Finding and Using Census & Demographic Data

Jennie Murack, Geospatial Data Librarian, murack@mit.edu

Nick Albaugh, Librarian for Innovation & Entrepreneurship and Economics, nalbaugh@mit.edu
US CENSUS GEOGRAPHY: HOW DATA ARE COLLECTED AND AGGREGATED
THE American Community Survey

This booklet shows the content of the American Community Survey questionnaire.

Start Here

Please print today's date.
Month: __  Day: __  Year: __

Please print the name and telephone number of the person who is filling out this form. We may contact you if there is a question.

Last Name: ____________________________________________
First Name: ____________________________________________

Area Code: +  Number: ________________________________

How many people are living or staying at this address?

INCLUDE: everyone who is living or staying here for more than 2 months.

DO NOT INCLUDE: people who are living here for less than 2 months.

If you have another place to stay even if you are here for 6 months or less, DO NOT INCLUDE anyone who is living elsewhere for more than 6 months or less.

Number of people:

Copyright 2016 by the Metropolitan Area Studies Group, Inc.

For more information on the American Community Survey, visit our website at: http://americancensus.gov/acs
Standard Hierarchy of Census Geographic Entities

NATION

REGIONS

DIVISIONS

STATES

Counties

ZIP Code Tabulation Areas
School Districts
Congressional Districts
Voting Districts
Traffic Analysis Zones
County Subdivisions
Census Tracts
Subminor Civil Divisions
Block Groups
Census Blocks

Urban Areas
Core Based Statistical Areas
Urban Growth Areas
State Legislative Districts
Public Use Microdata Areas

AIANNH Areas*
(American Indian, Alaska Native, Native Hawaiian Areas)

MIT Libraries, 2017
Census Geographic Areas

Blocks
- Smallest units of data tabulation
- Bounded by visible features or legal boundaries
- Size: average about 100 people.

Block Groups
- Groups of blocks within same census tract sharing the same first block number digit
- Smallest areas for which sample data are available
- Size: range between 600 to 3,000 people

Census Tracts
- Seamless national coverage
- 65,000 Census tracts across U.S.
- Boundaries try to be maintained over time for comparison
- Size: range between 1,200 and 8,000 people
US CENSUS SURVEYS
Surveys

• Decennial Census
  – Collected every 10 years
  – Basic demographic data (age, sex, race, Hispanic/Latino origin, household relationship, residence owned or rented)

• American Community Survey:
  – Collected continuously and aggregated into 1, 3, or 5 years estimates
  – Detailed data on social and economic characteristics of residents, physical characteristics of the housing unit, transportation, and financial characteristics of housing.
  – Larger margin of error, especially with the smaller, 1 year samples
American Community Survey Data
Availability

1-year estimates
- geographic areas with population \( \geq 65,000 \)
  - Nation, States, congressional districts, \( \sim 800 \) counties, and 500 metropolitan and micropolitan statistical areas

3-year estimates
- geographic areas with a population \( \geq 20,000 \)
  - Nation, states, congressional districts, \( \sim 1800 \) counties (~1400 counties NOT included), 900 metro and micropolitan statistical areas

5-year estimates
- Geographic areas with populations < 20,000
- Down to block group level geography

MIT Libraries, 2017
Margin of Error (MOE)

• The precision of the estimate at a given level of confidence.
• Census standard confidence level is 90%.
• Users can calculate MOE for other confidence levels.
• MOE is relative to the size of the estimate.
  – Larger sample size usually means lower MOE
• MOEs are provided in every ACS data table for each variable.
• In general, the larger the MOE, the less reliable the estimate.
<table>
<thead>
<tr>
<th>Estimate</th>
<th>Margin of Error (+/-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barnstable County</td>
<td>1,245</td>
</tr>
<tr>
<td>Berkshire County</td>
<td>1,590</td>
</tr>
<tr>
<td>Bristol County</td>
<td>758</td>
</tr>
<tr>
<td>Dukes County</td>
<td>7,379</td>
</tr>
<tr>
<td>Essex County</td>
<td>812</td>
</tr>
<tr>
<td>Franklin County</td>
<td>1,509</td>
</tr>
<tr>
<td>Hampden County</td>
<td>901</td>
</tr>
<tr>
<td>Hampshire County</td>
<td>1,680</td>
</tr>
<tr>
<td>Middlesex County</td>
<td>662</td>
</tr>
<tr>
<td>Nantucket County</td>
<td>8,075</td>
</tr>
<tr>
<td>Norfolk County</td>
<td>1,010</td>
</tr>
<tr>
<td>Plymouth County</td>
<td>884</td>
</tr>
<tr>
<td>Suffolk County</td>
<td>858</td>
</tr>
<tr>
<td>Worcester County</td>
<td>768</td>
</tr>
</tbody>
</table>
Changes over time

- ACS replaced the decennial census long form in 2010
- Changes to specific questions and answer choices and geographies
Public Use Microdata Samples (PUMS)

• Public Use Microdata Sample (PUMS): extracts of raw data from a small sample of Decennial Census long-form and ACS records

• Allow analysis of individual household records

• Available for only certain levels of geography (PUMAs): counties (or equivalent entities) and census tracts with population ≥100,000:
  – ACS PUMS: http://www.census.gov/programs-surveys/acs/data/pums.html
  – IPUMS-USA: http://usa.ipums.org/usa

• More detailed restricted census data is available to researchers under special arrangements via Research Data Centers
  – https://www.census.gov/ces/dataproducts/restricted_data.html

MIT Libraries, 2017
SOURCES OF US CENSUS DATA
Historical
• NHGIS
• Social Explorer

For mapping
• Social Explorer
• SimplyMap
• Business Analyst
• Geolytics

Current
• Social Explorer
• SimplyMap
• Factfinder

Change over time
• Geolytics
• Longitudinal Tract Database

Free
• Factfinder
• NHGIS
• Longitudinal Tract Database

Other
• State and city agencies may provide census data for a specific area (MassGIS, City of Cambridge)
• R ACS package

http://libguides.mit.edu/census

MIT Libraries, 2017
Working with Census Variables

• Census creates cross-tabulations and thematic tables (e.g. housing) of variables

• Different types of tables start with different letters.

• Some variables are presented as ranges and/or medians:
  – age 40-50 yrs: 5%, age 51-60 yrs: 6%
  – median age: 54.6 yrs
INTERNATIONAL DEMOGRAPHIC DATA
International Sources

• MIT Databases: China Data Online, China Geo Explorer II, IndiaStat
• National Statistical Office Websites
• MIT GeoWeb:
  – MIT’s GIS data repository
  – includes census data from India, down to ward and village for some areas
• China DVDs in the GIS lab (see GIS website)
China Data Online

• Statistics and reports on China's population, economy, and industries
• Available at the following levels:
  – National
  – Provincial
  – City
  – County
  – Industrial
• Some statistics update monthly
• Census Data: [China Geo Explorer II](http://example.com) (can download shapefiles)
IndiaStat

• Statistics available in over 16 categories including:
  – Demographics
  – Agriculture
  – Education
  – Labor and Workforce

• Only allows 2 simultaneous users

• Select the Login option
  – Through IP Login for access
National Statistical Office Websites

• Comprehensive list organized by continent/region
• Published by United Nations Statistical Division
• Generally in the native language of the country but often have an English option
• Information available will vary greatly by country
Mapping Census Data

• Social Explorer and SimplyMap have built in mapping features
• If you can’t get the census data you want in a shapefile, download the census boundaries as a shapefile and the census data as a table, and join them based on the FIPS code.

http://libguides.mit.edu/gis/tutorials

Stop by the GIS lab!
Resources

• ACS guide: http://www.census.gov/programs-surveys/acs/
• Library census guide: http://libguides.mit.edu/census
• GIS guide: http://libguides.mit.edu/gis
• Country Data & Analysis: http://libguides.mit.edu/country