Introduction to Canadian data and Odesi

SOC 3142
Susan Mowers
Data Librarian
• **Library services:**
  • Research and Data Support
  • SPSS labs and Libraries
  • Quality of data?

• **Data and Sociology / International Development**
  • Data options:
    • Canadian Survey data
    • Where can I get public use microdata?
      • Odesi
      • Course page
      • Videos
      • What is the must have data documentation?
  • Is there microdata with more detail?
  • Contacts
Services at your Library

Research and Data Support
Research Literature Reviews

Sociology!

Alain El Hofi

Appointments:  
aelhofi@uOttawa.ca  
613-562-5800 (7186)

Office:  
Room 2010B  
FSS Building

International Development!

Mish Boutet

Appointments:  
mboutet@uottawa.ca  
613-562-5800 (2721)

Office:  
Room 309C  
Morisset Library – GSG Centre

STAY TUNED!
Services at your Library

Research and **Data Support**
Data Support

- **Help you to find, access and understand** Statistics Canada data and other data to meet your information needs
- **Provide technical support** to make data formats analysis and visualization-ready
- To **provide** SPSSm Nvivo, an dGIS software (for stats, mapping, and qualitative data and visualization)
- Your contact with Statistics Canada etc. to ensure you get **timely data answers**

We are: **GeographicStatistical and Government Information Centre (GSG) / FSS**
Welcome to the **GSG Centre**
Morisset Library on the **3rd floor (309)**

And find Catie Sahadath at the FSS Library: FSS 2010!
Computer lab provides SPSS access
Morisset Library on the 3rd floor (308)
SPSS labs on campus

• **Morisset Library** – 65 University, room 308 **(all students, Mondays to Fridays 7:00 – 2:00 (a.m.), Sat/Sun. 10:00-2:00 (a.m.) & Exam period, 24/7)**

• **FSS Library** – 120 University, room 2010 FSS, **(all students, Mondays to Fridays: 7:30 a.m. to 10:00 p.m.)**

• **Vanier Hall** – 136 Jean-Jacques Lussier, rooms 2008, 2015, 2025 **(Faculty of Social Sciences, open 7:00 a.m. – 11:00 p.m.)**
FSS Library

- FSS, 2nd floor, Room 2010
- Librarian offices for Alain El Hofi, Catie Sahadath, Patrick Labelle
- Study space; and
- SPSS, nVivo and GIS software.
Quality of Data?

• Quality of data?
  • You can verify. What do the data collector’s say about their policies and data documentation ... about
    • their mandate,
    • their research and survey methodology and
    • who is sponsoring ($$) the data collection?

• This course focuses on Statistics Canada survey data
  • strong reputation!
    • Including transparent policies and documentation.

• Also reputable: ICPSR
  Inter-university Consortium of Political and Social Research
  • Largest archive of reputable social science survey data
  • Publish extensive documentation for every survey, e.g., from the U.S. General Social Survey.
Canadian Surveys - Data options
Access

Via the **Data Liberation Initiative (DLI)**

**Partnership between 75 + Canadian universities / colleges & Statistics Canada.** Funded and supported by your Library.

All uses permitted (academic, commercial ...) except cannot be redistributed outside the University of Ottawa

Find on **Odesi** (full and search modes)

- The Library also supports other types of (more detailed) data.
- Susan is your DLI Contact : [smowers@uottawa.ca](mailto:smowers@uottawa.ca) with Catie: [csahadat@uOttawa.ca](mailto:csahadat@uOttawa.ca)!
Canadian surveys

Example

What is Statistics Canada’s most important “survey”?

HINTS ...

- It’s “so” big, it can only happen every 5 years!
- It’s obligatory, again (the long version)
Census of Population 2016

- Portal to Census 2011 statistics also covers the voluntary NHS
  http://www12.statcan.gc.ca/census-recensement/index-eng.cfm

- Compare Census topics

<table>
<thead>
<tr>
<th>2011</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Age and sex</td>
<td>• Aboriginal peoples</td>
</tr>
<tr>
<td>• Families, households and marital status</td>
<td>• Age and sex</td>
</tr>
<tr>
<td>• Language</td>
<td>• Education</td>
</tr>
<tr>
<td>• Population and dwelling counts</td>
<td>• Ethnic origin and visible minorities</td>
</tr>
<tr>
<td>• Structural type of dwelling and collectives</td>
<td>• Families, households and marital status</td>
</tr>
<tr>
<td>Plus boundary mapping data</td>
<td>• Housing and shelter costs</td>
</tr>
<tr>
<td></td>
<td>• Immigration, citizenship and ethnocultural</td>
</tr>
<tr>
<td></td>
<td>diversity</td>
</tr>
<tr>
<td></td>
<td>• Income and earnings</td>
</tr>
<tr>
<td></td>
<td>• Labour</td>
</tr>
<tr>
<td></td>
<td>• Language</td>
</tr>
<tr>
<td></td>
<td>• Mobility and migration</td>
</tr>
<tr>
<td></td>
<td>• Place of work and journey to work</td>
</tr>
<tr>
<td></td>
<td>• Population and dwelling counts</td>
</tr>
<tr>
<td></td>
<td>• Structural type of dwelling and collectives</td>
</tr>
<tr>
<td></td>
<td>Plus boundary mapping data!</td>
</tr>
</tbody>
</table>
WHAT’S THE NEWS? From survey to data

Canadians answered the National Household Survey May through August 2011... to release of statistics

THE GLOBE AND MAIL

SEPTEMBER 11, 2013

The make-up of the (top-over $191,000) 1-per-cent club is overwhelmingly male, most-likely over the age of 45 and tends to work in management jobs. Canada’s top earners are also more likely to make money through investments and self-employment income.
Survey Microdata

- Already downloaded for you in SPSS format. DIY on Odesi
  - Individual respondent answers : Rows
  - Variables: Columns

<table>
<thead>
<tr>
<th>RECID</th>
<th>AGEGR10</th>
<th>SEX</th>
<th>MARSTAT</th>
<th>HSDSIZEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Respondents by rows

Information “Variables” by columns, e.g., Age group of respondent (groups of 10)

- Statistics Canada and Canadians depend on personal information privacy. Statistics Canada survey microdata is released while protecting privacy (next slide)
PUMFs

• **Public-use microdata files**
  
  • *Statistics Act*, Stats Can can’t divulge the identity of any survey respondents
  
  • “public-use” Anonymization necessary of responses with low case counts, e.g.,
    
    • Born in Morocco becomes → **Born in North Africa**
    • Haitian ethnic origin becomes → **Other Caribbean ethnic origins**

  • Statistics Canada publishes documentation for all PUMFs, including methodology
  
  • Free to the public
  
  • Researchers can use the microdata to analyse and group using statistical software e.g., SPSS
    
    • Hundreds of variables
    • Many thousands of cases
  
  • Geographic areas: usually = provinces/territories and fairly large sub-provincial areas like census metropolitan areas.

*S.C. 1970-71-72, c. 15, as am*
Data options

• Where can I get microdata?
Research Guide for Data

http://uottawa.ca.libguides.com/DataandStatistics-en - includes ODESI
Soc3142 data page

Get PUMFs and documentation at:

Must-have documentation

- More than the SPSS file
- Codebook – includes the literal survey questions ...
- Questionnaire
- User guide

- Example: General Social Survey, cycle 27, Social Identity
What is the GSS (General Social Survey)?

• From Definitions, Data sources and Methods

  • **Annual.** Established in 1985, telephone surveys from representative samples from the 10 provinces

  • Data collected over a 12 month period – population aged 15 years and older, living in private households

  • Note also, [http://www.issp.org/](http://www.issp.org/)
General Social Survey / Overview and Topics

Repeated cycles on key social issues usually every 5 years …

Appendix A
GSS topics according to series, year and cycle

<table>
<thead>
<tr>
<th>Topic</th>
<th>1st series</th>
<th>2nd series</th>
<th>3rd series</th>
<th>4th series</th>
<th>5th series</th>
<th>6th series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>1985 (1)</td>
<td>1991 (6)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Time use</td>
<td>1986 (2)</td>
<td>1992 (7)</td>
<td>1998 (12)</td>
<td>2005 (19)</td>
<td>2010 (24)</td>
<td>...</td>
</tr>
<tr>
<td>Victimization</td>
<td>1988 (3)</td>
<td>1993 (8)</td>
<td>1999 (13)</td>
<td>2004 (18)</td>
<td>2009 (23)</td>
<td>...</td>
</tr>
<tr>
<td>Education, work and retirement</td>
<td>1989 (4)</td>
<td>1994 (9)</td>
<td>2002 (16)</td>
<td>2007 (21)</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Access to and use of ICT</td>
<td>2000 (14)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Social engagement</td>
<td>2003 (17)</td>
<td>2008 (22)</td>
<td>2013 (27)</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Giving, volunteering, participating</td>
<td>19971</td>
<td>20001</td>
<td>20041</td>
<td>20071</td>
<td>20101</td>
<td>2013 (27)</td>
</tr>
</tbody>
</table>

.. not available for a specific reference period
1. Canada Survey of Giving, Volunteering and Participating (CSGVP)


Source: Statistics Canada, 2013, GSS: A New Data Overview
Open the *General Social Survey, Cycle 27, 2013: Social Identity*

1. Note variable ANS90D

   | ANS_90B | 1 | National symbols - Beaver | {1, Yes} |
   | ANS_90C | 1 | National symbols - Nature, geography, etc. | {1, Yes} |
   | ANS_90D | 1 | National symbols - Maple leaf | {1, Yes} |
   | ANS_90E | 1 | National symbols - Wildlife (excluding ... | {1, Yes} |
   | ANS_90F | 1 | National symbols - Canadian heritage | {1, Yes} |
   | ANS_90G | 1 | National symbols - Sports and leisure | {1, Yes} |
   | ANS_90H | 1 | National symbols - Multiculturalism | {1, Yes} |

2. Do you have enough information on what this variable is measuring as **Yes** versus **No**?
## Codebook (p. 69)

<table>
<thead>
<tr>
<th>Variable Name:</th>
<th>ANS_90D</th>
<th>Length: 1.0</th>
<th>Position: 217</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Name:</td>
<td>National symbols - Maple leaf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concept:</td>
<td>When you think of Canadian identity, what other symbol or image comes to mind? Maple leaf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question Text:</td>
<td>All respondents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source:</td>
<td>General Social Survey, Social Identity 2013</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Answer Categories</th>
<th>Code</th>
<th>Frequency</th>
<th>Weighted Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>1</td>
<td>4,088</td>
<td>4,042,404</td>
<td>13.9</td>
</tr>
<tr>
<td>No</td>
<td>2</td>
<td>17,375</td>
<td>18,632,661</td>
<td>64.0</td>
</tr>
<tr>
<td>Don’t know</td>
<td>7</td>
<td>3,734</td>
<td>3,747,970</td>
<td>12.9</td>
</tr>
<tr>
<td>Refusal</td>
<td>8</td>
<td>2,316</td>
<td>2,681,633</td>
<td>9.2</td>
</tr>
<tr>
<td>Not stated</td>
<td>9</td>
<td>21</td>
<td>27,246</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>27,534</td>
<td>29,131,913</td>
<td>100.0</td>
</tr>
</tbody>
</table>
• Confirm that your variable, maple leaf, is meaningful?
• If in doubt, should you add/use a stronger variable?
Questionnaire

Appreciation of national symbols (ANS)

ANS_Q10
When you think of Canadian identity, how important is:
the Canadian flag?
   Interviewer: Read categories to respondent.
   1. Very important
   2. Somewhat important
   3. Not very important
   4. Not at all important
   Don't know
   Refusal

ANS_Q30
(When you think of Canadian identity, how important is:)
   the Canadian Charter of Rights and Freedoms?
   1. Very important
   2. Somewhat important
   3. Not very important
   4. Not at all important
   Don't know
   Refusal

ANS_Q50
(When you think of Canadian identity, how important is:)
   the national anthem ("O Canada")?
   Interviewer: The Canadian Charter of Rights and Freedoms is a bill of rights entrenched in the Constitution of Canada. It forms the first part of the Constitution Act, 1982. The Charter guarantees certain political rights to Canadian citizens and civil rights of everyone in Canada from the policies and actions of all areas and level of government. It is designed to unify Canadians around a set of principles that embody those rights.
   1. Very important
   2. Somewhat important
   3. Not very important
   4. Not at all important
   Don't know
   Refusal

ANS_Q80
(When you think of Canadian identity, how important is:)
   hockey?
   1. Very important
   2. Somewhat important
   3. Not very important
   4. Not at all important
   Don't know
   Refusal

ANS_Q90
(When you think of Canadian identity, what other symbol or image comes to mind?)
   (80 spaces)
   Don't know
   Refusal

   e.g. maple leaf
User Guide

• Often but not always contains the codebook
• Key methodology including survey and sample design, weights etc.
You can get microdata after university for free

1. You can **e-mail** Statistics Canada and identify that you want the **public use microdata file (PUMF)** through the Open License (see [FAQ](#))

2. Validate what survey you want, e.g., 5024 - General Social Survey, Social Identity, 2013, cycle 27 as per: [Link](#)


How to access microdata → Access to Public Use Microdata Files (PUMFS) Collection → request by survey or subscription
Confidential data

• Why use confidential microdata?
  • Detail, e.g.,
    • Born in Morocco?
    • What is the respondent's actual age?

• Access via Research Data Centre (RDC)
  • Application process – within 4-5 weeks ....

• Access via Real Time Remote Access (RTRA)
  • Simpler application – within a week
Stay tuned for ...

- More! E.g., how to cite data and documentation (presentation)
For any assistance or more information about research and data, we’d love to hear from you!

- alain.elhofi@uottawa.ca Sociology
- csahadat@uottawa.ca Social Sciences Data Librarian
- mboutet@uottawa.ca International Development
- smowers@uottawa.ca Data Librarian

- Thank you!