

U.S. Census Microdata

Louise Klusek

Newman Library, Baruch College

May 15, 2007

What is Microdata?

- Original sample data containing individual responses to Census surveys
- Coded and edited in data files that can be analyzed with statistical software
- Available in Public Use Microdata Samples (PUMS) and in Restricted Access files

Is Microdata for you?

- You need large data samples
- You want to use particular units of analysis or population characteristics
- You need historical data for a longitudinal study
- You want to manipulate the data with statistical tools (SAS, SPSS)

Census Long Form Questionnaire

- Since 1960, sampled households and individuals in those households
- Sampled one in six households in the 2000 Census
- Asks for social & economic data about the household and its members
- Full questionnaire for Census 2000
www.census.gov/dmd/www/2000quest.html

Census Content

- **Demographic**
 - Sex, age, relationship, race, Hispanic origin...
- **Social**
 - Education, marital status, disability, school enrollment, language, citizenship status...
- **Economic**
 - Income and benefits, poverty, employment, occupation, industry, commute to work...
- **Housing**
 - Occupancy, structure, housing value and costs, rents, mortgage, utilities...

Public Use Microdata Files (PUMS)

- Subsets of the Census long-form questionnaire items selected to represent either 1% or 5% of U.S. housing units and the persons in the housing unit
- Based on large reporting areas called PUMAs or Super-PUMAs that mask identification of individuals or households
- PUMS data files are available for every U.S. Census since 1960

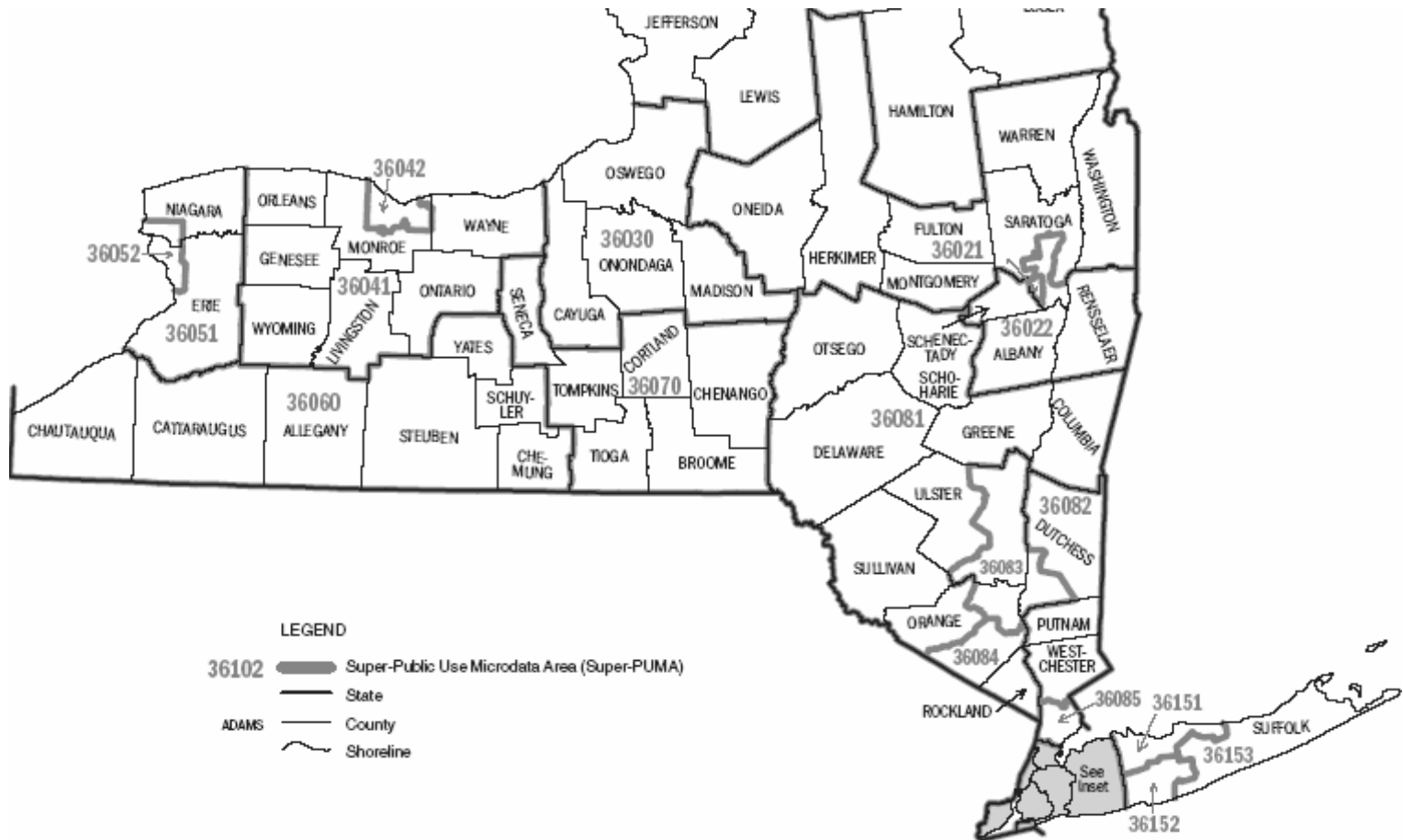
5% PUMS Geographic Reporting

- PUMA's must meet a population threshold of 100,000
- New York state PUMA boundaries are counties or contiguous Census tracts
- New York City PUMA boundaries approximate Community District boundaries

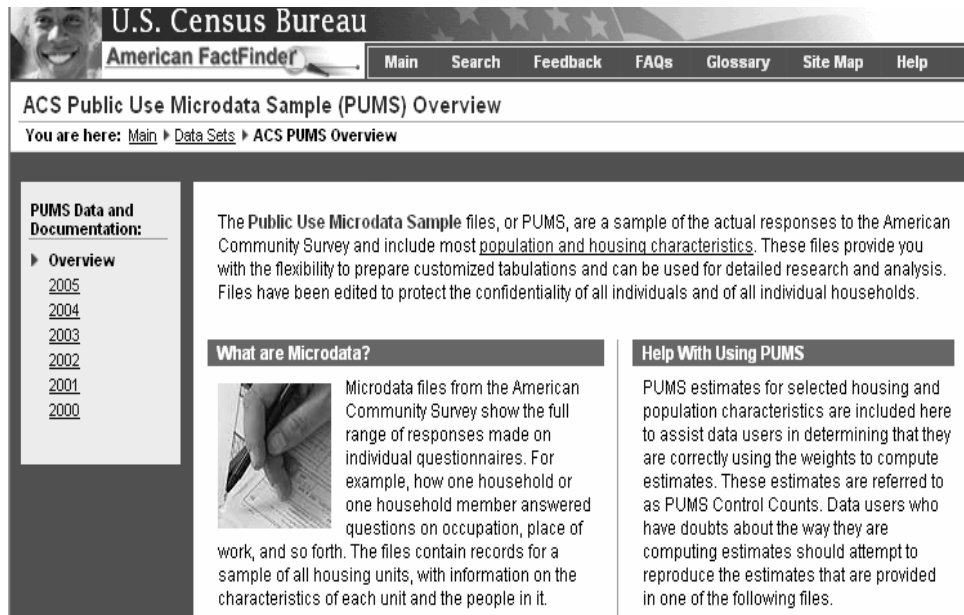
1% PUMS Geographic Reporting

- Super-PUMAs are contiguous PUMAs within state boundaries
- Super-PUMAs have a population threshold of 400,000

NY PUMA Map



Starting out with PUMS



The screenshot shows the U.S. Census Bureau American FactFinder website. The header includes the U.S. Census Bureau logo and the American FactFinder search bar. The main navigation menu contains links for Main, Search, Feedback, FAQs, Glossary, Site Map, and Help. The page title is "ACS Public Use Microdata Sample (PUMS) Overview". Below the title, a breadcrumb trail reads "You are here: [Main](#) > [Data Sets](#) > [ACS PUMS Overview](#)".

PUMS Data and Documentation:

- Overview
- [2005](#)
- [2004](#)
- [2003](#)
- [2002](#)
- [2001](#)
- [2000](#)

The Public Use Microdata Sample files, or PUMS, are a sample of the actual responses to the American Community Survey and include most population and housing characteristics. These files provide you with the flexibility to prepare customized tabulations and can be used for detailed research and analysis. Files have been edited to protect the confidentiality of all individuals and of all individual households.

What are Microdata?

Microdata files from the American Community Survey show the full range of responses made on individual questionnaires. For example, how one household or one household member answered questions on occupation, place of work, and so forth. The files contain records for a sample of all housing units, with information on the characteristics of each unit and the people in it.

Help With Using PUMS

PUMS estimates for selected housing and population characteristics are included here to assist data users in determining that they are correctly using the weights to compute estimates. These estimates are referred to as PUMS Control Counts. Data users who have doubts about the way they are computing estimates should attempt to reproduce the estimates that are provided in one of the following files.

http://factfinder.census.gov/home/en/acs_pums.html

Key Microdata Files

- Census 2000
- American Community Survey
- Current Population Survey
 - www.census.gov/cps/
 - <http://stats.bls.gov/cps/>
- American Housing Survey
 - www.huduser.org/datasets/
 - www.census.gov/hhes/www/housing/ahs/ahs.html

Census 2000 Gateway

U.S. Census Bureau



Your Gateway to Census 2000

Thank you, America, for your participation in Census 2000.
The population of the U.S. on April 1, 2000 was **281,421,906** [PDF 2M].

► [Census 2000 Gateway](#) | [Glossary](#)

- FTP access to 1% and 5% PUMS in state-level files
- Documentation and PUMS maps
- <http://www.census.gov/main/www/cen2000.html>

American Community Survey

- National Survey: ACS covers every state and all areas with a population of 65,000 or more
- Annual Data: Social, demographic, economic and housing
- ACS will be the basis for a re-engineered 2010 Census
- <http://www.census.gov/acs/www/>

Using the ACS through American Factfinder

ACS Public Use Microdata Sample (PUMS) 2005

You are here: [Main](#) ▶ [Data Sets](#) ▶ [ACS PUMS 2005](#)

PUMS Data and Documentation:

- [Overview](#)
- ▶ **2005**
- [2004](#)
- [2003](#)
- [2002](#)
- [2001](#)
- [2000](#)

Note: The population estimates of the American Community Survey (ACS) data are currently limited to the household population and exclude the population living in institutions, college dormitories, and other group quarters.

Download 2005 PUMS Data

■ Select a data type, data format, and state. Click 'GO'.

Data Type

- Population Records
- Housing Records

Data Format

- CSV (comma separated values)
- PC SAS Data Set
- UNIX SAS Data Set

State

United States

Documentation

- [Subjects available in PUMS files](#)
- [2005 PUMS top coded values](#)
- [2005 PUMS Code Lists](#)

[Ancestry](#)

[CMSA](#) (Note: CMSAs are not available on the ACS PUMS)

[Hispanic Origin](#)

[Industry](#)

[Language](#)

[Migration](#)

[Occupation](#)

[Place of Birth](#)

[Place of Work](#)

[Race 1 \(9 categories\)](#)

[Race 2 \(67 categories\)](#)

[Race 3 \(72 categories\)](#)

[State](#)

Integrated Public Use Microdata Sample (IPUMS)

MPC Minnesota Population Center
University of Minnesota

Integrated Public Use Microdata Series
census microdata for social and economic research

- High-precision U.S. population samples covering years 1850 to 2004 in IPUMS-USA
- Uniform coding across samples to facilitate use

<http://www.ipums.umn.edu/>

<http://usa.ipums.org/usa/>

Census 2000 from ICPSR



- ICPSR Direct is the world's largest archive of social science datasets
- ICPSR Census 2000 data and documentation - <http://www.icpsr.umich.edu/CENSUS2000/>
- SAS and SPSS commands to process Census data

Microdata from CRDC

- CRDC data is non-public microdata
- CRDC restricted access files contain all the long-form questionnaire responses
- The CRDC provides the only access to microdata from the Economic Census
- The CRDC ensures confidentiality of data by screening that certifies researchers and their research designs

Newman Library Can Help

- Identify relevant datasets
- Support ICPSR users
- Advise on mapping and use of geospatial data
- Help students with Census resources and databases