

Accessing Census Bureau Data

Presented By Howard Shih

howard.shih@aafederation.org



Asian American Federation

Major Transition Period

- Data Collection
 - Decennial Census
 - American Community Survey
- Data Access
 - New version of American FactFinder



Where to get community data from Census Bureau?

- Decennial Census
 - Complete count of all residents
 - Population items include sex, age, race, Hispanic or Latino origin, household relationship, household type, household size, family type, family size, and group quarters.
 - Housing items include occupancy status, vacancy status, and tenure
- American Community Survey
 - Survey of 3 million households annually
 - Includes everything in the Decennial Census
 - In addition, detailed social, economic, and housing characteristics



What is in the ACS?

Basic Demographics	Social Characteristics	Economic Characteristics	Housing Characteristics
<ul style="list-style-type: none"> •Age •Sex •Hispanic or Latino Origin •Race •Relationship 	<ul style="list-style-type: none"> •Ancestry •Citizenship Status •Disability Status •Educational Attainment •Fertility •Field of Degree •Grandparents as Caregivers •Language •Marital Status and Marital History •Place of Birth •School Enrollment •Residence 1 Year Ago/Migration •Veterans •Year of Entry 	<ul style="list-style-type: none"> •Class of Worker •Employment Status •Health Insurance Coverage •Income and Earnings •Industry •Journey to Work •Occupation •Poverty •Work Status 	<ul style="list-style-type: none"> •Food Stamps/ Supplemental Nutrition Assistance Program (SNAP) •House Heating Fuel •Kitchen Facilities •Owner Statistics •Plumbing Facilities •Renter Statistics •Rooms and Bedrooms •Telephone Service Available •Tenure •Units in Structure •Value of Home •Vehicles Available •Year Householder Moved Into Unit •Year Structure Built



Navigating the Decennial Census

- Key Data Products
 - Summary File 1 – available now. But only population counts for detailed ethnic groups. Full characteristics for major race groups.
 - Summary File 2 – due early next year. Full characteristics for detailed ethnic groups.
- Accessible through new American FactFinder
- Only available once every ten years



Navigating the American Community Survey

- Determine which version of the ACS to use
 - Currently only 2010 ACS data is on new American FactFinder
 - Previous data is in legacy American FactFinder, though due to transition “this fall”
- Select which data product contains the data you need.



Different Versions of the ACS

1-year estimates	3-year estimates	5-year estimates
12 months of collected data	36 months of collected data	60 months of collected data
Data for areas with populations of 65,000+	Data for areas with populations of 20,000+	Data for all Census Tracts
Smallest sample size	Larger sample size	Largest sample size
Least reliable	More reliable	Most reliable
Most current data	Less current data	Least current
Best used when	Best used when	Best used when
<ul style="list-style-type: none"> • Currency is more important than precision • Analyzing large populations 	<ul style="list-style-type: none"> • Analyzing smaller populations or geographies • Right now, most precise data set for selected population profiles 	<ul style="list-style-type: none"> • Precision is more important than currency • Analyzing very small populations and geographies • No select population data until next data release



ACS Data Products

Data Product	Description
Data profiles	Provide broad social, economic, housing, and demographic profiles.
Selected population profiles	Provide broad social, economic, and housing profiles for a large number of race, ethnic, ancestry, and country/region of birth groups.
Ranking tables	Provide state rankings of estimates across 86 key variables.
Subject tables	Similar to data profiles but include more detailed ACS data, classified by subject.
Detailed tables	Provide access to the most detailed ACS data and crosstabulations of ACS variables.
Geographic comparison tables	Compare geographic areas other than states (e.g., counties or congressional districts) for key variables.
Summary files	Provide access to the detailed tables through a series of comma-delimited text files on the Census Bureau's FTP site www2.census.gov .
Public Use Microdata Sample files	Provide access to ACS microdata for data users with SAS and SPSS software experience.



Before you dive in, ask these four questions:

- WHO do you want to learn about?
 - Race, ancestry, or ethnicity
 - Other demographics: age, gender, etc.
- WHAT is it you want to know about them?
 - Subject matter: demographics, social, economic, housing
- WHERE is the place that these people live?
 - Geography
- WHEN is the time period you are most interested in?
 - Most recent or historic



Who?

- If you are interested in the five major race groups, all Census Bureau products have data for them.
 - American Indian or Native Alaskan
 - Asian
 - Black
 - Native Hawaiian or Other Pacific Islander
 - White
 - Some Other Race is found in some products.
- Hispanic is its own category. Available in all Census Bureau Products
- Ethnicity data is only in decennial census and ACS (the check boxes and write-in categories for the Census and ACS, eg, Thai, Chilean, Sioux)
- Ancestry data is only in the ACS (examples are Scottish, Polish, Somalian)



What?

- Population and basic housing questions
 - Decennial census
- Detailed characteristics
 - American Community Survey (ACS)

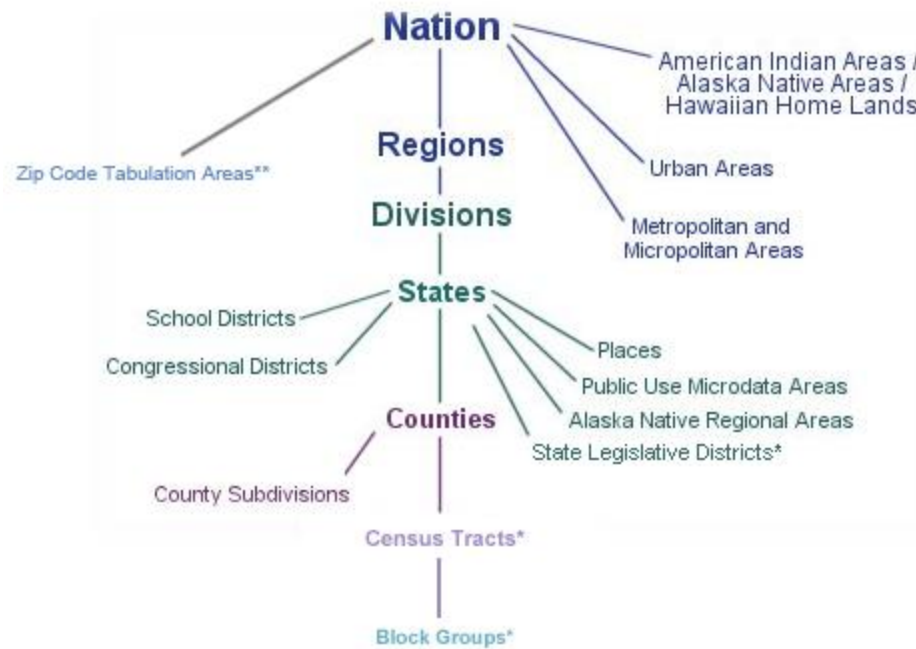


Where?

- Are you interested in national, state, county, city or neighborhoods?
- Neighborhood data is a challenge: What constitutes a neighborhood?
- Pre-defined areas versus creating your own
 - Pre-tabulated geographies: easy to download, but larger areas
 - Public Use Microdata Areas are close to New York City Community Districts.
 - Zip Code Tabulation Areas
 - Pre-defined geographies: provides a common frame of reference
 - New York City Neighborhood Tabulation Areas
 - New York City Health Districts
 - Make-your-own: Must decide what constitutes a neighborhood
 - Asian American Federation's Chinatown Brief



Hierarchy of Census Geographies



* 5-year ACS estimates only

** 5-year ACS estimates only, first release in 2012 for the 2007-2011 5-year estimates

When?

- Most recent data available?
 - Census is only once every 10 years
 - ACS is ongoing with data releases every year
 - 1-year data for population groups or geographic areas with more than 65,000 people
 - 3-year data for population groups or geographic areas with more than 20,000 people
 - 5-year data for Census tract level (neighborhoods)
- Must trade off between more current data versus more precise data



Major trade-offs between datasets (As of Sept 2011)

- Detailed Ethnic Groups
 - Can use 1 or 3-year ACS Selected Population Profiles or any Public Use Microdata Sample
 - Lose the ability to study smaller or custom geographies for now.
- Small or Custom Geographies
 - Must use 5-year ACS Summary File or Detailed Tables
 - Only available for major race and ethnic groups right now.
- Customized Cross-tabulations
 - Must use Public Use Microdata Sample
 - Lose the ability to study smallest or custom geographies



Challenges in comparing data sets

- Make sure you do not mix 1-year, 3-year, and 5-year ACS in the same analysis
- Most comparisons between 5-year ACS and 2000 Census Sample Data are valid. BEWARE of the following:
 - Migration
 - Marital Status
 - Disability Status
 - Contract Rent and Gross Rent
 - Gross Rent as a Percentage of Household Income
 - Real Estate Taxes
 - Telephone Service
 - Year Moved In
 - Year Structure Built



How to access data?

- American FactFinder
 - <http://factfinder.census.gov>
 - Downloading pre-formatted and pre-tabulated Census data
- Census Bureau FTP Site
 - <ftp://ftp2.census.gov>
 - Downloading large amounts of Census data for later data manipulation
- DataFerrett
 - <http://dataferrett.census.gov/>
 - Accessing Census and ACS PUMS files
- IPUMS
 - <http://usa.ipums.org/usa/>
 - Accessing Census and ACS PUMS data for comparing across time. Harmonized data.



Tips for Approaching New American FactFinder

- Answer the four questions above. This will help you narrow down the number of tables you will have to search through.
- Start with “Where?” Narrow down your choices by geography.
- “Who?”
 - For 2010 Census data, type in the search box for your ethnic groups. The filters leave out 2010 data currently.
 - For detailed characteristics from 2000, select filters from the Population Groups menu.
- “What?” Choose from Topics menu or Search box
- “When?” Often the previous three questions will limit your choices for what is the most recently available data.



Examples

- Using mapping function to define a neighborhood and download the number of Asian children in Manhattan's Chinatown from the 2010 Census.
- Compare the poverty rates for each State from the 2010 ACS and 2000 Census.



Helpful References

- American Community Survey
 - Documentation:
http://www.census.gov/acs/www/data_documentation/documentation_main/
- 2010 Census
 - Documentation for Summary File 1:
<http://www.census.gov/prod/cen2010/doc/sf1.pdf>
 - Data Products:
<http://www.census.gov/population/www/cen2010/glance/index.html>
- NYC Neighborhood Maps
 - Department of City Planning:
<http://www.nyc.gov/html/dcp/html/bytes/applbyte.shtml>
 - Neighborhood Tabulation Areas:
http://www.nyc.gov/html/dcp/html/census/demo_maps_2010.shtml
 - Community Districts:
http://home2.nyc.gov/html/dcp/pdf/census/puma_cd_map.pdf
 - ZIP Code Tabulation Areas: <http://www.census.gov/geo/ZCTA/zcta.html>





Handling Margin of Error

- ACS makes Margins of Error explicit in the published tables and provides a method to estimate Margins of Error for Public Use Microdata Sample
- Margins of Error existed for 2000 Census Sample Data, but not explicitly shown in data products. Technical documentation provided methodology to estimate Margins of Error for 2000 Census.
- Use statistical testing to verify the quality of data
 - Coefficient of Variation for individual estimates
 - Test statistical significance



Coefficient of Variation

Definition

The relative amount of sampling error associated with a sample estimate

Formula

$$CV = SE / \text{Estimate} * 100\%$$

$SE = MOE / 1.645$ (Census Bureau uses 90% confidence levels)

Unofficial Rule of Thumb

$CV > 15\%$ should be of concern



Statistical Significance for ACS Data

Formula

$$\left| \frac{\hat{X}_1 - \hat{X}_2}{\sqrt{SE_1^2 + SE_2^2}} \right| > Z_{CL}$$

where Z_{CL} is the critical value for the desired confidence level

Z_{CL} for 90% confidence level = 1.645



Example

- Has the poverty rate of Asian Americans in New York City changed since the financial crisis of 2008-2009?
 - In 2007, poverty rate was 17.1% with a 1.3% MOE
 - In 2010, poverty rate was 19.8% with a 1.3% MOE
- Coefficient of Variation
 - For 2007 estimate, CV was $1.3/17.1 = 7.8\%$
 - For 2010 estimate, CV was $1.3/19.8 = 6.6\%$
- Statistical Significance
 - Convert MOE to SE, $SE=0.79$

$$\left| \frac{19.8 - 17.1}{\sqrt{0.79^2 + 0.79^2}} \right| = 2.42 > 1.645$$

