

# SOCI 424 — Statistics and Microdata

## Research in the Social Sciences

### Microdata

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

### Statistics Canada

Odesi (<http://clues.concordia.ca/record=e1001350> )

Odesi is the self-extraction tool that provides users with microdata files from **Statistics Canada's** surveys and opinion polls from **Gallup, Léger Marketing, etc.** Users can retrieve entire files or select individual variables. Odesi also provides the complete documentation (user guides, codebooks). In *search* mode, choose Variable/Category Label from the drop down menu to search for keywords in survey questions and all the possible responses to these questions keywords as well as in the label of **individual variables**.

### Data Liberation Initiative (DLI) Website

(<http://0-www.statcan.gc.ca/mercury.concordia.ca/dli-ild/dli-idd-coll-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/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 an [account](#) to download datasets.**

### GESIS (Leibniz Institute for Social Sciences)

(<http://www.gesis.org/en/home/>)

Major European social sciences research institute. Datasets are available on a variety of socio-economic topics of German and European interest. Some datasets are publicly available, some are available on demand (after signing a license) and some are available at a cost.

## Main Sources of Aggregated Data

### Census Program (Census + National Household Survey)

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

#### Data Products

[NHS Profiles](#) present several variables from the NHS and the 2011 Census s at a given geographic level (Provinces, Census Metropolitan Areas, Census Subdivisions, Census Tracts, etc).

[Census of Population Program Datasets](#): powerful search tools that allows users to select precise datasets (topic based or profiles) based on census date, geographic level and variable selection.

#### Microdata

Public Use Microdata Files (PUMFs) are available for the National Household Survey and for pre-2011 censuses. You can download those from Odesi (see above).

### CANSIM

[CANSIM](#) on Statistics Canada's website is the largest Statistics Canada's database available to educational institutions. You can also use the [CHASS](#) interface which offers different search options. On CHASS, data can be downloaded in more formats (text, csv, excel, SPSS, SAS).

### Institut de la Statistique du Québec (ISQ)

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

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 database is the [Banque de données des statistiques officielles sur le Québec](#).

## Sources of International Statistics

**OECD iLibrary** (<http://clues.concordia.ca/record=e1000991~S0>)

Contains a very large number of publications and datasets released by OECD.

**UNData** (<http://data.un.org/>)

Free resource. Includes databases from several UN agencies. More data can be obtained for the individual agencies' websites.

**For more information, please contact me at:** [alex.quindon@concordia.ca](mailto:alex.quindon@concordia.ca)