International Statistical Agencies: What can we learn from other countries about how they are using administrative data to supplement, enhance, or create new statistical products?

Kate Lanman, https://orcid.org/0009-0007-1109-868X
Olivia Davis, https://orcid.org/ 0009-0005-3403-7009
Stephanie Shipp, https://orcid.org/0000-0002-2142-2136

Contact: sss5sc@virginia.edu

Social and Decision Analytics Division
Biocomplexity Institute & Initiative
University of Virginia

Contact: sss5sc@virginia.edu

Social and Decision Analytics Division
Biocomplexity Institute & Initiative
University of Virginia

Funding: This research was sponsored by the U.S. Census Bureau Agreement No. 01-21-MOU-06 and Alfred P. Sloan Foundation grant, G-2022-19536.

Acknowledgments: We would like to thank Deirdre Dalpiaz Bishop, Chief of the Geography Division at the U.S. Census Bureau, and Deborah Stempowski, Assistant Director for the Decennial Program at the U.S. Census Bureau, for introducing us to the staff at the Australian Bureau of Statistics, Office of National Statistics (England and Wales), Northern Ireland Statistics and Research Agency, and Statistics Canada. We would also like to thank the representatives at these agencies who spoke to us.

Lanman K, Davis O, Stephanie S. International Statistical Agencies: What can we learn from other countries about how they are using administrative data to supplement, enhance, or create new data products? Technical Report BI-2023. Proceedings of the Biocomplexity Institute, University of Virginia; 2023 October. https://doi.org/10.18130/2n54-sc22.

.

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

Abstract

The U.S. Census Bureau is transforming and modernizing its use of data across surveys and integrating it with administrative data to enhance the decennial census and current surveys and to create new statistical products. We examined statistical agencies in other countries to learn how they are modernizing their operations to take advantage of administrative and other data sources, such as private-sector data, to supplement, enhance, or create new data products. To do this, we summarized presentations by international statistical agencies in Australia, Canada, the United Kingdom, and New Zealand. In parallel, we interviewed representatives from a similar set of statistical agencies in Australia, Canada, the United Kingdom, and Northern Ireland.

International Statistical Agencies: What can we learn from the modernization of Statistical Agencies in other countries?

The Census Bureau is transforming and modernizing its use of data across surveys and integrating it with administrative data to enhance current surveys and create new statistical products. We examined statistical agencies in other countries to learn about how they are modernizing their operations to take advantage of administrative and other sources of data, such as private-sector data, to supplement, enhance, or create new data products. To do this, we summarized presentations by international statistical agencies in Australia, Canada, the United Kingdom, and New Zealand. In parallel, we interviewed representatives from a similar set of statistical agencies in Australia, Canada, the United Kingdom, and Northern Ireland.

We present findings from four webinars held by the Committee on National Statistics for the US Census Bureau. Statistical agency officials from Canada, the United Kingdom, New Zealand, and Australia. Each presenter discussed modernizing their population census data collection and quality using administrative data. Appendix 1 is a table that summarizes information from these webinars. We also interviewed four statistical agencies in Australia, Canada, the United Kingdom (England and Wales), and Northern Ireland. A summary of our findings and the list of representatives interviewed is presented in Appendix 2.

Collecting Population Censuses in Four Countries

Intro to Census Models

Based on data collection methods, we can classify census models into four categories: traditional, register-based, combined, and rolling. These models are defined by the United Nations Economic Commission for Europe (UNECE) as follows:

Traditional census (entire field enumeration, whether in person or online; registers and administrative sources may be used to support the enumeration but not directly to obtain census data)

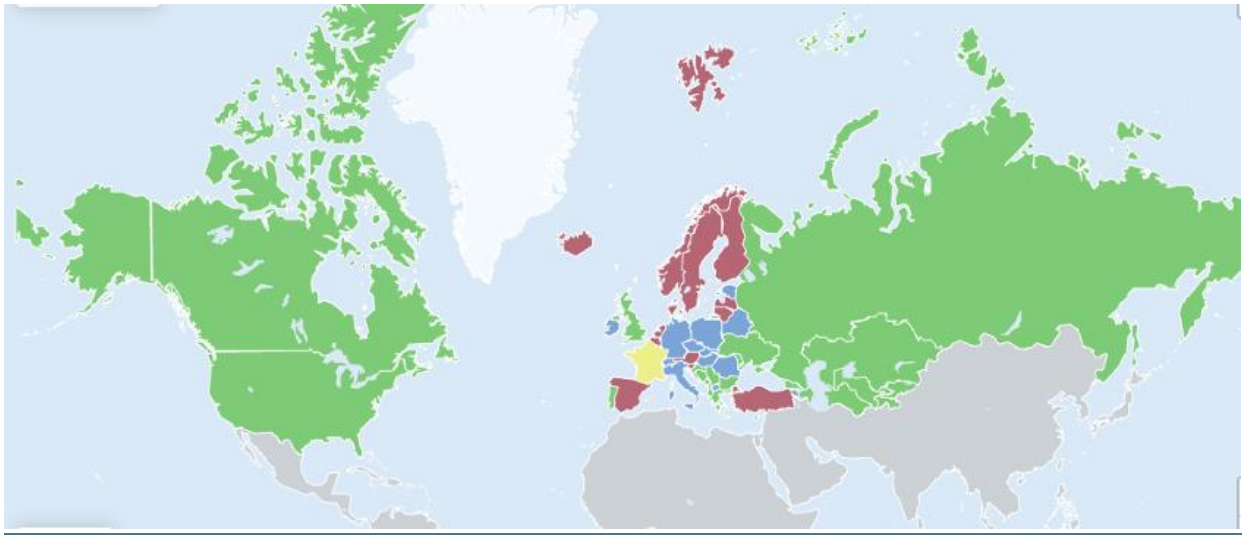
Register-based census (census based on data from registers and administrative sources, with no field data collection; may also include data from existing surveys not conducted for census purposes)

Combined census (some data obtained directly from registers or administrative sources, while other data are collected through field data collection conducted specifically for census purposes, covering the whole population or only a sample)

Rolling census (based on cumulative, continuous sample survey covering the entire country over an extended period).

The below figure shows how each UNECE country carries out its census. This section focuses on **Canada, the United Kingdom, New Zealand (not pictured), and Australia (not pictured).** Canada, the United Kingdom, and Australia perform traditional censuses, whereas New Zealand uses a combined model. As administrative data becomes increasingly available, all four countries are expanding how these data are used in their censuses. Canada and Australia are looking to move to a combined census approach for the coming censuses. The table below highlights in a little more detail how each census is collected in these countries. In the remainder of this paper, we examine the censuses of these four countries and how they are evolving to meet modern needs.

https://statswiki.unece.org/display/censuses/Censuses+of+the+2020+round



	Australia	Canada	New Zealand	UK	Northern Ireland
Frequency	Every 5 years	Every 5 years	Every 5 years	Every 10 years	Every 10 years
Most recent census year	2021	2021	2023 (results not finalized)	2021	2021
Current Model	Traditional	Traditional	Combined (as of 2023)	Traditional	Combined
Online Response	Yes	Yes	Yes	Yes	Yes

Note: Northern Ireland is included in the table since we interviewed a representative from its statistical agency. They were not part of the presentation series.

Modern Problems Require Modern Solutions

Censuses worldwide face common challenges, such as increasing nonresponse or the inability to count certain populations accurately. Many countries are turning to administrative data to supplement surveys to meet these challenges. These alternatives are becoming increasingly available

and expanding in scope. This section focuses on critical issues facing today's censuses and how Australia, Canada, New Zealand, and the United Kingdom use administrative data to address them.

Survey Response Quality

As administrative data becomes more extensive, it can be used to moderate data quality issues more effectively. Concerns with overall non-responses, item non-responses, and inaccurate or inconsistent responses are common with census surveys. The chance for respondent error has increased, especially with the increasing use of online surveys.

All four countries have begun to use or put plans in place to use administrative data to control for non-responses. New Zealand has struggled the most with overall non-response rates of the four countries. In 2018, unexpectedly low response rates pushed New Zealand to use a combined census. In 2023, their census was combined by design with plans to use administrative data post-collection to include people missed in field collection and determine missing characteristics. Administrative data will also be used as a quality check where census results can be compared to related items in this data. While the results from 2023 are not finalized, New Zealand anticipates that these contingencies will be necessary due to the impact of Cyclone Gabrielle.

Australia and Canada had similar low-response contingency plans as New Zealand for their most recent censuses, but neither needed to use them widely. However, Canada has used administrative data to completely replace income and some immigration questions on their short-form surveys and halt the delivery of surveys in areas affected by natural disasters by using administrative data population estimates for those areas. Canada also uses administrative data for various post-collection coverage studies, which help ensure data quality by correcting for misclassification of dwellings and estimating under and over-coverage of the population.

As with Australia and Canada, the United Kingdom has yet to face widespread non-response to their surveys. However, while the United Kingdom similarly focuses on moderating item non-responses through imputation, much of this work uses linked responses rather than administrative data. Compared to other countries, the United Kingdom's census asks far more about characteristics,

making it harder for administrative data to supplement this information. On the other hand, much like Canada and New Zealand, the United Kingdom uses administrative data to detect "item inconsistencies" or potential data quality issues.

Demands for More Timely Data

One of the key benefits of using administrative data is that it can produce much more frequent results. Australia, New Zealand, and the United Kingdom are all working on producing new administrative data tools to provide more timely outputs. The United Kingdom already produces administrative data-based midyear population estimates. These estimates are released in June annually, and they have been producing them for over a decade. The United Kingdom is also in the process of producing a Population and Migration Statistics System. This tool will use administrative data to provide annual information on migration, population subgroups and characteristics, housing and living arrangements, and life journeys and outcomes. New Zealand has produced a similar tool called the Administrative Populations Census (APC), which is in its experimental phase. The APC additionally notes indigenous descent.

Australia is also producing an administrative data snapshot tool, set to be released by the end of 2023. This tool will provide population and housing snapshot tables alongside a table builder tool to allow for more flexible outputs. Eventually, they hope to deliver a microdata product as well. Unlike the United Kingdom's statistic system and New Zealand's APC, this tool will be updated on the same cycle as the Australian census (every five years) despite its capability to produce annual outputs. Once the tool is published, Australia plans to gather public feedback to determine whether they should push for more frequent data releases.

Operational Efficiency

Administrative data is frequently used in less direct ways to achieve operational efficiencies. All four countries noted using administrative data to optimize data collection methods. Australia, New Zealand, and the United Kingdom all use administrative data to assist in determining

where to send surveys (i.e., what an actual dwelling is or which dwellings are occupied). Canada plans to implement this practice in its upcoming 2026 census.

Limitations of Administrative Data

Of course, using administrative data presents many challenges. This section will note how Australia, Canada, New Zealand, and the United Kingdom are facing limitations in the coverage and confidentiality of administrative data.

Coverage Issues

Administrative data coverage is not equal across all areas and demographics. This is especially an issue with indigenous populations in Australia, Canada, and New Zealand. These populations are often missed in administrative data and underrepresented in census surveys. To account for this, these countries are focusing on greater community outreach toward them. For example, in their most recent census, New Zealand increased the size of their community engagement teams and partnered with local indigenous communities to lead survey collection.

Likewise, migrant groups are under-covered by administrative data in all four countries. This is because it is difficult to capture events that occur overseas. Currently, New Zealand works with Australia to gather data on New Zealanders living in Australia, but this only moderates the issue.

Over-coverage problems with administrative data also occur often because there can be multiple sources for the same information. With data from outside organizations with different standards, sorting through and matching records can be challenging. Federal statistical organizations also have less control over these sources, making quality monitoring harder. However, Australia and the United Kingdom note that there are not typically issues pushing for technical change in their collaborators' data since most people want to address quality and coverage concerns.

Lastly, acquiring administrative data can be problematic. These data are often costly and not necessarily readily available because of how different parties store their data. Historical data, for example, can be difficult to access because it may not be digitized.

The concerns noted above lead to issues in population estimates based on administrative data.

As mentioned earlier, the United Kingdom has already begun to produce and publish annual administrative population estimates for the past decade. Despite administrative data coverage issues, the United Kingdom's administrative population estimates are close to the official estimates and are broken down by age and sex. Australia, Canada, and New Zealand have begun producing similar administrative population estimates. Australia's and New Zealand's administrative estimates slightly undercount the population, whereas Canada produces a slight overcount. However, in all three countries, the administrative estimate counts closer to the official population estimate than traditional census responses.

Threats to Confidentiality

The use of administrative data also brings about new data privacy concerns. While the timeliness of administrative data is seen as one of the main benefits of its use, this frequency makes it more difficult to ensure data remain confidential. The United Kingdom currently uses targeted method swapping on census data as one of its disclosure control methods. However, this method is much harder to implement with administrative data because the frequency and quantity of data make it easier to reverse engineer the method. Currently, the United Kingdom is looking into alternative disclosure controls that can be used to replace this method.

Summary

Administrative data use also brings with it questions of social acceptability. All four countries currently have legislation that allows their federal statistical agencies access to administrative data without express permission from the general population. The laws for each country are noted below.

- Australia Census and Statistics Act: https://www.legislation.gov.au/Details/C2020C00296
- Canada Statistics Act: https://laws-lois.justice.gc.ca/eng/acts/S-19/FullText.html
- New Zealand Data and Statistics Act: https://www.legislation.govt.nz/act/public/2022/0039/latest/whole.html
- United Kingdom Digital Economy Act: https://bills.parliament.uk/bills/1859

There is a significant focus on building public trust that statistical agencies and researchers will use administrative data ethically. All four countries are working to implement formal testing to gauge public reaction to administrative data usage and emphasize transparency throughout this process.

References

Admin-based population estimates: provisional estimates for local authorities in England and Wales, 2011 to 2022. Office for National Statistics, 2023, https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/adminbasedpopulationestimates/provisionalestimatesforlocalauthoritiesinenglandandwales2011to2022#:~:text=The%20DPM%20uses%20this%20information,on%20the%20June%202021%20estimate. Accessed August 2023.

Census and Statistics Act. Department of the Treasury of Australia, 1905. https://www.legislation.gov.au/Details/C2020C00296. Accessed August 2023.

Censuses of the 2020 Round. United Nations Economic Commission for Europe, 2020, https://statswiki.unece.org/display/censuses/Censuses+of+the+2020+round. Accessed August 2023.

Data and Statistics Act. Statistics New Zealand, 2022. https://www.legislation.govt.nz/act/public/2022/0039/latest/LMS418574.html. Accessed August 2023.

Digital Economy Act, United Kingdom House of Commons. 2017, c.30, https://bills.parliament.uk/bills/1859. Accessed August 2023.

Item editing and imputation process for Census 2021, England and Wales. Office for National Statistics, 2022, https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/itemeditingandimputationprocessforcensus2021englandandwales#:~:text=Historically%2C%20donor%2Dbased%20methods%20have,properties%20of%20the%20data%20accurately. Accessed August 2023.

Statistics Act. Royal Society of Canada. 1972, c. 15, s. 1, https://laws-lois.justice.gc.ca/eng/acts/S-19/FullText.html. Accessed August 2023.

The appendices provide more information.

- Appendix 1 provides a table that summarizes the four CNSTAT webinars.
- Appendix 2 summarizes the interviews with representatives from four statistical agencies.

Appendix 1: CNSTAT Public Webinar Series: : International Approaches to Conducting a Census in the 21st Century.

Webinar abstract: The challenges facing the U.S. federal statistical system are also being faced by other developed countries in the world, including declining cooperation with surveys and censuses, increasing availability of administrative and other alternative data to surveys, greater demands for more granular and timely data, and growing threats to the confidentiality of new and existing data products from censuses, surveys, and other sources. This series of seminars explored how several other prominent National Statistical Offices are addressing these common challenges and describe their plans for using administrative data in conducting a periodic census of their populations. Representatives from the United Kingdom, Canada, Australia, and New Zealand statistical offices gave presentations and answered questions.

UK - https://www.nationalacademies.org/event/05-18-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-uk

Canada - https://www.nationalacademies.org/event/05-24-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-canada

New Zealand - https://www.nationalacademies.org/event/06-07-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-new-zealand

Australia - https://www.nationalacademies.org/event/05-31-2023/cnstat-public-webinar-series-international-approaches-to-conducting-a-census-in-the-21st-century-australia

Appendix 2 Summary of CNSTAT webinars				
	Canada	**New Zealand**	**Australia**	**UK**
Frequency	**Every 5 years**	**Every 5 years**	**Every 5 years**	**Every 10 years**
Most recent census year	**2021**	**2023**	**2021**	**2021**
Current Approach	**Traditional**	**Combined**	**Traditional**	**Traditional**
Current Uses of Admin Data	● Coverage studies: ○ Dwelling classification survey: estimates and corrects for misclassification of dwellings on census day. ○ Undercoverage study: estimates population missed by the Census Bureau. ○ Overcoverage study: estimates population counted more than once in the census. ● Direct use: ○ Enumerate income and some immigration variables. ■ Income is determined through tax records. ■ Participants were previously notified of the usage of admin data to	● Framework for the assessment of the quality of admin data: ○ Relevance ○ Accuracy ■ Representation: coverage, unit error ■ Measurement: validity error, measurement error ○ Timeliness ○ Accessibility ○ Coherence and consistency ○ Interpretability ● Direct use: ○ Include people missed in field collection and determine missing characteristics. ■ Admin data was used in this way in 2018 to make up for unexpectedly low response rates (combined model	● Indirect use (provided support but not used for imputation/supplementation): ○ Planning and data collection - help maximize census responses. ■ Finding unoccupied dwellings (cost efficiency) ○ Data processing - improve census count. ○ Preparing low response contingency - repair of data for unexpected response impacts ● Admin data sets: ○ Admin population data set ■ Age, sex, indigenous status, … ○ Admin dwelling asset ■ Dwelling info, electricity usage, rental vacancies, … ● Admin population estimate. ○ Gets very close to official estimate. ○ Counts more accurately	● UK has been using admin data for their midyear population estimates for decades. ● Indirect use (provided support but not used for imputation/supplementation): ○ Quality assurance in processing ○ Determine addresses to send surveys to (occupied dwellings) ● Dynamic population model ○ Estimates demographic account by combining data sources. ○ Produces admin based population estimates. ■ Clear need for robust coverage adjustment ○ Captures dynamics of population change through cohort component method of demographic accounting (i.e., adding births & immigrants, taking away deaths & emigrants) ○ Inputs:

	determine income, but it was decided to be no longer necessary to state explicitly. ○ Replace field responses when emergencies impact traditional collection (ex: wildfires in northern region) ○ Impute non-responses post-collection. ■ In 2021, this strategy was used only in localized areas where response rate was less than 90%	for mitigation) ○ Admin resident population ■ Small undercount but less than undercount from traditional census ○ Mitigate for limited field collection and low response rates due to natural disasters. ■ Anticipating that this will be necessary for the 2023 census because of Cyclone Gabrielle ● For 2023, <u>priority was still on full field enumeration</u>, but this will be supplemented with admin data for units and attributes	than traditional census responses.	■ Counts of births and deaths from registration data ■ Population stock data ■ Smoothed super-population rates of births (by mother's age), deaths, in-migration, and out-migration
Future Plans	● Looking to move to a combined census approach as early as 2031. ○ Possible models: ■ No follow-up: admin data replaces non-response follow-ups for short form responders. ■ Admin-zone: create admin zones (no short form collected) and traditional zones	● Improved field collection model: ○ Choice of online or paper census ○ Increased assistance for completing forms. ■ More field agents ■ Earlier delivery of materials ■ Community outreach ● Experimental administrative population census: testing how admin data	● Working toward more direct usage of admin data in 2026 ● Possible large-scale direct use or combined census by 2031 ● Moving from a focus on operational efficiencies and data quality to looking more at data content ● Planning to release data tools from admin data. ○ Fixed tabular output ○ Microdata for research ○ Table builder for flexible tabular output	● Working on a population and migration statistics system ○ To provide more frequent, timely, and inclusive statistics on the population ○ To use admin data at the core ○ Outputs: ■ National to local population and migration stats ■ Population sub-groups and characteristics ■ Housing, accommodation, and living arrangements. ■ Population life journeys

	(short form collected) based on the quality of admin data available in different areas. ■ Admin-first with extended long form: use admin data to replace short form collection – possible extension of long form. ■ Long form collection to remain the same since the nature of the data means it cannot be replaced by available admin data. ● In 2026, plans to use more admin data for: ○ Operational efficiencies (reduce field efforts and optimize non-response follow-ups) ■ ex: use models to better predict what is/is not an actual dwelling. ○ Imputing non-responses post-collection ○ * Essentially: could they stop collecting sooner?	can be used for a full field enumeration. ○ *Age, sex, address, usual residence, number of children, ethnicity, Maori descent, birthplace, years since arrival in NZ, total income by income source, highest qualification, field of study, study participation, employment indicators and status, industry, sector* ○ Pushing out an interactive web app which allows users to explore this data. ■ Customized table downloads. ■ Limited crosstabs ■ Maps to small geography level ■ Timeseries ● Working toward an admin-first model ○ Work with agencies to get new data sources. ○ Continue to improve admin population census. ○ Improve population estimation models. ■ Adjust for over/under coverage. ■ Adjust for		and outcomes ● Produce a 100% longitudinal study from 2021 census. ● Future focuses: ○ Data quality ○ Sustainable platform for development, testing, and production ○ Sustainable data flows → make sure there is enough of a continuous admin data flow into the system to meet needs. ○ Transparency/ reproducibility → should be clear how outputs have been produced. ○ Efficiency ○ Improving existing outputs ○ Meet unmet user needs. ● Next steps: ○ Public consultation ○ Engagement with the public and third party sectors ○ Recommendations

	● Leading up to the 2026 census: ○ Social acceptability testing on the use of admin data ○ 2024: plans to send a small group an alternate census letter giving them the choice to participate in the traditional census or have their admin data used	misclassification of location. ■ Provide error bounds. ○ Determine the kind of surveying still necessary/what cannot be obtained by admin sources		
Limitations of Admin Data	● Coverage is not equal across all areas and demographics. ● Currently overcounts population slightly BUT this overcount is less than what is undercounted in the traditional census	● Census questionnaire is main source for Iwi (indigenous) affiliation. ● Multiple sources for the same information ● Cannot easily capture historical data (pre-digitization) ● Cannot easily capture events that occur overseas (i.e., hard to gather history on someone who recently migrated to New Zealand) ● Positive identification only ● Relying on multiple agencies with different possible agency collection issues ● Still remaining missing data	● Admin data does not pick up all groups (ex: migrants) ● Less control over these data sources, so harder to quality monitor ○ BUT haven't run into many issues pushing for technical change in collaborators data (most people want to correct quality/coverage issues)	● Costly and hard to acquire. ○ Not necessarily readily available because of how different parties store and organize their data. ● Data quality issues ● Targeted method swapping (one of the UK's disclosure control methods) is harder to implement with admin data because of the frequency of the data and the chance that the methods can be reverse engineered given the quantity of the data
Additional	● Canada considers third-	● New Zealand has no	● The Census and Statistics	● The UK asks a lot of questions

Notes	party/private sector data as part of their admin data. ○ This is already being used to help enumerate dwellings. ○ *However*, for more extensive use of this data, need separate social acceptability testing. ● Canada is not experiencing declining census cooperation at nearly the same rate as other countries (98% response rate in 2021 → only 0.4% decline from 2016) ○ Canadians like participating in the census, so part of the admin data social acceptability testing will measure sentiment towards less opportunities to participate	national individual identifier. ● New Zealand conducted their first combined (by design) census this year, so while they talked about their approach to the 2023 census, results are not yet known. ● NZ increased their field staff for 2023	Act gives their Bureau of Statistics the authority to use admin data without asking, but there is a big focus on building public trust to do so right now. ● There is little public awareness on the use of their admin data right now → hard to gauge public reaction. ● Currently uninterested in seeking usage of private data minus maybe utility data	about characteristics in their census survey → many of these characteristics cannot be gathered from admin data.

Appendix 2: Interviews with Four Statistical Agencies Australia, Canada, United Kingdom (England and Wales) and Northern Ireland[1]

We conducted one-hour interviews with representatives from four statistical agencies. We primarily asked them how they use administrative data to supplement or enhance existing censuses and surveys. They also discussed legislative changes that allowed these agencies to obtain and use administrative data.

Use of Administrative Data

Like the US Census Bureau, other international, governmental, and statistical organizations have been exploring how to use administrative data to improve the relevance, timeliness, and accuracy of their surveys and census. We explore how four countries, England/Wales, Canada, Northern Ireland, and Australia, are using administrative data in their initiatives and other work, as well as how legislation impacts their ability to acquire administrative data. Some statistical agencies are focusing on using administrative data to improve their census, while others are making strides in replacing surveys and census questions with more timely and granular information. Although there is often a tradeoff between accuracy and timeliness when collecting census and survey data, administrative data can improve quality and timeliness.

The UK Office for National Statistics (ONS) collects data for England and Wales. It is working to combat the declining quality of statistics seen between Censuses by using administrative data to provide more timely data that is of more consistent quality. Using statistical population datasets with administrative data to show how populations change and migrate, the ONS has been able to compute admin-based population estimates. These admin-based population estimates (ABPEs) can give population statistics four months earlier than the current system. Despite coverage

[1] We would like to acknowledge Olivia Catherine Davis, a third year undergraduate at University of Virginia. She is majoring in Statistics and Data Science, who summarized the interviews.

adjustments still being in the works, these ABPEs serve as granular case studies studying variables like ethnicity, income, and housing in a region.

Additionally, the ONS is exploring the use of administrative data in studying longitudinal life journey outcomes, migration statistics, and population subgroups, to name a few. This administrative data is also being complemented with surveys and commercial data, such as phone provider data, to give more informed and accurate statistics. As these initiatives are improved, the ONS will get public consultation and engagement from outside organizations to strengthen these systems built upon administrative data estimates.

STAT Canada is increasingly using administrative data in its census and surveys. They have been able to replace some survey questions by using administrative data. Additionally, they have been using administrative data to improve non-response rates on the Census, imputing values sourced from administrative data when people do not answer. In particular, Canada is focusing on using administrative data to improve the relevance and representativeness of its quality-of-life framework indicators. With 85 indicators on the well-being of Canadians, they are improving the quality of these social statistics by supplementing the data with administrative data. For example, healthcare and educational administrative data are used to create new statistical methodologies for those indicators. Private data such as retail, banking, and other commercial companies have also provided data to STAT Canada.

Northern Ireland uses administrative data as part of a data linkage and targeted enumeration strategy for their census. They use administrative data to improve aspects like deleting duplicates based on nicknames and maiden names, false people, keeping track of students moving, imputing errors, counting infants, and figuring out marital status. Additionally, administrative data were used to create an address registrar as the sample frame for social surveys. Validating addresses by performing data linkage with address-centric data sets helps the statistical association to plan enumerations. Based on the kind of property (rental or owned) and who lives there (student, elderly, etc.), the Northern Ireland Statistical Agency can preemptively predict census engagement and make conscious

choices on whether to send an online or paper version of the census to the household. They have developed strategies to follow up with those who do not respond, using contact approaches and targeted enumeration strategies. With these improvements using administrative data, Northern Ireland is working towards dropping census questions that can be derived from administrative data.

The **Australian Bureau of Statistics (ABS)** has been improving and enhancing census data with administrative data. Mainly, they have focused on improving the quality and accuracy of census counts by improving occupancy determination. They are using administrative data to get information on rental vacancies and other conditions that could make a house empty on census night. Additionally, the ABS has been using administrative data for imputations, adjusting missed people with data from similar households. This technique has proved especially useful for imputing ages. Administrative data has also enhanced the census with superior income data from tax information and has generally been a good tool in preparing for unexpected events.

Another initiative that uses administrative data is the development of labor statistics with more granular demographic and employer characteristics. By enhancing single-touch payroll data (that feeds through the tax office) with administrative data, ABS has created weekly indices of payroll jobs and wages. ABS has also created monthly spending indicators from private sector data such as credit card institutions. These indicators mirror the retail trade monthly survey but are not yet up to a standard where they can replace the survey. These private institutions share their data, although differences in quality can make combining their data with other resources difficult. They are exploring how to use administrative data to impute missing survey data.

Legislative Frameworks

Appropriate legislative frameworks are vital for allowing statistical organizations access to administrative and private-sector data. The Digital Economy Act in the UK means public organizations can share administrative data for statistics and research purposes. Additionally, the

ONS has data-sharing agreements with many organizations, such as commercial companies and local governments. Under the European Union's General Data Protection Regulation (EU's GDPR),[2] however, one can only seek data that they need and then must get rid of data they don't use or after it fulfills its need.

STAT Canada has extensive access to and authority over data (such as taxes) as a centralized authority. However, differing ideas of privacy have been a major legislative barrier. The media was upset when they tried to use credit and banking data. Although the accusations were unfounded, STAT Canada has had to be more cautious of data privacy, getting only what it needs to advance statistical purposes. They have been using sampling to get the administrative data they need without upsetting people, and they have been using checks and balances on information privacy, like deleting data after they use it.

Commercial companies, in particular, are not obligated to give STAT Canada complete data access but have cooperated in providing scanner data. The main cooperation issues arise within provincial jurisdictions. For example, health data must be collected to assess the impact of government funding, but there is resistance because provinces do not want their data used against them.

Northern Ireland has no legislative interference and can generally obtain and use supplier data under UK legislation. The Australian Bureau of Statistics has had little to no issues with acquiring data. They do not have to pay for data from private companies, recognizing this as a contribution to the public good for national statistics. They also have data agreements in place with the private sector that are not legally binding but based on good faith. Although ABS has legislation that can compel

[2] The General Data Protection Regulation is a European Union regulation on information privacy in the European Union and the European Economic Area. The GDPR is an important component of EU privacy law and human rights law, in particular Article 8 of the Charter of Fundamental Rights of the European Union. See Bowen, C. M. (2021). Protecting your privacy in a data-driven world. CRC Press.

companies to provide their data, they have not needed to implement it. Overall, the legislation surrounding data acquisition in a nation can be a critical component for advancing the use of administrative data in national statistics.

Interviewees at Statistical Agencies

Statistical Agency	Name and Title	Date
ABS	Anders Holmberg, Chief Methodologist, Methodology & Data Science Division	10 April 2023
ABS	Marcel van Kints, General Manager, Statistical Production and Digital Services Division	10 April 2023
ABS	Bruce Fraser, Director, Collection Management Unit	10 April 2023
ABS	Ross Watmuff, Statistician	10 April 2023
NI	David Marshall, Director of Census and Population Statistics NOTE: Dr. Marshall was named the new chief electoral officer on June 5, 2023. *See link below for article.*	12 April 2023
CA	Stephane Dufour, Chief Statistician of Canada	17 April 2023
CA	Lynn Barr-Telford, Assistant Chief Statistician, Strategic Engagement	17 April 2023
ONS	Sonia Carrera, Deputy Director, Social Statistics Transformation, Analysis and Research	26 April 2023
ONS	Becky Tinsley, Deputy Director, ONS Local and Coherence	26 April 2023
ONS	Jennet Woolford, Director of Population Statistics Directorate	26 April 2023

ABS - Australian Bureau of Statistics
NI – Northern Ireland Statistical and Research Agency - https://www.irishnews.com/news/northernirelandnews/2023/06/05/news/former_census_director_revealed_as_ni_s_new_chief_electoral_officer-3326112/
CA – Statistics Canada
ONS – UK Office of National Statistics