



Chris Leowski
University of Toronto, Canada
Faculty of Arts & Science



National and International Economic and Social Data in Teaching and Learning

Repositories, Web Interfaces, Analytics

4th QS-Maple Abu Dhabi, UAE, May 2014

DATA IN TEACHING AND LEARNING

- Why should we teach students to use real and complex data?
- . . . and how should we do it?

DATA “LIFE CYCLE”

- **Data collection / compilation**
- **Data storage / repository**
- **Data search and retrieval**
- **Data presentation**
- **Data analysis**

DATA PRESENTATION

- **Time series view - an array of data points**
- **Simple multi-dimensional view**
- **OLAP cubes and hyper-cubes**
- **Frequencies and cross-tabs for survey data**

EXAMPLES OF WEB BASED DATA REPOSITORIES AND SEARCH AND RETRIEVAL SYSTEMS

University of Toronto CHASS Data Centre

Wharton Research Data Services (WRDS)

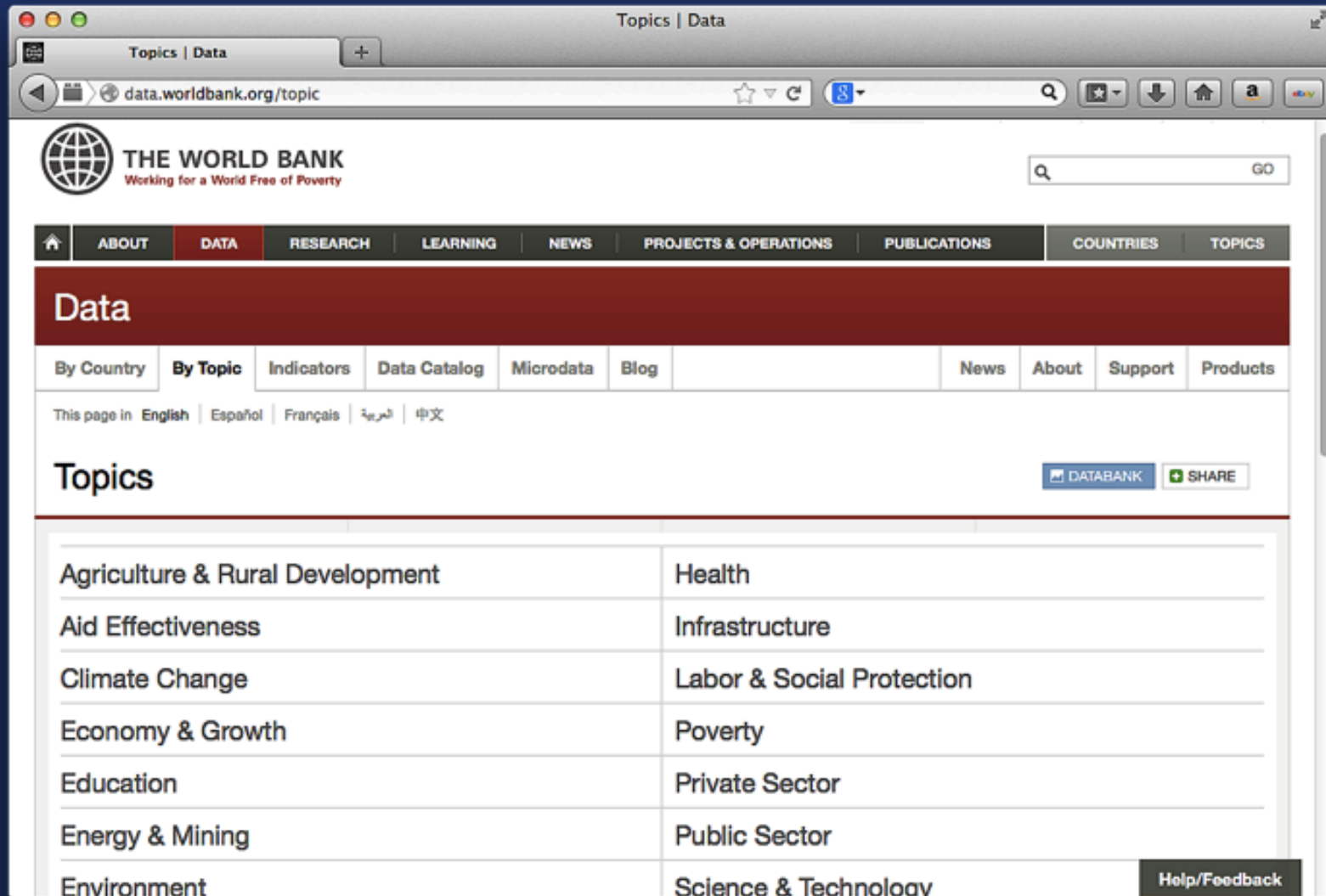
ODESI - Ontario, Canada

Organization for Economic Cooperation and Development

World Bank Data

WEB BASED DATA REPOSITORIES

THE WORLD BANK



The screenshot shows a web browser window displaying the World Bank Data website. The browser's address bar shows the URL `data.worldbank.org/topic`. The website header includes the World Bank logo and the tagline "Working for a World Free of Poverty". A navigation menu is visible with tabs for ABOUT, DATA (selected), RESEARCH, LEARNING, NEWS, PROJECTS & OPERATIONS, PUBLICATIONS, COUNTRIES, and TOPICS. Below the navigation menu, there is a search bar and a "GO" button. The main content area is titled "Data" and features a sub-menu with options: By Country, By Topic (selected), Indicators, Data Catalog, Microdata, Blog, News, About, Support, and Products. Below the sub-menu, there is a language selection bar with options: English, Español, Français, العربية, and 中文. The "Topics" section is displayed, showing a list of topics in two columns. The topics listed are: Agriculture & Rural Development, Aid Effectiveness, Climate Change, Economy & Growth, Education, Energy & Mining, Environment, Health, Infrastructure, Labor & Social Protection, Poverty, Private Sector, Public Sector, and Science & Technology. A "Help/Feedback" link is located in the bottom right corner of the page.

Agriculture & Rural Development	Health
Aid Effectiveness	Infrastructure
Climate Change	Labor & Social Protection
Economy & Growth	Poverty
Education	Private Sector
Energy & Mining	Public Sector
Environment	Science & Technology

WEB BASED DATA REPOSITORIES

OECD

The screenshot shows the OECD Statistics website interface. At the top, there is the OECD logo with the tagline "BETTER POLICIES FOR BETTER LIVES" and a navigation menu with options: OECD Home, About, Countries, Topics, Statistics, and Newsroom. A search bar and social media links are also present.

The main content area features the "OECD Data Lab" section, which includes a description: "OECD's laboratory for data visualisations. We welcome your feedback." Below this is a "Go to Data Lab" button and a visualization titled "Health data" with the text: "Compare life expectancy, health expenditure, alcohol consumption and smoking." The visualization shows a line graph with data points for the years 2007, 2012, 1992, and 199.

To the right of the Data Lab is a "Resources" section with buttons for "List of all statistics", "Glossary", "OECD Factbook Indicators", and "Tools for statisticians".

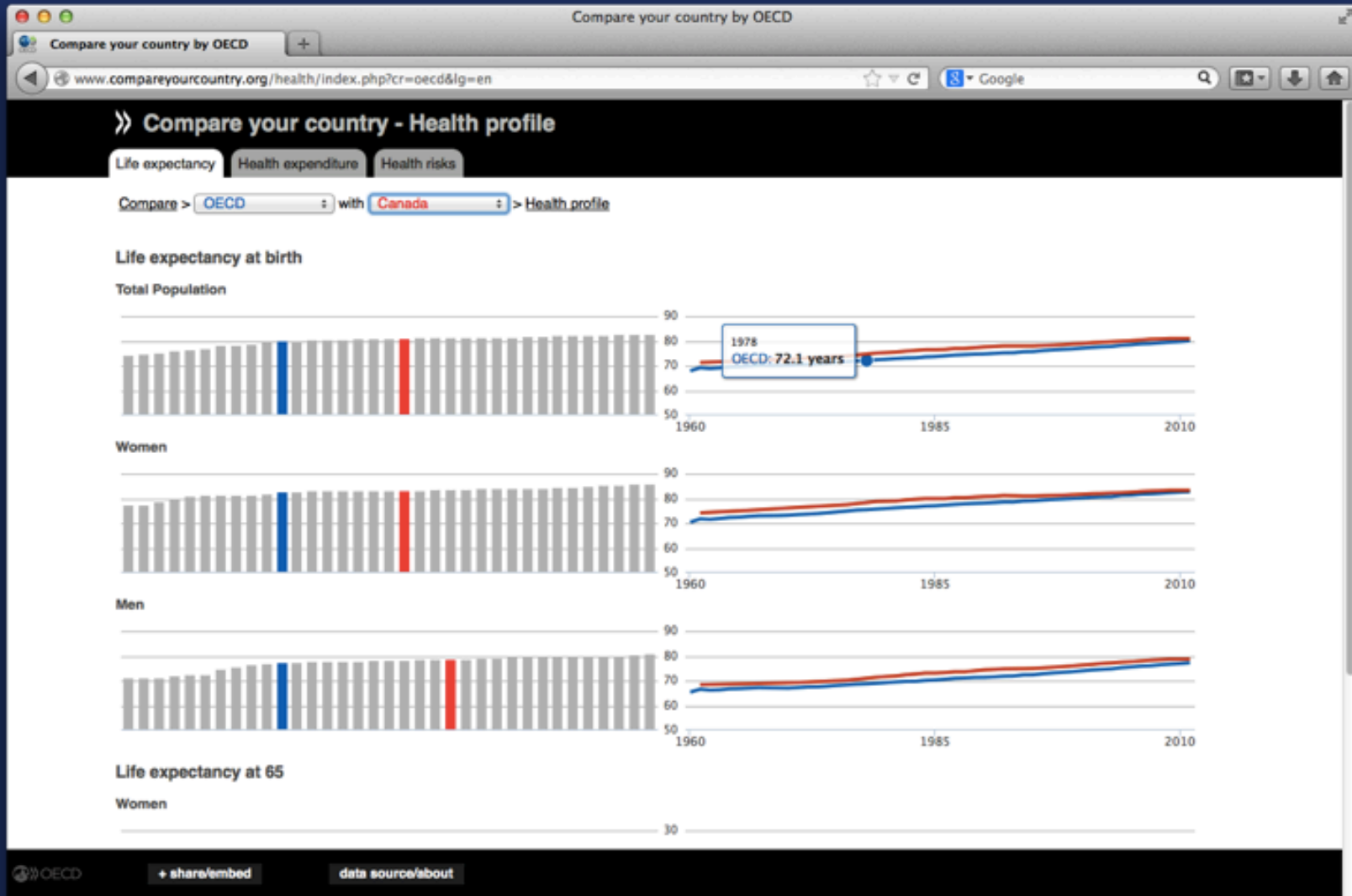
Below these sections is a "List of key indicators" section with a search bar and a table of indicators. The table shows the following data:

Keyword	Indicator	Table	Real-time dataset
Africa	African economies: basic indicators	Download Excel	
Agriculture	Consumer support estimates (subsidies)	Download Excel	Get real-time data
	Environment: agriculture	Download Excel	
	OECD Producer single commodity transfers	Download Excel	
	OECD Total support estimates	Download Excel	
	Producer support estimates (subsidies)	Download Excel	Get real-time data
	Wheat production	Download Excel	Get real-time data
	World cereal projections	Download Excel	

WEB BASED DATA REPOSITORIES

OECD

Life Expectancy (OECD vs. Canada)



4th QS-Maple Abu Dhabi, UAE, May 2014

WEB BASED DATA REPOSITORIES

<ODESI>

ODESI - Home

search1.odesi.ca

Google

Contact <odesi> support
Get data help at your institution

English | Français

About <odesi>

- Background
- Contributors
- <odesi> News
- <odesi> Repository (Nesstar)
- Contact Information

Accessing Data

- Who can Access it
- Terms of Use

Learning Tools

- Video Tutorials
- User Guide (In Progress)
- FAQ
- Instructional Resources

Related Resources

- <odesi> wiki
- Best Practices Document
- Other Data Portals & Networks
- Tools & Standards
- Protocols for Depositing Data

What is <odesi>?

Established in 2007, <odesi> (Ontario Data Documentation, Extraction Service and Infrastructure) is a digital repository for social science data, including polling data. It is a web-based data exploration, extraction and analysis tool that uses the [Data Documentation Initiative \(DDI\)](#) social science data standard.

<odesi> provides researchers the ability to search for survey questions (variables) across thousands of datasets. There are both microdata and aggregate data available, in a range of formats.

Search <odesi>

Search: in

Add Row >>

Refine Search

By Collection

- ODESI (Data Download Available)
 - Statistics Canada (Microdata)
 - Statistics Canada (Aggregate Data)
 - Public Opinion Polls
 - Other Canadian Data
- Additional Collections (No Direct Download Available)
 - CORA
 - ICPSR

By Date

New Data

[RSS Feed](#)

2014.02.24
Enquête de surveillance de l'usage du tabac au Canada, 2010 : Cycle 1, Fichier des ménages
[view survey](#)

2014.02.24
Enquête de surveillance de l'usage du tabac au Canada, 2010 : Annuel, Fichier des personnes
[view survey](#)

2014.02.24
Enquête de surveillance de l'usage du tabac au Canada, 2010 : Cycle 2, Fichier des ménages
[view survey](#)

2014.02.24
Enquête de surveillance de l'usage du tabac au Canada, 2010 : Annuel, Fichier des ménages
[view survey](#)

2014.02.24
Enquête de surveillance de l'usage du tabac au Canada,
[view survey](#)

WEB BASED DATA REPOSITORIES

WHARTON RESEARCH DATA SERVICES

The screenshot shows the Wharton Research Data Services (WRDS) website. The browser address bar displays "wrds-web.wharton.upenn.edu/wrds/". The page header includes the Wharton University of Pennsylvania logo, the "wrds" logo, and the text "WHARTON RESEARCH DATA SERVICES The Global Standard for Business Research". A navigation menu contains links for HOME, RESEARCH, SUPPORT, E-LEARNING, COMMUNITY, NEWS, ABOUT, and MyWRDS. A search bar is labeled "Search WRDS". A user is logged in as "Elenal" with a "Log Out" link. The main content area features a "Select a Data Set:" dropdown menu, a "Current Subscriptions" list, and an "Important WRDS Alert" dated Wednesday, January 29, 2014, at 01:24 PM. The alert states that WRDS users have access to an open trial of Morningstar CISDM data. Below the alert, there are sections for "Welcome to WRDS!", "Latest WRDS News" (with dates from January 28, 2014, to December 23, 2013), and "Financial Headlines from Moody's" (with dates from February 25, 2014, to February 25, 2014). A "What Type of Data Are You Looking For?" section lists categories like Company Financials, Financial Statements, Industry-Specific data, and Point-In-Time Accounting Data.

4th QS-Maple Abu Dhabi, UAE, May 2014

WEB BASED DATA REPOSITORIES

CHASS - UNIVERSITY OF TORONTO

The screenshot shows a web browser window with the address bar displaying "dc.chass.utoronto.ca". The page header includes the CHASS logo and the text "Data Centre Faculty of Arts & Science, University of Toronto". A navigation menu on the left lists various datasets: CANSIM, Canadian Census, SDA @ CHASS, Trade Analyser, CFMRC TSX Annual, CFMRC TSX Quarterly, TSX eReviews, D&B INKBR, Financial Post (Corporate), Fundata (MF), Basic Economics (A), Basic Economics (M), CRSP, Compustat, Penn World Tables, World Bank Databases, and OECD. The main content area features a "Welcome to the CHASS Data Centre" message, a description of the on-line databases, and a detailed section for the CANSIM dataset, including its purpose and a list of uses.

CHASS Data Centre
Faculty of Arts & Science, University of Toronto

Advancing Knowledge through Technology [What is new at Data Centre?](#)

CANSIM

Canadian Census
SDA @ CHASS
Trade Analyser
CFMRC TSX Annual
CFMRC TSX Quarterly
TSX eReviews
D&B INKBR
Financial Post (Corporate)
Fundata (MF)
Basic Economics (A)
Basic Economics (M)
CRSP
Compustat
Penn World Tables
World Bank Databases
OECD

Welcome to the CHASS Data Centre

- Mouse over the menu on the left to see the description for each dataset -

This collection of on-line databases and custom built search and retrieval programs are maintained by Computing in the Humanities and Social Sciences (CHASS) at the University of Toronto.

CANSIM

The Canadian Socio-Economic Information Management System (CANSIM) is Statistics Canada's key socioeconomic database. Updated daily, CANSIM provides fast and easy access to a large range of the latest statistics available in Canada. CANSIM comprises over 40 million time series and thousands of tables.

CANSIM can be used to:

- Track trends
- Analyze market potential
- Study economic activity
- Make investment decisions
- Evaluate social conditions
- Conduct feasibility studies
- Forecast economic conditions
- Plan programs or services
- Profile demographics

CANSIM @ CHASS features:

- Most frequently retrieved series by Academia
- Search by text, table number or time series number
- Browse by subject, survey number, table titles and labels or by Statistics Canada's keywords classification for tables

DATA PRESENTATION EXAMPLES: CANADIAN NATIONAL DATABASE (CANSIM)

The screenshot shows a web browser window with the title "CANSIM at CHASS - University of Toronto". The address bar contains the URL "dc2.chass.utoronto.ca.myaccess.library.utoronto.ca/chasscansim/". The main content area features a banner for "CANSIM @ CHASS" with the tagline "Advancing Knowledge through Technology". Below the banner, a text box states: "Starting January 4th, 2010 the flagship Canadian database CANSIM hosted at CHASS will be updated daily via a direct data feed from Statistics Canada." Two columns of links are provided: "CANSIM Multidimensional View", "CANSIM Time Series View", and "Historical versions of CANSIM" on the left; and "CANSIM Multidimensionnel", "CANSIM Séries Chronologiques", and "Versions historiques de CANSIM" on the right. A note below the links states: "Note: Cansim Multidimensional View is the recommended interface! Cansim Time Series View uses the same data feed as Cansim Multidimensional, but displays data in the former time series view." At the bottom, there is a "Data source" section with the text "Source de données" and logos for Statistics Canada and Canada. A footer contains the text: "The use of CANSIM @ CHASS is restricted to current faculty, staff, and students of [subscribing universities](#). L'utilisation de CANSIM @ CHASS est limitée au corps professoral, aux étudiants et au personnel [des universités participantes](#)." and navigation links: "Data Centre | CHASS Home | Email Us".

LIVE DATA PRESENTATION EXAMPLES: CANSIM OLAP CUBES AND SDA@CHASS

Sales of Natural Gas (4 dim)

Postsecondary Enrollment (6 dim)

Teaching Staff at Canadian Universities (7 dim)

Community Health Survey via SDA@CHASS

DATA PRESENTATION EXAMPLES:

CANSIM 1290003: SALES OF NATURAL GAS

CHASS Data Centre

clouddc.chass.utoronto.ca.myaccess.library.utoronto.ca/ds/cansim/olap/displayC

FACULTY of ARTS & SCIENCE, UNIVERSITY OF TORONTO

CHASS AGGREGATE MICRO META Data Centre VISUALIZE RESEARCH ANALYZE

en français
Welcome, University of Toronto

CANSIM Table 1290003: Sales of natural gas

Tips: Click on dimension to select members. Drag dimensions on axes. Click arrow to select next member on filter.

Columns: **Geography (9)**

Rows: **Time (71)**

Filter: **Sector (Total)** **Estimates (Revenue (Dollars))**

Table Information

Number:	1290003
Dimensions:	4
Freq:	monthly
UoM:	
Data points:	62,667
Last updated:	2014-01-27 09:16 EST

Data display
[Show table "as is" \(OLAP View\)](#)

Advanced Options

	Geography						
Time	Canada	Nova Scotia	New Brunswick	Quebec	Ontario	Manitoba	Saskatchewan
2008-01	3,101,171,000	16,689,000	34,322,000	293,458,000	1,329,424,000	109,086,000	176,141,000
2008-02	2,961,987,000	17,097,000	24,927,000	277,522,000	1,340,642,000	106,542,000	177,921,000
2008-03	2,885,065,000	14,459,000	25,631,000	268,155,000	1,290,140,000	85,552,000	171,546,000
2008-04	2,238,424,000	9,739,000	17,706,000	182,334,000	900,054,000	55,900,000	124,700,000
2008-05	1,815,642,000	13,854,000	16,963,000	144,689,000	704,126,000	44,142,000	129,735,000
2008-06	1,495,662,000	10,365,000	18,419,000	127,568,000	538,878,000	35,278,000	104,541,000

ON LINE DATA ANALYSIS

SDA@CHASS

A perfect example of academic collaboration

Software: University of California at Berkeley

Host: University of Toronto, CHASS

Data: Statistics Canada; U of Toronto Library

Users: 15 Canadian universities

ON LINE DATA ANALYSIS

SDA@CHASS

F
E
A
T
U
R
E
S

- Web based and interactive
- Data search and discovery
- Statistical analysis, including:
 - cross-tabulations
 - comparison of means
 - correlations
 - multiple and logistic regression
 - all standard measures of significance
- Graphic visualization of the results

ON LINE DATA ANALYSIS

SDA@CHASS - cross-tabulation

University of Toronto Data Lib... Tables - SDA
 sda.chass.utoronto.ca.myaccess.library.utoronto.ca/cgi-bin/sda/hstda3
 Canadian community health survey cycle 0-1, 2000 - common and optional contents
 Feb 25, 2014 (Tue 05:07 PM EST)

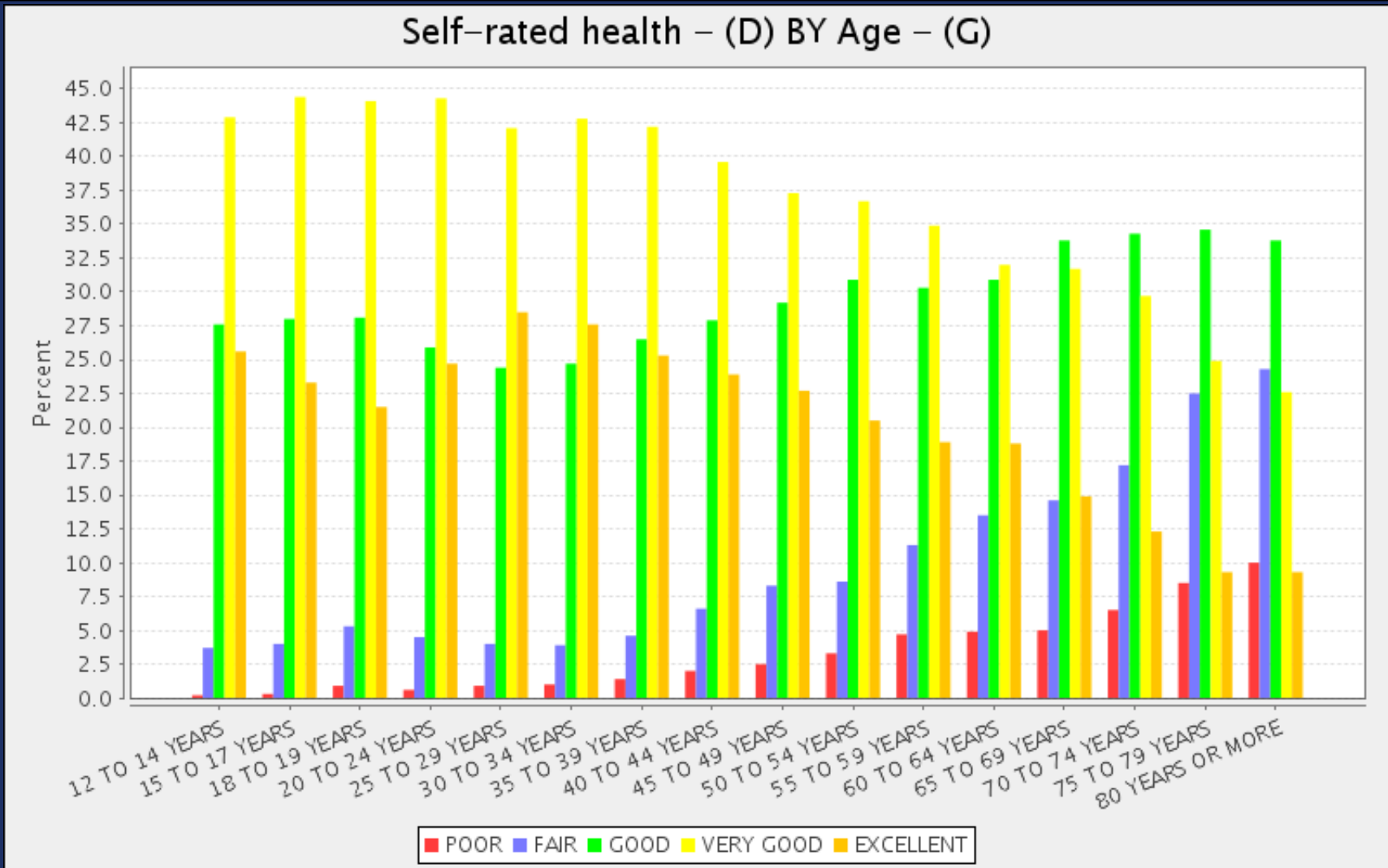
Variables					
Role	Name	Label	Range	MD	Dataset
Row	genedhdi	Self-rated health - (D)	0-4	6-9	1
Column	dhhegage	Age - (G)	1-16		1
Weight	wtse_m	Weights - Master	2.01-4,741.47		1

Frequency Distribution												
Cells contain: -Column percent -Weighted N		dhhegage										
		1 12 TO 14 YEARS	2 15 TO 17 YEARS	3 18 TO 19 YEARS	4 20 TO 24 YEARS	5 25 TO 29 YEARS	6 30 TO 34 YEARS	7 35 TO 39 YEARS	8 40 TO 44 YEARS	9 45 TO 49 YEARS	10 50 TO 54 YEARS	11 55 TO YEARS
genedhdi	0: POOR	.2 2,789.6	.3 3,668.9	.9 7,542.0	.6 13,315.4	.9 19,922.4	1.0 20,973.9	1.4 30,892.9	2.0 55,958.6	2.5 62,472.4	3.3 74,450.4	92,3
	1: FAIR	3.7 45,867.2	4.0 51,492.1	5.3 43,258.6	4.5 100,967.0	4.0 84,841.8	3.9 81,590.5	4.6 105,233.3	6.6 185,698.5	8.3 209,038.7	8.6 192,451.2	222,6
	2: GOOD	27.6 340,680.8	28.0 365,202.8	28.1 228,822.7	25.9 579,495.1	24.4 514,082.7	24.7 512,799.9	26.5 604,522.7	27.9 780,821.7	29.2 739,690.7	30.9 693,968.0	597,1
	3: VERY GOOD	42.9 528,952.2	44.4 579,085.5	44.1 358,490.7	44.3 991,113.8	42.1 885,661.7	42.8 889,840.1	42.2 962,858.7	39.6 1,109,045.4	37.3 944,589.9	36.7 826,466.7	688,4
	4: EXCELLENT	25.6 315,180.9	23.3 303,404.5	21.5 174,842.0	24.7 553,182.8	28.5 598,364.5	27.6 574,029.3	25.3 577,012.9	23.9 669,877.4	22.7 575,076.5	20.5 461,835.0	372,1
	COL TOTAL	100.0 1,233,470.7	100.0 1,302,853.9	100.0 812,956.0	100.0 2,238,074.1	100.0 2,102,873.1	100.0 2,079,233.8	100.0 2,280,520.6	100.0 2,801,401.8	100.0 2,530,868.1	100.0 2,249,171.3	1 1,972,7

4th QS-Maple Abu Dhabi, UAE, May 2014

ON LINE DATA ANALYSIS

SDA@CHASS - cross-tab graph



4th QS-Maple Abu Dhabi, UAE, May 2014

CONCLUSIONS

- Provide access to many academic and commercial data repositories.
- Form data consortia with other academic centres.
- Teach students how to navigate those databases, search for data, and retrieve data.
- Base all student assignments on real economic and social questions.
- Teach analytics on line and off line.



Chris Leowski
University of Toronto, Canada
Faculty of Arts & Science



Contact Information

chris.leowski@utoronto.ca

This presentation in PPT, PDF, Keynote
and video formats, and the full transcript,
can be downloaded from:

<http://chass.utoronto.ca/chris/QS-MAPLE-2014>

THANK YOU!

4th QS-Maple Abu Dhabi, UAE, May 2014