



Statistics
Canada

Statistique
Canada

Canada



Statistics Canada
www.statcan.gc.ca



Administrative Data Initiatives at Statistics Canada

Julie Trépanier, Jean Pignal and Don Royce

2013 Federal Committee on Statistical Methodology Research Conference
November 5, 2013



Outline

- Use of administrative data at Statistics Canada
- Creation of the Administrative Data Secretariat
- Administrative Data Inventory
- International review of legal, policy and organizational frameworks
- Research into a Canadian Statistical Demographic Database

Use of Administrative Data at StatCan

- For the statistical community, administrative data are seen more and more as data collected by sources external to statistical offices (not restricted to data from government programs)
 - Alternative (or secondary) data sources
- The authority of Statistics Canada to obtain administrative data is governed by Section 13 of the *Statistics Act*

Use of Administrative Data at StatCan

Statistics Canada has been using administrative data for a long time

- as a stand-alone source to produce national statistics, e.g.,
 - provincial and territorial vital statistics (since 1921)
 - customs' import and export data for international trade (since 1938)
- to support surveys, e.g.,
 - creating the Business Register using Canada Revenue Agency (CRA) data (since 1988)
 - creating the Address Register using files from CRA, the Canada Housing and Mortgage Corporation, selected telephone and cellphone companies, Info-direct, Canada Post Corporation, etc. (since 1991)

Use of Administrative Data at StatCan

- to partially replace survey data, e.g.,
 - Survey of Employment, Payrolls and Hours (since 1994): using number of employees and gross payrolls from CRA payroll deduction accounts
 - business surveys: using CRA income tax in annual surveys (since 1997) and goods and services tax sales in monthlies (since 2004)
 - In 2010/2011: 22% of business surveys used administrative data, representing over 300,000 fewer records to collect from respondents
 - for revenues in household surveys (since 1995): Using CRA income tax
- to provide new analytical opportunities, e.g.,
 - Longitudinal Immigration Database (built in 2001): using Citizenship and Immigration Canada data (landed immigrants since 1980) and CRA income tax data (since 1982) to study the economic behaviour of the immigrant taxfiler population

Creation of the Administrative Data Secretariat

Rational:

- Statistics Canada is mandated by the *Statistics Act* to promote the avoidance of duplication in the information collected by government
- Statistics Canada is committed to reduce response burden imposed by its surveys
 - Part of wider government initiatives to reduce burden on businesses
 - Difficulty in reaching household survey respondents has been identified as one of the six corporate risks for Statistics Canada



Creation of the Administrative Data Secretariat

Rational (continued):

- Increasing budget restrictions have resulted in more focus on using existing sources
- Previous committees at StatCan concluded that to increase the use of administrative data even more, better coordination was needed.
- StatCan's Tax Data Division provides that coordination role for data received from the CRA and used by many of its statistical programs. There is no equivalent organizational unit for other administrative data.

Creation of the Administrative Data Secretariat

Rational (continued):

- Statistics Canada's vision is that it will use administrative data in its statistical programs when it leads to a better outcome — that is, a better balance between relevance, quality, costs and respondent burden.
- An Administrative Data Secretariat is needed to support the vision stated above.
- It was created in September 2012 with the mandate to develop and implement a corporate approach to increasing administrative data use.

Creation of the Administrative Data Secretariat

The Secretariat has 3 key objectives:

- Put in place a **governance structure** that supports statistical programs in the efficient acquisition, management and use of administrative data
- Launch initiatives that look into **optimization** of methods and processes surrounding administrative data
- Supports statistical programs' **research** of new data sources and methods and facilitates the sharing of information and practices

Administrative Data Inventory

- Contains metadata on files received by StatCan, such as providing organization, statistical program using it and how, existence of a formal agreement, costs etc.
- First inventory of 2012-2013 data holdings was completed in June 2013. In need of improvement.

Source of File	#	Usage by Statistical Program				
		Economic & Environ.	Socio-economic	Censuses	Statistical Infrastruct.	Cost-recovery
Federal	187	148	51	3	37	2
Prov./Terr./Mun.	135	97	14	2	22	0
Private	166	140	1	2	18	5
Foreign	11	11	0	0	0	0
Others	13	2	6	2	3	0
Total	512	398	72	9	80	7



International review of legal, policy and organizational frameworks

- Six countries reviewed: Australia, Canada, Ireland, New Zealand, United Kingdom and United States
- No attempt to compare to countries with fundamentally different statistical or administrative infrastructures (e.g., Nordic countries, Netherlands)
- Also examined admin data-related work of international organizations: United Nations, UNECE/CES, and Eurostat/ESS

•

International review of legal, policy and organizational frameworks

Legal framework:

- For legal authority to influence, access and use admin data, there is a lot of variation across NSOs, e.g.,
 - UK must get Parliament's approval for any new source
 - Ireland has the strongest legislation – authority to influence as well as access and use admin data
- All countries give administrative data the same confidentiality protection as survey data
- All countries have privacy legislation with very similar principles

-

International review of legal, policy and organizational frameworks

Policy framework:

- Reflects overall government intent. Easier to modify. It typically includes:
 - Most countries have a form of national code of practice for official statistics
 - Many examples where countries take a more government-wide approach to the development of statistical policy for using admin data
 - Several countries are taking a more government-wide approach to “statistical data integration”
 - Several countries have external bodies that advise the government directly on statistical issues

International review of legal, policy and organizational frameworks

Organizational framework:

Three common themes emerged:

1. NSOs need to establish effective processes for identifying, influencing and accessing admin data, and for ensuring that data are of sufficient quality for statistical purposes
2. NSOs need to create the proper data stewardship environment to ensure that admin data are treated with proper care and in accordance with legal requirements
3. NSOs must ensure long-term stakeholder support by being transparent about what admin data are used and how they are used

Research into a Canadian Statistical Demographic Database

- Censuses in Canada are quinquennial
 - 2011: Short mandatory census + voluntary National Household Survey
- Final Report on 2016 Census Options
 - A traditional census is the only viable option for 2016
 - Pursue research into administrative data
 - <http://www12.statcan.gc.ca/census-recensement/fc-rf/index-eng.cfm>
- Research: Study the feasibility of building a “virtual population register” from various administrative data files
 - Micro level database
 - Basic demographic information (name, date of birth, sex, address of the usual place of residence)
 - Quality indicators for each of these variables



Base Population

COMPONENT OF CHANGE	SOURCE	MEASUREMENT
Natural Growth	Provincial and territorial Vital Statistics databases shared through the NRS, adding births subtracting deaths Births abroad and foreign adoptions can be obtained through CIC	Deaths (right arrow) Births (left arrow)
International Migration	CIC , Administrative database adding new immigrants, new non-permanent residents and subtracting non-permanent residents who left CRA , IDENT Administrative database used to identify emigrants and former emigrants who returned	Entering NPR (left arrow), Exiting NPR (right arrow) Permanent Immigrants (left arrow) Emigration (right arrow), Returning Canadians (left arrow)
Internal Migration	CRA , IDENT Administrative database used to identify migrants within Canada for provincial and territorial reapportionment	Origin/destination matrix (double-headed arrow)
Data Quality Adjustment	Survey based and/or Statistical Process to measure coverage errors (aka error of closure) from a Reverse Record Check and/or match to an alternate source	Net Temporary Emigration Coverage adjustment (double-headed arrow)

Next Base Population

CRA = Canada Revenue Agency
 CIC= Citizenship and Immigration Canada

Research into a Canadian Statistical Demographic Database

- Phase 1: Create the May 2011 prototype CSDD
 - Starting with the 2006 Census as the base population
- Phase 2: Assess quality against the 2011 Census and recommend improvements (March, 2015 – Interim report)
- Phase 3: Implement improvements and create a 2016 prototype
 - Two paths: With and without a rebasing with the 2011 Census
- Phase 4: Assess quality against the 2016 Census (January, 2017)
- Phase 5: Produce report and recommend uses (March, 2017 – Final Report)



Research into a Canadian Statistical Demographic Database

Potential uses and benefits:

- Support for future census planning, non-response follow-up, processing, quality assessment
- Support for coverage studies
- Support for the Population Estimates Program
- Person level frame for household surveys
- Alternative to a traditional census



Final thoughts

- Other initiatives:
 - Work to strengthen policies, directives, guidelines and tools related to admin data, especially their acquisition, and to better communicate them
 - Work to investigate optimisation opportunities
 - Framework to assess the quality of admin data and determine their statistical usability
 - Monitoring Big Data work internally and externally
- Limited budget: About 5-6 employees, including the census research project



Contact

For more information,
please contact:

Pour de plus amples
renseignements, veuillez
contacter :

Julie Trépanier

613-951-9437

Julie.Trepanier@statcan.gc.ca