Identifying the human resources required and determining how they should be deployed and developed to create the competencies needed to execute the strategy. For a full field enumeration, this might focus on determining the large numbers of enumerators and supervisors required, along with their skills development; For register-based approaches, with fewer individuals involved, the focus might be on institutional knowledge retention and succession planning;

iii. Financial planning: Preparing a budget outlining the funding required to implement the strategy;

iv. Contingency planning: Identifying operational risks and developing control measures and mitigating strategies to address those risks.

(b) Translating the strategy into action through:

i. Defining the scope of work, including key performance indicators, outputs to be delivered, targets and milestones to achieve the strategic objectives;

ii. Defining the method of work, such as developing a value chain at both conceptual and operational levels, including a quality management process to ensure data users' value is materialized;

iii. Defining the organization of work, including designing an organizational structure to effectively implement the strategy.

(c) Managing the implementation process and overseeing strategic change.

2.26. Strategy *monitoring and review* involves monitoring and reporting on the progress, achievements and challenges in the programme, taking corrective action where required, and evaluating the impact of any changes or improvements. For successful implementation, continuous monitoring is essential. As external and internal conditions constantly evolve, census managers must continuously review these environments, recognizing new strengths, weaknesses, opportunities and threats that may arise. The key element in strategy monitoring is obtaining relevant and timely information in order to take corrective actions where required.

C. Making censuses more inclusive

2.27. Censuses are the most important source of information on specific population groups, including small and marginalized groups. Census data are invaluable for understanding these groups in greater detail and at various levels of geographic aggregation. These groups may include national or ethnic, religious, linguistic, racial, or gender identity minorities, women-headed households, children and youth, older persons, persons with disabilities, international migrants, stateless and forcibly displaced people and nomadic populations, depending on the national context.[57] The need for comprehensive and granular data for these groups is discussed in detail in paragraphs 3.530-3.562.

2.28. While census data based on full counts generally exhibit little to no sampling errors, stereotypes related to specific population groups can introduce biases in census data collection and presentation, affecting coverage and quality. This risk is particularly high when multiple identities intersect, leading to

[57] See also the list of difficult-to-enumerate groups in paragraphs 5.46-5.47.

compounded biases. For example, if enumerators or respondents assume that a woman cannot be the head of a household, this may distort the identification or selection of the appropriate reference person, resulting in inaccurate or biased reporting.[58] These biases are discussed in more detail in Part Five (see for example paragraphs 5.141–5.152 and 5.363 relating to household relationships and economic characteristics, respectively).

2.29. Census authorities must remain vigilant to the possibility of stereotypes and biases affecting data on these population groups. Actions can be taken at different stages of the census process to help minimize these biases (see Figure 2.2.). For example, carefully designing census content - including what information is collected, the definitions and classifications used, and how data are gathered, tabulated and disseminated – is critical for generating data to inform policies and programme to address the specific challenges that certain population groups are facing. Census planners must ensure that the census design minimizes the influence of stereotypes on both respondents and enumerators. For example, using inclusive language in census questionnaires can help avoid perpetuating such stereotypes.[59] Strategies for designing an inclusive census are further discussed in paragraphs 3.166-3.179 and 3.534-3.562.

2.30. At the census operation stage, particular attention should be paid to the selection, training and supervision of field staff. This involves ensuring that both women and men from various population groups and communities are represented among enumerators and supervisors. Census manuals and training materials should address how stereotypes may cause potential errors in data collection, production and dissemination.

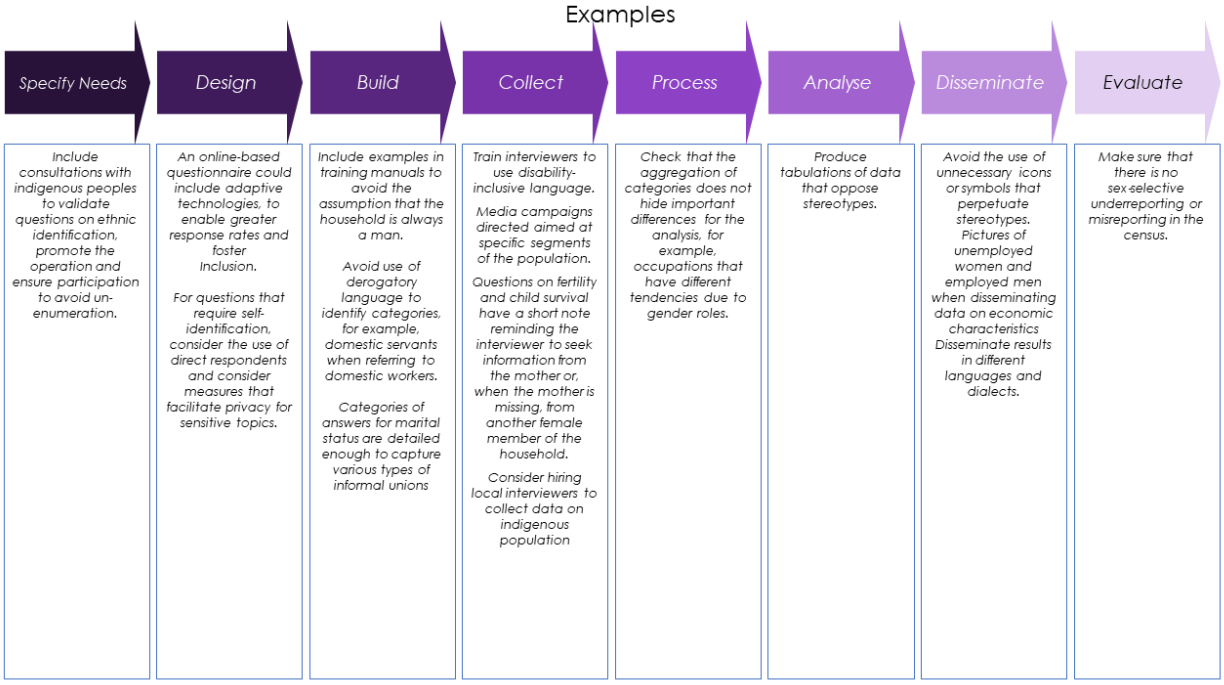
2.31. The entire process of census planning should also be gender-sensitive. For example, identification of strategic objectives should intentionally have a gender lens; project management team should intentionally aim for gender parity in terms of representation (e.g., composition of the committees, field staff, and leadership roles representation); user consultation, communication and publicity should ensure gender sensitivities. The aim is not to just have gender-sensitive census data but parallel to this, the entire process, operations and management is likewise gender-sensitive.

2.32. Engaging with women's groups and other civil society organisations concerned with specific population groups is recommended. Representatives of these groups can often provide valuable insights and information relevant to both census content and operational issues. Thus, special efforts should be made to consult with such groups throughout the census process, particularly during the planning phase. In the case of indigenous and populations living in isolated settlements or enclaves, such consultations are often critical for minimizing under-enumeration.

[58] See also paragraphs 5.141-5.148 on gender-biases in the selection of the reference person.

[59] Further discussions on disability-related wording can be found in paragraphs **Error! Reference source not found.**-5.233. Additionally, references can be made to the United Nations guidelines for gender-inclusive language https://www.un.org/en/gender-inclusive-language/ and for disability-inclusive language https://www.ungeneva.org/sites/default/files/2021-01/Disability-Inclusive-Language-Guidelines.pdf .

Figure 2.2. Best practices to make censuses more inclusive



IV. Units, place and time of enumeration

A. Units of enumeration

2.33. The units of enumeration should be clearly defined in the beginning of census planning as individual enumeration is an essential feature of a population and housing census. In population censuses, the primary unit of enumeration is the person. Individuals are identified and enumerated through (a) households and (b) collective living quarters or institutions.

2.34. Households are also counted as a unit of enumeration sand the majority of the population are assigned to such a unit; such persons constitute the “household population”. In a full field enumeration, careful identification of households is critical to facilitate the efficient enumeration and data collection, leading to better coverage. For register-based methods, the source data are often at the individual level, but the data requirements of uses will usually necessitate the construction of households from the individual-level source data.

2.35. Collective living quarters or institutions are another unit of enumeration which is also relevant to population censuses. People in these living quarters constitute the “institutional population”.

2.36. In housing censuses, there are three units of enumeration: (a) households, (b) living quarters, including housing units, collective living quarters or institutions, and (c) buildings. It is important to recognized that these three units are conceptually distinct from one another. There is no inherent identity or exact correspondence between these concepts, and the terms are not interchangeable. For example, multiple households may share the same set of living quarters, while a single household may occupy multiple living quarters. Similarly, multiple living quarters may be housed in a single building.

2.37. It is recognized that some countries may face challenges in maintaining independent concepts of "household" and "housing unit". However, the advantages of preserving these separate concepts, especially in terms of the usefulness of the resulting data, usually outweigh the additional effort required in maintaining them.

2.38. It is essential that the definitions of the units of enumeration be included in census manuals and reports to provide clear guidance to both enumerators and users for accurate data collection and utilization. To reduce potential difficulties in applying the recommended definitions, countries may need to tailor these definitions to the national conditions and circumstances. Post-enumeration field checks can be a useful means for assessing to what extent the national definitions of the units of enumeration were applied in the field and the consequent effects on the census results.

1. Person

2.39. For census purposes, the term "person" denotes each individual included within the scope of the census. A person can belong to either the household population or the institutional population. Although every person must be counted in the population, there may be some variation in the information collected on specific topics, typically depending on the person's:

(a) **Age** (e.g., economic activity, which may be influenced by national legal minimum working age laws);
(b) **Sex** (e.g., questions relating to women with children born); or
(c) **Labour force status** (for which no information on employment status, occupation, etc. is collected for the unemployed and persons outside the labour force).

2.40. Thus, it may be recommended that certain topics need not necessarily be investigated for the entire population. The group of persons from whom information on a given topic is to be collected is indicated in the definitions and specifications of such topics in Part Five, Chapter III, section D. Similarly, the scope of the census should clearly indicate which persons are to be included and which are to be excluded.

2. Household

2.41. The concept of "household" is based on the arrangements that a person or a group of persons make to provide themselves with food and other essentials for living. A household may be either:

(a) **a one-person household**, where a person who makes provision for their own food and other essentials for living without combining with any other person to form a multi-person household; or
(b) **a multi-person household**, which consists of a group of two or more persons living together who make joint provision for food and other essentials for living. Members of such a group may pool their resources, share a common budget, and may be related or unrelated to each other.

2.42. The above concept of a "household" is known as the "housekeeping" concept. This does not assume that the number of households and housing units are, or should, be equal. A housing unit, as defined below, is a structurally separate and independent space of abode that have been intended for

habitation.[60] It may be occupied by one household, by more than one household (e.g., two nuclear households that share one housing unit for economic reasons) or by a part of a household (e.g., one household in a polygamous society routinely occupying two or more housing units).

2.43. Some countries use the "household dwelling" concept, which considers all persons living in a housing unit as belonging to the same household. According to this concept, there is one household per occupied housing unit, meaning the number of occupied housing units and the number of households occupying should always be equal, and their locations are identical. However, this concept may obscure some information on living arrangements, such as multi-household or shared dwellings, which is relevant for evaluating housing needs. For more discussion of household occupancy, see paragraphs **Error! Reference source not found.**–5.532 and of household and family characteristics, see paragraphs 5.133-5.140.

2.44. In the register-based census approach, determining individuals' relationships based on the housekeeping concept can be challenging. Most countries that use this approach rely on the household dwelling concept instead.

2.45. Households usually occupy the whole or a part of one housing unit, but they may also be found in camps, boarding houses or hotels, or as administrative personnel in institutions, or they may be homeless. Households consisting of extended families that make common provision for food, or of potentially separate households with a common head resulting from polygamous unions, or households with vacation or other second homes, may occupy more than one housing unit. For more discussion of household occupancy, see paragraphs 5.528-5.532.

2.46. A household may consist of one or more homeless people. The definition of homelessness can vary from country to country because it is essentially a cultural concept influenced by ideas such as "adequate housing", "minimum community housing standard", or "security of tenure",[61] which can be perceived in different ways. The following two categories or degrees of homelessness are recommended:

(a) **Primary homelessness (or rooflessness)**: This category includes persons living on the streets or without any form of shelter that would fall within the scope of living quarters;

(b) **Secondary homelessness (or rootlessness)**: This category includes:

   i. Persons with no place of usual residence who frequently move between various types of accommodation (including dwellings, shelters or other living quarters);

   ii. Persons usually resident in long-term or "transitional" shelters or similar arrangements for the homeless.

2.47. For some topics in housing censuses, households may serve as a more efficient unit of measurement than living quarters. For example, topics such as tenure and rent should be collected with reference to households rather than living quarters. Similarly, information about household possessions that might, in some cases, by considered as part of the living quarters' equipment, such as radio, television

[60] See the detailed definition and classification of living quarters, including housings units and collectives living quarters in Part Five, Chapter III

[61] For the definition of tenure, see paragraphs 5.614-5.617.

receivers and laptops, should be collected with reference to households. On the other hand, type of ownership is more relevant to the condition of living quarters.

3. Population in collective living quarters

2.48. Collective living quarters or institutions are another unit of enumeration in population censuses. People residing in these quarters constitute "institutional population", distinct from "household population". This population type includes persons living in military installations, correctional and penal institutions, dormitories of schools and universities, nursing homes, religious institutions, hospitals, and similar settings.[62] Employees living together in such institutions should also be included in the institutional population. Older persons who receive care from a nursing facility but live in an independent living quarter are considered as part of the household population, not the institutional population. It is important to note that as more countries adopt self-response modes in censuses, each respondent in an institution may be enumerated as a single member household. This shift may pose challenges for ensuring intertemporal comparability of census data.

4. Living quarters

2.49. The principal unit of enumeration in housing censuses is living quarters. Precise identification of these units is essential for obtaining data that provide a meaningful description of the housing situation and serve as a basis for the development of housing programmes and policies.

2.50. Living quarters are structurally separate and independent places of abode. They may:

(a) have been constructed, built, converted or arranged for human habitation, provided that they are not being used wholly for other purposes at the time of the census, and in the case of improvised housing units and collective living quarters, they are occupied; or

(b) have not been intended for habitation but have been in use for that purpose at the time of the census.[63]

5. Housing unit

2.51. A housing unit is a separate and independent place of abode intended for habitation by a single household, or one not intended for habitation but occupied as living quarters by a household at the time of the census. Thus, it may be an occupied or vacant dwelling, an occupied non-conventional housing unit, or any other place occupied as living quarters by a household at the time of the census.

2.52. Key characteristics of a housing unit are described in paragraphs 5.485-5.488. It should be noted that any housing units located on the grounds or within the buildings of an institution, camp, and so forth should be separately identified and counted as such. Specific instructions for identifying housing units may vary based on local context and census requirements.

---

[62] For more detailed definition and specifications of institutions as a subset of collective living quarters, see paragraphs **Error! Reference source not found.**-5.519.

[63] For a more detailed discussion of the definition of living quarters and of the concepts of separateness and independence as used in the definition, see paragraphs **Error! Reference source not found.**-5.480.

2.53. The enumeration of housing units provides essential data on the living conditions and housing characteristics of the population. This data includes information on the number of rooms, type of dwelling, ownership status, and availability of basic amenities such as water, electricity, and sanitation facilities. By collecting detailed information on housing units, censuses can offer insights into housing quality, overcrowding, and the availability of housing resources. See Section III or Chapter II in Part Five for as the definitions and specifications of the information recommended to be collected on housing units. This information is vital for planning and policymaking, helping governments and organizations to address housing needs, improve living conditions, and allocate resources effectively.

2.54. Persons living in hotels or boarding houses are not part of the institutional population. Instead, they should be distinguished as members of one-person or multi-person households, on the basis of the arrangements that they make for providing themselves with the essentials for living.

6. Building

2.55. Buildings are regarded as an indirect, but nevertheless important, unit of enumeration for housing censuses. Information concerning the building – such as its type, construction material and other characteristics – is required for properly describing the living quarters located within it and for formulating housing programmes. In housing censuses, the questions on building characteristics are normally framed in terms of the building in which the enumerated living quarters are located, with the information recorded for each housing unit or other living quarters within that building.

2.56. A building is generally defined as any independent free-standing structure comprising one or more rooms[64] or other spaces, covered by a roof and usually enclosed within external or dividing walls[65] that extend from the foundations to the roof. However, national definitions of buildings may vary across countries: a building may consist solely of a roof without walls in tropical areas, while a roofless space enclosed by walls may be considered a building in some countries.[66]

2.57. In some countries, “compounds" can be used as a unit of enumeration, either in addition to or as a substitute for the building. In certain areas, living quarters are traditionally located within compounds, and this grouping of living quarters may have certain economic and social implications that are useful to study. In such cases, it may be beneficial to identify compounds and record information suitable for linking them to the living quarters during census operations.

B. Place of enumeration

1. Concepts relating to the place of enumeration

[64] For the definition of rooms, see paragraph 5.546.
[65] The term dividing walls refers to the walls of adjoining buildings (for example, of row houses) that have been constructed so as to be contiguous.
[66] For a more detailed discussion of the definition of building and related concepts, see paragraphs **Error! Reference source not found.**-5.591.

2.58. In population censuses, a country may choose to enumerate all persons present in its territory and/or those who potentially belong to the population of interest. The *population to be enumerated* refers to the group of persons whom the country decides to cover, regardless of their later inclusion in the official population count. For example, a country may enumerate everyone present at a specific point in time, but produce official statistics only for those who are usually resident.

2.59. The place of enumeration is either the place where the person is found (referred to as a *de facto* enumeration) or the place of usual residence of the person at the census reference moment (referred to as a *de jure* enumeration). It should be ensured that each person is enumerated in only one place to avoid double-counting. Countries should document the definition of the place of enumeration adopted for their census. In field enumerations, explicit instructions should be provided to enumerators on how to apply this definition during interviews, or to respondents when filling in self-response questionnaires.

2.60. For census purposes, *usual residence* is defined as the place at which the person lives at the time of the census and has been there for some (specified) time or intends to stay there for some (specified) time. This refers to the geographic place where the person usually spends their daily rest, assessed over a defined period of time, including the census reference time. For most individuals who have not changed residence during this time period, determining their place of usual residence is unambiguous. However, for others—particularly those who move frequently — applying this definition may lead to varying interpretations.

2.61. It is recommended that countries apply a 12-month threshold when determining a person's place of usual residence, according to one of the following two criteria:

(a) The place at which the person has lived continuously for most of the 12 months before census reference day (that is, for at least six months and one day, not including temporary absences for holidays or work assignments) or intends to live for at least six months;
(b) The place at which the person has lived continuously for at least 12 months before census reference day (not including temporary absences for holidays or work assignments) or intends to live for at least 12 months.

2.62. For register-based censuses, the place of usual residence could be taken to mean the place of legal or registered residence, defined with reference to a qualification period of 12 months or assessed with reference to the 12-month criterion.

2.63. Persons who move frequently and do not have a place of usual residence should be enumerated at the place where they are present at the time of the census.

2.64. Regardless of the criteria used to define the 12-month period, countries should ensure that each person has one *and only one* place of usual residence (or registered place of residence for census purposes).

2.65. There are various population groups for whom some uncertainty may arise about their inclusion in the usual resident population. The following persons should generally be considered as part of the usually resident population:

(a) Persons found at the moment of enumeration who cannot identify a place of usual residence, such as those who change residence often;
(b) National military, naval and diplomatic personnel and their families, located outside the country, irrespective of their duration of stay abroad;

(c) Foreign citizens working within the country (but not including foreign diplomats or military forces) and their families, provided that they meet the criteria for the usual residence in the country;
(d) Merchant seafarers and fishery workers usually resident in the country but at sea at the census reference time (including those who have no place of residence other than their quarters aboard ship);
(e) Persons who are irregular or undocumented migrants, as well as asylum seekers and persons who have applied for, or been granted, refugee status or similar types of international protections, provided that they meet the criteria for usual residence in the country (the intention is not to distinguish these persons separately, but rather to ensure that they are not missed from the enumeration);
(f) Persons who cross a frontier daily or weekly to work or study in another country, provided that they meet the criteria for usual residence in the enumeration country;
(g) Children born in the 12 months before the census reference time and whose families are usually resident in the country at the census reference time; and
(h) Persons who regularly live in more than one country during a year, if they reside in the country conducting the census most of the time, regardless of whether they are physically present in that country at the census reference time.

On the other hand, the following group of persons should generally be excluded from the usual resident population:

(a) Foreign military, naval and diplomatic personnel and their families, located in the country, regardless of their duration of stay and/or place of usual residence;
(b) Third-level students who are, or intend to be, absent from the country for one year or more;
(c) Persons who regularly live in more than one country during a year, if they reside in the country conducting the census the least amount of time, regardless of whether they are physically present in that country at the census reference time.

2.66. The concept of usual residence is often treated as synonymous with "*de jure*" residence. However, in certain circumstances, the term "*de jure*" may imply that a person's residence in an area must have a legal basis under the applicable legal system. This would suggest that people without such a legal basis should not be enumerated in that area. It is not recommended that population and housing censuses limit their enumeration only to individuals with a legal right to exist or reside in an area. Instead, as described in Section 2 below, censuses should include either all those present in the area or all usual residents at the census reference time.

2. Operational issues relating to the place of residence and the place of enumeration

2.67. In field enumeration censuses, each person's place of enumeration is entered as either the location where they are (or were) present on the day of the census or their place of usual residence.

2.68. When compiling census results by geographic areas, each person who is part of a household can be included in either:

(a) the household (and hence the geographic area) where the person was present on the day of the census; or
(b) the household (and the geographic area) where they usually reside.

The same applies to the institutional population.

2.69. If a "present-in-area" population distribution is wanted, it is logical to enumerate persons at the place where they are (or were) present at the time of the census. If a distribution by usual residence only is required, it is preferable to collect information about each person at the person's place of usual residence. It should be noted, however, that it is not always possible to collect information about each individual at their usual residence – such as when an entire household is away from its usual residence at the time of the census. Therefore, some provision must be made to collect information about these individuals at the place where they are found during the census and collect information on their place of usual residence.

2.70. If the objective is to obtain information on *both* the usually resident population and the present-in-area population, countries can enumerate all persons present as well as all persons usually residents but temporarily absent in each household or institution. A clear distinction must then be made in the questionnaire, as applicable, between

(a) persons usually resident and present on the day of the census;
(b) persons usually resident but temporarily absent on the day of the census; and
(c) persons not usually resident but temporarily present on the day of the census.

2.71. Census data collection should include the usual residence of individuals who are temporarily present at another location on census day. Similarly, for those temporarily absent, their usual residence should be recorded. This information helps allocate persons to the appropriate household or institution, as well as the geographic location at which they should be counted, ensuring that no person is counted twice. The enumeration procedures and the subsequent allocation of persons must be carefully planned, clearly documented and strictly adhered to ensure accuracy.

2.72. With the exception of mobile housing units, living quarters and buildings have a fixed location, so their place of enumeration is unambiguously identified in housing censuses. However, for field-based enumeration, information on households and their members can be collected and entered according to either where they are (or were) present on the day of the census or at their usual residence. If the population and housing censuses are conducted simultaneously, the procedure used in the housing census should align with that of the population census. If the housing census is conducted independently, the procedure must be carefully considered, since it may have a significant effect on the validity of its results.

2.73. Persons and households should be assigned to both their usual place of residence and their usual living quarters. If a household is occupying living quarters that differs from their usual place of residence at the time of the census, these living quarters should be counted as vacant if they are conventional dwellings or excluded from the census if they are non-conventional dwellings.[67]

2.74. Mobile housing units represent a special case. They should be enumerated where they are located on the day of the census. However, the occupants of such units should be allocated to the area where they usually reside, provided that the unit is their usual living quarters. If the mobile unit is not their usual

[67] To be considered as living quarters, non-conventional housing units and collective living quarters are required to be occupied in order to be included in the census.

living quarters, the occupants should be allocated to the place where they usually live, and the mobile housing unit should be excluded from the census.

C. Time of enumeration

2.75. One of the essential features of population and housing censuses is that every person and every set of living quarters should be enumerated at the same well-defined point in time (or with respect to the same point in time). This is usually accomplished by designating a specific moment (e.g., at midnight) at the beginning of the well-defined period of enumeration. This moment is known as the "census reference moment".

2.76. In population censuses, each person who is part of the target universe and alive at the census reference moment should be included, although the collection and processing of census forms or administrative data takes place later. Therefore, infants born after the census reference moment should not be included, even if they are living at the time of enumeration. Similarly, persons who were alive at the census reference moment but deceased before the enumeration should be included. In both field-based and register-based censuses, strict adherence to the census reference moment is important to produce high-quality data, especially since the field-based enumeration may extend beyond the designated enumeration period. In register-based censuses, modelling techniques may be required to harmonize administrative data with different reference dates to ensure that they refer to the same census reference moment.

2.77. In housing censuses, all the living quarters that have reached a specified stage of completion and are not scheduled for or in the process of demolition should be counted as a part of the total count of living quarters. Living quarters that would reach the prescribed state of completion after the census reference moment should not be included in the count, unless special instructions are issued for recording living quarters under construction.

D. Time reference period for data on the characteristics of the population and of living quarters

2.78. Census data on population and housing characteristics should refer to a well-defined reference period, but the time reference need not be the same for all the data collected. For most data, the reference will be the census reference moment. In contrast, the reference period of certain questions – such as those related to labour force status, rental arrangements, or fertility – may refer to a certain period prior to the census reference moment.

V. Legal basis

2.79. Legal authority for conducting a census is required to:

- Regulate primary administrative responsibility;
- Obtain necessary resources;
- Determine the general scope and timing of the census;
- Impose a legal obligation upon enumerators, supervisors and census officials to record responses faithfully and to fulfil their assigned responsibilities.
- Place a legal obligation upon the public to cooperate and provide truthful answers;
- Provide a mandate to collect and use administrative data for census purposes;

- Establish the obligations of register-holders to provide the requisite data and to inform about any changes to the content and/or structure of the relevant registers; and

2.80. In some countries, the legal basis for conducting a census may be specified in legislation that is separate from that relating to the NSO's responsibilities more generally. In this case, linkages should be established between census legislation and other relevant statistical legal frameworks.

2.81. Additionally, the confidentiality of the individual information should be clearly established in the census legislation and guaranteed by appropriate sanctions to ensure public confidence in the process, assuring the public that their data – whether collected directly or from administrative sources – are handled responsibly.

2.82. In countries that lack permanent legal authority for conducting periodic censuses, it is important to establish such authority early on, whether through *ad hoc* legislation or, preferably, through legislation that mandates a system of periodic censuses. In case of an *ad hoc* legal authority, steps should be taken to safeguard the independence of the implementing agency and delegate the necessary technical and administrative powers to it.

2.83. While the content of the census legislation will inevitably depend on national legal practices and procedures, as well as on the organization of the national civil service, the following components are usually prescribed:

- the purpose of the law;
- the coverage of the census;
- assigning the mandate for conducting a census to a specific institution;
- purpose of the census;
- obligation and rights of the citizens;
- modes of financing the census;
- organization of the census;
- the bodies responsible for administering the census;
- the mode of data collection;
- access to, and use of, record-level administrative data;
- the technology to be used;
- rights and obligations of enumerators and supervisors;
- census data dissemination and exploitation;
- treatment of personal data;
- confidentiality and privacy of respondents and their data;
- disposal of the census equipment and materials; and
- archiving.

2.84. A comprehensive and well-timed legislative framework is of utmost importance for ensuring the legality and authority to conduct a census. In cases where census legislation is enacted before each census on an *ad-hoc* basis, it may include specific census topics, thus providing additional legal weight to the questionnaire's composition and the census content.

2.85. The legal basis for a census should be continuously improved based on past experiences. After each census evaluation, legislative adaptations should be considered to strengthen confidentiality protections, foster the transparency and accountability of the census organization, improve public engagement and ensure the quality of census outputs.

2.86. When drafting such census legislation, the principle of conceptual and organizational flexibility should be observed, especially to ensure the NSO's adaptability to rapidly changing external conditions. Legislative provisions should guarantee data security and confidentiality, but should not be overly rigid regarding the type of data to be collected or the structure and relationships within the census organization. Instead, necessary operational details should be contained in regulations promulgated by the census organization. Moreover, provision should be made to simplify administrative procedures, including delegating authority for the procurement of equipment and supplies, as well as the recruitment of personnel during the operational phase.

In addition to legislation authorizing the conduct of a census, there may be legal requirements for specific census data elements to serve as the primary source of data for certain purposes, such as delineating electoral boundaries or determining revenue allocation among central and regional governments.

VI. Financial management

A. Financial basis for censuses

2.87. A census is the primary source of data about the size and characteristics of the population disaggregated into lower geographic levels; it provides a demographic profile of a country and is the basis for developing area sampling frames for use in surveys. A census, however, is usually one of the largest and costliest statistical activities that governments and their national statistical offices undertake. As a result, some countries have been forced to delay or even cancel a census owing to funding constraints. Countries that have been able to secure partial funds or secure funds but at a late stage of their census preparation have been forced to compromise their data collection, data processing, and dissemination of census results. It is therefore recommended that all census operations, including planning, cartography, enumeration, processing, analysis and dissemination, be fully budgeted from the beginning and efforts made to mobilize the required resources, for the entirety of the census cycle (i.e., for all fiscal years). Inflation should be taken into account, keeping in mind that duration has an impact on cost.

2.88. Hence, there is growing pressure to look into the solutions to census funding, taking into account the role of key stakeholders, namely governments and their statistical agencies, and the greater involvement of national and international development partners and stakeholders including the private sector. Concurrently, cost-effective strategies need to be put in place that would reduce census costs without compromising the quality of census data.

2.89. It should be emphasized, however, that censuses cannot be carried out merely by national statistical and census offices alone. Rather, conducting a census should be seen as a national endeavour involving many stakeholders. Thus, government departments, non-governmental organizations and private sector end users should be consulted at all stages to ensure the legitimacy of, and need for, conducting the census and, at the same time, to improve advocacy for sufficient funding. Although conducting a census is principally financed by the government, it should be designed in partnership with all political players so as to obtain their support and involvement in the census process.

2.90. NSOs need to advocate the importance of investing in censuses within their own governments. It is also important for NSOs to ensure continuous feedback and promote the use of statistical data from previous censuses, in order for users to recognize and appreciate the importance of the population census

as a unique source of statistical data and give their support. The possibility of reducing the financial costs through collaborations with other government departments and local authorities with existing devolved infrastructure, such as education and health ministries, should be further explored. These institutions could be supportive in providing administrative data and/or logistics arrangements for the census, such as the use of existing infrastructure, transportation, and communications facilities, and sharing of employees of other government departments.

2.91. Excellence in planning is an essential prerequisite not only for achieving a cost-effective census but also for securing comprehensive financial support for its funding. Technologies and methods that will be used in mapping, data collection, administrative data intake and management, processing, questionnaire design, and other activities must be decided upon in advance, as these have an influence on costs. Census planning must bring out the links between the various components, which will include types of resources (such as personnel, cost of equipment, stationery or printing) as well as tasks (including data collection and capture, data processing, data management, dissemination and archiving). Costs must be determined for each of these components together with a justification. Experience and lessons learned from past censuses or similar activities must be considered when estimating costs for the next census. Where multiple modes of data collection and new technologies are being used for the first time, these must be tested for viability, data quality and cost implications.

2.92. For each stage of the census process the costs must be optimized. A careful choice of the appropriate technology will greatly assist in this. Recent advances in technologies throughout the field-based census process, such as digital mapping, computer-assisted or Internet data collection, scanning, data processing (including machine learning technologies for coding) and data management and archiving, and census data analysis and dissemination, may be of assistance in achieving significant efficiencies, allowing more cost reductions or increased productivity. In addition, the proper selection and use of such technologies will speed up the computation of results and enhance their preservation. However, the choice of technology should be made only after carefully evaluating the costs and benefits of possible options. Some potential risks to canvass include the following:

- some approaches only become cost-effective for large operations;
- some are dependent on expensive and scarce inputs (for example very high-quality satellite images or paper for scanning);
- some are dependent on services that may not be available throughout the country (for example Internet access); and
- others require significant investments in high-quality computers and upfront investment in human resources.

The options examined in the cost–benefit analysis could incorporate consideration of leasing (rather than purchasing) equipment or sharing/borrowing it between countries that are undertaking censuses at convenient times.

2.93. Outsourcing or collaborating with the private sector could be considered as another potential cost-saving option, particularly in the context of publicity and advocacy or for systems development for data collection, transmission, processing, dissemination and archiving. Outsourcing can contribute technical expertise or resources not readily available within the national statistical office or census implementing agency. (Section XIV below discusses in more detail the issues to address when considering the outsourcing of census services and processes.)

2.94. It is anticipated that international donors will continue to play a pivotal role in supporting funding census costs in many countries. Technical cooperation and assistance from international agencies have

also contributed greatly to the success of censuses in many countries. It is worth noting that a population and housing census has some intangible positive values. It is an opportunity for mobilizing the whole country and reaching even the most remote corners of it. In the life of many citizens, a regular census is often the only time that the State reaches out to them and asks them some questions. Successfully conducting a census is a matter of pride in many countries and a welcome opportunity to recruit a massive labour force and generate jobs and train people in valuable tasks (such as data entry) or in other ways to add to the national infrastructure.

2.95. There is opportunity for cost savings through the use of administrative data to replace or support field-based data collection operations, usually the most expensive phase of census-taking. As countries increasingly integrate administrative data into their censuses, this can be an important driver of change. However, considerations should be made for the added expense associated with the initial acquisition, intake, processing and other management (e.g. modelling) of administrative data used for census-taking. Also, if administrative data cannot entirely replace what is collected traditionally, good financial planning will be required to supplement any continued surveying that must continue, post-transition to administrative data.

2.96. As the census is of interest and relevance to a wide range of national and international bodies, both in the public and private sectors, all activities related to funding need to be elaborated, documented, justified and presented to all stakeholders in a transparent and comprehensive manner.

B. Budget and cost control

2.97. While no universal system of census budgeting and cost control can be suggested - since financial practices vary greatly among countries - a few generally accepted principles can be noted. First and foremost, effective planning and control of the various census operations is not possible without a very careful financial estimate of the cost of each census operation, including all of its components, no matter how small. It is recommended that NSOs draft a detailed list of activities related to censuses with timelines and, as far as it is possible, to determine a budget in such a way that it corresponds to this list of activities. Secondly, it is critical for this census plan and budget to be presented by NSOs to their respective governments with adequate lead time, to facilitate the timely appropriation of sufficient resources from national budgets or, where required, from the international development community. Moreover, funding of the census must be accompanied and developed on a sound and adequate legal basis if effective national census operations are to be enabled.

2.98. Information on expenditures from the previous census classified by census phases, starting with the expenditure for different elements of the preparatory work and ending with expenditure for the dissemination and evaluation of the census results, provides an important basis for estimating the overall budget of the census. Figures from the previous census will, of course, have to be reviewed and updated in order to take into account:

- quantitative and qualitative change in hardware and software;
- changes in wage rates;
- the costs of equipment, supplies and other materials;
- inflation generally;
- planned changes in census content, methods and procedures; and
- anticipated changes in the population itself (for example, total size, percentage urban, and average household size).

2.99. All of the above may affect the cost structure of the census. In most countries, several cost elements tend to increase (for example, wage rates and the size of population) so that there is considerable pressure to achieve economies in other items of the census budget. In addition, time must be spent to identify new features of the next census, for which a past expenditure cannot provide the basis of a cost estimate. For example, future censuses might need to consider a new budget for cybersecurity, and/or the acquisition and use of administrative data.

2.100. NSOs need to implement transparent accounting procedures and financial management systems to ensure speedy disbursement of funds, proper receipting of expenditures and an efficient audit. This would enable prompt release of periodic allocations of census funds by national governments. A clean outcome from a financial audit adds credibility to the census process so that the government and civil society are more likely to accept the final results.

2.101. In the case of external or donor funds, the required conditions should be established well in advance through dialogue with the relevant parties. This will avoid delay in the release of such funds for census operations.

2.102. Control measures and monitoring systems must be developed for cost-effectiveness. Activities to be outsourced should be clearly defined and contracts for outsourcing should be well prepared with clear deliverables and timelines.

2.103. For planning the costs of a census, detailed and precise data will be required on the following:

(a) number and cost of census staff classified by function and manner of payment;
(b) types of equipment and material used for the census, manner of acquisition (purchase or rental) and cost;
(c) office space requirement (surface measurement), classified by use and type of cost (that is, for construction or for rent); and
(d) type of services used for census operations.

2.104. The usefulness of the above information would be enhanced if the information could be recorded by source of funding, in other words, in terms of whether the expenditure has come from:

(a) the official census budget;
(b) other funds of the census office (for example, a regular annual budget not specifically intended for census purposes, or general funds of the governmental agency or department of which the census office is a part);
(c) other parts of the government;
(d) non-governmental organizations; or
(e) international donors.

2.105. This information is needed not only for fiscal planning and control but also in order to examine the trade-offs in terms of costs and benefits among alternative ways of carrying out various census operations. Although cost experience from a previous census in a country may provide a useful benchmark for planning the next census, much more caution should be exercised in using the cost parameters from other countries. Differences in census content, organization and operations, as well as in cost accounting, can introduce serious incompatibilities into such country-to-country cost comparisons.

2.106. It is important that the persons at the administrative and supervisory levels who will be responsible for the execution of each operation should participate in estimating the budget items. Such

an organization of the work pre-supposes detailed advance planning and "cost consciousness" on the part of those responsible for a census.

2.107. The census plan as executed will certainly change in a number of respects after the making of the original cost calculations. Consequently, a perfect correspondence between the estimates and the final costs is not to be expected. Changes in the prices of major components of census costs should be monitored on a regular basis with either the census budget adjusted accordingly, or the census plans modified. Indeed, the development of the census budget is usually an incremental process in which rough initial estimates are replaced by more detailed and precise statements of resource requirements. Throughout the period of census taking and compilation of census results, the budget will have to be re-examined and performance compared with plans. With detailed information on expenditure, the governmental and census authorities will be better able to exercise control over keeping the development of census operations within the census budget and to assess and control the effectiveness and efficiency of these operations. This information is also very useful for studying possible improvements in census techniques and census methodology.

2.108. As with any project, particularly ones as large and complex as the census, it will not, as a general rule, go according to plan and there will be difficulties and deviations. Accepting this at the outset and making sure there are arrangements in place to deal with delays, changes or other unforeseen issues is essential. In particular, there must be resources set aside to enable such issues to be dealt with quickly. Therefore, some contingency funding should be included within the overall costs of the census, and some controls put in place to monitor and allocate the contingency pot. Different methods exist for estimating the cost of the contingency budget, such as risk modelling, but a good starting point might be to allocate a percentage of the annual budget (say 15 per cent) each year for such contingency.

VII. Resource mobilization

2.109. A census is a massive undertaking that requires significant resources for its planning and execution, particularly for countries with full field enumeration censuses. These resources include financial resources to cover various activities and procurements, human resources with diverse skillsets and physical assets such as census materials and equipment. The specific types of resource mobilization may need to go beyond financial and personnel resources. Depending on the national context, this can include the mobilization of resources not available within the NSO, such as technology, data infrastructure, and cybersecurity. Therefore, the NSO must develop a comprehensive resource mobilization strategy before the census begins. This strategy should detail the required resources, timelines, and mechanisms for mobilizing and monitoring these resources. It is important to note the need for resource allocation to ensure adequate support for all phases of the census (pre-, during-, and post-enumeration).

2.110. A resource mobilization committee may need to be established in advance to develop, implement and monitor the resource mobilization strategy. The committee can also facilitate collaboration and resource mobilization among government bodies and solicit international and regional cooperation. Where pertinent, the mobilization of resources should include pooling resources at the regional level. This can facilitate the purchase and sharing of tablets and GIS experts among countries.

2.111. As censuses are repeated periodically, the resource mobilization strategy should be sustained and be adapted with each cycle. Resource mobilization should not be treated as a one-time action but should

aim for sustainable and impactful longer-term outcomes. For example, in countries where external census funding is required, diversifying donors and partners is a best practice to minimize the impact of changes in funding. In addition, building staff capacity through trainings can strengthen the NSO's capacity to mobilize resources for future censuses.

VIII. Project management

2.112. Effective project management is the cornerstone of any successful census project. It ensures that all activities are carried out efficiently and effectively, from planning to execution and evaluation. At its core, project management involves the application of knowledge, skills, tools, and techniques to meet the project requirements. This includes defining the project scope, setting achievable objectives, and managing resources to achieve these objectives within the constraints of time, cost, and quality.

2.113. Project management should serve as an overarching function, encompassing all aspects of the census plan, including scope, budget, quality, and information technology. Utilizing a work breakdown schedule can effectively link scope, cost, and schedule, dividing the project into manageable sections. By integrating multiple management plans, organizations can manage the various components and layers of the census programme both independently and collectively.

2.114. A well-defined schedule with a clear critical path is vital for ensuring that the census programme meets critical milestones and is implemented according to plan. This schedule should be regularly reviewed and updated to reflect changes in the project's scope or timeline, ensuring that the census remains on course and achieves its goals. The use of project management software tools can significantly enhance project management capabilities. These tools help track progress, manage risks, and improve communication among team members.

2.115. Establishing effective monitoring and evaluation frameworks, along with early planning and execution of all census phases, is crucial for ensuring success. These frameworks provide a structured approach to assess the progress and performance of the project, allowing for timely adjustments and improvements.

2.116. By employing certified project management personnel and involving highly skilled NSO staff, organizations can navigate the complexities of census operations with confidence. To maintain continuity and expertise, it is crucial to retain key personnel, including the census project manager, throughout the census and post-census evaluation phases. This retention helps ensure a smooth transition and effective knowledge transfer, which are vital for the project's success.

2.117. Establishing an external project board, comprised of international experts, can provide valuable outside perspectives, oversight, and support. Such a board also serves as an assurance mechanism, ensuring the project stays on track and meets its objectives.

2.118. Effective project management integrates all these elements to ensure that the census project is completed on time, within budget, and to the required quality standards. It is a dynamic process that requires continuous improvement and adaptation to changing circumstances

A. Development of workplans

2.119. The Project Management Institute defines projects as "temporary efforts to create value through unique products, services, and processes"[68]. Censuses – whether field-based or register-based - are major projects which require very careful, advanced planning and governance. Their complexity necessitates a large number of people working together, with various skill types and levels. At the centre, orchestrating it all, should be a census manager and governance that pulls the necessary team members together to make planning and execution decisions.

2.120. During the planning phases of a census, the management team needs to develop detailed workplans. While learning from past censuses is valuable, each census is unique. Therefore, the team should carefully consider the specific requirements and objectives of the upcoming census and articulate a clear vision. This high-level plan will necessitate adapting existing workplans or developing new ones. These workplans should outline the planned activities, required resources (human, financial, technical, etc.), and timelines for planning, developing, and implementing the census.

2.121. Multiple workplans will be required. Each should be prepared by an assigned team member with the greatest expertise, after they are fully acquainted with the appropriate inputs from as many elements of the census programme as necessary. The assigned team member who prepares the workplan should also be the same person who presents updates on how plans are developing. In other words, it is the person accountable for planning and executing the workplans.

2.122. To ensure clarity, census project teams developing their workplans should have "charters" that describe their function within the census programme. These charters should describe roles and responsibilities that make it sufficiently clear as who should be the responsible team member taking the lead on the development of each workplan. The charter should also integrate with the overall census project document that provides the basic parameters for the next census, developed from previous cycles and adapted to incorporate the vision for the next.

2.123. A key product of the workplans should be a "critical path" of milestone dates for the completion of work to prepare for the next census. This critical path represents key dates when if one activity is not completed, it would have downstream impacts on other parts of the census plan. Other dates should be tracked as well, but the critical path defines the most important dependent dates that the census team should be reviewing regularly and with greater frequency at key points during the census cycle (for example, in the months leading up to field operations of a traditional census).

2.124. It is recommended that the NSO appoints as a member of staff, at least one certified project management expert, who can develop project plans and other elements crucial to the successful completion of the census.

B. Emergency preparedness, risk management and contingency planning

[68] Deguire, M. (2012). In the eyes of the beholder. Paper presented at PMI® Global Congress 2012—North America, Vancouver, British Columbia, Canada. Newtown Square, PA: Project Management Institute.

2.125. The proper conduct of a census requires thorough planning, often spanning several years of work by dedicated and knowledgeable staff. However, these plans can be interrupted by unexpected and severe events, such as natural disasters, pandemics, political and civil unrest, insecurity, war, data breaches or cybersecurity threats. These serious events can lead to the failure, delay or even cancellation of a census (in particular where a full field enumeration is being carried out) if there is insufficient time or capacity for a country to respond properly. Even less catastrophic incidents, such as labour disputes, supply chain problems or technical issues, can pose significant challenges for census managers due to the inherent complexity of census operations.

2.126. Unfortunately, the context for census-taking is getting increasingly complex and risk-prone. New technologies and techniques being adopted to enumerate populations more efficiently may introduce complexities and risks. For example, in the 2020 round of censuses, countries were faced with a world-wide pandemic, numerous natural disasters (some associated with climate change), new cybersecurity threats and tight labour markets. While some countries had time to manage and adapt to this rapidly changing context, others were forced to adapt "on the fly", with varying levels of capability to do so.

2.127. In response, countries require new support to manage issues and risks related to all phases of the census, but especially at high-risk periods such as during the field operations of a traditional census.

2.128. Emergency preparedness and contingency planning are broad topics which can mean different things to different people. The following guidance demonstrates how to avoid issues in the first place by effectively using the planning phase for risk mitigation, but also to support NSOs when the inevitable happens during the execution phase. Also, the following is designed to provide support and ideas for when, despite good planning and emergency management, a census has a major break-down of its intended plan.

1. Documentation of risks and mitigation plans

2.129. To avoid or manage risks, NSOs must first identify them. Census managers should be aware of their potentially most damaging risks and set up processes to document them by creating a risk register. As defined by the International Standards Organization (ISO) in their risk management standard ISO 31000, a risk is "…anything that generates uncertainty or creates a deviation from the expected".[69] Risks should be identified for all phases of the census: pre-enumeration, during enumeration, and post-enumeration. Given the lengthy period of a census programme, risk planning must be forward-looking and encompass all potential contingencies throughout the census lifecycle. Moreover, it should draw upon lessons learned from past experiences to avoid repeating previous mistakes.

2.130. In a census context, these are issues that may threaten the proper execution of a census, the quality of the data, the security of the information collected, or the availability of human or financial resources. As a country plans to conduct a census – either field-based, register-based or combined – time should be dedicated to on-going discussions with census management teams, in order to identify and

[69] International Standards Organization. (2021, December 10). Effective risk management. New international guidance on how to get it right. https://www.iso.org/news/ref2773.html

manage risks. When attempting to identify risks, census managers should build upon previous experiences by asking themselves such questions as: “What are the threats to the census”?; “What happened the last time?”; and “Can it happen again?”.

2.131. However, it is important to recognize that some incidents cannot be anticipated based on previous experience. This underscores the importance of considering the current context, which is constantly evolving. Unforeseen challenges, such as pandemics or natural disasters, can emerge at any time. Therefore, a proactive approach incorporating scenario planning and contingency measures is essential. A robust and effective risk identification and mitigation process should conduct regular environmental scans to identify emerging threats, and regularly review the risk register to reassess exposure and ensure adequate mitigation plans are in place.

2.132. Risks can be further assessed using categories such as “low”, “moderate”, “high” and “extreme” to characterize their inherent likelihood of occurrence and the level of impact if the risks were to materialize. With this approach, NSOs should be better aware of the vulnerability of the census to various threats.

2.133. Census managers and subject experts should then discuss mitigations, or ways to lower the likelihood of occurrence and/or the impact of any threat to a census. Management must often prioritize this work, focussing on mitigations for the most significant risks, and/or those risks which are most likely to affect the census in the near term.

2.134. The best way to deal with a risk is to avoid it altogether through mitigating steps, but not all risks are fully within the control of the NSO. It may be important to discuss risk mitigations with external partners, such as IT contractors, postal agencies, printing firms, labour providers, administrative data providers, or any other key stakeholder involved in the conduct of a census.

2. Emergency management, governance and decision-making

2.135. A census will have a large number of “moving parts” subject to events not controlled or even anticipated by the NSO. Natural disasters are a good example. Although census managers understand that natural disasters can affect their operations, there is no predicting when, where or to what extent. Even with the best risk management and mitigation plan in place, a census is too complex to anticipate every potential issue. As a result, the statistical agency must be prepared to respond to the unexpected (such as the COVID-19 pandemic during the 2020 round). With proper emergency management, governance and decision-making, censuses can weather difficult storms effectively.

2.136. Every census should have a process in place to identify, classify and respond to an inevitable emergency. Key members of the census team should know how to respond to emergencies during crucial phases of their census. They need to be able to determine if an emergency is underway, how to classify the severity of the emergency, who to contact when it occurs and how contact can be made, so that solutions can be put into place.

2.137. Census team members must be equipped with the information necessary to detect new issues and to respond effectively using communication lines to other team members, senior officials and perhaps the general public. For example, field-based censuses are making greater use of Internet questionnaires and real time metrics which can be monitored to ensure the proper functionality of that service to citizens.

As well, for both field-based and register-based censuses, massive databases require constant monitoring to detect anomalies and cyberthreats. Should there be an issue detected during such monitoring, the person viewing that information must be prepared and competent enough to interpret what is happening, classify the degree of the emergency and communicate immediately to other team members to establish a resolution plan.

2.138. When any emergency is detected, it should be assessed for its severity. Census team members responsible for the component of the census impacted by the event should be trained on various severity definitions during the planning phase of the census. For example, when an event impacting the Internet response service is detected, the information should be shared with the census manager who is responsible for that service, and who is trained to determine the emergency's severity level by deciding, for example, whether the event being detected is creating a slowdown of the service that might be of minor annoyance to the user, or is a broader interruption of Internet response, such as a cyberattack.

2.139. The severity level of the emergency as determined by the responsible manager should trigger a pre-determined response. For example, a high severity event could require an immediate meeting of the full census management team (no later than one hour after detection, for example), and require a faster resolution compared to a low severity event. At such a meeting, the situation should be described by the responsible manager, who should then present information on the cause of the emergency and any identified remedies. If possible, there should be an estimate of when impacted census operations will be fully operational again, which will be particularly valuable for communications and any down-stream activity that may be subject to a subsequent impact.

2.140. To detect and respond to an emergency, a census should have an Operations Centre where employees have access to tools that track relevant operational data. When a census is underway, a number of operational metrics are available to staff, who should be presented with those data in real-time and with an historical archive to reference as well. Such data can be specified in advance of the census, using expertise from previous cycles. The resulting "Management Information System" (MIS) will gather information from various systems and present them in a way that allows managers to detect and respond to incidents.

2.141. In the event of an emergency, census managers should be prepared to communicate to others within the NSO. As noted at paragraph 2.139 above, priority should be taken to communicate first with the census team, to ensure a timely response. However, internal communication to senior leaders in the organization and political staff may be required, depending upon the severity and visibility of the incident. Also, should they be involved in the response or impacted by the emergency, external contractors may need to be notified.

2.142. In addition to providing tools for employees tracking MIS systems for signs of trouble, the physical Operations Centre can also be a place to meet to discuss the issues and responses. The Operations Centre should be a place where communication crises are managed and should provide access to facilities and equipment to record and develop other content to be broadcast on the Internet and other media. To monitor any ongoing situations and detect any new ones, the facility should, in turn, have access to television, radio and Internet broadcasts.

2.143. It is recommended that the Operations Centre be both a physical and virtual space. Should a physical space not be available (which was the case for many countries during the COVID-19 emergency), staff should have all the same tools that exist in the Operations Centre available to them while working

remotely. From a location outside the NSO the fundamental requirements for staff should be to have access to MIS and other reports and to have the ability to communicate via phone and video calls. Any capacity to respond to an emergency should not be impeded by the closure of a census management office location.

3. Disinformation and misinformation threats to the census

2.144. Censuses face a new threat of dis/misinformation, which is unlikely to go away anytime soon. Exploiting weakness in Internet literacy and mistrust in government, sophisticated players can launch campaigns to divert attention away from a census or embarrass governments by disrupting operations. Censuses now need to invest in tools to monitor social media, the most common way that dis/misinformation campaigns are conducted.

2.145. Specialized "social listening" tools can comb through large databases of publicly available social media posts to find incorrect and mischievous information being shared about censuses. When detected, the NSO can decide to intervene as appropriate with responsive communication tactics, to correct the record and re-establish the facts regarding the census. Not all situations will be the same, and some responses will require close working relationships with major social media companies. If a major dis/misinformation campaign is detected, the NSO should be able to quickly contact the platform(s) on which the incorrect information is circulating, and request that posts be removed before more damage is inflicted on the census. Other interventions may be necessary, such as countering the offending posts with information posted on the ONS's own social media accounts (see paragraphs 2.200-2.201 for discussion on how to communicate such issues).

2.146. While powerful and effective, such social listening tools should, however, be used cautiously. While reviewing only publicly available information, there is a risk that the NSO could be perceived as invading privacy. As a result, before these tools are used, privacy experts should be consulted and offered an opportunity to approve or reject the proposed social listening tool.

4. Crisis response training for census staff

2.147. While it is impossible to anticipate all possible scenarios, key census staff should be prepared for all significant risks. After establishing a risk register and emergency management protocols, prior to the enumeration of a census, census teams should practice what they might encounter and how the issue would be mitigated. These practice sessions can take various forms and can be adapted for different types or levels of census staff. For example, one might be targeted toward IT staff, while another might be for the management team.

2.148. These exercises are an effective method to condition census staff on emergency management. The exercise would involve staff members who should be made aware of all (generic) emergency management protocols already developed. With that knowledge in hand, participants to the exercise will be introduced to various scenarios that should be based on potential issues identified in the risk register. When the scenarios are presented, staff should reflect on the best next steps that should be taken. Everyone at the sessions should be expected to contribute and provide feedback on the reflections of other team members to the scenarios.

2.149. During the conduct of such practice exercises, team members may find themselves uncomfortable or uncertain about the best course of action. This is expected, as the purpose of these exercises is to familiarize the team with challenging situations and develop their ability to respond effectively under pressure. The practice sessions should include scenarios of increasing difficulty, simulating realistic challenges that census managers might encounter. While it's impossible to anticipate every possible situation, these exercises should draw upon the previously developed risk register to cover a wide range of potential issues. The goal is not to memorize specific steps for every emergency, but rather to foster a general ability within the census team to react quickly and confidently when unexpected problems arise during the census operation. By working through challenging scenarios, participants will gain valuable experience and develop the skills needed to navigate difficult situations effectively.

5. Ensuring fast workflow during an emergency

2.150. Census teams understand the importance of their work and will respond effectively during emergencies. However, without proper structure and management, some census teams may feel compelled to do "whatever it takes" to ensure success despite setbacks. It is important to note that techniques beyond hard work and extended hours can be employed in emergency situations to ensure the completion of the work without causing unnecessary detrimental effects on the people involved.

2.151. One likely requirement of a census emergency will be fast workflow and strong communications to ensure the responsible teams are aware of what is changing, so they can adapt as necessary. During a census emergency, solutions must be found quickly and acted upon appropriately. As noted at paragraph 2.137 above, once an issue is detected a quick response will be required.

2.152. The census manager has an obligation to orchestrate the census team such that the resolution to the problem is efficient and effective. Management techniques can be applied which ensure fast workflow, such as Kanban boards[70] and quick, frequent "stand up" or "sprint" meetings. Management techniques such as these should be applied in order to ensure that the full census team is working on the proper priorities, and that such priorities are communicated to the affected team members.

2.153. One way that census managers can ensure fast workflow and project visibility, is to create virtual and real Kanban boards during an emergency. The Kanban technique will have the team develop and act upon problem statements describing the response to an emergency. For example, if there has been an attempted cyberattack, a problem statement might declare an issue to be resolved, such as "Re-open Internet portal" and a team is assigned to resolve the issue. There might be related issues created by the cyberattack for the communication team or the downstream operations. After the creation of the various issue statements, they can be placed into a "to-do" list or Kanban board, which provides a visual representation on the work created by the emergency. The census manager can meet with the team regularly, perhaps in the form of quick sprint meetings, to review the Kanban board and make sure the work to re-establish regular operations is moving at a fast pace. As well, the Kanban board is an effective

[70] Kanban boards are a visual project management tool used to track and manage work. They are based on the Kanban method, which originated in Toyota's production system.

way for other members of the team to see if there are any implications of the work of another team, on their area of responsibility.

6. Contingency planning for main plan failure

2.154. Despite the NSO's best efforts, censuses can fail to reach their full objectives. Disruptions to the execution of a census are inevitable and may severely impact the expected data quality. When risks are looming, NSOs can respond in a number of ways. The best scenario is that the risk is fully mitigated such that it does not develop into an issue. A more serious situation might require the NSO to cancel or delay a census, as was the case for many countries in the 2020 round, impacted by the COVID-19 pandemic. However, if a census risk cannot be fully mitigated or avoided, the issues that are created may not always be manageable. So much can be beyond the control of the NSO, and with censuses becoming increasingly complex (as, for example, with combined censuses or those using large amounts of administrative data from multiple sources), there is always the risk that the main plan fails. What does a statistical agency do in that case?

2.155. When facing potential disruptions to a planned census, options for mitigation might be limited. For countries relying on field-based censuses, accelerating the use of administrative data could be considered, but with important caveats. While administrative data may not be a perfect substitute for all planned census variables within the original timeframe, it might allow NSOs to fulfill the basic census function of producing a population count. However, it's crucial to acknowledge that integrating administrative data requires extensive pre-planning and might not be feasible as a quick fix in the face of an immediate risk. If the risk of the main census plan failing is high, NSOs could consider developing a parallel administrative count as a contingency. This would ensure the availability of basic population figures even if the primary census operation encounters significant challenges. However, this approach requires careful planning and resource allocation to ensure the quality and timeliness of the administrative data-based count.

2.156. While it may not be considered as such, a robust administrative data plan for field-based census takers can be an effective "insurance plan" and therefore an incentive for investment in administrative data technology to manage it, and expertise within the census team on how to use administrative data properly for census taking.

2.157. For register-based censuses, a disruption of the census would likely involve a communication plan and an adjustment of the timeline of the census, to allow a re-start of the process when the disrupting event has passed. If this is the case, there would be an impact on other programmes dependent upon the census, such as an extension or temporary introduction of intercensal population estimates.

C. Change management

2.158. In each census round, when planning for a new population and housing census begins, various innovations derived from technological developments, lessons learned from the previous census, and the ongoing need to improve the balance between budget and data quality, are evaluated. When a new census involves significant conceptual changes, from a *de jure* to a *de facto* enumeration for example, or technological changes, such as transitioning from paper questionnaires to electronic forms, or to a combination of collection methods, or methodological changes from a field-based to a register-based

census, then change management becomes essential to ensure that these innovations are implemented effectively, and the desired results are achieved.

2.159. Change management is grounded in a set of principles, approaches, and strategies designed to facilitate the successful transition of an organization or project from its current state to a desired one. There are various models of change management and its components. One option that a census might consider, is the model presented by the Harvard Business School:[71]

- Prepare the organization for change;
- Develop a vision and associated plan;
- Implement the vision and plan;
- Embed the changes within the culture of the organization; and
- Review progress and analyse results.

2.160. To achieve this, it is important to understand the reasons for the change and assess the prevailing conditions, challenges, and opportunities, in order to properly prepare the organization for change and to develop a vision and plan that is achievable. For instance, when transitioning from a paper-based census to one that is more electronic, it is imperative that staff appreciate the comprehensive benefits and impacts on the census process. Therefore, it is essential that senior management actively supports and effectively communicates the change vision and is actively engaged in the implementation process. A high-level "vision" paper prepared by census management, describing the proposed changes and their reasons can be an effective reference for the census team executing the vision. The vision should be circulated and discussed, so that employee insights and concerns can be taken into consideration, and so they should be provided with training in the new skills and knowledge necessary to adapt to the change.

2.161. In order to implement the envisioned changes, detailed planning is essential, encompassing clear objectives, timelines, responsibilities, allocated resources, and risk mitigation strategies, as well as other elements described in the previous sections on resource mobilization, project management and emergency preparedness/contingency planning. This plan should also maintain flexibility to address any deviations that may occur.

2.162. Ideally, the need for change from one census cycle to the next should be sufficiently well communicated to ensure that it becomes part of the culture of an organization. That is, the driver for change is so well understood by the census team that they focus completely on the new change element, and it changes the way they think and behave. For example, in a field-based census whose past cycles have provided insufficient respondent support and burden management, the change may involve considerable new investment in respondent materials or services, and changes to questionnaire content. With a major re-focus towards making a census friendlier for the population, the result may be a major change in employee perspective, shifting it from an NSO-centric and results-oriented organization towards one which ***also*** expects a high level of service to the population and where staff are highly focused on the new service elements of the next census.

[71] Harvard Business School Online (2020, March 19), 5 Critical Steps in the Change Management Process. https://online.hbs.edu/blog/post/change-management-process

2.163. During the execution of any census, it is essential to stay on-scope with the original vision. An important part of change management is to ensure that the census team does not deviate from the original plan. To do so, census management must keep tight control over the implemented change. The original vision should be associated with a series of "baselined" assumptions and activities, any deviations from which there should be a required "change request". For example, if in the previous paragraph's scenario there was a baseline assumption that there be a reduction in the amount of time it takes to complete a questionnaire, any deviation from this plan that holds constant or increases questionnaire time, can only occur if a request for change is approved by census management. Not only does this approach manage scope creep[72], but it can also reinforce the staff cultural elements described above.

2.164. Finally, when a significant change in the census is executed, time should be taken to re-evaluate how the change management process worked. That evaluation is best considered at the end of a census cycle, when other elements are being reviewed in preparation for the next cycle. Improvements to the change management approach should be treated like other lessons learned of a cycle.

IX. Administrative organization

A. Overall overview

2.165. In planning the organization and administration of a census, particularly for a field-based census, it is important to consider the role and relationship of the various executive and advisory stakeholders. National, subnational and local commissions and committees are frequently useful in the planning and preparations of a census. Such bodies may be composed of representatives of governmental agencies, community leaders with due representation of all sections of society, and of non-governmental users of the census data, particularly those involved in policy-oriented analysis of census results and analytical studies of the social, economic and demographic situation of the country. This ensures broad-based and complete participation of the people to enable proper canvassing of sensitive issues such as ethnicity, gender, disability, fertility and migration, and of marginalized groups. It is important, however, that their executive and advisory functions be clearly defined and that the final responsibility for planning and execution rests with the executive agency.

2.166. There are well-documented and proven advantages in having an office continuously responsible for census work established as an integral part of the statistical system of a country. Such an office (whether it is a constituent part of the NSO or a separate government department) assures continuity in census work and is the principal centre for the formulation of the programme and the initiation of preparatory work for the next census. Its permanence permits the development of specialized and experienced personnel and the maintenance of statistical and cartographic information, including cross-cutting issues such as information technology, essential for planning the next census.

2.167. At the pre-enumeration stage, the census office will need to be expanded to form the nucleus of the full census organization, which must be capable of directing the field organization during the

[72] Scope creep in project management is continuous or uncontrolled growth in a project's scope, generally experienced after the project begins.

preparatory work as well as during the enumeration and processing stages. In order to provide immediate supervision in each area, field offices at various levels are needed for the later part of the preparatory work, including staff recruitment and training, as well as for the enumeration period. Supervisory personnel in such offices should be persons who, being familiar with the particular area and the local language, are able to deal with local problems. This does not mean, however, that all supervisory positions need necessarily be filled by persons from the area. Personnel may be transferred from the central office or from other areas as the need arises. Prior training of all such available personnel is necessary, so that they have a working knowledge of all aspects of the current census programme. An essential part of the preparatory work is local administrative planning, which would set out the likely problems and challenges specific to the local area and how they are to be dealt with. Adequate coordination with local public authorities is always important so that the enumeration work is not interrupted by other activities.

2.168. Subsequent to the enumeration, the census organization may be readjusted and staff re-assigned to meet the needs involved in compiling, evaluating, analysing and publishing the results and to provide the continuity desirable for promoting the continued use of census materials. Census organizations need to pay special attention to continuity of knowledge and skills from one census to the next, since the intervening gap, which is usually a whole decade, is likely to cause loss of institutional memory and attrition of qualified personnel. Comprehensive documentation of census activities while they are being carried out is essential, as is training of younger personnel to create a pool of knowledgeable and experienced persons by the time the next census comes.

B. Roles and responsibilities of key stakeholders in the census process

2.169. Population and housing censuses require the involvement of several stakeholders. The roles and responsibilities of key stakeholders in various phases of the census process should be clearly defined and adapted to the local situation.

2.170. The role of each of the key stakeholders can be summarised as follows:

Government

- **Proclamation and resource allocation**: Issue a formal proclamation or decree to initiate the census, providing sufficient lead time for preparatory activities. Ensure the necessary allocation of financial, human and physical resources at all phases of the census process.
- **Leadership**: Ensure that census data serve the public good by contributing to the improvement of people's wellbeing. Facilitate collaboration among government agencies.
- **Supportive environment**: Create a conducive environment for the census, reinforcing its strategic objectives and protecting it from political or other influence that might compromise its integrity. Implement security measures to safeguard personnel and physical assets involved in field operations.
- **Data provision and security**: Provide essential administrative data for the census. Ensure the security of data collection, transfer, and storage in collaboration with data protection agencies.

NSOs

- **Coordination and engagement**: Ensure effective coordination across government agencies. Engage with various stakeholders throughout the census process.

- **Census design and execution**: Choose appropriate enumeration methods based on the national context. Make efficient use of the allocated resources. Ensure proper management of human resources, logistics and procurement.
- **Capacity building**: Strengthen its capacity and technical resources to ensure high data quality and compliance with international standards.
- **Data dissemination, utilization and security**: Promote the effective dissemination and use of census data, ensure public access while maintaining the safety and security of data.

The United Nations and other international organizations

- **Standards**: Coordinate the development and promote adoption of international principles, recommendations, standards and classifications for population and housing censuses. Collect national experiences and lesson learnt in census taking and evaluation to inform the development of these standards.
- **Technical assistance:** Provide technical assistance to NSOs. Identify and prioritize data gaps. Contribute to capacity-building and knowledge sharing, including South-to-South and triangular cooperation[73].
- **Advocacy**: Advocate for the census as a public good that contributes to improving people's wellbeing. Advocate for evidence-based policy making.

Donors

- **Financial support**: Provide financial assistance to NSOs within agreed-upon scopes.
- **Technical assistance**: Provide technical assistance to NSOs. Identify and prioritize data gaps. Contribute to capacity-building and knowledge sharing, including South-to-South and triangular cooperation.

The private sector

- **Technical assistance**: Supply goods and services essential for census implementation. Promote the adoption of new technologies in the census process, in accordance with industry best standards. Identify and prioritize data gaps. Contribute to capacity-building and knowledge sharing.
- **Financial support**: Provide financial assistance to NSOs within agreed-upon scopes.
- **Data provision**: Where appropriate, and with necessary safeguards in place, explore opportunities for collaboration with the private sector to access data that can support census operations. This might involve purchasing or accessing through data-sharing agreements, private sector data such as mobile location data or utility usage records, which can be valuable for understanding population mobility or identifying signs of life.

Academic and research institutions

[73] **South-South cooperation** is a broad framework for collaboration among countries of the South in the political, economic, social, cultural, environmental and technical domains. Involving two or more developing countries, it can take many forms, including the sharing of knowledge, skills, expertise, and resources. **Triangular cooperation** is collaboration in which traditional donor countries and multilateral organizations facilitate South-South initiatives through the provision of funding, training, management and technological assistance. More detailed information and resources are available on the UNOSSC website: https://www.unsouthsouth.org/

- **Data utilization**: Analyse census data and produce knowledge products to inform policies and programmes for sustainable development. Identify and prioritize data gaps.
- **Technical assistance**: Provide technical support to NSOs for data analysis and report writing. Participate in independent evaluation for census process and quality.
- **Data provision**: Where appropriate, share administrative data on school attendance and educational attainment with NSOs to produce census data on educational characteristics.

The media

- **Advocacy and communication**: Raise public awareness on the value and purpose of the census. Encourage public participation in the census. Publicize the needs of personnel to be recruited and the field operation. Provide mechanisms for the promotion, dissemination and use of census data. Contribute to the national discussion to identify and prioritize data gaps.

Civil society organizations relevant to specific population groups

- **Advocacy**: Advocate for the data needs and awareness for specific population groups to ensure inclusivity in the census. Promote the responsible and effective use of census data. Encourage public participation in the census.
- **Data provision**: Where appropriate, share administrative data on specific population groups (e.g., the location of homeless individuals, as provided by organization supporting homeless people) to ensure that these groups are enumerated in the census.

The general public

- **Participation and feedback**: Provide complete and accurate responses to the census questionnaire(s) and demand accountability for the use of census data to improve population well-being. Offer feedback to NSOs on data management, the acceptability of census content, and perceived trust in the organization.

Users of census data

- **Role of data users:** Data users play a crucial role in the census process. While they may also belong to other stakeholder groups, their unique perspective as data users provides valuable input, particularly in shaping census content and determining how data is disseminated. Their needs and priorities should be carefully considered throughout the census lifecycle. (Further details on the role of data users are provided in Section X.)

C. Statistical leadership

2.171. Every phase of the census process presents a unique opportunity to demonstrate leadership in promoting the use of statistics for people's wellbeing through improving service delivery and policy development. Statistical leadership should be developed and anchored in international standards and guidelines, such as the Fundamental Principles of Official Statistics[74], national statistical legislation and the national statistical code of ethics.

[74] United Nations General Assembly. (2014). Fundamental Principles of Official Statistics. Resolution 68/261. New York: United Nations.

2.172. The leadership of the census organization should ensure a successful census programme that delivers results for evidence-based policymaking, research, and decision-making. It is therefore recommended that the leadership undertake the following activities:

(a) Facilitate the establishment of statistical legislation that mandates the census;
(b) Set policy and strategy by defining targeted outputs and outcomes for the programme;
(c) Engage strategically with stakeholders, mobilizing participation across governments, stakeholders and the public;
(d) Raise the profile of, and commitment to, the use of statistical information, creating opportunities for engagement on key policy issues and strengthening relationships between information providers, policymakers and opinion leaders;
(e) Adhere to the Fundamental Principles of Official Statistics and ensure best practices are embedded in statistical procedures;
(f) Align with international standards, classifications and frameworks;
(g) Establish the statistical infrastructure and resources for conducting the census;
(h) Set up the management structure to plan, implement and monitor the census.

X. User consultation, communication and publicity

2.173. The extent of public consultation, communication and publicity can vary enormously according to the type of census being undertaken. Most of this section relates to activities associated with a full field enumeration or combined census with a field operations component. Countries conducting fully register-based censuses that require little or no direct participation by the general public should implement a communication and publicity strategy with messaging that aligns with the register-based approach.

2.174. A comprehensive programme of communications and engagement for a population and housing census covers three distinct audiences: (a) major users of census data, (b) persons and institutions participating in the census operations and (c) the general public, including particular population sub-groups. Since the field-based census is a national activity that is completely dependent for its success upon the cooperation and support of the general public and many governmental and local organizations, the entire communications and engagement effort should be developed as a coordinated activity in close coordination with the other substantive preparations for the census. These communications and engagement activities are valuable not only for informing stakeholders about the census but also for providing census authorities with early and continuing information about the reactions to census plans and activities of the general public in various parts of the country and of key population groups and institutions. Many NSOs use design thinking approaches to create better respondent experiences. This involves developing user personas to segment population groups and better understand differing needs, expectations and behaviours.

Consultation with data users

2.175. Extensive consultation with users and advocacy groups with an interest in census data on topics, definitions and, particularly, planned tabulations and other outputs, and the development of the census database is an indispensable step in the preparations for the census, which should be conducted early in the planning stage. These consultations will assist the census authorities in designing a census that, within the resources available, is as responsive as possible to user needs in terms of the collection, processing,

tabulation, storage and availability of meaningful statistics. Such consultations can also serve to foster a wider and more informed understanding of, and support for, census plans and activities. The users to be consulted should be from governmental departments, ministries, universities and other research institutions, the private sector, and other organizations (or individuals) representing the economic, social, educational, commercial and cultural life of a country. Many countries will want to include among the groups to be consulted organizations or bodies representing culturally diverse communities, ethnic and indigenous communities, refugee and internally displaced community groups, and religious groups, those with accessibility requirements and people with disabilities, housing associations and agencies supporting homeless people. Other key stakeholders may include partners with whom the census office collaborates for the provision of specialist services, and donors who may help fund elements of the census operation.

2.176. Taking into account the importance of the census in providing data for local planning and administration, it is also often advisable to have consultations with users in provincial and local governments and institutions in various parts of the country, particularly in remote rural areas. In large countries or countries where the provincial or local governments have a comparatively high degree of autonomy, consultation with users at the subnational level is essential if the full potential of the census is to be achieved. Strategies should be chosen according to the target group.

2.177. The consultation process can take many forms. If done in the form of meetings, it is often more useful to hold separate consultations with different types of users with common interests, such as administrators, policymakers, planners, demographers, academic researchers, users in the business community and so forth, rather than conduct a simultaneous consultation with all data users. Consultations involving different types of stakeholders in the same setting frequently prove frustrating to participants because there are substantial opposing differences among users in their data requirements, technical background and in their concern with the details of census content and operations.

2.178. Meeting data users in person is very informative but imposes physical and budgetary limitations. Broad consultation can be conducted online and via social media, e-newsletters, and digital stakeholder outreach. The strategy can be used both to identify users' requirements and also to provide transparency in the census preparation activities. Other forms of technology can be considered for holding decentralized or remote consultations. Users may be sent an electronic questionnaire to collect their priority data requirements, or invited to submit such requirements online. Digital channels, including social media engagement, offer the opportunity to reach a substantial number of users if accounts have adequate reach. Video content may be employed to support or promote consultation materials, and webinars can be an effective way to meet with large numbers of users remotely.

Communication with other stakeholders

2.179. In the context of conducting a census, it is essential to establish two-way relationships with key stakeholders at the lowest territorial units such as municipalities. These relationships will facilitate continuous communication and awareness throughout the entire implementation and development process of the census. Additionally, a meticulously crafted message tailored to the social context of the region should be developed to introduce the census to various interest groups effectively. This lays the foundation for ensuring the entry of field personnel into information sources, which is essential for successful awareness-building and data collection.

2.180. Once these relationships have been established and the census operation has been effectively presented, it becomes necessary to focus on the structured dissemination of census information. This

becomes a pivotal element in mobilizing society towards the production, understanding, and utilization of statistical information that will benefit the country. To achieve this purpose, efforts to raise awareness must be strengthened, aiming to prevent disinterest and disengagement and ensuring the collection/acquisition of information with the required coverage and quality. In this way, the objectives of collecting complete, accurate and valuable data for the benefit of all can be achieved.

2.181. In order to complete the preparatory work for a census involving a field enumeration and to carry out the enumeration itself, large numbers of temporary personnel will have to be recruited and trained and the contributions of a diverse group of national and local organizations will have to be effectively mobilized. A well-planned recruitment campaign and programme of engagement with external organizations can contribute to both efforts.

Communication with field personnel

2.182. The training and capacity development of field personnel play a pivotal role in the census process and necessitate effective communication, in much the same way as with various data users. To achieve this, it is imperative that data collectors comprehend three key aspects. Firstly, they must grasp the significance of statistical information production for the nation and the importance of censuses, along with a comprehensive understanding of the context in which field operations will take place. Secondly, those responsible for data collection must receive guidance, conveyed through communication, emphasizing the importance of a commitment to data collection quality. Lastly, communication must ensure that census enumerators are well-versed in:

(a) the content of the census questionnaires,

(b) general statistical operation concepts, and specifically, those related to cartography, and

(c) the appropriate use of mobile data capture devices in countries employing such data collection technology.

Therefore, it is necessary to prepare a field data collection guidebook with sufficient teaching materials and targeted teaching strategies.

Public outreach and information campaigns

2.183. An effective communications strategy, together with far-reaching publicity and information campaigns, play an essential role in ensuring the success of the census. This is especially so for those countries adopting a field enumeration methodology, either wholly or in part, where the general public is expected or required to participate actively in the census activities as respondents. Particularly in the case of countries that undertake a significant field operation, public acceptance and cooperation is essential to ensure the success of the census. It is crucial to underscore the vital importance of securing active participation from each data source or qualified informant in the census process. This translates into an imperative need to avoid definitive refusals in order to ensure the collection of information with the required coverage and quality. In specific cases where conventional communication channels may not be accessible, establishing a more personalized form of doorstep communication with the individuals to be enumerated becomes necessary. The aim is to ensure the presence of qualified informants when delivering the questionnaire.

2.184. As part of the communications strategy, countries should consider both proactive and reactive media management as well as planning engagements across multiple channels. The use of a call centre to deliver help and support via telephone and providing chat support for online enquires should be

considered. Consideration should also be given to the appointment and training of census spokespeople to communicate key messages consistently in local media and through community engagement activities.

2.185. While there may be occasions when door-to-door communication is necessary, a large-scale publicity and information campaign is recommended to inform the population about the census and to explain its purpose. Opportunities may be taken throughout the campaign to monitor feedback from the general public, particularly through social media, both to help shape future messaging and where applicable to inform the census operation. Designing and implementing the publicity programme is best undertaken by professional experts in the field of public relations, advertising and sociology. Such expertise is frequently not found within the NSO itself, and it may therefore be appropriate to outsource some or all of this work. The publicity programme may include:

(a) The development of a census slogan and logo
(b) A public relations campaign;
(c) A community liaison (or outreach) programme;
(d) An advertising campaign;
(e) Monitoring of public opinion;
(f) Media relations, including monitoring of the mass media;
(g) A social media strategy.

Key messages for public campaigns

2.186. There are several main messages that census agencies will need to communicate to the public in order to optimize outcomes for the census. Census publicity campaigns should encompass a wider set of messages, whose components might include:

(a) making the public aware of the census;
(b) educating the public about the benefits (to them and to the country) of the census, in order to improve sentiment towards participation;
(c) reminding people about their legal obligation and duty to take part in the census and (where appropriate) the penalties for refusing to do so;
(d) mitigate the risks of digital exclusion by explaining to the public what to do and when, and providing information on the practical support available to those who may face barriers to responding;
(e) informing and reassuring the public that privacy and confidentiality will be protected; and,
(f) expressing thanks to the public for taking part in the census.

2.187. Care is necessary in finding the correct balance between these different messages. For example, an overemphasis on the obligatory nature of the census (and any penalties for non-compliance) may serve to reinforce negative perceptions that the census is an imposition by the State on the population, rather than an activity for the common good.

Objectives of public campaigns

2.188. Publicity for a census operation entails an educational campaign, the purpose of which is to ensure the awareness, interest and cooperation of the general public, particularly population groups that may be reluctant to be enumerated. The aims, as a general rule, are not only to dispel any anxiety regarding the purposes of the census but also to explain the reasons for the various questions in the questionnaire, including the value census information provides by contributing to an informed data driven society, and to offer some guidance on how questions should be answered. The publicity campaign may

also be an important tool for increasing the completeness of census coverage, particularly among groups with lower coverage rates in previous censuses.

2.189. Planning for the general publicity campaign should start as soon as the census is authorized. The campaign itself should be closely synchronized with other census activities and full-scale publicity should not begin too far in advance of the date on which the enumeration is scheduled to start. However, plans for the publicity programme should be formulated early enough to take effect for the purposes of any census tests. In addition, the programme can use these tests to study the impact of alternative publicity materials and methods. If either the cartographic or house-listing operations require extensive fieldwork and widespread contacts with the public, it should be recognized that the personnel involved in these activities often provide the public with its first impression of the census. Training and publicity programmes should take this into account.

Strategies for reaching specific population groups

2.190. To the extent that the census communications budget allows, the general campaign should be directed to all sections of the country and all segments of the population through the use of all available publicity media, with special emphasis on the use of digital media channels. The general campaign may be supplemented by a number of specialized campaigns aimed at specific segments of the population to raise awareness on specific subjects, such as gender, migration, ethnicity and disability, in which the quality of response may depend on the level of prior awareness among the public at large or among the specific groups concerned. In multilingual countries, creating campaigns in the various languages used within the country (both official and non-official) may be necessary in order to ensure all communities have the opportunity to understand and participate in the census.

2.191. Publicity can be secured at low or no cost through social media and through public relations with national and local mass media, such as newspapers, television and radio. It is recommended that a range of spokespeople from the census organization are trained and made available to speak to national and local media, and as well as for community engagement activities. Paid-for advertising through various forms of media can also be effective. Interactive media may also be used such as a toll-free helpline and text messaging. Local events, conducted in local languages, where the public can participate can also improve public awareness and build trust.

2.192. Disseminating information about the rationale of the census and its utility helps allay possible misconceptions among the general public, thus increasing participation and coverage. Outreach campaigns involving a range of organizations and enlisting the support of local leaders and opinion makers to spread the word about the census in their area of influence are also good strategies. Maintaining ongoing relationships with community organisations and representatives in the period between censuses can be of mutual benefit, with the potential to improve engagement not only during the census operation but also during the consultation and dissemination phases. In addition to recruiting such organizations and leaders as partners, the census organization may develop key messages, web content and links, printed material and other material to support them in their activities. The use of publicity may also be considered to support the recruitment of field personnel though it is important to clearly differentiate between a recruitment campaign to hire field staff from among the general public and a general campaign to encourage participation discussed in paragraphs 2.183-2.189 above.

Branding and inclusiveness

2.193. For countries running a field-based census operation, it is recommended that a census "brand" be established, including a logo and tag line. Census branding is important for establishing awareness and trust. A simple but effective slogan and distinct logo can be used in all national and local advertising campaigns and in all types of media, booklets, posters, brochures and other promotional materials. A slogan and logo that are well recognized from the initial stages of the publicity campaign may serve to improve "brand recognition" for the census. The aim should be to encourage the respondent to feel more reassured that the census is an inclusive and beneficial activity.

2.194. An important first step is to identify population groups or communities that may be less likely to take part in the census without additional interactions. These population groups will vary from country to country, but examples could include transient groups, recent migrants, young people, older people, people living with disabilities, the digitally excluded, the homeless, people living with literacy and language difficulties, and inhabitants of high-rise apartments and dense urban areas. Having identified such population groups, NSOs should consider research and insight-gathering to understand the reasons that may prevent communities from taking part and design strategies to address these. An approach that has been successful in many countries is for NSOs to speak to organizations and charities that support the interests of these communities. Such reasons are likely to be able to be classified as motivational or practical.

2.195. In rural areas where digital channels may be limited, weekly markets, fairs and public festivals provide good opportunities to publicize the census message among people who may not have much exposure to mass media. An excellent opportunity exists to create widespread awareness of the census through a campaign targeted at schools. These campaigns can be targeted to areas where high proportions of people face barriers to taking part. Census organizations can develop free education resources (in multiple languages, as appropriate) which focus on educating pupils on the census, the use of census data to shape services locally and nationally, and the importance of everyone taking part. They can also include take-home materials and activities providing children with the opportunity to tell family and guardians what they have learned. These campaigns have proven to increase awareness of the census among parents in some countries. Other kinds of local-level publicity, such as wall writing and village announcements, can be planned according to local circumstances.

Communication on the use of online self-response mode

2.196. When adopting an online self-response questionnaire, the census organization must communicate and engage with the public effectively. For example, a webpage enabling online self-response should be developed and tested through user consultations to ensure a positive user experience, which will contribute to higher response rates. The website must also be equipped to securely handle a large number of completed submissions concurrently. Access codes for the online questionnaire must be distributed to all households and collective living quarters, accompanied by clear instructions on when and how to respond, along with reassurances about the security and privacy of their responses. The general communication/publicity campaign should reinforce these messages. Additionally, it is important to provide clear guidance to form-fillers on how to handle technical difficulties that may arise during the online census process, including information on available support channels and alternative methods for data collection.

Communication for changing census processes

2.197. Where there are significant process changes to the census operation such as the introduction of an online response, the public communications campaign should be tailored to reflect such changes. The changes should be communicated through a variety of traditional and social media channels in an accessible manner such that reach all sectors of society. The messages should address a range of issues including: the benefits of moving to an online census for both the respondent experience and the quality of the data; the reduced impact on the environment; and the increased the efficiency of the data collection operation. Clear information on the sequencing of collection activities should be given from initial communication (for example via publicity, mailout, or the enumerator) of details on: how to access an online or paper form; how to complete an online return or to receive support in completing such a return; or how to request a paper census form. The public should also receive assurances about the security of their data completed using the online mode. Communication strategies should also address changes beyond data collection methods, such as the introduction of new census topics or modifications to the questionnaire due to resource constraints.

Support for completing the questionnaire

2.198. Support should be made available to anyone who may face barriers to completing the questionnaire, regardless of the reason. Providing detailed explanations on how to fill out the questionnaire will be very helpful for both paper-based and online respondents. This can help prepare respondents for the enumeration process and reduce potential difficulties. Materials supplied to households should clarify the various routes available to individuals (such as completion online, by telephone, or the option to request a paper questionnaire). Websites containing frequently asked questions, or chatbot [75] services, can be a cost-effective way to answer common questions from respondents in the first instance.

2.199. Additionally, NSOs may wish to consider providing tailored support such as webchat services, social media, SMS, telephone advisors, or email. Consideration should be given, where appropriate, to alternative language options for all such resources. In countries where censuses are primarily conducted online, in-person support centres may assist people without Internet access. All guidance should be updated in response to emerging themes and decisions before and during the live census operation. Real-time management information (MI) from online responses (and slower MI from paper questionnaires) can be used to deploy the field force more flexibly and responsively towards non-responding households. It is advised that the field force be provided with guidance in responding to common questions and issues, to support households to respond.

Managing misinformation and incidents

[75] A chatbot is an advanced computer program that offers automated responses to frequently asked questions, while the live chat connects users with an agent for assistance when the chatbot cannot address their specific query. These features aim to improve and supplement the existing services provided to citizens by creating a positive, secure online experience that assist respondents with their census questions. The chatbot provides faster response times by answering simple FAQs quickly and automatically while reducing the burden on help desk staff. Extensive usability tests should be conducted to assess the quality and effectiveness of the chatbot and live chat options.

2.200. NSOs should develop a strategy for managing misinformation and other incidents that may impact on census operations (see paragraphs 2.144-2.146 for disinformation and misinformation threats to the census). Communications staff should be included in strategic discussions during operations to support the organizational response to issues as they arise. During operations public opinion and mass media should be monitored to assess the effectiveness of publicity campaigns and identify such issues. Public opinion could be monitored through surveys that can provide information on public attitudes to the census, or through interaction on digital channels including social media. Media monitoring refers to tracking, analysing, and measuring various media channels. These channels can include both traditional and digital media. Analysing the data from these sources helps NSOs understand public perception and identify emerging challenges related to the census. Such monitoring implies an ongoing accumulation of information, detection and prevention of the development of negative published comments on the census, and preparation of adequate responses to negative reports and information. Increasingly the media has a significant influence on people's behaviour and even minor distractions and mistruths can have a detrimental effect on the outcome of the census. Therefore, in developing their publicity campaigns, NSOs should give particular attention to preparing for unexpected events (such as negative attitudes, malicious lobbying, technical difficulties, delays and misleading information; see paragraphs 2.137-2.143 for discussion on preparation for such events).

2.201. 'Lines to take' should be developed and updated to respond to emerging issues, and may be used on social media directly in response to misinformation on such platforms. Direct communication with social media companies and establishing procedure to work with these companies, as well as points of contact, is also encouraged in advance of a census operation, to develop processes for identifying and managing misinformation. Public relations with mass media can also be used to challenge or alert the public to misinformation. It is also recommended that all official participants involved in census operations know their roles in the communication process both with the media and with the public at large.

Post-census communication

2.202. An integral part of census communication and publicity is informing key census data users and the general public about the availability of the census results and their utility. Awareness about the range of census data and products to be made publicly available should be raised before the enumeration so as to ensure that the public recognizes the importance of the census and appreciate statistics that are generated from it. It is critical that such communication strategies be developed as an integral part of census planning and not left as an optional add-on. It has been the experience of quite a few countries that the engagement of professional media and communication personnel adds value to the campaign. While a dialogue with users is necessary throughout the entire census process, consultation with users is particularly important as part of the concluding evaluation programme to gather feedback and inform future censuses.

XI. Census calendar

2.203. A calendar or timetable indicating the sequence and estimated duration of each component of the census process is indispensable in census planning. A provisional calendar with selected key dates should be prepared as a starting point of the census preparation. This calendar should be shared with stakeholders in advance for advice and support. As planning progresses, the calendar will inevitably be revised and made more detailed, with the aim of establishing final dates. However, adherence to major

milestones and deadlines is crucial, and any necessary changes must be communicated transparently to the public and the stakeholders, with clear justifications for any adjustments.

2.204. Such calendars are essential because they indicate the start and completion dates for the numerous components that make up a census, serving as both a pacemaker and a performance indicator for census operations. Key dates on a census calendar are sometimes referred to as the 'Census Critical Path,' a series of critical milestones that significantly impact other programme elements. Serious delays can be detected by comparing the target dates with the actual achieved dates. When a modification to a specific activity become necessary, all related activities will be impacted and should be correspondingly adjusted to avoid further disruptions to the whole census programme.

2.205. For censuses conducted with a field-based enumeration, the calendar usually groups operations into three broad phases: pre-enumeration, enumeration and post-enumeration. Similarly, the calendar may refer to stages (input, process, output) in the context of censuses compiled from administrative registers. Dissemination activities are the primary goal and ultimate outcome of the entire census operation. Release schedules should be a more visible and explicit part of the census calendar. The calendar should include the planned release dates of census outputs, as well as the data processing and dissemination activities. A critical date on the census calendar, upon which the scheduling of all other operations hinges, is the starting date for the general enumeration of the population. For purposes of control, many overlapping operations can be shown separately in the calendar. Census calendars can take the form of charts or graphs, in addition to detailed checklists of operations. Project management software may help in the preparation of the census calendar.

2.206. In establishing the census calendar, it is necessary to consider the relationship of the population and housing censuses to one another as well as to other statistical projects or other large-scale national activities. Although a joint population and housing census operation is likely to constitute, for the period of its duration, the major statistical undertaking of the government, care should be taken that it does not interfere unduly with the other regular national statistical activities that may be going on at the same time. A balanced statistical programme should avoid having too many simultaneous competing enquiries, which might place too heavy a burden on the statistical services and on the public, with a possible resultant loss of both administrative efficiency and public cooperation. Moreover, other events that may conflict with the field enumeration such as national, local, or regional elections, and public or religious holidays, should also be avoided so that the general population is not distracted.

2.207. It is often useful to draw up a comprehensive diagram showing the sequence, interrelationship and timing of all the various steps in the census programme – a Gantt chart and a critical path would be a good example.[76] This type of analysis often reveals the consequences of a delay at one step in terms of delays to other steps in the programme. It can therefore be a useful instrument against which the actual progress of the census preparations may be compared. Indeed, some countries have attempted to use such critical path analyses not only as an aid to census planning but also as a tool for the ongoing management of their census operations. In these instances, it is essential to establish procedures for

[76] The Gantt chart was developed around 1910 by Henry Gantt of the United States, based on the work of Karol Adamiecki of Poland. It is a type of bar chart that illustrates a project schedule and is available in a number of office software packages.

revising the critical path analysis in response to actual progress. It should be stressed, moreover, that the usefulness of such devices depends on how soundly they are designed, applied and understood.

2.208. Project management software can be useful in linking the diagrammatic structure of census operations with information about nodes or centres of responsibility for individual broad or detailed operations so as to control the chain of responsibility. Alternatively, event calendars can provide a broad view of the steps of the census programme and allow follow-up. Different tools can be found on the Internet for download or online use. Online versions allow immediate update and make it easier for staff to work in a group but are dependent on Internet access. Other tools, commonly referred to as groupware and collaboration software, as well as Internet and social media forums, can support census operations by providing an environment for exchange of information, files, and data among dispersed teams. A clear monitoring and evaluation plan of the calendar should be in place and regularly reviewed.

2.209. For countries transitioning from a field-based census with paper questionnaires to one with electronic questionnaires, the census calendar should take into account the need for a sufficient period in the pre-census stage for system development. This is because the process of designing and testing the collection systems is more complex and lengthier than that for paper-based questionnaires. Therefore, it is important to allocate enough time for this stage to ensure a smooth transition to electronic questionnaires. In addition to the added time needed for electronic questionnaire design and testing, there are other aspects of electronic data collection that require time, such as the implementation of robust data security measures and the evaluation, selection and procurement of the necessary technology. Finally, electronic questionnaire use may require legal evaluation, added communication to the public and increased training for staff also using the platform.

2.210. The census calendar must also allocate sufficient time for the preparation of geospatial information that will support the census operation (see Part Three, Section IV). The preparation of up-to-date census maps is a critical and time-consuming process that requires careful planning and execution. Sufficient time allocated for map preparation ensures the accuracy, completeness, and effectiveness of the census process, leading to reliable and actionable data for decision-making.

XII. Human resources management and training

2.211. Field-based census taking requires a large number of people to function properly. Early arrangements and good human resource information systems are necessary to secure the proper number and type of personnel required for each of the various census operations. While the preparatory and processing work generally calls for office employees possessing, or able to learn, certain specialized skills (cartographers, coders, data entry operators, programmers and so on), in a field-based census the enumeration stage usually demands a large number of persons capable of going to their assigned urban or rural EAs and collecting the information according to specific definitions and instructions. The number of enumerators required being quite high compared to normal staff strengths, and the period for which their services are needed being rather short, demands that the method of recruiting them needs to be worked out carefully in advance to facilitate quick, simultaneous and transparent recruiting, and subsequently remunerating them and relieving them of their duties promptly and efficiently. Good human resource management of temporary census staff can benefit other NSO programmes or subsequent censuses, as they are already trained and can sometimes be retained for non-census projects and may form part of the next census team.

2.212. Human resource information systems can greatly assist the management of the large temporary workforce of a full field enumeration census. These systems can help streamline and optimize recruitment and deployment of new staff and can help identify personnel with relevant skills who can be cross-trained for additional tasks, or for redeployment for other non-census activities.

2.213. When electronic means of enumeration are utilized, careful consideration should be given to the computer skills of potential census personnel. This might involve prioritizing candidates with existing computer skills to ensure they can effectively utilize the technology and/or providing comprehensive training to equip all personnel with the necessary computer skills for electronic enumeration, regardless of their prior experience. It is essential that the enumerators and, to the extent possible, their immediate supervisors be conversant with the languages or dialects of the area in which they will be working, and the knowledge of these languages should extend beyond any official languages in which the nation provides services. In addition, attention should be paid to efficiency, economy, experience, ability to read maps and communication skills in general. It is only prudent to recruit and train sufficient reserves of people for a variety of temporary census positions, in order to make contingencies any attrition that may occur in the process.

2.214. Once the cartographic preparations are substantially complete and the questionnaire has been sent for printing, if applicable, perhaps the single most important means that the census authorities have for influencing the success of the census is its training programme. The contribution that a well-planned and executed training programme can make to the quality of the census results cannot be stressed too strongly. Such a training programme must of course focus on the widely dispersed and difficult-to-supervise field staff (namely, the enumerators and their immediate supervisors) but it must also cover others (for example, the higher-level supervisors, editors, coders and computer operators).

2.215. Giving office employees who are working with the census preparations a brief, uniform basic training of all aspects of the census has two prime advantages: firstly, all personnel can understand the importance and the context of their part of the task; and secondly, since they are conversant with the basics, they can be more efficiently deployed in the field for supervision or coordination during the actual census operations whenever needed. The nature of the training can be traditional (classroom style), e-learning (self-training or live training via the Internet) or combination of these approaches.

2.216. The entire census training programme should comprise different courses that are designed to cover each phase of the work and provide an efficient and consistent means of effectively equipping large numbers of fresh employees with the necessary skills. The programme will need to correspond closely to the needs of the various operations and, where appropriate, may include both theoretical and practical instruction, with emphasis on the latter. In the case of the enumerators and their immediate supervisors, the training is most effective if it includes several opportunities for the trainees to participate in practice interviews and role-playing exercises, including the use of adopted IT solutions, if any. In countries in which multiple languages are used, the method and content of the enumerator training programme will need to be suitably adjusted. For example, if the questionnaire is printed in another language, provision will have to be made for instructing enumerators on the correct formulation of the census questions in the vernacular. In countries with large geographies or a variety of regions, it may be necessary to prepare training material for local areas, which is specific to the issues that are faced in those regions.

2.217. Enumerators and supervisors should be trained as close to the field operations as possible so as to avoid recall lapses. This leaves very limited time for conducting the training. Therefore, the logistics need to be worked out carefully in advance. The training programme for editors, coders, operators of data

recording equipment and so forth should also provide opportunities for the trainees to practise under the supervision of the trainers. The intermediate- and higher-level technical staff, such as programmers and system analysts, should also be given special training with emphasis on recent technical developments of relevance to the forthcoming census and on the interrelationships among the various aspects of census plans and operations. Thorough training in census practices is an extremely important component of quality assurance. Detailed and clear documentation of instructions with appropriate illustrations is a basic requirement in this regard. A proper training methodology and a variety of training aids will go a long way in enhancing the training effort.

2.218. The organization and conduct of training courses should be entrusted to those having the necessary qualifications to carry out this task successfully, taking into account not only their professional abilities but also their ability to teach. This means that staff in charge of training should have certain qualifications that will enable them to stimulate the interest of trainees and to transfer the required knowledge, since otherwise well-qualified technical personnel who are unable to transfer their knowledge to the trainees in a satisfactory manner will be unsuitable as instructors for group training activities. This must be taken into consideration when selecting instructors and it is recommended that objective criteria should be used. In practice, however, it is difficult to find the necessary number of instructors who have both the professional and the teaching qualifications; for this reason, any outsourced instructors selected should themselves undergo training in how to organize and conduct training courses. The use of professionally designed training guides can add immense value to the training effort. The involvement of experienced professional experts in the design and delivery of training programmes is also very useful. It should, however, be noted that the content should be the responsibility of the NSO and not that of outsourced instructors.

2.219. It is important that training manuals for each training programme are made available to the census managers and training instructors. Such manuals would be a valuable guide and would help considerably in the efficient training of census staff. They would also contribute to the uniformity of training, which is an essential factor for a successful enumeration, taking into account the large number of census instructors required. Simple audiovisual aids (for example videos, posters, audio recordings) can also be used to help make the training more effective and uniform throughout the country. If available, new web-based or multimedia technologies can facilitate the provision of training at distant locations and be effective and efficient supplementary tools to face-to-face training. Standardized training may also be provided in e-learning format on the Internet and on handheld devices.

2.220. It is very important to determine the time required to train staff for the various aspects of the census. This depends on several factors, such as the task for which they are being trained, the complexity of the content, the educational level of trainees, the number of instructors available and the funds available. Apart from fixing the number of days for training, it is also important to allocate appropriate time for each subject. Experience from previous censuses has shown that it is very often the case that the time required to train field staff adequately is severely under-estimated. There is, however, a balance to be made between the allocation of sufficient time and the cost of training. Drawing up lesson plans for each session of training is an effective way of ensuring that all subjects are covered, with the right amount of time being devoted to each. Enumerator training should also include how to handle distrustful respondents, data confidentiality and privacy.

2.221. Register-based censuses may have a human resource challenge very different from the scale-related issues of managing a large workforce needed in field-based censuses. Register-based censuses can involve relatively few people with specialized skills. Therefore, knowledge retention and succession

planning, as well as the need for good documentation of processes, as well as specialized training, may be a greater challenge for countries with this type of census.

XIII. Logistics management

2.222. A field-based population and housing census requires efficient logistics management, covering procurement, storage, distribution, recollection and disposal of materials and equipment.

2.223. Logistics management is the process of planning, implementing and controlling the flow of census materials and equipment needed for implementation of census operations. Effective logistics planning requires careful coordination between different phases of the census operation, such as mapping, training, field enumeration, data processing and dissemination. The scope of the logistics programme usually differs from one country to another, but generally covers the following activities:

(a) renting central and field offices;
(b) installing furniture and equipment;
(c) providing help desk support;
(d) mobilizing enumerators and supervisors, particularly in rural and dangerous areas, and
(e) distributing, recollecting and disposing of materials and equipment.

2.224. The NSO may need to establish a dedicated team for planning, implementation and controlling the logistics programme. The functions of this team should be clearly determined to avoid overlaps or omissions of any activity. During the planning phase outsourcing certain activities should be carefully examined as an option as further discussed in Chapter XIV. Procurement plays a particularly critical role in census logistics and should be prioritized throughout the entire operation.

2.225. The complexity of logistics management can be significantly reduced by implementing a logistics management information technology system. Such a system can assist in managing and monitoring logistics by providing necessary information to the logistics team on supply availability, distribution to enumerators, and the return of questionnaires and materials after operations are completed. Additionally, the integration of GPS tracking and real-time monitoring capabilities can further enhance logistics efficiency by providing visibility into the location and status of supplies and equipment throughout the supply chain.

A. Procurement management

2.226. Developing a strategic approach to procurement is a key element for the successful implementation of a field-based enumeration census. Given the complexity of the process, procurement planning requires logistical coordination with multiple census activities and counterparts. Proper planning contributes to efficient procurement processes and reduces the risk of confronting problems that may lead to additional costs and delays. While last-minute operations and emergency work are usually unavoidable, the benefits of early procurement planning usually outweigh these disadvantages.

2.227. Procurement planning entails the process of assessing and projecting the procurement needs of census operations. Needs assessment, cost estimation and requirement definition are essential components in procurement planning. The purpose of the requirement definition is to identify the precise needs of the census operation and to search for the best solution to meet those needs. The needs must be formulated through consultations with census subject-matter experts (e.g., demographers, IT

specialists, field operations managers) and be clearly described in the requirement definition to facilitate the procurement process. This definition is also developed in parallel with supplier sourcing and market research, allowing industrial information to tailor the requirement definition and to ensure best decision-making.

2.228. Procurement practices vary greatly among countries; therefore, no universal system of procurement management can be recommended. However, several generally accepted procurement principles are worth noting. First is the principle of *best value for money*. This refers to an optimal combination of technical and financial attributes – that is, the balance between price and performance that provides greatest overall benefit under the specified selection criteria. This does not necessarily mean selecting the lowest initial price option, but rather the option that offers the best return on the investment after a proper evaluation of offers based on the appropriate criteria in the solicitation documents. It requires an integrated assessment of technical, commercial, organizational and pricing factors in light of their relative importance. The assessment can include non-cost factors, such as fitness for purpose, adherence to quality standards, the vendor's reputation, service and support, as well as cost-related factors, such as price, life cycle costs and transaction costs associated with acquiring, using, holding, maintaining and disposing of the goods or services. This should be applied throughout the procurement process in order to attract the offer that most effectively meets the stated requirements of the census operation.

2.229. Second is the principle of *effective competition*. Effective competition is best explained as a situation in which at least three independent contractors acting on their own (that is, not in collusion with each other) effectively compete for the same business opportunity and each submit a responsive bid. The procurement processes should foster effective competition as a means of ensuring fairness, integrity, transparency and achieving best value for money. The competitive process should, as necessary, include:

(a) procurement planning for developing an overall procurement strategy;

(b) market research for identifying potential suppliers;

(c) consideration of prudent commercial practices and applicable national regulations, rules and procedures relating to procurement; and

(d) formal methods of solicitation, utilizing invitations to bid or requests for proposals on the basis of advertisement or direct solicitation of invited suppliers; or informal methods of solicitation, such as requests for quotations.

2.230. Another important principle is *fairness*. The manner in which the procurement process is carried out must give all stakeholders the assurance that the process is fair. The concept of fairness includes that the procurement process should be free from favouritism, self-interest or preference in judgment. The assurance of a fair process promotes transparency, a principle that ensures that timely information about existing conditions, decisions and actions relating to procurement activities and about procurement policies, procedures, opportunities and processes are clearly defined and made known simultaneously to all interested parties. A transparent system has clear rules and mechanisms to ensure compliance with those rules and, furthermore, ensures that procurement records are open, as is appropriate, to inspection by auditors. In line with the procurement principles of transparency, every step in the procurement process should be documented and kept on file, preferably electronically as well as in hard copy.

2.231. It's important to recognize that standard public sector procurement regulations might not always be perfectly suited to the unique needs and scale of a census. In some cases, adapting or streamlining certain regulations might be necessary to ensure efficient and effective procurement for census operations. Strengthening both procurement and technical teams is important for successful

procurement in censuses. By enhancing the capabilities of these teams, NSOs can ensure they are well-equipped to handle the technical specifications in contracts and effectively manage them.

2.232. Register-based and combined censuses require acquiring administrative data from various sources. This involves establishing and formalizing relationships with administrative data providers through agreements. While legislation often mandates the provision of administrative data for statistical purposes, practical considerations may arise. For instance, administrative data providers might require compensation for expenses incurred in preparing and delivering the data to the NSO. This necessitates financial agreements and arrangements for fund settlements. Furthermore, while most administrative data resides within the public sector, some might be held by private sector organizations. In such cases, depending on national procurement rules, the NSO might need to issue tenders to ensure fair pricing and adhere to established procurement principles.

B. Forward and reverse logistics

2.233. The type of census materials required for the census will differ depending on census methodologies and technologies used for the enumeration and data processing. However, in a full field enumeration, or for the field enumeration element of a combined census, any kind of materials related to the fieldwork has to be supplied to the field staff and returned. Strategies for distribution and return of the materials should be carefully planned according to the type of materials, volume and final destination of delivery.

2.234. As a first stage in this process, decisions should be made concerning the nature and responsibilities of the centrally controlled distribution and return operation. For example, a decision should be made about the geographic levels to which the materials will be distributed – regional office, local census committee, supervisors or other as appropriate. These decisions must be made keeping in mind the impact of the amounts of material to be transported, the transport facilities available to field staff and the condition or existence of roads or other means of transport. Once these decisions have been made, the key inputs to the dispatch and return of materials are as follows:

(a) Workload estimates from the mapping programme to establish packing volumes for transport requirements; and
(b) Name and address details for delivery and pickup points.

2.235. During the design of EAs and mapping activities, an estimate will be made of the number of EAs, and the amount of work in each. This information can be used to calculate how much material will be needed by each enumerator, supervisor, manager and so on. This method should provide a more accurate estimate of the total volume of all the materials.

2.236. The majority of these tasks are usually carried out under contract by a government transport service or commercial operator. The contractor will use specifications and consignment details provided by the NSO. The postal service may also be a feasible method of distribution.

2.237. A significant task in planning field operations is establishing the specifications for the packing and transport of materials. These specifications need to be developed regardless of whether these activities are carried out by the NSO itself or contracted out to another government agency or private company.

2.238. The role of the NSO with regard to dispatch and return tasks is primarily one of liaison and monitoring. For the most part, the contractor will contact regional managers and supervisors directly

about the delivery or pickup of material. The NSO can expect to be involved as a liaison between the contractor and field staff in the early stages of the operation or if there are any particular problems experienced by either field staff or the contractor. Census management staff should meet frequently with the contractor to discuss the operation and liaison arrangements. Part of the planning of the operation will include arrangements to enable the ~~national statistical authority's~~ management staff to monitor the delivery and return of materials. In particular, when material is picked up from supervisors, the census managers should maintain a close watch over what is taking place in the field as the transport of completed census forms is involved.

2.239. Distribution and collection of census materials can be monitored through use of an MIS (see paragraphs 3.209–3.213). Timely information can be produced about the progress in delivering the census materials and questionnaires and collecting the completed questionnaires and the other return materials. Depending on the logistics programme, the information needed for monitoring the logistics activities will differ. In general, the following information would be needed for both forward and reverse delivery:

(a) type of materials;
(b) timing of delivery;
(c) number of delivered materials; and
(d) name of the persons involved in the delivery.

The periodic reports produced from the MIS will be crucial to ensure the timeliness of the fieldwork across the country by giving an alert if there is any delay or any other problems regarding the delivery of census materials.

XIV. Contracting out

2.240. It is a customary practice in many countries to contract out (outsource) some of the tasks or activities of the population and housing census as a way of increasing efficiency by utilizing expertise and technologies not necessarily available within the NSO. At the same time, costs reductions may be achieved through a competitive selection process. However, not all census tasks are appropriate for outsourcing, and doing so will not necessarily bring the desired benefit of strengthening national capacities. Census activities may be broadly classified as core or non-core activities, which some may define as those outside of the critical path of the census. As a general rule of thumb, core activities should not be contracted out. If for some reason core activities need to be outsourced, then it is essential that the strategic control of such activities should remain firmly within the NSO at all times.

2.241. In the context of contracting out components of census operations, the NSO would need to build the capacity to ensure proper outsourcing. This is of primary importance at the preparatory stages, as outsourcing requires a solid and comprehensive knowledge of contemporary technologies and their advantages and challenges, as well as past experiences at home or in other countries. Consequently, the NSO would need to plan and develop a dedicated unit for the purpose of ensuring adequate and efficient outsourcing well in advance of the census itself, as there would be a need for extensive testing of the products and services that are to be contracted out, including during the pilot census.

2.242. The terms of engagement (scope of work) together with the deliverables and the timelines should be clearly laid down with definite dispute redresser mechanisms. Illustrative examples of areas of work that may be contracted out in the context of a full field enumeration census are as follows:

(a) Layout and printing of census questionnaires;
(b) Development of electronic questionnaire systems;

(c) Packaging of census questionnaires;
(d) Dispatch and delivery of census material;
(e) Census mapping (see, in particular, paragraphs 3.132-3.138);
(f) Communications, public relations and publicity;
(g) In-person and online training;
(h) Return collection of census questionnaires and other material;
(i) Inventory, storage and disposal of completed questionnaires;
(j) Scanning and data entry;
(k) Publication and dissemination (also relevant for administrative censuses).

2.243. Countries may find it advantageous to have the same organization take care of more than one of the steps above. That is, the list is not intended to suggest that each element to be contracted to a single, separate organization. However, experience has shown that if questionnaire design and printing and data processing are both to be outsourced there are clear benefits if these tasks are undertaken by the same contractor to ensure complete compatibility.

2.244. Time is of the essence in all these activities, and it is vital that adequate time is allocated. At the same time, backup plans should be in place in order to deal with any failure on the part of the vendors. As has been previously noted, census operations are time critical and commercial compensation is secondary. Depending on whether an activity is on the critical path or not, adequate flags should be provided. Milestones and timelines are also essential. The moment there is a failure in achieving any milestone, alerts should automatically be raised. Risk assessment represents a critical component for outsourcing; the risk of failure, and the costs involved in developing contingencies in case of failure, require particular consideration.

2.245. The appropriateness of contracting out should be determined step by step and after subdividing the overall census tasks into stages. In the context of quality management, the outsourcing of components of census operations still requires the NSO to take full responsibility for, and manage the quality of, the census data. Throughout the overall process, activities should be conducted by a method (considering accuracy and timeliness of the results) that can best satisfy the general public. No part of the work tasks should be done by a method that may result in loss of trust of the general public. When outsourcing, the NSO needs to ensure that it continues to be in a position to understand and manage elements that contribute to final data quality. So, in judging the propriety of contracting out, it is recommended that NSO should carefully consider the following criteria:

(a) Strict protection of data confidentiality;
(b) Method of confidentiality assurance that satisfies the general public;
(c) Guaranteed measures of quality assurance;
(d) The ability to manage and monitor the outsourced census tasks or activities;
(e) The competence of the contractor to stay within the planned budget; and
(f) The need to maintain control over its core responsibilities and appropriateness of judgement, taking into account the specific situation of each country.

2.246. Confidentiality assurance is the first and most important issue that should be considered. NSOs are responsible for data confidentiality, in terms of both perception and reality. It is extremely high risk for NSOs to have to manage leakage or misuse of confidential information. Consequently, contracting out of tasks that have the risk of such an incidence should be avoided. For example, in the phase of data collection, it is highly recommended that contracting out any tasks that risk eroding the trust and confidence that the public has in the NSO should be avoided. It is particularly important that where

temporary field staff are engaged under contract, this should be done in such a way that they are subject to strict measures of monitoring and control by the NSO. Such field staff should be engaged in such a way that their activities are governed by the relevant statistical legislation to preserve the confidentiality of the data they collect.

2.247. The second important and related issue that should be considered carefully is conveying confidentiality assurance to the general public. A census should be undertaken by the method that can produce the most reliable results and in a manner that ensures the trust of the general public in terms of both perception and reality. If either one of these attributes is not met, then the method used as well as the results obtained may not meet the approval of users and the general public and could result in the census itself being questioned. Thus, protecting data confidentiality refers not just to the *actual* protection of confidential data, but also to protecting the *perception* of confidentiality among the general public and providing a sense of corporate security.

2.248. The third significant issue to be considered in outsourcing is the guarantee of quality assurance in the outsourcing environment. The key point is that the NSO should be satisfied that the goods or services paid for are provided to the requisite standard. Cost should not be the first priority in considering and judging the successful bidder in this respect unless prescribed by procurement rules. Although it is desirable to engage in fair competition among several companies to reduce costs, it is worth mentioning that merely considering low-price bidding as a determinant factor may adversely affect the quality of the job to be done by the successful bidder. Low-quality work could cause a significant loss of trust among the general public. To assess the quality of work, as part of the contract allocation process, potential contractors should be required to provide samples of their work (for example, for printing, manufacturing enumerators' satchels, and other work), or if this is not possible, to list referees who could be contacted to verify the contractor's claims or allow sites at which previous work has been carried out to be inspected. The contracting process should state all the key requirements for the services sought and bidders should be measured against these. Although not a guarantee of quality, it will minimize surprises. Once the contract has been awarded, continuous monitoring of the progress of work entrusted to the selected company is necessary and the NSO should ensure that a system for monitoring quality is built into the contract. Consequently, in considering the proper contracting-out procedures, the NSO should also take into account the costs for constructing a system of surveillance for monitoring progress of the work being contracted out.

2.249. In addition to monitoring the providers of goods and services, NSOs need to plan for continuous interaction with contractors. This implies an additional step to monitoring and amounts to a necessity to work side by side on a regular basis in order to ensure the best quality of the products and services and to meet the standards and needs of the census operations. This coordinated work refers to providing technical and technological advice, as well as following the development of the services and applications from the substantive point of view. While the NSO may not have the full capability to develop certain products or applications, it certainly possesses considerable technical experience and understanding of producing statistics on a regular basis. Therefore, planning and implementing for a regular and continuous interface with providers when parts of the operations are outsourced needs to be incorporated in the overall planning from the beginning.

2.250. The fourth major issue in outsourcing census activities is the procedure of assessment and evaluation of the capabilities of the candidate providers. A quality assurance framework (for a detailed discussion on quality assurance, see Part IV) and implementation should be established in a first phase of outsourcing. Through this procedure the NSO should fully assess both the capabilities and the limitations

of companies in order to select the winner to which the activities in question are to be outsourced. It is highly recommended that practical and financial particularities of providers should be considered only after the assessment of their capabilities as any private company has a potential risk of bankruptcy or of changing the field of its activity. It should be kept in mind that if a selected company is unable to fulfil the assigned tasks, any consequent problems might not be resolved by applying penalties. A possible consequence is that users might not be able to make use of accurate and timely census results and the NSO might lose the trust of stakeholders in the census and even in future censuses or other routine statistical projects that it conducts.

2.251. Some approaches to outsourcing put an emphasis on a "turnkey" arrangement, by which contractors deliver the system according to a set of predetermined client specifications with the expectation that the client focuses solely on the outputs and not the internal working of the system. This assumes that the NSO completely understands and can fully anticipate all data quality issues that might arise during the census and has included these in the specifications. The client is not expected to have any understanding of how these systems work or how they might contribute to the final outputs. Any changes to the system typically require cumbersome processes to determine contractual responsibilities and heavy financial costs. This sort of approach is not recommended as it effectively hands over the quality of the census data to the contractor, while the risks associated with intervention remain with the census agency. It removes any flexibility and greatly restricts the ability of the NSO to react to quality problems that emerge during processing.

2.252. Suppliers should be made fully aware of the quality targets at the outset of the census programme, and the quality requirements of the outsourced components that enable the overall census quality target to be achieved. Operational quality control should apply to outsourced services in the same way as those that are provided internally. For some outsourced operations (such as the printing of questionnaires), NSOs may want to consider having staff on-site, to quickly judge and correct any quality issues as they arise.

2.253. In addition to managing outsourced activities or tasks, the ability to adapt to sudden or unforeseen changes is crucial. While outsourcing can offer potential benefits, it is essential to recognize that it does not automatically lead to cost savings. The overhead associated with monitoring contractors, addressing unforeseen challenges, and managing other complexities can offset expected cost reductions. As a result, it is recommended that NSOs should carefully assess whether or not outsourcing specific census activities aligns with their overall goals and capacity. This evaluation should consider factors such as cost-effectiveness, control, and the ability to manage potential risks.

2.254. Coding tasks, such as those for education, occupation, and industry classifications, are critical aspects of census data processing. Whether these tasks are performed by in-house staff or contracted out, ensuring adequate training is crucial. Thorough training is essential due to the intricacies of coding, which involve:

i) distinguishing between general and detailed classifications according to various coding standards.
ii) properly interpreting and applying the coding manual.
iii) considering the educational background and skill level of coders.
iv) ensuring the training effectively equips coders for the task.

Comprehensive coding manuals should be prepared and provided in advance, especially for contracted services. If automated coding methods are used, collaboration with the contractor is necessary to ensure thorough testing and validation before fieldwork begins. This proactive approach helps maintain coding accuracy and consistency throughout the census operation.

2.255. Censuses are large operations that collect massive quantities of data that require coding and editing. To reduce the staff resources required and to improve timeliness, uniformity and accuracy, automated coding procedures may be employed. Some countries have already implemented automated coding procedures for recording addresses, countries, education levels, occupation and industry. The development of the application software could be contracted out although the rules to be followed must be carefully specified by the NSO, which should always retain responsibility for implementing the system. The software application can often be used for other statistical collections undertaken by the NSO. When outsourcing, the NSO staff should be able to modify the parameters of such operations themselves at little cost and in a timely manner. By having this ability, the NSO can manage the appropriate balance between data quality, cost and timeliness.

XV. Choice of technologies

2.256. Technological advances and expanded access to modern technology have enabled significant streamlining of the way in which the business of a population and housing census is conducted in all its phases. With modern tools, every stage of the census—from planning and monitoring to implementation and evaluation—can be seamlessly integrated. Many facets of census activities can benefit from the use of technology. For example, remote sensing and imaging technology can generate maps with coordinates against which enumeration activities can be tracked. Fieldworkers can be paid their salaries and stipends using mobile device technology. Handheld devices equipped with editing functions and global navigation satellite system can improve the consistency of responses while tracking geo-activities, including those that can report on spatial coverage. Given the breadth and extensive capabilities of today's technology, it is important to select carefully at the planning stage the technological innovation elements that will be adopted in the census value chain.

2.257. The comprehensive integration of geospatial technology at every relevant stage of the census process has the potential to bring about a transformative shift in census planning, data collection, analysis and dissemination. In the planning phase, geographical information systems (GIS) can aid in the delineation of EAs, ensuring that no areas are overlooked or double-counted. This precise demarcation facilitates the efficient allocation of resources and fieldworkers. During data collection through direct enumeration in the field, geospatial technology enables real-time tracking of enumerators, ensuring their safety and optimizing their routes for efficiency. The geolocation data can be cross-referenced with responses to validate the data collected and to ensure full coverage, thereby enhancing the accuracy of the data and aiding in real-time monitoring and evaluation of census activities. In the data analysis phase, the overlay of census data onto geospatial maps can allow for a deeper understanding of demographic patterns, population densities, and other socio-economic indicators in a spatial context, allowing integration with other geo-coded datasets even from other sectors and providing richer dissemination products for data users. A fuller discussion on the use of geospatial information technology in the census is set out in Section IV.C in Part Three.

2.258. Rapid innovation has led to exponential growth and cost reduction in technology. These advancements have increased the adoption of technology across many aspects of the census in variety ways. While this brings considerable benefit, it also increases dependence on technology providers and introduces new challenges and risks. For example, pursuing the use of safe, secure and trustworthy artificial intelligence (AI) can be difficult due to the closed-source nature and limited-access licensing of many advanced algorithms, as well as the lack of a track record for such recent developments. To address

these challenges, it is important to implement safeguard mechanisms that evaluate AI-driven decisions, alert human operators to potential issues, and allow manual overrides in cases where AI suggestions or conclusions appear to deviate from expectations. The key to the successful use of technology in a census is not only understanding the rationale or objective for introducing the technology but also navigating the trade-offs between various success factors as discussed below. In many cases, no single technology option may fully satisfy all these factors, and trade-offs between performance, reliability, sophistication, and security must be carefully considered. These decisions often involve weighing the costs of information, the value of accuracy, the probability of certain risks, and the availability of skills. Budgetary constraints may ultimately dictate the optimal decision. Therefore, justifying not only the use of technology but also the decision not to adopt certain technologies or to limit their use is equally important.

2.259. Understanding of the true value of the technology is critical to formulating a business case and assessing whether or not to proceed with the project, and, if so, what technology choices to make. As the introduction of technology can be an expensive and risky exercise, it is important to ensure that there is sufficient value in its introduction for each specific census, and that the most suitable solution is selected.

2.260. The most common reasons for the introduction of technology in census operations are:

- **Efficiency and reduction of costs.** Technology provides an opportunity to reduce the number of personnel involved in different aspects of the census. For example, scanning and character recognition can reduce manual data entry and increase data accuracy; satellite imagery can reduce manual mapping; and Internet self-response can reduce fieldwork. It can potentially reduce respondent burden, shorten interview times and enhance communication with respondents. Technology can also reduce other non-labour expenses such as printing, freight and travel. In some cases, technology can simplify business processes and thus reduce cost or risk. For example, enumeration workflows can be simplified through the use of secure mobile technology that allows the enumerators to submit their availability, receive assignments and navigation instructions, collect and transmit data, and communicate with their supervisors via their mobile devices.
- **Data quality and consistency.** Technology, and in particular the automation of processes, can increase the consistency of census data and reduce data errors – for example, scanning and character recognition reduces data entry errors, and automated validation rules or edits ensure that data are checked and changed in consistent ways rather than relying on manual, dispersed field processes.
- **Timeliness.** Technology can reduce the time needed to carry out a field-based ~~census~~ enumeration, data processing, data analysis and preparation of results for publishing. The faster the census data are released, the more valuable the data are to census users, and thus the use of technology for data scanning, recognition, processing and publishing should be considered for its ability to advance publication dates.
- **Public expectations and confidentiality.** The field enumeration census relies on the general public to provide the data. Both public and user expectations, and in some cases legal requirements, may make it advisable or necessary to implement specific technologies to support and ensure their convenience of use and confidentiality protection features. This can manifest itself in various ways, such as offering online census questionnaires, facilitating online job applications for census roles, or providing census results digitally.
- **Data integration and dissemination.** For census statistics to hold value, they must be accessible and usable by a diverse group of data users, and, wherever possible, meet the principles of findability, accessibility, interoperability, and reusability. This includes everyone from the general public, private sector and media to government officials, civil society and

researchers, spanning a variety of levels of expertise and data needs. By leveraging various technologies, including GIS, the collection and dissemination processes can be enhanced. This allows for seamless integration of census data with other sectors, offers data aggregation and visualization capabilities, and ensures the availability of data in diverse formats tailored to user needs.

- **Event management.** The coordination and monitoring of a census enumeration across the complete nation or area has always been challenging and has relied on dispersed accountability and manual processes. Technology provides better visibility, oversight and capacity to monitor performance indicators and respond to enumeration events.
- **Data retention and utility.** The costs of conducting a census are significant and thus attempts should be made to optimize the value achieved from the census results. The use of technology enables census data to be safely managed, secured and retained, as well as opening up opportunities for statistical data integration and data reuse, such as for time series analyses and other analytical needs.
- **Assurance and anticorruption.** The adoption of technology can reduce the risk of fraud or corruption through providing more standardized, controlled and auditable records of actions, for example for the records of financial expenditure.

2.261. Considering the reasons listed above in a business case for a technology project often reveals inherent tensions and necessitates trade-offs that should be considered openly and transparently. For example, a solution that aims to fulfil a broader range of user expectations might entail higher costs. While the approach may benefit user satisfaction, it may negatively impact overall programme efficiency. It is possible to consider various tiers of technology use and functionalities and adjust the scope of technology to meet as many goals as appropriate (step approach).

2.262. While the conduct of significant technology projects is becoming increasingly commonplace across government, there is limited evidence to suggest that the mere introduction of new technology alone guarantees that projects will be completed on time and within budget. Careful planning and management and testing are critical to the success of these projects. The unique size, timing and nature of a census throw up unique challenges that require careful consideration. The following factors should be taken into consideration in the context of optimizing the use of contemporary technology for census operations, each of these factors should be validated through specific testing processes:

- **Suitability.** The functions and benefits of any new technology must be assessed thoroughly and objectively through tests that ensure it meets the needs of the census. Additionally, it is important to evaluate whether the technology is user-friendly for its intended users.
- **Scalability.** Technology solutions should be designed, implemented and tested to handle the expected number of users and of data items. Hardware and software must be scalable to meet the actual load of each business process during the census project, and this scalability should be tested in advance to prevent bottlenecks.
- **Security.** Ensuring the security of census data is critical to protect the confidentiality of personal information and maintaining public confidence for the census. Census systems, especially online ones, must secure the confidentiality, integrity and availability of census information. Given the increasing threat of cybersecurity risks such as data loss, fake data injection or spoofing, national statistical offices with a digital presence must invest appropriately to minimize these risks.
- **Stability.** A field-based census relies on completing the enumeration within a tight timeframe. Key technologies – such as the census call centre, self-response portal or self-help website –

may be designed and tested to ensure that they do not negatively affect response rates, particularly in countries relying on self-response.

- **Skill.** Whether the NSO plans to outsource or internally deliver the solution, it must ensure that it has the adequate skills, knowledge and capacity to lead and manage the project.

2.263. Census technology projects require strong project management expertise and thus adequately competent, experienced, motivated and knowledgeable staff need to be assigned to this role.

2.264. Detailed business requirements (meaning what the technology needs to do and how it needs to do it) need to be developed by a suitably skilled business analyst in order to ensure that the technical specifications and implementation are aligned with the census need.

2.265. Rigorous testing and piloting by the NSO is essential when integrating new technology into its processes. It is crucial to understand the various stages impacted by technological changes. Separate tests should be conducted to prove the efficacy of new technologies and to identify potential problems linked to their implementation. Depending on the extent and characteristics of IT, these tests should include all information technology components related to the fieldwork, data transfer or entry and processing well ahead of the actual census. Testing and piloting should evaluate the application systems (software) and the equipment (hardware) under real-world conditions, so as to identify and mitigate any potential malfunctions. Continuity planning should also be integrated into the testing phase, ensuring that there are fallback measures in place to counteract any equipment or system failures that could disrupt a smooth census operation.

2.266. A pilot census (in addition to all its other purposes) ensures that technology can be tested as part of a complete, end-to-end business process and highlight major implementations with functionality or capability. The pilot test should be accompanied by activities to test whether or not technology solutions are scalable, secure, accessible and robust enough for the real census. Stress tests should be conducted during the testing phase to ensure that the use of such technology can handle the maximum load of each module or business case and at the same time sustain data integrity.

2.267. Once the implementation of technology for the census is completed, a comprehensive evaluation should be conducted. This evaluation serves a dual purpose: to identify lessons learned and define areas for improvement to enhance future census operations; and maximize the value of the technology. During this phase, the results should be analysed, and the efficiency and accuracy of the technology assessed, taking into account the costs and benefits associated with the implementation. Feedback from enumerators and those members of the general public who participated in the process should be collected and considered. This will help identify specific issues that require solutions and contribute to the continuous improvement of census operations in the future.

2.268. There is an increasing adoption of technology in every NSO across the world. A number of nations have been developing their own technological solutions to support different aspects of the census, with some of these products being made available for NSOs in other countries. When making decisions around technology selection or development approaches there should be strong, favourable consideration of existing products within the international census community, and where products do not already exist the preference should be to undertake co-development with other members of the statistics community. The exchange of best practices and lessons learned is key to streamlining these processes.

2.269. There are, however, a number of risks or drawbacks associated with the introduction of technology that should be considered and managed, otherwise these could lead to increases in expenditure, delays to the census timetable or impacts on the quality of the census data. These include:

- Incompatibility or other integration issues between different hardware and software applications;
- Solution outage or failure (which could be for a variety of reasons – lack of connectivity, hardware failure, battery life, GPS black spots, software bugs, device theft);
- Lack of skills or knowledge among system users, particularly temporary field staff;
- Insufficient or inadequate communication between technology staff and business staff, particularly leading to misunderstanding of requirements and specifications;
- Hacking, online attack or other information technology security events;
- Maintaining, upgrading or decommissioning old or legacy systems;
- Lack of documentation or over-reliance on a small number of key people;
- Huge amount of digital data available, creating a potential distraction for staff.

2.270. To mitigate technological risks in census operations, NSOs should (as noted above) implement robust technology planning and standardization, conduct pilot testing, and ensure redundancy. Comprehensive training and a centralized knowledge base are essential for staff. Clear requirements, joint planning, and strong communication enhance collaboration. Information security measures and regular audits protect against threats. Technology lifecycle management, including regular updates and retirement planning, ensures systems remain secure. Detailed documentation and knowledge sharing prevent over-reliance on key individuals. Finally, data management and governance policies maintain data quality and integrity. These measures collectively reduce risks and support successful census operations.

2.271. A wide range of technologies, covering all stages of the census from planning to data dissemination, is presented throughout many of the sections of these Principles and Recommendations. However, the integration of technologies in census operation requires taking into account various specific national needs, and the value each technology would bring to a specific aspect of the census.

2.272. The utilization of technology is expected to vary considerably in NSOs across the world, considering the need and the resources that are available to them. Ultimately, technology is an enabler that can enhance efficiency, accuracy, speed and transparency of census operations. It can also lead to optimization of costs, depending on the circumstances of each country. It is to be reiterated, however, that the choice of technology and its level of deployment should be diligently assessed by each NSO before introduction.

XVI. Innovation to optimize census processes

2.273. Census organisations should constantly test and introduce innovative approaches to modernize census activities and meet the ever-changing needs of data users and other stakeholders. Changes in society bring new public expectations, while technological advancements (such as those discussed in the previous Section) lead to new opportunities. Establishing an ongoing innovation programme provides the platform necessary to respond to these evolving demands and opportunities.

2.274. Innovations can differ in scale. Some are adaptive, helping censuses adjust to new public expectations, data user needs and technological upgrades, among other factors. Other innovations can be transformative with significant implications for the entire census operation. For example, adopting

web-based self-response modes or transitioning from field-based to register-based approaches would be transformative.

2.275. Large-scale and complex census operations benefit from longer planning periods than other statistical programmes. This extended timeframe allows for the adoption of new innovations that can be identified through an ongoing process of researching, documenting and implementing new innovative ideas or during post-operation evaluations.

2.276. Census innovation can have a significant impact on the NSO's wider statistical activities. Census programmes are often relatively well-funded compared to other statistical projects, allowing for the early adoption of new innovations. These innovations can later be applied to non-census programmes, potentially leading to increased efficiency and cost savings across the NSO if managed properly. It is therefore recommended that census innovations be shared within the NSO, and that strong lines of communication are established to ensure that, conversely, relevant innovations from other areas of the NSO can also inform census taking.

2.277. NSOs should capitalize on the increased visibility and funding that often accompany census programmes to invest in innovation. While initial investments in innovation can be substantial, they often yield long-term efficiencies that justify the upfront costs throughout the census lifecycle. Cost-benefit analyses are valuable tools for assessing the worthiness of these investments. Fostering a culture of innovation within the NSO is therefore valuable.

2.278. Innovation should, however, not take place in national isolation. The NSO should engage with international partners to share innovative ideas and lessons learned from testing and operational experiences. Domestically, it is worth exploring how other organisations undertake similar large-scale activities. Establishing a local culture of innovation, and investing in long-term research activities, is necessary for ensuring the ongoing modernisation and future-proofing of the census.

PART THREE. CENSUS OPERATIONAL ACTIVITIES

I. Introduction

3.1. Part Three of these *Principles and Recommendations* focuses on an elaboration of census operational activities pertaining to the census methods described in Part One. While the population and housing census is in essence a statistical data collection exercise, it still incorporates components that are not present in a routine statistical survey conducted within the frame of the national statistical system. In consequence, the layout of this part follows the frame and the logic of the Generic Statistical Business Process Model (see Figure 2.1);[77] but it also provides particular methodological and operational guidelines relevant to the population and housing census. This part of the *Principles and Recommendations*, therefore, starts with a discussion on developing the census questionnaire and building census infrastructure before going on to discuss: geospatial information and mapping; testing; establishing a census frame; field enumeration; data processing; dissemination and the utilization of data; documentation of the census experience; and archiving.

3.2. It is important to note that Part Three primarily focuses on field-based censuses. As such, some of the operational activities discussed in this Part, particularly those related to field enumeration, may not be directly applicable to register-based systems. Operational activities pertaining to administrative register-based censuses are elaborated in the UN Handbook on Registers-based Population and Housing Censuses[78].

II. Census questionnaires: content and design

3.3. The preparation of the census questionnaire to be used in a full field enumeration or combined census refers to a well-designed process that should start at a very early stage of the census operation. This process includes developing various methods and tools that will respond to national priorities and data user needs from different sectors, such as governmental agencies, research institutions, the private sector, development partners, civil society, non-governmental organizations, and the general public. In addition to consulting with users to determine data needs, other factors that would have an impact on the selection of census topics, include:

- quality of data collected through the previous census;
- timeliness of data production and dissemination;
- national sensitivity for specific topics;
- the feasibility of collecting reliable data on a particular topic;
- the length and complexity of the questionnaire;
- the availability of resources (financial, personnel and equipment); and
- the burden on the responding public.

3.4. Taking into consideration that the basic principle is to meet user needs and to make census statistics as useful as possible, the content of the census questionnaire, therefore, should be

[77] United Nations Economic Commission for Europe (UNECE), on behalf of the International Statistical Community, *Generic Statistical Business Process Model (GSBPM*), Version 5.1 (January 2019).
[78] UN Handbook on Registers-based Population and Housing Censuses. Link: https://unstats.un.org/unsd/demographic-social/publication/handbook-registers-phc.pdf

determined through consultations with census data users (see paragraphs 2.175-2.178). However, the consolidation of the results of the consultations with all stakeholders needs to be balanced by the other factors noted above.

3.5. The final content and design of the questionnaire will also be the result of rigorous and meticulous testing. It is necessary that all aspects of the census questionnaire, such as wording, structure and order of the questions, and the design, be carefully tested[79] to ensure the successful application of the questionnaire in the field. Questionnaire content should also be examined with respect to the proposed mode(s) of data collection to ensure that formatting and wording are appropriate to the particular mode(s).

3.6. Preparation of the census questionnaire(s) requires the most careful consideration, since the consequences of a poor design and content cannot be overcome during and after enumeration. To ensure inclusivity, census questionnaire design must prioritize accessibility and cultural sensitivity. This involves offering questionnaires in multiple languages to accommodate diverse linguistic communities and implementing design features to meet the needs of individuals with disabilities. Successful implementation of this process will have significant impacts on the quality of collected data and the subsequent census outputs. The benefits of evaluating the efficacy of the questionnaire during a pilot census cannot be over-stressed.

a) Selection of census topics

3.7. As a first step in determining the content of the census questionnaires, the topics that will be covered in the census should be selected taking into consideration the priorities of national needs, regional and international recommendations,[80] historical comparisons, regional and international comparability, suitability of topics for collecting reliable information and resources available for the census undertaking. Census takers should evaluate the national needs in the light of possible new topics and the needs for continuation of the topics covered in the past censuses and/or surveys. Topics that are no longer necessary in the current census due to evolving data needs and availability of alternative data sources, as well as new topics for which data is needed, should be carefully reviewed in this process. More information on the factors determining the selection of census topics is given in Chapters I and II of Part Five.

3.8. A further factor that should be considered in the process of selecting census topics is related to the decision on whether to use a single census questionnaire for all respondents (all households and individuals therein) or adopt a two-questionnaire approach – a combination of a short-form and ata long-form questionnaire. Using a single questionnaire consists of a standard set of questions for all individuals and housing units covered in the census. In the latter approach, countries use a short-form questionnaire with a limited number of basic questions for enumerating most of the population, while a long-form questionnaire is applied to a sample of the population for collecting more detailed information (on the same topics as contained on the short form and on additional select topics). The advantages and challenges of adopting a two-questionnaire approach are summarised in paragraph 1.154. but the following paragraphs provide a more in-depth elaboration on this approach.

[79] For more information about census tests, see Chapter V.

[80] Part Five of these *Principles and Recommendations* focuses exclusively on core and non-core topics for population and housing censuses.

b) Use of short and long questionnaires

3.9. With each new census and the advancement in processing and utilising census data for informed decision making, there is an increased interest in adding topics to the population and housing census, in addition to those historically covered. Because of additional costs, the time required to collect and process the information, and the attendant burden on the respondents, imposing a long questionnaire on the total population, in many cases, may not seem to be appropriate or necessary. Hence, some countries may decide to broaden the scope of the census by covering additional topics through the use of a sampling methodology. In this approach, two questionnaires are used:

(a) a short questionnaire containing only those questions intended for universal coverage; and

(b) a long questionnaire typically including all the questions from the short questionnaire plus additional questions to gather more detailed information from a sample.

3.10. Using a sampling methodology in conjunction with full enumeration requires careful planning for two key aspects:

(i) Determining the additional topics for the long questionnaire.

(ii) Selecting the sample of households to complete the long questionnaire: This involves determining the sampling method and sample size to ensure the data collected from the sample is representative of the entire population.

While this approach can be cost-effective by reducing the overall data collection time, it does introduce some complexity in field organization. Careful planning is needed to manage the logistics of deploying two questionnaires and ensuring proper completion by the selected sample.

3.11. The ever-expanding needs in most countries for extensive and reliable data have made the use of sampling a cost-effective part of census taking. Sampling is increasingly being used to broaden the scope of the census through asking additional questions of only a sample of the population. The use of sampling makes it feasible to obtain urgently needed data of acceptable precision when factors of timing, respondent burden, and cost would otherwise make it impractical to obtain such data on a complete count basis.

3.12. The suitability of particular questions for a sample enumeration depends on the precision with which results are needed for small areas and small population groups, and on the enumeration costs involved. Collecting data on particular questions for smaller areas or groups requires careful consideration of the sample size. While moving the questions to full enumeration may initially appear to be the only solution, exploring the option of increasing the sample size can offer advantages. A thorough cost-benefit analysis will ultimately determine the most efficient and effective approach.

3.13. It is important to bear in mind, however, that national legal requirements may make it mandatory to collect certain information on a complete count basis. Legislation in many countries prescribes complete population enumerations at particular times or makes certain political or administrative dispositions dependent on particular results from a full enumeration. For example, the apportionment of seats in the legislature among the civil divisions of a country often depends on the number of persons actually enumerated in each division. The data needed for this, and for similar purposes, cannot be obtained using a sampling methodology. Moreover, one of the essential elements of the census – universality – will be forfeited.

3.14. Census information that is collected for only a sample of the population or housing units is usually obtained by one of two different methods. The first predesignates a systematic subset of census households to receive a so-called “long” questionnaire. Depending on the sample requirements, which in turn take account of considerations of cost and precision, the systematic subset that is designated for the long questionnaire may represent, for example, a 10, 20 or even 25 per cent sample of all households. Under such a sampling scheme, all other non-sampled households in the census will receive a short questionnaire containing only those questions intended for universal coverage. If countries choose this option, it is recommended that the selection of the sample households that are to receive the long questionnaire be carried out at a central location by supervisory statistical staff, since it has been shown that when the enumerators themselves actually identify the sample households in the field, the results are often biased. Such selection can be done after the pre-enumeration household listing is completed.

3.15. The second method of sampling often used involves designating a sample of EAs in which the long questionnaire will be administered. In this approach, all households in the designated EA the long questionnaire will be enumerated with the long form while the short questionnaire will used for all households in the remaining unsampled EAs.

3.16. The advantage of the first method over the second is that the sampling precision of the census results is greater in the former because clustering effects increase the sampling variance when whole EAs are used as sampling units. On the other hand, the advantage of the second method is that field staff may be trained more easily, since one set of enumerators need be trained only for the long questionnaire and the other set only for the short questionnaire. Sampling the entire universe often requires a significant sample size if there are questions on the long-form questionnaire that target a small population group. To reach these populations, NSOs must weigh the costs of increasing sample sizes against the increased variance of a clustered sample. However, the clustered sample would only be effective if the rare population is known to be clustered in the EAs, which is more often not the case.

3.17. It is important to make certain that the inclusion of questions intended only for specific subsets of the population does not create legal, administrative, or political challenges. Census legislation often imposes obligations and potential penalties for non-response, necessitating careful consideration of question design to avoid complications.

c) Sections of the questionnaire

3.18. Clearly distinguishing between the units of enumeration is an essential element when designing the census questionnaire, as each question aims to collect data for a specific unit. The census questionnaire can be clearly and intuitively structured so that the design reflects the unit of enumeration. This structure has a positive impact in better understanding the meaning of the questions and in conducting the interview smoothly. This also facilitates the successful completion of the questionnaire for self-enumeration (see Section IV.A in Part Two).

d) Questionnaire design

3.19. Although the majority of countries that collect data in a field enumeration are using face-to-face interviews with electronic questionnaires to do so, many other countries are using other methods of collecting census. information, including the use of multi-mode methods. Some of these methods are face-to-face interview with electronic questionnaire, telephone interviewing, self-enumeration with a paper questionnaire collected by enumerators or returned by mail, self-enumeration via the Internet, as well as the use of population registers-and other pre-existing administrative records, either alone or (the case of a combined census methodology) in combination

with other sources. The methods of enumeration and technology used for data capture are among the main factors affecting the design of the questionnaires. For example, the structure of questionnaires that will be used for the face-to-face interview and for self-enumeration will differ, as the former will be administered by trained enumerators while the latter will be completed solely by the respondents. Whatever methods are chosen, these need to be tested and assessed in advance for data quality and feasibility. It is also important that data security and confidentiality is maintained whatever modes or approaches are used.

3.20. The design of the questionnaire should be based on the type of data collection mode and approach used. Questionnaire design should also be based on the technology being adopted for data processing, for example whether data processing will be done through scanning, manual entry or electronic transmission of data to a database. It should be noted, however, that regardless of the mode of enumeration used, it is advisable that there is compatibility in the design among the modes in order to enhance data comparability (for more information on designing multi-mode data collection, see paragraphs 3.180-3.191).

3.21. Where countries utilize the Internet or handheld devices to collect their census information, or a portion of that, the layout and organization of the data collection instrument may differ from that of the paper questionnaire, although it is important that the content be the same for all the modes of data collection used. Comparability of content is important to ensure that the same information is collected and that entries are similarly checked across each mode of data collection in order to enhance comparability of data. It is important to note, however, that most often, adopting an Internet-based data collection approach also means moving from an enumerator-based approach to a self-completion approach. Consequently, the questions should be worded and structured in a way that makes completion by the respondent (without the assistance of an interviewer) more intuitive. In this case, care should be given to the design of the questions so that respondents can easily provide the required information without assistance. Therefore, it is imperative that census management involve the information technology team, as well as cognitive testing experts at the start of the questionnaire preparation stage. While many of the same principles (for example clarity of wording, omission of unnecessary material) will apply also to an Internet-based or handheld device-based collection of information, specialized advice should be sought and provided regarding such issues as:

(a) the technology employed to present the questions to the respondent;
(b) the method of capturing the response; and
(c) quality assurance checks employed during data capture process.

It is also important to evaluate mode bias resulting from differences in responses to the same question being collected using different methods (such as Internet-based vs. enumerator interview).

3.22. A fundamental principle is that questionnaire design must be regarded as part of an integrated process of satisfying user demands by collecting, processing and disseminating high quality information provided by respondents. Questionnaire design must be driven by a planning process based upon dialogue between the NSO and the data users. Information to be collected should respond to user needs both at national and international levels and therefore user consultation is crucial in this regard. Previous census questions that are no longer relevant should be dropped as these do not add any more value. This is essential if the questionnaire is to be designed to provide only the information needed by users, as well as to manage respondent burden.

3.23. It is important also that questions and response options are free from ambiguity. Moreover, questions should not be offensive nor biased and should use inclusive language to the extent possible; in many cases this can be avoided by excluding extremely sensitive topics from the census questionnaire, but care must always be taken to consider the reaction of respondents when

designing all questions. In addition, it should be noted that the quality of information collected in a census will be reduced if the questionnaire is excessively long. These issues should be carefully assessed during the testing programme, including any cognitive tests and a pilot census (see paragraphs 3.147–3.148) since poorly worded questions will not only collect poor quality data, but, by confusing respondents and/or enumerators, may also impact on the responses to subsequent questions in the questionnaire.

3.24. If questions include 'write-in' response options, especially for new questions, data processing can be time-consuming and resource intensive. This is due to the wide range of possible responses and the need for coders to research unfamiliar terms and assign appropriate codes. To address this, NSOs should consider carefully when to include write-in options and provide clear instructions to respondents to limit the variety of responses. This can help streamline the coding process and reduce the overall workload.

3.25. Skip patterns and the form and format of the response options require careful consideration. A question can be perfectly worded, but if the response options are not presented in a way that is meaningful, unambiguous, and easily understood, resulting data may be of poor quality.

3.26. Special provisions will have to be made if two or more languages are used in the country. Several methods have been used to deal with this situation, such as:

(a) a single, multilingual questionnaire; or

(b) one version of the questionnaire for each major language; or

(c) translations of the questionnaire in the various languages being made available in the enumerators' manual or on the Internet site for the census.

In order to ensure consistency and to avoid interviewer bias if option it is important that, if option (c) is adopted, translations of the questionnaire into various languages that are used in the country be centralized and not be undertaken in the field by interviewers. Information on the distribution of languages in the country is important for sound census planning and, if not already available, should be determined at some stage of the census preparations. Staff recruitment and training procedures (see paragraphs 2.211-2.221) will also have to take language issues into account.

3.27. If the housing census and the population census constitute a single operation, it is necessary to consider whether a single questionnaire should be utilized to collect information on both population and housing topics. If separate questionnaires are used, they should be uniquely identified in a way that links the component forms so as to permit subsequent matching, both physical and automated, of the data for each set of living quarters with the data that refer to the occupants thereof. This consideration is particularly important when employing separate personal forms for each individual in the household.

3.28. When paper questionnaires are used for data collection, the use of processing techniques, such as optical mark reading or intelligent character recognition, will have a significant effect upon the questionnaire design (see paragraphs 3.268–3.270). In the case of optical mark reading, it is necessary both to provide spacing of response areas and to ensure printing is undertaken to precise tolerances so that the data capture software is able to capture all required data but not any of the material around the designated response areas. With regard to intelligent character recognition (ICR), it is crucial to allow sufficient space for response boxes and to ensure that these are designed according to the requirements of the processing system so that each response box contains only one character. Enumerators should be trained to print neatly and correctly (usually in upper case). To better handle the entries by self-respondents, the ICR technology has to be built to interpret various forms of manual entries. Where the scanning process requires that a booklet questionnaire is separated into component pages, it is important that some form of linking (for example by serial

numbers or barcodes) is employed to ensure that all the information is assigned to the correct individual or household in the computer records.

3.29. The final questionnaire must be drafted in time to allow for: (a) printing and developing the data collection application, in the case of using electronic questionnaire (making allowance for the many contingencies, such as industrial action, breakdown of printing equipment or delay of programming activities, that can arise in these processes); and (b) undertaking quality assurance checks to ensure the printing is of sufficient quality to be used in the data capture regime and to ensure that the data collection application is running correctly under the data entry rules.

3.30. As some countries are also utilizing Internet portals and handheld devices for conducting the census, sufficient time must be given to design, develop, test and implement bug-free e-questionnaires and related software systems. Last-minute inclusion or changes in the questionnaire may affect the overall quality of the data collection, and in turn the census results.

3.31. In view of the many issues to be addressed in designing a census questionnaire, it is not feasible to suggest specific model questions for the census topics that are covered in Part Five of these Recommendations. However, images of all census questionnaires that have been made available to the United Nations Statistics Division have been placed on the Division's website[81] together with research papers relating to questionnaires used to collect information on the various topics recommended for collection and also using the different technologies.

e) Design of questionnaires for electronic data collection (CAPI, CAWI, CATI)

3.32. The design of electronic questionnaires for data collection via Computer-Assisted Personal Interviewing (CAPI), Computer-Assisted Web Interviewing (CAWI)[82] and Computer-Assisted Telephone Interviewing (CATI) methods require additional considerations to make the data entry process intuitive for the enumerator or respondent. Some essential functional features that should be used in the design of the electronic questionnaires include:

(i) **Questionnaire navigation** should allow enumerators/respondents to move sequentially through the questionnaire in order to enter responses in the most effective way, giving the ability to pause and resume at the last answered question with a "save and continue later" functionality. On the other hand, the design should impose some restrictions on navigation, for example, by preventing enumerators/respondents from entering certain questions without having first obtained responses from other, earlier, questions.

(ii) **Skipping/automated routing** is one of the most important error reducing features in electronic questionnaires. It obviates the need to respond to questions that should be skipped. It also avoids the converse - skipping questions that should be asked, thus minimizing the need to impute for missing responses. Basic skips allow the response to *a particular question* to determine whether or not the next question is relevant, while complex skips are those that use responses *from several previous questions* to determine whether or not subsequent questions are relevant.

(iii) **Pre-coding** allows relevant questions to be answered from pre-coded drop-down menus. In some cases, drop-down menus could be altered dynamically, depending on previous responses, so that the interviewer is never presented with an impossible response code.

[81] See http://unstats.un.org/unsd/demographic/sources/census/censusquest.htm.

[82] Also referred to as Computer-Assisted Self Interviewing (CASI).

(iv) **Validation.** Real-time data validation checks can correct invalid or inconsistent responses that could be the result of either interviewer or respondent error, thus reducing the amount of post-enumeration data edits.

(v) **Support** should be available to both enumerators and self-enumeration respondents to address a wide range of issues, including technical problems and difficulties understanding the census questionnaire. Support can be provided through various channels, such as help buttons, online resources, additional instructions, and dedicated help hotlines.

III. Building census infrastructure

3.33. The population and housing census requires a concerted effort in building an intricate and complex infrastructure that is suitable for conducting many disparate activities simultaneously. This infrastructure refers to:

- the development of the census instruments/questionnaires;
- logistic schemes for disseminating and collecting returns;
- facilities and systems for capturing, processing, editing, and validating response data; and
- the systems and tools for dissemination of statistics and data products. It also requires well-defined workflows and testing of the production system as well as of the whole statistical business process.

As most censuses are periodic with as many as ten years between them, the infrastructure is also likely to be temporary, so care must be taken to ensure the appropriate balance between what is needed for a given census and what is a reasonable expenditure of resources for a temporary infrastructure.

3.34. The necessary infrastructure for a census will vary significantly based on the design of the census and the data collection mode(s) being employed. For a field-based paper-based census, the infrastructure will need to support the paper-based activities such as:

- printing questionnaires and materials;
- space for storage of paper materials;
- mailing and distribution plans for distributing paper materials to field operational areas and returning completed questionnaires for processing;
- systems and equipment for capturing data from paper questionnaires; and
- space/procedures for long-term storage of paper materials based on the organization's records retention policy.

In addition, special procedures should be developed and distributed to all office and field staff on the appropriate security measures necessary to ensure the safety and integrity of the paper materials and the confidentiality of sensitive information such as addresses and personal data.

3.35. For electronic data collection, on the other hand, the infrastructure will need to include facilities, equipment, software, and information technology security measures involved in the electronic data entry, transmission, and storage. In this case, cloud infrastructure can be utilized to achieve the necessary flexibility with the performance and reliability of the provisioned resources. The infrastructure for a register-based census will need to also include extensive storage for data files and tools and systems to facilitate the necessary linking and processing of the administrative data. Censuses that employ a hybrid approach using a combination of two or more of these approaches may require a very complicated infrastructure that must support all of the data collection modes being used.

3.36. A census infrastructure must be put in place well before the data collection exercise itself, as all of the components need to be extensively tested in circumstances that are as real as possible

and require reasonable time for any corrections needed as the testing may indicate. The importance of testing cannot be over-stated, especially in the context of the need to build the components for processing and analysing data. The preparation of the infrastructure does not usually translate into immediate procurement decisions, as the technological process results in better performance, richer functionality, and lower cost of the equipment with time. Thus, purchasing equipment or devices (such as tablets for data collection) several years in advance of the fieldwork brings the risks that such equipment will become obsolete or unsupported by the time of the census. Instead, a limited sample of the equipment can be acquired at the time of building the infrastructure to facilitate testing, with the bulk being acquired later, when confirmed suitable by the tests, and in order to take advantage of the most recent technology available.

3.37. Building census infrastructure should be based primarily on the experiences and lessons learned from the previous census. If this took place a long time previously and the documentation and institutional memory are not sufficient, putting together census components should rely on statistical practice in conducting large surveys and on the schemes for developing and implementing a statistical business process. Lessons learned from conducting similarly demanding operations (for example, an agricultural census, or census of establishments) can also be considered. Collaboration and consulting with other NSOs that have more recently conducted a large-scale survey or a census (for example, in a neighbouring country) can also be a valuable source of information and guidance.

3.38. Particular components of the census infrastructure may already be available within the NSO – for example, the statistical network of field offices covering the country's territory and staff experienced in collecting and producing statistics. Other components will need to be established from the beginning, such as the training synopsis and schedules, recruitment procedures and logistical arrangements. Methodical planning is critical in this process of building census infrastructure, and developing detailed project schedules early will significantly facilitate managing the process in a holistic manner.

3.39. In that context, clear and unambiguous delineation of management responsibilities is equally important to the building process. Managers and supervisors need to have a comprehensive assignment from the very beginning and at all levels of management. Along with roles and responsibilities, methods and channels of communication must also be clearly defined to ensure the different operational units are collaborating and coordinating their assignments. While some improvisation may be necessary in certain circumstances, all efforts should be put into anticipating risks and developing and testing strategies to allow managers to mitigate them and have prescribed alternatives in place when needed (see Section VIII.B in Part Two).

3.40. In building new components of the census infrastructure – for example, outsourcing – it is necessary to liaise with other parts of the government that may have more extensive experience in that respect. In most cases, other government departments may not only have extensive experiences – such as subcontracting for defence purposes – but also might have the capacity to provide the goods and services that are needed. Therefore, the process of building infrastructure needs to first assess existing government capacities, and then extend beyond them.

3.41. The process of building the census infrastructure has to be completely finished before the pilot census takes place. Certainly, parts or perhaps even whole components of the infrastructure may need to be adjusted and altered as a result of the pilot census. Taking this process into account, the schedule for the preparations of the census must allocate sufficient time for the modification of systems necessary at the end of the pilot test.

3.42. After the census is completed, all steps involved in building the census infrastructure and its testing should be documented in detail for future reference. Certain components, such as computers, networks, and servers, may remain part of the broader national statistical infrastructure. Comprehensive documentation is essential not only for future censuses but also for a better understanding of the recently conducted census (see Section X below on documentation of census experience).

IV. Geospatial information

A. Strategic planning

3.43. Everything happens somewhere, at a given point in time. Geospatial information provides one of the most crucial aspects of census taking – location, essential at all its stages. The fundamental role of geospatial information in census operational activities is to support the enumeration and to present aggregate census results in geographic and cartographic form. Very few enumerations during the last several census rounds were executed without extensive geospatial support.

3.44. There is widespread recognition of the importance for NSOs to develop a long-term geospatial data strategy and build the necessary infrastructure and human resources to support the nation's data needs. Leveraging geospatial capabilities can significantly enhance population and housing censuses and other statistical activities. A dedicated geospatial function within the NSO can contribute to census analysis, data visualization, and a range of geospatial applications and services. This rich data source can serve as the foundation for a robust National Spatial Data Infrastructure (NSDI)[83], promoting standardized geographic units and facilitating location-based services across government levels. While building internal geospatial capacity is essential, NSOs should also explore opportunities to collaborate with existing national geospatial information agencies (NGIAs) or research institutions. The collaboration should be of mutual benefits with NGIAs supporting the NSOs with base maps for censuses, and where NSOs can offer updated content in exchange if helpful. Partnerships and knowledge-sharing can optimize resource utilization and accelerate geospatial integration within the statistical system.

3.45. Geospatially enabled census data is indispensable for effective decision-making. Location-specific information provides unparalleled insights compared to aggregated national data. By integrating census data with other statistical information through geographic units, policymakers can access actionable knowledge to inform policy and implementation.

3.46. A census limited to aggregated data offers limited value. Understanding population distribution at a granular level is essential for comprehensive analysis and evidence-based decision-making. Strategic planning for the integration of census statistics and geospatial information yields interoperable datasets that can be analyzed at various spatial scales, ensuring data relevance for diverse policy and decision-making needs.

[83] The National Spatial Data Infrastructure (NSDI) is a framework of policies, institutional arrangements, technologies, data, and people that enables the sharing and effective usage of geographic information by standardizing formats and protocols for access and interoperability. The development of a NSDI is a multi-agency and multi-partner activity.

3.47. The integration of statistical and geospatial information is increasingly relevant and necessary. It serves as a crucial bridge enabling the production of harmonized, standardized, and integrated geospatially-enabled statistical data. This, in turn, facilitates data-driven decision-making by providing valuable insights into spatial patterns and relationships. In particular with regards to population censuses and monitoring of internationally agreed development agendas, the United Nations Statistical Commission, at its fifty-first session, in March 2020, approved the Global Statistical Geospatial Framework (GSGF).

3.48. The GSGF should be considered a critical component of census planning and operations, aligned with the broader context of the GSBPM. The GSBPM provides a structured approach for integrating geospatial considerations into the overall statistical production process. The GSBPM sets out typical activities and stages that statistical organizations undertake when producing statistics and this provides a common thread for documenting activities related to the geospatial domain so that relevant actions are taken at the correct stage of the census. Moreover, geospatial information, including a GIS and mapping, can be applied in a census structured according to the GSBPM.

3.49. In summary, the GSBPM geospatial view[84] will enable the description of geospatial activities, essential to producing geospatially enabled statistics, using the GSBPM framework. By considering the GSGF Principles, these geospatial activities will contribute to increased data standardization, flexibility, and integration.

3.50. The use of geospatial information and mapping methods can range from printed maps to Global Navigation Satellite System (GNSS)-equipped handheld devices and geoportals according to the technological and financial capacity of each country (see paragraph 3.74(d) for a description of GNSS). Adding geospatial information and technology to the census may, at first, seem costly for some NSOs, however from the experiences of many countries there can be gains in efficiency, increased coverage and accuracy and return on investment into the census activities.

3.51. Major technological advances include the widespread availability of: handheld devices, GNSS and GIS) software; low-cost aerial, drone, and satellite acquired imagery; and open and free earth observation data[85]. These advances have put new tools in the hands of NSOs to enable them to collect more accurate and timely information about their national populations. Adopting such new methods can require significant planning and lead times for building new enterprise capacity and organizational restructuring, but this allows for the development of an incredibly powerful geospatial data system that should then be leveraged to support a variety of services and applications beyond census activities.

3.52. Geospatial support services for census activities can vary widely. Traditional mapping techniques remain relevant in some countries, particularly in remote areas. However, advancements in technology have enabled countries to adopt digital mapping techniques and improve census operations. While implementing new technologies requires careful planning and resource assessment, the costs of hardware and software solutions have decreased significantly. The availability of handheld devices, off-the-shelf software, and free or open-source software has further reduced the financial burden. Outsourcing geospatial support services can be a viable option when

[84] Geospatial View of Generic Statistical Business Process Model - GeoGSBPM.
[85] Open and free data from programmes such as Landsat and Copernicus, are relevant for several applications including the census.

internal capacity is limited. However, a thorough evaluation of costs, benefits, and potential risks is essential before making any outsourcing decisions.

3.53. It is recommended that all NSOs prioritize digital collection of information for full field enumeration censuses. Additionally, location data through coordinates and/or addresses should be considered core census information. For countries where printed maps remain necessary, greater emphasis should be placed on registering addresses and locational information (such as geographical names). This will allow for better supervision of the operation, data disaggregation, and the creation of a solid location database for future use. In many countries, a combined solution using paper and digital maps can be useful, especially during the data collection stage.

3.54. In the case of register-based or combined census, location must also be considered core, especially due to the ability to allow cross-referencing with other georeferenced statistical databases. It is also an ally to assess the territorial coverage of administrative records and any need for field collection. In these types of censuses, addresses tend to be more frequently recorded than coordinates, so attention must be paid to situations involving population in the remote areas, indigenous populations, and residents of informal settlements whose addresses may not be clearly defined or may be unknown. The same applies to countries that use mixed-mode data collection, making sure that in all the modes used it is possible to collect precise locational information.

3.55. A range of techniques and technologies exist for incorporating geospatial information into census-mapping exercises. However, the financial and technological realities, technical capacity, and team experience of each NSO must be considered when selecting techniques and technologies. These recommendations aim to present and develop the essential principles behind geocoding of census information. They emphasize the importance of developing and improving the utilization of geospatial information and its integrating with census statistics.

B. The role of geospatial information in the census

3.56. Mapping has been an integral part of census taking for a long time. The fundamental act of taking a census is inherently a geospatial endeavour; enumerating a population in each area, at a particular location, and in a specific dwelling, is fundamentally geographic in nature. Over the years, census maps have played a critical role in all processes from preparation to dissemination of the census results. More recently, the use of geospatial information has been increasing in various stages of censuses around the world.

3.57. In general, mapping, including its digital version, serves several purposes in the census process, as follows:

(a) Maps ensure coverage and facilitate census operations (pre-enumeration). The census office needs to ensure that every household and person in the country is counted and counted only once. For this purpose, census geographers partition the national territory into contiguous small data collection units. Maps showing such EAs thus provide an essential control device to ensure full coverage of the census.

(b) Maps support data collection and can help supervise census activities (during enumeration). During field-based census, maps ensure that enumerators can easily identify their assigned geographic areas. In a more technologically mature census, it is possible to load these maps onto handheld data collection devices equipped with GNSS and issue alerts if the enumerator leaves his work area and encroaches on to another's work area (geofencing).

(c) Maps are also issued to the census supervisors assigned to enumerators to support census operational activity planning and control tasks. Maps can thus also play a role in

supervising the progress of census operations. This allows supervisors to strategically plan, make assignments, identify problem areas, and implement remedial action quickly.

(d) More recently, digital maps or geoportals can be used to display the progress of the data collection in real time. In particular, these maps, when used with handheld equipment with GNSS, will be able to accurately display the houses visited with coordinates. It is possible to view this data at different scales, such as territories, regions and EAs. Base layers such as satellite images and Open Street Map can be incorporated to help identify areas not covered or even not previously mapped.

(e) Digital maps can also be used to verify spatial patterns of refused, absent or unoccupied housing units, directing surveillance efforts or even localized advertising campaigns to increase response rates.

(f) Maps are essential tools for presenting, analysing, and disseminating census results, offering a visual representation of demographic and social patterns. Geospatial information plays a key role in enabling the production of geocoded census outputs, including grid-based data, which enhances the ability to analyse and visualize census results at various spatial scales. This supports evidence-based decision-making across the public and private sectors.

(g) Digital maps loaded with high-resolution satellite images, household coordinates and aggregated demographic data have also proven useful for disaster response action. The confidentiality and granularity or geographic disaggregation of this data should always be considered.

(h) Digital maps loaded with recent high-resolution satellite images can also be useful to verify the migratory dynamics of nomadic populations or the expansion of informal settlements, ensuring the count of these population groups.

3.58. The census enumeration team needs to have a set of unique maps covering the entire country that accurately defines the boundaries of each EA. Therefore, the quality of maps used in the census has a major influence on the quality and reliability of census data.

3.59. The types of maps required for census management include the following:

(a) small-scale reference maps for use in the national statistical office or census agency to manage the overall operation;

(b) large-scale topographical maps for use by enumerators; and

(c) maps of the subregions or administrative areas, for the use of managers, showing the location of small population settlements and dominant physical features, such as roads, rivers, bridges and the type of terrain.

3.60. The use of satellite images and earth observation (EO) data to generate these maps is now increasingly common in many countries (see paragraph 3.74(a) and 3.74(c) for a description of EO). NSOs with fewer resources can leverage readily available free and open global human settlement maps. While such maps offer a valuable starting point, it is important to recognize their potential limitations in terms of currency and accuracy. Regular updates and ground-truthing[86] are essential to ensure map reliability. Collaborating with NGIAs to develop and maintain accurate cadastres is vital for supporting census operations and broader geospatial data needs.

[86] Ground-truthing is a verification process that involves comparing map data with real-world observations to ensure that the map accurately represents the physical landscape.

3.61. Careful consideration should be given to organization and management of geospatial support activities during the census planning and preparation phases. The lead time necessary for the implementation of a mapping and geospatial support solution for a country will be determined by a wide range of factors, including the number of maps to be produced, the technology available to produce and assign them, the availability of funds to acquire additional resources and the time required for distribution of the maps to field staff.

3.62. It is still the case that in many countries there is only a limited range of maps available and that these often do not show sufficient detail to enable the boundaries of small areas to be clearly defined. This is particularly likely to apply in areas of unplanned or informal settlement. However, with high resolution satellite imagery of sufficient quality and other current EO data sources this is no longer the case. Unplanned areas in some countries can be mapped precisely thus facilitating more efficient field activity planning. These spatial data sources can be further enriched with other material, such as:

(a) lists of households, preferably compiled by NSO staff as part of the process of delineating EAs, but on occasions provided by local leaders (see paragraphs 3.150–3.156); or
(b) a textual description of the boundary, including roads, railway lines, power lines, rivers and other physical features. This description may also include obvious landmarks on the boundary (such as school buildings, water points and other reference features).

3.63. This can create the base of a rich geographic framework for use as a reference moving forward. In this way, census activities can be used to enhance or update the national map or provide an important source for the update of an existing national geospatial database. The geospatial data collected during planning and field activities should be treated as an extremely important and rich data source for further geospatial processing.

3.64. Previous experience has shown that relying entirely on a list of households, written or verbal descriptions and directions or local knowledge of the area boundaries, will often lead to confusion and error because people tend to have mental images of places that may not align with the area as it is really reflected in the design of the EA. Similarly, the supervisor's mental map of an EA may differ markedly from that of an enumerator, and from enumerator to enumerator. To overcome such problems, it is important that the best possible quality maps be the basis for census enumeration operations and that the data collection staff receive comprehensive training in the correct use of use of the maps, digital maps, GNSS handheld devices and any associated documentation provided.

3.65. Despite the great value that paper maps have had in previous censuses, technology today allows many of the limitations of this tool to be overcome. The widespread use of handheld devices equipped with GNSS positioning technology allows census takers to locate themselves, within their allocated EA by receiving alerts in case of invasion of other EAs, among other functionalities. Furthermore, collecting coordinates of the units investigated in the census allows for a significant improvement in supervision activities.

3.66. Where a digital base map is prepared, this may be used in conjunction with a series of geospatial layers as the basis for coding information collected in the census. This could apply to address of usual residence now and/or in the past, or place of work or education.

3.67. The use of paper maps is, however, still valuable, especially in areas not or inadequately covered by GNSS or where the use of technological equipment poses risks, such as theft or damage. Many countries have successfully implemented a hybrid approach, combining paper and digital maps to optimize census operations. Paper maps offer flexibility for on-the-ground adjustments and

data recording, while digital maps provide precise location information and enable advanced spatial analysis. A hybrid approach, combining both formats, can optimize data collection and management efficiency.

3.68. Statistical agencies should, however, avoid a paper-only solution. This approach greatly limits exploring the potential of location data at all stages of the census and reduces the possibility of integrating statistical and geospatial data. The subsequent costs in time and financial resources to standardize and process paper-collected data are unlikely to be justified. A paper-only solution should only be considered as a last resort and in very limited circumstances.

3.69. With the rapid growth in freely accessible high and medium resolution EO data, specifically satellite and aerial imagery, there are increasingly few, if any, circumstances where hand-drawn maps or hand-collected data should be incorporated as an acceptable standard solution for a modern census. A paper-based collection solution is not subject to the same level of quality assurance and quality control that can be built into a digital capture solution. In the case of capturing geographic data (such as the location of a housing unit) on paper, the basic geospatial reference data is not captured and may require additional time and research in the central office or necessitate additional field visits.

3.70. The next two sections cover geospatial information technology and GIS, both fundamental elements for building a geolocated database in a census. These recommendations speak directly to Principle 1 of the GSGF (Use of fundamental geospatial infrastructure and geocoding) whose objective is to obtain high quality, standardized location references such as physical addresses, property or building identifiers, or other location descriptions, in order to assign accurate coordinates, and/or a small geographic area or standard grid reference, to each statistical unit at the microdata/unit record level.

3.71. Time and date stamping locations is vital to clearly place the statistical unit both in time and in space. By grounding census data in a strong geospatial foundation, NSOs can more effectively develop additional statistical products, integrate administrative records, and disseminate census results at granular levels. This enables a wide range of data-driven applications and insights.

C. Geospatial information technology for census

3.72. Before census mapping commences, the census agency needs to determine the appropriate technology for doing so. Countries need to choose technologies to improve efficiency of census operations, data quality and timeliness, balancing these with cost factors within the context of their national needs and circumstances. The application of technology must also ensure that the confidentiality of data is maintained.

3.73. In general, countries need to approach the use of technology for mapping as a continuous process rather than merely a sequence of mapping and dissemination operations. Use and application of geospatial technologies and geographic databases are very beneficial in improving the overall quality of census activities at all the stages. Major technological advances include the widespread availability of personal computers, handheld devices, and personal digital assistants, GNSS and GIS software, and low-cost or open-source aerial and satellite imagery. These advances would be of interest to NSOs in collecting more accurate data in a timely manner.

3.74. Where accurate and current maps at relevant scales are not available for a country, or part of a country, the technological alternatives described in the following paragraphs could be employed subject to consideration of some constraining factors:

a) ***Satellite images*.** A satellite image typically covers a large area and can be cost-effective compared to other sources. Imagery should be pre-processed by the supplier so that it is rectified and georeferenced (by which a known scale and orientation, with some latitudes and longitudes, is printed on the face of the image). Satellite image data have gained in volume, popularity, and ease of use. Satellite imagery, if used pragmatically, can save countless person–hours by focusing attention on critical areas. Remote sensing data can be used as an independent check on the field verification process.

b) ***Aerial photography*.** Acquisition of aerial photographs for large areas of a country may be expensive. However, existing archives of photographs can be an excellent resource for preliminary counts of dwellings and as a base for basic maps, provided they are dated. Knowing the specific time period when the photographs were taken is essential for their accurate use. In many cases digital images captured from an aircraft or drone-based platforms are a cost-effective way of incorporating components of a GIS. However, NSOs should not regard satellite images and aerial photography as an end product in themselves. The information contained in both requires further processing and interpretation to be useful for census purposes. This adds to the burden/cost of its use. Sometimes, where there is no capacity for additional field operations, or no current digital information is available, open data of human settlements can be a rather good asset to fill geospatial data gaps (see point (e) below).

c) ***Other Earth Observation (EO) Data.*** EO is the process of gathering information about the Earth's surface, waters, and atmosphere via ground-based, airborne and/or satellite remote sensing platforms. Unmanned Aerial Vehicles (UAVs), also known as drones, can be used to collect high-resolution imagery and data for specific areas, providing a more flexible and cost-effective option for Earth observation. UAVs can be particularly useful in areas with limited satellite coverage or for monitoring changes in rapidly developing regions. The acquired data, usually in the form of digital imagery, are processed and analysed to extract different types of information that can be used to monitor and assess the status of – and changes in – both the natural and human-made environments. For situations where there is no digital map available, or specific acquisition and processing of satellite imagery or aerial photography is unpractical, open and free global human settlement maps can fill data gaps and serve as information baseline for census activities. These products delineate built-up areas and human settlements based on open satellite imagery data. These resources are well documented and widely used to delineate human settlements. For some regions of the globe there are also building footprints which increasingly available for free and which can provide an initial georeferenced database of buildings. Digital Surface Models (DSMs), which represent the Earth's surface with all the objects on it, and Digital Terrain Models (DTMs), which represent the bare surface of the ground without any objects like plants and buildings can also allow EO information to be incorporated into census mapping.

d) ***Global Navigation Satellite System (GNSS).*** The standard generic term for satellite navigation systems that provide autonomous geospatial positioning with global coverage. This term includes for example, the Global Positioning System (GPS), Global Navigational Satellite System (GLONASS), Galileo, Beidou and other regional systems. As laptop computers and other portable computing devices equipped with GNSS are becoming less expensive, integrated field mapping systems are becoming a viable option for collecting accurate field data (in terms of latitude and longitude) for census purposes. Advances in technology, including GNSS, wireless communication, and computer miniaturization, have made numerous new applications for portable GIS possible, particularly the development of specialized software for census fieldwork. The use of this equipment combined with digital maps or geoportals will bring

significant gains to the supervision of activities, including the possibility of recording enumerators' routes.

e) ***Georeferenced address register.*** A high-quality, comprehensive, updated, and georeferenced address register of each building and dwelling can lend great support in planning and organizing a census. A georeferenced list of addresses can play a central role in many fieldwork operations and will provide the key to accurate delivery, collection, and follow-up of questionnaires. The best way to associate each address with a location in physical space of a map is to specify its coordinates in a proper geographic reference system. With geographic coordinates addresses can be entered in available maps. Ideally, addresses and coordinates should be collected in combination, as they are complementary information. In rural or remote areas, it may not be possible to collect the address, making it important to collect the coordinates.

3.75. The implementation of strategies using such technologies must be thoroughly planned with the guidance of qualified staff or external experts with formal qualifications in the use of advanced mapping technology. It is particularly important that the cost of acquiring and maintaining the hardware required to use this technology is factored into the budget (and a sound cost–benefit analysis undertaken to support such changes), and that adequate plans are made to ensure the availability of sufficient quantities of hardware in time for the census.

3.76. It should be noted that there may be additional risks due to the need for equipment to be operated in suboptimal conditions, including poor weather, dusty conditions, or poor lighting. Despite its versatility, GNSS may not be able to differentiate the coordinates of overlapping or closely located dwellings in multistorey buildings, and, in this circumstance, should only be regarded as providing coordinates for the building rather than the dwelling units within it.

3.77. Software and applications that can read both static and interactive maps should be considered, as they will greatly help in deciding the type of maps to produce. Interactive maps are very useful in the identification of EAs in countries where structures have no addresses. Software and application selection should consider a balance of cost-effectiveness, compatibility, and usability. In addition to the geospatial team's expertise and the need of enumerators, input from other relevant agencies, such as national cartography and geospatial information authorities, is important in order to align technology choices with broader national geospatial strategies.

3.78. It is also important to ensure that where such systems are employed, they are clearly understood by field staff. This can be achieved by ensuring that the staff, whether at the mapping update (pre-enumeration) stage or enumeration stage, are given adequate training in the interpretation of the maps. Should the maps be incorporated into digital devices such as personal digital assistants, the staff should be trained in the use of both the hardware and the software.

3.79. As with all other significant changes to census procedures, it is crucial that census geographic and mapping processes are successfully included in tests prior to being used in the main operation. This is particularly the case where a change in level of technology is being considered.

D. Geographic information systems

3.80. A geographic information system (GIS) can be seen as a system of hardware, software (commercial and/or open source) and procedures designed to support the capture, management, processing, analysis, modelling, and display of spatially referenced data. In practical terms, a user’s

interaction with such a system may range from a simple hand-held or desktop mapping facility or a web-based mapping and digital visualization application.

3.81. Careful planning should be undertaken to implement a complete GIS system that can solve complex census operational planning, providing solutions for: managing operations; identifying, and resolving enumeration coverage problems; monitoring field operations; enforcing quality controls; producing detailed georeferenced inventories; and supporting a variety of dissemination platforms, products, and formats. The ability to use geographic space to integrate and manipulate data sets from heterogeneous sources makes a modern geospatial system a vital and the fundamental tool for planning and managing the modern census.

3.82. For example, a GIS provides functions for the areal interpolation of statistical data in cases where the administrative unit boundaries have changed between censuses. While the development and implementation of such a repository of georeferenced data can be daunting and take extensive planning, simple desktop mapping systems that generate thematic maps from a database of base maps and indicators will satisfy the needs of most NSOs. The census should be seen as an opportunity to make strides and build upon previous geospatial operations. Once collected, a plan for geospatial data maintenance can be developed, and each census can provide the opportunity for an update and advancing the development of the NSDI.

3.83. GIS technology should be considered only at a level appropriate to the skills and resources available and should constitute an integral part of the overall work of the organization. Cooperative arrangements with other agencies – especially with the national geospatial information agency or national mapping agency – should be pursued particularly regarding the acquisition and maintenance of base map data, which usually is not the responsibility of the statistical organization. The United Nations Integrated Geospatial Information Framework (UN-IGIF) provides a basis and guide for developing, integrating, strengthening, and maximizing geospatial information management and related resources supporting NSOs in seeing themselves in the national geospatial data production ecosystem.

3.84. Given the great potential of geospatial information for censuses, NSOs should pursue the development or implementation of GIS and integration of geospatial information whenever possible, including for intercensal surveys. The integration of statistical and geospatial information must be an objective of NSOs, in their role as important institutions that produce or consume geospatial information. Therefore, NSOs must pay attention to the five GSGF Principles for internal structuring.[87]

3.85. NSOs must also understand GIS as an important tool for spatial analysis and dissemination of census data. The use of interactive maps, geoportals, and user-centred geography construction tools[88] tend to be more direct and simple ways to interpret information than tabular data, especially for the general public. NSOs provide vital information about current demographic conditions and future trends to policymakers in a range of sectors, such as health care, education, infrastructure

[87] See https://unstats.un.org/unsd/statcom/51st-session/documents/The_GSGF-E.pdf

[88] See https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/buildacustomareaprofile/2023-01-17.

planning, agriculture, and natural resource management; and the availability of spatially referenced census databases is an essential prerequisite to facilitate the use of demographic data.

3.86. In this regard, it should be noted that the GIS should be capable of generating additional geographic delimitations beyond those used in the census, such as school districts, water catchment areas, or power service units. These entities will have to be constructed from the smallest geographically identified units available in the census (for example block faces, grid squares or EAs). If (as is the case in most developing countries) EAs are the smallest unit, this will have important implications for the establishment of EA boundaries. Cooperation with the authorities responsible for these geographic entities before the boundaries of enumeration areas are drawn can reduce later problems.

3.87. With the rise of free and open-source software, the cost-benefit for rapid deployment and implementation of a GIS programme in a national census is irrefutable. However, in some developing countries with limited capacity and resources, GIS may need to be introduced carefully and gradually, given that is a rather complex and resource-consuming technology. As an alternative to immediately launching full-scale GIS applications, NSOs may wish to start with a simple and robust design that is likely to be understood and maintained by a wide array of users, transferable to a wide range of software packages and independent of any hardware platform.

3.88. GIS implementation in a developing country may follow a hierarchical strategy, with the NSO employing a high-end commercial GIS with extensive capabilities for handling and analysing large amounts of spatial data. Widespread dissemination of databases can then be achieved by creating a version of the finished databases using a low-end mapping software format for distribution at low cost and through web dissemination of macro information in an online GIS.

3.89. In recent years, many countries have adopted GIS to facilitate census mapping in the production of EA maps and outreach products. As the cost is decreasing, geospatial technologies and inputs are being made available free of charge or openly, and the basic technology is now well established, its use can be expected to continue to expand in future census rounds. It is likely that the census could help to increase the capacity of the statistical service. The adoption of GIS should therefore be seen as an important strategic decision, with impacts that go beyond the operation of the census, and many issues need to be considered.

3.90. A geospatial database at the EA or grid level is a cornerstone for the national statistical system, providing integrated statistical and geospatial data essential for managing, analysing, and disseminating census information and monitoring changes over time. It also serves as a fundamental building block for a national spatial data infrastructure, enabling the integration of socioeconomic and environmental data for evidence-based decision-making. A prerequisite for building a geographic database at the EA level or in grid format is the development of a geocoding scheme, in which each EA has a unique code - an administrative identifier which can be used to link geographic features to the attributes registered for them.

3.91. The potential benefits and challenges of GIS are summarized as follows:[89]

(a) Benefits

(i) It provides efficiencies and insights for field preparation for data collection;

(ii) It facilitates closer linkage between maps for enumerators and map-based products for users;
(iii) It enriches the dissemination of census data as they can be visualized in geographic areas for easy understanding by users;
(iv) The cost of intercensal updating of the base map will be less with a digital base map, enabling among other things the construction and updating of sampling frames;
(v) Producing duplicate maps may be less expensive with a GIS solution;
(vi) GIS will have increased ability to undertake quality assurance of geographic boundaries;
(vii) The census agency will have a greater ability to perform spatial queries and advanced analysis under GIS, including the application of the degree of urbanisation method; and
(viii) The space needed to store input maps for digital purposes will be far less.

(b) Challenges
(i) It requires additional technical expertise;
(ii) It requires a higher level of computing infrastructure;
(iii) A clerical census system can proceed on the basis of basic maps. However, use of GIS in this task requires that a digital map base exists. If it is necessary to create the digital map base, significant lead times are required as well as significant funding. In both cases, more experienced technical staff are required;
(iv) In most cases, the preparation of maps or GIS will not be the core business of the statistical agency.

E. Census geography

3.92. Prior to developing a geospatial support strategy for the census, consideration needs to be given to the geographic classification to be used and the geospatial infrastructure necessary to carry out the required tasks. As the geographic unit used for basic census collection usually[90] determines the minimum geography on which the census data can be disseminated, a geographic classification should be devised in conjunction with the development of census mapping.

3.93. The publication and utilization of the geographic classification system developed for the census across the statistical system, the NSDI, and administrative domains will enhance the value and utility of census data by improving its compatibility and integration with other data sources. The details of designing a general geographic classification, including the definition of the various areas of the geographic classification and their relationship to one another, are more complex than those involved in census geospatial activities and will not be covered further in this chapter.[91] However, the design of EAs and the careful planning of their relationship to other census management areas is of vital importance for the census and is outlined in the following paragraphs.

3.94. It is of critical importance to ensure that the boundaries of various administrative units are available in a geospatial format (a shapefile), and frozen at least six months before the census reference date so that no further jurisdictional changes are affected until the enumeration is over.

[90] In some cases, the geographic unit for collection slightly differs from the tabulation/dissemination geographical unit.
[91] For a full elaboration and details, please see *Handbook on Geospatial Infrastructure in Support of Census Activities,* United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

This would be of considerable help in delimiting EAs and minimizing the chances of omission or duplication of enumeration units (persons and household). To this end, the geospatial data infrastructure must be organized with nested hierarchical levels, while the more detailed census geography is a subdivision of the higher-level geography without gaps or overlaps (for example, state, region, province or municipality, district, neighbourhood, EA). To be useful, administrative boundaries must be stored as geographies in geospatial datasets containing identifiers (geocodes) and their attributes.

3.95. If point-based geospatial data is collected in conjunction with each individual enumeration, then the dissemination unit can be easily conformed to different geographic layers. In other words, the availability of a precise coordinate for the location of each household can tie each individual enumeration response to a precise location where it was recorded. Then there is a much greater flexibility in the eventual dissemination units, and those point locations can be aggregated to any geography.

3.96. Although this Section deals more directly with census geographies for census planning, organization, and management, such geographies are not disconnected from census dissemination geographies, as many serve both purposes. In this sense, it is essential that there is a holistic view that sees geography as a tool for integrating data. It is important to consider that, in many countries, census data generate functional areas[92] (such as, electoral districts, commuting areas, etc.) and that census geographies themselves are geostatistical functional areas. Functional areas form the link between data collection and the subsequent use of census data to inform policy decisions, resource allocation, and other actions. They can be used to visualize data, but also for analytical purposes and, if stable, trends over time.

3.97. Point-based (geolocated) information must be considered with the enumeration, dissemination, administrative and functional geographies so that aspects of spatial disaggregation, sample expansion and confidentiality issues such as spatial differentiation can be assessed.

3.98. A common set of geographies ensures that statistical data is geospatially enabled in a consistent manner and is capable of being integrated at the aggregate level; and ensures that users can discover, access, integrate, analyse, and visualise statistical information seamlessly into the geographies of their interest. This speaks directly to the GSGF Principle 3 (common geographies for dissemination of statistics), the goal of which is to support the provision of a common set of geographies that ensure the consistent geospatial aggregation and dissemination of statistical data, irrespective of whether they are in gridded or administrative boundaries.

(i) Administrative hierarchy

3.99. One of the earliest decisions in census planning pertains to the administrative areas for which census data will be reported. Linking data to administrative geographies provides the spatial component that further enables data integration and broader comparability. Administrative areas can be any special geographic unit, but mainly they are units of administration, that is, some governmental authority has jurisdiction over the territory. Census preparation involves creating a list of all administrative and statistical reporting units in the country. The relationships between all types

[92] Functional areas relate to, and support, the organisation and management of people, communities, society, and their activities in geographic space. For a full elaboration and details, please see The Global Fundamental Geospatial Data Themes, United Nations, 2019.

of administrative and reporting unit boundaries should be defined. Every country has its own specific administrative hierarchy, that is, a system by which the country and each lower-level set of administrative units (except the lowest) are subdivided to form the next lower level. Census operations provide an opportunity to verify and update these boundaries.

3.100. Only some of these hierarchical levels may have actual administrative roles; for example, the province, district and locality levels may have capitals with local government offices that are responsible for those areas. Other units may have statistical roles alone; that is, they are designed for the presentation of data and not for administering the territory. In some instances, however, administrative units may not be completely nested. Especially when considering both administrative and other statistical reporting units, the NSO may need to deal with a very complex system of geographic regions.

3.101. In some sub-national regions, the establishment of a definitive list of units is a major operation because of difficulties arising from the frequent fragmentation, disappearance, or combination of small localities, and from changes in name, variations in spelling, the existence of more than one name for the same place or the use of identical names for different places. This listing should be held as a formal database or as an integral component of the databases forming part of a census fundamental geospatial infrastructure.[93] This should be incorporated into basic census planning to allow for enough advance planning for the geospatial activities.

(ii) *Delineation of enumeration areas*

3.102. The design of EAs should take various criteria into account. Correctly delineated, EAs will:

a. Be mutually exclusive (non-overlapping) and exhaustive (cover the entire country);
b. Have boundaries that are easily identifiable on the ground;
c. Be coterminous with the administrative hierarchy;
d. Be compact and have no pockets or disjoined sections;
e. Have populations and surfaces areas of approximately equal size;
f. Be accessible enough to be covered by a single enumerator within the enumeration period;
g. Be homogeneous in aspects of morphology and function (such as surface area, land cover and use, and rural and urban);
h. Be planned as a (potential) common geography for dissemination
i. Be appropriate for other types of censuses and data collection activities as well.

3.103. For the delineation of EAs, population size and number of households are generally the most important criteria, but surface area and accessibility must also be taken into account to ensure that an enumerator can service an EA within the time allotted. Another factor that is sometimes

[93] For further details on GIS mapping, see *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009.

taken into consideration is a hard-to-count index[94] which assigns a score to each EA taking account certain population characteristics, such as the proportion of immigrant populations, multi-occupied dwellings and other factors which make the enumeration difficult. Testing should be conducted to determine enumeration capabilities and capacity. While it may be difficult to assign a hard-to-count index through testing alone, such an index can often be devised using data from the previous census. Population density will vary between rural and urban areas and have an impact on the size of EAs, since enumeration can proceed more quickly in densely populated areas towns and cities than in the sparsely populated countryside. Given these considerations, EAs that are significantly larger or smaller than the national average may have to be defined.

3.104. As a basic reference, before the delineation of EA boundaries, the number of persons living in an area and their geographic distribution need to be estimated. Unless there is up-to-date information from an exhaustive registration system or some other accurate and recent information source (administrative records, georeferenced address register etc.) these numbers will need to be determined by counting the housing units, the extent of residential built-up surfaces, the number and the size of households. This is usually done through fieldwork, cooperation with sub-national governments or, more recently with the extrapolation from previous census results by means of aerial photographs, satellite imagery or derived products such as building footprint and settlement data. There are also freely available high-resolution population estimates produced using geospatial data and population from sub-national projections or surveys to approximate current population distribution. But this data should be used only if the accuracy is of proven acceptable quality.

3.105. EA boundaries need to be clearly observable on the ground. Even if they do not have considerable geographic training, all enumerators need to be aware of the boundary limits of the area for which they are responsible. Thus, population sizes between EAs may be varied to produce an easily identifiable delineation. Natural features that can be used for this purpose are fences, landmarks, roads, railroads, streams and rivers, lakes, boundaries, or any other feature that defines an observable area. If handheld equipment with embedded GNSS technology is used in the enumeration, alerts through geofencing should be used.

3.106. Procedures for the delineation of EAs should be developed that will allow, as far as possible, comparability of areas from one census to the next. Change analysis at the local level is greatly facilitated if the units of enumeration remain compatible between censuses. A unique code should be assigned to each EA and the changes (for example in the case of splitting areas with a high level of growth) should be tracked. The NSO is often the custodian of coding schemes in the country and should also be the focal point for the design of the census-mapping codes. In cases where this is not possible, the criteria can outline design principles that will allow users to easily compare EA-based data across censuses. When the population and housing census are conducted separately, effort should be made to use the same EA frame for both censuses as far as possible. The delineation of

[94] The "hard-to-count index" identifies areas or populations likely to be undercounted in a census. It considers factors like high poverty rates, low education levels, linguistic isolation, unstable housing conditions, and historically undercounted minority groups. These factors correlate with low census response rates and help target efforts to improve accuracy. The U.S. Census Bureau offers detailed documentation on calculating a hard-to-count index. For more information, refer to: Understanding Hard-to-Count and Historically Undercounted Populations (link: https://www.census.gov/newsroom/blogs/random-samplings/2023/10/understanding-undercounted-populations.html) and,
Counting the Hard to Count in a Census (link: https://www.census.gov/content/dam/Census/library/working-papers/2019/demo/Hard-to-Count-Populations-Brief.pdf)

EAs that remain consistent between censuses is primarily valuable for ensuring comparability of outputs, especially when EAs are used for dissemination purposes. However, the choice of data collection geography can vary depending on the methodology and technology adopted.

3.107. In addition to the limits and physical-natural features to support localization, the delimitation of EAs must pay special attention to building blocks and block faces, as these will be essential for the enumerator to move within the EA and locate the households. These spatial elements must also have their unique geocode following the spatial hierarchy. In censuses using handheld GNSS equipment, the enumerators' routes can be tracked to improve and update the internal mapping of EAs, with potential use for national mapping.

3.108. In censuses that simultaneously use short and long questionnaires (see paragraphs 3.9-3.17), special attention must be paid to the distribution of the statistical weights of the sample in the EAs to guarantee adequate territorial coverage of the sample and not over-burden census enumerators. It is also necessary to ensure that the areas of dissemination of sample results, generally aggregations of EAs, include all EAs without gaps or overlaps to guarantee coverage and expansion of the sample throughout the territory.

(iii) Delineation of supervision areas

3.109. Supervisory areas provide the means for a group of enumerators to be effectively managed. After delineation of EAs, the design of supervisory maps is usually straightforward. Supervision areas consist of groups of contiguous EAs that share some of the same characteristics of the constituent EAs. The EAs assigned to the same supervisor should be compact, to minimize travel times, and of approximately equal size.

(iv) Delineation of census management areas

3.110. Census management areas will consist of aggregations of supervisory areas brought together for ease of managing the enumeration staff. Where existing government staff and structure are used for enumeration purposes, the census management areas may be the same as the administrative regions. It should be noted that this may be a matter of administrative convenience and the hierarchy (or way of aggregating EAs into larger areas) for this purpose need not necessarily be the same as that for the dissemination phase, which should be driven by the needs of users. However, the greater the congruence between management areas and pre-existing administrative boundaries the easier the task of conducting the census in the field.

(v) Other census geographies

3.111. Progressively and with the use of geospatial technology and GIS, NSOs have been incorporating other geographies into the classification of census geographies, taking advantage of the ability to overlay spatial layers. These geographies are often used for tabulating results and presenting data, but their potential has increasingly been used for the planning and enumeration phases.

3.112. Many of these geographies arise from the need to disseminate results for specific population groups, such as indigenous populations. These geographies are often used for populations that lack recognized or delimited territories or are dispersed across various areas. The primary objective is to obtain statistical data for these populations within well-defined geographic areas that have a significant concentration of their members. Traditionally, data collection from specific population groups has involved either separate questionnaires or incorporating relevant questions into the main questionnaire. However, leveraging geospatial technology and targeted data collection

methods can provide a more efficient and effective approach. Geospatial technology (GNSS positioning) can be used to geo-enable specific census questions pertaining to these population groups. By integrating location data, NSOs can tailor data collection efforts to target specific areas where these populations reside. This can help avoid the need for separate questionnaires or specific questions in the main census questionnaire, while still capturing the unique characteristics and needs of these populations.

3.113. In other words, areas with characteristics of specific population groups need to be pre-identified and mapped. These areas can be either incorporated areas – legally incorporated under state law, with a legally defined boundary – or census designated places – statistical geography representing closely settled, unincorporated communities that are locally recognized and identified by name. These areas may be composed of aggregations of EAs or be contained within EAs. Enumeration of households within these areas of interest will allow for targeted collection of information relevant to those specific population groups[95].

(vi) *Geocoding*[96]

3.114. A geocoding scheme for the census needs to be determined on a country-by-country basis, ensuring codes are unique, unambiguous, and should preferably be designed in collaboration with the national geospatial information agency. The most important principles in the design of a coding scheme are flexibility, expandability, and compatibility with other coding schemes in use in the country. As set out in Principle 2 of the GSGF (geocoding unit record data in a data management environment), this includes the use of data management tools, techniques, standards, and good practices to facilitate the linking and management of geocodes within statistical datasets. This also serves to ensure that privacy and confidentiality requirements are correctly managed for the disseminated data.

3.115. The coding of each housing or population unit to a specific longitude and latitude, or to the smallest geographic unit, preferably smaller than the EA, allows for a flexible production of different geographic tabulations of outputs, and production of comparable area-based geography over time. Two somewhat different approaches are available for coding the location of housing or population units. The first approach is to code all units to the lowest-level EA. The second approach, which permits finer geographic specificity, is usually based on some coordinate, address, or grid system. In either situation, careful planning should allow for maximum flexibility in future aggregation and disaggregation of the data and geographic units.

3.116. A digital geographic database organised in geospatial layers[97] in vector format consists of a structured set of points, lines and polygons, or grid cells. Each geographic feature — point, line, or area — has a unique identifier that is used by the system internally. This internal identifier is not usually accessible by the user and should not be modified externally. A more meaningful identifier is needed that can be used to link the geographic features to the attributes recorded for them. For the

[95] Brazil has applied it for two population groups of the Brazilian Traditional Peoples and Communities, the indigenous people and the quilombolas in the 2022 census.
[96] This section draws heavily upon material in *Handbook on Geospatial Infrastructure in Support of Census Activities*, United Nations publication, Sales No. E.09.XVIII.8, United Nations, 2009 and in the *Global Statistical Geospatial Framework,* United Nations, 2019.
[97] Geospatial layers can also be in raster format, with regular grids of defined resolution.

EAs and administrative units, this link is the unique EA or administrative identifier, which is listed in the master file of all geographic areas relevant in the census.

3.117. Indeed (as noted at paragraph 3.106), a unique code needs to be assigned to each EA that is then used in data processing to compile enumerated information for households in the EA and to aggregate this information for administrative or statistical zones for publication. This is the numerical code that provides the link between the aggregated census data and the digital EA boundary database stored in the case of using GIS.

F. Implementation of a census geospatial information programme

3.118. The main goal of a census geospatial information programme is to structure a location-based conduct of census operations and activities and to explore the possibility of integrating statistical and geospatial information. The development within the NSOGSGF requires the coordination of a series of complex tasks with relatively long lead times. It is important that project plans are established to manage this process. In the implementation of a census geospatial information programme, NSOs must consider standardisation and interoperability of data which will lead to improved efficiency and simplification in the creation, discovery, integration, and use of geospatially enabled statistics and geospatial data. This will increase the potential application of a larger range of data and technologies, and thereby enable a wider range of information to be available and accessible for use in decision making. It will also facilitate better cooperation between all stakeholders producing and using statistical and geospatial information.

3.119. NSOs should consider a long-term investment in geospatial information technology that can be fully applied in the statistical business processes. Principle 4 of the GSGF (statistical and geospatial interoperability) defines the preconditions for statistical and geospatial data to work as a data ecosystem, in which those involved interact with each other to exchange, produce, and consume data. Interoperability between statistical and geospatial data and metadata standards is needed to overcome structural, semantic, and syntactic barriers between data and metadata from different communities and providers. Furthermore, it is necessary to enhance the effectiveness of the discovery of, access to, and use of, geospatially enabled data. Often, full interoperability of data first requires the removal of obstacles in country-level laws, policies, and organisations that hamper cooperation between stakeholders and create barriers between producers and end users. The UN-IGIF can provide NSOs with valuable resources and tools to advocate for the necessary changes and facilitate greater data interoperability.

3.120. The main activities to be reflected in implementing a census geospatial information programme are discussed below:

(a) Establish a geospatial unit. A geospatial information programme requires a specialized project team. With each round of censuses, these activities go beyond census mapping, including: the collection of coordinates and addresses; the spatial supervision of data collection; and the development of geoportals for dissemination, among other activities. Therefore, it is important to have an interdisciplinary team of

geographers, cartographers, data scientists[98], statisticians and other specialists that allow for an integrated vision. When geospatial information incorporation and mapping activities are outsourced (see Section G below), this team will be responsible for specifying census requirements for products and coordinating agreements with the service provider.

(b) ***Developing a timetable*.** The critical date is the date on which geospatial inputs and maps must be delivered or made available to teams in the field. The geospatial information programme should begin early in the census cycle to allow sufficient time to produce national mapping coverage well in advance of the census date and before field staff are trained. In censuses where handheld equipment and geoportals are used, sufficient time must be given for IT development. Hiring an updated high-resolution satellite imagery service is also an activity that will take considerable time to complete.

(c) Sourcing of basic mapping and digital geographic data. A major step in the mapping project is establishing a base map of the country, including digital mapping data if required. If a census-mapping project already exists, the NSO may still require updates to their existing map holdings. In this context, it is important that NSOs see themselves within the geospatial data production ecosystem as an important stakeholder and seek exchanges and partnerships in the context of the NSDI and the implementation of the UN-IGIF national action plan.

(d) Integrating statistical and geospatial information: The census geospatial information programme must be guided by the five principles and key elements of the GSGF to ensure the production of harmonised and standardised geospatially enabled statistical data and that the census results can be integrated with statistical, geospatial, and other information to inform and facilitate data-driven and evidence-based decision making for supporting local, subnational, national, regional, and global development priorities and agendas.

***(e)* Disseminating information geographically:** It is essential that the census geospatial information programme includes geographic, cartographic, and geospatial products, platforms and applications resulting from the census. This topic is discussed below (in Section IX.C4), but planning must be carried out in an integrated manner. In general, maps and geoportals are powerful, simple language tools for interpreting data and information.

(f) **Continuing capacity development plan.** The NSO should keep an ongoing capacity development plan for census teams encompassing: the conduct of periodic training in the use of geospatial information technology, GIS, integration of statistical and geospatial information; developing projects; and participating in local and international forums and meetings.

(i) Sources and types of hard copy maps

3.121. Where a hard copy base map is to be used, official published maps may be available from national or provincial government mapping agencies, the local government, or municipal bodies. Other sources of maps may be other government agencies or private companies. Where the maps

[98] Data scientists are individuals who use scientific methods, algorithms, and systems to extract knowledge and insights from data. They combine skills from various fields, such as statistics, computer science, and domain expertise, to analyse large and complex datasets.

are obtained from sources outside the NSO, permission to use the maps collected should first be sought from the original source and any copyright issues addressed.

(ii) Digital mapping data

3.122. When establishing a digital geographic database, a major consideration is the determination by the NSO of its data requirements. With increasing amounts of digital spatial data becoming available, it is also important that standards and a common data specification be produced to ensure data validity and consistency.

3.123. The key rules to be followed in selecting data items for inclusion are to question whether (a) the data item will be useful to enumerators in navigating their way around their enumeration area; and (b) the data item is relevant to users. Assessing the utility of data items to users in a census-mapping context must place significant emphasis on the user needs for small or customized areas. Data items that meet neither of those criteria should not be included in the database.

(iii) Updating maps or digital mapping data

3.124. Preparing or updating base maps, or the base map digital data, requires substantial resources. The final content of base maps will have a major bearing on the accuracy and completeness of EA maps and, subsequently, the effectiveness of the census enumeration. The updating of base maps should be scheduled according to priorities, based on areas in which changes to the number or characteristics of the people require the maps to be updated. Important features to be updated include:

(a) accurately named and presented roads and waterways;
(b) administrative boundaries; and
(c) landmark features, such as schools, place of worship, post offices, parks, and large buildings.

(iv) Operational design for enumeration and supervisory areas

3.125. Whether a hard copy or digital base is employed, an EA design manual should be produced that contains the design criteria and the procedures to be followed when designing the EAs. The manual can be used as a basis of training for those involved in the design process.

3.126. EA design should be conducted by regional statistical office staff, where these exist, to leverage local knowledge and expertise. While the specific approach may vary across countries, it is necessary to ensure that EA design is carried out using consistent criteria nationwide. This involves gathering information on population and boundary variations to determine the most effective way to design and update EAs. The output of this process should be a list that provides the enumeration phase with all relevant field data for each EA, and the dissemination phase with relevant output geographic data.

3.127. The design of field supervisor and management area boundaries can be determined at the completion of the EA design process through the aggregation of EAs, and the allocation of geocodes.

3.128. Quality assurance measures should be implemented to ensure that data are correct to a minimum standard, both for field navigation and for technical correctness in cases where a digital base is to be used as an output medium.

(v) Printing and content of field maps

3.129. Careful consideration should be given to the (considerable) time required for printing maps when establishing the project plan for census mapping. Digital options should be given primary consideration and accepted as the default format for any static reference maps. In this case, attention should be paid to IT development requirements.

3.130. Maps should be provided to every level of field staff. If paper maps are used, at least one map should be printed for every EA in the country. However, it is recommended that two copies of the map be produced, one copy to be used by the enumerator and the other by the field supervisor for training and reference purposes. Both copies should be retained by the NSO so that the information aggregated by the field team can be incorporated into the territorial base for the regular surveys and the following census cycle. The use of computer technology for data collection, such as tablets, laptops and other handheld devices with GNSS technology should be considered as the default solution and suitable for displaying maps available to field staff.

3.131. Other considerations for the preparation of enumeration maps, also considering digital versions available on handheld devices, tablets or computers, include the following:

(a) Enumerators may be required to navigate in poor lighting conditions and thus details should be easily read;
(b) The maps must be easily interpreted with text and symbols readily identifiable and correctly placed, along with the information being presented in a standard format compared to other source maps;
(c) All maps must include basic cartographic elements such as title, legend, scale and a compass setting for orientation and navigation purposes;
(d) Boundaries (such as EA boundaries) must be clearly and unambiguously delineated on maps;
(e) EAs must be distinguishable from the surrounding area, both visually and geographically;
(f) Folding or refolding of large paper maps (larger than A2 in size) is inconvenient for field staff;
(g) Paper and digital maps need to facilitate the addition of the enumerator's written comments;
(h) Production of the maps should be cost-effective;
(i) The EA maps should be suitable for dissemination purposes where this reflects user demands.

3.132. Maps for supervisors or regional managers should be of smaller scale, providing sufficient detail to identify major features. When using paper maps those should not be so large as to be difficult to handle easily in the field. In many cases, the use of inset or supplementary maps may be required if the map is to cover a relatively large area. For all levels of senior field staff, the maps should show the boundaries of all subsidiary units for which they are responsible.

G. Contracting out and partnerships for census geospatial operations

3.133. The development of a census geospatial information programme beyond rudimentary clerical systems requires considerable knowledge of mapping, cartography, and geographic systems. Furthermore, the census geospatial information program increasingly requires the integration of the resulting geospatial data with statistical data, as well as the incorporation of big data, machine learning and data sciences. If a census agency cannot draw on such skills internally, it may be required to contract out some or all the elements of the preparation of geospatial input preparation. A partnership with the national geospatial information agency (NGIA) can fully or partially meet the NSO's needs.

3.134. Mapping for field purposes under a contract or agreement basis requires the NSO to specify its requirements and prepare clear terms of reference to the contractor. These may include the following:
(a) acquiring the base map data;
(b) creating (or obtaining) the statistical boundaries and aligning them to the base map;
(c) providing a process for the EA designers to advise on changes to boundaries (and updates to associated spatial data);
(d) producing hard copy maps or soft copy using handheld devices such as a smartphone or tablet computer as specified for fieldwork; and
(e) ensuring that digital maps are compliant with standards (such as file formats that ensure compatibility, portability, and exchange).

3.135. The NSO should undertake the EA design work and validation of the associated geospatial data, as well as take delivery of the hard copy maps for quality assurance checks and subsequent delivery into the field. The NSO must also accept full responsibility for the quality standards and delivery of the maps to field staff as required. After the census, any feedback received from enumerators about the base map should be communicated to the mapping agency.

3.136. In general, NSOs have progressively gained autonomy in census mapping, eliminating the need for full outsourcing of this service. The access to updated high-resolution satellite images, however, is an increasing benefit in the production of census mapping and in the generation of input for digital maps and should be considered in census planning from the beginning. The acquisition of handheld GNSS equipment for collecting coordinates and tracking the enumerator's route must also involve the participation of the geospatial team to guarantee the quality of the equipment and, consequently, the results.

3.137. The acquisition or incorporation of georeferenced databases such as an address register, electricity company databases, mobile phone databases, administrative records, or EO data can be very useful for the census geospatial information programme. In many countries, these databases are available in government agencies and can be incorporated through partnerships. In some countries, this data is commercialized by these government agencies or is privatized, requiring the expenditure of financial resources. The NSO should evaluate the need for this data and the most viable options within the national context.

3.138. There are good options for open-source and commercial GIS programmes. In general, it is recommended that NSOs should adopt open solutions, both due to the cost and the option of developing their own solutions, with a vast library in forums, which can be complemented by commercial solutions when necessary. This assessment should be made jointly by the census geospatial information team.

3.139. Mapping for dissemination purposes may be more challenging because the outputs will involve the representation of statistical information (with, or as part of, a map) and will often be accompanied by analysis or commentary about that information. Advances in mapping software have made it easier for NSOs to produce a wide variety of standard thematic maps (such as point maps, proportional symbol maps, choropleth maps, cartograms, etc.). However, advanced mapping products (such as geoportals and dashboards) may require more expertise. In these cases, the NSO may first seek support from the NGIA and the country's NSDI to find solutions. The United Nations Integrated Geospatial Information Framework (UN-IGIF) can be of great use in establishing this partnership between statistical and geospatial institutions. If such a partnership is unfeasible, the NSO should evaluate the contracting out of a service considering the quality of the geospatial

information technology, knowledge transfer, internalization capacity, customization, statistical disclosure control procedures and the confidentiality of the respondents' information.

V. Census tests

3.140. The success of a census depends substantially not only on the effort spent in designing and building the census operations and infrastructure, but also on the effort spent on testing and piloting all components of the census design. Thus, having systematic testing procedures in place is vital for the efficiency and effectiveness of the data collection and processing operations and for producing high quality data. In order to identify problems and to implement appropriate improvements, the use of pre-field tests (testing under laboratory conditions) and field tests (testing under field conditions), including a pilot census, is indispensable.

3.141. The testing of various aspects of a census plan prior to the enumeration is of critical importance for all countries, and an essential one for countries without a long history of census taking and for those in which fundamental changes in census methods or use of new technologies are being considered. Census tests can be designed for different purposes and in different ways. To yield full benefits, tests should be employed for all stages of the census, including enumeration, data processing, and methods for evaluating census results. Testing should be conducted in phases to build up to the full test of the holistic design, normally done with a pilot census. The testing phases should begin with testing individual components of the census design. For example, individual software components can be testing during a unit testing phase and the wording of questions for the census questionnaire may be tested in specialized content testing such as cognitive testing. The testing will build to eventually focus on the integration of all of the components together.

3.142. Testing of the individual components, such as specific technological solutions (for example, Internet data collection or the use of handheld devices), ensures that these solutions are working as intended and allows for efficient identification and resolution of errors specific to the individual solutions. However, the larger scale testing of the integration of all components is of utmost importance to ensure that all of the components work together successfully. Testing phases should be planned and conducted with both pre-field testing and field testing. The work to design all testing efforts, whether pre-field or field, should include the development of test objectives and measures. The test design should reflect those objectives and include the ability to collect the necessary data to demonstrate success, or failure, in meeting the predetermined objectives.

3.143. Much testing can be conducted without requiring field work – pre-field testing. Technological solutions can be tested using data from previous censuses or surveys in order to simulate data collection activities. If data processing methods are being tested, for example, those methods can be tested against existing data to determine if the processing generates the expected results. Using existing census or survey data for this purpose is especially helpful as the outcome is already known. In addition to cognitive research conducted in a laboratory setting with small numbers of participants, real-world observations can provide valuable insights into how respondents interact with, and react to, a questionnaire. Laboratory studies can also be used to examine how interviewers interpret and implement questionnaire instructions. Findings from these cognitive research events are particularly valuable in probing for weaknesses in the questionnaire, in understanding of definitions and concepts, in the instructions, or in enumeration procedures that might affect the quality of the data. Cognitive research should be designed to ensure that the subjects participating in the testing represent groups and populations of interest to ensure the results are relevant for all expected census respondents.

3.144. Field testing is an important stage to ensuring readiness of all components of the census design. Bringing all of the components together in field test creates the ability to identify problems linked to the integration of the data collection application design and architecture, the data transfer system, and the integrity and security of data transferred. Such tests can also yield important information on the adequacy of the field organization, training programme, extent of respondent burden, processing plan, budget, and other important aspects of the census. They can be designed to provide information on the relative efficacy of alternative methods of enumeration and technology, and on the average time required for enumerating a single household or a single set of living quarters. Such information is useful in estimating staff and cost requirements. In addition, census tests serve as practical training for the nuclear staff of supervisors and other officials. If resource limitations make it impossible to test all components of the census design, the testing that is feasible should be prioritized based on those components that present the most risk to the full census, such as major changes or innovations to the census design, new or emergent technologies, or design elements that have not performed well in previous censuses.

3.145. When carrying out census tests, probability samples of geographic areas or units are not always necessary. Since the purpose of the pilot census and pretests is to judge the operational feasibility of a proposed course of action for the main census rather than produce population estimates, purposive samples can often be used for such tests. Purposive selection of one or a few geographic areas is generally preferable for such feasibility testing. Purposive samples are also particularly useful when it is necessary to test particular census questions, concepts and methods in areas with particularly difficult conditions, or where population sub-groups or housing conditions of particular interest are predominant. On the other hand, random sampling procedures are often necessary, depending on the objectives of the test. For example, when overall quantitative measures are needed for comparing efficiencies of different procedures (for instance, in examining the anticipated response errors arising from different systems of enumeration), NSOs may need to design the test to ensure adequate representation of the different systems in the resulting data.

3.146. A critically important test to be carried out during census preparations is questionnaire testing. The purpose is to assess the suitability of intended census questions, including their formulation and the instructions provided, as well as the viability of the questionnaire design. Such tests can be particularly helpful in assessing the suitability of the proposed material for enumerating specific population groups, as well as the general public. For example, these tests can determine if the terminology used in a particular question is familiar to potential respondents, or if the sequence of questions is easy to follow for respondents of varying levels of literacy. Such tests are also used for estimating the time requirements for the enumeration. It is practical to carry out questionnaire tests on a small scale in a laboratory environment or in several purposively selected places, for example, a location that is known to be populated with particular demographic groups whose interaction with the questionnaire is particularly important. Because they are relatively inexpensive, repeated rounds of questionnaire tests may be carried out until a satisfactory questionnaire design has been achieved. Questionnaire testing can be conducted with a wide range of sophistication. Such testing can be as simple as conducting cognitive research in a laboratory environment where potential respondents interact with a paper questionnaire and share their thoughts and reactions to the content. Or, more sophisticated testing, especially with electronic data collection modes, can use leading-edge methods such as eye-tracking technology. The full range of methods is beneficial to design questionnaires more scientifically as reading patterns of respondents can be recognized technically through these types of tests.

3.147. One vital form of field testing is a pilot census that provides a final opportunity to bring together all aspects of the census design on a smaller scale to ensure the design is complete and can be scaled up in readiness for the full census. The results of the pilot census can be used to ensure

accurate coverage of the target populations and to identify any gaps in the design that must be addressed before the census itself. Such a pilot should ideally be designed and managed to thoroughly test all census components and the entire census infrastructure, and effectively serves as a dress rehearsal for the main event. Essential features of a pilot census are coverage of one or more sizeable administrative divisions and encompassment of the preparatory, enumeration and processing stages of a census, by which the NSO is able to assess the adequacy of the entire census plan and of the census organization. In order to best serve this purpose, care should be taken to ensure that conditions in the pilot census are as close to the conditions that would be present during the actual enumeration as possible. For this reason, it is often taken exactly one year before the planned census so as to simulate the expected seasonal patterns of climate and activity and to be conducted as close in time to the production census as possible. It is generally unwise to use the pilot census a source from which to derive usable substantive data. Apart from the sampling problems involved, such a use inevitably detracts from the central purpose of the pilot, which is to prepare for the main census. Nevertheless, a pilot census can be an ideal opportunity for testing the tabulation and analysis processes necessary to generate statistical data products.

3.148. The area(s) where the pilot census is conducted must not, of course, be excluded from the actual census itself. However, it should be anticipated that when the full census enumeration participants in the pilot census may believe that they already responded. So, appropriate publicity notifications should attempt to address this potential issue and include guidance to respondent to the pilot census that their involvement does not exempt their participation in the actual census itself in due course.

3.149. It is critically important to undertake a set of tests of the information and communication technology (ICT) solutions and production systems that are planned to be applied in the census (see paragraphs 2.265-2.266 on considerations for testing technology as part of the pilot census). Depending on the extent and characteristics of ICT, these tests should include all relevant components related to the fieldwork and to data transfer, entry and processing well ahead of the census itself. This is particularly important if a new technology is being introduced, such as the Internet, handheld devices or other electronic collection tools, and scanning the questionnaires as a means of capturing data. Tests should cover applications, systems and the equipment itself, as well as the underlying circumstances necessary to avoid equipment malfunctioning, such as climate, or significant delays due to inadequate quality of paper causing paper jams or unexpected problems in programming activities. In the context of new approaches using electronic data collection, testing should include data transfers, and the relevant frequency of those data transfers, to the major depository of data and validation of those data deliveries. Testing the efficiency of data entry rules, coding, editing and tabulation applications should be done based on results collected by the pilot census.

VI. Establishing the census frame

3.150. Prior to implementing a census, it is important to establish the frame or universe for the census. In most cases, this frame will consist of all living quarters and/or structures containing living quarters. The frame can be:

(a) constructed before the census using listing activities in the field;
(b) developed from existing address and/or person registers; or,
(c) built from available information such as aerial imagery and GIS data files.

3.151. Field listing activities generally require traveling all roads and neighbourhoods within the area to record the address or location of all living quarters and/or structures. This listing activity can be done with or without contact and interviewing of residents or representatives of the living

quarters/structures, but contact is recommended to ensure the most accurate information about the living quarters/structure is recorded. Field listing can be conducted using both traditional paper-based methods and advanced automated tools. While paper address listing tools and maps remain valuable in certain contexts, automated solutions can offer increased efficiency and accuracy.

3.152. For automated address listing activities, field workers/listers use mobile devices equipped with pre-loaded, georeferenced lists of addresses and locations within their assigned enumeration areas (EAs). Field workers/listers begin by identifying the boundaries of their EAs and then proceed to record information for every address or location within those boundaries. This means updating the list of addresses/locations that they contain. Changes in the list are recorded on the map of the EA that has been given provided for this purpose. However, field workers/listers may also be asked to update the addresses and locations in their EA by specifying their geographical coordinates using the mobile devices' GPS functionality.

3.153. The frame that is available at the start of the census is an instrument for the control of the enumeration, particularly in the absence of adequate and updated maps. The frame is also useful for:

(a) estimating the number of enumerators and the number of schedules and quantities of other census materials needed in an EA;
(b) estimating the time required for the enumeration;
(c) determining the boundaries of the EAs and for establishing the necessary links between population and housing units if the two censuses are carried out separately: and,
(d) for compiling provisional results of the census.

Finally, as the frame represents the full universe of living quarters or households that should be enumerated, it can be used as guide to monitor the completeness and quality of the enumeration in a given area.

3.154. Consideration should be given to the importance of permanent identifiers to streets and buildings, which can be used not only for successive censuses but also for other statistical purposes. A listing of sets of living quarters, particularly in densely populated places, cannot be made unless streets have names and buildings have unique numbers. Individual apartments in multi-dwelling buildings need to be numbered or otherwise unambiguously identified. Where these prerequisites do not exist, numbering immediately prior to the census would prove useful, but is likely to be outside of the authority of the NSO.

3.155. Having a complete and accurate address frame ensures the ability to provide the enumerators with assistance in the form of lists of addresses to visit. Having the full list enables the enumerators to recognize existing structures and identify new structures that may need to be added to the list and enumerated. Address lists will be essential if self-enumeration, where questionnaires are sent to the households by mail, is adopted as a method of data collection. Some countries have and maintain address or population registers that allow more or less complete address lists to be generated relatively simply. The census can then not only use these lists, but also assist in further improving the population register by reporting any discrepancies found in the field. Where official population registers are not available, or insufficiently complete, it may be possible to obtain additional address lists from postal authorities, utility companies, or the private sector (for example, mail order companies). A definitive list for the enumerators could then be prepared by the census office by merging and rationalising the lists obtained from these various sources.

3.156. Where a functioning population register exists, it may be possible to prefill the household questionnaires with information such as the names of the persons expected to be members of a household that are already recorded in the register (see paragraph 1.252. for example). This may reduce the response burden, accelerates the information-gathering process, and helps to pinpoint

deviations. On the other hand, it might have a negative psychological effect if respondents believed that the authorities were monitoring them too closely. To combat this negative effect, the census organization should develop proactive messaging and notification to the respondent universe about the intent to use register data and clearly articulate the benefits of such an approach. Using one or several registers as the point of departure for a census that still includes full coverage field enumeration is an approach applied in some countries; differences between the register(s) and the field situation will necessarily come to light, and rules will be required to deal with such differences. (See Section IV, C2 in Part One for a discussion on the advantages and challenges of such an approach.)

3.157. Any differences in the data from registers and the data from field data collection can be caused by one of several reasons. They could be due to a difference in reference dates where the register has not been updated for some time prior to the census, or where there are differences in definitions and concepts, such as when the register records only persons who are "legally registered" at an address but do not necessarily correspond to the persons who are actually members of the household at the time of the census. An understanding of the cause of any such difference is critical to making appropriate decisions on which source to trust and use, and for ensuring that the merging of the records data and the field-collected data leads to accurate and meaningful census results.

VII. Field enumeration

3.158. Recent years have witnessed significant advancements in technology that have transformed how countries conduct field enumeration. While face-to-face interviews remain a valuable method, there has been a growing trend towards self-enumeration, encompassing both online data collection via the Internet and completion of paper questionnaires. An increasing number of countries have successfully implemented electronic data collection during their censuses, utilizing tablets or portable devices equipped with electronic questionnaires. However, online data collection is often limited to countries with high levels of information technology penetration and is frequently used in conjunction with traditional field-based methods. While electronic data collection offers several advantages, it is important to recognize that it may not be suitable for all populations or regions. Face-to-face enumeration remains a crucial method, especially in areas with limited technological infrastructure or where there may be concerns about respondent literacy or internet connectivity.

A. Method of enumeration

3.159. There are two major methods of field enumeration:

(i) the face-to-face (or interview) method, whereby information for each individual (in a population census) and for each set of living quarters and the occupants thereof (in a housing census) is collected and entered in the questionnaire (either in paper form or electronically via a tablet) by an enumerator; and

(ii) the self-enumeration method, whereby the major responsibility for entering the information is given to a designated respondent in the unit being enumerated (usually the reference person of the household) and where the questionnaire is usually distributed, collected and checked by an enumerator, or mailed out and mailed back, or completed via the Internet.

3.160. The face-to-face method relied on paper questionnaires in most countries until recently. However, electronic collection via tablets has become the dominant method and the new norm in many developing countries. Paper questionnaires might still be needed in these countries for parts of the population or areas with limited electricity and Internet connectivity. For self-enumeration,

countries with up-to-date address/population registers and high literacy rates have traditionally utilized the approach of mailing out census forms and having the public mail them back completed. Recently, Internet data collection has become a common method for self-enumeration, especially in countries with widespread telecommunication services. Some countries use a combination of Internet data collection and postal distribution of the questionnaire (with or without return postage) for self-enumeration. Self-enumeration approaches, using methods such as mailing, telephone, and Internet data collection, can also be combined with face-to-face interviews. All these methods can be employed independently or combined with verification by a census official.

3.161. Whatever approach is used, the complete enumeration plan should be prepared and adequately tested and piloted well beforehand. This involves:

(a) the determination of the enumeration method(s) to be used, and whether or not there should be differences for specified geographic areas and/or population groups;

(b) the basic procedures to be followed in the collection of the data and the control of the enumeration;

(c) the procedures for quality control of the data; and

(d) an estimation of the number of sets of living quarters and the probable size of the population to be enumerated so that the number of questionnaires, handheld devices and other materials required for the enumeration, and the number of enumerators and supervisors needed, as well as the cost of enumeration can be properly ascertained.

If more than one enumeration method is offered, it should be decided whether the methods will be offered in phases or simultaneously. A phased approach may be used if the level of self-enumeration is expected to be significant.

3.162. Each method has its own advantages and limitations. The face-to-face method is the only method that can be used in largely illiterate populations or in other population groups that may be unwilling to complete the census forms themselves or find it difficult to do so. On the other hand, in countries where literacy is virtually universal and educational attainment relatively high, the self-enumeration method may often yield more reliable results at substantially lower costs, particularly if Internet data collection or a mail out/mail back procedure can be adopted. However, the postal services may be used to distribute the census forms only when a comprehensive and up-to-date list of addresses is available or can be prepared.

3.163. Another consideration is the emphasis to be placed in the census on obtaining responses, whenever possible, directly from the person concerned. The self-enumeration method allows for - and its instructions may encourage, at no extra cost to the NSO - consultations among family members when they complete the census form. In contrast, with the face-to-face method it may be prohibitively time-consuming to encourage enumerators to attempt to interview persons in addition to the designated responsible adult[99] in order to collect the requisite information on the other household members. In the light of these considerations, it may sometimes be desirable to rely on one method for enumerating most of the population and to use another method in certain areas or for special groups of the population.

3.164. With the advance of IT, the penetration of the Internet has increased in recent decades. In these circumstances, it is recommended that Internet collection methodology should be explored

[99] In some countries the census legislation prescribes the person who is to be legally responsible for providing the information.

depending on national circumstances. This method can be cost-effective as the expense of printing questionnaires and the field staff wage bill can be greatly reduced. Also, self-enumeration through the Internet can help secure the privacy of respondents. However, a combination of face-to-face and Internet modes can result in duplication during the enumeration. Therefore, careful consideration of the management of the dwelling/household list is essential. Overly complex designs should be avoided, and adequate quality checks introduced to avoid duplications and frauds.

3.165. The decision regarding the method of enumeration to be employed should be taken at an early stage on the basis of thorough testing of the various alternatives in terms of the costs involved, the quality of the data produced and their operational feasibility. Even where a method has been followed traditionally, it is best to periodically reassess its relative advantages in the light of current census needs, changing techniques, and emerging technologies. An early decision is required because the method of enumeration used affects the budget, the organizational structure, the publicity plan, the training programme, the design of the questionnaire and, to some extent, the range of data that can be collected.

3.166. Challenges that affect or hinder the ability to achieve a response should be fully considered when developing the enumeration design and methods. The design should, as far as possible, reflect particular activities or actions that seek to increase the likelihood of receiving a response. Challenges can be grouped into two types: people and physical challenges. People challenges tend to focus on particular subgroups of the population that tend to be difficult to enumerate for a number of reasons. Physical challenges are those related to the type of environment in which the people live. The two are not necessarily mutually exclusive, but each require different methods to overcome the difficulties that they present.

3.167. The section on population sub-groups that are difficult to enumerate (see paragraph 5.46) gives a thorough overview of the more common examples and how such difficulties might be addressed during the enumeration. Population groups that are typically difficult to enumerate include people with language difficulties and Indigenous Peoples, nomads, migrants, students and older persons.

3.168. ***People with language difficulties and Indigenous Peoples*.** Not all respondents will speak or understand the language(s) in which the census is being conducted. Therefore, consideration needs to be given to the provision of translation services and materials, taking into account the types and concentrations of languages required. To the extent possible, decisions around languages/translations to support the enumeration should be data driven to avoid any perception of bias or discrimination.

3.169. To facilitate the participation of **Indigenous Peoples**, it is essential to involve members of their communities in the enumeration process. This can be achieved by: employing enumerators from within Indigenous communities; providing training and support to enumerators working with Indigenous populations; and, engaging translators or community leaders to facilitate communication and ensure cultural sensitivity. Consideration must be given to the different socio-territorial contexts to identify monolingual and multilingual areas and address the needs for translation, enumerator training, and supervision. Particular attention should be paid to areas with limited access where

Indigenous peoples often reside, to ensure their inclusion in the census. However, it is crucial to respect the rights of Indigenous peoples in voluntary isolation or initial contact. [100]

3.170. ***Nomads.*** To successfully carry out the enumeration of nomads, it is particularly necessary to pay full attention to the preparatory work in order to determine suitable enumeration techniques. It should be pointed out that there is no absolute recommended methodology for the enumeration of nomads since conditions will vary from country to country. The particular method suitable for a country undertaking to enumerate nomads as part of the census should be determined only after a detailed preliminary study and after undertaking field tests.

3.171. Some of the methods used to enumerate nomads and semi-nomads may be classified as follows:
(a) In the ***group assembly approach*** nomads are asked to assemble at particular interview sites on certain fixed dates. This method can be adopted only through the administrative or tribal authorities.
(b) The ***tribal or hierarchical approach*** is a favourite method, since the nomads usually follow what is dictated by the tribal or hierarchical chief. The enumeration can be carried out as a kind of administrative census by contacting the tribal chief and collecting, sometimes from memory and sometimes from a register, all the needed information on the chief's followers. The other approach is to contact those followers with the assistance of the chief or a representative and to collect the necessary data directly from the household. In this case, the unit of enumeration is not areal but tribal.
(c) The ***enumeration area approach*** presupposes creating conventional census EAs and then contacting each nomadic household that happens to be staying in a particular EA during the census.
(d) In the ***water point approach***, a list of all water points available to the nomads during the period of enumeration is prepared. Since numerous temporary water points are created during the rainy season, a meaningful list of water points may be prepared with reference only to the dry season. The enumerator is given the task of locating and visiting every nomadic household that may be using a certain water point.
(e) In the ***camp approach,*** a list of camps is prepared together with the approximate location of each within the country, and enumerators are sent to visit all the households in each camp.
Sometimes a combination of two or more of these methods may be appropriate.

3.172. ***Migrants and persons with an irregular migration situation.*** Effective communication is essential for engaging all immigrants in the census, including recent arrivals, those who have entered through irregular channels, and those who may have overstayed their visas. This can be done by providing materials and assistance in multiple languages, working with community organizations that work with migrants, and addressing concerns about immigration status. As part of developing the enumeration design consideration needs to be given to communicating with these groups, particularly about the benefits of the census, to encourage their participation. For example, materials can be translated into the languages that are most commonly spoken by migrants in the area. Community organizations can help to spread the word about the benefits of the census and enumerators can be trained to answer questions about immigration status and to assure migrants that their data will be kept confidential. Flexible enumeration methods, such as online

[100] Comisión Económica para América Latina y el Caribe (CEPAL), "Recomendaciones para los censos de población y vivienda en América Latina. Revisión 2020", Documentos de Proyectos (LC/TS.2021/150), Santiago, 2021.

questionnaires, phone interviews, and in-person visits, can also be used to accommodate the diverse needs of migrants. And ensuring that enumeration venues are safe and comfortable will further contribute to a positive experience for migrants.

3.173. ***Refugees and internally displaced persons (IDPs) in camps***. In many censuses several questionnaires are utilized, including one for conventional/private households and another for institutional or collective living quarters. These institutions often comprise camps for IDPs or refugees. However, institutional questionnaires may not be suitable for camps for the displaced, particularly where these facilitate more than short-term arrangements. As the question content is often much shorter in the institutional questionnaire than in the household questionnaire, the information collected will be insufficient to provide for the requirements of users interested in displaced people. Wherever possible the camps for refugees and IDPs should be enumerated using the questionnaires designed for private households. The census communication plan should include messages to convey that the census will include displaced people, refugees and IDPs and that their responses are important, and to assure the confidentiality of data collected. It is important that such communication is in the languages relevant to displaced populations, as refugees and IDPs may speak a different language than the local population. The Compiler's Manual[101], developed by Expert Group on Refugee, IDP and Statelessness Statistics (EGRISS) primarily aimed at technical personnel in NSOs who want to include forcibly displaced populations in official statistics, provides additional insights. The *Use Case A* from the *Compiler's Manual on Forced Displacement Statistics* describes how displaced populations can be included in a population census.

3.174. ***Students*.** Tailored approaches are needed to reach students, such as collaborating closely with schools, colleges, and universities. This can help to ensure that students are aware of their legal obligation to participate in the census and that they have the opportunity to do so. Clear and concise communication materials that resonate with student life and their priorities are essential. For example, materials could be developed that focus on the benefits of the census for students, such as how census data can help to improve education funding and opportunities. Appointing student ambassadors can also be an effective way to promote the census within student peer groups. Leveraging online platforms, social media, and mobile applications is crucial for engaging technically aware students, who are more likely to respond through digital channels. For example, the census could create a dedicated website or social media page for students. And students could be encouraged to share information about the census with their friends and classmates through social media.

3.175. ***Older persons***. Additional support and tailored strategies are essential for engaging older persons in the census. Effective engagement includes providing supplementary materials and support mechanisms to assist with the census process (for instance, a national events calendar may be provided to assist older persons to recall or estimate their age). It can also involve supporters, such as family members, village elders, or residential home staff, in the enumeration process. For example, family members could help to translate census materials or to answer questions from older relatives. Enumeration venues, where used, should be accessible to all, including those with mobility

[101] Expert Group on Refugee, Internally Displaced Persons and Statelessness Statistics (EGRISS) (2023). Compilers' Manual on Forced Displacement Statistics. https://egrisstats.org/wp-content/uploads/2023-EGRISS_CM_Final-web.pdf

limitations. They should also provide a comfortable environment for older participants. Assistance could be provided through call centres as well. [102, 103]

3.176. During the development of field enumeration methods, several physical challenges need to be considered. These include:

- access-controlled properties such as gated communities, apartment buildings with secured entrances, or private residences with restricted access which can pose challenges for enumerators trying to reach residents;
- rural areas that are sparsely populated with limited infrastructure, long distances between residences, and difficult terrain can complicate the logistics of enumeration;
- communal institutions such as prisons, hospitals, or military bases that often have specific protocols and access restrictions that require special arrangements for enumerating residents; multi-occupied dwellings can make it challenging to ensure all residents are counted, especially if there's no easy way to identify separate living units; and
- areas that are insecure/violent which may pose a risk to enumerators.

3.177. ***People living in access-controlled properties*.** Some properties, establishments, communities or compounds have controlled access, presenting difficulties in gaining access to undertake an interview or to deliver or follow-up on a questionnaire. Access control mechanisms may include locked gates with an intercom to each individual dwelling, commercial farms where the gates are unmanned and are some distance from the dwelling units or gates or doors managed by a concierge or security guards. When developing enumeration procedures, advice needs to be given as to how to gain access and actions to take if access proves difficult. Some of the activities may include building a relationship with the owner of the properties to approve access to engage with residents; using the postal service to deliver questionnaires to these properties; and additional communication methods (such as a letter informing residents about the census and how to complete their questionnaire or inviting them to arrange a particular time to complete their questionnaire via interview. The use of mail out/mail back or online questionnaires that do not require the intervention of enumerators can be employed to resolve this problem.

3.178. ***People in remote rural areas*.** Understanding the extent of population in remote rural areas and the associated logistical and management challenges of enumerating them needs careful consideration. Lack of infrastructure, such as access to electricity and Internet connectivity, and transport facilities, may limit the modes of data collection that could be used in these areas.

3.179. The enumeration of other groups, such as the ***homeless*** and ***institutional populations***, also requires special attention and tailored approaches. Due to their transient nature and lack of fixed addresses, traditional enumeration methods might be ineffective for the homeless. Solutions include collaboration with homeless shelters, soup kitchens, and outreach programmes to establish trusted points of contact; and deploying enumerators with experience and sensitivity to engage with this

102 Comisión Económica para América Latina y el Caribe (CEPAL), “Recomendaciones para los censos de población y vivienda en América Latina. Revisión 2020”, Documentos de Proyectos (LC/TS.2021/150), Santiago, 2021.
103 Comisión Económica para América Latina y el Caribe (CEPAL), “Recomendaciones para los censos de población y vivienda en América Latina. Revisión 2020”, Documentos de Proyectos (LC/TS.2021/150), Santiago, 2021.

population. Solutions for institutionalized populations might include establishing partnerships with institutional administrators to facilitate data collection; utilizing administrative records from these institutions with proper consent and privacy safeguards; and, deploying trained enumerators to conduct interviews within the institutions, following established protocols. These targeted approaches can help achieve a more complete and accurate count.

B. Multi-mode data collection

3.180. Countries are becoming more innovative in how they conduct population and housing censuses in order to improve response rates while also reducing census costs. One of these innovations is the exploration of alternative approaches to the traditional method of face-to-face canvassing of the whole country to enumerate the total population. Such alternative approaches can incorporate the use of two or more data collection methods in combinations of (i) interviewer-administered modes, such as face-to-face interviews or interviews over the telephone, and (ii) self-administered modes, whether online or by mail.

3.181. Interviewer-based data collection modes include: the traditional face-to-to-face personal interview with a paper questionnaire (PAPI), computer assisted personal interviewing (CAPI), and computer assisted telephone interviewing (CATI). On the other hand, self-administered data collection modes include paper questionnaire with self-enumeration (PASI) and computer assisted (online) self-interviewing (CAWI). For more detailed information on these modes of data collections, see Chapter E of the *Guidelines on the use of electronic data collection technologies in population and housing censuses*.[104]

3.182. Application of multi-mode data collection entails the uses of two or more of the modes of data collection listed above, either concurrently or sequentially. In concurrent multi-mode data collection, respondents can choose between one of two or more modes, such as completing a paper questionnaire or providing the information online via the Internet. In sequential mixed-mode data collection, all respondents are first requested to provide information in one mode but then, should they not choose to do so, are offered other modes to increase the response rate. Sequential mixed-mode designs may start with the mode that is least costly to implement and progress to more persuasive modes even if these are most costly. Both options aim to maximise coverage/response rates and completion rates, both for the general population and for special groups. It should be noted, however, that while mixed mode data collection methods are effective at reducing non-response, they could lead to measurement error due to response bias arising from the use of the various modes (see paragraph 3.184).

3.183. Multi-mode data collection offers several benefits over the traditional single-mode. Firstly, offering different modes of data collection, to different populations, representing a broader cross-section of the population, has the potential to increase coverage/response rate and improve data quality. For example, some individuals may prefer to complete the census questionnaire online, while others may prefer a telephone call or a paper-based questionnaire. Providing various modes of data collection will enhance the ability to reach respondents who are difficult to enumerate with the use of a single mode, such as persons living in remote areas or those with no Internet access. Secondly, multi-mode data collection can help to reduce field costs, especially when self-enumeration methods are utilized to a significant degree. Self-enumeration, whether conducted

[104] https://unstats.un.org/unsd/demographic/standmeth/handbooks/data-collection-census-201901.pdf

online or using paper questionnaires, can minimize the need for face-to-face interviews, thereby reducing the time and resources required for field operations. Additionally, automating tasks such as data entry through electronic data collection can further contribute to cost reduction.

1. Management of multi-mode data collection operations

3.184. There are some potential challenges to the use of multi-mode data collection. It is more complex to implement than it is for single-mode data collection. This is because multi-mode collection requires more intricate planning and coordination than a single mode. Also, once data has been collected from multiple modes, it needs to be integrated into a single dataset, and this can be challenging, as different modes may produce data that are not be directly comparable. When integrating census data from different collection modes, it is important to be aware of the potential for bias. This can be due to a number of factors, such as differences in response rates by mode of data collection, and differences in the wording or routing of questions. As a result, comparability of data from multiple modes of data collection may be a challenge due to "mode effect"[105] and has implications for the quality of data that has been collected. Understanding any such mode effect biases on coverage, response, and measurement can help develop methods to minimize them.

3.185. Multi-mode data collection often relies on technology, such as the Internet and mobile devices, and this can be a challenge in areas with limited access to such technology. Use of multiple modes of data collection can be more expensive than single-mode data collection, especially if new technologies are used. It is worth noting, also that not everyone has access to the Internet or other technologies needed to participate in online data collection. This can lead to under-coverage of certain groups. Despite these challenges, multi-mode data collection is becoming increasingly common because it offers several advantages over a single-mode approach, and offers the possibility to compensate the weaknesses of one mode with the strengths of another. By carefully planning and implementing multi-mode data collection, NSOs can improve the accuracy and completeness of collected data.

3.186. However, the management of multi-mode data collection operations is a difficult task due to the complexity of designing questionnaires for interviews and for self-completion, and of monitoring the enumeration, as well as the validation and integration of data collected by different modes. The more options that have been employed for data collection, the more complex the operation becomes. Consequently, it is particularly important that an integrated management information system (MIS) should be developed for executing, monitoring, and reporting functions related to operational procedures, including scheduling, contact management, case status management, avoiding duplication of responses, managing switching/transfer of a case between data collection modes, response management, assignment of field staff, and response integration, requiring a dedicated management team to monitor and control the multi-mode collection operation.

2. Designing a census multi-mode data collection

3.187. A crucial part of the implementation of a multi-mode census data collection relates to decisions on the modes of data collection and how they are to be implemented. It is the case that

[105] See Chapter E of the Guidelines on the use of electronic data collection technologies in population and housing censuses. https://unstats.un.org/unsd/demographic/standmeth/handbooks/guideline-edct-census-v1.pdf

Such decisions should be based on national circumstances given the diversity of national experiences and capacities to successfully implement the various options. Hence, there is no “one-size-fits-all” design that can be recommended for international use.

3.188. However, regardless of the modes included in the combination for a multi-mode census data collection, there are pertinent factors that should be considered in its design. First is the need to have a clear understanding of the target population as well as its characteristics, and how easily the modes under consideration can be implemented. In addition to, it is also the case that even within the same country there may be variations in ability based on the levels of literacy and education of different population groups as well as differences by geographic area of residence of the population. For example, self-enumeration by mail or by use of the Internet may not be appropriate where a large proportion of the population is not literate, or are living in remote areas where access to the Internet or to postal services may be limited or even non-existent.

3.189. In order to make informed decisions about the modes of data collection to include in a census, it is essential to have an understanding of the characteristics of the national population, based on data from previous censuses, surveys, or other relevant sources. This information should include details on age, literacy levels, geographic distribution, and other key characteristics. This would enable the selection, development, and tailoring of multiple modes to include only those that can easily be implemented in the country. A further step would be to develop a mode assignment strategy, or for modes to be concurrently assigned to different sub-groups based on well establish criteria on the ability of the corresponding population groups to easily and accurately provide the information that is being requested. In a sequential design, the NSO needs a strategy, not only on the various modes of data collection to be adopted, but also on the order that the different modes will be offered in order to maximise the convenience for the population to take part in the census.

3.190. Another issue to which attention should be given in designing a multi-mode data collection is ensuring that the questionnaires are designed in a compatible way across the modes of data collection to be used. For example, question wording and format, response options, as well as the layout and design of the appropriate questionnaires should be consistent and comparable across modes to be used. This would enhance the comparability of the collected data across the various modes by minimising response bias.

3.191. It is, therefore, important to test and pilot the various modes of data collection process in order to identify any problems and make adjustments to the questionnaires and the data collection process. For example, the pilot census would help to ascertain the compatibility the collected data, as well as ~~to~~ assess whether or not the questionnaires in the different modes are easy to understand and complete.

C. Use of administrative data during field-based enumeration

3.192. The use of administrative data can facilitate the enumeration in a field-based census in a number of ways. As noted at paragraphs 1.252. and 3.156 for example, pre-filling electronic census questionnaires is one of the ways in which administrative data could improve enumeration. Such an approach can reduce the workload of enumerators (by, for example, reducing data entry for basic demographic information). Leveraging existing administrative data can potentially reduce respondent burden. When data items are taken from administrative sources to pre-fill the questionnaire, the data should be checked or validated by the respondent. The NSO should establish clear protocols on which information to accept in instances when there are inconsistencies between the prefilled administrative data and the respondent’s responses. However, this approach may have negative effect on respondents that have privacy concerns.

3.193. In certain cases, necessary information for particular sub-populations, such as the members of institutional establishments (such as students or workers dormitories, old people's homes, assisted-living facilities and welfare institutions, military barracks, correctional and penal institutions, religious institutions) may be available from the administrative data sources. Using these data has the potential to effectively mitigate or substantially decrease the extent and complexity of field operations, especially when direct data collection for a sub-population is problematic.

3.194. There are also cases where the availability of administrative data assists in facilitating the field operation in other ways. For instance, as noted earlier, sometimes prior to the field operation, households that are likely to be missed by enumerators (such as the ones that have recently relocated, or comprise older persons or migrants, or that are located in remote or hard-to-reach areas, or within indigenous communities) can be identified in order to devise appropriate strategies to cover them. Additionally, administrative data may be used to determine whether an address is occupied or not, which will assist in planning follow-up visits to non-respondents.

3.195. Furthermore, when administrative data sources provide an accurate aggregate estimate of the number of housing units/households/individuals in a given area, this information can be used as benchmark to monitor and quality assure actual performance in the enumeration.

3.196. In the combined census approach (see Section IV.C in Part One), data for specific high-quality variables can be directly obtained from administrative sources, eliminating the need to collect them from individuals during the field enumeration. This is particularly beneficial for: sensitive topics for which administrative data can provide information without requiring direct questioning, such as infant mortality and income thereby reducing respondent burden and potential disclosure risks.

D. Timing and length of the enumeration period

3.197. The choice of the time of year in which the census is to be taken is of great importance. The main consideration should, clearly, be to select a period in which the census is likely to be most successful and yield the most useful data. This may depend on a number of factors. Firstly, it is necessary to avoid those seasons in which it will be difficult to reach all inhabited areas because of rains, flooding, snow and other inclement conditions in which the work will be particularly arduous, as is also the case during extremely hot weather. Secondly, a time should be chosen when most people are staying at their place of usual residence; such a choice will simplify the census operations both in a *de jure* and in a *de facto* enumeration, and it can make the results of a *de facto* enumeration more meaningful. Seasons of peak agricultural activity should be avoided because it is difficult to interview persons who work late every day and who may even stay nights on their land if that is distant from home. Traditional secular and religious festivals, pilgrimages and fasting periods are also unsuitable times for census field work. Since in many developing countries the bulk of the field staff is recruited among schoolteachers and older students, the conduct of the census may be feasible only during school vacation periods, though, as already indicated, the days of major festivals should be avoided. On the other hand, enumerating tertiary students in vacation time is problematic as many will not be present at their place of usual residence (their term-time address) and are likely to be undercounted. Another criterion to consider in the timing of the field work is to avoid the period when hours of daylight are at their minimum, as interviews should, preferably, not be conducted at night.

3.198. In a country that includes areas of sharply contrasting seasonal patterns of weather or activity or in which potential census personnel are in very short supply, it may be necessary to

enumerate different parts of the country at different times or to enumerate the nomads or other special population groups at a different time from that established for the settled population. This, however, is generally not a very desirable solution because nomads cannot always be clearly differentiated from settled population and there may also be mobility among the settled inhabitants. Furthermore, such a solution erodes one of the key elements of the census, that of simultaneity. The census design should consider these complications and risks, and attempt minimize negative impacts on, or biases in, the data.

3.199. When a census has been undertaken and the census date is found to have been, on the whole, satisfactory, subsequent censuses should be taken using the same time of the year, unless there are strong overriding reasons for not doing so. A regular census date enhances the comparability of the data over time. The tradition of a fixed census period within the calendar also provides administrative discipline, motivating and enabling all those involved in the census to make necessary preparations in a timely manner.

3.200. It is desirable to keep the census enumeration period as short as possible in order to avoid double counting and omissions, which can occur in spite of applying a single reference date. On the other hand, the shorter the enumeration period, the greater the number of field staff that have to be recruited, trained and supervised. This increases the cost and may lower the quality of the data. If the census is using automated methods for enumerator interviews, the costs are further increased by the need for additional devices for the additional field staff. How these different considerations should be reconciled depends on the size and nature of the country and on the resources at the NSO's disposal. The length of school holidays is sometimes a restricting factor, although governments of several developing countries, recognizing the national importance of a census, have prolonged the school holidays in the census year in order to allow teachers and students to work on the census for as long as required.

3.201. While it is desirable to design the census enumeration period to be of optimal length and to conduct the enumeration at the optimal time of the year, the design must also include flexibility and contingency measures to enable adjustment in the case of an unexpected emergency or disaster. If an unexpected event makes it impossible or impractical to conduct the enumeration period as planned, an alternative timing or length must be implemented. In this case, it is imperative to attempt to understand and quantify the possible impacts on data quality and the comparability of the census data over time. Lengthening the enumeration period introduces the risk of recall bias, duplicate enumerations, and/or omissions. Changing the timing of the enumeration period can increase the risk of omissions if some populations (such as migratory workers) are no longer present. It should be additionally noted that any change to the census date will usually require an amendment to the relevant legislation, which itself can be a lengthy process.

3.202. In recent censuses, most developing countries have allowed about one week to 10 days for the training of enumerators, while the enumeration period has generally varied from a few days to two weeks. In the case of using a self-response method through the Internet, the enumeration period can take longer, and enumerators may require less training if they are simply going to follow up on non-respondents. Short periods are often feasible in small countries while longer periods may be necessary in large countries with poor communications, widely dispersed populations, or where multi-mode data collection strategy is adopted.

3.203. One method sometimes used to allow sufficient time for enumeration and yet make the census simultaneous is first to enumerate the population over a longer period, say a week or more, and then, on one single day, to revisit all households, deleting and adding persons as needed to update the census return. This procedure is, however, not practicable in very sparsely settled areas

and can be very costly to deploy a large number of staff on the single re-enumeration day. It is important to note that utilizing a post-enumeration survey (see paragraphs 24.119-24.126) can eliminate the need for this second visit for the purpose of assessing coverage errors.

3.204. While simultaneous data collection is generally preferred in a census, exceptions may be necessary in countries with diverse geographical or cultural contexts. These exceptions require careful planning and execution to minimize potential biases from variations in data collection timing. Additionally, the increasing practice of allowing census responses to be submitted before the official census reference date, often through self-enumeration options, introduces new challenges. These challenges include changes in household composition or individual circumstances between the time of response and the census reference date, which could affect data accuracy. Clear communication is crucial, ensuring respondents provide information as of the census reference date, even if they submit their responses earlier. Addressing these challenges through careful planning and communication helps maintain data quality and minimize potential biases associated with non-simultaneous data collection.

E. Management and supervision

3.205. A comprehensive and elaborate management system is necessary for resource management and providing timely instruction and advice to the field staff. There is no one universal model for such a system; however, it is important to develop a hierarchical and geographically dispersed system to build a direct and effective communication mechanism between two key groups: field staff (this includes enumerators and census supervisors) and central management (this refers to the permanent managers within the NSO). The system should facilitate a smooth flow of information and support.

3.206. Adequate supervision of the enumeration is essential for ensuring the quality of the data collected. Many countries use a field supervisor to enumerator ratio ranging from 1:10 to 1:15. Frequent, if not continuous, assessment of the quantity and quality of the work accomplished by enumerators and other field staff is recommended, in order to facilitate the correction of inefficiencies and to maintain satisfactory progress during the enumeration period. Such assessment should be carefully organized to ensure the quality of the work and to gather appropriate information about the the progress of the fieldwork.

3.207. Each staff member involved in the management and supervision system should have a clear role and responsibility and should be fully trained to deal with, and resolve, possible problems occurring during field enumeration. For an efficient system, it is important to give clear instructions to the field staff, as part of their training, on how to carry out their own duties and responsibilities. It is important to note that methods and technologies used during the enumeration have a direct impact on the roles of managers and supervisors; consequently, a complete understanding of the characteristics and operational aspects of both the enumeration method and technology is a prerequisite for an efficient supervision of the enumeration phase.

3.208. Depending on the communication facilities and other infrastructure available in a country, different mechanisms for exchanging information among managers and field staff need to be developed. These mechanisms are important for ensuring consistent dispatch of field instructions and also sharing best practices, particularly for finding solutions to unexpected problems. The use of mobile phones and accompanying technologies, such as Short Message Service (SMS), Unstructured Supplementary Service Data (USSD) and other enterprise level communication platforms significantly increases communication capabilities, more especially when such communication needs to broadcast simultaneously to allow all field staff to get the same message from the source.

a) Management information system

3.209. A robust management information system (MIS) is essential for optimizing the efficiency and effectiveness of census field operations. By providing real-time progress information, data, and analytics to central offices, regional managers, and field staff, an MIS streamlines administrative tasks, facilitates supervision, and enables timely identification and resolution of operational challenges. This, in turn, can improve data quality, increase productivity, and ensure the successful completion of the census. Furthermore, integrating communication tools within the MIS allows for seamless information sharing and coordination between all levels of field staff and management, further enhancing operational efficiency.

3.210. An MIS should be established to collect information needed for timely management and supervision of field operations. To establish this system, the following steps can be considered:

i. Determining information needed for supervising and managing fieldwork;
ii. How and when each piece of information/data should be collected;
iii. How and by whom each piece of information/data will be used.

It is important to collect the amount of information that can reasonably be assimilated with good quality and used effectively, otherwise every additional datum with low priority will affect the cost of collecting the requisite information.

3.211. The advent of real-time data analysis in censuses has resulted from the technological advancements and practices that enable the collection, processing, and dissemination of information on field activities in real time as they happen, or at least as soon as they are recorded. This concept has gained significant importance of late due to its potential to provide immediate insights, enhance decision-making, and improve overall efficiency.

3.212. The following information can be collected through such a system:

i. Information about particular activities that are implemented before the enumeration, such as establishment of local census commissions and training of census field staff;
ii. Information about the field staff needed for administrative tasks, such as recruitment and hiring field staff, bank account information for payment, work accomplished, dynamic fieldworker payment advices especially when utilising electronic data collection methods;
iii. Progress of the enumeration of population and housing units to evaluate whether or not the field operation proceeds according to schedule;
iv. Information about logistics issues, such as: the shipment, storage and maintenance of questionnaires and other census materials; the quantities and types of such materials required; and the timing of receiving and despatch of such materials.

3.213. Paradata - which is data automatically collected about the census field process captured during computer assisted data collection - can become a vital source of information for the MIS. Examples of these include interviewer observations, number of keystrokes, data time stamps to record responses and various other data captured during the process. This information is increasingly being used in real time for both monitoring purposes and to better manage the process of collecting large amounts of data

b) Supervising the enumeration

3.214. A robust supervision system is a critical component of the MIS described at (a) above, enabling real-time monitoring of field operations and facilitating the correction of errors and omissions. This system can also be used to make any necessary adjustments during the fieldwork. In

countries that utilize technology for data collection, such as the Internet or handheld devices, a computerized online system can be developed, incorporating automated procedures for supervision.

3.215. The key to rapid quality control of the enumeration is the fast flow of information from the supervisors in the field upwards to the managers in the local or regional statistical offices and/or to the central census office. The most efficient way of exchanging this information is via the Internet. If field supervisors have Internet access, information can even be submitted through a password-protected database interface (a web-based application).

3.216. Close monitoring during the enumeration phase is essential to ensure coverage, quality and compliance with deadlines. It must be ensured that all staff involved in the data collection have access to up-to-date reports with relevant information. These reports should be made available regularly in digital form.

3.217. For successful monitoring of the field enumeration, actual performances should be evaluated against the set targets. The following inputs are recommended to achieve the desired outcomes.

3.218. ***Using historical data***. Data from previous censuses and other relevant data sources such as household surveys and administrative registers can be used as a benchmark to determine the data needed for monitoring the performance of enumerators against the set goals and targets. Examples of historical information include total population of previous censuses; population growth rate; sex ratio; proportion of the population in urban areas; proportion of vacant dwellings; and occasional use dwellings in relation to those occupied.

3.219. ***Setting goals and targets***. As a population census is a time-bound project, any extension can be considered as failure. Setting goals and targets will be very important for measuring whether the series of activities is under control or not. Goals and targets for measuring the quality of enumeration and for systematic monitoring of the enumeration can be set based on the experience of previous censuses and other relevant data collection processes. The following indicators can be used for monitoring the enumeration:

(a) the proportion of occupied and vacant dwellings;
(b) the average number of residents per dwelling;
(c) response and refusal rates;
(d) population size; and
(e) population growth rate.

Significant deviation between the target values and the enumerated values may indicate a problem in the collection process. Estimation of the number of housing units and population size – if available – based on census maps and the listing of living quarters and households can also be used as information for monitoring the enumeration.

3.220. ***Preparing policies and procedures.*** Preparing policies and procedures. Policies and procedures for monitoring census enumeration should be defined early in the planning process, with endorsement from the highest levels of management to ensure proper support. To inform this process, it is essential to review the evaluation of the previous census field operation, including lessons learned and recommendations for improvement. This evaluation should be conducted immediately after the enumeration of the previous census and documented in a comprehensive report. By reviewing the successes and challenges of past censuses, NSOs can refine control procedures and improve the execution of the current enumeration.

F. Security during data collection

3.221. The census is a monumental undertaking which collects sensitive personal information from the population. To maintain public trust, it is important to ensure the data collected is not only accurate but also secure and protected from unauthorized access and misuse. Census data should be safeguarded throughout the entire lifecycle of the census, but particularly during the field operation.

3.222. In the *pre-enumeration phase*, plans need to be drawn up for assuring the security and confidentiality of the information to be collected. These safeguards should aim to protect the privacy of individuals and prevent unauthorized access to, or disclosure of, their personal information. Information security measures include:

(a) robust data encryption methods to secure data both during storage and transmission;
(b) strict access controls and authentication protocols to limit access only to authorized personnel;
(c) regular security audits and assessments to identify vulnerabilities; and,
(d) comprehensive training programmes to educate field staff on the importance of upholding data privacy and security practices, and their legally prescribed responsibilities.

In parallel, it is essential to establish strong physical security measures, such as limiting access to census facilities and using secure data storage systems to safeguard the physical infrastructure and prevent unauthorized tampering or theft of data.

3.223. During the *enumeration phase*, various information security safeguards should be implemented to ensure the protection and confidentiality of the collected information. First and foremost, robust encryption protocols should be used to secure data transmission between the data collection devices and the central database. Secondly, data accessibility should be limited to census officials. Strict confidentiality agreements should be upheld by all personnel involved in data collection, with severe penalties for any unauthorized disclosure.

3.224. Census data collection with electronic questionnaires (whether with handheld/mobile devices or via the Internet) should be subjected to a thorough security review to confirm that the overall design of the system is robust to shocks (connectivity disruptions, power failures, etc) and security vulnerabilities (unauthorized access to data). The review process must concentrate on as many risks as it is possible to formulate in advance and include the common risk points: data at rest on a mobile device; data in transit to the intermediate or central location (via networks or via physical media, or both depending on the system); and data at store in the central location.

G. Technologies for field enumeration

3.225. Technology, namely the use of computer-assisted methods for census planning, organization, data collection and data processing, has become indispensable for improving the cost-benefit, quality and efficiency of the population and housing census. Rapidly expanding worldwide mobile and Internet connectivity coupled with advancements in areas such as cloud computing, smart mobile devices, GPS, GIS, natural language processing and AI provide new opportunities for further increasing the quality and speed of census data collection. Furthermore, the disruptive effects of the COVID-19 pandemic precipitated more widespread adoption of online tools and methods, as social distancing measures necessitated a shift towards remote solutions for data collection. The technological tools and instruments described below are well documented in national practices in conducting the population and housing censuses in the most recent census rounds.

a) Electronic questionnaire

3.226. Combined with, or completely replacing, the paper questionnaire, an electronic questionnaire can be used in either the face-to-face or self-enumeration method. Electronic forms can provide improved data quality and operational efficiencies by:

- implementing validation rules on individual questions;
- cross-validation between questions or with other records;
- automatic sequencing (leading the operator to the next appropriate question);
- offering more options in pull-down lists;
- capturing more detailed data;
- providing computer-assisted coding; and
- facilitating the ability to ask tailored supplementary questions.

Electronic questionnaires can give access to guidelines, illustrations, help material and even videos to provide instruction to the interviewer or household. Electronic questionnaires can also be pre-filled with geocodes, which enumerators can validate against GPS information collected from the field and update, if necessary. Pre-filled administrative data could be validated by respondents.

3.227. The use of electronic questionnaires provides a potential opportunity for audit and quality control: in particular, a confirmation of the visit and contact with the household. This can be done with a combination of capturing an image, capturing the geo-location during the interviewing process, and a range of operational information that can be used to monitor operations and analyse responses (often referred to as paradata – see paragraph 3.213), including the time taken to complete the form, the date and time the form was completed, and the device used to complete the form.

3.228. Electronic questionnaires can also help the NSO to provide census results more quickly by facilitating the transfer of data relating to a household to a central database immediately or soon after its enumeration. This can be achieved through various methods:

- real-time connectivity whereby data is transmitted electronically as soon as it is collected;
- periodic/on-demand synchronization when a connection becomes available; and
- offline data collection with later uploads to a local centre via cable or Bluetooth connectivity.

Electronic questionnaires reduce the amount of material (such as paper questionnaires) to be printed, distributed and returned, and may reduce data scanning and capture errors (and costs). Yet the use of electronic questionnaires may introduce new expenses such as the costs of the electronic devices used for data entry in the field, and the transmission of such data.

i. Electronic questionnaire: interviewer-administered

3.229. Census interviews can be undertaken using an electronic questionnaire on smart phones, tablets, laptops or other portable devices using a computer-assisted personal interviewing (CAPI) methodology. Electronic questionnaires may also be employed in computer-assisted telephone interviewing (CATI) in which the interviewer records the respondent's answers to questions asked over the telephone. Each enumerator can be assigned one or several EAs so that the records are tagged with the respective EA identity and enumerator's/supervisor's name. The device may also be able to capture information on the location of the interview, time of day and other metrics that may be useful for quality control. A CAPI system integrated with GIS and operational management applications which captures geo-coded data together with operational information (such as date and start and end times of the interview) can improve the monitoring of data collection operations, and the coordination of field operations, logistics, and communications.

3.230. For areas with lack of Internet connectivity, systems with local direct data exchange between the interviewers and supervisors should be considered in order to facilitate a review of the completed interviews by the teams' supervisors even before the data from such interviews are sent to the central database.

ii. Electronic questionnaire: self-enumeration method

3.231. Computer-assisted self-interviewing (CASI), whereby respondents access and self-administer the electronic questionnaire by means of an Internet browser application, can substantially reduce operating costs in areas where a high rate of coverage can be expected using this method. Electronic questionnaires for households should be implemented via a publicly-accessible Internet portal. Households are usually provided with a unique identifier (often delivered by mail, email, by enumerator or by self-registration system) that is used to initiate their questionnaire, or resume a partially completed questionnaire, via the Internet. Householders may prefer to respond using an electronic rather than a paper questionnaire for its convenience. If the census information is collected in a multimodal approach, for example by offering both electronic and paper questionnaire options, a system will be required to track the status of each dwelling (questionnaire) throughout data collection to assess the completeness of coverage and to ensure non-response follow-up is not attempted with households that have already responded.

3.232. The design of a system implementing self-enumeration should demonstrate effective solutions for dealing with the following common problems:

(a) duplication of submission of information (for example by different members of the same household);

(b) fake or bogus submissions/digital vandalism (for example, the submission of data for a non-existing address or in the name of another household by hackers); and

(c) attempts to obtain confidential data by gaining access to the unique identifier used to access the form.

3.233. However, total self-enumeration remains a challenge: access to the Internet is not universal, and there are groups of persons in the population who are not frequent users (older persons in particular) and may have technological barriers; recent immigrants may have language barriers; and persons with disabilities may experience difficulties reading the questions or supplying the answers to them. Thus, this method is often implemented as complementary to, but not entirely replacing, the interview-administered enumeration.

b) Handheld or mobile devices

3.234. Mobile devices such as smartphones or tablets can greatly improve the efficiency of census field operations by facilitating: the interview; the communication and supervision of fieldwork; data entry and validation; and data transmission. GPS-enabled handheld devices can assist enumerators with navigation by identifying buildings for each EA using GPS coordinates and capturing the interview location information.

3.235. When deciding to use mobile devices in the census process, it needs to be determined whether the NSO will purchase and provide the device, or the field officers will use their own devices. As the availability and proliferation of devices increases there can be financial benefits, as well as reduced training needs, if field officers can utilize their current device rather than be provisioned with a new device. This, however, presents a range of technical, support, security and legal considerations, and for this reason this deployment option should require detailed specifications on the compatibility/ platform of the device, minimal requirements (in terms of the

storage, memory, performance, communication capabilities, battery capacity and life, as well as screen size and resolution). Another significant consideration is the fact that mobile devices operate on different platforms, and thus developing applications for different platforms increases the costs and efforts unless such cross-platform portability is already provided by the electronic system chosen for the data collection and census operations management.

c) Geographic information system (GIS) and geospatial information

3.236. GIS may be used for several major purposes during a census:

i. to place the enumerator in the context of the environment, by showing his/her location, the location of all or surrounding dwellings in the EA (as points or where available as footprint polygons), and whether or not each has already been enumerated;
ii. routing capabilities for enumerators to navigate to the dwelling (for the purpose of making return visits for example) and/or for supervisors (for audit visits);
iii. to provide controls for the supervisory staff, especially with respect to the coverage control of the EA;
iv. to produce various dashboards and GIS-based analysis of the progress of the fieldwork for the census managers.

3.237. Such uses rely on the location-enabled mobile devices (GPS or analogous positioning system) and availability of digital maps (sourced from the Internet in the areas with connectivity or stored on the device in areas where the connectivity is absent or sporadic).

3.238. The choice of technology for production of EA-level maps – which is dependent on the choice of data capture software - can greatly affect the time it takes to produce maps. These must be taken into account during the planning stage in order to establish an accurate timeline for the census operation. An extensive elaboration on the use of both GIS and GPS is presented in Section IV above.

d) Dashboards

3.239. Dashboards present aggregate information about the overall progress of the census field operation in a concise, visually compelling and efficient manner. The key indicators to be tracked with the dashboards are to be chosen and defined at the planning stage. A dashboard is not a substitute for detailed reports, but serves as a supplementary tool that provides supervisors and managers with dynamic, possibly real-time, updated information.

3.240. A layered approach to the construction of the dashboard presents the possibility to visualize different indicators to various members of the management team. While the managers may be more interested in the overall progress of the census data collection, the supervisors in the field may be getting the indicators derived from the data collected by their specific team.

3.241. Access to the dashboard information should be well-protected and limited to only those team members that need such information for their operational duties.

e) Contact centre

3.242. The contact centre, also known as a call centre, is a valuable component of field operations. It can be utilized at every stage of the census, but it is particularly helpful in supporting fieldwork and contributing to the efficiency and accuracy of the enumeration. A contact centre can employ

various communication channels, including phone calls, chats, and web-based FAQs, to provide assistance to respondents. Interactive voice response (IVR) technology can be used to direct calls to specific agents based on the options selected by the caller. It can also be used to provide consistent answers to frequently asked questions, thereby reducing the workload of field and management staff. Website call-back and chat features can be implemented to assist respondents while they fill out the online e-questionnaire in the portal. Contact centres can therefore offer support to both the public and census personnel, ensuring that everyone involved in the census process has the resources they need to complete their tasks effectively.

f) Situation room

3.243. A census situation room serves as a centralized operational hub for coordinating and managing the multifaceted operations of a census field operation. Its primary purpose is to ensure the smooth execution of enumeration activities, swiftly addressing challenges and optimizing resources in real-time. In leveraging geospatial technology, the situation room can: (a) provide real-time visualizations of field operations across regions thereby enabling decision-makers to monitor the progress of data collection; and (b) cross-check enumerator locations with predefined areas to predict potential bottlenecks and allocate resources as and where needed, and ensure enumerator safety. In the event of unforeseen incidents or challenges, such as natural disasters, civil unrest, or other disruptions, geospatial data can help the situation room quickly assess the impact on census operations and devise strategies to mitigate the challenges.

g) Short messaging service (SMS)

3.244. Short messaging service (SMS) offers a versatile tool for the census field operation. It can be used throughout various stages to: i) share information among field personnel and respondents (such as passwords and guidelines); ii) send alerts, marketing messages, and reminders; and iii) integrate with the central database for critical event notifications (such as low coverage or violations). In addition to SMS, other communication media like USSD and media platforms can further facilitate communication. These platforms can be used for sharing information, receiving feedback, and supporting both respondents and field staff.

VIII. Data processing

3.245. No matter how thorough and accurate the census enumeration is, the usefulness, quality and timeliness of the census tabulations will suffer unless the collected data are properly processed. An important element of a successful processing operation is the close and continuing collaboration, at all levels, between the data-processing staff, the subject matter specialists and the general statistical staff. As a minimum requirement, the subject matter and general statistical staff will need to become familiar with, and take a continuing interest in, the processing plans and operations, while the processing staff will need to become similarly familiar with, and involved in, the core elements of the census.

3.246. Plans for data processing should be formulated as an integral part of the overall plan of the census, and those responsible for such processing should be involved from the inception of the planning process. Data processing will be required in connection with the results of census tests, the compilation of preliminary results, the preparation of the output tabulations, the evaluation of census results, the analysis of census data, the arrangements for storage in and retrieval from a database, and the identification and correction of errors. In addition, data-processing technologies are playing an increasing role in the planning and control of field operations and other aspects of census administration. Data processing has an impact on almost all aspects of the census operation

ranging from the selection of topics and the design of the questionnaire to the analysis of the final results. Therefore, data-processing requirements in terms of personnel skills and knowledge, space, equipment and software (computer programs) need to be considered from the point of view of the census as a whole and at an early stage in the planning.

3.247. When processing census data collected through various different methods and modern practices, several considerations come into play. Firstly, integrating data from diverse collection modes into a cohesive dataset is essential, but it can be challenging due to potential differences in formats and structures. Therefore, establishing a robust integration process is vital for achieving accuracy and completeness. Additionally, assessing the quality of data collected through different modes is imperative. For instance, the responses given in self-enumeration questionnaires may be less reliable, necessitating thorough error identification and correction (see Part Four for a thorough discourse on assuring the quality of the census). Furthermore, ensuring the security and confidentiality of the data, especially when collected electronically, is paramount. Implementing encryption and other security measures safeguards against unauthorized access, while anonymization techniques protect individual identities.

3.248. To address these considerations, employing a versatile data management system is advised - one capable of seamlessly integrating and handling data from various modes. Establishing specific data quality checks and validation rules aid in error identification and rectification. Encryption and other security measures are crucial for safeguarding data integrity, and anonymization techniques help preserve confidentiality.

3.249. To successfully integrate new computer hardware or software into the census, it is necessary to expand and enhance the skills of the existing data-processing staff. This is because unfamiliar technology often demands new proficiencies for operation and maintenance. Given the intricate nature of the census process, it is important that the data-processing team be well-versed in the specific requirements of the census for efficient data processing. Early training is essential in order to integrate the staff into census planning and operations in a timely manner, enabling them to acquaint themselves with the new technology and contribute valuable insights to the process. This early preparation ensures that the team is poised to process the collected data promptly. The training programme should encompass a comprehensive range of skills necessary for the staff to be able to execute their roles effectively, including proficiency in operating the new hardware and software, processing census data, conducting quality checks, generating census reports, troubleshooting hardware and software issues, and maintaining the technology infrastructure. To cater for diverse learning styles, the training should be delivered through various formats, such as classroom instruction, hands-on practice, and online learning modules. Additionally, it is valuable to evaluate the effectiveness of the training programme through methods such as surveys, interviews, and observations, using the results for refining future training endeavours.

3.250. There is a potential downside to training census staff on the software too early. Staff might leave before the census even starts due to various reasons (such as finding other job opportunities, retiring, relocating, or facing personal commitments). This exodus could adversely affect census quality and efficiency as replacement staff may lack adequate training and require additional resources. To mitigate this risk, it is recommended that NSOs train only committed staff, offer attractive incentives, provide avenues for professional growth, foster a positive work environment, and have a contingency plan for replacement training. Factors to consider include the cost of training, availability of qualified trainers, and the importance of maintaining consistency in the census process. Ultimately, the decision to conduct early training should be tailored to the specific circumstances of each case.

3.251. Decisions will need to be made concerning the location of the various data-processing activities within the country, including the extent to which the processing work is to be decentralized. This decision should be partly based on the ability to recruit the required personnel for the processing operations. Acquisition of both equipment and supplies can require long lead times; estimates of both data capture and computer processing workloads must be made early to enable timely procurement. Closely related to the question of equipment is that of the provision of adequate space. Although the maintenance of most personal computer equipment no longer requires adherence to rigid standards in terms of temperature, humidity, dust and so on, attention to issues related to power supplies is still important. Inevitably, more important is the attention to be devoted to the maintenance of servers (especially heavy-duty servers), where most of the information processing is likely to take place and saved, as well as the data transmission infrastructure.

3.252. Ensuring smooth and uninterrupted Internet and web communications is vital for the success of census operations, as timely and accurate data transmission between different units is crucial. To achieve this, it is advised that NSOs employ diverse Internet connections from different providers, implement load balancing for improved performance, utilize content delivery networks to ease server load, and maintain vigilant network monitoring to promptly address any issues. In addition, for traditional archiving, creating a secure and controlled space to store completed census forms is imperative due to the sensitive personal data they contain. This involves employing physical security measures, logical security measures such as encryption and access controls, environmental controls to protect against damage, and implementing audit trails for accountability.

3.253. When preparing for census data editing and tabulation, decisions regarding the software to be employed are pivotal. This choice hinges on several key factors, encompassing the census's scale and complexity, the proficiency of the data-processing team, budgetary constraints, and the availability of technical support. Various software options are available, falling into three primary categories:

(a) census-specific software tailored explicitly for census processing, such as CSPro;
(b) general-purpose software applicable for various tasks including census processing, such as spreadsheets, databases, and statistical analysis software; and
(c) customized software developed specifically to align with the unique needs of a particular census organization.

3.254. When opting for editing and tabulation software, it is imperative to weigh factors such as:

- user-friendliness for the data-processing team;
- requisite functionality encompassing data entry, editing, imputation, tabulation, and reporting;
- scalability to manage extensive datasets;
- robust security features to safeguard data confidentiality, affordability; and
- the availability of reliable technical support.

3.255. Subsequently, the selected software will necessitate customization to align with the precise requirements of the census organization, potentially involving code modification, report creation, or procedure development. Allocating ample time for customization is crucial. Additionally, the data-processing team should undergo comprehensive training on the software, covering all essential features.

3.256. Outsourcing some of the predominantly IT-related operations may be considered. Outsourcing should be implemented in such a way as to bring immediate economic and quality advantages to census operations. However, it is vitally important to approach outsourcing in a

manner that upholds data confidentiality. To safeguard this, NSOs should rigorously select reputable contractors known for their commitment to data confidentiality, and contracts should encompass provisions mandating confidentiality protection. Continuous monitoring of contractors' adherence to these provisions is imperative. Employing encryption and other security measures during data transmission, and limiting access to the data, both internally and externally, are essential steps in preserving data confidentiality. Lastly, it is paramount that the NSOs retain control over the data, ensuring that outsourcing operations do not entail any relinquishment of data control. (See Section XIV in Part Two for a more detailed discussion on the appropriateness of outsourcing).

A. Method of processing

3.257. The choice of an appropriate processing method depends on various factors specific to each country, including the data collection methodology employed, the size and complexity of the census, the availability of technology and resources, and the desired level of detail in the census results. Rapid advances in data-processing technology have greatly increased the speed and reliability of producing detailed tabulation, thereby making computer processing the standard method of processing around the world. Furthermore, an alternative to mainframes, whose computational power was necessary before the advent of lighter and more scalable IT hardware solutions and more recently cloud-based computing, is the use of a client-server environment.

3.258. Cloud-based computing, often referred to as cloud computing, is a technology model that allows individuals and organizations to access and use computing resources over the Internet, as opposed to owning and managing physical hardware and software on their local computers or data centres. In cloud computing, these resources are hosted and managed by third-party service providers in data centres located around the world. Cloud computing offers several benefits for conducting a census: it makes the process more efficient, cost-effective, and scalable; and offers the ability to store, process, and analyse data while providing flexibility, scalability and cost savings compared to traditional on-premises infrastructure efficiently and securely. These advantages can lead to a more accurate and timelier census. However, it is important to acknowledge the potential challenges associated with cloud computing, particularly the perceived risk of data disclosure to third-party service providers. To mitigate this risk, NSOs should carefully evaluate the security measures and data protection policies of potential cloud providers, and ensure that data confidentiality and privacy are maintained throughout the census process.

3.259. Several lighter tasks, including the editing and tabulation of data files, can very well be done on small-sized desktop systems that can be placed in substantive departments and in field offices. On the server side, most of the heavier computing operations, such as scanning, aggregation and analysis of large sets of microdata, coordination of data transmission, Intranet web hosting and so forth, can be executed more reliably than on microcomputers. However, a client-server environment to handle census data must operate over a robust and secure local area network (LAN) or wide area network (WAN) into the cloud environment. Therefore, computer work is not necessarily dependent on a centralized data-processing facility, provided that a robust LAN or WAN interconnects workstations dispersed over various offices, buildings and different parts of the country.

3.260. In a NSO that utilizes a networked computer environment, the central file or database servers allow both data and program files to be stored in a central location. This system economizes specifications of client computers and removes the need for much physical movement of programs and data on other computer media. Data storage requires frequent backups of the system information to avoid major data loss due to hardware or software faults. Thus, servers have a strategic importance, and their location and administration must be well defined and secure enough

to ensure data protection. Also, it is recommended that proper business continuity and security policies duly certified by the competent authorities should be in place.

3.261. In determining the type of equipment to be employed and the advisability of a new machine installation (either complete or partial), or of additions or upgrades to existing equipment, consideration should be given to all the processing requirements of the data collection programme for which the population and housing census is but one part. Only on this basis can a reasonable decision be made. Decisions on the type of data-recording equipment and computer equipment should be made at least one year in advance of the scheduled date of enumeration in order to allow for:

(a) appropriate questionnaire design and proper preparation of instructions to field staff;
(b) the development of coding schemes, specification of data-handling controls and procedures; and
(c) the recruitment and training of data processing personnel.

3.262. Rapid processing of a pretest or pilot census that covers end-to-end census operation, including enumeration, initial census result, output dissemination, and handing over and closure procedures, is particularly important for identifying any improvements needed in the census questionnaire, instructions to enumerators, computer systems or whatever other preparations may be needed. It is recommended, therefore, that arrangements for using appropriate equipment and software be made well in advance of such tests. It is also recommended that all systems used to support census operation be thoroughly tested in advance of operations to ensure that they function as intended and that they are secure (that is, they will not lead to loss or unauthorised disclosure of data).

B. Preparation for data capture

3.263. In the case of paper questionnaires, the most common procedure is to have the census documents arrive in the processing centre in batches by EA. Maintenance of these batches throughout the data processing is recommended, since documents for a given EA reflect the work of one enumerator and may contain a series of errors systematic of that person. To ensure the integrity of the batches, the questionnaires should be stored in a specially designed storage facility. The batch for each EA should first be checked for completeness, geographic identification codes and other characteristics of acceptability before being sent to the next stage of data processing. Transcribing all coded data onto another sheet (for example, the coding form) should be avoided since it may add transcription errors. The same considerations apply to the case of electronic transmission of questionnaires or when the first phase of data processing consists of the scanning, and text or image recognition, of the questionnaires. In the case of self-completion questionnaires transmitted electronically via Internet or a mobile devices, it is appropriate to set up a metadata model where the EA can be recorded. As far as storage is concerned, if paper questionnaires are scanned, secure media storage for any backup copies as well as the originals should be planned for.

3.264. If the enumeration is conducted in a multimodal approach, for example using self-response by Internet and field follow-up of non-respondents, it may not be possible to batch questionnaires for processing by EA. A master control system will be required to track the status of each questionnaire throughout the collection and processing operations in order to ensure completeness of coverage. This is closely aligned to any master control system that is being utilised to manage dashboards and monitor the progress of the data collection (see paragraphs 3.239-3.241).

C. Data capture

3.265. Data capture involves converting census data, collected through various methods including paper questionnaires, electronic data collection devices, or online self-enumeration, into a computer-readable format. This process can occur both during and after the enumeration period.

(i) Data capture during enumeration

3.266. Electronic data collection (including collection by Internet or using handheld or laptop computers) has been widely used by countries. Electronic data collection means the integration of interviewing and the data entry process including data capture, coding and consistency checks. Electronic data collection with handheld devices, Internet or telephone allows the capture of information with relevant codes (there might be some exceptional variables (such as occupation and industry) that may require more detailed coding centrally. It also allows the identification of potential errors during the interview with pre-programmed consistency checks. Because the consistency checks are performed in real-time, errors and inconsistencies can be resolved, and corrective action can be taken by the respondent or the enumerator. However, introducing editing rules into the data collection application has to be carefully considered so as not to (a) affect its performance significantly in the field, (b) introduce a bias in the responses, and (c) affect the quality of the interview for questions that may not be answered properly, especially in case of information provided from proxy respondents.

(ii) Data capture after enumeration

3.267. The more traditional methods of processing census data occur after the enumeration phase. They include keyboard data entry, optical mark reading, optical character reading or image-processing techniques, such as intelligent character recognition. Computer-assisted keyboard data entry is usually carried out using personal computer data entry programs with built-in logic controls. Some of the tasks accomplished by the programs are:
(a) verifying that EA codes are valid, and copying them automatically from one record to the next;
(b) assigning a number to each person in a household automatically (and perhaps to each household within an EA);
(c) switching record types automatically if the program's logic requires it;
(d) checking that variable values are always within predetermined ranges;
(e) skipping fields if the logic indicates doing so;
(f) supporting keyboard verification of the information entered earlier; and
(g) generating summary statistics for the operator and the batch.
In order not to delay the data capture task, data entry applications should limit checking to problems that are either very serious (for example, wrong EA code), or likely to be caused by a simple misread or key entry mistake. More sophisticated checking is deferred until the editing stage (discussed at Section E below).

3.268. Optical mark reading (often called optical mark recognition) (OMR) equipment has been available for many years and has nowadays reached good levels of reliability. OMR is the simplest of the commonly available form data capture technologies. Owing to relatively stringent requirements for the successful data capture of the paper, countries with very dusty environment or humid climates and poor transport infrastructures are discouraged from using OMR as it demands the necessity to heed special questionnaire design restrictions, to consider the quality of the paper, and to adhere to precise specifications regarding the printing and cutting of the sheets. In some developing countries, this may mean that local production of the questionnaires will be problematic. The need to reserve a relatively large space for marking areas and to adhere to other limitations imposed by OMR equipment sometimes make it difficult to design the best questionnaire from the point of view of the enumeration process.

3.269. OMR questionnaires can be marked by the respondent or by the enumerator. Marking by respondents is attractive from a cost perspective, but it depends on the presence of a cooperative spirit and relatively high levels of literacy. A practical problem is that most OMR devices put restrictions on the writing instrument and the colours that can be used in the marking. Assuming that the rules are followed, the rejection rate for marked forms is often low, especially if the forms have been inspected visually before being fed into the readers. Converting a manually completed census questionnaire to an OMR format after it has been received in the census office is inefficient and becomes a source of errors and should therefore be avoided.

3.270. Optical character reading (also called optical character recognition) (OCR) and intelligent character recognition (ICR) consist of the use of special equipment to read characters at specific locations in the questionnaire. These two methods involve very similar technological approaches. Specialized sources tend to identify OCR with the capability of recognizing printed characters only, whereas ICR would extend this capability to handwritten text. There is no agreed definition of ICR. In the context of censuses, therefore, this would require that handwritten text in the filled-in questions be as standard as possible so as to enable efficient recognition. In general, recognition of numerals is more efficient in an uncontrolled environment, that is to say, where the machine has not been adapted to the writing style of a particular person. OCR and ICR technology has matured considerable with sophisticated recognition algorithms and the use of neural networks for self-learning.

3.271. Imaging techniques and scanning devices, together with OCR and ICR software, have been used by several countries for data capture. Experience shows that significantly low error rates are achieved at an optimum cost using these techniques. The efficiency is greater in the case of numerical and alphanumerical characters written by trained enumerators. However, alpha-numerical characters are prone to higher error rates. Extensive testing must be conducted well in advance to determine the best type of equipment and paper. The use of imaging techniques is also dependent on the availability of local maintenance and support capabilities. Whatever methods of coding and data capture are chosen, it is essential that they be carefully tested before final adoption. Recognition engines can be customized to recognize various sets of characters and scripts, but unless good experience is available at the census office, careful planning and preliminary work are needed in conjunction with the OCR or ICR system providers. A combination of ICR (for numerical characters) with computer-assisted coding (for alpha characters) is also an effective method used by some countries.

3.272. In addition to the benefits of the scanning technology for capturing the information, an important by-product of scanning census questionnaires is that this allows for the possibility of digitally filing and naming the scanned questionnaires. This increases the efficiency of storage and retrieval of the questionnaires for future use, particularly during subsequent data-editing operations.

3.273. The quantity and type of data entry equipment required will depend on the method of data capture selected, the time available, the size of the country, the degree of decentralization of the data capture operations, and a number of other factors, such as the use of digital enumeration approaches. For keyboard data entry, the average input rates usually vary between 5,000 and 10,000 keystrokes per hour. Among the factors that affect operator speed are:
(a) the supporting software and program with easily navigable screens, spell checker on the description fields if any, keyboard shortcuts throughout the program, less utilization of the computer mouse, and so forth;
(b) the complexity of the operators' tasks;

(c) the ergonomic characteristics, reliability and speed of the equipment;
(d) the question whether or not work is always available;
(e) the training and aptitude of the recruited staff; and
(f) the motivation of the workers.

3.274. Several options are available to help ensure that data entry operations are completed in a timely manner. They include (a) procuring more equipment; (b) increasing the number of working hours by working double or even triple shifts and during weekends; and (c) applying independent verification to varying extents. In the case of keyboard data entry, with the increasing safeguard of data quality by data entry programs, complete verification has become less necessary. Full independent verification may be applied only in the initial stage of data entry and may be reduced when each worker has achieved an acceptable level of quality. After that, a sample verification plan can be applied. Operators may be assigned to sample verification depending on their observed error rate. The work of reliable operators need be verified only for a small sample of the EAs, while more extensive verification is continued for the more error-prone operators.

D. Coding

3.275. Whenever possible, pre-coded responses should be used in census questionnaires with numerical or alpha-numerical codes. While computer processing of write-in responses to open-ended questions has improved in recent years, it can still be time-consuming and prone to errors. Therefore, using pre-coded responses can significantly streamline data processing and improve accuracy. The incorporation of artificial intelligence (AI) and machine learning (ML) technologies would be ideal. These advanced technologies offer the potential to automate critical tasks, including the classification and coding of open-ended responses. Such integration stands to enhance the overall efficiency and accuracy of the data collection process. While the use of pre-coded responses in census questionnaires is encouraged, there will inevitably be instances where respondents provide written responses necessitating coding. In such cases, dedicated computer programs or skilled coders can be employed to translate these written responses into codes. Although automatic coding is the preferred approach, certain situations may call for manual intervention. While coding directly during interviews can be advantageous, it is often logistically challenging as enumerators may not always possess the requisite training or resources for accurate coding. Therefore, in most cases, responses will be gathered in text format and subsequently coded by proficient coding experts. Given the scale of coding operations in a census, optimizing automatic coding processes is paramount to minimize the need for human intervention. Techniques such as natural language processing (NLP) and deep learning can be employed to facilitate this optimization. Additionally, AI and ML technologies can be harnessed to effectively classify occupations and industries based on text descriptions, a valuable tool for coding open-ended responses pertaining to occupation and industry.

3.276. Automatic or computer-assisted coding will efficiently support the coding activity, reducing coding errors and speeding up the coding process. When required, a coder normally works with one or several codebooks for various items in the questionnaires. This technology expedites coding tasks, liberating coders to concentrate on more intricate responsibilities, such as coding open-ended responses (to the occupation question for example) or addressing coding discrepancies. Additionally, automatic or computer-assisted coding provides invaluable feedback to coders, enhancing the overall accuracy of the coding process by aiding in error identification and correction. By automating laborious coding tasks, this technology alleviates coder fatigue, thereby elevating the overall quality of coding work. Lastly, it ensures the consistency of the coding process by enforcing uniform adherence to coding guidelines among all coders, leading to enhanced data quality.

3.277. ***Computer-assisted coding*** uses personal computers to assist the coders. The process requires that all the codes be stored in a database file and be accessed by coders during the coding operation. Computer-assisted coding is based on at least two general approaches. In the first one, coded answers are matched to a set of keywords. Textual information from the census questionnaire is parsed and compared to an indexed list of keywords, and then the likelihood of matching between found keywords and coded answers is measured and scored. If the score results are over a certain (high) threshold and there is no ambiguity, a sorted list of coded answers is presented to the coder, who retains the ultimate decision of accepting or refusing the system's proposed answers. In using this method, it may be advantageous to change the order of activities so that the capture of pre-coded information in the questionnaire occurs first, followed by the capture and computer-assisted coding of the remaining information.

3.278. In the second approach, which is mainly used in image processing of data (the ICR method) for non-Latin languages or multilingual countries, because of the difficulty and existing problems in character (alpha-numerical string) recognition, the procedure is as follows. After the scanning and during the coding operation phase, the image of the text will be shown on the monitor and, at the same time, a pull-down menu from a coding database will present the coder with the ability to enter as few key entries as possible to get to the full textual and coding content of a specific case. When the coder selects a code, it will be allocated and saved in the database for that specific case. Although this approach is more time consuming and costly in comparison to the first approach, the quality of coding is much higher than in the more traditional way of coding.

3.279. On the other hand, both techniques have several similar advantages:

(a) capturing the pre-coded information at an early stage leads to some data files becoming rapidly available, which opens up the possibility of generating and releasing preliminary census results;

(b) the computer-assisted coding process provides an opportunity for a computer system to alert the operator to problems with data supposedly already captured, for example, missing information for a fully pre-coded variable;

(c) the coder works directly on the computer screen; and

(d) information from other variables may be helpful in determining applicable codes for write-ins.

3.280. ***Automatic coding*** is a process in which the decision about the code to be assigned is delegated to a computer program. The main difference from computer-assisted coding consists in the automatic acceptance of the answer if its score is (as with computer-assisted coding) over a predetermined threshold and relatively higher than possible identified alternatives. Both computer-assisted and automatic coding systems may exploit self-learning capabilities of neural networks to fine-tune their capacity of detection. A human operator becomes involved only in those cases where the software cannot resolve the issue. While computer coding may use, in addition to the written response for the item in question, other relevant information available in the record or the questionnaire, automatic coding is more applicable in cases where the data-capturing process has already been completed, by Internet, handheld devices or other forms of electronic data collection, manually or by some form of automatic reading. Developing computer software for automatic coding is a complex task. Automatic coding methods need to be complemented by computer-assisted or conventional coding methods for unresolved responses.

E. Data editing[106]

3.281. Raw data files can contain various types of errors, including those introduced by respondents, enumerators, data processing operations, coding, data entry, and transcription. Respondent misinterpretation, enumerator misunderstandings, typographical errors, software-related issues, and transcription errors are all potential sources of errors in raw data. From an operational point of view, such errors are of two types: (a) those that have the potential of blocking further processing (critical errors); and (b) those that introduce distortions into census results without interrupting the logical flow of subsequent processing operations (non-critical errors). Throughout the correction process, backup copies of the original data file should be consistently created. The timely identification and correction or errors is imperative to prevent inaccuracies that could profoundly impact policy-making, resource allocation, and crucial decision-making processes. Various error-checking and correction methods, such as range checks, consistency checks, logic checks, and manual review, can be employed to enhance data accuracy, with manual review offering the highest accuracy albeit with increased time investment.

3.282. Since for large censuses manual correction is rarely economically feasible, the conditions for such corrections are usually specified in specially designed computer programs for automatic error scrutiny and imputation based on other information for the person or household or for other persons or households. Whenever imputation is used, a flag should be set so that analysts are able to distinguish between reported information and that imputed by the editing system. In cases where insufficient information hinders error correction, methods such as hot deck imputation can be deployed. This technique uses information obtained from previously processed persons or households with similar characteristics as the "best fit" value in replacing missing values or values that have failed processing edits. However, this technique requires careful programming work, considering that the search for appropriate information in the census database would slow down computer program execution.

3.283. A prudent approach in data processing involves isolating out-of-range or blatantly inconsistent values prior to editing or classification, mitigating the risk of introducing statistical biases. But precautionary measures should also be defined and set for the fact that overambitious automatic editing programs may cause the so-called "corrected" data to be significantly flawed. In this respect, it would make sense to have an acceptable cut-off value for error rates at the EA level. If a data scrutiny program finds that more than a certain percentage of the records in a particular batch have one or more serious problems, the whole batch should be rejected and subjected to human or fieldwork verification.

3.284. Developing editing and imputation rules should be entrusted to subject matter experts rather than computer programmers. This is because specialists possess a profound comprehension of the data and potential sources of errors, enabling them to identify likely error types and establish effective correction protocols. An error scrutiny and editing plan should be crafted early in the census, delineating specific rules and their implementation procedure. This plan should also incorporate a mechanism for periodic review and adjustment of the rules. These rules must be meticulously documented and shared with programming staff to ensure accurate and consistent implementation, utilizing clear and straightforward language for easy comprehension. While

[106] For further details on census data editing, see *Handbook on population and housing census editing*, *Revision 1,* United Nations publication, Sales No. E.09.XVII.11, United Nations, 2010.

programmers bear the responsibility for preparing the coding rules, subject matter experts should verify the accuracy of the codes. Thorough testing by both subject matter and software testing experts is imperative before applying the editing programs to the census dataset, detecting and rectifying any program flaws under diverse conditions in order to ensure comprehensive functionality. Adhering to these guidelines yields numerous benefits, including heightened data accuracy, reduced error risks, enhanced transparency in the editing process, improved collaboration between subject matter experts and programmers, and better quality of census results. However, challenges persist, encompassing the time-consuming nature of rule development and program testing, and the potential complexities in coordinating the efforts of subject matter experts and programmers, along with the need to mediate disagreements between them.

F. Validation

3.285. The outcome of editing is a set of records that are internally consistent and in which person records relate logically to other person records within the same household. This process does not, however, provide the full range of assurance necessary to accept the data set as the best possible. A range of conditions could create errors that cause the data to be consistently wrong: for example, perhaps a condition in the editing suite itself is set incorrectly; proportions in an imputation program may be set wrongly; or enumerators may complete a collection control panel incorrectly. To identify such consistent errors, it is necessary to critically review some key aggregated tables to isolate outlying numbers and identify the cause of the unusual values. These key tables may be a subset of those intended for output or may be tables specifically designed for this purpose.

3.286. It is recommended that a bottom-up approach be used in this process. That is, the tables should first be examined for a selection of EAs, then for the next level up and so on up to the first set of national tables. There are two reasons for this:

(a) The first EA will complete the processing cycle well before any other geographic level. Thus, commencing at this level gives the earliest possible warning of any problem, enabling corrections to be made before a large amount of reprocessing is required.
(b) It is far simpler to examine a few hundred records within an EA than to attempt to resolve the problem in the millions of records in a national file.

3.287. A crucial stage in the process is designing the analytical tables. One way of approaching this could be to identify a set of variables that are conceptually consistent with those in the previous census (or a major survey) or in administrative records from various authorities in the country, such as expatriate visas issued, national ID programme or number of registered establishments. Thus, a set of benchmark values could be constructed before the field operation commences and then compared with those from the actual enumeration. The content of the benchmark set will depend upon the content of the enumeration and much of this must therefore be determined by each country. However, any census will include the variables age and sex so a comparison of the age pyramid and sex ratio for each 10-year age cohort would be basic elements of such analysis.

3.288. A second component of the analysis is the compilation of a set of data regarding expected changes since the benchmark survey. For example:

(a) It is possible that in the time since the previous collection, improvements in maternal health care programmes have led to an increased survival rate for women, in which case cohort survival ratios for females should be higher for younger women than older ones.
(b) If literacy is included in the analysis and government policy has been to strongly support increased school attendance, an increase in the proportion of literate people could be expected.

3.289. There will be a need for careful judgement when the analytical tables show a significant and unexpected difference from the benchmarks. While it may be found that the difference is due to a problem with the current collection, it could also be due to:

(a) a problem in the collection that has generated the benchmarks;

(b) a genuine and previously undetected social change that is being correctly revealed by the current census data.

3.290. In the latter two cases it would be wrong to make any change to the current data set. However, it is important that details of the investigation are made known to users (by preparing suitable metadata) so that they would be able to interpret and analyse the data correctly. If such analysis indicates that there is a problem with the census data, it will also be a matter of judgement on how to react to it. One solution might be to revise the input processing system to prevent the problem from being continued. After applying such changes, and to avoid introducing further problems, it is essential that the input processing system be fully tested and accepted. The second possibility is to decide as to whether to reprocess the records that have already been processed. This decision should be guided by the following:

(a) the significance of the error;

(b) the number of questionnaires that have already been processed;

(c) the time required for the reprocessing;

(d) the impact of such a decision on other subsequent phases of the census (such as tabulation and dissemination); and

(e) the cost implications.

3.291. The elements involved in data processing require joint and integrated work between census topic analysts (such as demographers, geographers and economists) and technical personnel (such as programmers). This is especially so for validation where thematic and comparability (with other sources) indicators are generated separately from tabulations to improve data coherence. Before and after comparations (original and validated data) should be included as part of the validation. In the same way, knowing the data dissemination plan helps organize and prioritize both editing and validation.

G. Processing control

3.292. Careful planning and control are required to ensure an uninterrupted flow of work through the various stages from receipt of the census questionnaires through to the preparation of the database and final tabulations. For the data from questionnaires completed through hand-held devices and transferred to a server, a clear data extraction strategy must be put in place with consideration to the scale of the census operation. The plan should provide for the computer edit to follow closely the coding, checking and recording of the data so that errors can be detected while knowledge related to them is fresh and appropriate remedial actions may be taken.

3.293. Countries may wish to establish a computer-based processing management and control system to check individual forms or groups of forms for each EA or for other geographical units. Such a system should link the databases for EAs and other geographic entities with the control information.

3.294. The system would check and manage progress from process to process so as to ensure the completeness of records at each stage of the processing operations. As specified earlier (at paragraph 2.114), project management software may support the formal description of different processes and provide an environment to control the execution of all operations connected to an individual phase or status of the census. This system should be fed into the overall quality assurance and improvement system, the management of which is elaborated in paragraphs 24.19-24.27. If a

computer-based processing system is established, a close and real-time communication between the headquarters, local offices and enumerators should also be established. This is beneficial for the control of field staff and the management of the logistics of enumeration materials. Also, as each problem occurs in the field, a solution can be shared through the bulletin board in the system, which can greatly reduce non-sampling error.

H. Master file

3.295. When data editing and validation are in progress, new files consisting of clean data records for each person are produced; these can be assembled to build a master file for later tabulations and census indicators (often called the microdata file). This master file, like the raw data files, can have a simple rectangular sequential format. There is usually no need for having the master file organized with a database structure with index files (but neither should it be discouraged). However, the master file should usually be maintained in geographic hierarchy, starting with the lowest geographic entity, sorted by housing unit or household. Another method commonly used to generate tabulations involving both the person, the household or housing unit, is to include in the household reference person's record (or head of household's record) selected characteristics of these latter units. Alternatively, a single hierarchical file can be created involving, for example, person and housing unit records. Whatever the chosen structure, the master file must allow for easy checks, controls and computations to be performed. The hierarchical model should be the best shape to generate this master file including all in one.

3.296. Census master data files are usually very large and require powerful servers to process them or specialized software with those capabilities. Well-equipped desktop systems also have higher computational power and are equipped with much bigger and cheaper mass storage devices. Nonetheless, the hardware infrastructure available to many countries is older, thus two strategies are applied to reduce file size and to make data management simpler. The first involves working with the next lowest geographic entity as a basis, processing the data on this level and aggregating later to obtain national results. The second is to apply on-the-fly compression and decompression to the storage medium. Census files can be compressed quite significantly to less than 20 per cent of their original size. Since tabulation programs access the data in sequential order, using the compressed data will result in a faster reading process. New compression and storage technologies can be incorporated as well as encryption methods.

I. Methods of tabulation

3.297. Preparing the tabulation plan is the substantive responsibility of the demographers and other subject matter specialists who have the necessary expertise in interpreting the census results. This will require consultation prior to the census with principal users of the census information (see paragraphs 2.175–2.178). The duties of the data-processing department should be limited to checking the logic of the various aggregations, designing the required programs and producing good quality results within the shortest possible time. It is possible that the need for initially unanticipated tables will become apparent, so the NSO should always be prepared to produce additional aggregations as required. This may involve newly defined classifications for certain variables, new types of cross-classifications, differently defined geographic subdivisions, and so on. If the master file is organized according to the principles of relational databases in a relational database management system, original and additional aggregations can be designed according to relatively easily structured query language statements.

3.298. Alternatively, if the master file is a list of records with a rectangular structure, online analytical processing (OLAP) tools can be leveraged for creating multidimensional tabulations. It is

imperative to ensure that all requisite information for tabulation production is readily available in the master file, as adding it later may prove to be prohibitively expensive. The tabulation plan should meticulously outline the tables to be produced, the variables encompassed in each table, and the cross-classifications to be applied, also specifying the level of detail to be provided in each table. It is essential for the NSO to review and endorse the tabulation plan prior to data processing, with provisions for revisions as necessary throughout the data processing phase. The tabulation plan stands as a pivotal component of the census process, guaranteeing that census data are presented in a format conducive to user needs, while also ensuring the accuracy and reliability of the data.

3.299. Because of its efficiency in processing large datasets, specialized census tabulation software can ensure accuracy through error-checking features, and offering flexibility in generating customized tables and reports. Its user-friendly interface and low cost make it accessible and cost-effective. The software also includes various functionalities like data import/export, data cleaning, table generation, charting/graphing, and report generation, further enhancing its utility. Overall, specialized census tabulation software is an invaluable tool for NSOs worldwide, significantly improving data processing efficiency, accuracy, and cost-effectiveness.

3.300. Tabulation work can also be easily done by software belonging to either one of two other classes: statistical analysis and database software. However, these packages have not been designed with large-scale sequential or geographic processing in mind. They may require substantially more computer time than a specialized census tabulation system. In countries with a limited capacity of powerful computers, this can be an important consideration.

3.301. Other factors that should be taken into consideration when selecting software packages for tabulation work include:

(a) **Expertise**: The NSOs existing expertise and familiarity with different software systems should be taken into account. A minor improvement in software performance may not justify a significant retraining effort.
(b) **Compatibility:** Compatibility with existing resources, such as operating systems and database software, is important in order to minimize additional costs associated with acquiring new hardware or software.
(c) **Cost-effectiveness:** The software's cost, including licensing fees, maintenance, and support, should be evaluated in relation to its capabilities and benefits.
(d) **Vendor support:** Robust vendor support, including technical assistance, documentation, and training, is essential for successful software implementation and ongoing use.
(e) **Scalability:** The software should be capable of handling large datasets efficiently, ensuring that it can meet the demands of the census tabulation process.
(f) **Security:** Robust security features are vital in order to protect sensitive census data from unauthorized access or breaches.

J. Use of administrative data during data processing (for field-based censuses)

3.302. Administrative data may be leveraged during the data processing of a field-based census, mainly for:

(a) editing of missing or implausible data or non-respondent households: or
(b) validation of results. It may also inform the development of proper edit rules such as checks on the plausibility, or automatic coding systems for specific variables.

Before using administrative data, two key conditions should be in place: for (a) there must be a legal framework that allows such use and for (b) there must be a proven method of linking data between different sources at the unit record level.

3.303. Census data often suffer from undercounting or missing values due to various reasons, such as non-response or failures in data collection. Administrative data can play a key role in imputing missing or incomplete census data. When the census form is incomplete, information from administrative sources, such as records of government benefits, healthcare records, or school enrolments, may be used to either replace - when microdata integration is possible - or to impute missing values.

3.304. In cases of non-response, administrative data sources can be used, such as data on electricity consumption from electricity distributors, to identify occupied dwellings. Data on electricity consumption may also aid in estimating the number of individuals residing in non-respondent dwellings. This will enhance the imputation procedures, resulting in improved census count.

3.305. When administrative data is of high quality, it can serve as an invaluable tool for validating the coverage and content of census data. By cross-referencing the census records with various administrative records such as tax filings, birth and death registrations, or immigration records, discrepancies and inconsistencies can be identified. This validation process helps to ensure that the census data are reliable, providing a more accurate picture of the population. For example, administrative data can be used to verify that individuals listed in the census are indeed living at their reported address and that demographic information aligns with official records. However, as mentioned earlier, this is contingent upon the quality of administrative data and the degree of coherence with census definitions, concepts and timeframes. In cases where the quality of administrative data at the micro-level is not high yet satisfactory at the aggregates, it may be used in the validation of aggregated census outputs.

3.306. The use of administrative data in census, either to provide data on specific variables or sub-populations can significantly affect data processing. This is because using administrative data often introduces a "pre-processing" step that involves cleaning, standardizing, and transforming the data to ensure it aligns with census requirements and definitions. This pre-processing step is in addition to other data processing elements that may be affected, such as data editing, imputation, and tabulation. Data-processing staff will certainly need to develop skills to deal with challenges involved in the use of administrative data sources. A close and continuing collaboration, between the data-processing staff, the subject matter staff and relevant administrative agency is essential in mitigating and tackling the challenges.

K. Security during data processing

3.307. Security concerns during census data processing are paramount due to the sensitive and confidential nature of the information collected. Ensuring the confidentiality, integrity, and availability of census data is critical for maintaining public trust and complying with privacy regulations. Secure physical and/or virtual environments should be used for census data processing, analysis, and storage. In these environments, strict measures such as encryption, access controls, and firewalls should be implemented to prevent unauthorized access. Additionally, regular security audits and assessments should be conducted to identify and address potential vulnerabilities. Trained and authorized personnel must follow strict data handling and confidentiality protocols, adhere to ethical guidelines, and sign confidentiality agreements. Data anonymization or pseudonymization techniques may also be utilized to further protect the confidentiality of census data.

3.308. Some of the key security concerns during census data processing are noted here:

a) ***Data privacy and confidentiality***

i. Data privacy laws: Census agencies must adhere to data privacy laws and regulations that govern the collection, storage, and processing of personal data. Violations can lead to legal consequences and damage public trust.

ii. Anonymization: Techniques such as data anonymization and de-identification are necessary to protect personal confidentiality. Personally identifiable information should be removed or transformed to prevent data from being linked to identifiable individuals.

b) ***Data encryption***

i. Data in transit: Data should be encrypted while it is being transmitted over networks (for example, from data collection points to data processing centres) to help protect against interception and eavesdropping.

ii. Data at rest: Data at rest, whether stored in databases or as backups, should also be encrypted to prevent unauthorized access in case of data breaches or physical theft.

c) ***Access control***

Role-based access should be implemented to restrict access to census data based on job roles and responsibilities. Only authorized personnel should have access to specific data sets.

d) ***Data leakage and disclosure prevention***

Solutions to monitor and prevent the unauthorized sharing or leakage of sensitive data, either intentionally or unintentionally should be implemented.

e) ***Security auditing and monitoring***

i. Logging and auditing: Detailed logs of data access and processing activities should be maintained and reviewed regularly for suspicious activities and potential security breaches.

ii. Intrusion detection and prevention systems: Intrusion detection and prevention systems should be deployed to detect and respond to potential threats and attacks in real-time.

f) ***Physical security***

Census agencies should ensure that data centres housing census data are physically secure with access controls, surveillance, and environmental protections to prevent unauthorized physical access or damage.

g) ***Incident response plan***

Preparations for dealing with security incidents should be developed by creating a well-defined incident response plan that should outline the steps to take in the event of a breach and how to communicate with affected parties.

h) ***Compliance and audits***

Regular security audits should be conducted and assessed to verify compliance with security policies and regulations. These should include third-party audits and assessments for independent verification.

i) ***Data retention and disposal***

Data retention policies should be defined and enforced to ensure that data are retained only for the required duration and are securely disposed of when no longer needed.

j) ***Employee training and awareness***

i. Security training: Regular security awareness and training programmes should be undertaken for employees and contractors to educate them about security best practices and the importance of data protection.

ii. Vendor security: Similarly, if third-party contractors are involved in data processing, NSOs should ensure that they adhere to stringent security standards and practices in order to safeguard the census data.

k) ***Public communication***

A clear communication strategy for addressing security incidents should be developed, ensuring transparency, and to maintain public trust.

3.309. In summary, maintaining security and confidentiality during census data processing requires a comprehensive approach that includes technical, organizational, and procedural measures. It is crucial to prioritize privacy and security to protect sensitive information and maintain the integrity of the census process.

IX. Census products, dissemination and utilization

A. Introduction

3.310. The population and housing census is a statistical operation of exceptional value to every country. It is the primary source of basic national population data for administration and for many aspects of economic and social planning. Consequently, data from national censuses represent a valuable public good that should be widely promoted by NSOs in order to enhance its utilization by as wider range of users as possible. Thus, the census should not be an end in itself but should be backed by the value of the results, in terms of utilization, by its various users.

3.311. While the importance of disseminating census results to the fullest extent possible should be stressed, there are, nevertheless, some factors that are essential for the NSO to keep in mind. Among these is the prerequisite to ensure openness and transparency in the way the results are disseminated. It is equally important that NSOs maintain professionalism and demonstrate neutrality and objectivity in the presentation and interpretation of the results, and are free from real or perceived political interference, so that the objectivity and impartiality of the statistics is assured. This in turn will build trust in, and acceptance of, the results. Furthermore, the disseminated census results should be of sufficient quality to meet user's stated needs, and safeguards should be in place to ensure individual information is kept confidential.

B. Plans for census products and data dissemination

3.312. In order to maximize the utilization of results from their population and housing censuses, NSOs should have a sound dissemination programme whose objective is to promote the benefits and applications of census data. The NSO should develop and implement an effective strategy for producing and disseminating output products and providing related services based on the demonstrated needs of the diverse user community. What follows are some salient issues for an effective dissemination programme.

1. Developing a dissemination strategy

3.313. A census is not complete until the information collected is made available to users in a form and scale suited to their needs. In order to fulfil this requirement, it is essential to develop a strategy for producing and disseminating outputs taking into account all potential users of the data. This in turn requires identifying such users and their demonstrated needs at a very early stage in the census planning programme so that appropriate products and related services can be developed in good time (see paragraphs 2.175-2.178 of Section X in Part Two).

3.314. An effective dissemination strategy should have a diversified approach for meeting the requirements of different users. Appropriate technologies and media need to be identified for effective and easy dissemination of census data and any accompanying metadata. When developing a dissemination strategy for census data, it is important to take into account geospatial information in order to provide a spatial context to the data. Geospatial data can be used to create thematic maps that visualize census information to make it more user friendly. Geospatially-enabled statistics,

in particular at a high level of resolution, can allow statistical organisations to produce new analytical outputs and conduct a wide range of spatial analysis (see paragraph 3.338-3.351 below).

3.315. A number of key elements should be taken into account in the development of a strategy for census data dissemination, including:

a. identifying the diverse categories of users and their data needs and uses through user consultation and stakeholder engagement;
b. the media of dissemination;
c. metadata to aid in the interpretation of the results;
d. legal and ethical consideration, confidentiality and privacy measures;
e. assessing the required technologies to meet user needs;
f. a dissemination policy;
g. quality assurance;
h. budget and human resources;
i. the need for continuous improvement along with a feedback mechanism for dissemination products;
j. the alignment of census outputs with international standards; and
k. a census release calendar.

These elements are summarized below, or a cross-reference to another location in the document is provided.

a. Identifying the diverse categories of users and their data needs and uses through user consultation and stakeholder engagement

3.316. The demand for, and use of, statistical products and services must drive all census operations. The user consultation process in terms of census products is a major factor in the development of a dissemination programme. The type of consultation discussed in this section complements the consultation that is undertaken to determine census content (see paragraphs 2.175–2.178). The work done at this stage of the census is important in achieving the objective that the census is relevant to users, which is a major indicator of the quality of the census (see paragraph 24.10). The selection of suitable census data products and related services should be guided by a detailed assessment of user requirements with due consideration given to budgetary and human resource constraints.

3.317. Plans for what and how products will be disseminated should be made early enough in the planning process and shared with potential users in order to get their feedback. Based on this feedback, the national statistical or census office can tailor its data dissemination programme to suit the requirements of the users. Maintaining continuous communication and obtaining feedback from users is also important for making modifications to products and services, including being able to respond to user requests that become known later in the programme.

3.318. Based on the foregoing, it is important to note that the supply of census products and services goes far beyond the first couple of years after the census. It is important, therefore, that budget and human resources are available for many years after the end of the census collection activities.

3.319. A wide range of statistical products can be made available to different users. A detailed plan for producing different census outputs should be guided by early user consultations to ensure data and information requirements will be met in a format commensurate with user needs and demands; such a plan will also be a useful guide to prioritizing data processing and tabulations. To help manage expectations regarding data dissemination, a provisional calendar outlining the anticipated release dates for preliminary results, final results, and other census products should be published. However,

it should be noted that these dates may (and often do) change due to unforeseen circumstances, including issues arising during the data processing and analysis phases.

3.320. The definition of geographic output products and services and the scheduling of their release should be closely coordinated with the timetable for the overall census project. The tabulation of census data may require information from the census geographic unit, and thematic maps and digital geographic databases can only be completed once census data processing has been completed.

3.321. Some data users will need specialized products that the NSO is not initially planning to produce as part of the general census programme. It is recommended that the NSO establish a service to meet such specialized requests. The pricing of such special products and services may be included in any pricing policy where cost-recovery measures are adopted.

b. The media of dissemination

3.322. With the rapid development of technology, census data users have an increasing interest in a broad range of products and services from the census organization. The types of output that census offices may produce and disseminate must be current and may include printed products, static electronic products, interactive electronic products, application programming interfaces (APIs), customized products, user-interactive products and special audience products and services. Partnerships with key stakeholders are encouraged in the development of the various census products and the media for their release.

3.323. Printed publications are becoming less popular as many countries move towards digital outputs. For those countries that continue to produce printed publications, target dates for publication should be determined well in advance and processing and printing programmes should be planned accordingly. In addition to traditional methods of printing, there are various methods of reproduction available that are fast, more economical and of good quality, and these should be investigated. For an increasing number of users, environmental concerns and a lack of searchability means online electronic data dissemination is preferred over printed paper. Electronic dissemination reduces cost, and unlike weighty paper publications, supports the direct availability of the data for further computer processing. In addition to pre-processed tabulations, more and more countries are offering interactive dissemination tools enabling users to obtain customisable outputs such as tabulations, maps, etc. Finally, sample data at the unit record level (more commonly referred to as microdata samples) are also provided by some countries for permitted research purposes. In such cases, the sample should be carefully drawn to ensure an adequate level of representation while at the same time ensuring that anonymity is not compromised. Some countries have also adopted very creative techniques for data dissemination, cross-tabulations, infographics, story maps and visualization to allow users to better understand relationships and patterns within the data. NSOs should consider developing visualization products incorporating storytelling—a technique where narratives unfold as the user scrolls down a webpage—since it offers a compelling method to communicate key insights from censuses to non-expert audiences. The development of such dissemination products should be part of the planning process of the census.

3.324. The information stored in the census database allows fast and relatively inexpensive production of additional tables, maps and other output products. Traditionally, countries have often offered on-demand services to provide census information to users who require tables or other outputs that would not otherwise be produced, or aggregates not available, through other means. Now that electronic dissemination is the new standard, often including user-friendly interactive dissemination tools, customized tabulations and applications can often be extracted directly by end

users. In this case, the NSO should prepare in advance, and then implement, an authorization and a data security and disclosure control strategy, so that additional disclosure risks entailed by providing customized outputs are accounted for.

c. Metadata to aid in the interpretation of the results

3.325. See section on metadata (Section B6 below).

d. Legal and ethical consideration, confidentiality and privacy measures

3.326. See section on confidentiality and privacy (Section B5 below).

e. Assessing the required technology to meet user needs

3.327. Given the widespread availability and use of technology for easier production and access to census products, NSOs should evaluate which ones are best suited to the needs of their data users taking into account budgetary and human resource constraints. These technologies include use of GIS and thematic map generators, new sophisticated database systems and, interactive web access, including user-customised and user-friendly table builders. This may be complemented by APIs to enable standardised automated data retrieval for heavy users. A more detailed discussion on the choice of technologies to support the conduct of the census more generally is set out at Section XV in Part Two.)

3.328. Interactive online tools are becoming the main dissemination channel in more and more countries. A high level of user customisation and automation of output retrieval may also increase efficiency, as it reduces the need for staff-supported on-demand services. On the other hand, very flexible tools that serve highly customized user requests by direct queries to the underlying microdata entail specific additional disclosure risks that need to be accounted for by dedicated disclosure control strategies (see paragraphs 3.482-3.484). When developing websites or web-based applications aimed at disseminating the census data, it is recommended to ensure inclusivity and accessibility for all users by considering the adoption of the Web Content Accessibility Guidelines (WCAG)[107], an internationally recognized ISO standard. Following WCAG guidelines will enhance the usability and navigability of the websites for individuals with disabilities, ensuring equal access to information.

f. Dissemination policy

3.329. When planning the dissemination programme, a dissemination policy should be established as well. This policy should cover issues such as ways of marketing the census products, which in most cases means mainly how to inform a wide range of potential users about the availability of the products. Where applicable, a clear pricing policy should also be determined in countries where not all outputs are made available free of charge, and a decision made as to the conditions under which external distributors are permitted to further disseminate, and charge for, census data. The dissemination policy should be transparent and cover relevant issues connected with the protection of the confidentiality and privacy of personal data, and the broad measures that will be used for each of the different products. Disseminating census data (other than microdata) under a clear,

[107] https://www.w3.org/WAI/standards-guidelines/wcag/

open license such as the Creative Commons Attribution 4.0 International (CC BY 4.0)[108] is recommended to ensure that users have a precise understanding of the terms of use, including redistribution and transformation rights. CC-BY 4.0, for instance, mandates attribution, which is necessary for acknowledging the original data source and maintaining transparency. Moreover, it requires users to indicate any modifications or derivatives, preventing misrepresentation. However, if the government has already established an open data license, it should be utilized. This approach promotes responsible and ethical data practices.

g. Quality assurance

3.330. Quality refers primarily to user needs and satisfaction. Even if data are accurate, they do not have sufficient quality if they are produced too late to be useful, or cannot be easily accessed, or conflict with other credible data. Therefore, quality is increasingly approached as a multi-dimensional concept. As is discussed in more depth at paragraphs the output of any statistical exercise should possess the following attributes: accuracy, relevance, reliability, timeliness, punctuality, accessibility, clarity, coherence, comparability and metadata.[109] Management of quality in census dissemination is driven by concerns to deliver output that adhere to each of these dimensions within agreed cost constraints. (See paragraphs 24.8-24.17 for more details).

h. Budget and human resources

3.331. Two obvious key elements (usually constraints) in the development of strategies for census data dissemination are the budget that can be allocated and the availability of human resources. With the high relevance of new technologies in all the census stages, and in particular for data dissemination, this is a factor that needs to be carefully analysed when deciding about the specific strategy of census data dissemination. For instance, interactive user--customised dissemination tools may increase efficiency (see sub-section (e) above). The alternative to the NSO's recruitment of human resources may be contracting out some of the dissemination activities, in particular those connected with the development of more sophisticated systems. However, this solution needs to be carefully considered (see paragraphs 2.240-2.255), and it is extremely important to ensure that the contractor is committed to the census project until its very end, and that at least some of the new abilities remain in the organization for further use.

i. The need for continuous improvement along with a feedback mechanism for dissemination products

3.332. See the section on Dissemination at paragraphs 24.81-24.85.

---

[108] The Creative Commons Attribution 4.0 International (CC BY 4.0) license is a copyright license that allows others to distribute, remix, adapt, build upon your work, and commercially exploit it, as long as they give you credit. This is the most permissive license offered by Creative Commons.

[109] United Nations Statistics Division (2012): *Guidelines for the template for a generic national quality assurance framework (NQAF)* – available at http://unstats.un.org/unsd/dnss/docs-nqaf/GUIDELINES%208%20Feb%202012.pdf.

j. Alignment of census outputs with international standards

3.333. See Section C below on Census data dissemination: products and services.

k. Census release calendar

3.334. To help communicate the timeline for dissemination, a calendar with the expected release dates for the preliminary results, final results, as well as the timing of the release of the various census products should be published. See paragraph 3.319 above and paragraphs 3.399-3.400 for discussion on the benefits of a release schedule).

2. Tabulation programme

3.335. In most countries, the tabulation programme represents a compromise between the full range of tabulations desired by users and the limits imposed by practical circumstances. To ensure that this compromise is made transparently and efficiently it is important that, as previously emphasised, planning the census dissemination task is started at the earliest stage of the census development cycle by a round of user consultations. Once the census-testing programme has identified a practicable range of data items to be included in the questionnaire, data users should again be consulted on the specific cross-tabulations required and the relative priority for their production. It is essential that the programme be outlined sufficiently early so that the procedures and costs involved are investigated thoroughly before final decisions are reached. The type of questionnaire and the method of enumeration may limit the kinds and amounts of data that it is possible to collect. Publication time and costs, and the data-processing resources available, will determine the number and complexity of the tabulations that can reasonably be produced. This will enable prospective census data users to make firm plans and the census data processing staff to complete all systems analysis, programming and testing work in a timely manner.

3.336. The tabulations presented on the website of the United Nations Statistics Division are those fulfilling the most essential, or generally required, information[110]. The databases of census information can be used throughout the intercensal period to address the needs of specialist users for whom these tabulations are not adequate.

3.337. It is important to plan the tabulation programme in such a way that final results can be issued within a reasonable period of time after the enumeration and before the information has become out of date for current needs. It is desirable that the details of the tables be prepared, and the order of their preparation be decided early in the planning so that the processing of the data is not delayed.

3. Geospatial information for analysis and dissemination

3.338. **Geospatial** information can unlock the full potential of census data and facilitate its use for territorial analysis, to describe and investigate phenomena, such as demographic trends and spatial inequalities in accessing services (such as schools, health facilities, etc.), that unfold across time and space. A more general discussion on the use of geospatial information, not only for analysis and dissemination but in other stages of the census process, is presented at Section IV in Part Two.

[110] Link to UNSD website containing the tabulation shells.

(a) Dissemination geography

3.339. Maps, which are now commonly found in the form of digital products, play an increasingly important role in the dissemination phase of the census. Statistics compiled from census data can be geographically referenced and provide for methods of analysing the geographic characteristics of those statistics. Maps may then be used effectively to relate statistical data to the geographic area to which the census results refer. This makes the statistics easier to understand and more readily usable by both expert users and the general public.

3.340. If a complete digital census geographic database has been created, then statistical databases for administrative or statistical units can be produced through aggregation. For the countries that do not use digital techniques for the production of EA[111] maps, options still exist to develop a digital geo-referenced census database at this stage for producing publication-quality maps to accompany census reports, for distribution to outside users who want to analyse census data spatially, or for internal applications. This database can be compiled for a suitable level of the administrative hierarchy or for other aggregated statistical regions. At that level of aggregation, the resources required for producing a digital database are much less than those necessary for a complete digital EA map database.

(i) Linking collection geography to dissemination geography

3.341. An essential feature of the census is its diversity in terms of the geographic level at which data can be disseminated. This is due to the ability of the census to produce statistics that can be disseminated at the lowest geographic level (small area), through a geographic hierarchy up to the country level and by geocoding at the lowest level of collection to support the production of grid-based outputs. Consequently, one of the earliest decisions in census planning relates to the administrative and geographic areas for which census data on diverse socioeconomic characteristics of the population will be reported and disseminated in order to satisfy the needs of the various data users, taking into consideration statistical confidentiality, in particular where non-nested small areas (such as smallest administrative unit and grid) are disseminated.

3.342. In addition to administrative units, most countries will have a number of other sets of areas that are used for different purposes and for which census data will need to be compiled.[112] Such areas, which have special uses, include health regions, electoral districts, urban agglomeration or metropolitan areas, grids, and utility zones (such as water or electricity supply districts). It should be noted that some of these areas may not fit perfectly into the administrative hierarchy of the country. It is important, therefore, that to the extent possible these reporting units are taken into account when designing EAs in order to facilitate generation of census data for such geographies. This draws attention to the fact that when delineating the geographical collection units (EAs), it is essential that dissemination geography and disclosure risks due to overlapping geographies are kept in mind.

---

[111] The Enumeration Area or District is only relevant in a traditional or combined census methodology. Where data is derived wholly from administrative sources in a register-based census, such areas do not have any meaning and small area geography must be determined by other boundaries.

[112] *Handbook on Geospatial Infrastructure in Support of Census Activities*, Studies in Methods, No. 104 (United Nations publication, Sales No. E.09.XVII.8).

3.343. Two somewhat different methods are available to provide the census with a flexible capability for generating tabulations in terms of a wide variety of geographic aggregations, including those needed for public and private sector data uses at the local level. The first method simply extends the traditional hierarchical system for coding all major and minor civil divisions so as to cover at the lowest level the EA (sometimes referred to as the "enumeration district"). The second method, which at greater cost permits finer geographic specificity, is usually based on some coordinate or grid system, such as latitude and longitude. This method is often referred to as a "geocoding system". Particularly in the absence of a comprehensive system of street names, numbers or geocoded coordinates of housing and building units, the first method, which uses the enumeration area as the key unit to produce small area data, is to be preferred.

3.344. Where buildings or housing units have been geocoded, these geocodes can be used to directly allocate each household to the correct EA. The fact that census data are available by EA provides for considerable flexibility which can be of much value given that the geographic divisions used by various branches of the administration or by other data users do not always coincide and may therefore require different regroupings. Moreover, when changes are planned in administrative boundaries, tabulation of census data by the planned new entities can also be facilitated through the EA approach. However, if these changes cross EA boundaries and it is decided to try to retabulate the census according to the new boundaries, very complex recoding of individual records may be involved. As an alternative, statistical concordances, showing the quantitative relationship between the previous and current classifications, could be used.

(ii) Uses of small geographic data

3.345. Censuses provide data from the highest to the lowest geographic levels of aggregation. Tabulations from census results yield relevant statistics for any reasonable combination of characteristics for the country as a whole, regions or provinces, down to small areas such as localities, and EAs and geographic grids. This important feature of the census makes the data amenable to the development of estimates of variables of interest for small and local areas in two major ways: directly from the production of tables from the microlevel data for the required characteristics; and indirectly from applying estimation techniques by combining other sources, such as sample surveys and administrative data to the population and housing census results.

3.346. Data for small areas enable the user to obtain statistical information about any number of local areas of interest, in addition to showing variations among small areas in individual parts of the country. The increased capabilities of data processing systems greatly facilitates the utilization of census results for analysing the information for small areas, limited only by issues of collection design and confidentiality and statistical disclosure when cell entries in cross-tabulations become very small. For example, the analysis of whether or not programmes have improved educational attainment for women and girls at a regional level may be carried out by analysing data from the smallest administrative units so as to observe local variations and produce more accurate assessments of cause and effect.

3.347. Implementation of various national social and economic development programmes is a function of the state, province or lower levels of government in many countries. Results of population and housing censuses are useful for planning and monitoring development at the local area, small town level or small area. Small-area data are also important for private businesses in developing their distribution and marketing strategies. For example, information on housing demand from the population and housing census may be used by local authorities, local real estate companies, building and housing development contractors, and manufacturers of construction materials, among others.

3.348. Census data have been traditionally aggregated by various types of administrative units (for example, towns, villages, provinces and electoral units). In addition, other types of small areas are sometimes used in the census that are essentially statistical in nature (for example, census tracts and grid squares that do not change from census to census, and very small units such as city blocks or block faces). There have also been increasing demands for small-area data that cut across the local administrative boundaries. Population and housing censuses provide a powerful tool for assessing the impact of population on the environment, for example on drainage basins and on water resource management systems. The spatial units for such a study may combine a group of local administrative areas. In this situation the availability of census databases with mapping capability (see paragraphs 3.436–3.439) is of great importance.

3.349. Tabulations for small areas may be prepared on the basis of the resident (*de jure*) population of each area or on the basis of the population present (*de facto*) in each area at the time of the census. Tabulations relating to the resident population are produced for the apportionment of representation in legislative bodies, the measurement of internal and international migration, the computation of measures of fertility and mortality by place of residence, and the planning and administration of such services as schools and housing, which have relevance only to the resident population. Tabulations based on the population present in the area at the time of the census are useful where this population is considerably larger than the resident population and thus raises the demand for products and services above the level required by the resident population alone. The combined population and housing census may also be used to make comparisons of resident and daytime populations in specific localities, if an item on place of work is included in the population census. It is therefore important that users express their needs for particular data disseminated in a given format, based on the usual residence or place of enumeration, at an early stage of census preparations.

3.350. Population and housing census plays an essential role in the economic, social and environmental components of the national statistical system and also serves as a sampling frame for sample surveys. Another significant way in which the census results complement survey statistics is in small-area estimation, whereby models constructed from survey data are applied to census results for any specified geographic area. This estimation approach may be used for generating such indicators as employment, poverty and other economic indicators, for which measurement is required at the local area level.[113] The application of small-area estimation techniques to poverty measurement and mapping is an important extension of the use of census results.

(b) Spatial analysis

3.351. Spatial analysis on census data is key to presenting census data in an understandable and accessible way as decision makers are able to relate more easily to information presented in maps than with data tables. For example, the SDGs vision of leaving no one and no place left behind requires, at a minimum, analyses and presentation of trends at the highest spatial resolution possible – to show existing (spatial) inequalities among places and people. Census data, because their high resolution is able to accurately showcase prevailing inequalities and, in turn, inform decisions and actions to ensure equitable development. In addition, census data are crucial to

[113] *Measuring the Economically Active in Population Censuses: A Handbook*, United Nations publication, Sales No. E.09.XVII.7.

producing data disaggregated by location (such as city, towns and suburbs, villages, and rural areas), sex, income groups, and persons with disabilities (among others) as required for monitoring global trends. For example, understanding the share of different groups of urban populations who are within a proximal distance of access to public transport (SDG 11.2.1), open spaces (SDG 11.7.1), or the share of rural populations with access to all-season roads (SDG 9.1.1) or the computation of ratio of land consumption rate to population growth rate (SDG 11.3.1) requires a clear analysis of census data to depict location of populations across the urban-rural continuum.

4. Mode of dissemination of outputs

3.352. As has already been mentioned, a census is not complete until the information collected is made available to potential users in a format suited to their needs (paragraph 3.313). Consequently, meeting the needs of data users means that the NSO should not only provide data products to the users, but should also provide them in formats that are suitable to the needs of the users. The information in the products may be included in published tables and reports for general distribution, produced as tables in unpublished form for limited distribution or stored in a database and supplied upon request, or disseminated online either as static or interactive products.

3.353. It should be noted, however, that regardless of mode, all dissemination is subject to issues of (a) quality assurance; (b) possible disclosure of information about identifiable respondents; and (c) copyright and ownership. In addition, the issue of cost recovery has become important to many statistical organizations. Each medium of dissemination has its advantages and limitations, and the choice of one or more of them depends on the context, and on the intended categories of users. In most instances, these methods complement each other and can provide effective ways to reach out to the public and private sectors.

3.354. When data are provided in electronic form, special attention should be given to providing users with easy means of data and metadata retrieval. The options for obtaining the relevant metadata and the data should be accessible in standard and contemporary formats.

(a) Publication of printed tables and reports

3.355. Nearly all countries now disseminate their census results online, and printed publications have become a secondary choice for the dissemination of the main census results.

3.356. When printed publications are still used, the choice of how the actual printing is to be done entails in fact a trade-off involving quality, cost and speed. The best results can usually be obtained by sending the documents in computer-readable format to a professional printing plant[114]. This will allow high-quality typesetting and the use of supporting colours. Alternatively, master printouts can be made in the NSO and sent to the printer for cheaper duplication or offset printing. There are also affordable high-speed printing systems that can be directly controlled by the computers in the census office. Publications can also be distributed in electronic format (pdfs) as part of the NSO commitment to reducing paper waste.

[114] In some countries the printing of census reports is often managed by an official government publications agency that is responsible for maintaining quality standards

3.357. As noted at paragraph 3.334, to ensure timely dissemination, target dates for online and printed publications should be established well in advance, and processing and reproduction programmes should be planned accordingly. In addition to traditional printing methods, NSO may wish to explore alternative reproduction techniques that offer speed, cost-effectiveness, and high readability. These may include digital printing, online publishing platforms, or other modern methods that can efficiently produce and distribute census materials.

(b) Online dissemination

3.358. For most users, online is now the preferred channel of to access census outputs. NSOs are increasingly using cloud-based solutions for the storage, security, management, and dissemination of census data to users. (See also paragraph 3.425 summarising the advantages and challenges of using cloud-based data storage facilities.) Online dissemination of all kinds of information, including statistical information, has increased with innovative formats for displaying census data. The advantages of online dissemination are found primarily in terms of speed, flexibility and cost, and making results accessible to a wide range of data users. Online census dissemination should feature a modern interface and API as well as good search and filtering abilities.

3.359. Online dissemination of data was common well before the Internet gained prominence. The same website could be used for both internal and broad community communication, with the granting of access rights in certain areas to privileged users only. Security measures, including passwords and callback procedures, can be used to exclude unauthorized users from reaching these areas. This is however risky, since resourceful hackers may find their way around the barriers and gain access to confidential information. The NSO's website is probably the first place where Internet-connected users would look for census information. It is, however, recommended It is also recommended that a powerful firewall constitute a security layer between the website that is visible to the public and the working network of the NSO. Firewalls are hardware or software security systems that limit the exposure of a computer or network to malicious infiltration from an external location. Websites of public administrations are under constant attack from hackers and very sophisticated security measures should be adopted when disseminating interactive census products on the Internet. Cyber security, despite being an issue of a technical nature, has to be mandated, demanded and resourced by the highest levels of management of the census office.

3.360. Flexible table builder software provides the potential for users to access census microdata and create their own multivariate output tables at different levels of geography. The software runs statistical disclosure control while the request is being processed to ensure protection of individual identities in detailed statistics, making sure that no individual can be picked out in the tables. It allows NSOs to release billions of anonymised census statistics far more quickly than ever before. A key design choice with profound implications for data security and confidentiality is whether such flexible table builders query the underlying microdata in the background or utilize a set of very detailed pre-tabulated queries. Direct microdata queries may lead to additional disclosure risks even with automatic disclosure control methods running in the software.

3.361. A website can be used not only to make information available as soon as it has been cleared, but also for other forms of communication with users. Possibilities include online ordering of any printed publications and one or more receiving areas for questions that would be answered later through the same medium by appropriate specialists. One such area could be a census forum or "chat room".

3.362. Internet websites may support "door" or "gateway" applications that allow users to run outside programs on the computer on which the Internet web server operates. Interactive access to

census outputs can be offered to most types of databases and census products, including reports, publications, tables, maps and graphs. For example, there may be a database of aggregated census data for small areas or a microdata database that users can access in this way. When the required data are not readily available, users could run an on-the-spot query to obtain and retrieve results that satisfy their needs. This can be done by offering to Internet users census microdata samples and an interactive tabulation system. Users can then select records from these data sets that satisfy certain parameters and compute statistical information, such as two-dimensional cross-tabulations of either original or recoded variables. Program execution by users, however, raises important questions of cost, efficiency and confidentiality, which have to be resolved. For reasons of efficiency, it is recommended that information that is provided, or likely to be heavily requested by users accessing the census website, be made available in a static format, which is faster to download. Letting the user run data extraction on online databases, which would be a dynamic way of accessing the census information, is more resource consuming and should be the second choice for those users needing more detailed data than those available through static pages.

3.363. Other outlets such as social media are useful in disseminating census information targeted at different sectors of the population. More generalist media, such as the radio, television programmes, newspapers and press conferences, offer the possibility of reaching out to sectors of the population not otherwise reachable.

3.364. A hybrid solution for data dissemination that appears to combine the advantages of several approaches is one whereby the NSO makes basic data available to users online, while additional information may be provided by some other online protocols, such as file transfer protocol sites. This will usually take the form of a package that contains basic data, metadata and data browser software. The basic data may contain existing time series, report files and the like, as well as national and regional maps that can be used to generate thematic maps with various indicators. Maps made available to general users need not ensure the same geographic detail as maps used by enumerators. Lighter versions of maps at any subnational level may be provided to the general public, and more sophisticated and detailed ones to those fewer users who would actually need an increased level of detail. It is thus important that the NSO website should specify the instructions on how to contact those responsible for providing any special dissemination services.

5. Confidentiality and privacy

3.365. According to principle 6 of the Fundamental Principles of Official Statistics, “Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes”.[115] Maintaining data confidentiality is an indispensable element for retaining the trust of respondents. If respondents believe or perceive that the NSO does not protect the confidentiality of their data, they are less likely to cooperate or provide accurate responses to the census questionnaire. This in turn inevitably affects the quality of the statistics.

3.366. All the information stored in the census database allows the production of tables both for very small areas (such as enumeration areas or villages) and for all individual units in these areas. Therefore, two key considerations guide the construction of a census database: technical aspects and ensuring confidentiality. Protecting individual privacy is paramount, and it should influence the

[115] http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

design of both data collection and data processing programs. Accordingly, microdata, such as name and local address, or any other characteristics that permit the identification of individual respondents, must be removed from the database or encrypted.

3.367. The same care must be taken if a transcription of information from original questionnaires (that is to say, from a representative sample) is needed for use by qualified agencies and research institutes engaged in special studies beyond the purview of the regular census programme. Such needs have sharply decreased with the almost universal use of computer technology. However, when such a procedure is possible under the census law, individual privacy should be ensured in all cases.

3.368. The ever-increasing demand from users for more data, especially microdata and at lower geographic levels, and also with more technological advancement for data linkage, particularly over the Internet, has created more challenges for managing data confidentiality. As a result, NSOs should examine the data and make modifications, when necessary, prior to its dissemination, the objective of which is to prevent identification of individual respondents and of intentional or inadvertent disclosure of their personal information. This is particularly the case when microdata is disseminated and when data are linked to location, such as with the use of GIS.

3.369. Statistical confidentiality methods range from traditional methods such as cell suppression, rounding, Barnardisation[116] or recoding, to more modern methods based on the introduction of random noise. Depending on the method, the implementation complexity also ranges from a few simple processing steps to elaborate statistical applications for entire databases[117]. The approaches used to limit disclosure are tailored according to the type of data and the product to be disseminated.

3.370. Different techniques are also employed depending upon the type of data product to be released (for example. microdata files, tables, interactive data explorers, maps).[118] For example, disclosure methods for **microdata** include (i) suppression of direct identifiers, (ii) recoding or local suppression for quasi-identifiers, (iii) noise addition, and (iv) the suppression of sensitive variables and/or outlying values. More recently, differential privacy and synthetic data have been proposed as safer approaches.

3.371. For **tabular/query data**, more traditional methods aiming to mask confidential information (e.g. cell suppression, recoding) have been widely used over the past decades. However, these methods have meanwhile been found to suffer from additional loss of data security and utility in modern dissemination scenarios with very detailed tables particularly when combined with various other data products. Emerging alternatives are predominantly based on adding some uncertainty to the outputs in a consistent manner (noise injection).

3.372. **Interactive data explorers** are often based on flexible rules for users to submit custom table queries (table builders), where additional disclosure risks are present in implementations where

[116] A method of statistical disclosure control for tables of counts. It involves adding +1, 0 or -1 to some or all of the internal non-zero cells in a table.
[117] Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Schulte Nordholt, E., Spicer, K., de Wolf, P. (2012). Statistical Disclosure Control. Germany: Wiley.
[118] de Wolf, V.A. (2003). Issues in accessing and sharing confidential survey and social science data. *Data Science Journal* 2(17): 66–74.

highly flexible queries are evaluated directly from the microdata (as opposed to querying a curated fixed set of detailed tabulations stored in the background).

3.373. Finally, **maps** are becoming a standard data product from censuses, more and more often with interactive functions. Specific additional disclosure risks may emerge from very detailed geographical outputs[119] when these are not nested (geographic differencing – for example between very small administrative or enumeration units and grids).

3.374. The disclosure of confidential information in the context of the dissemination of census data integrated with geospatial information requires careful attention. Geocoded census data are becoming available at increasingly finer resolutions, increasing the risk of data disclosure. Geospatial information adds a new dimension to data with which an individual statistical unit can be more easily identified in combination with other information.

3.375. When an interactive mapping application is used, the tool should be configured in a way that users are prevented from drilling down to the spatial resolution level at which unit record data might be disclosed. When microdata are released with geographic coordinates, geomasking techniques should be used to anonymize their precise location and prevent identity disclosure. Since aggregated data with spatial attributes provide coarsened location information of the respondents, their privacy can be protected by data suppression on the areas with low values. Geospatial products should be cross-checked with other dissemination products (such as tabular aggregates or micro datasets) before release so that they do not breach confidentiality on their own as well as in combination with other outputs.

6. Metadata

3.376. In order to assist data users to better understand and interpret the data, it is important that there is adequate documentation providing a complete and clear description of the production process, including data sources, concepts, definitions and methods used[120]. This information represents metadata that, it is recommended, should accompany all census products. Providing metadata promotes transparency and credibility of census results and facilitates one of the key dimensions of quality - interpretability. Additionally, dissemination of census products with accompanying metadata ensures harmonization and comparability of census data with other data sets. International metadata standards (such as the DDI Codebook standard for microdata or the ISO19139 for geographic datasets) provide a convenient solution to ensure completeness of the metadata and to make metadata machine readable (which can help make the data more discoverable, visible, and reusable). SDMX[121] (Statistical Data and Metadata eXchange) is an example of an exchange and data warehouse tool (see paragraph 3.433).

(a) Definition and content

3.377. Metadata comprise descriptive and structured information or documentation about data that informs users about the content, quality and condition of the data. In this context, metadata

[119] Where maps illustrate percentage distributions rather than absolute numbers, there are far fewer risks of disclosure.
[120] See Annex for an example of a metadata checklist (Census_Metadata_Checlklist.docx)
[121] SDMX stands for Statistical Data and Metadata eXchange. It's an international standard (ISO 17369) used to describe and exchange statistical data and metadata.

offer guidance on the proper usage or interpretation of data by providing information on the processes of production and describing the structure of data sets, thereby making it easier to retrieve, use or manage the data. Metadata constitute a standardized way of organizing data and can be categorized as follows: (a) reference metadata, which allow understanding and interpretation of the corresponding statistical data by describing the concepts, definitions, methodology and quality of data, production and dissemination processes, data access conditions, etc.; and (b) structural metadata, or "data about data", which provide information about the structure of the data set and act as identifiers and descriptors of the data, making it possible to properly identify, retrieve, browse and further process the data.

(b) Uses of metadata

3.378. Comprehensive and readily accessible metadata foster the comprehension of the statistical data being disseminated, public trust in the data, and informed decision-making. Detailed metadata shed light on underlying concepts, methods, and definitions, thereby fortifying the accurate interpretation of the census results. Rich metadata also bolster data assessment and discoverability, enabling users to navigate and locate pertinent information efficiently, and to evaluate the fitness of the data for their specific purposes. Moreover, by increasing transparency and credibility, meticulous metadata empower both individuals and electronic systems to reliably capture, interpret, and present statistical and geospatial information.

3.379. All tabulations should include the following metadata or references to where this information can be obtained:

- the relevant census questions and (at least some of) the reasons why they are included;
- conceptual definitions (census dictionary);
- geographic hierarchies used;
- changes since the previous census with regard to content, operational methods or geographic boundaries; and
- quality indicators such as coverage rates and item non-response.

Data files also should be accompanied with metadata, including names and codes for common variables, personal files and household files. If a long-form sample is used in the census, metadata should also provide information on the sampling variability of the results. When the census tabulations include suppressed data cells due to small numbers, the metadata should also include a methodological note on the rules and methods of suppression[122]. Metadata should be preserved for future reference. With the increased use of technology, properly designed metadata systems for machine-readability and web-based applications are recommended.

(c) Geospatial metadata

3.380. Geospatially integrating social and demographic data necessitates the inclusion of additional metadata to the statistical record, preferably in the form of x- and y-coordinates but other forms of geographic information, through a "geocode" such as an address or locality, can be used. Furthermore, additional metadata including details about the coordinate system, projection, the precision of the geocoded locations, as well as data confidentiality measures, such as any spatial

[122] The rules and methods of suppression should be described in a manner that protects against the possibility of disclosure by reversing the process and deriving the unsuppressed data.

aggregation or masking techniques used to protect sensitive locations, should be included to facilitate interoperability, appropriate use, and correct interpretation of the data.

3.381. Compared to statistical metadata, there is limited awareness and understanding in statistical organisations on the metadata associated with geospatial information and services. Geospatial information is in the form of various types, and this great variety as well as methodologies/ technologies involved in producing the data adds challenges for NSOs to standardise geospatial metadata and systematically manage it.

3.382. Given the wide scope of geospatial metadata, it is important to first investigate crucial metadata elements (for example, data type (point, line or polygon), time stamp, and coordinate system) needed for different stages of production and determine a core metadata set and standards (for example ISO 19115, ISO 19119, GeoDCAT)[123] to follow at the corporate level. After priorities are defined, a continuous improvement process could be put in place to gradually improve the scope covered by the metadata. It is important to have a corporate strategy in place to build a consistent metadata system in order to avoid compliance issues with other existing metadata systems. As for the statistical metadata, the geospatial metadata should be managed and updated continuously throughout the production process as the changes affect downstream tasks and influence the final outputs.

3.383. Alignment harmonisation of geospatial metadata concepts with those of statistical metadata in existing metadata systems is critical, and there has been an increasing effort to connect statistical metadata with geospatial metadata (for example technical specification expanded for geospatial metadata in SDMX 3.022).

7. Promotion of, and training on, uses of census data

3.384. The main purpose of a census is to collect, process and disseminate information that will be used as the basis of informed, evidence-based decision-making. The benefits of this approach to such decisions are not always apparent to users, especially in situations where other approaches may have been used in the past. It is therefore important to promote such uses of census results among users.

3.385. In other cases, users may be willing to use the information but require additional training to more fully understand the data. Such training may be usefully combined with training in statistical dissemination techniques or uses of more advanced data products. At a very basic level, some users may require guidance on such mundane issues as how to contact the NSO, or how to find the information they require within the systems of that office, or how to use the website and other electronic applications and tools. Appropriate training can be delivered via video tutorials, podcasts or e-learning platforms.

3.386. Whichever approach is taken to enhancing promotion of, and training in, the use of statistical data, a number of strategic issues need to be addressed. These include:

[123] GeoDCAT is an extension of DCAT (Data Catalog Vocabulary) specifically designed for describing geospatial datasets, dataset series, and services. It provides a standardized way to represent metadata about geospatial resources, making them easier to discover, access, and use.

a. Ensuring that the needs for training are identified early in the census planning process and that required funds are included in the census budget. In this regard it should be noted that in many cases the courses requested by users will be specific to those users; in such cases it may be desirable to request the user to provide funds to cover the marginal (or full) costs of the course.
b. The proposed courses or materials should be fully integrated into the overall census advocacy or training programme. It is essential that messages about the use of data fully reflect the messages given when initially advocating the census or seeking public cooperation with, and participation in, the collection phase.
c. If the training facility is itself promoted properly, it is highly likely that the demand for training will far outstrip the ability of the NSO to deliver it. In this case it will be necessary for the NSO to have prepared transparent strategies that (i) identify those areas in which the statistical office wishes to participate (for example, dealing with lifeline clients, and topics on which the statistical office has particular knowledge or expertise); (ii) establish partnerships with other bodies to provide training; (iii) use approaches other than classroom training to provide learning-at-a-distance opportunities (for example, e-learning); and (iv) have a pricing regime to cover costs where this is seen as necessary.

3.387. The list of target audiences and topics for such training must clearly be determined by the NSO. It should be noted, however, that basic training in the use and interpretation of the results of one census is a very strong method of advocating support for future censuses. It is thus recommended that countries consider development of a basic course in

(a) potential uses of census data;
(b) how to access census data;
(c) interpretation of census data at the broadest level, including the interpretation of its completeness and level of accuracy; and
(d) spatial analysis.

The target audience for such training should be, in the first instance, key decision makers in the political and administrative hierarchy of the country. It should be outlined that the uses of census data at the local or small areas level offer potential for better decision-making for the benefit of local communities; the spatial distribution of population by age and sex, for example, provides an ideal framework for local officials to address the most pressing issues of their constituents, such as location of schools, utilities, service delivery and so forth.

3.388. A second group of key importance are members of the mass media, such as press, radio and television journalists. A focus on training such personalities is important because they can carry the message to a wide audience. This will assist in the general raising of awareness in the population at large, as well as in generating an awareness of the census among the governmental, academic and business users who may not have contact with the NSO on a regular basis. Such training should be completely integrated with the overall public relations and advocacy work.

3.389. A third group to be targeted are schools - both students and teachers. In addition to the aim of creating more widespread awareness of the census for the purpose of maximising participation, a focus on training of teachers to use census data in the mathematics, social sciences and geography curricula creates awareness among children about using statistics in decision-making and allows them to develop numerical skills using real data (see paragraphs 3.506-3.507 below).

3.390. A fourth group to be targeted are geographers, with the aim of integrating census and survey data with GIS shapefiles in order to perform spatial statistical analysis. This training will enable specialists to better present statistics geographically.

3.391. A fifth group to be targeted is the research community and civil society organisations. Researchers will benefit from focused training on the application of various statistical techniques to census data to enhance data utilization. Civil society organizations will require briefings on key census findings relevant to their constituencies to support advocacy and program development.

3.392. A sixth group to be targeted are potential new users who have historically been less likely to engage with census data. A number of stakeholders are unaware of how census data can be used in their area of work to make evidence-based decisions.

C. Census data dissemination: products and services

1. Provisional and final results

3.393. Some countries release provisional results very soon after enumeration is completed. Subject to change once the full data-processing and verification operations have been completed, they nevertheless provide a general picture of population trends. Provisional census results may be processed manually or by computer. For reasons of efficiency and quality, the use of computers is always preferable. The ability to verify data quality during the enumeration phase with the help of validation programs, quick indicator reports, data consistency reports and tabulations, greatly increases the confidence with which provisional results can be announced. Provisional results will normally cover information only on total population by sex and by major division. The number of households and housing units may also be derived easily from this exercise. The provisional result of the census can be reported right after the completion of the enumeration by utilizing the summary of household lists without individual data processing. This can be possible as these summaries usually include the total population, households and housing unit in each major division.

3.394. An interim type of outputs are ‘preliminary results’. These are distinct from ‘provisional’ results in that they comprise tables of fully processed and quality assured data but are often in the form of simple one-dimensional counts that are published quickly before the more detailed multi-variate, topic-oriented, tabulations can be made available.

3.395. The final census results will be the output of the main tabulation programme. Tabulations may be based on all of the returns or on a sample. If some of the topics are collected on a sample basis only, proper weights will have to be applied in the tabulation stage to produce valid national estimates. In addition, the NSO should be prepared to facilitate the production of customized outputs requested by researchers and other users (see paragraphs 3.489–3.495).

3.396. To ensure data accuracy and minimize potential confusion, users should be made aware of the potential differences between provisional and final results. The implications of using provisional population counts should be clearly outlined, and users should be advised to exercise caution when relying on these figures. Quality assurance processes should be implemented to minimize discrepancies between provisional and final results. Once final results are released, users should be encouraged to discontinue the use of provisional data to avoid ambiguity and ensure consistency in analysis. To further safeguard data integrity, the NSO should consider removing provisional data from its website once final results are available.

3.397. The final census results should be published as soon as possible. Countries may aim to publish the basic, essential results within one year of enumeration. The use of technology may reduce the time between the release of the provisional and final results, which may over time render provisional results obsolete. The dissemination of the final census results should be part of a comprehensive dissemination strategy and plan. A description and planned schedule of upcoming

releases of final results and products should be made public early in the data processing phase to maintain interest by the public in the census (see also release calendar, paragraphs 3.319 and 3.334). The releases can be staggered, from simple, descriptive one-page summary fact sheets covering a country's major geographic divisions initially, to more comprehensive tabulations and descriptive reports later on.

2. Census reports

(a) Basic reports

3.398. Every effort should be made to publish the principal results of a population census (such as those on age, sex and geographic distribution of the population) and of a housing census (such as a geographic distribution of sets of living quarters, households and population by type of living quarters) as soon as possible after the enumeration, otherwise their usefulness and the extent of their interest to the public will be diminished. With technological advancements, the time required for processing and tabulating results has been significantly reduced. As a result, collection restrictions, in terms of cost and accuracy of the data, have a greater relative weight in determining the number and complexity of the tabulations that can be produced and disseminated.

3.399. As noted at paragraph 3.397 above, preparing a provisional release schedule/calendar for census outputs and making it public can provide the NSO with the opportunity to influence the reception of statistical releases. It ensures that stakeholders know when to expect particular outputs of statistical releases. It ensures that stakeholders know when to expect particular outputs, which may impact their plans for making decisions, producing complementary publications, or updating information they hold. It can also ensure that the media is prepared to publish news on specific topics, and advance press releases for each topic can help messages land through the media with the necessary context and sensitivity. Releasing univariate outputs at a steady pace, instead of in batches, can improve the ability of NSOs to guide their audiences through the key news and information in census outputs. This can help ensure that the most important stories are given appropriate prominence.

3.400. Consultation and engagement with users should be considered when determining the timing and prioritization of different outputs in the release schedule. NSOs should, however, be cautious of optimism bias when assembling a release schedule. They should carefully consider how and when to communicate the release schedule to users before making it public. There is a significant risk of failing to meet users' expectations relating to the timeliness or quality of census outputs, which can impact trust. In some cases, it may be necessary to deviate from a planned release schedule to respond to emerging priority user needs for information.

3.401. A successful approach could be to announce the broad periods in which different types of releases will be released, such as provisional results, preliminary results, univariate topic summaries, and multivariate data with detailed analysis. Then, release detailed plans with specific dates for each phase as early as possible on a rolling basis as those dates become finalized.

3.402. NSOs should consider which tabulations and analysis should be produced and in which format, such as paper reports or online. "Flexible table builder" technologies can be used to complement fixed multivariate tabulations and analyses that may be produced for priority user needs. Subject to relevant statistical disclosure control measures, flexible table builders enable users to create cross-tabulations across variables and geographies of interest. This can reduce the resource and time requirements for meeting a wider range of needs.

3.403. As with communications during the census operation, accessibility requirements should be considered at the outputs phase, ensuring that outputs can be accessed by a variety of audiences. This could entail producing outputs in a variety of formats.

3.404. The population and housing census tabulations presented and illustrated at the website of the United Nations World Population and Housing Census Programme are intended to provide, in tabular form, the most important census information needed as a basis for programmes of economic and social development and to be used for research purposes. They do not in any way represent all of the tabulations that a given country may publish and certainly not all of the tabulations that may eventually be prepared for special purposes. Nor do these tabulations take into account the form in which information may be entered into a database, which may be more detailed than that required for these illustrative census tabulations.

3.405. Three categories of tabulations are described below: (a) basic or essential, (b) recommended, and (c) optimum tabulations.

(i) Basic or essential tabulations

3.406. These are tabulations that are deemed of top priority for production by countries. They are also regarded as essential for countries in difficult circumstances, such as those that have emerged from a conflict or those that have not carried out a census in a long time, in terms of providing minimum statistics to meet their basic data needs.

(ii) Recommended tabulations

3.407. Recommended tabulations are those that are considered adequate for meeting the essential data needs for evidence-based planning, monitoring and implementation of national policies because of their perceived relevance at both the national and the international levels. These tabulations are also designed with the potential for producing statistics at the lowest geographic level and are expected to be produced by each country.

3.408. The recommended set of tabulations also includes the basic or essential tabulations discussed above. Schematic presentations of all tabulations are presented online at the website of the United Nations World Population and Housing Census Programme[124].

3.409. Associated with the recommended tabulations are the core topics that go into their production. Core topics are the primary variables for the recommended tabulations. For a more detailed discussion of these topics, please refer to the paragraph references listed under the 'List of Topics' section in both Chapter 1 and Chapter 2 of Part Five.

3.410. As stated in paragraph 5.1, the aim of the recommended tabulations is to permit national and international comparability of data due to use of common concepts and definitions of the core topics. For each of the recommended tabulations, the core topics that it represents should be listed as part of the metadata. Other metadata that are presented for each of the recommended tabulations should include:

[124] Link to UNSD website containing the tabulation shells.

l. the source of statistics, that is to say, whether from a field enumeration, administrative registers, a survey or a combination of sources;
m. the type of population count, that is to say, whether a *de jure* or *de facto* population or a combination of these; and
n. the definition of urban and rural areas used.

(iii) Optimum tabulations

3.411. The optimum set of tabulations includes the basic or essential and the recommended tabulations discussed above, as well as additional tabulations, and is designed to meet the needs of most of the users at the national and the international levels. This set can be viewed as being equivalent to the complete set of tabulations that could be generated from a population and housing census.

3.412. To avoid excessively large or sparse census tabulations, some countries might opt for a more restricted geographic classification for data presentation, rather than including all available geographies. This could involve aggregating data at a higher geographic level or focusing on specific areas of interest. For example, basic facilities such as piped water or electricity may be almost completely lacking for large areas of some countries. Under these circumstances, tabulation of the relevant data for small geographic areas would not be appropriate. The geographic classification to be utilized needs to be carefully considered, taking into account the type of information being tabulated, its probable frequency distribution and the uses to which the data are likely to be put. Privacy and confidentiality of individuals and households must always be protected (paragraphs 3.365–3.375).

3.413. As well as collecting data on core topics, some countries may also collect data on additional topics (non-core topics) in the census questionnaire to address specific concerns, for example, whether or not the birth of an individual is registered, the age at which a woman first marries, or an individual's ethnicity or religion. In other cases, detailed tabulations for specific population sub-groups may be required for use in planning or evaluation of programmes. Tabulations for the non-core topics may be done after the basic tabulations are completed, but priorities should be based on user needs. Consequently, consultations with user groups at both the national and local levels may be helpful in determining the most suitable tabulation plan and method of dissemination.

(b) Thematic statistical or analytical reports

3.414. Many countries prepare different types of thematic or analytical reports. These reports should be planned and scheduled during the preparatory phase and published according to the release calendar in order to avoid outdated reports. The reports may range from volumes presenting extensive and detailed statistical tabulations, particularly cross-tabulations, to more analytical reports that combine tabular materials with some interpretative or analytical text. This latter group of reports might include, for example, volumes of regional analysis on such subjects as population or housing conditions of urban areas, major metropolitan areas or big cities, and regional distributions; locality reports on infrastructure; and comparisons of key social indicators such as education, living arrangements, housing conditions, sanitation, disability and economic activities. Other such reports might include community profile analysis of, for example, indigenous populations, and profiles of specific population groups, such as families, children, youths, persons with disabilities, forcibly displaced and older persons. Reports on population growth and distribution that examine changes in the demographic characteristics of the country's population with breakdowns by two or three levels of administrative areas would be very useful. Such reports might focus on the growth, location and mobility of the population at the national and regional levels, and administrative areas. It is important that appropriate language is used to correspond to the target audience (lay or expert

audience) for each thematic report. It is recommended that multidisciplinary task teams be established, including line ministries and agencies, for the preparation of thematic and analytical reports in line with agreed guidelines. Partnerships and external cooperation with academic institutions and other specialists in subject matter, which can facilitate such work and strengthen collaborations, may be sought whenever possible.

3.415. Thematic and analytical reports must be based on user needs and respond to a country's specific development needs and emerging issues. These reports can also be used to show time series and trend analyses of socioeconomic and demographic indicators and may combine census data with other data sources to provide a more comprehensive and current outlook.

(c) Methodological reports

3.416. Other published reports may include: an exposition of the census design, including, for example, the data collection methodology and encompassing, if applicable, sampling design; and a census quality and/or evaluation report, which may include estimates of census coverage and the methodology used for their preparation.

3.417. Many countries publish a census dictionary, which contains comprehensive definitions of terms and concepts and detailed classifications used to present census outputs. Some countries also publish geographic classifications and codes and the definitions of areas used in the census and their relationships with the administrative areas. Explanations of user-defined areas for specific census tabulations and the type of format available (printed or electronic) may be provided.

(d) Administrative report

3.418. One of the most important reports in the publication programme is the administrative or procedural report, which is a record of the entire census undertaking, including problems encountered and their solutions (see also paragraphs 3.576–3.580 on systematic recording and documentation of the census experience). The report may include the following topics:

- a brief history of the census in the country and the legal basis for conducting the census
- budget requirements, expenditure, and control
- source and allotment of funding
- census committees and their activities
- stakeholder management
- census organization and personnel structure
- staff management
- quality control procedures
- census calendar
- census cartographic work
- development and design of the questionnaires
- enumeration methodology
- census promotion, publicity, and communication;
- field organization;
- manual editing and coding;
- data-processing development and organization
- data capture
- computer editing and imputation procedure
- capital, equipment and infrastructure management
- computer hardware and software used
- census evaluation

- publication and data dissemination programme
- and archiving.

The census administrative report is very useful both for the users and for the NSO's own staff. The administrative report is an essential product for the planning of future censuses (see also paragraphs 3.578–3.580 and paragraph 3.575 about the value of having documented evidence of the how the census was undertaken.

3.419. With developments in information technology, the census data files and publications have become increasingly available in electronic formats. A description of the procedure in the development of these data files may also be included in such a procedural report. Consideration of a separate volume of the procedural report for the processing and dissemination phases may be considered to enable a review of the planning and field operations phases to be prepared and made available immediately after the census enumeration is competed.

3. Databases

3.420. To maximize the utility and lifespan of census data, and as a complement to the standard production of tables, NSOs are encouraged to store the census data in various computerized database forms to better satisfy the full range of needs of internal and external data users. Census databases assist data users by providing easy access to a wide range of census data.

3.421. The establishment of such databases can enhance the dissemination of the census results as well as increase their usefulness by combining census data together with related information from other demographic enquiries in a common format. (An important special case is bringing together the data from prior censuses into a single database.) In addition, such databases can improve the coherence of the input and output processing systems.

3.422. Given the wide-ranging needs of users, there is no preferred one-size-fits-all approach to setting up a census or population database. Decisions range from the type of data to provide (microdata, sample microdata, aggregated data or a combination of these options) to whether to try and incorporate the new census results into existing database structures from previous census rounds[125] or establish new ones. In the latter case, the issue to be considered is whether the new database(s) will be exclusively in the form of a census database or constitute instead the nucleus of one or more population databases incorporating data from other sources. Consideration will also have to be given to such issues as:

- identification of the different types of users
- their information requirements
- types of information to be stored in the database
- sources of information
- maintenance and update of information
- processing of user queries
- identification of the appropriate commercial software or
- alternatively, whether it is feasible to develop such software, and
- ~~and~~ selection of the appropriate hardware capable of supporting the current database and its anticipated growth.

[125] Or in some cases existing survey or population databases used by the NSO for other collections.

3.423. Since building a census or population database requires careful planning and can be time consuming, such implementation should fit within the global statistical framework of the organisation, and be seen as an ongoing process both complementing the data dissemination strategy and strengthening the NSO's statistical capacity.

(a) ***Microdatabases***

3.424. Microdata (records of individual persons and households) collected in the census can be stored either in their raw form, or in their final edited form, or in a file that combines both raw and edited records. Digital technologies for mass storage present two issues for census managers and technicians: (a) the issue of when it will be appropriate to adopt a new technology as the standard; and (b) that of the need to convert materials stored in older media to the new standard or otherwise provide accessibility to the older materials.

3.425. Cloud-based data storage is becoming increasingly pervasive, and while it offers several advantages, including scalability, enhanced accessibility, disaster recovery, automatic updates, and seamless integration with other services, it also presents significant challenges. Data security remains a primary concern, especially with sensitive census information. Additionally, issues related to data sovereignty, potential cost overruns, data transfer bottlenecks, and the complexity of managing cloud configurations can arise. While the cloud provides a flexible and modern storage solution, its application to storing census microdata should be carefully evaluated.

3.426. With technological advances in mass storage devices and media, it is now feasible to store the full census data file as a single large rectangular file. After adding a data dictionary that describes the data format and a tabulation module, one obtains a set that could be described as a census database. The microdatabase requires a cross-tabulation program, which can be either part of the package or external. The software normally used for census tabulation still requires some prior training and may be confusing to inexperienced users. More intuitive tabulation software is available but may be either too slow in processing or too limited in its options to be fully satisfactory.

3.427. The organization of the microdatabase may take several formats, for example the software may allow for reorganizing the data in a transposed format (for example, one separate file per variable). This can substantially reduce the need for storage space and increase the speed of tabulations. However, establishing this kind of database is more complex, technically demanding and time consuming. There would be advantages in storing census microdata with standard commercial databases. This approach has the advantage that many users are already familiar with such software and so it is easier to find programmers and system analysts in the labour market. Even though the storage space required would be comparatively larger, today's market for mass storage has made available very large and fast hard disks at much cheaper prices than a few years ago and the hardware market seems to continue to follow this trend.

3.428. One of the main advantages of a microdatabase is that it permits the retrieval of data, at least in principle, at any level of detail. Since microdata could be used to obtain information on individual persons, families, households or family enterprises, privacy concerns must always be taken into consideration. In most countries, the use of census data to identify individuals is prohibited by law. Moreover, the long-term reputation of the NSO may well be jeopardized if such disclosures occur. Full elaboration of principles and protocols for disseminating microdata is elaborated below in paragraphs 3.468–3.488.

3.429. As presented in this subsection, there are methods (such as sampling, introduction of random disturbances, recoding and aggregation) that can be used to make such microdata available

while still protecting individuals' rights to privacy. All have in common the fact that they sacrifice some information in order to eliminate or greatly reduce the risk of disclosure. However, it is important that NSOs that plan to disseminate microdata to outside users should take the appropriate precautions to protect privacy and confidentiality.

(b) ***Macrodatabases***

3.430. Aggregated census data can be stored in many formats, either as the results for one census, as a database covering more than one demographic enquiry, or in a broad database of statistical and geospatial information. Whereas microdata are saved to allow aggregations to be made that were not programmed initially, macrodata are stored to preserve earlier aggregations, offering the general public readily usable information and avoiding redundancy for those who may find that the summary data they require have already been produced.

(i) Publication equivalents

3.431. The simplest form of what could be called a database for macrodata is a straight copy of a publication on the NSO's website. A machine-readable publication-equivalent database may have the advantage of being less expensive to prepare than its hard copy counterpart. In addition, electronic or paper copies can be made quickly, copying only the part of the publication that is required. However, accessing this information requires a computer, potentially with specific software to interpret the census information.

(ii) Table-oriented databases

3.432. More advanced users, especially those engaged in spatial analysis, may require a macrodatabase that transcends the scope of the printed publication. They might like to be able to manipulate the tables and shapefiles themselves in various ways in order to obtain views or results that represent their specific requirements more precisely (see paragraphs 3.462-3.467 below). Standard formats should be used to provide (partly) customized outputs, tables or georeferenced information to advanced users.

3.433. Census offices should be aware of the potential offered by extensible markup language (XML). XML is not, as a matter of fact, a language itself but rather a metalanguage system designed to be used on the Internet. With XML, users can define their own "tags" to structure the information within a document. XML thus offers the potential of precisely describing all elements composing a statistical table: title, subtitle, units of measure, indicators, values, the time dimension and footnotes, and, in short, the metadata. One extension of XML is the Geography Markup Language (GML), which is a modelling language for expressing geographical features. Other solutions, such as EDI/EDIFACT (electronic data interchange for administration, commerce and transport) and SDMX (Statistical Data and Metadata eXchange), are a set of internationally agreed standards, directories and guidelines for the electronic interchange of structured data between independent, computerized information systems.

(iii) Time series and indicator databases

3.434. Databases can also cover more than one demographic enquiry, and census results can be integrated with various other data sets, including the results of earlier censuses. In developing databases that are aimed at serving a heterogeneous user community, the issue of a number of basic trade-offs will have to be addressed. For example, on the one hand, the number of variables should be kept as small as possible to make the database easy to use; on the other hand, it should

be as comprehensive as possible to address the broadest possible requirements. A minimum data set of versatile indicators should consist of those variables that are useful for a wide range of applications and consistently available across space and time, and whose characteristics are clearly defined. In developing such a database, not only storage of the key indicators and variables themselves, but also the inclusion of some basic figures (absolute numbers or basic data) as a way of standardizing the basic statistical framework, is recommended.

3.435. With regard to time series and indicators, standardization and interoperability of data and metadata can be achieved via the Statistical Data and Metadata eXchange (SDMX)[126] international standard. In SDMX, Data Structure Definitions (DSDs) define structures that describe the data for each statistical domain, specifying the dimensions, attributes and measures, together with associated code lists. Metadata is similarly structured using the Metadata Structure Definition. Data and metadata adhering to the SDMX encoding standard may be stored in several formats, including SDMX-ML (an extension of XML) and comma separated values. Data and metadata represented in SDMX and conforming to the relevant DSD can be easily read and processed by any system or tool that supports SDMX, as well as stored in traditional databases, file systems, or other storage solutions.

(c) ***Graphing and mapping databases***

3.436. Databases equipped with graphing and mapping capabilities, especially those powered by GIS, significantly enhance their utility and relevance. GIS databases are specialized systems that store, retrieve, manage, display, and analyse geographic and spatial information. Unlike traditional databases, these are designed to handle spatial data, which are associated with geographic locations, and can visualize these data in the form of maps.

3.437. A GIS database typically consists of layers, each representing a particular type of data, such as roads, rivers, population density or land use. Each layer contains features, such as particular roads or rivers, which, in turn, have attributes, which are the specific pieces of information associated with that feature, stored in tables. For example, a road feature may have attributes such as its name, type, length, and surface material. Users can query the database, overlaying multiple layers, including geocoded census data, to see relationships between different types of data and perform spatial analysis. Users can also generate custom maps that highlight trends or illustrate spatial relationships.

3.438. Many users want data for relatively small areas in relation to such matters as home ownership, educational profiles and the labour market. While the database may relate to one census, some historical information can be included to allow users to observe prevailing trends over time.

3.439. Both microdata and macrodata can form the basis of these dissemination products. However, owing to disclosure problems as well as the need to increase processing speed, some form of prior aggregation is usually applied, for example by using summary data. Such summary data could also be combined with the general-purpose graphing and mapping software. However, this would result in a reduction of the user community to those able to handle rather more complicated processing jobs. Making available a census database with tightly integrated graphing and mapping

[126] https://sdmx.org

capabilities (which usually implies a tabulation function) is an excellent way to improve the effectiveness of census information dissemination.

4. Geographic products

a. *GIS for census data dissemination*

3.440. A geographic information system (GIS) typically embodies hardware and software configurations designed to support the capture, management, analysis and dissemination of spatially referenced data. Applied to census activities and outputs, such systems facilitate census cartography and data capture, and by linking population data (demographic, social and socio-economic) to geographic areas, a GIS provides very powerful data management functionalities in allowing users to explore, analyse, describe and communicate population census information according to their own data and information demands. Readers may want to refer to paragraphs 3.80-3.91 for a more general exposition of GIS systems.

3.441. In practical terms, such systems may range from simple desktop mapping facilities to complete systems that are capable of solving complex planning and management problems, producing detailed georeferenced inventories and spatial statistical analysis. The ability to use space to integrate and manipulate data sets from heterogeneous sources can make its application relevant to planning and managing the census process itself. For example, as noted at paragraph 3.82 a GIS provides functions for the areal interpolation of statistical data in cases where the boundaries of administrative units have changed between censuses.

3.442. Every NSO should have organizational competency and staff with competences in geospatial analysis and should develop GIS applications with population data and other georeferenced data from other sources for more advanced forms of spatial analysis. The NSO should then be in a position to supply census data at the right level and in the right format to users. NSOs provide vital information on current demographic conditions and future trends for policymakers in a range of sectors, such as health care, education, infrastructure planning, agriculture and natural resources management; and the provision of spatially referenced census databases is an essential prerequisite for the facilitation of the use of demographic data in these fields.

3.443. To achieve maximum efficiency gains, GIS applications should also be capable of generating additional geographic information beyond those used in the census, such as school and health districts, and water supply, power and utility service units. These entities will have to be constructed from the smallest geographically identified units available in the census, such as census blocks, grid squares (see paragraphs 3.456-3.458) or EAs. If, as is the case in most developing countries, EAs are the smallest units, this will have important implications for the design of EA boundaries. This requires close collaboration between NSOs and national mapping and survey agencies on the one side, and school, health, and utility service authorities on the other, when EA boundaries are drawn or modified, to avoid potential discrepancies later on.

3.444. Apart from providing NSOs with a very effective means/tools to disseminate and increase the utilization of census data, a GIS, more than any other data/information management system, provides easy and user-friendly access to census data in user-relevant formats. This allows analysts and planners to undertake policy analysis, planning and research that can more readily identify thematic and geographic priority areas and thus contribute to evidence-based and better-informed policy and decision-making at different levels of geography. Some of the spatial statistical analysis includes clustering, Moran's spatial autocorrelation, Anselin outlier analysis, Getis-Ord hotspot

analysis, ordinary least squares regression and geographically weighted regression[127]. A GIS allows governments to effectively monitor development progress across different sectors at village, municipality and subregional levels; it raises awareness about the importance of census and other socioeconomic data; and it increases the institutional capacity of NSOs and social and economic planning agencies to engage in more in-depth analyses of social and economic data and deliver information products in even more user-friendly formats.

b. ***Basic maps***

3.445. NSOs should take advantage of emerging GIS technologies to make the census results more understandable and easier to use. The purpose of statistical maps is to present the results in terms of their geographic distribution and also to make it easier for the general public to understand census results than when information is presented only in the form of statistical tables. There is special interest in the current pattern of the geographic distribution and also in changes in the patterns that have occurred over time, particularly since the last census. Harmonization of the boundaries between the censuses is essential for comparability of data.

3.446. The provision of maps serves two purposes: firstly, census area identification maps locate and show the boundaries of all administrative areas and units for which data are reported in census publications; and secondly, statistical or thematic maps and graphs present the significant results of the census, thus allowing the general user to visualize the geographic distributions and patterns inherent in the data. Well-designed and attractive maps will interest the users of census reports and may raise questions that may encourage them to investigate the statistical tables for further details.

3.447. There are three major types of area identification maps that are commonly used in most census publications in printed or GIS shapefile formats:

a. national maps showing the boundaries of the first- and second-order geographic divisions and of the major cities or metropolitan areas;
b. maps of each first-order division showing the boundaries of the second- and third-order divisions for which statistical tables will be prepared; and
c. urban or metropolitan and degree of urbanisation maps showing small sub-area boundaries as well as general streets, roads and rivers.

c. ***Thematic maps***

3.448. A comprehensive map publication programme should be developed as part of the overall population and housing census publication programme in order that the needed resources may be provided within the budget at the initial planning stages. In addition to preparing maps for the census tables and reports, many countries have also found it useful to produce a population atlas (dynamic or static) as a distinct census output. Collaboration with other departments and interested agencies might be sought to facilitate the production of such an atlas that would include maps depicting population and housing characteristics, as well as other data influencing the growth, composition and distribution pattern of population and housing.

3.449. As regards thematic maps, priority indicators for a census are: total population and its distribution by sub-area; population density; urban and rural population by degree of urbanisation

[127] Cressie, N. (2015). Statistics for Spatial Data (Revised ed.). Wiley.

or proportion of metropolitan and non-metropolitan population; and percentage changes in the population totals since the last census. Other topics for which key indicators can be presented include age, sex, fertility, mortality, migration, educational attainment, employment, household size, type of housing, ownership, number of rooms and sanitary facilities, a broad range of household amenities, and also, more recently, population-based development indicators such as household access to safe water, and household incomes. This list of indicators is merely an illustration of the type of thematic maps individual countries might find useful to produce. Producing maps using a common set of indicators enables countries to meaningfully compare their results over time and with international or regional norms. Also refer to the list in the table noted at paragraph 3.452 below.

3.450. Maps are an invaluable aid in meaningfully comparing subnational results with national values or with other regional and international norms. Emerging technologies provide great flexibility in composing informative and visually appealing maps. Often several maps can be combined on a single page to show one indicator, for example, for the urban and the rural population by degree of urbanisation to provide a more nuanced representation in the three classes of the methodology (cities, towns and semi-dense areas, and rural areas). Also, combining maps and statistical charts is an effective means of presenting census information. Dashboards with different methods of presenting census data on maps, charts, tables and infographics can provide users with comprehensive but easy to understand message about relevant phenomena.

3.451. By having associated graphing and mapping capabilities, databases will greatly increase their usefulness. Ideally, users should be able to generate the graphs or maps for their own needs. Several NSOs have developed the facility to produce this kind of product, sometimes in cooperation with a commercial company. However, it is recommended that NSOs develop mapping capabilities as a core competence for statistical production. Many users require small-area data concerning such matters as home ownership, educational profiles, and the labour market. As noted at paragraph 3.438, while the database may serve one census only, some historical information can be included to allow users to observe prevailing trends over time. As with all time series-type data, it is important to maintain consistency in both definitions and spatial representations to ensure comparability.

3.452. The list at Table 3.1 presents some suggested topics for census maps. The list is not exhaustive: most topics that appear in the questionnaire, as well as derived topics covered in Part Two, can be presented in cartographic form. In some countries, special themes such as population distribution by ethnic or language group may be appropriate. Conversely, some of the listed maps present information on the same topic in somewhat different form, so that the NSO may wish to select the most suitable indicator for the needs of the country.

Table 3.1 Illustrative list of thematic census maps

Population dynamics and distribution	***Socioeconomic characteristics***
Percentage population change during intercensal period(s)	Percentage of children not in primary school
Average annual growth rate	Percentage of youths neither in education training nor employment (15-24) NEET
Population density (persons per square kilometre)	Adult literacy rate (aged 15 and over)
Urban population as percentage of total population	Mean years of schooling (aged 25 and over)
Distribution and size of major cities and towns	Illiteracy rate of population aged 15 and over
In-migration, out-migration and net migration rates	Illiterate population aged 15 and over (total number)
Percentage born in country and foreign born	Educational level of population aged 10 and over
Percentage born in another division of the country	Labour force as percentage of total population
	Women's share of adult labour force
	Percentage of labour force by economic sector, type of occupation and status in employment
	Poverty mapping

Demographic characteristics	***Households and housing***
Sex ratio (males per 100 females), possibly by age groups	Average number of persons per household
Percentage of population aged 0–14	Percentage of households headed by women
Percentage of population aged 15–24	Average number of dwelling rooms per household
Percentage of population aged under 18	Percentage of households by tenure status
Percentage of population aged 15–64	Percentage of dwellings by type of construction material
Percentage of population aged 65 and over	Percentage of population with access to adequate shelter
Percentage female population in childbearing ages 15–49	Percentage of population with access to safe water
Total dependency ratio (population aged 0–14, and 65 and over, as percentage of population aged 15–64)	Percentage of population with access to electricity
	Percentage of population with access to sanitation
Percentage of population by marital status	Percentage of population with access to health services
Birth rate	
Total fertility rate	
Mean age at first marriage	
Death rate	
Infant mortality rate	
Under-five mortality rate	
Life expectancy at birth	
Percentage of persons with disabilities	

3.453. Where appropriate, the indicators can be presented disaggregated by sex as well as by urban or rural area (for example, where the rural population is greater than about 25 per cent of the total population). If information about an indicator is also available from a previous census, it is often very informative to produce maps showing change over time or to present maps for both time periods.

3.454. Also, where appropriate, countries are encouraged to perform spatial statistical analysis by producing maps showing spatial clustering and outlier analysis of the variables of interest, such as electricity and water supply.

3.455. The development of maps showing population size of locality (such as village, town, city, community, small area) by region is of particular value. These maps combine two types of information: locality population statistics and locality locations within each region or subnational area. More information can be presented on, for example, the locality location within the district and the region, habitable and non-habitable areas, densely populated localities, areas with no localities, and the proximity of localities to each other. Locality population size maps can also serve as base maps for displaying additional information about locality services and activities. This can include visualizing the location and distribution of services like primary schools, dispensaries, and piped water access points. However, it's important to note that determining the precise service areas of specific facilities, such as which localities are served by a particular school or dispensary, might require additional data sources and analysis beyond the scope of the census.

d. ***Grid based census outputs***

3.456. Grid-based[128] census outputs are a type of ~~data~~ presentation where census data are aggregated and organized within a grid system, usually a rectangular grid of equal-sized cells. This

[128] The grid square system is a method of dividing a geographic area into smaller, regularly shaped units for statistical purposes. These units are typically squares defined by lines of latitude and longitude, forming a grid that covers the entire area of interest.

system replaces the traditional method of presenting ~~census~~ data based on administrative boundaries, offering several advantages and facilitating new applications.

3.457. Grid-based census outputs are an increasingly valuable tool for researchers, policymakers, and the public alike. They provide a powerful way to analyse, visualize, and understand census data, offering significant benefits for decision-making, planning, and resource allocation across various sectors. As technology advances and data integration techniques improve, NSOs and users alike can expect grid-based outputs to play an even greater role in future data analysis and research.

3.458. The benefits of grid-based census outputs are:

Spatial analysis: Grids provide a consistent and easily interpretable spatial framework for analysing census data. This enables researchers and policymakers to identify spatial patterns and trends, visualize data variations across space, and conduct advanced geospatial analyses.

Data integration: Grids facilitate the integration of census data with other spatial datasets, such as environmental data, infrastructure maps, and land use data. This allows for a more comprehensive understanding of the relationships between population characteristics and other factors in the environment.

Data accessibility and transparency: Grid-based outputs can be easily visualized and shared through maps and other interactive tools, making census data more accessible and transparent to the public.

Flexibility: Unlike traditional area-based geographies, the grid system does away with the need for predefined local and national boundaries. This makes it a highly flexible tool for analysing data at any level of detail, from square metre grids to grids spanning tens of square kilometers.

Standardization: Grid-based outputs often use standardized grid systems, such as the Universal Transverse Mercator (UTM) grid, allowing for data comparison across different regions and time periods and between countries with vastly different administrative geographies.

e. ***Innovative geospatial products***

3.459. Other innovative geographic products for census data which utilize the emerging and still relatively new technologies of AI and ML include:

(i) Smart maps and geoportals

3.460. Smart maps is a term used to describe maps that are more interactive, intuitive, and informative than traditional maps. They may use AI and ML to improve the accuracy and usability of traditional maps and are designed to help users visualize data in a more meaningful way and make better decisions based on that data. They can be used to analyze large datasets, identify trends and patterns, and create maps that are more informative and engaging.

(ii) GeoAI and GeoBlockchain

3.461. GeoAI is the application of AI technology to enhance the intelligence of GIS software which enables the analysis of massive amounts of geospatial data (such as satellite imagery), whereas GeoBlockchain is a combination of blockchain technology and GIS that can be used to support the analysis of spatial-temporal trends. When combined, GeoAI and GeoBlockchain can be used to create a powerful platform for analysing geospatial data and generating insights that can drive better decision-making.

5. Interactive electronic outputs

3.462. It is of paramount importance that census data and any information produced to support them are widely disseminated and communicated, and that NSOs have a clear customer, client and stakeholder focus. That means that NSOs should place more emphasis on providing a service and creating partnerships than on merely providing products and should be guided by user-relevance and user-friendliness in all their operations, rather than by simply providing the tables, graphs and reports that they have traditionally produced.

3.463. Given its importance and widespread use, the web has emerged as the primary medium for providing general access to census statistics. Many NSOs have utilized the Internet as the principal channel for data communication, enabling users to have access to all published data online. When developing new census products, and when reviewing existing products, NSOs should consider all ways and means of making census statistics accessible, giving high priority to dissemination on the web. The advantages of online dissemination are primarily in terms of speed, flexibility and cost, as well as in providing accessibility to census results to a wide range of data users and allowing the delivery of data to be tailored to the level of expertise of the user.

3.464. Making a census database available online along with integrated searching, tabulating, graphing, mapping and analytic capabilities is an important way to improve the effectiveness of census data dissemination. Most NSOs provide user access to electronic databases and data files through their websites, satisfying the full range of needs of internal and external data users. This is a valuable service that allows users to access and display census data instantaneously and interactively. The establishment of such databases can enhance the dissemination of census results as well as increase their usefulness by allowing user interaction with census data. Such interaction is a key concept whereby users are enabled and empowered to access and explore census data themselves, and build their own customized tables or spatially configure data outputs according to their own requirements.

3.465. Interactive web-based data tools provide a user-friendly entry point to the entire range of census outputs. Basic design considerations of web-based interactive tools should factor issues such as identification of the different types of users, their information requirements and the types of information to be stored in the database. Content should be organized so that it can be easily understood and found, with an overview or tutorial to provide orienting information to users about the data products that can be obtained using the tool(s). Context should always be provided to all outputs through metadata, links to related information, and cross-referencing to glossaries, publications and other background material.

3.466. In practical terms, interactive web-based data tools should enable users to access census data themselves, and build their own customized tables or spatially configure data outputs according to varying spatial requirements. The tools should allow users to visualize and explore the data in various standard formats such as column charts, line graphs, maps and scatterplots. The table-building functionality should also have the ability to sort and order tabular results, and more easily select census dates and indicators. Tools should also be provided for downloading, conducting analysis or retrieval for use in other software. Design considerations to improve the interactivity of data interfaces should include the provision of user support designed to help users to anticipate, interpret and evaluate results. Support to users should include demonstrations and tutorials intended to describe how to perform the various functions related to the interactive web-based tools.

3.467. Especially for highly flexible interactive online tools, there are disclosure risks in addition to the risks present with traditional data releases, and a corresponding protection strategy must be a key part of the design of such tools from the outset (see the section on confidentiality and privacy at paragraphs 3.365-3.375) In addition to web-based tools, some interactive electronic products may also be published through other media.

6. Microdata dissemination[129]

3.468. The provision of census microdata, including subsets of the microdata based on a random sample, allows for further analysis and research, especially in the context of analysis of SDGs where further data dissemination by various characteristics is required. Also, longitudinal microdata files linking a sample of persons on census files over time have been produced and are widely used among the research communities in several countries.

(a) Definition of microdata

3.469. In general, when NSOs conduct censuses, surveys or collect administrative data, they gather information from each unit of observation. Such a unit can, for example, be a household, person, firm or enterprise, agricultural holding, school, or health facility. In this context, microdata are the electronic data files containing the information about each unit of observation. Microdata differ from macrodata or aggregated data, which provide a summarized version of this information in the form of counts, means, ratios, frequencies or other summary statistics.

3.470. Typically, microdata are organized in data files in which each line (or record) contains information about one unit of observation. This information is stored in variables that can be of different types (for example, numerical or alphanumerical, discrete or continuous). They can be obtained directly from the respondent via a questionnaire or by observation or measurement (for example, by GPS positioning), or imputed or calculated.

3.471. In the context of the population and housing census, microdata refer to electronic files consisting of individual records on persons, households and housing units. More specifically, microdata would typically be organized in multiple files: one with records on households, another with records on individuals, and yet another with records on housing units.

(b) Core principles for disseminating census microdata[130]

3.472. The United Nations Fundamental Principles of Official Statistics[131] provide unambiguous guidance in administering official statistics at national and international levels. A particular emphasis of these principles is on confidentiality of information collected for statistical purposes. In the

[129] The elaboration on the dissemination of microdata is largely based on the *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Dupriez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

[130] The elaboration of core principles for dissemination of microdata is quoted from *Managing Statistical Confidentiality and Microdata Access: Principles and Guidelines of Good Practice*, United Nations Economic Commission for Europe, Conference of European Statistician, United Nations publication, Sales No. E.07.II.E.7, United Nations, 2007.

[131] Presented at the United Nations Statistics Division website at: http://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

context, the sixth principle, governing international statistical activities, states: "Individual data collected by statistical agencies for statistical compilation, whether or not they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes."[132]

3.473. Any principles for microdata access must be consistent with this recommended principle and the principles contained in the legislation pertaining to the national statistical authority. The following principles should be considered for managing the confidentiality of microdata:

Principle 1: Appropriate use of microdata

3.474. It is appropriate for microdata collected for official statistical purposes to be used for statistical analysis to support research as long as confidentiality is protected.

3.475. Making microdata, or random samples of microdata, available for research is not in contradiction with the sixth United Nations Fundamental Principle as long as it is not possible for such data to refer to an identifiable individual. Principle 1 does not constitute an obligation to provide microdata. The NSO should be the one to decide whether to provide microdata or not. There may be other concerns (for example, quality) that make it inappropriate to provide access to microdata. Or there may be specific persons or institutions to whom it would be inappropriate to provide such data.

Principle 2: Microdata should only be made available for statistical and research purposes

3.476. For Principle 2, a distinction has to be made between statistical or research uses and administrative uses. In the case of statistical or research use, the aim is to derive statistics that refer to a group (be it of persons or legal entities). In the case of administrative use, the aim is to derive information about a particular person or legal entity to make a decision that may bring benefit or harm to the individual. For example, some requests for data may be legal (such as a court order) but inconsistent with this principle. It is in the interest of public confidence in the official statistical system that these requests are refused. If the use of the microdata is incompatible with statistical or research purposes, then microdata access should not be provided. Ethics committees or a similar arrangement may assist in situations where there is uncertainty whether to provide access or not.

3.477. Academics may wish to access microdata for research purposes but to support this research they may need to compile statistical aggregations of various forms, compile statistical distributions, fit statistical models or analyse statistical differences between subpopulations. These uses would be consistent with statistical purposes. To the extent that this is how the microdata are being used, it could also be said to support research purposes.

Principle 3: Provision of microdata should be consistent with legal and other necessary arrangements that ensure that confidentiality of the released microdata is protected

3.478. With respect to Principle 3, legal arrangements to protect confidentiality should be in place before any microdata are released. However, such arrangements have to be complemented with administrative and technical measures to regulate the access to the microdata and to ensure that

[132] Ibid.

information relating to an identifiable individual cannot be disclosed. The existence and visibility of such arrangements (whether prescribed in law or supplementary regulations, ordinances, and so forth) are necessary to increase public confidence that microdata will be used appropriately. Legal arrangements are clearly preferable, but in some countries, this may not be possible, and some other form of administrative arrangements should be put in place. The legal (or other arrangements) should also be cleared with the privacy authorities of countries where they exist before they are established by law. If such authorities do not exist, there may be non-governmental organizations that have a "watchdog" role on privacy matters. It would be sensible to get their support for any legal or other arrangements, or at least to address any serious concerns that they might have. In some countries, authorizing legislation does not exist. As a minimum requirement, release of microdata should be supported by some form of authority. However, an authorizing legislation is the preferable approach.

Principle 4: The procedures for researcher access to microdata, as well as the uses and users of microdata, should be transparent and publicly available

3.479. Principle 4 is important in order to increase public confidence that microdata is being used appropriately and to show that decisions about microdata release are taken on an objective basis. It is up to the NSO to decide if, how and to whom microdata can be released. But any such decision should be transparent. The NSO's website is an effective way of demonstrating compliance and also for providing information on how to access research reports based on released microdata.

Principle 6: Confidentiality

3.480. Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

(c) Microdata anonymization[133]

3.481. When disseminating census microdata files to the public, researchers or other agencies, the NSO faces conflicting objectives. On the one hand, it aims to release microdata files supporting a wide range of statistical analyses; on the other, it must safeguard the confidentiality of respondents' identities. Processes aimed at the latter are referred to collectively as statistical disclosure control (SDC) or anonymization.

3.482. A disclosure occurs when a person or organization recognizes or learns via released data something they did not know about another identifiable person. There are two types of disclosure risk: identity disclosure and attribute disclosure.[134] The former occurs when a respondent's identity is directly associated with a disseminated data record. This can occur easily when the data record includes variables unambiguously identifying the respondent – for instance, the respondent's name, address, passport or identification number, or telephone number. It is essential that such identifying

[133] The elaboration of procedures for anonymizing microdata draws extensively from *Handbook on Statistical Disclosure Control, Version 1.0*, Anco Hundepool, Josep Dominho-Ferrer, Luisa Franconi, Sarah Giessing, Rainer Lenz, Jane Longhurst, Eric Schulte Nordholt, Giovanni Seri, Peter-Paul de Wolf, Centre of Excellence for Statistical Disclosure Control, Eurostat project, December 2006.

[134] *Measures of Disclosure Risk and Harm*, Diane Lambert, Journal of Official Statistics, Volume 9, Number 2, Statistics Sweden, 1993.

variables be removed from any microdata files before dissemination. Attribute disclosure occurs when attribute values (or estimates thereof) in the disseminated data are associated with a particular respondent.

3.483. A combination of variables in a microdata record that can be applied to reidentify a respondent is referred to as a "key". Reidentification can occur (a) when a certain key value is rare or unique in the population with respect to a specific respondent; and (b) when this key can be used to match a microdata file to other data files that might contain direct or other identifiers such as voter lists, land registers or school records (or even publicly accessible Internet search engines).

3.484. The essential component of dissemination of census microdata files is avoiding both identity and attribute disclosures. In that respect, there is a need to strictly apply SDC or anonymization techniques to census microdata files. The first key step in anonymizing a microdata file is to remove (or encrypt) all direct identifiers – variables that unambiguously identify the respondent. Thereafter, a microdata file can be anonymized further by applying SDC techniques.[135]

(d) Protocols for dissemination of census microdata

3.485. Disseminating census microdata may be an unprecedented activity for the NSO. In that case, there is a need to develop particular protocols that would comply with the essential principles for disseminating microdata, as described above, and would also spell out the necessary requirements and components of such dissemination. Hence, such a protocol would need to take into account the following components and requirements:[136]

Openness. Access should be provided on equal terms for the research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination.

Flexibility. The means of access should take into account the rapid and often unpredictable changes in information technologies, the characteristics of each research field and the diversity of research systems, legal systems and cultures of each member country.

Transparency. Information on research data and data-producing organizations, documentation on the data and specifications of conditions attached to the use of these data should be internationally available in a transparent way, ideally through the Internet.

Legal conformity. Data access arrangements should respect the legal rights and legitimate interests of all stakeholders.

Protection of intellectual property. Data access arrangements should consider the applicability of copyright or of other intellectual property laws that may be relevant to publicly funded research databases.

---

[135] For full elaboration of these techniques, refer to Chapter 7 of the *Dissemination of Microdata Files: Principles, Procedures and Practices*, Olivier Dupriez and Ernie Boyko, IHSN Working Paper No. 005, August 2010.

[136] As presented in OECD Principles and Guidelines for Access to Research Data from Public Funding, Organisation for Economic Co-operation and Development, 2007.

Formal responsibility. Access arrangements should promote explicit, formal institutional practices, such as the development of rules and regulations, regarding the responsibilities of the various parties involved in data-related activities. These practices should pertain to authorship, producer credits, ownership, dissemination, usage restrictions, financial arrangements, ethical rules, licensing terms, liability and sustainable archiving.

Professionalism. Institutional arrangements for the management of research data should be based on the relevant professional standards and values embodied in the codes of conduct of the scientific communities involved.

Interoperability. Technological and semantic interoperability is a key consideration in enabling and promoting international and interdisciplinary access to and use of research data.

Quality. The value and utility of research data depend, to a large extent, on the quality of the data. Data managers, and data collection organizations, should pay particular attention to ensuring compliance with explicit quality standards.

Security. Specific attention should be devoted to supporting the use of techniques and instruments to guarantee the integrity and security of data.

Efficiency. One of the central goals of promoting data access and sharing is to improve the overall efficiency of publicly funded data collection to avoid the expensive and unnecessary duplication of data collection efforts.

Accountability. The performance of data access arrangements should be subject to periodic evaluation by user groups, responsible institutions and funding agencies.

Sustainability. Due consideration should be given to the sustainability of access to publicly funded research data as a key element of the research infrastructure. This means taking administrative responsibility for the measures to guarantee permanent access to data that have been determined to require long-term retention.

(e) Dissemination of population and housing census microdata in practice

3.486. It can be expected that the 2030 round of population and housing censuses, taking into account the contemporary development of processing technology and power, will put considerable pressure on NSOs in respect of the dissemination of the census microdata. In the context of implementing the principles and protocols described in the preceding paragraphs, NSO would need to ensure such dissemination in at least two different settings.

3.487. For the purpose of public dissemination, either online or via electronic media, only a representative sample of the individual records should be made available after ensuring the confidentiality or non-disclosure of individual information as elaborated above. The size of the sample would depend on the capacity and resources of the NSO. When generating public-use or research-use microdata samples from census data, not only the sample size but also the sample design should be optimized, recognizing that while simple random sampling methods may often suffice, consulting a sampling expert is recommended to ensure the appropriateness and effectiveness of the chosen approach.

3.488. As for the access to the complete census master file, it should only be made available to users using the model of the data laboratory or enclave[137]. This is a facility equipped with computers not linked to the Internet or an external network and from which no information can be downloaded via USB ports or other external drives. Users interested in accessing a data enclave will not necessarily have access to the full census data set – only to the particular data subset they require. They will be asked to complete an application form demonstrating a legitimate need to access these data to fulfil a stated statistical or research purpose and be briefed on the legal responsibility and repercussions related to maintaining the confidentiality of individual information. The outputs generated need to be scrutinized by way of a full disclosure review before release and they can contain only aggregates.

7. Customized products

3.489. The increasing activity in the field of economic and social planning and the attention of such planning to subnational areas are placing new demands on statistical information in general and on population and housing censuses in particular. There is an increasing need for tabulations and mapping not only by major and minor civil divisions and by other units of analysis such as metropolitan areas but even, beyond these, by small local areas.

3.490. Therefore, it is useful to establish an "on request" customer service for users who require aggregates not available through other means. This will be especially relevant in situations where such users cannot obtain census microdatabases. In essence, the service would require that the users provide the NSO with the details of the tables or other aggregates requested so that the NSO could fulfil the request, for payment of a cost-recovery customer service fee (where appropriate and permitted). Offering and promoting this service, especially online, would place the statistical service in a more favourable proactive position, rather than a static one, and could be a strong catalyst for closer cooperation with census product users.

3.491. The cost of such special purpose tabulations, which require computer programming, could be high, especially for academic institutions and other users who do not have access to a large budget. Some NSOs allow the users to do the necessary work using a user-friendly kind of software. A clearly written manual is required to guide the users in using the software, including the contents of the census data dictionary and other relevant information. The resulting tables are checked for any possible breach of confidentiality, in particular table cells with very small values.

3.492. Many NSOs provide services for special requests for census products, such as thematic databases, tables, and graphic and mapping outputs that can be designed for small, medium and large businesses, communities or special interest groups. These services are normally provided to meet the increasing demand of data users for a wide range of applications, such as monitoring trends, analysing unmet needs, identifying market potential, segmenting markets, identifying service areas and priority zones, determining optimum site locations, and designing and advertising new

[137] A data enclave is a secure network through which confidential data, such as identifiable information from census data, can be stored and disseminated. In a virtual data enclave, a researcher can access the data from their own computer but cannot download or remove it from the remote server. Higher security data can be accessed through a physical data enclave where a researcher must access the data from a monitored room where the data is stored on non-network computers.

products and services. Each category of products should also be made available on various media (typically, paper, disk, online) for dissemination according to user requirements.

3.493. Once the databases are created and have served the policy needs, they can serve other data users if they have market value. Since the NSO is normally the only source of many geographic databases related to census data applications, market demand for these products is increasing, particularly in the geographic and population-related areas. In such cases, census products could be governed by a licence. The licence permits the users to use the product without transferring ownership, since the ownership remains with the government agency. Either of two different licensing arrangements may be applied. The first is offered to organizations that use the data for their own needs and the other is offered to those that redistribute data or provide analytical services using census data to third party persons or organizations for a fee (see section on dissemination policy at paragraph 3.329).

3.494. Customized services of data on computer media are differentiated in terms of the forms of the data. Census products may be distributed in their original form, with or without other related information, or they can be distributed after making certain value-added modifications to meet the need of the users. Examples of such value-added activities include converting the data into another format (for use by other software packages), making the data more useful by creating subsets of the original data sets, merging the data from other sources and bundling with software. NSOs are encouraged to disseminate census content in formats that comply with the requirements assessed by the Open Knowledge Foundation.[138] According to these requirements, the openness of the contents is assessed within the Open Knowledge conceptual framework and concerns the possibility to reuse, revise, remix and redistribute data. In cases where copyright laws protect census data ownership, some royalty fees and data usage fees may be charged to the distributors to ensure a minimum return. However, it's important to ensure these charges are reasonable, as excessively high fees can hinder access to and use of census data, ultimately limiting its impact and value.

3.495. Some NSOs may assist their users by merging selected variables with the GIS shapefiles as a customized product. This has proven to be beneficial for schoolchildren in particular.

8. General interest products and special audience products

3.496. Information generated by a census is extensively used by a wide range of users with a variety of expertise. With the increase of demand for census products, efforts must be made by NSOs to produce a variety of products for various user groups, including special interest groups. In order to address their disparate needs, NSOs should segment stakeholders and potential audiences into groups to better know, understand and respond to their needs. This will help form the design and content of the various census products. In response to the needs of special audiences, NSOs may prepare products for key variables such as: policy summary reports; thematic and analytical reports; key findings reports; fact sheets; posters, brochures and flyers; detailed tables and spreadsheets; articles; and videos and social media products. Specialised analytical or thematic reports incorporate a high level of very sound analysis undertaken by staff who have a solid foundation in analytical techniques as well as the topic being analysed. In some cases, NSOs may undertake the analysis in collaboration with academic institutions or other specialists. Criteria used

[138] https://okfn.org/opendata/

in establishing the topics chosen will have to be set by the NSO, that may include: particularly interesting facts shown by the census data (perhaps confirming or rebutting conventional theories; confronting census data with material from other sources; or responding to issues raised by the public during user consultations on the content of the census).

3.497. It should be noted that the following products can only be effective in encouraging the use of census information if they are prepared and promoted in a timely and professional manner. This will require specialist skills from people familiar with communicating to the target audiences. These resources are costly and, therefore, NSOs should censure an adequate budget for these products and campaigns.

(a) Posters

3.498. One of the most common ways to disseminate census information consists of publishing posters highlighting key facts such as: How many are we? Where do we live? and summarizing a profile for the major civil divisions of a country. Posters might also be prepared addressing issues that are particularly relevant to, and designed to engage the attention of, different population groups such as teenagers, adults, indigenous populations, seniors and women.

3.499. Since the objective of a poster is to catch the eye at a distance, relatively few facts should be presented in a way so that the key message is immediately visible. Posters can be greatly enhanced by the addition of well-designed graphs, infographics and maps to increase the readability and comprehensibility of the key message. Posters and banners are short-term communication products and should be used to communicate key findings.

(b) Brochures and flyers

3.500. Professionally designed brochures and flyers are another way to disseminate basic census data. These brochures should be written in a very easy and comprehensible language indicating the demographic profile of the country and illustrated with suitable graphics and explanatory material. They are particularly suitable for preparation as promotional materials for people attending events and exhibitions, such as the launch of more traditional materials, or for inclusion on display racks in public libraries.

(c) Videos and sketches

3.501. In order to create a better understanding among certain interest groups, the use of other communication media are recommended, including videos and sketches. In order to promote the story behind the numbers and increase the use of census data, graphics such as charts or maps could be included on CD, DVD, or memory sticks. These might indicate how census data can assist policymakers, planners and people in general to better understand their societies and how census data can assist in identifying the main problems and the evaluation of solutions.

(d) Instructional materials for the general public

3.502. Instructional materials in an easy-to-understand form can be prepared for the general public, indicating the advantages and limitations of census data. Such material can often form the basis of information campaigns as part of the advocacy material for the next census.

3.503. A particular implementation of instructional materials can be the preparation of a kit for use in schools. Not only will this provide high-quality information for the students but, by including

exposure to the use of statistical materials in the school process, it will encourage the use of evidence-based analysis throughout society. It should be noted that professional input should be sought in ensuring that these materials follow sound educational practices and can be accommodated within the appropriate curriculum.

(e) Social media

3.504. Internet-based social media have become an indispensable tool in disseminating information and marketing statistical products. Various social media platforms have been successfully used by countries as part of the dissemination of their census results. Interacting with followers and users on these platforms provides the NSO with an opportunity to disseminate information, build relationships with established and new users, and engage the public on a regular basis.

3.505. Social media platforms such as Facebook, X (formerly Twitter) and online video sites can be used to post all census-related advertisements and other videos. It is important to recognize, however, that social media can be used to spread misinformation, which may negatively impact public trust in the NSO. To mitigate this risk, the NSO should actively monitor social media platforms and develop strategies to address and counter harmful misinformation (see paragraphs 2.144-2.146).

(f) Education material for schools

3.506. Learning about the census and about how to use census and other statistical data effectively is a key part of a country's educational programme. It helps students to better understand society and their role as responsible citizens in participating in that society. It helps them to acquire some very important skills. For instance, learning how data is collected, analysed and reported in a meaningful way not only helps to develop statistical, presentational and descriptive skills, but also helps them to understand the data that is presented to them. It gives them the skills to ask questions to find the facts to support their understanding of what is happening in the world around them in a meaningful way. Students will use these skills in many of their future studies whether these are in the areas of mathematics, statistics, demography, business, economics, history, geography or in conducting scientific projects.

3.507. In addition to equipping students of all ages with essential data literacy skills, the inclusion of census-related activities in the school curricula can provide a good opportunity to encourage use of census data, particularly among hard-to-reach groups. It is advisable to get practicing schoolteachers involved in developing appropriate census content that aligns with the school curricula. This should include extremely accessible website content featuring videos, infographics and practical exercises[139].

D. Census data utilization

1. General uses of population and housing censuses

3.508. As has been noted at the very outset of these *Principles and Recommendations*, population censuses are traditionally used for public and private sector policymaking, planning, administrative

[139] https://www.cso.ie/en/census/censusforschools/censusinschools2022/.

and research purposes at national and subnational level. One of the most basic of the administrative uses of census data is in the demarcation of constituencies and the allocation of representation on governing bodies, as well as allocation of resources within a country. Certain aspects of the legal or administrative status of territorial divisions may also depend on the size of their populations. Housing censuses are used to develop benchmark housing statistics and to formulate housing policy and programmes, and in the private sector to assist in site selection for industrial, retail and service facilities, as well as for the commercial development of residential housing. Data from the housing census are very useful in helping national and local authorities assess vulnerability and infrastructure as well as the overall effects of climate change on the sustainability of dwellings/buildings.

3.509. As has been noted several times in these *Principles and Recommendations*, information on the size, distribution and characteristics of a country's population is essential to describing and assessing its economic, social and demographic circumstances and to developing sound policies and programmes aimed at fostering the welfare of a country and its population. The population and housing census, by providing comparable basic statistics for a country as a whole and for each administrative unit and locality therein, can make an important contribution to the overall planning process and the management of national development. The availability of information at the lowest levels of administrative unit is valuable for the management and evaluation of such programmes as education and literacy, employment and human resources, reproductive health and family planning, housing and environment, maternal and child health, disability, rural development, transportation and highway planning, urbanization and welfare. Population and housing censuses are also unique sources of data for producing relevant social indicators to monitor the impact of such-cross cutting government policies and programmes (see paragraphs 3.530–3.532).

(a) Uses of population censuses

3.510. The uses of population census results and the associated tabulations described in the following paragraphs relate to the topics recommended for inclusion in the census as listed in Table 3 in Part Five. An exposition of the uses of the population census more generally is set out in paragraphs 1.21. -1.51. .

3.511. Reporting on the total population and its distribution among major and minor territorial divisions and localities is a fundamental requirement of the census because these data are used for determining the apportionment of representation in legislative bodies, for administrative purposes and for planning the location of economic and social facilities. Internal migration, one of the major components of population change, frequently affects the trends in population distribution. Data on internal and international migration, together with fertility and mortality, are needed to prepare population estimates and projections for planning purposes and for determining policies on migration and assessing their effectiveness. Data on forcibly displaced status, both for refugee and internally displaced populations allows to measure integration and policy responses in this area[140].

3.512. The household - a basic socioeconomic unit in all countries - is often central to the study of social and economic development. The number, size and structure of households and changes in the rate of household formation are useful for planning and developing special policies formulated for

[140] IRRS see paragraph 4, more info https://egrisstats.org/recommendations/international-recommendations-on-refugee-statistics-irrs/. Also, IRIS see paragraph 3, more info on https://egrisstats.org/recommendations/international-recommendations-on-idp-statistics-iris/

selected groups of the population, such as children, youth, older persons, and those with disabilities. Thus, the distribution of individuals within households is used to determine the living arrangements of families, the patterns of family structure observed, the time when new families are formed and changes in family structure due to death, divorce or separation, migration, or the departure of children to form their own households. The relationship among household members can be used to determine family structure and the existence of households composed, partially or completely, of unrelated persons.

3.513. Traditionally defined demographic, social and ethnocultural characteristics collected from the population census include sex, age, marital status, religion, language, nationality, and ethnicity and indigenous identification. Sex and age are fundamental to the majority of the characteristics collected in the census, which provides more data than any other single source on the demographic and socioeconomic differences between women and men.

3.514. Depending on national circumstances, cultural diversity may be indicated by language spoken in the home or community, religion and national or ethnic group. For countries that are not homogeneous in terms of one or more of these variables, linguistic, religious and national or ethnic groups provide the basic information for a quantitative assessment of the relative size and age-sex distribution of this diversity.

3.515. Although census data on fertility and mortality cannot serve as a substitute for reliable birth and death statistics derived from civil registrations, they are particularly valuable for countries where birth or death registration is lacking or incomplete and where, as a result, vital statistics are unavailable. Even in countries with complete registration of these events, the population census is useful as a supplement to such registration data because the responses to fertility questions provide input for calculating lifetime fertility of the female population or cohort fertility.

3.516. Education has historically been one of the key factors determining the quality of life, and interest in education continues today in most countries of the world, with emphasis on improving access to, and the quality of, education, as well as broadening the scope of basic education.[141] Education is also considered a major tool in closing the gap between women and men in respect of socio-economic opportunities. Benchmark data obtained from national population censuses will therefore be of considerable importance towards fulfilling this objective. Census data reveal the disparity in educational opportunities between the sexes, age cohorts or generations, and urban–rural populations, and provide important indications of the capacity of the country for economic and social development. They furnish material for the comparison of the present educational attainment of the adult population with current and anticipated requirements of educated human resources for various types of economic activities. Such a comparison may serve as a guide both for national policy in terms of the development of the educational system, and for the planning of the economic development programmes that it will be feasible to undertake in view of human resource requirements.

3.517. Census information on the economic characteristics of the population focuses on enumerating the labour force so as to provide benchmark data for current studies of employment and labour underutilization, in particular unemployment and the potential labour force. It provides

[141] Education for All Summit of Nine High-Population Countries, New Delhi, 12–16 December 1993: Final Report (Paris, UNESCO, 1994).

information on the growth, composition and distribution of the labour force for use in policy formulation and the appraisal of human resource utilization. Economic data from censuses can also provide some input into statistics on the distribution of income, consumption and accumulation of households, and participation in agriculture and non-agricultural activities. Furthermore, the data on the labour force may give an approximate indication of the number of workers who are responsible for the support of dependants.

3.518. Statistics obtained from different sources (for example surveys on the labour, agriculture, and establishments) rely on the census for sampling frames, and the use of common concepts in the different sources helps in securing comparability when multiple sources for changing patterns of economic activity are being relied upon.

3.519. As interest in the movement of people across national boundaries has grown steadily among countries, and census tabulations relating to such international migration have grown in importance. Such tabulations are designed to assess the impact of migration on receiving countries, to understand patterns of diversity and develop programmes for the adaptation of immigrants, and to serve as a source of information on emigration. The census also plays a crucial role in measuring internal migration within a country. Historically, censuses were primarily conducted to assess population movements, particularly the large-scale migration from rural areas to urban centres during the industrial revolution. By providing detailed demographic data, censuses offer valuable insights into migration patterns, helping to inform policies related to urban planning, resource allocation, and social services. Additionally, census data on refugees and internally displaced persons is essential for understanding the impact of forced displacement and informing policies to support these populations.

3.520. The census is also an important source of information on household poverty and living conditions. In particular it is an important source of data on persons with disabilities. Census data help to monitor the social and living conditions of such persons in terms of school attendance, educational attainment, employment, marital status and household arrangements. The data also provide a basis for developing policies to meet the needs of persons with disabilities and for evaluating the effectiveness of these policies.

3.521. Population censuses also provided historical data for historians, and researchers studying social, economic, and demographic changes over time.

(b) Uses of housing censuses

3.522. The primary uses of information from housing censuses include (among many others):

- the development of a basis for planning housing and human settlement programmes and policies;
- public and private sector studies of urban and other non-agricultural land use;
- evaluation of the adequacy of housing stock and assessment of the need and market for new housing;
- measuring access to amenities and facilities; and
- studies of the living conditions of the homeless and those living in temporary or substandard housing.

Information collected on the number of sets, type and characteristics of living quarters and their occupants is necessary from the point of view of monitoring housing conditions and needs of the population. Combined with the information collected by regular annual statistical programmes on housing construction, data from the housing census provide a basis for identifying national, regional and local housing patterns, which are needed for the development of a rational housing market

aimed at stimulating various types of housing construction. The type and quality of shelter in which people are housed, that is to say, the space, degree of crowding, facilities, surroundings and available transportation facilities, affect their economic activity, health, social intercourse and general outlook. It also affects their security and ability to combat the effects of climate change. The supply, characteristics and costs of housing are therefore subjects for which the housing census is an important source of information.

3.523. Housing census data provides valuable insights into poverty levels, and along with data on other socio-economic characteristics of the population (derived from the population census), can be used to compute poverty indices. By analysing information on housing characteristics, amenities, and living conditions, it is possible to identify areas with high rates of poverty and assess the impact of housing-related factors on overall well-being. Housing census data can also be used to develop targeted poverty reduction strategies and monitor the effectiveness of interventions aimed at improving housing conditions and reducing poverty. A more general description of the uses of housing censuses is set out at paragraphs 1.52. -1.60. .

2. Analysis of the results

3.524. In order to ensure the fullest possible utilization of census results by national and local governmental authorities, by academic researchers and a diversity of other users, NSOs should draw up a comprehensive and coordinated programme of analytical studies, phased over a period of several years. This will help allocate effort and resources in such a way as to ensure that important policy needs are adequately met, undue duplication of research effort is avoided, and priorities are observed as far as possible. In these studies, the data of the current census should be examined not only by themselves but also as complemented by relevant data from other sources and from earlier censuses, in order to obtain a broader context, improve the estimates and establish trends.

3.525. The analytical studies to be included in such a programme will vary according to the needs and circumstances of the country. The programme may include descriptive summaries of results, policy-oriented analyses of census results and detailed analytical studies of one or more aspects of the demographic and social situation of the country. Some of these studies may be undertaken by the NSO itself, but others, particularly the more time-consuming studies, can most effectively be carried out in cooperation with specialists in different subjects having experience in in-depth analytical studies from universities or other research centres. In any case, it is desirable to invite specialists from other governmental offices and experts outside government to take part in drawing up this programme of studies, and it is natural that they would play an important part in the execution of various parts of the analytical programme.

3.526. One important aspect to be considered in establishing a programme of analysis is the possible use of census results in achieving the goals and objectives of population, human settlements or similar policies and strategies at the national and local level, and in applying available resources effectively towards the improvement of conditions in these fields. For this purpose, it will be necessary to analyse the census results within a wider framework provided by other available sources of information so as to achieve an integrated approach to the solutions of population, human settlements and similar problems. (See paragraphs 1.109. -1.130. for an insight into the benefits of integrating the data from the census and other sources.)

3.527. In addition to the studies that are part of the overall census programme, additional analyses carried out on their own initiative by research organizations, universities or other experts should be encouraged.

3. Intercensal and post-census population estimates

3.528. Intercensal population estimates (more often referred to as 'population projections') are estimates of the future population using the most recent census population as the base and applying assumed levels of births, deaths and net migration for any year in within specified period of time subsequent to that census. The projection for each year will, in due course, be revised when data on the actual levels of natural increase and net migration for that year are made available and the official estimates for that year are published. Post-census estimates (defined here to mean back-estimates for each year prior to the census) are used to revise previously made annual estimates for the years following the previous census.

3.529. Intercensal population estimates can be calculated using a combination of data sources and statistical methods; the actual methods used vary by country and organization[142]. Inputs for intercensal and post-census estimates can include the base population from the most recent census, estimates of the number of births and deaths, both international and domestic migration, as well as adjustment for specific sub-groups such as specific age groups[143]. For analytical and planning purposes, results of the census evaluation should be taken into account in preparing the intercensal and post-censual population estimates in order to correct deficiencies in the data collection phase, and in the base population in particular.[144] It is recommended that the NSO (or any other national statistical institution) in charge of producing intercensal and/or post-census population estimates make available information on the data, methods and adjustment (if any) used to update these population estimates. Such information is important if national official population estimates are to be considered and used in processes at the global level.

4. Cross-cutting and emerging social issues

3.530. Reflecting the concerns and priorities among countries around the world, the United Nations convened a series of global conferences: on children, education, environment and climate change, human rights, population, sustainable development, women and human settlements. Each of these conferences recognized the importance of adequate information for formulating policy and monitoring progress in the achievement of conference goals and called on countries and international organizations to develop and improve the requisite statistics and indicators. These recommendations are reflected for example in the Vienna Declaration and Programme of Action of

[142] Spoorenberg, T. (2020). "Data and methods for the production of national population estimates: An overview and analysis of available metadata" United Nations, DESA, Population Division, Technical Paper No. 2020/01
https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2020/Sep/un_2020_techpaper1.pdf

[143] European Commission (2003). Basic Methodology for the Recalculation of Intercensal Population Estimates. Working Papers and Studies (3/2003/E/no 27), Luxembourg: Office for Official Publications of the European Communities; Statistics Canada (2012). Population and Family Estimation Methods at Statistics Canada. Ottawa: Statistics Canada, Demography Division, Catalogue no. 91-528-X; U.S. Bureau of the Census (2012). Methodology for the Intercensal Population and Housing Unit Estimates: 2000 to 2010. Washington, D.C., U.S. Bureau of the Census, Revised October 2012.

[144] United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022: Methodology of the United Nations population estimates and projections. UN DESA/POP/2022/TR/NO. 4

the World Conference on Human Rights;[145] the Programme of Action of the International Conference on Population and Development;[146] the Copenhagen Declaration on Social Development and the Programme of Action of the World Summit for Social Development;[147] the Platform for Action[148] adopted by the Fourth World Conference on Women; the Sustainable Development Goals adopted by the General Assembly; and the Climate Change Declaration. The programmes of action adopted by these international conferences targeted many interrelated areas of concern and called for improved statistics to monitor progress. In deciding which social groups merit monitoring in regard to measuring the disadvantages suffered by particular groups of people, each country should determine which groups within it need special attention. Some of the common factors leading to social disadvantage are gender, age, physical or mental impairment, race and creed. The disadvantaged are not necessarily small in number; they may constitute the majority of the population.[149]

3.531. To meet the need for statistics on gender equality, many activities have been undertaken during the last two decades at the national and international levels to improve concepts, definitions and classifications for collection of statistics related to women and men. In this present document, the importance of the population and housing census as a data source has often been stressed. The population and housing census is also the principal, or sometimes the only, comprehensive national data source with respect to meeting the need for statistics on children, youths, older persons and persons with disabilities in the development of policies and programmes at the national and international levels. Therefore, it is important that countries identify data requirements concerning various population groups of particular interest when planning their censuses and ensure that the definitions and classification to be followed in censuses are appropriate and also consistent with those in use for the entire population.

3.532. Furthermore, the census tabulation plan should ensure the inclusion of all relevant details about special population groups and a range of cross classifications for each group, with a view to analysing its social and economic conditions. The concepts and methods for collecting, processing and tabulating census data for special population groups should take account of the needs of the users concerned with such groups. In the case of some groups, for example persons with disabilities, a special set of questions is required to identify members of this population group. In the case of others, standard questions, for example on age, are sufficient to identify groups such as children, youths and older persons. In both cases, most variables needed for cross-tabulations are already provided for in the international recommendations and many national censuses. Throughout the census operation, however, attention will often need to be given to improvement of coverage, quality-of-data issues and avoidance of stereotypic treatment. The present section deals with gender issues, a few special population categories such as children and youths, older persons, and persons

---

[145] A/CONF.157/24 (Part I), Chapter III.

[146] *Report of the International Conference on Population and Development, Cairo, 5–13 September 1994* (United Nations publication, Sales No. E.95.XIII.18), Chapter I, Resolution 1, Annex.

[147] *Report of the World Summit for Social Development, Copenhagen, 6–12 March 1995* (United Nations publication, Sales No. E.96.IV.8), Chapter I, Resolution 1, Annexes I and II.

[148] *Report of the Fourth World Conference on Women, Beijing, 4–15 September 1995* (United Nations publication, Sales No. E.96.IV.13), Chapter I, Resolution I, Annex II.

[149] *Note by the Secretary-General transmitting the report of the Expert Group on the Statistical Implications of Recent Major United Nations Conferences* (E/CN.3/AC.1/1996/R.4), annex, paragraphs 68–69. Presented to the Working Group on International Statistical Programmes and Coordination at its eighteenth session, New York, 16–19 April 1996.

with disabilities, so as to assist in the preparation of detailed tabulations and databases according to international standards.

(a) Statistics on gender equality

3.533. The global conferences on women have contributed to an increased awareness of the importance of statistics not only on women and girls but, more broadly, on gender-related issues. For example, in developing census plans in a number of countries, efforts have been made to review and assess the adequacy of statistics for understanding the diversity of both women's and men's lives. Improvement of statistics and statistical methods related to gender issues should be an important priority in all stages of work on planning the collection, analysis and dissemination of data on topics for which tabulations should be presented cross-classified by sex.

3.534. In addition to the more general problems of the quality of census data, other types of problems that apply particularly to women and stem from gender-based stereotypes and biases have been noted. Similarly, the notion that only men can be heads of the household affects the way questions have been designed and asked in censuses. Such stereotypes also affect the way respondents reply to the questions. A common problem, for example, is classifying women automatically as outside the labour force without asking whether they perform any work for pay or profit, even as a part-time or secondary activity.

3.535. Another problem relates to biases in the collection, processing, compilation, and presentation of data. For example, when census tabulations are prepared for the employed population by occupation, they may be presented for males only on the assumption that information on the occupational pattern of women is of little use. And even when tabulations of the employed by occupation are disaggregated by sex, main gender differences in occupations may be missed if the occupation data are presented only at the two-digit level of the classification. Rather, special tabulations showing, for example, the 10 or 20 detailed occupational groups with the highest concentrations of women or men would be needed to render visible gender-based occupational segregation.

3.536. During the past few decades, considerable effort has been devoted, on the one hand, to reviewing such bias and its impact on statistics concerning gender equality and, on the other hand, to improving the concepts and methods involved in the collection of data in censuses and surveys. Related improvements in the revised System of National Accounts and recommendations concerning statistics of work, employment and labour underutilization adopted by the International Conference of Labour Statisticians in their 19th session[150] in 2013 and 21st session[151] in 2023 are also of importance to the population census. They are intended to overcome the above-mentioned conceptual deficiencies and to identify and provide measures for all productive activities (that is, all forms of work) performed by women and men, whether paid or unpaid. Similarly, efforts at the national level have been focused, for example, on eliminating biases in concepts, classifications and definitions of head of the household.

[150] https://webapps.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_220535.pdf
[151] https://ilostat.ilo.org/about/standards/icls/icls-documents/

3.537. Important statistical series and measures on the status of women can be readily obtained based on the topics and recommended tabulations from censuses. Furthermore, in the case of most topics, the primary unit of classification is the individual and therefore a vast array of indicators may be obtained by devising appropriate additional cross-classifications for the female and male populations separately. For an illustration of those census topics and tabulations that are useful for developing comprehensive statistics on women and men, see the Minimum Set of Gender Indicators,[152] *Integrating a Gender Perspective into Statistics,*[153] and *Methodological Guidelines for the Gender Analysis of National Population and Housing Census Data.*[154] The household and family status classifications presented in paragraph 5.162 are appropriate for analysing the living situation of women and men, with specific reference to single mothers and fathers and older women and men living alone.

3.538. It should be emphasized that while all data collected at the individual level can be presented by sex, this is not always done. Cross-classifications by sex tend to be suppressed when they become complex with multiple-variable tables. However, in order to satisfy one basic condition for gender statistics, which is that all statistics on individuals (subject to one or two exceptions such as those relating only to women) should be presented by sex, sex should be considered the overriding variable in all tables relating to persons.

3.539. Another important consideration is to broaden the target of dissemination and use of census data by popularizing the statistics that are published. One approach to achieving this wide outreach is to present statistics in the form of charts and simplified tables, with a simple and clear interpretation of the data (see paragraph 3.496-3.501). Countries planning to issue an analytical report might wish to consider using such formats. The analytical publication could cover the main census topics or alternatively those topics that are especially important to understanding the relative position of women and men in society.
`

(b) Statistics on children and youths

3.540. Extensive data on children and youths are available in censuses but may need improvements in terms of: coverage to ensure all such persons are counted; the quality of information on specific characteristics of these sub-populations; and the range of tabulations available.

3.541. For statistical purposes, “children” are defined as persons under 15 years of age[155], and “youths” are defined as those aged 15–24. However, it is useful to further divide these special groups by 5-year age groups (or, nationally, by groups of specific school ages) because of the rapid changes in characteristics in this age range, such as in school attendance, marital status and economic activity status. Also, because of differences by sex in the age at marriage or family formation, family or household status and entry into the labour market, data should be classified not only by age but also by sex (as recommended at (a) above). While single-year age distributions for young children under 5 might not always be feasible or necessary, distinguishing between infants (under 1 year) and those aged 1-4 can provide valuable insights for specific demographic analyses

[152] Available at: https://gender-data-hub-2-undesa.hub.arcgis.com/.
[153] United Nations ST/ESA/STAT/SER.F/111, Sales No. 13.XVII.9.
[154] UNFPA, 2014.
[155] Note that UNICEF uses under 18 to define children based on the Convention of the Rights of the Child. The Convention defines a child as any human being below the age of 18 years unless the law of a specific country determines an earlier age of majority.

and policy planning. For youths aged 15–19, it would be desirable to distinguish between those 15–17 years of age and those 18–19 years of age, or to have a distinction corresponding to the age below which the country considers an individual to be a minor.

3.542. For the purpose of developing statistics on children, the principally recommended topics include (a) sex, (b) age, (c) school attendance (for school-age children) and (d) relationship to the reference member (or head) of the household.

3.543. Given the priority on the girl child highlighted by the World Summit for Children (1990)[156], the International Conference on Population and Development (1994) and the Fourth World Conference on Women (1995), special attention needs to be given to improving and disseminating statistics on children by sex. Of particular concern is the situation of the girl child with respect to school attendance, mortality, early marriage, early fertility and other issues of importance. A basic problem with statistics on the girl child is that data on children ever born and children surviving tend not to be disaggregated by sex at either the questionnaire design or the tabulation stage. These data can therefore only be used for deriving indirect estimates of child mortality.

3.544. The principal topics of investigation identified for children apply also to youths, with the following additions: (a) marital status, (b) literacy, (c) educational attainment, (d) economic activity status, (e) number of children born alive and (f) age at marriage or family formation, though the latter two topics will only apply to female youths.

3.545. Some of the useful statistics and measures can be readily compiled based on the abovementioned topics, while any additional indicators can also be obtained based on more detailed cross-classifications using other recommended census topics or tabulations.

(c) Statistics on older persons

3.546. Older persons are defined by the United Nations as all persons aged 60 years and over. For purposes of classification, depending on the national situation, it is useful to tabulate data by 5-year age groups up to age 100, instead of including them in the single broad age category 60 and over.

3.547. For such older persons, extensive data are collected in population and housing censuses but may need detailed age-sex classification, as described below. For the purpose of developing statistics and indicators on older persons, the recommended topics that are particularly relevant include (a) sex, (b) age, (c) marital status, (d) labour force status, (e) income, (f) type of household, (g) type of living quarters, (h) institutional population and (i) disability.

3.548. The statistics needed for studies of older persons are disparate, depending as they do on national policies and circumstances. Internationally, no illustrative list of indicators is available to ensure appropriate census tabulations.

(d) Statistics on persons with disabilities

3.549. The census can provide a valuable source of information on the frequency and distribution of disability in the population, at national, regional and local levels. Experience shows that although

[156] https://documents.un.org/doc/undoc/gen/n90/267/21/pdf/n9026721.pdf

an increasing number of countries ask questions about disability in their censuses, the presentation of disability data has often been limited to tabulations showing the number of specific severe disabilities present in the population. Unfortunately, cross-tabulations with other characteristics are not usually provided.

3.550. A great deal of work on concepts, classifications and development of statistics on persons with disabilities has been undertaken in recent years, particularly through the work of the Washington Group on Disability Statistics[157]. For the third time, recommendations are being made on including disability questions in a population census. The following two paragraphs highlight issues involved in preparing detailed census tabulations on persons with disabilities, but a more detailed discussion on the concepts and methods used to measure disability from census questions is set out in paragraphs **Error! Reference source not found.**-5.236.

3.551. For the purpose of developing statistics on the situation of persons with disabilities the principal topics in census recommendations that would be necessary for the assessment of equality of opportunities include (a) sex, (b) age, (c) place of residence, (d) type of household, (e) marital status, (f) educational attainment and school attendance, (g) labour force status, (h) status in employment, (i) industry and (j) occupation.[158]

3.552. Not only should the tabulation plan for the disability data include prevalence rates by sex and age, but it is also very important that tabulations comparing persons with and without disabilities on key social and economic characteristics be presented.

(e) Ethno-cultural characteristics

3.553. Collecting information on the ethnic composition of the population informs a better understanding of the ethnic background of a country's population, especially with respect to indigenous populations, international migrants, nomads, stateless and forcibly displaced populations, and other specific population groups of interest.

3.554. There are some difficulties in collecting this information since some population groups may report their ethnic identification based on its local meaning, and so, in order to correctly allocate these persons to their particular ethnic group it is necessary for the NSO to compile a list of nationally recognised ethnic groups, sub-ethnic groups and local definitions of small ethnic population groups. This will facilitate accurate and standardised data on the ethnic composition of population. It would also be useful if demographers and specialists in the field of ethnography, as well as organizations involved with indigenous people, are consulted in the creation of such a list. A detailed discussion on the concepts and methods used to collect census information on ethnicity and indigenous populations is set out in paragraphs 5.201-5.213.

3.555. In order to obtain comprehensive information characterizing ethnic composition of the population, it would be useful to tabulate data by (a) sex, (b) age, (c) place of residence, (d) marital

[157] For more information on the Washington Group on Disability Statistics, see: http://www.cdc.gov/nchs/citygroup.htm.

[158] This list of topics is essential for monitoring the implementation of the Convention on the Rights of Persons with Disabilities, especially concerning education, work and employment.

status, (e) births and deaths within specific ethnic groups, (g) educational attainment, (h) labour force status (i) status in employment (j) industry, (k) occupation, and (l) type and size of household.

3.556. It is important to obtain comprehensive information on indigenous populations in order to have statistics on the number as well as the demographic and socio-economic structure of the given population group. These data would provide valuable information to support the development of programmes for social support of such people.

3.557. Statistics about the ethnic composition of international migrants together with information about country of birth and citizenship will help to more precisely determine the inward flow and volume of international migration.

3.558. Population censuses are also the sources of information about religious identification of the population. It would be useful to obtain this information by (a) sex, (b) age, (c) ethnic group, (d) place of residence, and (d) place of birth. This information would be useful to study distribution of religious affiliations. (See also paragraphs 5.192-5.196.)

3.559. Information about knowledge of languages is widely used. Countries find it useful to study the official language of the country as well as the use of mother tongues or some other languages. In any case it would be useful to tabulate this information by (a) sex, (b) age, (c) ethnic group, (d) place of residence, and (e) place of birth.

3.560. Information about knowledge of the official language of the country would be very useful for studying the integration of international migrants and may be used, for example, for development of programmes to learn the language. (See also paragraphs 5.197-5.200.)

3.561. Information about knowledge and usage of the mother tongue of indigenous populations is very important. This information could allow obtaining statistics of “indigenous” languages and would be very useful for development programmes to support the development of such languages.

(f) Statistics on poverty

3.562. The census can provide a valuable source of information on the living conditions of households as a proxy measure of poverty to complement quantitative survey data. Census data can provide a quantitative approach to measuring poverty. In some cases, countries may compile multiple deprivation indices using census data such as tenure, economic activity status, income, and household amenities.

5. Development indicators

3.563. Indicators are a good means by which countries may track the progress of various developmental goals and, as such, efforts should be made by NSOs to produce relevant indicators to meet this need. In the 2020 round of population and housing censuses, many countries produced indicators based on the Sustainable Development Goals (SDGs) as was then recommended. The type of indicators required to meet international and national reporting requirements need to be taken into account early in the planning phase of the census.

3.564. It should be emphasized that both global and national reporting and monitoring require reliable and comparable national data for the compilation of indicators. In this regard, it is of paramount importance that countries have the statistical capacity to produce, analyse and disseminate the requisite data for these indicators. The availability of reliable statistics and the

capacity of governments to systematically measure and monitor indicators is a critical success factor for the achievement of development goals. The lack of statistical capabilities in some developing countries makes it difficult to obtain good and reliable data. Many countries do not have a sustainable, coherent programme of household surveys, or administrative data systems that can be used to produce basic statistics routinely. Where basic statistical systems are not available, the global monitoring may have to rely on national and international estimates of widely varying quality and reliability. This may lead to misjudgements regarding progress and may undermine the effectiveness of policy interventions at national and subnational levels.

3.565. The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, provides a shared blueprint for peace and prosperity for people and the planet, now and into the future. At its heart are the 17 SDGs, which are an urgent call for action by all countries – both developed and developing - in a global partnership. They recognize that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests.

3.566. The 17 SDGs are divided into 169 accompanying targets and over 230 indicators, many of which are subject to disaggregation, and which pose a major challenge to the data generation capabilities of NSOs. In some cases, SDG indicators can be directly generated from the population and housing census which allows for disaggregation of the indicators at the sub-national level and for sub-populations that are harder to report on through sample surveys. (For reference and further detail see *Technical Report: Measuring Sustainable Development Goals Indicators through Population and Housing Censuses and Civil Registration and Vital Statistics Data* – UNSD Oct 2022).

X. Overall evaluation of the census

A. Importance of evaluations

3.567. A population and housing census consists of a complex series of interrelated steps, and constitutes perhaps the single most extensive, complex and expensive statistical operation that a country undertakes. Like any other project, the census needs to be evaluated to ensure that the effort and investment of resources have been worthwhile. Evaluation of the census for coverage and also to assess the quality of the data is covered in Part Four, Chapter II, and quality assurance for the census in Part Four, Chapter I. This section is concerned with the overall evaluation of the processes and procedures of the census operation.

3.568. Evaluation of the census is important for a variety of reasons, including for providing transparency, accountability, and valuable lessons learned for future censuses. For this reason, evaluation is generally regarded as the last stage of the census cycle or the first step in the next census cycle. Evaluation assesses the effectiveness of operations, systems and processes and their likely impact on data quality. In this connection, it is particularly important to assess various aspects of the census operation, particularly when changes have been introduced compared to the previous census. Assessments provide valuable information on the strengths and weaknesses of past operational procedures, which should be carefully reviewed prior to the development of the next census.

3.569. In addition to the evaluation of the conduct and operational elements of the census, it is valuable to evaluate the use and value of census data and products. Evaluation of the outputs of the census can be conducted through the conduct of user surveys to gain subjective feedback or through looking at metrics around product use, for example number of visits to web products or the number

of publications requested or purchased. Evaluation of census products can be used to measure whether the data and the selected topics are sufficiently accessible, timely, relevant, accurate, coherent, trusted and interpretable to meet user requirements. The outcomes of the evaluation might provide information that can be immediately acted upon and remedied in the intercensal period, and should certainly feed into the preparations for the subsequent census.

3.570. In recent years, countries have introduced new methodologies and technologies in conducting censuses in order to reduce census costs and also to improve the quality and timeliness of census data. Plans should be made for thorough evaluations of the effectiveness of the new technologies so additional improvements can be implemented for the next cycle. There is use of new methodologies based on administrative registers and combinations of sources to produce information throughout all phases of the census. Plans should be developed to assess the quality of potential administrative sources. The added value of including these new administrative sources and the potential improvements in quality should be assessed alongside the risks associated with these new sources (for example, the status of the data supplier and the level of commitment of the data suppliers to the dimensions of quality). Other changes relate to the use of multi-modal enumeration methods as well as substantial outsourcing of some aspects of census operations. Each new methodology, technology, and administrative source needs to undergo quality assessment so that a comprehensive risk profile and risk management plan can be established for the census design.

3.571. The evolving nature of census methodologies, technologies and operations underscores the need for rigorous evaluation to assess the impact of recent innovations. Furthermore, evaluations are necessary in order to provide lessons learned not just for the countries concerned but also for those who want to adopt similar processes for their future censuses.

B. Planning for the evaluation

3.572. Evaluation of census processes should not be undertaken on an ad hoc basis. The evaluation programme should be included in the overall census plan and be appropriately budgeted for. Lack of proper budgeting and planning for activities that come after the enumeration phase, such as the post-enumeration survey, has in some cases led to financial shortages with negative consequences for the activities concerned. The success of the programme of evaluation depends to a large extent on setting, early enough, clear and unambiguous objectives to be served by the evaluation. Having clear objectives would help to design the best way to undertake the evaluation.

3.573. Ascertaining the objectives of the evaluation encompasses other aspects as well. It is important to establish the scope of the evaluation in terms of intended objectives, keeping in mind that the wider the scope the more complex the evaluation is likely to be. It should be noted also that evaluation of census processes and procedures could last several years and may cover different aspects of the census operation, as necessary, thereby resulting in a series of reports. These reports should be shared externally, as appropriate, to maintain transparency about the evaluation results and how those results can be used to improve the future design. As for other census activities, there is a cost associated with the evaluation and the more complex the undertaking or the more processes get evaluated, the higher the cost is likely to be.

3.574. In addition to the financial cost, human resources and skilled staff to undertake the evaluation should also be planned for. Depending on what aspects of the census operation will be evaluated, NSOs should ensure that they have adequate personnel (in type and quantity) to undertake and complete the evaluation. A related issue that should be taken into account is the feasibility of the staff to undertake the evaluation. Lack of requisite skilled staff may limit the ability

of the country to undertake some or all of the planned components of the evaluation. An alternative is to hire temporary staff with the required technical skills to conduct the evaluation or to collaborate with other agencies such as research institutions and international organizations. NSOs should consider engaging external experts with open access to census data, and records to conduct the evaluation.

3.575. As has already been stated, plans for the evaluation of census processes and procedures should be an integral part of the overall census plan and be planned for from the start of census activities. In addition, documentation of the census experience should have the provision of information for evaluation as one of its objectives. As presented in paragraphs 3.576–3.580 on "Documentation of census experience", the census organization should have a knowledge management system to document and archive complete records on plans, activities, and decisions taken during the entire census operation, including on problems faced and how they were resolved. Documented evidence on how the census was undertaken provides valuable input for the evaluation programme. It should be noted, however, that depending on what has to be evaluated, some of the information may not be available until the end of the census operation.

XI. Documentation of census experience

3.576. The cumulative experience of previous censuses in any country is very useful in the preparation of a new census. Because of the lapse of time between censuses (generally 10 years) and the likelihood that experienced staff may have left the census office during the intercensal period, it is essential that there is a comprehensive record of how the census was planned, organized and conducted.

3.577. The NSO should, therefore, plan for and implement a knowledge management system to assemble complete records on plans, activities, decisions taken, challenges faced and lessons learned during the entire census operation. This would entail documentation and archiving of information related to plans and their implementation, as well as problems encountered and how they were resolved at each stage of the census cycle. It is recommended that documentation of census experience be undertaken at each stage of the census operation and not be left until the end of the census process. This would include plans, decisions and activities related to preparatory activities, the methodology of the census, fieldwork or other data collection activity, data processing, analysis, dissemination, cost and implementation of the census budget, and evaluation of performance of each of these activities. Examples of items to track or monitor include implementation of activities, time taken to complete an activity, resources used and cost. All these should be assessed against set goals so that changes to plans can be recorded, including information on what changed and why. Tracking and systematically recording the census experience should also take into account risks encountered and how these risks were managed. For more information see paragraph 3.570 in Part Three on "Overall evaluation of the census".

3.578. Use of knowledge management tools and techniques is thus beneficial for preserving institutional memory in a codified way so that lessons learned from the past may be used for better management of future census planning and execution. Records in the system should be arranged in such a way that information on each aspect of the census operation is found easily.

3.579. Systematic recording of census experience is not an end in itself. It is recommended that every country prepare and, if possible, publish an administrative report, as a census "historical memory", based on information that has been recorded in the knowledge management system (see paragraphs 3.418–3.419). Depending on the methodology of the census, the administrative report should contain information on the method(s) by which the census was planned, organized and

conducted, as well as important methodological and other problems encountered at various stages of the programme. As appropriate, the report should provide specimens of the census questionnaires and forms, instructions for the enumeration, and detailed information on the cost of the census and on the implementation of the census budget, as well as points to be considered in future censuses.

3.580. The structure of the report could be similar to the structure of the project plan. It is important that the report be as comprehensive as possible, covering all stages and aspects of census planning and operations, including fieldwork, processing, analysis, dissemination and evaluation. It is important to note that while such a report would be based on items and information in the knowledge management system, it may not necessarily contain detailed descriptions of all the processes or information as some of these may be for internal use only. Such a report would both assist the users of the census results in appraising and interpreting the data and facilitate the proper planning of future census programmes.

XII. Archiving individual records

A. Purpose of archiving individual records

3.581. The focus of this Section is on the archiving of individual census records irrespective of their format – paper questionnaires or electronic records. The discussion on the value and process of retaining other census documents relating the systematic recording of the census experiences is presented in paragraphs 3.576–3.580.

3.582. Individual census records refer to the actual questionnaires used to record the information directly from the population whether they were in either paper or digital format and where the information therefore includes individual identifiers, such as name and address. In the case of bimodal or multimodal data collection, that is, a combination of paper and non-paper questionnaires as, for example where Internet forms and mail out/mail back paper questionnaires were used, the resulting collection of individual records would also comprise a combination of paper and digital recordings.

3.583. In the case of digital records, some accompanying documentation becomes an indispensable part of the archiving process. As a number of variables in the digital record are presented as codes, it is necessary to archive all the codebooks and all the other documentation, such as the data collection instrument, that are needed for fully utilizing the value of each of the variables in the record. While this accompanying documentation is also valuable in the case or archiving paper questionnaires, these are by nature visual, thus requiring only reading skills and the knowledge of the language initially used for completing them to understand the content, providing that they are in a good state of preservation.

3.584. Whatever the means, the essential purpose of archiving individual census records is to keep them safe for future use, primarily in the domain of genealogical research and longitudinal social and anthropological studies, as well as for use by historians and demographers. The release of archived individual census records is subject to a passage of time prescribed in census legislation and that usually encompasses many decades, thus ensuring that the use of individual information would not endanger the confidentiality and the privacy of the respondents.

3.585. Consequently, the importance of providing detailed guidance on the process of archiving individual census records in the census legislation cannot be overstated. These provisions provide the legal basis for maintaining the archives and procedures related to the release of archived

records. The time lag between the data collection and the release of the archived records needs to be clearly indicated – internationally it varies from 72 years (in United States of America for example) to 100 years (in the UK). In some cases, the original questionnaires are only temporarily stored before being fully destroyed before the next census takes place (one year before as in the case of India for example).

B. Procedures for archiving

3.586. Archiving a vast number of records represents a considerable challenge in all circumstances. In the case of individual census records it may be compounded by the sheer number and format. However, in all cases the NSO needs to develop an institutional strategy for archiving, based on three components: organizational infrastructure, technological infrastructure and resources.

3.587. **Organizational infrastructure** refers to the arrangements that need to be put in place within the NSO in such a manner so as to ensure the efficiency of the archiving and eventual retrieval process. In most cases a centralized unit within the office that is responsible the archiving, maintenance, secure storage and eventual release[159] of individual records. Once the time lapse mandated by the law for the release of records expires, the actual release to the public may be implemented by dispatching relevant batches to the libraries covering parts of the country to which the records refer and/or to a central depository such as a National Record Office.

3.588. **Technological infrastructure** refers to the actual technology used for archiving. In the contemporary circumstances, storing huge numbers of paper questionnaires would prove not to be cost-effective, as it would require not only a significantly large space but also a physically secure structure, regulated temperature and humidity, and a host of other requirements, including protection from fire hazards, floods and extreme weather events. Consequently, in most cases the paper questionnaires are scanned and images of them stored in various electronic storage devices.

3.589. The technological infrastructure does not refer only to the actual process used in the archiving process – it also consists of a series of protocols for archiving and establishing cross-references that enable the successful retrieval of records. Therefore, the technology should be built around a well-developed archiving scheme that enables efficient identification and retrieval of the individual records.

3.590. In the case of some more recent censuses, in countries where digital and paper questionnaires have been available interchangeably to respondents, data from paper questionnaires have been stored digitally and the paper forms destroyed, with the resulting digital census data subsequently kept in secure cloud storage alongside accompanying metadata and documentation. It is best practice for such digital archives to be backed up in physical and cloud formats, with regular checks to ensure the records remain stable, accessible and usable over time.

3.591. **Resources** must be allocated for effective archiving. Effective archiving requires long-term planning and careful consideration of the technological and organizational infrastructure needed as part of the overall census strategy. Assessing the necessary funding needs, strategically, to recognize that archiving, maintaining, and releasing census records is an ongoing activity that will continue for

[159] Or the transfer of individual records to the National Archive or Public Records Office for public release.

as long as censuses are part of the national statistical system. This necessitates independent, long-term funding in addition to the budget allocated for any one particular censuses.

C. Archiving microdata

3.592. Individual census records as described above for archival purposes differ from the census microdata in a most significant manner: they retain the direct identifiers – name, address, enumeration area – as these very identifiers represent the essential information necessary for genealogical, anthropological, historical and longitudinal social studies. In the case of microdata, these identifiers would be removed, as well as any other information from which any individual can be directly or indirectly identified.

3.593. Microdata are defined as electronic records pertaining to each unit of observation; in the case of the population and housing censuses, it would be individuals, housing units and households. While names and addresses are the most obvious identifiers, certain housing unit characteristics, especially in combination with other information, could increase the risk of re-identification. NSOs should carefully consider the level of detail provided in housing unit data within microdata files to ensure confidentiality. This may involve techniques like data aggregation or suppression to minimize the risk of re-identification.

3.594. It is expected that the use of anonymized microdata will continue to be a standard feature of census data dissemination for the 2030 round of censuses. Consequently, paragraphs 3.468–3.488 present a comprehensive elaboration of principles and protocols for the dissemination of microdata files.

PART FOUR. QUALITY ASSURANCE

I. Quality assurance

A. Plans for quality assurance

4.1. Regardless of census methodology used – full field enumeration, register-based or combined – the visibility and significance of a census plays a large role in the reputation of a NSO's capacity to produce and disseminate high-quality data. All countries use the census as the base to project future population for planning purposes. Irrespective of the type of data collection the census also provides the sampling frame for future surveys. Furthermore, a high-quality census increases the trust of data users from all sectors. Thus, the quality of a census has serious implications for long-term data quality, planning and evidence-based policy making as well as upholding the reputation of the NSO.

4.2. However, since most countries conduct population and housing censuses only once every 10 years, retaining institutional knowledge from one census to the next becomes challenging. But the experience gained from previous population and housing censuses as well as other censuses, such as an agricultural census, is very useful in planning for a quality assurance and improvement programme for the current census. It is of critical importance that errors detected from previous censuses or similar activities are documented and used as the basis for developing quality assurance measures for the next census. Each country must have a quality assurance and improvement programme in place to measure the quality of every phase of the census operation. Such a programme should also be viewed as a quality *improvement* programme. Errors and inconsistences detected in such a programme can then be addressed promptly. If data are of poor quality, decisions based on them can lead to costly mistakes. Eventually, the credibility of the census and even the national NSO itself may be called into question.

B. Quality assurance components

4.3. Quality assurance[160] should be viewed not only as an important element of any census, regardless of the data collection methodology but also within the broader context of managing both the statistical system and the institutional environment. The quality assurance and improvement programme should be developed as part of the overall census project and integrated with other census plans, schedules and procedures. The programme should be established for all phases of census operations (including planning, mapping, questionnaire development, data collection, processing, evaluation, analysis, and dissemination) based on the United Nations National Quality Assurance Framework (UNNQAF)[161]. Establishing a quality assurance and improvement system and setting quality standards at the planning stage is crucial to the success of the overall census operation.

[160] Quality assurance refers to all the planned and systematic activities implemented that can be demonstrated to provide confidence that the census will meet requirements for quality. Quality control is also referred to in this document and relates to the operational techniques and activities used to meet requirements for census quality.

[161] https://unstats.un.org/unsd/methodology/dataquality/un-nqaf-manual/

4.4. The UNNQAF for Official Statistics provides a framework for addressing quality within the context of official statistics and contains 19 principles that are relevant for quality management in population and housing censuses. These quality principles and associated requirements consist of four levels, ranging from overarching institutional and cross-institutional management through to the management of statistical production processes and statistical outputs:

1. Managing the statistical system.
2. Managing the institutional environment.
3. Managing statistical processes.
4. Managing statistical outputs.

1. Managing the census within the national statistical system

4.5. Three preconditions for the efficient production of high-quality census statistics are: a national statistical system that supports the coordination of census activities with other parts of that system; the management of relationships with stakeholders; and an enabling legislative environment. Coordination of this system and managing relations with all stakeholders is a precondition for the quality and efficient production of official statistics. Ensuring the use of common statistical standards throughout the system is an important part of this management. Due consideration should be given to the following principles:

i. Coordinating census activities within the broader context of the national statistical system.
ii. Managing relationships with data users, data providers, and other stakeholders.
iii. Managing statistical concepts, definitions, classifications, methods, and procedures.

2. Managing the institutional environment

4.6. Quality will be better supported by sound institutional arrangements that assure the following principles:

iv. *Professional independence*. Census results should be developed, produced, and disseminated based on technical decisions without any political or other interference or pressure.
v. *Impartiality and objectivity*. The census should produce and disseminate statistics in a way that is, professional, impartial, and unbiased.
vi. *Transparency*. Census policies and management practices should be well-documented and publicly available.
vii. *Statistical confidentiality and data security*. Information provided by respondents should be kept confidential and will be used for statistical purposes only.
viii. *Commitment to quality*. Census managers should be dedicated to assuring quality in their work, and the work of their teams, and systematically and regularly identify strengths and weaknesses to continuously improve census processes.
ix. *Adequacy of resources*. The financial, human, and technological resources should be adequately planned for both in magnitude and quality and be sufficient to meet census objectives regarding the development, production, and dissemination of census-based statistics.

3. Managing census processes

4.7. Quality is the outcome of sound statistical processes, and deficiencies in quality (for example, delays in processing or lack of accuracy in the results) are usually the result of deficiencies in the process rather than the actions of individuals working in that process. Census processes should be based on the following principles:

x. *Methodological soundness*. This encompasses adherence to professional methods and (internationally) agreed standards, guidelines, and good practices.
xi. *Cost-effectiveness*. Resources should be effectively and efficiently used, and census objectives achieved at a reasonable cost.
xii. *Appropriate statistical procedures*. Effective and efficient statistical procedures should be implemented throughout the census production chain.
xiii. *Managing respondent burden*. The level of burden put on respondents who provide census data (where these are collected in the field) has an impact on the quality of such data. Minimizing respondent burden is important, requiring a careful balance between data needs and questionnaire length. Not only do register-based censuses reduce direct respondent burden, but maximizing the use of administrative and other data sources is a key element in optimizing data collection efficiency.

4. Managing census outputs

4.8. It should be the aim of any NSO that the outputs of any census should comply with all the following quality dimensions:

xiv. Relevance and completeness
xv. Accuracy
xvi. Timeliness and punctuality
xvii. Accessibility and clarity
xviii. Coherence and comparability
xix. Interpretability
xx. Linkability (for data derived from any administrative source)

These dimensions are expanded upon below in paragraphs 4.104.17.

4.9. In the census context, the importance of some dimensions of quality assurance may be emphasized over the rest. For example, the census should always produce statistics that are relevant to data users who may consider coherence and comparability with other data sources less vital. A census with full field enumeration is a particularly expensive exercise to undertake and creates a burden on respondents. Therefore, it is essential to ensure that any demand for data is met to a minimum acceptable level and that questions related to topics for which there is little demand are not included on the census form. As noted at paragraphs 2.175-2.178, consulting with users of census data as one of the first steps in designing the census process is a positive public relations undertaking and an efficient, transparent means of determining the demand for potential census data.

4.10. The ***relevance*** of data or of statistical information is a qualitative assessment of the value contributed by these data. Value is characterized by the degree to which the data or information serve to address the purposes for which they are produced and sought by users. Value is further characterized by the merit of those purposes, in terms of the mandate of the agency, and legislated requirements. In the context of a census the concept of fitness for purpose as a measure of relevance is important. If it is only necessary that data on a specific variable are available at the broad geographic level (for example, national or major civil division), user requirements could be met more cheaply and effectively through a sample survey or aggregated from administrative records.

4.11. The ***completeness*** of data (defined as the degree to which statistics fully cover the phenomenon or population they are supposed to describe) means that statistics should not only cover the target population as fully as possible but that they should also satisfy the user needs as completely as possible.

4.12. Statistical agencies should develop, produce and disseminate statistics that accurately and reliably portray reality. ***Accuracy*** of data or statistical information is the degree to which those data correctly estimate or describe the quantities or characteristics that the census was designed to measure. Accuracy has many attributes, and in practical terms there is no single aggregate or overall measure of it. Of necessity these attributes are typically measured or described in terms of the error, or the potential significance of error, introduced through individual major sources of error, for example coverage, sampling, non-response, poor response, processing and dissemination. When assessing the accuracy administrative data, representation errors and measurement errors should be considered.

4.13. ***Timeliness*** of information reflects the length of time between its availability and the event or phenomenon it describes, but considered in the context of the time period in which the information can still be regarded as relevant. It typically involves in a trade-off with accuracy in that the quicker the data is disseminated the less time is spent on processes to ensure its accuracy. ***Punctuality*** is defined by the degree to which the availability of the data meets planned and pre-announced release dates.

4.14. ***Accessibility*** reflects the ease of availability of, and access to, the information held by the NSO, also taking into account the suitability of the form in which the information is made available, the media of dissemination, and whether or not the user has reasonable opportunity to know they are available and how to access that information. Any undue cost to the user in acquiring the data from the NSO may also impact on its accessibility. ***Clarity*** refers to the degree to which the data are understood (regardless of how accessible they are), particularly for non-experts.

4.15. ***Coherence*** reflects the degree to which the census data can be combined with other statistical information within an integrated framework over time. The use of standard concepts, definitions and classifications promotes coherence. Equally important is the internal coherence of data, referring to the consistency of information across different topics of the census and census outputs. This coherence is usually addressed through the meticulous development of data edits. Data are most useful when ***comparable*** across different geographic domains, such as between countries or between regions within a country, and over time. More and more emphasis is also put on enabling comparison of geography over time, as well as maintaining consistency and comparison of census topics from one census to another. Another aspect of comparability that is sometimes overlooked is the degree to which data referring to the same characteristics derived from different sources (for example, rates of unemployment as measured by the census and the labour force survey) are comparable.

4.16. For the purposes of achieving ***interpretability***, metadata should be provided covering: the underlying concepts; the definition of variables and the classifications use; the methodologies of the data collection and processing employed; and indications of the quality of the statistical information, to enable the user to fully understand census statistics and any limitations that they may have.

4.17. When assessing the quality of administrative sources for use in censuses, the principle of ***linkability*** (meaning the ability to link records from different sources that relate to the same enumeration unit, and the efficacy of any conflict resolution process) should also be considered. The six standard quality dimensions capture many relevant aspects of administrative data quality, however, alone they are not always sufficient to fully assess the quality of such data.

C. Other international standards, tools and references

4.18. It should be noted that there are other international standards and quality frameworks and tools that may be valuable for quality assessment work:

a) The Handbook on Management and Organization of National Statistical Systems[162], the 4th edition of a series of publications on the management of statistical organizations, was endorsed by the 52nd session of the UNSC in March 2021 as a non-prescriptive compendium targeting chief statisticians, senior managers, and staff members of statistical organizations. The Handbook provides guidance on how to develop and maintain national statistical capacity that is fit for purpose and apprises critical issues and topics, including the integration of innovative data sources and technologies in the production of statistics and indicators. The Handbook includes a dedicated chapter on quality which describes a number of national quality frameworks, guidelines, and tools for potential use by countries.

b) Also, for reference, UNSD maintains an extensive list by country of quality references and tools[163]. In addition, the Maturity Model on Quality Culture in Official Statistics[164] can be used by NSOs in the context of censuses to assess and improve their processes and overall performance.

c) The Generic Statistical Business Process Model (GSBPM) is a model for the business processes that NSO should follow to produce official statistics, conceptualised in terms of phases and associated sub-processes. In order to monitor the quality of the statistical production process for each of these phases, a set of quality indicators have been developed in collaboration by representatives from the NSOs. This work maps quality indicators for surveys to the structure of the GSBPM.[165] (See paragraphs 2.5 and 3.48.)

d) Independent observation of census implementation by experts drawn from other NSOs, regional intergovernmental institutions, and international agencies supporting censuses can provide an additional layer of quality assurance. Independent observations can provide a comprehensive report to underscore the credibility and transparency of the census process.

e) Countries conducting a combined census should identify from the planning stage the elements to be collected from the field enumeration, from registers and the methods for linking this information. As a consequence, certain stages and components could benefit from different quality frameworks such as the Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses.[166]

f) For countries that conduct register-based censuses, the Principles and Recommendations for a Vital Statistics System, Revision 3, contains information for assessing quality. It provides recommendations on some of the practices adopted to assess the quality of civil registration data and the quality of vital statistics based on those data.[167] Additionally, UNSD maintains an inventory[168] of resources on the use of administrative data for statistical purposes. The aim of this inventory is to make materials on the use administrative data for statistical purposes more readily available and easy to find. The inventory includes a variety of materials, ranging from recommendations to practical examples in different contexts.

[162] https://unstats.un.org/capacity-development/handbook/
[163] https://unstats.un.org/unsd/methodology/dataquality/quality-references/
[164] Maturity Model on Quality Culture - 2024-04-23.docx (live.com)
[165] Quality Indicators - Generic Statistical Business Process Model - UNECE Statswiki
[166] https://unece.org/sites/default/files/2021-03/03_CensusAdminQuality_forConsultation_0.pdf
[167] M19Rev3-E.pdf (un.org)
[168] https://unstats.un.org/UNSDWebsite/capacity-development/admin-data/inventory

Furthermore, the UNECE Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses provide a framework and practical tools for evaluating the suitability of administrative data for use in population and housing censuses.[169]

D. The role of managers

4.19. Managers have a vital role in setting quality standards and quality management as a priority in the census (See Section III of Part Two on Strategic objectives and management). The biggest challenge to managers is first to establish a culture within the NSO that has a focus on overall quality management issues and to obtain the commitment of staff to strive to achieve high-quality goals. Ensuring that high grade managers from the NSO are employed within the census and providing the appropriate responsibility and resources can contribute to this culture. At the same time, such managers need to be aware that to achieve high-quality outcomes they need to give their staff responsibility to achieve these outcomes. Managers who do not delegate responsibility will find it difficult, if not impossible, to establish teams that strive for high-quality outcomes.

4.20. The project manager is responsible for the project work from the initial kick-off through to closure. Only the primary responsibilities are given in the list below, which can be expanded considerably and are covered in more detail in Section VIII of Part Two on project management. The responsibilities of the project manager include:

a. Using quality project management tools and techniques;
b. Identifying and managing the project stakeholders, both those inside the NSO as well as external groups and organizations that can impact the project outcomes;
c. Creating the conditions for good team working;
d. Establishing a project support office;
e. Identifying and managing the project risks by developing a risk register, establishing a risk prioritization and risk mitigation plan, and monitoring and managing these risks;
f. Working within a changing environment and managing unpredictable events as they occur, including assuring professional independence of the project staff and demonstrating objectivity through transparency;
g. Implementing improvement and change initiatives to ensure the soundness of methodologies, the cost-effectiveness of resources, the appropriateness of statistical procedures and the proactive management of respondent burden;
h. Delivering the project deliverables and benefits;
i. Leading the project team;
j. Evaluating and closing the project.

4.21. Firstly, managers must ensure that staff understand the philosophy behind the approach to quality. As noted above in paragraph 4.19, staff involvement is a vital ingredient to quality improvement. Therefore, an environment needs to be established in which staff contributions are expected and staff members are supported to develop an understanding of the linkages between their functions and those of others with the census project.

4.22. The second part of their role is to ensure that users' data requirements are understood, and that these are met (particularly in terms of relevance, accuracy, and timeliness) and built into planning

[169] United Nations Economic Commission for Europe (UNECE) (2023). Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses. https://unece.org/statistics/publications/CensusAdminQuality

objectives and into the systems that are to deliver them. Managers may have to work with multiple users simultaneously to satisfy project objectives and this complexity should be considered when building systems and during project planning.

4.23. Thirdly, processes need to be documented and understood by the staff implementing them. Systems and processes for implementing the "quality assurance circle"[170] also need to be documented and put in place. Questions such as how quality is going to be measured, who is involved in identifying root causes of problems with quality, and how the process improvements are going to be implemented need to be answered. These will vary greatly depending on the nature of the process. Appropriate quality assurance techniques for each phase of the census are summarized Section E below.

4.24. These processes and metrics should also be shared with staff as appropriate, to demonstrate how quality expectations are being met (see also paragraph 4.27). The greatest test of management commitment to genuine quality improvement lies in how management approaches problem-solving. Ultimately, staff will observe and emulate management's behaviour, particularly in their approach to problem-solving, more than they will adhere to stated policies or verbal commitments to quality improvement.

4.25. Managers who always react to problems by seeking out people to blame, or who establish systems that focus disproportionately on the merits or demerits of individuals at the expense of the team, are sending messages that are contrary to the thrust of quality improvement. An environment where the emphasis is on fault finding, rather than on finding solutions to problems, or on excessive competition, will ensure that staff ceases to be part of the solution and becomes part of the problem. Managers therefore need to take upon themselves the responsibility for problems, as they are ultimately responsible for the systems or processes that caused them. They should not seek to transfer the problems to lower-level staff.

4.26. However, even in the best-managed processes, there are circumstances where individuals can be justifiably responsible for negatively impacting on quality. These may be individuals who do not possess adequate skills for performing their duties, and even deliberately flout procedures. These individuals need to be dealt with decisively, first and foremost by providing additional training and guidance all the way to administering disciplinary measures. Managers must deal promptly with these cases and act in a consistent manner. By doing so, managers will demonstrate to all other staff their commitment to quality.

4.27. To be successful, it is necessary to create a culture in which everyone has the opportunity to contribute to quality improvement. Most of the staff engaged in census operational work undertake routine tasks, and it is up to management to help them see the bigger picture, to motivate them and to encourage them to assume ownership of their work. This can be done by promoting a commitment to quality improvement and by adopting a consistent approach to management.

[170] The quality assurance circle is a continuous feedback loop designed to prevent errors and improve processes, essential for maintaining data quality in census operations. It involves detecting errors early through various checks, analysing their root causes, providing feedback to workers, implementing corrective actions, and continuously monitoring the process. This cycle ensures that errors are identified, corrected, and used as opportunities for learning and improvement, maintaining high data accuracy and consistency.

E. Quality improvement

4.28. The following sections outline the way in which the concepts of quality relate to the different phases of census activities[171]. As a basic premise, it is critical, in order to build trust and confidence with stakeholders, for documentation on the various quality considerations to be provided as soon as feasible. This can be done on a public website, presented during a presentation, or other common methods used to provide information about the census. Quality considerations are discussed in relation to:

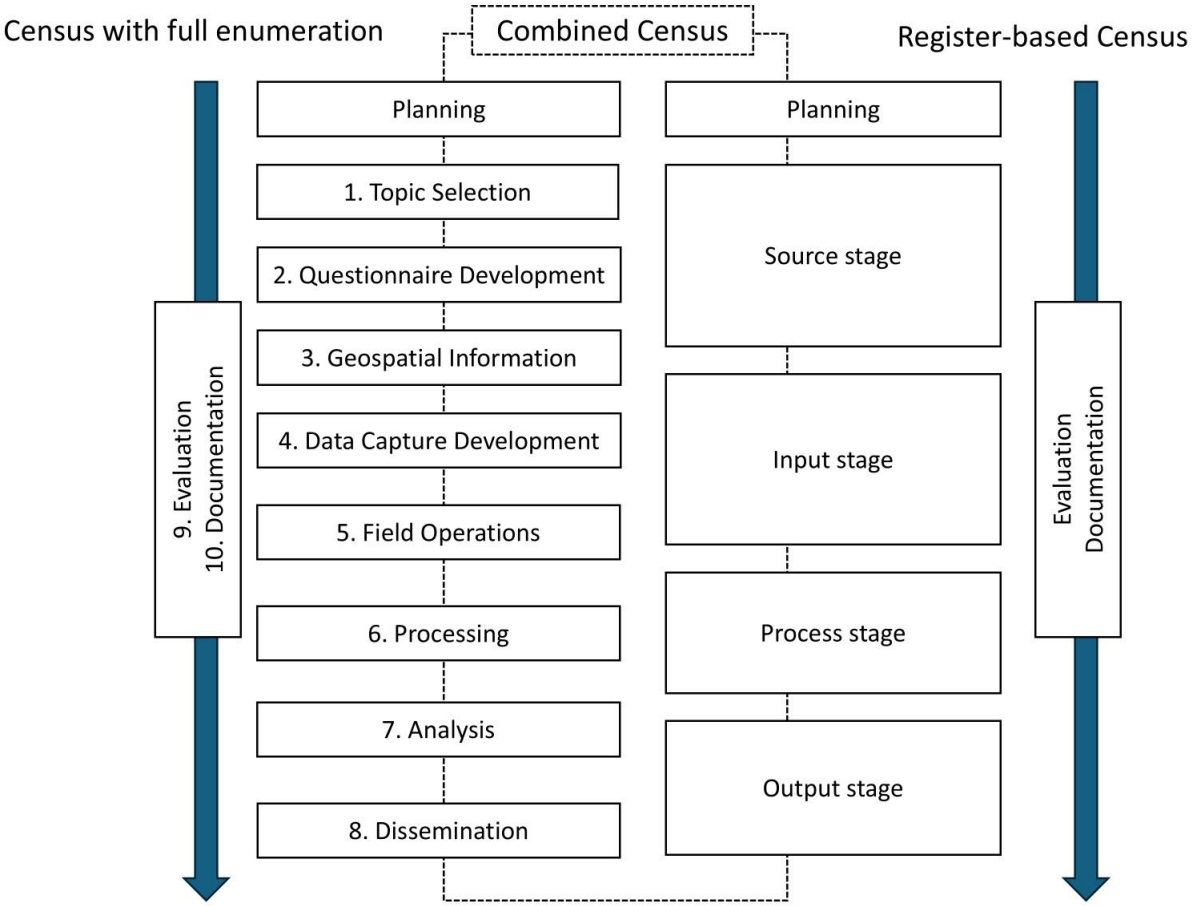
1. Topic selection;
2. Questionnaire design, development, and processing system testing;
3. Geospatial information;
4. Development of a data capture system; (including CAPI and CAWI);
5. Field operations (including pilot testing);
6. Processing;
7. Analysis;
8. Dissemination;
9. Evaluation; and
10. Documentation

4.29. While the quality considerations for these 10 stages are relevant to a full field enumeration census, some of the stages also apply to combined and register-based censuses (see Figure 4.1). The additional quality considerations related to data derived from administrative sources are outlined in the Handbook on Registers-based Population and Housing Censuses[172] and are covered in Section F below. In some contexts, the first step in ensuring a high-quality census is the development of the census project document. Quality considerations should begin at the earliest stage of planning.

[171] The Generic Statistical Business Process Model (GSBPM) also addresses the various stages of the statistical production process. There is a significant overlap between the steps outlined in this section and the GSBPM.

[172] https://unstats.un.org/unsd/demographic-social/publication/handbook-registers-phc.pdf

Figure 4.1: Overview of Quality Processes Outlined in Part 4, Chapter I, Sections E and F



1. Topic selection

4.30. The first step in managing the quality of the product (namely, the statistical outputs to be produced) is to ensure that it will be relevant to the users' requirements. The key process is extensive consultation with actual and potential users of census data. Potential topics for inclusion in the census should be agreed early in the census planning and questionnaire development phases. The key success factor in this process is full, frank and open consultation with users, and with subject matter and classification experts to ensure that census outputs address well-defined user needs. If their needs are not met from the current census products, users may be reluctant to propose their needs for a future census. It is important to note that such needs will change over time, and users should be kept informed of changes to variable definitions and output classifications. A review of the extent to which user needs have previously been met should be seen as part of the evaluation process of the preceding census feeding into the current cycle, and the first step of quality management.

4.31. When considering the topics to include in the census countries should consider their context, the availability of alternative sources, the burden on respondents, privacy concerns, and the sensitivity of questions. Operational considerations such as the cost of collection, the ability to generate high-quality data, and operational feasibility should also be taken into account.

4.32. For further details and a description of the processes involved in topic selection, see paragraphs 3.7-3.8 and the list of recommended topics in Part Five.

2. Questionnaire design and development

4.33. The next quality management task concerns the testing of each census question and of the design of the form (paper or electronic version according to instrument(s) used). The development and testing of paper forms and of electronic versions (including apps) should be treated as separate, but parallel, processes each with their own set of quality indicators. The results of each test should be analysed and evaluated before being subject to further design and testing. For further details and a description of the processes involved in questionnaire design and development, see Section II of Part Three.

4.34. The following teams are the key internal stakeholders of the form design process:

(a) The analysis and dissemination teams who will ensure that the questions asked will aim to deliver the data to meet the needs of users);
(b) The subject matter specialist team, who will provide expertise on the topics covered in the census and ensure the questions are clear, relevant, and culturally appropriate;
(c) The team responsible for development of the data capture system. For example, positioning of the text and the delineation of response areas may be dependent on the data capture and processing methodology to be adopted. It is important that there is ongoing coordination between the form design and processing areas;
(d) The field operations team that is responsible for training the enumeration workforce and printing the form.
(e) The translation team (for countries with multiple official languages)

4.35. The process of testing the response options employed in the questionnaire is pivotal to ensure their alignment with both national and international requirements. This alignment guarantees that the subsequent output classifications can be effectively compared and seamlessly linked to other databases. This meticulous testing procedure is imperative to guarantee the accuracy of the data. By harmonizing the classifications, the foundation is laid for cross-referencing and integration with external datasets thereby maximising the coherence of the data. Consequently, this endeavour facilitates a comprehensive analysis and provides valuable insights into various quality dimensions of the collected data. Furthermore, the rigorous testing of these classifications ensures the apt design of the data structure for the census outputs. A well-structured data framework enhances the usability and interpretability of the census results, and facilitates accessibility by ensuring data is provided in formats suitable for diverse users. This approach ensures that the data outputs are not only accurate but also optimally organized for analytical purposes, enabling more informed decision-making processes.

3. Geospatial information

4.36. Many of the quality assurance components listed earlier (in paragraphs 4.3-4.17) are also applicable for mapping operations, geospatial products, and their downstream uses in the census. For example, processes that improve the completeness or accessibility of geospatial data (such as identifying new housing units or enabling user download of geographical areas) support quality. The dual role of geospatial data in providing the operational framework for pre-census operations and the means to associate statistical observations with recognizable areas during dissemination, leads to some noteworthy quality considerations. These considerations relate specifically to the capacity of geospatial data to:

(a) Deploy and route field staff and support material efficiently;
(b) Define assigned collection and supervisory areas unambiguously; and `
(c) Associate tabulated census data with administrative and other useful geographic areas.

4.37. For further details and a description of the processes involved in geospatial information and mapping see Section IV in Part Three.

4.38. Ensuring the quality of the demarcation of areas requires rigorous testing during census planning and design that assesses the efficiency of statistical unit design for operational purposes. The exact rules used to define physical features (such as a river or road) and the target number of households or persons for each EA will necessarily vary by country. A high-quality demarcation helps to predict the rates at which enumerators should collect data in different types of terrain during the field operation. These rates are estimated during census pre-tests and a full census pilot that use planned census modalities within varied geographic areas.

4.39. The quality of spatial attributes associated with geospatial data supports quality by improving interoperability and reducing duplication of effort. The use of harmonized classification and coding systems for qualitative characteristics (such as road quality, river type, urban status) and the administrative geocoding naming scheme (that is, a gazetteer) within the NSO and other agencies with geospatial functions supports data coherence in cartographic products used to support both enumeration and dissemination.

4.40. The sophistication of the technology suite used during a census does not singularly guarantee the quality of geospatial data produced. Progressively more advanced technologies may support improved quality as they allow geospatial data to achieve improved relational and positional accuracy. However, the technological capacity to associate geographic coordinates with housing units, using CAPI for example, does not solely ensure the integrity of the census or that a household frame can be generated from the census. Coordinate capture must be part of a coherent, reproducible census methodology with measured positional accuracy and coverage error rates to meaningfully support census data integrity. NSOs must also have the financial and human resources to maintain the frame during the intercensal period, otherwise the potential quality gains for future statistical processes from increased technological sophistication during the census are greatly diminished.

4.41. Programs that reduce ambiguity when interpreting boundaries contribute to quality by improving accuracy and enhancing reproducibility during data collection (so that, for example, a field worker assigned to a statistical unit based on census geography unambiguously knows which housing units are contained within, or lie outside of, their unit with as little re-canvassing as possible). An ongoing program to harmonize geographic data should coordinate between the NSO, other custodians of geospatial data, local governments, and any local statistical offices. Since most NSOs are not the custodians of all administrative boundaries, this coordination ideally takes place within an NSDI framework.

4.42. An ongoing and consistent approach to geographic area and boundary maintenance is not always possible in resource-constrained environments. The census may be a rare opportunity to improve the completeness and congruence of administrative/statistical geographies and physical features. Collection and collation of metadata during the census process regarding how cartographic staff and field personnel resolve boundary questions will support both the integrity of the census and data maintenance efforts during the intercensal period. These measures can be as simple as encoding the rules used to prioritize physical features or ensuring that census field notes are digitally captured and saved rather than discarded once an issue is resolved.

4.43. Programmes that improve data user familiarity with geographic areas, especially functionally defined areas, improve quality by making the data more relevant and interpretable for their specific needs. NSOs may even solicit data user feedback to ensure that the geographic areas used for census tabulation continue to meet their needs. Census GIS systems may be used to provide data for other

customized geographies required by other government entities. To the extent possible, these requirements should be accounted for when creating collection or dissemination geographies.

4. Development of the data capture system

4.44. The transition from paper to digital data collection introduces a new operation, that of developing an electronic questionnaire. In this section, specific quality concerns related to the development of Computer Assisted Personal Interviewing (CAPI) applications and Computer Assisted Web Interviewing (CAWI) applications are discussed.[173, 174] Additional material on quality considerations when capturing data from administrative records is covered in section F below.

4.45. For further details and a description of the processes involved in the development of the data capture system see Section VIII.C in Part Three.

a) Computer Assisted Personal Interviewing (CAPI)

4.46. The development of a CAPI application is frontloaded. If such an application is poorly designed, insufficiently implemented, and inadequately tested it will collect low quality data. Several software development methodologies can be applied to CAPI development. These include, creating a questionnaire specification, managing change control, and establishing a cut-off date after which modifications should not be made to the application. Establishing a change control mechanism enables documentation of why and when a specific change is made and who should make it. Revision control software can help facilitate change control. It is important to select a date when all development should stop and after which only necessary fixes to coding and software errors are accommodated. Unlike the traditional paper-based approach, CAPI applications tend to be seen as an ever-evolving piece of software. However, even a simple change can have unintended side effects, be incomplete, or simply be wrong, which in a CAPI environment can result in a small change having a disproportional negative impact. Thus, testing after modifications and retesting after fixing coding issues is crucial in the CAPI application development process. Test cases should be designed to test all potential branches in the questionnaire. Automated testing tools loaded with these cases can reduce the workload of such testing.

4.47. The following are best practices that will improve the quality of the data collected by the CAPI application:

i. Consideration should be given to the development of a field management system to handle tasks that are not directly related to the interview. This will include logging in to the system, the ability to assign and receive assignments, complete and review questionnaires, track progress, and the ability to synchronize data. Once an enumerator is logged into such a system and has selected an assignment, the questionnaire can be prefilled with data such as the enumerator and geographic codes which will reduce respondent error. Additionally, it can be helpful if the field management system is uploaded with paradata such as device ID, interview start time, and software versions.

[173] CAPI involves an interviewer asking respondents questions and CAWI involves self-response.

[174] Many quality assurance considerations for CAPI and CAWI also apply if Computer Assisted Telephone Interviewing (CATI) is used for data collection.

ii. As a minimum requirement each CAPI question will comprise question text and response options. The question text should include personalized elements, such as the name of the household member, customized for each interview to reduce error. This means that the CAPI system would automatically insert the previously recorded name of the household member into relevant questions to ensure accuracy and clarity. Text boxes that allow open-ended responses should be kept to a minimum and used for fields such as name. More often, a discrete number of pre-coded response options should be shown as a dropdown menu. Ideally, the dropdown options include human-readable responses, meaning they are presented in plain language rather than numeric codes. Numeric codes can be recorded in the background without the enumerator having to be familiar with them. Furthermore, response options can be created dynamically, so that an enumerator is only presented with valid responses. Also, help and instruction text can be included to assist the interviewer. The help text will include supplemental information such as definitions, while the instruction text will include reminders such as probing techniques specific to the question.

iii. The flow of the CAPI application and use of consistency checks are critical considerations that will impact data quality. The CAPI application must be able to enforce the sequencing of questions, so that they cannot be asked out of order or missed. The CAPI application also needs the ability to skip questions that are no longer relevant to the current interview based on the responses to previous questions. Consistency checks are helpful for identifying grossly inconsistent data while the enumerator is in the field, but should not prohibit the enumerator from entering data in a timely and unobtrusive manner.

b) Computer Assisted Web Interviewing (CAWI)

4.48. When designing a census using Computer Assisted Web Interviewing (CAWI)[175] it is important to consider factors that will maximize data quality, efficiency for respondents, respondent satisfaction, and a positive census brand experience. Additionally, it will be equally important to consider factors that minimize respondent burden, measurement error, and design inconsistencies across data collection. The size of the screen to display the questionnaire is an important consideration and testing should be conducted across all types of devices that will be used for collection (tablets, smartphones, laptops).

4.49. Improving the quality of data collected in CAWI involves meticulous questionnaire design, robust pre-testing, and proactive data validation to ensure clear, accurate, and reliable responses from participants. Minimizing the time taken to complete the questionnaire means optimizing the interface with the respondent and providing user-friendly features to streamline the completion process. Elevating respondent satisfaction in CAWI can be achieved by prioritizing intuitive design and ensuring transparent data handling practices to enhance respondents' overall experience and trust in the census process. An additional priority is strengthening the census branding experience in CAWI to align the data collection platform with the official branding and messaging of the census. This alignment communicates a consistent and trustworthy identity to respondents.

4.50. Decreasing measurement error in CAWI involves using clear and unambiguous questions, employing thorough pre-testing and validation checks, and ensuring that data quality control measures are in place to enhance the accuracy of collected data. To minimize design inconsistencies across data collection in CAWI, consideration should be given to: maintaining a standardized

[175] Alternatively described as Computer Assisted Self Interview (CASI).

questionnaire template; providing detailed design guidelines; and conducting regular reviews of question order, response options, and visual elements.

4.51. Reducing respondent burden in CAWI is achieved by designing concise questionnaires that request only the required information and employing skip logic to personalize the questionnaire.

5. Field operations and pilot testing

a) Quality considerations during field operations

4.52. The quality management process continues throughout the design of the census field operations. These are tested as far as possible in conjunction with questionnaire design testing. The key internal client of field operations is the data processing team. However, field operations can also impact other areas, such as dissemination and the work of teams responsible for maintaining standard classifications. This is because certain concepts, such as the definition of a dwelling, are implemented during the field operations phase, and these concepts need to be aligned with how data is processed, disseminated, and classified. For further details and a description of the processes involved in field operations see Section VII in Part Three).

4.53. The quality of the field operation depends on the quality of activities that precede it, such as:

i. Demarcation of EAs;
ii. Map production;
iii. Form printing, where a sample of forms is rigorously tested for adherence to required standards.
iv. Development of applications to ensure data quality, such as:
    - *a collection application* with automatic data storage (in the device and online), the necessary control specifications (consistency and completeness checks, automatic skips and alert messages), and GPS coordinates of the household dwelling to check the enumerator's presence;
    - *a control application* for the enumeration supervisor;
    - *a data transfer application* at central level, with all the necessary security features (data encryption, data bandwidth fluidity) to avoid data loss in the field;
    - *a web application/dashboard* for monitoring progress and data quality via real-time indicators for management teams in the field and at central level;
v. Installation of a real-time data centrally-based back-up server and a second replication server at another site as a risk mitigation to ensure against losing data and facilitating data collection continuity in the event of any technical malfunction;
vi. Installation of a call-centre with a toll-free number and a team of teleoperators to answer people's questions about the census in real-time, or the implementation of a website with support chat. Teleoperators can register and report the various problems encountered by field agents. These problems can be reviewed and analysed to identify improvements to the data collection operation.

4.54. All systems supporting data collection must be thoroughly tested and piloted before the actual census. This is especially critical when collection technology is used, such as handheld devices or online forms for self-response. All data quality benefits of using such technology could be compromised if problems arise during the enumeration. Quality monitoring should be established and mechanisms put in place to ensure that the outcomes of the monitoring are used to improve processes. It is more difficult to implement quality improvements during the enumeration itself owing to the very tight time constraints. However, this can be achieved by:

i. Clearly establishing the aims of the field operations phase;

ii. Applying thoroughly documented procedures;
iii. Ensuring that the enumerators understand their role through appropriate training;
iv. Training must consider the monitoring and evaluation of quality indicators during enumeration.
v. Providing opportunities for field staff to be observed on the job so that feedback can be given;
vi. Establishing communication and feedback loops with the general public through helplines, online forums, social media, etc., so that problems in the field can be detected and corrected.
vii. Establishing communication with supervisory field staff through digital field management systems.

4.55. However, it has to be acknowledged that during the actual execution of the enumeration this approach tends to identify "problem enumerators" rather than systemic or process errors. This means that an evaluation following the data collection process is vital. Such an evaluation should attempt to capture the experiences and suggestions of a selection of enumerators and other field staff so that improvements can be made to the subsequent census.

4.56. A general overview of the quality of the enumeration can be obtained through:
i. Use of techniques such as post-enumeration surveys to measure the level of under-enumeration or over-enumeration of people, households or dwellings;
ii. Overall response from the target population or level of non-response at the individual question level;
iii. Consistency and results of the imputation and editing processes observed during the processing stage at the questionnaire and variable level.
iv. Feedback from field staff;
v. Monitoring of the performance of field staff by their supervisors, (or independent observers) especially early in the data collection operation;
vi. Measures of the quality of any coding undertaken by field staff;
vii. Mechanisms that may be in place to handle queries from the public;
viii. Analysis of administrative data used to assist the field operation and historical trends.

4.57. In certain situations, the enumeration schedule may have to be extended in order to achieve improved coverage. Using quality monitoring processes and mechanisms that are already in place, the potential impacts on data quality should be monitored carefully during an extended enumeration period. Shifts in schedule dates could potentially impact on quality in a number of ways:
i. **Transient populations:** The inability to enumerate people in one place before they move to another place, introduces the risk of missing them in the census altogether or counting them in both places, based on the timing of their move. A longer data collection period increases the risk that people move.
ii. **Recall bias:** Any timing shift for data collection activities increases the risk related to recall bias. For example, a longer enumeration period increases the risk that respondents misreport the persons who were living in the housing unit on the census reference day.
iii. **Vacant Units:** Any timing shift in data collection activities increases the risk of a change of occupancy status for housing units, be it vacant, occupied or demolished.
iv. **Training:** Shifts in training to accommodate schedule changes increase the risk that field staff workers would be under-prepared to do their work which could result in data quality issues. Changes in the schedule may increase the need to develop further training to avoid such issues.

b) Pilot testing

4.58. A pilot census is a comprehensive trial or rehearsal that tests all census procedures including personnel systems and procedures and plays a crucial role in census quality assurance. It provides an opportunity to fine-tune procedures, identify and rectify potential issues, and ultimately ensure the accuracy of the census data (see paragraph 3.147).

4.59. A pilot census allows NSOs to:

1. Test all procedures, methodologies, and technologies that will be used in the actual census, including data collection methods, questionnaire design, interview techniques, and data processing systems;
2. Identify any potential problems or challenges that may arise during the actual census such as logistical issues, misinterpretation of questionnaire wording, and to anticipate difficulties in data collection in certain areas;
3. Assess the effectiveness of the training of census staff and enumerators, and their familiarity with expected roles and responsibilities; and
4. Test sampling methods and procedures where relevant.

4.60. Also, the data collected during the pilot census may be compared against known population statistics or other sources of demographic data to assess the efficacy of the census methods. Any discrepancies or errors can then be addressed and corrected before the actual census takes place. Overall, a properly executed pilot census contributes to the improvement of the accuracy and consistency of data collection during the actual census.

4.61. In the case of multi-mode data collection (field, Internet, and telephone for example), it is necessary that all applications to be adopted in the census be tested in the pilot to demonstrate that they are interoperable. Controls should be established for integrating information according to the collection mode to ensure the integrity and uniqueness of the data, allowing for proper tracking of the coverage.

6. Processing

4.62. Within the NSO, the key recipients of the processed census data are the teams responsible for: the dissemination of the census results; ensuring the adoption and consistent application of standard classifications; and preparing specialist subject matter documentation. Effective dissemination requires the processing team to produce data in an agreed format and compiled to agreed quality standards. This is necessary so that the data can be used in dissemination systems. For further details and a description of the processes involved in data processing see Section VIII in Part Three.

4.63. Details on the quality control processes that are necessary when integrating administrative data from different sources is covered in Section F below and have been further elaborated by the UNECE. [176]

[176] Also see chapter 6 (Process Stage) of The UNECE *Guidelines on Assessing the Quality of Administrative Sources for Use in Censuses.*

4.64. Since the census is part of an overall national statistical system, data from it are likely to be used in conjunction with data from other sources. Thus, the classification and subject matter specialist teams that are responsible for those other data collections, need to be satisfied that the coding, editing and other data transformation processes are conceptually sound and deliver data of acceptable quality.

4.65. Extensive testing of processing systems should be undertaken in advance of the census. Coding processes and training packages need to be prepared and tested utilising the type of staff likely to be involved in the census itself. The processing phase gives the ample opportunity for implementing quality improvement techniques, as many processes within this phase are repetitive and take a significant amount of time to complete. It is vital that structures are put in place not only to monitor quality but also to involve processing staff in the identification of problems with quality and in proposing solutions.

4.66. Processes such as editing and imputation may redress some of the inconsistencies and omissions within the data, thus improving data quality.

4.67. However, if a strategy of adjusting census data items after collection to eliminate identified errors and inconsistencies is adopted, then the following paragraphs provide examples of how *data cleansing*, *auditing*, and *output consistency checks* can be implemented. Information about the nature of the checks undertaken and the proportion of records subjected to editing should be published to facilitate interpretability and to maintain user confidence in the census results and the methodologies adopted for such adjustments.

4.68. ***Data cleaning*** is an important step in ensuring the accuracy of the census dataset. This process involves several key components, including:

i. A data validity review to confirm that all entered responses are accurate. This review ensures that numeric fields contain only numerical data, and text fields do not contain numerical data or other inappropriate information.
ii. Methods for handling missing data to ensure that all variables pertaining to the characteristics of the census units are complete. In cases where data is missing, appropriate imputation techniques can be used to fill in the gaps and maintain data completeness (see paragraphs 3.281-3.284 on data editing).
iii. The correction of values outside electronic validation rules. It is important to check that any values outside the scope of the electronic validation rules have been rectified. These values may arise due to defects in the electronic survey during data collection or errors made in the survey process.
iv. The detection of duplicates and a set of rules to support a decision for deleting them.

4.69. ***Validation*** is an important process that involves the review of various aspects of the data to ensure accuracy and consistency of the data. The key components of data validation include:

i. Validation of the coding during the conversion of textual data into numeric codes.
ii. Validation of the data collected to ensure that they adhere to international standards and classifications so as to ensure consistency in data representation.
iii. Range checks to ensure that all answer values of variables are examined to confirm that they fall within predefined limits or ranges. Range checking can identify any outliers or values that may be outside of the expected bounds.
iv. Applying validation rules to review the logical relationships and correlations between variables.
v. Dealing with extreme values by applying various reviews and audits to ensure that any illogical values have been identified and corrected. This involves deciding on the treatment of

unrealistic or unreasonable values and making a determination if values should be edited or replaced with more realistic data points.

vi. Verification that all targeted variables and geographic distributions have been thoroughly and comprehensively covered during the census. This means ensuring that all relevant data elements have been collected and included in the dataset to provide a comprehensive and accurate representation of the population or sample.

4.70. Maintaining the **consistency** of the main census outputs is essential to ensure the coherence of data and is vital for making informed decisions, conducting meaningful analyses, and ensuring that the data accurately represents the real-world context it is intended to describe. It helps build trust in the data and usability for various applications. This involves considering various aspects, including:

(i) Historical comparability: Ensuring consistency with historical data to enable comparisons over time and track trends.

(ii) Geographic uniformity: Using consistent geographic areas and definitions to ensure uniformity across the country and facilitate spatial analysis.

(iii) Integration with other data sources: Aligning concepts and definitions with other data sources to enable data integration and comprehensive analysis.

4.71. To ensure data accuracy it is important to check and cross-check the data collected in the field with reference data from administrative sources, or data previously collected by the NSO. This involves the following steps:

i. *Establishing data validation procedures.* Define specific validation rules or checks for each type of data collected. For example, NSOs might validate numerical values within a certain range, verify consistency between related fields, or check for missing or incomplete responses.

ii. *Match collected data with reference data.* Compare the collected data with the corresponding reference dataset. This can involve matching identifiers, survey responses, or other relevant data points.

iii. *Conduct data reconciliation.* Identify any discrepancies between the collected data and the reference data. Investigate the reasons behind the discrepancies and resolve any inconsistencies or errors.

iv. *Document data validation and reconciliation process.* Keep a clear record of the steps taken to validate and reconcile the data. This documentation helps ensure transparency and reproducibility of the data analysis process.

v. *Define data integration processes.* Determine how the collected data will be integrated with the reference data. This can involve data mapping, transformation, or merging procedures to ensure compatibility and consistency between the datasets.

vi. *Automate data validation and reconciliation.* Whenever possible, automate the data validation and reconciliation processes to minimize human error and increase efficiency. This can be achieved through scripting or programming techniques that compare and reconcile the data automatically.

vii. *Perform data analysis and verification.* Once the data has been validated and reconciled, perform thorough data analysis to identify patterns, trends, and insights. Additionally, verify the results against the reference data and ensure consistency between the analysed data and the original source.

viii. *Conduct data reconciliation in real-time.* If possible, perform data reconciliation in real-time or near real-time. This allows for immediate identification and resolution of any discrepancies between the collected data and the reference data. Real-time reconciliation ensures that any errors or inconsistencies are addressed promptly, reducing the risk of relying on inaccurate or outdated data.

By following these practices, the accuracy of the collected data can be enhanced. These data validation steps also improve confidence in the alignment of the data collected in the field with data taken from administrative data.

7. Analysis

4.72. Processing and analysis can be iterative and parallel. Therefore, some of the quality considerations discussed in the processing section are reiterated here. Analysis can provide a deeper understanding of the data, which could indicate the need for additional processing. In the analysis step, statistics are produced, validated, interpreted, and made ready for dissemination. Also, during tabulation, SDC methods are applied to ensure data privacy and confidentiality, and statistical content to support the interpretation of census outputs is prepared.

4.73. For further details and a description of the processes involved see Part Three, Chapter VIII and Chapter IX.D.2

4.74. Quality assurance should be applied for each of these activities to guarantee interpretability and clarity of the statistical outputs, maintain accuracy, ensure comparability, address coherence with outside sources, guarantee relevance, and ensure confidentiality.

4.75. The interpretability of the statistical outputs is assessed by checking the existence of metadata, commentaries, technical notes and any other necessary supporting statistical information. In order to interpret and explain census statistics, it is necessary to view them from multiple perspectives and understand the extent to which these statistics meet the needs of the intended audience by conducting in-depth statistical analyses, including consistency and comparability analysis.

4.76. An assessment of the impact of any measures taken to ensure statistical confidentiality and data privacy should consider three quality dimensions, (i) accuracy, (ii) relevance and (iii) clarity.

i. Accuracy is assessed by analyzing the impact of SDC methods on the quality of the output and by verifying that such methods do not affect significant relationships in the data. SDC methods must inevitably, by intention, affect the accuracy of the data, but should not do so to an unacceptable degree.
ii. Relevance is assessed by analyzing how well the output meets user needs after applying SDC methods.
iii. Clarity is assessed by describing the SDC method(s) applied (but only to a level of detail that does not itself pose a disclosure risk) and any restrictions on the use of the dataset.

4.77. SDC methods and the importance of confidentiality are discussed in detail in Section IX.B5 in Part Three). The impact of SDC on data quality should be considered when determining the methods to use for census outputs. Minimizing information loss due to suppression, perturbation, or other SDC techniques should be prioritized, while still adhering to the privacy and confidentiality principles outlined in the Fundamental Principles of Official Statistics.

4.78. After assessing the impact of SDC on data quality and determining the appropriate techniques to apply, pre-release discussions and review with appropriate internal subject matter experts with in-depth knowledge of specific statistical domains should be undertaken. In addition, consistency checks should be implemented to ensure relevance and confirm that census outputs are fit for purpose and ready for dissemination. One of the main quality assurance measures is having a thorough review system in place for data products to make sure that the data or indicators being released are of high quality.

8. Dissemination

4.79. Census dissemination can easily be overlooked in the chain of providing a quality outcome for the census as management attention may be focused more on the far more costly and risk prone data collection and processing operations. The dissemination area is responsible for the timely delivery of products and services to the census data users. Therefore, insufficient planning and resources for this phase can have the effect of delaying the release of the data and thus adversely impacting on the dimensions of timeliness and punctuality thereby compromising the overall achievement of the census objectives. The dissemination phase should also be regarded as an ongoing process that will service the needs of users over a period of time that will extend well into the intercensal decade.

4.80. For further details and a description of the processes involved in dissemination see Section IX.C in Part Three.

4.81. Management of quality in census dissemination is driven primarily by concerns to (a) deliver relevant products and services; (b) maintain accuracy of the data; and (c) ensure timeliness and punctuality of data release within agreed cost constraints. Effective management of quality in census dissemination should also aim to prevent miscommunication and misunderstanding between the NSO and its data users.

4.82. The first of these objectives is to provide relevant products and services. This can only be done by reviewing the experiences of the previous census products and services and through a continuous consultation process with both current and potential users of census data.

4.83. The second objective is to ensure that the data supplied from the processing system are accurately transformed into the required output products. A quality assurance strategy to ensure that data tabulations and transformations are carried out accurately needs to be documented and followed. The quality assurance circle approach to these processes needs to be applied and any gaps identified and corrected through extensive testing prior to the census and ongoing process improvement during the analysis phase before the dissemination of census outputs.

4.84. The third objective is the timely and punctual release of data from the census. While this an the objective of all phases of the census programme, its role in dissemination is crucial. The dissemination team needs to be realistic about release dates and ensure that these are communicated to users. The staff actually responsible for the dissemination phase should be closely involved in determining these dates where this is possible. Dissemination systems and processes need to be documented and tested prior to the release of data from the processing and analysis phases.

4.85. A release calendar needs to be prepared to keep the user community informed about the likely month of release so that their use of data can be planned in advance (see paragraphs 3.334 and 3.399-3.401). A mechanism to provide metadata on census indicators and the geographic level at which these are made available needs serious consideration. Every country should assess the requirements and put in place a dedicated customer service team to assist data users. The services of call centres may be used if the number of data seekers cannot be handled in-house.

9. Evaluation

4.86. An evaluation of the overall census operation is vital for identifying strengths and weaknesses of each of the census phases, including planning, data collection, data processing and dissemination, and also for the purpose of assessing the quality of census statistics. Together with the quality

assurance and improvement programme, the main objective is to ensure that quality assessment is consistently incorporated in all phases of the census, focusing on efforts in controlling the occurrence of errors and taking actions to ensure the highest quality of both the processes and their outcomes. Yet, errors appear to be inevitable in such a complex undertaking – consequently, there needs to be a mechanism put in place to determine the deficiencies and their quantitative impact on the census results.

4.87. An evaluation of the entire census operation, encompassing all dimensions of quality requires a comprehensive programme for assessing and documenting the outcomes of each process using appropriate and customized methodologies that should be planned well in advance as a component of the planning phase of the census. The evaluation process should cover all phases, from the planning stage through the completion of census operations. The insights and lessons learned should be considered the crucial first step in the next census cycle, ensuring ongoing learning and improvement. See Section X of Part Three for a discussion on overall evaluation of the census.

4.88. The evaluation process for measuring the coverage and quality of census data is discussed in more detail in Section II below but the key elements are noted in the following paragraphs.

4.89. Evaluation of the accuracy of the census data should also be undertaken, to the extent possible, by: conducting a post-enumeration survey (PES) for measuring coverage and content errors; by comparing the census results with similar data from other sources; and by applying demographic analysis. Such other sources may include surveys and administrative records (in particular, a population register) relating to a similar time frame, and previous census results. The purposes of evaluating the accuracy of the data are to inform users of the quality of the current census data and to assist in future improvements, which may be achieved by both improving processes and establishing performance benchmarks against which the quality of the data from subsequent censuses can be measured.

4.90. Evaluation of data accuracy may be undertaken in two stages. Preliminary evaluation will enable the identification of any problem areas that have not been previously detected through the quality management processes in earlier phases of the census. More extensive evaluation should then be undertaken on data items where problems have been identified or where the data have been collected through new questions or processes.

4.91. The census evaluation programme should be undertaken by subject matter specialists according to the agreed goals and methodologies covering each of the dimensions of quality noted at paragraph 4.8. The following are just some examples:

i. Identification of the deficiencies and achievements in data capture, coding and editing (through mechanisms developed for checking the quality of process and the work of personnel);
ii. Relevance of census data to user needs and satisfaction of users with dissemination tools and products (based on information collected through user consultation and, specifically, through a user satisfaction survey);
iii. Achievements and difficulties in the use of new technologies and methodologies and the identification of possible improvements for the next census;

4.92. The results of evaluations of census operations for both operational aspects and the quality of data should be made available to the stakeholders promptly as they become available. A commitment to transparency will ensure confidence in census results and provide a basis for improvements for the next census cycle.

4.93. Depending on the national context, census results by default may be presented without adjustments. But for analytical and planning purposes, results of the census evaluation should be taken into account in preparing intercensal and post-censual population estimates in order to correct any deficiencies in the data collection and subsequent phases.

10. Documentation

4.94. Through all the census phases discussed in Sections E.1-E.9 above, documentation is extremely important for improving the quality of future censuses. But that is not enough. It is recommended that once census data are evaluated and disseminated, NSOs should develop various reports to retain institutional knowledge and memory. NSOs should ensure that managers responsible for each of the census processes should not be transferred, or moved to other duties until the evaluation of these processes have been documented. Administrative and methodological reports should be written to document planning and implementation phases of the data collection, while a procedural history report should cover the product and dissemination phases. [177] It is also recommended that the NSO should document the SDC measures that have been applied as these too will affect the quality of the data. However, NSOs should avoid providing very specific details on SDC methodology so as to prevent unwanted attempts to reverse engineer the SDC methods.

4.95. For further details and a description of the processes involved in documentation, see Section XI in Part Three.

F. Ensuring and measuring the quality of data derived from administrative sources

4.96. As has been often noted throughout this document, there has been a clear trend towards increased use of administrative data in censuses, in line with a more generalized trend towards increased use of administrative data in all statistical domains. Their use can: provide more frequent and timely statistics about the population; facilitate improvements in accuracy and reliability; and result in significant reductions in costs and respondent burden. Nevertheless, there are many key quality considerations that should be assessed and evaluated before incorporating the data from any administrative source into a census.

4.97. The quality of input data used in the process of compiling census statistics strongly affects the quality of the statistical output produced. Thus, the quality of data from administrative registers and other administrative sources is a key element that should be considered in the decision-making process on any move to the use of administrative data in the production of statistics[178]. It is necessary to prepare and implement a standard method of assessing the quality of administrative data as potential sources for the census. Such assessment should be used to evaluate the usability of the data source. To this end a set of indicators should be developed by which each source may be analysed. Based on these indicators, the NSO can decide whether or not to use a particular source.

[177] https://www2.census.gov/programs-surveys/international-programs/events/stic/planning/improve-future-data-quality.pdf

[178] The advantages and challenges of a register-based approaches are outlined in paragraphs 1.228. -1.240.

4.98. In general terms, based on user needs identified for the census, the NSO should identify administrative data sources that may meet those needs. It is important to set out what the expected or required outcomes of using the source would be. This could include improvements to the efficiency of the census operation in terms of reductions in cost and respondent burden, improvements to the quality of the census, or the delivery of new or enhanced census outputs.

4.99. The following key quality considerations should inform decisions about the use of administrative data in a census:

i. Administrative data have, in general, not been collected for the purpose of a census. Therefore, such sources: may have adopted different concepts, definitions and classifications from those required by the census; may refer to different reference periods; be subject to time lags in the updating of information; and may have limited coverage of the census target population. Additionally, the accuracy and completeness of the data will be highly dependent on the importance of the data to the data supplier's administrative function. The administrative sources may also be subject to changes over time and inconsistencies in the way the data are collected across segments of the population.
ii. The use of administrative data by NSOs for purposes other than those for which the data were originally collected raises privacy, security, and legal concerns. An assessment of public acceptability of the use of administrative data for statistical purposes should be undertaken, and such usage should only be considered if the NSO has a legal authority and is publicly acceptable. It is particularly important that the public and data users alike understand how and why administrative data are being used in the census. The NSO should be transparent about their use and provide a clear justification of the benefits against any risks and costs. This can be achieved through good communication, including the publication of the procedures and policies that support the effective use and protection of the data.
iii. The NSO has limited control over the way the administrative data are collected and processed. Administrative data can be subject to changes in population coverage and population measures over time due, for example, to legal, policy, procedural, or system changes affecting the data and/or their delivery. This limited control and the associated risks to the census processes can be mitigated by working with data suppliers on potential or planned changes, by being flexible and responsive to change, and by reducing reliance on any single data source or item where possible.
iv. The complexity of administrative data, including its structure and the availability of comprehensive metadata, can significantly impact an NSO's ability to understand, access, and utilize it effectively. For example, complex data structures may pose technical challenges for assessment and integration. This complexity can also affect the accessibility and clarity of the data for users, hindering their ability to understand and utilize the information.

4.100. Assessing the quality of administrative sources requires recognizing inconsistencies and errors, then determining how these errors can be mitigated. But the quality assessment should not be limited to assessing the error. More generally, it should include an assessment of the impact of the integration of administrative data into the census in terms of the extent to which it adds error or uncertainty to the outputs, vis-à-vis of the advantages gained from integration.

4.101. A framework for assessing the quality of administrative sources for use in censuses can be built in stages: the Source stage, the Input stage, the Process stage, and the Output stage (See Figure 4.1). The first stage deals with the assessment of the quality of the administrative source itself, while the second stage focuses on the quality of the acquired data. The last two stages are concerned with the quality of the transformed and/or integrated data. For each stage, the key quality dimensions

should be identified, along with the respective tools and indicators for the quality assessment[179,180]. The four stages of the framework, broadly related to the lifecycle of using administrative data in the census, are applicable regardless of census type. A brief description of each stage is included below.

i. ***Source stage.*** The quality of administrative data sources being considered for use in the census should be assessed during the planning stage before the supply (or re-supply) of any data to the NSO. Such assessment should be carried out in a metadata-based approach in which the key parameters such as concepts, definitions, reference dates, revisions and delivery times provided by the data holders can be reviewed.

ii. ***Input stage[181].*** This refers to the quality assessment of raw administrative data as supplied to the NSO, with reference to the expectations and requirements established through the metadata-based assessment at the Source stage. The quality of administrative data at this stage is to be assessed against several dimensions including readability and validity, accuracy, timeliness and punctuality, and linkage ability.

iii. ***Process stage.*** Once administrative data are received and the quality assessed, the data will require processing to be usable in the census. The administrative data will need to be integrated into the census design and any quality issues will need to be addressed. Some of the most common processes required for using administrative data in the census are: linking records, assessing signs-of-life, assessing coverage errors, resolving inconsistencies from different sources, editing and imputation.

iv. ***Output stage.*** Measuring output quality cannot be reduced to only the estimation of uncertainty (the accuracy of estimates); rather, it should include an assessment across all other quality output dimensions. In this, such an assessment is fundamentally no different to that carried out for data collected in a field-based census. The introduction of administrative data will, however, likely lead to gains in some dimensions and losses in others. Achieving the right balance across the quality dimensions is the key to best meeting user needs. Updated and new administrative data can be used to improve census statistics by revising previous estimates. However, this needs to be balanced against user needs for comparability with respect to revisions.

4.102. It's important to document and publish the results of quality assessment and assurance activities throughout the census production process. This transparency fosters ongoing dialogue between data producers and users, facilitating a shared understanding of data quality and helping determine if the appropriate balance has been achieved across various quality dimensions. This collaborative approach promotes trust in the data and supports informed decision-making.

II. Methods to evaluate the coverage and quality of census results

A. Purpose of evaluating the quality of census data

4.103. Determining the quality of population and housing census data is very important for many reasons, including building public trust in and understanding of the national statistical system. The purpose of census evaluation is to provide users with a measure of the level of accuracy and confidence when utilizing the data, and to explain any errors in the census result. It is therefore

[179] See UNSD Handbook on Registers-based Population and Housing Censuses (2022).

[180] See UNECE, Guidelines for assessing the quality of administrative sources for use in censuses, 2021.

[181] Also referred to as “data stage” in the UNECE Guidelines for assessing the quality of administrative sources for use in censuses (2021).

important to choose an appropriate way of sending out these messages to the relevant group of stakeholders.

4.104. The evaluation methods discussed here are those that apply to field-based censuses. To some extent they also apply to register-based censuses and other census methodologies, but these also present their own particular challenges and solutions[182].

4.105. It is universally accepted that a population and housing census can never be perfect and that errors can and do occur at all stages of the census operation, but these errors should be measured. Errors in the census results are classified into two general categories – coverage errors and content errors. *Coverage errors* are the errors that arise due to omissions or duplications of any of the enumeration units – persons, households or housing units – during data collection. The sources of coverage error may include :

- Incomplete or inaccurate maps or lists of EAs;
- Failure on the part of enumerators to canvass all the units in their assignment areas;
- Duplicate counting of persons who have two or more places of residence or who move during the enumeration period;
- Persons who for one reason or another do not allow themselves to be enumerated;
- Erroneous treatment of certain categories of persons such as visitors or non-resident aliens; and
- Loss or destruction of census records after the enumeration.

Content errors are errors that arise in the incorrect reporting or recording of the characteristics of persons, households and housing units enumerated in the census. Such errors may be caused by several factors, including:

- Poorly phrased questions or instructions, or enumerator errors in phrasing the census questions;
- Inability or misunderstanding on the part of respondents in respect of answering specific items;
- Deliberate misreporting;
- Errors due to proxy response; and
- Coding or data entry mistakes.

4.106. Many countries that conduct a field-based census have recognized the need to evaluate the overall quality of their census results and have employed various methods for measuring census coverage as well as certain types of content error. Comprehensive evaluation should also include an assessment of the success of all aspects of the census operation, in each of its phases, including activities such as designing the geographic frame, the census publicity campaign, field staff training, data collection, data processing, data dissemination and data utilization. Countries should ensure, therefore, that their overall census evaluation effort addresses the census process, as well as the results. The present section is devoted to evaluation of the results. However, the section on the quality improvement (Section E, paragraphs 4.28-4.95) provides further recommendations relating to controlling and assessing the quality of the census operations.

[182] See, for example, *Census Quality Evaluation: Considerations from an international perspective*, note produced by the UNECE Secretariat for the 11th Joint UNECE/Eurostat Meeting on Population and Housing Censuses, Geneva, 13–15 May 2008, http://www.unece.org/stats/documents/2008.05.census.html.

4.107. Evaluation efforts that are focused on the census results should generally be designed to serve the following objectives: firstly, to provide users with some measures of the quality of census data to help them interpret the results; secondly, to identify, as far as is practicable, the types and sources of error in order to assist the planning of future censuses; and thirdly, to serve as a basis for constructing a best estimate of census aggregates, such as the total population, or to provide census results adjusted to take into account identified errors at national or subnational levels if some errors, such as coverage error, are substantial and the validity of census results would otherwise be questionable.

4.108. As the decision to adjust census figures is sensitive, it is bound to be of interest at the highest levels of government responsibility. However, matters relating to the quality of statistical outputs must be the sole responsibility of the NSO. Some critical statistical considerations should be very carefully weighed. Consideration must be given to what geographic domains the adjustment would cover, knowing that such adjustments have an effect on demographic distributions. Transparency in the methodology used for such adjustment is critical to maintain trust with all stakeholders but particularly data users. In cases where census results are presented without adjustments, the results of the census evaluation should nevertheless be taken into account in the preparation of post-census population estimates and for analytical and planning purposes.

4.109. The final publication of census results should include an estimate of coverage error, together with a full indication of the methods used for evaluating the completeness of the data. The publication should also provide users with some cautions or relevant metadata about the results, in addition to some guidance on how they might use the evaluation results. It is also desirable to provide, as far as possible, an evaluation of the quality of the information on each topic and of the effects of the editing and imputation procedures used.

4.110. The range and quality of the editing of incorrect or inconsistent data and the imputation of missing data items are greatly enhanced by the use of computer edit programs that permit inter-record checks (for example, the replacement of missing values based on one or more items on the basis of reported information for other persons or items in similar households). If any imputation is made, the topics affected, the methods used, and the number of cases affected should be documented and clearly described in the census evaluation report.

4.111. The results of the evaluation of census results should be made available to users with a measure of the quality to help them interpret the results.

4.112. As discussed in the following subsection, a number of methods exist for carrying out a census evaluation. In practice, many countries use a combination of such methods.

B. Methods of census evaluation

4.113. The choice of evaluation methods to be used depends upon the evaluation objectives. These, in turn, depend on national census experience in terms of past and anticipated errors, user and public concerns, and the financial and technical resources available for evaluation. The decision needs to be made whether to measure coverage error, content error or a combination of the two. In addition, both gross and net error must be considered in developing the overall evaluation plan. Gross coverage error in a census is defined as the total of all persons omitted, duplicated or erroneously enumerated. Net coverage error takes into account the underestimates due to omissions and the overestimates due to duplications and erroneous inclusions. When omissions exceed the sum of duplications and erroneous inclusions, as is usually the case in most countries, a net undercount is said to exist; otherwise, a net overcount results. Similarly, both gross and net content errors have to be considered in the evaluation design.

4.114. Numerous methods are available to estimate the coverage and content error of censuses. Methods for evaluating the quality of censuses can be divided according to their complexity, the type of measures used, or data sources considered in the evaluation process. One of the most commonly used classifications is based on what sources of information are taken into account in the process of census quality assessment. According to this classification, there are three types of methods for assessing the quality of census data: 1) those based solely on data from the census itself (single source); 2) those in which census results are compared with data from other statistical surveys or from administrative data sources (multiple data sources); and 3) those based on results from independently conducted surveys. Each of these methods can be used to assess the size of coverage and content errors. These include simple techniques of quality assessment, such as internal consistency checks.

4.115. Comparisons of results with other data sources, including previous censuses, current household surveys and administrative records, are also useful techniques. Such comparisons may be made in aggregate by comparing the overall estimates from two sources (net error only). Differences in conceptual frameworks and quality assessment for all these sources should be considered when using them as references to assess the coverage and content error of censuses.

4.116. Alternatively, record checking, whereby individual census records are matched against alternative sources and specific items of information are checked for accuracy, may be used. Both gross and net errors can be estimated in record checks, which may involve field reconciliation of differences, a costly exercise but one that that should not be overlooked. An important but complicating factor in the use of record checks is the requirement for accurate matching. It is essential to plan carefully for this aspect, since the operation can be not only costly but tedious. It should be noted that record checks are best employed to study the coverage of certain segments of a population, such as children whose birth records are complete, since these checks are, by definition, limited to sub-populations with complete, accurate records.

4.117. Demographic analysis and a PES are two very important methods for evaluating census data, and these are discussed in further detail in the following two sub-sections. There are many other methods for testing the validity of population counts and to verify the completeness (coverage) of the census estimates, such as: the Census Coverage Survey (CCS), Reverse Record Check (RRC)[183], Dual System Estimation (DSE), and adjustments for any residents living temporarily overseas.

4.118. The evaluation of register-based and combined censuses includes measures such as the rates of record linking and imputation, the use of methods such as the signs of life and their impact on census coverage, and evaluation of the impact of incorporating administrative data on census outputs. Understanding these measures is especially important for countries transitioning to an administrative-based data collection methodology in their census.

[183] A frame containing all persons who should be enumerated in the current census is built up from the returns of the previous census and intercensal birth and immigration registrations. A random sample is selected from each source and each selected person is traced to his/her current census address (since the address obtained at the time of selection is usually out of date, a tracing operation must be undertaken to determine the address of each selected person as of census day) census forms are then checked to determine whether or not the selected person was enumerated. See e.g. the Canadian Census Undercoverage Survey (CUS) methodology. (https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3902).

C. Post-enumeration survey

4.119. The PES can be defined as the re-enumeration of a representative sample of the census population and matching each individual who is enumerated in the survey with information from the main enumeration. The objectives of the PES can be summed up as follows:

a. To assess the degree of coverage during census enumeration;
b. To examine the impacts of coverage deficiencies, if any, on the usefulness of the census data;
c. To examine the characteristics of persons who may have been missed during census enumeration.
d. To obtain information for the design of future censuses and surveys;

4.120. While a PES can be designed to provide a comprehensive evaluation of coverage and content error, especially when supplemented by, and integrated with, detailed demographic analysis of census quality, the methodology of a sound PES is complex, so that countries should accordingly weigh with care the demanding technical requirements and the costs of conducting a successful survey, and elaborate a clear statement of its objectives, before deciding to undertake such a survey.[69] Careful advance planning is vital. To be valid, a PES has to function within a number of operational and statistical constraints. These include the requirement that the survey be carried out within a few months of the end of the census, and that changes in transitory conditions (such as employment status) and lapses in respondent recall do not over-complicate the exercise.

4.121. Another basic feature of a PES design and execution involves matching and reconciliation. Matching the unit records from the PES with the corresponding census records is an operation whose performance must be of very high quality to ensure that inaccuracies in the PES itself do not effectively ruin the estimate of coverage error. Matching is especially difficult in countries where many surnames are identical or where individuals are known under more than one name, and where well-defined street addresses do not exist. Part of the matching operation usually involves a field visit to reconcile differences between the census and the PES with regard to either coverage or content. Reconciliation of course adds another dimension of cost and complexity, since it entails a second PES visit in the field. When forcibly displaced populations are part of the census, special attention should be given to households that may have experienced displacement between census and the PES[184]

4.122. Clearly defining the objectives of a PES is the first and most crucial step in planning the survey. These objectives should align with the broader goals of assessing census coverage and data quality, as outlined in paragraph 4.119. This might include:

(i) Estimating coverage error at the national level.
(ii) Estimating coverage error for major subnational domains or population subgroups, each with its own specified level of precision.
(iii) Measuring content error for specific census items.
(iv) Examining the impacts of coverage deficiencies on the usefulness of the census data.
(v) Obtaining information to improve the design of future censuses and surveys.

[184] https://egrisstats.org/activities/compilers-manual/ (see paragraphs 49 – 53)

Additionally, a secondary objective may be to compare the distributions of key characteristics between the census and the PES. For this purpose, measures of consistency, such as the concordance correlation coefficient (CCC) or Cohen's Kappa coefficient,[185, 186] may be used.

4.123. As already noted, the design of a PES is complex and there are various alternatives. A number of references are available that set out highly detailed procedures for designing a PES and the conditions under which they may or should be considered. However, in general, when designing such a survey, the following key considerations should be taken into account:

i. The time between the census and the PES should be minimized to avoid as much recall error as possible and the impact of population changes (births, deaths and migration).
ii. The PES must be independent of the census. Interviewers must not have census information about the areas where they are working. When interviewers have knowledge of census responses, bias in introduced from the tendency to confirm only what the census recorded.
iii. To preserve the independence of the PES, its data collection and processing operations must be completely separate from those of the census.

4.124. The sample design for a PES must be based upon sound probability sampling methods taking account of the measurement objectives of the evaluation study. These usually include the need to estimate census coverage with a certain degree of reliability. In addition, the estimates of coverage may be wanted for geographic areas such as provinces or states and large cities, for urban–rural comparisons and so forth. Such requirements also greatly affect the sample design of a PES, as the necessary sample size is increased substantially when estimates of subnational coverage (or under-coverage) are wanted.

4.125. Sometimes a PES is designed to measure content error only, in which case it is sometimes referred to as a reinterview survey. The advantage of a well-designed reinterview survey is that the results are more accurate than those of the census insofar as the operation is much smaller and can be more effectively controlled. Estimates of relative response bias can be obtained from a reinterview survey, which is generally considered the more accurate standard for measuring content error, as it typically involves better-trained interviewers and more intensive survey procedures, leading to more reliable results than the census itself.

4.126. As part of the design of some PESs, a sample of the original census EAs, blocks or other areas is chosen and recanvassed for the survey. As regards methodology, this constitutes a useful reinterview technique for measuring content error, and such an element in the design is often put into practice because the matching operation between survey and census records is then dramatically simplified.

D. Demographic analysis for census evaluation

4.127. Demographic analysis offers a powerful methodology for evaluating the quality of a census, and NSOs are encouraged to use it as part of their overall census evaluation methodology. A wide variety of demographic techniques have been developed and used, ranging from visual inspection of

[185] Cohen's Kappa coefficient is a statistical measure used to assess the agreement between two raters or methods that classify items into categories. It's particularly useful when dealing with categorical data and wanting to determine if the agreement between the two raters is due to actual agreement or just chance.
[186] Statistics Poland (2023), National Population and Housing Census 2021. Assessment of the data quality, Warsaw.

census data to comparative analysis of two census age distributions. A basic procedure for assessing census quality on age-sex data is graphical analysis of the population pyramid. Age heaping, or the tendency of respondents to report a particular ending digit (typically '0'), is a useful check of the quality of age reporting, as are sex ratios by age and certain summary indices of age-sex data, including the United Nations age-sex accuracy index, which extends age-sex ratio analysis by observing deviations of the observed age-sex ratios from the ones expected for each five-year age group and combining the results into a single score.[187] Other summary indices are Whipple's index and Myer's blended index, used for judging age heaping[188].

4.128. The PES method, while useful in providing an overall assessment of census quality, cannot differentiate the sources of census error in terms of the relative contributions from under-coverage (or over-coverage) or content error. Better information about coverage error, through demographic analysis, derives chiefly from comparative analysis of data from successive censuses, in which four methods are used. The four methods are:

i. derivation of an expected population estimate taking account of vital registers of births, deaths and net migrants between censuses, as compared with the latest census[189];
ii. population projections based on the results of the prior census plus data on fertility, mortality and migration from various sources and comparing the projected estimates with the new census results (cohort component method)[190];
iii. comparison of two consecutive census age distributions based on intercensal cohort survival rates and evaluation of the population age distribution for two or three census dates by year of birth of each cohort[191];
iv. and estimates of coverage correction factors using regression methods to make the age results from the two censuses mutually consistent (cohort survival regression method).

It should be noted that the first two methods would probably have to be restricted to evaluation studies of coverage at the national level, especially in countries that do not have good subnational data on internal and international migration.

4.129. A fifth method focuses on the evaluation and potential correction of under-enumeration of young children. The population at younger ages is often presenting a pattern of under-enumeration[192]. The observed population of children at age 0, 1, and/or 2 may be lower than expected, in particular in

[187] See *Methods of Appraisal of Quality of Basic Data for Population Estimates: Manual II* (United Nations publication, Sales No. E.56.XIII.2).

[188] See Johnson, P. and others (2022). *Method protocol for the evaluation of census population data by age and se*x. United Nations, DESA, Population Division, Technical Paper No. 2022/05. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2022_tp-methodprotocol.pdf

[189] Himes, C. L., & Clogg, C. C. (1992). An overview of demographic analysis as a method for evaluating census coverage in the United States. *Population index*, 587-607.

[190] See United Nations, Department of Economic and Social Affairs, Population Division (2022). *World Population Prospects 2022: Methodology of the United Nations population estimates and projections*. UN DESA/POP/2022/TR/NO. 4

[191] see GRPOP-YB workbook from the U.S. Census Bureau Population Analysis System (PAS) (https://www.census.gov/data/software/pas.html) and newer birth Cohorts () R implementation from Demo Tools (Riffe, 2021). Riffe, T. (2021). DemoTools: Standardize, Evaluate, and Adjust Demographic Data. https://rdrr.io/github/timriffe/DemoTools/

[192] Ewbank, D. C. (1981). *Age Misreporting and Age-Selective Underenumeration: Sources, Patterns and Consequences for Demographic Analysis*. Washington, D.C., National Academy Press, Committee on Population and Demography, Report 4.

the developing countries with high fertility, where a continuous increase in the population is expected at younger ages. But patterns of under-enumeration of young children is found in more statistically advanced countries as well[193]. Some of these under-enumeration problems also affect PESs because the information provided on the household composition, and children in particular, is provided by the same proxy respondent.

4.130. Patterns of child under-enumeration can be corrected through demographic analysis by estimating the population at ages 0, 1-4 and 5-9 based on the smoothed adult female population from a census and estimates of fertility and mortality for the various periods preceding the census. In other words, the population below age 10 can be reconstructed based on estimates of past fertility and mortality (based on, for example, either vital registration or recent surveys).[194]

4.131. Further evaluation and potential adjustments of age misreporting or exaggeration at older ages (for example, for those aged 70 or 80 years and over[195] and especially for centenarians) require the use of additional administrative data sources (see section E below) or registration data (such as birth records) with more accurate recording of the true year of birth rather than self or proxy reporting of age, but the availability and accuracy of vital records for older birth cohorts can be limited due to low completeness of birth registration more than 50 years ago in many locations. Several demographic methods can also be used using additional assumptions on expected survival at older ages between censuses, the expected age distribution of the population at older ages based on the underlying mortality[196], or methods based upon the availability and sufficient completeness of registered deaths[197] [198]

E. Comparison with administrative sources

4.132. Comparing census data with administrative data in a country or specific area can provide valuable insights into changes since the last census. By analysing administrative data (including data on development trends), it is possible to assess how well the new census data captures these changes and whether or not the census accurately reflects the characteristics of the underlying population.

4.133. Administrative data sources, such as population registers, records of birth, death, migration, etc. provide estimates of population changes. By comparing census counts with these administrative

[193] Special issue on undercount of children in censuses from the Statistical Journal of the IAOS, Volume 33, issue 2, https://content.iospress.com/journals/statistical-journal-of-the-iaos/33/2 and Tomkinson, J. (2023). The Omission of Young Children in the French Census: What Can Linked Census Data Reveal? | Cairn International Edition https://www.cairn-int.info/journal-population-2023-2-page-229.htm

[194] see BASEPOP and BPSTRNG workbooks from the U.S. Census Bureau Population Analysis System (PAS) and newer basepop_five() R implementation from DemoTools (Riffe, 2021).

[195] Del Popolo, F. (2000). *Los problemas en la declaración de la edad de la población adulta mayor en los censos*. CEPAL.

[196] see OPAG workbook from the U.S. Census Bureau Population Analysis System (PAS) and newer R extension from DemoTools implementation (Riffe, 2021).

[197] Jdanov, D. A., Scholz, R. D., & Shkolnikov, V. M. (2005). Official population statistics and the Human Mortality Database estimates of populations aged 80+ in Germany and nine other European countries. *Demographic Research*, 13, 335-362.

[198] Terblanche W, Wilson T (2015) An Evaluation of Nearly-Extinct Cohort Methods for estimating the Very Elderly Populations of Australia and New Zealand. PLoS ONE 10(4): e0123692.

data sources, census under- or over-coverages can be assessed. However, in these comparisons, the alignment of geographical boundaries, reference periods, and the target population covered by administrative data and the census should be considered. The coverage of census at the household, and dwelling level can also be assessed if relevant administrative data exist.

4.134. Availability of administrative data enables comparisons at the micro-level as well. Such comparisons can be facilitated if national unique identification numbers exist and have been included in the census questionnaire. This will not only assist with locating coverage errors but also enables evaluation of the content. Administrative data sources, such as healthcare records, school enrolments, and tax filings, social welfare programmes, and employment record, etc. can be compared to census data to identify content errors related to demographic or socio-economic information. Additional information from utility services, such as electricity, water, sanitation, phone, etc. with sufficient household coverage can provide complementary data sources useful to evaluate the occupancy of housing units.

4.135. The comparison with administrative data involves the following stages:

i. *Identify the relevant administrative data sources:* Administrative data can include records from government agencies, local authorities, educational institutions, healthcare systems, and other organizations.
ii. *Gather administrative data:* Collect the administrative data that aligns with the relevant timeframe and census topics. This data may include population records, building permits, school enrollment figures, healthcare statistics, or employment data.
iii. *Analyze administrative data trends:* Examine the administrative data to identify any significant changes or trends that have occurred. Look for relevant indicators such as population growth or decline, changes in employment sectors, infrastructure development, or changes in school enrollment.
iv. *Compare the census data with the administrative data (aggregate data or unit records):* Compare the findings from the analysis the administrative data with the census data and evaluate the level of alignment and consistency between these sources. Comparing census data with administrative data can help identify potential discrepancies and provide insights into the accuracy of the census data.

4.136. When comparing census data with administrative data it is important for the NSO to:

- assess the quality of administrative data;
- understand differences in data definitions, concepts and reference dates;
- consider spatial variations by analyzing the data at different geographic levels;
- conduct additional research to validate the findings; recognize that both census data and administrative have limitations;
- seek stakeholder perspectives; document findings; and
- report the results of the evaluation.

F. Adjusting census data

4.137. Achieving full coverage of the population is one of the major challenges in carrying out any census, and some coverage error is unavoidable. Hence it is important to measure, analyse and report on such error.

4.138. Population under-coverage is regarded as one of the most significant sources of error affecting census data. This results from missing housing units, households, or persons during the data collection phase. Over -coverage, on the other hand, results from duplications (occurring when persons, households or housing units are enumerated more than once) or from erroneous inclusions (occurring

where persons, households or housing units are enumerated in the census when they should not have been or when persons are enumerated in the wrong place[199]). These two sources of error (resulting in census counts under-estimating or over-estimating the true population counts) can also distort the distribution of population characteristics estimated from census data when the over-covered and missed persons do not have the same characteristics as those who were enumerated correctly.

4.139. Both gross and net error must be taken into account in developing the overall evaluation plan (see paragraph 4.113 for definition of gross coverage error and net coverage error). A net undercount occurs when omissions outweigh duplications and erroneous inclusions, while a net overcount results from the opposite. Evaluation efforts should assess both gross and net content errors.

4.140. In some countries the evaluation of census coverage forms an integral part of the census quality assessment and improvement component of the quality management framework, and feeds directly into the published census results. In other countries, it is a separate exercise, used either to measure the census coverage error or to adjust the census results.

1. Adjustments for full field enumeration censuses

4.141. In the case of a census employing a full field-based enumeration, the most common problem is under-coverage [200]. The evaluation is usually done via an independent PES of a sample of enumerations areas (as described in Section C above) or via an RRC methodology. In the case where the latter method is used to estimate the number of persons missed in the census, this estimate is combined with the estimate of the number of persons enumerated more than once from a census coverage study in order to calculate net under-coverage[201].

4.142. In both cases the purpose is to measure the accuracy of the census. By independently surveying a sample of the population, the aim is to estimate the proportion of people and housing units potentially missed or counted erroneously in the census. On the basis of net coverage rates, adjustments may be made to population census results. Using results from a carefully designed and implemented PES, under- or over-counts can be converted into adjustment factors, that is, a correction factor attached to each record, so that the census population is increased or decreased accordingly by such factors. The adjustment is usually done at the national level or for other sufficiently large geographic areas.

2. Adjustments for register-based censuses

4.143. In the case of census data derived mainly or wholly from administrative sources, a common issue that can arise is over-coverage. For example, if individuals who have emigrated do not officially declare their departure, their records might remain unchanged in administrative registers, leading to

199 Of course, counting people in the wrong place may only seriously affect the coverage at the local area level; at higher levels of geography the effect will be minimised and at the national level the effect will be zero.

[200] Of the 1,758 census populations between 1950 and 2021 used for the production of the estimates of the *World Population Prospects 2022*, 310 had associated PES results available and the mean net error was -3.4%. United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022

[201] Net coverage error takes into account the underestimates due to omissions and the overestimates due to duplications and erroneous inclusions. When omissions exceed the sum of duplications and erroneous inclusions, a net undercount is said to exist; otherwise, a net over-count results.

over-coverage in the census.[202] To address this issue, the census can implement a post-enumeration survey (PES) or ad hoc surveys. These surveys can help identify and exclude individuals who have emigrated but are still included in the administrative data. Another strategy is to employ dependent interviewing techniques. This involves verifying administrative records with individuals in the field to identify and correct any discrepancies. However, it's important to note that dependent interviewing raises ethical and privacy concerns, and may not be feasible in all countries.[203]

4.144. Another possible means of avoiding coverage errors is to create a residency index on the basis of the records held in multiple registers in order to determine a so-called 'signs of life' (SoL) score. The approach is to define for all possible residents the 'sign of life' as a binary score (with a value 0 or 1) for each record in each register. Using these signs of life as explanatory variables, it is possible to build a model forecasting the size of under and over-coverage.[204]

4.145. However, even though register-based censuses are usually concerned mainly with over-coverage, in principle they are also affected by under-coverage, that is, there might be constituents of the population not included in any register. Another way to account for coverage errors in statistical registers is based on the correction of the population register through estimates obtained from two coverage samples for each area. More precisely, a field (area) sample of people living in the area on census day is used for estimating the register under-count (the "U sample"), whereas a sample of people registered in the same area is used for estimating the register over -count (the "O sample"). The U sample can also be used for collecting information on census variables not available in registers.

4.146. Some countries are using a combination of the methods above to investigate and refine methodologies for using of administrative sources as a means to understanding coverage issues in the census and make adjustments to census data. This is an emerging area of statistical practice. The availability of administrative data sources is changing rapidly as more data is captured and stored electronically. Therefore, approaches differ from country-to-country and often need to be tailored based on national circumstances and requirements.

3. Adjustments for combined censuses

4.147. There are several options to check the coverage errors in the case of a combined census. One is to use the methodology commonly adopted for full field enumeration censuses (for the field enumeration component), by organizing a PES and estimating both over-coverage and under-coverage statistically (see section C). For example, the Census Coverage Survey (CCS) is a PES that many countries conduct after the decennial population census. The CCS plays a crucial role in providing accurate population estimates and helps to identify any biases or errors in the census data. The main goal of this survey is to measure the coverage of the census by identifying any undercounts or

[202] UNECE: Guidelines for assessing the quality of administrative sources for use in censuses, United Nations, Geneva (2021), http://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf
[203] A discussion of dual- and multisystem estimation methods for tackling coverage errors is provided in Brown J., C. Bycroft, D. Di Cecco, J. Elleouet, G. Powell, V. Račinskij, P. Smith, S.-M. Tam, T. Tuoto, and L.-C. Zhang (2020), Exploring developments in population size estimation, The Survey Statistician 82, 27-39.
[204] UNECE: Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, United Nations, New York and Geneva (2018), http://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf.

overcounts of the population. It is often based on dual system estimation[205] whereby data from the census and the coverage survey are matched to estimate the population size. After combining census and CCS results, the capture-recapture methodology is used to assess the numbers of households/individuals missed by the census enumeration.

4.148. The 'signs of life' approach (noted at paragraph 4.144 above), that compares records from different registers, can then be adopted to provide information on the level of under-coverage and over-coverage of the register-based component of the census.

4. The decision to adjust, or not to adjust, census figures

4.149. After measuring the accuracy of the population count, some countries adjust census figures while others choose to leave census data unadjusted. In fact, many countries, even though they may undertake a PES, demographic analysis or other means of assessing coverage, do not adjust their census figures on the basis of the results of these. This is due to various reasons. For example, adjusting census figures can be impossible due to national circumstances or legal reasons; or be too politically sensitive (due to the impact of adjustment on budget allocation or the size of electoral districts/constituencies); or because coverage error assessment is performed more for quality reasons (that is, for measuring the level of accuracy of the population count) rather than in order to adjust census figures.

4.150. Countries may choose to decide to make adjustments to the census results only when coverage errors are substantial, and the validity of census results is questionable. However, it is necessary to evaluate whether it is possible to apply adjustments and, if so, consideration must be made to what level of geography the adjustments would cover knowing that such adjustments impact demographic distributions. Input from stakeholders, including experts in demography, statistics, and affected communities, may also need to be taken into account.

4.151. In those countries where the total population figure has been adjusted for under-enumeration or over-enumeration, both the enumerated figure (the population *count*) and the adjusted population figure (the population *estimate*) should be published for the sake of transparency and ensure that the method used should be clearly understood by data users.

4.152. Whatever decision is made (either to adjust or not to adjust census data), it is important that any information about the accuracy of census data (for example, PES results and/or demographic analysis evaluation) is fully documented alongside the census results.

G. Acceptance of results

4.153. In countries with limited prior census experience and without a well-functioning civil registration system, where population data are based largely on estimates, it is important to inform the users, particularly the relevant governmental authorities, that the census results could differ from such estimates and to explain the reasons for these differences. In some cases, there may be doubts expressed about the reliability of census results; usually those doubts focus narrowly on the total

[205] See the UNSD Technical Report on Post-Enumeration Surveys (2010) for more detail on dual system estimation.

population of the country, major subdivisions or population subgroups, rather than on the bulk of the census data relating to characteristics of the population or on the data for local areas.

4.154. In this situation, it is advisable to modify the census evaluation programme to better document differences and/or by adding appropriate qualifications to the text of the census reports or in tabular footnotes. The other statistical programmes (such as population estimates) might also provide documentation to explain differences to the census results and help users navigate their way across the different sources.

4.155. Nevertheless, the government may proceed with the utilization of the census results for official purposes. In any case, every effort should be made to process and evaluate the complete census operation and to make appropriate use of as many of the census tabulations as possible.

4.156. More generally, in order to ensure transparency about data quality and enhance users' confidence in census results, it is important to assess the quality of census data as much as possible prior to their release, and to make this information publicly available, even though, in order to provide detailed coverage results and information on how accurately the census counted certain population groups, it might be necessary to wait for the results of the PES or other coverage surveys.

4.157. One of the primary methods of evaluating the quality of a census is comparing the results to other sources, mainly population benchmarks. While other sources certainly offer views on the quality of the census, there are limitations to the information they provide and the comparisons that can be made according to the sources. Therefore, data users should be guided in exercising caution in how they interpret differences. NSOs should explain which population comparisons are possible, and provide guidance on how to interpret differences between demographic benchmarks and census results. For example, the results from the Labour Force Survey data are generally used for making comparisons about the population labour market (economic activity) status. In this case, users should be helped in interpreting any differences arising from data collection and questions design.

4.158. As another example, doubts may be expressed with regard to how well specific ethnic groups or religious groups were counted in the census, by comparing census data with data derived from other sources. In such cases as well, it is important to explain the concepts and definitions used in the census, and how these might differ from the one used in other sources (such as, in administrative records). To assist with making this comparison, NSOs might associate the categories from the census with those that appear in data from the administrative records and explain the methodology used for this.

4.159. In providing documentation on the main comparisons possible between census data and other sources, NSOs should advise data users in being cautious when they interpret differences and to recognize that errors in the census are only one possible source of difference between the census and the population estimates or other sources.

H. User satisfaction

4.160. In many countries, NSOs conduct user satisfaction surveys on products and services on a regular or occasional basis, in order to identify the gaps between what the census has achieved and what users actually needed or expected to receive. In general, the main aim of user satisfaction survey is to gather opinions of census users about the quality of data produced and disseminated. Some NSOs collect information from users in an anonymous continuous form, using a facility accessible via the institutional website or dissemination web platform, after consultation or downloading of statistical data. Periodically, this information is processed for internal use and in some cases presented in

technical reports available on the web. The most important aspects that are measured through user ratings are timeliness or relevance of the data offered. Other statistics collected in this way can relate to:

(i) the purpose of using NSOs data;
(ii) the quality of such data compared to statistics produced by other organisations;
(iii) users' trust in the NSOs statistics;
(iv) the frequency of tailored requests; and
(v) the frequency of accessing the NSOIs website.

In other countries, specific feedback surveys are conducted after the data are released among samples of users from different sectors, including central or local government, academia and businesses. For impartiality purposes, NSOs may wish to consider engaging an independent organization to conduct such user satisfaction surveys.

4.161. Such a survey should address questions such as whether or not:

(a) the disseminated census data fully satisfies users' stated requirements with regard to the topics covered and the tabulations and other products provided;
(b) the level of territorial detail allowed by the census methodology and privacy constraints meet the user's needs;
(c) the timeliness of dissemination fully satisfies users' expectations;
(d) the census data have been integrated with other sources available to the users and whether or not they have been used through with GIS technologies; and
(e) the data have been used to carry out the studies, research, analyses, planning activities to be conducted within the scope of their institutional mission etc.

Attention should also be paid to the accessibility, with regard to the suitability of the form or medium through which census information can be accessed both with regard to specialized users and to the public at large, and to interpretability of the census information, that is, the availability of supplementary information and metadata necessary to explain and use it, and indications of the accuracy of the information. Furthermore, in those censuses where there have been changes in the data collection methodology that may have affected the scope and geography of disseminated data, the impact on the use of that data should be investigated.

PART FIVE. POPULATION AND HOUSING CENSUS TOPICS

Chapter I: Population census topics

I. Factors determining the selection of topics

5.1. In line with the overall approach to this revision of *Principles and Recommendations for Population and Housing Censuses*, the selection of census topics is based on the outputs expected to be produced by the census. Therefore, the first step involves clear identification of user requirements for data; the core and additional topics are then decided on that basis. For each of the core topics there is a recommended tabulation.[206] It is recommended that countries collect data on the core topics and also produce the recommended tabulations, as this would improve the international harmonization and comparability of statistics through the use of common concepts, definitions and classifications. Use of an agreed international approach would also enhance the capacity of countries to generate statistics for monitoring the socioeconomic situation of their populations, including for the provision of data for the internationally agreed development goals.

5.2. The topics to be covered in the census (that is, the subjects regarding which information is to be sought for each individual or household) should, however, be determined upon a balanced consideration of:
(a) **National priority**: It is essential to consider the diverse needs of data users in the country at both the national and local area levels to ensure that the census generates relevant information to effectively address these needs.
(b) **International comparability**: It is important to ensure that the data collected is comparable at regional and global levels through aligning with international standards. This would facilitate inter-country analyses and monitoring.
(c) **Suitability**: For information collected through questionnaires, the NSO should consider the sensitivity of the topics and the respondent burden, that is, the willingness and ability of the public to provide accurate information on these topics. Similarly, with respect to information to be collected through administrative registers, the NSOs should consider the sensitivity of the topics and the willingness of the public to have their personal information linked and used, in other words assess the acceptability of the use of existing administrative information for statistical purposes. It is also important to consider the technical competence of NSO staff in regard to obtaining the requisite census information on the topics through data linkage.
(d) **Resources**: NSOs should consider the total national resources available for conducting the census, including financial, human, and technological assets.
(e) **Alternative sources**: It is important to assess the availability of relevant information from alternative data sources to reduce duplication and improve the efficiency of data collection.

5.3. Such a balanced consideration will need to take into account the advantages and limitations of alternative methods of obtaining data on a given topic within the context of an integrated national programme for gathering demographic and related socioeconomic statistics (see paragraphs 1.131. –1.136.).

[206] Link to UNSD website containing the tabulation shells.

5.4. In selecting the population topics, careful consideration should also be given to the usefulness of historical continuity, which provides the opportunity for comparison of changes over a period of time. NSOs should avoid, however, collecting information that is no longer required by users. Information should not be collected simply because it has been traditionally collected in the past, bearing in mind changes in the socioeconomic circumstances of the country. It becomes necessary, therefore, in consultation with a broad range of users of census data, to review periodically the value of even long-standing topics and to re-evaluate the need for their continued inclusion, particularly in the light of new data needs and alternative data sources that may have become available for investigating topics hitherto covered in the population census. Each of the five key factors (i.e., national priority, international comparability, suitability, resources, and alternative sources) that need to be taken into account in reaching a final decision on the content are briefly reviewed in the following paragraphs.

A. National priority

5.5. Prime importance should be given to the fact that population censuses should be designed to meet national needs. In determining such needs, the full range of national uses (for example, policymaking and monitoring, administration and research) and national users should be consulted. The prime consideration is that the census should provide information on those topics that are of greatest value to the country, with questions framed, or appropriate administrative data sources identified, so as to elicit data of maximum utility.

5.6. Each country's decision with regard to the topics to be covered should depend upon a balanced appraisal of how urgently the data are needed and whether the information could be equally well or better obtained from alternative data sources such as administrative records or household surveys (see Section E below). Experience has shown that national needs will best be served if the census includes topics generally recognized as being of basic value and defined in accordance with regional and global standards. The recommendations set out in this document can help in this appraisal by providing information about core and additional census topics and their definitions and concepts, emanating from national experiences. It is recognized however that countries that rely more on administrative records as their primary data source may be more limited in the precise detail of the information that can be collected on certain topics.

5.7. Many countries may find it necessary to include topics of national or local interest in their census, in addition to the topics suggested in these recommendations. Labour force surveys and/or other survey data may supplement census data so as to obtain information on topics that cannot be included in the population census for whatever reason. It is possible that some countries may omit from their census certain recommended topics either because there is no longer a national need to collect the data or because there are legal barriers or particular sensitivities in collecting information on such topics (for example, fertility, ethnicity and religion).

B. International comparability

5.8. The desirability of achieving international comparability should be another major consideration in the selection and formulation of topics to be included in the census. National and international objectives are usually compatible; since international recommendations are based on a broad study of country experiences and practices, the definitions and methods contained in international recommendations have successfully met general national needs in a wide range of circumstances. Furthermore, the analysis of census data for national purposes will often be facilitated if, through the use of international recommendations, it is possible to compare the data with those of other countries on the basis of consistent concepts, definitions and classifications. The 2030 Sustainable

Development Agenda, which places increasing demand for expanded data collection, is also another determining factor that countries should take into consideration.

5.9. If the particular circumstances within a country require a departure from international standards, every effort should be made to explain these departures in the census publications and to indicate how the national presentation can be adapted to the international standards.

C. Suitability

5.10. A prerequisite for the inclusion of topics in field-based censuses should be the willingness and ability of respondents to provide accurate information on them, or, for register-based censuses, the acceptance of the public to have their personal information linked and used for statistical purposes. It is advisable to avoid topics that could increase the burden on respondents and those that are likely to arouse fear, local prejudice or suspicion, or that might be used to deliberately promote political or sectarian causes, as these are likely to have a detrimental effect on response rates and support for the census. In an interview-based census or when the enumerators need to obtain information through observation, consideration needs to also be given to the level of knowledge and skill of the enumerators and whether they can be adequately trained to collect this information accurately. Topics that are too complicated or difficult for the average respondent to answer, or the enumerator to record, quickly should not be included. The exact phrasing of a question that will obtain the most reliable responses may depend on national circumstances and should be well tested prior to the census (see paragraphs 3.140–3.149).

D. Resources

5.11. The selection of topics should be carefully considered in relation to the total resources available for the census. An efficient collection of accurate data for a limited number of topics, followed by prompt tabulation and publication, is more useful than the collection of data for an overambitious list of topics that cannot be properly processed and disseminated in a timely, reliable and cost-effective manner. In balancing the need for data against available resources, the extent to which questions can be pre-coded is an important consideration. Additionally, information from studies on data users' needs and how they utilize census data can help determine the economic feasibility of including certain topics. For example, if a particular topic is only relevant to a small subset of users, or if the data is unlikely to be widely used, it might not be cost-effective to include it in the census. In the case of those topics on which data is taken from registers, the additional resources required to transform the administrative records into useable statistical data should also be noted.

E. Alternative sources

5.12. In the selection of topics to be investigated in a census, consideration should be given to whether or not data are available from alternative sources, such as administrative records or household surveys, taking into account their relative advantages and limitations. While these sources offer valuable information, they often have limitations. Household surveys, for example, typically rely on samples and might lack the geographic coverage needed for detailed analysis of small areas or certain small population groups. Censuses, on the other hand, can collect comprehensive data through direct interviews or self-enumeration. This allows for probing questions and clarification, which may not be possible with self-enumeration or administrative records. Those topics for which no alternative sources exist should be given higher priority for inclusion in a census, while those for which alternative sources are readily available may be accorded a lower priority.

II. List of topics

5.13. The list of topics included in these recommendations for population censuses are based on the global and regional census experiences in the last several decades. Most of the topics included here are, with some revisions, generally the same as those included in the previous United Nations population census recommendations.[207] However, the concepts and definitions for some of the topics relating to economic characteristics have been substantially revised to reflect the more recent recommendations of the International Conference of Labour Statisticians[208] and other recommendations, while new topics on reason for change of residence, ever resided abroad, reason for international migration, and registration of vital events and assignment of legal identity have been introduced.

5.14. It should be emphasised that no country should necessarily attempt to cover all the topics included in the list of population topics (see Table 5.1). Rather, countries will need to make their selection of topics in the light of the five key considerations discussed in paragraphs 5.1–5.12 above, taking into account any global or regional recommendations relevant to their specific context. In using the classifications of different topics presented in this document, it should be noted that all the one- and two-digit classification levels are recommended, while those at the three-digit level are provided for reference and should be applied only when specific data requirements necessitate a greater level of detail. Ultimately, the decision on which topics to include in a census should be based on a careful assessment of data availability, cost-effectiveness, and the specific needs of the country.

5.15. Over the past several decades, global and regional census experiences have highlighted a set of topics on which there is considerable agreement on their importance and feasibility for data collection in a census. However, the specific topics included in a census will vary depending on a country's unique circumstances, data needs, and available resources. Some countries with highly developed statistical systems may rely on non-census sources like registers and administrative records for certain topics, while others with limited data collection opportunities might use the census to gather information that might otherwise be collected through alternative means.

5.16. Although the set of topics covered in these recommendations is quite comprehensive in terms of those generally considered suitable for inclusion in a population census, it is also recognized that some countries may find it necessary to include additional topics on which information is of particular national or local importance. However, before the final decision is made to include any such additional topics, their suitability should always be carefully tested, and the quality of the resulting data evaluated beforehand.

5.17. To assist countries in using the present publication and in determining their own priorities, lists of recommended population topics are given in Table 5.1, with the core topics shown in boldface. These core topics correspond to those that were included as “priority topics” in the majority of the regional recommendations in previous census decades.

---

[207] *Principles and Recommendations for Population and Housing Censuses Revision 2*, Statistical Papers No. 67/Rev2 (United Nations publication, Sales No. E.07.XVII.8).

[208] For more detail see the Resolution concerning statistics of work, employment and labour underutilization, adopted by the 19th International Conference of Labour Statisticians (Geneva, 2013).

5.18. The topics listed in Table 5.1 are grouped into ten categories:

(a) geographic and internal migration characteristics;
(b) international migration characteristics;
(c) household and family characteristics;
(d) demographic and social characteristics,
(e) fertility and mortality;
(f) registration of vital events and assignment of legal identity;
(g) educational characteristics;,
(h) economic characteristics; and
(i) agriculture.

5.19. Within each category, a distinction is made between topics collected directly (either from responses given in a census questionnaire or taken from an administrative register) or derived from such responses by which is meant that although data for such derived topics also come from information on the questionnaire or the register(s), they do not necessarily come from replies to a specific question. “Total population”, for example, is derived from a count of the persons entered on the questionnaires as persons present or resident in each geographic unit, and the concept of a “Family” is derived from the information collected on relationship to the household reference person. Such derived topics may perhaps be more correctly considered as tabulation components, but they are listed as topics in order to emphasize the fact that the questionnaire must in some way yield this information.

5.20. The paragraph numbers in parentheses after each entry in Table 5.1 refer either to the paragraphs in which the group of topics as a whole is discussed in section IV below or to the paragraphs in which the definition and specifications of individual topics are discussed.

5.21. In the following list of population census topics, core topics are shown in bold and are represented by ♦ for topics that are collected directly, and by □ for those that are derived. Additional topics are represented by ○, and additional topics derived from a core topic are indicated with Δ.

Table 5.1. List of population census topics

A. Geographic and internal migration characteristics (paras. 5.49–5.103)		
(1)	**Place of usual residence** (paras. 5.51–5.56)	♦
(2)	**Place where present at time of census** (paras. 5.57–5.62)	♦
(3)	**Place of birth** (paras. 5.63–5.70)	♦
(4)	**Duration of residence** (paras. 5.71–5.73)	♦
(5)	**Place of previous residence** (paras. 5.74–5.75)	♦
(6)	**Place of residence at a specified date in the past** (paras. 5.76–5.80)	♦
(7)	Reason for change of residence (paras. 5.81–5.84)	○
(8)	**Total population** (paras. 5.85–5.92)	□
(9)	**Locality** (paras. 5.93–**Error! Reference source not found.**)	□
(10)	**Urban and rural** (paras. 5.97–5.103)	□
B. International migration characteristics (paras. 5.104–5.132)		
(1)	**Country of birth** (paras. 5.109–5.113)	♦
(2)	**Country of citizenship** (paras. 5.114–5.119)	♦
(3)	Acquisition of citizenship (paras. 5.120)	○
(4)	**Year or period of arrival** (paras. 5.121–5.125)	♦
(5)	Ever resided abroad (paras. 5.126–5.127)	○
(6)	Reason for international migration (paras. 5.128–5.132)	○
C. Household and family characteristics (paras. 5.133–5.161)		
(1)	**Relationship to the reference person of household** (paras. 5.141–5.152)	♦
(2)	**Household and family composition** (paras. 5.153–5.160)	□
(3)	Household and family status (para. 5.161)	○
D. Demographic and social characteristics (paras. 5.162–**Error! Reference source not found.**)		
(1)	**Sex** (para. 5.165–5.168)	♦
(2)	**Age** (paras. 5.169–5.180)	♦
(3)	**Marital status** (paras. 5.181–5.189)	♦
(4)	Ethnocultural characteristics (paras. 5.190-5.191)	○
(5)	Religion (paras. 5.192–5.196)	○
(6)	Language (paras. 5.197–5.200)	○
(7)	Ethnicity (paras. 5.201–5.207)	○
(8)	Indigenous peoples (paras. 5.208–5.213)	○

(9)	**Disability status** (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	♦
	E. Fertility and mortality (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	
(1)	**Children ever born alive** (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	♦
(2)	**Children living** (paras. 5.260–**Error! Reference source not found.**)	♦
(3)	**Date of birth and sex of last child born alive** (paras. **Error! Reference source not found.**–5.266)	♦
(4)	Births in the past 12 months (paras. **Error! Reference source not found.**–5.269)	Δ
(5)	Deaths among children born in the past 12 months (paras. **Error! Reference source not found.**–5.272)	Δ
(6)	Age, date or duration of first marriage (para. **Error! Reference source not found.**–**Error! Reference source not found.**)	○
(7)	Age of mother at birth of (date or time when) first child born alive (para. **Error! Reference source not found.**)	○
(8)	**Household deaths in the past 12 months** (paras. **Error! Reference source not found.**–5.280)	♦
(9)	Maternal or paternal orphanhood (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	○
	F. Registration of vital events and assignment of legal identity (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	
	G. Educational characteristics (paras. **Error! Reference source not found.**–5.333)	
(1)	**Literacy** (paras. 5.304–5.310)	♦
(2)	**School attendance** (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	♦
(3)	**Educational attainment** (paras. **Error! Reference source not found.**–5.325)	♦
(4)	Field of education and training, and educational qualifications (paras. 5.326–5.337)	○
	H. Use of information and communication technology at individual level (paras. **Error! Reference source not found.**–5.339)	
(1)	Individuals using the Internet (paras. 5.338)	○
(2)	Individuals owning mobile cellular phones (paras. **Error! Reference source not found.**)	○
	I. Economic characteristics (paras. **Error! Reference source not found.**–5.447)	
(3)	**Labour force status** (paras. **Error! Reference source not found.**–5.384)	♦

(5)	**Status in employment** (paras. **Error! Reference source not found.**–5.412)	♦
(6)	**Occupation** (paras. **Error! Reference source not found.**–5.416)	♦
(7)	**Industry** (paras. **Error! Reference source not found.**–5.420)	♦
(8)	Place of work (paras. **Error! Reference source not found.**–5.426)	○
(9)	Institutional sector of employment (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	○
(10)	Working time (paras. **Error! Reference source not found.**–5.436)	○
(11)	**Participation in own-use production of goods** (paras. 5.437–5.442)	♦
(12)	Income (paras. **Error! Reference source not found.**–5.447)	○
J. Agriculture (paras. 5.448–**Error! Reference source not found.**)		
(2)	Own-account agriculture production (paras. **Error! Reference source not found.**–**Error! Reference source not found.**)	○
(3)	Measure of farmland and number of livestock (paras. 5.452–5.456)	○
Legend: ♦ Core topic, collected directly; □ Core topic, derived; ○ Additional topic; and Δ Additional topic, derived from a core topic.		

III. Population count

5.22. One of the main objectives of a population census is to provide an accurate count of the population of a country at a specific point in time. An accurate population count is essential for, inter alia, the efficient planning and delivery of services, distribution of resources, defining boundaries for electoral representation, policy development and a wide range of other administrative, commercial, and statistical purposes.

5.23. A clear distinction should be made at this point between the concepts of a 'population count' and a 'population estimate'. As noted at paragraph 4.151, a *population count* refers to the enumerated population and may relate to the whole, or a subset, of that population. A country may therefore produce one or more population counts, all derived from the enumerated population. A *population estimate* on the other hand refers to the size of the population after the data have been statistically adjusted to account for under- and over-enumeration and other shortcomings in the enumeration.

5.24. Countries are usually most interested in the count and distribution of usual residents because usual residence is generally the best indication of where people will demand and consume services. A count (or estimate) of usual residents is therefore most relevant for many common planning and policy purposes.

5.25. Some countries may supplement the population count from their census with information from other sources, such as administrative records. This can be used to estimate the population of usual residents who were temporarily outside the country at the time of the census. Alternatively, the census itself may collect data on these groups, resulting in an enumerated population that includes individuals who are usually resident but temporarily absent.

5.26. In countries whose censuses use direct enumeration, information about each person can be collected and entered on the census questionnaire either where the person is (or was) present on the day of the census or at their place of usual residence. Paragraphs 2.58–2.74 describe the basis for deciding the place of enumeration in the census.

5.27. As noted at paragraph 5.23 above, a country may need to produce one or more different counts for different national purposes. Population counts may be required on a population present basis, or for the usually resident population, or for the service population (see paragraphs 5.42-5.44). The choice of population count(s) required will depend on national circumstances; some countries will require more than one. The information collected about each person by the census will need to enable the required population counts to be derived. In many cases, for international and regional comparison purposes, the population count based on the generally accepted concept of usual residence will need to be produced.

5.28. A census aims to achieve a full and unduplicated coverage of the population. In practice, NSOs may face a range of challenges in enumerating the population at the place they decide (where present on census day or where usually resident), and in producing the population counts they require. Many of these challenges relate to the difficult-to-enumerate/reach groups of the population and persons for whom the concept of "usual residence" is not easily defined. The latter present an increasing problem as populations become more mobile (nationally and globally) and household and family structures more dynamic.

5.29. In developing strategies for enumerating the population and collecting information to support the required population counts, it is important to consider consistency with the standards for international migration statistics described in paragraphs 5.104–5.108.

A. Population present count

5.30. A population present count (often referred to as the *de facto* population) is a census count that records individual based on their location on the census reference date, regardless of their usual residence. In a questionnaire-based census where no reference is made to usual residence, people are enumerated at the place where they spent the census night. Foreign residents who are in the country at the time of the census will be included but usual residents of the country who are absent at that time will be excluded.

5.31. A population present count removes complications associated with the application of the concept of place of usual residence, and can reduce the incidence of double counting or missing people if the enumeration is carried out in a single day or if reference is made to the same census moment for the whole population. Apart from the benefit of simplicity, a population present count offers a cost advantage because the census does not need to collect additional information about usual residents not at their usual place of residence at the time of the census.

5.32. The major disadvantage of a population present count is that it does not enable a full count of usual residents to be derived, and does not provide a true geographic distribution of usual residents for effective planning and policy purposes.

5.33. A population present count may be a good proxy for a count and distribution of usual residents, particularly if nearly all the population will be at their usual residence at the time of the census, or if the characteristics of those persons present are very similar to the characteristics of usual residents. However, in many countries significant numbers of people will not be at their usual residence at the time of the census, and the characteristics of absent usual residents will be systematically different from non-residents present. In such cases, a population present count will not be a good proxy for a count of usual residents. Large seasonal movements of people due to weather changes, agricultural employment, education, holidays and other factors can result in these systematic differences between the population present and the usually resident population. The ability to produce accurate information on families and households is also reduced when persons temporarily absent from their usual residence (e.g. seasonal workers or students) are not enumerated with their families or households.

5.34. To produce a population present count, information is required on all persons present and the address at which they are enumerated. Additionally, while it is not necessary to produce the population present count, it might be valuable to collect data that identify individuals who are not at their usual residence or who are not usual residents of the country. By collecting this additional data, census agencies can create a more comprehensive and accurate snapshot of the population at a specific point in time, enabling better analysis of population dynamics, migration patterns, and service needs.

5.35. Ideally a population present count should include all the persons present at the census reference moment, regardless of the difficulty of their enumeration. For some of these groups the concept of "at the time of the census" may need to be extended to allow a full enumeration to be completed. When, however, the enumeration is extended over a period of time, the risk of either overcount or undercount may increase. In fact, persons who are at multiple locations during this extended period may be counted at more than one location, or alternatively they may not be counted at any location. Those risks increase further when reference is made to a *census period* rather than to a *census moment*.

B. Usual resident population count

5.36. Countries increasingly prefer a usual resident population count (often referred to as the *de jure* population) because this count offers better information for planning and policy purposes on service demand, households, families and internal migration.

5.37. A **usual resident population count** is a count of all usual residents of a country at the time of the census. Although countries will determine the definition of a usual resident according to their own particular circumstances, it is recommended that in defining a usual resident and the place of usual residence, countries apply the definition contained in paragraphs 2.60 and 2.61. Usual residents may or may not necessarily have citizenship of the country, and they may also include undocumented persons or irregular migrants, applicants for asylum or refugees. For statistical purposes, usual residents then may include foreign citizens who reside (legally or illegally), or intend to reside, in the country continuously for either most of the last 12 months or for 12 months or more, depending on the definition of place of usual residence that is adopted by the country. Persons who may consider themselves usual residents of a country because of citizenship or family ties, but were absent from the country for either most of the last 12 months, or for 12 months or more, depending on the definition

adopted, should be excluded. Conversely, persons who are normally resident in the country but who are temporarily absent should be included in the usually resident population. Countries applying a different definition of a usual resident for national purposes should produce a usual resident population count using the recommended 12-month definition for the purposes of international comparability.

5.38. Whether enumeration takes place on a “place where present” basis or on a “where usually resident” basis (as described in paragraphs 2.58–2.66), a count of usual residents can be produced provided that the necessary information is collected. In which case, information is required on all usual residents and the address of their usual residence, with sufficient detail to be able to record usual residence at the lowest geographic area level required for tabulation. If the census is taken on a population present basis, then the information collected needs to differentiate clearly among persons enumerated at their usual residence, persons usually resident who were elsewhere at the time of the census, and persons present who are usually resident elsewhere. Information should also be collected to identify those persons present on census day who are not usual residents of the country. If, however, the census is taken on a usual residence basis, then, in order to ensure full coverage, information about all usual residents needs to be collected with respect to their place of usual residence, regardless of whether they are present at the time of the census or not.

5.39. It may be difficult to obtain information on those usual residents who are absent from the country at the time of the census, particularly where no other person is present at the place of usual residence at the time of the census to provide the relevant information. In such cases, imputations, using data from the census itself or other sources (such as a PES) and as long as the quality of these sources are appropriate, could be utilised regarding the number and characteristics of these usual residents to estimate the usually resident population count.

5.40. Determining the "place of usual residence" can be challenging when a person is considered to have more than one residence, within or outside the census country. This is particularly the case for people who may spend parts of the time in communal establishments or institutions, such as schools/dormitories or military camps. Additionally, some individuals might not have a conventional usual residence, such as nomadic peoples or persons sleeping rough.[209] In such cases, the place where they are enumerated on the census date can be considered their place of usual residence for census purposes. To ensure consistency and minimize discrepancies, countries should develop clear operational rules for determining usual residence in ambiguous cases. These cases may include situations where it is unclear whether a person is a usual resident of the country or where their place of usual residence within the country is uncertain.

5.41. There are population groups for which some uncertainty may arise in defining their place of usual residence within the country. The recommended conventional treatment of these cases is as follows:

a) For persons who work away from home during the week and return at weekends, their place of usual residence should be their weekend residence.

[209] People ‘sleeping rough' are those living in the streets or public spaces, without a shelter that can be defined as living quarters. The group sleeping rough is a subset of the homeless. Not everyone who is homeless is sleeping rough. There are many other forms of homelessness, including those living in shelters, cars, or couch surfing.

b) For persons of minor age in primary and secondary education who are away from home during the school term, their place of usual residence should be their family home/residence.

c) For students in tertiary education who are away from home while at college or university, their place of usual residence should be their term-time address regardless of whether this is an institution (such as a boarding school) or a private residence.

d) The institution should be taken as the place of usual residence of all occupants who at the time of the census have spent, or are likely to spend, six months or more in that institution. Examples of occupants of institutions include patients in hospitals or hospices, older persons in nursing homes or convalescent homes, children in orphanages, prisoners and those in juvenile detention centres.

e) Where a person regularly lives in more than one residence within the country during the year, the one where he or she spends the majority of the 12 months before the census should be taken as his or her place of usual residence. These persons are not considered to be persons with no usual residence.

f) For the (national) military, naval and diplomatic personnel and their families located outside the country the following classification rules should be applied:
i) If they are residing abroad for less than 12 months and they intend to return to the place/country of departure, they should be allocated within the country in accordance with the rules for usual residence. In particular, they could be allocated to (by decreasing order of priority):
- The family home address within the country, if any; or
- The duty station within the country to which they were attached before leaving.
ii) If they are residing abroad for at least 12 months or if they do not intend to return to the place/country of departure (although returning in, or have visited the country, within a 12-month period), they should be attributed to a "virtual region" (extra-region) of the country of departure.

g) The place of enumeration should be taken as the place of usual residence of homeless or roofless persons, nomads, vagrants and persons with no concept of usual residence.

h) For a child who alternates between two households within the country (for instance when his or her parents are not living together because of divorce, legal separation, or other reasons), the household where he or she spends the majority of the 12 months before the census should be considered as his or her place of usual residence. Where an equal amount of time is spent with both parents, the place of usual residence should be with the parent or household with whom the child is staying at the census reference time.

C. Supplementary population counts

1. Service population count

5.42. Counts of usual resident and/or population present are the most commonly produced census population counts, since in most cases they provide a good indication of the geographic distribution of, demand for, and patterns of, the use of services. However, in some countries or specific parts of some countries, the demand for, and use of, services, as well as other considerations such as exposure to risks and opportunities, is shaped significantly by temporary movements. This is the case, for example: in some agricultural areas which may have very few usual residents but large temporary populations during harvest season; in tourist resorts with large numbers of visitors but relatively few usual residents;

or in urban centres where few people live but many are present as workers during the daytime. In such cases the NSO may wish to produce supplementary population counts.

5.43. Service populations are relevant where a significant proportion of the population providing or using services in an area are not usual residents of that area. Types of service population counts include daytime populations, workplace populations and visitor populations. In some countries there may also be an interest in foreign service populations, consisting of foreign citizens who cross the border regularly to provide or consume services. This is particularly important in the planning and provision of transport services.

5.44. To produce a service population count, in addition to an estimate of usual residents, information is required about where people provide or require services. For seasonal populations (holiday, resort), information is needed on the destination and timing of seasonal trips. Some countries will produce service population counts by supplementing the population present count or usual resident population count with information from other sources, such as visitor information from hotels and resorts, to produce visitor population counts. Alternatively, such additional information may be collected by the census itself.

2. Counts of population subgroups

5.45. Accurate population counts, necessary for the efficient planning and delivery of services, distribution of resources, defining of boundaries for electoral representation, policy development and the design and analysis of household surveys, are required for various population subgroups within a country. These subgroups are typically based on geography, age and sex. There may also be a need to identify other populations such as the school population, working population, indigenous population or population groups in vulnerable situations in order to enable more informed policy formulation and better targeted service provision. A range of characteristics will be required to identify such population subgroups, depending on the services being planned, the resources to be distributed and so on. The need for population counts for particular subgroups will determine some of the questions asked in the census or the information collected from registers.

D. Difficult-to-enumerate groups or difficult-to-reach groups

5.46. Difficult-to-enumerate groups can be defined as those for whom a real or perceived barrier exists to full and representative inclusion in the data collection process[210]. In fact, some populations present special challenges of various sorts that make them harder to enumerate than the general population. The definition of difficult-to-enumerate groups can vary between countries as the reasons why members of a population group are hard-to-reach (or difficult-to-enumerate) can vary according to the

[210] Tourangeau R., Brad Edwards B., Johnson T. P., Kirk M. Wolter K.M. Bates N. (eds), Hard-to-Survey Populations, Cambridge University Press (2014), DOI: https://doi.org/10.1017/CBO9781139381635.003; U.S. Census Bureau: Counting the hard to count in a census (2019), Available from: https://www.census.gov/content/dam/Census/library/working-papers/2019/demo/Hard-to-Count-Populations-Brief.pdf

context of each national, geographic, or social environment. Nonetheless, countries should ensure that the following difficult-to-reach/enumerate groups are included in the population count:

(i) **Nomads and persons living in areas to which access is difficult.** Making contact with these groups to enumerate them can be difficult, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups. For example, countries might consider asking those who provide services to these groups to assist with their enumeration. Seasonal movements may be identified in advance and this information can be used by collectors to enable contact. There needs to be planning and consultation, particularly with influential members of these groups, prior to the census to organize for their enumeration. Communications that publicize the benefits of the census and engaging appropriate leaders in support of the census may assist coverage. Awareness of cultural issues relevant to specific groups should also be considered in developing enumeration strategies. (More detail on some of the methods than may be adopted to enumerated nomads is given at paragraph 3.170.)

(ii) **Civilian residents temporarily absent from the country.** Usual residents absent from the country at the time of the census should be included in the usual resident population count, despite their absence. To produce a usual resident count, countries may collect information on these people from another family or household member present at the time of the census, but (as noted at paragraph 5.39) where a complete family or household is outside the country at the time of the census, it may not be possible for the census to collect information about these people. Estimates for usual residents temporarily absent from the country based on other sources may be required to produce reliable estimates of usual residents for planning and policy purposes.

(iii) **Civilian foreign citizens who do not cross a frontier daily and are in the country temporarily.** These may include undocumented persons, or transients on ships in harbour at the time of the census. If these groups are present in the country on census day, they should be included in the population present count (they would typically be excluded from the count of usual residents). It is important to include these groups in the population count if their demand for services is to be considered for planning and policy development purposes. However, these groups may prefer not to be counted, either because they fear ramifications from being counted or because they do not identify themselves as part of the population of the country. Language and communication may present challenges. Countries need to develop strategies, appropriate for their context, to include these groups in their enumeration.

(iv) **Refugees, asylum seekers and internally displaced persons.** Refugees, asylum seekers[211] and internally displaced persons[212] (in and outside camps) should be enumerated in both the population present count, and in the usual resident count if they meet the established duration criteria. To facilitate analysis, these groups should be reported separately, to allow the calculation of a country's population excluding these categories when required for non-demographic purposes.

(v) **Military, naval and diplomatic personnel and their families located outside the country and foreign military, naval and diplomatic personnel and their families located in the country.**

[211] https://egrisstats.org/recommendations/international-recommendations-on-refugee-statistics-irrs/
[212] https://egrisstats.org/recommendations/international-recommendations-on-idp-statistics-iris/

Apart from the difficulties mentioned in (b) and (c) that are common to groups who are absent from their own country, enumeration of these groups is subject to diplomatic protocols. Detailed counts and characteristics of these groups may be considered sensitive on security grounds in some countries. Counts of these groups may be available from administrative records.

(vi) **Civilian foreign citizens who cross a frontier daily to work in the country.** This group should be excluded from a usual resident population count. The practice of counting people where they spend census night removes much ambiguity and reduces possible duplication. The difficulty then is trying to include them in a service population if countries want to consider this group in policy development and in planning service delivery.

(vii) **Civilian national citizens who cross a frontier daily to work in another country.** These persons are usual residents of the country and should be included in the population count.

(viii) **Merchant seafarers and fishermen resident in the country but at sea at the time of the census.** This group includes those who have no place of residence other than their quarters aboard ship. Identifying that the ship will be at sea at the time of the census may be problematic, so countries will need to develop strategies to ensure inclusion of this group in the population count. This may include providing this group with census forms before their ship goes to sea or enumerating the ship before the census date.

(ix) **Homeless or roofless persons, vagrants and persons with no concept of usual residence.** These should be included in the population count, and the NSO should work with local government agencies, charities and other supporting bodies that provide support for these population groups to identify the best method of collecting census information from them.

(x) **Persons living in buildings with restricted access.** Access to persons living in gated communities, condominiums and apartment buildings in order to enumerate them could be difficult, particularly as part of a point-in-time count. Enumeration may need to be done at a different time, over an extended period, or by using alternative methods to enable contact with these groups.

(xi) **Stateless persons[213].** These are individuals who are not considered as nationals by any State under the operation of its laws. They are often undocumented and may not wish to be enumerated. However, every effort should be made to include such persons - both those with recognised stateless status and those without a recognised nationality status[214] - in the census. The NSO should work with responsible government agencies, non-governmental organizations familiar with this population group and the United Nations High Commissioner for Refugees (UNHCR) to establish the best method for including stateless persons in the population count. Information on country of citizenship is generally essential for the identification of this group, though the NSO should consult with relevant ministries and agencies, in addition to the UNHCR, to determine whether additional information (such as residence history or identity documentation) may be required to establish the status of a stateless person.

[213] https://egrisstats.org/recommendations/international-recommendations-on-statelessness-statistics-iross/

[214] IROSS, see paragraph 172, more info https://egrisstats.org/recommendations/international-recommendations-on-statelessness-statistics-iross/.

(xii) **Persons with an irregular migration situation.** Persons moving outside regular migration channels[215] as well as persons who entered regularly but overstayed the permitted duration. The means of ensuring that migrants in general are included in the census are covered at paragraph 3.172.

(xiii) **Persons who refuse to participate in the census due to political reasons.** Some individuals may refuse to participate in the census due to political reasons, such as distrust of the government. Others may have cultural or religious beliefs that prevent them from participating. NSOs should work closely, via a strong communication/publicity team, with the major political parties and relevant cultural organizations to encourage universal participation.

(xiv) **Indigenous peoples**. See paragraphs 5.208-5.213. The indigenous population may tend to live in sparsely populated remote areas which are usually difficult to access and enumerate. In some countries, these areas are managed by a leader and councillors who maintain up to date records on basic demographic and social data (including address) on each family/household. In such cases, the NSO could access these records and work in collaboration with the respective leaders to mobilize and or locate the actual families/households for enumeration purposes. In cases where the administration of such areas is the responsibility of a particular government ministry, agency or department, the NSO could access these population with the assistance of the responsible organisation.

(xv) **Seasonal workers**. These are persons who are employed in a temporary employment situation that repeats yearly around the same time. Some of the most common industries for seasonal workers are farming and agriculture, tourism, and construction. This group should be included in the usual resident population count if they meet the 12-month criteria described in paragraph 5.37 and counted where they spend census night. This will remove ambiguities and limit possible duplications.

(xvi) **Older persons and those with cognitive impairments**. Older people living in institutions (relevant for both enumeration-based and register-based censuses), and persons living in private households but without the cognitive capability to complete a questionnaire-based census may require additional support and tailored strategies (as noted at paragraph 3.175) to ensure that they are included in the census.

(xvii) **People with language difficulties**. To ensure inclusivity and accessibility, it is essential to consider the language needs of respondents. As noted at paragraph 3.168, not all individuals will speak or understand the languages used in the census. Therefore, providing translation services and materials in various languages is crucial. Decisions regarding languages and translations should be data-driven to avoid any perception of bias or discrimination. While immigrants often face language barriers, it is important to recognize that other groups, such as ethnic and linguistic minorities or those in rural or remote areas, may also have language-

[215] See: https://www.iom.int/key-migration-terms. Although a universally accepted definition of irregular migration does not exist, the term is generally used to identify persons moving outside regular migration channels. Moreover, categories of migrants who may not have any other choice but to use irregular migration channels can also include refugees, victims of trafficking, or unaccompanied migrant children. Persons who entered a country's territory without authorization as well as persons who arrived in a country legally for work or study and overstayed their permits or visas, as well as persons who received a negative decision in an asylum determination procedure or who lose their residency status because of a conviction for a criminal offence or following a review of their refugee or subsidiary protection status fall under this category.

related challenges. Providing materials in appropriate languages is essential for ensuring that all individuals can participate effectively in the census.

(xviii) **Students.** Tailored approaches are needed to reach students, such as collaborating closely with schools, colleges, and universities. This can help to ensure that students are aware of their legal obligation to participate in the census and that they have the opportunity to do so. Clear and concise communication materials that resonate with student life and their priorities are essential. More details on the strategies for ensuring that students are counted in the census are set out at paragraph 3.174.

(xix) **Persons living in remote rural areas.** As noted at paragraph 3.178 understanding the extent of population in remote rural areas and the associated logistical and management challenges of enumerating them needs careful consideration. Lack of infrastructure, such as access to electricity and Internet connectivity, and transport facilities, may limit the modes of data collection that could be used in these areas.

5.47. The same difficulties apply to register-based censuses. While administrative data can enhance coverage for certain hard-to-reach populations, it may also miss other 'hidden' groups within different population segments. For instance, undocumented migrants are not automatically included in population registers, and consequently, they are absent from statistics derived from administrative records. However, many other difficult-to-enumerate groups might be included but not easily identifiable within administrative data. It is important to recognize these potential limitations when relying primarily on administrative sources.

IV. Definitions and specifications of topics

5.48. This section contains the recommended definitions and specifications of all topics presented in the order in which they appear in Table 5.1 above. It is important that census data be accompanied by the corresponding definitions adopted in carrying out the census. It is also important that any changes in definitions that have been made since the previous census are reported in the metadata and, if possible, accompanied by an assessment of the effect of such changes on the relevant data, in order to ensure that users will not confuse valid changes over a period of time with apparent increases or decreases resulting from changes in such definitions.

A. Geographic and internal migration characteristics

5.49. It should be noted that "place of usual residence" and "place where present at time of census" may be considered alternative topics when countries do not have the resources to investigate both topics for general census purposes. Some countries, however, will want to investigate both topics. The relationship between the two topics and their further relationship to the topic of "place of enumeration" are set out in paragraphs 2.58–2.66).

5.50. It is recommended that countries investigating only "place where present at time of census" for general purposes should also obtain information on "place of usual residence" for all persons who do not usually reside in the household where they were enumerated, to be used in connection with the information on "place of birth", "duration of residence", "place of previous residence" or "place of residence at a specified date in the past" for the purposes of determining internal migration status. If, in the compilation of the population of geographic units, persons are allocated to the place where they were present at the time of the census, information on the four above-mentioned migration characteristics will be irrelevant for persons who were only visiting, or transient in the place at which

they were present. Since such persons must, in any case, be identified in the questionnaire as non-residents so that they will not be erroneously classified as recent immigrants, information on place of usual residence should be collected, which will make it possible to include the entire population in the tabulation of internal migration characteristics.

1. Place of usual residence (core topic)

5.51. Information on the number of people usually residing in an area is basic to most informed decision-making about the area, whether it be a country, an urban agglomeration or a civil division. The number of residents determines the levels of most services required in an area.

5.52. The *place of usual residence* may be the same as, or different from, the place where the enumerated person was present at the time of the census or his or her legal residence. For a definition of place of usual residence, see paragraphs 2.59–2.61.

5.53. Although most persons will have no difficulty in stating their place of usual residence, some confusion is bound to arise in a number of situations where persons have more than one residential address. These cases might include persons who maintain two or more residences, students living at school, members of the armed forces living at a military installation but still maintaining private living quarters away from the installation, and persons who sleep away from their homes during the working week but return home for several days at the end of each week (see paragraph 5.41). In some other circumstances, referring to the person's intentions for the future may assist the determination of the place of usual residence.

5.54. Problems may also arise with persons who have (a) been residing at the place where they are enumerated for some time, but do not consider themselves to be residents of that place because they intend to return to their previous residence at some future time; or (b) left the country temporarily but are expected to return after some time longer than 12 months from the departure. In such instances, clearly stated time limits of presence in or absence from a particular place must be based upon the 12-month limit and used to determine whether or not the person is usually resident there. The 12-month criterion is necessary for determining whether or not a person is usually resident in the country. It is also required for international comparability of migration statistics.

5.55. If each person is to be recorded in the questionnaire only at his or her place of usual residence, the topic need not be investigated separately for each person, because the information will be available from the location information entered for the questionnaire as a whole.

5.56. Information on the place of usual residence should be collected with enough locational details to enable tabulations to be made for the smallest geographic subdivisions required by the output plan and to meet the requirements of the database within the cost limits and operational procedures required to code to a fine degree of detail.

2. Place where present at time of census (core topic)

5.57. In cases where the census is taken on the basis of "place where counted" (often referred to as a *de facto* count), this topic may fulfil some of the functions of place of usual residence.

5.58. The *place where present at time of census* is, in theory, the geographic place at which each person was present on the day of the census, whether or not this was his or her place of usual residence. In practice, the concept is generally applied to the place where the person was present at the moment of the census, because many persons may not be physically present at the place of enumeration during most of the day.

5.59. The concept is sometimes further extended to apply to the night preceding the day of actual enumeration in cases where the enumeration extends over a long period of time and persons are not likely to be able to supply information relating to a single moment in this period. Other departures from the definition may be necessary to deal with individual cases, such as persons travelling during the entire night or day of the census and persons who work night shifts.

5.60. If each person is to be entered on the questionnaire only at the place where he or she was present at the time of the census, the topic needs to be investigated separately for each person.

5.61. Information on the place where each person was present should be collected with enough locational details to enable tabulation to be made for the smallest geographic subdivisions required by the tabulation plan and to meet the requirements of the database within the cost limits and operational procedures required to code to a fine degree of detail.

5.62. For countries relying on administrative data sources for their census, the concept of "present at the time of the census" may not be relevant. In such cases, and especially when the census population has legal implications, it is recommended to define the population based on usual residence.

3. Place of birth (core topic)

5.63. Information on the place of birth is a major input to the development of policies relating to migration and the related issues of service delivery to migrants. For the purposes of measuring internal migration, migrants are defined as those persons who are usually residing in a civil division of the country at the time of the census, but were previously resident outside that division. That is, movements within the civil division should not be regarded as being migratory.

5.64. The *place of birth* for those persons born within the country is the civil division in which the person was born; for those born in other countries, it is the country of birth. For persons born in the country (the native-born population), the concept of place of birth usually refers to the geographic unit where the mother of the individual resided at the time of the person's birth and not the location of the hospital/health facility in which the birth occurred. In some countries, however, the place of birth is defined as the geographic unit in which the birth actually occurred. It is recommended to use the place of birth as where the mother of the individual resided at the time of the person's birth. If this concept is not used, the NSO should explain the definitions it uses in both the census enumerator instructions and in the census reports to aid the interpretation of the data.

5.65. The collection of information distinguishing between the native-born population and those born elsewhere (foreign-born) is necessary where any enquiry on place of birth is made. Even countries where the proportion of foreign-born population is insignificant, and who may only be interested in information on the place of birth of the native-born population, must first separate the native-born from the foreign-born population. It is therefore recommended that place of birth be asked of all persons. In

countries that combine the questions on place of birth and country of birth (where the latter is used to measure international migration), the guidance on the country of birth (see paragraphs 5.109–5.113) should apply.

5.66. Information on the place of birth of the native-born population is usually used primarily for the investigation of internal migration. For countries that have been recently formed from parts of previously separate entities, however, such information may be of use in assessing the relative size of the population segments from each of those entities and their distribution throughout the country.

5.67. Information on whether or not a person is "born in the country" captures the population according to the boundaries at the time of the census. Using the "born in the country" concept would account for individuals who may have been affected by changes to a country's boundary.

5.68. For the purposes of measuring internal migration, it is usually sufficient to collect information only on the major civil division (state, province or department, for example) in which the place of birth is located. If desired, more detailed information on the subdivision of a specific locality can be collected and used for accurate coding of the major division or for presenting data for smaller areas.

5.69. However, for more detailed studies of internal migration, data on the place of birth of the native-born population even in terms of major civil divisions may not be adequate. For better understanding of the movements of people since birth it ~~is~~ may be necessary to collect information at the smallest possible geographic level, bearing in mind that:

(a) The boundaries of administrative units such as cities and other civil divisions will change over time, which may give rise to ambiguity in the data reported;

(b) The costs of coding the reported data to these smaller units may be prohibitive, especially where there are many units and the population is highly mobile.

To overcome the first problem, to the extent possible, both national and subnational boundaries should refer to the boundaries applying at the time of the census. Countries should address the second problem in the light of their own circumstances, bearing in mind the reduced value of place of birth as a measurement of internal migration in a very mobile population.

5.70. It is recommended that, for the study of internal migration, the data on place of birth be supplemented by information collected on duration of residence (see paragraphs 5.71–5.73) and place of previous residence (see paragraphs 5.74–5.75) or of residence at a specified date in the past (see paragraphs 5.76 and 5.80).

4. Duration of residence (core topic)

5.71. The *duration of residence* is the interval of time up to the date of the census, expressed in complete years, during which each person has lived in (a) the locality that is his or her usual residence at the time of the census; or (b) the major or smaller civil division in which that locality is situated.

5.72. In collecting information on duration of residence, it should be made clear that the interest is in length of residence in the major or smaller civil division, or the locality, but not in the particular housing unit. The concept of duration of residence also relates to the most recent move to the current place of usual residence.

5.73. Data on the duration of residence have only limited value in themselves because they do not provide information on the place of origin of in-migrants. Therefore, when the topic is investigated, the place of previous residence should also be collected so that the data can be cross-classified.

5. Place of previous residence (core topic)

5.74. The *place of previous usual residence* is the major or smaller civil division, or the foreign country, in which the individual resided immediately prior to migrating into the civil division of present usual residence.

5.75. As with the topic of duration of residence, data on the place of previous residence have only limited value in themselves because they do not provide information on the time of in-migration. Therefore, when the topic is investigated and included in the census, the duration of residence (see paragraphs 5.71–5.73) should also be included so that the data can be cross-classified. Alternatively, countries may choose to include a question on place of residence as a specified date in the past (see paragraphs 5.76–5.80 below).

6. Place of residence at a specified date in the past (core topic)

5.76. The *place of usual residence at a specified date in the past* is the major or smaller division, or the foreign country, in which the individual resided at a specified date preceding the census. The reference date chosen should be that most useful for national purposes. In most cases, this has been deemed to be one year or five years preceding the census (or both of these time frames in cases where internal migration is of particular importance to users and resources are sufficient to code the data).

5.77. The former reference date provides information for statistics of both recent internal and international migration during a single year, while the latter may be more appropriate for collecting data for longer-term analysis of migration. When selecting the reference date, the ability of individuals to recall with accuracy their usual residence one year or five years earlier than the census date should be considered. For countries conducting quinquennial censuses, the date of five years earlier can be readily tied in, for most persons, with the time of the previous census, but it should be noted that a one-year recall is likely to result in more accurate information than a five-year recall.

5.78. In countries where using a time reference of either one year or five years preceding the census may present recall difficulties, NSOs may choose to adopt a reference period that can be associated with the occurrence of an important event that most people will remember.

5.79. For foreign-born persons, the collection of information on year of first or last arrival in the country is recommended (see “International migration characteristics”, paragraphs 5.121–5.125).

5.80. However, no matter what previous date is used, provision must be made for the treatment of infants and young children who are resident at the time of the census but would not have been born at the earlier date. Tabulations of the data should indicate the nature of the treatment of this group.

7. Reason for change of residence

5.81. Some countries may also wish to collect information on the *reason for change of residence* of persons who have reported a change of usual residence in accordance with conditions set out in paragraphs 5.74-5.80. The information should refer only to the *main* reason for such a change of residence. If a person had moved from the place of previous residence for the purpose of education, for example, but at some subsequent point in time had obtained employment, the reason for change of residence would be 'education'.

5.82. As different members of the household might change residence for different reasons, the information should be collected for each member of the household.

5.83. There are no recommendations on the specific reasons, and countries should devise a classification in accordance with their own needs. However, examples of some common reasons[216] for change of residence include:

a) work or employment (including military service);
b) education and training;
c) marriage, family reunification or family formation;
d) forced displacement (conflict, persecution, human rights abuses, natural disasters, etc.); or
e) other specified reason.

5.84. In collecting information on reason for change of residence, it should be made clear that it is with respect to the topic of 'place of previous residence' (see paragraphs 5.74-5.75) or as a sub-topic of the item on 'ever resided abroad' (see paragraphs 5.126-5.127).

8. Total population (core topic)

5.85. For census purposes, the *total population* of the country consists of all the persons falling within the scope of the census. In the broadest sense, the total may comprise either all usual residents of the country or all persons present in the country at the time of the census. The total of all usual residents is (as noted at paragraph 5.35) generally referred to as the de jure population, while the total of all persons present is (as noted at paragraph 5.58) referred to as the *de facto* population.

5.86. In practice, however, countries do not usually fully achieve either type of count, because one or more groups of the population are included or excluded, depending on national circumstances. The general term used to describe the total might imply a treatment opposite to the one given to any of

[216] Expert Group on Refugee, Internally Displaced Persons and Statelessness Statistics (EGRISS) (2023). Compilers' Manual on Forced Displacement Statistics. Chapter 3: Forced Displacement Data Sources. https://egrisstats.org/wp-content/uploads/2023-EGRISS_CM_Final-web.pdf, p. 155. And, United Nations Department of Economic and Social Affairs, Statistics Division (2022). Handbook on Measuring International Migration. Chapter 7: Demographic and Social Characteristics of Migrants. https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf, p. 316.

these groups. It is recommended, therefore, that each country describe in detail the figure accepted officially as the total, rather than simply label it as "*de jure*" or "*de facto*".

5.87. The description should show clearly whether each group listed below was or was not included in the total. If the group was enumerated and identified as a separate group, its magnitude should be given; if it was not enumerated, an estimate of its size and the method of estimation should be given, if possible. If any of the groups listed below is not represented at all in the population, this fact should be stated, and the magnitude of the group should be shown as "zero". This may occur particularly with groups (a), (b, (d) and (n) described below (see also paragraph 5.46).

5.88. The groups to be considered are:

a) Nomads;
b) Persons living in areas to which access is difficult (e.g., remote areas);
c) Military, naval and diplomatic personnel and their families located outside the country;
d) Merchant seafarers and fishermen resident in the country but at sea at the time of the census (including those who have no place of residence other than their quarters aboard ship);
e) Civilian residents temporarily in another country as seasonal workers;
f) Civilian residents who cross a frontier daily to work in another country;
g) Civilian residents other than those in groups (c), (e) or (f) who are working in another country;
h) Civilian residents other than those in groups (c), (d), (e), (f) or (g) who are temporarily absent from the country;
i) Foreign military, naval and diplomatic personnel and their families located in the country;
j) Civilian foreign citizens temporarily in the country as seasonal workers;
k) Civilian foreign citizens who cross a frontier daily to work in the country;
l) Civilian foreign citizens other than those in groups (i), (j) or (k) who are working in the country;
m) Civilian foreign citizens other than those in groups (i), (j), (k) or (l) who are in the country temporarily;
n) Refugees, asylum seekers and internally displaced persons in camps;
o) Transients on ships in harbour or at airports at the time of the census;
p) Persons living in slums and informal settlements.

5.89. In the case of groups (h) and (m), it is recommended that an indication be given to describe the criteria used (including with respect to the 12-month rule) in determining the presence in, or absence from, the country as temporary.

5.90. In those countries where the total population figure has been adjusted for under-coverage or over-coverage (or net under-coverage), both the enumerated figure and the estimated adjusted population figure should be shown and described. In such cases, documentation should be provided for users explaining how the total population figure from the census has been adjusted. Ideally, the detailed tabulations should be consistent with the adjusted population figures. However, where this is not possible, due, for example, to the costs of the methodology for undertaking these adjustments being prohibitive, the detailed tabulations will, of necessity, be based only on the actual enumerated population.

5.91. The use of both adjusted and non-adjusted population figures may be confusing for some users, particularly those unfamiliar with the adjustments made. For example, policy planners may find it more helpful to receive only the adjusted population estimates, while researchers are likely to require both figures for their analysis.

5.92. The population of each geographic unit of the country, as with the total population of the country may comprise either all usual residents of the unit (see paragraph 5.52) or all persons present in the unit at the time of the census (see paragraphs 5.57–5.58).

9. Locality (core topic)

5.93. For census purposes, a *locality* should be defined as a distinct population cluster (also designated as inhabited place, populated centre, settlement and so forth) in which the inhabitants live in neighbouring or contiguous sets of living quarters and that has a name or a locally recognized status. It thus includes fishing hamlets, mining camps, ranches, farms, market towns, villages, towns, cities and many other population clusters that meet these criteria. Localities can be defined using a national methodology and/or the methodology associated with the Degree of Urbanisation[217], particularly for international statistical comparison purposes. The census reports should explain which methodology has been applied, and if it departs from the concept of a distinct population cluster, provide an explanation to aid in the interpretation of the data.

5.94. Localities as defined above should not be confused with the smallest civil divisions of a country. In some cases, the two may coincide. In others, however, even the smallest civil division may contain two or more localities. On the other hand, some large cities or towns may contain two or more civil divisions, which should be considered as subdivisions of a single locality rather than separate localities.

5.95. A large locality (that is to say, a city or a town) is often part of an urban agglomeration, which may comprise the city or town proper together with a suburban fringe or heavily populated area lying outside, but adjacent to, its boundaries. The urban agglomeration is therefore not coterminous with the locality but is an additional geographic unit, which may include more than one locality. In some cases, a single large urban agglomeration may comprise several cities or towns and their suburban fringes. The components of such large agglomerations should be specified in the census results.

5.96. To classify localities by size, it is recommended to use a subset of the following size classes and to place an open-ended category at the two extremes:

- 49 inhabitants or less
- 50-99
- 100-249
- 250-499
- 500-999
- 1,000-2,499
- 2,500-4,999
- 5,000-9,999

[217] The Degree of Urbanisation is a harmonised methodology to facilitate international statistical comparisons and to classify the entire territory of a country along an urban-rural continuum. The Degree of Urbanisation classification defines cities, towns and semi-dense areas, and rural areas. See https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/ks-02-20-499.

- 10,000-24,999
- 25,000-49,999
- 50,000-99,999
- 100,000-249,999
- 250,000-499,999
- 500,000-999,999
- 1,000,000-2,499,999
- 2,500,000-4,999,999
- 5,000,000 or more inhabitants

10. Urban and rural (core topic)

5.97. Urban and rural areas can be defined using a national methodology, which can be established in accordance with a country's own needs. National methodologies for classifying urban and rural areas often incorporate criteria such as population size, density, land use patterns, administrative boundaries, and access to basic services. However, for international statistical comparisons, urban and rural areas should be delineated using the Degree of Urbanisation (DEGURBA) methodology which was endorsed by the UN Statistical Commission in 2020. For a full elaboration on the DEGURBA methodology, please see the manual[218].

5.98. A classification of areas as urban or rural should be done at the smallest administrative unit of the country, or the smallest census data collection unit. The classification should be made, in addition to the national methodology, first and foremost, on a measure of population density and population size.

5.99. Some of the information required for classification as urban or rural may be provided by the census results themselves, while other information may be obtained from external sources. The use of information provided by the census (as, for example, the size class of the locality or the percentage of the population engaged in agriculture), whether alone or in conjunction with information from other sources, means that the classification will not be available until the relevant census results have been tabulated. If, however, the census plans call for the investigation of a smaller number of topics in rural areas than in urban areas or for a greater use of sampling in rural areas, the classification must be available before the enumeration takes place. In these cases, reliance must be placed on previous censuses and/or external sources of information, even if only to bring up to date any urban–rural classification that was prepared at an earlier date.

5.100. The usefulness of housing census data (for example, the availability of electricity or piped water) collected simultaneously with, or not too long before, the population census should be kept in mind. (Housing census topics are discussed in Chapter II.) For assembling information from more than one source, the importance of a well-developed system of geocoding should not be overlooked.

[218] EC, FAO, ILO, OECD, UN-Habitat and the World Bank (2021). Applying the Degree of Urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons — 2021 edition https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-02-20-499

5.101. The traditional distinction between urban and rural areas within a country has been based on the assumption that urban areas, no matter how they are defined, provide a different way of life and usually a higher standard of living than are found in rural areas. In many developed countries this distinction has become blurred, and the principal difference between urban and rural areas in terms of the living standards tends to be the degree of population concentration or density. On the other hand, the differences between urban and rural ways of life and standards of living remain significant in developing countries, but even there, rapid urbanization in these countries has created a great need for information related to different types of urban areas.

5.102. Hence, although the traditional urban–rural dichotomy is still needed, a classification by Degree of Urbanisation can usefully supplement the dichotomy as it captures characteristics related to the rural-urban continuum from the most sparsely settled areas to the most densely built-up localities with seven classes. When using the DEGURBA methodology, it is recommended to use the three classes at level 1, which include:

(i) **Cities:** Densely populated areas with a high concentration of economic and social activities.
(ii) **Towns and semi-dense areas:** Intermediate areas with a mix of urban and rural characteristics.
(iii) **Rural areas:** Sparsely populated areas with predominantly agricultural or natural landscapes.

5.103. To effectively define urban and rural areas, the DEGURBA methodology should be applied first to a population grid and then to smaller administrative or statistical areas. Ideally, a population grid should be created using point-based census data. If such data is unavailable, the DEGURBA methodology can be used to estimate population distribution. NSO should, however, recognize the limitations of population grid estimation, as it may not perfectly align with actual population distribution, especially in areas with significant seasonal variation or specific land use patterns (e.g., non-residential resort areas).

B. International migration characteristics

5.104. Interest in the movement of people across national boundaries - international migration -has steadily grown among countries concomitant with the increase in such movement. The decision to collect and disseminate information on international migration in a census is dependent upon a number of considerations and national circumstances, primarily the national need for such information. Data on international migrants can provide information on the diversity of a population and can serve to identify particular population subgroups. This present section on international migration supplements and expands the topic “geographic and internal migration characteristics”, covered in the previous Section A.

5.105. The revised conceptual framework on international migration and mobility statistics[219] deals with both international migration flows and international migrant stocks, identifies population censuses as being one of the main sources for collecting data on international migrants and their characteristics, and recognizes the value of the population census to produce an estimate of annual flows when a question on place of residence one year before the census is included. It emphasizes consistency and

[219] Conceptual frameworks and Statistical Concepts and Definitions on International Migration (2021). https://unstats.un.org/unsd/demographic-social/migration-expert-group/task-forces/TF2-ConceptualFramework-Final.pdf

alignment between such stocks and flows and can determine temporary mobility and temporary populations for improved accuracy in international migration data.

5.106. This section is concerned chiefly with the topics relating to measurement of international migrant stocks and flows as derived from field-based population censuses.

5.107. In the revised conceptual framework on international migration statistics, four subgroups of the resident population critical to understanding immigrant and emigrant populations are defined: native-born citizens, native-born foreign citizens (including stateless or those without citizenship), foreign-born citizens and foreign-born foreign citizens (including stateless or those without citizenship).

5.108. Consequently, for the study of international migration and migrants, identifying all four subgroups of the resident population, and producing statistics for all four subgroups, is necessary. In order to identify members of those groups, two key items must be recorded in the census: (a) the country of birth, and (b) the country of citizenship. In addition, it is important to record the year of arrival in the country so as to establish the length of stay of resident international migrants.

1. Country of birth (core topic)

5.109. *Country of birth* is the country in which the person was born. This information is generally collected from the same question used to determine 'place of birth' covered at paragraphs 5.63-5.70 above. The concept of country of birth may refer to either the country in which the mother was usually residing at the time of the person's birth or the country where the birth actually took place. Either concept can be used depending on the information needs of the country; but for the purpose of measuring international migration, it is recommended to use the country where the mother of the individual resided at the time of the person's birth. Each country should explain which definition it used in the census. It should be noted that the country of birth of a person is not necessarily the same as his or her country of citizenship, which is a separate census topic dealt with at paragraphs 5.114-5.119 below. It is recommended that country of birth be asked of all persons in order to distinguish the native-born from the foreign-born population. The collection of this information is necessary even in countries where the proportion of foreign-born population is small. For any foreign-born respondents who cannot identify their country of birth, at least the continent or region where that country is located should be ascertained.

5.110. For purposes of both internal consistency and international comparability, it is recommended that information on the country of birth be recorded according to national boundaries existing at the time of the census. Information on the year of arrival in the country (see paragraph 5.121 below) can be used to identify persons who owe their status of foreign-born to changes in national boundaries.

5.111. The country of birth should be coded in sufficient detail to allow for the identification of all relevant countries of birth, and for such purposes, it is recommended that countries use the numerical

coding system presented in the UN's *Standard Country or Area Codes for Statistical Use*.[220] The use of standard codes for classification of the foreign-born population by country of birth will enhance the usefulness of such data, including an international exchange of foreign-born population statistics among countries. The option "Unknown" should be added to the codes for country of birth. If countries decide to combine countries into broad groups, it is recommended that the standard regional and subregional classifications identified in the UN publication be adopted.

5.112. Countries with a significant number of international migrants may wish to collect, additionally, information on the country of birth of parents (both father and mother), in which case the information should be asked of all respondents following the same guidelines. The decision to collect and disseminate such information in a census is dependent upon a number of considerations and national circumstances, including, for example, the suitability, sensitivity and relevance of seeking information relating to persons who may not be residents in the country.

5.113. However, information on the country of birth of parents can be used, in combination with information on the person's own country of birth, to identify native-born children of the foreign-born population (the so-called "second generation") and to study the integration processes and outcomes of migrants and their descendants. Moreover, in countries that have experienced return migration, information from this topic allows the identification of foreign-born children of native-born parents.

2. Country of citizenship (core topic)

5.114. *Country of citizenship* is defined as the country of which a person is a citizen and with which that person enjoys a particular legal bond, acquired by birth, naturalization, marriage or some other mechanism. For the purposes of the census a 'citizen' is a person who is legally recognised as a subject of the country of enumeration; thus, a non-citizen may be a foreign-citizen (that is a citizen of another country), or a stateless person or a person with unknown or undetermined citizenship status. Because the country of citizenship will not always be the same as the country of birth, information on both topics should be collected in a census.

5.115. Information on the country of citizenship is particularly important for foreign citizens. It is important to record country of citizenship as such and not to use another concept (such as nationality), since some of these concepts may also be used to designate ethnic group (see paragraph 5.202).

5.116. As with country of birth, it is important that NSOs code the information on country of citizenship in sufficient detail to allow for the identification of all individual countries, using the same UN guidance (see footnote 216 at paragraph 5.111). The use of standard codes for classification of the foreign population by country of citizenship will enhance the usefulness of such data and permit an international exchange of information among countries on their foreign populations. If countries decide to combine countries of citizenship into broad groups, it is recommended (as with country of birth) that the UN's standard regional and subregional classifications be adopted. Two additional categories should

[220] United Nations, *Standard Country or Area Codes for Statistical Use*, http://unstats.un.org/unsd/methods/m49/m49.htm.

be listed: 1) stateless or no citizenship/nationality status, and 2) unknown citizenship/nationality status. IROSS[221] provides a suitable framework for estimating the combined category of 'stateless persons and those without a recognized nationality status'. However, clearly differentiating between individuals who are stateless and those whose citizenship is unclear presents a significant challenge in most census operations.

5.117. The reliability of reported citizenship may be doubtful in the case of persons whose citizenship has recently changed as a result of territorial changes, or among the emigrants of some newly independent countries where the concept of citizenship may have only recently become important. Clear guidelines issued by the NSO can help improve the accuracy of the data collected. As an aid to the analysis and interpretation of the results, notes on the likelihood of these and other possible causes of misstatement should accompany any tabulations based on citizenship.

5.118. Enumeration and processing instructions should provide clear guidance on the treatment of stateless persons, persons with dual nationality, persons in the process of naturalization and any other groups with ambiguous citizenship. The treatment of these groups should be described in the census reports and included in the metadata for accompanying tabulations.

5.119. In cases where people have more than one citizenship and where this information is useful for decision-making, details may be collected on whether the person holds one or multiple citizenship. If this information is to be published, care should be taken to explain how the possibility of people being included in the table more than once affects the marginal totals on the table. Usually, however, it may be more practicable for tabulations by citizenship to refer to one citizenship only. Thus, persons with multiple citizenships should be allocated to a single “primary” citizenship, for example by giving precedence to the citizenship of the “home” country.

3. Acquisition of citizenship

5.120. In addition to collecting information on country of citizenship, for countries where the population includes a significant proportion of naturalized citizens it may be important to collect information on the method of acquisition of citizenship so as to enable the classification of the population into:

(a) citizens by birth;
(b) citizens by naturalization whether by declaration, option, marriage or other means; and
(c) non-nationals or foreign citizens (including stateless or those without citizenship.

In such countries it may also be useful to ask questions on previous citizenship and year of naturalization.

4. Year or period of arrival in the country (core topic)

[221] IROSS, see paragraphs 154 and 165, more info https://egrisstats.org/recommendations/international-recommendations-on-statelessness-statistics-iross/

5.121. *Year or period of arrival in the country* refers to the calendar year and month of arrival to the country of enumeration of any person who has been resident abroad. This information enables the calculation of the number of completed years between the time of arrival in the country and the date of the census.

5.122. Information on the month and year of arrival also provides the flexibility of classifying foreign-born persons by period of arrival in terms of any prespecified period, such as 1995–1999, 2000–2004, 2005-2009 and so forth. It is thus recommended that the period of arrival be shown, in any tabulation in which the variable appears, in terms of the actual year of arrival.

5.123. It is possible to collect information on either the date of first arrival in the country or the date of the most recent arrival in the country. Each has its own advantages and disadvantages, and the choice should primarily align with the country's specific policy needs and data user requirements. In making the choice of which information to collect, countries should be guided first and foremost by their policy and user needs.

5.124. Information on time since arrival can also be collected by asking how many years have elapsed since the time of arrival, instead of in what calendar year and month the person arrived. However, use of such a question is not recommended because it is likely to yield less accurate information.

5.125. Note that information on the year and month of arrival is focused mainly on persons born outside the country of enumeration, that is to say, persons who must have arrived in that country at some time after their birth. However, it should be noted that the phenomenon of "international return migration" is becoming increasingly common, and countries that have population groups that maintain links to other countries, migrating to or from another country at different life stages (for example, as students or pensioners), may have an interest in collecting information on returning migrants: in this case, the question on year and month of arrival could also be asked of native-born respondents who have ever lived in another country. In addition, it might also be important to collect information on previous country of residence for persons who have ever lived abroad.

5. Ever resided abroad

5.126. The topic "ever resided abroad" focuses on those persons resident in the country at the time of the census who have ever been a resident in another country, regardless of country of birth or citizenship. To collect information on this topic, individuals should be classified depending on whether or not they have ever had a usual residence abroad using the 12-month rule described in paragraph 5.37. Countries may choose to collect information on this topic from: (a) citizens only; (b) native born persons only; or (c) all persons. The choice depends on which population subgroup's arrival or return is most relevant to policy making, and on whether a topic on country of birth or country of citizenship has been included in the census.[222] When asked of all persons, the topic allows the identification of all

[222] From 2022 UNSD Handbook on Measuring International Migration through Population censuses, para. 313 https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

international migrants.[223] When asked of native-born persons, it can serve more specifically to measure the stock of returned international migrants, defined as “persons who previously resided in the country of enumeration who emigrated and subsequently came back to live in the country and stayed or intended to stay for the minimum duration required for establishing residence”.[224]

5.127. For those who have ever resided abroad, the year of arrival (see previous topic) should also be collected.[225] Often, a question about whether an individual has ever resided abroad is asked first. If the response is "yes," a follow-up question about the year of arrival is then asked. Through inclusion of a question on the respondents’ most recent year or period of arrival or return, most countries differentiate between migrants arriving or returning recently and those who arrived or returned at a more distant time in the past. The year of arrival or return can also be matched with the date of birth to permit calculation of the age at which the person arrived or returned to live in the census country, information that can be of great value for research and policy.

6. Reason for international migration

5.128. This topic refers to the main reason for the person’s most recent migratory move.[226] When included in the census, it may be most appropriate to include this topic as a part of the question on ever resided abroad (see paragraphs 5.126-5.127). The question should be asked with explicit reference to the main reason for (last/most recent arrival (or return) to the country, to avoid potential conflation with any subsequent internal migration moves which would be captured when asked in reference to a move from the previous place of usual residence.

5.129. In all cases, it is recommended that only the main reason for migration, as self-declared by the respondent be recorded. The information should be collected at the individual level, for each household member, as the main reason for migration may differ among the members. For example, one member may have moved for employment reasons, while another may have moved as an accompanying family member.

5.130. There are no internationally agreed recommendations on the main reasons to include as answer categories. Countries considering this topic for inclusion should establish a classification that is

[223] From UNECE CES 2020 Census recommendations, para 662. https://unece.org/DAM/stats/publications/2015/ECECES41_EN.pdf

[224] From UN Expert Group on Migration Statistics. https://unstats.un.org/unsd/demographic-social/migration-expert-group/task-forces/TF2-ConceptualFramework-Final.pdf

[225] From 2022 UNSD Handbook on Measuring International Migration through Population censuses, para. 313 https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

[226] From UNECE CES 2020 Census recommandations para 670. https://unece.org/DAM/stats/publications/2015/ECECES41_EN.pdf

appropriate to their national needs. Examples of some of the more common reasons for migration include[227]:
(a) work (including to invest or start a business) or employment (including military assignments)
(b) education and training
(c) marriage, family reunification or family formation
(d) forced displacement (refugees, asylum seekers, temporary protected status, others)
(e) deportation/involuntary return
(g) retirement
(e) other specified reason

5.131. While the main reason for migration can be a topic of interest to governments and researchers, this information alone may not be sufficient to assess the various drivers of migration, as persons may migrate for multiple reasons that include both push and pull factors. This should be taken into consideration when using information on this topic for assessments of population movements due to complex factors such as economic conditions, climate change or forced displacement.

5.132. When a question on main reason for international migration is included in the census, information on the year of arrival (see paragraphs 5.121-5.125) should also be collected to enable analysis of changes across periods and/or over the life-course.

C. Household and family characteristics

5.133. In considering the topics related to household characteristics, it is important to be aware of the differences between the concepts of a "household" and a "family" as used herein.

5.134. A *household* may be:

(a) A one-person household, that is to say, a person who makes provision for his or her own food or other essentials for living without combining with any other person to form part of a multiperson household; or

(b) A multi-person household, that is to say, a group of two or more persons living together under the same roof who make common provision for food, housing or other essentials for living. The persons in the group may pool their resources and have a common budget; they may be related or unrelated persons or a combination of persons both related and unrelated. This arrangement exemplifies the "housekeeping" concept.

This definition is commonly referred to as the "housekeeping" concept. Some countries, however, adopt a different concept, namely, the "household dwelling" concept, which regards all persons living in a single housing unit as belonging to the same household. According to this concept, there is one household per occupied housing unit, that is the number of occupied housing units and the number of households occupying them are equal and the locations of the housing units and households are the same whereas in countries that

[227] Categories in line with EGRISS recommendations, para 155; EGRISS Compiler's Manual (Box A.2, pp. 20); UNSD 2022 Handbook on Measuring International Migration, para 316: https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2022-UNSD-Handbook-Meas-Intern-Migration-E.pdf

use the “housekeeping” concept, the number of households may be greater than the number of dwelling units.

Countries should specify in their census reports whether they used the “housekeeping” or the “household dwelling” concept of a private household.

In countries that use register-based census methods, the concept of "household dwelling" is generally used.

See paragraphs 2.41-2.47 for further details.

5.135. A household may be located in a housing unit (see paragraph **Error! Reference source not found.**) or in a set of collective living quarters such as a boarding house, a residential hotel or camp, or a mobile structure such as a caravan, or may may comprise the administrative personnel in living accommodation within an institution,

5.136. The *family* within the household is defined as those members of the household who are related, to a specified degree, through blood, adoption, or marriage. The degree of relationship used in determining the limits of the family in this sense is dependent upon the uses to which the data are to be put and so cannot be established for worldwide use. See paragraph 5.154 for a definition of the family nucleus.

5.137. Although in practice most households are composed of a single family consisting of a couple with or without children or of one or both parents and their children, it should not be assumed that this correlation always exists; census tabulations should therefore clearly indicate whether they relate to households or to families within households.

5.138. From the definitions of “household” and “family”, it is clear that these are different concepts that cannot be used interchangeably. The key differences between the household and the family are that:

a) A household may consist of only one person but a family must contain at least two members;

b) The members of a multiperson household need not be related to each other, while the members of a family must – by definition – be related.

5.139. Therefore, a household can contain more than one family, or one or more families together with one or more non-related persons, or it can consist entirely of non-related persons. A family typically will not comprise more than one household. However, factors like polygamy, shared child custody arrangements, and support arrangements provided by family members living in other households (for example, grandparents) can introduce complexities. One approach to deal with such complexities is, for the purpose of the census, to define a family as comprising only those related persons living in the same household. Thus, for example, a child being shared between two separated parents will only be included as a member of the family in the household where he/she is enumerated. However, individual countries should carefully consider their specific circumstances and determine the most appropriate way to derive and report data on families. This may involve alternative definitions or approaches that better reflect the unique family structures and dynamics within their populations.

5.140. It is recommended that the household be used as the unit of enumeration (as defined in paragraphs 2.41–2.47) and that the family be a derived topic only. The place of usual residence is

recommended as the basis for assigning persons to households where they usually reside. Where the *de facto* approach is used as the method of enumeration (see paragraphs 2.58–2.74), household lists should, where feasible, also include usual residents temporarily absent. In published reports, NSOs should indicate whether or not household information refers to usual residents and also what the time limits are in respect of being included or excluded as a usual resident. For a more detailed discussion on the difficulty of collecting information on place of usual residence, see paragraphs 5.51–5.56. In countries that produce household information based on register data, definitions can be made by setting limits on the minimum ages of the people residing (for example, in cases where three children under the age of 18 live together, further investigation may be needed to determine if they constitute a household or require alternative living arrangements).

1. Relationship to the reference person of household (core topic)

5.141. In identifying the members of a household (as defined in paragraph 5.134), it is useful to identify first the household reference person (HRP) and then the remaining members of the household according to their relationship to that person. Countries may use the term they deem most appropriate to identify this person (such as, for example, household reference person, head of household or householder, among others) if the person so identified is used solely to determine relationships between household members. It is recommended that NSOs present, in published reports, the concepts and definitions that are adopted. It is noted that when administrative data sources are used, deriving the relationship of household members to the reference person may present several challenges.

5.142. With respect to the selection of the HRP, NSOs should specify criteria for choosing that person from among the other household members, especially in polygamous, multi-family and other households, such as those composed only of siblings without a parent and those composed entirely of unrelated persons. This information should be included in training materials and instructions to enumerators.

5.143. The traditional concept of 'head of household' assumes that most households are family households and that one person in such family households has primary authority and responsibility for household affairs and is, in the majority of cases, its chief economic support. This person is then designated as the HRP.

5.144. The selection of one reference person in a household to whom all other persons in the household report, or designate, their relationship requires careful consideration. In the past the person considered to be the "head of the household" was generally used as the reference person, but this concept is no longer considered appropriate in many countries. It has also sometimes been proposed that the person designated as the reference person should be the oldest person in the household or the one who contributes the most income. However, given that the primary purpose of the question is to assign family status and to assign individuals into families, both of these approaches have weaknesses. The automatic selection of the oldest person may be undesirable because in multigenerational households many explicit kin relationships can be reported where the reference person is selected from the middle generation. Similarly, the selection of the person with the highest income may be a person who will not solicit the broadest range of explicit kin relationships.

5.145. Where more than one member of the household consider themselves to be equal in household authority and responsibility and may share the economic support of the household, the concept of a single head of household may no longer be considered valid even for family households. For the

relationship among members of the household to be determined under these circumstances, it is necessary that either:

(a) the members of the household designate one among them as a reference member with no implication of headship; or

(b) provision be made for designation of joint headship where desired (in such cases, only one of the joint heads—it doesn't matter which one—should be designated as the household reference person).

In any case, it is important that clear instructions be provided in the census as to how this situation is to be handled.

5.146. Even in countries where the traditional concept of head of household is still relevant, it is important to recognize that the procedures followed in applying the concept may distort the true picture of household composition, particularly with regard to female heads of households. The most common assumption that can distort the facts is that no woman can be the head of any household that also contains an adult male. Enumerators and even respondents may, in some countries, simply take such an assumption for granted.

5.147. This common gender-based stereotype often reflects circumstances that may have been true in the past but are no longer valid, insofar as the household and economic roles of women continue to change. It is therefore important that clear instructions be provided as to who is to be treated as the sole reference person of the household so as to avoid the complications of enumerator or respondent preconceptions on the subject and the bias that such preconceptions may create. The procedure to follow in identifying the HRP when the members of the household are unable to do so should be clear and unambiguous and should avoid gender-based bias. Where alternative definitions are used, this should be made explicitly clear in the census questionnaire and in the tabulated census results.

5.148. Given below is some guidance on the selection of the HRP, which will yield some explicit kin relationships:

(a) Either spouse of a married couple living in the household (preferably from the middle generation in a multigenerational household);

(b) Either partner of a consensual union couple living in the household where there is no married couple present (where applicable);

(c) The parent, where a lone parent lives with his or her children of any age; or

(d) Where none of the above conditions apply, any adult member of the household may be selected by the other household members.

Note that these categories are neither comprehensive nor mutually exclusive.

5.149. After identification of the HRP, the relationship of each of the remaining members of the household to the HRP should be recorded as one of the following:

(a) Spouse,
(b) Partner in consensual union (cohabiting partner), where applicable,
(c) Child,
(d) Spouse of child,
(e) Grandchild or great-grandchild,
(f) Parent (or parent of spouse),
(g) Other relative,
(h) Domestic worker, or
(i) Other person not related to the head or other household member.

Where this classification is considered too detailed for successful collection of the information, categories (f) and (g) may be consolidated as "Other relative" and (h) and (i) can be consolidated as "Other unrelated person".

5.150. As an aid to the identification of family nuclei (as defined in paragraphs 5.154–5.156) within the household, it might be helpful if persons were recorded on the census questionnaire to the extent possible in the order of nuclear relationship. Thus, the first person entered after the head or other reference person would be the spouse or cohabiting partner of that person, followed by unmarried children and then by married children, their spouses or cohabiting partners and children. For polygamous households, the order of entry could be such that each wife or cohabiting partner and her unmarried children appeared in succession.

5.151. For estimating fertility by the *own children* method (see paragraph 5.240), the natural mother of each child under 15 years of age should be identified if both are living in the same household. One way of doing this is to link the child's record with their mother's record in the census database, if both are living in the same household. The information is not relevant for stepchildren, adopted children or foster children under permanent or temporary care.

5.152. In order to meet increased data needs on households and families, countries may wish, while conducting their population censuses, to collect more detailed information on relationships. In households where the relationship structure is complex, including those with foster children, obtaining accurate information on the relationships *between* household members may be difficult. Some countries may supplement the information on relationship to the HRP with information on direct relationships between each household member thereby recording the relationship of a child to its parents even when neither parent is the HRP. Enumerators should, in such circumstances, be encouraged to probe for a clear relationship (such as child, niece or aunt). The recording of non-specific responses such as "relative" should be avoided. It is recommended that specific guidance be provided on acceptable responses, that relationships be specified completely in the census questionnaire, and that any precoded categories used should be sufficiently detailed to produce desired outputs.

2. Household and family composition (core topic)

5.153. Household and family composition can be examined from different points of view, but for census purposes it is recommended that the primary aspect considered should be that of the family nucleus.

5.154. A *family nucleus* is of one of the following types (each of which must consist of persons living in the same household[228]):

(a) A married or cohabiting couple without children;
(b) A married or cohabiting couple with one or more unmarried children;
(c) A lone father with one or more unmarried children;
(d) A lone mother with one or more unmarried children.

[228] In countries where a different definition of family nucleus is used, it should be clearly stated in the census report.

5.155. The concept of family nucleus as defined above limits relationships between children and adults to direct (first-degree) relationships, that is to say, between parents and children. In some countries, numbers of skip-generation households, that is to say, households consisting of grandparent(s) and one or more grandchildren with no parent of those grandchildren present, are considerable. Therefore, countries may include such skip-generation households in their family nucleus definition, and the census report should clearly state whether or not skip-generation households are included in the family nucleus definition. In some countries, the term "couple" includes both opposite-sex and same-sex couples.

5.156. The family nucleus is identified from the answers to the question on relationship to the HRP (topic 1 above), supplemented where necessary by information on name and marital status. The identification of offspring and their mother and the order in which persons are entered in the questionnaire may be of additional assistance in this respect. The identification of family nuclei is likely to be more complete in *de jure* than in *de facto* enumerations, because the latter do not take account of temporarily absent household members who may constitute part of a family nucleus.

5.157. For census purposes, a "child" is defined as any unmarried individual, regardless of age, who lives with his or her parent(s) and has no children in the same household. Consequently, the definition of a child is primarily a function of an individual's relationship to other household members, regardless of age. In accordance with this definition, a household consisting of a married or cohabiting couple with two never-married children, divorced son, and a married daughter and her husband would be considered to be composed of two family nuclei, with the divorced child being regarded as a member of the parents' family and the married daughter and son-in-law as a second family. As used here, the term "child" does not imply dependency, but rather is used to capture household living arrangements of persons who are in a parent-child relationship. Countries need to be clear in their metadata how they treat foster and adopted children.

5.158. The family nucleus does not include all family types, such as brothers or sisters living together without their offspring or parents, or an aunt living with a niece who has no child. It also excludes the case of a related person living with a family nucleus as defined above, for example, a widowed parent living with her married son and his family. The family nucleus approach does not, therefore, provide information on all types of families. Countries may extend the investigation of families beyond that of the family nucleus, in accordance with their own interests.

5.159. Households could be classified by type according to the number of family nuclei they contain and the relationship, if any, between the family nuclei and the other members of the household. The relationship should be through blood, adoption, or marriage, to whatever degree is considered pertinent by the country (see paragraph 5.152). Given the complexity of this topic, it is important that information on relationship to the household reference person be properly processed. The types of household to be recommended are:

- *(a)* *One-person household*;
- *(b)* *Nuclear household*, defined as a household consisting entirely of a single family nucleus. It may be classified into:
  - *(i)* Married couple family:
    - a. With child(ren);
    - b. Without child(ren);
  - *(ii)* Partners in consensual union (cohabiting couple family):

a. With child(ren);
b. Without child(ren);

(iii) Lone father with child(ren);
(iv) Lone mother with child(ren);

(c) *Extended household*, defined as a household containing one or more family nuclei and their relatives: either one or more family nuclei (related to each other), with other persons related to any of the family nuclei in the same household; or two or more persons related to each other, none of whom constitutes a family nucleus (see paragraph 4.146 in Revision 3 of the Principles and Recommendations). The subdivisions of this category may be modified to suit national circumstances.

(d) *Composite household*, defined as a household consisting of one or more family nuclei (related or unrelated to each other) plus other persons, some of whom are related to at least one of the nuclei and some of whom are not related to any of the nuclei; or two or more persons related to each other but none of whom constitute a family nucleus, plus other unrelated persons; (see paragraph 4.146 in Revision 3 of the Principles and Recommendations). The subdivisions of this category may be modified to suit national circumstances.

(e) Other;

(f) Unknown or not stated.

5.160. In the census tabulations, all countries should at least distinguish between one-person, nuclear, extended and composite households. Where feasible, some or all of the subcategories shown above should also be distinguished, although NSOs may find it appropriate to modify the classification according to national circumstances. For example, in countries where almost all households contain only one family nucleus at most, the distinction between nuclear, extended and composite households may be applied only to households containing one nucleus or no nucleus; multinuclear households may then be shown as an additional category without any further classification by type. In countries where multinuclear households are comparatively common, further breakdowns of extended and composite households, distinguishing between those with three, four or more family nuclei, may be helpful.

3. Household and family status

5.161. For purposes of determining household and family status and identifying how a person relates to other household or family members, persons may be classified according to their position in the household or family nucleus. Classifying persons according to household and family status has uses in social and demographic research and policy formulation. Census data could be presented according to both household and family status for a variety of purposes. Although status itself is based on information derived from responses to the question on relationship to the HRP and other items, the classification of persons by their household and family status is a different approach from the traditional one of classifying household members solely according to their relationship to the HRP. The following household and family status classifications illustrate how such an approach may be used.[229] Care should be taken at the planning stages to relate this classification to that of households by type as recommended in paragraph 5.159 above.

[229] To date, only the population and housing census recommendations for the Economic Commission for Europe region contain household and family status classifications.

Persons living in households are classified by household status as:

1 Person in a household with at least one family nucleus
- 1.1 Married spouse
- 1.2 Partner in consensual union (cohabiting partner)
- 1.3 Lone mother[230]
- 1.4 Lone father[231]
- 1.5 Child living with both parents
- 1.6 Child living with lone mother
- 1.7 Child living with lone father
- 1.8 Not a member of a family nucleus
    1. Living with relatives
    2. Living with non-relatives

2 Person in a household with no family nucleus
- 2.1 Living alone
- 2.2 Living with others[232]

1. Living with sibling(s) only
2. Living with other relatives only
3. Living with non-relatives only
4. Living with a combination of siblings, other relatives, and non-relatives

Persons are classified by family status as:[233]

1 Spouse
- 1.1 Husband
    - 1.1.1 With child(ren)
    - 1.1.2 Without a child
- 1.2 Wife
    - 1.2.1 With child(ren)
    - 1.2.2 Without child
- 1.3 Cohabiting partner (of either sex)
    - 1.2.1 With child(ren)
    - 1.2.2 Without child

2 Lone parent
- 2.1 Male
- 2.2 Female

3 Child
- 3.1 With both parents
- 3.2 With lone parent

[230] Person living with children, without spouse or partner.
[231] Person living with children, without spouse or partner.
[232] The subdivisions in this category should be modified to suit national circumstances.
[233] The subdivisions in this category should be modified to suit national circumstances.

3.2.1 With lone father
3.2.2 With lone mother
4 Not member of a family nucleus
4.1 Relative of husband or wife
4.1.1 Parent of husband or wife
4.1.2 Sibling of husband or wife
4.1.3 Other relative of husband or wife
4.2 Non-relative

D. Demographic and social characteristics

5.162. Sex and age are considered to be the most basic of all demographic variables. Of all the topics included in population censuses, *sex* and *age* are more frequently cross-classified with other characteristics of the population than are any other variables. Apart from the importance of the age-sex structure of the population in itself, accurate information on these two topics is fundamental to the great majority of the census tabulations.

5.163. A very important use of census data on the sex and age composition of the population is the evaluation of the data, especially with respect to coverage. The variables are therefore vital, and it is important that this information be reported in respect of every person for whom census information has been collected.

5.164. It is therefore recommended that where this information is incomplete it should be imputed for census purposes rather than being reported as "not stated". Possible difficulties in securing accurate age data are often not recognized because the topic appears to be a simple one. The difficulties associated with this topic are therefore highlighted in paragraphs 5.169–5.180 below.

1. Sex (core topic)

5.165. The sex of every individual should be recorded on the census questionnaire for those countries that collect their census information in this way. The disaggregation of data by sex is a fundamental requirement for gender statistics. For many socioeconomic and demographic characteristics that are collected through a census, such as education, economic activity, marital status, migration, disability and living arrangements, there are generally variations by sex. The successful planning and implementation of gender-sensitive policies and programmes requires the disaggregation of data by sex to reflect problems, issues and questions related to both men and women in society.

5.166. Sex, together with age, represents the most basic type of demographic information collected about individuals in censuses and surveys, as well as through administrative recording systems, and the cross-classification of these data with other characteristics forms the basis of most analyses of the social and demographic characteristics of the population, as it provides the context within which all other information is placed.

5.167. It should be emphasized here that sex and gender are different concepts, although the two are often conflated or used interchangeability. Sex is a biological attribute whereby persons are traditionally classified as either male or female, while gender is a social construct that refers to the roles, behaviours, activities, and attributes that a particular society associates with being a man or a woman.

5.168. To reflect social and cultural changes and, in some cases, to meet legislative requirements, some countries are now collecting data on gender identity in addition to information about the sex of individuals. This allows respondents to express their identity beyond traditional binary options. Including questions on gender identity in a census requires careful consideration and testing, as there are no international standards for measuring this fluid and evolving concept.

2. Age (core topic)

5.169. *Age* is the interval of time between the date of birth and the date of the census, expressed in completed solar years. Every effort should be made to ascertain the precise and accurate age of each person, particularly of children and older persons.

5.170. Information on age may be secured either by obtaining the date (year, month and day) of birth or by asking directly for age at the person's last birthday.

5.171. The first method yields more precise information and should be used whenever circumstances permit. It also allows for the calculation of age at a reference date other than census day for the purposes, for example, of deriving annual census-based mid- or end-year population estimates. If neither the exact date nor the month of birth is known, an indication of the season of the year can be substituted if this information can be easily recorded. The question on date of birth is appropriate wherever this is known, whether in accordance with the solar calendar or a lunar calendar, or whether years are numbered or identified in traditional folk culture by names within a regular cycle. It is extremely important, however, that there should be a clear understanding between the enumerator and the respondent about which calendar system the date of birth is based on. If there is a possibility that some respondents will reply with reference to a calendar system different from that of other respondents, provision must be made in the questionnaire for noting the calendar system that has been used. It is not advisable for the enumerator to attempt to convert the date from one system to another. The necessary conversion can be best carried out as part of the data-editing work.

5.172. Where the information is taken from administrative data sources, date of birth is usually more accurately recorded.

5.173. The direct question on age is likely to yield less accurate responses for a number of reasons. Even if all responses are based on the same method of reckoning age, there is the possibility of a misunderstanding on the part of the respondent as to whether the age wanted is that at the last birthday, the next birthday or the nearest birthday. In addition, asking a direct question on age can result in occurrences, with comparative ease, of rounding to the nearest age ending in zero or five (a phenomenon referred to as 'age heaping'), providing estimates not identified as such and deliberate misstatements. Difficulties may arise in the reporting, or in the recording, of the information for children under 1 year of age, which may be given erroneously as "1 year of age" rather than "zero years of age". These difficulties may be mitigated by collecting information on the date of birth of all children reported as "1 year of age", while using only the direct age question for the rest of the population. Another possible approach is to obtain age in completed months for children under 1 year of age. This method, however, can give rise to another type of recording error, that is to say, the substitution of years for months, so that a 3-month-old child, for example, might be reported as being 3 years of age.

5.174. Some countries have made improvements in the quality of age data by asking both questions on age and date of birth, although this imposes an additional burden on respondents.

5.175. An additional complication may occur with the use of the direct question if more than one method of calculating age is in use in the country. In some countries, certain segments of the population may use an old traditional method whereby persons are considered to be 1 year of age at the time of birth and everyone advances 1 year in age at the same fixed date each year. Other segments of the population in the same countries may use the Western method, in which a person is not regarded as being 1 year of age until 12 months after the date of birth, and advances 1 year in age every succeeding 12 months. If there is a risk of different methods of age calculation being used by respondents, operational rules should be developed to address such issues based on national contexts.

5.176. In spite of its drawbacks, the direct question on age is the only one that should be used when people cannot provide even a birth year. As regards persons for whom information on age is unavailable or appears to be unreliable, an estimated age may have to be recorded. This may occur for example, where responses are to be given by proxy, or in general in cultures where there is little awareness of individual age and no interest in it. In the latter circumstances, criteria for making estimates should be provided in the instructions for the enumerators.

5.177. One of the techniques that have been used to aid enumerators is provision of calendars of historic events of national or local significance to be used either in probing questions or in identifying the earliest event that the respondent recalls. Another technique consists of pre-identifying locally recognized age cohorts in the population and then asking about membership in the cohorts. Enumerators may also ask if the person in question was born before or after other persons whose ages have been roughly determined. Furthermore, use can be made of age norms for weaning, talking, marriage and so forth. Whatever techniques are used, enumerators should be impressed with the importance of securing age data that are as accurate as possible within the amount of time that they can devote to the topic.

5.178. In view of the possible difficulties in the collection of age data, census tests should be used, as appropriate, to determine the difference in results with the use of a question on age as compared with a question on date of birth, what calendar or method of age reckoning most people use, and in what parts of the country age will have to be estimated for the majority of the population and what techniques to use as an aid in such estimation. Testing of the calendar or method of age reckoning that most people use is particularly important where an official change from one calendar or method of reckoning to another calendar or method has taken place recently enough so that the new calendar or method of reckoning may not yet be in common usage among some or all of the population.

5.179. Enumerators who are likely to be called upon to estimate age in a substantial number of cases should be given training in the applicable techniques as part of their general training.

5.180. As noted in paragraph 5.164 it is recommended that where this information is incomplete it should be derived or imputed for census purposes rather than being reported as “not stated”.

3. Marital status (core topic)

5.181. Despite the changing nature of marriage, marital status remains a useful demographic variable. The direct relationship between marriage and fertility is still recognized, as is the indirect relationship with other demographic, social and economic characteristics. Numerous variations exist in many

countries, but it is important that marriage be defined in terms of the laws and customs of individual countries.

5.182. *Marital status* is the personal status of each individual in relation to the marriage laws or customs of the country. The categories of marital status to be identified should at least include the following:

(a) Never married;
(b) Married;
(c) Separated (living apart from a spouse while still married);
(d) Divorced and not remarried'
(e) Widowed and not remarried.

5.183. In some countries, category (b) may require a subcategory of persons who are contractually married but not yet living as spouses. In all countries, category (c) should comprise both the legally and the *de facto* separated, who may be shown as separate subcategories if desired.

5.184. In some countries, it will be necessary to take into account customary unions, such as registered partnerships and consensual unions, which are legal and binding under law. In countries with legal provision for registered or legal partnership (for opposite-sex couples or same-sex couples), or where same-sex couples can legally marry, subcategories may either be included in the category (b) Married or in a legally registered partnership, namely (b)(i) "Opposite-sex marriage/partnership", (b)(ii) "Same-sex marriage/partnership".

5.185. The treatment of persons whose only or latest marriage has been annulled is dependent upon the relative size of this group in the country. Where its size is substantial, the group could constitute an additional category; if its size is insignificant, however, the individuals in the group should be classified according to their marital status before the (annulled) marriage took place.

5.186. At times countries have experienced difficulties in distinguishing between (i) formal marriages and *de facto* unions (see paragraph 5.189 below) and (ii) persons legally separated and those legally divorced. Information on de facto unions is very useful in studies of fertility. It is suggested that countries wishing to investigate these relationships should consider collecting separate data for each person on de facto unions and on the duration of each type of union (see paragraphs 5.273–5.274). If either of these circumstances necessitates a departure from the recommended classification of marital status, the composition of each category shown in the tabulations should be clearly stated.

5.187. If complete information on marital status is needed, then this information should be collected and tabulated for persons of all ages, irrespective of the national minimum legal age, or the customary age for marriage, because the population may include persons who were married in another country with a different minimum marriage age. In most countries, there are also likely to be persons who were permitted to marry below the legal minimum age because of special circumstances. In order to permit international comparisons of data on marital status, however, any tabulations of marital status not cross-classified by detailed age should at least distinguish between persons under 15 years of age and those 15 years of age and over. If circumstances necessitate a departure from the recommended classification of marital status, the composition of each category shown in the tabulations should be clearly stated.

5.188. The collection of additional information related to customs in particular countries (such as concubinage, polygamous or polyandrous marital status, or inheritance of widows) may be useful in meeting national needs. For example, at times countries may wish to collect data on the number of spouses of each married person. Modifications of the tabulations to take account of such information should be made within the framework of the basic classification in order to maintain international comparability as far as possible.

5.189. The concept of marital status and the marital status categories described above should not be confused with the concept of de facto union status, which describes extra-legal unions (including some consensual unions) of varying degrees of stability common in some countries. Information on these relationships may also be derived from information collected on the relationship to HRP or other persons in the household, in order to distinguish between people who are living in either a consensual union or marriage, and those who are not.

4. Ethno-cultural characteristics

5.190. Countries with a culturally diverse population may wish to collect information on the ethnic identity (or composition) of the population, on mother tongue, on the knowledge and practice of languages and on religious communities and denominations. They are all characteristics that allow people the flexibility to express their ethnocultural identity in the way that they choose. Data on such characteristics of the population are of increasing relevance to countries in the context of migration, integration and minority policies.

5.191. Ethnocultural characteristics generally have a subjective dimension. Moreover, different countries will adopt different concepts. Such characteristics can also be politically sensitive and may apply to very small, yet identifiable, population subgroups. The free and open declaration of the respondents is therefore of essential importance. Members of certain minority groups may be particularly vulnerable to discrimination on the grounds of ethnic group or religion. Special care, therefore, may be required in census procedures and outputs relating to ethnic group and religion in order to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. In some cases, countries may even wish to collect such data on a voluntary basis if this is permitted by national legislation.

5. Religion

5.192. Each country that investigates religion in its census should use the definition most appropriate to its needs and should include the definition that has been used as part of the metadata in the census publications and dissemination programme.

5.193. For census purposes, *religion* may be defined as either:

(a) Religious or spiritual belief of preference, regardless of whether or not this belief is represented by an organized group; or

(b) Affiliation with an organized group having specific religious or spiritual tenets.

5.194. The decision whether to collect and disseminate information on religion in a national census is dependent upon a number of considerations and national circumstances, including, for example, the strength of the user requirement for such data, and the suitability and sensitivity of asking a religion question in a country's census. Owing to the sensitive nature of the topic, special care may be required

to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. It is important that the responding public be informed of the potential uses and needs for this information.

5.195. The level of detail collected on this topic is (as with any other topic) dependent upon the users' requirements. It may, for example, be sufficient to enquire only about the religion of each person; on the other hand, respondents may be asked to specify, if relevant, the particular denominations to which they adhere within a religion. In countries where a large number of sects or denominations exist there will be implications for space on any census questionnaire and implications for data capture, especially in cases where "write-in" responses are offered. In an effort to ensure international comparability as far as possible, it is recommended that religion or religious affiliation should be measured directly by a question that asks "What is your religion?" rather than use of a filter question that asks for example "Are you religious?" and if so "What is your religion?" Response categories should include "No religion/religious affiliation" together with a "Religious but prefer not to answer" or "Not stated" category, in effect making responses to such a question voluntary.

5.196. For the benefit of users of the data who may not be familiar with all of the religions or denominations within the country, as well as for purposes of international comparability, the classifications of the data should show each denomination as a subcategory of the religion of which it forms a part. A brief statement of the tenets of religions or denominations that are not likely to be known beyond the country or region would also be helpful.

6. Language

5.197. There are four types of language data that can be collected in a census, namely:

(a) Mother tongue, defined as the language usually spoken in the individual's home in his or her early childhood;

(b) Main language, defined as the language that the person commands best;

(c) Usual language, defined as the language currently spoken, or most often spoken, by the individual in his or her present home;

(d) Ability to speak one or more designated languages, including the country's official language(s) and/or indigenous languages.

5.198. Each of these types of information serves a very different analytical purpose. Each country should decide which, if any, of these types of information is applicable to its own needs. International comparability of tabulations is not a major factor in determining the form of the data to be collected on this topic.

5.199. In compiling data on the usual language or on the mother tongue, it is desirable to show each language that is numerically important in the country and not merely the dominant language.

5.200. Information on language (including any sign language) should be collected for all persons. In the tabulated results, the criterion for determining language for children not yet able to speak should be clearly indicated.

7. Ethnicity

5.201. The decision whether to collect and disseminate information on ethnic or national groups of a population in a census is (as is the case with religion – see paragraph 5.194) dependent upon a number of considerations and national circumstances, including, for example, the strength of the users' requirement for such data, and the suitability and sensitivity of asking ethnicity questions in a country's census. Owing to the sensitive nature of questions on ethnicity, special care may be required to demonstrate to respondents that appropriate data protection and disclosure control measures are in place. It is important that the responding public be informed of the potential uses and need for data pertaining to ethnicity, as this improves public support for the census exercise. Data on ethnicity provide information on the diversity of a population and can serve to identify subgroups of a population. Some areas of study that rely on such data include demographic trends, employment practices and opportunities, income distribution, educational levels, migration patterns and trends, family composition and structure, social support networks, and health conditions of a population.

5.202. Broadly defined, *ethnicity* is based on a shared understanding of history and territorial origins (regional and national) of an ethnic group or community, as well as on particular cultural characteristics such as language or religion.

5.203. Respondents' understanding or views about ethnicity, awareness of their family background, the number of generations they have spent in a country, and the length of time since immigration are all possible factors affecting the reporting of ethnicity in a census. Ethnicity is multi-dimensional and is more a process than a static concept, and so ethnic classifications should be treated with movable boundaries. This may mean that classifications of ethnic categories will change between censuses which, while mirroring society at any one time, may lead to a degree of non-comparability between one census and another.

5.204. Ethnicity can be measured using a variety of concepts, including ethnic ancestry or origin, ethnic identity, cultural origins, nationality, race, colour, minority status, tribe, language, religion or various combinations of these concepts. Because of the interpretative difficulties that may occur with measuring ethnicity in a census, it is important that, where such an investigation is undertaken, the basic criteria used to measure the concept are clearly explained both to respondents during data collection and to users in the dissemination of the resulting data. The method and the format of the question used to measure ethnicity can influence the choices that respondents make regarding their ethnic backgrounds and current ethnic identification.

5.205. The subjective nature of the concept of ethnicity (not to mention increasing intermarriage among various groups in some countries) requires that information on ethnicity be collected through self-declaration of a respondent and also that respondents have the option of indicating multiple ethnic affiliations. Data on ethnicity should not be derived from information on country of citizenship or country of birth. The classification of ethnic groups also requires the inclusion of the finest levels of ethnic groups, self-perceived groups, regional and local groups, as well as groups that may not be considered as ethnic groups at all, such as religious groups and those based on nationality. Countries collecting data on ethnicity should note that the precoding or pre-classification of ethnic groups at the time of data capture may have a tendency to lose detailed information on the diversity of a population unless space to record write-in, free-form responses is provided. If implemented, the precoding or classification of ethnic groups should be based on field research and self-declared categories in previous census data collection.

5.206. Respondents should be free to indicate more than one ethnic affiliation or a combination of ethnic affiliations if they wish to do so. For such cases, separate questions for the main ethnicity and secondary ethnicities should be included. Countries should explain in the census instructions and the census documentation how the ethnicity of children from mixed couples is to be reported (for example, explicit instructions to allow respondents to provide multiple responses or to allow for responses such as “Bi-racial”). Also, to guarantee the free self-declaration of ethnicity, respondents should be offered response option such as “Not declared” or “Prefer not to answer” in effect making responses to such a question voluntary.

5.207. Because the ethnocultural composition of a country can vary widely from country to country and due to the diversity in the approach and the various criteria for establishing ethnicity, it is recognized that there is no single definition or classification that could be recommended that would be applicable to all countries. However, countries should document the basic criteria and classification procedures for ethnicity and inform the data users about the concepts on which they are based.

8. Indigenous peoples

5.208. Facilitating the collection of data on indigenous peoples for national and international needs can serve to improve socioeconomic and active participation of such people in the development process for many countries. As with ethnicity, the potentially sensitive nature of questions pertaining to the indigenous population requires care in assuring the public that the appropriate disclosure and data protection methods are being enforced. The responding public should be informed on the potential uses and need for such data to improve public support for the census exercise.

5.209. Dissemination of census data pertaining to indigenous peoples contributes to research in areas such as the socioeconomic conditions of the indigenous population, trends, causes of inequities, and the effectiveness of existing policies and programmes. Availability of these data can also assist indigenous peoples in assessing their living conditions and give them the information they need to participate and advocate in the development of programmes and policies affecting their communities, such as those impacting health systems, models of economic production, environmental management and social organization. In addition, the development of indicators relevant to the indigenous population and the measurement of such indicators in the data collection process can be used to monitor the socioeconomic development of indigenous peoples.

5.210. Generally, *indigenous peoples* of a particular country are social groups with an identity that is distinct from the social and cultural identity of the dominant society in that country. Questions on indigenous identity should abide by the principle of self-identification.

5.211. It is important that, where such an investigation is undertaken, multiple criteria are developed to accurately capture identity and socioeconomic conditions of indigenous peoples. Defining such a population can be done in many ways, such as through a question on ethnic origin (that is to say, ancestry) or on indigenous identity. Identifying the indigenous community also requires recognition of the diversity in this subpopulation, including nomadic, semi-nomadic and migrating peoples, peoples in transition, displaced persons, indigenous peoples in urban areas, and particularly vulnerable subgroups. It is important to point out that there is no single term among countries to describe the indigenous population. Consequently, countries tend to use their own national concepts to identify the indigenous population.

5.212. Differing national contexts also imply that enumerating the indigenous population can be done in multiple ways, for example, by way of specific questions on the census form, with specialized questionnaires for the indigenous population, or with follow-up or complementary surveys. In Canada, for example, identification of the indigenous population comes not only from its national census, but also from a post-censal survey. In Australia, in addition to the national census, there is the National Aboriginal and Torres Strait Islander survey, while in Argentina there is a complementary survey after the census targeting indigenous peoples. In addition to a general census, Paraguay also administers a specific census in the same year to identify the indigenous population.

5.213. Involvement of the indigenous peoples themselves in the data development and data collection processes provides the arena for capacity-building and helps to ensure the relevance and accuracy of the data. Using local indigenous languages, employing local indigenous people (as interpreters, for example), and training and building the capacity of local indigenous people in data collection processes can facilitate the collection and dissemination of this information.

9. Disability status (core topic)

5.214. A census can provide valuable information on disability in a country. For countries that do not have regular special population-based disability surveys or disability modules in ongoing surveys, the census may be the only source of information on the prevalence and distribution of disability and functioning in the population at national, regional and local levels. Countries that have a health registration system providing regular data on persons with the most severe types of impairments may use the census to complement these data with information related to selected aspects of the broader concept of disability and functioning based on the International Classification of Functioning, Disability and Health (ICF).[234] Census data can be utilized for:

[234] Adopted in 2001, the ICF is the international standard for describing and measuring health and disability at both the individual and population levels. More information on the ICF framework is available at: http://www.who.int/classifications/icf/en/.

- general planning programmes and services (prevention and rehabilitation);
- monitoring selected aspects of disability trends in the country;
- the evaluation of national programmes and services concerning the equalization of opportunities; and
- international comparison of selected aspects of disability prevalence in countries, including in the context of internationally-agreed monitoring frameworks such as the Sustainable Development Goals (SDGs) or their successors.

Census data on disability can also be used by countries to produce and disseminate the statistics required in order to comply with the Convention on the Rights of Persons with Disabilities.

a. Disability characteristics

5.215. *Disability status* characterizes the population into those with and those without a disability. Persons with disabilities are defined as those persons who are at greater risk than the general population for experiencing restrictions in performing specific tasks or participating in activities. This group would include persons who experience limitations in basic activity functioning, such as walking or hearing, even if such limitations were ameliorated by the use of assistive devices, a supportive environment or plentiful resources. Such persons may not experience limitations in specifically measured tasks, such as going to working or shopping, because the necessary adaptations have been made at the personal or environmental levels. These persons would still, however, be considered to be at greater risk of restrictions in activities or participation than the general population because of the presence of limitations in basic activity functioning, and because the absence of necessary accommodations would jeopardize their current levels of participation.

5.216. A comprehensive measure to determine disability would include the following six domains of functioning in a way that can be reasonably measured using a census and that would be appropriate for international comparison:

(a) Walking
(b) Seeing
(c) Hearing
(d) Cognition
(e) Self-care
(f) Communication.

5.217. The first four domains (a) to (d) are to be considered essential in determining disability status. The additional domains (e) and (f) comprise a more comprehensive measure for determining disability. Not including these items will affect the population with disabilities identified.

b. Use of the census to measure disability at the aggregate level

5.218. A census format offers only limited space and time for questions on any one topic such as disability. Since ICF offers several dimensions for use to develop a census measure, it is best to focus on a few of those dimensions, leaving the remaining dimensions for use in more extensive household-based disability surveys. Short sets of disability questions, which can be included in censuses and extended sets to be recommended for inclusion in surveys, have been developed and tested by the Washington Group

on Disability Statistics.[235] The aim of the recommended sets is to improve comparability of disability and functioning data across countries.

5.219. Given the complexity of disability definition and measurement and, in certain cultures, the sensitivity attached to people identifying as having a disability, it is recommended that several functional activity domains be defined whereby people can respond to questions about their difficulty in performing those activities rather than enquiring directly whether or not they have a particular disability or "any" disability.

5.220. Despite being widely included in population censuses, obtaining accurate and objective responses to inquiries about disability can be challenging. When responses are typically provided by proxy household members and enumerators lack specialized training, the quality of census data on disability may lack objectivity and accuracy compared to that from specialized surveys. International experiences suggest that that censuses may under-estimate true levels of disability, particularly among young children.

(i) **Essential domains**

5.221. It is suggested that the domains to be included in a short set of questions recommended for use in censuses should satisfy a set of selection criteria that include-cross-cultural comparability, suitability for self-reporting, and space on the census form. Other suggested criteria include the importance of the domain in terms of public health problems. Based on these criteria, the Washington Group on Disability Statistics has developed a Short Set on Functioning (WG SS-F) questions in the six domains for the purpose of measuring disability within a census format. The four basic domains - walking, seeing, hearing and cognition - are (as noted at paragraph 5.217) considered to be the most essential. In addition, if space permits, two other domains - self-care and communication - have been identified for possible inclusion. However, to obtain a more comprehensive assessment, it is strongly encouraged that the two other functioning domains, self-care and communication, be included. Deviating from collection of data using the standard set of domains will impact comparability when making international comparisons.

5.222. *Walking* fulfils the criteria of cross-cultural applicability for comparable data and space requirements since walking is a good indicator of a central physical function and is a major cause of limitation in participation.

5.223. While *seeing* also represents a public health problem, the collection of the relevant information on this domain through self-reporting can be problematic if the subject uses glasses to correct visual impairments. Similar difficulties are associated with including a question on *hearing* activity.

[235] The Washington Group on Disability Statistics, a United Nations City Group that focuses on proposing international measures of disability, has developed these questions. For updates on the question wording and more information supporting the collection and use of data on disability, see www.cdc.gov/nchs/washington_group.htm.

5.224. Assistive devices, such as glasses, and to a lesser extent, hearing aids, provide almost complete accommodation for a large proportion of those with impaired functioning. It is often argued that asking about seeing without the use of glasses greatly increases the number of persons with disabilities and makes the group too heterogeneous, that is, the group would include persons at very little risk of participation problems along with those at greater risk. As a result, questions on difficulty in seeing or hearing should be asked with the use of glasses or hearing aids if they are usually worn.

5.225. Of the four essential domains, *cognition* is the most difficult to operationalize due to its complexity. Cognition includes many functions such as remembering, concentrating, decision-making, understanding spoken and written language, finding one's way or following a map, doing mathematical calculations, reading and thinking. Deciding on a cross-culturally similar function that would represent even one aspect of cognition is difficult. However, remembering, concentrating and making decisions are related to the other aspects of cognition and meet the criteria of cultural compatibility.

(ii) **Additional domains**

5.226. In additional to the four essential domains, two other have been identified for possible inclusion: self-care and communication. The *self-care* domain is intended to identify persons who have some problems with taking care of themselves independently. Washing and dressing represent self-care tasks that occur on a daily basis and are considered to be basic activities.

5.227. The purpose of the *communication* domain is to identify persons who have some problems with talking, listening or understanding speech such that it contributes to difficulty in doing their daily activities. Two aspects of communication should be considered: understanding others (receptive communication) and being understood by others (expressive communication). Communicating (understanding and being understood) refers to the exchange of information or ideas between two people through the use of language. Communication is a key domain of functioning and should be included to better identify the population and produce more comprehensive statistics.

5.228. Beyond the six domains identified above, there are other functioning domains that could be included in a set of census questions depending on the space available. The Washington Group Extended Set on Functioning (WG ES-F) includes questions on domains that address functioning of the upper body (the arms, hands and fingers), psychological functioning (such as anxiety and depression), and pain and fatigue.

5.229. A set of questions on child functioning has been developed through a collaboration between the WG and UNICEF.[236] The Child Functioning Module (CFM) comprises two questionnaires, one with 16 questions for children aged 2 to 4 years and another with 24 questions for children aged 5 to 17 years. The questions are to be administered to the mother or primary caregiver of the child in question and assesses difficulties in various domains of functioning. In addition to collecting data on domains related to walking, hearing, seeing and cognition, the CFM includes questions on difficulties in self-care, communication, learning, interacting with peers, coping with change, focusing attention and psychosocial

[236] Loeb, Mitchell, et al., 'The Development and Testing of a Module on Child Functioning for Identifying Children with Disabilities on Surveys. I: Background', Journal of Disability and Health, vol. 11, no. 4, 2018, pp. 495–501.

functioning. Using a shortened version of the CFM or using other instruments like the WG-SS that collect data based on a restricted number of functional domains is not recommended to collect data on children. Such choices will not identify all children with all types of difficulties in basic activities and would result in a severe underestimation of the proportion of children with disabilities. A more accurate measure of disability among children requires the full range of functional domains included in the CFM and administration of the questions to a child's mother (or primary caregiver). In censuses, the child's mother or primary caregiver is unlikely to be the respondent to the questionnaire. For these reasons, the CFM is best suited for use in surveys.

c. Census question wording

5.230. It is recommended that special attention be paid in designing census questions to measure disability. The wording and design of the questions greatly affect the precision in identifying persons with disabilities. Each domain should be asked through a separate question.[237] The language used should be clear, unambiguous and simple. Negative terms should always be avoided. The disability questions should be addressed to each single household member aged five years and above and general questions on the presence of persons with disabilities in the household should be avoided. If necessary, a proxy respondent can be used to report for the family member who cannot respond for him/herself. The important thing is to account for each family member individually rather than ask a blanket question. Scaled response categories can also improve the reporting of disability. The census questions on disability endorsed by the Washington Group include four response categories:

(a) No difficulty;
(b) Yes – some difficulty;
(c) Yes – a lot of difficulty;
(d) Cannot do [the activity] at all.

Disability prevalence is determined based on any response that is "a lot of difficulty" or "cannot do at all" for any of the domains.

5.231. The information that results from measuring disability status as defined here is expected to:

(a) Represent a large proportion of, but not all, persons with limitation in basic activity functioning in any one country (only the use of a wider set of domains would potentially cover all such persons, but as noted earlier this would not be practicable in a census context);
(b) Represent the most commonly occurring basic activity limitations within any country;
(c) Collect information that is internationally comparable.

5.232. The questions identify the population with limitations in basic activities that have the potential to restrict independent participation in society. The intended use of these data would be to compare levels of participation in employment, education, or family life for persons with a disability with those without such a disability to determine the extent that persons with disabilities are socially included. In addition, the data could be used to monitor prevalence trends for persons with limitations in the particular basic activity domains selected.

[237] When domains are combined, such as asking a question about seeing or hearing, respondents frequently are confused and think they need to have difficulty in both domains in order to answer yes. In addition, having the numbers with specific limitations is useful for both internal planning and for cross-national comparisons.

5.233. Because disability is a complex concept, it is necessary to adopt an explicit definition based on the ICF domains used when developing census (or survey) questions that will be used to identify disability status. The recommended set of questions for censuses is based on such an explicit definition (as described above). It is essential that estimates or tabulations based on the recommended set be accompanied by information on how disability is defined and how the questions are asked. This information should be included as part of the metadata associated with the questions and data set and it should be included as a footnote to tables that include these estimates.

d. Use of the census to screen for disability and follow-up with other surveys

5.234. Countries that are planning specialized surveys on disability may want to use the census to develop a sampling frame for these surveys and include a screening instrument to identify persons who will be interviewed subsequently. The main purpose of a screening instrument is to be as inclusive as possible in order to identify the largest group of people who could be further studied. The screening question should be designed so that false negatives[238] are minimized, while false positives[239] should be less of a concern.

5.235. Within the framework of the ICF, the census screening may include all of the three main dimensions of body structure and function, activity and participation. This will allow for keeping a broad approach to the follow-up survey where the different aspects of disability can be better studied. The same recommendations highlighted in paragraphs 5.230–5.233 should also be considered when a screening module is designed.

5.236. Before embarking on using the census to develop a frame for a follow-up survey, it is important that the legal implications of using the census data for this purpose are fully considered. Respondents to the census should be informed that the data may be used for follow-up studies, and the national authorities responsible for ensuring the privacy rights of the population may need to be consulted in order to obtain their approval.

E. Fertility and mortality

5.237. The investigation of fertility and mortality in population censuses is particularly important in countries lacking a timely and reliable system of civil registration and vital statistics; a census can offer the opportunity to provide data for estimating vital rates that would not otherwise be available. Even in countries with complete birth and death registration, some of the topics (such as “children born alive”, “children living”, “age at marriage or union” and “age at first birth”) are equally appropriate because they provide data that are not easily available from registration data but which are necessary for the computation of cohort and period fertility tables. The census provides an opportunity to collect data for estimating fertility and mortality at national and subnational levels in a cost-effective manner. Additionally, the investigation of these topics offers an opportunity to estimate differentials in fertility

[238] Persons who have disabilities but are not identified in the census as having disabilities.

[239] Persons who are identified with disabilities in the census but in reality do not have disabilities (as assessed in the largest instrument used in the follow-up survey).

and mortality considering other socioeconomic characteristics investigated in the censuses. The inclusion of these topics in population censuses for the purpose of estimating fertility and mortality rates and other related indicators is both prudent and cost-effective, particularly in countries where civil registration and vital statistics systems are not fully developed and costs of conducting large periodic demographic surveys are high. Nevertheless, it is important to note that census information is a poor substitute for complete and reliable vital registration data. If countries desire accurate and detailed estimates of fertility and mortality, they must establish, and need to maintain, comprehensive and universal civil registration systems.

5.238. To obtain data on fertility, as a minimum requirement, information should be collected on the core topics "children ever born", and "date of last child born alive". In addition questions on "age of mother at birth of first child born alive" and "age, date or duration of marriage or union" may improve fertility estimates based on children ever born. For the collection of reliable data, some of the topics may require a series of probing questions that, because they are time consuming, are more suitable for use in sample surveys than in censuses.

5.239. The target population for which data should be collected for each of the topics included in this section consists of women 15 years of age and over regardless of marital status. However, it may be appropriate in some countries to reduce the lower age limit by a few years. It is acknowledged that in some countries, certain cultural sensitivities (for example, regarding the collection of information on childbearing from never-married women) exist to collecting information from all women aged 15 years of age and over without regard to marital status. In such cases, every effort should be made to collect the requisite information. In countries that do not use the data for women 50 years of age and over, it may be appropriate to limit data collection to women under that age, thereby allowing more concentrated effort on data collection for the requisite age group.

5.240. In addition to the characteristics indicated above, another useful variable that allows the estimation of fertility is the "own children" method[240] and birth history reconstruction.[241] The application of these methods requires the identification of the natural mother of each child where both are members of the same household. In cases where it is difficult to ascertain the identity of the natural mother, one may use as a proxy the relationship to the HRP (see paragraphs 5.141–5.152) or the number of children living (see paragraphs 5.260–5.262) to establish the identity of the natural mother. In essence, information on the child's age and the mother's age are used to estimate a series of annual

[240] For methodological details, see United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods. Series F, No.92, Chapter V, section B (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); United Nations (1983). *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm), Chapter VIII, section C; and Cho, L.-J., Retherford, R. D., & Choe, M. K. (1986). *The own-children method of fertility estimation*. Honolulu, HI: Population Institute.

[241] For methodological details, see United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods. Series F, No.92, Chapter V, section C (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); and Luther, N.Y., Cho, L-J. (1988). Reconstruction of birth histories from census and household survey data. *Population Studies*, 42: 451–472.

fertility rates for years prior to the census. The reliability of the estimates produced depends, among other things, on the proportion of mothers enumerated in the same questionnaire as their own children, the accuracy of age reporting for both mothers and their children and the accuracy of available estimates of mortality for women and children.

5.241. In the case of subnational fertility estimates, it's important to consider the impact of migration patterns. When mothers leave their children with family members to work elsewhere, it can lead to an underestimation of fertility rates in their area of origin. On the other hand, accurate and timely birth registration, including for children born to migrant women, can improve the accuracy of these estimates. Therefore, understanding both migration patterns and the quality of birth registration is crucial for obtaining reliable subnational fertility estimates.

5.242. A more parsimonious estimation method of fertility, the reverse survival method, allows estimating fertility from data collected in a census even if no questions have been asked about fertility directly.[242] Based on population data by single age and sex, the method consists in 'reverse surviving' those no longer present in the population of a given age in order to derive the number of births that occurred in the past, usually for a period of 15 years preceding the data collection. The consistency of the fertility estimates obtained from the reverse survival method depends on the errors in the age distribution (e.g., omission of infants and young children and heaping on specific ages).

5.243. Mortality topics include infant and child mortality, obtained from data on "children ever born alive" and "children living", and adult mortality, obtained from "household deaths in the past 12 months" and "maternal or paternal orphanhood". The extent to which mortality (particularly infant and child) can be adequately measured from population census data is largely dependent on the quality of the training of field staff to minimize non-response to the relevant questions on and the reporting of erroneous information. Enumerator manuals should include the measures that are needed to minimize such errors. Accurate responses to these questions are often difficult to obtain, thus resulting in imperfect data. Nevertheless, it is often possible to derive usable adjusted estimates from this information.

5.244. As far as possible, efforts should be made to obtain information on fertility, child mortality (or survival) and marriage or union directly from the woman or mother involved, because she is more likely to recall correctly the details of her fertility, the mortality of her offspring and her marital experiences than any other member of the household. Information on household deaths, by date, sex and age, in the 12-month period prior to the census should be collected from the HRP. Information on maternal orphanhood and paternal orphanhood should be collected for each person in the household regardless of age.

[242] For methodological details, see Moultrie, T.A., Dorrington, R.E. , Hill, A.G., Hill, K.H. , Timæus, I.M. , and Zaba, B. (eds.) (2012). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population; Spoorenberg, T. (2014), "Reverse survival method of fertility estimation: An evaluation", *Demographic Research* 31(9): 217-246.

5.245. A number of countries have restricted the collection of data from fertility and mortality questions in the census to a sample of enumeration areas,[243] entailing the introduction of more vigorous training and permitting the selection of more suitable field staff. When those items are included in the census, certain precautions to ensure accuracy and completeness should be observed. As noted in the previous paragraph, every effort should be made to collect all relevant information directly from the woman concerned. To reduce underreporting of events and to improve the accuracy of responses to questions on fertility and mortality, enumerators need to receive specific training on probing questions that highlight common causes of errors and omissions.

5.246. The limitations of the data collected and of the estimates based on them should be made clear in the census reports. Furthermore, since some of the estimation procedures are only suitable for use in certain circumstances, it is important that NSOs consult topic specialists or carefully evaluate the methodologies for estimating the indicators for their appropriateness in a given situation. In general, the data in the basic tabulations resulting from these questions should not be used for the direct calculation of fertility and mortality rates. Reliable estimation of fertility and mortality levels using census data requires adjustment based on methods of demographic analysis.[244]

5.247. In countries where civil registration of births and deaths is incomplete or unreliable, it is recommended that a subset of the remaining items should be included as well. Among these, one item ("date of birth of last child born alive") is useful for the indirect estimation of current fertility levels. Two additional items ("children living" and "household deaths in the past 12 months") are especially important, as they allow for the indirect estimation of mortality levels.

5.248. The three remaining items have lower priority: "age, date or duration of first marriage/union"; "age of mother at birth of first child born alive"; and "maternal or paternal orphanhood". However, in situations where a country has included one of these items in consecutive previous censuses, it may be useful to collect comparable information to measure continuing changes over time and because cohort analysis, particularly of the prevalence of orphanhood, can be useful in assessing levels of mortality.

5.249. For countries with low fertility and mortality, and where the civil registration is not reliable, further consideration should be given to the item "age of mother at birth of first child born alive", since it improves the timeliness of estimates of child mortality based upon children ever born and children surviving.

---

[243] For the use of sampling in the enumeration, see Part One, Chapter VI.

[244] *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (United Nations publication, Sales No. E.83.XIII.2); National Academy of Sciences, Committee on Population and Demography, *Collecting Data for the Estimation of Fertility and Mortality*, Report No.6 (Washington D.C., National Academy Press, 1981), p. 220; *Handbook of Population and Housing Censuses, Part II*, Studies in Methods, No. 54 (United Nations publication, Sales No. E.91.XVII.9), Chapters III and IV; *Step-by-Step Guide to the Estimation of Child Mortality*, Population Studies, No. 107 (United Nations publication, Sales No. E.89.XIII.9); Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/.

5.250. In countries with high adolescent birth rates and common child or early marriages (and with a large proportion of women married or in union before age 18), it is recommended that the questions on “age or date of first marriage/union” and “age of mother at birth of first child born alive” are included, since the information improves understanding of the spatial distribution of extent of early childbearing and marriage that surveys cannot provide.

5.251. It is worth emphasizing that all estimates of fertility and mortality derived from census data are approximate and subject to various sorts of error. Therefore, in the absence of complete and reliable civil registration data, it may be desirable to collect more than one type of census information on each topic (for example, both household deaths in the past 12 months and maternal or paternal orphanhood for the purpose of estimating adult mortality). Lastly, it should also be born in mind that while fertility surveys can provide data on current fertility, they cannot provide the small-area level data that the census can. Therefore, a fertility question in the census can still be a priority for many countries.

1. Children ever born alive (core topic)

5.252. Information on number of *children born alive* (lifetime fertility) should include all children born alive (that is to say, excluding foetal deaths) during the lifetime of the woman concerned up to the census date. The number recorded should include all such children, whether born in or out of marriage, whether born in the present or a prior marriage, or in a *de facto* union, or whether living or dead at the time of the census. In the event of multiple births (for example twins), each child should be counted as an individual birth.

5.253. Data on the total number of live-born children should preferably be collected for all women 15 years of age and over, regardless of marital status. If, from a cultural standpoint, it is not acceptable in some countries to attempt to obtain the information for single women, it should be collected for all women 15 years of age and over who are or have been married or in a union (in other words, all ever-married or ever cohabiting women), a group that also includes all widowed, divorced and separated women. Depending on national needs and circumstances, data on their live-born children can be collected for girls under 15 years old as well. In either case, the group of women for whom the data have been collected should be clearly described in the census report so as to avoid ambiguity in the analysis of the results.

5.254. In some countries, there can be substantial misreporting of ages or dates in the census, which distorts fertility and mortality estimation based on children ever born and children living cross-tabulated by age or by years since the woman’s first birth.

5.255. In order to improve the completeness of coverage and to assist the respondent in recalling her children ever born alive, it is recommended that a sequence of questions be included in the following order:

(a) "total number of sons ever born alive during the lifetime of the woman";
(b) "total number of sons living (surviving) at the time of the census"; and
(c) "total number of sons born alive who have died before the census date";

and then

(d) "total number of daughters ever born alive during the lifetime of the woman";
(e) "total number of daughters living (surviving) at the time of the census"; and
(f) "total number of daughters born alive who have died before the census date".

5.256. The responses to topics (b), (c), (e) and (f) allow for a checking of the consistency of the responses to (a) and (d). Inconsistencies in the figures, if any, can sometimes be resolved during the interview. In obtaining this information and ensuring its accuracy, more than one filter question can be asked to serve as a reminder for the respondent. The responses can be confirmed by using additional tables/lists containing the names and birth years of the children she has given birth to.

5.257. The number of sons and daughters should comprise all children ever born alive whether born of the present or a prior marriage or union[245] and should exclude foetal deaths and adopted children. Also, the number of children, male and female, who are alive at the time of the census should include those living with the mother in the household and those living elsewhere, no matter where the latter may reside and regardless of their age and marital status.

5.258. The collection of data on children ever born specified by sex not only improves the accuracy of information but also provides data for the indirect estimation of sex ratio at birth as well as the calculation of infant and child mortality, in combination with data on children living (surviving) by sex (see paragraph 5.260). If the information on "children ever born alive by sex" is collected for only a sample of women, the data on "children living by sex" should also be obtained for the same sample.

5.259. Collecting data on the "total number of children ever born alive by sex" is desirable because it can enhance the quality of demographic analysis. This can be achieved by ensuring that sex ratios of births follow an expected pattern and do not exhibit unusual behaviour. While this point does add value by addressing the importance of sex-specific data collection, it could benefit from providing more context on how anomalies in sex ratios might impact demographic analysis and planning.

2. Children living [246] (core topic)

5.260. Data on *children living*, in conjunction with those on children ever born, are used in indirect estimation of infant and child mortality in situations where there are no reliable data from a civil registration system.

5.261. It is expected that improved coverage and quality of data on the total number of children ever born will be achieved if additional questions about the current residence of children ever born are asked, in terms of the following:

(a) “Total number of sons living in the household”;
(b) “Total number of sons living elsewhere”;
(c) “Total number of sons born alive who have died before the census date”;
(d) “Total number of daughters living in the household”;
(e) “Total number of daughters living elsewhere”;
(f) “Total number of daughters born alive who have died before the census date”.

[245] As indicated in paragraph 5.159, couples living in consensual unions may, where appropriate, be regarded as married.

[246] For methodological details on the uses of the data, together with data on live-born children, see the publications mentioned in footnote 120.

These questions not only give a more complete and accurate reporting of children ever born alive specified by sex but also increase the suitability of the information collected for subsequent analysis.

5.262. The identification of the natural mother of each child under 15 years of age in the same household, to be used in the "own children" method of estimating fertility (see paragraph 5.240), should be made by asking each woman who reports one or more of her children as being born alive and living in the household to identify these children in the census questionnaire. The section of the questionnaire on "relationship to the household reference person" may be used for identifying the natural mother of each child living in the household.

3. Date of birth and sex of last child born alive (core topic)

5.263. Information on date of birth (day, month and year) and sex of the last child born alive is used for estimating current fertility. Data on the sex of the child can also be used to evaluate the sex ratio at birth and to detect potential sex-selective birth omissions, misreporting or coding errors. This information can be useful as a means of deriving both national and subnational fertility estimates. In countries lacking adequate data from civil registration, sample surveys have become a major source of information for estimating national fertility levels, but such surveys usually do not permit the derivation of reliable estimates at subnational levels.

5.264. An estimate of the number of live births during the 12 months immediately preceding the census date can be derived at the data-processing stage from the information collected. For estimating current age-specific fertility rates and other fertility measures, the data provided by this approach are more accurate than information that may have been collected in earlier censuses from a question on the number of births to a woman during the 12 months immediately preceding the census.[247] Information on the date of birth of the last child born alive provides the number of women who had at least one live-born child during the 12-month period, not the number of births during the 12-month period. However, generally only a very small proportion of women will have had more than one child in a year and hence that omission will not significantly affect the fertility estimate derived from it.

5.265. The information needs to be collected only for women between 15 and 49 years of age who have reported having at least one live birth during their lifetime. For countries where childbearing starts early, the minimum age can be lowered according to the national circumstances and needs. Also, the information should be collected for all the marital or union status categories of women for whom data on children ever born by sex (see paragraph 5.252) are collected. If the data on children ever born are collected only for a sample of women, information on date of birth for the last child born alive should be collected for the same sample.

5.266. A census question on the date of birth of the last child born alive should always be paired with a simple follow-up question about whether or not the child is still alive, from which information on deaths

[247] The approach to calculating fertility rates from these data is described in Moultrie T.A. 2013. Evaluation of data on recent fertility from censuses, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/evaluation-data-recent-fertility-censuses (accessed 06/11/2013).

of children born in the last 12 months can be rederived (see paragraph 5.270 below), and which yields data that can be used for studying child mortality (see paragraph 5.276).

4. Births in the past 12 months

5.267. Births in the past 12 months refers to the total number of children born alive to the woman concerned during the 12 months immediately preceding the census. The information about births in the past 12 months is useful for calculating fertility indicators either directly or indirectly.

5.268. Because of errors and omissions commonly encountered in the reporting of live births within a 12-month retrospective period, this topic by itself cannot generally be relied on to generate accurate estimates of current fertility. It is recommended that an estimate of the number of live births during the 12 months immediately preceding the census date be derived from information on "date of birth of last child born alive" (see paragraph 5.264). However, for estimating current age-specific fertility rates and other fertility measures, the data provided by this approach are more accurate than information that may have been collected in earlier censuses from a question on the number of births to a woman during the 12 months immediately preceding the census.[248] Only if a country's population is characterized by low levels of date numeracy – that is, there is a strong indication of deficient accuracy of the replies – should a direct question on number of births in the last 12 months be asked. In all other cases, the core question on date of last child born alive should be sufficient.

5.269. If the topic needs to be introduced into the census, information should be collected for all the marital status categories of women for whom data on children born alive are collected. In countries in which current births to mothers at ages below 15 years are statistically significant, the age limit for the investigation of current fertility may be lowered to include mothers at appropriate younger ages.

5. Deaths among children born in the past 12 months

5.270. Deaths among children born in the past 12 months refers to the number of deaths that occurred among the live births within the 12 months preceding the census reported for the woman concerned. Hence, the topic should be investigated only if live births within this period are also investigated.

5.271. It is recommended that information on the topic be derived from a pair of census questions on "date of birth of last child born alive" and the follow-up question about "whether or not the child is still alive" (see paragraph 5.266). Although this pair of questions does not produce a valid estimate of the infant mortality rate (since the numerator excludes infant deaths occurring below age 1 in the past 12 months among children born 1–2 years before the census date), they can provide useful information on differences in child survival by age of mother or other socioeconomic characteristics.

---

[248] The approach to calculating fertility rates from these data is described in Moultrie T.A. 2013. Evaluation of data on recent fertility from censuses, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation.* Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/evaluation-data-recent-fertility-censuses (accessed 06/11/2013).

5.272. The information needs to be collected only for women for whom at least one live birth during the reference period was reported.

6. Age, date or duration of first marriage or union

5.273. *Date of first marriage* comprises the day, month and year when the first marriage took place. In situations where date of first marriage is difficult to obtain, NSOs should collect information on age at marriage or on how many years ago the marriage took place (duration of marriage[249]). The information should relate to all types of marriages such as contractual first marriages and *de facto* unions, customary marriages and religious marriages.

5.274. For women who are widowed, separated or divorced at the time of the census, information on the "date of/age at/number of years since dissolution of first marriage" should be collected. Information on dissolution of first marriage (if pertinent) provides data necessary to calculate "duration of first marriage" as a derived topic at the data-processing stage. In countries in which duration of marriage is reported more reliably than age, tabulations of children ever born by duration of marriage yield better fertility estimates than those based on data on children born alive classified by age of the woman.[250] Data on duration of marriage can be obtained by subtracting the age at marriage from the current age, or directly from the number of years elapsed since the marriage took place. With regard to de facto unions, it's often difficult to determine precise start and end dates. For women whose de facto unions have terminated, there may be no clear way to ascertain their duration.

7. Age of mother at birth of (date or time since) first child born alive[251]

5.275. Date of first birth comprises the day, month and year when the woman's first live birth took place. In situations where date of first birth is difficult to obtain, it is advisable to collect information on age of mother at first birth or on how many years ago the first birth took place (time since first birth). In countries in which time since first birth is reported more reliably than age, tabulations of children ever born and children surviving by time since first birth yield more timely child mortality estimates than

[249] While this provides some insights into marital history, it doesn't correctly capture the duration of the marriage if it has ended prior to the census date.

[250] See United Nations (1983). *Manual X: Indirect Techniques for Demographic Estimation*, Population Studies, No. 81 (http://www.un.org/esa/population/publications/Manual_X/Manual_X.htm), Chapter II, section D; and United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*, Studies in Methods Series F, No. 92 (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf), Chapter V, section D.

[251] Ibid., Chapter II, section B.3.

those based on data on children born alive classified by age of the woman.[252] If the topic is included in the census, information should be obtained for each woman who has had at least one child born alive.

8. Household deaths in the past 12 months[253] (core topic)

5.276. Information on household deaths in a specified period (usually 12 months) prior to the census date, classified by age and sex of the deceased, is used to estimate the level and pattern of mortality in countries that lack satisfactory continuous death statistics from civil registration. However, exceptional circumstances, such as during times of disaster or pandemics, may necessitate extending the specified period for measuring deaths in the household beyond 12 months to accurately capture mortality patterns.[254] In order for any estimation derived from this item to be reliable, it is important that all deaths to household members occurring during the specified period preceding enumeration be reported as completely and as accurately as possible.

5.277. Census reports often underestimate the true number of deaths. This happens when a death breaks up a household, and the surviving members don't report it, especially if they've moved or formed new households. Deaths of people who lived alone are even less likely to be reported, as there's no one to provide the information. However, sometimes deaths can be over-reported. This can happen when a household splits up before the census because of a death, and then multiple new households report the same death. To make sure the count is accurate, it's important to stress that only the household where the person lived at the time of their death should report it. Clear instructions and training for those collecting census information can help avoid both undercounting and overcounting deaths. Nevertheless, provided that there are no serious errors in the reporting of age at death, estimates of

[252] See Hill K. 2013. Indirect estimation of child mortality, in Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality); and Rajaratnam, J.K., L.N. Tran, A.D. Lopez, & C.J.L. Murray (2010). Measuring under-five mortality: validation of new low-cost methods. *PLoS Med*, 7(4), e1000253. doi: 10.1371/journal.pmed.1000253http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality (accessed 13/12/2013).

[253] See United Nations (2004). *Handbook on the Collection of Fertility and Mortality Data*. Studies in Methods, Series F, No.92 (http://unstats.un.org/unsd/publication/SeriesF/SeriesF_92E.pdf); United Nations (2002). *Methods for Estimating Adult Mortality.* ESA/P/WP.175 (http://www.un.org/esa/population/techcoop/DemEst/methods_adultmort/methods_adultmort.html); and Dorrington, Rob E. (2013). The Brass Growth Balance Method and the Preston-Coale Method for one census, and the Generalized Growth Balance Method and Synthetic Extinct Generations Methods upon the availability of deaths from two censuses, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population. http://demographicestimation.iussp.org.

[254] The DESA-WHO group, established during the COVID-19 pandemic, on mortality estimates has recommended extending the reference period for household death inquiries from one year to a flexible duration, potentially up to two years or longer in the event of crises like pandemics or natural disasters. This change should be implemented following rigorous testing and evaluation of its impact on data quality.

completeness of death reporting can be derived via indirect estimation and adequate mortality estimates obtained.[255,256]

5.278. Ideally, information on mortality should be collected for each household in terms of the *total number of deaths in the specified period* prior to the census date. For each deceased person reported, name, age, sex and date (day, month and year) of death should also be collected. Care should be taken to clearly specify the reference period to the respondent so as to avoid errors due to its misinterpretation. For example, a precise reference period could be defined in terms of a festive or historic date for each country. It is also recommended for a country to use specific reference period e.g. between DD/MM/20YY and DD/MM/20YY.

5.279. When information is collected on household deaths in the previous specified reference period, NSOs may wish to include a pair of follow-up questions concerning cause of death. After ascertaining the name, age and sex of the deceased person and date of death, two additional optional questions could be asked[257]:

(a) Was the death due to an accident, violence, homicide or suicide?

(b) If the deceased was a woman aged 15[258] to 49 at the time of death, did the death occur while she was:

b1. Pregnant? Yes/No

b2. During childbirth? Yes/No

b3. During the six weeks after the end of pregnancy? Yes/No

5.280. Data derived from such questions can help to assess trends and levels, and some causes, of adult mortality and maternal mortality. According to the Tenth International Classification of Diseases, (ICD-10, Class XV, codes O00–O99) there are several concepts regarding maternal death, such as:

- Maternal death: the death of a woman while pregnant or within 42 days of the termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental causes.
- Pregnancy-related death: the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of cause.
- Late pregnancy-related death: the death of a woman while pregnant or within 12 months of termination of pregnancy, irrespective of cause.

[256] See Chapter 4 on methods for data evaluation and adjustment in WHO (2013). *WHO guidance for measuring maternal mortality from a census*. World Health Organization, Geneva (http://www.who.int/reproductivehealth/publications/monitoring/9789241506113/en/index.html); and Moultrie, T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.) (2013). *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org).

[257] See: WHO, 2013 and US Census Bureau, 2017.

[258] It may be appropriate in some countries to reduce the lower age limit by several years.

9. Maternal or paternal orphanhood[259]

5.281. Some countries may also wish to collect information on maternal or paternal orphanhood in another attempt to ascertain the level and patterns of mortality in the population. Census data from these two topics are intended for indirect estimation of mortality by sex. Estimates are based on the proportion of persons classified by age whose natural mothers or fathers are still alive at the time of the census.

5.282. For the collection of information on orphanhood, two direct questions should be asked, namely:

(a) Whether or not the natural mother of the person enumerated in the household is still alive at the time of the census,

(b) Whether or not the natural father of the person enumerated in the household is still alive at the time of the census,

regardless of whether or not the mother and father are enumerated in the same household.

5.283. The investigation should secure information on biological parents. Thus, care should be taken to exclude adopting and fostering parents. It should be kept in mind, however, that overcounting may occur in the case of parents with more than one surviving child among the respondents, particularly in high fertility societies.

5.284. It is preferable for these questions to be collected from every person in the household regardless of age (not just children under 18, which would otherwise make the information useless for estimating adult mortality). Not only is this important for estimating mortality at older ages, but also for estimating the extent of age exaggeration at the older ages. Whenever the context allows, the date of death should be collected to help to improve knowledge of the timing of death, and in other contexts a simple follow-up question about whether the parent was still alive five years ago can help to narrow down the timing of death and to improve adult mortality measurement for recent years by analysing these data as successive cross-sectional enquiries.[260]

F. Registration of vital events and assignment of legal identity

5.285. The collection of data relating to the registration of vital events and the assignment of legal identity is recommended for countries with incomplete civil registration and vital statistics (CRVS)

---

[259] For methodological details on the uses of the data, see Timæus, I.M. (2013). Indirect estimation of adult mortality from orphanhood, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/orphanhood); and United Nations (2002). *Methods for Estimating Adult Mortality*. ESA/P/WP.175 (http://www.un.org/esa/population/techcoop/DemEst/methods_adultmort/methods_adultmort.html).

[260] See Timæus, I.M. (2013). Indirect estimation from orphanhood in multiple inquiries, in T.A. Moultrie, R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus & B. Zaba (eds.), *Tools for Demographic Estimation.* International Union for the Scientific Study of Population (http://demographicestimation.iussp.org/content/orphanhood-multiple-inquiries).

systems, especially those working to improve such systems. Including questions on these topics in population censuses provides valuable data for assessing the quality of civil registration and legal identity management systems. These questions offer insights into the completeness and accuracy of the information recorded in these systems across sub-national regions, population sub-groups, and socio-economic segments. Data from such census questions are crucial for aligning with the principles of universal legal identity and "Leaving No One Behind" as outlined in the 2030 Sustainable Development Agenda. By identifying disparities and inequalities, targeted interventions can be implemented to strengthen civil registration and vital statistics systems and ensure everyone has access to legal identification.

5.286. Census data on these topics can provide valuable insights into societal vulnerabilities. For example, by cross-referencing birth registration data with labor force information, child labor patterns can be identified. Similarly, analyzing marriage registration data can help uncover instances of early marriage. Linking legal identity information with demographic and socio-economic data can highlight disparities in access to essential services such as education, healthcare, and financial systems.

5.287. The decision whether to include questions on registration of vital events and assignment of legal identity requires careful consideration. Factors to weigh include the development stage of the civil registration system, census costs and feasibility, the availability of recent data on the civil registration system's completeness and quality, and the sensitivity of the topic. Ultimately, the specific reasons for including these questions will vary depending on each country's unique context and needs.

5.288. *Civil registration* is the permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population in accordance with the law [261]. There are five types of registration of vital events that can be collected in a census, namely:
i) *registration of birth,* the permanent and official record of a person's existence which provides legal recognition of the person's identity;
ii) *registration of death*, the official recording of the occurrence and characteristics of a death;
iii) *registration of marriage*, the process by which the legal relationship of spouses is constituted;
iv) *registration of divorce,* the legal dissolution of a marriage or registered partnership, according to national laws, conferring on the parties the right to enter into another partnership or marriage; and,
v) *assignment of legal identity,* the issuance or possession of legal identity documents.

[261] UN Principles and Recommendations for a Vital Statistics, Revision 3 (2014).

5.289. The registration of vital events is usually a multi-stage process. In most countries it involves three basic stages[262]: notification, registration, and certification[263].

5.290. Some census respondents and enumerators may not be clear on the differences between notification, registration, or certification of a vital event. And some may even conflate official civil registration with (non-official) recognition of a vital event by other important entities, such as religious institutions or community leaders. Hence, the design and wording of census questions should be carefully reviewed by multiple stakeholders (across the government entities involved in civil registration and identity management processes) as well as community institutions involved in ritual and documentation that are often carried out around the time of a birth, marriage or death.

5.291. Thorough testing of the questions prior to finalization is also recommended. To ensure good quality data, it is important for these concepts to be clarified to field workers during training on data collection.

5.292. Ideally countries should strive to assess levels of registration as evidence of permanent recording of the occurrence of vital events.

5.293. Additional questions on date of occurrence of events can be included. Date of notification and date of registration may be asked in countries where notification does not immediately result in registration. This is largely where the country's legislation provides longer time frames, or no legal period for registration.

5.294. Alternatively, countries may opt to skip registration questions and ask about notification along with possession of a birth/death/marriage certificate. In these instances, possession of a certificate is regarded as proof for registration. In such cases, questions can be added on both date of notification and date of issue of a certificate.

5.295. One of the key considerations is the reference period to which such questions relate. Countries can align registration questions to the census reference period for both births and deaths. For censuses with a reference of the preceding twelve months for household deaths, questions can be in relation to registrations within that same period. This reference period is ideal for countries with a legal framework that stipulates registrations within one to twelve months after occurrence. In the case of birth registration questions, alignment can be with questions on births in the last twelve months.

---

[262] Ibid.

[263] *Notification* is the process of providing all the information on and all the characteristics of an event that is to be legally registered by the registration authority. The informant reports or notifies the vital event to the civil registrar. *Civil Registration* is the permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population in accordance with the law. A certificate is an official document containing all or part of the exact information contained on the original vital record, issued by the registration office, and has the full force and effect of the original vital record. *Certification* is thus a product (output) of registration. It is when the civil registrar issues an official copy containing some or all the vital event information contained in the register.

5.296. Identifying the target population for questions on vital events' registration and the assignment of legal identity documents is crucial. This applies primarily to birth registrations, marriage registrations, and, to a lesser extent, identity card possession. The age range of individuals about whom about birth registration questions should be asked and the relevant time period depend on the purpose, such as:

(i) Assessing recent interventions: Focus on children born within the last twelve months to evaluate the effectiveness of timely birth registration initiatives.
(ii) Tracking progress over time: To assess long-term birth registration trends, consider data spanning several recent years.
(iii) Evaluating policy changes: If a major policy shift or increased investment in CRVS systems happened five years previously, analyze data for the last 10-15 years. This allows comparison of birth registration completeness before and after the change.

5.297. If the goal is to assess birth registration at the national level due to efforts to improve registration for all, questions can apply to all individuals in a household. Where legislation mandates birth registration before issuing legal IDs (linked CRVS-ID systems), then careful consideration should be given to the age range for questions. The target population depends on the legal age for obtaining a national ID card in each country. Countries with separate CRVS and ID systems might assess the extent of linkage by asking everyone in a household about birth registration and ID possession.

5.298. The concept of marriage can vary significantly across cultures and societies. To ensure accurate data collection, the census questionnaire design and enumerator training should clarify whether the question on marriage registration refers to the most recent marriage or the latest marriage that is still ongoing. Only "married" persons should then answer the marriage registration question.

5.299. This clarification will help enumerators to accurately record data and avoid confusion for respondents with diverse marital histories. Furthermore, there may be divergence, within and among communities, between the legal concept of marriage, as per the Family Code and/or civil registration laws, and how nuptiality is practiced (sometimes entailing a wide array of marriage types - such as religious marriage, civil marriage, traditional marriages, and polygamous marriages). Assessment of marriage registration completeness using census data should be well-defined as to the specific phenomenon being assessed, the target population and time period to be considered.

5.300. Regarding the assignment of legal identity, it is important to consider that eligibility for inclusion in identity management systems varies by country. Some countries establish eligibility for a national identity card/credential at birth, whereas others establish it for persons above 16 or 18 years. Countries also have different regulations and processes around naturalization of foreign citizens. These fundamental differences are important considerations, when designing census questions on legal identity.

5.301. Questions on legal identity should be carefully considered, especially in relation to privacy concerns. While some countries might ask about national identification numbers, this should be done cautiously and with respect for data privacy. The collection of national identification numbers can facilitate the linkage of census data with administrative records, potentially supporting a transition towards register-based censuses.

5.302. It is recommended that, as with all parts of a census enumeration, the collection of data on registration of vital events and the assignment of legal identity is carried out with careful consideration

of privacy and confidentiality of individuals, national data protection standards and the UN Fundamental Principles of Official Statistics.

5.303. Civil registration and legal identity systems are usually the responsibility of more than one authority. Such systems often are coordinated across multiple ministries and agencies, and this inherent nature of the systems can affect access to, and inclusion in, public benefit systems such as health, education, and social protection. Therefore, it is recommended, when designing and implementing census questions on civil registration and legal identity systems, that key stakeholders (such as relevant line ministries, civil society, religious and local community leaders) be consulted. Consultation ensures alignment between the chosen measurement approach and the legal framework and operational practices of the civil registration and legal identity systems.

G. Educational characteristics

1. Literacy (core topic)

5.304. *Literacy* has historically been defined as the ability both to read and to write, distinguishing between "literate" and "illiterate" people. A literate person is one who can both read and write, with understanding, a short, simple statement on his or her everyday life. An illiterate person is one who cannot, with understanding, read nor write such a statement. Hence, a person capable of reading and writing only figures and his or her own name should be considered illiterate, as should a person who can read but not write as well as one who can read and write only a ritual phrase that has been memorized. However, a more modern understanding referring to literacy as a continuum of skills, levels, domains of application and functionality is now widely accepted.

5.305. The notion of literacy applies to any language insofar as it exists in written form. In multilingual countries, the census questionnaire may also enquire into the languages in which a person can read and write. Such information can be essential for the determination of education policy. This item would, therefore, be a useful additional subject of enquiry.

5.306. It is recommended that data on literacy be collected for all persons 10 years of age and over. In a number of countries, however, some children may only become literate through school between the ages of 10 and 14 years. The literacy rate for this age group may be misleading. Therefore, in any international comparison of literacy, data on literacy should be tabulated for all persons 15 years of age and over. Where countries collect the data for younger persons, the tabulations on literacy should at least distinguish between persons under 15 years of age and those 15 years of age and over.

5.307. Straightforward operational criteria and instructions for collecting literacy statistics should be clearly established on the basis of the concept given in paragraph 5.304, and applied during census taking.[264] Accordingly, although data on literacy should be collected so as to distinguish between persons who are literate and those who are not, consideration should be given to distinguishing broad

[264] Depending on the need for small-area data and the circumstances in a country, literacy may best be measured through surveys.

levels of literacy skills. Simple questions with response categories that reflect different levels of literacy skills should be used. In addition, since literacy is an applied skill, it needs to be measured in relation to a particular task, such as reading (and understanding) personal letters and newspapers or magazines, or writing a personal letter or message. Respondents may be able to do so easily, with difficulty or not at all, reflecting the different levels of literacy skills. Reading and writing may be measured separately to simplify the questions.

5.308. It is preferable to use standardized questions, harmonized across countries to ensure comparability. The United Nations Educational, Scientific and Cultural Organization (UNESCO) has developed a reference database of model questions. In addition, UNESCO recommends that literacy tests should be administered, in order to verify, as well as improve, the quality of literacy data. Nevertheless, administering such a test to all household members in the course of a census enumeration may prove impractical and affect participation, thereby limiting the utility of the results. Instead, administering such a test to a sample of respondents may be considered either in the census itself or in a targeted post-enumeration follow-up survey. Some countries have regularly used simple self-assessment questions within a census to provide an indication of literacy rates at the small-area level. An evaluation of the quality of statistics should be provided with census statistics on literacy.

5.309. The collection and tabulation of statistics on literacy during the population census should not be based on any assumed linkages between literacy, school attendance and educational attainment. In operational terms, this means systematically enquiring about the literacy status of each household member irrespective of school attendance or highest grade or level completed.

5.310. The literacy question currently varies across countries and, as a result, the data based on it are not always internationally comparable. Literacy should not be derived as an educational attainment proxy because although the two are related, there are substantial differences. For example, there are numerous cases where people leave school with only partial literacy skills, or lose them because of a lack of practice. Therefore, educational attainment is not necessarily a good proxy measure of literacy skills. For countries adopting a register-based census, deriving levels of literacy from administrative data sources poses some difficulties.

2. School attendance (core topic)

5.311. *School attendance* is defined as regular attendance at any accredited educational institution or programme, public or private, for organized learning at any level of education at the time of the census or, if the census is taken during the vacation period, at the end of the school year or during the previous school year. According to the International Standard Classification of Education (ISCED), education is taken to comprise all institutionalized, intentional and planned activities designed to meet learning needs. Instruction in particular skills that is not part of the recognized educational structure of the country (for example, in-service training courses in factories) is not normally considered “school attendance” for census purposes.

5.312. Information on school attendance should, in principle, be collected for persons of all ages. It relates in particular to the population of official school age, which ranges in general from 5 to 29 years of age but can vary from country to country depending on the national education structure. In the case where data collection is extended to cover attendance in pre-primary education or other systematic educational and training programmes organized for adults in productive and service enterprises (such as

the in-service training courses referred to in paragraph 5.311), community-based organizations and other non-educational institutions, the age range may be adjusted as appropriate.

5.313. Data on school attendance should be cross-classified with data on educational attainment, according to the person's current level and grade (see paragraph 5.318). This cross-classification can provide useful information on the correspondence between age and level or grade of educational attainment for persons attending school.

5.314. The issue surrounding the number of out-of-school children has grown in importance within recent decades, particularly within the context of the UNESCO 'Education for All' goal with regard to achieving universal primary education. The target year for Education for All was 2015 and new goals for the post-2015 period were not yet defined at the time when this document was prepared. The census thus offers an opportunity to measure the number of "out-of-school" (reciprocal of attendance) or "ever-in-school" children.

5.315. There is a difference between "attending school" and "enrolled in school", therefore results from censuses and administrative data may differ. A child can be enrolled in a school but not necessarily be attending it. It is recommended that these concepts be clearly defined so that countries can determine which variable they wish to collect via the census. It is also recommended that NSOs consider the need for internationally harmonized questions in order to measure school attendance and school enrolment. Field-based censuses will usually collect data on 'attendance' while administrative data will only refer to 'enrolment'.

5.316. For purposes of international comparison, data on school attendance should be presented by the ISCED-P (or ISCED-Programmes) levels listed below, which are used for the classification of education programmes in ISCED 2011. Correspondence between a national education system and ISCED can be established through mapping of national education programmes to the following ISCED classification:[265]

ISCED level 0: Early childhood education/less than primary education
ISCED level 1: Primary education
ISCED level 2: Lower secondary education
ISCED level 3: Upper secondary education
ISCED level 4: Post-secondary non-tertiary education
ISCED level 5: Short-cycle tertiary education
ISCED level 6: Bachelor's or equivalent level
ISCED level 7: Master's or equivalent level
ISCED level 8: Doctoral or equivalent level

3. Educational attainment (core topic)

5.317. The recommendations on "educational attainment" (see paragraph 5.318) and "educational qualifications" (see paragraph 5.332) make use of categories of the 2011 revision of ISCED, issued by

[265] The UNESCO Institute for Statistics maintains a database with ISCED mappings at http://www.uis.unesco.org/ISCED.

UNESCO.[266] In accordance with national conditions and requirements, many countries can continue to apply national classifications of levels and grades of education and of fields of education when collecting and tabulating statistics from population censuses. Special attention needs to be paid to establishing appropriate level or grade equivalence for persons who have received education under a different or foreign educational system. These national classifications, however, should be able to be converted or mapped to the ISCED 2011 classification system, this typically being achieved during post-census processing.

5.318. *Educational attainment* is defined as the highest ISCED level successfully completed by an individual. Educational attainment is usually measured with respect to the highest education programme successfully completed, which is typically certified by a recognized qualification. Some countries may also find it useful to present data on educational attainment in terms of the highest grade completed. For international purposes a "grade" is a specific stage of instruction usually covered in the course of an academic year, though this may vary from country to country. Information on educational attainment should preferably be collected for all persons 5 years of age and over.

5.319. To produce statistics on educational attainment, a classification is needed that indicates the qualifications certifying the successful completion of levels of education. Since the educational structure may have changed over time, it is necessary to make provisions for persons educated at a time when the national educational system differed from that in place at the time of the census. In addition to focusing attention on the collection of educational attainment data, both the enumerator's instructions and the coding and data processing systems need to be designed in a way that will take account of any changes in the educational system of a country over the years and of those educated in another country, as well as those educated in the current system.

5.320. Information collected on the highest level of education successfully completed by each individual, typically certified by a recognized qualification, facilitates flexible regrouping of the data according to various kinds of aggregation. Recognized intermediate qualifications are classified at a lower level than the programme itself. Information on intermediate qualifications or on the highest grade completed can be used to distinguish between persons who did or who did not complete each level of education.

5.321. For international comparison, data from the population census are needed for all levels of education defined in ISCED. To the extent possible, countries should classify statistics on educational attainment by the individual ISCED-A (or ISCED-Attainment) levels listed below, which are used for the classification of educational attainment in ISCED 2011:

ISCED level 0: Less than primary education
ISCED level 1: Primary education
ISCED level 2: Lower secondary education
ISCED level 3: Upper secondary education
ISCED level 4: Post-secondary non-tertiary education
ISCED level 5: Short-cycle tertiary education

[266] UNESCO Institute for Statistics (UIS). 2012. *International Standard Classification of Education: ISCED 2011*. Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf.

ISCED level 6: Bachelor's or equivalent level
ISCED level 7: Master's or equivalent level
ISCED level 8: Doctoral or equivalent level
or by their equivalent as set out according to the national classification of levels of education.

5.322. For the classification of educational attainment, ISCED level 0 has a different meaning in ISCED 2011 than for the classification of education programmes (see paragraph 5.316): it means not having successfully completed ISCED level 1. This includes individuals who have never attended an education programme, who have attended some early childhood education (defined as ISCED level 0 in the classification of education programmes), or who have attended some primary education but have not successfully completed ISCED level 1. Any differences between national and international definitions and classifications of education should be explained in the census publications in order to facilitate comparison and analysis.

5.323. NSOs could consider asking a question that captures levels of education not successfully completed, should this be of interest to policymakers or other users. This could be in the form of a direct question asking if a person has some education at the relevant level or via a question asking the last grade or year completed from any given level of education.

5.324. Data on school attendance, educational attainment and literacy status should be collected and tabulated separately and independently of each other, without (as elaborated in paragraph 5.309) any assumption of linkages between them.

5.325. In order to ensure continued and improved international comparability of census data by level of education, it is recommended that countries continue to ensure that the educational attainment variable can be mapped into the ISCED 2011 classification. This is typically achieved during data processing.

4. Field of education and educational qualifications

a. Field of education and training

5.326. Information on persons by field of education and training is important for examining the match between the supply and demand for qualified workers with specific specializations within the labour market. It is equally important for planning and regulating the production capacities of different levels, types and branches of educational institutions and training programmes.

5.327. A question on field of education and training needs to be addressed to persons 15 years of age and over who attended at least one grade in secondary education or who attended other organized educational and training programmes at equivalent levels.

5.328. The ISCED Fields of Education and Training 2013 (ISCED-F 2013) distinguishes between broad fields (two-digit codes), narrow fields (three-digit codes) and detailed fields (four-digit codes) of education and training.[267] The broad fields are listed here:

00 Generic programmes and qualifications
01 Education
02 Arts and humanities
03 Social sciences, journalism and information
04 Business, administration and law
05 Natural sciences, mathematics and statistics
06 Information and communication technologies (ICTs)
07 Engineering, manufacturing and construction
08 Agriculture, forestry, fisheries and veterinary
09 Health and welfare
10 Services
99 Unknown

5.329. Additionally, NSOs may wish to consider collecting data on more narrow and detailed fields of education and training. For this, countries should make use of the classification and coding of fields of education and training of ISCED.

5.330. NSOs that code field of education and training according to a national classification should establish correspondence with ISCED, either through double coding or through conversion from the detailed national classification to ISCED. A problem may arise in identifying the exact fields of education and training of persons with interdisciplinary or multidisciplinary fields of specialization. In these cases it is recommended that NSOs follow the procedure of identifying the major or principal field of education and training of those with multidisciplinary specialization.

5.331. In order to ensure continued and improved international comparability of census data by field of education and training, it is recommended that the classification structure for the fields of education and training continue to be based on the most recent version of ISCED.

b. Educational qualifications

5.332. *Educational qualification* is the official confirmation, usually in the form of a document, certifying the successful completion of an education programme or a stage of a programme. Qualifications can be obtained through (a) successful completion of a full education programme; (b) successful completion of a stage of an education programme (intermediate qualifications); or (c) validation of acquired knowledge, skills and competencies, independent of participation in an education programme.

[267] UNESCO Institute for Statistics (UIS). 2014. *ISCED Fields of Education and Training 2013 (ISCED-F 2013).* Montreal: UIS. http://www.uis.unesco.org/Education/Documents/isced-fields-of-education-training-2013.pdf.

5.333. According to national needs, information on qualifications may be collected from persons who have reached a certain minimum age or level of educational attainment. Such information should refer to the title of the highest certificate, diploma or degree received.

H. Individual use of information and communication technology

5.334. In today's digital world, access to and use of Information and Communication Technologies (ICT) by individuals has become a defining feature of modern life. Measuring this phenomenon goes beyond simple device ownership, encompassing the variety of ways people utilize technology for communication, information access, and participation in the digital economy. Understanding these patterns of individual ICT use is crucial for policymakers. It provides insights into factors influencing digital inclusion, highlights potential inequalities in access and skills, and ultimately informs the development of effective policies to bridge these gaps.

5.335. These data are also essential for monitoring progress towards internationally agreed development goals, such as those outlined in the UN's 2030 Agenda for Sustainable Development (indicators on individual Internet use and mobile phone ownership are each referenced through SDG indicators 17.8.1 and 5.b.1, respectively). Mobile phone ownership, in particular, is important to track gender equality since the mobile phone is a personal device that, if owned and not just shared, provides women with a degree of independence and autonomy, including for professional purposes. Several studies have highlighted the link between mobile phone ownership and empowerment, and productivity growth. Existing data on the proportion of women owning a mobile phone suggest that fewer women than men own a mobile phone. This indicator highlights the importance of mobile phone ownership to track and improve gender equality, and monitoring will help design targeted policies to overcome the gender digital divide. By tracking individual ICT usage, policymakers can assess the effectiveness of initiatives aimed at promoting digital literacy, fostering inclusive growth, and ensuring everyone benefits from the opportunities offered by the digital revolution.

5.336. The 3rd Revision of the Principles and Recommendations recommended collection of information on communication device ownership and internet access at the household level. However, the rapid evolution of ICT and the widespread use of mobile devices by individuals means this approach may miss crucial details about a country's digital landscape. Specifically, relying solely on household data overlooks the significant mobile phone penetration and diverse usage patterns within a population. Mobile phones have transcended communication purposes, enabling internet access, streaming services, and other functionalities that may not be captured through traditional household surveys. This proliferation of individual mobile phone use necessitates a shift in data collection.

5.337. The UN's International Telecommunications Union (ITU) recommends[268] collecting information on access to and use of ICTs at both the household and individual levels as that would provide a more accurate picture. At the individual level, the ITU specifically recommends collecting data on Internet use and mobile phone ownership.

[268] *Manual for measuring ICT access and use by households and individuals, 2020 Edition.* https://www.itu.int/en/ITU-D/Statistics/Documents/publications/manual/ITUManualHouseholds2020_E.pdf

1. Individuals using the Internet

5.338. Use of the Internet[269] by individuals is tracked by policymakers and analysts as a key indication of a country's progress towards becoming an information society. The recommended question formulation by the ITU[270] is "Have you used the Internet from any location in the last three months? With response categories of "Yes" or "No". This question is asked of all individuals in the target population, not only those who have used a computer. Access to the Internet can be via a fixed or mobile network. The reference period for individuals who used the Internet should be in the last three months and include use by individuals from any location. Classificatory variables for individuals using the Internet – such as age, sex, education level or labour force status – can help identify digital divides in individuals using the Internet. This information can contribute to the design of targeted policies to overcome such divides.

2. Individuals owning mobile cellular phones

5.339. Mobile cellular telephones have become the predominant method of communication in many countries. Indicators related to mobile telephony are therefore fundamental indicators of the information society. Mobile cellular telephones[271] refer to portable telephones using cellular technology that provides access to Public Switched Telephone Network (PSTN). An individual owns a mobile cellular telephone if he/she has a mobile cellular phone device with at least one active SIM card for personal use. It includes mobile cellular phones supplied by employers that can be used for personal purposes (to make personal calls, access the Internet, etc.) and those who have a mobile phone for personal use that is not necessarily registered under their own name. It excludes individuals who have only active SIM card(s) and not a mobile phone device. Users of both post-paid subscriptions and pre-paid accounts are included. The equipment should be in working condition at the time of the data collection. While these detailed qualifications are important for accurate data collection, they might not all be included verbatim on the questionnaire. Instead, enumerators should be trained to understand and apply these definitions during the interview process. For countries collecting this data, further analysis can be conducted across various dimensions: region (urban/rural), sex, age group, educational level, labour force status, and occupation.

I. Economic characteristics

1. Introduction

5.340. Statistics on the economic characteristics of persons are needed from population censuses for many reasons. Information on the productive activities of persons is vital to establish a comprehensive

[269] The Internet is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer - it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.).

[270] Manual for Measuring ICT Access and Use by Households and Individuals, 2020 Edition. Link: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/manual/ITUManualHouseholds2020_E.pdf

[271] This includes analogue and digital cellular systems and technologies such as IMT-2000 (3G) and IMT-Advanced.

picture of the economic structure of a country, and the work patterns, labour market participation and extent of labour underutilization of its population. This information, when combined with other personal and household characteristics collected in the census, enables assessments of the socioeconomic situation of persons and households, which are essential to inform the planning of a wide range of economic and social policies and programmes related to such areas as employment creation, poverty reduction, work–life balance, vocational education and training, provision of social security and other social benefits, gender justice and social inclusion.

5.341. Such statistics can be obtained from other sources, such as household-based surveys or administrative records, but these have certain limitations. Household surveys, especially labour force surveys, are particularly well suited for generating a broad range of statistics on the economic characteristics of the population at aggregate levels, such as national and broad regional groupings. Data obtained from labour force surveys, however, are subject to sampling error and, therefore, rarely provide reliable estimates for small areas, small population groups, or detailed groups of industries and occupations. In contrast, population censuses can provide certain core statistics at the lowest levels of aggregation for such small population groups and for detailed occupation and industry groups. Administrative records may not have the same quality of occupational and industry coding, nor have the same comprehensiveness in population or activity coverage, generally excluding productive activities that are informal or unpaid. Furthermore, for many register-based census approaches, data from administrative sources may not fully align with ILO concepts, definitions, or classifications.

5.342. The population census also provides benchmark information to which statistics from other sources can be related. Population censuses likewise provide the sample frames for most household-based surveys, including labour force surveys. In countries with a limited or infrequent household survey programme, the population census may represent the main or only source of information on the economic characteristics of the population.

5.343. In deciding which topics relating to the economic characteristics of the population to include in the population census, NSOs will need to assess the existence of other data sources and their complementary uses. The aim should be to cover the core topics needed as benchmark information for the preparation of sample frames, and to provide essential statistics for small areas and small population groups, and for detailed occupation and industry groups, as relevant in the national context.

5.344. International resolutions and guidelines to produce statistics relating to the economic characteristics of the population are adopted by the International Conference of Labour Statisticians (ICLS) and endorsed by the Governing Body of the International Labour Organization.[272] Recommendations on topics amenable for inclusion in population censuses are discussed in general in Box 1.

Box 1. International recommendations concerning statistics of work, employment and labour underutilization

[272] See http://www.ilo.org/stat.

In October 2013, the Nineteenth International Conference of Labour Statisticians (19th ICLS) adopted the *Resolution concerning statistics of work, employment and labour underutilization*.[273] The resolution replaced the previous international recommendations relating to the measurement of the economically active population, employment, unemployment and underemployment dating from 1982 (the 13th ICLS) and related guidelines.

The standards adopted at the 19th ICLS introduced a number of important revisions, among which were a conceptual framework for work statistics consistent with the System of National Accounts; guidelines for separately measuring different forms of work, including a more targeted definition of employment as *work for pay or profit*, and for expanding the range of measures of labour underutilization beyond the traditional unemployment. New terminology was also introduced, as relevant, and terms considered to be out of date, particularly "economically active/inactive", were replaced with "in the labour force/outside the labour force."

Important elements from the 13th ICLS resolution of the economically active population, employment, unemployment and underemployment essential to the internal consistency of the statistics remained unchanged. The refinements of the definition of employment and introduction of measures of labour underutilization resulted, however, in breaks in the historical series of statistics of the economically active population, employment, unemployment and underemployment. In particular, productive activities carried out without pay, such as those listed below, were no longer included within the scope of employment:

- Production of goods intended for own final use by the household;
- Unpaid work by apprentices, interns and trainees;
- Organization-based volunteer work;
- Direct volunteering to produce goods for other households.

With the adoption of the 20th ICLS resolution concerning statistics on work relationships in 2018, a new classification of status in employment, ICSE-18, was introduced. ICSE-18 includes 10 detailed categories for classifying all jobs, providing a more nuanced understanding of different work relationships. These categories can be organized according to the type of economic risk associated with the job and the worker's level of authority over the economic unit and their own work organization. The recommendations for the 2030 round of censuses reflect this change to the status in employment classification.

Countries are encouraged to develop their statistical system to cover work statistics, including statistics on the labour force, based on their specific national needs and resources. In the case of the measures affected by the 19th and 20th ICLS resolutions, the updated international standards would ideally be implemented over time, in a way that is feasible for national statistical systems. It is of utmost importance that the institutions and persons responsible for planning and managing the production of statistics on the economic characteristics of the population develop a strategic and coordinated approach that takes

[273] The Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, October 2013), is available at: http://www.ilo.org/global/statistics-and-databases/meetings-and-events/international-conference-of-labour-statisticians/19/WCMS_230304/lang--en/index.htm.

into account all official sources of statistics, including the population census, labour force survey and other household-based surveys and administrative records. Data users will need to be kept well informed of the process, including by widely disseminating the relevant metadata and by maintaining parallel series for a specified period following their implementation.

2. Conceptual framework for work statistics

a. Work

5.345. Measurement of the economic characteristics of the population is based on the conceptual framework for work statistics (see Box 1). In this framework, work is defined for reference purposes as "any activity performed by persons of any sex and age to produce goods or to provide services for use by others or for own use".

5.346. The concept of *work* is aligned with the general production boundary as defined in the System of National Accounts 2008 (2008 SNA), enabling full integration between work statistics and production statistics. All *work* or *productive* activities are thus included, irrespective of their formal or informal character, or whether paid or unpaid. Excluded are activities that do not involve producing goods or services (for example begging and stealing), self-care (for example personal grooming and hygiene), and activities that cannot be performed by another person on one's own behalf (for example sleeping, learning and activities for own recreation).

5.347. *Work* can be performed in any kind of economic unit, including market units (for example corporations, quasi-corporations and household unincorporated market enterprises); non-market units (for example government and non-profit institutions serving households); and households that produce goods or services for own final use by the producers.[274]

5.348. The conceptual framework for work statistics identifies five mutually exclusive *forms of work* for separate measurement (see Figure 5.1):

(a) *Own-use production work*, comprising production of goods and services for own final use;

(b) *Employment work*, comprising work performed in exchange for pay or profit;

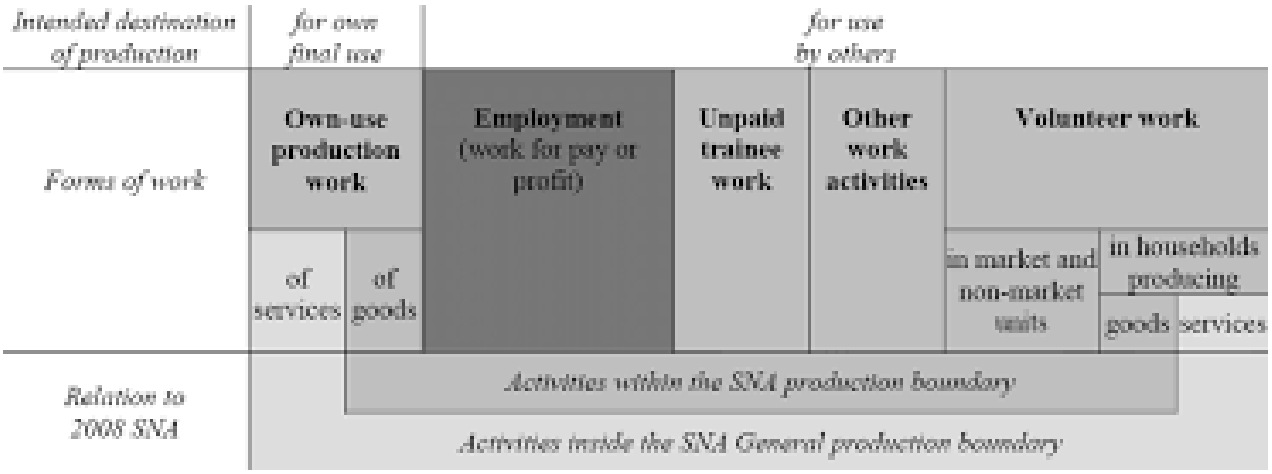
(c) *Unpaid trainee work*, comprising work performed for others without pay to acquire workplace experience or skills;

(d) *Volunteer work*, comprising non-compulsory work performed for others without pay;

[274] United Nations, *System of National Accounts 2008* (New York, 2008).

(d) *Other work activities*, including unpaid compulsory work performed for others, such as community service and work by prisoners, when ordered by a court or similar authority, and unpaid military or alternative civilian service.

Figure 5.1 Forms of work and the System of National Accounts 2008



5.349. During a given reference period, persons may engage in one or more forms of work in parallel or consecutively, that is, persons may be employed, volunteering, doing unpaid trainee work or producing for own final use, in any combination.

5.350. To meet different objectives, NSOs may measure the economic characteristics of the population with respect to their participation in one or in several forms of work. In particular, in the population census, this may include measurement of the following:

(a) Information on *persons in employment* is essential as part of the preparation of labour force statistics that include unemployment and other measures of labour underutilization. It is needed to assess the labour market participation of the population and to classify the population according to their labour force status in a short reference period (see paragraphs 5.358–5.377).

(b) Information on *persons in own-use production of goods* is especially important in countries where particular groups of the population engage in agriculture, fishing or hunting and gathering for own final consumption, including for subsistence (see paragraphs 5.437–5.442), and to enable integration of the population census with the agricultural census (see paragraphs 1.71. –1.77.).

(c*)* Collecting information on *persons in unpaid trainee work* is recommended: where unpaid apprenticeships, internships and traineeships may be a main mechanism of labour market entry for particular groups such as youth or for specific occupations, given their likely overall small size in the country; and where there is a limited availability of alternative statistical sources.

5.351. Given the need for detailed probing, measurement of participation in *own-use provision of services, unpaid trainee work* and *volunteer work* is more appropriate through household surveys or, if desired, through the population census by means of a long form applied to a sample of the population.[275]

5.352. Additional information may also be collected in the population census in order to classify the population according to their main form of work based on self-declaration, in a short or long reference period.[276]

b. Working time

5.353. The concept of working time comprises the time associated with productive or work activities and the arrangement of this time during a specified reference period.[277] Working time relates to each form of work.

5.354. The number of persons engaged in a given form of work provides only a very rough estimate of the volume of work performed, particularly when the work is performed on a part-time, casual or occasional basis. Information on working time is necessary to prepare estimates of the volume of work or labour input for complete national production accounts. It is also essential to support the design, monitoring and evaluation of economic, social and labour market policies and programmes targeting labour market flexibility, work–life balance and conditions of work, including situations of underemployment due to insufficient working time (that is, time-related underemployment) and of excessive working time.

5.355. The population census can serve to provide information on two measures of working time in particular: *hours usually worked,* and *hours actually worked*. Where the census is the only available data source it may as a minimum incorporate a single question on *hours usually worked* for persons in employment (see paragraphs 5.430–5.436) and for persons in own-use production of goods, as relevant (see paragraph 5.442).

c. Population coverage and age limits

5.356. Information on the economic characteristics of the population should in principle cover the entire population, regardless of country of origin or birth, citizenship or geographic location of their place of work. In practice, a lower age limit is usually set in accordance with the conditions and practices in the country. Where national programmes of statistics on the working-age population or on child labour exist, the statistics derived from the population census will serve to complement those bodies of statistics. For purposes of compiling statistics on the working-age population, international standards recommend that countries set the lower age limit taking into consideration the minimum age for employment and

[275] For more details see Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraphs 22(c) and 37–39.
[276] Ibid, paragraph 17.
[277] Resolution concerning the measurement of working time, adopted by the Eighteenth International Conference of Labour Statisticians (Geneva, 2008).

exceptions specified in national laws or regulations, or the age of completion of compulsory schooling.[278] For compiling child labour statistics, the relevant international standards identify the target population as all persons in the 5 to 17 years age group.[279] Countries in which many children participate in employment or in other forms of work (particularly in agriculture) may need to select a lower minimum age than countries where child labour is uncommon. Census tabulations of economic characteristics should at least distinguish between persons under 15 years of age and those 15 years of age and over.

5.357. In general, an upper age limit is not recommended, so as to provide a comprehensive coverage of work activities of the adult population and to examine transitions between employment and retirement. Many people continue to be engaged in employment and in other forms of work beyond retirement age, and the numbers involved are likely to increase as a result of factors associated with the "ageing" of the population. Countries may, however, wish to balance the cost of collecting and processing information relating to the productive activities of older persons (those aged 75 years or more) and the additional response burden imposed on them against the significance and reliability of the information provided.

3. Labour force status (core topic)

5.358. Persons may be classified in a short reference period according to their *labour force status* as being employed, unemployed, or outside the labour force as defined below in paragraphs 5.364–5.384.

5.359. A classification of persons by their *labour force status* provides important information about their relation to the labour market, in particular to *work for pay or profit*, in a short reference period. The three categories of *labour force status* i.e. employed, unemployed or outside the labour force, are mutually exclusive and exhaustive. While even during a short period persons may be engaged in multiple activities, to establish their *labour force status*, priority is given to employment over other forms of work, and over unemployment; and to unemployment over outside the labour force (see Figure 5.1). Thus, for example:

- a volunteer worker who also has a part-time employee job should be classified as *employed*,
- a student who is also seeking and available for employment should be classified as *unemployed*, and
- a person who has a part-time job working only a few hours for pay and who is also seeking another job should be classified as *employed*.

The sum of persons in employment plus persons in unemployment comprises the labour force.

5.360. The *labour force status* of persons is established with regard to a short reference period of seven days or one week, which may be the last seven days prior to the specified census day, the last completed calendar week or a specified recent fixed week. For comparability purposes, it is particularly useful to apply a short reference period with the same total duration (such as a reference week) for the census as for the national labour force survey, if any. This short reference period serves to provide a snapshot picture of labour market participation in the country around the time of the census. As such, the *labour*

[278] Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraph 65.
[279] Resolution concerning statistics of child labour, adopted by the Eighteenth International Conference of Labour Statisticians (Geneva, 2008).

force (that is, persons in employment plus persons in unemployment) reflects the supply of labour for the production of goods and services in exchange for pay or profit at a specified point in time. Seasonal variations in employment and unemployment levels, which may be significant both in industrialized and in developing economies, will not be captured. Assessments of such temporal variations in work patterns are more adequately captured through sub-annual household surveys (for example monthly, quarterly).

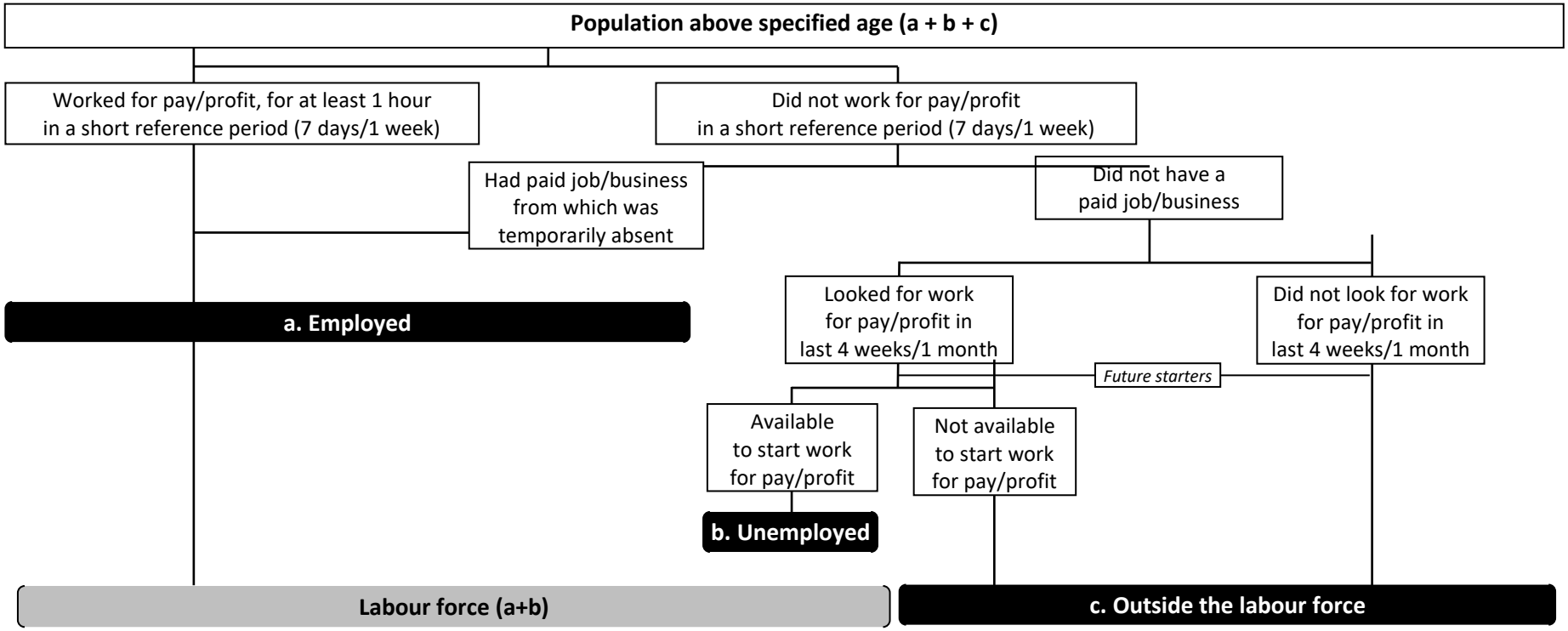
5.361. It is recognised that there is a growing trend for census returns (whether on paper questionnaires or online) to be completed and submitted to the NSO before the official census reference date. Where such circumstances are permitted there should be clear guidance on the questionnaire as to the reference period to which any employment-related activities should refer. This practice may introduce some degree of non-comparability with data collected using the census reference period. Therefore, NSOs should develop procedures for adjusting data, if necessary, to account for potential discrepancies arising from early submissions.

5.362. Depending on the way the relevant parts of the census questionnaire have been constructed, the determination of the labour force status of a person may be influenced by respondents' or enumerators' subjective understanding of the concepts of employment and unemployment. In this regard, particular attention should be given to special groups for which the determination of the labour force status may be difficult. These groups include youths, women and older persons after the normal age of retirement, in particular those working as contributing family members. Their participation in employment and job search activities is frequently overlooked and needs close attention. In particular, a problem in classifying women as outside the labour force, or cultural perceptions relating to gender roles, can result in serious omissions with respect to measuring women's participation in employment and job search activities. To reduce underreporting, enumerators need to be explicitly instructed, or the questionnaires specifically designed, to ask about the possible jobs, including part-time, casual, temporary and informal jobs, or job search activities of every woman and man above the specified age in the household.

5.363. The addition of probing questions in an interview, or more detailed questions in a self-administered questionnaire, may increase the burden on the respondent by lengthening the time required to complete the questionnaire. Accordingly, it will be necessary to balance the gains in terms of minimizing response errors when such questions are used against the added costs associated with their inclusion. Given the importance of reliable data on labour force status, however, serious consideration should be given to minimizing classification errors. To this end, the training of enumerators should highlight likely sources of omission or gender bias leading to underestimation of participation in employment.

Figure 5.2. Classification of working age population by labour force status

***Note: Figure will be reformatted*.**



a. Employed persons

5.364. *Employed persons* are all those above the specified age who during a short reference period of seven days or one week were engaged in any activity to produce goods or provide services for pay or profit. The notion "for pay or profit" refers to work done as part of a transaction in exchange for remuneration payable in the form of wages or salaries for time worked or work done or in the form of profits derived through market transactions from the goods and services produced. It includes remuneration in cash or in kind, whether actually received or not, payable directly to the person performing the work or indirectly to a household or family member.

5.365. Two categories of persons in employment are

(a) employed persons "at work", that is, who worked for pay or profit for *at least* one hour; and

(b) employed persons "not at work" due to working-time arrangements (such as shift work, flexitime and compensatory leave for overtime) or to "temporary absence" from a job for pay or profit.

5.366. Use of the one-hour criterion serves to ensure coverage of all types of jobs engaged in, including part-time, temporary or casual jobs, thereby supporting identification of all persons in employment and analysis of their working conditions. This criterion is also essential in order to ensure that unemployed persons refer to those without *any* employment, who are seeking and available for work for pay or profit. Moreover, this criterion is a prerequisite for the consistency of employment statistics with national accounts data on production. When information on working time is also collected (see paragraphs 5.430–5.436, it is recommended that employed persons be classified by specified bands of working time so as to enable identification of persons with both short and excessive working time.

5.367. Persons on "temporary absence" from a job at the time of the census, including ~~as~~ employees or self-employed, should be considered as in employment, provided that they were "not at work" for only a short duration and maintained a labour attachment during the absence. The existence of a labour attachment should be established on the basis of the reason for the absence and, in the case of certain reasons, the continued receipt of remuneration or the total duration of the absence (in general not greater than three months). Reasons for absence include:

(a) those where labour attachment is generally maintained and thus do not require further probing including: sick leave due to own illness or injury (including occupational injury); public holidays, vacation or annual leave; and periods of maternity or paternity leave as specified by legislation; and

(b) those requiring further assessment of continued receipt of remuneration or total duration including: parental leave, educational leave, caring for others, other personal absences; strikes or lockouts; reduction in economic activity (for example temporary lay-off, slack work); and disorganization or suspension of work (for example due to bad weather, mechanical, electrical or communication breakdown, problems with ICT, and shortage of raw materials or fuels).

Treatment of specific groups

5.368. According to international standards, the following groups of persons should be classified as in employment:

(a) Persons with a job for pay or profit who, during the reference period, were on training or skills enhancement activities required by their job or for another job in the same economic unit;
(b) Apprentices, interns or trainees who receive remuneration in cash or in kind;
(c) Persons who work for pay or profit through employment promotion programmes;
(d) Persons who work in their own economic units to produce goods intended *mainly* for sale or barter, even if part of the output is consumed by the household or family;
(e) Persons with seasonal jobs during the off season, if they continue to perform some tasks and duties of the job, excluding, however, fulfilment of legal or administrative obligations (for example paying taxes);
(f) Regular members of the armed forces and persons on military or alternative civilian service who perform this work for pay in cash or in kind.
(g) Persons who either work in a market unit operated by a family member living in the same or in another household (that is, contributing family workers) or perform tasks or duties of an employee job held by a family member living in the same or in another household should also be classified as in employment.

5.369. These groups of workers are included in employment, regardless of the number of hours actually worked, as they contribute their labour to produce goods and services for pay or profit, payable to the household or family.

5.370. In accordance with the priority rule to establish their *labour force status* (see paragraph 5.355), persons who during the reference period were primarily students, homemakers, pensioners, registered unemployed, or engaged in other forms of work, and at the same time were engaged in employment, as defined above, should be classified as in employment.

5.371. Excluded from employment are:
(a) Apprentices, interns and trainees who work *without* pay in cash or in kind (that is, unpaid trainee work);
(b) Participants in skills training or retraining schemes within employment promotion programmes, when *not* engaged in the production process of an economic unit;
(c) Persons who are required to perform work as a condition of continued receipt of a government social benefit such as unemployment insurance;
(d) Persons with seasonal jobs during the off season, if they *cease* to perform the tasks and duties of the job;
(e) Persons who retain a right to return to the same economic unit but who were absent for reasons specified in paragraph 5.367(b), when the total duration of the absence exceeds the specified threshold or if the test of receipt of remuneration is not fulfilled;
(f) Persons on indefinite lay-off who do not have an assurance of return to employment with the same economic unit;
(g) Persons who work to produce goods intended *mainly* or *exclusively* for consumption or use by the household or family, even if a surplus or part of the output is sold or bartered (that is, own-use production of goods, see paragraphs 5.437-5.442);
(h) Household members who provide *unpaid* services for consumption or use by their household (that is, own-use provision of services);
(i) Persons who work voluntarily and *without* pay to produce goods or services through or for other economic units, including market, non-market units and households (that is, volunteer work).

5.372. Information should be given in the census reports describing how the above-mentioned groups and other relevant groups were treated. Consideration should also be given to the desirability

of identifying some of the groups (for example paid apprentices, interns and trainees) separately in tabulations.

b. Unemployed persons

5.373. *Unemployed persons* are all those above the specified age who (a) were not in employment, (b) carried out activities to seek employment during a specified recent period and (c) were currently available to take up employment given a job opportunity.

5.374. To be classified as unemployed, a person must satisfy all of the three criteria, where:

(a) "Not in employment" (that is, not engaged in work for pay or profit) is assessed with respect to the short reference period for the measurement of employment as defined in paragraph 5.364.

(b) To "seek employment" refers to any activity when carried out, during a specified recent period comprising the last four weeks or calendar month prior to enumeration, for the purpose of finding a job or setting up a business or agricultural undertaking. This also includes seeking part-time, informal, temporary, seasonal or casual employment, paid apprenticeships, internships or traineeships, within the national territory or abroad. Examples of such activities are:

- arranging for financial resources;
- applying for permits or licences;
- looking for land, premises, machinery, supplies or farming inputs;
- seeking the assistance of friends, relatives or other types of intermediaries;
- registering with or contacting public or private employment services;
- applying to employers directly, or checking at worksites, farms, factory gates, markets or other assembly places;
- placing or answering newspaper or online job advertisements; and
- placing or updating résumés on professional or social networking sites online.

(c) "Currently available" serves as a test of readiness to start a job in the present, assessed with respect to the same short reference period that is used to measure employment. Depending on national circumstances, the reference period may be extended to include a short subsequent period not exceeding two weeks in total, so as to ensure adequate coverage of unemployment situations among different population groups.

5.375. Unemployment has been one of the most widely used measures of labour underutilization. However, it only captures persons in situations of complete lack of work for pay or profit, and where opportunities for job search exist. In circumstances where there are few channels for seeking employment or where labour markets are limited in scope, or when labour absorption is inadequate, unemployment will not capture fully all persons with an unmet need for employment, as persons will take any available jobs, create their own jobs, often as own-account workers, become discouraged, or engage in production of goods for own final use (for example, subsistence agriculture or fishing). Additional measures of underutilized labour include the *potential labour force*,[280] defined in paragraph

[280] Replaces the optional relaxation of the "seeking work" criterion in the previous International Conference of Labour Statisticians standards.

5.381, and persons in *time-related underemployment*.[281] Separate identification of these groups of persons supports better assessment of the different types of underutilization affecting labour markets across settings, and for more targeted policymaking.

5.376. It may be useful to distinguish first-time jobseekers, who have never worked before, from other jobseekers in the classification of the unemployed. Such a separation would be useful for policy purposes as well as in improving the international comparability of employment statistics. To do so, however, may require an additional question regarding previous work experience, which may impose too much of a burden for a population census.

Treatment of specific groups

5.377. Also classified as unemployed according to international standards are:

(a) *Future starters*, defined as persons "not in employment" and "currently available" who did not "seek employment" because they had already made arrangements to start a job within a short subsequent period, set according to the general length of waiting time for starting a new job in the national context but generally not greater than three months;

(b) Participants in skills training or retraining schemes within employment promotion programmes, who, on that basis, were "not in employment", not "currently available" and did not "seek employment" because they had a job offer to start within a short subsequent period, generally not greater than three months;

(c) Persons "not in employment" who carried out activities to migrate abroad in order to work for pay or profit but who were still waiting for the opportunity to leave.

(d) In accordance with the priority rule to establish their *labour force status* (see paragraph 5.359), persons who during the reference period were mainly students, homemakers, pensioners, registered unemployed or engaged in forms of work other than employment (for example, own-use production work, volunteer work), and who at the same time were "not in employment", carried out activities to "seek employment" and were "currently available", as defined above, should be classified as in unemployment. Information should be given in the census reports on how persons in these and any other specific groups were treated.

c. Persons outside the labour force

5.378. *Persons outside the labour force* comprise all those in the working-age population who in the short reference period were neither employed nor unemployed as defined above.

5.379. Different classifications of persons outside the labour force may be used for analytical purposes. Particularly useful for informing labour market and social policies and programmes are classifications by *degree of labour market attachment* and by *main reason for not entering the labour force*. These alternative classifications can be derived from the same questions used to identify the unemployed and may be used separately or in combination to enable further analysis.

5.380. Persons outside the labour force may be classified by the *degree of labour market attachment* into the following groups:

[281] See Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013).

(a) *Unavailable jobseekers*, that is, those "seeking employment" but not "currently available";

(b) *Available potential jobseekers*, that is, those not "seeking employment" but "currently available";

(c) *Willing non-jobseekers*, that is, those neither "seeking employment" nor "currently available" but who want employment;

(d) *Others*, that is, persons neither "seeking employment" nor "currently available" who do not want employment.

5.381. The classification of persons outside the labour force by degree of labour market attachment allows identification of the *potential labour force*, computed as the sum of (a) *unavailable jobseekers* plus (b) *available potential jobseekers*. Together with unemployment, the *potential labour force* is a key measure of labour underutilization, relevant both in more and less developed settings, especially where the conventional means of seeking employment are of limited relevance, where the labour market is largely unorganized or of limited scope, when labour absorption is, at the time, inadequate, or where persons are largely self-employed.

5.382. Although not a part of the potential labour force, the group (c) *willing non-jobseekers* represents another group of persons outside the labour force with an expressed interest in employment and is particularly relevant for social and gender analysis in specific contexts.

5.383. Persons outside the labour force may also be classified by their *main activity or reason for not entering the labour market* into the following groups. Some persons may be classifiable in more than one category. In such situations, priority should be given to the possible categories in the following order:

(a) *Attending an educational institution* refers to persons outside the labour force, who attended any regular educational institution, public or private, for systematic instruction at any level of education, or were on temporary absence from the institution for relevant reasons corresponding to those specified for employed persons "not at work".

(b) *Performing unpaid household services* refers to persons outside the labour force engaged in the unpaid provision of services for their own household, such as spouses and other relatives responsible for the care and management of the home, children and older persons. (Domestic and personal services provided by domestic employees working *for pay* in somebody else's home are considered as employed in line with paragraph 5.364 above).

(c) *Retiring on pension or capital income* refers to persons outside the labour force who receive income from property or investments, interests, rents, royalties or pensions from former employment.

(d) *Other reasons* refers to all persons outside the labour force who do not fall into any of the above categories (for example, children not attending school, those receiving public aid or private support and persons with disabilities unable to work).

5.384. Additional reasons for not entering the labour force that are considered particularly important at national or regional level, such as "engaged in own-use production of goods" (for example, foodstuffs from agriculture, fishing, hunting or gathering) should also be taken into account in the classification of population outside the labour force.

4. Characteristics of jobs and establishments

5.385. Once the labour force status of persons has been established, additional important topics regarding the labour market participation of the population relate to the characteristics of their jobs and of the establishments in which they work. These include, in particular, status in employment, occupation, industry, place of work, institutional sector, working time, participation in own use production of goods, and income. Each of these characteristics are discussed in the topics that follow.

5.386. A *job* is defined as the set of tasks and duties performed or meant to be performed by one person for a single economic unit. Persons may have one or several jobs during the reference period. The *main* job is that with the longest hours *usually* worked even if the employed person was not at work in the reference period.[282]

5.387. Job-related characteristics are generally collected in reference to the main job for persons in employment, and may also be collected in reference to the *last main job* (if any) for persons not in employment (that is, unemployed or outside the labour force). This allows for classification of the labour force (that is, employed persons and unemployed persons) and of persons outside the labour force by characteristics of their (last) main job. Once the (last) main job is identified, it is essential that all subsequent questions refer to that same job, even if the respondent was not at work in the reference period. The census questionnaire or the census information taken from registers should be designed in a way that will ensure that the variables "status in employment", "occupation", "industry", "place of work" and "institutional sector" relate to the same job. This should be a central concern also for countries that rely on the use of administrative registers for the capturing of the information on these variables.

5.388. The collection of data on characteristics of the *last main job* of unemployed persons, especially occupation, industry and status in employment, may be useful in order to inform policies aimed at promoting employability and job creation. To serve this purpose, it is generally recommended to set a time limit for past employment experience (for example, during the last five or ten years) and only collect information on the characteristics of the last main job if it was held within this time limit.

5.389. When secondary jobs held in the reference period are also identified, the questionnaire should be designed so as to enable clear and separate identification of characteristics relating to main and secondary jobs. Identification of secondary jobs is particularly important in countries where multiple job holding is commonplace, particularly in agriculture, and when collecting information on income from employment and working time, in order to support analysis of the relationship between employment, income and poverty.

5. Status in employment (core topic)

[282] Resolution concerning statistics of work, employment and labour underutilization, adopted by the Nineteenth International Conference of Labour Statisticians (Geneva, 2013), paragraph 12(b).

5.390. *Status in employment* refers to the type of explicit or implicit contract of employment with other persons or organizations that the employed person has in his or her job. The basic criteria used to define the groups of the classification are: the *type of authority* over the organization of the work and over the economic unit for which the work is performed; and the *type of economic risk* to which the worker may be exposed to the loss of financial or other resources in pursuance of the activity and the unreliability of remuneration. Care should be taken to ensure that an employed person is classified by status in employment on the basis of the same job used for classifying the person by "occupation", "industry" and "sector".

5.391. In October 2018, the Twentieth International Conference of Labour Statisticians (20th ICLS) adopted a new standard defining the classification of work relationships. The Resolution concerning statistics on work relationships, replaced the previous Resolution concerning the International Classification of Status in Employment (ICSE-93) set in January 1993 by the Fifteenth International Conference of Labour Statisticians (15th ICLS) regarding the International Classification of Status in Employment.

5.392. The new International Classification of Status in Employment (ICSE-18)[283], defined in the resolution provides improved boundaries and definitions and facilitates more detailed statistics on various types of work relationships, helping to describe the structure and changes in labour markets more accurately.

5.393. ICSE-18 introduces ten detailed categories that apply to activities defined as employment (see paragraph 5.390 below). These categories can be organized in two ways according to:

a) *type of authority* that the worker has to exercise control over the economic unit for which the work is carried out and for the organizing the work, and the extent to which the worker is dependent on another person or economic unit for organization of the work and/or for access to the market (ICSE-18-A); and

b) *type of economic risk* the worker is exposed to in relation to a possible loss of financial or other resources related to the activity and unreliability of receiving remuneration (ICSE-18-R).

5.394. When organized by authority, ICSE-18-A distinguishes between independent workers and dependent workers, making it useful for various types of labour market analysis such as the impact of economic cycles and government policies. When organized by economic risk, ICSE-18-R categorizes workers into those employed for profit and those employed for pay. This hierarchy is typically preferred for national accounts, identifying wage employment, and producing statistics on wages, earnings, and labour costs.

5.395. Correctly identifying the ten detailed categories in ICSE-18 may require additional questions. Accordingly, it will be important to balance the benefits of providing detailed statistics with the extra burden and cost of including these questions in a census. A recommended approach to minimize burden and costs while still obtaining key statistics is to limit the identification of ICSE-18 to selected aggregated categories. This includes identifying employers, independent workers without employees, and contributing family workers, as defined by ICSE-18-A based on authority.

[283] The new international classification of status in employment (ICSE-18) adopted in a resolution at the 20th ICLS. https://www.ilo.org/publications/international-classification-status-employment-icse-18-manual

5.396. For more detailed categories, such as identifying workers employed for pay, or distinguishing dependent contractors, these can be considered based on the country's objectives, keeping in mind the feasibility and relevance of such data as further discussed now in paragraph 5.397-5.412.

5.397. **ICSE-18-A (according to type of authority):**

Independent workers

A. Employers
- *11. Employers in corporations*
- *12. Employers in household market enterprises*

B. Independent workers without employees
- *21. Owner-operators of corporations without employees*
- *22. Own-account workers in household market enterprises without employees*

Dependent workers

C. Dependent contractors
- *30. Owner-operators of corporations without employees*

D. Employees
- *41. Permanent employees*
- *42. Fixed-term employees*
- *43. Short-term and casual employees*
- *44. Paid apprentices, trainees and interns*

E. Contributing family workers
- *51. Contributing family workers*

5.398. **Employers** are persons who own and manage their own business activities and have control over their operations. This can be done individually or with one or more partners. Employers hire one or more people as employees on a *regular basis*. Hiring employees on a *regular basis* includes engaging temporary employees (with an agreement of three weeks or more including the reference period), engaging different employees with shorter durations in subsequently weeks and employees that are temporary absent. It excludes the employers themselves, their partners, and contributing family workers.

5.399. **Independent workers without employees** are persons who operate a business alone or with partners but do not hire anyone regularly, except themselves, their partners, and contributing family workers. This category corresponds to "own-account workers" in ICSE-93.

5.400. Employers and independent workers without employees can, if feasible and relevant, be further categorized based on whether they operate a corporation or a household market, leading to the identification of the detailed categories (as listed in the classification at paragraph 5.397): *employers in corporations*, *employers in household market enterprises*, *owner-operators of corporations without employees*, and *own-account workers in household market enterprises without employees*. This would require determining if the enterprise is incorporated. An incorporated enterprise holds a distinct legal identity, and its owner(s) bear limited liability for example in relation to any debts incurred by the enterprise. Incorporated enterprises can potentially be identified by adding an extra question in the census or by utilizing administrative records. The identification of these

detailed categories would enhance the accuracy in differentiating between workers in employment for pay and workers in employment for profit, as defined in ICSE-18-R[284].

5.401. **Dependent contractors** are persons whose activities depend on another entity that exercise operational and/or economic control over the activities. They are not employees of that entity as they have a commercial agreement but are dependent on that unit for organization and execution of the work, income, or for access to the market. By definition, dependent contractors lack regular employees and do not own or operate an incorporated enterprise. Dependent contractors are a new category in ICSE-18, sharing characteristics of both independent workers and employees.

5.402. The separate identification of dependent contractors typically requires additional questions to establish whether or not there is a dependent relationship and if control is exercised. In a census, this might not be relevant unless there is a strong national need. If not separately identified dependent contractors will be categorized as independent workers or employees depending on the perception of the person as dependent contractors are to be found among both groups. The impact of this treatment could be estimated using a more suitable statistical source, such as a labour force survey, by for example cross-tabulating ICSE-18 with ICSE-93.

5.403. If deemed relevant, a minimized approach in a census could be used to get a broad estimation of the share of dependent contractors in the country. The measurement should be restricted to non-employees with commercial agreements that do not have regular employees and do not own an incorporated enterprise, focusing on identifying whether these persons have a main client or an intermediary client that sets the price for the goods and services produced[285].

5.404. **Employees** are persons who are in employment for pay, on a formal or informal basis, and who do not hold controlling ownership of the unit (enterprise, a non-profit institution, government unit or household) in which they are employed. Employees would typically be remunerated in cash or in kind for time worked but may also be paid for each task or piece of work done or for service provided. Employees would be accountable to a third party within the unit such as a person or board.

5.405. ICSE-18 provides for detailed categories of employees as noted in the classification at paragraph 5.393: *permanent employees*, *fixed-term employees*, *short-term and causal employees* and *paid apprentices, trainees and interns*. If feasible and relevant these four detailed groups can be identified by establishing whether the employee has a permanent agreement until retirement (or until further notice) or a fixed-term agreement, and whether the fixed term-agreement is less than three months or three months or more. Paid apprentices, trainees and interns are employees who work to acquire workplace experience or skills in a trade or profession and receive payment in return for the work performed. They can potentially be identified separately through a pre-coded category.

5.406. **Contributing family workers** are persons who assist a family member or a household member in a market-oriented enterprise operated by the family or household member. They do not receive regular payments, such as a wage or salary, in return for the work performed, but may receive some

[284] See paragraph 24 in the 20th ICLS resolution concerning statistics on work relationships for the classification of status in employment according to economic risk (ICSE-18-R), adopted by the Twentieth International Conference of Labour Statisticians (Geneva, 2018).

[285] For the recommended approach for identifying dependent contractors see *ICSE-18 manual* at https://www.ilo.org/ilostat-files/Documents/ICSE-18_manual.pdf

irregular payment. They do not make the most important decisions affecting the enterprise or have any responsibility for it.

5.407. Member of a producers' cooperative no longer constitute a status in employment category. In general, such persons are to be consider independent workers in ICSE-18 but may under some circumstances be regarded as dependent contractors if they depend significantly on the cooperative in terms of access to market, organization of pricing of the work, and fulfil the criteria to be classified as dependent contractors.

5.408. Seasonal workers, domestic workers, home-based workers, and workers in multi-party work relationships (including agency workers, and employees providing outsourced services) are, according to the Resolution concerning statistics on work relationships, so called cross-cutting categories as these would be groups that consists of several different status in employment categories. The definition and treatment of these groups can be found in paragraphs 73-124 of the Resolution.[286]

5.409. When members of the armed forces paid in cash or in kind are counted among the employed, they should be included in the category of employees. However, because of the wide range of national practices in the treatment of the armed forces, it is recommended that census tabulations and related notes provide an explicit indication of the status in employment category in which they are included.

5.410. The treatment of additional groups such as work gang (crew) members, franchises, sharecroppers, communal resource exploiters, workers in cooperatives, outworkers and crowd workers can be found in the conceptual framework for statistics on work relationships as these groups are no longer addressed explicitly in the Resolution.

5.411. In most census questionnaires, the information concerning status in employment will be captured through pre-coded alternatives where only a few words can be used to convey the intended meaning of each category. This may mean that classification of some of the situations on the borderline between two or more categories will be carried out according to the subjective understanding of the respondent rather than according to the intended distinctions. This should be kept in mind in designing the questionnaire and also when presenting the resulting statistics. NSOs that rely on the direct use of administrative records for the classification of persons according to status in employment may find that the groups “contributing family workers” and “dependent contractors” cannot be separately identified. Those who would have been classified as being in these groups when information is collected using a questionnaire may either be classified as part of one of the other groups or, in the case of contributing family workers, excluded from persons in employment.

5.412. Activities in areas such as agriculture, fishing, hunting, and gathering intended mainly for own consumption by the household are not considered employment and are therefore not included in ICSE-18. Instead, participation in these productive activities should be measured through the separate concept of “own-use production of goods” (see paragraphs 5.437–5.442) and categorized according to the International Classification of Status at Work (ICSaW-18) as defined in the Resolution concerning statistics on work s relationships[287].

[286] The Resolution concerning statistics of work relationships, adopted by the twentieth International Conference of Labour Statisticians (Geneva, October 2018).
[287] The Resolution concerning statistics of work relationships, adopted by the twentieth International Conference of Labour Statisticians (Geneva, October 2018), Paragraph 59.

6. Occupation (core topic)

5.413. *Occupation* refers to the type of work done in a job by the person employed (or the type of work done in the last job held, if the person is unemployed), irrespective of the industry or the status in employment in which the person's job should be classified. Type of work is considered in terms of the main tasks and duties performed in the job.

5.414. For purposes of international comparison, it is recommended that countries make it possible to prepare tabulations involving occupations in accordance with the latest revision available of the International Standard Classification of Occupations (ISCO). At the time the present set of census recommendations was approved, the latest revision was the one adopted by a Tripartite Meeting of Experts in Labour Statistics in 2007 and endorsed by the Governing Body of the International Labour Organization in 2008[288] and generally known as ISCO-08. Countries coding occupation according to a national standard classification should establish a correspondence with ISCO either through double coding or through mapping from the detailed groups of the national classification to ISCO.

5.415. NSOs should code the collected occupational responses at the lowest possible level of ISCO or a related national classification supported by the information given in each response. In order to facilitate detailed and accurate coding, it would be useful for the census questionnaire to ask each employed person for both the occupational title and a brief description of the main tasks and duties performed on the job. Information provided in response to the industry questions (see paragraphs 5.417-5.420) may also be used to assist in the coding of occupation data, where the occupation response on its own is insufficient to assign a detailed occupation classification code.

5.416. In preparation for the coding of the occupation responses, the NSO should prepare a coding index reflecting the type of responses that will be given by the respondents. This should be constructed by occupational classification experts on the basis of responses to similar questions in other data collections, such as previous censuses, census tests and labour force surveys, as well as input from job placement officers of the employment service and the content of newspaper advertisements of vacant jobs. The coding index should clearly distinguish between responses belonging to "not elsewhere classified" categories and responses that do not provide enough information to determine an occupational group.

7. Industry (core topic)

[288] International Standard Classification of Occupations (ISCO-08), Volume 1, Structure, Group Definitions and Correspondence Tables (Geneva, International Labour Office, 2012). The ISCO companion guide is also available to assist countries using ISCO-08.

5.417. *Industry* (branch of economic activity) refers to the kind of production or activity of the establishment or similar unit in which the job(s) of the employed person was located during the time reference period established for data collection on economic characteristics.[289]

5.418. For purposes of international comparison, it is recommended that countries compile information on industry according to the most recent revision of the International Standard Industrial Classification of All Economic Activities (ISIC) available at the time of the census. At the time this present set of census recommendations was prepared, the fourth edition of ISIC, adopted by the United Nations Statistical Commission at its thirty-seventh session in 2006, was the latest revision available. NSOs coding industry according to a national standard classification should establish correspondence with ISIC either through double coding or through mapping from the detailed groups of the national classification to ISIC.

5.419. NSOs should code the collected industry responses at the lowest possible level of ISIC or a related national classification supported by the information given in each response. In order to facilitate detailed and accurate coding, for each job to be coded the census questionnaire should ask for the main products and services produced, or the main functions carried out at the establishment or enterprise in which the person was employed. It is recommended that the name and address of the establishment should also be collected (see also paragraph 5.424). In countries with business registers that are complete and up to date, the NSO can then use this response as a link to the register in order to obtain the industry code given there to the establishment.

5.420. In preparation for the coding of the industry responses that cannot be matched to a pre-coded register, the NSO should create a coding index that reflects the type of responses that will be given on the census questionnaire. This coding index should be constructed by industry classification experts on the basis of available lists of enterprises, establishments, businesses and so forth, as well as from responses to similar questions in other data collections, including previous censuses, census tests and labour force surveys. As in the coding of occupation (see paragraph 5.416) the industry coding index should clearly distinguish between responses belonging to "not elsewhere classified" categories and responses that do not provide enough information to allow for the coding of a detailed industry group.

8. Place of work

5.421. Two main topics related to the place of work of persons in employment are the *type of workplace* and its *geographic location*. The type of workplace refers to the nature of the place where the person performed his or her *main job* and distinguishes between the home and other workplaces, whether fixed or otherwise.

[289] For those persons who are recruited and employed by one enterprise but who actually work at the place of another enterprise (called "agency workers" or "seconded workers" in some countries), there would be user interest in gathering information about the industry of the employer as well as the industry of the place of work. However, the collection of both would be more appropriate in a labour force survey rather than in a population census. The industry of the actual place of work may provide more reliable reporting of the "industry" variable in a population census. Any such choice should, however, be consistent with the treatment of this group in the System of National Accounts.

5.422. Three main categories, or a variation thereof necessitated by national circumstances, are recommended for classifying the *type of workplace*:

(a) *Work at home*. This category includes those who perform the tasks and duties of their main job from within the home, such as farmers who work and live on their farms, homeworkers, self-employed persons operating (work)shops or offices inside their own homes, and persons working and living at work camps.

(b) *No fixed place of work*. This category should be restricted to persons who, in performing the tasks and duties of their main job, travel in different areas and who do not report daily in person to a fixed address as a work base, for example, travelling salespersons, long-distance commercial vehicle drivers, seafarers, fishermen and own-account taxi drivers. It also includes itinerant vendors, operators of street or market stalls that are removed at the end of the workday, construction workers working at different sites during the reference period and push-cart operators.

(c) *With a fixed place of work outside the home*. All other persons in employment should be included in this category, including persons who move around in their job but have a fixed-base location to which they report daily, such as bus and taxi drivers (with a base), train and airline staff, and operators of street and market stalls that are not removed at the end of each workday. This group may also include individuals who travel to work, on a regular basis, across the national border to a neighbouring country. A fixed place of work outside the home includes: client's or employer's home; employer's workplace or site; own business premises; or, client's workplace or site.

5.423. It is possible that for some jobs, performance may be at more than one location (for example, at home some of the time or season and in a fixed location outside the home at other times) or the category cannot be clearly distinguished. One approach, in the case of the former, would be to select the place where the individual spends or spent (during the reference period) the major part of his or her working time. Where the distinction between categories is blurred, as is the case for work done, for example, on a rented plot of land adjacent to the person's home, it would be useful to identify borderline cases, according to national circumstances. Specific instructions should be given to the enumerators on how to select between two or three possible responses to classify borderline cases.

5.424. The *geographic location* of the place of work can provide useful information for planning when used together with information on place of residence. To this end, NSOs should collect, for employed persons with a fixed place of work outside the home, information on the location of the place of work (or the reporting place) during the reference period. The information collected should relate to the smallest civil division in which the job is performed, for example, in order to establish commuter flows from the place of residence to the place of work. Some NSOs investigating this topic in the population census collect the actual address of the place of work, allowing detailed tabulations and mapping of place of residence by geographic location of place of work. Information on actual address of the place of work can also be useful for industry coding (see paragraph 5.419) in countries where a business register has been developed that shows the industry code of each recorded establishment.

5.425. In some countries there may be concerns about the sensitivity of questions on the address of place of work owing to suspicions that there may be follow-up to a respondent's employer. In many developing countries, it may not be possible to gather information on actual address of place of work because street addresses do not exist, and for proxy responses, the address may not be known. In those situations, it would be useful to consider collecting information on the village, suburb, or similar low level of geography.

5.426. Additional questions may also be asked on the method of travel to work in order to produce statistics on travel-to-work patterns, valuable as basis for transportation planning. Persons not travelling to work should be classified as "no travel to work."

9. Institutional sector of employment

5.427. The *institutional sector of employment* relates to the legal organization and principal functions, behaviour and objectives of the enterprise with which a job is associated.

5.428. Following the definitions provided in the System of National Accounts, distinction should be made between the following institutional sectors:

(a) *Corporation*, comprising non-financial and financial corporations (in other words incorporated enterprises, private and public companies, joint stock companies, limited liability companies, registered cooperatives, limited liability partnerships, and so forth) and quasi-corporations (that is to say, an unincorporated enterprise that is managed as if it were a corporation, in that a complete set of accounts is kept), as well as non-profit institutions, such as hospitals, schools and colleges that charge fees to cover their current production costs;

(b) *General government*, comprising central, state and local government units together with social security funds imposed or controlled by those units, and non-profit institutions engaged in non-market production controlled and financed by government, or by social security funds;

(c) *Non-profit institutions serving households* (for example, churches, professional societies, sports and cultural clubs, charitable institutions and aid agencies) that provide non-market goods and services for households (that is to say, free or at prices that are not economically significant) and whose main resources are from voluntary contributions;

(d) *Households* (including unincorporated enterprises owned by households) comprising unincorporated enterprises directly owned and controlled by members of private and institutional households (made up of persons staying in hospitals, retirement homes, convents, prisons and so forth, for long periods of time), either individually or in partnership with others. Partners may be members of the same household or from different households.

5.429. In most census questionnaires, the information concerning institutional sector of employment will be captured through precoded alternatives where only a few words can be used to convey the intended meaning of each category. This may mean that classification of some units on the borderline between two or more categories will be carried out according to the subjective understanding of the respondent rather than according to the intended distinctions. This should be kept in mind when presenting the resulting statistics.

10. Working time

5.430. The number of employed persons provides only a very rough estimate of the volume of work performed, especially when such persons have non-standard working hours. Inclusion in the census of an item on time worked helps to ensure a more accurate measurement of the concept by capturing the full contribution of persons who were in and out of the workforce or who worked only for a brief time during the year (for example, women).

5.431. To provide a comprehensive measure of working time in employment that will best inform policy and analytical needs, it may be preferable to collect information about the total hours worked in all jobs rather than to limit the information to hours worked in the main job.

5.432. Information on two distinct concepts of working time can be collected in a population census: *hours actually worked* and *hours usually worked*.

5.433. *Hours actually worked* is defined as the time spent in a job for the performance of activities that contribute to the production of goods and/or services during a specified reference period. It covers the time spent in "direct hours", in "related hours", "down time" and short "resting time". "Direct hours" is the time spent carrying out the tasks and duties of the job – and may be performed in any location. "Related hours", while not leading directly to goods produced or services provided, is the time spent maintaining, facilitating or enhancing productive activities, including upkeep of the workplace, changing time or decontamination of work clothes, purchasing or transporting materials, waiting for business, customers or patients, on-call duties, travelling between work locations, and work training or skills enhancement required by the economic unit. In practice, "down time" includes unavoidable, temporary interruptions to work (for example machinery or Internet breakdown, lack of supplies). "Resting time" is inactive time for short rest or refreshment in the course of performing job-related activities, (for example coffee breaks). Longer breaks for meals, time spent not working because of vacation, holidays, sickness, industrial disputes, etc., commuting to work (if not also performing job tasks or duties) and educational leave even if paid, are excluded from hours actually worked.

5.434. Measurement of hours actually worked in employment, in the context of the population census, is usually collected using one direct question; it is optimally measured using a set of questions, requesting hours separately for each day of the week. For employed persons not at work in the short reference period, it is possible to have a value for the number of hours actually worked of zero (for persons away on leave) or some other reduced number of hours (if a part of the reference period was taken off for sickness, holiday, or other purpose).

5.435. *Hours usually worked* is defined as the typical value of the hours actually worked in a job per short reference period (for example one week) over a long observation period (month, quarter, season, year) that includes the short reference period itself. This "typical value" of time worked during a normal or typical week may be the modal number of the hours actually worked in the short period as distributed over the long period. This would include overtime hours regularly worked whether paid or unpaid. Days and hours not usually worked and unusual periods of overtime are not included.

5.436. Measurement of hours usually worked in employment relating to the short reference period of one week can be done with one direct question: how many hours do you usually work per week (in your main job or in all jobs)? For persons with more than one job during the reference week, to record both working time in the main job (for which the other descriptive variables are collected) and total working time (sum of working time in all jobs) the questionnaire would require, at minimum, two questions.

11. Participation in own-use production of goods (core topic)

5.437. Countries where production of goods for own final use (such as the production of foodstuffs from agriculture, fishing and hunting and of other household goods, and the gathering of water and firewood) represents an important component of the livelihood of a part of the population, whether as a main or secondary activity, will need to consider collecting information in the population census on the number of persons engaged in this form of work (previously included within the concept of employment). Such information is essential for benchmarking purposes, especially where household surveys are not frequent, for comprehensive sectoral analysis, particularly of work in agriculture, forestry and fishing, and to enable integration of the population census with the agricultural census (see also paragraphs 5.448–5.456).

5.438. *Persons in own-use production of goods* are all those above the specified age who, during a specified reference period, performed "any activity" to produce goods for own final use. The notion "for own final use" is interpreted as production where the intended destination of the output is *mainly* for final use by the producer in the form of capital formation, or final consumption by household members, or by family members living in other households.

5.439. According to international standards, "any activity" to produce goods (within the 2008 System of National Accounts production boundary) covers work performed for at least one hour in the following activities, when the intended destination of the output is *mainly* for own final use, as specified above:

(i) Producing and/or processing for storage agricultural, fishing, hunting and gathering products;
(ii) Collecting and/or processing for storage mining and forestry products, including firewood and other fuels;
(iii) Fetching water from natural and other sources;
(iv) Manufacturing household goods (such as furniture, textiles, clothing, footwear, pottery or other durables, including boats and canoes);
(v) Building, or effecting major repairs to, a person's own dwelling, farm buildings, etc.

5.440. For measurement purposes, the intended destination of the output is established in reference to the specific goods produced, based on self-declaration (that is, mainly for own final use). In the case of goods from agriculture, fishing, hunting or gathering intended mainly for own consumption, a part or surplus may nevertheless be sold or bartered.

5.441. Persons may engage in own-use production of goods as a main or secondary activity, throughout the year or on a seasonal basis. To ensure complete coverage, a dedicated census questions on participation in own-use production of goods should be applied to all persons above the specified age for collecting information on the economic characteristics of the population, *irrespective* of their labour force status. The reference period may refer to the last 12 months, calendar year, agricultural year or season, as relevant to national circumstances. Where pertinent, the choice of reference period should promote coherence with the agricultural census (see also paragraphs 1.71. – 1.77.).

5.442. For assessments of the volume of work performed by persons in own-use production of goods, particularly when using a long reference period, it may be useful to include a question on working time, in particular hours usually worked (see paragraph 5.435), or based on broad categories such as part time or full time, part year or full year, number of months, as feasible and relevant to the main uses of the statistics.

12. Income

5.443. Countries may wish to collect information on the amounts of income received by individual persons or households during a specified reference period, from any source. If this topic is included in the census it is recommended that data be obtained for all persons above a specified age, whether they are employed or not. The NSO may then report income at the household level, or for each individual in the household.

5.444. *Income* may be defined as all receipts whether monetary or in kind (goods and services) that are received by the household or by individual members of the household at annual or more frequent intervals, but excluding windfall gains and other such irregular and typically one-time receipts. Household income covers (a) income from employment (both paid and self-employment); (b) income from the production of goods for own final use; (c) income from the provision of household services for own final use; (d) property income; and (e) current transfers received.[290]

5.445. The collection of reliable data on income, especially income from self-employment, property income and investments, is extremely difficult in general field enquiries, particularly population censuses. The inclusion of non-cash income (such as food or housing benefits) further compounds the difficulties. Collecting household income data in a census, even if limited to monetary income, poses challenges related to privacy concerns, respondent burden, and the potential for inaccurate responses. Therefore, this topic is generally considered more suitable in a sample survey of households or from administrative data sources such as tax or social security records. But, depending on the national requirements, it might still be useful to get some basic income information in a census. This could mean asking only about wages, keeping the time period short (like one month), and focusing only on money received as currency, electronic transfers, or checks. Even this limited information can be helpful for many purposes.

5.446. According to international standards on the subject, the income from employment of employed persons should include wages and salaries of employees, income of members from producers' cooperatives and the mixed income of employers and own-account workers operating business and unincorporated enterprises. In addition to the income from employment of employed household members, the total income of the household should include, for example, the interest, dividends, rent, social security benefits, pensions and life insurance annuity benefits of all its members. The *Handbook on household income statistics*[291] provides further guidance on concepts and methods related to this topic.

5.447. The concepts involved in determining income are not simple to grasp and respondents may be unable or unwilling to provide exact information. For example, income should include social security benefits and pension fund contributions and should be calculated before any direct taxes have been deducted from employees' salaries. Significant items of total household income may also be excluded or misstated. Despite instructions given to enumerators, the data collected can therefore only be expected to be approximate. Accordingly, in the presentation of results it is usually appropriate to band income or earnings into broad classes. As an aid to the interpretation of the results, tabulations of the data should be accompanied by a description of the items of income assumed to be included and, if possible, an estimate of the accuracy of the figures.

[290] See Resolution concerning household income and expenditure statistics, adopted by the Seventeenth International Conference of Labour Statisticians (Geneva, 2003), paragraphs 4–5. Available at http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087503.pdf.

[291] *Canberra Group Handbook on Household Income Statistics*, second edition, 2011, available at http://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canbera_Handbook_2011_WEB.pdf.

J. Agriculture

1. Introduction

5.448. This section presents two non-core topics related to agriculture. These topics are helpful for countries aiming to gather information in their population census that can be used to create a frame of agricultural holdings in the household sector for designing a future agricultural census (see paragraphs 1.71. -1.77.).

5.449. Some countries use population and housing census data on occupation and employment status to identify individuals involved in agriculture. However, as noted in paragraph 1.73. , this approach has limitations. The short reference period (often a week) might miss people involved in seasonal agricultural activities or those for whom agriculture is not their main activity. On the other hand, the population census could include people working in agriculture as paid employees, which are outside the scope of the agricultural census. To address this issue, two non-core topics at the household level are proposed.

2. Own-account agriculture production

5.450. Some countries might use the population census to identify households engaged in "own-account agricultural production." This refers to households raising livestock or using land (wholly or partly) for agricultural purposes. Households with members engaged in agricultural activities only as paid employees would not qualify. The definitions used here align with those outlined in Resolution I of the 19th ICLS, which includes own-use production of goods (see paragraph 5.438). Paragraph 5.384 highlights the importance of capturing information on own-use production, especially "in countries where populations rely on agriculture, fishing, or hunting and gathering for their own consumption." This data is crucial for integrating population and agricultural censuses. Including this topic can facilitate analysis of the population census data related to agriculture and its use as a frame for a subsequent agricultural census or surveys. However, comprehensive data collection would require measuring both employment and own-use production in agriculture.

5.451. The question on own-account agriculture production gathers information at the household level. It asks if any household member participated in own-account agricultural production activities during the year preceding the census, either at their usual residence or elsewhere. As noted at paragraph 5.441, a year-long reference period is necessary to capture seasonal variations, which would not be possible with the shorter reference period used in other topics pertaining to economic characteristics.

3. Measure of farmland area and number of livestock

5.452. The second topic collects information at the household level on the size of the farm (land area or number of plots) used for agriculture and the number of livestock reared (for the main livestock species) on the census day. This helps assess the extent of a household's agricultural activities and to define a suitable threshold for the census of agriculture (e.g. a minimum size limit for inclusion). For countries seeking a more comprehensive picture of agricultural activities within population and housing censuses, gathering data on employment in agriculture and own-use production of goods is recommended.

5.453. "Agricultural activities" encompass all the major agricultural production activities relevant to the country (including crops, livestock, and related activities listed at paragraph 5.456). Information on aquaculture, forestry, and fishing can also be collected if significant for the country.

5.454. Countries aiming for more extensive agricultural data can include an agricultural module with selected data items suggested by the FAO World Programme for the Census of Agriculture 2030[292] and the FAO/UNFPA Guidelines for Linking Population and Housing Censuses with Agricultural Censuses.[293]

5.455. For countries where household-level aquaculture production is important, information can be collected on whether or not any household member participates in any form of own-account aquacultural production activities.

5.456. Agricultural production activities refer to groups 011, 012, 013, 014, and 015 of ISIC (Rev. 5.0):

- Group 011: Growing of non-perennial crops
- Group 012: Growing of perennial crops
- Group 013: Plant propagation
- Group 014: Animal production
- Group 015: Mixed farming

Aquacultural production activities refer to group 032 of ISIC (Rev. 5.0):

- Group 032: Aquaculture

[292] Food and Agriculture Organization of the United Nations (FAO) (2026). World Programme for the Census of Agriculture 2030 (WCA 2030). [Forthcoming]

[293] Food and Agriculture Organization of the United Nations (FAO) and United Nations Population Fund (UNFPA) (2012). Guidelines for Linking Population and Housing Censuses with Agricultural Censuses. [Link: https://www.fao.org/4/i2680e/i2680e00.htm]

Chapter II. Housing census topics

I. Factors determining the selection of topics

5.457. In line with the overall approach to this revision of *Principles and Recommendations for Population and Housing Censuses*, the selection of housing census topics, as with the population topics described in Chapter I, is based on the outputs expected to be produced. Therefore, the first step involves the clear identification of such outputs; the core and additional topics are then decided on that basis. It is recommended that NSOs collect data on the core topics and also produce the recommended tabulations,[294] to improve the international harmonization and comparability of statistics through the use of common concepts, definitions and classifications. Use of an agreed international approach would also enhance the capacity of NSOs to generate statistics for monitoring the socio-economic situation of national populations, including for the provision of data for the internationally agreed development goals.

5.458. In reference to the selection of topics to be included in a housing census, limiting statistical enquiries to the collection of data that can be processed and published within a reasonable period of time is very important. Such cautions are especially applicable to a housing census, since it is customary to conduct housing and population censuses simultaneously or as consecutive operations. There is a high probability that the amount of data required from a census may be beyond the capacity of enumerators to collect or the census agencies to process. It may be sufficient in some developing countries, for example, to ascertain only the number of housing units and other sets of living quarters of various types, the number and characteristics of the occupants thereof and the availability of a water supply system.

5.459. In this context, NSOs should, however, not attempt to collect housing data that are so incomplete that they fail to serve the principal purposes for which they are required. It is important, therefore, for NSOs to consult closely with the principal users at an early planning stage in order to identify the data that are of highest priority and the means of supplying them in the most useful formats.

5.460. The topics, therefore, to be covered in a housing census (that is to say, the subjects for which information is to be collected regarding buildings, living quarters and households) should be based on a balanced consideration of:
(a) **National priority**: It is essential to consider the diverse needs of data users in the country at both the national and local area levels to ensure that the census generates relevant information to effectively address these needs.
(b) **International comparability**: It is important to ensure that the data collected is comparable at regional and global levels through aligning with international standards. This would facilitate inter-country analyses and monitoring.
(c) **Suitability**: For information collected through questionnaires, the NSO should consider the sensitivity of the topics and the respondent burden, that is, the willingness and ability of the public to provide accurate information on these topics. Similarly, with respect to information to be collected through administrative registers, the NSOs should consider the sensitivity of the topics and the willingness of the public to have their personal information linked and used, in other words assess the acceptability of the

[294] Link to UNSD website containing the tabulation shells.

use of existing administrative information for statistical purposes. It is also important to consider the technical competence of NSO staff in regard to obtaining the requisite census information on the topics through data linkage.
(d) **Resources**: NSOs should consider the total national resources available for conducting the census, including financial, human, and technological assets.
(e) **Alternative sources**: It is important to assess the availability of relevant information from alternative data sources to reduce duplication and improve the efficiency of data collection.

5.461. Such a balanced consideration will need to take into account the advantages and limitations of alternative methods of obtaining data on a given topic within the context of an integrated national programme for gathering housing statistics.

5.462. In selecting the housing topics, regard should also be given to the usefulness of historical continuity, which provides the opportunity for comparison of changes over a period of time. NSOs should avoid, however, collecting information that is no longer required by users. Information should not be collected simply because it has been traditionally collected in the past, bearing in mind changes in the socioeconomic and housing circumstances of the country. It becomes necessary, therefore, in consultation with a broad range of users of census data, to review periodically the value of even long-standing topics and to re-evaluate the need for their continued inclusion, particularly in the light of new data needs and alternative data sources that may have become available for investigating topics hitherto covered in the population and housing census. Each of five key factors that need to be taken into account in reaching a final decision on census content are briefly reviewed in the following paragraphs.

A. National priority

5.463. The priority of designing a housing census should (as with the design of the population census) be to meet national needs. The prime consideration is that the census should provide information on those topics that are of greatest value to the country, with questions framed, or appropriate administrative data sources identified, so as to elicit data of maximum utility..

5.464. Each country's decision with regard to the topics to be covered should depend upon a balanced appraisal of how urgently the data are needed and whether the information could be equally well or better obtained from other sources such as administrative records of housing surveys (see Section E below). Some countries may omit from the census certain recommended topics because there is not a national need to collect the data. For example, a particular amenity, such as electricity or toilet facilities, might be available virtually everywhere in a country, and, consequently, there may be no need to collect such information in a census at all. Conversely, some topics may not be included in a census because of the almost total absence of certain amenities, particularly in the rural areas of some developing countries.

5.465. The importance of involving data users and policy makers in the process of identifying priorities and policy needs has to be taken into consideration early in the process of designing the housing census. The topics that are of particular interest to policymakers need to be carefully assessed in terms of applicability, reliability of data and census limitations (such as respondent burden). More detailed information on involvement of stakeholders is presented in the section on "User consultation,

communication and publicity" at paragraphs 2.173–2.202, and also in the *Handbook on Census Management for Population and Housing Censuses*.[295]

B. International comparability

5.466. The desirability of achieving regional and worldwide comparability should be another major consideration in the selection and formulation of topics to be included in the census. National and international objectives are usually compatible, since broad studies of countries' experiences and practices are the basis of international recommendations.

5.467. If particular circumstances within a country necessitate a departure from international standards, every effort should be made to explain these departures in the census publications and to indicate how the national presentation can be adapted to the international standards.

C. Suitability

5.468. A prerequisite for the inclusion of housing topics in the census should be the willingness and ability of respondents to provide accurate information on them, or, for register-based censuses, the acceptance of the public to have their housing information linked and used for statistical purposes. It is advisable to avoid topics that could increase the burden on respondents and those that are likely to arouse fear, local prejudice or suspicion or which might be used to deliberately promote political or sectarian causes, as these are likely to have a detrimental effect on response rates and support of the census. In an interview-based census or where the collector needs to obtain information through observation, consideration needs also to be given to the level of knowledge and skills of the interviewers or collector and whether or not they can be adequately trained to collect this information accurately. Topics that are too complicated or difficult for the average respondent to answer or the enumerator to record quickly should also not be included. The exact phrasing of questions that will obtain the most reliable responses may depend on national circumstances and, should be well tested prior to the census (see paragraphs 3.1403.149).

D. Resources

5.469. The selection of topics should be carefully considered in relation to the total resources available for the census. An efficient collection of accurate data for a limited number of topics, followed by prompt tabulation and publication, is more useful than the collection of data for an overambitious list of topics that cannot be properly processed and disseminated in a timely, reliable and cost-effective manner. In balancing the need for data against available resources, the extent to which questions can be precoded is yet another consideration. This may be an important factor in determining whether or not it is economically feasible to include certain topics in the census. In the case of those topics on which data is taken from registers, the additional resources required to transform the administrative records into useable statistical data should also be noted.

E. Alternative sources

[295] United Nations publication, Sales No. E.00XVII.15 Rev.1.

5.470. In the selection of topics to be investigated in a housing census, consideration should be given to whether data are available from other sources, such as surveys and administrative records, taking into account their relative advantages and limitations. While these sources offer valuable information, they often have limitations. Housing surveys, for example, typically rely on samples and might lack the geographic coverage needed for detailed analysis of small areas or certain types of dwellings. Censuses, on the other hand, can collect comprehensive data through direct interviews. This allows for probing questions and clarification, which may not be possible with self-enumeration or administrative records. Those topics for which no alternative sources exist should be given higher priority while those for which alternative sources are readily available should be accorded lower priority.

II. List of topics

5.471. The units of enumeration for housing censuses are buildings, living quarters, households and occupants. The building is often an indirect but important unit of enumeration for housing censuses since the information concerning the building (building type, material of construction of external walls and certain other characteristics) is required to describe properly the living quarters located within the building and for the formulation of housing programmes. In a housing census, the questions on building characteristics are normally framed in terms of the building in which sets of living quarters being enumerated are located, and the information is recorded for each of the housing units or other sets of living quarters located within it. Administrative registers may contain useful information concerning buildings.

5.472. The principal direct enumeration unit in a housing census is the living quarters. Only by recognizing this as such can data be obtained that will provide a meaningful description of the housing situation and a suitable basis for the formulation of housing programmes.

5.473. The second direct unit of enumeration is the households occupying the living quarters. For each household, it is often useful to collect information on the characteristics of the head or reference person, tenure in the housing unit, and other relevant characteristics.

5.474. The final units of enumeration are the occupants within households. However, the detailed characteristics of each individual household member are collected in a population census and are covered in Chapter I.

5.475. The list presented below (Table 5.2) is based on the global and regional census experience of the last several decades. The topics included are those on which there is considerable agreement on their importance and feasibility for inclusion in a census for the purpose of measuring and evaluating housing conditions and formulating housing programmes. Those that are likely to present difficulties and require time-consuming questioning can probably best be investigated in a separate housing survey of a sample of living quarters.

5.476. Core topics are those of common interest and value to countries and also of importance in enabling comprehensive comparison of statistics at the international level. Other topics refer to data that need to be collected in order to meet the additional requirements of national users.

5.477. It should be emphasized that the topics or variables on housing contained herein are for tabulation and production of outputs as this is the overall purpose of this part of the document. Issues

that pertain to data collection are addressed in other parts of these recommendations and other relevant United Nations handbooks.

Table 5.2. Housing census topics by unit of enumeration

No.	Topic	Living quarter		Building	Household
		Housing unit	Collective living quarter		
1	**Living quarters – type of** (paras. 5.479–**5.519**)	□	♦		
2	**Location of living quarters** (paras. 5.520–5.527)	□	♦	□	□
3	**Occupancy status** (paras. 5.528–5.532)	♦			
4	**Occupancy by one or more households** (paras. 5.533–5.537)	□			♦
5	**Occupants – number of** (paras. 5.538–5.539)	♦	♦		□
6	**Ownership – type of** (paras. 5.540–5.545)	♦			
7	**Rooms – number of** (paras. 5.546–**5.548**)	♦			□
8	Bedrooms – number of (paras. 5.549–5.550)	○			○
9	Useful floor space – (paras. 5.551–5.552)	○	○		○
10	**Water supply system** (paras. 5.553–5.556)	♦	○		□
11	**Drinking water – main source of** (paras. 5.557–5.558)	♦	○		□
12	**Toilet – type of** (paras. 5.559–5.562)	♦	○		□
13	**Sewage disposal** (para. 5.563)	♦			□
14	**Solid waste disposal – main type of** (paras. 5.564–5.565)	♦			□
15	**Bathing facilities** (paras. 5.566–5.568)	♦	○		□
16	**Kitchen – availability of** (paras. 5.569–**5.572**)	♦	○		□
17	**Fuel used for cooking** (para. 5.573)	♦			□
18	**Energy used for lighting – type of** (paras. 5.574–5.575)	♦	○		□
19	Heating – type and energy used (paras. 5.576–5.578)	○			○
20	Hot water – availability of (para. 5.579)	○			○
21	Piped gas – availability of (para. 5.580)	○			○
22	Use of the housing unit (paras. 5.581–5.582)	○			○
23	**Building – type of** (paras. 5.583–5.591)			♦	
24	Year or period of construction (paras. 5.592–**5.596**)	○		○	
25	Dwellings in the building – number of (para. 5.597)	○		○	
26	Position of dwelling in the building (paras. 5.598–**5.600**)			○	
27	Accessibility to dwelling (para. 5.601)	○			
28	**Construction material of outer walls** (paras. 5.602–5.605)	♦		♦	
29	Construction material of floor and roof (para. 5.606)	○		○	
30	State of repair (paras. 5.607–5.608)	○		○	
31	Elevator – availability of (paras. 5.609–5.610)	○		○	
32	Farm building (para. 5.611)	○		○	
33	**Age and sex of the reference person of the household** (paras. 5.612–5.613)				♦
34	**Tenure** (paras. 5.614–5.617)				♦
35	Rental and housing costs (paras. 5.618–5.620)				○
36	Furnished/unfurnished (para. 5.621)	○			○

No.	Topic	Living quarter		Building	Household
		Housing unit	Collective living quarter		
37	ICT devices – availability of (paras. 5.622–5.626)				○
38	Cars – number of available (para. 5.627)				○
39	Durable household appliances – availability of (para. 5.628)				○
40	Outdoor space – access to (para. 5.629)				○
	Legend: ♦ - Core topic □ - Core topic, derived ○ - Additional topic				

III. Definitions and specifications of topics

5.478. Paragraphs 5.479–5.629 below contain the recommended definitions and specifications of the housing topics. It is important that census data be accompanied by the definitions used in carrying out the census. It is also important that any changes in definitions that might have been made since the previous census be indicated and, if possible, accompanied by estimates of the effect of such changes on the relevant data. In this way, users will not confuse valid changes over time with increases or decreases that have occurred as the result of changed definitions.

1. Living quarters – type of (core topic)

i. Definition of living quarters

5.479. *Living quarters* are structurally separate and independent places of abode. They (a) may have been constructed, built, converted or arranged for human habitation, provided that they are not at the time of the census used wholly for other purposes and that, in the case of non-conventional housing units and collective living quarters, they are occupied at the time of the census; or (b) though not intended for habitation, were in use for such a purpose at the time of the census.

5.480. In a census with a field enumeration, instructions should be issued to field staff so that it is clearly understood at what stage of completion living quarters should be in order to be included. Living quarters being demolished or awaiting demolition should normally be excluded. The definitions and criteria used in the field enumeration should align with those employed for the system of current housing statistics to avoid double counting, especially when construction statistics are used to update census data. Special instructions will need to be issued concerning “core dwellings” in countries where these are provided within a preliminary phase of dwelling construction (see paragraphs 5.496–5.499).

ii. Classification of living quarters

5.481. Living quarters are either conventional dwellings, other occupied housing units or collective living quarters. Normally, the collection of information concerning buildings and housing units located within those buildings is of prime importance in a housing census, since it is in such units that the majority of the population permanently lives. Furthermore, housing units are intended for occupancy, or are occupied, by households, and it is with the provision of accommodation for households that housing programmes and policies are mainly concerned. However, certain types of “collective living quarters”

are also of significance with respect to the housing conditions of households; these include hotels, rooming houses and other lodging houses and camps occupied by households. Housing units should be classified so as to distinguish conventional dwellings from other types of housing units. It should be emphasized that without an adequate classification of living quarters, no meaningful analysis of housing conditions based on housing census data is possible.

5.482. The classification outlined below (see also Figure 5.3) and a system of three-digit codes have been designed to group in broad classes housing units and collective living quarters with similar structural characteristics. The distribution of occupants (population) among the various groups provides valuable information about the housing accommodation available at the time of the census. The classification also affords a useful basis of stratification for sample surveys. The living quarters may be classified into the following categories:

1 Housing units
- 1.1 Conventional dwellings
  - 1.1.1 Has all basic facilities
  - 1.1.2 Does not have all basic facilities
- 1.2 Other housing units
  - 1.2.1 Semi-permanent housing units
  - 1.2.2 Mobile housing units
  - 1.2.3 Informal housing units
  - 1.2.4 Housing units in permanent buildings not intended for human habitation
  - 1.2.5 Other premises not intended for human habitation

2 Collective living quarters
- 2.1 Hotels, rooming houses and other lodging houses
- 2.2 Institutions
  - 2.2.1 Hospitals
  - 2.2.2 Correctional institutions (prisons, penitentiaries)
  - 2.2.3 Military institutions
  - 2.2.4 Religious institutions (for example, monasteries and convents)
  - 2.2.5 Retirement homes, homes for older persons
  - 2.2.6 Student dormitories
  - 2.2.7 Staff quarters (for example, hostels and nurses' homes)
  - 2.2.8 Orphanages
  - 2.2.9 Other institutional places
- 2.3 Camps and workers' quarters
  - 2.3.1 Military camps
  - 2.3.2 Worker camps
  - 2.3.3 Refugee camps
  - 2.3.4 Camps for internally displaced people (IDPs)
  - 2.3.5 Other camps and worker's quarters
- 2.4 Other collective living quarters

5.483. Not all the categories in the above classification are of importance under all circumstances. For example, in some countries certain categories may not need to be considered separately, while in others it may be convenient to subdivide them. However, some of the categories are of special significance for assessing the housing situation and should be distinguished even where a simplified classification is employed. The distinction between conventional and informal housing units is referred to particularly.

iii. Definitions of each type of living quarters

5.484. A description of the categories listed in at paragraph 5.482 follows.

1. Housing units

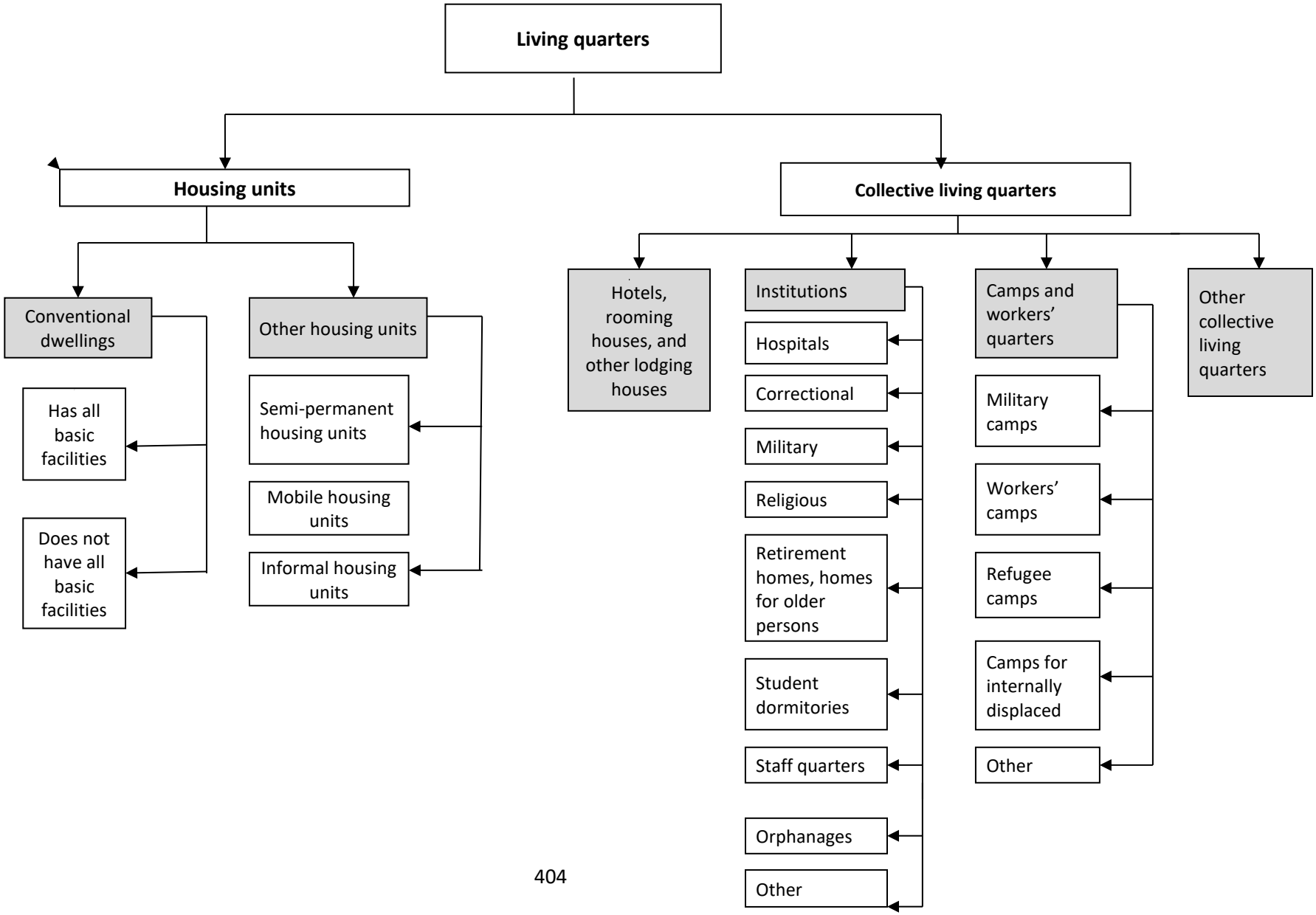
5.485. A *housing unit* is a separate and independent place of abode intended for habitation by a single household,[296] or one not intended for habitation but occupied as living quarters by a household at the time of the census. Thus, it may be an occupied or vacant dwelling, an occupied non-conventional housing unit or any other place occupied as living quarters by a household at the time of the census. This category includes housing of various levels of permanency and acceptability and therefore requires further classification in order to provide for a meaningful assessment of housing conditions.

5.486. The essential features of housing units are separateness and independence. An enclosure may be considered separate if surrounded by walls, fences, or other barriers, and whether or not covered by a roof so that a person or group of persons can isolate themselves from other persons in the community for the purposes of sleeping, preparing and taking their meals, and protecting themselves from the hazards of climate and environment. Such an enclosure may be considered independent when it has direct access from the street or from a public or communal staircase, passage, gallery or grounds, in other words, when the occupants can come in and go out of their living quarters without passing through anybody else's premises.

5.487. Attached rooms having an independent entrance, or detached rooms for habitation that clearly have been built or rebuilt or converted for use as part of living quarters, should be counted as part of the housing units. Thus, housing units may comprise rooms or groups of rooms with independent entrances, or separate buildings.

[296] Although intended for habitation by one household, a housing unit may, at the time of the census, be occupied by one or more households or by a part of a household.

Figure 5.3. Classification of living quarters



5.488. It should be noted that any housing units located on the grounds, or within the buildings, of an institution or camp should be separately identified and counted as such. For example, if, within the grounds of a hospital, there is a separate and independent house intended for the habitation of the director and his/her family, the house should be counted as a housing unit. In the same way, self-contained apartments located within hotel buildings should be counted as housing units if they have direct access to the street or to a common space within the building. Similar cases will need to be identified and described in the instructions for the enumeration.

1.1 Conventional dwellings

5.489. A "conventional dwelling" is a room or suite of rooms and its amenities in a permanent building or structurally separated part thereof which, by the way it has been built, rebuilt or converted, is intended for habitation by one household and is not, at the time of the census, used wholly for other purposes. It should have a separate access to a street (direct or via a garden or grounds) or to a common space within the building (staircase, passage, gallery and so on). Therefore, there are four essential features of a conventional dwelling:

(a) It is a room or suite of rooms;
(b) It is located in a permanent building;
(c) It has separate access to a street or to a common space;
(d) It is intended to be occupied by one household.

5.490. Examples of conventional dwellings are houses, flats, suites of rooms and apartments. Although a conventional dwelling is a housing unit originally intended, that is to say, constructed or converted, for habitation by one household, it may, at the time of the census, be vacant or occupied by one or more households. It may be noted that the terms 'dwelling', 'dwelling unit', 'dwelling house', 'residential dwelling unit', 'family dwelling', 'house', '*logement*', '*vivienda*', '*unidad de vivienda*' etc. have been used indiscriminately to refer to housing units of any type. The referent of the term "dwelling" is here limited to a housing unit located in a permanent building and designed for occupancy by one household.

5.491. A "permanent building" is understood to be a structure that is not intended to be moved and that may be expected to maintain its stability for 15 years or more, depending on the way countries define durability. It is recognized that the criterion of permanency or durability (particularly with respect to the significance of materials and methods of construction) may be difficult to apply either in the field or from information held in administrative records and that its adaptation to local conditions would require considerable study and experimentation by the NSO. In some cases, it may be of greater significance nationally to apply the criteria of construction materials and methods of construction directly in order to establish whether or not the building containing the housing unit is of permanent construction, rather than to translate these criteria into a time period.

1.1.1. Conventional dwelling – has all basic facilities

5.492. A conventional dwelling with all basic facilities refers to a unit that meets all the needs of the household, such as protection from elements, cooking and maintaining hygiene. Thus, in addition to the four essential features of a conventional dwelling described in paragraph 5.489, all of the following facilities must be available for a dwelling to fall in this category:

(a) Piped water within the dwelling;
(b) Toilet within the dwelling;
(c) Fixed bath or shower within the dwelling;
(d) Kitchen or other space for cooking within the dwelling.

1.1.2. Conventional dwelling – does not have all basic facilities

5.493. The conventional dwellings that fall in this category are dwellings that have the essential features of a conventional dwelling (see paragraph 5.489) and some, but not all, of the basic facilities described in paragraph 5.492.

5.494. With increased urbanization, the need for building low-cost housing units within the city limit has been increasing in many countries. This housing most frequently consists of buildings containing a number of separate rooms whose occupants may share some or all facilities (bathing, toilet or cooking facilities). Such units do not meet all the criteria of a conventional dwelling with all basic facilities available within the dwelling, especially from the point of view of maintaining health standards and privacy. For example, these units are known as *casa de palomar* in some countries of Latin America.

1.2 Other housing units

1.2.1. Semi-permanent housing unit

5.495. The term “semi-permanent housing unit” refers to a structure that, by the way it has been built, is not expected to maintain its durability for as long a period of time as a conventional dwelling, but has some of the main features and facilities of a conventional dwelling. As noted at paragraph 5.491, durability needs to be specifically defined on the basis of national standards and practices. The number of these units in some countries may be substantial. Semi-permanent housing is not to be confused with informal housing units (see paragraph 5.502 below).

5.496. Many countries with insufficient resources to fully meet their housing needs have implemented programmes to improve the living conditions of populations in informal/squatter settlements or inadequate housing. These programmes often involve providing "core" or "starter" housing units. These core units typically consist of a small, basic structure (with walls, a roof, and a floor) that can be expanded over time as the household's needs and resources grow. This approach aims to provide a more stable and secure foundation for housing, allowing residents to gradually improve their living conditions.

5.497. A core dwelling is sometimes only a sanitary unit containing bathing and toilet facilities, to which may be added, in subsequent phases, the other elements that will finally make up the completed dwelling. Such units do not fall within the definition of a conventional dwelling. However, although the household obviously continues to occupy its original shelter (which would probably be classified as an “improvised housing unit”), its housing situation is a vast improvement over that of households remaining in the squatter areas and the provision of the cores is a significant step towards the alleviation of housing shortages.

5.498. The problem is thus one of reflecting in the statistics the improvements brought about by programmes such as those described above without distorting the data that refer to fully constructed conventional dwellings. It is recommended, therefore, that core dwellings should be counted as dwellings

in the census if at least one room[297] in addition to the sanitary facilities, is completed, and also that those dwellings that have not reached this stage of completion should be recorded as cores. Arrangements should be made so that the facilities available in the core dwelling can be related during data processing to the households for whose use they have been provided.

5.499. In other countries, the population may have developed, over time, a traditional and typical type of housing unit that does not have all the characteristics of conventional dwellings but is nevertheless considered suitable from the point of view of climate and tradition. This is especially the case in many tropical and subtropical rural areas where housing units have been constructed or built with locally available raw materials such as bamboo, palm, straw or any similar materials. Such units often have mud walls and thatched roofs, and may be expected to last only for a limited time (from a few months to several years), although occasionally they may last for longer periods. This category is intended to cover housing units that are typical and traditional in many tropical rural areas. Such units may be known, for example, as cabins, *ranchos* or *bohíos* (Latin America), *barastis* (Bahrain), or *bahay kubo* (the Philippines).

1.2.2. Mobile housing units

5.500. A "mobile housing unit" is any type of living accommodation that has been produced to be transported (such as a tent) or is a moving unit (such as a ship, boat, barge, vessel, railroad car, caravan, trailer or yacht) occupied as living quarters at the time of the census. Trailers and tents used as permanent living quarters may be of special interest in some countries.

5.501. Although mobile housing units are significantly different from other housing units in that they can be readily moved or transported, mobility in itself is not necessarily an indicator of low quality. For the assessment of housing conditions in countries with a substantial number of mobile units, it may be useful to classify them further as tents, wagons, boats, trailers or other unit types.

1.2.3 Informal housing units

5.502. The term "informal housing unit" refers to a unit that does not have many of the features of a conventional dwelling and is generally characterized as unfit for human habitation, but that is used for that purpose at the time of the census. Therefore, it is neither a permanent structure nor one equipped with any of the essential facilities. The definition of informal housing varies significantly across countries. Many of its characteristics are not directly related to the physical structure of the unit itself. Additional factors, beyond the lack of public services, should be considered. Depending on national circumstances, NSOs should develop detailed instructions to distinguish between informal and semi-temporary housing units. When collecting information on the type of housing unit, it is important to specify whether the data should be recorded by the householder or the enumerator. This will ensure more accurate and consistent reporting.

5.503. Informal housing units comprise three subgroups: "improvised housing units", "housing units in permanent buildings not intended for human habitation" and "other informal housing units". These units are characterized by the fact that they are either makeshift shelters constructed of waste materials and

[297] For the definition of "room", see paragraph 5.546.

generally considered unfit for habitation (squatters' huts, for example) or places that are not intended for human habitation although in use for that purpose at the time of the census (barns, warehouses, natural shelters and so on). Under almost all circumstances, such places of abode represent unacceptable housing and they may be usefully grouped together in order to analyse the housing conditions of the population and to estimate their housing needs. Each subgroup is defined below.

1.2.3.1 Improvised housing units

5.504. An improvised housing unit is an independent, makeshift shelter or structure, built of waste materials and without a predetermined plan for the purpose of habitation by one household, which is being used as living quarters at the time of the census. Included in this category are squatters' huts, *poblaciones callampas* (Chile), *hongos* (Peru), *favelas* (Brazil), *sarifas* (Iraq), *barong barong* (the Philippines) and any similar premises arranged and used as living quarters, though they do not comply with generally accepted standards for habitation, and do not have many of the characteristics of conventional dwellings. This type of housing unit is usually found in urban and suburban areas, particularly at the peripheries of the principal cities.

5.505. There is a wide variation in the procedures and criteria used in classifying these units. There are many borderline cases, and NSOs will need to make decisions and issue detailed instruction on how to enumerate and classify such housing units.

1.2.3.2 Housing units in permanent buildings not intended for human habitation

5.506. Included in this category are housing units (in permanent buildings) that have not been designed, constructed, converted or arranged for human habitation but that are actually in use as living quarters at the time of the census. These include housing units in stables, barns, mills, garages, warehouses, offices, and other non-residential buildings.

5.507. This category also may cover units and their occupants in buildings initially built for human habitation but later abandoned with all services cut because of deterioration. These dilapidated buildings can be found, especially in large cities, still standing, although marked for demolition. They should be included in this category if inhabited thereby providing an exception to the general recommendation at paragraph 5.480 that living quarters being demolished or awaiting demolition should normally be excluded from the census.

5.508. Premises that have been converted for human habitation, although not initially designed or constructed for this purpose, should not be included in this category, but classified as "other informal housing units".

1.2.3.3 Other informal housing units

5.509. This category refers to living quarters that are not intended for human habitation or located in permanent buildings but that are nevertheless being used as living quarters at the time of the census. Caves and other natural shelters fall within this category.

2. Collective living quarters

5.510. *Collective living quarters* include structurally separate and independent places of abode intended for habitation by large groups of individuals or several households and occupied at the time of the census. Such quarters usually have certain common facilities, such as cooking and toilet installations, baths, lounge rooms or dormitories, which are shared by the occupants. They may be further classified into hotels, rooming houses and other lodging houses, institutions and camps.

5.511. Housing units located on the grounds, or within the buildings, of an institution, camp, hotel and so forth should be separately identified and counted as such.

5.512. The criteria established for the identification of collective living quarters are not always easy to apply and it is sometimes difficult for an enumerator to decide whether living quarters should be classified as a housing unit or not. This is particularly true in the case of a building occupied by a number of households. Enumerators should be given clear instructions as to whether the premises occupied by a group of people living together are to be considered a housing unit or as collective living quarters. This may be less of a problem where census information is collected from administrative data sources and where such buildings are registered as being for communal living. The several types of collective living quarters are listed at paragraph 5.482 and shown Figure 5.3, and described as follows:

2.1 Hotels, rooming houses and other lodging houses

5.513. This group comprises permanent structures that provide lodging on a fee basis and in which the number of borders or lodgers exceeds five.[298] Where there are less than five, the living quarters should be classified as a housing unit. Hotels, motels, inns, boarding houses, pensions, lodging houses and similar structures fall within this category. If there is any accommodation within a hotel or similar establishment that is occupied by a household and which fulfils the requirement of a conventional dwelling it should be classified as such.

2.2 Institutions

5.514. This group covers any set of premises in a permanent structure or structures designed to house (usually large) groups of persons who are bound by either a common public objective or a common personal interest. Such sets of living quarters usually have certain common facilities shared by the occupants (for example baths, lounges and dormitories). Hospitals, military barracks, boarding schools, convents and prisons fall within this category (see the categories in paragraph 5.482).

5.515. It may be useful, depending on national needs, to require that an institution be used as the principle usual residence of at least one person at the time of the census.

2.3 Camps

[298] The threshold of five lodgers is the one most used. However, depending on national circumstances, this number might be adjusted accordingly.

5.516. Camps are sets of premises originally intended for the temporary accommodation of persons with common activities or interests. Included in this category are military camps, refugee camps and camps established for the housing of workers in mining, agriculture, public work, seasonal works, other types of enterprises, or victims of disasters such as earthquake and floods. While some refugee and IDP camps meet the criteria for collective living quarters, others do not. These camps often lack collective facilities for cooking or sleeping, distinguishing them from more traditional collective living arrangements such as military quarters or workers' accommodation. Many refugee and IDP camps function more like cities or villages, with families and individuals living in semi-permanent, mobile, or informal housing units. These camps often resemble conventional residential areas rather than more conventional collective living quarters. NSOs should consider these exceptions when classifying refugee and IDP camps.

2.4 Other collective living quarters

5.517. This is a residual category for collective living quarters that may not conform to the definitions of those included in groups 2.1-2.3 above. It should be used only when the number of units in question is small. Where the number is substantial, additional groups of living quarters that have common characteristics and that are of significance for an improved appraisal of housing conditions should be established.

5.518. In some countries, it seems that certain types of multi-household living quarters have emerged in response to the particular circumstances of the population, and that the characteristics of these quarters enable them to be readily identified in the field. It may be useful in these countries to provide a separate subgroup for any such special types.

5.519. It should be stressed that the types of living quarters to be included in this category are those intended for communal habitation by several households, that is to say, constructed or converted for this purpose. Such units intended for occupancy by one household, but those at the time of the census are occupied by several households, are not to be included as collective living quarters because this obscures the identification of households doubling up in dwellings (an important element in estimating housing needs). It is suggested that, in carrying out the census, a strict distinction be maintained between a housing unit occupied by more than one household and living quarters constructed or converted for communal habitation by several households.

2. Location of living quarters (core topic)

5.520. The concepts of "locality" and "urban and rural" (see paragraphs 5.93-5.103) provide the basis for recording valuable information for locating living quarters. Understanding these geographic concepts is necessary for housing censuses, impacting on both the collection and dissemination of data. When a housing census is combined with, or closely linked to, a population census, careful coordination is essential. The geographic areas used in both censuses should be defined in order to optimize the value of each operation. For countries conducting separate population and housing censuses, ensuring consistent recording of geographic location for dwellings and resident households is critical. However, in countries with a single data collection operation, this information need only be recorded once.

5.521. Information on location should be collected in sufficient detail to enable outputs to be produced for the smallest geographic subdivisions required by the tabulation plan. To satisfy the requirements of the recommended geographic classifications, information is needed on: whether the living quarters are

located in an urban or rural area; the major and minor civil division; and, for living quarters located in principal localities, the name of the locality.

5.522. Where a permanent system of house or building numbers does not already exist, it will be necessary for the NSO to establish a numbering system so that the location of each set of living quarters can be adequately described. Similarly, in cases where streets do not have names or numbers properly displayed, such identification should be provided as one of the pre-census field operations. Adequate identification provides the basis for the preparation of census control lists (see also 3.150–3.156); it is required in order to monitor and control the enumeration, and to identify living quarters for possible callbacks and post-enumeration evaluation surveys (PES) as well as for other post-censal enquiries that use the census as a sampling frame or other point of departure. Ideally, each building or other inhabited structure should be provided with a number, as should each set of living quarters within buildings or structures. In preparing a census control listing, it is the practice to identify further each household within the living quarters. In register-based censuses, it is very important that the frame used is not subject to any under-coverage or over-coverage. Therefore, periodic quality controls must be carried out.

5.523. Living quarters that are not located in areas with a conventional pattern of streets, such as those in squatter areas or in some places not intended for habitation, may require special identification. Since it may not be possible to describe the location of these units by use of a formal address, it may be necessary to describe them in terms of their proximity to natural or created landmarks of various kinds or in relation to buildings that are located in areas where a formal address exists. If possible, coding "slums" separately is of critical importance for countries with prominent slum populations.

5.524. The various geographic designations that together define the location of living quarters are discussed below.

i. Address

5.525. Information that describes the place where the living quarters are to be found and distinguishes them from other living quarters in the same locality falls within this category. As a rule, the information includes the name or number of the street and the number of the living quarters; in the case of apartments, the building number and the apartment number are required.

ii. Locality

5.526. For the definition of “locality”, see paragraphs 5.93–5.96.

iii. Urban and rural

5.527. For the definition of “urban and rural”, see paragraphs 5.97–5.103.

3. Occupancy status (core topic)

5.528. *Occupancy status* refers to whether or not a conventional dwelling is occupied at the time of the census. For those dwellings not occupied (because they are vacant or in secondary use), the reason for not being occupied should be classified.

5.529. Information should be obtained for each conventional dwelling to show whether the dwelling is occupied or vacant at the time of the census. For vacant units intended for year-round occupancy, the reason for the vacancy (for rent or for sale, for example) should be reported if this is known. Occupancy status applies only to conventional dwellings, since all other types of living quarters are required by definition to be occupied in order to be enumerated in the census.

5.530. The enumeration of vacant conventional dwellings is likely to pose difficult problems, but at least a total count should be made for purposes of controlling the enumeration. The type of vacancy is frequently indicated by "for sale" or "for rent" signs posted on the dwelling. Although it may not be feasible to investigate all of the topics included in the census for vacant units, as much information as possible should be collected, including information on whether the living quarters are vacant seasonally or non-seasonally.

5.531. Vacant units intended for seasonal or secondary occupancy may represent a substantial proportion of the housing stock in resort areas and in areas where large numbers of seasonal workers are employed. The separate identification of such categories may be necessary for the correct interpretation of the overall vacancy rate, as well as for an evaluation of the housing situation in the area concerned.

5.532. Whether living quarters whose occupants are temporarily absent or temporarily present are to be recorded as occupied or vacant will need to be considered in relation to whether a *de jure* or *de facto* population census is being carried out. In either case, it would seem useful to distinguish as far as possible conventional dwellings that are used as a second residence. This is particularly important if the second residence has markedly different characteristics from the primary residence, as is the case, for example, when agricultural households move during certain seasons of the year from their permanent living quarters in a village to rudimentary structures located on agricultural holdings. It may be the case that a dwelling is occupied by persons who are not included in the census - such as diplomatic staff. It may be important in some countries to identify such accommodation separately as a distinctly different component of the housing stock. The recommended classification of occupancy status for conventional dwellings is as follows:

1 Occupied
2 Vacant / not occupied
- 2.1 Seasonally vacant
  - 2.1.1 Holiday homes
  - 2.1.2 Seasonal workers' quarters
  - 2.1.3 Other
- 2.2 Secondary residences
- 2.3 For rent/sale
- 2.4 For demolition
- 2.5 Other

4. Occupancy by one or more households (core topic)

5.533. For the purpose of a housing census, each household must be identified separately. With respect to housing programmes, the use of the separate concepts of "household" and "living quarters" in carrying out housing censuses permits the identification of the persons or groups of persons in need of their own dwellings. If the household is defined as a group of persons occupying a set of living quarters, the number of households in the living quarters and the number of sets of occupied living quarters will always be equal

(see the definition of the 'household dwelling' concept at paragraphs 2.42 and 5.134) and there will be no apparent housing need as reflected by the number of "sharing" households that require their own living quarters. Conversely, if living quarters are defined as the space occupied by a household, the number of households in living quarters will again be equal to the number of sets of living quarters, with the added disadvantage that there will be no record of the number of structurally separate living quarters.

5.534. Occupancy by more than one household is a useful topic for assessing the current housing situation and measuring the need for additional housing. For countries relying on the housekeeping concept, the number of households occupying a housing unit is needed to understand the extent of shared housing.

5.535. In countries where it is traditional to count families, the family in the broad sense of the term may be adopted as an additional unit of enumeration; in the great majority of cases the composition of this unit will coincide with that of the household.

5.536. A household and family should be defined in the same way for housing census purposes as for population censuses (see paragraphs 5.133–5.140 and 5.153–5.160).

5.537. For the definitions of "household", "reference person of household" and "persons living in institutions", see paragraphs 5.133–5.160 and 2.48–2.54.

5. Occupants – number of (core topic)

5.538. Each person usually resident in a housing unit or in collective living quarters should be counted as an occupant. Therefore, the unit of enumeration for this topic is the living quarters. However, since housing censuses are usually carried out simultaneously with population censuses, the applicability of this definition depends upon whether the information collected and recorded for each person in the population census indicates where he/she was present on the day of the census or whether it refers to the usual residence (see paragraphs 5.51–5.62). Care should be exercised in distinguishing persons occupying mobile units, such as boats, caravans and trailers, as living quarters from persons using these units merely as a means of transportation.

5.539. Depending on the national requirements for information (such as the need to measure levels of overcrowding), some NSOs may wish to distinguish between those occupants that are usually resident and those that are not usually resident in the living quarters for the purposes of better understanding the housing conditions and living arrangements of non-residents.

6. Ownership – type of (core topic)

5.540. This topic refers to the type of ownership of the housing unit itself and not of that of the land on which it stands. Type of ownership should not be confused with tenure, which is a characteristic of the household and is covered in paragraphs 5.614–5.617.

5.541. Information should be obtained to show:

(a) Whether the housing unit is owned by the public sector (central government, local government, public corporations);

(b) Whether the housing unit is privately owned (by households, private corporations, cooperatives, housing associations and so on).

The question is sometimes expanded to show whether the housing units are fully paid for, being purchased in instalments or mortgaged. The classification of housing units by type of ownership is as follows:

1 Owner-occupied
2 Non-owner-occupied
- 2.1 Publicly owned
- 2.2 Privately owned
- 2.3 Communally owned
- 2.4 Cooperatively owned
- 2.5 Other

3 Othe types of ownership

5.542. Housing units are defined as owner-occupied if used wholly or partly for own occupation by the owner. In principle, if a housing unit is being purchased in instalments or mortgaged according to national legal systems and practices, it should be enumerated as being owned. Instructions should also cover other arrangements, such as housing units in cooperatives or housing associations.

5.543. The information on ownership may be classified, as a minimum, into two main groups, namely "private ownership" and "other ownership". Depending upon the prevalence of various types of ownership and their significance with respect to housing conditions and the formulation of housing programmes, it may be useful to further classify the category "other ownership" into the relevant examples of the subgroups shown. The categories used should be consistent with those employed in the system of national accounts of the country concerned and in accordance with the recommendations contained in the *System of National Accounts, 2008*.[299]

5.544. It has been observed that the collection of information on type of ownership in a census may be hampered by the fact that the occupants might not know who the owner of the property is. Furthermore, there are numerous cases of borderline and co-ownership, which make the topic difficult for nationwide enumeration. This is one of the topics for which more accurate information might be obtained through a housing survey or administrative records in registers such as land use and property gazetteers which usually record ownership.

5.545. In countries where there is a substantial amount of employer-issued housing, it would be useful to include the subcategories "issued by the employer" and "not issued by the employer" under the category "privately owned" (or publicly owned where the employer is a public sector entity). It is important that such information be known from the point of view of assessing the impact of job loss, in order to gauge the magnitude of the population whose housing circumstances are dependent on their employment.

[299] *System of National Accounts 2008*, United Nations publication, Sales No. E.08.XVII.29.

7. Rooms – number of (core topic)

5.546. A *room* is defined as a space in a housing unit enclosed by walls reaching from the floor to the ceiling or roof covering, or to a height of at least 2 metres, of an area large enough to hold a bed for an adult, that is, at least 4 square metres. The total number of rooms usually includes bedrooms, dining rooms, living rooms, studies, habitable attics, domestic workers' rooms, kitchens, rooms used for professional or business purposes, and other separate spaces used or intended for dwelling purposes, so long as they meet the criteria concerning walls and floor space. Passageways, verandas, lobbies, bathrooms and toilet rooms should not be counted as rooms, even if they meet the criteria. Separate information may be collected for national purposes on spaces of less than 4 square metres that conform in other respects to the definition of "room" if it is considered that their number warrants such a procedure.

5.547. Rooms used exclusively for business or professional purposes should be counted separately, as it is desirable to include them when calculating the number of rooms in a dwelling but to exclude them when calculating the number of persons per room. This procedure allows density levels to be studied according to the number of rooms available for living purposes in relation to the number of occupants. In any event, the NSO should indicate the procedure that has been followed in disseminating the results of the census.

5.548. It is recommended that kitchens be included in the count of rooms provided they meet the criteria concerning walls and floor space. Kitchens or kitchenettes that have an area smaller than 4 square metres or that have other characteristics that disqualify them should be excluded. For national purposes, countries may wish to identify and count kitchens within a separate group that may be analysed with respect to size and utilization, and to consider separately those used exclusively for cooking (see paragraphs 5.569-5.572).

8. Bedrooms – number of

5.549. In addition to recording the total number of rooms, some NSOs may wish to collect information on the number of bedrooms in a housing unit in order to provide a measure of overcrowding.

5.550. A *bedroom* is defined as a room equipped with a bed and used for night rest and excludes makeshift and temporary sleeping quarters. Rooms where sleeping areas are partitioned with curtains or blinds or any similar makeshift arrangements cannot be counted as a bedroom. The count of bedrooms includes spare/guest bedrooms not occupied at the time of the census. As defined at paragraph 5.546, there must be some permanency in relation to the walls enclosing the bedrooms. A room used for other activities by day and sleeping by night should not be counted as bedroom.

9. Useful floor space

5.551. This topic refers to the useful and liveable floor space in housing units, that is to say, the floor space measured inside the outer walls of housing units, excluding non-habitable cellars and attics. In multiple-dwelling buildings, all common spaces should be excluded. Although information on this topic is often collected to supplement that on the number of rooms, in some countries it is the primary focus.

Floor space is generally considered a more accurate measure of housing density than the number of rooms divided by occupants, as room sizes can vary significantly.

5.552. Collecting information on the floor space available to occupants of housing units may prove to be difficult; occupants often may not know the exact or even the approximate area of the housing unit they occupy, and training enumerators to calculate the floor space would be complicated and costly, and would result in inaccuracies. In this context, and taking into account the importance of the information concerned, NSOs should take into consideration developing detailed instructions on proper procedures for collecting data on floor space.

10. Water supply system (core topic)

5.553. A basic housing characteristic to be determined in the census is whether housing units have or do not have a piped water installation, in other words, whether or not water is provided to the housing unit by pipes from a community-wide system or a private installation, such as a pressure tank or pump. The unit of enumeration for this topic is the housing unit. It is also recommended that census should collect information to indicate whether the unit has tap water inside or, if not, whether it is within a certain distance from the door. The recommended distance is 200 metres, assuming that access to piped water within that distance allows occupants of the housing unit to obtain water for household needs without being subjected to extreme efforts. In addition to the location of the tap water relative to the housing unit, the source of water available to households is also of interest. Therefore, the recommended classification of housing unit by water supply system is as follows:

1 Piped water inside the housing unit
  1.1 From the community scheme
  1.2 From an individual source
2 Piped water outside the unit but within 200 metres
  2.1 From the community scheme
    2.1.1 For exclusive use
    2.1.2 Shared
  2.2 From an individual source
    2.2.1 For exclusive use
    2.2.2 Shared
3 Other (see category 3 of the classification in paragraph 5.558 for more details)
4 No piped water

5.554. A community scheme is one that is subject to inspection and control by public authorities. Such schemes are generally operated by a public body but, in some cases, they are operated by a cooperative or private enterprise. An individual source of water refers to a source that is not part of a community scheme, such as an individual or shared water reservoir.

5.555. As noted above the unit of enumeration for this topic is the housing unit. However, some NSOs may find it useful to collect information on the availability of piped water for the use of occupants in collective living quarters. Such living quarters are usually equipped with multi-facilities for the use of large groups, and information on the water supply system in relation to the number of occupants would be significant in respect of analysing housing conditions. The water supply system in collective living quarters constitutes an additional topic.

5.556. The most significant information from a health point of view is whether or not the living quarters have piped water within the premises. However, a category may be added to distinguish cases where the piped water supply is not within the living quarters but rather within the building in which the living quarters are situated. It may also be useful to collect information that would show whether the water supply is for the sole use of the occupants of the living quarters being enumerated or for the use of the occupants of several sets of living quarters, as indicated in the above classification at the three-digit level. Where there is a large proportion of housing units with no piped water, this category may be expanded to specify sources commonly used in a country.

11. Drinking water – main source of (core topic)

5.557. Having enough water for drinking and personal hygiene is essential, but quantity by itself is not sufficient. The quality of the water is a key factor to the assessment of any health risk arising from poor housing conditions. Consequently, one of the targets of the water supply, sanitation and hygiene (WASH) post-2015 recommendations[300] proposed by the WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (JMP) is "universal access to basic drinking water, sanitation and hygiene", assessed in part by having access at home to safely managed drinking water. A safely managed drinking water service is defined as one that reliably delivers water that is sufficient to meet domestic needs and does not represent a significant risk to health. This implies a system that delivers water to the household and includes measures to prevent risks and to verify water quality through compliance monitoring. An improved water source (such as piped water, public tap or standpost, tubewell or borehole, protected dug well, protected spring, rainwater in collection tank) can be safely managed. Unimproved sources, which by definition are not safely managed, include unprotected dug well, unprotected spring and surface water from a river, stream, dam, lake, pond, canal or irrigation channel. Delivered water (for example, through trucks, carts, sachets or bottles) can potentially be safely managed, but if these are the primary drinking water sources, other improved sources of water must be accessible to the household for other domestic uses (such as washing or bathing).

5.558. Countries are encouraged to collect the information on the main source of drinking water for the household, particularly where there is a considerable difference between the sources of water for general household use and for drinking. For those countries wishing to collect this information, the following categories of main source of drinking water are recommended:

1 Piped water inside the unit
- 1.1 From the community scheme
- 1.2 From an individual source

2 Piped water outside the unit but within 200 metres
- 2.1 From the community scheme
  - 2.1.1 For exclusive use
  - 2.1.2 Shared
- 2.2 From an individual source
  - 2.2.1 For exclusive use

[300] See WHO/UNICEF Joint Monitoring Programme on Water Supply and Sanitation (JMP) report entitled *WASH post-2015: proposed targets and indicators for drinking-water, sanitation and hygiene* (April 2014), accessible at http://www.wssinfo.org/.

2.2.2 Shared

3 Other sources of water

- 3.1 Borehole/tubewell
- 3.2 Protected dug well
- 3.3 Protected spring
- 3.4 Rainwater collection tank
- 3.5 Delivered water – bottled, sachet
- 3.6 Delivered water – tanker trucks, carts
- 3.7 Unprotected dug well/spring/river/stream/lake/pond/dam/canal/irrigation channel
- 3.8 Other

12. Toilet – type of (core topic)[301]

5.559. A *toilet* may be defined as an installation for the disposal of human excreta. A flush toilet is an installation provided with piped water that permits humans to discharge their wastes and from which the wastes are flushed by water. The unit of enumeration for this topic is the housing unit.

5.560. For housing units reported as having a toilet, additional information may be sought to determine whether the toilet is used exclusively by the occupants of the living quarters being enumerated or is shared with the occupants of other living quarters. For living quarters reported as having no toilet, it would be useful to know if the occupants have the use of a communal facility and the type of facility, or if they have the use of the toilet of other living quarters and the type, or if there is no toilet of any kind available for the use of the occupants.

5.561. Some countries have found it useful to expand the classification for non-flush toilets so as to distinguish certain types that are widely used and indicate a certain level of sanitation. The recommended classification of housing unit by toilet facilities is as follows:

1 With toilet within housing unit

- 1.1 Flush/pour flush[302] toilet
- 1.2 Other

2 With toilet outside housing unit

- 2.1 For exclusive use
  - 2.1.1 Flush/pour flush toilet
  - 2.1.2 Ventilated improved pit latrine[303]
  - 2.1.3 Pit latrine without ventilation with covering
  - 2.1.4 Holes or dug pits with temporary coverings or without shelter

[301] It is also necessary to distinguish between conventional dwellings with all main facilities and other conventional dwellings.

[302] A pour flush toilet uses a water seal, but unlike a flush toilet, a pour flush toilet uses water poured by hand for flushing (no cistern is used).

[303] A ventilated improved pit latrine (VIP) is a dry pit latrine that uses a hole in the ground to collect the excreta and a squatting slab or platform that is firmly supported on all sides, easy to clean and raised above the surrounding ground level to prevent surface water from entering the pit. The platform has a squatting hole, or is fitted with a seat.

2.1.5 Other

2.2 Shared

2.2.1 Flush/pour flush toilet

2.2.2 Ventilated improved pit latrine

2.2.3 Pit latrine without ventilation with covering

2.2.4 Holes or dug pits with temporary coverings or without shelter

2.2.5 Other

3 No toilet available

3.1 Service or bucket facility (excreta manually removed)

3.2 Use of natural environment, for example, bush, river, stream

3.3 Other

5.562. Although the unit of enumeration for this topic is the housing unit, as with the water supply system noted at paragraph 5.555, some NSOs may find it useful to collect information on the availability of toilet facilities for the use of occupants in collective living quarters. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number and type of toilets in relation to the number of occupants would be significant in terms of analysing housing conditions. The availability of toilets for collective living quarters represents an additional topic.

13. Sewage disposal (core topic)

5.563. Information on toilets should be combined with that on the sewage disposal system to which they are connected in order to determine the adequacy of sanitation facilities of the housing unit. To be considered adequate sanitation, toilets or latrines have to be connected to non-clogged sewage disposal systems. The information on housing units by type of sewage disposal system may be classified as follows:

1 Empties into a piped system connected to a public sewage disposal plant

2 Empties into a piped system connected to an individual sewage disposal system (septic tank, cesspool)

3 Other – such as toilet empties into an open ditch, a pit, a river, the sea

4 No disposal system

14. Solid waste disposal – main type of (core topic)

5.564. Securing sustainable development and, in this context, the usual manner of treatment of solid waste (garbage) generated by the household, has prompted the incorporation of this topic in a number of national housing censuses.

5.565. This topic refers to the usual manner of collection and disposal of solid waste or garbage generated by occupants of the housing unit, including e-waste[304] disposal. The unit of enumeration is the housing unit. The classification of housing units by type of solid waste disposal is according to the following guidelines:

1 Solid waste collected on a regular basis by authorized collectors
2 Solid waste collected on an irregular basis by authorized collectors
3 Solid waste collected by self-appointed collectors
4 Occupants dispose of solid waste in a local dump supervised by authorities
5 Occupants dispose of solid waste in a local dump not supervised by authorities
6 Occupants burn solid waste
7 Occupants bury solid waste
8 Occupants dispose solid waste into river, sea, creek, pond
9 Occupants compost solid waste
10 Other arrangement

15. Bathing facilities (core topic)

5.566. Information should be collected on whether or not there is a fixed bath or shower installation within the premises of each set of housing units. The unit of enumeration for this topic is the housing unit. Additional information may be collected to show whether or not the facilities are for the exclusive use of the occupants of the living quarters and where there is a supply of hot water for bathing purposes or cold water only. In some countries the distinction proposed above may not be the most appropriate for national needs. It may be important, for example, to distinguish between whether, in terms of availability, there is a separate room for bathing in the living quarters, a separate room for bathing in the building, an open cubicle for bathing in the building, or a public bathhouse. The recommended classification of housing units by availability and type of bathing facilities is as follows:

1 With fixed bath or shower within housing unit
2 Without fixed bath or shower within housing unit
  2.1 Fixed bath or shower available outside housing unit
    2.1.1 For exclusive use
    2.1.2 Shared
  2.2 No fixed bath or shower available

5.567. Alternatively, and in line with the elaboration in the preceding paragraph, the following classification may be more appropriate in certain circumstances:

1 Separate room for bath or shower within the housing unit
2 No separate room for bath or shower but bathing space available within the housing unit (for example, in an open area around the well within the housing unit, in the courtyard)

[304] E-waste has become an important environmental and public health issue, not only at global scale but also in low- and middle-income countries as well. Many of Electrical and Electronic Equipments (EEE) contain hazardous chemicals and materials. For this reason, inadequacies in recycling and disposing of electronic waste (e-waste) can cause serious health and environmental pollution problems. Great care must be taken to prevent unsafe exposure in recycling operations and prevent contamination of the environment.

3 Bathing room available but outside the housing unit for exclusive use
4 Shared bathing room outside the housing unit
5 No specific bathing room available

5.568. As noted above the unit of enumeration for this topic is the housing unit. However, as with the type of toilet, some NSOs may also find it useful to collect information, as an additional enquiry, on the availability of a bath or shower for the use of occupants in collective living quarters as well. Living quarters of this type are usually equipped with multi-facilities for the use of large groups, and information on the number of fixed baths or showers in relation to the number of occupants would be significant in terms of analysing housing conditions. The number of fixed baths or showers in collective living quarters would represent an additional enquiry.

16. Kitchen – availability of (core topic)

5.569. Information should be obtained on whether the housing unit has a kitchen or some other space is set aside for cooking, such as a kitchenette, or there is no special place set aside for cooking. The unit of enumeration for this topic is the housing unit.

5.570. A *kitchen* is defined as a space that conforms in all respects to the criteria for a room, and is equipped for the preparation of the principal meals of the day and intended primarily for that purpose.

5.571. The kitchen is counted as a room if it confirms to the specifications set out in paragraph 5.546. However, any other space reserved for cooking, such as a kitchenette, should not be counted as a room even though it may be equipped for the preparation of the principal meals of the day and is intended primarily for that purpose. The collection of data on the availability of a kitchen may provide a convenient opportunity to collect information on the kind of equipment that is used for cooking, for example, a stove, hotplate or open fire, and on the availability of a kitchen sink and a space for food storage so as to prevent spoilage. The recommended classification of housing units by availability of a kitchen or other space reserved for cooking within the housing unit is as follows:

1 With kitchen within housing unit
  1.1 For exclusive use
  1.2 Shared use
2 With other space for cooking within housing unit, such as kitchenette
  2.1 For exclusive use
  2.2 Shared use
3 Without kitchen or other space for cooking within housing unit
  3.1 Kitchen or other space for cooking available outside housing unit
    3.1.1 For exclusive use
    3.1.2 Shared
  3.2 No kitchen or other space for cooking available

5.572. As noted above the unit of enumeration for this topic is the housing unit. However, some NSOs may find it useful to collect information on the availability of kitchen facilities for the use of occupants in collective living quarters. Living quarters of this type are usually equipped with shared, multi-facilities for the use of large groups. it would be necessary to establish procedures for determining the appropriate number of kitchens or kitchenettes in relation to the number of occupants. This information is significant for analyzing housing conditions and represents an additional topic to consider.

17. Fuel used for cooking (core topic)

5.573. The proportion of households using solid fuels is one of the indicators used in monitoring internationally agreed development goals. There are important linkages between households' use of solid fuel, indoor air pollution, deforestation and soil erosion and greenhouse gas emissions. The type of fuel used in cooking tasks is an important predictor of exposure to indoor air pollution. NSOs should therefore collect information on the fuel used for cooking by each housing unit. Fuel used for cooking refers to the fuel used predominantly for preparation of principal meals. If two fuels (for example, electricity and gas) are used, the one used most often should the one that is reported. The classification of fuels used for cooking depends on national circumstances and may include electricity, gas, oil, coal, firewood and animal dung. It would also be useful to collect this information for collective living quarters, especially if the number of sets of collective living quarters in the country is significant. The classification of fuel used for cooking is as follows:

1 Gas
2 Electricity
3 Liquefied petroleum gas (LPG)
4 Kerosene/paraffin (petroleum-based)
5 Oil (including vegetable oils used as fuel)
6 Coal
7 Firewood
8 Charcoal
9 Animal dung
10 Crop residues (for example, cereal straw from maize, wheat, paddy rice, rice hulls, coconut husks, groundnut shells)
11 Other

18. Energy used for lighting – type of (core topic)

5.574. Information should be collected on the main type of energy used for lighting in the housing unit, such as electricity, gas or oil lamp. If the source of energy for lighting is electricity, some NSOs may wish to collect information showing whether the electricity mainly comes from a community supply, private generating plant or some other source (industrial plant, mine and so on).

5.575. As noted above the unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of electricity for the use of occupants in collective living quarters. Such living quarters are usually equipped with multi-facilities for the use of large groups, and information on electricity would be significant in terms of analysing housing conditions. The availability of electricity in collective living quarters is defined as an additional topic. No classification is specifically recommended.

19. Heating – type and energy used

5.576. This topic refers to the type of heating of housing units and the energy used for that purpose. The unit of enumeration is the housing unit. This topic may be less relevant for a number of countries where, owing to their geographic position and climate, there is no need to provide energy for heating.

5.577. *Type of heating* refers to the kind of system used to provide heating for most of the space: it may be central heating serving all the sets of living quarters or serving a set of living quarters, or it may not be central, in which case the heating will be provided separately within the living quarters by a stove, fireplace or some other heating body.

5.578. As for the energy used for heating, it is closely related to the type of heating and refers to the predominant source of energy, such as solid fuels (coal, lignite and products of coal and lignite, wood), oils, gaseous fuels (natural or liquefied gas), electricity, solar energy or other sustainable energy sources. No classification is specifically recommended. Census reports or relevant metadata should specify how the primary energy source was selected in housing units using two or more heating types.

20. Hot water – availability of

5.579. This topic refers to the availability of hot water in housing units. Hot water denotes water heated to a certain temperature and conducted through pipes and tap to occupants. The information collected may indicate whether there is hot water available within the housing unit, or outside the living quarters for exclusive or shared use, or not at all. No classification is specifically recommended.

21. Piped gas – availability of

5.580. This topic refers to whether piped gas is available in the housing unit or not for heating and/or cooking purposes. Piped gas is usually defined as natural or manufactured gas that is distributed by pipeline and whose consumption is recorded. This topic may be irrelevant for a number of countries where there is either a lack of sources of natural gas or no developed pipeline system. No classification is specifically recommended.

22. Use of the housing unit

5.581. Use of the housing unit refers to whether the housing unit is being used wholly for habitation (residential) purposes or not. The housing unit can be used for habitation and for commercial, manufacturing or some other purposes. In a number of countries, houses are used simultaneously for more than one purpose. For example, the lower floor is used as a store or workshop, and the upper floors for habitation. This type of mixed use should still be considered a housing unit, even if part of the space is used for non-residential purposes.

5.582. The recommended classification of the use of the housing unit is as follows:

1 Used solely for habitation
2 Used for habitation and other purposes
- 2.1 Economic or business purposes
- 2.2 Other purposes

23. Building – type of (core topic)

a. Definition of building

5.583. A *building* is any independent free-standing structure comprising one or more rooms[305] or other spaces, covered by a roof and usually enclosed within external walls or dividing walls[306] that extend from the foundations to the roof. However, in tropical areas, a building may consist of a roof with supports only, that is to say, without constructed walls; in some cases, a roofless structure consisting of a space enclosed by walls may be considered a "building" (see also "compound" in paragraph 5.591).

5.584. In defining a "building", particular care should be given to differentiating this from "type of living quarters" (see paragraph 5.479). Type of living quarters refers to structures that are designed for residential habitation or are being used for residential habitation. A building could comprise a number of living quarters, a commercial premises not meant, or being used, for habitation, or a mix of the two.

5.585. A building may be used or intended for residential, commercial or industrial purposes or for the provision of services. It may therefore be a factory, shop, detached dwelling, apartment building, warehouse, garage, barn and so forth. In some exceptional cases, facilities usually provided by a set of living quarters are located in two or more separate detached structures, as when a kitchen is in a separate structure. In the case of living quarters with detached rooms, these rooms should be considered separate buildings. A building may therefore contain several sets of living quarters, as is the case of an apartment building or duplex; it may be coextensive with single detached living quarters; or it may be only part of the living quarters, as is the case, for example, for living quarters with detached rooms that are clearly intended to be used as part of a household's living quarters.

5.586. The concept of a building should be clearly defined and, in a census with a field enumeration, the instructions should indicate whether all buildings are to be listed and enumerated or only those used in whole or in part for residential purposes. Instructions should also indicate whether or not buildings under construction are to be recorded and, if so, at what stage of completion they are to be considered eligible for inclusion. Buildings being demolished or awaiting demolition should normally be excluded.

b. Classification of buildings by type

5.587. The following classification of buildings (or of living quarters) by type of building is recommended:

1.0 Residential buildings
- 1.1 Buildings containing a single housing unit
  - 1.1.1 Detached
  - 1.1.2 Attached
- 1.2 Buildings containing more than one housing unit
  - 2.2.1 Up to 2 floors
  - 2.2.2 From 3 to 4 floors
  - 2.2.3 From 5 to 10 floors
  - 2.2.4 11 floors or more
- 1.3 Buildings for persons living in institutions

[305] For the definition of "room", see paragraph 5.546.

[306] The term "dividing walls" refers to the walls of adjoining buildings that have been so constructed as to be contiguous, for example, the dividing walls of "row" houses.

1.4 Other residential buildings
2.0 Non-residential buildings

5.588. It should be noted that, for the purpose of the housing census, the above classification refers to the building in which the sets of enumerated living quarters are located and that usually it will be the living quarters, not the buildings themselves, that will be tabulated according to the classification.

5.589. Category 1.1 provides separate subgroupings for "detached" and "attached" buildings because, although many single-unit buildings (suburban homes, villas, and so forth) are detached, in some countries a substantial number may be attached (as in a row of terraced houses, for example) or semi-detached, and in such cases it may be useful to identify these separately. According to the definition of "building" in paragraph 5.589 above, a group of, for example, a row of three terraced houses is considered to be three separate buildings if their "external walls or dividing walls" extend from "the foundations to the roof". Buildings containing more than one housing unit (category 1.2) will usually be apartment buildings, but they may also be other types of buildings, for example, buildings that are structurally subdivided so as to contain more than one housing unit. Buildings under the latter category should be subdivided into the following: up to 2 floors, from 3 to 10 floors and 11 floors or more. Category 1.3, "buildings for persons living in institutions", includes hospital buildings, prisons, military establishments, and so on. On the other hand, a structurally separate housing unit (a house or apartment intended for the occupancy of staff of the institution) or one that is either within a building of the institution or detached but within the grounds, belongs in category 1.0; if the housing unit is coextensive with a building, it belongs in category 1.2.

5.590. In addition to the above, and for subsequent analysis of housing conditions, the NSO will find it useful to provide for separate identification of the special types of buildings that are characteristic of the country concerned. These can be classified as category 1.4.

c. Compound

5.591. In some countries, it may be appropriate to use the "compound" as a unit of enumeration. In these places, housing units are traditionally located within compounds, and the grouping of sets of housing units in this way has economic and social implications that need to be studied. A compound, in these circumstances, becomes a distinct unit of enumeration, on a par with a housing unit. For purposes of international comparability, a compound should be classified according to the main features and facilities it displays and classified with housing units.

24. Year or period of construction

5.592. This topic refers to the age of the building in which the living quarters are located. It is recommended that the exact year of construction be sought for buildings constructed during the decennial period immediately preceding the census. The year of construction refers to the year the building was considered complete and not the year construction started. For buildings constructed before that time, the information should be collected in terms of longer periods that will provide a useful means of assessing the age of the housing stock. Difficulty may be experienced in collecting data on this topic in a field enumeration because in some cases the occupants may not know the date of construction, particularly for older buildings. However, more accurate information is more likely to be available where NSOs use housing registers or other administrative data sources for the census.

5.593. The collection of data for single years during the most recent intercensal period can be used for the purposes of comparing census information with construction statistics to assess under-coverage of the construction statistics or of the census.

5.594. Rather than collecting single years of construction, if this is seen to be too burdensome on the respondent, longer periods of construction should be recorded. Such periods could be defined in terms of events that have some special significance in the country concerned, particularly with regard to the effect on the condition of the housing stock; examples would be the period since the Second World War; the period between the First World War and the Second World War; and the period before a major earthquake, flood or fire. Alternatively, the response ranges could be equal to intervals from one census to the next, such as ten- or five-year periods depending on the frequency of census collection. This allows for comparisons across the same periods and across censuses. To facilitate cohort analysis, the reference periods for collecting period of construction should be as consistent as possible. The specific period selected should be concomitant with the materials and construction methods used in the country and the typical lifespan of buildings.

5.595. Where parts of buildings have been constructed at different times (such as the construction of extensions or annexes), the year or period of construction should refer to the major part. Where living quarters comprise more than one building (living quarters with detached rooms, for example), the age of the building that contains the major part of the living quarters should be recorded.

5.596. In countries where a significant number of households construct their own living quarters, it may be useful to collect additional information that will distinguish the living quarters according to whether or not they were constructed by the households occupying them. The information should refer only to living quarters constructed during the preceding intercensal or decennial period, and it should be made clear in formulating the question that it refers to living quarters constructed mainly by households (with or without the help of other households in the community) and not to construction executed by enterprises on behalf of households.

25. Dwellings in the building – number of

5.597. This topic refers to the number of conventional dwellings in the building. This topic is applicable in cases where there is a possibility to have unique identifier for the building itself. If a census can establish such an identifier (building number, for example, linked to the address) then it would be viable to include this topic. This information is best collected during the listing exercise or by enumerators recording this information from their own observations.

26. Position of dwelling in the building

5.598. Some countries may want to collect information on the position of the dwelling or housing unit in the building. This information can be used as an indicator of accessibility to dwellings, possibly in conjunction with information on the accessibility to the dwellings.

5.599. The following classification of dwellings by position in the building is recommended:

1.0 Dwelling on one floor only

- 1.1 Dwelling below the ground floor
- 1.2 Dwelling on the ground floor of the building
- 1.3 Dwelling on the 1st or 2nd floor of the building

1.4 Dwelling on the 3rd or 4th floor of the building
1.5 Dwelling on the 5th floor of the building or higher
2.0 Dwellings on two or more floors
2.1 Dwelling on the ground floor of the building or below ground level
2.2 Dwelling on the 1st or 2nd floor of the building
2.3 Dwelling on the 3rd or 4th floor of the building
2.4 Dwelling on the 5th floor of the building or higher

5.600. For dwellings on two or more floors, information should be provided with reference to the lowest floor level of the dwelling.

27. Accessibility to dwelling

5.601. The following classification of accessibility to the front door of the dwelling or housing unit is recommended, based on the presence or absence of steps or stairs, ramps, and lifts/elevators:

1 Access without any aids
2 Access with steps/stairs only
3 Access with ramp only
4 Access with elevator/lift only
5 Access with steps/stairs and ramp
6 Access with steps/stairs and elevator/lift
7 Access with ramp and elevator/lift
8 Access with steps/stairs, ramp, and elevator/lift
9 Access to a disability lift (with or without other means of access)

28. Construction material of outer walls (core topic)

5.602. This topic refers to the construction material of external (outer) walls of the building in which the living quarters are located. If the walls are constructed of more than one type of material, the predominant type of material should be reported. The types of materials distinguished will depend upon the materials most frequently used in the country concerned and on their significance from the point of view of permanency or durability. The following classification of construction materials is recommended:

1. Burnt clay (bricks, blocks, panels)
2. Unburnt clay (mud, earth) - may be omitted by countries where this is not important
3. Wood
4. Bamboo, trees, grass
5. Corrugated sheets
6. Steel frame
7. Stone
8. Concrete (including in situ cast concrete, reinforced concrete, blocks, panels, etc.)
9. Prefabricated units – generally factory-constructed and brought to the site and erected
10. Mixed materials (that is, a combination of building materials)
11. Other materials (to be specified)

5.603. In some countries, the material used for the construction of roofs or of floors may be of special significance for the assessment of durability and, in such cases, it may be a particular need to collect specific information on this characteristic as well as on the material of the walls. Durability refers to the

expected lifespan of a structure under regular maintenance. A durable structure should remain habitable for a considerable period, such as 15 to 20 years. Durability depends not only on construction materials but also on adherence to building standards and regulations. Technological advancements in treating traditional materials, like bamboo, can extend their durability. To assess the quality of the national housing stock, durability can be measured by considering both the materials used and compliance with construction standards. Specific instructions for enumerators should be developed based on national building practices.

5.604. While the material of construction is a useful addition to data collected on the type of living quarters, it should not be considered a substitute for the latter topic. Wood, for example, may be the material of both a poorly constructed squatter's hut and a durable and well-constructed dwelling. In these cases, information on the type of living quarters adds significantly to the value of the census in assessing the quality of a country's housing stock.

5.605. Information on building thermal insulation can be used to assess energy efficiency and well-being. Some countries may collect data on the thermal insulation of external facades, top-story ceilings, basement ceilings, or ground-level floors. Additionally, information on green roofs and walls can be valuable for assessing both climate indicators and well-being. Some countries may choose to collect data on these features of buildings.

29. Construction material of floor and roof

5.606. In some cases the material used for the construction of roofs and floors may be of special interest and can be used to further assess the quality of dwellings. This topic refers to the material used for roof and floor (although, depending on the specific needs of a country, it may also refer to other parts of the building in which the housing unit is located, such as the frame or the foundation). Information on the predominant material only should be collected. The following classification of construction materials is recommended:

1. Tile
2. Concrete
3. Metal sheeting
4. Wood
5. Bamboo
6. Palm, straw
7. Mud
8. Plastic sheeting
9. Other materials

Where the materials used on floor and on roof differ, classification specific to each should be used.

30. State of repair

5.607. This topic refers to whether the housing unit or the building in which the housing unit is located is in need of repair and to the kind of repair needed. The following classification is recommended:

1. Repair not needed
2. In need of repair
    1.2 Minor repair
    1.3 Moderate repair

1.4 Serious repair
2 Irreparable

5.608. Minor repairs refer mostly to the regular maintenance of the building and its component housing units, such as repair of a cracked window. Moderate repairs refer to the correcting of moderate defects such as missing gutters on the roof, large areas of broken plaster or stairways with no secure handrails. Serious repairs are needed in the case of serious structural defects of the building, such as missing shingles or tiles on the roof, cracks and holes in the exterior walls, and missing stairways. The term "irreparable" refers to buildings that are beyond repair, that is to say, with so many serious structural defects that it is deemed more appropriate to demolish the building than to undertake repairs; most usually this term is used for buildings with only the frame left standing or without complete external walls or roof.

31. Elevator – availability of

5.609. This topic refers to the availability of an elevator (or lift) in a multi-storey building (categories 2.2.3–2.2.4 of the classification of type of buildings). It is recommended that the information should be collected on the availability of an elevator that is operational for most of the time, subject to regular maintenance.

5.610. This topic can be useful for providing further information for indicating the accessibility to the building or the housing unit. This is of particular relevance for older persons and persons with disabilities. In this context it could also be useful to collect information on the size of the lift (for the handicapped persons and ambulance transport), if the lift goes to the ground floor, and whether or not the lift stops on the same floor as the dwelling.

32. Farm building

5.611. Some national censuses may collect information to determine if a building or dwelling is located on a farm. A farm building may be considered as being one that is part of an agricultural holding whether it is residential or not, that is, whether it is used for agricultural or housing purposes. All the information that is relevant to other buildings and dwellings should also be collected.

33. Age and sex of the reference person of household (core topic)

5.612. From among the topics recommended for inclusion in the population census, age and sex of the household reference person (HRP) is considered as being of most significance in relation to housing conditions. For the housing census, the data usually relate only to the housing units or buildings in which the housing units are located, but some of the characteristics of households that live in the housing units are also relevant to the housing condition can usefully be presented cross-classified by the age and sex of the HRP.

5.613. While this information will usually be collected in a country's population censuses and, if the population and housing censuses are conducted simultaneously, as is the practice in the majority countries, then information on age of the HRP will have already been collected together with other relevant demographic characteristics in the population part of the census. If, however, the housing census is collected independently of the population census, then there should be a separate provision for collecting this information.

34. Tenure (core topic)

5.614. Tenure refers to the arrangements under which the household occupies all or part of a housing unit. The unit of enumeration is the household occupying the housing unit. The classification of households by tenure is as follows:

1 Household owns the housing unit
2 Household rents all or a part of the housing unit
    2.1 As a main tenant
    2.2 As a subtenant
3 Household occupies the housing unit partly free of rent
4 Household occupies the housing unit wholly free of rent
5 Household occupies the housing unit under some other arrangement

5.615. National circumstances can dictate the need to assess the number of households occupying the housing unit free of rent to further distinguish whether such arrangement is with or without the consent of the owner. However, this information regarding the consent of the owner is subject to special scrutiny in terms of reliability. Furthermore, in countries where communal ownership is significantly represented, this topic on tenure might be further expanded in order to capture tenure arrangements of communally owned housing. Likewise, the category “other arrangements” can be extended to capture forms of tenure specific to some countries.

5.616. The information on tenure collected in the census needs to be clearly distinguished from the information on ownership (see paragraphs 5.540–5.545) and is one that should be asked of all households, otherwise there is a danger that it may be omitted in cases where more than one household occupies a single housing unit. Under some circumstances, it may be useful to identify separately those households that, although not subtenants in the sense that they rent from another occupant who is a main tenant or owner-occupant, rent part of a housing unit from a landlord who lives elsewhere. These households and subtenant households may be of special significance in formulating housing programmes. However, in countries where subtenancy is not usual, information on subtenants need not be collected in the census or, if collected, may be tabulated only for selected areas.

5.617. In countries where the land and the living quarters are frequently occupied under separate tenure, the topic may be expanded to show separate information for the tenure under which the household occupies the living quarters and for the tenure of the land upon which those living quarters are located.

35. Rental and housing costs

5.618. Rent is the amount paid periodically (for example weekly, monthly, quarterly, annually) for the space occupied by a household. Information may be collected on the basis of a scale of rents rather than on that of the exact amount paid. The data may be related either to household characteristics or to the characteristics of the living quarters. In the latter case, where more than one household occupies a single set of living quarters, the rents paid by all the households will need to be summed in order to obtain the total rent for the living quarters. In the case of living quarters that are partly rented and partly owner-occupied, it may be necessary to impute the rent for the owner-occupied portion.

5.619. In countries where rent for the housing unit is paid separately from rent for the land upon which the housing unit stands, separate information may need to be collected reflecting the amount of ground rent paid.

5.620. In addition to the amount of rent paid by renting households, it may be useful to collect information on the housing costs, which could include information on monthly mortgage payments (for owner-occupiers), taxes and cost of utilities.

36. Furnished/unfurnished

5.621. Countries may wish to make some provision for indicating whether the housing units covered by the rent are furnished or unfurnished and whether utilities such as gas, electricity, heat and water are included.

37. Information and communication technology – availability of (at the household level)

5.622. As noted at paragraph 5.334, information and communication technology (ICT) devices play an increasingly crucial role in modern society. Modern technology, with its ever-evolving array of devices, has fundamentally reshaped how we communicate and consume media. This constant innovation disrupts traditional structures and patterns, creating entirely new ways for societies to engage with information and connect with each other. Population censuses offer a valuable opportunity to assess household access to ICT devices. The chosen categories of devices to be included in the census should provide sufficient insight into the role of ICT within households and support government and private sector planning. This data can help optimize service delivery and understand the societal impact of technology.

5.623. The International Telecommunication Union (ITU) recommends including two topics in population censuses (collected at the household level):

(i) Household having a computer (desktop, laptop, or tablet)
(ii) Household with Internet access

Paragraphs 5.334-5.339 discuss the use of ICT at the individual household member level.

5.624. In designing the questions, NSOs should differentiate between two distinct viewpoints, namely (a) the availability of ICT devices to the households; and (b) use of ICT devices by the household members. The distinction is important, since households need not necessarily own, but may still have access to, computers, mobile phones and the Internet through devices provided by school, university or work to household members. It also means that NSOs interested in collecting information on ICT use, particularly of the Internet and mobile phones, would need to include a relevant question in the individual module of the population census questionnaire. The rational for adopting either viewpoint, or even a combination of both, is not necessarily only technical, but rests more on the prevailing conditions in the society, and on how the information will be used to characterize the socioeconomic profile of households of a country. Detailed usage statistics, including the intensity (frequency) of use and the full range of activities performed, are preferably obtained using household surveys.

5.625. In the context of this topic, a ‘computer’ refers to a desktop computer, a laptop (portable) computer or a tablet (or similar handheld computer). They exclude terminals connected to mainframe computers for data processing, and midrange multi-user systems that are primarily intended for shared

use. Devices with some embedded computing abilities, such as smart TV sets, and devices with telephony as their primary function are also excluded. 'Household with a computer' means that the computer is generally available for use by any member of the household at any time, regardless of whether or not it is actually used. The computer may or may not be owned by the household, but should, in either case, be considered a household asset.

5.626. 'Internet access' from home refers to the ability of the household to connect to the Internet using TCP/IP protocols[307]. Internet connections may be classified according to the technology employed, devices used, communication medium, or connection bandwidth (speed). Internet access at home is meant to include both narrowband and broadband connections. Access to the Internet is measured irrespective of the type of access, device used to access the Internet, or the method of payment. 'Households with Internet access' means that the Internet is generally available for use by any member of the household at any time, regardless of whether it is actually used. The connection and devices may or may not be owned by the household but should be considered household assets. If one member of the household has a mobile phone with connection to the Internet and makes it available for all members, then it should be considered that the household has access to the Internet.

38. Cars – number of available

5.627. This topic refers to the number of cars or vans normally available for use by members of the household. Motorcycles are relevant in some countries. The term "normally available" refers to cars and vans that are either owned by occupants or are under some other more or less permanent agreement, such as a lease, and includes those provided by an employer if available for use by the household, but excludes vans used solely for carrying goods or other commercial purposes.

39. Durable household appliances – availability of

5.628. The unit of enumeration is the household occupying the housing unit and information may be collected on the availability, within the housing unit, of durable appliances such as washing machines, dishwashing machines, refrigerators, deep freezes and microwave cookers, depending on national circumstances and user needs.

40. Outdoor space – access to

5.629. This topic refers to the ready access to an outdoor space intended for the recreational activities of the members of a household occupying a housing unit. The classification can refer to any outdoor space that is available:

1 As part of a housing unit (for example, a garden or backyard)

2 Adjacent to the building (for example, playgrounds placed next to the apartment building)

[307] The Internet protocol suite, commonly known as TCP/IP, is a framework for organizing the set of communication protocols used in the Internet and similar computer networks according to functional criteria. The foundational protocols in the suite are the Transmission Control Protocol (TCP), the User Datagram Protocol (UDP), and the Internet Protocol (IP).

3 As part of common recreational areas within a walkable distance from the housing unit (for example, parks, lakes, sports centres and similar sites)
4 Beyond a 10-minute walk.

Glossary of definitions and acronyms

Accessibility: A measure of data quality relating to the conditions and modalities by which users can obtain, use and interpret the data. (See paragraph 5.14)

Accuracy: A measure of data quality relating to the degree to which those data correctly estimate or describe the quantities or characteristics that the statistical activity was designed to measure. (See paragraph 4.13)

Administrative data: Data holdings that contain information collected primarily for administrative (not research or statistical) purposes. This type of data is collected by government departments and other organizations for the purposes of registration, transaction and record keeping, usually during the delivery of a service.

Address register: an up-to-date comprehensive list of all known physical addresses (usually with geo-locations) within a country. (See paragraph 1.94)

AI: Artificial intelligence

Anonymization: The process of protecting the confidentiality of personal information be removing all unique identifiers from the unit records.

Barnardisation: A method of statistical disclosure control for statistical outputs involving the addition of +1, 0 or -1 to some or all of the internal non-zero cells in a table.

Base register: A register whose function is to keep stock of the population and to maintain identification information. It may fundamental geospatial data such as addresses, geocoded buildings, establishments and dwellings, functional areas, and geographical names. (See paragraph 1.195)

Bedroom: A room equipped with a bed and used for night rest and excludes makeshift and temporary sleeping quarters. (See paragraph 5.550)

Blockchain technology: A decentralized, distributed ledger that stores the record of ownership of digital assets.

Building: A unit of enumeration defined as any independent free-standing structure comprising one or more rooms or other spaces, covered by a roof and usually enclosed within external walls or dividing walls that extend from the foundations to the roof. (See paragraphs 2.43-2.44 and 5.589)

Building register: A building register is a comprehensive database of buildings within a specific area. It contains detailed information about each building, often including its location (address, geographic coordinates), characteristics (type, size, number of stories, construction materials), ownership, and usage.

Business register: A business register is a comprehensive database of businesses within a specific area or jurisdiction. It contains detailed information about each business, which can vary but often includes: the business's name and address, industry classification, ownership details, and employment size.

CAII: Computer assisted Internet interview

CAPI: Computer assisted personal interview

CASI: Computer assisted self-interview

CATI: Computer assisted telephone interview

CAWI: Computer assisted web interview

Clarity: A measure of data quality that refers to the degree to which the data are understood (regardless of how accessible they are), particularly for non-experts. (See paragraph 4.15)

Coherence: A measure of data quality relating to the degree to which the census data can combined in different ways and for various purposes with statistical information from other sources.(See paragraph 4.16)

Collective living quarters: Structurally separate and independent places of abode intended for habitation by large groups of individuals or several households and occupied at the time of the census. (See paragraph 5.517)

Comparability: A measure of data quality relating to the degree to which statistics are comparable between geographic areas and over time.(see paragraph 5.16)

Combined census: A census in which some information on the numbers and characteristics of the population are derived from information taken from administrative data sources held for non-statistical purposes, but where other information that not available from such sources are collected directly from individual persons and households by means of full or partial field enumeration or from other sample surveys.

Completeness: A measure of data quality relating to degree to which statistics fully cover the phenomenon or population they are supposed to describe. See paragraph 4.12)

Conventional dwelling: A room or suite of rooms and its amenities in a permanent building or structurally separated part thereof which, by the way it has been built, rebuilt or converted, is intended for habitation by one household and is not, at the time of the census, used wholly for other purposes. (See paragraph 5.495)

Country of birth: Either the country in which a person's birth occurred or the country in which the mother was usually residing at the time of the person's birth. (See paragraph 5.111)

Country of citizenship: The country of which a person is a citizen and with which the person enjoys a particular legal bond, acquired by birth, naturalization, marriage or some other mechanism. (See paragraph 5.116)

CVRS: Civil registration and vital statistics

De facto census: A census based on a count of persons at where they were present on the reference date.

DEGURBA: Degree of Urbanisation (See paragraph 1.40)

DSE: Dual system estimation

De jure census: A census based on a count of persons at their place of usual residence on the reference date.

Deterministic method: A process that uses known input data to produce a specific output without any randomness or uncertainty

DSE: Dual System Estimation: A statistical method, based on a capture-recapture technique, applied to estimate the size of any undercount or overcount of the population. (See paragraph 4.145)

Enumeration: The act of collecting data about a person (or household) irrespective of whether this occurs with the direct participation of that person (or household) through a field operation, or indirectly using data that is already recorded in administrative registers.

EO: Earth observation

Family: those members of the household who are related, to a specified degree, through blood, adoption, or marriage.
FAO: Food and Agriculture Organization of the United Nations

Gender: A social construct that refers to the roles, behaviours, activities, and attributes that a particular society associates with being a man or a woman.

GeoAI: The application of AI technology to enhance the intelligence of GIS software which enables the analysis of massive amounts of geospatial data. (See paragraph 3.459)

GeoBlockchain: A combination of blockchain technology and geographic information systems (GIS) that can be used to support the analysis of spatial-temporal trends. (See paragraph 3.459)

Geocoding: The process of geospatially enabling statistical unit records so that they can be used in geospatial analysis. (See paragraph 1.125)

Geodata: See Geospatial information.

Geospatial information (also referred to as geodata): Data that describe the location of objects, events, or other phenomena within a country's land surface. (See paragraph 1.116)

GNSS: Global navigation satellite system

GSBPM: Generalized Statistical Business Process Model. (See paragraph 2.5)

GSGF: Global Statistical Geospatial Framework (see paragraph 1.116).

Housing census: The total process of planning, collecting, processing, analysing, disseminating and evaluating statistical data relating to the number and condition of housing units and facilities as available to the households pertaining, at a specified time, to all living quarters and occupants thereof in a country or in a well-delimited part of a country. (See paragraph 1.6)

Household: A unit of enumeration based on the arrangements made by persons, individually or in groups, for providing themselves with food and other essentials for living. (See paragraphs 2.35 or 5.135 for a fuller definition.)

Household reference person: The sole person in the household to whom all other household members report their relationship. (See paragraphs 5.142-5.145)

Housing unit: A unit of enumeration defined as a separate and independent place of abode intended for habitation by a single household, or one not intended for habitation but occupied as living quarters by a household at the time of the census. (See paragraph 5.491)

ICLS: International Conference of Labour Statisticians

ICSaW: International Classification of Status at Work

ICSE: International Classification of Status in Employment

ISCO: International Standard Classification of Occupations

ISIC: International Standard Industrial Classification of All Economic Activities

ICT: Information and communication technology

IDP: Internally displaced person

IGIF: Integrated Geospatial Information Framework

ILO: Internation Labour Organization

Imputation: The process of replacing missing data with substituted values.

Interoperability: The ability of different operating systems or software to exchange and make use of the same information.

Interpretability: A measure of data quality relating to the availability of supplementary information and metadata necessary to understand and use that data. (See paragraph 4.17)

ISCED: International Standard Classification of Education

IT: Information technology

LA: Local authority (England and Wales)

LAU: Local administrative unit. The classification of local administrative areas used by Eurostat. Levels 1 and 2 equate to those areas that were previously classified, respectively at the NUTS 4 and 5 levels.

LFS: Labour Force Survey

Living quarters: Unit of enumeration defined as are structurally separate and independent places of abode. (See paragraphs 2.47 and 5.485)

Locality: A distinct population cluster in which the inhabitants live in neighbouring or contiguous sets of living quarters and that has a name or a locally recognized status. (See paragraph 5.94)

Metadata: Information about the content, structure, quality and other relevant characteristics of a register or output product. (See paragraph 3.371)

Microdata: Information in a register or statistical database relating to a single entity or entities.

Migrant: A person who has changed his/her place of usual residence in a specified period prior to the census

NGIA: National Geospatial Information Agency. (See paragraph 1.124)

NSDI: National Spatial Data Infrastructure is a framework of policies, institutional arrangements, technologies, data, and people that enables the sharing and effective usage of geographic information by standardizing formats and protocols for access and interoperability. (See paragraph 1.121)

NSO: National Statistics Office

OLAP: Online analytical processing

Paradata: Formalized data on methodologies, processes and quality associated with the production and assembly of statistical data.

PES: Post-enumeration survey, defined as the complete re-enumeration of a representative sample of the census population and matching each individual who is enumerated in the PES with information from the main enumeration. (See paragraph 4.117)

PIN: Personal identification number

PHC: Population and housing census

Place of birth: The geographic area in which either the birth of an individual person actually occurred, or where the mother of the individual person resided at the time of the person's birth.

Place of usual residence: Either the place where the enumerated person was present or usually resident at the time of the census, or his or her legal residence. (For a fuller definition see paragraphs 2.50–2.52.)

Population: Any set of persons assigned to a geographic area who meet the defined criteria at the census reference time.

Population census: The total process of planning, collecting, processing, analysing, disseminating and evaluating demographic, economic and social data at the smallest geographic level pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country. (See paragraph 1.4)

Population count: the aggregate obtained from the simple addition of individual records from the either to the whole, or a sub-set, of the enumerated population.

Population estimate: the aggregate obtained as an outcome of a statistical method of estimating the target population.

Population present count; A count of population enumerated at the place where they are found in the census, usually the dwelling where they spend census night.

Population register: A system that records all residents in the country, together with their basic demographic information, and updates for births, deaths, changes of marital status, international and internal migration. (See paragraph 1.90)

PR: Population register

Probabilistic method: A mathematical technique that uses probability as a means of deriving statistical values.

Process quality: A measure of how well a methodological or statistics processes are implemented and followed to produce the desired results.

Punctuality: A measure of data quality relating to the delay between the date of the release of the results and the target date (the date by which the data should have been delivered. (See paragraph 3.14)

Register: A systematic collection of unit-level data organized in such a way that updating is possible. Updating is the processing of identifiable information with the purpose of establishing, bringing up to date, correcting, or extending, the register, that is, keeping track of any changes in the data describing the units and their attributes.

Register-based census: A census in which the data on the numbers and characteristics of the population are derived from information taken from administrative data sources held for non-statistical purposes. No information is collected directly from individual persons or households.

Relevance: A measure of data quality relating to the degree to which statistics meet current and potential needs of the users. See paragraph 4.11)

Rolling census: A census in which information on individuals and households is collected through a continuous cumulative survey covering the whole country over a period of time (generally years), rather than a particular day or short period of enumeration. (See Table 1)

Room: A space in a housing unit enclosed by walls reaching from the floor to the ceiling or roof covering, or to a height of at least 2 metres, of an area large enough to hold a bed for an adult, that is, at least 4 square metres. (See paragraph 5.546)

Rural area: Defined using a national methodology established in accordance with a country's own needs. (See paragraph 5.99)

SAE: Small area estimation is a means of estimating indicators when critical data are not available at lower levels of geography. (See paragraph 1.45)

SBR: Statistical Business Register - a register that provides the infrastructure to support the collection of high quality, reliable and consistent economic data and the production of economic statistic', (See paragraph 1.98)

Sex (at birth): The biological attribute whereby persons are traditionally classified as either male or female for the purposes of statistical analysis.

SDG: Sustainable Development Goal

Simultaneity: That feature of the census that requires that each person and each set of living quarters should be enumerated at the same well-defined point in time and the data collected should refer to a well-defined reference period. (see paragraph 1.11)

SMS: Short Message Service

SNA: System of National Accounts

SOL: Signs of life

SPD: Statistical population dataset

Specialized register: A register that hold information relating to a specific theme or topic such as vehicles, social security, education, or health. (See paragraph 1.198)

Statistical population register: A systematized and indexed collection of individual records for every resident (including nationals and foreign citizens) of the country. (See paragraph 1.201)

Statistical register: A register created for statistical purposes. They are typically created by transforming data from registers and/or other administrative data sources.

Synchronisation: The transmission of data recorded on handheld devices in the field to a central server.

Timeliness: A measure of data quality relating to the period between the availability of the information and the event or phenomenon it describes. (See paragraph 3.14)

Traditional census: A census based on the direct count of all individuals and the collection of information on their characteristics through the completion of either a self-completion or interview-based questionnaires, either in a paper or electronic format.

UNECE: United Nations Economic Commission for Europe

UNESCO: United Nations Educational, Scientific and Cultural Organization

UNFPA: United Nations Population Fund (formerly the United Nations Fund for Population Activities)

Unique ID/key: A single alpha numeric identifier that relates a characteristic or variable to a particular entity (person, household or dwelling) across a range of different registers or administrative data sources.

Universality: that feature of the census that requires that it should cover a precisely defined territory (for example, the entire country or a well-delimited part of it) and should include every person present and/or residing within its scope. (see paragraph 1.10)

Urban area: Defined using a national methodology established in accordance with a country's own needs. (See paragraph 5.99)

Usual resident population count: A count of all persons assigned to the geographic area where they are usually resident at the time of the census

USSD: Unstructured Supplementary Service Data

Work: Any activity performed by persons of any sex and age to produce goods or to provide services for use by others or for own use. (See paragraph 5.341)

Working time: The time associated with productive or work activities and the arrangement of this time during a specified reference period. (See paragraph 5.349)

XML: Extensible markup language. In computing a language that that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.