

ECO3148 A

Assignment 1

K. Day

Fall 2015

Your task for this assignment is to retrieve relevant data and prepare a brief report on some aspect of regional or urban disparities. I recommend using Excel to prepare graphs and do calculations because it integrates well with MS Word, but you can use any software that gets the job done. Your written report should be about 4-6 pages in length, **double-spaced**, excluding tables, figures, and references.

Due dates:

- **September 28, 2015, 11:30 am:** Team and topic. A maximum of four teams will be permitted to work on any given topic. I encourage you to clear your topic with me as soon as possible rather than waiting till September 28th, because popular topics may no longer be available by then. If you fail to inform me of your team and topic by the end of the day on September 28th, a penalty of 5 marks out of 100 will be assessed against your team's mark.
- **October 8, 2015, 1:00 pm:** Hand in a printed copy of your report at the beginning of class. You should also submit an electronic copy in case your printed copy is in black and white, in which case graphs may be difficult to read, and an electronic copy of the Excel file you used to do calculations and prepare graphs. If you did not use Excel that is okay, as long as the TA and I can read your data manipulation file. Submission of electronic files can be done through the course website. Only one member of each team should submit each file.

Teams

This assignment is to be a team project. When employers today check references, one of the things they regularly inquire about is whether or not the candidate is able to function well as part of a team. Doing assignments in groups will thus help to prepare you for the working world. For this assignment, teams should consist of two individuals. You are free to form your own teams; if you do not know too many other students in the class, post a message on the Discussion Board on the course website asking who else is still looking for a partner. If there is a particular topic that interests you, you can post a message asking who else is interested in that topic. Each team must provide me with a list of members and their topic by the September 27th deadline.

What to do

Each team should begin by selecting a topic, either from the list below or a topic of your own choice that has been approved by me. No more than 4 teams will be permitted to work on the same topic, so the sooner you let me know which topic you would like to pursue, the better. Note that collaboration between different teams working on the same topic is not permitted. Once you have chosen a topic, you will need to locate and download the necessary data to answer the question that needs to be addressed. You should retrieve data for as long a time period as possible. Each report should include the following elements:

- Summary statistics measuring the *magnitude* of regional or urban disparities (e.g., SD, CV, MAD, etc.)
- Graphs or tables illustrating the *pattern* of regional or urban disparities
- A clear explanation of all formulas used and calculations carried out (i.e., state which measures were used and provide mathematical formulas)
- A discussion of your findings
- A complete list of all data sources, including date of access (see *Data Documentation* below)
- A complete list of all references consulted, if any.

Note that in your discussion, you are welcome to compare your results to those of published studies, if you can find any, but it is not necessary to do so. Your discussion should focus on your own findings and the answer to the question you are addressing.

As far as the format of your report is concerned, please double-space and use a 12 point font. Please indent the first line of each paragraph by 0.5 inches (12.5 cm), and do not insert any extra blank lines between paragraphs. Page margins should be 1 inch (2.54 cm) on all sides. A title page with your names and student numbers would probably be a good idea. If your report includes a list of references, please use the format provided in the [Writing and Style Guide for University Papers and Assignments.pdf](#). However, individual entries should be single-spaced, not double-spaced as recommended in the guide, with a hanging indent and a blank line between each entry. Please do not forget to number the pages of your report. Finally, it is okay if your printed copy is double-sided.

Data sources

Since most reports will likely be dealing with Canada, most of you will probably need Canadian data. The principal producer of Canadian data is Statistics Canada. The following Statistics Canada resources are likely to be particularly useful:

CANSIM: Statistics Canada's primary database of socio-economic data. Most of the data in CANSIM are times series data, but some CANSIM tables contain information from recent censuses and other surveys. It is updated frequently. CANSIM can be accessed directly through Statistics Canada's website. When downloading data, I recommend you use the .csv format that is readable by Excel and many other spreadsheet software programs.

Canadian Census: Census data is available at 5-year intervals from 1976-2006. Unfortunately the 2011 census collected only basic information about population characteristics rather than the detailed economic data that had been collected in previous years, so it does not provide much information about recent trends with respect to human capital characteristics or income. However, the Census Profiles are a useful source of information about sub-provincial and urban areas up until 2006. For 2011, one can use data from the National Household Survey, but you should do so with some caution because no one knows exactly how comparable data from this voluntary survey are to data from past censuses. Census Profiles can be accessed through the Canadian Census Analyzer at CHASS, which is accessible through the Library website; the ODESI database, also accessible through the Library; and Statistics Canada's website. Some census data is also available in CANSIM.

While these are the main electronic data sources that you are likely to find useful, the following data sources may also be helpful, depending on the topic:

- the websites of government departments and agencies, both federal and provincial

- summary tables available on Statistics Canada's website
- Statistics Canada print publications, all of which are available in the Morisset Library. Early data for some variables may only be available in print.
- the Conference Board of Canada's e-library and e-Data, which contains a variety of provincial indicators (available through the Library website)
- other information that may be available via the library's research guide for economics, which has more information on where to find economic data:
<http://uottawa.ca.libguides.com/Economics-en>

If you are having trouble locating data on your chosen topic, ask the reference librarians in the Morisset Library for help. I may also be able to help you. If you are looking for data for other countries I will probably be of less help; for European data the Eurostat or OECD websites may be helpful. For U.S. and international data, the *Resources for Economists on the Internet* web page (<http://rfe.org/>) is a good place to start. Again, through the library website you may be able to access certain U.S. and international internet data sources that are not free of charge to the general public.

Documentation of Data

Your report must include detailed documentation of your data sources. Documentation of data sources is often a problem in economics, despite the fact that conclusions can be altered dramatically by data-related mistakes.¹ You must provide enough information that someone else could easily retrieve **exactly** the same data that you did. For example, for CANSIM data you should report the table number and/or the data series numbers, and the name of the series retrieved; for print publications you should report the table, page number, date of publication, etc. For electronic data, it is important that you report the date the data were retrieved, or the date the data were last updated by the statistical agency (the date access is always provided by Statistics Canada when data are downloaded from its website).² You can include this information on the final page(s) of your report. Do not forget to describe any data manipulation you had to do before calculating measures of regional disparities. See the "How to cite" link on the library's Economics research guide page for more advice on how to cite data sources.

Data Retrieval Tips

As previously noted, when you retrieve data from an electronic source I recommend that you choose a comma-delimited (.csv) format if possible. This format is easily readable by all spreadsheet packages. Graphs can be copied and pasted into your word processing software. I also recommend that if you are retrieving multiple variables, you retrieve each variable in a separate file. For example, if downloading data on the unemployment rate for men, I would set up my file with provinces or urban areas as columns and years as rows; I would then download data on the unemployment rate for women in a separate file. The different .csv files can then be loaded into different pages of the same spreadsheet, and calculations can be performed more efficiently. Otherwise the pages of your spreadsheet will contain a mixture of variables for different regions, and likely be more difficult to work with.

¹ For a discussion of some of the problems resulting from poor data documentation, see Dewald, W., J. Thursby, and R. Anderson. "Replication in Empirical Economics: The *Journal of Money, Credit, and Banking* Project." *American Economic Review* 76 (1986): 587-603.

² When you retrieve data from CANSIM, you will find that a citation for the data source is usually included at the bottom of the page.

As noted in class, while spreadsheet packages such as Excel include functions to compute the unweighted sample mean and sample standard deviation, they generally do not include functions to compute weighted measures, so you will have to program the calculation yourself. I will post at least one of my Excel files on the course website as an example for you to look at.

Another thing you should bear in mind when retrieving data is that when agencies such as Statistics Canada introduce major changes in their data collection methods, breaks are introduced into data series. For example, in the mid-nineties Statistics Canada introduced a major revision to the National Accounts and Provincial Economic Accounts, but only revised the provincial data backward to 1981. While the old data are still available in CANSIM in matrices that say "Terminated," they are not always directly comparable, in that if you had overlapping observations they would not be the same. This does not mean that you cannot use the old series; however, you should not simply splice the old and new series together. You can either present separate lines for different time periods, or make some adjustment to the old data to make them more comparable to the new data before splicing the old and new series together. One type of adjustment procedure that is quite commonly applied when overlapping observations are available is to multiply the old series by the ratio of the new to the old value for the overlapping observation, so that growth rates are preserved but the levels of the variable are adjusted. Do not forget to properly document all your calculations.

Finally, do not forget to report dollar values in **real** terms. This means that if you are working with income or wage data, you may also have to download a price series in order to deflate the data, unless constant dollar measures are directly available. Remember that it is easy to rebase a price index if your real values do not all have the same base year.

Possible Topics

1. How does the Canadian tax-transfer system affect regional income disparities in Canada? (i.e., how do regional disparities in after-tax incomes compare to regional disparities in before-tax incomes?)
2. How have regional disparities in human capital in Canada evolved over time? (Human capital is usually measured in terms of level of education and/or amount of work experience.)
3. Are there regional disparities in health outcomes within Canada? (There are many different measures of health outcomes one could look at here.)
4. How big are income disparities between urban areas within the province of Ontario, and how do they compare in magnitude to interprovincial income disparities? (Look for data at the CMA or census agglomeration level.)
5. How big are income disparities between urban areas within the province of Quebec, and how do they compare in magnitude to interprovincial income disparities? (Look for data at the CMA or census agglomeration level.)
6. How important are rural-urban differences in unemployment rates in Canada? (This topic is likely to require comparisons between provinces.)
7. How important are rural-urban differences in incomes in Canada? (This topic is likely to require comparisons between provinces.)
8. How much do regional differences in the industrial composition of employment affect regional disparities in wage rates? (i.e., if the employment share of each industry was the same in each province, would regional wage disparities be larger or smaller? I recommend you work with broad industry categories.)
9. How much do regional differences in the occupational composition of employment affect regional disparities in wage rates? (i.e., if the employment share of each occupation was the same in each province, would regional wage disparities be larger or smaller? I recommend you work with broad occupational groups.)
10. How important are differences in crime rates across urban areas in Canada? (You may choose which types of crimes to examine.)
11. Are there provincial differences in the degree of income inequality? (I suggest you look for data on Gini coefficients.)
12. How do poverty rates vary across urban areas in Canada? (I suggest you look for data on "Persons in low income.")
13. How would regional disparities in unemployment rates in Canada differ if the age distribution of the population was the same in each province? (i.e., if the population share of each age group was the same in each province, would regional unemployment rate disparities be larger or smaller?)
14. How does the cost of living differ across Canadian cities? (For this topic you need to find an inter-city price index. You should probably also look at more than one category of spending.)
15. Some other topic of your choice that has been approved by me.