

SOCI 612 - Selected Statistical Sources

NOTE: This handout will be available online on the DLI page:

<http://library.concordia.ca/about/partnerships/dli.html>

Data Liberation Initiative (DLI)

Microdata

Consists of the data directly observed or collected from a specific unit of observation (individual cases, answers from one individual, from one respondent). One needs to use statistical analysis software (SPSS, SAS, Stata) to make sense of these datasets.

Equinox (<http://clues.concordia.ca/record=e1000885~S0>)

Equinox is the **self-extraction** tool that provides users with microdata sets. Users can retrieve entire files (Public Use Microdata Files –PUMF) or select individual variables. Equinox also provides complete documentation (user guides, codebooks) as well as syntax files for SPSS, SAS and Stata. The *advanced search* mode allows to search for keywords in **file** name or description as well as in description, name and label of **individual variables**.

DLI Website

(<http://0-www.statcan.gc.ca/mercury.concordia.ca/dli-ild/dli-idd-eng.htm>)

Information on datasets available from DLI as well as release dates for new products. The survey documentation (user guides, codebooks, etc) is publicly available but **data can only be downloaded by the DLI contact**.

ICPSR (U.S. and international microdata)

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/access/index.jsp>

ICPSR is a statistical data repository covering fields such as sociology, political science, economics, demography, education, child care, health care, crime, minority populations, aging, terrorism, substance abuse, mental health, public policy, and international relations. The collection focuses on U.S. issues but also includes several datasets that deal with international and Canadian topics. **You need to create a [MyData](#) account to download datasets.**

Main Sources of Aggregated Data

Census of Canada

(<http://www12.statcan.ca/census-recensement/index-eng.cfm>)

All of the release components are now available: age & sex, education, ethnic origin, families & households, etc. The microdata (individual file, PUMF) is also available.

Data Products

Profiles present all variables by release components at a given geographic level (Provinces, Census Metropolitan Areas, Census Subdivisions, Census Tracts, etc). There is one *cumulative profile* that includes all the release components.

- [Community Profiles](#) and [Census Tract Profiles](#)
Census profiles presented at the community (census subdivision) and census tract levels. The data is available through a dynamic web interface.

Topic-Based Tabulations: present *selected* variables on a given topic. *Variables are cross-tabulated.*

CANSIM @ CHASS

(<http://clues.concordia.ca/record=e1000249~S0>)

CANSIM is the largest Statistics Canada's database available to educational institutions. The CHASS interface is a good starting point to retrieve data tables on all subject-areas. It is **updated daily** and contains more time series than E-Stat. Data can be downloaded in various formats (text, csv, excel, SPSS, SAS). For more graphical output options, use the [E-Stat interface](#).

Institut de la Statistique du Québec (ISQ)

(<http://www.stat.gouv.qc.ca/default.htm>)

A lot of the data is derived from Statistics Canada, but original data is produced especially for areas of provincial responsibility like culture and health. The main ISQ databases is the [Banque de données des statistiques officielles sur le Québec](#).

For more information, please contact me at: alex.quindon@concordia.ca