

IN THIS ISSUE

Human Subject Research	1
Off-Campus Access to Roper <i>Express</i>	2
Summer Research Program at ICPSR	2
Counting Words and Names in Creative Data Displays . . .	3
SSML Mobile Lab	4
FactFinder Download Center .	4
Crossroads Corner	4

.....

DPLS News contains articles about local, national and international data issues. It is published twice a semester by the library staff and sent to faculty, graduate students, and librarians, with a focus on the Social Science departments.

Editor: Joanne Juhnke, Special Librarian

Contributors: Cindy Severt, Senior Special Librarian, and Lu Chou, Senior Special Librarian

.....

University of Wisconsin-Madison
1180 University Drive
3308 Social Science Building
Madison, Wisconsin 53706 USA

Phone: (608) 262-0750

E-mail: dpls@mailplus.wisc.edu

Web: <http://dpls.dacc.wisc.edu/>

Hours: Monday-Friday 8:30-4:30

.....

This issue of DPLS News is online at:

<http://dpls.dacc.wisc.edu/pubs/Newsletters/feb06news.html>

HUMAN SUBJECT RESEARCH— PRIVILEGE AND RESPONSIBILITY

by *Cindy Severt*

The last time *DPLS News* visited the topic of human subject research was in 2002 in an article titled “So You Want to Use Restricted Data...” Much has changed since then, so we decided it was time to revisit this important issue.

One key fact remains the same—secondary data analysis demands consideration of human subject issues. At UW-Madison, it is important to know about the following eight data repositories (including DPLS). Research that analyzes public-use secondary data from these repositories does not require prior Institutional Review Board (IRB) approval:

Secondary data analysis demands consideration of human subject issues.

- Better Access to Data for Global Interdisciplinary Research (BADGIR)
- Center for Demography and Ecology—UW-Madison
- Data and Program Library Service (DPLS)—UW-Madison
- Inter-University Consortium for Political and Social Research (ICPSR)
- National Center for Health Statistics
- National Center for Education Statistics
- National Election Studies
- Roper Center for Public Opinion Research
- U.S. Bureau of the Census

If the data you plan to analyze does not come from one of these repositories, or if you plan to merge more than one dataset such that individuals might be identifiable, you will have to seek approval by submitting a protocol to the Social & Behavioral Science Institutional Review Board (IRB), one of five campus IRBs.

What is the reasoning behind the protocol requirement?

Key to determining the reason for submitting a protocol is the definition of *research* which is defined in the federal Common Rule as “a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge” (45 CFR 46.102(d)); and *human subject* which the federal Common Rule defines as “a living individual about whom an investigator (whether professional or student) conducting research obtains (1) data through intervention or interaction with the individual, or (2) identifiable private information” (45 CFR 46.102(f)).

Continued on p.3

OFF-CAMPUS ACCESS TO ROPER EXPRESS

by Lu Chou

DPLS is happy to announce that RoperExpress is available via the UW-Madison Libraries web proxy server, at <http://digital.library.wisc.edu/1711.web/rcpor>. If you have a valid UW-Madison campus ID, you can use this link to download studies from the Roper Center for Public Opinion Research from anywhere in the world. The proxy server is set up to verify your affiliation with UW-Madison. To learn more about the proxy server, visit <http://www.library.wisc.edu/help/remote/>.

SUMMER RESEARCH PROGRAM AT ICPSR

by Lu Chou

Does spending part of your summer in Ann Arbor, Michigan sound appealing to you? The Summer Research Program at the Inter-University Consortium for Political and Social Research (ICPSR) offers an array of classes in research design, statistics, data analysis, and social science methodology. The program emphasizes methods of quantitative analysis within a broader context of substantive social research.

There are two four-week Summer Program sessions, June 26 - July 21 and July 24 - August 18. Course information and online registration for 2006 is available online at <http://www.icpsr.umich.edu/training/summer/>.

COUNTING WORDS AND NAMES IN CREATIVE DATA DISPLAYS

by Joanne Jubnke

Sometimes a picture is worth a thousand words, or several thousand points of data. And then again, sometimes the words themselves are the picture. The web sites WordCount™ and Baby Name Wizard NameVoyager provide two different experiments in aesthetics-driven displays of frequency counts.

The WordCount site, at <http://www.wordcount.org/>, takes as its data source the British National Corpus (BNC), a massive collection of written and spoken texts designed to represent the current state of British English. WordCount takes a frequency count of all the words used at least twice in the BNC database, and displays the words end-to-end, from most-used to least, in a Flash-based program. The font-size of each word relative to its neighbors gives another visual representation of each word's relative popularity in the BNC. Users can scroll from one word to the next, search the collection by word, or select a section of the display by rank. The image below is the result of a search on the word "data." In an interesting twist, users can see an analogous display called QueryCount, whose data source is the search terms used by previous visitors to the site. One frustrating gap: no actual frequency count numbers are displayed.



The Baby Name Wizard NameVoyager site, at <http://www.babynamewizard.com/namevoyager/>, takes the challenge one step further by incorporating name popularity over time. The Java-based program takes baby-name data from the U.S. Social Security Administration and displays a graph representing the 1000 most popular names by decade starting in 1880. Boy names are displayed in shades of blue, girl names in shades of pink, with deeper shading for more popular names. Frequency of use is depicted by the width of the pink or blue stripe associated with the name, and a mouseover function identifies the popularity ranking. Select a specific name to see its graph alone, or compare "neighboring" names beginning with the same letters. *The Baby Name Wizard* is also a book written by Laura Wattenberg, whose fascination with baby-name statistics extends to a companion blog at <http://www.babynamewizard.com/blog/>.

HUMAN SUBJECT RESEARCH

Continued from p.1

Is it a lengthy process?

Not necessarily. Chances are your protocol, should you need to file one, falls under the expedited category, but even so, every application consists of six steps:

1. Determine if your research is covered by federal regulations
2. Take the required training (the online tutorial described below)
3. Select an IRB (Social & Behavioral Sciences for most of our users)
4. Plan for informed consent and HIPAA authorization
5. Decide what kind of application to submit: full, expedited, exempt
6. Prepare and submit an application

These steps and additional supporting information can be found at <http://info.gradsch.wisc.edu/research/compliance/humansubjects/>.

Datasets that require a protocol

If in doubt about needing to submit a protocol for a dataset you wish to obtain, ask us—especially if the dataset you seek specifies restrictions and/or requires a signature. Be aware that such a signature has to be someone from Research and Sponsored Programs who can legally represent UW.

Tutorial

The most informative method of learning about the rules and regulations governing the professional privilege of conducting research involving human subjects at UW is to take the online tutorial at <http://info.gradsch.wisc.edu/research/compliance/humansubjects/tutorial/>.

The tutorial consists of four modules and includes numerous examples to show practical applications of the concepts.

SSML MOBILE LAB

by Doug Hemken

The Social Science Microcomputing Laboratory (SSML) Mobile Lab provides web access and statistical programs anywhere in the Social Science Building. The Mobile Lab has up to 40 laptop computers and a wireless network hub. SSML staff bring the Mobile Lab to your classroom, set it up, and collect it when your class is over. Many classes are eligible to use SSCC's Room 2470. To ask about reserving the Mobile Lab for your class, contact Ann Lewis in the SSML, atlewis@wisc.edu, 262-0862.

Note: DPLS and SSML together make up the Data and Computation Center (DACC) on the third floor of the William H. Sewell Social Science Building.

FACTFINDER DOWNLOAD CENTER

by Joanne Jubnke

The U.S. Census Bureau continues to fine-tune the American FactFinder interface at <http://factfinder.census.gov/>. One recent feature, the Download Center, allows users to download tables for an entire category of geographic areas, such as all the counties in a state or all the ZIP Code Tabulation Areas (ZCTAs) in the country. In the past, the interface would only let you select a single area at a time—unless you wanted to download massive FTP files to work with on your own!

The Download Center is listed as a link in the left-hand column of the American FactFinder home page. From the Download Center page, you first choose a dataset. Available data includes files from Census 2000, the 2004 American Community Survey, and the 2004 Population Estimates. After you choose the dataset, the next step is to choose geography and then decide whether you want selected tables (up to 50) in comma-delimited format, or all the tables in the Census' own summary-file format.

Users should be aware that the Download Center does not provide for all combinations of geography: for example, one can get all ZCTAs in the country, but not for an individual state. The interface also does not allow for selection of multiple individual geographies. Still, the Download Center does represent a useful step forward.

Crossroads Corner highlights web sites recently added to the searchable Internet Crossroads in Social Science Data, available on the DPLS web site.

by Joanne Juhnke

ETI Drill Down Tool Kits

No, this isn't a dental-association site; it's a nicely-packaged collection of U.S. Census data broken down to the census tract level, examining business and worker characteristics and purchasing power estimates. The Employment and Training Institute at the University of Wisconsin-Milwaukee presents the information at <http://www.uwm.edu/Dept/ETI/drilldowns/index.html> with an eye toward business plans, economic development and academic research. Six collections of "drill down" reports are available on the site:

- Business Place-of-Work Drill Downs examine the characteristics of jobs located in each U.S. census tract.
- Employer Diversity Drill Downs identify the race/Hispanic origin composition of the workforce employed in each U.S. census tract.
- Neighborhood Workforce Drill Downs describe the type of jobs held by local residents.
- Purchasing Power Profiles examine retail potential for 16 different types of consumer expenditures for census tracts and residential ZIP codes.
- Urban Markets Retail Sales Leakage/Surplus Drill Downs show the difference between the purchasing power of residents compared to the retail sales estimated to result from retail employees in the same neighborhood.

The data are compiled from Census Transportation Planning Packets and decennial data from 2000, Consumer Expenditures Surveys, and ZIP Code Business Patterns.

National Data Analysis System (NDAS)

The National Data Analysis System, at <http://ndas.cwla.org/>, provides data and

information about child welfare in the United States. Produced by the Child Welfare League of America, the site aims to support an information-based grounding for children's programs and policies. Data topics include child abuse and neglect, adoption and foster care, child health, juvenile justice, and child welfare administration. Users can create their own tables and graphs for a single state or groups of states. Tables may be downloaded as ASCII files to be read in Excel. Many topics only cover the latest year's data, generally two years behind due to reporting cycles. However, some topics have data over time going back a decade or more. The site also provides fact sheets and data trends reports for the fifty states plus the District of Columbia.

MIPT Terrorism Knowledge Base

The Memorial Institute for the Prevention of Terrorism (MIPT) was founded in 1995 after the bombing of the Murrah building in Oklahoma City. MIPT is a non-profit organization focused on the prevention of terrorism in the United States, currently funded by the U.S. Department of Homeland Security. The Terrorism Knowledge Base (TKB), online at <http://www.tkb.org/Home.jsp>, covers terrorist groups worldwide. The types of groups in the database range from nationalist to anti-abortion, from racist to animal-rights. The database includes terrorism incident data and hundreds of group and leader profiles and trials. Also included are interactive maps, statistical summaries, and analytical tools for creating custom graphs and tables.

Terrorism incident data covers 1968-1997 for non-U.S. incidents, and the entire world from 1997 to the present. The site also houses the Worldwide Incidents Tracking System (WITS) from the National Counterterrorism Center, a separate and more extensive accounting of terrorist incidents from 2004 to March 2005.