



## Tapping into administrative data in census-taking: an emerging trend in Asia and the Pacific

An increasing number of countries in Asia and the Pacific are exploring the use of administrative data sources as an addition or alternative to full field enumeration in censuses. While some are aiming to eliminate field enumeration altogether, others are using administrative data to improve the quality and efficiency of a traditional census approach. This Stats Brief overviews how some of the National Statistical Offices in Asia and the Pacific are integrating administrative data into their censuses.

This Brief is part of the ESCAP's work on population registers and censuses under the Bloomberg Philanthropies Data for Health Initiative (D4H) Project: "Supporting improvements to civil registration and vital statistics (CRVS) systems in Asia and the Pacific (Phase III)".

*This Stats Brief was prepared by Stein Terje Vikan, Consultant, ESCAP Statistics Division, with valuable inputs from Petra Nahmias, Chief of Population and Social Statistics Section, ESCAP Statistics Division and Afsaneh Yazdani, Statistician, ESCAP Statistics Division. The author prepared this Stats Brief based on the report on "Emerging trends in census approaches in Asia and the Pacific".<sup>1</sup> The author extends his gratitude to Seiffe Tadesse, Meryem Demirci and Srdjan Mrkic, United Nations Statistics Division, Alison Culpin and Michael Sharp, the Pacific Community, Gloria Mathenge, the Economic Commission for Africa, as well as National Statistical Offices of Armenia (Karine Kuyumjyan), Australia (Ross Watmuff, Gemma Van Halderen, and Georgia Chapman), Indonesia (Ali Said, Nashrul Wajdi and Diah Ikawati), Mongolia (Batchuluun Dorj), Republic of Korea (Kang Jeong-hee, and Byungsik Lee), Thailand (Hataichanok Puckcharern), and Turkey (Mehmet Doğu Karakaya, Metin Aytaç, Mehmet Şaban Ucari, Alper Acar, Önder Değirmenci, and Osman Orhanli).*

### Introduction

Population and housing censuses are the primary source of data needed for formulating, implementing and monitoring a range of policies and programmes within a country and comparing between countries.<sup>2</sup> The main aim of a population and housing census is to fully enumerate a country's population and housing conditions to provide detailed demographic and socio-economic data at highly disaggregated geographical and substantive levels. For a long time, the only approach to providing this information was to dispatch a large number of enumerators with a questionnaire to collect information on each individual from each dwelling in the country. This approach is referred to as a 'traditional census'.

However, several trends have put pressure on countries to modernise their censuses and look for alternative approaches to the traditional census. Firstly, hiring

such a large number of temporary staff is challenging and the operation itself is very costly. Further, reduced willingness amongst people to respond to census questionnaires has led to challenges of increasing non-response. Since the purpose of a census is to count everyone, increasing non-response is cause for concern. And lastly, there is a push to process census data faster in order to release timelier results thus increasing their relevance for policymaking.

During the 2020 census round, another strain on censuses emerged: the challenge of conducting vast field-based data collection exercises during a highly infectious coronavirus disease (COVID-19) pandemic. The COVID-19 pandemic caused delays and rescheduling of censuses across the globe, in addition to an urgent rethinking of data collection approaches in many countries.

<sup>1</sup> See <https://unescap.org/kp/2022/emerging-trends-census-approaches-asia-and-pacific-country-examples>.

<sup>2</sup> See <https://unstats.un.org/unsd/demographic-social/census>.

Broadly speaking, there were two types of responses to the challenges posed by traditional censuses in recent decades. The first was the application of new technologies to make traditional censuses more effective. This included the introduction of different types of electronic data collection and the use of the internet for online questionnaires. Electronic data collection allowed for quicker capturing and processing of data, whereas online questionnaires reduced the need for hiring enumerators. The use of the internet for data collection during the COVID-19 pandemic also reduced the need for face-to-face interviews.

The second type of response was to explore the use of administrative data and registers (Box 1) as an alternative to canvassing the entire country through field enumeration. This development started in the Scandinavian countries, with Denmark as the first country to conduct a complete population and housing census based on administrative data in 1980.<sup>3</sup> Several other countries have followed, not only in Europe but also in other parts of the world.

#### Box 1: Definitions of registers, administrative data sources and statistical registers<sup>4</sup>

- A *register* is defined as a systematic collection of unit-level data organized in such a way that updating is possible.
- *Administrative data sources* are data holdings that contain information collected primarily for administrative purposes. This type of data is collected by government departments and other organizations for the purposes of registration, transaction and record-keeping, usually during the delivery of a service.
- *Statistical registers* are registers created for statistical purposes. They are typically created by transforming data from registers and/or other administrative data sources.

### Register-based versus combined census

The Scandinavian countries developed fully *register-based censuses*, meaning that all information at the individual level is collected from registers or administrative data sources, with no field enumeration. This requires that all of the variables for a full census

be available at the unit level in up-to-date registers or administrative sources in the country. Box 2 provides an overview of the most commonly used registers in a census, which are referred to as “base registers”.

#### Box 2: Base registers commonly used in a register-based census<sup>5</sup>

- *Population register* – A register of all residents in the country.
- *Address/Dwelling register* – A register of all (whether occupied or not) addresses or dwellings in the country.
- *Business register* – A register of all businesses or legal entities in the country.
- *Activity register* – A register of activities that residents are engaged in, such as (un)employment/work, benefits and pensions or studies.

However, not all countries that want to move away from a traditional census have access to all the necessary data from administrative sources. Many countries have therefore applied a combination of methods to use register data for some parts of the census and enumerated data for others. This is referred to as a *combined census*.

There are various types of combined censuses, with two main types commonly recognized. The first is register data combined with full enumeration, and the second is register data combined with one or more sample surveys. The second option significantly reduces the need for field-based data collection, but it requires that some core census variables be produced solely from registers. The first option still needs a large-scale field-based data collection and is often used as a step towards greater reliance on register data.

In Asia and the Pacific, we see another type of combined census, where traditional data collection (albeit in a modernized way, for instance, through the use of online questionnaires) is the default option for unit-level data, but administrative data is used in a variety of ways to support the census. We can refer to these as full enumeration supported by register data. These have a lot in common with censuses that combine register data with full enumeration, but often

<sup>3</sup> See [https://unece.org/fileadmin/DAM/stats/publications/Register\\_based\\_statistics\\_in\\_Nordic\\_countries.pdf](https://unece.org/fileadmin/DAM/stats/publications/Register_based_statistics_in_Nordic_countries.pdf).

<sup>4</sup> See <https://unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf>.

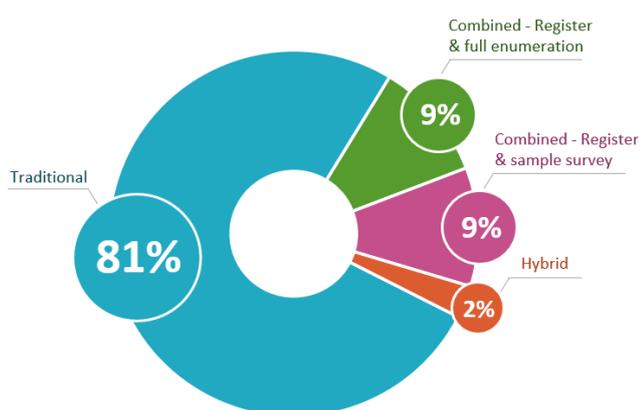
<sup>5</sup> Ibid.

the immediate goal is not to move to a register-based census. This can be either because a country lacks sufficient register infrastructure to make the full transition but still wishes to use registers to improve its census operations, or that the country has a strong register infrastructure and wants to take advantage of it but does not intend to move away from enumeration as the cornerstone of its census.

## Use of administrative data in censuses in Asia and the Pacific

Recent research on census-taking in Asia and the Pacific<sup>6</sup> indicates that 47 of the 58 ESCAP member States and associate members have used or are planning to use a traditional approach in their 2020 census round. This represents 81 per cent of all members. A total of ten countries, representing 17 per cent of all members, are planning to or have carried out a combined census, with five countries combining register data with sample survey data and five combining with full enumeration (see figure below). In addition, one country, Afghanistan, has carried out a hybrid census, which is a combination of a sample survey with satellite images. This was due to long-standing security and accessibility issues in the country.

### Types of censuses in Asia and the Pacific, 2020 round<sup>7</sup>



Note: Percentages do not add up to 100 due to rounding.

To combine register data with sample surveys, it needs to be possible to establish a statistical population register based on either a population register or from a combination of administrative data sources. The Republic of Korea, Singapore and Turkey are good

examples of this category, as they base a key set of census indicators on register data, and then complement this with a large-scale survey for the remainder of the indicators. This reduces both the cost of the operation and the response burden. For countries that are either not yet able to do so or choose not to, there is this option to combine the register data with full enumeration.

In the next pages we will present four country case studies, showcasing how some countries in Asia and the Pacific have approached exploiting administrative data in their census. These cases illustrate some of the different ways that administrative data can be used to improve the quality and efficiency of census-taking.

### Australia – Using administrative data to support census operation

Up until 2011, the Australian Bureau of Statistics (ABS) primarily carried out traditional censuses. However, in the 2011 Census, ABS faced severe challenges with data collection. Less than half of the doors were answered when field enumerators attempted to visit the households<sup>8</sup>. As a consequence, enumerators had to visit a large number of households multiple times. This, in turn, led to higher costs and still did not prevent dwindling response rates. To counter this, ABS started exploring the use of administrative data to improve the efficiency, cost-effectiveness and quality of the census operation itself.

Thus, ABS introduced three improvements to its census strategy in 2016. Firstly, a national Address Register was developed to support mailing out of materials to households across Australia. This register was formed using available administrative data in conjunction with a large-scale canvassing exercise. Secondly, an instruction letter detailing how to complete the census online or how to request a paper form was mailed out. Thirdly, a smarter online form was introduced. Many enhancements were made to the online form to improve quality and make it easier for respondents to complete.

<sup>6</sup> See <https://unescap.org/kp/2022/emerging-trends-census-approaches-asia-and-pacific-country-examples>.

<sup>7</sup> Ibid.

<sup>8</sup> See <https://abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2900.0~2016~Main%20Features~How%20we%20collected%20your%20information~2>.

## Expansion of the use of administrative data for the 2021 Census

This strategy was further developed for the 2021 Census, with three new innovations. Firstly, ABS uses administrative data to *help communities during the census*. They use various administrative sources such as health, social security, migration and taxation, often in combination with previous census data, to better identify and understand communities that in the past have had lower than average response rates. Once these communities have been identified, strategies for assisting them, for instance, through information material, can be devised.

Secondly, similar administrative data sources can be used to *help improve the census count* by determining whether a dwelling was occupied on census night or not. Australia conducts a *de facto* census, meaning that they count where people actually are on census night. For dwellings where no information was obtained through regular enumeration, then administrative data can be utilized to determine whether anyone from the address in question had accessed any government services recently, indicating it the likelihood that the house is occupied. This exercise is done with the strictest data privacy procedures.

Thirdly, administrative data is being used to *be prepared for unexpected events*. Sometimes unexpected events such as bush fires make data collection impossible in certain areas of the country. Administrative data can then be utilized to fill in some of the gaps which this creates.

### Comparing census data with administrative data

In addition, ABS conducts research or quality control projects in which the census results are compared to administrative data. Such a comparison was made for the 2016 Census, and a similar exercise is planned for the 2021 Census. This research has documented that there is very little difference between administrative data and census data, providing reassurance that administrative data sources are fit for purpose to support the census.

Even though enumerated census data is the preferred choice, administrative data play an important role in the

Australian census. The various ways ABS uses administrative data to support the census may serve as a very useful illustration for countries considering leveraging their administrative data for census work.

## Indonesia – Using administrative data for pre-listing households

Up until 2010, the census approach in Indonesia was traditional. However, for the 2020 Census, the potential use of population data from its Population Registration System to modernize its census approach was explored.

A key challenge for Statistics Indonesia was to secure access to administrative data for use in statistics. Thus, to facilitate such access, several legislative instruments were either introduced or amended, with a focus on providing Statistics Indonesia with the necessary access to data from the Population Registration System.

### Civil Registration data as input in the pre-listing phase

Civil registration is an integral part of the Indonesian Government's poverty-reduction policy, with a focus on providing identity documents to all citizens.<sup>9</sup> However, there are still many challenges when it comes to the registration of vital events in Indonesia.<sup>10</sup> The proportions of events such as births, deaths, marriages and divorces that gets registered are still relatively low. Reasons for this are mostly structural barriers including prohibitive costs, travel distance, and lack of knowledge of the registration process.

Even if the population data from the Population Registration System was not of sufficient quality to be used for creating census variables, the data was considered to be useful for the census operation, particularly for the pre-listing phase. The administrative data was combined with previous census data to create a statistical population register. Then an online census was carried out, where the questionnaire was populated with the administrative data. Respondents were asked to verify or correct the pre-filled information or fill it out if it was missing.

<sup>9</sup> See <https://getinthepicture.org/country/indonesia>

<sup>10</sup> See <https://getinthepicture.org/resource/civil-registration-and-vital-statistics-crvs-indonesia>.

In subsequent stages, the results of the online census were verified by local leaders, and a field enumeration was carried out to fill the gaps. A post-enumeration survey was also conducted as part of the quality control of the administrative data. More detailed information such as fertility, mortality and migration were collected through a sample survey.

### *Towards a combined approach*

Indonesia has been able to take advantage of the fairly new existence of administrative population data to reduce costs and improve the effectiveness of the census operation. While the administrative data was not considered of sufficiently high quality to be able to provide census variables directly, it was used to improve and speed up the household listing. In combination with online census taking, it significantly reduced the need for field enumeration. Based on the experiences from 2020, Indonesia is well placed to produce some census variables from administrative data for the 2030 census cycle.

---

## **The Republic of Korea – Introducing a combined census approach**

---

Up until 2000, the Republic of Korea conducted traditional censuses, with full enumeration of a short form and a sample survey to collect more detailed information in a long-form. After the census in 2000, the Government of the Republic of Korea started raising issues around the census methodology. Motivating factors for exploring this shift were the quality, the cost-efficiency and the timeliness of the census.

A starting point for Statistics Korea (KOSTAT) was that the Republic of Korea had several well-established administrative data sources, including the Resident Registration Register (RRR). In 2010, KOSTAT chose to conduct another traditional census and then use the result of this census to carry out comparisons to data from administrative sources. This work, in turn, prepared them for moving to a combined census in 2015.

### *Combining data at the unit level from multiple sources*

In addition to the RRR, the Republic of Korea had several administrative data sources that contained information relevant to a census. In the end, KOSTAT established a statistical register based on information from 25 administrative registers from 14 public institutions. This data was combined to create a microdata set for a register-based census. The merging of data at the unit level from several sources was possible as all individuals have a unique Personal Identification Number. To ensure data privacy, KOSTAT created a virtual identification number for all individuals, which was used as the matching key.

### *Comparative study of administrative data and census data*

To assess the readiness of the administrative data sources to replace full enumeration for the 2015 Census, a couple of studies were carried out after the completion of the 2010 Census. The basis for these studies was comparing an administrative population register based on the RRR with the actual 2010 enumerated Census Population. The results were used to identify some key sources of discrepancies.

### *The 2020 Census and beyond*

The Republic of Korea has implemented a transition towards a combined census where a core set of 13 census items is now produced from register data on an annual basis, resulting in reliable and timely population statistics for planning purposes. In 2020, KOSTAT also carried out a sample survey, covering 20 per cent of the population to gather data for more detailed and additional items. They also included general items that are produced based on register data to enable quality control and identify ways to improve the register-based census data.

The approach taken by the Republic of Korea is a good example for countries with well-established administrative data sources that are looking to move towards a combined census.

## Turkey – A planned transition towards a register-based census

The last traditional census in Turkey was carried out in 2000. Key motivating factors to explore the use of administrative data sources were the logistical challenges as well as the cost of carrying out a traditional census in such a large country.

In the early 2000s, the Turkish Ministry of Interior established the Central Civil Registration System. During this project all citizens were given a unique 11-digit identification number. Together with a newly established National Address Database (NAD) and a once-off full field enumeration to establish each individual's usual place of residence, this formed the basis for the Address Based Population Registration System (ABPRS). This population register was instrumental in enabling the Turkish Statistical Institute (TurkStat) to move away from full census enumeration.

### *The 2011 Census – Introducing the combined approach*

The 2011 Census was the first opportunity to use the newly established ABPRS for census operations. The system was able to produce approximately 20 per cent of the census variables directly from the register, the remaining 80 per cent had to be collected using a large-scale sample survey. Full enumeration was thus not necessary anymore.

In addition to using register data from ABPRS to produce census variables, register data from NAD was also used to create the household listing and the enumeration areas for the field operations.

### *Expanding the use of register data for the 2021 Census*

In the years prior to the 2021 Census, several additional registers were either established or further developed in Turkey. To facilitate the use of these additional administrative sources for statistical purposes, TurkStat is developing the Population Characteristics Database. This database brings together data from several administrative sources which enables TurkStat to prepare an increasing number of census variables directly from registers. Key census information that can be produced from this database includes age, sex,

nationality, disability, education, employment status, property, migration and family structure. The National identification number enables the combining of all data sources. TurkStat aims to produce approximately 70 per cent of the census variables for the 2021 Census from register data by integrating all these supplementary registers in the Population Characteristics Database with the ABPRS. For the remaining 30 per cent of census variables on building and dwellings characteristics, TurkStat has carried out an ad-hoc survey using Computer-Assisted Telephone Interviewing (CATI).

Prior to integration into the Population Characteristics Database, the quality of all administrative sources has been thoroughly assessed. The insights gained from quality checks will be fed back to the register owners to improve their data quality in the future. TurkStat thus plays an important and active role in improving and sometimes establishing new registers in Turkey.

### *Register dominant census*

Turkey is an example of a country that is in the midst of a planned transition from a traditional census to a fully register-based census. Turkey still implements what TurkStat refers to as a “register dominant” combined census in 2021 but is likely to move towards a full register-based census in the next round. It is important to note that this is only possible due to the existence of several well-functioning registers within the Turkish government system.

## Growing trend – Exploring the use of administrative data

Australia, Indonesia, the Republic of Korea and Turkey are examples of countries that have, in various ways, introduced the use of administrative data to improve their censuses. Whereas Turkey and the Republic of Korea are producing significant parts of their census from administrative data alone, Australia is using administrative data in innovative ways to improve an essentially traditional census approach. Indonesia has explored the possibility of using its newly established Population Registration System to simplify the household listing phase, but without being able to produce census variables from it yet.

As mentioned earlier, ten countries in the region have moved away from the traditional census approach and implemented a combined census for the 2020 round. Several of these countries are at similar stages as the examples presented above.

For its upcoming census in 2022, the Statistical Committee of the Republic of Armenia (Armstat) is planning to use administrative data for a small set of variables available in the State Population Register (SPR), whereas a supplementary household sample survey will be conducted by census enumerators using tablets. The sample survey will collect all census variables and be used to verify and improve the SPR, in addition to being the primary source for many of the census variables.

Mongolia is another country combining a sample survey with register data in their 2020 Census. In 2014, the National Statistics Office of Mongolia set up the National Population and Household Registration Database (PHRD) as a repository for information stemming from various government administrative systems. The PHRD is a statistical database that is only used for statistical purposes. A sample survey was then administered to 10 per cent of the households, selected from the PHRD to collect more detailed census information.

New Zealand is a country that is using administrative data in its census without moving away from the full enumeration. In 2012 Statistics New Zealand (Stats NZ) launched a new programme called the Census Transformation Programme, first implemented for its 2018 Census. Administrative data was used to provide information on individuals not reached through traditional enumeration, with approximately 89 per cent of the population counted through census responses and 11 per cent through administrative data. Stats NZ is continuing its census transformation as they plan for the next census in 2023.

The National Statistical Office of Thailand introduced several changes to its census for the 2020 round, including an online census. In addition, they conducted two pilots to assess the feasibility of using administrative data for its census. A key finding from these pilots is that the administrative data is not yet of sufficient quality to be used in census work, but this will be explored in the future.

### Box 3: The Pacific Island Countries and Territories – Increasing interest in administrative data

The Pacific Island Countries and Territories (excluding New Zealand and Australia) still employ a traditional approach to their censuses, although there is growing interest in applying new technologies and approaches. For instance, both Cook Island and Samoa are having discussions around establishing civil registration systems that may be used in future census work.

## Concluding remarks

Over the recent decades, there have been, increasing challenges to the traditional census across the world. Cost, non-response and timeliness issues put pressure on National Statistical Offices to make improvements. Moving towards increased use of administrative data is a key strategy for many countries in meeting these challenges.

This global trend is also present in Asia and the Pacific. As shown in this Stats Brief, countries are using a variety of strategies on integrating administrative data into their censuses, depending on availability of quality data, amongst other factors. In this regard, it should also be noted that replacing enumeration with administrative data without proper quality control procedures may significantly reduce the quality of the census outputs. Some caution is warranted when it comes to implementing such a transition too quickly.

Given that administrative data is not initially generated for the purpose of a census, National Statistical Offices need to be in close collaboration with administrative authorities and may want to be closely engaged in initiatives that support developing and improving administrative data in the country. While civil registration and vital statistics systems are especially important for Population and Housing Censuses, better administrative data on addresses, dwellings/buildings, health, education, taxation and social security, to name a few, would also be extremely beneficial.

For more information regarding ESCAP's work in statistics development please visit: [www.unescap.org/our-work/statistics](http://www.unescap.org/our-work/statistics)

Previous issues of Stats Brief: [www.unescap.org/resource-series/stats-brief](http://www.unescap.org/resource-series/stats-brief)

Contact us: [stat.unescap@un.org](mailto:stat.unescap@un.org)