



Regional Analytics
Laboratory

PRESENTS

February 2022

DATA TO DECISIONS: Data Collection

Written by Meghan Eibner and Jamie Ward



Foreward

Data to Decisions is an initiative of the Regional Analytics Laboratory (RAnLab) and is sponsored in part by Future Skills Centre.

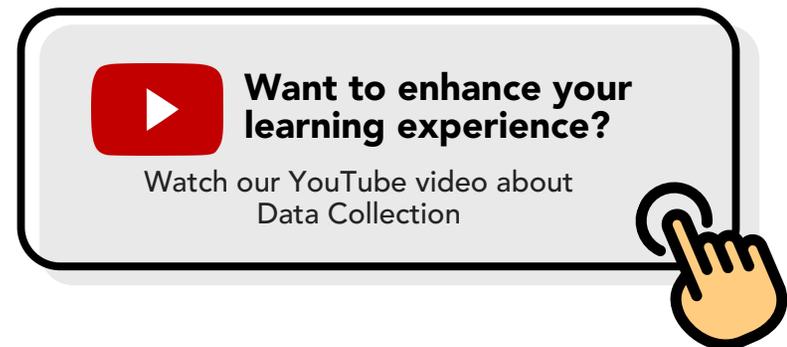
RAnLab is the data and analytics unit of Memorial University's Leslie Harris Centre of Regional Policy and Development. RAnLab analyzes data and geography to provide insight into, and projection modeling for, things like demographics, labour supply, service demands, commodity prices, and other socio-economic indicators. Some examples of their work include producing long-term community and regional population projections, assessing local housing demand, and providing detailed local data analysis and modelling to municipalities and regions—providing critical information for evidence-based decision-making.

The purpose of Data to Decisions is to help Canadians from varying backgrounds learn how to apply data to their own projects. Data to Decisions uses plain language, examples, and video presentations to enhance the learning experience.

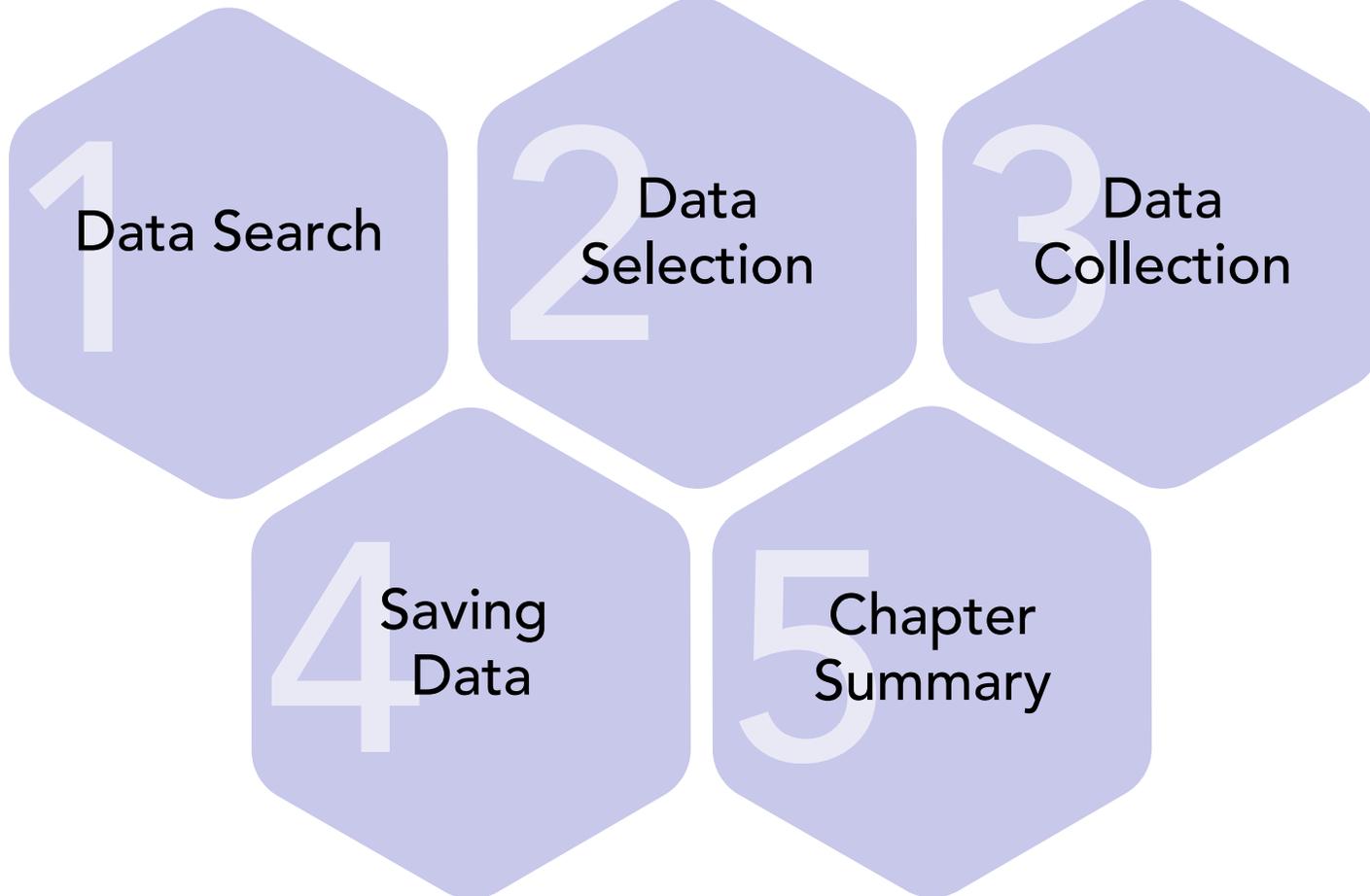
If you have any questions, please contact Meghan Eibner (meibner@mun.ca) or Jamie Ward (jward@mun.ca).

Throughout this chapter, you will see clickable buttons that link to webpages with additional information.

The button on the right will lead you to the Data to Decisions presentation on Data Collection.



Contents



Introduction

Data can be found from numerous sources in numerous formats. One of the most comprehensive and current databases is at Statistics Canada.

This guide will focus on how to find, select, collect, and save data from Statistics Canada.

Click here to visit the Statistics Canada data centre



Data Search

Data can be searched by topic, keyword, description, or table number. Filters for subject, geography, frequency, survey or statistical program, and content exist to help narrow down the search results.

Example: a search for "population" returns over 4,000 results.

Narrow the search by selecting some filters.

Geography: Province or territory. Frequency: Annual.

This reduces the search results to less than 200.

- Province or territory (179)
- Annual (179)

A description of each search result is presented to give more detail about the data.

Click the linked title of the dataset you want to select.

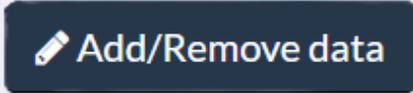
Example: select the first dataset titled "Population estimates on July 1st, by age and sex."

Data Selection

The Population estimates on July 1st, by age and sex data table automatically presents population data for:

- Canada
- 2016-2020
- Both sexes
- Age by 5-year groups

This data can be customized by clicking the "Add/Remove data" button. This allows the selection of different geographies, sexes, age groups, and reference periods. You can also customize the table layout by changing the data presented in rows and columns.

A dark blue button with a white pencil icon and the text "Add/Remove data" in white.

Example:

1. Click the "Add/Remove data" button
2. Under the geography tab, select "Canada" and "Newfoundland and Labrador"
3. Under the sex tab, select "Both Sexes"
4. Under the Age Group tab, select "All Ages", "Median Age", and "Average Age".
Note: The "All Ages" option has a + beside it. Click the "+" to reveal more options. For this example, deselect all the preselected age groups.
5. Under the Reference Period tab, select from 2010 to 2020.
6. Under the Customize Layout tab, set the "Geography", "Sex", and "Age Group" as Columns and set the "Reference Period" as Row.
7. Press the "Apply" button.

This returns a customized table presenting the selected population data for Canada and Newfoundland. The data table not only allows for comparative analysis between two geographies but also includes time series data to show how the selected indicators have changed over time.



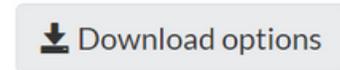
Want to see how to do this step by step?

Watch our YouTube video on Collecting Data

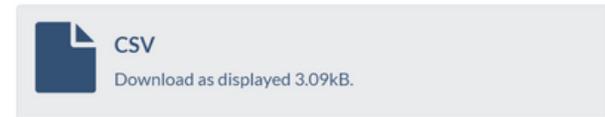


Data Collection

Once you have searched and selected the desired data, there are a couple ways of collecting it. The first is to simply highlight and copy the data table and paste it into another program (i.e. Microsoft Excel). The second option is to click the "Download Options" button above the data table. Several options for download as available.



- The simplest option will likely be the first ("CSV, download as displayed") which will download the selected data in the selected layout.
- The "CSV, download selected data" (for database loading) option will download the selected data without the customized layout. This format allows the selected data to be integrated into another database.
- There is also an option to download the entire "Population estimates on July 1, by age and sex" table which will download a zip file with the full dataset in an unformatted/uncustomized format.



Saving Data

Once the data has been copied or downloaded, some simple steps can be taken to help cite the data for future use. Every data table in Statistics Canada has an associated table number. This can be found below the title of the data table. When saving your data table in Excel, it is always a good idea to also reference the table number. This allows for quick retrieval of the data from the Statistics Canada website.

Example: The table number associated with "Population estimates on July 1st, by age and sex" is Table: 17-10-0005-01. This table number can be searched directly in the data search bar and return the corresponding data table.

Population estimates on July 1st, by age and sex^{1, 2, 3, 4}

Frequency: Annual

Table: 17-10-0005-01 (formerly CANSIM 051-0001)

Release date: 2020-09-29

Geography: Canada, Province or territory

Similarly, the URL can also be copied and saved with the data table in order to quickly and easily link back to the Statistics Canada data table. The URL for "Population estimates on July 1st, by age and sex" can be seen below.



The customization selected in your data table will not be displayed when following the link but it will take you directly to the original data table which can then be manipulated using the "Add/Remove Data" options.

Chapter Summary

Data to Decisions: Collecting Data

- Data on a wide range of topics can be collected from a multitude of different sources. Statistics Canada is one of the most comprehensive data sources for Canadian statistics.
- When collecting (downloading) data it is important to properly cite the data source. This can be done by noting the source website (including a link), the name of the data series, and the data table reference number.