

DATA LINKS

University of Guelph Data Resource Centre

<http://drc.uoguelph.ca>

Fall 1997

DRC Opens Its Doors

The success of the Data Resource Centre pilot project has resulted in an endorsement by the Vice President's Advisory Council for the Centre to operate as an open service to the University community.

Students and researchers requiring data now have access through a centralized point.

DRC Services

The DRC homepage, at <http://www.drc.uoguelph.ca>, provides access to an extensive collection of data through a variety of sources.

Data, purchased by the University and made available locally, includes such items as *Canada Census*; *Census of Agriculture*; Stats Canada special surveys, including *Aboriginal Peoples Survey*, *Survey of Consumer Finance*.

Data is also available from external sources through consortial arrangements. Specific data can be obtained from these sources upon request. Catalogues are available for viewing in the DRC.

In this issue...

DRC Services	1
Search Engines.....	1
Canada Census of Agriculture	2
Workshops.....	2
Collaborative Ventures.....	3
List of Available Data Sets.....	3
Data applications on campus.....	4
Summary.....	4

The Centre is collecting a master list of links to data mounted world-wide. Much of this data is available on-line, while some sites provide on-line catalogues of data available for purchase. The Centre provides initial support accessing remote sites, however, users may be required to contact these sites directly for assistance.

The Centre will operate a limited walk-in service with hours:

Monday - Friday: 10:00 a.m. - 12:00 a.m.
1:30 p.m. - 3:30 p.m.

In-depth personal assistance will be provided by appointment. Users requiring detailed assistance are requested to contact the Centre by phone(x6410) or by e-mail (drc@uoguelph.ca) to ensure that appropriate instruction can be arranged.

Search Engines

The DRC is in the process of developing a searchable index for our site. In the meantime the following services are very useful in locating sources of data.

Thematic Search Tool

Statistics Canada has mounted a search tool on its home page which allows users to search for social statistics by keyword. While this tool currently covers social statistics, plans are underway to include economic, business and trade sources.

This tool is intended to guide users to data and variables from Statistics Canada data files dating back to 1985. The files are organized in broad themes such as childcare, health, etc. By limiting searches to key

words and subjects, users are able to find the best source of data for their specific research needs. Detailed descriptions of individual surveys are available including information on data gathering, methodology used, and output format.

Once data sets are identified, users can click on source options. Information concerning codebooks, questionnaires and record layouts are also available on-line.

For University of Guelph users, it is then possible to link back to the DRC home page and locate data previously mounted. If this data is available locally, users can then choose to work directly from the raw data, or, where applicable, choose the SAS-to-Web option for obtaining statistical summaries.

If the data is not available locally, options can be arranged with Statistics Canada for obtaining data. Many of the data sets are available free of charge through our current agreements. Pricing and ordering information is available on-screen at the Stats Canada site.

CESSDA

The Cessda database (<http://www.nsd.uib.no/cessda/IDC>) allows searches of holdings of major European and North American data centres. Although not all information sited is freely available, staff in the DRC can inquire about acquisitions of specific surveys.

Feature Article

SAS To Web Project

Our initial newsletter featured an overview of the DRC SAS-to-Web project. This initiative provides ease of access to statistical data in a format which does not require the user to have extensive knowledge of statistical packages. Each issue of the newsletter will feature a data set which has been mounted in the SAS-to-Web format.

Census of Agriculture

The 1986 and 1991 Census of Agriculture data has recently been mounted in the SAS-to-Web format and are among our most heavily used data sets. Access can be gained through the DRC home page (<http://drc.uoguelph.ca>) by selecting the U of G web-retrieval option. After reading through the licencing agreement covering this data, *licenced* users may then select 'Canadian Census' from the list of options. Click on the 'Submit Request' button. Choose the set of data required for research.

Users may then select from a list of pre-determined variables to subset data. These pre-determined variable selections are based on the most commonly requested variables for subsetting (eg. subset by specific provinces, or by all provinces). This list can be easily expanded upon request. Also included is a list of all variables contained within the data set from which users can choose one or more items.

Once the variables are selected, users submit the request and a unique file is created with the results of the search.

Data gathered for the agricultural census focuses on aggregate output statistics for an exhaustive list of crops and livestock. As with many Statistics Canada data sets, in order to protect anonymity the smallest level of aggregation is the township. If more detailed information is required it may be possible to obtain a custom tabulation.

A wide variety of output options are available to users. Options include Lotus, Quattro pro, Padox, Excel, SAS and SPSS to mention a few. For those researchers preferring to use raw data, all raw data files are available at the top of the screen under Program files.

Codebooks, questionnaires and user guides can be accessed electronically wherever possible. Options are also available to summarize the data using means, cross-tabs, frequencies and simple regression. Paper copies of codebooks are available in the DRC.

Workshops

This past summer staff of the DRC attended a workshop in Ottawa which was sponsored by Statistics Canada. It was designed for data librarians involved with disseminating information from the DLI (Data Liberation Initiative). It was an excellent opportunity to share ideas with people from across Canada.

Workshops began with the basics of what data is, how to find and order data, and how to disseminate it to the user community. Demonstrations were given on use of specific data sets, such as Sabal, which uses the 20/20 interface from Ivision. This interface is being used on several Statistics Canada data sets and is an excellent system for displaying and manipulating data (see next newsletter).

Staff also attended the Canadian Association of Public Data Users (CAPDU) meeting, held in conjunction with the Learned's in St. John's Newfoundland. A brief introduction to the activities of the DRC was presented. Sessions were attended on the 1996 and 2001 census. There was also a great deal of discussion concerning the Canadian data community and future opportunities to develop cooperative data resource sharing. Feedback from both sessions, on efforts at Guelph, were extremely positive and encouraging.

Collaborative Ventures

Discussions are underway with the University of Waterloo and Wilfrid Laurier University for the establishment of a joint Data Resource Service centred around the University of Guelph's web retrieval system.

The increase in the amount of data available electronically and changes in dissemination tools make this an excellent opportunity for collaborative ventures. An update will be presented in the next newsletter

The extensive updating of these lists since our last newsletter is an example of how quickly new information can be made available using our SAS-to-Web program.

List of Available Data-Sets: SAS-to-Web

(as of November 1997)

STATISTICS CANADA

- Aboriginal Peoples Survey
- Family Expenditure Survey
- General Social Survey - Cycles I-X
- Health & Activities Limitation Survey
- Nat'l Longitudinal Survey of Children & Youth
- National Population Health Survey & Supplement
- Survey of Consumer Finance (HIFE)
- Survey of Labour & Income Dynamics - Job and Person Files
- Violence Against Women
- Work Arrangement Survey

CANADIAN CENSUS

- Agricultural Census - 1986, 1991
- Census - family, household, individual (PUMF) - 1986, 1991
- Census - postal codes - 1991
- Census - EA profiles Part A - 1991
- Census - EA profiles Part B - Atlantic, Quebec, Ontario, West - 1991

CLIMATE DATA

- ERS - 0800, 1600. Hourly Data, - Rain data
- GTI - 0800, 1600, 2400, Hourly Data, - Rain data

ICPSR (U. S.)

- Childbirth & Adoption Survey
- Ego-Alter File
- International Social Survey - Environment
- IMF - Government Finance Statistics
- National Hospital Discharge Survey
- National Ambulatory Medical Care Survey
- National Medical Expenditure Survey
- National Health Interview Survey

NOTE:

See DRC web site or information sheets for listings of years held for each data set. This list will be continually updated as new data arrives.

New Data Sets from Statistics Canada (CD-ROM Products)

- E - Stat
- Extraction System of Agricultural Economics
- Georef
- Importance of Wildlife to Canadians
- Labour Force Historical Survey
- Postal Code Counts
- Small Area Business & Labour (SABAL)
- TIERS - Trade Information

The above data sets are available on CD ROMs mounted locally in the Data Resource Centre. Help sheets are available outside the DRC.

Data Applications on campus

Each newsletter will feature an application associated with data available in the DRC.

Projected Needs for Physicians in Kansas:

Professor B. Ferguson of the Economics Department recently acquired a series of studies from ICPSR (see Summer 1997 newsletter) related to health care in the United States. These data were freely available through the consorcial agreement.

There is an extensive array of information collected from hospital discharge and medical expenditure records, as well as ambulatory records and health interviews.

This information is being used as part of a project looking at the projected needs for physicians in Kansas. They are specifically looking into the demographic pressure associated with an aging population.

Summary

Staff at the DRC are pleased to offer reference assistance in support of the data needs of the University community. We are working to provide a wider range of accessibility through the development of information sheets, detailed on-line help screens as well as walk-in and phone-in service.

Through a joint initiative with the Library, the DRC will be providing public terminals to be located outside the DRC office. This will increase access to those users who do not have direct access to computers with a graphical browser.

We welcome suggestions and comments regarding the services we offer. We will continue to add data sets to the web site as they become available.

Our web pages also include numerous links to data available from outside sources. These pages are reviewed and revised on an on-going basis in order to provide current and accurate information.

The next issue of DATA LINKS will appear in the Winter 1998 semester.

NEXT ISSUE:

In the next issue we will be focusing on the services offered to on-campus users. Several new initiatives will be in place by the Winter semester which will provide increased access for users.

Our feature article will once again highlight a data set of interest to many researchers on-campus. This Statistics Canada product, called SABAL, has many new capabilities which allow users increased flexibility in tailoring search results.