## The Canadian Census of Population Administrative data usage

**CNSTAT Public Webinar Series:** 

International Approaches to Conducting a Census in the 21<sup>st</sup> Century May 24, 2023





1





## Statistics Canada: Who we are and what we do

- Canada's national statistical agency
- One of the leading statistical organizations in the world
- More than 5,000 employees across Canada
- More than 350 active surveys
- Trusted source of information accurate, objective and unbiased
- Privacy and confidentiality are paramount





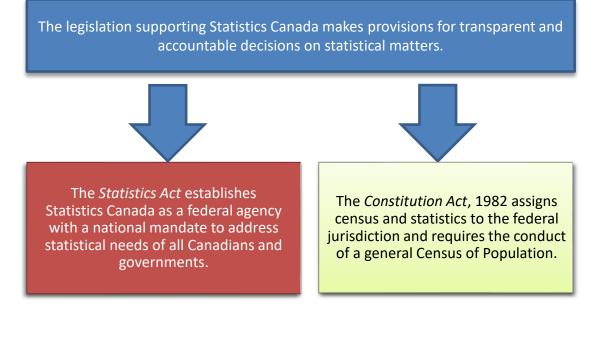


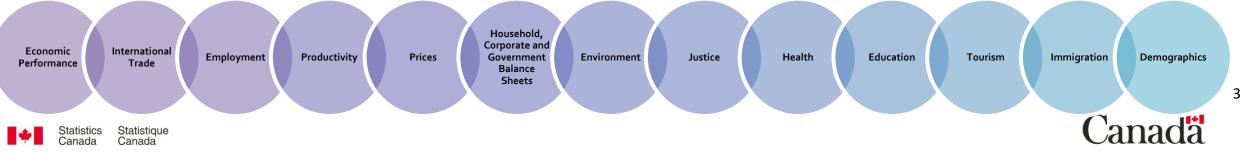
### The organization

### PROVIDING INSIGHT THROUGH DATA

Statistics Canada is an independent, arm's-length agency with a national mandate to **provide statistical information** on every aspect of our country and its communities, including its population, economy, society and environment.

As the **National Statistical Office**, Statistics Canada provides Canadians with vital information to help monitor inflation, promote economic growth, plan cities and roads, adjust pensions, and develop employment and social programs. The agency **provides data and insights to help governments, businesses and individuals make informed decisions**.





### CENSUS · RECENSEMENT

## Census Around the World – 2020 Round

- Of the 52 UNECE Countries with information available, less than half are conducting a traditional Census in the 2020 round
  - 23 traditional census
  - 13 register-based census
  - 16 combined approach

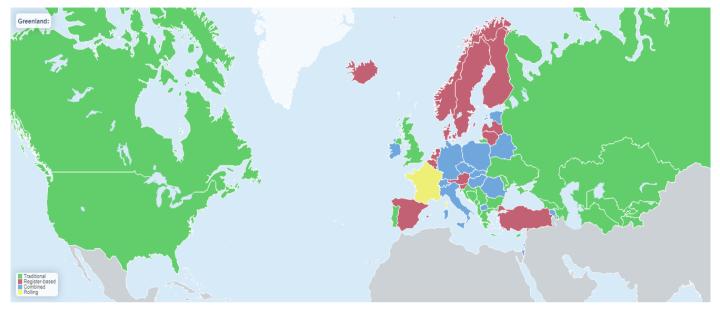
#### Pages / UNECE census wiki / Censuses in UNECE Countries

### Censuses of the 2020 round

Created by Paolo Valente, last modified by Kristen Jeffers on 07 Jun, 2022

#### Plans and practices of UNECE countries for the population and housing censuses of the 2020 round.

The information is collected directly from national statistical offices and census agencies, with additional information gathered by UNECE where appropriate. All the information presented should be considered as j have taken place. See bottom of page for definitions and instructions. The map can be zoomed and scrolled. Click on a country for more information or see the full entry in the table below.



Traditional; Register-based; Combined, Rolling



Q Search





## **Introduction to the Canadian Census of Population**

- The Canadian Census of Population aims to enumerate and to collect information on the entire population of Canada every five years
  - Full field enumeration is conducted  $\rightarrow$  Traditional Census
- Short form and long form questionnaires
  - Short form questions on basic socio-demographic information and languages (100% of population)
  - Long form (25% sample in the 2016 and 2021 censuses): dwelling characteristics and additional demographic and socio-economic characteristics on individuals
- Collection:
  - Collection modes: "mail out" and "list leave"
  - Various field operations conducted to ensure appropriate dwelling classification and response level
  - Non Response Follow Up (NRFU)





# 2021 Wave Methodology

Statistique

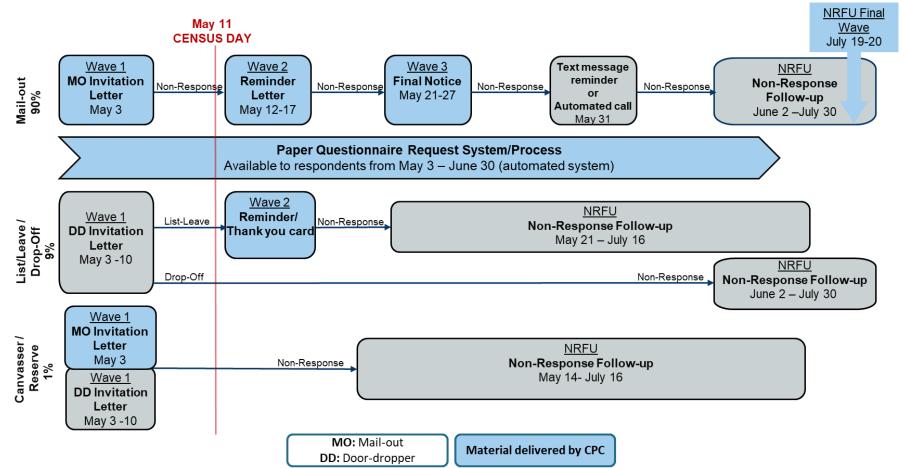
Canada

Statistics

Canada



**CENSUS** • **RECENSEMENT** 



#### Wave 1

15.6 million one-day delivery of self-mailers (including 225,000 for Census of Agriculture)

#### Wave 2

11.2 million self-mailers 1.1 million ad cards

#### Wave 3

5.7 million letters in envelope

### NRFU Final Wave

1.9 million letters in envelope





CENSUS · RECENSEMENT

•





## **Global Collection Response Rates - Highlights**

Collection Rate	2016 Census	2021 Census	
	Actual	Planned	Actual
Global Census Collection Rate	98.4%	98%	98.0%
Long Form Collection rate	97.9%	96%	97.4%
Internet	68.3%	80%	84.1%
Paper	20.5%	10%	4.1%
Self-Response	88.8%	90%	88.1%
NRFU	9.7%	8%	9.9%
Workload at start of NRFU	3.7 M	4.0M	4.6M





## **Introduction to the Canadian Census of Population**

- Coverage Studies: measured through 3 statistical activities
  - Dwelling Classification Survey (DCS): follow-up operation to estimate and correct for misclassification of dwellings on Census Day
  - Census Coverage Measurement Program: Main sources of information are administrative data
    - Census Undercoverage Survey (CUS) which was formerly known as the Reverse Record Check: estimates the population missed by the Census using various administrative data sources and to create an independent frame and follow-up with survey information
    - Census Overcoverage Study (COS), estimates the population enumerated more than once mostly a linkage exercise
- Census of population direct use of administrative data
  - Income and some immigration variables
  - Disruption of collection, e.g. wildfires in northern region

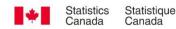




## **Statistical Contingency Plan for the 2021 Census**

- In the event that COVID-19 or other disasters prevented the traditional census from a satisfactory enumeration
  - Statistics Canada developed (and eventually implemented) a "statistical contingency plan" based on the use of administrative data for post-collection imputation of 2021 Census non-responding households
    - Note that the contingency plan used data already provided to Statistics Canada, and meets the highest standards of privacy, confidentiality and data security.
  - Statistical models (see paper for details on the models) were used extensively to form households from individual government-provided administrative information (e.g., tax records) and quality indicators were developed to determine household eligibility for inclusion in imputation processes.

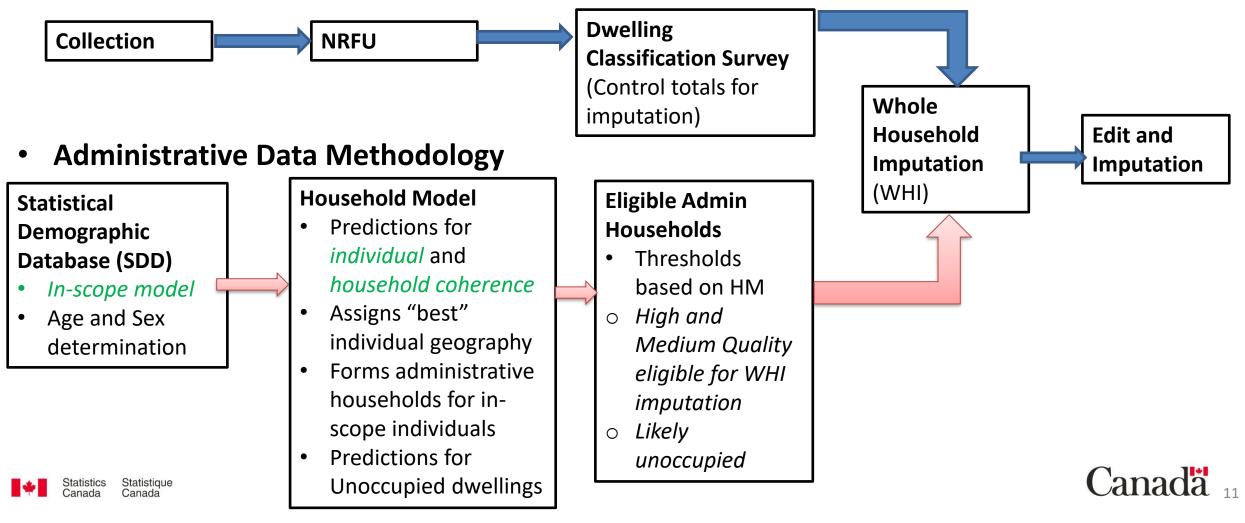






### Administrative Data Contingency Methodology Census 2021 (Post-collection)

Traditional Census Collection Methodology





## Statistical Contingency Plan for the 2021 Census

- 2021 Census had a successful enumeration, with 98% of Canadians responding to the census.
  - Widespread use of the contingency plan was not necessary
  - Administrative data were used to support the imputation of non-responding households for some localized areas of the country that had response rates well below the national rates.
    - 1,045 collection units (out of 49,000 in Canada) showed a response rate below 90%.
    - Dwellings that had good quality administrative data were in-scope.
    - About 12,000 non-responding households representing less than 0.1% of occupied private dwellings in Canada.
    - However, there were 9.2 million dwellings eligible for imputation using administrative data







## Increased Usage of Administrative Data for the 2026 Census

- How ?
  - Operational efficiencies during the NRFU
    - Various models could be used to reduce field efforts away from some in-person activity (classification of dwelling)
    - Optimizing NRFU activities by using signal from good administrative data
  - Post-collection: Imputation of non-responding households
    - Mainly used in mail out areas in 2021; extend to non-mailout in 2026





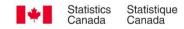


## Combined Census research: Could we conduct one in 2031?

A "combined census" combines information from self-response with administrative data to generate Census counts.

- Simulation and testing of various combined census models
- Research and testing using various admin data sources
- Research and testing of new types of post-censal coverage studies
- Admin data acquisitions







CENSUS · RECENSEMENT

### Research on administrative population counts: What we have learned so far

- We have access to plenty of good quality administrative data, but not everywhere in the country or for all population groups
- We can produce population counts from administrative data
  - However, the administrative data over-count somewhat. That over-count is smaller than the undercount from the current traditional census...
- Of the individuals in those files, we are getting good at predicting if they are in-scope for census enumeration:
  - Current "in-scope" model accurately predicts if in/out of scope 88.4% of the time
- Our models are getting good a placing individuals from administrative data into households
  - "Household model" places individuals in correct addresses 80% of the time.
- The large majority of census households can be exactly re-created by administrative data

- 60.6% of derived households match exactly to the Census

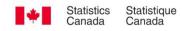




## Combined Census models under consideration

In order of increasing degree of admin data usage:

- **1.** No follow-up model: No follow-up for short form dwellings after Wave 1 letter, admin data replaces NRFU. Long-form sample continues as usual, with follow-up if needed.
- 2. Admin-zone model: Determine "admin" and "traditional" geographic zones based on the quality of available admin data. No short-form in "admin zones". Short-form as usual in "traditional zones". Long-form sample as usual.
- **3.** Admin-first model with extended long-form: Admin replacement for short-form (no census invitation sent). Long-form sample maintained or increased in some areas as needed.
- \* Assumption that long-form sample will persist as usual throughout all combined census options. Short-form characteristics not available in admin data collected on the long-form.







## Social acceptability testing

Any major change to census collection methods will be predicated on a solid foundation of ongoing research, consultation, and transparent communication with internal and external stakeholders, and the Canadian public

- Research, testing, and consultation roadmap to 2027
- Qualitative and Quantitative social acceptability testing opportunities
- 2024 Census Behavioural Tests (BT4)

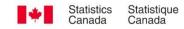






Why measure the social acceptability of a combined census?

- Provides strategic information for the use of admin data in 2026
- Key input into decision in 2027
- Strong evidence basis required for 2031
- CSAC, OPC, and SMC have stated the need for social acceptability measures
- General public, researchers, data users, provinces and territories are all stakeholders







## Consultation, outreach, communication

A proactive and reactive communications strategy to mitigate the risks of negative public reaction to combined census research and re-use of admin data

- Qualitative testing of communications materials
- Consultations with key stakeholders
- Privacy Impact Assessment (PIA), Necessity & Proportionality review (N&P), and consultation with Office of the Privacy Commissioner (OPC)
- Census Futures section on the Census website
- Media lines and Q&A





# Thank you / Merci

## QUESTIONS / COMMENTAIRES?



