

DATA LINKS

TriUniversity Data Resources <http://tdr.tug-libraries.on.ca> Vol. 5 Issue 1

Editor: Shabiran Rahman srahman@library.uwaterloo.ca

NEW TDR LOGIN FOR ENHANCED SERVICE

Starting in Spring 2002 TDR will be adding a new service that will allow us to notify users when data sets and documentation have been updated. In order to accomplish this, users will be required to login to the TDR web site.

Currently users are informed of changes and/or data updates via the readme file that accompanies each data set. It is the users' responsibility to be aware of any changes that may have occurred since they downloaded their data. By developing a database of user e-mail addresses, we are becoming more proactive and are providing users with the changes directly. All information will remain confidential and users will have options to choose when and if they are contacted. This model is similar to what has been adopted at other data providers, such as ICPSR and IPUMS. We appreciate your comments and suggestions.

Dr. Michelle Edwards
Academic Services, CCS
University of Guelph

ICPSR DIRECT

ICPSR Direct is a new service providing direct access to the data holdings of the Inter-university Consortium for Political and Social Research (ICPSR) in Ann Arbor, Michigan. Currently, ICPSR provides access to approximately 10,000 files of US and international (including some Canadian) quantitative research data.

CONTENTS

New TDR Login
ICPSR Direct
Featured Data Set
Data Documentation Initiative (DDI)
Graduate Students Use TDR Data
Contacts
SWORDC Opens
News
New Acquisitions

With ICPSR Direct, any member of the University of Guelph or University of Waterloo academic communities, with an authorized campus IP address, can now download data directly from ICPSR. Wilfrid Laurier University members should continue to request data through their library data centre until the ICPSR Direct service is implemented on that campus.

Please note that ICPSR Direct is an additional service, not a replacement for access to ICPSR data provided by library staff. One major advantage to using ICPSR Direct is that this service provides for immediate delivery of data. You can access ICPSR Direct in several different ways:

- a) visit the ICPSR web site: www.icpsr.umich.edu
- b) visit the TUG TriUniversity Data Resources web site: tdr.tug-libraries.on.ca and click on TDR Data—> ICPSR Direct

c) visit the University of Waterloo's EDS web site:
www.lib.uwaterloo.ca/uweds/icpsr/direct.html

Once you are on the ICPSR web site, begin by conducting a search for data using the search tool at the top of the home page. If subsequently you click on a "downloads" link and if you are using a computer in the uwaterloo.ca or uoguelph.ca domains (or using the TUG proxy service), you will then see an Authorized Download page. This page describes conditions of data use and procedures for obtaining assistance.

Most ICPSR technical documentation, including codebooks, is available in electronic downloadable format from the ICPSR web site. For further information about access at the three universities, please refer to the "Contacts" information included in this issue.

Richard Pinnell
University of Waterloo

FEATURED DATA SET:

2000 General Social Survey – Cycle 14 Access to and Use of Information Communication Technology

How "connected" are Canadians? To find this out check out the data in Cycle 14 of the General Social Survey. This cycle is the first cycle to collect detailed information on access to and use of information communication technology. The survey includes use of Automated Teller Machines (ATMs), fax machines, DVD players, and cellular phones, but it has a specific focus on computer and Internet use.

The GSS program, established in 1985, conducts telephone surveys across the 10 provinces using a Random Digit Dialing approach to select households. From January to December 2000 25,090 people were successfully interviewed.

Questions were asked about the location of computer use and Internet access including home, work, school, and other public locations. The questionnaire collected data about the impact of technology on work, e-mail as a communication tool, and time displacement. Those surveyed who do not have computers and do not use the Internet were asked questions to find out why they don't. Parents with children were asked whether they purchased computers for their children's use and if they monitored their use. Respondents were asked whether they had concerns about providing

credit card numbers and other personal information over the Internet?

The questionnaire also included information about the respondents' work, educational background, volunteer activities, and income. Responses to these questions will support the analysis of the social impact of technology on Canadians.

Statistics Canada has published some results based on the findings of this survey. Both are available on-line at www.statcan.ca.

Access to and Use of Information Communication Technology (56-505-XIE)

General Social Survey – Internet Use in *The Daily*, March 26, 2001.

To access the 2000 General Social Survey – Cycle 14 data, please visit the TDR web site:
tdr.uoguelph.ca

Diane Boyd
University of Guelph

DATA DOCUMENTATION INITIATIVE (DDI)

The purpose of the DDI is to create international criterion and methodology in the preservation and use of "metadata". Metadata (data about data) are the documentation that accompany data files. Through the establishment of DTD or Document Type Definition, metadata can be created in a completely searchable format. DTD employs eXtensible Markup Language or XML, which creates a uniform structured format for marking up metadata.

The DDI is now being used in such international projects as Networked Social Science Tools and Resources - NESSTAR (www.nesstar.org/) and the Inter-university Consortium for Political and Social Research – ICPSR (www.icpsr.umich.edu/). Harvard-MIT and the University of California, Berkeley are just two of the universities now beginning to employ DDI criterion in the presentation of data.

TDR is in the early stages of testing DDI criterion, and local applications to our data files have been developed. The end product will mean that users can search data sets in order to find specific variables, the questions associated with the variables, the responses, and the frequency of the

responses to the questions. Testing of the project began in early January 2002.

For more information on DDI and its possibilities, see the ICPSR DDI site at www.icpsr.umich.edu/DDI/

Carol Perry
University of Guelph

GRADUATE STUDENTS USE TDR DATA TO LEARN APPLIED STATISTICAL ANALYSIS

As an integral part of the learning experience in their multivariate statistics class, graduate students in the Department of Recreation and Leisure Studies use a variety of data sets made available through the TDR site. Along with the supporting documentation that accompanies the data, the TDR site provides a comprehensive and rich resource for students to turn basic concepts and procedures in statistics, into application.

Among the data sets that the graduate students have been using to explore various issues related to leisure include: the 1994 and 1996 National Population Health Surveys; the 1998 Canadian Travel Survey; the 1997 National Survey on Giving, Volunteering, and Participating; the 1991 Health and Activities Limitations Survey; the 1996 National Survey on the Importance of Nature to Canadians; and several of the different cycles of the General Social Survey, including Cycles 7 (1992) and 12 (1998) on Time Use of Canadians, and Cycle 11 (1996) on Social and Community Support. Typically based on large national surveys of several thousand Canadians and often incorporating several hundred variables, these data sets provide tremendous flexibility for analysis and interpretation.

As the course instructor I use classroom time to discuss broader conceptual and measurement issues and statistical applications relevant to leisure studies. Then, the students apply what they have learned by testing out their own ideas on a data set they have chosen. I provide the students with individualised CD-ROMs that include the data set in an SPSS savefile downloaded from the TDR site as well as any supporting documentation pertaining to the study, such as the user guide and questionnaire. The students can then access and work with these files in the computing labs using the statistical analysis package, SPSS for Windows.

There are several advantages to using a variety of different data sets. Leisure Studies embraces a wide array of topic areas and is influenced by a number of disciplinary perspectives, such as psychology, sociology, economics, and geography. Focusing on data drawn from one subject area limits the relevance of the examples. By using data drawn from the TDR site, students are able to select a data set on a topic that coincides with their own research interests and disciplinary orientation, and they can explore ideas and problems similar to those they might face when undertaking their thesis work. Even though the data they use are different, all of the students must still organise their data and construct composite measures using similar strategies, and eventually analyse the data and interpret their results according to the same statistical principles. Consequently, they learn more about the very diverse areas of research that their peers are pursuing while at the same time seeing how issues of conceptualisation, measurement, analysis, and interpretation are common across interest areas.

The process has been extremely successful and gives the students an opportunity to apply what they have learned in the classroom to “real life” data. In this way, they are able, too, to experience the sense of discovery that comes from exploring ideas of their own.

Dr. Bryan Smale
Department of Recreation and Leisure Studies
University of Waterloo

NEWS

Aggregate data files from Statistics Canada are frequently presented in IVISION B20/20 format for easy manipulation and use.

Users should ensure that they occasionally download the latest available version of B20/20 software. Some newer files such as *Income Trends in Canada 1999*, require the use of version 5.2 or higher to open the files.

Each introductory or read-me file accompanying a data set using B20/20 software contains a link to the latest version available for download.

SOUTHWESTERN ONTARIO RESEARCH DATA CENTRE (SWORDC) OPENS

The official opening of the Southwestern Ontario Research Data Centre, or SWORDC, was held on December 7, 2001. Among those present were Dr. David Johnston, President of the University; Dr. Paul Guild, Vice-President for Research; Dr. Bob Kerton, Dean of Arts; and Dr. Gustave Goldmann, Statistics Canada's manager of the research data centres project, along with a large group of researchers from all across campus.

UW is collaborating in the SWORDC venture with Wilfrid Laurier University, University of Guelph, Brock University, University of Windsor, and University of Western Ontario. Four of the five partner universities were represented at the opening.

A presentation on how to use the SWORDC was given by John Goyder (Co-Director of SWORDC along with Keith Warriner) and Pat Newcombe-Welch, the Statistics Canada Analyst who runs the Centre.

As of the week of January 21, 2002, the Centre became fully operational following a delay while final Statistics Canada requirements for the security of data were put into place. The Centre has available a wide variety of advanced statistical packages such as SUDAAN.

The SWORDC is located at Room 2228 of the PAS (Psychology, Anthropology, Sociology) building at the south end of the UW campus. The Centre can obtain virtually any survey collected by Statistics Canada, but unlike library data facilities such as TDR, access is by approved-project only. The essential requirement is that master files with full, un-collapsed, codings of variables be necessary for the project. Projects not meeting that criterion are not RDC-eligible and in any case would be better served using public use version data sets.

Details on procedures are outlined on the SWORDC web site at www.rdc.uwaterloo.ca. Researchers who would like to explore the suitability of their project for the SWORDC are cordially invited to contact John Goyder (x2643) or Keith Warriner (x3678). The STC Analyst, Dr. Pat Newcombe-Welch (x5504)

panewcombe@uwaterloo.ca is the best source for details about data sets. For copies of the SWORDC brochure, contact Pat Newcombe-Welch at:
SWORDC, c/o Sociology Department, University of Waterloo

Dr. John Goyder
jgoyder@watarts.uwaterloo.ca

A FEW OF OUR NEW ACQUISITIONS

- Labour Market Activity Survey 1988-1990
- Occupational Prestige Ratings Survey for 1989 GSS - U.S.
- Public Service Employee Opinion Survey 1999
- Household Internet Use Survey 2000
- Homeowner Repair & Renovation Survey & Home Energy Retrofit Survey 1994 & 1995
- Income Trends in Canada 1980-1999

For other recently added data sets please visit
tdr.tug-libraries.on.ca/NEWS/mainnews.htm#newacq

YOUR CONTACTS FOR DATA SERVICE AT OUR THREE LOCATIONS

University of Guelph
Michelle Edwards
edwardsm@uoguelph.ca
519-824-4120 x4539

University of Waterloo
Susan Moskal
srmoskal@library.uwaterloo.ca
519-888-4567 x2890

Shabiran Rahman
srahman@library.uwaterloo.ca
519-888-4567 x2882

Wilfrid Laurier University
Wray Roulston
wroulsto@wlu.ca
519-884-0710 x3743