

Introduction to U.S.
Census/ACS Data at CSSCS

Updated by Tina Tian

Center for Social Science Computation & Research
110 Savery Hall
University of Washington
Seattle WA 98195 U.S.A.
(206)543-8110

August 2010

<http://julius.csscr.washington.edu/pdf/census10>

Required by the U.S. Constitution, the U.S. decennial Census has been conducted every ten years since 1790 to count the population and housing units for the entire nation. The American Community Survey (ACS) is a new approach to producing critical information about the characteristics of local communities. It will eliminate the need for a long form in the 2010 Census and is a key part of the Census Bureau's Decennial Census Program. Started in 2006, every year the ACS can support the release of single-year estimates for geographic areas with populations of 65,000 or more; also accumulate sample over 3-year produce estimates for geographic areas with populations of 20,000 or more, and 5-year intervals to produce estimates for smaller geographic areas including census tracts and block groups.

The Census/ACS provides population counts that determine how seats in the U.S. House of Representatives are apportioned. Census/ACS data are also widely used in drawing congressional and state legislative boundaries, allocating federal and state funds, formulating public policy, and assisting with planning and decision-making in the public and private sectors.

Census/ACS Files Available at CSSCR

As one of the distribution centers for the Census Bureau's electronic products, CSSCR houses Census data from 1980 - 2010 and ACS 2005 - Present, including Summary Tape Files (STF), Summary Files (SF/ACS-SF) and Public Use Micro Samples (PUMS/ACS-PUMS) in various of statistical software formats. On request, we also help retrieve historical U.S. Census data from the IPUMS-USA database available at <http://www.ipums.org>. IPUMS-USA is a comprehensive national census database containing integrated microdata from 1850 to the current. We host a website for the Washington State decennial Census data files and documentation at <http://julius.csscr.washington.edu/Decennial%20Census.htm>, and ACS data files and documentation. http://julius.csscr.washington.edu/american_community_survey.htm. In addition, CSSCR also distributes some TIGER/Line files, a Census geographic product.

STF/SF and PUMS are two major types of Census/ACS data products. STF/SF contains tabulated summary statistics for different levels of Census geography from the nation, the state, the county, down to the city block. STF/SF 1 and 2 present tabulated data from the Census short-form (100%) questionnaire. STF/SF 3 and 4 present cross-tabulations of information from the long-form (sample) questionnaire. Tables in STF/SF 2 and 4 are iterated for many detailed racial groups, as well as American Indian and Alaska Native tribes. In SF 4, many data are also tabulated by detailed ancestry groups.

PUMS files present a sample of individual records of responses to the long-form questionnaire with unique identifiers (such as addresses, names, etc.) removed to protect individual confidentiality. Different from the STF/SF which provides summary data information for a specific geographic entity, in PUMS, the basic unit of analysis is a housing unit or a person. PUMS files allow users to prepare their own customized tabulations of most population and housing subjects. There are two sets of PUMS: 5-percent sample file (PUMS-A file) and 1-percent sample file (PUMS-B file). Besides the obvious difference in sampling size, the 5-percent and 1-percent files differ in the geography around which the files are constructed. The 5-percent sample is basically a county level file. The 1-percent sample is basically a metropolitan file. The Public Use Microdata Area (PUMA) is the lowest level of geography identified on any PUMS file. For the 5-percent file, the PUMA can be a single county (or county equivalent), a group of counties, a place, or county/place parts if that county had more than 100,000 persons; for the 1-percent file, the PUMA is an MSA, groups of MSAs, parts of MSAs when the MA is larger than 400,000 persons, and groups of non-metropolitan

areas. For ACS-PUMS, the single year ACS-PUMS is 1-percent file, while the 3-year ACS-PUMS is 3-percent file and the 5-year ACS-PUMS is 5-percent file. The PUMA is also the lowest level of geography identified on any ACS-PUMS files, which is equal to 5-percent sample's PUMA with 100,000 persons+. At CSSCR, 1990 PUMS data files are available in SPSS format for all states.

TIGER is the acronym of Topological Integrated Geographic Encoding & Referencing. TIGER/Line files are extracts of geographic and cartographic information from the TIGER system, describing line segments that represent physical features of the entire U.S. such as roads, railroads, rivers, lakes, political boundaries, census statistical boundaries. The database contains information about these features such as their location in latitude & longitude, name, type of feature, address ranges for most streets, the geographic relationship to other features, and other related information. These files are not graphic images of maps, but rather digital data describing geographic features. With appropriate mapping or GIS software that can import TIGER/Line data, users can produce maps ranging in detail from a neighborhood street to the entire U.S.. The latest version of Census 2000 TIGER/Line files are the 108th CD Census 2000 TIGER/Line files which contain the 108th Congressional Districts. The 108th CD Census 2000 TIGER/Line files also contain the corrected Census 2000 Urban Areas and the 1990 Urban Areas, redefined based on the Census 2000 Urban and Rural criteria. CSSCR website hosts the 108th Census 2000 TIGER/Line files for the State of Washington. The later version of TIGER/Line files are accessible from Census Bureau website <http://www.census.gov/geo/www/tiger/>.

Data Access Media

Most of our Census/ACS data for Washington State and some of national level's data are accessible through our Web at <http://julius.csscr.washington.edu/data.htm> ; and they are available in different formats: ASCII, SAS, SPSS and STATA. We also have free FTP access to the Census files archived by the Inter-University Consortium for Political & Social Research (ICPSR) (<http://www.icpsr.umich.edu/access/subject.html#I>). Most online Census data are flat ASCII files which require programming to become usable. For some files, SPSS or SAS data definition statement files are provided along with the data.

To access Census/ACS data for nation or other states, please contact with Data Archivist at txtian@u.washington.edu .

CensusCDs from Geolytics

Geolytics publishes detailed Census demographic, housing and geographic data on CDROM with innovative and easy-to-use data extract and mapping software. It allows users to choose geographic entities and data items and easily create reports, dBase files, tab- or comma-delimited ASCII files, thematic maps, and Arcview files.

CSSCR currently has following several sets of CensusCDs/DVDs from Geolytics and plans to purchase more when relevant Census/ACS data become available in the future. The Geolytics products available at CSSCR are: CensusCD+Map, CensusCD Blocks, CensusCD 1980, StreetCD, StreetDVD 2007, Historic CensusDVD, 2000 CensusDVD, and Normalized CensusDVD.

CensusCD+Map contains data from U.S. 1990 Census STF 3 A, B, C & D (3400 variables), 1999 Population Estimates and 2004 Projections, 1999 Consumer Expenditure data, County Time Series

Data back to 1970 for crime, industry, federal spending, vital statistics, etc., historical county level population since 1790, and boundaries for all geographies.

CensusCD Blocks contains demographic and housing data and map boundaries for 7 million blocks from the 1990 Census STF 1B and PL94-171 files, the TIGER boundaries, and over 50 geographic identifiers including the 1980 FIPS codes.

CensusCD 1980 contains 1980 Census STF 1 & 3 data and boundary files from the block group level and up.

StreetCD contains all layers of map data from TIGER/Line 1999. Its built-in Windows software let users easily select the layer and geography, view output data and maps, and export the selected data into documents or a database. Users can generate layer data in two formats: Arcview shape files and MapInfo MID/MIF files. This integrated data product significantly reduces the difficulty of acquiring and converting TIGER 1999 street and boundary data.

StreetDVD 2007 is an integrated data product that has all the layers of the 2007 TIGER/Line® Shapefiles. It combines geographic data with a user friendly interface for quick and easy data extractions, that can be instantaneously exported to ArcGIS and Mapinfo formats.

Historic CensusDVD contains 1960 long form, 1970 long form, 1980 long form and 1990 long form.

- **1960 CensusCD** Urban Long Form includes data for 22,871 tracts from the urban parts of the country. The US Census Bureau did not fully tract the US until 1990 so there is only data available for the urban areas (or what was urban in 1960).
- **1970 CensusCD** contains Long Form data with over 5,000 variables. It has data and mapping capabilities for counties and tracts, and data only for MCDs and places. The country was not fully tracted in 1970; there is no tract data available in untraced areas.
- **1980 CensusCD** contains 1980 Census Long and Short Form data sets. This CD includes 1500 variables: 1200 from the Long Form and 300 from the Short Form; data and mapping capabilities for county, MCD, place, tract and block group. The country was not fully tracted in 1980; there is no tract or block group data available in untraced areas.
- **1990 CensusCD** contains Long Form (SF3) in 11 boundaries such as tracts, block groups, zip codes, MSA and more, and Short Form (SF1) blocks data.

2000 CensusDVD contains Long Form (STF3), Short Form (STF1), Redistricting, and both block level products.

2010 CensusDVD only contains short form questionnaires, Summary File 1 (SF 1) and Summary File 2 (SF 2), with only seven questions. The lowest level of geography is the block for SF1 and the census tract for SF2 as in Census 2000.

2005-2009 ACS in 2000 Boundaries

2006-2010 ACS in 2010 Boundaries

2007-2011 ACS in 2010 Boundaries

Normalized CensusDVD contains Neighborhood Change Database (NCDB), 1990 Long Form in

2000 Boundaries and 1980 Long Form in 2000 Boundaries

- **Neighborhood Change Database (NCDB) CD** contains 1970, 1980, 1990, and 2000 Long Form data with details. It gives users instant access to U.S. Census data from 1970, 1980, 1990 and 2000 at the census tract level, and allows users to analyze changes that have occurred in U.S. neighborhoods over four decades.
- **1990 Long Form in 2000 Boundaries CD** allows users to access U.S. Census data from 1990 and easily compare it with the 2000 Census data in 2000 boundaries.
- **1980 in 2000 Boundaries CD** contains both the 1980 Long Form (STF-3) and Short Form (STF-1) datasets, it allows users to access U.S. Census data from 1980 and easily compare it with the 2000 Census data.
- **1990 Long Form in 2010 Boundaries CD** allows users to access US Census data from 1990 and easily compare it with the 2010 Census data in 2010 boundaries.
- **2000 Long Form in 2010 Boundaries CD** allows users to access US Census data from 2000 and easily compare it with the 2010 Census data in 2010 boundaries.

All the Geolytics products are accessible from the computer in room 119 at CSSCR.

LandView 6

CSSCR also archives a TIGER related product on DVD called LandView, which is a federal geographic data viewer. LandView 6 contains database management software and mapping software that allows users to create thematic maps of Census 2000 data, browse and query the Census, EPA and USGS databases and show the query results on the map. It also provides the capability to locate a street address or intersection on a map based on Census 2000 TIGER/Line road features and address ranges. With its mapping software, users can create large scale maps showing Census 2000 legal and statistical entities, EPA regulated site locations and USGS GNIS features. It allows users to customize the maps by varying the scale and controlling which map layers are shown. The product also provides search capability for map objects based on radius or map layer and offers tools to users to add information to the maps.

To access Census CDROM/DVDs or Census files provided by ICPSR, please contact CSSCR Archivist at (206) 543-8110 or email txtian@u.washington.edu.