

CENSUS • RECENSEMENT

# Researching Administrative Data Sources and Geospatial Datasets for Canada's 2021 Census

Presentation by Marc Hamel  
Census Manager, Statistics Canada  
UNECE Expert Meeting on Census  
September, 2019





CENSUS • RECENSEMENT

# Introduction

- This presentation is on research done on the use of administrative data to replace content, add new content and improve existing content for the Canadian Census, for:
  - Immigration
  - Housing and collective dwellings
  - Commuting to work

2



CENSUS • RECENSEMENT

# Administrative data

- Research is underway to assess the strengths and weaknesses of replacing questionnaire content with administrative data
- Exploring the potential of replacing: immigrant status, year of immigration and housing information
- Effort also underway to assess use of administrative data to improve concepts such as commuting to work, and simplify enumeration of individuals living in collective dwellings

3





CENSUS • RECENSEMENT

# Immigration data in Canada

- **Census of Population**

- Key data source on the socio-economic outcomes of immigrants in Canada
- Immigration questions: immigrant status, year of immigration, place of birth, place of birth of parents and citizenship

	<b>Immigrant</b>	<b>Non-immigrant</b>
Canadian by birth		Canadian by birth
Canadian by naturalization	Canadian by naturalization	
Not a Canadian citizen	Permanent resident	Non-permanent resident



CENSUS • RECENSEMENT

# Immigration data in Canada

- **Immigration, Refugees and Citizenship Canada (IRCC) administrative data**
  - Immigration landing file (1980 to present)
  - Historical landing file (1952 to 1979)
  - Asylum claims and temporary resident permits (1980 to present), visitor records (2006 to present), and citizenship records (2005 to present)
- **IRCC administrative data was used to enhance 2016 Census**
  - Used to improve 2016 Census data processing
  - Used to add admission category variables onto the 2016 Census

5





CENSUS • RECENSEMENT

# Immigration

- For 2021, Canada is testing replacing questions on the year of immigration and immigrant status using integrated IRCC data to:
  - Reduce response burden
  - Improve data quality
  - Potentially add new content to the 2021 Census from administrative data

6





CENSUS • RECENSEMENT

# Challenges of administrative data

- Administrative data are not updated to reflect deaths or emigration
  - The data need to be linked to estimate the current population of immigrants
- No records for immigrants who landed prior to 1952
  - 111,000 immigrants or 1.5% of immigrants in 2016
- Weaker linkage quality for the pre-1980 file
  - Especially records from 1961-1972 due to incomplete date of birth (only year of birth is available)
- No records for dependants of temporary resident permit holders, visitors, or other non-permit non-permanent residents

7





CENSUS • RECENSEMENT

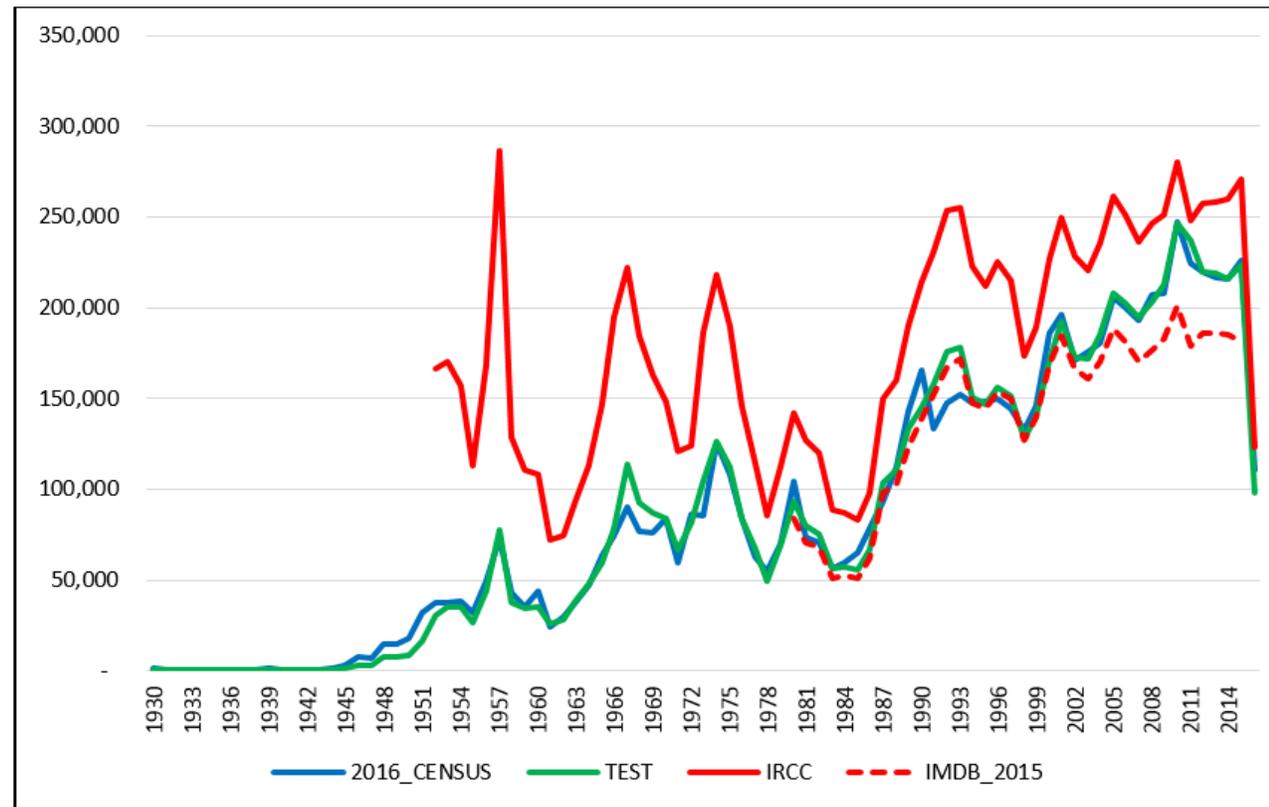
# Research: How to address missing integrations or inconsistencies

- Record linkage to administrative data identifies immigrants and non-permanent residents
  - Year of immigration for immigrants
- Similar edit and imputation methods used in the past can be applied. Absence of a value could be due to missing administrative data or respondents out-of-scope for linkage.



CENSUS • RECENSEMENT

## Preliminary Test Results by single year of immigration, 2016 Census and test results



**Source:** Statistics Canada, 2016 Census and administrative data replacement results, Census of Population and Longitudinal Immigration Database (IMDB) for tax year 2014; and Immigration, Refugees and Citizenship Canada (IRCC).





CENSUS • RECENSEMENT

# Housing

- In previous censuses, Canadians have been asked questions about their shelter costs, including the amount spent on utilities such as electricity; oil, natural gas and other fuels and water and other services.
- For 2021, Statistics Canada is researching the feasibility of using information obtained from electrical suppliers for two purposes:
  1. Assisting in establishing the presumed occupancy status of a non-responding dwelling. The desired outcome would be to reduce follow-up efforts.
  2. Using the electricity bill as a replacement for the electricity payments component of the utility cost question.

10



CENSUS • RECENSEMENT

# Housing

- Preliminary results from the 2019 Census Test show that not asking respondents about electricity costs:
  - Reduced completion time
  - Increased presence and increased amounts of fuel payments and water and other services payments
- Additional challenges when matching the electric meter addresses on billing files to census dwellings

11



CENSUS • RECENSEMENT

# Limitations of using Administrative Data Housing

- The general caveats apply for integrating administrative data. They must be solved or brought to an acceptable level of risk:
  - compatibility of concepts
  - sufficient coverage
  - data supply risk
  - cross-country uniformity
  - mode uniformity

12





CENSUS • RECENSEMENT

# Collective dwellings

- A dwelling of a commercial, institutional, or communal nature.
- Statistics Canada employs a number of mechanisms to manage the response burden, such as encouraging facility administrators to help us use administrative records to enumerate those living in institutional collective dwellings.
- Since 2016, Statistics Canada has been exploring new uses of administrative data for certain types of collectives.

13



CENSUS • RECENSEMENT

# Collective dwellings

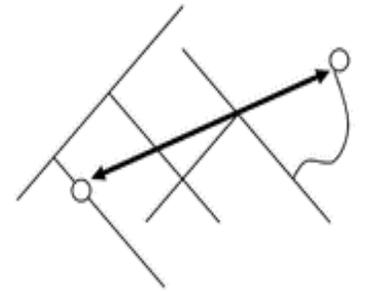
- Investigated reusing data from the Canadian Forces to see whether Statistics Canada could do records integration to count individuals living on military bases.
  - Will likely continue to enumerate military bases with field visits (as in 2016), given concerns with the quality of the records. Other uses may be possible.
- Intend to explore the use of data from public guardians and trustees to supplement data collection for some residents of collective dwellings (Nursing Homes, Residences for Senior Citizens and Residential Care Facilities).
- Attempting to replace field enumeration of Correctional Facilities for the 2021 Census with re-use of administrative data from our Justice programs.

14



CENSUS • RECENSEMENT

# Commuting to work



- One derived variable that the Census disseminates is the distance from home to work: the straight-line distance between the blockface edge of the home address and the blockface edge of the workplace address.
- Major limitation: few commuters take a straight line from home to work. Road networks and land features (rivers, mountains, etc.) all alter our journey to work trajectory.
- Therefore, this measure underestimates actual distance.

15





CENSUS • RECENSEMENT

# Commuting to work

## Using latitudes and longitudes

- Various geospatial datasets are being evaluated, including DMTI Route Logistics, the National Road Network File (NRN), open data from transit authorities, and Google Maps API.
- The tools used to assess these datasets include ArcGIS, Python, Google Maps API, Open Street Map and Azure Server.
- The idea is to use the actual latitude and longitude of the location, rather than the latitude and longitude of the blockface edge of the location.

16





CENSUS • RECENSEMENT

# Limitations of using Administrative Data Commuting to Work

- There are a few limitations to using this new approach to calculate distance from home to work.
- The first limitation is resource-based: human and financial resources.
- Secondly, route calculations are not 100% accurate. Further restrictions for calculating routes are needed to reduce errors.

17





CENSUS • RECENSEMENT

# Conclusion

- Administrative data offer an opportunity to improve the quality of census data, reduce burden and expand content.
- There are challenges to overcome in order to successfully take advantage of administrative records for these purposes.
- One challenge is acquiring the data.
- Sometimes the quality of the administrative data are not sufficiently high for substitution.

18





CENSUS • RECENSEMENT

# Conclusion

- Statistics Canada has several examples of very successful data replacement (income and immigration).
- For the 2021 Census, the use of immigration data records for immigration information will be expanded, subject to testing and analysis.
- Administrative data sources also provide an opportunity for innovation (journey to work , collective dwellings)

19





CENSUS • RECENSEMENT

# Questions?

Thank you!

**Marc Hamel**

**Census Manager, Statistics Canada**

**[marc.hamel@canada.ca](mailto:marc.hamel@canada.ca)**