Recovering Microdata of the 1950's through the 1980's from the Census Bureau's Last Unisys Mainframe

Date and Time: Thursday, October 22nd, 2015
12:30 – 3:00 p.m.

Moderator: Pam McGovern, National Agricultural Statistics Service

Location: Bureau of Labor Statistics Conference Center

To be placed on the seminar attendance list at the Bureau of Labor Statistics, you need to e-mail your name, affiliation, and seminar name to wss_seminar@bls.gov (underscore after 'wss') by noon at least two days in advance of the seminar." Please bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Use the Red Line to Union Station. Parking in the area of BLS is available at Union Station. For parking information see http://www.unionstationdc.com/parking. No validation is available from BLS for reduced parking rates.

Sponsor: WSS Methodology Section

WebEx event address for attendees:

https://dol.webex.com/dol/j.php?MTID=m934dd06d6115c4ea842545c3182d29db

For audio: Call-in toll-free number (Verizon): 1-866-747-9048 (US)
Call-in number (Verizon):   1-517-233-2139
(US) Attendee access code: 938 454 2

Note: Particular computer configurations might not be compatible with WebEx.

Schedule

Time	Speaker	Affiliation	Point of Contact
12:30	Pam McGovern	NASS	Pam.McGovern@nass.usda.gov
12:40	Mark Mildorf	Census (retired)	mark.f.mildorf@gmail.com
1:20	Randy Becker and	Census	randy.a.becker@census.gov
	Cheryl Grim	Census	Cheryl.Ann.Grim@census.gov
2:00	Todd K Gardner	Census	Todd.K.Gardner@census.gov
2:40	Questions and Answers		

Abstracts

Files Retrieved: A Brief Look Inside the Recovery Mission
Speaker: Mark Mildorf, Retired from U.S. Bureau of the Census, Chief Planning Officer in the Research and Methodology Directorate

In early 2010, the Census Bureau decommissioned its last remaining Unisys mainframe, ending a thread of computer usage that can be traced back to the first commercial digital computer, UNIVAC I, acquired by the Census Bureau in 1951. In the months leading up to decommissioning, a small team in the Census Bureau's Center for Economic Studies (CES) led an effort to retrieve hundreds of data files that remained on this computer. These files included data from Economic and Demographic surveys and censuses, several Decennial Censuses and other files of interest from the late 1950s through the 1980s. This retrieval effort was necessary due to the unique characteristics of the Unisys mainframe and the data files stored on it; once this mainframe was decommissioned, the Census Bureau would have no ability to read or access these files.

Public Use Microdata Sample of 1950
Speaker: Todd K Gardner, U.S. Bureau of the Census

In the last few years the Census Bureau has devoted considerable resources to recovering internal historical microdata files for research purposes. One example of this is an enhanced version of the 1950 Public Use Microdata Sample (PUMS). The 1950 PUMS is a one-percent sample of the 1950 decennial census that was produced in the 1980s in a collaboration between the Wisconsin Center for Demography and Ecology (CDE) and the Census Bureau. A newly created internal version of the 1950 PUMS combines three files into one: 1) the original 1950 PUMS that was released to the public, 2) the IPUMS version of this dataset with its harmonized coding, and 3) the original data entry files from the CDE project, which contain detailed geography and the original alpha strings for fields like birthplace and occupation. Using the geographic information in the alpha files, and with the help of a genealogical website run by Steve Morse, we were able to identify all census tracts as they were defined in 1950.

Research Opportunities
Speakers: Randy Becker and Cheryl Grim, U.S. Bureau of the Census

We will offer an overview of the research opportunities available through the Federal Statistical Research Data Center program, including the restricted-use Census Bureau microdata currently available. We then discuss the wide variety of business data recovered in recent years and the opportunities and challenges associated with their use. We then highlight some of the progress made with the recovered data, with a particular focus on manufacturing data. We also discuss ongoing efforts to recover data from legacy tapes.