



**IN THIS ISSUE**

Yearning for 100%-Count Data,  
Eight Years After Census  
2000 ..... 1  
Researcher's Notes ..... 2  
BADGIR Update ..... 2  
For Undergrads: News From  
ICPSR ..... 3  
Crossroads Corner ..... 4

.....

DISC 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  
Staff contributors: Lu Chou, Senior  
Special Librarian

.....

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

Phone: (608) 262-0750  
E-mail: [disc@mailplus.wisc.edu](mailto:disc@mailplus.wisc.edu)  
Web: <http://www.disc.wisc.edu/>  
Hours: Monday-Friday 8:30-4:30

.....

This issue of DISC News is online at:  
[http://www.disc.wisc.edu/pubs/  
Newsletters/nov08news.html](http://www.disc.wisc.edu/pubs/Newsletters/nov08news.html)

**DISC will be CLOSED:**

Thursday-Friday, Nov. 27-28  
Wednesday-Friday, Dec. 24-26  
Wednesday-Thursday, Dec. 31-Jan. 1

**YEARNING FOR 100%-COUNT DATA,  
EIGHT YEARS AFTER CENSUS 2000**

*by Joanne Juhnke*

We live in a world where news coverage is updated around the clock, and today's print edition is passé long before nightfall. Electoral poll-watching addicts, tracking the U.S. elections just past, had multiple daily trackers over which to obsess (see Crossroads Corner for related sites). The expectation that all data is just as up-to-the-minute begins to seep into our consciousness—until we have the need for census-tract-level demographic data, and we run squarely into the fact that 100%-counts only happen once every ten years. And the 2010 Census won't happen for another year and a half, with the inevitable ensuing time lag prior to the release of that data.

**The Census Bureau  
has attempted to  
address the insatiable  
urge for the latest  
to-the-minute data.**

Even in the absence of 100%-counts, however, there are additional options. This article summarizes several Census Bureau projects and also highlights SimplyMap, a market-research demographic mapping product from Geographic Research, Inc., available on the UW-Madison campus.

The Census Bureau has in fact attempted to address the insatiable urge for the latest to-the-minute data. Two Census programs of long standing, that run hand in hand, are the Population Estimates program (<http://www.census.gov/popest/estimates.php>) and the U.S. Population Projections (<http://www.census.gov/population/www/projections/>). The Population Estimates address present and past data, while the Population Projections reference future probabilities.

U.S. Census population estimates are calculated annually, with a date of reference of July 1, and are released to the geographic level of cities and towns. The estimates include total population; population by age, sex, race, and Hispanic origin; and the number of housing units. When each new estimate is released, the previous estimates back to the prior decennial census are revised as well, and old estimates are archived. The population estimates are available at the above URL, and also through American FactFinder at <http://factfinder.census.gov/>.

The U.S. Census population projections speak to considerably larger geographic areas than the estimates, being offered only at the national and state levels. The latest state-level projections were released in 2005 and look as far

*Continued on p. 3*

## RESEARCHER'S NOTES

by Benjamin Cowan

I am a dissertator in the Department of Economics, and my research concerns the effects of college expectations on teenage behavior. To estimate these effects, I need access to variables that change college expectations without changing other factors that determine teenagers' use of drugs and alcohol, study habits, and other behaviors. One such variable is the cost of attending college in an individual's state of residence. Because in-state college tuition is more affordable in some states than it is in others, teenagers living in different states may have different expectations about the likelihood of attending or finishing college, which may be reflected in their behavior.

Data on tuition and fee rates for state colleges and universities across the country is collected by the Washington Higher Education Coordinating Board (HECB). Because it was not immediately obvious that I could procure this data directly from the source, the staff at DISC was very helpful in helping me determine which third parties (for example, the NCES Digest of Education Statistics) may also have access to the same kind of data. Eventually, I was able to get access to the data through HECB itself. As I seek to move my research in this area forward, I look forward to more interaction with the staff at DISC, as I have been impressed with their resources, knowledge, and willingness to help up to this point in my graduate studies.

## BADGIR UPDATE

by Lu Chou

DISC is pleased to announce that four important studies have been added to our BADGIR (Better Access to Data for Global Interdisciplinary Research) catalog, at <http://nesstar.ssc.wisc.edu>. BADGIR is powered by the Nesstar software suite, and allows our users to search, browse and analyze data online.

### *National Health Measurement Study (NHMS), 2005-2006*

NHMS surveyed older US adults with a suite of health-related quality-of-life (HRQoL) indexes like EuroQol EQ-5D, Health Utilities Index, SF-36v2TM, and QWB-SA. It oversampled African Americans and older individuals (ages 65+) to allow subgroup analyses. NHMS offers a unique source in comparing how these instruments describe and measure health. Downloadable documentation and data in SAS format, together with a link to BADGIR, are also available from the DISC online archive, at <http://www.disc.wisc.edu/NHMS/>.

### *Puerto Rican Elderly: Health Conditions (PREHCO) Wave 1, 2002-2003 and Wave 2, 2004-2006*

The first wave of the PREHCO Project investigated the characteristics of older non-institutionalized adults (aged 60+) in Puerto Rico through an island-wide, cross-sectional sample survey of target individuals and their spouses. Topics include self-reported health conditions, physical and mental impairment, housing arrangements, functional status, transfers, labor history, migration, income, childhood conditions, health insurance and use of health services, marital history, sexuality, and others. PREHCO2 interviewed survivors of the sample interviewed in PREHCO1. It collected information on changes in health, family and residential arrangement transitions, familial and non-familial transfers, fluctuations in income and assets, and labor force status changes. PREHCO data is useful for researchers examining issues affecting the elderly population in Puerto Rico.

### *National Survey of Families and Households (NSFH) Wave 3, 2001-2002*

NSFH investigates the causes and consequences of changes in American family and household structure over time (1987-2002). NSFH Wave 3 data files have been available from the NSFH site, <http://www.ssc.wisc.edu/nsfh/>, as SPSS sav files. With the addition of Wave 3 to BADGIR, NSFH users can now view variables, check summary statistics and run simple cross-tabulations online.

Users must be registered and have received an e-mail confirmation before they can analyze or download data using BADGIR. To register, first-time users can visit the BADGIR registration page (<http://www.ssc.wisc.edu/cdha/badgir/terms.htm?submit2=Register>). Viewing documentation and univariate statistics in BADGIR does not require registration.

## YEARNING FOR 100-COUNT DATA...

*Continued from p. 1*

forward as 2030. National-level projections were most recently released on August 14, 2008, and look forward to 2050. Projections are always superseded by estimates when the estimates become available for a formerly-projected year.

The newest and most sweeping Census Bureau attempt to address the intercensal data gap, however, has been the American Community Survey (ACS), online at <http://www.census.gov/acs/www/> and through American Fact-Finder. Since its initial pilot data collection in 1999, the ACS has been moving toward serving as a replacement for the decennial census long-form data in 2010 and future decades.

Instead of a snapshot, the ACS is better described as a continually-updated moving video image. ACS data collection is an ongoing process rather than a one-time event, and data is presented in rolling figures of one-year, three-year, or five-year estimates. Currently, 2007 ACS one-year estimates are available for geographic areas with at least 65,000 people. Starting in December 2008, three-year estimates will become available for areas with at least 20,000 people, and by 2010 there will be 5-year estimates available for areas down to the tract-level and block-level. Large areas such as states will have multiple estimates: ACS estimates for 1, 3, and 5 years, and Population Estimates program estimates as well. The Census Bureau has recently begun to release handbooks for various user groups of the ACS, at [http://www.census.gov/acs/www/UseData/Compass/handbook\\_def.html](http://www.census.gov/acs/www/UseData/Compass/handbook_def.html). As of this writing, there are handbooks available for general users and the business community.

At this point, however, the 5-year ACS estimates that address the tract level do not yet exist. In the absence of Census products, marketing research firms have created their own proprietary estimates to fill the gap—for a fee.

The UW-Madison Libraries began a subscription to such a product earlier this year, called SimplyMap, at <http://digital.library.wisc.edu/1711.web/simplymap>. SimplyMap draws on Census 2000 data, 2006 & 2007 Census estimates, and 2011 & 2012 Census projections, but takes the estimates and projections out to the ZIP code, tract, and block group level, a much more detailed level than the Census Bureau provides. Other groups of variables from additional sources include consumer expenditures, consumer price index, business counts, market segments, retail sales, and sales potential.

UW-Madison's SimplyMap subscription can be accessed from any computer on campus, or off-campus with NetID and password. Users must also create a "personal workspace" with their e-mail address and an additional password to be able to use SimplyMap. Once a workspace is created, users may select geographic locations and variables to create and export tables and maps. Online tutorial videos are also included.

## FOR UNDERGRADS: NEWS FROM ICPSR

*by Joanne Jubnke*

ICPSR has announced two undergraduate research paper competitions for 2009. The first competition, sponsored by the general archive at ICPSR, invites research papers involving quantitative analysis of data from ICPSR. The second competition, sponsored by ICPSR's Minority Data Resource Center (MDRC), invites papers addressing issues relevant to underrepresented minorities in the United States, with data to be drawn from the MDRC.

Each competition carries identical awards: first prize is \$1000, second prize is \$750, and third prize is \$500. Full details are online at <http://www.icpsr.umich.edu/ICPSR/prize/>; submission deadline is May 31, 2009.

ICPSR has also announced their summer undergraduate internship, which runs from June 8 to August 14, 2009. Interns will gain experience using statistical programs such as SAS, SPSS, and Stata for data checking and processing. Interns also attend courses in the ICPSR Summer Program in Quantitative Methods of Social Research, in addition to a weekly Lunch and Lecture series. Applicants must have completed their sophomore year in a social science major by the time of the internship. Full details are available at <http://www.icpsr.umich.edu/ICPSR/pdf/internship.pdf>; application deadline is February 2, 2009.

*Crossroads Corner highlights web sites recently added to the searchable Internet Crossroads in Social Science Data, available on the DISC web site at <http://www.disc.wisc.edu/newcrossroads/index.asp>.*

---

### Google Flu Trends

As the flu season nears, the online search giant Google is now offering a novel take on tracking influenza outbreaks in the United States, based on aggregated data from certain flu-related Google search terms. Google reports that their numbers match up remarkably well with flu surveillance data from the Centers for Disease Control (CDC flu data available at <http://www.cdc.gov/flu/weekly/fluactivity.htm>). Not only that, but since the CDC's numbers must be reported by actual physicians nationwide and compiled before release, Google's instantly-available search numbers can anticipate CDC reports of flu outbreaks by one or even two weeks. The Google data, broken out weekly by state and region back to 2003, can be downloaded in CSV at <http://www.google.org/flutrends/>.

---

### CIRCLE

The run-up to the 2008 U.S. presidential election brought much speculation regarding the enthusiasm and likely turnout of young voters. According to CIRCLE, the Center for Information and Research on Civic Learning and Engagement at Tufts University, around 23 million Americans under the age of 30 voted in 2008, up 3.4 million from the 2004 turnout. CIRCLE's research focus is "the civic and political engagement of Americans ages 15 to 25."

The data section of the CIRCLE website, at <http://www.civicyouth.org/?author=10>, carries several surveys for download or online analysis, including:

- The Civic and Political Health of the Nation survey (2002, 2004, & 2006)
- National Youth Survey, 2004

- YouthVote's June 2002 Survey of Young Americans (18-24)

The site also features factsheets, reports, and links to related sites.

---

### Real Clear Politics

#### FiveThirtyEight

In November 2006, we highlighted Pollster.com, a political-polling analysis site co-founded by UW-Madison's own Charles Franklin. The 2008 election saw an explosion of new political polling, particularly at the state level. In the wake of this month's elections, here are two additional polling-analysis sites to which poll-tracking aficionados can turn.

Real Clear Politics, whose polling section may be found at <http://www.realclearpolitics.com/polls/>, is a right-leaning political analysis site founded in 2000 by John McIntyre and Tom Bevan. The polling section of the site includes links to national and state polls, polling averages, and electoral college maps, including maps from elections back to 1968. The site also featured a fantasy marketplace game based on the Intrade online futures trading market, where participants could bet pretend-money on their political picks, at <http://fantasy08.realclearpolitics.com/>.

FiveThirtyEight by Nate Silver, at <http://www.fivethirtyeight.com/>, boasts the tag line "Electoral Projections Done Right" and takes its name from the total number of electors in the US electoral college. Silver's politics lean to the left, and his methodology leans to the complex; his site not only differentiates itself from other poll compilations by weighting polls for sample size, recentness, and pollster reliability, but also announces "we simulate the election 10,000 times for each site update in order to provide a probabilistic assessment of electoral outcomes based on a historical analysis of polling data since 1952."