

## Tutorial for downloading and analyzing data from the Atlantic Canada Opportunities Agency


The agency, which goes by the acronym ACOA, is one of many federal institutions that uploads data to the federal government's [open data site](#), which is a repository for hundreds of datasets, some useful for journalists, some not. "Open data" is a policy adopted by developing countries and a growing number of Third World Nations. The policy, which has been dubbed a "movement", contains two important principles for journalists: a government's commitment to sharing data should be a default position; and that open data is equated to open government where citizens possess the right to gain easy access to downloadable data, and even the right to demand datasets absent from the website.

The good news for journalists is that open data is not only a global phenomenon, but one that has grown roots closer to home at the provincial and municipal levels. To date, news outlets like the Toronto Star have used it to tell stories about [missed garbage complaints](#), and [suicide calls](#).

You'll find a list of government open data websites by clicking [here](#).








Before we shift to the Atlantic Opportunities Agency, a word about the datasets you'll encounter on these websites. The datasets in question are in what are called "machine readable format", meaning they can be downloaded and opened in a spreadsheet such as Excel. For the most part, the files are in one of three formats: csv (comma separated value, text file and an Excel). You'll learn more about these file formats on pages 43 to 45 in chapter three of our textbook, [Computer-Assisted Reporting: A Comprehensive Primer](#), which you'll find in Carleton's bookstore.

Now, let's shift to our tutorial. Go to the federal [open data site](#), and select the "Open Data" icon.



## Open Data

Search open data that is relevant to Canadians, learn how to work with datasets, and see what people have done with open data across the country.

Follow:       

### [Search Through our Open Data Portal](#)

Looking for data about Government of Canada services, financials, national demographic information or high resolution maps? Discover that and more through our Open Data portal, your one-stop shop for Government of Canada open datasets.

### [Working with Data and Application Programming Interfaces](#)

Learn more about how to use a variety of data formats, including Application Programming Interfaces (APIs).


### [Apps Gallery](#)

Browse and download apps that were built both by the Canadian public and the Government of Canada that use open data.

### [G8 Open Data Charter](#)

Review the [G8 Open Data Charter](#) and its [technical annex](#) which commit Canada and the other G8 nations to the proactive release of more information.

Search through our Open Data Portal.



### [Open Data 101](#)

Learn what open data is, how to get the most out of it, and explore the principles that are guiding the Government of Canada's Open Data initiatives.

### [Suggest a Dataset](#)

Can't find the dataset that you are looking for? Let us know!

### [Submit Your App](#)

Have you made an application using GC Open Data or Open Information? If so, let us know and we will share it with Canadians in our [Apps Gallery](#).

### [Canada's G8 Open Data Action Plan](#)

See Canada's plan to implement the principles laid out in the G8 Charter.

Select the first link to the top left. On the right-hand side, you'll notice a list of federal organizations, click on the "Show more" link to obtain a full list.

The screenshot displays the 'Open Data Portal' interface. At the top left, the title 'Open Data Portal' is visible. Below it, a search bar labeled 'Search Datasets' contains the text 'Search...'. To the right of the search bar is a 'Suggest a Dataset' button. Below the search bar, it indicates '244,631 datasets found' and an 'Order by' dropdown menu set to 'Last Modified'.

Three dataset cards are shown:

- Integrated pan-Arctic Snow Melt Onset from Satellite Measurements**: Description: 'The integrated pan-Arctic snow melt onset dataset was generated by combining active and passive microwave satellite derived estimates for northern high latitude land surface...'. Organization: Environment Canada. Resource Formats: **gif** (highlighted with a red box).
- Government of Canada Spend Data by Department**: Description: 'The Government of Canada Spend Data by Department dataset is a summary of procurement expenditure data at the Government of Canada level. It is derived from expenditure data...'. Organization: Public Works and Government Services Canada. Resource Formats: **CSV**, **XML**, **HTML** (all three highlighted with a red box).
- Facts & Figures 2014: Immigration Overview - Temporary Residents**: Description: 'Facts and Figures 2014: Immigration Overview - Temporary Residents presents the annual intake of permit holders and refugee claimants from 1994 to 2014. The report also shows...'. Organization: Citizenship and Immigration Canada. Resource Formats: **XLS**, **doc** (both highlighted with a red box).

On the right side, there is a 'Search Filters' section with a red border. It includes a dropdown menu for 'Organization' and a list of various Canadian government departments and agencies, each followed by a number in parentheses indicating the count of datasets. The list includes: Aboriginal Affairs and Northern Development Canada (2), Agriculture and Agri-Food Canada (1821), Atlantic Canada Opportunities Agency (2), Canada Border Services Agency (6), Canada Council for the Arts (1), Canada Economic Development for Quebec Regions (1), Canada Mortgage and Housing Corporation (31), Canada Revenue Agency (44), Canada Science and Technology Museum (1), Canadian Institute of Health Information (3), Canadian Radio-television and Telecommunications Commission (4), Canadian Space Agency (6), Citizenship and Immigration Canada (48), Correctional Service of Canada (3), Defence Research and Development Canada (1), Department of Finance Canada (130), Department of Justice (1), Elections Canada (5), Employment and Social Development Canada (71), Environment Canada (269), and Financial Transactions and

Please take a minute to notice the file formats we discussed above. In addition to the csv formats, the icons also indicate XML, HTML and doc. As long as you download a file with a “csv”, “txt”, “xls” (the older, pre-2007 version of Excel), or xlsx (any version from 2007 and on), you don’t have to worry about these other formats. However, it’s worth knowing that they exist.

The number beside each institution indicates the number of datasets. It’s also worth noting that this portal is a living document, meaning that departments continually add new data.

Select the Atlantic Canada Opportunities Agency.

**Open Data Portal**

**Search Datasets**

Search... [Suggest a Dataset](#)

2 datasets found Order by Last Modified

Atlantic Canada Opportunities Agency

**ACOA - Disclosure of Contracts Over \$10,000**

This dataset provides information on contracts issued by or on behalf of the Atlantic Canada Opportunities Agency. On March 23, 2004, the Government announced a new policy on...

Organization: Atlantic Canada Opportunities Agency

Resource Formats: [CSV](#)

**ACOA Project Information**

This dataset contains information about projects that have been approved by the Atlantic Canada Opportunities Agency since 1995. Note: When the Atlantic Canada Opportunities...

Organization: Atlantic Canada Opportunities Agency

Resource Formats: [CSV](#) [TXT](#)

You can also access this registry using the [API](#) (see [API Docs](#)).

**Search Filters**

**Organization**

[Atlantic Canada Opportunities Agency \(2\)](#)

**Data Type**

[Raw Data \(2\)](#)

**Tags**

[ACOA \(2\)](#)  
[Atlantic Canada \(2\)](#)  
[Atlantic Canada Opportunities Agency \(2\)](#)  
[proactive disclosure \(1\)](#)  
[procurement \(1\)](#)

[Show more](#)

**Subject**

[Economics and Industry \(1\)](#)  
[Government and Politics \(1\)](#)

Most, if not all, federal institutions disclose the contracts they award that are worth more than \$10,000. These, too, is datasets rich with story ideas about who's getting federal money and how much. There is also a lot of cash to follow in the second link, "ACOA Project Information". It contains data in the "csv" and "txt" format. While both can be opened in Excel, you're better off selecting the csv file.

Click on the "csv" download button, open your download folder, move the file to another folder and open it in Excel.





There are two ways to determine the number of rows or records in your table. Highlight a column to activate the number count feature on the border below the table. If a number is absent, click on the border to obtain a menu and select COUNT, which counts the number of rows in the table. Some versions of Excel allow you to select a number of these features. Others only allow one selection at a time.

The second way to determine the number of rows is to use the vertical scroll bar on your right to navigate to the bottom of the table and read the row number to the right. Again, after reading chapter three, and with a bit of practice, navigating a worksheet will become second nature.

This version of the dataset contains 30,770 records.

Navigate up and down: write the names of the column names, and describe the information they contain. As we mentioned in class, recording information about the data you've just downloaded is a good way to slow yourself down to find out what the data can tell you, what it can't, what's unclear and in need of follow-up. Also pay attention to whacky dates or other bits of information that appear to be mistakes. To use the old saying that has become cliché among data journalist, "all data is dirty". So assume that it contains mistakes.

Now let's sort the data fields to determine the dataset's age. There are three date columns. We determined that the best one to select was column O, the "Public Access Date".

O	P	Q
Public Access Date	Estimated Commencement Date	Estimated Completion Date
25/08/2015	01/08/2015	31/03/2016
25/08/2015	01/07/2015	31/12/2015
25/08/2015	01/06/2015	30/09/2015
25/08/2015	15/06/2015	31/03/2016
25/08/2015	01/09/2015	31/08/2016
25/08/2015	31/08/2015	29/02/2016
25/08/2015	01/06/2015	31/10/2015
24/08/2015	01/06/2015	31/03/2016
24/08/2015	01/07/2015	30/06/2016
24/08/2015	01/07/2015	31/03/2016
24/08/2015	01/08/2015	31/07/2016
24/08/2015	01/04/2015	31/03/2018
23/08/2015	01/04/2015	31/03/2016
23/08/2015	01/07/2015	31/12/2015
23/08/2015	31/07/2015	30/11/2016
22/08/2015	01/08/2015	31/12/2015

Since we downloaded this data on Sept. 2, the dataset is fairly recent, though it's worth asking how frequently it is updated. This is another important query to make before ever using data.

Now apply filter by clicking on the filter icon – the funnel -- to the far left located on the menu's "Data" section. You will get a dialog box that looks something like this. Select currency.

The screenshot shows an Excel spreadsheet with the following data in columns J through O:

J	K	L	M	N	O
istar	ACOA A	Total Government Fundir	Eligible Amou	Total Project Co	Public Access Da
conditic	6,484.00	6,484.00	12,967.00	12,967.00	25/08/2015
n-Repa	26,000.00	42,000.00	52,000.00	52,000.00	25/08/2015
n-Repa	93,760.00	187,520.00	201,367.00	210,520.00	25/08/2015
n-Repa	99,750.00	199,750.00	207,761.00	207,761.00	25/08/2015
nditiona	##				
n-Repa	70				
conditic	26				
n-Repa	50				
n-Repa	74				
n-Repa	##				
n-Repa	50				
n-Repa	##				
n-Repa	##				
n-Repa	##				
conditic	43				
n-Repa	##				
conditic	50				
n-Repa	##				
n-Repa	76				
n-Repa	##				
n-Repa	##				
n-Repa	85				
n-Repa	##				
n-Repa	##				
n-Repa	50				
conditic	##				
n-Repa	##				
conditic	##				
n-Repa	50				
conditic	#####	500,000.00	1,542,365.00	2,014,365.00	17/08/2015

The 'Format Cells' dialog box is open, showing the 'Number' tab. The 'Category' list on the left has 'Currency' selected. The 'Sample' text box displays 'Total Government Funding'. The 'Decimal places' is set to 0, and the 'Symbol' is 'None'. The 'Negative numbers' list shows four options: '-1,234', '1,234', '-1,234', and '-1,234'. The 'OK' and 'Cancel' buttons are at the bottom right.

L	M	N
Total Government Funding	Eligible Amount	Total Project Cost
\$6,484	\$12,967	\$12,967
\$42,000	\$52,000	\$52,000
\$187,520	\$201,367	\$210,520
\$199,750	\$297,761	\$297,761
\$155,175	\$206,900	\$206,900
\$90,500	\$94,000	\$94,000
\$44,358	\$52,185	\$52,185
\$50,000	\$66,667	\$66,667
\$74,217	\$98,956	\$101,056
\$106,522	\$257,462	\$257,462
\$50,000	\$66,667	\$66,667
\$366,000	\$537,000	\$537,000
\$631,152	\$640,818	\$640,818
\$43,500	\$87,000	\$87,000
\$182,069	\$216,399	\$216,399
\$68,375	\$100,000	\$100,000
\$200,000	\$266,666	\$266,666
\$91,153	\$106,153	\$106,153
\$572,052	\$572,052	\$572,052
\$805,000	\$600,000	\$1,861,500
\$164,740	\$190,544	\$190,544
\$1,105,352	\$1,381,692	\$1,381,692
\$410,100	\$614,000	\$614,000
\$50,000	\$100,000	\$100,000
\$722,500	\$965,000	\$1,400,000
\$217,950	\$424,950	\$424,950
\$724,250	\$729,500	\$825,000
\$50,000	\$70,000	\$70,000
\$500,000	\$1,542,365	\$2,014,365
\$490,000	\$727,604	\$727,604
\$50,000	\$66,666	\$66,666

This is a rich dataset that contains many options for filtering. We chose column J, “Assistance Type”.



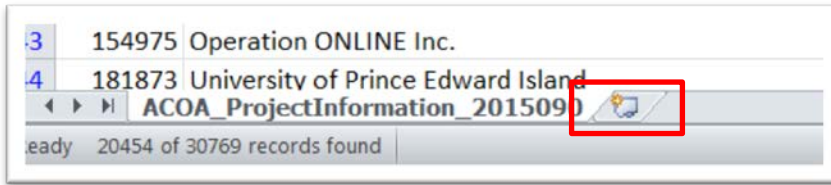
Connections			Sort & Filter			Data Tools		Outline	
Assistance Type									
D	E	F	G	H	I	J	K	L	M
Client C	Client P	Project	Project	Project	Program	Assistar	ACOA A	Total Government Fundin	Eligible Amou
Conceptio	A1W 3J1	Purchase	Sort A to Z					\$6,484	\$12,967
Sydney	B1P 1C6	Hire a co	Sort Z to A					\$42,000	\$52,000
Miramichi	E1V 2N5	Upgrade	Sort by Color					\$187,520	\$201,367
St. John's	A1E 1W6	Undertak	Clear Filter From "Assistance Type"					\$199,750	\$297,761
Saint John	E2L 4V1	Impleme	Filter by Color					\$155,175	\$206,900
Dartmouth	B3B 1T5	Develop	Text Filters					\$90,500	\$94,000
Elmsdale	COB 1K0	Purchase	Search					\$44,358	\$52,185
Halifax	B3H 0A8	Hire a Ch	<input checked="" type="checkbox"/> (Select All)					\$50,000	\$66,667
St. John's	A1C 5S7	Strategic	<input checked="" type="checkbox"/> Action Loan					\$74,217	\$98,956
St. John's	A1C 5S7	Build inte	<input checked="" type="checkbox"/> Conditionally Repayable Contribution					\$106,522	\$257,462
Rocky Har	A0K 4N0	Hire a cu	<input checked="" type="checkbox"/> Grant					\$50,000	\$66,667
Charlottet	C1A 7N7	Support t	<input checked="" type="checkbox"/> Interest Contribution/Buydown					\$366,000	\$537,000
Mulgrave	B0E 2G0	Provide s	<input checked="" type="checkbox"/> Loan Insurance/Guarantee					\$631,152	\$640,818
Grand-Sau	E3Y 1B5	Equipme	<input checked="" type="checkbox"/> Non-Repayable Contribution					\$43,500	\$87,000
Annapolis	B0S 1A0	Undertak	<input checked="" type="checkbox"/> Unconditionally Repayable Contribution					\$182,069	\$216,399
Saint John	E2L 1H3	Purchase						\$68,375	\$100,000
Halifax	B3J 3S9	Impleme						\$200,000	\$266,666
Fredericto	E3B 5H1	Attend th						\$91,153	\$106,153
St. John's	A1A 1W7	Support I						\$572,052	\$572,052
Trinity, Tri	A0C 2S0	Produce						\$805,000	\$600,000
Appleton	A0G 2K0	Develop						\$164,740	\$190,544
Gander	A1V 1G7	Enhance						1,105,352	\$1,381,692
Halifax	B3J 1E7	Fund star						\$410,100	\$614,000
Bedford	B4B 0L9	Pursue n						\$50,000	\$100,000
St-Prosper	G0M 1Y0	Upgrade a						\$722,500	\$965,000
Dartmouth	B2Y 4P5	Support se						\$217,950	\$424,950
St. John's	A1C 1C3	Purchase						\$724,350	\$720,500

We were interested in the “Non-Repayable Contribution” category. De-select by clicking on the “(Select All)”, and then select the one you want. You can also make many selections such as all the assistance types that appear to be grants that institutions don’t have to pay back.

Now sort column K in descending order to find out which institution received the largest loan that it didn’t have to repay.

A	B	C	D	E	F	G	H	I	J	K	L
Project	Client Name	Client A	Client C	Client P	Project	Project	Project	Program	Assistar	ACOA Assistance	Total
165558	CASH MANAGEMENT - EDA	Halifax	Halifax	In/a	Cash Manu	HALIFAX	1209021	Cash Manu	Non-Repa	28,100,000.00	
165969	Technology & Science Secretariat (NS)	Halifax	B3J 3C8	B3J 3C8	IEI - Inform	NOVA SCO	1200000	Technolog	Non-Repa	23,520,631.00	
803905	Sydney Ports Corporation Inc.	Sydney	B1 Sydney	B1P 6H2	Dredging c	SYDNEY	1217014	Enterprise	Non-Repa	19,000,000.00	
159684	SOURIS FOOD PARK DEVELOPMENT CORPORATION	Charlottet	Charlottet	C1A 7L9	Creation o	SOURIS	1101036	Eastern Ki	Non-Repa	17,000,000.00	
115012	Memorial University of Newfoundland	St. John's	St. John's	A1B 3K5	Pan Atlant	ST. JOHN'S	1001519	AIP - Atlan	Non-Repa	15,242,000.00	
181935	Memorial University of Newfoundland and Labrador	St. John's	St. John's	A1C 5S7	Inco Innov	ST. JOHN'S	1001519	AIP - Atlan	Non-Repa	13,113,492.00	
171223	Minister of Finance-N.B.	Fredericto	Fredericto	E3B 5H1	Base Clos	MIRAMICHI	1309050	Base Clos	Non-Repa	10,347,500.00	
203544	Forest Protection Limited	Lincoln	E3 Lincoln	E3B 7E6	Early Inter	FREDERICT	1310032	Atlantic In	Non-Repa	10,000,000.00	
194460	Regional Development Corporation	Fredericto	Fredericto	E3B 5H4	Francopho	FREDERICT	1310032	Business D	Non-Repa	10,000,000.00	
190260	Regional Development Corporation	Fredericto	Fredericto	E3B 5H4	Provide wr	FREDERICT	1310032	Business D	Non-Repa	10,000,000.00	
185598	Memorial University of Newfoundland and Labrador	St. John's	St. John's	A1C 5S7	Inco Innov	ST. JOHN'S	1001519	Business D	Non-Repa	10,000,000.00	
201276	Tourism Industry Association of Prince Edward Island	Charlottet	Charlottet	C1A 7N7	2012-2015	ATLANTIC	1400000	BDP - Tour	Non-Repa	9,975,000.00	
194701	Tourism Industry Association of Prince Edward Island	Charlottet	Charlottet	C1A 7N7	2009 - 201	ATLANTIC	1400000	BDP - Tour	Non-Repa	9,975,000.00	
190023	Tourism Industry Association of Prince Edward Island	Charlottet	Charlottet	C1A 7N7	2006 - 200	CHARLOTT	1102022	BDP - Tour	Non-Repa	9,975,000.00	
197160	The German Marshall Fund of the United States	Washington	Washington	20009	Deliver m	HALIFAX	1209034	Business D	Non-Repa	9,960,000.00	
203545	Condo Trail Development Authority Inc.	Gr. Martin	Gr. Martin	E3B 1B7	Franchisat	CT. MARTIN	1301003	IFC - Grate	Non-Repa	9,800,000.00	

Of course, we could filter the data in a number of ways. If you are happy with this subset of the data, select and copy it, create a new work sheet by clicking on the blank worksheet at the bottom of the workbook.



You can call it something like “Non-repayable\_loans”. You’ll notice that names in databases are devoid of spaces for reasons we’ll explore later. If you want to separate words in a title, use an underscore, which Excel reads as a character. As well, keep the titles concise because you can only use a certain number characters when naming worksheets.

Return to our original worksheet, and get rid of the squiggly lines that border the table by scrolling to the far right, activating a cell outside the table and tapping your space bar.

Brace yourselves, now it’s time for a bit of math: Subtraction.

If you can remember, create two new columns, one that calculates the number of days between the “Estimated Commencement Date” and “Estimated Completion Date” columns. Remember, every calculation in Excel begins with an “=” sign. Copy the formula to the bottom of the table. Pages 84 to 86 of the textbook covers working with dates.

And, finally, create a new column which will calculate the number of years by dividing by 365. Copy the formula to the bottom of the table and filter out the error messages.

Questions to answer when working with specific datasets from open-data websites:

1. How often is the data set updated?
2. How many records or rows does it contain?
3. How does the institution use the data?
4. Who inputs the data?
5. What time period does it cover?
6. What information does it exclude and is it possible to obtain that information in the spirit of open data = open government?
7. What stories have been done using this data?
8. Is there at least one story idea in this dataset?

#### **To-do list:**

Now that you’ve mastered the steps outlined above. Now it’s time to look for story ideas beyond the ones we discussed in class.

1. Continue to navigate the dataset, perusing information in the columns we did not discuss in class.
2. Apply the filter once again to narrow the selections in the columns that were not part of the tutorial above.
3. Copy and paste subsets of the data in two more worksheets and think about possible stories the data could tell.
4. Be prepared to discuss the ideas in class.