UNECE

Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses

UNITED NATIONS

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses

Prepared by the Conference of European Statisticians Task Force on Assessing the Quality of Administrative Sources for Use in Censuses

United Nations
Geneva, 2021

United Nations publication issued by the Economic Commission for Europe

ECE/CES/STAT/2021/4

ISBN: 978-92-1-117276-8
eISBN: 978-92-1-000090-1
ISSN: 0069-8458
Sales No. E.21.II.E.13

Preface

This publication's main purpose is to provide guidance to the producers of population and housing censuses on how to assess the quality of administrative data for use in a census. The Guidelines cover the practical stages of assessment, from working with an administrative data supplier (or administrative authority) to understand a source, its strengths, and limitations, all the way to the receipt and analysis of the actual data. The Guidelines cover key quality dimensions on which an assessment is made, using various tools and indicators. For completeness, the Guidelines also include information about the processing and output stages of a census, with respect to the use of administrative sources.

The publication was prepared by a Task Force established by the Conference of European Statisticians (CES), composed of experts from national statistics offices (NSOs), and coordinated by the United Nations Economic Commission for Europe (UNECE).

Acknowledgements

These Guidelines were prepared by a UNECE Task Force consisting of the following individuals:

Steven Dunstan (Chair), United Kingdom
Katrin Tschoner, Austria
Christoph Waldner, Austria
Josée Morel, Canada
Lionel Espinasse, France
Stefan Dittrich, Germany
Tobias Kahlenberg, Germany
Thomas Körner, Germany
Ingeborg Vorndran, Germany
Sheelagh Bonham, Ireland
Brendan Murphy, Ireland
Alaa Atrash, Israel
Yael Feinstein, Israel
Gerardo Gallo, Italy
Donatella Zindato, Italy
Snezana Remikovic, Montenegro
Eric Schulte Nordholt, Netherlands
Christine Bycroft, New Zealand
Abby Morgan, New Zealand
Janusz Dygaszewicz, Poland
Krzysztof Woźnica, Poland
João Farrajota, Portugal
Sandra Lagarto, Portugal
Paula Paulino, Portugal
Dmitrii Calincu, Republic of Moldova
Valentina Istrati, Republic of Moldova
Marina Pérez Julián, Spain
Alberto Salcedo, Spain
Şebnem Beşe-Canpolat, Turkey
Muharrem Gürleyen Gök, Turkey
Mehmet Şaban Ucari, Turkey
Louisa Blackwell, United Kingdom
Adriana Castaldo, United Kingdom
Sara Correia, United Kingdom
Sara Haylock, United Kingdom
Charlotte Hillyard, United Kingdom
Jack Sim, United Kingdom
Stephan Tietz, United Kingdom
Clare Watson, United Kingdom
Marina Wright, United Kingdom
Tom Mule, United States of America
Eduard Jongstra, UNFPA
Diana Beltadze, Eurostat
Sorina Vâju, Eurostat
Ian White, Independent expert
Fiona Willis-Núñez, UNECE

The Guidelines were developed and agreed upon by the entire Task Force. Each chapter was drafted by a team under the leadership of one or more individuals: Census methodologies and uses of administrative data for censuses: Sara Correia and Ian White; Quality framework: Sara Correia and Sorina Vâju; Source Stage: Sorina Vâju, Josée Morel, Diana Beltadze and Steven Dunstan; Data Stage: Christoph Waldner, Tobias Kahlenberg and Sara Correia; Process Stage: Sara Haylock, Abby Morgan, Adriana Castaldo and Steven Dunstan; Output Stage: Sara Correia, Marina Pérez, Sandra Lagarto and Steven Dunstan. The Guidelines were edited by Arona Pistiner.

The European Free Trade Association (EFTA), an intergovernmental organization comprising Iceland, Liechtenstein, Norway and Switzerland, provided financial support for the editing of these Guidelines. The EFTA Statistical Office provides statistical assistance to countries in cooperation with other international organizations.

The Task Force extends particular thanks to the United Kingdom's Office for National Statistics (ONS) for the invaluable contributions of many of its staff to these final Guidelines, most notably Steven Dunstan, Sara Correia and Adriana Castaldo.

Contents

List of boxes

List of figures

List of tables

List of country case studies

Acronyms and abbreviations

ABPE	Administrative Data-based Population Estimate
ABS	Australian Bureau of Statistics
AIDA	Archivio Integrato Dati Amministrativi (Integrated Administrative Data Base, Italy)
CAPI	Computer-assisted Personal Interview
CATI	Computer-assisted Telephone Interview
CAWI	Computer-assisted Web Interview
CAxI	Computer-assisted Multi-mode Interview
CDB	Central Database (Austria)
CES	Conference of European Statisticians
CIS	Customer Information System
CPR	Central Population Register
CT	Census Test (Portugal)
DSE	Dual System Estimation
DA	Demographic Analysis
ESS	European Statistical System
ESSnet	European Statistics System Network
FDP	Final Data Pool (Austria)
FPC	Fichero Precensal (pre-censual file, Spain)
GSBPM	Generic Statistical Business Process Model
GMTMM	Generalized Multi-trait Multi-method Model
INSEE	Institut national de la statistique et des études économiques (National Institute of Statistics and Economic Studies, France)
ISO	International Organization for Standardization
LAU	Local Administrative Unit
LMS	Legal Marital Status
MOU	Memorandum of Understanding
MSE	Mean Square Error
NHS	National Health Service (United Kingdom)
NIP	Numer Identyfikacji Podatkowej (tax identification number, Poland)
NRFU	Non-response Follow-up
NSO	National Statistical Office
NUTS	Nomenclature of Territorial Units for Statistics
NZ	New Zealand
ONS	Office for National Statistics (United Kingdom)
PBR	Population Base Register (Italy)
PE	Population Estimate
PES	Post-enumeration Survey
PESEL	Powszechny Elektroniczny System Ewidencji Ludności (Universal electronic system for registration of the population, Poland)
PHC	Population and Housing Census
PIA	Privacy Impact Assessment
PII	Personally, Identifiable Information
PIN	Personal Identification Number

PPHC	Permanent Population and Housing Census (Italy)
PR	Patient Register
QA	Quality assessment
RBI	Registro base degli individui (population base register, Italy)
REGON	Rejestr Gospodarki Narodowej (Business identification number, Poland)
ROC	Receiver Operating Characteristic
RSBL	Registro statistico base dei luoghi (statistical base register of addresses, Italy)
SCD	Statistical Census Dataset (Poland)
SDC	Statistical Disclosure Control
SE	Statistic Estonia
SIR	Sistema Integrato dei Registri (integrated system of statistical registers, Italy)
SOL	Signs of Life
SP	Statistics Portugal
SPD	Statistical Population Dataset (Portugal)
Stats NZ	Statistics New Zealand
TSE	Total Survey Error
UK	United Kingdom
UKSA	United Kingdom Statistics Authority
UNECE	United Nations Economic Commission for Europe
UPRN	Unique Property Reference Number
US	United States
VOA	Valuation Office Agency
VQS	Variable Quality System

Executive summary

The use of administrative data in censuses continues to increase across the countries of the UNECE region and beyond, whether it be to support a traditional census, under a combined census, or register-based census methodology whereby the population is enumerated and/or the census variables populated using administrative data. It is important that NSOs understand the strengths and limitations of administrative data for use in their censuses to ensure that the right decisions are made about the use of such data.

These Guidelines aim to provide census producers with a practical guide for assessing the quality of administrative data, through a series of assessment Stages. The Guidelines draw on quality frameworks and best practices adopted by NSOs across the world, including the widely used framework of Statistics Netherlands (Daas et al. 2012), the New Zealand Total Error Framework (Zhang 2012), and the deliverables from the Statistical Network Methodologies for an Integrated Use of Administrative Data in the Statistical Process project (Eurostat ESSnet MIAD 2014).

The Guidelines are based on four Stages: Source, Data, Process, and Output, with the first two Stages being the principal focus of the Guidelines, providing an assessment of input quality (i.e., the quality of administrative data sources set against their use in a census).

The Source Stage covers the assessment of the administrative source through working with the data supplier and reviewing relevant metadata. This Stage includes an assessment of whether the source can meet the needs of the census, under the quality dimensions of relevance, accuracy, timeliness, coherence, and comparability. An assessment is also made of the accessibility and interpretability of the administrative source, covering any restrictions on access and use, and public acceptability. Finally, an assessment is made of whether the data supplier can meet the needs of the NSO, considering factors such as the strength of the relationship with the supplier and the status of the supplier.

The Data Stage covers an assessment based on an analysis of the actual data (as transmitted by the data supplier) and through comparisons with other sources. This Stage includes the validation of data on receipt, an assessment of accuracy and reliability (including coverage and measurement errors), timeliness and punctuality, and an assessment of linkability. For the Source and Data Stages, the assessment is against key data quality dimensions, for which various tools and indicators are provided.

The **Process** and **Output Stages** are provided for completeness and give the reader information about the key processes and considerations for transforming administrative data for use in a census, and for assessing the quality of census outputs that are based on administrative data.

The experiences of several countries are included throughout the Guidelines, using basic illustrations or more detailed case studies.

The Guidelines also provide suggested areas for further work in the concluding chapter and a set of key recommendations for NSOs to consider:

1. Identify administrative sources against specific use cases, to assess the expected or required outcomes of using the source for the identified use case.
2. Build and support relationships between NSOs and data suppliers, with a legal basis for supply and use of data, and collaborative feedback mechanisms.
3. Harness supplier relationships to ensure a comprehensive understanding of source metadata.
4. Assess the coherence and compatibility of the administrative source to the census, to understand differences, if any, between the required populations, concepts, definitions, and time-related dimensions.
5. Understand restrictions and challenges to acquiring an administrative source and integrating it into a census by comparing the value of overcoming these challenges against the effort and risk entailed in doing so.
6. Assess and manage the risk implied by use of an administrative source.
7. Be transparent in communication with data users and with the public about how and why administrative data are used for the census, emphasizing procedures for ensuring effective use of the data, and for data protection.
8. Undertake feasibility research as a 'proof of concept' and test runs with real data prior to including administrative data in census production.
9. Make use of expert review and conduct comparisons between sources over time to identify quality concerns in each source.
10. Record and publish results of quality assessment at all Stages.
11. Develop an NSO-specific quality assurance framework and strategy, supported by clear and comprehensive documentation and training procedures, with a focus on continuous assessment and communication between the NSO, data users, and data suppliers.

Chapter 1. Introduction

1.1 Background

1. In 2017, the UNECE Task Force on Register-based and Combined Censuses prepared the *Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses*[1]. The Guidelines included a section on "data sources and their quality" with a general discussion of this topic. Experts at the UNECE-Eurostat Expert Meeting on Population and Housing Censuses (Geneva, 4-6 October 2017) identified the quality of administrative sources as a topic of primary importance for many countries. Consequently, the Expert Meeting called for the establishment of a new UNECE Task Force on Measuring the Quality of Administrative Sources for Use in Censuses, building on the work of the previous Task Force.

2. The Task Force was established in 2018, with its Terms of Reference[2] approved at the February 2018 meeting of the Bureau of CES in Helsinki (14-15 February 2018). The Task Force reported to the UNECE Steering Group on Population and Housing Censuses, which in turn reports to CES and its Bureau.

3. The objective of the Task Force was to develop guidance on the measurement of the quality of administrative sources for use in censuses[3]. The terms of reference stipulated that the Task Force should develop guidance that is relevant to all UNECE countries, and that it should build on the work of Eurostat's ESS.VIP ADMIN project[4] on the use of administrative sources in the production of official statistics.

4. The Task Force met in-person during the 2018 and 2019 UNECE-Eurostat Expert Meetings on Population and Housing Censuses and held an additional in-person meeting in Geneva, Switzerland on 5-6 March 2020.

5. The Task Force presented annual reports of its progress to the UNECE-Eurostat Group of Experts on Population and Housing Censuses in 2018, 2019 and 2020. A full draft of these Guidelines was circulated in advance of the 2020 expert meeting (online, 30 September – 1 October 2020). Feedback received from participants was used to refine the Guidelines.

6. These Guidelines serve as a practical toolkit for the assessment and measurement of the quality of administrative sources for population and housing censuses.

[1] Available at: http://www.unece.org/index.php?id=50794

[2] Available at: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2018/February/06Add.1-TF on quality of admin data for censuses ToR apr.pdf

[3] The Task Force subsequently decided to adjust its title and the corresponding objective to 'assessing' rather than 'measuring' the quality of administrative sources for use in censuses.

[4] More information on this project can be found at: https://ec.europa.eu/Eurostat/cros/content/ess-vision-2020-admin-administrative-data-sources en

1.2 Use of administrative data in censuses

7. Administrative data sources are data sets that contain information collected primarily for administrative purposes[5]. This includes, but is not limited to, data collected by government departments, public bodies, and other organizations for the purposes of registration, transaction, and record-keeping, usually during the delivery of a service. They include administrative registers (with a unique identifier) such as a country's population, business, address, education, health, employment, and tax registers, as well as other administrative sources (without a unique identifier). Administrative registers and/or other administrative sources are used to create statistical registers, which are specifically used for statistical purposes, including a census. The administrative sources, most used in censuses, are outlined in Chapter 2 of the Guidelines.

8. The use of administrative data sources in censuses varies across countries. Such sources may be used to enhance or to supplement a traditional census, to conduct a combined census, or in the construction of a fully register-based census. 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. This has been motivated by the benefits administrative data can bring, including reduced cost and respondent burden, improved timeliness and frequency of results, improvements to quality, and greater flexibility to respond to user needs (see, for example, Section 4.1 of UNECE 2018). Furthermore, the conditions within many countries have changed to support and facilitate the use of administrative data throughout national statistical systems (c.f. Section 4.2 of UNECE 2018). This has occurred due to changes in legislation, public and stakeholder acceptability, and through developments in technology and statistical methodologies.

9. The importance of administrative data has been highlighted by the challenges that NSOs are now facing when it comes to collecting data directly from the population, whether due to a reluctance of the public to engage with the census or their ability to do so. This challenge was emphasized at the onset of the COVID-19 crisis in 2020, when both the public's ability to engage with NSOs and at the same time NSOs' ability to engage with the public were affected significantly. The various ways in which administrative data are used in censuses are covered in Chapter 2 of these Guidelines.

1.3 Key risks to quality

10. For all the benefits that administrative data can bring, there are many key quality considerations that must be assessed and evaluated before incorporating an

[5] The UNECE guide *Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices* (UNECE 2011, pp.1-3) discusses the evolution in understanding of what 'administrative purposes' mean. The guide concludes that a wide and inclusive definition encompassing private-sector data is increasingly favoured.

administrative source into a census. First, the NSO will have only limited control over the way the data are collected and processed. There is significant dependency on the authorities creating the administrative data. For example, if the data supplier is unable to meet the NSO's requirements to provide the right data at the right time, this will impact the timeliness of the census results. Similarly, if the data supplier does not adequately engage with the NSO on any potential changes to the source, this could impact coherence and comparability.

11. Second, the use of administrative data by the NSO for purposes other than those for which the data were originally collected raises privacy, security, and legal concerns. The NSO must assess public acceptability of the data, guaranteeing the required assurances are in place and that they are communicated to the public (and to the data supplier). The use must also be lawful. Without acceptance or agreement both from the public and from the data supplier, or a credible legal basis for the use of the administrative source, there would be significant risk to the reputation of the NSO and its ability to deliver a high-quality census. This can arise if the public changes their behaviour in the way they interact with the data supplier or the NSO, due to concerns over the way the NSO is using their data.

12. Third, administrative data have (in general) not been collected for statistical purposes. Consequently, the data sources may have adopted different concepts, classifications and definitions from those required by the census, they may refer to different reference periods, be subject to lags in the updating of information, and may have limited coverage of the census population. Additionally, the accuracy and completeness of the data will be highly dependent on the importance of the data to the data supplier's 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. Furthermore, the data sources may not have the necessary identifiers or variables to permit the record linkage required for the census.

13. Finally, the complexity of the administrative data and the availability and completeness of the associated metadata will impact the ability of an NSO to understand, access, and use an administrative source. For example, administrative data can be held in large, complex data structures, posing significant technical challenges for the NSO to assess, and overcome. The complexity of administrative data may also impact the accessibility and clarity output quality dimension from a data user's perspective. That is, users of the census data may find it difficult to understand the use of administrative data in the census and the impact this use has on the quality of the census outputs.

14. These key quality considerations will inform decisions about the use of administrative data in a census. The Guidelines address each of the considerations in detail.

1.4 Scope and structure of the Guidelines

15. The focus of the Guidelines is on the assessment of the quality of administrative data sources for use in the census (i.e., input quality). They do not cover other sources per se (e.g., Big Data, commercial data). Nevertheless, much of the material within the Guidelines is applicable beyond administrative data (guidance on the quality assessment (QA) of Big Data can be found in UNECE 2014).

16. The Guidelines begin by providing information about the different census methodologies and how administrative data can be used under each of these methodologies, including the types of data sources used. The aim is to provide information that may be useful for NSOs that wish to incorporate new administrative data sources into the design of their censuses (Chapter 2). The impacts of the COVID-19 pandemic on the use of administrative data in censuses are briefly considered in this chapter.

17. Chapter 3 outlines the overall quality framework on which the Guidelines are based. The framework is built around four assessment Stages. The Stages broadly relate to the lifecycle of using administrative data in the census:

    (a) Understanding, evaluating, and working to acquire a source (the Source Stage),

    (b) Receiving the actual data and assessing its quality (the Data Stage),

    (c) Processing the administrative data for use in the census (the Process Stage),

    (d) Assessing the quality of the census outputs that use administrative data (the Output Stage).

18. The chapter also outlines the quality dimensions assessed within each Stage and the associated errors (e.g., representation and measurement errors). The chapter concludes with an outline of the importance of carrying out feasibility research on the use of administrative data, explaining how the Stages within the Guidelines can be used for this purpose.

19. Chapter 4 covers the first assessment Stage (the Source Stage), where information is gathered about an administrative source through communication with the data supplier and by reviewing existing metadata. At this Stage, the focus is on assessing the relevance of the source against the needs of the census, covering accuracy, timeliness, coherence and comparability, accessibility and interpretability. An assessment is also made about the institutional environment, including whether the data supplier can meet the needs of the NSO, considering factors such as the strength of the relationship with the supplier and the status of the supplier.

20. Chapter 5 covers the Data Stage of the assessment, where data are received from the data supplier and are assessed through analysis of the data and through comparisons with other data sources. During both the Source and Data Stages, the assessment and

measurement of quality is set against many data quality dimensions, using various tools and indicators. The two Stages together provide an assessment of input quality.

21. The information and insight gained through the Source and Data Stages are useful not only to determine whether a particular source could be used in the census, but also to determine the necessary processing of the administrative data for use in a census. In general, administrative data cannot be used directly in a census, due to conceptual and definitional differences. There are also limitations of coverage, completeness, and accuracy. It is necessary to transform the data from administrative sources (including registers) using the information gained at the Source and Data Stages. Some of the most important processes and the associated quality considerations are covered in Chapter 6 of the Guidelines.

22. The Source, Data, and Process Stages relate directly to the quality of census outputs in accordance with the European Statistical System (ESS) output quality dimensions. Conversely, the assessment of the census outputs will provide valuable information about where there may be limitations or concerns about the administrative data, or the processing of these data, that were not identified initially at the Source, Data, and Process Stages. There is an iterative process of assessment, which can inform both ongoing improvements to the administrative sources (working with the data supplier to improve the source), and improvements to the processing of the administrative data by the NSO. The assessment of census outputs quality which use administrative data is covered briefly in Chapter 7 (Output Stage).

23. Various country-specific examples and case studies using basic illustrations or more detailed case studies are provided throughout the Guidelines. The chapters illustrate the application of the Stages of QA in practice.

24. Finally, Chapter 8, Conclusions and recommendations, provides a summary of the recommendations that are presented throughout the earlier chapters. The final chapter proposes further internationally coordinated work on the QA of administrative data.

Chapter 2. Census methodologies and uses of administrative data for censuses

25. This chapter summarizes the range of census methods and uses of administrative data in censuses, which are common throughout UNECE countries. This will help NSOs within the UNECE region and beyond when using administrative data in their censuses – regardless of data collection methodology adopted.

2.1 Census methodologies

26. As has been noted in previous UNECE publications (UNECE 2015; UNECE 2018) there are several different ways to undertake the data collection process in a population and housing census. This section provides an overview of census methods and where these Guidelines may be useful to NSOs.

27. For the sake of simplicity this chapter summarizes only the three main categories of census data collection methods:

    (a) Traditional census,

    (b) Register-based census, and

    (c) Combined census.

28. The UNECE Census Wiki[6], which compiles information on the 2020 round of censuses as reported by member countries, indicates that the trend away from the traditional census is continuing rapidly. Out of 52 UNECE countries for which information is available, fewer than half (23) are conducting a traditional census in the current 2020 census round (with 13 countries planning to conduct a register-based census and 16 countries planning a combined approach). Nevertheless, as discussed below, there are still opportunities and benefits for NSOs that conduct a traditional census to utilize administrative data.

29. The key features of the three census methods identified are summarized below. A more detailed discussion of the various census methodologies, including the necessary prevailing conditions, advantages, and challenges, is given in the *CES Recommendations for the 2020 Censuses of Population and Housing* (UNECE 2015). For a detailed description of the essential features of a census and how these relate to the different census methodologies, see Chapter 3 of UNECE (2018).

[6] Available at bit.ly/UNECECensusWiki2020

2.1.1 Traditional census

30. The term 'traditional census' refers, in the broadest sense, to a census based on a direct count of all individuals, households and housing units, and the collection of information on their characteristics through the completion of census questionnaires, either on paper or in an electronic format. The information is collected in the field by means of a full enumeration across the whole country in a relatively short period of time.

31. The information can be collected by one or more of the following methods:

    (a) Directly from households (with delivery and collection of paper forms undertaken by enumerators, the postal service, or other means),

    (b) Online, using electronic questionnaires, or

    (c) By enumerators during a face-to-face interview of the household using either paper or electronic questionnaires.

32. Since 2001 some countries have made significant changes to their data collection operations, while still falling within the definition of a traditional design. For example, in the United States, the Census Bureau focuses on collecting only short form data (10 questions, mainly demographic data with three housing questions) in a full enumeration in the decennial census year (with a reference date of Census Day, 1 April). A large sample household survey, the American Community Survey (ACS), collects the more detailed long form data (demographic, social, economic, and housing) every month. New data from the ACS are released annually throughout the decade, replacing the need for a census 'long form' that had previously been sent out to a sample of the population.

33. In contrast, France's INSEE has adopted a different approach. A rolling census is conducted by means of a cumulative continuous sample survey, covering the whole country over the decennial period rather than an enumeration carried out simultaneously, in all areas, on a specific reference date. The French rolling census is also, in large municipalities (more than 1,000 inhabitants), based on an exhaustive register of dwellings. This register is updated using administrative data (building permits) and with checks by municipalities. Tax data files are also used for estimates for 40 per cent of the small municipalities (less than 10,000 inhabitants) each year. An annual survey, such as that used in France, may be conducted over the course of a year, in a particular month, or a shorter time frame. With such an approach it is possible to build a sample framework to produce:

    (a) National results with a single annual survey,

    (b) Regional results by cumulating data from several consecutive annual surveys, and

    (c) Small-area results by cumulating data from a more substantial number of years.

2.1.2 Register-based census

34. The register-based census is a totally different approach initially developed by the Nordic countries in the 1970s, among which Denmark was the first to conduct a fully register-based census in 1981. Under this approach, there is no direct collection of data from the population, and the traditional enumeration is replaced by administrative data held in various registers (such as a population register, building or address register, social security register, tax records, etc.) through a matching process, usually making use of personal identification numbers (PINs). Once a good quality system of statistical registers has been established, this approach permits the (often more frequent) production of census data at a greatly reduced cost and with far less human effort.

35. This methodological approach clearly demands the greatest use of administrative sources and is, therefore, heavily dependent on establishing and ensuring the highest levels of quality of data from such sources.

2.1.3 Combined census

36. Since the 1990s, several other countries within the UNECE region and elsewhere have developed innovative methods to conduct their census, combining the use of administrative data with the collection of an often-reduced set of data from a direct field enumeration of the population. The field enumeration may still be the primary method for collecting census data. However, administrative data are used where available to reduce response burden and add extra information not collected in the census (e.g., income-related questions). The field enumeration aims to derive specific variables for which the relevant data are not readily available from any administrative source. Under this combined approach, the field data collection may cover the whole population or just a sample.

37. This methodological approach has been used recently by several NSOs in their transition from a traditional to a wholly register-based census. These Guidelines have been written primarily to help statistical producers through such a transition, or when running a combined or register-based census. Nonetheless, they will also support the assessment of administrative data used in a primarily traditional census.

2.1.4 Impacts of the COVID-19 pandemic on census methodologies

38. The distribution of the different census methodologies among CES countries was already undergoing a general shift towards combined or fully register-based censuses, prior to 2020. During CES-wide consultation on a draft of these Guidelines, many countries indicated that this shift had been accelerated and/or that they were making increased use of administrative data to support traditional censuses, as a result of the pandemic. The key impacts of the pandemic on the use of administrative data in censuses, as identified by countries in the consultation, are summarized in Box 1.

Box 1: Impacts of the COVID-19 pandemic on censuses and administrative sources

The COVID-19 pandemic has had a profound impact on censuses across the world and on the use of administrative data (see, for example, UNECE 2021). It has affected the way census collection activities have been carried out; has increased or accelerated moves towards the use of administrative data (particularly where field data collection has not been possible); and has even delayed the census in many countries. The crisis has also demonstrated the need for more frequent and timely statistics about the population during times of unprecedented change, as decision-makers seek information on where and how people live, study and work, and on health and mortality outcomes.

Where the shift towards more use of administrative data has been expediated, this has required rapid changes and improvements to data collection and processing systems across both administrative organizations and NSOs. Specifically, new procedures, protocols and even legislation have been required to facilitate the collection, sharing and use of data. This has required effective collaboration between administrative organizations and the NSO (see section 4.1.5).

For countries that have delayed their census beyond 2021, administrative data will be important to support the production of census statistics for previous reference years. For example, in order to meet Eurostat requirements, both Germany and Hungary, which have postponed their census to 2022, will make use of administrative sources to produce statistics for the 2021 reference year from their 2022 national census.

The pandemic has also had a significant impact on the quality and content of administrative sources. For example, some types of interactions with health services may have decreased (with people avoiding health care services due to concerns about catching the virus) while others have increased (with people registering to be tested, treated or vaccinated). These changes will impact the coverage of health registers.

Furthermore, the need to deliver new services and support to the public has led to the development of new administrative processes and systems. In the UK, for instance, this has included systems to support those out of work due to the pandemic (through furlough payments) and to support the roll-out of testing, tracing and the vaccination programme (providing new sources of data).

Finally, the pressures on the public and administrative organizations due to the pandemic have impacted the timeliness and accuracy of administrative data. Examples include delays in registrations of births and reductions in the level of quality assurance (with resources being temporarily diverted elsewhere).

2.2 Uses of administrative data

39. The extent of using data from administrative sources for the purposes of carrying out a population and housing census clearly will depend on the type of methodology used in the data collection operation.

40. Across the different types of census methodology described above, administrative data can be used in a variety of ways. Among these, the following use cases emerge as key:

    (a) In the construction and optimization of census sampling frames and field operations (as adopted by the US and Canada),

    (b) To enable the QA of census estimates[7] by comparison with administrative sources and to inform adjustments through, for example, editing and imputation (as adopted in Estonia for its 2011 census),

    (c) To derive existing census variables and to add new census variables (as adopted by the UK),

    (d) In the construction of statistical population registers[8] and direct use of administrative data-based enumerations for the census (as seen in Spain and New Zealand, respectively), and

    (e) In a full administrative data census (such as in the Netherlands).

2.2.1 Construction and optimization of census sampling frames and field operations

41. The first use case is for administrative data to construct and optimize census dwelling/address frames and field operations. This includes assessing the quality of a census sampling frame constructed from administrative data. The use case also establishes if the administration data can determine whether an address is likely to be occupied and by whom, or whether a certain address is likely to be 'hard-to-reach'[9], thereby optimizing census field operations (see Box 4 and section 6.3).

42. For those countries where some element of a field enumeration is retained – either in a fully traditional census or where a combined approach is adopted – data from administrative sources can be used to support the field operation(s). Many such countries may, for example, use information from address or building registers to construct consistently sized enumeration areas that contain broadly the same numbers of households or dwellings.

43. Alternatively, such information can be used to select appropriate household or housing unit samples where a full dataset is not collected from the whole population.

[7] See glossary for clarification of what is meant by the term 'census estimates'
[8] See glossary for a definition of a statistical register.
[9] See glossary for a definition of hard-to-reach.

44. The quality of administrative data-based census frames will benefit from an assessment of data sources at the Source, Data, and Process Stages proposed in these Guidelines. However, given the iterative nature of a field operation (i.e., the census frame improves throughout collection), such an assessment may emphasize aspects of coverage (linked to relevance) over the accuracy dimension.

2.2.2 Replacing and/or adding new census variables

45. The second use case is concerned with assessing the quality of administrative data used to replace and add new variables to the census.

46. When countries decide to reduce the size (and the associated cost) of a full field enumeration by adopting a combined census approach, data from appropriate administrative sources can be used to replace the information collected from a household questionnaire. For example, reliably accurate information on marital and employment status, or the year of immigration may be readily available from administrative registers, thus eliminating the need to collect such data directly from individuals.

47. Alternatively, a valid case may be made by users for the NSO to collect information in the census either that has been shown to be publicly sensitive or that requires a level of detail many individuals may be unable to report accurately on a traditional census questionnaire. For example, information relating to infant deaths may be culturally sensitive in some countries, while data on household income may often require potentially confidential information to be shared among other household members. In these cases, the equivalent data relating to the linked individual may be obtained from administrative sources (such as vital registration or tax records).

48. Quality assessing administrative data at the Source Stage can aid in the decision of selecting administrative sources to use in such cases. In addition, assessing the chosen source(s) at the Data and Process Stages may ultimately ensure the quality of the outputs.

2.2.3 Construction of statistical registers and the direct use of administrative enumerations

49. The third use case relates to administrative sources to enumerate the population (see also section 6.3). The whole population may be enumerated through an administrative list (e.g., a population register), or administrative data may be used to enumerate part of the population, for example, those missed by field enumeration[10]. A distinction is drawn between situations when NSOs can rely on unique identifiers to integrate multiple sources into one register, or identifiers do not exist (and when reliance on

[10] Case study 6.7.4 from Italy provides an example of the use of administrative data to adjust for survey undercoverage under a complex estimation system using various administrative sources and surveys.

deterministic/probabilistic methods for entity resolution and to link sources on variables such as name, date of birth, address, etc. are required).

50. The United Nations (2014) has noted that population registers are now well established in several countries, especially those in the UNECE region, where they have been effectively used as a statistical data source for decades. Registers may be considered the logical product of the evolution of a vital statistics system. They have become an important source of information for various statistical surveys and, in many cases, for population and housing censuses.

51. 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), and marital status. Moreover, if complete, population registers can produce data on both internal and international migration through changes of residence, as well as international arrivals and departures. Registers can be used as the direct base for an 'administrative enumeration' to replace a traditional field data collection operation.

52. As with the previous use case, quality assessing source data at the Source and Data Stages will be essential in designing a methodology for the construction of statistical population registers. Ultimately, this will be an iterative design process, where QA at the Output Stage may reveal issues to be addressed at earlier Stages. It is suggested that when constructing registers, NSOs follow all the quality Stages proposed in this guide.

2.2.4 Quality assessment and adjustments

53. The fourth use case relates to the quality of the data source to be used for enhancing existing census variables. In this type of use case, administrative data are used for the editing and imputation of an existing census variable, as opposed to the direct/complete replacement of a traditional collection.

54. Even in those countries that continue to carry out a traditional census, data from administrative sources can be used to either quality assure the information collected from households, or to adjust such data where it can be shown that there are errors or omissions to the data collected in the field.

55. Moreover, where the reported data in a traditional census contains errors of substance or omission, the incorrect responses may be edited, and/or the missing responses imputed using either the information recorded in the census itself from similar households or the data relating to the variable and individual in question in a corresponding register.

56. When using administrative data to quality assess census data (collected in the field), the Source, Data, and Process Stages are key. In addition, while outside the scope of these Guidelines, it is important to consider issues of circularity with respect to the overall design of the census. For example, when an administrative data source has been used

to impute missing values in the census data, or replace a census variable, it should not also be used in its QA.

2.2.5 Full register-based census

57. Finally, the last use case concerns measuring the quality of sources where the entirety of the census is conducted based on an administrative-based population register, instead of a traditional census methodology.

58. Clearly, the most widespread use of data from administrative sources occurs, by definition, when NSOs undertake a wholly register-based census. In the context of a full register-based census, assessing quality at each of the proposed Stages is vital.

59. The quality of census outputs is particularly dependent on the continuous improvement of quality during the Source, Data, and Process Stages. Depending on the possibility of proper linking with other registers, much additional information, although not recorded in the population register itself, may be added to a single record, such as language(s), ethnicity, educational attainment, parity, activity status, and occupation. In countries where register-based censuses are conducted, the quality and stability of the underlying administrative sources at these earlier Stages is such that register-based census results are considered the 'gold standard'. The collection of census data in this way does not, however, preclude the NSO from undertaking a field-based post-enumeration survey (PES) as a means of independently assessing the quality of the coverage or content of the counts in the resulting census database.

2.3 Types of administrative sources

60. As CES has noted (UNECE 2015) the development of a register-based population census system (whether within the context of a full register-based or combined approach) is a long process, which might take many years. Many countries will choose to continue to retain elements of a traditional data collection in some way even when they start to use administrative registers as an alternative data source.

61. This section of the Guidelines briefly discusses some of the administrative source types from which data are more commonly used by an NSO for the purpose of the census and the uses to which the data from each can be used. Where appropriate, these uses refer to the topics currently recommended by CES to be included in a census (UNECE 2015).

2.3.1 Use of administrative sources to support a traditional census

62. The extent an NSO uses administrative source data increases progressively from census to census as the move from a traditional field enumeration, through a combined approach, to a full register-based census develops. But even those countries that

continue to adopt a traditional census are likely to use administrative data increasingly to support their census operations.

63. Use of *address registers* is now commonly made by NSOs to create lists of dwellings and households. The registers can construct and map enumeration areas, resulting in balanced workloads for enumerators, or provide stratified sampling designs for post-enumeration or other sample surveys. The creation of a purpose-built address list, by an NSO, may involve the amalgamation of data from several separate and independent registers (that possibly constructed for different administrative purposes) to minimize under- or over-enumeration.

64. For example, lists of registered electors used for national and local voting purposes, or lists of dwellings used by local authorities for assessing rateable values, may not include all postal addresses used by national or commercial mail carriers. Moreover, buildings identified by a national mapping agency for the purposes of producing accurate large-scale official maps may identify the location of addresses that are not used for residential purposes, which are often excluded from the census address database.

65. Those NSOs undertaking a traditional census may use data from administrative sources to assess the quality of the data collected on the household questionnaire. Data from a *national vital registration system*, for example, can provide accurate information on the numbers of births and deaths during successive 12-month periods before the census. These data on the ages of young children can be compared and benchmarked to the census data. Similarly, data on address changes required to be reported to local authorities, for the purposes of maintaining *population registers,* can be used to validate the migration information since the previous census.

66. However, it should be noted where data are used to *assess the quality* of information provided on the census questionnaire and to *supplement* the census data to account for missing or incorrect responses, then the census can be considered to have progressed from a traditional to a combined approach methodology.

2.3.2 Use of administrative sources to derive populations or census characteristics

67. One of the uses of administrative sources for censuses is to provide data to derive the required output variables without having to collect the relevant information directly from the public. The type, structure, and content of such administrative sources will, of course, vary from country to country depending on the administrative purposes for which the data are used by the data providers. The most common generic types of registers used for this purpose are summarized below:

68. *Population registers* are registers (often held by a national government department and/or appropriate local authorities with responsibility for internal security matters) provide a frame of persons usually resident in a country. These registers are typically maintained to fulfil a legal requirement that both nationals and foreigners residing in the country should register with the local authorities. Aggregation of these local

accounts result in a record of population and population movement at the national and local level. Additionally, they often record information on some characteristics of individuals from which data on several core census topics can be derived, such as date and place of birth, gender, date of arrival/departure, citizenship, and marital status for each resident person by place of usual residence (however that may be defined).

69. *Social security registers* are registers held by official bodies typically for the purposes of national contributory social insurance programmes administration and benefits and allowances allocation (i.e., the unemployed, families, pensioners, and the disabled and long-term sick). The data from such registers may be used to derive census attributes for such topics as sex, age, marital status, unemployment status, income, and disability/health status.

70. *Tax registers* are registers held by national and local tax authorities for the purposes of the administration and collection of income tax, purchase taxes, building rates, and other national and locally levied taxes. The data from such registers may be used primarily to derive census data on personal or household income that might otherwise be difficult, or too sensitive, to collect directly on a household questionnaire. Other information held on such registers may include sex, age, marital status, employment status, occupation, place of work, and place of usual residence.

71. *Employment registers* are the registers from which the country's official employment and unemployment figures are derived. The data recorded may enable the NSO to derive census figures relating to the key socio-economic topics of economic activity, employment status, occupation, hours of work, and place of work (the latter two enabling analysis on travel-to-work patterns).

72. *Business registers* hold information to underly the provision of a range of services that can vary from country to country. Principally, their aim is to register, monitor, and store corporate information, such as a company's legal status, its headquarters, capital, and legal representatives. The NSO may be able to use this information to derive census data on economic topics, particularly industry.

73. *Education registers* are maintained both centrally and by individual educational and academic establishments, for the purpose of registering admissions and the performances of students, as well as the employment of teaching staff. The data held may be used by NSOs to create census statistics on attendance, literacy, and highest level of educational attainment – though it should be recognized that such data may often only refer to the current student population. Data on persons no longer formally attending an educational establishment must be obtained from other sources.

74. *Health registers* are maintained by locally based health authorities, for the purposes of providing health-related services, whether these are within the context of national health service or provided by insurance-based private agencies. The raw information they contain are usually treated as confidential but can be anonymized to a sufficient extent to allow them to be used by the NSO to create data on health status, geographic area, level of disability, and parity.

75. *Building and dwelling registers* are registers held usually by land and property valuation agencies and by local authorities responsible for the development of housing policies and urban planning. They may include information relating to the ownership, size, and physical construction of individual housing units, but may not necessarily relate these to the persons living in them. The data held may enable NSOs to obtain data to create census statistics relevant to the needs of a housing census, such as type of dwelling, floor space, floor level, construction materials, and period of construction, and may also distinguish between residential and non-residential buildings.

76. NSOs may also be able to access data from other administrative sources to provide topic-oriented census outputs. For example:

(a) *Registers of motor vehicles* may allow the collection of data on car availability,

(b) *Registers of foreign nationals* may provide information on migrants, year of entry into the country, citizenship, and asylum seekers,

(c) *Lists of military service personnel* may (if access by the NSO is permitted) indicate employment within the armed forces,

(d) *Prison registers* can provide some basic information on members of a population group that is particularly difficult to enumerate in a traditional census operation, and

(e) *Registers held by public facility service providers* (may offer information on the availability of household amenities such as piped water supply, electricity and/or piped gas, and sewage and waste disposal facilities.

Chapter 3. Quality framework

77. The quality of statistics depends on whether the statistical output satisfies its intended use. For example, the ESS definition of quality is derived from the ISO 9000 family of standards, "the degree to which a set of inherent characteristics of an object fulfils requirements" (ISO 2015). In official statistics, the object may include "a statistical product, service, process, system, methodology, organisation, resource, or [data] input" (Eurostat 2020, p.17). In a census context, the quality of administrative data used should therefore be considered in relation to the ways data are collected and processed by data suppliers and NSOs, through to the final census outputs.

78. Throughout the process, errors may occur which will compromise quality. Here, error is understood as the difference between a final estimate and the true population parameter it represents. This is highlighted in the Generic Statistical Business Process Model (GSBPM), which provides a standard structure to describe most statistical processes and includes 'quality' as an aspect which cuts across all its stages (Eurostat ESSnet MIAD 2014). In addition, Lothian et al (2019) argued for the need to understand the whole statistical production process when dealing with alternative data sources such as administrative data. Assessing the quality of administrative sources requires mapping the errors which may occur before and after the data is supplied to NSOs and determining how any such errors can be mitigated (e.g., through changes to collection, processing and/or integration with other sources). These Guidelines identify four broad Stages of census production: Source, Data, Process and Output. They then set out how the quality of administrative data may be assessed, by identifying the key quality dimensions at each Stage and the respective tools and indicators for QA.

79. As well as drawing on the GSBPM, this approach also draws on Daas et al. (2009), who identified cross-cutting areas which concern quality or 'views' of quality which they call 'hyper-dimensions', relating to the source, metadata and data (2009 p. 3). Each of these views comprise several data quality dimensions, assessed via quality indicators. In line with this approach, these Guidelines also identify quality dimensions, indicators and methods used in the assessment of administrative data, with a particular focus on censuses. At the same time, the focus is on census production stages which would be more intuitive for NSOs, for whom these Guidelines were written. Focusing on production stages highlights that quality is an inherent part of statistical design and enables NSOs to focus on the part(s) of the Guidelines which are most relevant to their use-case and/or current production stage.

3.1 Quality and error in censuses

80. Where official statistics are produced using a sample survey methodology, survey questions are designed and tested to reduce measurement errors, ensuring maximum accuracy and reliability. The error of the estimates produced are assumed to be caused

by deficient sampling and are typically measured and communicated using the Mean Squared Error (MSE) framework and/or through confidence intervals. However, such measurements do not capture non-sampling errors. These are particularly important in the context of censuses, where the aim is to the capture the full population. Generally, for statistics produced with administrative data, the key sources of error in the context of censuses are not sampling errors, but representation (coverage) and measurement errors (Zhang 2012). A common practice is to adjust census estimates based on the results of a survey (see Chapter 7).

81. Where administrative data and other alternative data sources such as Big Data are used in censuses, the range of possible errors is greater than in a traditional census, because data collection processes are not controlled by NSOs. Zhang (2012), drawing on Groves et al. (2004), distinguishes between two broad types of error in statistics produced using administrative data: measurement and representation errors. The first relates to errors in the measurement of characteristics (e.g., age, gender etc.), while the second to errors in the coverage of population units or objects (e.g., individuals or households in a census).[11] Zhang also distinguishes between the quality of single sources as provided by data suppliers and the quality of transformed and/or integrated sources, after processing by the NSO. This approach is mirrored in the Guidelines which assess the quality of single administrative sources (see Source and Data Stages below) and integrated sources (see Process and Output Stages), with a particular emphasis on identifying measurement and representation errors.

82. Furthermore, the total survey error (TSE) framework has also been adapted to assess the quality of administrative data. In contrast to MSE, TSE identifies a wider range of errors including validity, frame/coverage, nonresponse, measurement, processing, and model errors. TSE frameworks have sought to capture how a variety of errors accumulate throughout the statistical design and methodology, resulting in the final error of any given estimate. This approach has been adapted to report the quality of statistics which integrate administrative data (e.g., Reid, Zabala and Holmberg 2017; Rogers and Blackwell 2020). At the same time, the quality of statistics cannot be reduced to assessing error alone. When considering the integration of data from an administrative source into the census design, the impact of such integration on quality should be assessed in terms of the extent to which it adds error or uncertainty to the outputs, vis-à-vis the advantages of integration e.g., reducing response burden, increasing timeliness, reducing costs. These Guidelines identify additional dimensions

[11] Based on Zhang (2012), In relation to input data, measurement errors relate to differences between supplied and target characteristics (e.g. gender, sex, age, ethnicity, occupation etc.) and include several types of error within variables including relevance (definition misalignment), mapping (errors in the re-classified measures due to poor equivalence between supplied and target classifications which may therefore require adjustments, e.g. through imputation) and comparability errors (errors between the re-classified and adjusted measures). Representation errors relate to the difference between the units supplied and the target units. They include errors relating to over and under-coverage (lack of alignment with target population), identification (errors in classifying a unit based on inconsistencies across multiple sources) and unit errors (errors in the statistical creation of statistical units of interest where they do not exist in any available data source).

which can affect the overall quality of census outputs including the Institutional Environment and the need to balance quality dimensions to meet user needs.

83. Following these Guidelines will help ensure that census estimates are based on the most appropriate sources and methods and are not misleading. At the same time, consideration should also be given to the way administrative sources are intended to be used in the census design (see Chapter 2). Given the variety of possible uses, this framework should be used flexibly and adapted to the level of quality required by different uses of administrative data by the NSO and different statistical requirements from the users of census statistics including the generally public, organizations, local and national governments. Inevitably, QA relies on skilled professional judgement throughout the entire statistical production process, from collection to publication, to meet the data users' needs.

3.2 Measuring quality

84. The quality of census estimates produced using administrative sources is particularly difficult to assess and/or measure due to the complexity and multi-dimensionality of the data used. As noted above, many factors affecting quality are not quantitatively measurable. Moreover, what constitutes 'fitness for purpose' and high-quality statistics will necessarily vary from one user to another (e.g., some users may prioritize timeliness over accuracy). It is important to assess/measure administrative data quality across the key dimensions which will be of interest to NSOs and data users. What is meant by assessment and measurement needs further clarification.

85. These Guidelines distinguish between *assessing quality*, meaning a qualitative evaluation, and *measuring quality* – meaning attaching a quantitative metric to this evaluation of quality. Where it is not possible to produce indicators for quantitative measurement, or where they have not yet been developed, these Guidelines recommend a qualitative assessment of their impact on quality. In addition to these, there are several additional principles which guide the production of official statistics (UNECE 1992), and which are applicable throughout the full statistical process and the wider NSO environment (e.g., commitment to quality, independence, data protection, statistical confidentiality, etc.). These themes are relevant for all statistical processes and are not fully covered within the scope of the present Guidelines. However, it must be acknowledged that a census that uses administrative sources usually relies on data that were produced outside of the statistical system, in a different organization over which the NSO usually has no control.[12] For this reason, the impact of using these outside sources on these principles, must be considered carefully.

[12] In some cases, the NSO has some control over the register. In Switzerland for example, the Federal Register of Buildings and Dwellings or the Enterprise Register are part of the Federal Statistical Office. It might be feasible in a long-term perspective to integrate certain suitable registers within NSOs. Implications/advantages of this are briefly discussed in Section 4.2.5.

3.3 Stages of quality assessment

86. To ensure these Guidelines are easy to follow, the QA of administrative sources is considered across four broad stages of the census lifecycle. These are applicable regardless of census type (see Chapter 2). While statistical design is never entirely linear, thinking of how to carry out QA in this way should enable statistical producers to quickly identify the key quality considerations most relevant to their own circumstances. The Stages are:

    (a) **Source Stage**: A metadata-based QA of new or re-supplied administrative sources to be used in the census. This Stage does not require NSOs to be in possession of the actual data, but it is crucial for the Stages that follow,

    (b) **Data Stage:** The QA of the raw administrative data supplied to NSOs by administrative authorities (data suppliers). This will require NSOs to validate the data supplied against the lessons learned from the Source Stage. As well as basic validation, this Stage includes any processing required to establish the quality of the data supplied vis-à-vis what was expected and comparisons with alternative sources,

    (c) **Process Stage:** The processes often carried out on administrative data sources, in censuses, to transform the data for census use and/or to improve quality. The processes identified include:

        (i) Record linkage,

        (ii) Statistical registers and the 'signs-of-life' methodology,

        (iii) Enumeration of population units: administrative data-based models,

        (iv) Conflict resolution/decision between sources, and

        (v) Editing and imputation.

    (d) **Output Stage:** The overall QA of the census outputs produced using administrative data. While this is not conceptually that different from the assessment of the outputs of a traditional census, these Guidelines attempt to identify where this may differ.

87. The figure below summarizes the Quality Assessment (QA) Stages:

Figure 1: The Quality Assessment Stages

STAGE 1: SOURCE – A metadata-based QA of new or re-supplied administrative sources to be used in the census.	**STAGE 2: DATA** – QA of the raw administrative data supplied to NSOs by administrative authorities.
STAGE 3: PROCESS – The processes carried out on administrative data sources, in censuses, to transform the data for census use and/or to improve quality.	**STAGE 4: OUTPUTS** – The overall QA of the census outputs produced using administrative data.

88. These Guidelines are focused primarily on input quality of administrative sources and the Source and Data Stages. However, Process and Output quality are included for completeness. Ultimately, the question of whether the administrative data are good enough for census purposes can only be answered based on their planned use or the census output they generate. The four Stages cannot meaningfully be separated. For the first two Stages, the Guidelines identify in detail the key data quality dimensions for assessment, the key tools used in completing their assessment and where possible, set out the criteria against which the assessment may be carried out. In addition, key issues in the QA of Process and Output when census estimates are produced using administrative data are briefly reviewed. Key Recommendations are provided across each of the Stages, which are summarized within Chapter 8 of the report, along with suggestion for areas of further work.

3.4 Quality dimensions

89. As previously noted, the quality of statistics and of administrative data is understood to encompass multiple dimensions which are not reducible to coverage or measurement errors. For instance, statistics which are accurate but out-of-date are of limited use. The quality dimensions identified by ESS include:

(a) Relevance,

(b) Accuracy and reliability,

(c) Timeliness and punctuality,

(d) Accessibility and clarity, and

(e) Coherence and comparability.[13]

90. However, for assessment of administrative data these "standard quality dimensions [alone] are not always applicable" (Daas et al 2008, p.2). At the same time, they do capture many of the relevant aspects of administrative data quality. This report, therefore, includes dimensions beyond those included in the ESS dimensions, whilst also capturing the important aspects of the ESS dimensions with respect to the assessment of data sources. The following tables set out the assessment dimensions of administrative sources for each of the Stages described in these Guidelines.

[13] Alternative dimensions are used by various NSOs (e.g., Statistics Canada 2017, Australian Bureau of Statistics 2009). Overall, these alternative frameworks cover approximately the same content albeit using different terminology or classifications.

Table 1: Quality dimensions at Source Stage

	Quality Dimension	Definition
SOURCE STAGE	Relevance and Accuracy	The degree to which the administrative data source meets the needs of the census. This covers the overlap between the census target population, concepts, and definitions (relevance) and the degree to which the data correctly describe the phenomena they were designed to measure (accuracy).
	Timeliness	The lapse between the end of the reference period to which the information pertains and the date on which the information becomes available to the NSO.
	Coherence and Comparability	The degree to which the administrative source can be successfully combined with other sources used in the census, including linkability.
	Accessibility and Interpretability	The ease in which the NSO can obtain the administrative data, covering the impact of any restrictions, privacy and security, public acceptability of the use, the ease of data transfer and receipt, and the availability of metadata.
	The Institutional Environment	Organizational factors affecting the data supplier's capacity to supply data to the quality expected. Covering the strength of the relationship, previous experience, existence of formal agreements, risks associated with the status of the supplier and the supplier's quality standards.

Table 2: Quality dimensions at Data Stage

	Quality Dimension	Definition
DATA STAGE	Validation and Harmonization	The data files provided to the NSO are in a readable format. Further data validation and harmonization arrangements are in place upon data transfer to the NSO. This is done to confirm that the expected variables/units/reference period/formats have been supplied and to ensure data processing by the NSO is consistent across census use cases.
	Accuracy and Reliability	The accuracy, completeness (for variables and population coverage) and coherence of the data supplied matches the requirements of the specific census use case for which it will be used. Comparisons with alternative sources reveal acceptable levels measurement or representative errors.
	Timeliness and Punctuality	The timeliness and punctuality of the data supplied matches the requirements of the specific census use-case for which it will be used.
	Linkability	Adequate linkage variables are available (i.e., either common unique identifiers or a combination of variables which enable identification) and these are of sufficient quality to enable data linkage.

Table 3: Quality dimensions at Process Stage

	Quality Dimension	Definition
PROCESS STAGE	Accuracy of record linkage	Where multiple sources are linked (to each other or census responses), the linkage is accurate and unbiased, thereby improving the overall quality of the census methodology and/or dataset.
	Coverage and coherence of statistical registers and admin-based enumerations	Where census (sub)population registers are constructed, or when administrative data are used to supplement census collection, they adequately cover the target population/variables, thereby improving the overall quality of the census methodology and/or dataset.
	Accuracy of conflict resolution	Where different sources are linked and the same attributes are available in them, methods for deciding between sources improve the overall quality of the census methodology and/or dataset.
	Accuracy of editing and imputation	Where census variables/units are derived/constructed through imputation or modelling techniques, this derivation is accurate and unbiased, thereby improving the overall quality of the census methodology and/or dataset.

Table 4: Quality dimensions at Output Stage

	Quality Dimension	Definition
OUTPUT STAGE	Relevance	The degree to which statistical outputs meet current and potential data user needs.
	Accuracy & Reliability	The closeness between an estimated result and the unknown true value – and how reliable these are over time and geographies.
	Timeliness & Punctuality	The lapse of time between publication and the period to which the data refer, and the time lag between actual and planned publication dates.
	Accessibility & Clarity	The actions taken to help the data user find and understand the data in which she or he is interested.
	Coherence & Comparability	The degree to which data can be compared over time and domain. The degree to which data that are derived from different sources or methods, but which refer to the same phenomenon, are similar.

Source: Eurostat 2013 and 2018

3.5 Feasibility research

91. It is unlikely that new administrative data sources will be integrated into census production without prior feasibility research by NSOs. The quality of a data source may be established by acquiring test data and assessing its quality at the various stages suggested in these Guidelines. This will aid design thinking, that is, designing a census methodology that makes the most of the available administrative data and considers the impact of its use on the quality of the census overall.

92. Firstly, feasibility research involves developing a detailed understanding of the data supplier's collection processes, the population covered, and variables included within the source as well as data accessibility (the Source Stage, Chapter 4). Secondly, supply, acquisition and ingestion of test data should be rehearsed, and test data examined, in detail, to identify quality issues and define cleaning and harmonization, along with validation checks (the Data Stage, Chapter 5). Thirdly, when data from multiple registers are combined, they can be used for verifying data quality and for selecting the most reliable variables and values, in accordance with the developed methodological rules (the Process Stage, Chapter 6). Finally, estimates produced using test data can be compared with previous census estimates or another such 'gold standard', contributing to an assessment of the overall quality of the output (the Output Stage, Chapter 7).

93. Generally, census characteristics cannot be acquired directly from administrative data sources because they have been designed for other, non-statistical purposes. Most of the definitions and classifications used by administrative authorities are different from standard statistical definitions. Data from multiple registers may be used to construct or derive certain census characteristics, while other characteristics may be covered by duplicate information in several registers. This makes feasibility research key for developing methods for the derivation of census characteristics.

94. Census methodologists should address the following main challenges when deriving census characteristics:

    (a) Ascertaining the international census standard (definition, classification, etc.) applicable to the target census characteristic,

    (b) Comparing and contrasting census definitions and classifications with the definitions and classifications used in the administrative source,

    (c) Testing the accuracy of the administrative data against alternative sources and work collaboratively with data suppliers to eliminate/mitigate any shortcomings,

    (d) Determining which and how many sources are required, to derive and quality assure each target census characteristic,

    (e) Establishing optimal rules for deriving each census characteristic and developing the necessary data processing software, optimized for the quality of outputs sought, and

    (f) Taking steps to ensure creation of the necessary register or register part (e.g., suggest amendments in register procedures, the legal environment, etc.), where characteristics are not covered by any administrative sources.

Box 2: Feasibility research in Estonia

In 2016, a pilot Population and Housing Census (PHC) was conducted in Estonia. Data for the mandatory census variable "Year of arrival in the country", was available in the country's administrative population register. However, following an analysis of distributions, the variable in the register could not be directly used for the census. In the first half of the 1990s (when the register was first established), the year 1994 or 1995 was recorded as the year of arrival in the country for many persons. Comparing the distributions of arrival year in the register to alternative migration data sources, showed immigration, in Estonia in the 1990s, was too high. To address this issue, data from PHC 2011 and different population register variables (e.g., entry creation date and country of birth) were used, so that the derived census variable could correspond, as closely as possible, to the definition in the UN Principles and Recommendations for Population and Housing Censuses (2008, Revision 2).

Box 3: Feasibility research in Israel

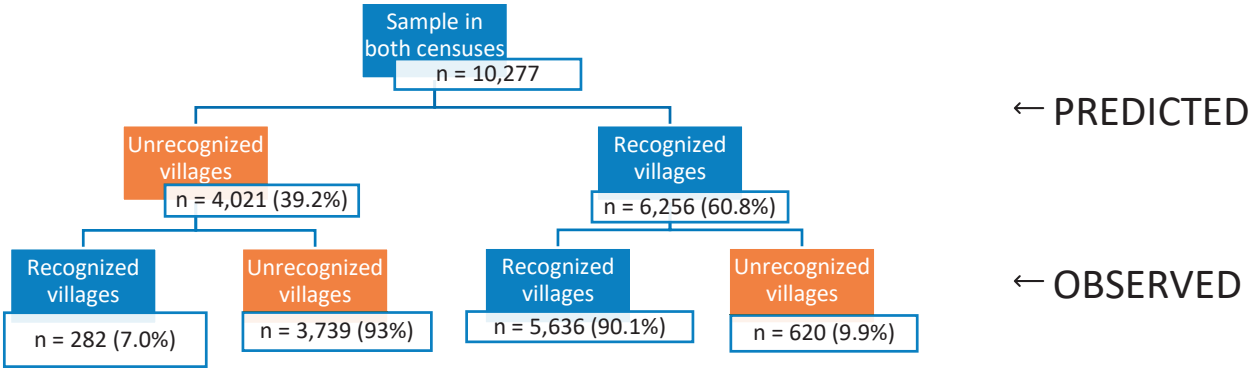
In Israel, feasibility research has been undertaken to develop methods for choosing the best address for the hard-to-reach Negev Bedouin population, by comparing estimates produced using administrative data, to those produced with the last traditional census in 1995. The Negev Bedouin is an ethnic group that includes approximately 283 thousand Arab Muslims, living in the Negev Desert. They are a unique population as they traditionally live as nomadic tribes with a distinctive culture (e.g., 16 per cent of men are polygamous). In the traditional 1995 census, Bedouin households were interviewed, and their places of residence marked on maps. However, this population is considered hard-to-reach as about one third of this population lives in unrecognized villages, which are not connected to public infrastructure like electricity, water, or paved roads. Additionally, they have low levels of engagement with government agencies.

Research was conducted to explore the potential of administrative data in determining the geographic location for this population, based on the Central Population Register (CPR). Within it, every person has a unique Personal Identification Number (PIN), which is linked to individuals' day-to-day interactions with government agencies and services. Additionally, each CPR record contains links to the records of the person's father, mother, spouse, and demographic variables.

By comparing the CPR to previous census data, it is known this register carries inherent errors and incompatibilities with census definitions, including omission of residents (foreigners), inclusion of non-residents (emigrants) and purposely incorrect address registration – 20 per cent of the population does not report their most recent address. Furthermore, there are limitations particular to the Negev Bedouin population. Bedouins, in the Negev Desert, are expected to be registered in the CPR, their address registration does not allow for an accurate location. This is especially true for individuals living in unrecognized villages, registered under tribe names rather than the geographic area they live in, because the tribes could be scattered throughout the geographic area of the Negev Desert. Moreover, Bedouins in the unrecognized villages intentionally register themselves in the CPR, as if they live in recognized villages to obtain services, such as educating their children in schools at the recognized villages. Finally, even Bedouins, who have a "real address" in one of the recognized villages may not be recorded with sufficient accuracy.

In our research, the first stage (initial location) was to use the current CPR address and their 1995 CPR address, to locate individuals at the reference day. For example, if their address has not changed in the CPR between the years 1995-2019, it means that they still live in the same geographic area reported in the 1995 census with their offspring. The second stage was to use a 'signs of life' methodology (see Chapter 6) based on other administrative sources (e.g., marriage records, address changes, local tax, water service, students at school and electric service), improve the accuracy of location data. The results were then compared to those obtained from the traditional 1995 census. This method was tested and re-evaluated by replicating the methodology with 2008 census data (Figure). It was found that approximately 90 per cent of the sample found to be living at the same geographic area was predicted using this method. This result was further confirmed through a small field test (n = 110).

Figure 2: Results predicted through the administrative method (level 1) versus the observed method (level 2) on the 2008 census in Israel



95. The challenges referred to above are best addressed through feasibility research, as in the examples from Estonia (Box 2) and Israel (Box 3). The second example in particular highlights both the challenges and opportunities administrative data can present in the production of statistics about hard-to-reach populations (see Box 4). Reaching an adequate harmonization of register and census concepts can be a complex and time-consuming activity which should not be underestimated. It is recommended that the inclusion of administrative data sources into census production should be preceded by adequately resourced feasibility research which provides a 'proof of concept' for the planned integration of administrative data into the census production. Additionally, making the four Stages of QA an integral part of feasibility research will enable NSOs to directly apply the lessons learned from feasibility research to the census production context and to better inform users on the quality of data sources.

96. Drawing on a review of the literature and the experience of NSOs, the remainder of these Guidelines focuses on tools and indicators for assessing the quality of administrative data sources, against each of these dimensions. In the chapters that follow, in addition to the work of Daas and colleagues, these Guidelines also draw on comprehensive sets of quality indicators for administrative data developed by others (e.g., Iwig et. al. 2013; Eurostat ESSnet MIAD 2014; Eurostat ESSnet KOMUSO 2016, 2019).

Box 4: Hard-to-reach population groups

Both traditional census collections and administrative-based enumerations may fail to reach certain groups of the population (United States Census Bureau 2019). Conversely, administrative data may include individuals who have not been captured via a traditional census collection, such as those who are unwilling or unable to participate in the census, but who have interacted with an administrative process.

Hard-to-reach population groups are often of significant interest for policy reasons. They may include young and mobile individuals; immigrants, asylum seekers and refugees; those belonging to certain communities or ethnic groups (see Box 3); and vulnerable people, such as the homeless and people living in poverty or with poor living conditions. Even if these groups have been captured in the administrative data, their information may be significantly out of date or poorly recorded.

Chapters 4, 5 and 6 provide approaches for assessing the coverage of administrative sources and population registers, but it is important to consider the 'hidden' populations and to develop strategies to reach these individuals. Such strategies may include working with community groups, non-governmental organizations, housing associations etc. which hold information about these groups, or who can advise on other ways to reach the population of interest (Statistics Canada 2020). This may also entail exploring new sources, including commercial sources (where this is legally permissible and acceptable under ethical standards) that could provide information about individuals missing from government administrative systems.

Another approach is to encourage hard-to-reach groups to interact with the administrative sources, which may include building trust and addressing concerns about the use of their data; or it could involve incentivizing people to interact with an administrative system. The response to the COVID-19 pandemic is an example: education, social and health care organizations have taken steps to work with the public in some countries to engage with administrative systems to ensure they can adequately support the whole population, including through testing and vaccination programmes.

Members of hard-to-reach groups can be more likely to be associated with conflicting information across administrative sources, particularly for their name and address. For example, members of the highly mobile student population may have their parents' address on a tax register, their (out-of-date) first year term-time address registered with the education authority and their (current) second year term-time address registered with the health authority. In such instances, it is important for the NSO to develop an in-depth understanding of how different groups might interact with the different administrative sources, to be able to make decisions about which information is likely to be relevant or correct (see Chapter 6).

…

Differences in names can be particularly problematic for members of minority national, linguistic or ethnic groups and communities, where different names or different renderings of a name may have been provided for the same individual across different sources. This can significantly increase the likelihood of false or missed linkages when building a statistical population register from various administrative sources, which can lead to coverage error (see chapter 6). To address this, it is important to understand the different naming practices used by minority communities and ethnic groups, which can then be incorporated into the linking methodologies.

Reaching hard-to-reach populations and capturing their data accurately remains an important challenge for NSOs across the UNECE and beyond (Section 8.2).

Chapter 4. Source Stage

97. This chapter provides a guide to the key quality dimensions, the process of assessment, and associated tools and indicators for evaluating the quality of administrative data sources to be used in the census production – both at first acquisition and when they are regularly re-supplied to the NSO. Normally, no data are accessible during this stage. The search for information about the administrative data sources begins, most likely through communications and exploratory meetings between the NSO and the administrative data supplier.

98. The evaluation in this stage should lead to a recommendation on whether to proceed with the acquisition initiative (or continue the re-supplying of the data). If the decision is to go ahead, the administrative data supplied will undergo a more detailed evaluation at the Data Stage.

99. It is necessary to assess source quality both at the first acquisition of an administrative data source and in each instance when it is re-supplied to the NSO. This is because the characteristics of any data set that has previously been supplied might differ due to changes in concepts, classifications, collection methods, and so forth.

4.1 Source quality dimensions

100. The data quality dimensions to consider at this stage of QA are:

(a) Relevance and Accuracy,

(b) Timeliness,

(c) Coherence and Comparability,

(d) Accessibility and Interpretability, and

(e) The Institutional Environment.

The dimensions are described below, with the processes, tools, and indicators for assessment provided in the sections that follow. It should be noted that failure to reach minimum acceptable quality against any of the dimensions cannot be compensated by success in the other dimensions.

4.1.1 Relevance and Accuracy

101. Relevance reflects the degree to which an administrative data source meets the needs of the NSO with respect to the intended use. To be deemed relevant, the administrative data source must fulfil the reasons for its acquisition. This could be with respect to reduced costs or respondent burden, improvements to the quality of census outputs,

or through the delivery of enhanced or new census outputs. To achieve this, the administrative source should be representative of the population of interest for the census (the target population) and the measurements from the population should align with the needs of the census. A key part of the assessment of Relevance is understanding the context in which the administrative data have been collected.

102. As part of the assessment of Relevance, the Accuracy of the administrative data is also considered. Accuracy refers to the degree to which the data correctly describe the phenomenon they were designed to measure. It is important to understand how the collection, processing, and QA carried out by the data supplier might affect the Accuracy of the resulting data and their usefulness.

4.1.2 Timeliness

103. Timeliness refers to the period between the date to which the information pertains and the date on which the information becomes available to the NSO. The timeliness of the information will affect its relevance.

4.1.3 Coherence and Comparability

104. Coherence reflects the degree to which the administrative data can be successfully combined with data from other sources used by the NSO, i.e., census data, within a broad analytical framework, over time. The use of standard concepts, classifications, and target populations promotes coherence within and between censuses. A clear understanding of the operational definitions used by the data supplier, the purpose of data collection, and the impact on comparability of changes in an administrative source over time and across geography factors into assessing coherence.

105. It is often a requirement to link an administrative source at the level of the census statistical unit to integrate the data into the census design. The Comparability of identifiers, across the different data sources to be linked, is a consideration under coherence.

4.1.4 Accessibility and Interpretability

106. Accessibility and Interpretability refer to the ease with which the NSO can obtain (and understand) the relevant administrative data in their entirety. This includes understanding any restrictions (legal and those imposed by the supplier), privacy and security, public acceptability, the ease of data transfer and receipt (suitability of the medium for transferring data and costs), and the availability and clarity of documentation and metadata. It is crucial that the use of the administrative data source is based on a legal framework that gives the NSO the unequivocal right to access and use the data and the metadata for statistical purposes.

4.1.5 The Institutional Environment

107. The Institutional Environment refers to the organizational or institutional factors that may have an impact on the data supplier's capacity to supply data to the quality expected and to the agreed timetable (punctuality). Components of this include:

    (a) Strength of the relationship with the data supplier, comprised of the effectiveness of communication channels and how responsive the supplier is to the NSO's requests,

    (b) Existence (or potential for) formal agreements and risks associated with the status and complexity of the supplier organization, and

    (c) Quality standards and procedures adopted by the supplier organization(s).

4.2 Tools and indicators

108. The quality of an administrative source should be assessed against the quality dimensions outlined in the section above. The following section provides guidance on the process of assessment, including tools and indicators for evaluating an administrative source for use in the census.

4.2.1 Relevance and accuracy

109. An understanding of the differences between the administrative population and the required census population, and between the measures/variables in the administrative source and the required census characteristics is important to assess Relevance and Accuracy. The errors arising from these differences are referred to respectively, as representation and measurement error (Zhang, 2012). At the Source Stage of assessment, it is possible to gain some understanding of these errors and their impact on Relevance (as outlined in the subsections below) based on metadata about the data supplier's purpose and methods of data collection. The impact of representation and measurement errors on Accuracy and reliability are also considered at the Data and Process Stages (Source Stage - Chapter 5 and Process Stage - Chapter 6).

4.2.1.1 The Census target population (representation)

110. To assess Relevance, the NSO must determine whether the set of objects in an administrative data source align with the population units of interest for the census (the target population). An object is the basic element of the population for which information is collected, for example, a person, household, dwelling, event, or transaction. The following indicators are proposed for establishing Relevance, with respect to representation. Against each indicator is a series of questions to help guide the assessment:

(a) Alignment (of the objects) with the census target units.

(i) How comparable are the administrative objects with the census target units?

(ii) What definitions, methods and processes are used to identify and include an object in the source?

(iii) Are there any laws or regulations that define the objects?

(iv) Are any checks carried out by the data supplier to ensure the definitions hold?

(v) In the case of misalignment with the census units, is a transformation possible that could meet the census needs?

(b) Coverage (of the set of objects) against the census target population.

(i) Does the coverage of the objects meet the needs of the census?

(ii) Is there evidence of undercoverage (objects that are missing from the source, but are part of the census target population) and overcoverage (objects that are in the source, but are not part of the census target population) that would impact the usefulness of the source?

(iii) Are there any differences across geographical areas due to differences in practices by the data supplier or due to legislation that need to be considered?

(iv) Are there any rules, legislative or regulatory requirements, including penalties for non-compliance that may impact the inclusion or exclusion of objects in the source?

(v) What methods and processes are adopted by the data supplier to include new objects that meet the required inclusion criteria/definitions (e.g., registration procedures) and to remove objects that no longer align with the target population for the administrative source (e.g., deregistration procedures)?[14]

(vi) In the case of coverage errors, are there other data sources that could be used in combination with the source to overcome under- or overcoverage in the source?

[14] An administrative source or register might be built by the data supplier through linkage of multiple sources. In such instances, it is important to understand the quality of linkage and any potential for error – including coverage errors (see Chapter 5 and Chapter 6 for details on data linkage and the associated error). For example, the Colombian Sisbén Database (System for the Identification of Beneficiaries of Social Programmes) is linked with the deaths database of the National Registry, however, there is evidence of missed links.

4.2.1.2 The census variables/characteristics (measurement)

111. To assess Relevance, the NSO must also determine whether the information collected from administrative data objects meets the needs of the census, with respect to the target concepts (e.g., employment status, household size, tenure status, etc.). The following indicators are proposed for establishing Relevance, with respect to measurement:

    (a) Availability of the target variables/characteristics.

        (i) Does the administrative source include the variables needed for the census and the intended use of variables in the census?

        (ii) Do the variables/characteristics broadly cover the relevant census reference period?

(b) Alignment of variable concepts, definitions, and classification with the census needs.

  (i) Are the administrative concepts, definitions, and classifications comparable with the census needs?

  (ii) Is there a difference between the data supplier's ideal target concepts and their operational target measure used in the collection?

  (iii) In the case of misalignment with the census concepts, definitions, and classifications, is a transformation possible to meet the census needs?

(c) Alignment/measurement error against the census reference period.

  (i) What is the collection frequency for a variable/characteristic?

  (ii) Are there known delays between an event or phenomenon occurring and being captured in the administrative source (e.g., parents may not have to register a birth for several weeks on a country's birth register)?

  (iii) Are there time stamps recorded on the data source to indicate what period a data item refers?

  (iv) Are there any incentives or disincentives for a data subject to update their information and when their circumstances/information changes on the administrative source (e.g., benefits or penalties for not doing so/or doing so)?

(d) Quality of collection and potential for measurement error against the census concepts.

  (i) What is the data supplier's purpose for collecting the data and how might this influence the quality of the data?

  (ii) Are there any legal obligations, targets, or incentives (or lack of incentives) that could influence the data collection quality?

  (iii) Does the data supplier's collection process raise any concerns about the quality of the variables, including the potential for any biases? This could include whether data are recorded by proxy and therefore not reported directly by the data subject (increasing the potential for misreporting).

  (iv) What procedures are in place to validate and check data on entry by the data supplier?

  (v) Are there any incentives or disincentives for data subjects to provide complete and accurate information to the data supplier?

(e) Quality of data processing and potential for processing error by the data supplier.

(i) Does the processing carried out by the data supplier suggest the resulting data quality will meet census needs?

(ii) What checks are carried out by the data supplier to assure quality?

(iii) Are data edited or imputed? If so, when, and how, and is there an indicator on the data source to identify when an edit and imputation has taken place?

(iv) Are there any rules, regulations or incentives on the data supplier that may impact on the way data are processed?

112. At the Source Stage assessment, the evaluation against the indicators is usually based on a qualitative assessment (e.g., indicating whether the need is fully met, partially met, or not met against each indicator, with an explanation of why, based on the answers to the question set). A quantitative assessment of representation and measurement error is carried out in the Data Stage (based on analysis of the data) under the dimension of Accuracy and Reliability (Source Stage, Chapter 5).

113. The assessment against the indicators should inform a decision (often based on experience and expert judgement) on the use (or continued use) of a source in the census. The decision should take account whether the data source can meet the needs of the census (e.g., reductions in costs and respondent burden, improvements, and enhancements to the census outputs), set against any costs or risks (referenced under the Institutional Environment and Accessibility dimensions below).

114. There are various quality frameworks described in the literature that provide similar indicators as in this chapter against the different dimensions of quality, along with question sets and scoring systems for informing the assessment [e.g., Daas et. al. 2009; Eurostat ESSnet MIAD 2014; Iwig et al 2013; Statistics Canada's *Administrative Data Evaluation Guide* (Lavigne and Nadeau 2014); Statistics Austria's *Quality Assessment of Administrative Data, Documentation of Methods Framework* (Statistics Austria 2019)]. The New Zealand case study (Section 4.4.1) provides a practical example of a framework being used to assess administrative sources for use in the census.

4.2.2 Timeliness

115. An administrative source may cover the relevant timeframe for the census, but to be useful it will also need to be available in time for the census. The following indicator may be used to assess Timeliness:

(a) Timeliness and frequency of supply against census needs.

(i) What is the time lag between the end of the reference period for the administrative data and the date the data are available to the NSO?

(ii) How frequently can the data be supplied to the NSO, to meet the needs of the census?

(iii) Are there any requirements, in terms of the delivery method and required formats and data structures, the NSO uses that could impact the data supplier's timeline?

(iv) Is there enough time from when the data become available, for the NSO to process the data for use in the census?

116. In instances when the data are unlikely to be available in time, the NSO may wish to establish whether a provisional version of the dataset can be made available ahead of schedule. In such cases, the dataset may be incomplete and subject to higher levels of error. There may be trade-offs to consider between the Timeliness of the data and Accuracy.

117. As referenced in the dimensions of Institutional Environment below, it is important to include the delivery dates with the data reference periods, in formal agreements with the data supplier. Although the data may be available on time to meet the requirements of the data supplier, they may not necessarily be delivered to the NSO in time, while the latter carries formal responsibility for timely delivery of the census.

4.2.3 Coherence and Comparability

118. It is important to assess the degree to which an administrative source can be successfully combined with other data sources for census use. The information gathered to assess Relevance can also be used to assess Coherence. This includes information about the differences between the underlying concepts, definitions, classifications, and methods between the administrative data source and the other data sources for combined use in the census.

119. For a full register-based census, it is important to analyze the census characteristics and administrative data source; mapping and ascertaining the extent to which the information in the administrative data source enables the derivation of the relevant census characteristics. In particular, the NSO should establish whether the data in the registers conform to the definitions of the census characteristics. In the case of partial or no conformity, the NSO should examine the causes of non-conformity between the census characteristics and the information available in the administrative data source.

4.2.3.1 Comparability

120. Administrative data are subject to changes and difference over time and across geography due to changes in legislation, regulation, and procedures, which can affect the concepts, definitions, classifications, and coverage of a source. More generally, the changes can impact all the indicators under representation and measurement, as outlined under the dimension of Relevance. This is of particular importance for the census, when stability over time can be a key concern. The following indicator is proposed to assess Comparability:

    (a) Comparability over time and domains

        (i) Are there any changes across time or differences across geographical areas affecting:

        - The definition and coverage of an object on the administrative source relevant to the census?
        - The concepts, definitions, and classifications associated with the variables on the administrative source relevant to the census?
        - The data collection, processing, and QA procedures that could impact the quality of the source data for census purposes?

4.2.3.2 Linkability

121. A consideration under Coherence and Comparability is the ease with which an administrative data source can be linked with other relevant datasets for the census. The Estonian case study (Section 4.4.3) provides an example of how it is possible to link different administrative data sources with a few different unique identifiers. The following indicators are proposed to access the Linkability of a source:

    (a) Presence of a unique key for linkage.

        (i) Does the source include a unique identifier that is common with the unique keys required for the census linkage?

        (ii) Is the identifier available for all the relevant objects on the source?

    (b) Presence of a unique combination of variables for linkage.

        (i) Does the source include a unique combination of variables (e.g., name, age, and address), which could be used for the census linkage?

        (ii) Are the unique combination of variables present for every object on the source?

122. The quality of linkage variables is also assessed at the Data Stage (Chapter 5) and the quality of the linkage process is covered as part of the Process Stage (Chapter 6).

4.2.4 Accessibility and Interpretability

123. The following indicators are proposed for the assessment of Accessibility and Interpretability:

(a) Restrictions on data access and use,

(b) Public acceptability,

(c) Ease of data transfer and receipts, and

(d) Interpretability of the source – clear and comprehensive metadata.

124. The sections below provide details of the relevant information for assessment against each of the indicators.

4.2.4.1 Restrictions on data access and use

125. It is important to identify any restrictions that may impact the NSO's ability to access and use an administrative source. For example, existing data protection restrictions embedded in legislation can impose certain limitations on the data acquisition and processing, especially when data are protected with extra security measures or laws at the national level. Legislation may be specific to particular data sources (for an example, refer to the case study from Estonia in Section 4.4.3) or may be more generic allowing the NSO access to the data sources as and when required, subject to the agreement of the data supplier. The data supplier may also impose further restrictions on the data and the permitted use(s). These may include:

(a) Suppression of records or variables,

(b) Disclosure treatments (pre-delivery), such as encryption of identifiers, perturbation, banding, or top coding of the supplied data,

(c) Restrictions on how the data can be used,

(d) Restrictions on the retention of data and deletion and disposal rules, and/or

(e) Rules on disclosure methods that must be applied by the NSO, affecting the census outputs.

126. **The NSO should establish a list that describes any restrictions that apply, so an assessment can be made on the impact (and risks) of the restrictions on the use of an administrative source in the census. As part of the assessment, the NSO should consider whether it has the capability to abide by the restrictions. This could include the** technical and procedural safeguards the NSO must adopt. The safeguards would generally form part of a Memorandum of Understanding (MOU) or Data Security Agreement with the data supplier. In particular, the MOU may describe how Personally Identifiable Information (PII) will be protected.

4.2.4.2 Public acceptability

127. Whether an NSO can access and use a data source in the census may depend on public acceptance. It is essential that the public understand and are supportive of, or at least not hostile to, new approaches and uses of their information. If the public or specific groups within it are opposed to the use of an administrative data source, there is a quality risk. For example, this could change the way the public interact with the census, or with an administrative source used in the census. The NSO should therefore be transparent about the use of administrative sources in the census, highlighting the benefits to the public, whilst providing assurances concerning privacy and data security.

128. To assess public acceptability, the following tools or processes can be used:

    (a) Public consultation or engagement,

    (b) Privacy Impact Assessment (PIA), or

    (c) Data Ethics Assessment.

129. A ***public consultation or engagement*** may be carried out by the NSO on the use of administrative data in the census (or for other statistical research or outputs). This can take various forms, including formal consultations, questionnaires (through surveys or the NSO inviting feedback on its website), qualitative research into public attitudes, or the use of Citizens' Panels. Citizens' Panels aim to bring together members of the public (to be representative of the population, or to reflect different population groups of interest) to assess their views and opinions. Engagement with key groups in the society, such as indigenous communities and ethnic minorities, is essential in determining and responding to their specific needs and concerns about the use of data pertaining to them, especially when the proposed uses are not those for which the data are/were initially collected.

130. A ***Privacy Impact Assessment (PIA)*** is a formal process resulting in a document that describes the processes, findings, and results that helps the NSO consider the effects of a new programme or service (or proposed policies and plans) on the privacy of individuals. As a risk management tool, used in the planning phase of a programme or service initiative, PIAs assist organizations to more fully consider the privacy implications of a given proposal. PIAs are also used to ensure data stewards can meet their obligations under the General Data Protection Regulation (under European Law). A PIA can be applied to the various ways an NSO may wish to use a data source in the design of the census. The New Zealand case study, in Section 4.4.2, describes the privacy risks involved and the mitigation measures used by the NSO to eliminate or reduce each of the risks.

131. A ***Data Ethics Assessment*** is carried out to establish whether the access, use, and sharing of public data for research and statistical purposes is ethical and serves the public good. NSOs may use an ethics self-assessment tool (e.g., UKSA 2020), and/or they may use a formal body to provide expert advice or endorsement, such as a Data Ethics

Advisory Committee (e.g., the UK National Statistician's Data Ethics Advisory Committee[15]).

132. The findings from public consultations and engagements, PIAs and ethics assessments, can help an NSO assess public acceptability of the use of administrative sources in the census (and for other statistics produced by the NSO).

Box 5: Statistics Canada's Trust Centre

Statistics Canada has a Trust Centre that outlines how information is protected, placing privacy as a top concern. This includes how societal needs for new data insights and the protection of privacy are balanced, using a modern "necessity and proportionality" framework. The Trust Centre provides clear and comprehensive information to reassure the public on the use of their data, including the use of infographics and short videos accessible via the website. One such video ('Joe Anonymous') explains how the public's data are used, including the importance of bringing together multiple sources. There is an emphasis on the work and culture within Statistics Canada to protect data, including a promise to protect the identify of people, their families, and their businesses.

Openness and transparency are at the heart of the Trust Centre and information about administrative sources to be acquired and used by Statistics Canada is published (and updated regularly) on the website. [Available at https://www.statcan.gc.ca/eng/trust]

4.2.4.3 Ease of data transfer and receipt

133. The data supplier might use quite different data models, formats, schemas, software, and hardware to those with which the NSO is familiar. This includes how data are held and transmitted, (including the security arrangements for transmission). The data structures could also be complex and file sizes extremely large (particularly for transaction data). It is important that the NSO understands such differences and complexities, to assess whether it is feasible to receive and ingest the datasets into the NSO's systems. This process can also include negotiations with the supplier on the development of processes and systems to facilitate the transmission of datasets in a format that meets the needs of the NSO. This can, however, be a time consuming and costly process.

134. More generally, cost is a key factor to be considered when assessing ease of access. This can include costs imposed by the data supplier, or costs incurred by the NSO in developing its capability to receive an administrative dataset (purchasing new software

[15] For more information see https://uksa.statisticsauthority.gov.uk/about-the-authority/committees/national-statisticians-data-ethics-advisory-committee/#:~:text=The%20National%20Statistician%E2%80%99s%20Data%20Ethics%20Advisory%20Committee%20%28NSDEC%29,advise%20the%20National%20Statistician%20on%20the%20ethical%20

or hardware). It is important to assess any costs against the expected value a new administrative source will bring.

135. In practice, detailed arrangements for the data transmission to the NSO, including the files structures, format (e.g., flat files, a relational database; SAS, Excel, or text formats, etc), the variables, the supply frequency, delivery dates, data standards, and agreed costs, would be included in Data Sharing or Delivery Agreements between the NSO and the data supplier. Such agreements would be signed by authorized managers at each of the organizations.

4.2.4.4 Interpretability of the source – clear and comprehensive metadata

136. An assessment of Interpretability relates to the existence and availability of comprehensive and clear metadata and documentation about the administrative source. Without this, it is not possible to understand and assess the administrative source against the intended use. The metadata should include details about:

(a) The administrative organization,

(b) Purpose of the collection,

(c) Concepts, definitions, classifications, and protocols used,

(d) Collection and processing of the data,

(e) Validation and quality assurance methods and procedures, and

(f) Reporting units and variables, including data dictionaries, file structures, formats, and relationships within the data.

137. This information is important for the assessment against the other quality dimensions outlined in this chapter. It will often be the case that clear and complete metadata will not exist for all aspects of an administrative source in the initial phase of exploring the source for use by the NSO. It may be necessary to work with the data supplier to build the relevant metadata. This relies on good communication with the data supplier and a willingness of the data supplier to work with the NSO (see Institutional Environment below). Depending on the complexity of an administrative source, an NSO may decide to set up secondments for staff to work within the administrative organization to develop an in-depth understanding of a source. Once located and understood, it is important that the metadata are recorded, stored, and maintained (a ‘metadata base’) to allow easy location in the future.

Box 6: Metadata templates for assessing administrative sources

New Zealand

Statistics New Zealand has a Guide to Reporting on Administrative Data Quality *(Stats NZ 2016), with an associated Metadata Information template for Admin Data (available at Stats NZ 2020). The template is a useful resource for capturing metadata about an administrative source, covering general information about the administrative organization, the data collection, population objects and variables, changes over time, and aspects of accessibility.*

The Netherlands

Statistics Netherlands Checklist for the Quality Evaluation of Administrative Sources (Daas et al 2009), provides a useful template (the Annex to the paper) for recording information and metadata about a source. The ordering of the dimensions and indicators within the template directs the user through the recording and assessment against the metadata efficiently – ensuring problems are revealed early on before moving on to later stages.

The Statistical Network on the Methodologies for an Integrated use of Administrative Data (SN-MAID) Project

Deliverable B2.3 (Source) and B2.4 (Metadata) (SN MAID, 2014) provide checklists including quality indicators and fields for recording metadata about an administrative source, which is used to assess the quality of the source for use in statistics. The checklists draw on the work of Daas et al. (2009).

4.2.5 The Institutional Environment

138. The NSO is completely reliant on the data supplier to collect, process, and deliver the administrative data to the quality expected and to the agreed timetable. The NSO is also reliant on the quality of the information the data supplier provides about the data (see Interpretability, Section 4.2.4.4 above) and about any unforeseen changes to the data. It is important to assess confidence in the data supplier's ability to meet these needs. The following indicators are proposed for the assessment of the Institutional Environment:

    (a) Strength of the relationship with the data supplier,

    (b) Previous experiences with the data supplier,

    (c) Existence of formal agreements,

    (d) Risk posed by the status of the data supplier, and

    (e) Data supplier's quality standards.

139. **The strength of the relationship.** There should be processes in place for managing the relationship with the data supplier; ensuring there is a continuous dialogue. These should include mechanisms for:

    (a) Communicating the NSO's requirements to the data supplier,

    (b) Communicating, in a timely manner (by the data supplier), any changes that might affect the source data (e.g., changes to the legal basis for the data, concepts and classifications, and the processes and procedures for data collection, management, and supply),

    (c) Raising any questions with the data supplier about the data source, and

    (d) Feeding back findings to the data supplier that could result in improvements to the source.

Box 7: A Quality Assurance Toolkit: Communication with data suppliers

The UK Statistics Authority's Administrative Data Quality Assurance Toolkit (UKSA 2015b) describes "practice areas" associated with data quality, including an area for communication with data suppliers. The area covers the importance of collaborative relationships with data collectors, suppliers, IT specialists, policy, and operational officials. Highlighting the importance of formal agreements detailing arrangements (see below), as well as regular engagement with all involved parties. There are three levels of assurance proposed, depending on importance: basic, enhanced, and comprehensive.

140. **Previous experiences.** This includes how responsive a data supplier has been to the NSO's requests, whether any issues have arisen with previous supplies of data (e.g., late delivery, unexpected errors), whether the data supplier has provided accurate information in the past about source data (this might have been established through checks at a later stage by the NSO).

141. Formal agreements. This includes whether written agreements (legal or otherwise) exist or can be developed, covering:

    (a) Roles and responsibilities of the NSO and data supplier. This could include whether the NSO has a role in the approval of any changes to an administrative source used (or to be used) in the census,

    (b) Legal basis for the data supply and any security and/or confidentiality requirements, and

    (c) Specification of requirements, according to the Data Sharing or Delivery Agreement referenced in Section 4.2.4.3 above.

Box 8: Statistics Netherlands System of Base Registers

In the Netherlands, a system of administrative base registers is adopted, comprising 13 registers on population (residents and non-residents), addresses and buildings, enterprises, real estate (boundaries, ownership, value, etc.), topography (maps: land, water, roads), cars (model, colour, ownership, etc.), taxable income, labour (wages, employers, social benefits, etc.) and subsoil (sewerage, cables, etc.). The system of base registers is based on legislation and supports the production of statistics (including the census) by Statistics Netherlands.

The use of data from base registers is compulsory for governmental agencies. The objective is that all users of the system contribute to the data quality. Therefore, users are obliged to notify the owners of the base registers if they have alternative data that are considered of better quality (except for the NSO, due to legal considerations). Users of base registers can rely on their validity. Statistics based on these registers require only a limited amount of data editing. The registers adopt standardized approaches and identification numbers for linkage, so the statistical data are generally coherent.

Each base register has its own project board. All stakeholders' groups are represented on these project boards. Project boards operate within the legal framework and see to it that the register data fulfil the legal requirements (quality, completeness, etc.) and that the data are correctly applied. Project boards act as advisory boards to the responsible cabinet ministers and meet a few times per year.

142. **The status of the data supplier**. The risk associated with the status of a data supplier should be assessed by the NSO, taking into consideration whether the data supplier is an established, stable, and reputable organization. The NSO should consider whether there is any legal or regulatory basis to the administrative function the data supplier carries out that would make the NSO confident about the sustainability and quality of the source. Risks associated with the complexity of the collection, processing, and delivery of the source should also be considered, since there may be multiple bodies or organizations involved in producing the source data, each impacting the quality of the final data supplied.

143. **Data supplier's quality standards**. An assessment of whether the data supplier can meet the quality expectations of the NSO should be made. The assessment should consider information on the principles, standards, and guidelines adopted by the data supplier for assuring quality, including the procedure(s) in place covering collection, processing, and the supply of data to the NSO. Evidence of how the data supplier checks whether the standards are being met is valuable. This can be accomplished through internal or external audits by regulators or professional bodies. The data supplier may also produce quality reports, which should be reviewed by the NSO. A more detailed assessment based on key aspects of the administrative source is included under the Relevance quality dimension above.

144. Following an assessment of the data supplier based on the quality criteria outlined above, the NSO can evaluate the risks associated with the data supplier delivering the administrative data on time and with the required quality.

4.3 Recommendations

(a) Identify relevant and promising administrative sources for use in the census (see Chapter 2).

(b) Set out clearly the required target population, variables, and concepts, along with the anticipated outcomes for using an administrative source in the census on which to base the assessment.

(c) Understand the restrictions and challenges to acquiring and integrating administrative sources into the census, including where changes may be needed to the NSO's methods, processes, and computing systems.

(d) Build and maintain clear and comprehensive metadata capturing all relevant quality information about a source (this will provide a valuable resource for the NSO). Structure metadata using an appropriate, agreed-upon metadata standard format is important (Cornell University Research Data Management Service Group 2020).

(e) Develop a good understanding of the data supplier, the context, and purpose of the data collection and the quality standards they uphold.

(f) Build strong relationships with the data supplier, to ensure effective sharing of information – building a common understanding of each other's needs.

(g) Put in place formal agreements, which outline clearly the NSO and data supplier requirements, roles, and responsibilities.

(h) Carefully assess the value of acquiring and using an administrative source, against any risks and costs. This can be with respect to the stability of a source over time and the risk of a data supplier failing to deliver data on time or to the expected quality.

(i) Ensure there is a sound legal basis to the receipt and use of an administrative source, with effective safeguards in place to protect the privacy of the data subjects.

(j) Be clear and transparent about the use of administrative data, showing evidence that the benefits outweigh any privacy concerns.

(k) Accept that objects, definitions, concepts, and time reference periods within an administrative source may not align with the census targets. It will,

therefore, be necessary to transform data and make judgements on what levels of misalignment are acceptable.

(l) Assess quality on a continuous basis (using the process and tools outlined) – responding to any anticipated or known changes to a source.

(m) Document and publish the strengths and weaknesses associated with administrative sources used for the census, so that data users have confidence in the data and can take account of any limitations.

(n) Be prepared that it will take time to understand and acquire administrative data sources for use in the census, particularly, where a programme of work is required to develop registers for use in the census (as per Estonia case study).

4.4 Case studies

4.4.1 New Zealand: Source assessment

145. In March 2012, the New Zealand Government agreed to a census transformation strategy. Part of the first phase of that programme was to complete a broad look at the potential for administrative data to produce the long-form (social and economic) information currently provided by the census (O'Byrne et al 2014). This process identified administrative data sources related to the census topics and used quality measures to assess how likely those sources were to satisfy the information needs previously met by the traditional census. The investigation did not include population counts and demographic breakdowns which were investigated elsewhere.

146. The purpose of this work was to provide an early indication of the likely ability of existing administrative data sources to produce census long-form information and to guide decisions about where to direct more in-depth analysis.

147. The steps in the process included:

(a) Identifying data sources – achieved through tapping into existing Stats NZ use of administrative sources, web searches, and contact with government agencies.

(b) Understanding the nature and content of potential administrative data sources – achieved through review of publicly available information, discussions with experts from Stats NZ, and the source agencies.

(c) Using five critical quality dimensions for QA.

(d) Assigning a quality rating - the likelihood that administrative data could satisfy a census topic.

148. The quality measures used in the assessment were adapted from existing quality dimension frameworks (such as the Stats NZ quality model, Eurostat, 2009 and 2011). The five measures identified as relevant for this assessment were: relevance, accuracy of coverage, accuracy of linkage, timeliness, and accessibility. These quality measures were chosen because they are strongly discriminatory, in the sense that they are essential for the use of administrative data for census information and are also measures for which reasonable judgements can be made from metadata.

149. This assessment was done by jointly assessing as many administrative data sources as may be needed to satisfy that census variable. For each variable, each quality dimension was rated as excellent, good, or poor, which determined an overall rating of 'likely', 'possible' or 'unlikely' to be satisfied by administrative data sources. The key questions considered for each dimension are outlined in Table 5.

Table 5: Key questions for each dimension

QUALITY MEASURES	MAIN QUESTIONS FOR ASSESSMENT
Relevance	How close are the administrative data to the statistical concept? (The census topic is used as a proxy for the statistical concept) Who/what should be included in these data (target population)? Who/what is included in these data (observed population)?
Accuracy of coverage	Are there missing people or responses (undercount)? Are there duplicate records or other people who should not be included (overcount)?
Accuracy of linkage	Is it possible to link the data to the census population or dwelling lists?
Timeliness	How frequently are the data supplied? How long after the reference date are the data available to Statistics NZ?
Accessibility	Are there privacy or legal issues around using these data? Are there any other barriers to access?

Source: Stats NZ

150. The study showed which administrative sources would be most important in providing census-type information and detailed analysis of most of the variables identified as being 'possible' or 'likely' has now been completed. One of the most important findings was that most of the current census variables were unlikely to be obtained from administrative sources, and a survey component would still be needed. The quality ratings used are shown in Table 6.

Table 6: Quality ratings

Quality measures	Definition of quality rating		
	Excellent	Good	Poor
Relevance	The data collected in the administrative sources are close to the statistical concept.	The data collected in the administrative sources are not the same as the statistical concept, but are close, or related to a similar statistical concept that might be acceptable.	The data collected in the administrative sources are not at all relevant to the statistical concept we are interested in.
Accuracy of coverage	The coverage (net, under and over) is like the census.	Most of the population is covered and those who are missing are 'missing at random'.	Coverage (net, under and over) is very low, or there is bias in the distribution of missing values.
Accuracy of linkage	Data have excellent individual identifiers that can link the units in one dataset to other datasets.	Data have good individual identifiers.	Data have no individual identifiers. Data linkage is not possible.
Timeliness	Data are updated at least every year and available to Statistics NZ within two years.	Data are updated at least every two years and available to Statistics NZ soon after.	Data are updated sporadically, or with delays of more than two years.
Accessibility	No privacy or legal concerns exist. Statistics NZ understands the data and has a good relationship with the administrative data owner.	Some privacy or legal concerns exist with one or more key datasets.	Serious privacy or legal concerns exist. No relationship with administrative owner or no history of using the data.

Source: Stats NZ

4.4.2 New Zealand: Privacy impact assessment

151. Privacy impact assessments (PIAs) are a useful tool when considering the accessibility dimension of quality, specifically the legal implications of administrative data use and for building public trust. In New Zealand, the Office of the Privacy

Commissioner provides guidelines and templates to support organizations completing PIAs. This guidance outlines 12 privacy principles (these principles are drawn from the Privacy Act 1993 and range from collection of data to use of unique identifiers) to be considered as part of a PIA. It also includes guidance on the key questions to ask during the process, some of the common risks to be aware of, as well as possible mitigation strategies to consider. Prior to the New Zealand 2018 Census, Stats NZ engaged an external organization to complete an independent PIA on the planned use of administrative data in the census. Stats NZ later completed and published an additional PIA covering the intention to extend the use of administrative data to mitigate the lower-than-expected response rate. The overarching goal of a PIA in this context is to bring together information about what, why, and how a NSO wants to use specified administrative data, and to assess the potential value gained against a range of privacy considerations.

152. Key topics covered in the second edition of the 2018 Census PIA include:

(a) Information about the benefits of using administrative data in the census and detail about how security is managed through the process of constructing the final census dataset,

(b) A summary of relevant legislation,

(c) A summary of the privacy assessment for each of the 12 privacy principles,

(d) Recommendations and action plan to minimize harm, and

(e) A risk and mitigation table containing risk ratings (consequences and likelihood) for each of the 12 privacy principles along with some additional principles to reflect obligations under the Statistics Act 1975.

153. The PIA concluded that using administrative data in the census is lawful, safe, and beneficial to New Zealanders.

4.4.3 Estonia: Improving data through legislation

154. Statistics Estonia (SE) carried out work during 2010-2013 in cooperation with data suppliers and scientific communities. The goal was a QA of administrative sources to be used in the census production. The requirements were analyzed for those census characteristics laid down in the regulations of the Council of Europe and the European Parliament, as well as the regulations of the European Commission (European Commission 2008). The coverage of each census characteristic was mapped, and suggestions were made for the formation of census characteristics in future and for quality analysis.

155. Based on this analysis, it was concluded that as many as 20 different administrative sources (held by nine different authorities or ministries) would be necessary to provide data of sufficient scope and quality. SE was given a mandate to determine the minimum

universal criteria for all those registers that were required to provide the data to meet the needs of users.

156. SE was made aware of the limitations in use of registers, the main cause of which was the lack of sufficient metadata information provided by register owners. The metadata that did exist had been compiled merely to satisfy the administrative purposes for which the data were collected and were often not relevant for the statistical use of the data. There were often conceptual inconsistencies between the definitions and classification adopted in the register and those necessary for use in the census. Nor was coverage of the base population or the availability of topic variables in the registers always compatible with national census requirements – particularly where variables related to self-defined statuses.

157. The target for 2014 was to agree on a package of legal and organizational measures to improve the quality, timeliness, and coverage of the dataset for the register-based census based on the bottlenecks pointed out in the methodological report.

158. Starting in 2014, SE participated actively in formal deliberations with the relevant authorities with the aim of making the necessary revisions to the legal acts governing the specific data sources required for census purposes. National authorities were requested to disclose in their legislative proposals if a new administrative data source was going to be established or an existing one modified. Any data sharing mode was also to be prescribed. Provisions for the scope to start or improve the data collection process were also covered in the legislation.

159. SE was charged with the responsibility for improving data quality in the registers. Accordingly, it devised a roadmap based on suggestions given by experts and prepared an improved business model to facilitate better cooperation between administrative registers. SE worked on an action plan up to 2020, which comprised different tasks for data source owners. The most urgent of which was to create a legislative environment for adding any necessary new characteristics to the registers (such as, occupation, industry, and place of work) and for updating these characteristics in the registers (including the tax board registers, planned working register, business register, etc).

160. The next critical task was to improve the accuracy of residence registration to gain better coverage for households, institutional populations, and tenants. SE initiated a national project, launched by the Ministry of Interior, for adding archival data on families and relationships between family members to the Population Register. This would improve several census characteristics (such as legal marital status and relationship within household).

161. Amendments to the legislation relating to foreigners has helped to improve data collected on the foreign-born population. This has allowed improvements to registration procedures to obtain more complete information on new arrivals (including characteristics on education, marital status, and relationships between family members).

162. Altogether, about 20 different suggestions were made to data source suppliers to improve data source quality using the legislative framework.

163. To create linkable data some basic rules, prescribed by special governmental regulations, were adopted by 16 register owners from 2016:

    (a) All data in registers for persons, enterprises, and dwellings must be identified (using unique codes),

    (b) Address data should be used in all registers according to the established standard, and

    (c) Metadata should be available and updated.

164. Another important aspect related to the source data quality used concerned data transfer. It is necessary to have a fool-proof and reliable environment for transferring data from different registers to the NSO. In Estonia, such an environment, named X-Road, facilitates the transfer of large quantities of data between institutions or the provision of individual persons with their personal data. Data capture for census purposes was allowed, according to the government regulation, through X-Road. Previously, data owners used e-mail or file transfer protocol (FTP), as encrypted comma separated value files (.csv) or Microsoft Excel files (.xlsx).

165. The quality standard was prepared for assessing data sources. In the quality standard, the numerical values were fixed for accepted biases in census variables and hypercubes, when the following quality dimensions of data were taken into consideration:

    (a) Relevance (coverage, conceptual differences, etc),

    (b) Timeliness & Periodicity (last date of record update, lags in supply, etc.), and

    (c) Accuracy: especially of linkage variables to assess linkability of source.

166. By 2020, SE had derived 38 different variables relating population and dwellings required by the current EU census programme from 26 different administrative sources (Statistics Estonia 2019).

Chapter 5. Data Stage

167. This chapter provides a guide to the key data quality dimensions, tools, and processes for the assessment of administrative data at the Data Stage of production. This refers to the QA 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 Source and Data Stages together provide an overall assessment of input quality, with respect to an administrative data source (see UNECE 2018, Chapter 6).

168. The quality of administrative data at the Data Stage is assessed against several dimensions including readability and validity, accuracy and reliability, timeliness and punctuality, and linkability. These dimensions are explored (Section 5.1), along with the tools and indicators for their assessment or measurement (Section 5.2).

169. At the Data Stage, it is possible to establish a baseline for the quality of the individual administrative datasets supplied, based on edit and validation rules. These should be developed based on the known properties of the data captured in the Source Stage assessment and the statistical design requirements. These may also be improved over time. To enable this baseline assessment, a level of data processing, including linkage to other sources may be necessary. This processing is limited to making the data usable to undergo validation checks and to establish their quality vis-à-vis other sources.

170. The results of the QA at the Data Stage inform the NSO of any corrections that are necessary (through the resupply of data by the supplier). They also inform the necessary processing of the data for use in the census design, through an understanding of the error that must be accounted for or adjusted (see Chapter 6). Furthermore, they provide information needed to understand the implications of any errors in the sources on the final census outputs (see Chapter 7), which would need to be communicated to data users.

5.1 Data quality dimensions

5.1.1 Harmonization and Validation

171. A general assessment of the accessibility of the data is part of the QA at the Source Stage (see Chapter 4). However, it is crucial for the NSO to ensure that the transmitted data files are in the required ‘readable’ format; the databases are structured in a way which can be ingested and read by the NSO’s systems. Where this is not the case, the NSO may be unable to process the transferred data files.

172. Additionally, further data harmonization and validation arrangements should be in place upon data transfer to the NSO, ensuring consistent use across census use cases. The Data Stage provides the opportunity to validate the dataset supplied against

metadata collected at the Source Stage, the reference period, and other data requirements for specific variables. To do this, some basic harmonization may be required, such as ensuring that all missing values are coded in the same way. Harmonization arrangements and validation checks may be developed based on previous experience working with test data (see Section 3.5 on feasibility research). They may be improved over time, as the NSO receives regular re-supplies of the same data source.

5.1.2 Accuracy and reliability

173. An assessment of the accuracy of the input data should be conducted to identify **measurement** and **representation errors** within the administrative dataset (see Chapter 3), as described in Zhang's (2012) two-phase life-cycle model and adopted in QA literature (Stats NZ 2016 and Eurostat ESSnet KOMUSO 2019).

5.1.2.1 Representation Errors

174. Representation errors (errors relating to the target units, see 0) might occur if data are not reported correctly to the data supplier resulting, for example, from non-registration or delayed self-registration on an administrative register (e.g., birth, death, or full population register). Some data records may not be transmitted to the NSO because of technical problems or be transmitted with errors if units are not maintained properly by the data supplier (resulting in duplicates). It should be noted that representation errors may cause measurement errors where the unit of statistical measurement changes. For example, a person missing in the administrative population register may lead to an understated value for the variable 'size of household'. For an overall coverage assessment of a dataset, an examination of both over- and undercoverage is needed. Undercoverage may be of particular importance with respect to 'hard-to-reach' populations (see 0, Box 4).

5.1.2.2 Measurement Errors

175. Implausible or missing values are indicative of measurement errors (that is, error within variables) and may reduce the accuracy of the raw data (see 0). To assess whether a value is implausible or missing, it is important to examine not only specific records, but also variable distributions for all records. Reasons for a lack of accuracy might be technical, such as errors in the process of data transfer. Or a lack of accuracy may be systematic. For example, this may result from an inadequate submission or maintenance on the part of the data supplier, particularly if the variable is not of administrative importance for the data supplier. Therefore, the variable is not systematically recorded (such as a person's occupation in the Austrian tax register)

(Eurostat ESSnet KOMUSO 2019). Missing values may also be due to an administrative source (or variables on a source) being only recently established[16].

5.1.2.3 Re-supplied data

176. In general, the data supplier will improve the quality of the data through regular maintenance and updating of the data source. Many registers, however, may be subject to changes in the structure and/or content resulting from internal administrative requirements and processes. These changes may in turn lead to a loss of quality, particularly regarding comparability. When data are being supplied periodically, there is a need for additional, longitudinal QA. Re-supplied data offers the opportunity to assess the reliability of specific variables, such as the closeness of initially supplied values to the subsequently re-supplied values within a dataset. Usually, it is assumed that more up to date values are more accurate.

5.1.3 Timeliness and punctuality

177. It is important that the difference between the reference date to which the data refer and the date they are supplied to the NSO is kept to a minimum. The longer the delay, the less relevant those data become, even though they may still be accurate (UNECE 2018, p.15). This gap between reference date and acquisition by the NSO is referred to as **timeliness**.

178. **Punctuality** is the difference between the expected date of delivery and the actual date of delivery. This is important as the NSO will usually have a responsibility for producing census outputs to an agreed schedule and would not want any a delay in the supply of the census data to affect this.

5.1.4 Linkability

179. Often, determining the quality of a dataset will require its linkage to another dataset for comparison. Also, if the NSO relies on more than one source of administrative data for its census, it is necessary to be able to link data from the different sources at the unit/record level (see Chapter 6). The degree of success of such linkage will affect both the accuracy and the relevance of the input data.

180. A common unique identifier reduces the effort required to link the data by making it easier to evaluate the completeness and accuracy of matching. In the absence of such an identifier, it is more difficult to link data reliably. In this case, record linkage using multiple variables that are common to the units in each data source (typically, name, date of birth, sex, and address) may be possible (see Chapter 6). In this case, the NSO needs to be assured that such 'matching' variables are of sufficient quality in all sources,

[16] The Register of Higher Education in Hungary only contains data on individuals that graduated after 2005. The Austrian Central Population Register was established in 2001, but the attribute legal marital status was not introduced until 2006, resulting in missing values on the register.

otherwise the quality of record linkage, and thus the reliability of the data, will suffer. Even where probabilistic matching methods are used, the quality of the linkage variables will ultimately have an impact on the risk of false matches and false nonmatches in later production stages (cf. Eurostat 2014, Section 3.5.2) (see also Chapter 6). Enhanced validation checks should be developed for variables to be used in linkage.

5.2 Tools and indicators

181. The following tools and indicators are useful to the NSO in assessing the quality of raw data against the dimensions discussed in Section 5.1 above. This application of the tools and indicators supports a consistent assessment across different sources, to decide whether administrative data are fit for purpose.

5.2.1 Harmonization and validity

182. To ensure the readability and validity of the transmitted data files, it is crucial to implement technical checks to validate the data files against the expected data format. If this validation fails, the NSO may require the data files to be re-submitted in the correct format. Before such checks can be carried out, data must often undergo a basic cleaning and/or harmonization process, so that they are comparable to other sources and are optimized for use with the NSO's statistical software.

183. Examples of the **harmonization** processes include consistent coding of missing values, formatting of date variable types, and removal of or otherwise dealing with duplicate records from the dataset. Data harmonization rules should be agreed within the NSO and applied consistently to the data, regardless of the different census use cases for which they are intended. Agreed upon, organization-wide harmonization standards will aid coherence and comparability. The data harmonization processes, and validation results should be documented.

184. Previous literature has identified specific indicators which can be used to assess the **validity** (e.g., Daas et al. 2009; Eurostat ESSnet MIAD 2014; Cerroni, Di Bella and Galiè 2014). These include:

    (a) The variables supplied are correctly named and formatted (e.g., numerical, categorical, data variables, etc.),

    (b) The correct reference period has been supplied,

    (c) The variables match the expected pre-defined content, established through the metadata collected at the Source Stage (and through feasibility research, where possible), and

(d) No unexpected differences between current and previous supplies of the data source are found with respect to number of records and variables (further examined below).

185. Linking records from the supplied data to another reliable data source at the unit level provides a tool for determining whether the correct reference date is supplied (Asamer et al 2016). It is also possible to check variables with a date specification to determine whether they are compatible with the census reference date. A correct reference date is important, especially for changeable variables such as current activity status for seasonal workers. When possible, any such inconsistencies should be corrected at the Process Stage (see Chapter 6).

186. These Guidelines are not prescriptive about how harmonization rules and validity checks should be applied, as these need to be developed based on production needs and specific planned uses of administrative data within the census design.

5.2.2 Accuracy and reliability

5.2.2.1 Representation Errors

187. A variety of indicators can be used to measure the accuracy of the supplied objects or units, providing an assessment of representation errors in the data (see Daas et al. 2009; Eurostat ESSnet MIAD 2014; Cerroni, Di Bella and Galiè 2014)[17]. Basic indicators include:

(a) Total number of objects or statistical units (for comparison against expected count),

(b) Per cent of duplicate objects or statistical units.[18]

188. Additional indicators suggested by Cerroni, Di Bella, and Galiè (2014, p.128) include:

(a) Per cent of 'inconsistent' objects or statistical units, i.e., involved in non-logical relations with other (aggregates of) objects or statistical units[19],

[17] As noted in the glossary, in some of the literature (e.g., Zhang 2012), the term 'object' is used to refer to the units within an administrative dataset. The term is used to distinguish between units in the administrative data and the statistical units after these data have been transformed in some way. This is particularly relevant in cases where the unit (or 'object') in the administrative register differs from the target statistical unit. For example, in a tax register where the units of a yearly tax return (i.e., the same person may make several returns in one or multiple years) are converted into individual 'people'.

[18] The percentage of duplicates identified might only be a lower bound due to undetected duplicates. If the share of undetectable duplicates is too high the indicator will lack precision.

[19] An example of such an error is an adult in a dataset that is assigned to multiple households as their main residence (the number of inconsistent units of this type divided by the total units would be calculated).

(b) Per cent of 'dubious' objects or statistical units, i.e., involved in implausible but not necessarily incorrect relations with other (aggregates of) objects[20].

189. A broad assessment of over- and undercoverage of the data can be made by computing and comparing the total number of objects, as well as cross-tabulations of frequency/per cent across characteristics (e.g., sex, age, geography, etc.) on an aggregate level, between the administrative source and other/alternative sources taken as reference or a comparative **'gold standard'** (e.g., a complete base register[21] or traditional census). The indicators suggested by Cerroni, Di Bella and Galiè (2014, p.129) include:

(a) Undercoverage:

(i) Per cent of objects of the reference source missing in the supplied source.

(b) Overcoverage:

(i) Per cent of objects in the source not included in the reference population, and/or

(ii) Per cent of objects in the source not belonging to the target population of the NSO.

190. The above indicators are subject to two key assumptions. First, a suitable 'gold standard' must be available to compute over- and undercoverage. For instance, deceased persons may be still (incorrectly) reported by a country's central population register but may be identified as deceased in a central social security register. Second, it should be clear which objects of the 'gold standard' dataset should be included to compute undercoverage. An example of this is school-aged children in the base register should be largely covered in the register of enrolled pupils.

191. Finally, comparisons can be made to determine the percentage of objects which are inconsistent within and across sources. An example of inconsistent objects might be where each row within an administrative dataset represents an event of registration (e.g., doctor visit) which includes, name, address code, date of registration and perhaps date of deregistration. Two objects relating to a single person are inconsistent if the period of registration of the objects at different addresses overlaps. The per cent of inconsistent objects provides an indicator of error. However, as noted by Zhang, object-level analysis has its limitations as sources may differ at the micro level but result in similar statistical measures such as means, medians, etc. Unit-level analysis "may fail to

[20] A soft rule can be defined to identify the occurrence of dubious objects in an administrative source. For example, the number of employees working for more than four employers during the same period was used to detect dubious objects within the Italian Social Security Data (Cerroni, Di Bella and Galiè 2014, p. 138).

[21] In the literature, base or core administrative registers are often distinguished from additional registers (e.g., Daas et al. 2009). Base or core registers are those assumed to have the most exhaustive coverage of the target resident population.

reveal such statistical equivalence" (Zhang 2012, p.45). In addition, where unit-level comparisons are made between multiple sources, it is important to note the possible impact of selectivity bias within the linkage process on any resulting differences.[22]

5.2.2.2 Measurement Errors

192. **Statistical techniques and metrics**, such as frequency distributions, can reveal unexpected patterns and outliers. For example, a cross tabulation of age and marital status may lead to the identification of implausible cases, such as a 5-year-old child that is married. Other examples include the comparison of date of birth with that of other events in the German case study in Section 5.4.1, and the cohesion analysis of address data in the Polish case study in Section 5.4.2. Observed patterns might indicate systematic measurement errors, as also illustrated in the case study from Germany (Section 5.4.1). Note that if inconsistencies are identified and the data supplier cannot fix such problems, then certain edits (as part of the Process Stage, Chapter 6) may be necessary.

193. As above, previous literature contains **basic indicators** to measure the completeness of the variables supplied within administrative datasets at the aggregate level (e.g., characteristic variables such as age, sex, ethnicity, etc.) (see Daas et al. 2009; Eurostat ESSnet MIAD 2014, Cerroni, Di Bella and Galiè 2014). These include:

    (a) Number and per cent of missing values within key variables,

    (b) Number and per cent of out-of-range values within key variables (for example a recorded age of 120 years),

    (c) Number and per cent of implausible values (based on, for example, cross-tabulations of different variables),

    (d) Prevalence of unexpected frequencies, patterns, or outliers, based on frequency/distributional analysis of key variables (aggregate comparisons with external sources, as well as expert knowledge can be used to identify data oddities also), and

    (e) Prevalence of rounding for the main variables of interest (can be detected by analysing the distributions).

194. The degree of consistency of the supplied data at the aggregate level, namely that relationships between related variables are consistent and not implausible, provide a measure of the accuracy of variables. However, to assess consistency at the micro level,

[22] Selectivity bias within linkage refers to situations where linkage variables or methods are more or less accurate for certain groups, particularly with respect to hierarchical and probabilistic matching. For example, foreign names may be misspelled with more frequency, resulting in more missed matches using match-keys which include name. In addition, if a native-language phonetic algorithm is used within the linkage method to identify matches between records where individuals' names were spelled in different ways (e.g., Steven and Stephen), this will result in less accurate matches for those with non-native-language names.

enhanced validation checks for related variables within a supplied data set should be carried out. Based on previous literature, key indicators include:

(a) Per cent of objects whose combinations of values for variables are involved in non-logical relations,

(b) Per cent of objects with dubious variable values or objects whose combinations of values for variables are involved in implausible but not necessarily incorrect relations (i.e., outliers),

(c) Per cent of objects with missing values for key variables that have different characteristics to complete objects, and

(d) Per cent of objects with values imputed by the data supplier for the main variables of interest.

195. Similarly, to the assessment of representation error, an efficient way to assess variable accuracy, especially in the preliminary analysis of data and the very first time the data are used, is the **comparison of data**; that is, the input data are checked by means of comparison with other independent sources that contain the same variable. Suitable independent sources for comparison could include a national survey (such as a labour force survey) or a previous census (Asamer et al 2016).[23]

196. More complex methods for assessing the accuracy of administrative data, where administrative data are linked to a comparative source (which includes the variable/concept of interest), are described in the literature. Bakker (2012) uses structural equation models to estimate the validity of administrative variables, using survey data. The model is applied to data on age, gender, educational attainment, and wages. Scholtus and Bakker (2013) also use a simulation study to test the robustness of the model to additional components of measurement error, to misspecification of the measurement model, and to small sample size. Oberski et al (2017) apply a generalized multi-trait-multi-method (GMTMM) model, under a general framework for evaluating the quality of administrative and survey data simultaneously. The framework allows both survey and administrative data to contain random and systematic errors and therefore does not assume the survey is without error as with other methods (Yucel and Zaslavsky 2005). Their approach accommodates common features of administrative data such as discreteness, non-linearity, and nonnormality and may improve on other models used (such as structural equation models).

[23] It should be noted that consistent values and cross-tabulations generate through different sources and methodologies (e.g., administrative data and survey data) suggest that both sources are likely to be correct. Inconsistent values leave an open question as to which result is most accurate, i.e., closest to the true population value. This depends how survey questions are answered, and how the administrative source is collected, which again highlights the importance of the Source Stage. It is not always true that the administrative data source will be less accurate (e.g., see literature on receipt of state benefits). A more sophisticated analysis is needed to determine the accuracy of both the administrative and external source to assess the cause of inconsistencies found.

5.2.2.3 Re-supplied data

197. Administrative data may be re-supplied to ensure NSOs have access to the most recent and relevant data for use in the census. As with the initial data supplied, the first step to assess the quality of re-supplied data is to perform a macro-level comparison of the main key metrics (such as total number of records, number of missing values, etc.), against what was expected to be received. For resupplied data, a comparison with previous supplies will identify any unexpected differences across the datasets that may indicate a quality concern. Furthermore, longitudinal comparison between the data supplied in the current and previous period is important for revealing possible quality changes, especially in terms of coverage, completeness, and linkability.

198. For key variables that are expected to be stable over time, it is possible to compare values for the same unit (e.g., a person) over time to check for unexpected changes. These checks are easier for 'invariant' variables, such as date of birth or place of birth, and for data where a unique key is available and stable over time. Even for changeable variables such as legal marital status or highest level of education, such checks may still be possible in a restricted form. Longitudinal comparisons can serve as an internal measure of the **reliability** of the data, by providing indicators such as the means or medians of differences or relative differences between the newest and previous data supplies.

199. If there is no key variable that is stable over time, then the distribution of the variables can be used to compare the time periods.

5.2.3 Timeliness and punctuality

200. Measures of **timeliness** and **punctuality** as defined in Section 5.1.3 can be determined by comparing the reference date, the agreed delivery date, and the actual delivery date of the data. The following indicators are suggested by Cerroni, Di Bella and Galiè (2014, p.130):

(a) Timeliness

(i) Time difference (days) = (Date of receipt by NSO) – (Date of the end of the reference period over which the data source reports).

(ii) Time difference (days) = (Date of receipt by user) – (Date of the end of the reference period over which the data source reports).

(b) Punctuality

(i) Time difference (days) = (Date of receipt by NSO) – (Date agreed upon, as stipulated in the contract).

(c) Overall time lag

(i) Total time difference (days) = (Predicted date at which the NSO declares that the source can be used) – (Date of the end of the reference period over which the data source reports).[24]

(d) Delay

(i) Contact the data source holder to provide their information on registration delays.

(ii) Time difference (days) = (Date of capturing the change in the source by the data source holder) – (Date the change occurred in the population).

201. The *Delay* indicator is dependent on information that may not be known or be available to the NSO. However, if some information is available about when data for an observation was updated in the source, this basic indicator can be calculated:

(e) Per cent of observations updated within the past year, counted from the date of delivery to the NSO.

202. The use and interpretation of this indicator is context-dependent, since in some circumstances there might be good reasons for a lack of update, e.g., if there has been no relevant event to trigger a change in the register since the last update for a given record.

5.2.4 Linkability

203. Often, administrative data sources will be linked to other sources be it the census enumeration itself or other administrative sources. A QA of the variables in each source used in the linkage provides general information to inform the design of a successful linkage process as described in Chapter 6.

204. Regardless of whether a unique key or identifier variable is available or whether several variables will be used in combination to identify matches in the linkage process, these indicators should inform the choice and evaluation of the quality of linkage variables supplied, including:

(a) Per cent of unique values, which can be calculated either where a single identification variable is available (e.g., a personal identification number), or a combination of variables to be used in linkage (e.g., a match-key combining age, date of birth, and address), and

(b) Prevalence of biased distributions with respect to accuracy indicators (as described in previous sections, including missing values, implausible values, etc.). Is there evidence of measurement errors, within linkage variables, that are not random? For example, is there a significantly higher proportion of out

[24] This indicators considers the time lag between data arriving at the NSO and it being available for production teams to use given the need to clean, harmonize, validate, ensuring analysts have the correct permissions to access the data etc.

of range or missing values for a key linkage variable such as date of birth, within certain geographies?

205. Finally, if the linkage variables have been provided to the NSO in an encrypted or 'hashed' form. Hashing is a practice that is often used in computer science to protect confidentiality of individuals or other entities in data. It involves applying an algorithm to every piece of information in the original data (e.g., a name) to create a string of characters that uniquely identifies that information and masks the original data. The NSO must verify that the hashing performed by the supplier matches the hashing algorithm used at the NSO. Otherwise, it will not be possible to link the data supplied to other data sources, undermining the relevance of the data. Chapter 6 provides further details about the linkage of encrypted keys.

5.3 Recommendations

(a) As noted in 0, before using an administrative data source within census production, at least one test run with real data is advisable, if not essential. Such a test should be carried out early enough to allow a readjustment of the technical infrastructure and processes to guarantee the readability, harmonization, and validation of the data.

(b) Check that the data supplied matches the metadata collected at Source Stage and that the correct reference date has been supplied.

(c) Compute and monitor basic indicators of the supplied data to gage possible representation and measurement errors.

(d) Verify consistency of related objects and variables within a supplied dataset through enhanced validation checks.

(e) Use statistical metrics to reveal unexpected patterns and outliers.

(f) Compare the total number of records and cross-tabulations with independent comparable sources, to assess accuracy.

(g) Ensure the NSO can clarify data queries with the data supplier. When queries regarding the data arise post-supply, there should be adequate mechanisms in place to ensure these can be resolved.

(h) To improve input quality and ensure consistency, provide feedback to the data supplier about any anomalies (such as inconsistencies within the dataset) found, at least on an aggregated level, providing that the relevant laws on data protection allow this.

(i) When data are being supplied periodically there is a need for additional, longitudinal, QAs.

(j) Determine the timeliness and degree of punctuality of data supplies.

(k) Determine the quality of linkage variables to guarantee the best possible linkage results (see Chapter 6).

5.4 Case studies

5.4.1 Germany: The quality of the data provided from the local population registers for the 2021 census

5.4.1.1 Introduction

206. The German National Census 2022[25] is a combined census using data from multiple sources. Data from all local population registers of the approximately 11,000 municipalities – administrated by around 5,100 local registration offices – is the fundamental source of data, but other information (not specifically relevant to this case study) is collected from a variety of other official sources such as the Federal Mapping Agency, the Federal Ministry of Defence, the Federal Foreign Office, and the Federal Ministry of the Interior, Building and Community. In total, six deliveries of data from local population registers are scheduled for the Census 2022. Since a person, in Germany, may notify a registration office of an address change, after the fact, there will be two different dates for data deliveries the population count is based on – one with a reference date equivalent to the census reference date and one delivery with a reference date three months after the census reference date.

207. This case study focuses only on the quality of German population registers data and problems that occurred during the delivery of that data in January 2019. The 2019 data simulated the largest dataset from the population registers that is to be delivered in the context of the census 2022. The data delivery in January 2019 was a test run to assess the quality of the raw data, test data transmission, optimize existing techniques of data processing, and test the transmission of historical data records. Note that some critics, of this approach, considered a test with anonymized data or a random sample to be sufficient.

208. The case study focuses only on the examination of input quality. For statistical purposes, in Germany, there is not a unique identifier available for a person.

209. In general, the dataset contains every person who was registered with a first or second residence for the reference date of 13 January 2019. The data includes historical records of recent changes in the registers close to the reference date.

210. Since the previous national census in 2011, measures have been taken to improve the quality of German population registers. When a person moves from one municipality to another, the registration offices in the two municipalities communicate this change

[25] The census was originally scheduled for May 2021 but was postponed to May 2022 due to the Covid-19 pandemic.

automatically. Local registration offices communicate any change in their population register to the Federal Central Tax Office, since every person has unique tax ID it is highly likely that the number of first residence duplicate records in the 2019 data has shrunk since 2011. However, this trend is still under examination.

5.4.1.2 Readability

211. In Germany, a standardized, universal format has been determined for the transmission and delivery of data from a local population register. The recipient (which, in this case is data for the census, is the Federal Statistical Office) only accepts the data if delivered in this format. This helps to improve the input quality of the data.

212. At least four municipalities tried to transfer some variables in a format that violated the general formatting rules. Therefore, they could not deliver the affected data records. This consequently led to an incomplete data delivery. For subsequent data deliveries, the format of these variables was broadened so that this problem should not occur again.

5.4.1.3 Accuracy

213. Recalling that population registers are administrated locally, it is not surprising that the accuracy of the data varies across municipalities. The following two examples illustrate the differences in the data accuracy.

214. In the first example, for more than 40 municipalities in one or more of the three variables 'date of moving to an address', 'date of moving to the municipality' or 'date of registration', some 75 per cent or more of all first residence records contain the same date. It can be assumed that this was an error made during a data merge necessitated by the consolidation of two or more municipalities. Such data anomalies can be critical for identifying first residence duplicate records.

215. In the second example, persons on about 120,000 data records had either or all the dates for the three variables earlier than their date of birth. One state had 60 per cent of these erroneous records in its registers.

216. To improve input quality, the municipalities received feedback about anomalies found in their data on an aggregated level and the plausibility checks on the data required expansion. An exchange with the software producers for the population register subsequently took place.

5.4.1.4 Completeness

217. During the 2019 data delivery, several technical problems arose which also had a negative impact on the completeness of the delivered data.

218. Due to an error in the software, the municipalities used to retrieve the data, approximately 1,200 municipalities transmitted files with missing data records. This error was detected only by accident. For some of those municipalities, the software

provider as well as the municipality, initiated a data delivery, due to miscommunication. (In some municipalities, the software provider holds an exact copy of the register with all their data). A comparison of these two deliveries showed that the software provider failed to transmit some data records. The software provider had to schedule a second delivery replacing the first one. The data delivered by the municipalities was deleted. Therefore, the technical infrastructure needs to block the integration of deliveries consisting of data combined from different senders.

219. Generally, it is hard to identify whether some records are missing since the recipient may have no information on the exact number of records that have to be delivered. The recipient can only compare the number of the transmitted data records on first residences in a municipality with its own projected population estimates. However, it is not uncommon for these two figures to differ by up to several percentage points.

220. Some municipalities did transmit, for every data record, missing values for some variables. This showed up as an incomplete data retrieval from the local databases. For instance, the variables 'most current date of moving to Germany' and 'country of origin' (which should be empty if it is Germany) were blank for all data records in approximately 1,200 municipalities. Prior to integrating the data into the database, it is important to check whether variables are missing throughout the data for an entire municipality due to technical problems.

5.4.1.5 Time-related dimension

221. Some municipalities were not able to compile their data until several days after the reference date. A person who reports a subsequent notice of departure in a municipality during the intervening period is not covered. To reduce the possible damage of such a mistake during future data deliveries, it is crucial that municipalities develop the capability to retrieve historical versions of their registers.

5.4.1.6 Conclusion

222. Technical problems during and before the delivery significantly lowered the quality of the data received from local municipalities' registration offices. Hence, the test run for the census 2022 in January 2019 was important to assess procedural and technical flaws. A test run with anonymized data, or a random sample would not have detected most of the described flaws. The test's timing more than a year prior to the Census 2022 data delivery provided enough time to analyze and eradicate errors and to optimize data processing capabilities on the central as well as the local level. Furthermore, municipalities were informed about data anomalies on an aggregated level, since it is legally forbidden for the Federal Statistical Office to return individual data records. This will, hopefully, help to improve the input data quality delivered from the population registers.

5.4.2 Poland: The Polish variable quality system

5.4.2.1 Introduction

223. For the censuses in Poland, data are collected from multiple sources, including administrative ones. Registries and database systems are characterized by a wide variety of content and complexity of structure, resulting from the fact that they are created for different purposes and are managed by different data suppliers. Accordingly, the standards of storage, accuracy, and recording methods adopted in each case also vary. The lack of uniformity exists not only between the registers but also with the data within any one register.

224. The data quality from administrative sources used affects the quality of the census results. Adequate input quality is a prerequisite (although not the only one) to obtain correct census results. When using administrative sources (not only in the context of census production) essential steps are to identify and understand problems and errors in the data and correct the data. For the QA of the input data, the first point is especially important.

225. Having assessed the viability of using administrative sources, the process of managing the quality of data collected from administrative sources in Poland is divided into three parts: input (equivalent to the Source and Data stages in these Guidelines), process, and output quality. The process of managing quality is monitored constantly. At Statistics Poland, the Variable Quality System (VQS) is used for this purpose. The VQS is a system for viewing, analysing, and reporting data from administrative sources.

226. At first, the VQS validates the data received. The process involves applying a set of rules assessing the dataset for completeness, consistency, and correct format for consumption into the system. A key consideration is the completeness and accuracy of the unique identifiers provided in the data supply – critical to ensure high quality data integration. Missing or erroneous values in the unique identifier field prevent records from being integrated effectively across multiple data sources. Data that do not pass the validation assumptions are set for correction – a harmonization process to align the data to the expected standard.

227. Following both the validation and correction steps, a data quality improvement report is produced to inform decisions on whether to approach the data supplier to improve the data quality at the point of supply, or to complete any additional data processing. It enables Statistics Poland to monitor closely all the changes that are taking place in administrative data sources used in our official statistics and permits the automation of the calculation of quality indicators for both input and output data. This case study focuses on the input QA.

5.4.2.2 Accuracy and Reliability

228. The VQS contains the results of the Polish data profiling of the raw data. Data profiling is a procedure with which the user obtains, among other things, information on the accuracy of the raw data. It provides a series of statistical metrics:

    (a) Ordinal position,

    (b) Data type,

    (c) Count (number of records),

    (d) Non-null count, and

    (e) Data length.

229. For numeric variables:

    (a) Minimum value,

    (b) Maximum value,

    (c) Mean, and

    (d) Median.

230. For character variables:

    (a) Pattern count,

    (b) Unique count,

    (c) Minimum length,

    (d) Maximum length,

    (e) Frequency distribution, and

    (f) Pattern frequency distribution.

231. Within the VQS, a cohesion analysis of address data is conducted to check their accuracy and consistency. The address consists of the following hierarchical levels of the territorial division:

    (a) Voivodship (or province, of which there are 16 in Poland),

    (b) Powiat,

    (c) Gmina,

    (d) Locality, and

    (e) Street.

232. Considered separately, the individual address field values may comply with the standard, but do not form a consistent address string appearing in the National Official Register

of the Territorial Division of the Country (TERYT). To consider the address valid, the correct parts are not, in themselves, sufficient. The logical structure must also be kept, that is the street must be in the locality, the locality in the gmina, the gmina in the powiat, and the powiat in the voivodship. Only addresses following this structure are considered consistent. Cohesion to the street is considered as full cohesion, cohesion to the level of the city (compatible sequence of the voivodship, the powiat, the gmina, the locality) or the gmina (compatible sequence of the voivodship, the powiat, the gmina) needs to be improved or supplemented by other available data. With respect to the cohesion analysis, the VQS generates the following quality indicators:

(a) TERYT dictionary comparability (number),

(b) Change of TERYT dictionary comparability (per cent),

(c) Conversion dictionary comparability (number),

(d) Change of conversion dictionary comparability resulting from various stages – input, output (per cent), and

(e) Level of cohesion of address variables (flag).

233. To check the completeness of a variable, the VQS generates the following quality indicators for every variable:

(a) Completeness (number), and

(b) Change of completeness (per cent).

5.4.2.3 Timeliness and Punctuality

234. Long-term, effective, and transparent cooperation with administrative data suppliers is crucial. In Poland, the acquisition of data for census purposes is supported by a legal framework including both a Statistics Act and a Census Act. The VQS records information on the reference date of the data and the date of data receipt by Statistics Poland.

235. Data are usually collected at the end of the year or according to the date of the relevant survey. Data for the decennial census are collected, as soon as possible, during its implementation, permitting the necessary time required to process the data. To maximize the relevance of the data, the collection should either be as close as possible to the reference date of the census or, if the receipt of data is a continuous process, as close as possible to the reference date of the data.

5.4.2.4 Linkability

236. Completeness and accuracy are crucial for unique identifiers such as:

    (a) The PESEL number[26] is widely used in population administrative registers; the number uniquely identifies a person and results in distinguishing between many people having the same first name and surname,

    (b) REGON: business identification number, and

    (c) NIP: tax identification number.

237. Identifiers should be characterized by the required number of characters and the compliance of the check digit. The high quality of identifiers is of utmost importance during data integration. Missing or erroneous values do not allow the same entities to be identified in different sources. The VQS generates the following quality indicators for identifiers:

    (a) Number of correct identifiers (number), and

    (b) Change of number of correct identifiers (per cent).

5.4.2.5 Conclusion

238. Within the methodological framework for improving the input, process, and output quality, the VQS is an important tool for controlling data quality, making quality comparable among different suppliers, and monitoring quality changes over time.

[26]The PESEL (universal electronic system for registration of the population - from the Polish Powszechny Elektroniczny System Ewidencji Ludności) number is an 11-digit permanent numeric symbol that uniquely identifies every person registered in the PESEL database.

Chapter 6. Process Stage

239. Once administrative data are received and the quality assessed by the NSO, 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 (e.g., conceptual misalignment with the census definitions and concepts, coverage, and measurement errors). The Process Stage of these guidelines provides an overview of key processes used to integrate administrative data into the census and the related quality concerns.[27]

240. The processing of the administrative data is informed by the findings obtained from the Source and Data Stages. For example, the assessment of the linkability of an administrative source informs how data are linked. An understanding of coverage error informs the processes for data integration to achieve the coverage needed for the census. The assessment of the administrative data accuracy will inform the editing and imputation stages and provide the insight to support decisions about how sources should be used together to construct the census variables. However, processing can introduce additional error (systematic or random), thus introducing bias or variance in the final estimates. For this reason, it is important that the processes are appropriately tested and evaluated. Error needs to be managed along the entire statistical production chain. This chapter focuses on some of the most common processes required for using administrative data in the census. These are:

    (a) Linking records,

    (b) Assessing coverage error in statistical registers or administrative data in the enumeration of population units,

    (c) Resolving inconsistencies in data item values from different sources, and

    (d) Editing and imputation.

241. Each of these processes is described in more detail in the following sections, along with the challenges associated with each, ways to assess the quality of the processed data, based on the available literature, and the experiences of different countries.

6.1 Record linkage

242. Almost every administrative data source requires some form of record linkage to other data sources, for validating data or for ensuring adequate coverage of the census population units and variables. For example, two or more data sources may need to be

[27] See KUMUSO, Quality Framework for Multisource Statistics, 2019 WP1 for quality indicators, measures, and methods for assessing process and output quality.

combined to achieve better coverage of the target population, including to adjust for potential overcoverage (see Section 6.2). Likewise, linkage of multiple sources may be necessary to provide complete and accurate data for the census variables (see Section 6.4).

243. Many countries integrate administrative data from multiple sources to create administrative-based statistical registers; these may include address, population, or business registers (see UNECE 2018, Chapter 8 and Section 6.2 below). Even countries without statistical registers are moving towards maximizing the use of administrative data in the production of their core population, social, and business statistics.

244. This makes record linkage one of the most important processes for using administrative data in the census. It is thus important and necessary to assess the linkage process quality, through an assessment of the linkage variables or keys (as described in the Source and Data Stages) and through an assessment of the process itself, as outlined in the sections that follow in this chapter.

245. The impact of linkage error (representation and measurement error) on the overall accuracy of population and census estimates should be considered (see Daan Zult et al 2019). For instance, missed and false links can lead to over- and undercoverage of the census population and can cause the wrong value to be assigned to a census variable for a given person or household. Address data often need particular attention, as addresses can be used for both linking data for an individual (e.g., as a linkage key in combination with name and date of birth) and linking individuals together to form households. However, people do not always alert a data supplier when they move. A registered address might also not be the primary residence address. Therefore, the accuracy of address data can be poor in administrative sources. Linkage error can also introduce bias in DSE (Abbott 2009).

246. Methods for linking data typically rely on the existence of common unique keys (or identifiers) across the sources to be linked. For example, Poland uses a unique identifier, the PESEL number (see Section 0).

247. In the absence of common unique keys, other common identifying variables, such as address, name, sex, and date of birth, may be used to link records from multiple sources. Although this is more challenging and subject to higher levels of error, as outlined below.

248. In some cases, the NSO may only have access to anonymized or 'hashed' identifiers in the administrative data (see Section 5.2.4). Hashing has some important quality implications for data linkage (see Shipsey and Plachta 2020 for a description of methods for linking with anonymized data, the challenges, and limitations).

249. Linkage methods are of two main types: deterministic, when matches are made based on a set of common identifiers, and probabilistic when matches are made based on model-based linkage weights (see Harron, Goldstein and Dibben 2015). Probabilistic matching does not require record values to be identical between two records but relies on similarity between records. One additional linkage method that can be applied to

unlinked records after deterministic and probabilistic methods are applied, is clerical linkage, which involves a manual inspection of the unlinked records. Clerical linkage is not possible to do when the data are hashed.

250. Linkage error can occur through unlinked records that should have been linked (also known as 'false negatives') and linked records that should not have been linked (also known as 'false positives').

251. Two common methods for assessing linkage quality are:

(a) Estimation of false positive and false negative rates, using a clerical review of samples of linked and unlinked records. Although, the clerical review can only be done when the data are not hashed. If the data are hashed, the NSO should try and obtain access to samples of the linked and unlinked records in their original state, prior to hashing, to assess the linkage,

(b) Comparison of the distributions of characteristics of linked and unlinked records (e.g., age, sex, and ethnicity). Differences in characteristics may suggest bias was introduced by linkage error. This implies certain types of records (e.g., individuals) may not be linked because they are more difficult to link.

252. The assessment of linkage error using the methods described above are presented in the United Kingdom and New Zealand case studies, see Sections 6.7.1 and 6.7.3.

Box 9: Methods for data linkage and the assessment of linkage quality: a UK cross-government review

The importance of linking administrative data for the public good (including for the census) is widely recognised and resulted in a cross-government review within the United Kingdom to develop guidance on data linkage methods, covering the QA of linkage. The review drew on the work of experts across government, academia, the private sector and internationally. The outcome was a series of articles covering: the future of data linking methods; QA in data linkage; longitudinal linkage (design principles and the total error framework); preserving privacy; linking with anonymised data; and procedures for improving efficiency (see ONS 2020).

6.2 Statistical registers and the 'signs of life' methodology

253. As mentioned in Section 6.1, integrating data from different sources for use in a census is becoming increasingly common; record linkage plays an important role in this process. Two key quality dimensions related to the integration of data from various sources are coverage and coherence. Integrating data is done to assess and possibly reduce coverage error. It also enables and requires assessment of information coherence across sources and over time.

254. One example of data integration for use in a population and housing census is the construction of statistical population registers. By linking information from the available sources at the record level, it is possible to determine individuals or households that are resident in a country and their characteristics. The integrated data from these sources become a statistical register, namely, a database that can be used for further processing and analysis to produce census-type outputs (see UNECE 2018, Chapter 8).

255. Some of the key processes involved in the construction of a statistical register are:

    (a) Identifying the data sources to be used,

    (b) Linking the sources (Section 6.1),

    (c) Developing and applying a set of rules to make decisions about which records should be included in the final estimates (Section 6.2),

    (d) Resolution of conflicting information (e.g., date of birth or address) between the linked sources (Section 6.4), and

    (e) Editing and imputation (Section 6.5).

256. The quality considerations and indicators suggested in Chapter 4 and Chapter 5 will help identify the data sources to be used in a statistical register. This section focuses on the application of decision rules and some quality considerations related to this process. The section also discusses other methods of coverage assessment that can be used in statistical registers along with decision rules.

257. Decision rules, or 'activity' rules, are inclusion criteria that are often applied when constructing statistical population registers to ensure that only individuals who meet some pre-defined usual residence criteria are included in the final estimates. This process is sometimes known as the 'signs of life' (SOL) method and is a widely used tool to reduce overcoverage in statistical registers (e.g., inclusion of records that are not part of the usually resident population).

258. Spain uses 'signs of presence' from four types of administrative sources: tax files, the social security database, labour market-related sources and the central registry for foreign nationals. Individuals who have reached the threshold level of signs of presence are considered 'active' and included within the population count, whereas all others, called 'inactive', are not included in the count (see Vega Valle et al 2020 and the case study from Spain, Section 6.7.2 for more details).

259. The UK uses a similar approach to decide which records from selected administrative sources should be included in their Administrative-Based Population Estimates (ABPEs) (see ONS 2019). In an earlier version of the ABPEs, a record was included in the population estimates if it was present on two of the selected administrative sources. In a subsequent version of the ABPEs, strict criteria for inclusion were applied to each source separately (where records were only included if there was a sign of activity within the last 12 months) and the rule of including records, only if present on two sources, was removed (with data linkage used to un-duplicate records that appeared

on multiple sources). The subsequent version of the ABPEs aimed at further reducing the overcoverage found in the previous version, at the expense of increasing undercoverage (records that are missed from the population estimates). The expectation was undercoverage would be addressed using a coverage survey combined with a DSE.

260. The success of a SOL method relies on the availability of good indicators of signs of activity in the individual or combined administrative data sources. The application of the method typically involves making some assumptions, which determine who is considered as active and who is not. NSOs should be clear about these assumptions and where possible provide relevant supporting evidence. In particular, the choice of signs of activity indicators (or decision rules) should be informed by QA at the Source and Data Stages (see Chapter 4 and Chapter 5), including consultation with data suppliers, cross-validation between sources and over-time, and expert knowledge.

261. As already mentioned in the case of the UK, the application of SOL methods may be combined with other methods to assess and account for coverage error in statistical registers. One of these is to conduct a survey that is independent from the statistical register and use the combined information from the survey and the register to estimate the number of records that are missed in the register (or in the survey) to improve the final estimates. This is a similar approach to conducting a PES after the traditional census and applying DSE methods (also known as capture-recapture) to assess the level of undercoverage in the census (Abbott et al 2020).

262. Overcoverage in statistical registers can also be assessed through linking the register to a survey through an approach called ‘dependent interviewing’, which aims at verifying administrative records in the field. This approach has been used in Italy and in some other countries (e.g. Israel) that have successfully transitioned to administrative data-based censuses. However, not all countries can carry out dependent interviewing, due to ethical and privacy concerns (see Chapter 4). Brown et al (2020) provides a discussion of dual- and multisystem estimation methods for tackling coverage errors.

263. In Italy, dependent interviewing with a sample of households drawn from the population base register (PBR) (also known as the Registro Base degli Individui or RBI) and a SOL methodology (using other administrative sources) is used in combination to estimate and adjust for overcoverage error in the PBR. In addition, a sample survey of addresses drawn from the statistical base register of addresses (RSBL) is used to adjust for undercoverage. As a result of this process, the population estimates are obtained by applying correction coefficients for both under- and overcoverage errors to individual data on the PBR. The Italian case study, in Section 6.7.4, provides details on the complete methodology.

6.3 Enumeration of population units: administrative data-based models

264. Related to the construction of statistical registers, administrative data can be used to enumerate population units (e.g., individuals, households, or occupied addresses), to support or supplement a census field collection. This approach was used in both New Zealand to address undercoverage in their 2018 Population and Housing Census and in the United States (US) to improve the efficiency of their field Non-Response Follow-Up (NRFU) operation.

265. The approach involves linking integrated administrative data sources to a 'gold standard' dataset (in this case the traditional census) to build models to assess the quality of the administrative data and to determine under which conditions the administrative data are used for the census. The approach allows for partial usage of administrative record information where they are believed strongest.

Box 10: Determining occupancy at an address (the United States Census field operation)

For the United States Census, the aim was to use administrative data to determine vacant and non-existent addresses and to enumerate occupied addresses as part of the Non-Response Follow-Up (NRFU) Operation. Where the administrative data predicted (based on defined cut-offs) that an address was un-occupied, the field contacts could be reduced, thus reducing costs, and improving efficiency. Predictive models were developed based on the relationships observed in 2010 between census outcomes (as a 'gold standard'), government administrative records and third-party data. The performance of the models was then tested as part of the 2015 and 2016 census tests, and via a retrospective evaluation using the 2010 Census. Multiple administrative sources (government and commercial) were used, including tax, social security, health, housing, and Postal Service data.

The performance of the models was used to determine cut-offs to guard against under-coverage (where addresses are incorrectly classified as vacant by the administrative-based model), while aiming to minimise NRFU workloads. Specific attention was paid to the performance of the model by different geographic areas, with different concentrations of population groups (e.g., Hispanic, Non-Hispanic, and Black populations). This resulted in further development of the strategy to protect against misclassification of addresses as un-occupied (Section 3 of Administrative Records Modelling Update for the US Census Scientific Advisory Committee, 2017, provides details on the QA that was carried out).

Box 11: Direct Enumeration (the New Zealand 2018 Census)

The New Zealand 2018 Census used administrative data to enumerate people that had been missed from the field collection. Census data (previous and current) were linked to administrative records to build models that were used to assess the quality of the administrative data and to determine how and when they would be used to include people, families, and households in the census.

The primary aim of the administrative enumeration was to target under-coverage in the census. The linkage method was designed to minimise false positives (i.e., to minimise the number of administrative records incorrectly excluded from the census dataset because they were wrongly linked). Furthermore, an adjustment was made as part of the final enumeration process to reduce false negatives (i.e., administrative records that were incorrectly not linked, and thus added to the census dataset in error, causing over-coverage).

The administrative records that were selected for inclusion following the linkage process, were divided into those to be included into dwellings (with families and households created), and those included at a small geography only (with no relationship to dwelling and no family or household created). This decision was driven by statistical models that were specifically developed to predict the reliability of administrative data for representing households. The models (which used census data) were assessed using receiver operating characteristic (ROC) curve analysis.

To assess the performance of the approach, an indication of the coverage patterns for the census after the administrative enumerations were included was carried out. A newly developed DSE benchmark population provided the most suitable estimate of the true census usual resident population available at that stage. Population distributions by age, sex, ethnicity, and geography were produced. The distributions showed that the 2018 Census dataset was largely consistent with the benchmark and in most cases, the inclusion of administrative records in the file greatly reduced (but not resolved all) under-coverage (Stats NZ, 2019a). These indicative findings provided confidence in the new methods when the census data were released. See case study 6.7.3 for more details on this approach.

6.4 Conflict resolution/decision between sources

266. As mentioned in Section 6.2, when combining administrative data to create statistical registers there may be inconsistencies in the values of key variables across different sources. For example, once a decision has been made on which administrative records (persons) to include in the usually resident population, if a person's address on two or more sources is different (e.g., due to delays in communicating a change of address, administrative processing delays, or second/multiple homes), then the NSO may need

to decide at which address the person should be included. Conflicting (or multiple) address information and any related decision may cause undercoverage in some areas and overcoverage in others. At an aggregate (e.g., national) level this may not be an issue because the person may only be counted once. However, at a small-area level this may matter, if the two addresses are in two different areas, as it could cause overcoverage in one area and undercoverage in the other.

267. Abbott et al (2020) describe three approaches for deciding between sources in the context of address conflict:

(a) Remove the record from the population,

(b) Split the record between the different locations according to weights (e.g., half if there are two locations),

(c) Choose which source is the most likely to be up-to-date based on the characteristics of the individual or the administrative variables. This approach could also use additional data sources when the same individual appears.

268. The first approach increases undercoverage in the population estimates. The other two approaches may produce acceptable population estimates at an aggregate level but may introduce significant biases due to coverage and linkage error in estimates at lower levels of disaggregation, such as, age and sex. These approaches, (b) and (c), have been tested in the UK as part of their development of the ABPEs; further research is ongoing (ONS 2016, Section 6).

269. Similar approaches to measure the quality of attributes in statistical registers when the same attribute is available in different sources, have been used in Austria and in Spain. In the Austrian full register-based census, a combined quality indicator is calculated using the Dempster-Shafer theory, also known as the theory of belief functions and a generalization of the Bayesian theory of subjective probability (Dempster-Shafer Theory: see Shafer 1992). A comparison with an external source is carried out to assess the associated statistical rules (Statistics Austria 2019).

270. The Spanish population register lacks information on individuals' LMS. To estimate LMS, several registers are used to obtain complete information (Argüeso 2019), including data from the tax agency, the civil register, the social security database, and the central register of foreign nationals. Since an individual may appear in multiple data sources with conflicting information, decisions rules are applied to determine the most plausible value. The decision rules are applied to each person record, after which a value for LMS may be given. If cases remain unassigned, a value is imputed based on age, information in past censuses, and the number of household members. The results generated by this method are promising; further research is ongoing.

271. To summarize, methods for deciding between sources, when the same attributes are available in different sources, typically rely on decision rules, like in the SOL methods (see Section 6.2). Different approaches should be considered and tested by NSOs,

according to their census specific needs and based on quality information gained at the Source and Data Stages (see Chapter 4 and Chapter 5).

6.5 Editing and Imputation

272. The QA at the Source and Data Stages (Chapter 4 and Chapter 5) will inform whether the administrative data requires editing (for incorrect/implausible values) and/or imputation (for missing values). Editing and imputation may be required both on the single source and the integrated data.

273. In the Austrian register-based census, seven 'base registers' are used to provide basic information on their respective census topics. These base registers are supplemented by eight 'comparison registers', which are used mainly for validation purposes. That is, one base register is selected to provide the value for a certain census variable, and the comparison registers are used to confirm these values (see Schnetzer et al 2015). However, in some cases, the comparison registers also provide data that are either fully or partly missing in the base registers. The combined dataset from the base and comparison registers, called the Central Database (CDB), is enhanced with imputations for item non-response and implausible values, creating the Final Data Pool (FDP). Quality is assessed throughout, from metadata and contact with data suppliers (e.g., to understand the reliability of the data for the intended purpose and how the data suppliers dealt with missing or implausible values) to checks on the register-based output through comparison to an independent external source (Statistics Austria 2019).

274. Three imputation methods have been applied in the Austrian register-based census: deterministic editing, statistical estimation (including hot-deck and logistic regression), and statistical matching. For example, hot-deck imputation has been used to impute LMS. This involves aggregating individuals into groups ('decks') by attributes which are strongly correlated with the target variable. The marginal distribution of the target variable within a deck (with existing values) is used to impute the target variable in the corresponding deck (with missing values). In the final data QA, in the FDP, a quality indicator for the imputation is computed.

275. Schnetzer et al (2015) suggests the use of classification rates to evaluate different imputation models. This involves applying the imputation method to already existing data and comparing the results of the imputation process with the true values for each unit. The classification rate is derived as the ratio between the values that match and the numbers of all compared units. The classification rate is like a hit ratio and can be used for categorial and numerical values.

276. Chambers (2001, cited in Schnetzer et al 2015) describes five quality-related properties that imputations should fulfil:

    (a) **Predictive Accuracy** – The imputed values should be as 'close' as possible to the true values,

(b) **Ranking Accuracy** – The imputation process should preserve the order of imputed values (for attributes which are at least ordinal),

(c) **Distributional Accuracy** – The imputation procedure should preserve the distribution of the true data values,

(d) **Estimation Accuracy** – The lower-order moments of the distribution of the true values should be reproduced by the imputation process (for scalar attributes),

(e) **Imputation Plausibility** – The imputation procedure should result in imputed values that are plausible.

6.6 Recommendations

(a) As mentioned in Chapter 5, the accuracy and completeness of linkage variables should be assessed prior to linking data from different sources.

(b) For the linkage method, overall linkage rates and false positive/negative rates should be assessed and reported against. Thresholds for linkage error should be pre-determined and the trade-off between minimising false positive or false negative links should be considered.

(c) Coverage error in the statistical population register should be assessed and accounted for. This can be achieved using comparisons with other sources, including the SOL methodology and/or using surveys (which can be specifically designed to adjust for over- and undercoverage).

(d) The choice of 'signs of activity' indicators (or decision rules) when constructing statistical registers should be informed by an assessment of quality at the Source and Data Stages and different methods (and underlying assumptions) should be tested.

(e) Models can be used to both assess the quality of administrative data for the purposes of enumerating population units (against a dataset that is taken as the 'truth') and to determine when and how to use the administrative data for this purpose.

(f) When deciding between sources, when the same attributes are available in them, different approaches should be considered and tested, according to census specific needs and based on quality information gained at the Source and Data Stages.

(g) The quality of editing and imputation should be assessed both on the individual sources and on the integrated data; different imputation models should be assessed.

6.7 Case studies

6.7.1 United Kingdom: measuring linkage quality when replacing a census variable with administrative data

277. The decennial Census of England and Wales is conducted by the ONS to enumerate the population, and record population and household characteristics. The ONS is looking to replace a census question on "number of rooms" for the 2021 Census using administrative data. Some elements of this work remain to be completed; however, linkage quality has been tested through using 2011 Census data (see ONS 2020b).

278. The 2011 Census asked the two questions "How many rooms are available for use only by this household?" and "How many of these are bedrooms?" The responses are used to derive occupancy rates by comparing the rooms/bedrooms that are available to the "rooms/bedrooms required." A negative occupancy rating implies there are fewer rooms/bedrooms available than required by the household (overcrowding). The information allows central and local governments to develop appropriate housing policies and plan future housing provision. The quality of responses to the 2011 Census number of rooms question, as measured by the 2011 Census Quality Survey, was considerably lower (67 per cent) than that of the number of bedrooms question (91 per cent). This and the motivation to reduce respondent burden led the ONS to consider administrative data, specifically Valuation Office Agency (VOA) data, as an alternative way to meet the information needs. The VOA is a government agency. It has been responsible for banding all domestic properties in England and Wales for the local council tax since the tax was first introduced in the early 1990s.

6.7.1.1 Linkage quality

279. The unique property reference number (UPRN), a unique alphanumeric identifier for every spatial address in the UK, was used to link VOA and census data. To ensure high quality linkage, the uniqueness of this variable was assessed in both the VOA and census data prior to linking the two sources. In census data, responses with a non-unique (duplicate) UPRN were treated as if they have missing number of rooms values, as these cases cannot be linked to the administrative data with certainty. Duplicate UPRNs in census data occurred if two or more different census addresses were assigned the same UPRN. An example of this might be where a ground floor flat and a first-floor flat are assigned the same UPRN but have different census address identifiers. This is likely to be due to matching error when census address records are linked to the address frame, as the method includes an element of probabilistic matching.

280. In the VOA data, records with a non-unique linkage variable (1 per cent) were excluded. This is like duplication in the 2011 Census data. Other VOA records are 'cleaned' prior

to data linkage (3 per cent), including removing records that had not been assigned a UPRN by GeoPlace[28] (0.2 per cent) and records with duplicate UPRNs (0.3 per cent).

281. The linkage rate of 2011 Census responses with administrative data by UPRN was also measured. A high linkage rate was important because unlinked census records to the VOA records result in missing values in the VOA 'number of rooms' variable. Excluding wholly imputed households (non-responses) and non-unique records, 96 per cent of 2011 Census households linked to the VOA property data.

282. An important assumption of the intended editing and imputation approach (namely, a donor-based imputation method) is that the pool of donors is as representative of the recipients and as large as possible. Therefore, prior to editing and imputation, the distributions of unlinked and linked census to the VOA records were compared on key household variables, such as accommodation type and number of usual residents. A similar comparison was carried out for missing number of rooms data in linked and unlinked datasets. Although some differences in distributions were observed, the number of available 'donor' records where number of rooms was non-missing was sufficient, when broken down by a single household variable and by local area the number of donors always exceeded those with missing values. The edit and imputation processes were tested for ten local authorities with the highest per cent of missing number of rooms.

6.7.2 Spain: Use of administrative data in the construction of a census data base for the 2021 Spanish Census: the 'signs of life' method.

283. The 2021 Population Census in Spain is viewed as a microdata database with approximately 47 million records, one for each resident. For the census enumeration, administrative records contain a vast amount of relevant information, despite being collected by authorities for purposes unrelated to population counts. Administrative sources are linked together to create a statistical population register to identify who is residing in the country and to produce population estimates.

284. The basic structure for the population count is based on *Padrón*, the Spanish population register where all residents in each municipality of Spain are recorded. Individuals are required to register in the municipality in which they live. Since registration brings many advantages, residents normally do register.

6.7.2.1 'Signs of life' method

285. When using the *Padrón* for census purposes, an adequate statistical register must be constructed. After receiving the original *Padrón* database referenced on 1 January of each year, a statistical treatment is carried out. Some assumptions are made around the presence of foreign nationals in Spain whose registrations have expired or about to

[28] GeoPlace is the central data source for UK street addresses. More information is available at https://www.geoplace.co.uk/about-us/who-we-are/our-story

expire. Moreover, population figures are statistically corrected to ensure they meet the 'usually resident' definition, applying the twelve-month residence concept. In short, population numbers are obtained from *Padrón,* but they are not exactly the result of counting the registered population since some individuals are excluded while others are added.

286. From the whole population register, approximately 1.7 per cent of the persons are excluded from the population counts, while approximately 0.15 per cent are added and included in the population counts.

287. To identify which individuals are usually resident, the SOL method is applied. All individuals are analyzed within the available administrative data sources and their movements are detected in *Padrón* for the months following the reference date (Vega Valle et al 2020). The four key administrative sources used to assess SOL are as follows:

    (a) Tax agency and local tax files,

    (b) Social Security Insurance Database comprised of individuals with insurance and beneficiaries (employees and pensioners),

    (c) Labour market-related sources including:

        (i) Unemployment National Service Database providing a job seekers file of unemployed individuals,

        (ii) Social Security Affiliation Registers of affiliation information of the employed population,

        (iii) Public Assistance Database containing information about benefit recipients.

    (d) Central Registry for Foreign Nationals Database provides supplementary information about foreign nationals living in Spain, including date of application for residence permit, licence or rejection of residence permit, expiration dates, and residence checks.

288. Through the SOL method, individuals who reach the threshold of presence signals within administrative data will be identified as 'active' and will be included in the population counts. Individuals not meeting the threshold will appear 'inactive' and will not be included. These SOL from administrative data can also be compiled at individual and household level; information is available on how many household members are 'active'.

289. For both Spanish and foreign nationals, the address changes in *Padrón* are considered in the months following the reference date. There are certain address changes that require the direct intervention of the person. There can also be a residence check made by a municipality, which generates a high probability that the person is residing in Spain on the reference date. Other changes are good indicators that a person is not residing

in Spain, on the reference date. These can be used to identify individuals that are 'usually resident' in the country.

290. For minors, a sign of presence is if an adult, in the same household, shows signs of presence themselves. Minors who do not meet this requirement are excluded from the population count. The possibility of using information about students enrolled in official studies is currently being analyzed.

6.7.3 New Zealand: Process QA when including administrative enumeration in the New Zealand 2018 Census

291. For the first time, the New Zealand 2018 Census dataset included administrative records for the direct enumeration of people who were missed by the census field collection, replacing the use of imputed records in previous censuses. These administrative enumerations are drawn from a resident population derived from administrative data which have already been assessed for input quality and quality limitations are known (Gibb et al 2016; Stats NZ 2017). The administrative enumerations are only added to the census data, if the individuals were in New Zealand on census night and were census non-responders (Stats NZ 2019a). This case study focuses on how we measured and assessed the accuracy of our linking and statistical modelling processes.

292. The administrative enumeration methodology was designed to achieve a final census dataset with the highest possible coverage of the census target population. We were most concerned with eliminating potential overcoverage due to the use of administrative records, both nationally, and for local areas, and expected that this would result in some remaining undercoverage. Linkage processes were designed to ensure that administrative records are added only for people who have not already responded to the census. Statistical models have been developed to manage the known quality limitations of the administrative resident population.

293. At the highest level, the process for including administrative records in the 2018 Census data involved linking the census responses with the administrative data, selecting administrative records to be included into dwellings (with families and households created), and those records included at a small geographic area only (with no relationship to dwelling and no family or household created). At each stage of the process, we assessed quality and decided if the methodology was acceptable.

294. The link between the census responses and administrative population was achieved using a fully automated probabilistic linkage process designed to minimize false positive linkages (Stats NZ 2019b). The quality of the linkage is assessed through estimating the false positive and false negative rates. The false positive estimate was derived from manually checking a small sample of linked records. The false negative estimate was based on an approach developed by Choi (2019), in which we estimated the missed matches from a subset of the census forms that met the criteria for inclusion in the administrative data with a high level of certainty (so we should be able to match records). The overall link rate achieved was high (97.7 per cent) with false positive links

estimated at 0.6 +/-0.3 per cent and false negative matches estimated at 1.21 per cent (Stats NZ 2019c). The high link rate coupled with low error rates gave us confidence that the linkage was of acceptable quality. We were mostly concerned with false negative matches and the potential for them to impact accuracy by contributing overcoverage to the 2018 Census data, so we included an adjustment for these false negatives later in the administrative enumeration process.

295. The methodology used for assigning administrative records to dwellings and the subsequent step into small geographic areas is designed to balance the quality limitations of the administrative data against the quality requirements of the 2018 Census data (Stats NZ 2019a). As stated previously, the driving dimension of quality is accuracy. To assess the quality of the administrative data for inclusion in census, we developed statistical models (using current and previous census data) to predict reliability of administrative data for representing an entire household (Gath & Bycroft 2018; Stats NZ 2019d). We used census data for testing and assessing the models (assuming census responding households represent the truth). A model score was generated for each administrative household, representing how reliable the administrative data was for the entire household in a particular dwelling. A model score cut-off determined which of the non-responding administrative households was added to the census data. The model was assessed using a ROC curve analysis and by analysing performance metrics such as sensitivity, specificity, and precision (Stats NZ 2019d) across a range of model score cut-offs. We saw medium to high scores on the sensitivity measure (the proportion of correct administrative households that we included) across the full range of cut-off scores giving us confidence we were able to correctly identify most of the high-quality administrative households. In contrast, we saw greater variability in the specificity measure (the proportion of incorrect administrative households that we excluded) indicating we were also likely to include some administrative households without the correct household composition.

296. With the remaining administrative population, to ensure we did not introduce people who should not be included, we first adjusted for potential overcoverage (using a strict SOL approach) and then adjusted for duplication caused by missed links between returned census forms and administrative data. A model, like the one used for inclusion of households, was applied predicting the likelihood that the administrative meshblock[29] reflected a person's true usual residence meshblock; people with scores greater than a cut-off were included.

297. Much of the QA involved determining where to set the model cut-off scores considering relevance, accuracy, coherence, interpretability of the methodology, and data produced. The cut-off for inclusion of administrative records in dwellings has been set as a balance between strict criteria of obtaining the same people in the household as

[29] A meshblock is the smallest geographic unit for which statistical data is reported by Stats NZ. It is a defined geographic area, varying in size from part of a city block to large areas of rural land. Meshblocks are contiguous, which means that each meshblock borders another to form a network covering the whole of New Zealand (which includes coasts and inlets).

we observe in the census and including administrative households that reflect similar household composition patterns as the census, even if we cannot guarantee that all household members are the same. The cut-off for inclusion of administrative records in small geographic areas once again represents a trade-off; between maximising the use of administrative data to improve national demographic counts and minimising the number of individuals enumerated in the wrong area.

298. The QAs outlined have several limitations due to subjectivity in judgements, statistical assumptions, and challenges with the underlying administrative data. The linkage error assessment of false positive links was dependent on the quality of judgements made by clerical reviewers, while the false negative link assessment relied on the assumption that the records used in estimation are representative of those not eligible. The modelling assessments were also limited by the subjectivity in setting an appropriate model cut-off score, robustness of underlying assumptions such as census response data represented the truth (which extended into assuming no within household non-response), and the lack of information available for determining when administrative data was incorrect. Future work on process QA will include further methodological development, testing of assumptions, and exploration of alternative QA tools for these processes.

6.7.4 Italy: The combined use of survey and register data for the Italian Permanent Population Census count[30]

6.7.4.1 From door-to-door enumeration to the Permanent Population Census

299. The Permanent Population and Housing Census (PPHC) has been designed based on the Italian National Institute of Statistics' (Istat') modernization program, which places the integrated system of statistical registers [Sistema Integrato dei Registri, (SIR)] at the core of statistical production. The role of field surveys in this system is to support registers, in the broad sense of assessing their quality and to add information that is missing, incomplete or of insufficient quality.

300. The 2011 Census, though being register-assisted, was still a traditional census, with an exhaustive field-collection.[31] The PPHC is based on a reverse relationship between field enumeration and registers, where register data are supplemented by field data collection.

[30] ISTAT (2020) Nota tecnica sulla produzione dei dati del Censimento Permanente: la stima della popolazione residente per sesso, età cittadinanza, grado di istruzione e condizione professionale per gli anni 2018 e 2019: Dalla rilevazione "porta a porta" al Censimento permanente [Technical note on the production of Permanent Census data: estimating the resident population by sex, age, citizenship, education level and occupational status for the years 2018 and 2019: From door-to-door survey to permanent census]. Rome: ISTAT. Available from https://www.istat.it/it/files//2020/12/NOTA-TECNICA-CENSIPOP.pdf

[31] Municipal Population Registers were used to guide field enumeration i.e., as enumeration lists to mail out questionnaires, while other administrative sources integrated into the Additional List of Auxiliary Sources were used to correct the list under-coverage i.e., to enumerate people usually resident but not yet registered.

301. At the core of the PPHC is the population base register (RBI), whose main administrative data source are the local population registers of Italian municipalities. Together with the statistical base register of addresses (RSBL) and with the thematic registers on education and employment, it provides the basis to produce population census data, while ad hoc surveys are used to measure the coverage errors of the RBI and to collect data for variables that are not available (or only partially available) from the registers.

302. Two separate sample surveys (Areal survey and List survey) are conducted annually in self-representative municipalities [32] and every four years, according to a rotation scheme, in non-self-representative municipalities, for a yearly total of approximately 1.4 million households (of which 450,000 households are in the Areal survey and 950,000 households are in the List survey).

303. In the Areal survey, a sample of addresses and/or enumeration areas (depending on the quality of addresses in a municipality) drawn from the RSBL is canvassed "blindly" (as in traditional censuses) to enumerate every household.

304. The List survey, based on a sample of households drawn from the RBI, is conducted with a mixed mode technique (CAWI, CAPI, CATI). The first phase is comprised of only "spontaneous responses," while in the second phase, there is also field follow-up of non-responding households by enumerators. For each non-responding household, a pre-coded 'outcome' is registered in the survey monitoring system at the end of the fieldwork.

305. The same questionnaire is used in both surveys (the only difference is that the list of household members is prefilled with RBI data in the List survey) and includes all census variables (including those available in registers) to test the quality and the coverage of data already available in registers against the data collected by the surveys.

6.7.4.2 The combined use of register and survey data for assessing and correcting for coverage errors of the RBI

306. With the aim of producing the population count, survey data are used to correct RBI data within a DSE model aimed at estimating the coverage errors of the register. In a traditional census, a PES is often used to measure the census under-count (with the PES being the second 'capture' while the census itself is the first 'capture'). In the PPHC, the RBI represents the first 'capture' whilst the annual sample surveys and the 'administrative SOL' represent the second 'capture'. Furthermore, differing from a typical PES, aimed at measuring undercoverage, in the PPHC design, the second 'capture' aims at measuring and correcting for both undercoverage and overcoverage of the RBI.

307. In the field, the second 'capture' is two-fold, with the Areal survey used for measuring the undercoverage error of the RBI, and the List survey used, together with information on "administrative signs of life" derived from the [Integrated Administrative Data Base

[32] Self-representative municipalities are those with a population over 17,800 inhabitants and smaller ones which do not rotate in the sampling scheme of the Labour Force Survey. All the others are non-self-representative.

(AIDA)], for measuring the overcoverage error of the RBI. As a result of this process, the population count is finally obtained by applying correction coefficients for undercoverage and overcoverage errors to individuals in the RBI.

308. Through the linkage with the RBI, the Areal survey allows Stat to estimate the number of individuals usually resident in the municipality who are not included in the RBI. Similarly, through the linkage with the RBI, the List survey allows Istat to estimate the number of individuals included in the register who are no longer usually resident in the municipality. For this purpose, non-responding households are classified according to their 'coverage status' based on the outcome registered in the survey monitoring system.

309. However, since the survey itself might be affected by undercoverage errors, failing to reach all usually resident individuals in a municipality, a further step is undertaken before calculating the overcoverage rate. Within the subset of 'potential overcoverage' individuals (individuals still present in the municipality according to the RBI and not found in the field), a distinction is made based on SOL in the municipality recorded in the AIDA. Non- responding household on the List survey are thus 'recovered' if they show strong evidence (i.e., of at least 8 months) of SOL in the same municipality where they are recorded in the RBI. While individuals lacking such SOL in the municipality are confirmed as overcoverage of the register. The SOL considered: are being a public servant, private employee or self-employed; receiving a retirement pension; attending school (including pre-primary) or university; perceiving an unemployment benefit or basic income; or being a fiscal dependent family member of an individual with strong evidence of SOL.

310. The correction coefficients applied to individuals in the RBI are obtained through the following steps.

(a) Calculation of the raw non-weighted rate of undercoverage per each profile[33] as the ratio between the newly enumerated individuals (i.e., individuals not expected according to the RBI), and the total number of individuals enumerated:

$$p_{ij,under} = \frac{Newly\ Enumerated_{ij}}{Total\ Enumerated_{ij}}$$

(b) Calculation of the raw non-weighted rate of overcoverage per each profile as the ratio between individuals expected according to the RBI and not found in the survey (or not 'recovered' according to the AIDA) and, in the denominator, the same individuals plus individuals expected according to the RBI and enumerated in the survey (or 'recovered' according to the AIDA):

[33] All individuals who have the same profile in the municipality i.e., the same citizenship ('Italian' or 'foreign') get the same corrector value.

$$p_{ij,over} = \frac{Expected\ and\ not\ found_{ij}}{Expected\ and\ not\ found_{ij} + Expected\ and\ Enumerated_{ij}}$$

(c) Calculation of the raw coverage corrector:

$$corr_{ij} = \frac{1 - p_{ij,over-coverage}}{1 - p_{ij,under-coverage}}$$

(d) Calculation of direct and indirect estimates – Direct estimates calibrated for over- and undercoverage for each profile are first calculated for the sampled municipalities. The calibration process constraints the survey sample weights to the known population totals derived from the RBI for each profile. Using small area estimation models, indirect estimates are then calculated to reduce the direct estimates variability for sampled municipalities and to derive estimates for non-sampled municipalities.

(e) Calculation of the average corrector 2018-2019[34] – For each municipality and separately for over- and undercoverage, the average of the 2018 and 2019 correctors is calculated and weighted with the respective demographic sizes. The estimate of the 2018-2019 average corrector is the ratio between the weighted averages of the estimates of the overcoverage corrector and the undercoverage corrector.

6.7.4.3 The population count based on the RBI coverage corrections

311. At the end of the process, a 'weight' is attached to each record in the RBI, which 'corrects' for the coverage errors of the register estimated for a given municipality. The weight applied to register records will be equal to one if the RBI, for a given municipality, is affected by neither overcoverage nor undercoverage errors (or if the two errors compensate each other).

312. If the estimated undercoverage of the RBI is greater than the estimated overcoverage, the corrector applied to each record in the RBI will be higher than one and the total population will be higher than the number of records in the RBI.

313. Conversely, if the RBI's estimated undercoverage is lower than the estimated overcoverage, the corrector applied to each RBI record will be lower than one and the total population will be lower than the number of records in the RBI.

314. After validating the population count, the data collected both in the Areal and the List surveys are used in conjunction with the RBI data and data from the thematic registers on employment and education, using predictive statistical models, to produce data on education, foreign country of citizenship, and labour force status.

[34] Due to insufficient stability of the estimates between 2018 and 2019, an average population corrector of the 2018 and 2019 data has been adopted for each municipality.

Chapter 7. Output Stage

315. This chapter provides a guide to the quality dimensions, some of the key tools, and processes used to assess the measurement of census output quality, where estimates are produced through the integration of administrative data sources into the census design (also see UNECE 2018, Chapter 9). Section 7.1 describes the output quality dimensions on which an assessment should be made and Section 7.2 details additional tools and processes that can be used to assess quality against the dimensions.

316. While measuring output quality moves beyond the quality of the sources *per se*, producing high quality estimates using administrative data is the goal. Therefore, these Guidelines would not be complete without considering output quality. At the same time, it must be emphasized that all the preceding quality stages contribute towards the quality of the outputs. In the case of a combined or full administrative data-based census methodology, a census design which is informed by the rigorous assessment of quality at the source, input, and process quality stages will ultimately result in high quality outputs (see for quality indicators, measures, and methods for assessing output quality).

317. Measuring output quality cannot be reduced to the estimation of overall uncertainty of the estimate (the accuracy dimension); rather, it should include an assessment across all other quality output dimensions. The introduction of administrative data will 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.

7.1 Output quality dimensions

7.1.1 Relevance

318. Relevance refers to the degree to which the census outputs meet the needs of users in terms of both coverage and content. Data are relevant when they relate to the issues data users care about most. This dimension may require NSOs to adjust the direction of their programmes over time, as needed. Assessing relevance is subjective because it often depends on varying user needs. The challenge, therefore, for a census programme is to balance any conflicting user requirements and to go as far as possible towards meeting the most important needs within resource and other constraints (UNECE 2015). Section 7.1.6 provides details on meeting use needs and balancing quality dimensions.

319. Various tools and approaches can be used to assess relevance, including the use of user needs surveys, consultations, and user satisfaction surveys; by building user feedback mechanisms into the census process and by analysing the usages of census data (see UNECE 2018 p28).

7.1.2 Accuracy and reliability

320. The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. More simply put, **accuracy** is the proximity between an estimate and the unknown true value. It is usually characterized in terms of error in statistical estimates and is traditionally broken down into bias and variance. In a census context, variance applies in situations where a portion of the questionnaire is used for a sample of persons or households, where only a sample of records is processed, or can be introduced during the processing stages (e.g., probabilistic imputation and linkage – see Chapter 6). Accuracy can also be described in terms of measurement and representation error, as defined in 0.

321. **Reliability** is the degree of closeness of initial estimates to subsequent estimated values (the concept is listed by the ESS together with accuracy; however, it is also related to comparability - see below). Administrative data, by nature, can be subject to improvements in accuracy over time (e.g., coverage can improve, as lagged registrations and de-registrations become available, and the quality of measurements can also improve). Therefore, an NSO can make use of "new" data to improve their census statistics, revising previous estimates. However, this needs to be balanced against user needs with respect to revisions. Methods for assessing the accuracy of census outputs are provided in the Sections 7.2.1 and 7.2.2 below.

7.1.3 Timeliness

322. **Timeliness** refers to the lapse of time between the period to which the census data refer (e.g., Census Day) and the date of publication of the data. A combined or register-based census often allows for census estimates to be produced in a more timely and frequent manner than a traditional decennial census – indeed, this is one of the greatly-hailed advantages of census transformation. Considering this, the timeliness of estimates that can be produced should be a key quality consideration, and improvements should be made to this aspect wherever possible. The timeliness of the data themselves is an important determinant of the timeliness of the output, thus linking back to the quality Stages discussed in the preceding chapters. There is often a trade-off between timeliness and accuracy. It may be the case that different data users will have different views on the balance between the two, and as such they may not have the same view on the effect of improved timeliness vis-à-vis accuracy (see Section 7.1.6).

323. Several straightforward timeliness metrics can be found within the literature. Quantitative indicators can be applied to measure the time lag for the final results, e.g. between data collection, data acquisition, data linkage, and publication of statistics. For example, the overall timeliness may be calculated as the time from the end of reference period to receiving administrative data supply, divided by the time from the end of reference period to publication date, multiplied by 100 per cent (Eurostat ESSnet KOMUSO 2016; Eurostat 2013; Eurostat 2014; UNECE 2018).

7.1.4 Coherence and comparability

324. The *ESS Quality Framework* (Eurostat 2019) defines **coherence** and **comparability** as the adequacy of statistics to be reliably combined in different ways and for various uses and the extent to which differences between statistics can be attributed to differences between the true values of the statistical characteristics. The ESS Quality Framework and the *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing* (UNECE 2018) expand the definition to include "the degree to which the census information can be successfully brought together with other statistical information within a broad analytical framework." Comparability can be seen as a special case of coherence, where **coherence** is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar, while **comparability** is the degree to which data can be compared across countries, regions, subpopulations, and time.

325. Measuring the extent to which estimates produced using administrative data are internally and externally coherent and comparable is a centrally important aspect of output quality for all census types, including those which make use of administrative data. Such estimates should be coherent with the known characteristics of the population, longitudinally, across geographies and population characteristics (see Section 7.2.2). Sub-totals should correctly sum to overall totals. In addition, it is important to assess the extent to which census integrated statistics are internationally comparable and to communicate this to data users.

7.1.5 Accessibility and clarity

326. **Accessibility** is defined generally as the ease with which data users can access and understand statistical outputs. **Clarity** relates to the availability of any supplementary information or metadata that may be necessary to help the data user interpret and understand the accompanying published data. The concept of 'clarity' is essentially the same as 'interpretability'. Section 7.2.6 provides details of quality reports and metadata that should be accessible and understood by data users.

7.1.6 Meeting user needs and balancing quality dimensions

327. Whether or not administrative data are used in statistical production, assessing the overall quality of estimates produced should consider all the above quality dimensions. This includes not only the accuracy dimension – the aspect which is most often reported in relation to survey methodologies – but also the remaining quality dimensions. In a census context, the overall quality of estimates is about establishing the balance across the quality dimensions which best meets the needs of census data users. To achieve this, it is necessary to not only consult data users throughout the census design process, but also to give them access to the general information and specific metadata they need to appraise quality decisions and to feed back on the QAs undertaken by the NSOs.

Quality reporting and quality metadata are essential (see Section 7.2.6). In addition, the continuous improvement of input and process quality will ensure that output quality also improves. The former will be aided by the implementation of the necessary supplier feedback mechanisms (Section 7.2.3) and the latter through independent expert review of methods (Sections 7.2.4 and 7.2.5).

7.2 Further tools and processes

7.2.1 Assessing the accuracy of population estimates (coverage error)

328. In several countries the overall accuracy of census population estimates has traditionally been carried out based on a DSE framework which involves conducting a traditional census (i.e., taking a census 'stock' at one point in time); followed by a large post-census coverage survey (also at one point in time); and then relying on the DSE which uses capture-recapture methods to estimate under- and overcoverage (O'Hare 2019). These estimates could then be adjusted based on administrative data on deaths, births, and migration flows, for each year between decennial censuses. Alongside this, in some cases (e.g., the UK 2011 Census Quality Survey) countries have carried out small post-census surveys, where data are collected on all census questions and then matched to census responses, to measure respondent error.

329. For some of the census types and use cases described in Chapter 2, the traditional methods for determining overall coverage and quality are still applicable. However, new, or revised methods are necessary in the case of population estimates produced primarily from administrative records, as is the case with a combined or full administrative data-based census. These methods, including the use of dependent interviewing and the "SOL" methodology, were described in Chapter 6. This continues to be an area of significant interest across NSOs, with ongoing developments from countries [see Brown et al (2020) for a summary of new and emerging methods].

7.2.2 Demographic analysis: comparison with alternative sources

330. Demographic analysis (DA)[35] can be applied to assess the accuracy and to understand the coherence and comparability of census outputs. DA involves systematic comparisons, establishing thresholds of acceptability, and understanding any significant discrepancies. It cannot be carried out without the conceptual research at the Source Stage or the validation and harmonization work at the Data Stage. It may also require multiple sources to be combined to meet the target population at the Process Stage.

[35] See O'Hare (2019) for an introduction to the method and its limitations.

331. The census estimates which integrate administrative data are validated against alternative sources – e.g., survey data, previous census data[36], or alternative sources. When using DA, it is important to keep in mind that estimates in two sources can be different across different sex-age or other breakdowns. These differences could be caused by different target populations, different reference dates, or population changes (when comparing to historical census data), by conceptual differences and variations in classification between the variables being compared across sources, and/or by differences in sampling, collection methods, and approaches to data processing. Any such comparisons must be made based on the results of the QA at the Source and Data Stages.

Box 12: Demographic analysis in Spain

In Spain, the pre-censual file (the FPC) is constructed based on the Spanish population register (Padrón) by applying a 'SOL' methodology to enumerate the census population. The population numbers obtained in the FPC are then compared at the minimum geographical level with the official population counts. The main objective is to detect and correct for possible under- and over-coverage problems.

To ensure the quality of the FPC population numbers, the population is disaggregated by the most relevant demographic variables and compared for each level of the variables: sex, age (year by year), nationality type (Spanish/foreign,) and nationality (disaggregated by countries). These micro comparisons help to establish the consistency of common variables.

Analysis of specific sub-populations is carried out to check for possible over-coverage problems. The most significant differences between the pre-censual file and the official population counts are due to the administrative nature of Padrón, as it is not a statistical register but an administrative one and, requires processing to add and remove units as necessary (for instance adding births or removing deaths).

On the other hand, to avoid possible under-coverage in the pre-censual file, all people listed in each of the available administrative sources (e.g., tax files, social security files, unemployment files etc.) that have not been found in Padrón on the reference date, are checked. If there is strong evidence that a person is residing in Spain (given their presence in several administrative sources) but is not registered in Padrón, this person is incorporated into the FPC.

A common example of this situation is people who have been removed from Padrón some months before the census reference date, 1 January, and who then appear on the Padrón again, a short time later, e.g., in February. This may be the case with foreigners whose registration has expired and for which their renewal takes a few months to be processed.

[36] As countries transition from traditional to register-based censuses, they may decide to link their administrative-based statistical population datasets to the traditionally-collected census data in order to provide an assessment of coverage and measurement error, which can inform decisions about the pace of transition to the new methodology.

Having considered the assessment of individual data sources at the Source and Data Stages and of sources combined into statistical registers at the Process Stage, it is possible to make professional judgements about whether differences found through DA are within an acceptable margin. This will vary from country to country and thus it is recommended that such standards are locally developed.

7.2.3 Supplier feedback mechanisms and data quality incentives

332. The continued improvement of census estimates which integrate administrative data relies on the continued improvement of the administrative data collected by the data supplier who supplies it (including the various organizations that might supply data for an administrative register, such as authorities of the municipalities). Achieving this requires adequate feedback mechanisms between the data supplier and the NSOs, and the existence of the right kind of incentives for both the data supplier collecting the data and the individuals whose data they collect.

333. It is often the case that a data supplier is also a data user and will have an interest in the quality of the census results, which can support the relationship between the NSO and the data supplier. Communication between the NSO and the various stakeholders was discussed in detail in Chapter 4. Good communication mechanisms will contribute towards closing the gap between operational and statistical quality, hence ensuring continuous improvement in the quality of the data used in the census and the estimates produced.

334. To support improvements in quality, the NSO can also work with the data supplier to develop suitable tools, systems, and standards (e.g., online interfaces, clear definitions, agreed areas of best practice, etc.) to improve the collection, processing, and transmission of the data.

Box 13: Demographic analysis in Canada

For the first time in 2016, the Canadian census programme gathered income information solely from administrative data sources. The estimates produced with these data were compared, to the extent possible, with other data sources. Comparison analysis focused on various topics including individual income by source, coverage issues, conceptual and processing differences, and regional differences. Given the sensitivity of most income indicators to such methodological differences, however, data users should use caution when comparing 2016 census income estimates to those produced using other household income surveys, administrative data, or earlier census data.

7.2.4 Independent review of methods

335. An independent review of the census design and methods will encourage the continued improvement of quality, i.e., achieving the best balance between quality dimensions to meet the needs of users. Such reviews should be carried out by population and methodology experts.

336. In August 2018, Stats NZ established a panel of experts to provide advice and guidance to Stats NZ on the methods used in creating the 2018 Census data, as well as to data users on the quality of the resulting data. The panel endorsed the statistical approaches used for including administrative enumerations in the data and concluded that the inclusion of those records improved the coverage and accuracy of population counts for the core demographics: age, sex, place of usual residence, and ethnicity (Stats NZ 2019g).

337. Similarly in the UK, the external Methodological Assurance Panel has three aims: 1) to provide external, independent assurance and guidance on the statistical methodology underpinning 2021 census estimates and those based on administrative sources, 2) identify significant gaps and risks in methods and make suggestions for mitigation, and 3) review administrative data methods and contribute to their continuous improvement (UKSA 2018b). The panel review is taking place between 2018 and 2023.

7.2.5 Sensitivity analysis

338. As well as having population experts reviewing the overall method, quality will be improved by engaging experts in an analysis of particularly concerning topic areas or quality decisions throughout all quality Stages, which we will call sensitivity analysis. Sensitivity seeks to establish the extent to which the method used can "count a population within a geographic region or demographic group", which "can be used to understand bias in census data, and plan for the next census by identifying the groups most difficult to count" (Stats NZ 2019e, p.5).

339. Stats NZ engaged external providers to assess both the methods used to add people to the 2018 Census data and the fitness of the data for three important use cases, including the determination of electoral boundaries. A sensitivity analysis of the methods used to add people to the 2018 Census file found that the threshold for inclusion in meshblocks had the most impact on who was included in the census file and that the threshold used was a sensible balance. Further sensitivity analysis determined that 2018 Census data was robust for the purpose of determining electoral boundaries and the electorate boundaries drawn using census counts were not likely to be impacted by the choice of threshold for adding administrative enumerations at the meshblock level (Stats NZ 2019d; Stats NZ 2019e). This was an important finding in support of the quality of the census data.

7.2.6 Quality reports and metadata

340. Within the last QA Stage, a report should be produced to document the results of the QA and assurance throughout the census production. This report should include information against each QA Stage as well as communicate to data users where and how each quality dimension was considered. This will enable the producers and data users to appraise and provide feedback on the quality decisions, to determine whether the right balance has been achieved across the quality dimensions, and if there is sufficient metadata for the QA.

Box 14: Quality metadata in Spain

In the work leading up to the first administrative-data-based census in Spain, an extra categorical variable providing a data QA based on the origin of each value is being developed. This will provide data users with a variable-specific quality indicator (Pérez Julián, Casaseca and Argüeso Jiménez 2018). As previously noted, in Spain, a population statistical register is created by linking the population administrative register (Padrón) with multiple administrative sources. This can be visualised as a huge matrix in which the census variables are considered columns and each person is presented by a row, so the matrix cells would contain values for each individual per variable. To help users, understand the quality of census data, for each census variable, another categorical one will be added to inform data users of the quality of each cell value. As explained below, this categorical variable is intended to inform data users on quality, either directly or indirectly.

The initial proposal to develop this quality measure, for each cell, is based on the type of methodology or source used to fill each cell value (see Table 7). Typically, a cell value derived from an up-to-date administrative source has the highest quality and one derived through deterministic imputation the lowest. In this way, the quality of each cell value can be understood by data users in an indirect way.

Additionally, the quality measure for each cell value depends not only on the nature of the underlying source and methodology, but also on the rest of the individual's characteristics. For instance, where a 20-year-old person is missing values for the variables LMS and his/her main industry and these are deterministically imputed to 'single' and 'Accommodation and food services' respectively, chances are the first imputed value is much more reliable than the second one. The relation between age and LMS is likely to produce good deterministic imputation estimates, while this is not the case when imputing a value for industry. Several such rules have been developed to inform the quality of imputations based on known individual characteristics.

Another proposal is a more direct way which would be to provide a quality punctuation variable, for example, on a scale from 1 to 4, where 1 would be the highest quality and 4 the lowest one. This would help data users understand how 'good' or 'bad' an imputed value can be considered.

Both mechanisms, the indirect or direct, offer enormous potential in the assessment of output quality in two dimensions: by variable and by unit or subpopulations. It is proposed that all data users should have free access to these quality variables in the census microdata release for 2021 (approximately 10 per cent of the whole census product) and would have specific methodological notes with explanations.

*Table 7: Initial proposal of categories indicating source quality by type**

DATA TYPE	DESCRIPTION	QUALITY**
DS	Information provided by direct sources up to date	Highest
DSN	Information provided by direct sources but not up to date	
CS	Past census information	
PI	Probabilistic imputation	
DI	Deterministic imputation	Lowest

*Adapted from Pérez Julián, Casaseca and Argüeso Jiménez (2018), p.4

** The actual order is not immutable and will depend on the variable, the sources used and the quality of the underlying imputation process.

7.3 Case studies

7.3.1 Portugal: quality assessing the population register

7.3.1.1 Background

341. The *Census Admin* project (short for *Census with Administrative Data*) is part of the framework for the development of a national data infrastructure which includes Statistics Portugal's (SP) strategy of data integration, from several sources, to respond to an increasingly complex society with new expectations towards statistics.

342. Central to the project is the creation of a resident population dataset (object type statistical population dataset or SPD), covering a set of characteristics (geographical, demographic, and socio-economic) of the resident population in Portugal. The SP´s goal is to report population statistics using the SPD from the 2021 Census onwards.

343. The SPD prototype was built in 2015 with reference to the 2011 population. Meanwhile, four new annual editions were created, with annual reference dates from 2015 to 2018.

344. For each annual edition, the consistency of the SPD results is evaluated by systematically comparing it against population estimates and known population characteristics. Additionally, comparisons with the census' test results have been considered to measure the quality of the SPD results.

7.3.1.2 SPD population counts by geographical level

7.3.1.2.1 Evaluation of 2018 SPD results at national and regional levels

345. The resident population in Portugal, estimated through administrative data by the 2018 SPD is 10,300,502 persons, representing a relative deviation of +0.2 per cent when compared to the 2018 population estimates (PE) released by SP. The PE provide the

official figures of the annual resident population in Portugal, using cohort components and the population census concept. The calculations are based on the natural and migratory demographics, with information from: live births, deaths, emigration, and immigration estimates.

346. The national level results obtained in the *Census Admin* project are very promising, considering the different assumptions, methodologies, and distinct sources of the two types of statistical output: SPD and PE. Consistently, across all the annual editions of the SPD, relative deviation between these two sources is less than 0.5 per cent (under- or overcoverage).

347. At a regional level (NUTS II), the 2018 SPD-PE relative deviation fluctuates between -0.4 (Centro) to 3.5 per cent (Algarve); Lisbon Region with -0.1 per cent.

348. The results of the Portuguese SPD are also promising at the municipality level: for 2018, more than 76 per cent of the 308 municipalities present levels of under or over coverage within 5 per cent, when compared to the PE; it should be noted that in 64 municipalities, the relative deviation SPD-PE is under 1 per cent (under- or overcoverage). Only a small number of, mostly less populated, municipalities (15), show relative differences greater than 10 per cent (higher or lower).

349. Along with the geographical distribution, the SPD generally aligns with the PE in terms of basic demographic and socio-economic dimensions. For example, the SPD-PE relative differences in the age structures of PE are very small for most age groups and across all SPD versions (the largest differences occur in elderly people).

7.3.1.2.2 Evaluation of 2015 SPD results at local level

350. Comparisons have also been carried out at a lower geographical level; the parish or Local Administrative Units – level 2 (LAU2). As detailed below, the 2016 Census Test (CT), (reference date 26 September), contributed to the evaluation of the 2015 SPD results (reference date 31 December) at the LAU2 level.

351. The analysis of the results of the 2016 CT showed that in four of the five parishes in the sample, where it was possible to guarantee exhaustiveness on data collection, 2015 SPD estimated more people than those that were enumerated. The relative deviations varied from -14.1 per cent to -5.7 per cent. Overall, the population counts in the 2016 CT, when compared to the estimated 2015 SPD, had a deviation of -8.8 per cent.

352. To further evaluate the 2015 SPD estimates, microdata from 2016 CT was linked to 2015 SPD results and for individuals who matched (about 80 per cent), their characteristics were compared. For place of usual residence, for example, 90 per cent of respondents were found in the same LAU2 as that registered in the 2015 SPD (quite satisfactory considering the 9 months' time lag between the CT and SPD reference dates).

353. If we take the place of usual residence at the municipality level as a basis for comparison, the equality rates are overall around 93 per cent, since 3.2 per cent of the individuals in

the 2016 CT matched to the 2015 SPD lived administratively in another parish of the same municipality.

7.3.1.3 Final observations

354. The focus of this work is to assess the quality of the Portuguese SPD to estimate the resident population.

355. For that purpose, we showed results of several comparisons: with the population estimates, disaggregating by geographical level (national to regional total) and with the census tests (finer geographical level).

356. The set of administrative information currently integrated in the SPD has a high potential for the transition to a registered-based or combined census model. At a national and regional level, the SPD results are very promising. However, at a lower geographical level, comparison with the census test showed that the SPD estimates can be improved. SP is looking to develop more robust estimation methods and to review the 'SOL' rules. Nevertheless, although the population counts at parish level present some differences, the structure and characterization of the parish population given by the 2015 SPD is very consistent with that collected in 2016 CT.

Chapter 8. Conclusions and recommendations

357. Administrative data can be used across the different census methodologies and to support all stages of the census process, including:

    (a) Constructing an address frame,

    (b) Supporting field operations,

    (c) Enumerating the population,

    (d) Collecting census variables,

    (e) Conducting QA,

    (f) Editing and imputation, and

    (g) Modelling and estimation.

358. Their use can provide more frequent and timely statistics about the population, improvements in accuracy and reliability, and significant reductions in costs and respondent burden.

359. However, there are significant quality challenges to assess and overcome before an administrative source can be used in a census. Most significant among these is that administrative data have, in general, not been collected for the purpose of a census. As such, the NSO may have little control over:

    (a) Concepts and definitions used,

    (b) Target population,

    (c) Data collection method,

    (d) Processing and QA procedures,

    (e) Data methods, and

    (f) Structures and systems used.

360. The Guidelines present the QA Stages, set against quality dimensions, with associated tools and indicators to lead the user through the process of assessment. The application of the Guidelines should help readers make decisions about the use of administrative data in the census, whilst supporting a process of continuous assessment and improvement. Throughout the Guidelines, proposals and recommendations have been made, which are summarized below.

8.1 Recommendations

(a) The NSO should **identify the administrative sources** that may be relevant to their census, set against different use cases. It is important to **set out what the expected or required outcomes of using the source would be**, against which an assessment of relevance can be made. 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. Central to such assessment is setting out what the administrative source needs to deliver in terms of the target population and the required measurements from this population for the census use case. Chapter 2 of the Guidelines and the case studies across the other chapters provide examples of how administrative data have been used in several different countries.

(b) The **relationship between the NSO and the administrative data supplier** is of critical importance (Chapter 4). This should be supported by robust mechanisms of communication, written agreements, and an excellent understanding of the needs of both parties. There must also be an agreed legal basis for the data supply and use of the data. To help build the relationship and secure a data supply, the NSO should identify areas of benefit to the data supplier. This could be feedback mechanisms to help the data supplier better understand their data through collaborations on areas of common interest, or by helping the data supplier (using their data in the census) to support the wider good. Of course, providing possible quality concerns with the data has the added benefit of facilitating ongoing quality improvements.

(c) **The NSO should engage with the data supplier to gain an in-depth understanding of the data source**. This should translate into the creation of clear and comprehensive metadata about the administrative source. The metadata will provide a useful reference both for the census and for any other surveys that might benefit from using the source. Chapter 4 provides details on the metadata that should be collected, along with various references to the relevant literature.

(d) Since administrative data are generally not collected for the needs of the census, it is important for the NSO to **understand and assess the differences between the required populations, concepts, definitions, and time-related dimensions**. More generally, a thorough assessment of the coherence and comparability of the administrative source, along with its limitations across the various quality dimensions, is essential. This includes the linkability of the source if this is a requirement for use in the census. This assessment will inform the processing stages, including editing and imputation, and the linkage and integration of sources, where decisions are made between and across sources based on their quality (see Chapter 6).

(e) The NSO must **understand any restrictions and challenges to acquiring and integrating an administrative source into the census** (Chapter 4). This could include resources and costs, risks associated with the data supplier's ability to deliver on time to the required quality standards, and whether the use of a source is acceptable to the public and census data users. There are important trade-offs that the NSO must consider. Specifically**, the value of the administrative source must be assessed against its usefulness for the census, set against the effort and risks of acquiring and using the data**.

(f) The NSO has limited control over the collection and processing of the administrative data, which can be subject to changes in population coverage and population measures over time. This can be due, for example, to legal, policy, procedural, or system changes affecting the data and/or their delivery (Chapter 4). **The NSO must assess and manage an acceptable level of risk**. The risk should be managed by working with the data supplier 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, whether using other data sources or by adapting processes/methodologies (Chapter 6).

(g) **It is important that the public and data users understand how and why administrative data are being used in the census** (Chapter 4). The NSO should, therefore, be transparent about the use; providing a clear justification of the benefits against any risks and costs (i.e., a strong proportionality case exists). 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**.

(h) The inclusion of administrative data sources in census production should be preceded by adequately resourced feasibility research which **provides a 'proof of concept' for the planned integration of administrative data into the census production**. It is advisable to carry out test runs (using real data) well in advance of the main census to ensure any unforeseen issues are identified, allowing enough time to correct or adjust the methods, processes, or systems (as described in Chapter 5 and Chapter 6).

(i) **Expert review** (working with data suppliers and subject experts) and **comparisons between sources over time are important to identify any quality concerns** with a source or register. The use of well-designed surveys (linked to administrative data or registers) can be particularly important in identifying and adjusting for coverage and measurement errors (Chapter 5, Chapter 6 and Chapter 7).

(j) **The NSO should record and publish the results of the QA and assurance throughout the census production**, including the Data, Process, and Output Stages. This will enable NSOs and census data users to appraise and provide feedback, supporting an ongoing dialogue. This is important to ensure that

data users understand the strengths and limitations of the data. This can also help determine whether the right balance has been achieved across the quality dimensions (Chapter 7).

(k) The NSO should **develop its own QA framework and strategy**, supported by clear, comprehensive documentation, and training procedures. These Guidelines provide a useful basis to support this, along with the reference material and case studies within the Guidelines. The strategy should build the continuous assessment and improvement of administrative data into the plans and procedures for the census. This should include the communication links between the NSO, the data users, and the data suppliers.

8.2 Areas for further development

361. The Guidelines have focused on the QA of administrative sources for use in censuses, while providing some information about the processes used to integrate and transform data to improve quality. The quality of census outputs that use administrative data is also covered briefly. However, the Guidelines do not provide a **wider total error framework or a model** for how the error from each source translates into the error in the final census estimates, taking account the changes in quality due to processing (which can reduce or increase error).

362. The development of such a model that takes into account all error sources is partly addressed by the total error framework adopted by Statistics New Zealand (Reid et al 2017). The framework builds on Zhang's (2012) extension of the Total Survey Error (TSE) paradigm (see Groves and Lyberg 2010; Biemer 2010). It has three phases covering:

(a) An assessment of the single sources,

(b) An integrated data set assessment, and

(c) An estimation and output assessment.

363. The work of the ESSnet KOMUSO on the quality of multiple source statistics (Eurostat ESSnet KOMUSO 2019) also provides a useful framework for assessing the quality of statistical outputs based on multiple sources (survey and administrative data).

364. This could be an area for further development and international collaboration with **a specific focus on how such a framework can be applied to censuses**. This could include examining how a total error framework or model can be developed and used to assess the quality of census outputs based on multiple sources. It could also include work to understand how the impact (and compounding impact) of various errors across the stages of the census can inform decisions about the best overall statistical design for the census.

365. The Guidelines have focused on the assessment of administrative data, but there are **other sources of commercial data that present opportunities for use to improve or**

enhance census statistics (e.g., geospatial data, mobile phone data). The quality Stages, dimensions, tools, and indicators within these Guidelines are to a great extent applicable to sources beyond administrative data. This too could be an area requiring further international work, with a specific focus on whether and how the tools and techniques for QA of such sources, for use in the census, differ from those identified here.

366. Finally, during the CES-wide consultation on the draft of these Guidelines, several countries suggested that there is a need for further work to examine countries' experiences and consider what constitutes good practice in assessing the quality of administrative sources as they relate to coverage of **hard-to-reach groups**. This may be an area that the international statistical community might wish to pursue in future. Such work could involve a wider community of experts beyond those in the census field, since the topic is one that is relevant to other statistical areas.

References

Abbott, O. (2009). 2011 UK Census Coverage Assessment and Adjustment Methodology. Population Trends 137, Autumn 2009.

Abbott, O., B. Tinsley, S. Milner, A. C. Taylor, and R. Archer (2020). Population statistics without a Census or register. Statistical Journal of the IAOS, 36(1), 97-105.

Argüeso Jiménez, A. (2019). Population and Housing Census in Spain will be fully register-based.

Asamer E.-M., F. Aztleithner, P. Ćetković, S. Humer, M. Lenk, M. Moser and H. Rechta (2016). Quality Assessment for Register-based Statistics - Results for the Austrian Census 2011. Austrian Journal of Statistics Vol. 45, No. 2, pp. 3-14

Australian Bureau of Statistics (2009). ABS Data Quality Framework, May 2009, cat. no. 1520.0, ABS, Canberra.

Bakker, B. F. M. (2012). Estimating the validity of administrative variables. Statistica Neerlandica (2012). Vol. 66, nr. 1, pp. 8–17.

Biemer, P. P. (2010). Total Survey Error: Design, Implementation, and Evaluation, Public Opinion Quarterly, Volume 74, Issue 5, pp. 817–848

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.

Cerroni, F., G. Di Bella and L. Galiè (2014). Evaluating administrative data quality as input of the statistical production process. In: Rivista Di Statistica Ufficiale. ISTAT.

Chambers, R. (2001). “Evaluation Criteria for Statistical Editing and Imputation.” National Statistics Methodological Series 28: 1–41.

Chieppa, A., G. Gallo, V. Tomeo, F. Borrelli and S. di Domenico (2018). "Knowledge discovery for inferring the usually resident population from administrative registers" in Mathematical Population Studies, Pages 92-106.

Choi, H. (2019). Adjusting for linkage errors to analyse coverage of the administrative population. Statistical Journal of the IAOS, 35(2), 253-259.

Cornell University Research Data Management Service Group (2020). Guide to writing "readme" style metadata. https://data.research.cornell.edu/content/readme

Crescenzi, F., G. Sindoni and D. Zindato (2014). Lessons learned from the 2011 Italian census and innovations leading towards a continuous census. Note by the National Institute of Statistics of Italy, presented at the UNECE/Eurostat Group of Experts on Population and Housing Censuses, Sixteenth Meeting, Geneva, 23-26 September 2014.

Daan Zult, P., P de Wolf, B. Bakker and P. van der Heijden (2019). A linkage error correction model for population size estimation with multiple sources.

Daas, P.J.H., J. Arends-Tóth, B. Schouten and L. Kuijvenhoven (2008). Quality Framework for the Evaluation of Administrative Data. Paper presented at the European Conference on Quality in Official Statistics.

Daas P., S. Ossen, R. Vis-Visschers and J. Arends- Tóth (2009). Checklist for the Quality evaluation of Administrative Data Sources. The Hague: Statistics Netherlands.

Daas, P., S. Ossen, M. Tennekes, J. Burger and F. Cobben (2012). Input Quality of administrative data (BLUE-ETS WP4). Presented at Quality 2012.

European Commission (2008). Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses as regards the technical specifications of the topics and of their breakdowns.

Eurostat BLUE-ETS (2011). List of quality groups and indicators identified for administrative data sources.

Eurostat (2013). Use of administrative and accounts data in business statistics

— (2014). ESS Handbook for quality reports.

— (2017). European Statistics Code of Practice, revised edition 2017.

— (2019). Quality Assurance Framework of the European Statistical System. Version 2.0.

— (2020). European Statistical System Handbook for Quality and Metadata Reports,

Eurostat ESSnet KOMUSO (2016). Checklist for Evaluating the Quality of Input Data https://ec.europa.eu/Eurostat/cros/system/files/essnet_wp1_report_final_version4.pdf

— (2019). Quality Guidelines for Multisource Statistics https://ec.europa.eu/Eurostat/cros/system/files/qgmss-v1.1_1.pdf

Eurostat ESSnet MIAD (2014). MIAD deliverable B2, B3 - Quality check list for the Source phase [Data]. Available from: https://ec.europa.eu/Eurostat/cros/content/miad-deliverable-b2_en

Falorsi, S. (2017). Census and Social Surveys Integrated System. Note by the National Institute of Statistics of Italy, presented at the UNECE/Eurostat Group of Experts on Population and Housing Censuses, Nineteenth Meeting, Geneva, Switzerland, 4–6 October 2017.

Gallo, G., A. Chieppa, V. Tomeo and S. Falorsi (2016). The integration of administrative data sources in Italy to increase Population Census data availability. Note by the National Institute of Statistics of Italy, presented at the UNECE/Eurostat Group of Experts on Population and Housing Censuses, Eighteenth Meeting, Geneva, Switzerland, 28-30 September 2016.

Gath, M., and C. Bycroft (2018). The potential for linked administrative data to provide household and family information.

Gibb, S., C. Bycroft and N. Matheson-Dunning (2016). Identifying the New Zealand resident population in the Integrated Data Infrastructure (IDI).

Groves, R. M., F. J. Fowler Jr., M. Couper, J. M. Lepkowski, E. Singer and R. Tourrangeau (2004). Survey methodology, Wiley, New York.

Groves, R. M. and L. Lyberg (2010). Total Survey Error: Past, Present, and Future, Public Opinion Quarterly, Volume 74, Issue 5, 849–879.

Harron, K., H. Goldstein, and C. Dibben (2015). Introduction. In K. Harron, H. Goldstein, & C. Dibben (Eds.), Methodological Developments in Data Linkage. New York: John Wiley & Sons.

International Organization for Standardization (2015). Quality management systems — Fundamentals and vocabulary. ISO 9000:2015(en)

Iwig, B., M. Berning, P. Marck and M. Prell (2013). Data Quality Assessment Tool for Administrative Data

ISTAT (2020). Nota tecnica sulla produzione dei dati del Censimento Permanente: la stima della popolazione residente per sesso, età cittadinanza, grado di istruzione e condizione professionale per gli anni 2018 e 2019: Dalla rilevazione "porta a porta" al Censimento permanente [Technical note on the production of Permanent Census data: estimating the resident population by sex, age, citizenship, education level and occupational status for the years 2018 and 2019: From door-to-door survey to permanent census]. Rome: ISTAT.

Lavigne, M., and C. Nadeau (2014). A Framework for the Evaluation of Administrative Data. In Proceedings of Statistics Canada Symposium.

Lothian, J., A. Holmberg, and A. Seyb (2019). An evolutionary schema for using "it-is-what-it-is" data in official statistics. Journal of Official Statistics 35, 137-165.

Oberski, D. L. A. Kirchner, S. Eckman & F. Kreuter (2017). Evaluating the Quality of Survey and Administrative Data with Generalized Multitrait-Multimethod Models, Journal of the American Statistical Association, 112:520, 1477-1489

O'Byrne, E., C. Bycroft and S. Gibb (2014). An initial investigation into the potential for administrative data to provide census long-form information: census transformation programme. Wellington: Statistics New Zealand.

O'Hare, W.P. (2019). Methodology Used to Measure Census Coverage. In: Differential Undercounts in the U.S. Census. Briefs in Population Studies. Springer, Cham

Office for National Statistics (ONS) (2016). Methodology of Statistical Population Dataset V2.0.

— (2019). Developing our approach for producing admin-based population estimates, England, and Wales: 2011 and 2016.

— (2020a). Joined up data in government: the future of data linking methods. Data and Analysis Method Reviews.

— (2020b). ONS working paper series no 20 – Feasibility of using donor-based imputation for census outputs on number of rooms using Valuation Office Agency data: Demonstration of using a donor-based imputation method (CANCEIS) to address missing values when replacing the number of rooms question on Census 2021. ONS working paper series 20, 31 July 2020, Newport: ONS.

Pérez Julián, M. P., C. Casaseca and A. A. Argüeso Jiménez (2018). Assessing quality in a register-based census. Paper presented at the European Conference on Quality in Official Statistics, Krakow.

Reid, G., F. Zabala, and A. Holmberg (2017). Extending TSE to Administrative Data: A Quality Framework and Case Studies from Stats NZ, Journal of Official Statistics, 33(2), 477-511.

Rogers, N., and L. Blackwell (2020). A statistical quality framework for longitudinally linked administrative data on international migration.

Schnetzer M., F. Astleithner, P. Cetkovic, S. Humer, M. Lenk and M. Moser (2015). Quality Assessment of Imputations in Administrative Data, Journal of Official Statistics, Vol. 31, No. 2, pp. 231–247.

Scholtus, S. and B. Bakker (2013). Estimating the Validity of Administrative and Survey Variables by Means of Structural Equation Models. *New Techniques and Technologies for Statistics* conference 2013.

Shafer, G. (1992). Dempster-Shafer theory. Encyclopaedia of artificial intelligence.

Shipsey, R. and J. Plachta (2020). Linking with anonymised data- how not to make a hash of it.

Statistics Austria (2019). Quality assessment of administrative data - Documentation of Method.

Statistics Canada (2017). Statistics Canada's Quality Assurance Framework

— (2020). 'Counting all Canadians' in Painting a Portrait of Canada: The 2021 Census of Population.

Statistics Estonia (2019). Pilot Census Report.

Statistics New Zealand (2017). Experimental population estimates from linked admin data: 2017.

— (2019a). Overview of statistical methods for adding admin records to the 2018 Census dataset

— (2019b). Linking 2018 Census responses to the Integrated Data Infrastructure

— (2019c). Dual system estimation combining census responses and an admin population

— (2019d). Electoral boundaries sensitivity analysis of 2018 Census data.

— (2019e). Predicting the quality of admin location information for use in the 2018 Census.

— (2019f). Population counts sensitivity analysis of 2018 Census data.

— (2019g). 2018 Census External Data Quality Panel: Assessment of variables.

— (2020). Guide to reporting on administrative data quality.

UK Statistics Authority (UKSA) (2015a). Quality Assurance of Administrative Data - Setting the Standard.

— (2015b). Administrative Data Quality Assurance Toolkit. Version 1 January 2015.

— (2018a). Code of Practice for Official Statistics (Edition 2.0).

— (2018b). Methodological Assurance Review panel – Census.

— (2019). Quality Assurance of Administrative Data (QAAD) toolkit.

— (2020). Ethics Self-Assessment Tool.

UNECE (1992). Fundamental Principles of Official Statistics. Available from https://unece.org/statistics/fundamental-principles-official-statistics

— (2011). Using Administrative and Secondary Sources for Official Statistics: A Handbook of Principles and Practices.

— (2014). A Suggested Framework for the Quality of Big Data.

— (2015). Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing.

— (2017). Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources. Version 2.0.

— (2018a). Guidelines on the use of registers and administrative data for population and housing censuses.

— (2018b). Annex F – Portugal Case Study in Guidelines on The Use of Registers and Administrative Data for Population and Housing Censuses. pp.64-67.

— (2021). Keeping Count: conducting the 2020 round of population and housing censuses during the Covid-19 pandemic.

United Nations (2009). Handbook on Geospatial Infrastructure in Support of Census Activities Studies in Methods: Series F No. 103, ST/ESA/STAT/SER.F/103 New York: United Nations Department of Economic and Social Affairs, Statistics Division.

United States Census Bureau (2009). History: 2000 Census of Population and Housing.

— (2019). Counting the Hard to Count in a Census. *Select topics in international censuses*.

Vega Valle, J.L., A. Argüeso Jiménez and Pérez Julián, M. (2020). 'Moving Towards a Register Based Census in Spain'. Statistical Journal of the IAOS.1 Jan. 2020: 187 – 192.

Yucel, R.M. and A.M. Zaslavsky (2005). Imputation of binary treatment variables with measurement error in administrative data. Journal of the American Statistical Association, Vol. 100, No. 472, 1123–1132.

Zhang, L.-C. (2012). Topics of statistical theory for register-based statistics and data integration. Statistica Neerlandica 66: 41–63.

Glossary of Terms

Accessibility: The dimension of quality that is defined generally as the ease with which users can access the data.

Accuracy: The dimension of quality that refers to the degree to which the information correctly describes the phenomena it was designed to measure. More simply put, accuracy is the proximity between an estimate and the unknown true value.

Address register: A register of residential addresses, often used for the purposes of creating enumeration areas comprising comparable numbers of dwellings. In cases of multi-occupied dwellings there can be more than one dwelling under a given residential address.

Administrative enumeration (New Zealand): The process of collecting data taken from an administrative source for the purpose of supplementing data recorded on questionnaires collected in a field enumeration.

Administrative data: Data held on registers and other administrative sources relating to information collected by government and/or other organizations primarily for administrative (not research or statistical) purposes, such as registration, transaction and record keeping, usually for the provision of public services.

Administrative (data) source: A data holding that contains information collected primarily for administrative (not research or statistical) purposes. Such sources include administrative registers (with a unique identifier) and possibly other administrative data without a unique identifier.

Administrative population: The population set of objects or units (e.g., people, dwellings, businesses) that is captured by the administrative source or register.

Administrative register: A systematic collection of unit-level data organized in such a way that updating is possible (where 'updating' is the processing of identifiable information with the purpose of establishing, bringing up to date, correcting or extending the register. Such registers are primarily used in an administrative information system in which the data are used in the production of goods and services in public or private institutions or companies. Administrative registers used for statistical purposes are normally operated by the state or jointly by local authorities, but some registers operated by private/commercial organizations may also be used.

Administrative unit: The units for which administrative data are recorded. These may or may not be the same as those required for the statistical output (which are referred to as statistical units).

Attribute: A socio-demographic or economic characteristic relating to an administrative or statistical unit for which information is required for the purpose of the census.

Base register: Registers upon which the whole system of register-based statistics depends. They include both administrative base registers and statistical base registers, with the former

being resources kept for public administration, and the latter being based on the administrative register, with the key roles of defining important populations and containing links to other base registers.

Benchmarking: Comparing data, metadata, or processes against a recognized standard.

Big data: Large, often unstructured data sets that are available, potentially in real time, but which are difficult both to process efficiently and quality assure using traditional methods and technologies. The amount and variety of data available is growing rapidly, and such data sets are available in many formats, including audio, video, computer logs, purchase transactions, sensors, and social networking sites. Some of these data are freely available on the web, whereas others are held by the private sector to which there may be no free access.

Census day: The date of the reference period for the census, irrespective of when the data are collected.

Census estimates: A term used by some countries to describe the census output data to reflect the fact that the published figures do not purport to be true counts and that there must always be some degree of uncertainty (however small) in the accuracy of the numbers.

Clarity: The dimension of quality that relates to the availability of any supplementary information or metadata that may be necessary of help the user to interpret and understand the accompanying data.

Classifications: Statistical classifications provide a set of related categories in a meaningful, systematic, and standard format e.g., the NSO's standard for classifying occupations. Classifications are generally developed to support policy making and because of that, to organize and present statistics.

Coherence: The dimension of quality that refers to the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar.

Combined census: A census based on a combination of data taken from administrative registers and collected on questionnaires.

Comparability: The dimension of quality that refers to the degree to which data can be compared over time and domain.

Daas hyperdimensions: High-level dimensions or 'views' of quality of an administrative source to be used for statistical purposes. The three key dimensions refer to: the source; the metadata; and the data itself.

Data controller: See 'Register owner'.

Data editing: The process by which data that exhibit errors, logical inconsistencies and spurious values are detected and corrected.

Dempster-Shafer theory: A generalization of the Bayesian theory of subjective probability.

Derived variable: A new variable formed by using the data from other variables**.**

Dual System Estimation (DSE): A statistical method, based on a capture-recapture technique, applied to estimate the size of a population.

Estimates: The term is used in these Guidelines to refer to the statistics produced in census outputs, and reflects the processes undertaken by NSOs to adjust the input data to take account of under- or overcoverage, errors, missing counts, and measures to control statistical disclosure.

Field enumeration: The process of collecting information on individual persons, households and/or housing unit covering the whole population (or a sample of it) using questionnaires.

Frame: Any list, material or device that delimits, identifies, and allows access to the elements of the target population. A statistical register is a specific example.

Hard-to-reach: Groups that tend to be under-represented either because they are numerically very small; because they are hard to identify, e.g. due to a lack of standardized definitions or because of a lack of data collection on the relevant variables; because they choose not to be identified, e.g. due to stigma associated with group membership; because they are systematically excluded from standard collection techniques and sampling frames, e.g. people living in institutions; because are physically hard to reach, e.g. those living in remote areas or without a fixed abode; or because they are hard to enumerate even once identified and sampled, e.g. people living with dementia, people who do not speak the national language.

Imputation: The process by which missing input data items are replaced with plausible and consistent values.

Input data: The data (sometimes referred to as 'raw data') derived from an administrative source before any processing or validation by the NSO.

Input quality: The quality of administrative data sources set against their use in a census. the Source and Data Stages together provide an overall assessment of input quality.

Linkability: The ability to link data from several different administrative data sources to the same unit, usually by means of a unique identification number or code.

Measurement error: error in the measurement of variables or characteristics (e.g., age, gender etc). They include several types of error within variables including relevance (definition misalignment), mapping (errors in the re-classified measures due to poor equivalence between supplied and target classifications which may therefore require adjustments, e.g., through imputation) and comparability errors (errors between the re-classified and adjusted measures).

Meshblock: The smallest geographic unit for which statistical data is collected and processed by Statistics New Zealand.

Metadata: Data that describe or define other data. This broadly refer to anything that users need to know to make proper and correct use of the real data, in terms of accessing, processing, interpreting, analysing, and presenting the information. Metadata include, for

example, file descriptions, codebooks, processing details, sample designs and fieldwork reports. Metadata should be distinguished from 'Paradata' which generally refer to the details that describe the process by which the census data are collected, either from administrative sources or a field enumeration/survey.

Objects: In some of the literature (e.g., Zhang 2012), the term 'object' is used to refer to the units within an administrative dataset. The term is used to distinguish between units in the administrative data and the statistical units after this data has been transformed in some way. This is particularly relevant in cases where the unit (or 'object') in the administrative register differs from the target statistical unit. For example, where a tax register, where the units of a yearly tax returns (i.e., the same person may make several returns in one or multiple years), is converted into individual 'people'.

Output data: The processed data as it is used in statistical outputs.

Output quality: The quality of the processed data as it is used in statistical outputs.

Padrón (Spain): The Spanish population register, usually compiled for each Municipality.

Paradata: See 'Metadata'.

Periodicity: Within the context of the supply of administrative data, this is the time period between reference dates for consecutive input datasets. For the census more generally, it is the time between the dates of consecutive censuses (census days).

Population register: A statistical register and a frame of persons usually resident (however defined) in a country. Additionally, it often provides some demographic characteristics of individuals.

Privacy Impact Assessment: A process which assists organizations in identifying and managing the risks to privacy arising from new projects, initiatives, systems, processes, strategies, policies, and business relationships.

Process quality: The effect of changes to the quality of data being used for the purpose of the census during the processing of the raw data by the NSO.

Punctuality: The dimension of quality that relates, when referring to data, to the time lag between the planned (and often pre-announced) publication dates and actual publication dates. In the context of the administrative source, it relates to the time lag between the expected (or contracted) date of the delivery of the data to the NSO and the actual date of delivery.

Raw data: See 'Input data'.

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, updating, correcting, or extending the register.

Register-based census: A census where all data is collected from administrative registers. A census based on combination of data taken from registers and questionnaires is called a 'combined census'.

Register keeper: See 'Register owner'.

Register owner: The authority responsible for keeping and maintaining an administrative register (also referred to as the 'Register keeper' or 'Data controller'.

Relevance: The dimension of quality which, when referring to data, refers to the degree to which they meet the needs of users in terms of coverage and content. When referring specifically to data sources, the dimension refers to the degree to which such sources contain data that meets the needs of the NSO with respect to their intended use.

Reliability: The dimension of quality that refers to the degree of closeness of data values to earlier or subsequent data.

Representation error: error in the representation of the intended population units or objects (e.g., individuals or households in a census). They include errors relating to over and undercoverage (lack of alignment with target population), identification (errors in classifying a unit based on inconsistencies across multiple sources) and unit errors (errors in the statistical creation of statistical units of interest where they do not exist in any available data source).

Rolling census: An alternative approach to the traditional model of census taking by means of a cumulative continuous survey, covering the whole country over a specific time, rather than on a particular day. There are two main parameters to consider in a rolling census: (a) the length of the periodicity, which itself is linked to the frequency of updating required; and (b) the sample size, which depends on the budget and the level of geographical analysis required for dissemination.

'Signs of life': An indicator used to minimize the overcoverage of persons recorded on different administrative registers derived by applying strict criteria or 'activity rules' to ensure that only living individuals who are usually resident are included in the census estimates.

Source quality: The quality of administrative sources from which data is supplied to NSO for the purpose of the census.

Statistical disclosure control: The process(es) by which the raw data taken from an administrative source or collected in the field is modified during data processing to avoid the disclosure of information about identifiable individual persons or households.

Statistical register: A register processed for statistical purposes. A statistical register could be based on one or several administrative registers. Statistical registers are also referred to as 'secondary registers.'

Target population: The universe for which information is required. The target population is the set of the statistical units.

Test data: smaller supplies of data from an administrative source/register shared with NSOs for the purposes of feasibility research and the testing of systems.

Timeliness: The dimension of data quality that refers to the lapse of time between the period to which the data refer (in the case of census data this is usually Census Day) and the date of publication of the data. In the use of administrative data, timeliness also refers to the length of time between the date of the event recorded in the data source and the date when the data are delivered to the NSO.

Unit: The smallest entity to which any administrative data item refers. For census purposes, units may refer to individual persons, households, buildings, or dwellings.

Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses

The use of administrative data for censuses continues to increase across the countries of the UNECE region and beyond. Administrative sources are used in a wide variety of ways, from supporting operations in a traditional census all the way to fully register-based censuses. It is important that NSOs understand the strengths and limitations of administrative data for use in their censuses, to ensure that they make fully-informed decisions about the use of such data.

Drawing on quality frameworks and best practices adopted by NSOs across the world, these Guidelines lead census practitioners through the practical stages of assessment; from working with an administrative data supplier to understand the strengths and limitations of a source, all the way to the receipt and analysis of the actual data. The Guidelines cover key quality dimensions on which an assessment is made, using a variety of tools and indicators. The Guidelines are based on four Stages: Source, Data, Process, and Output, with the first two Stages being the principal focus of the Guidelines.

The Guidelines are the result of work undertaken by a Task Force of experts from 20 countries and organizations, reflecting the wide diversity of census methodologies and uses of administrative data in the UNECE region. The experiences of many of these countries are included throughout the Guidelines as illustrative examples and case studies, showing how the tools presented can be put into practice.

The Guidelines were endorsed by the 69th plenary session of the Conference of European Statisticians in 2021.

Information Service
United Nations Economic Commission for Europe

Palais des Nations
CH - 1211 Geneva 10, Switzerland
Telephone: +41(0)22 917 12 34
E-mail: unece_info@un.org
Website: http://www.unece.org

ISBN 978-92-1-117276-8

Printed at United Nations, Geneva – 2114186 **(E)** – October 2021 – 642 – **ECE/CES/STAT/2021/4**