Canada case study

COUNTRY CONTEXT

This case study on administrative data systems and gender statistics is based on interviews held with technical staff from Statistics Canada during December 2019 and January 2020. Canada has highly sophisticated statistical and administrative data systems; however, it continues to face challenges associated with decentralised systems of governance and data management. The Government of Canada has demonstrated its commitment to advancing gender equality through significant investments in key programmes, policies and initiatives. While substantial gains have been made, the Government has reflected that more work is still needed, particularly in the fields of education and skills development, economic participation, leadership, elimination of gender-based violence, reduction of poverty and improved health outcomes, as well as contributing to gender equality globally.

Introduced in 2018, the Gender Results Framework (GRF) represents the Government of Canada's vision for gender equality. It is a whole-of-government tool designed to define what is needed to achieve gender equality, determine how progress will be measured, and track progress (1). The GRF defines the gender equality agenda for the Government of Canada: it ensures that gender is considered in relation to other intersecting identity factors and has legal authority through the Canadian Gender Budgeting Act, which ensures federal government's budgetary and financial management decision-making processes consider gender equality and diversity.

Within the Framework, six key areas – that require change – have been prioritised, and progress on performance indicators is routinely monitored and made available online at Statistics Canada's Gender, Diversity and Inclusion Statistics Hub and through the Women and Gender Equality (WAGE) GRF website.[2]

Many of the indicators, which form part of the Canadian Indicator Framework (CIF), align with other international frameworks, such as the Sustainable Development Goals (SDGs). Whenever possible, these indicators will be made available by intersecting identity characteristics such as disability, gender identity, sexual orientation, Indigenous identity, immigrant status and visible minority status. Data to support the CIF come from existing data collection mechanisms and reporting sources, including surveys, administrative records, monitoring networks, and other forms of open data (2). Such data are used for two purposes: to set Canada's goals for gender equality; and for ongoing gender budgeting and monitoring.

As set out in the Statistics Act, Statistics Canada is required to collect, compile, analyse, abstract and publish statistical information relating to the commercial, industrial, financial, social, and economic activities and condition of the people of Canada (3). It collects data directly through traditional means, such as by paper, telephone, or in person, and has over 350 active surveys on most aspects of Canadian life, along with the Census, which is collected every five years (4). Statistics Canada has also been using non-survey data in official statistics for around 100 years, with many programmes using administrative data from government agencies and private sector organizations, which are then integrated into official statistics to meet statistical and research requirements (5). Administrative data are used to complement and replace surveys (or components of

2 Available at: https://www.statcan.gc.ca/eng/topics-start/gender_diversity_and_inclusion and https://cfc-swc.gc.ca/grf-crrg/index-en.html

surveys) and are both an economical method and one that reduces the burden on respondents. Statistics Canada plays an important role in housing data collected from various surveys and administrative systems, and has very strict privacy, security, confidentiality, and transparency policies, ensuring that any data collected is proportional to the need.

Statistics Canada is the central focal point for reporting Canada's data on the SDG indicators, which are available online through the agency's Sustainable Development Data Hub.[3] Currently 9 of the 17 sub-indicators relating to Goal 5 (Achieve gender equality and empower all women and girls) are reported online, while appropriate data sources for another seven are still being explored, and data for one sub-indicator is not available. Subject experts were involved in a consultative process to help select the best data sources for each sub-indicator, with preference given to those sources that most closely match international metadata, which resulted in six of the nine available sub-indicators being sourced from survey data.

When considering gender statistics more broadly, although Canada has comparatively good data, significant gaps remain, particularly around emerging areas such as the representation of women in politics across the country's jurisdictions, and the gendered impact of climate change on diverse groups of people. Data that have traditionally been sourced from household surveys, including topics related to the environment, labour force, and income and tax, are also of particular concern as the 'gender element is missing' from many of these surveys, which have focused on data at the household, instead of the individual level. Respondents also commented that more data on access to sexual and reproductive health services would be beneficial, along with having access to additional data for cross-tabulation, such as usual residence, to assist in lower-level statistical analyses.

KEY COMPONENTS OF A SUCCESSFUL APPROACH TO USING ADMINISTRATIVE DATA FOR GENDER STATISTICS

While acknowledging the important and continued role that surveys will play in Canada's statistical system, respondents commented on the need to use multiple sources of data as part of good statistical governance. Within this context of moving away from relying primarily on surveys, respondents equally stressed the importance of needing to carefully consider the use of administrative data sources, and of being constantly mindful in understanding why the data is being collected, particularly on sensitive topics. One respondent, for example, commented that collecting data on sex and gender for display on official identity documents is very different to collecting the same data for a research project, thereby highlighting the importance of good communication about data collection and its use.

The success of Canada's approach to using administrative data is due in part to the strong national coordination mechanisms established between Statistics Canada and key partners including WAGE, the Department of Finance, and provincial ministries such as health and education. National coordination is supported by a strong partnership approach fostered by Statistics Canada to facilitate data sharing and to ensure data quality. As reflected by one respondent, "strong relationships [are] needed – otherwise data quality will suffer." Part of this work includes ensuring that data custodians, as producers of the data, can see the 'value-add' in investing in their administrative systems. For example, understanding the value of having data by sex and/or gender offers organization in terms of monitoring a policy or programme. This partnership approach is critically important, because even with legislation that mandates the sharing of data, sharing is not always guaranteed and nor is its quality.

Further, while a key role of Statistics Canada, or any National Statistical Organization (NSO), is to assist in improving the quality of data, it is the responsibility of the NSO to fully understand the data source, including any challenges and weaknesses. This process is facilitated by Statistics Canada's Quality Assurance Framework (QAF), which outlines the measures the agency has put in place to manage quality, and provides

---

3 Available at: https://www144.statcan.gc.ca/sdg-odd/index-eng.htm

unicef
data for children

guidance to statistical programme areas as they develop and implement quality management strategies to meet their users' needs (6). One respondent commented that working with administrative data is:

> "not free, [and] sometimes not useful, but that doesn't make it not worthwhile examining it – as it may be useful for something else; [or] for building a relationship with another entity and data source."

Another important lesson to come from Canada is the need to "start simply". Regarding the SDGs, for example, one respondent commented that if all indicators were disaggregated by all possible variables, it would produce over 700,000 data points, raising the new challenge of attempting to understand what all that data means. While countries with sophisticated statistical systems may, technically, be able to produce high levels of disaggregation, respondents commented on the need for data custodians and users to understand basic issues such as "what do you want to look at" and "what do you need disaggregation on." Focusing on a select number of high-priority data items, and doing them well, sets the stage to expand as well as helping to build strong relationships with data custodians for future collaborations.

Along with this central coordination and relationship-building role, Statistics Canada also provides specialized technical assistance on gender statistics. Using the GRF as a guide to help highlight gender equality issues, a training programme was developed to help organizations find gender-relevant data within Statistics Canada's website, both to help keep the data relevant and ensure organizations have access to up-to-date statistics. Training is tailored to each organization, so that staff can understand how to find the statistics, and to understand why it is important to them. Statistics Canada has also implemented standards for the transmission and composition of data messages from data custodians, to ensure that data arrives in the same format, improving the overall quality of the data and reducing the time needed to clean and validate the various data sources.

INNOVATIONS WITH DATA LINKAGE

Statistics Canada has a sophisticated data linkage programme, offering substantial insights into several areas of gender equality. Data from income tax files have been linked with business ownership data to get a better understanding regarding the proportion of women who own businesses: while data on mothers (collected through birth registration) have been linked with census data on additional characteristics and variables to provide a more complete understanding of the lives of Canadian mothers. In Ontario, graduate earnings (as measured through tax returns two years after graduation) have been used to demonstrate education outcomes for higher-education providers, while also offering an important insight into any persistent gender wage gaps. Preliminary work has also been done in the education sector in linking administrative data on public postsecondary enrolments and graduates with census data to improve coverage on Indigenous status, which is currently poorly reported, allowing for more detailed gender analyses. Similarly, plans are underway to link postsecondary education data to administrative data on disability support to provide an understanding of postsecondary participation for students with disabilities.

It is important to note that while several different identity numbers exist in Canada, there is no universal identifier, with most linkage being done via probabilistic methods, rather than directly. Further, all linkage is done in the context of strict privacy regulations, with one respondent commenting that the first question must always be "do we really need to link?", followed by "are we able to?" As part of ensuring confidentiality, analytical records are kept separate from personal identifiers, and Statistics Canada has sophisticated data validation processes, which include guidelines on linkage rates, and minimum acceptable coverage levels for linked data to be published. As reflected by one respondent, given the extensive initial investments and ongoing data validation processes required, data "linkage is good, just not timely."

CHALLENGES AND BARRIERS WHEN WORKING WITH ADMINISTRATIVE DATA SYSTEMS

Administratively, Canada is divided into 10 provinces and 3 territories, with provinces considered sovereign within certain areas based on the divisions of responsibility between the provincial and federal government. As with many countries that have decentralised systems of governance, this has a significant impact on operations of statistical and administrative systems. For ministries of health and education particularly, each jurisdiction (province or territory) does things differently. This means that every administrative data file may not be collected the same way, with certain jurisdictions asking questions in a slightly different way, making national-level analyses difficult. With regard to the counting of maternal deaths, one respondent commented on the ongoing difficulties in ensuring that each jurisdiction asks and records data relating to the length of pregnancy in a consistent way. For education, although gender is now being requested for postsecondary administrative data, the ability to provide it is not always possible because this information is not collected for administrative purposes by some institutions.

As reflected by one respondent, legislative change is also required for provinces to amend the wording of questions on various data collection forms, which is a very onerous process. For agencies that cannot see the direct benefit of the change, this makes the task even more difficult. Overall, this means that some changes may never happen, while others may not happen for many years: innovative approaches are required by Statistics Canada to develop ways to work with the datasets, given their limitations and weaknesses.

MOVING FORWARD

Given the significant investments in key programmes, policies and initiatives on gender equality in Canada, the Government is continuing to build on its successes in this area. The 2018 report, 'Modernizing the Government of Canada's Sex and Gender Information Practices', provided recommendations on ways to strengthen handling of information on sex and gender, as an initial step to embed gender diversity in business modernization (7). In response, Statistics Canada has developed new standards on sex and gender variables and classifications, as well as offering a non-binary gender option using the 'X' identifier when collecting information in certain administrative systems. One respondent reflected that this will improve data availability around gender-based violence, for which the experiences of people identifying as non-binary are currently missed. While data will continue to be collected on sex, this new policy direction will allow for a better understanding of gender issues by providing a two-step process for collecting data on sex and gender. One respondent commented that these "new and exciting" methods of data collection are needed to help "understand issues rich in experiences", a critical component of gender analyses.

ACKNOWLEDGEMENTS

REFERENCES

1. Government of Canada, 'Gender Results Framework: Status of women in Canada', 19 March 2019,   <www.cfc-swc.gc.ca/grf-crrg/index-en.html>, accessed 31 January 2020.

2. Government of Canada, 'Towards Canada's 2030 Agenda National Strategy', July 15, 2019, <www.canada.ca/en/employment-social-development/programs/agenda-2030/national-strategy.html>, accessed 6 March 2020.

3. Statistics Canada, 'Mandate and objectives', 30 March 2016, <www.statcan.gc.ca/eng/about/mandate?HPA=1>, accessed 31 January 2020.

4. Statistics Canada, 'About us', 5 June 2019, <www.statcan.gc.ca/eng/about/about?MM=as>, accessed 31 January 2020.

5. Statistics Canada, 'Where does data come from?', 15 November 2019, <www.statcan.gc.ca/eng/our-data/where>, 31 January 2020.

6. Statistics Canada, 'Statistics Canada's Quality Assurance Framework', Statistics Canada, Ontario, 2017.

7. Government of Canada, 'Modernizing the Government of Canada's Sex and Gender Information Practices: summary report', Government of Canada, 8 April 2019, <www.canada.ca/en/treasury-board-secretariat/corporate/reports/summary-modernizing-info-sex-gender.html>, accessed 3 February 2020.