Brief History of Public Use Microdata Areas (PUMAs)

Summary
Public Use Microdata Sample (PUMS) files contain individual records of the characteristics for a sample of persons and households.[1] PUMS files were created for the 1960, 1970, 1980, 1990, and 2000 decennial censuses. Samples from the 1970 through 2000 decennial censuses employed a 5-percent sample size; a 1-percent PUMS sample has been produced since 1980. In all cases, PUMS data are tabulated and published for selected geographic entities. Until 1990, the geographic units for which PUMS data were available included states, counties, county groups, metropolitan areas and urban/rural classifications. Since 1990, Public Use Microdata Areas (PUMAs) have been delineated by State Data Centers (SDCs) as a unique type of geographic area for the tabulation and dissemination of PUMS files. For Census 2000, the standard 5-percent PUMAs were defined along with the new 1-percent super-PUMAs, which were based upon aggregations of the smaller, 5-percent standard PUMAs. Provided below are the highlights of decennial PUMS and PUMA history and geography through the years.

1960
The Census Bureau first created PUMS files as 1960 Census data products.
PUMS files: Only one type of PUMS file was published for the 1960 census based on a 12-percent sample.
PUMS geography: The PUMS files identified states as the smallest geographical unit.

1970
PUMS files: Two PUMS files were created for the 1970 Census. The Census Bureau used two long-form questionnaires to survey the population- one for a 15-percent sample, and another for a 5-percent sample.
PUMS geography: Three geographic schemes were applied to publication of the two samples – state, metro (metropolitan areas[2] and county groups), and neighborhood areas. Both the 15-percent and the 5-percent PUMS products were published to each of the geographic schemes.

The PUMS data presented for states identified the state, as well as the urban/rural, metropolitan/non-metropolitan, and central city/outside central city for each state.

The PUMS data presented for metropolitan areas and county groups identified all Standard Metropolitan Statistical Areas (SMSAs) with a population of 250,000 persons or greater, aggregations of counties that met that minimum population threshold, and areas outside the SMSAs with a minimum population of 250,000 persons. In New

[1] Further information about microdata, PUMS, and confidentiality of PUMS data is provided on the Census Bureau website. Please refer to http://www.census.gov/main/www/pums.html and http://www.census.gov/acs/www/data_documentation/public_use_microdata_sample/.
[2] Refer to the Historical Metropolitan Area Definitions page for information regarding the metropolitan area definitions used in the provision of 1970, 1980, 1990 and 2000 PUMS publication: http://www.census.gov/population/www/metroareas/pastmetro.html

England, where SMSAs were defined based on cities and towns, the SMSA was approximated in terms of entire counties. In Hawaii, the city of Honolulu and the remainder of the state were identified. The PUMS data presented based on neighborhood area delineations were identified by census regions or divisions only. These neighborhood delineations could cross state boundaries, and were required to meet the minimum population threshold of 250,000 persons.

1980

PUMS files: For the 1980 Census, the Census Bureau published three different PUMS files - a 5-percent sample, and two 1-percent samples.

PUMS geography: The PUMS 5-percent sample identified each state and various "county groups" within the state with a minimum population of 100,000 persons, individual counties with at least 100,000 persons, and metropolitan areas or groups of places meeting the minimum threshold of 100,000 persons . SDCs grouped counties with populations less than 100,000 persons into analytical areas that frequently followed metropolitan areas or state planning district boundaries. In New England, aggregations of cities and towns were used instead of aggregations of counties.

The PUMS 1-percent metro sample individually identified Standard SMSAs with a population of 100,000 persons or greater , combined SMSAs with smaller populations (to meet the minimum 100,000 population threshold), and divided remaining state areas into county groups. Specifically, the metro sample identified 282 SMSAs and 20 states with county groups and places with populations of 100,000 persons. The remaining 31 states were not identified because they contained SMSAs that crossed state boundaries and did not meet the minimum threshold of 100,000 persons within a state. An additional 36 SMSAs were paired together so that metropolitan and non-metropolitan territory could be analyzed separately. Areas outside SMSAs were grouped according to a state planning district, or into other reasonable analytical areas with a minimum population of 100,000 persons.

The PUMS 1-percent urban/rural sample identified regions, divisions, and most states by type of area—central cities of urbanized areas, urban fringe (the remainder of urbanized areas outside central cities), other urban, and rural. The sample identified 27 states, the District of Columbia, 58 places, and 73 urbanized areas, each meeting the minimum population threshold of 100,000 persons. The remaining states were divided into 8 groups, each nesting within a census region or division boundary.

The state and metro sample groups (the 5-percent and 1-percent metro sample) were applied to place-of-work in 1980 and place of residence in 1975 to allow detailed analyses of migration and commuting patterns that included origins and destinations.

1990

PUMS files: For the 1990 Census the Census Bureau published three PUMS files—a 5-percent, 1-percent, and 3-percent sample.

PUMS geography: For the publication of 1990 Census PUMS, the Census Bureau created a new geographic entity type called the Public Use Microdata Area (or PUMA).

These PUMAs were delineated by the SDCs using Census Bureau criteria and guidelines[3]. Two levels of PUMAs were delineated: a PUMA to publish the 5-percent sample microdata, and a PUMA to publish the 1-percent sample microdata.

The 5-percent PUMAs (or PUMAs to publish 5-percent sample data) contained a minimum threshold of 100,000 persons and did not cross state boundaries. These PUMAs could be aggregations of counties and places. In New England, 5-percent PUMAs were defined by cities and towns instead of counties. PUMAs with populations of over 200,000 persons may have included parts of counties or places. In some sparsely populated areas, 5-percent PUMAs included one or more noncontiguous parts to meet the minimum population criteria. Where the combinations of state and metropolitan areas would enable the identification of areas smaller than 100,000 persons, the 5-percent state sample suppressed the metropolitan area information.

The 1- percent PUMAs (or PUMAs to publish 1-percent sample data) were built primarily on metropolitan areas, groupings of non-metropolitan counties, and other large areas with at least 100,000 persons and were permitted to cross state lines. They could be composed of whole central cities, whole metropolitan areas and metropolitan areas outside the central city, groups of metropolitan areas, and groups of areas outside metropolitan areas. Additionally, many large cities, groups of cities and counties were identified within large metropolitan areas. When areas contained more than 200,000 persons, 1-percent PUMAs could comprise parts of central cities or metropolitan areas. Outside metropolitan areas, counties were grouped according to state planning districts or other analytic units with populations of 100,000 persons or more. Where combinations of state and metropolitan areas could enable the identification of areas smaller than 100,000 persons, the 1990 metro sample suppressed state information.

The 3-percent elderly sample was based on the same geographical components as the 5-percent PUMAs. This sample exclusively surveyed households containing at least one person of age 60 or older. In addition to 5-percent PUMAs, the elderly sample used State Planning Service Areas (PSAs).

2000

PUMS files: For Census 2000, the Census Bureau published two PUMS files (a 5-percent and a 1-percent) extracted from Census 2000 long-form questionnaire responses. The 1-percent PUMS files provided a full range of detailed characteristics for larger geographic areas (super-PUMAs). The 5-percent PUMS files provided less characteristic detail for smaller geographic areas (PUMAs).

PUMS geography: To define the two geographic areas, the Census Bureau offered the SDCs the opportunity to follow bureau criteria and recommendations to delineate, or coordinate, the delineation of PUMAs. The SDCs in 48 states, the District of Colombia, and Puerto Rico participated in the delineation program. The Florida and Rhode Island SDCs chose not to participate; therefore, the Census Bureau delineated super-PUMAs and PUMAs for those two states.

[3] The Census Bureau delineated 1990 PUMAs for Georgia, Indiana, and Oregon because the SDCs in those states declined to do so.

The 5-percent PUMAs were used to present the 5-percent PUMS files, were required to contain a minimum population of 100,000 persons, and had to nest within states. PUMAs could consist of a single county or an aggregation of one or more counties, census tracts, or minor civil divisions (MCDs) in the New England states. Additionally, an incorporated place with a minimum population of 100,000 persons could be defined as a PUMA.

The 1-pecent super-PUMAs were used to present 1-percent PUMS files, were required to contain a minimum population of 400,000 persons, and had to nest within states. These super-PUMAs were a new geographic entity for Census 2000 and were aggregations of the smaller, 5-percent PUMAs.

Migration (MIG) PUMAs and place-of-work (POW) PUMAs were created for use in the publication of MIG and POW microdata files. These geographies were based on one or more 5-percent PUMAs. In cases where 5-percent PUMAs encompassed one or more whole counties, the MIG and POW PUMAs were equivalent to the 5-percent PUMA geography; however, when 5-percent PUMAs contained other types of geographic entities, the MIG and POW PUMAs were based on aggregations of two or more 5-percent PUMAs in order to encompass whole counties. Within the six New England states (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, and Connecticut), MIG and POW PUMAs encompassed minor civil division (city and town)-based 5-percent PUMAs.

Use of PUMAs in the American Community Survey (ACS)[4] post-Census 2000

The ACS was fully implemented in the United States and Puerto Rico in January 2005. Census 2000 PUMAs have been used in the tabulation and dissemination of ACS PUMS data and to present ACS period estimates. In 2008, the Census Bureau released its first 3-year PUMS and estimates based on ACS data collected from 2005 through 2007. The first 5-year PUMS and estimates, based on ACS data collected from 2005 through 2009, were released on December 14, 2010.

[4] Information about the ACS can be found at: http://www.census.gov/acs/www/. The ACS is conducted in the United States and in Puerto Rico. In Puerto Rico the survey is called the Puerto Rico Community Survey (PRCS).