

## Spotlight – UC DATA at UC Berkeley (from the March 2012 CSDC Network News)

The UC Data Archive at UC Berkeley is a university-based social science data archive, providing access, support, and preservation for a broad range of machine readable data, primarily data collected by the Census Bureau. UC DATA was founded in 1962 to assist social scientists in obtaining, processing, and analyzing domestic and foreign survey materials, and has adapted to the development of technology and explosion of available resources over the last 50 years.

UC DATA began providing assistance and access to census data with the release of the summary data from the 1970 Census. At Berkeley, a Census Service Facility was established as a cooperative effort between the Institute of Governmental Studies and the Survey Research Center, providing a large set of standardized and customized tabulations to meet individual user requests. These 1970 census summary data became what is now known as the “thin edge of the wedge”, paving the way for much broader distribution of machine-readable data from the census and other sources. When the State Data Center network was founded in 1979, UC DATA joined with the primary responsibility of serving academic data users.

By the early 1990's, UC DATA had partnered with the Lawrence Berkeley Laboratories and the UC Berkeley Library to provide the first public Internet access to the 1990 Census summary tape files. It subsequently built on these collaborations to design and provide access to the subject summary detail data for race and ethnicity groups for 1990. More recently, UC DATA has partnered with the Census Bureau to host one of the twelve Research Data Centers (RDCs) in the U.S.

RDCs are highly secure Census Bureau facilities which were established to address a pair of complementary research needs. On the part of the Census Bureau, there is a need for innovative research and expertise to improve the quality, utility, and understanding of Census Bureau products and collections. On the part of researchers, there are questions and analyses that cannot be adequately explored and addressed using existing published microdata or tabulations. The Census Bureau evaluates proposals by qualified researchers that address both of these needs. Approved projects are provided secured access (fee-based) to restricted-use microdata for statistical purposes only in the secure RDC setting.

Data in the RDCs include both household surveys -- such as the decennial censuses, American Community Surveys (ACS), Current Population Surveys (CPS), and Surveys of Income and Program Participation (SIPP) -- as well as business and establishment surveys, like the Economic Census and annual surveys conducted by the Census Bureau. Other federal agencies also permit access to the restricted-use holdings in the RDC setting, such as the National Center for Healthcare Statistics (NCHS), the Agency for Healthcare Research and Quality (AHRQ) and the Bureau of Justice. More information about the RDC network can be found at <http://www.census.gov/ces/rdcresearch/>.



Jon Stiles

Jon Stiles, the Executive Director for the Census RDC and an archivist for UC DATA, sees the partnerships between the SDC and RDC networks as two halves of the data access puzzle. “Every user’s needs differ to some extent. Data are increasingly ubiquitous, and the challenge is to provide the right level of access and assistance to users, while preserving the confidentiality of survey respondents and the integrity of future data collection. This requires a tailored approach – and that is what we strive to provide.”