

Roadmap

Enhanced access, use and management of administrative data in statistical programs

Prepared as part of the
PRASC Regional Seminar on Administrative Data
Trinidad, Jan 26-28, 2016



Funded by the
Government
of Canada

Canada 

ROADMAP FOR ENHANCED ACCESS, USE AND MANAGEMENT OF ADMINISTRATIVE DATA IN STATISTICAL PROGRAMS

National Statistical Offices (NSOs) should influence, have access to and make maximum possible use of administrative data in their statistical programs, resulting in reduced response burden, lower costs, improved quality and new data series for users, while maintaining public trust.

The steps outlined in this document are provided as a guideline for NSOs who want to enhance the access, use and management of administrative data in their statistical programs. For each step, a series of tasks is provided for consideration by the NSO and the benefits to the statistical program are listed. For some steps, tools and templates are included. The order of the different steps presented is provided as a general guideline – some tasks can be conducted in parallel.

Administrative data sources are data holdings containing information which are collected by organizations other than the NSO for their own purposes.

This document was prepared as part of the regional seminar on Administrative Data held in Trinidad, January 25-28, 2016, as part of the Project for the Regional Advancement of Statistics in the Caribbean (PRASC). More details on the different steps of the roadmap can be obtained by referring to the presentation slides. Any questions or assistance request can be directed to the PRASC email: statcan.prasc-prasc.statcan@canada.ca

National Statistical Offices should influence, have access to and make maximum possible use of administrative data in their statistical programs, resulting in reduced response burden, lower costs, improved quality and new data series for users, while maintaining public trust.

1. Identification of Administrative Data Coordinator

A first step can be tasking someone in the NSO to take charge of the administrative data portfolio. This can be done by identifying an administrative data coordinator (one person or a small team) who will move the project forward and will be the go-to person for all aspects of administrative data (strategy, access, use and management).

Description of tasks:

- a) Name a person in the NSO to act as the Administrative Data Coordinator, whose role will be to:
 - a. *Validate the priorities of the NSO in terms of administrative data by liaising with all areas of the NSO and articulate a vision statement;*
 - b. *Lead the review and development of an enhanced governance framework, i.e., with respect to the legislation, policies, directives, organizational structure and centralized functions;*
 - c. *Identify and lead the development of tools to support the governance framework, such as templates for data acquisition agreements and a tool to register these, an administrative data inventory, templates to consult the NSO on needs for new administrative data sources and to assess such new sources, tool to grant access to administrative data once in the NSO; and*
 - d. *Communicate the above to the NSO staff and offer training.*
- b) Communicate the name and role of the person to the NSO staff.

Benefits:

- a) A person to determine and implement a strategy for enhancing the access, use and management of administrative data.
- b) One go-to person for all questions regarding the administrative strategy and data holdings of the NSO.
- c) Information kept centrally makes for easier knowledge retention.

2. Review of legislation

An important second step can be tasking someone in the NSO to provide the legal services function. This person or a small team will ensure that the NSO has the necessary legislation that favours access and use, as well as protects the data. This legal services function will also ensure that access and statistical use and management of the data is not impeded by other legislations of the country.

Description of tasks:

- a) Review the NSO's current statistical act in light of current and planned access to and use of administrative data.
- b) Determine if the NSO has the legal authority to access, use and manage (including protect) administrative data and document this authority.
- c) Determine if there are any barriers to access, use and management, and document these barriers.
- d) Review other acts if applicable (privacy, access to information) and document how they impact access, use and management of administrative data.
- e) Establish a strategy for the creation and/or modification of statistical act, if needed.

Ideal provisions in a statistical act:

To favor access:

- i. *Provide the NSO access to administrative data, including identifiers*
- ii. *Provide the NSO access to administrative data from public authorities free of charge*
- iii. *The NSO's right of access applies notwithstanding other enactments, other than those pertaining to the protection of public order or the security of the state*
- iv. *Provide the NSO the right to use the data to produce and support any official statistics, including the right to match data from different sources*
- v. *Compel the public authorities to consult with the NSO when introducing, revising or extending an administrative data system*

To enforce use:

- vi. *Instruct the NSO to first use administrative data when possible*
- vii. *Mandate the NSO to coordinate the national statistical system and to promote the avoidance of duplication in data collection by governments*

To protect data (part of the data management):

- viii. *Protection of the confidentiality of administrative data, i.e., the principle of "one-way traffic"*
- ix. *Privacy rights of individuals with respect to collection, use, disclosure, disposal, retention, accuracy and protection of their personal information*

Benefits:

- a) Favor access
- b) Enforce use
- c) Protect data

Tool:

The attached document provides an annotated version of the statistical model bill developed for the Caribbean Community Secretariat.¹



(right-click on icon and Open File)

¹ Participants to the Regional Seminar on Administrative Data can make use of the output of Exercise 1: Analysis of Country-Specific Legislation which provides a good starting point in the assessment of the strengths and limitations of the NSO's legislation in terms of access, use and management of administrative data.

3. Inventory of existing administrative data sources

Making a thorough inventory of the current sources being received and used by the NSO will provide the starting point for developing a strategy for enhancing the access, use and management of administrative data.

Description of tasks:

- a) Take an inventory of existing sources. The inventory should contain the following elements for each data source:
 - *Information about the data provider (e.g., ministry and contact name)*
 - *Information about the data source (e.g., name, topic, frequency of delivery, costs, record layout)*
 - *Existence of an agreement between the NSO and the data provider*
 - *Existence of quality assessment made on the data*
 - *Identification of the custodian of the data within the NSO (e.g., division, name of employee)*The inventory could also include which divisions and which employees have access and use the data.
- b) Make the inventory accessible and available to all NSO staff.
- c) Include in the inventory a list all sources that were explored and state all analyses that may have been made to evaluate the data source.

Benefits:

- a) Knowledge of what is received by the NSO
- b) Allows connection to supporting documentation such as the agreement, the record layout, the list of those who have access
- c) Avoids duplication of efforts.

Tool:

An inventory can be as simple as an excel file such as this one.



(right-click on icon and Open File)

4. Policy and directives for use (access, privacy, security)

Setting out clearly, in the NSO's own set of policies and directives, any legislative requirements and policies and directives of the central government that pertain to access, use and management of administrative data by the NSO will ensure a common understanding of requirements and a consistent rule.

Description of tasks:

- a) Set up policy for use of administrative data for statistical purposes (high-level principles and requirements)
- b) Set up directives for access, use or management of administrative data (more detailed requirements)
- c) Share with staff
- d) Share with current and future stakeholders (data providers)

Benefits:

- a) Translate legislative requirements and central government policies and directives into requirements and responsibilities
- b) Communicate requirements and responsibilities to NSO managers and employees
- c) Include clear provisions for non-compliance
- d) Can be shared with data providers to show good practices of NSO

Tool:

As an example, the Statistics Canada policy on informing data users is provided at this link.

<http://www.statcan.gc.ca/eng/about/policy/info-user>

5. Strategic collaboration with data providers

Collaboration with data providers needs to be fostered through regular contacts and transparency of the use of their data in the statistical office.

Description of tasks:

- a) Prepare communication package for data providers, which may include
 - NSO enabling legislation (own and any others that may support access, use and protection of administrative data)
 - NSO policy and directives for access, use and management (especially protection) of administrative data
 - NSO intended use of specific data holdings
 - Benefit of using administrative data to produce statistics (concrete examples of user needs, reduced burden)
 - Any offers to improve the data holding (returning assessment of quality, statistical products, etc.)
 - Template of data acquisition agreement, including protection of the data
- b) Identify other stakeholders who may see the benefit of the NSO accessing and using the administrative data and get their support. Communicate this support to the data provider.
- c) Set up regular bilateral meeting (annual or semi-annual) with current partners to keep abreast of changes in administrative data collection (e.g., concepts, classification, processing). For each data provider or set of data providers:
 - Determine mandate/objective of meetings
 - Determine membership, frequency and dates
 - Determine generic agenda topics
 - Hold meeting
 - Document decisions in minutes
- d) Consider establishing a committee of heads of relevant departments that will discuss ways to value the use of administrative data for statistical purposes and provide leadership in their respective departments.

Benefits:

- a) Collaboration is key to increase the statistical use of administrative data in a NSO
- b) It can lead to better data and better methods and processes to use the data.
- c) The strategy is four fold:
 - Communicating clearly the NSO intention
 - Finding win-win strategies to secure access to quality data
 - Partnering with other interested stakeholders
 - Maintaining long term partnerships
- d) Bilateral committees with important data providers will prevent interruption in delivery, unplanned changes to the data, and ensure constant discussion to improve the data

6. Exploration of new administrative data sources and identification of administrative data custodian for all data sources entering the NSO

Before investing too much time and effort into obtaining a data source, a preliminary assessment should be conducted to explore its potential for the statistical programs.

Description of tasks:

- a) Start with a preliminary high level consultation on potential uses of the administrative data source in the NSO
 - Information about the data provider
 - Information about the data product
 - What are the potential uses for the data product throughout the NSO
- b) Assess quality (with no data at hand)
 - To determine if the administrative data can fulfill the needs of the NSO using metadata and information available from the data provider
 - To provide a decision point to investigate further and acquire the administrative data or at least a sample or test version of the data

Elements of quality to assess when no data at hand:

- **Institutional environment** (of the data provider) is the organizational or institutional factors that may have an impact on the provider's capacity to supply quality administrative data over time such as Legislative or regulatory terms of access and Reliability of the data provider or data source
- **Relevance** is the extent to which the administrative data meet the agency's statistical needs (Alignment with desired statistical population, variables,
- **Timeliness** is the period between the reference point to which the administrative data pertains and the date on which the information becomes available to the NSO
- **Accessibility** is the ease with which the NSO can use the administrative data
- **Interpretability** is the availability of information and metadata needed to interpret and use the administrative data appropriate
- **Coherence** is the degree to which the administrative data can be used with the statistical agency's other data within a broad analytical framework over time

- c) Decide if the administrative data source is worth pursuing
- d) Determine who is the administrative data custodian, i.e.,
 - who will liaise with the data provider to negotiate access to the data and put in place an agreement
 - who will conduct the assessment once the data is received
 - who will provide access to the administrative data to others in the NSO once fitness for use is confirmed
 - who will continue to liaise with the rest of the NSO and the data provider (e.g., to respond to questions, to ensure timely delivery of quality data)

Benefits:

- a) Good understanding of the potential uses of the administrative data source across the NSO
- b) Low effort assessment to determine if data source is worth pursuing and what specific data elements to request

Tools:



(right-click on icon and Open File)

7. Agreements

Agreements can be formal (such as a Memorandum of Understanding) or informal (email). They could be contracts with the private sector, when costs are involved. They provide written trace of the exchange of data between the data provider and NSO.

Description of tasks:

- a) Design new data acquisition agreements
- b) Sign data acquisition agreement with data provider
- c) Store signed data acquisition agreements in central repository at NSO

Benefits:

- a) Clear requirements on the acquisition, use and management (including protection and disposal) of the data source

Agreements:

At a minimum, any agreement should mention:

- *The NSO's legal authority to obtain the administrative data*
- *The NSO's intended uses of the administrative data, e.g.,:*
 - *"For statistical purposes only as legally permitted under the Statistics Act"*
 - *"For uses by the Census and Social Statistics Programs"*
 - *When the administrative data contain mainly tombstone information on individuals or businesses: "For uses that improve the quality and efficiency of the statistical processes leading to the production of statistical outputs in the Census and Social Statistics Programs"*
 - *Avoid statements that limit the use to a specific survey, division or type of uses!*
- *The NSO's legal requirement to protect the administrative data*
- *The NSO's legal requirements to protect the administrative data*

Tool:

Proposed template (annotated and unannotated) for an informal letter of agreement.



(right-click on icon and Open File)

Proposed template (annotated and unannotated) for a formal letter of agreement.



(right-click on icon and Open File)

8. Assessment of fitness for use

Once an agreement is in place and data has been received (either a partial data file or full data for a trial period), fitness for use must be assessed before deciding to use the data in the NSO statistical program and to request this data on a regular basis. The quality of the data will also dictate the types of uses that can be considered.

Description of tasks:

- a) Assess accuracy of the data source
 - Evaluate coverage
 - Evaluate nonresponse or reporting errors in the data
- b) Document quality assessment and store centrally (administrative data inventory)
- c) Recommend uses of the data based on its quality (high quality data warrant direct use whereas lower quality data can still supplement statistical programs through indirect use)
- d) Decide if the NSO should obtain this data on a regular basis and if so, meet with data provider to discuss a longer term agreement.

Benefits:

- a) Administrative data is used knowingly of its limitations
- b) It forms the basis to provide a quality statement or measure to data users for the product in which the administrative data was used
- c) Maximizing use of administrative data will result in lower costs, lower burden for respondents, higher quality of data and new, more frequent or more detailed data products

Tools:

See table below for the types of assessment that should be considered for each type of usage

Use	Description	Assessments
DIRECT	Stand-alone use of the administrative data source Direct computation of totals, means, percentiles, etc.	Coverage of the statistical population: Requires the object set in the administrative data source to align closely with the target statistical population (via full object set or a representative sample of it) Variables of interest: Requires attributes that match the concepts being measured and that are of sufficient accuracy
DIRECT	Substitution and supplementation for direct survey collection – for certain variables, for all or part of the sampled units	Identification variables: Require attributes to link the right element of the object set to the statistical population units for which substitution occurs Variables of interest: Require attributes that match the concepts being measured by the survey and that are of sufficient accuracy Variables of interest: Require attributes that match the concepts needed for the analysis and that are of sufficient accuracy
DIRECT	Production of analytical outputs – The administrative data are linked to other data to conduct analysis	Statistical population units: Require units in the other source (e.g., survey) to be found as elements in the object set provided by the administrative data source Identification variables: Require attributes to link the right element of the object set to the unit in the other source Variables of interest: Require attributes that match the concepts needed for the analysis and that are of sufficient accuracy
Use	Description	Assessments
INDIRECT	Creation and maintenance of survey frames or base statistical registers	Coverage of the statistical population: Requires the object set in the administrative data source to align closely with the target statistical population Identification variables: Require attributes that allow statistical population units to be uniquely identified Classification variables: Require attributes to partition the statistical population into subpopulations needed for survey design or collection Contact variables: Require attributes to contact survey respondents
INDIRECT	Construction of sampling designs – stratification, sample allocation, sample selection	Variables of interest: A sampling design can be improved by using attributes correlated to the survey variables to further stratify the survey frame (e.g., age for social surveys, size measure for business surveys) A sample can be more efficiently allocated by having measures of variability coming from attributes correlated to variables of interest (e.g., variability of income by stratum) Some sample selection methods (e.g., proportional to size) require attributes correlated to variables of interest Coverage of the statistical population: Requires the object set in the administrative data source to align closely with the target statistical population Classification variables: Requires attributes to partition the statistical population into strata of interest Identification variables: Require attributes to link the right elements of the object set to the survey frame
INDIRECT	Editing and imputation in surveys	Variables of interest: Collected survey data can be verified (“edited”) against attributes in the administrative data source, e.g.,: <ul style="list-style-type: none"> <i>A claimant of employment insurance (administrative source) should report to be looking for a job in labour force survey)</i> Data found in error or missing in a survey can be imputed directly or indirectly from attributes correlated to the variables of interest being imputed (or that explains the nonresponse), e.g.,: <ul style="list-style-type: none"> <i>Total number of employees reported in a business’ payroll remittances to the revenue department can assist the imputation of number of employees by category (e.g., full time, part time)</i>

		<p>Coverage of the sample: Requires units in the survey sample to be found as elements in the object set provided by the administrative data source</p> <p>Identification variables: Require attributes to link right elements of the object set to the survey sample</p>
INDIRECT	Estimation – population benchmarks, calibration estimation, etc. (also for non-response adjustments)	<p>Coverage of the statistical population: Requires object set in the administrative data source to align closely with the target statistical population so that reliable control totals (benchmarks) can be derived</p> <p>Variables of interest: Estimates can be strengthened by integrating in the estimator attribute(s) that are correlated to the survey variables (e.g., income for a survey on household expenditures)</p> <p>!Caution: Accuracy of attributes is important to avoid presence of outliers that distort estimation model</p> <p>Identification variables: Requires attributes to link the right elements of the object set to the survey sample in the case of calibration-type estimators</p>
INDIRECT	Data validation / confrontation	<p>Coverage of the statistical population: It must be understood how the object set in the administrative data source aligns with the target statistical population of the survey one wants to validate</p> <p>Variables of interest: It must be understood how attributes align with the variables of interest one wants to validate</p>

9. Incorporation of administrative data in programs

Once the data has been deemed fit for use, it can be incorporate in the statistical programs.

Description of tasks:

- a) The administrative data custodian establishes rules to receive, capture, format, standardize, code and validate the administrative data. Determine if these steps are done by a centralized function (administrative data operations) or by the administrative data custodian
- b) The administrative data custodian provides access to statistical programs in the NSO on the need-to-know principle
- c) Statistical programs incorporate the data
- d) Statistical programs provide an assessment of quality of administrative data for external users

Benefits:

Maximum use of administrative data will result in:

- a) Lower costs
- b) Lower burden for respondents
- c) Higher quality of data
- d) New, more frequent or more detailed data products

10. Management of use of administrative data

An important part of the management of administrative data is ensuring that the NSO is respecting policies and directives that apply to the administrative data.

Description of tasks:

- a) On regular basis (or when new questionnaires or new data comes available), re-assess use of administrative data to ensure that it complies to the conditions set forth in the agreement
- b) On a regular basis, assess if those with accountabilities, as set forth in the policies and directives are compliant with the requirements of these governing instruments – this may involve self-checks, audits, or other means to assess compliance.

Benefits:

- a) Ensures that staff is aware of conditions for using administrative data
- b) Enhances trust in NSO (for data providers and public)

Summary checklist

Step		Status	Priority	Required Actions	Timeline	Comment
1	Identification of Administrative Data Coordinator					
2	Review of legislation					
3	Inventory of existing administrative sources					
4	Policy and directives for use (access, privacy, security)					
5	Strategic collaboration with data providers					

6	Exploration of new administrative data sources and identification of administrative data custodian for all data sources entering the NSO					
7	Agreements					
8	Assessment of fitness for use					
9	Incorporation of administrative data in programs					
10	Management of use of administrative data					