An Introduction to the US Census
Understanding, Finding, & Accessing Data

Frank Donnelly, Geospatial Data Librarian

Baruch College CUNY
francis.donnelly@baruch.cuny.edu

March 2019
Introduce census concepts, data, and sources so that librarians can assist patrons with finding and using data.

1. Learn about census concepts (geography and subjects)
2. Gain familiarity with two fundamental census datasets (the decennial census and the American Community Survey)
3. Explore several different sources of data so you can connect each person to the best source
Questions for the Librarian

Where can I find:

- Statistics about my neighborhood?
- Home ownership and internet access data by county?
- Number of young adults in NYC by ZIP Code?
- Data about poverty in the US?

How do I:

- Use the Census Bureau’s website?
- Interpret estimates from the ACS?
- Understand the difference between a ZIP Code and ZCTA?
- Identify census tracts that are not residential?
Project Assistance

- Policy: create business profiles for community districts
- Journalism: reporting on a county that “flipped” in 2016
- Business: identify areas to locate a new coffee shop
- Social Science: factors that lead to heat-related death
Today’s Topics

Introduction to the Census

Data Profiles

Comparison Tables

Review - What to Use and When to Use it
The US Census

It’s *not* just the count that takes place every 10 years:
- Rich collection of several different datasets
- Summarizes people, houses, and businesses by place and subject
- Uses consistent definitions and terminology
- Collected over varying periods of time
Applications of the Census

All datasets and questions have some basis in federal law

- Apportioning congressional seats
- Drawing congressional and voting districts
- Apportioning billions of dollars in federal funds
- The basis for creating thousands of derived datasets
- Policy-making and research in public, private, non-profit, and academic sectors
- Describes what our country is like
Decennial Census
Dataset #1

- 100% count of the population
- Taken every ten years on April 1st
- US Constitution, Article I Section 2
- Individual responses summarized by category and geography
- Individual responses confidential
- 2010 Census: basic demographic variables

https://www.census.gov/programs-surveys/decennial-census.html
2010 Census Data

- Population
- Age and gender
- Race and Hispanic origin
- Households and families
- Family relationships
- Housing units
- Housing occupancy and tenure
American Community Survey
Dataset #2

- Launched in 2005
- Rolling sample survey of 3.5 million addresses
- Estimates with margins of error
- 1 year averages for geographies with 65K+ people
- 5 year averages for all geographies down to census tracts
- Released in autumn each year for the previous year

https://www.census.gov/programs-surveys/acs/
ACS Data

- All variables published in the decennial census
- Educational enrollment & attainment
- Income & poverty
- Ethnicity & language
- Labor force & journey to work
- Home value, rent, & other characteristics
- And a lot more...
The Decennial Census vs the ACS
Census Tract 68, New York County, NY

2010 Census
- 7,614 residents
- 5,414 White
- 330 Black
- 560 Hispanic
- 1,080 Asian
- 230 Other

2008-2012 ACS*
- 6,383 residents +/- 817
- 4,716 White +/- 735
- 297 Black +/- 187
- 469 Hispanic +/- 186
- 725 Asian +/- 295
- 176 Other +/- 156

* 90% Confident that population is between 5,566 and 7,200
Coefficient of Variation
Evaluate How Good an Estimate Is

CV measures relative amount of sampling error associated with an estimate.

Evaluating Estimates for Race Census Tract 68

CV = ((MOE / 1.645) / Estimate)*100
(1.645 is Z value for 90% confidence interval)

White population: 9 (0-15 highly reliable)
Hispanic population: 24 (16 to 34 medium reliability)
Black population: 38 (35 + low reliability)
### Derived Estimates

**Example: Sum Population of Two Tracts**

<table>
<thead>
<tr>
<th>Sum of Total Population for 2008-2012 ACS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tract 68 (6,383 +/- 817) and Tract 64 (7,439 +/- 621)</td>
</tr>
<tr>
<td>6,383 + 7,439 = 13,822</td>
</tr>
<tr>
<td>SQRT((817^2)+(621^2)) = 1,026</td>
</tr>
<tr>
<td>Tract Combo Population at 90% confidence is:</td>
</tr>
<tr>
<td>13,822 +/- 1,026</td>
</tr>
<tr>
<td>(90% confident that pop is between 12,796 and 14,848)</td>
</tr>
</tbody>
</table>
Choices

- Dataset
- Geography
- Variables
Census Geography

ANSI / FIPS Code (for tract)

GEOID2: 06037206031

06: California
037: Los Angeles County
206031: Census Tract 2060.31

TIGERweb

Application for viewing census geography

https://tigerweb.geo.census.gov/tigerweb/
Neighborhoods?

PUMAs

ZIP / ZCTAs

Census Tracts

NTAs
Boroughs?
Federal Data Sources Use County Names

- Bronx = Bronx County
- Brooklyn = Kings County
- Manhattan = New York County
- Queens = Queens County
- Staten Island = Richmond County
Census Universe
Population Groups and Subgroups

Population

Household
- Family
- Non-Family

Group Quarters
- Institutional
- Non-Institutional

Housing Unit
- Occupied
- Owner
- Renter
- Vacant
- Vacancy Status
Sources

The Census Bureau’s website provides data, but serves many other purposes. Since the data is in the public domain, you can download census data from many places. Choose the best source for the job:

- Range of datasets, variables, and geographies: everything and the kitchen sink, or just the basics?
- Interface that’s quick and easy or one that gives you the most flexibility and options?
- Additional tools for browsing, graphing, mapping?
- Downloading a little bit of data or a lot in bulk?
- Data profiles or comparison tables?
Today’s Topics

Introduction to the Census

Data Profiles

Comparison Tables

Review - What to Use and When to Use it
Data Profiles

Provide a wide range of variables for one place. Good for general overview, fact finding, and gaining familiarity with the census. Sources that provide profiles typically draw data from five tables published by the Census Bureau:

- **Decennial Census**
  - DP-1 Profile of General Population and Housing Characteristics

- **American Community Survey**
  - DP02 Selected Social Characteristics
  - DP03 Selected Economic Characteristics
  - DP04 Selected Housing Characteristics
  - DP05 Demographic and Housing Estimates
Profile Exercise

NYC Population Factfinder
Map-based interface for finding profiles for neighborhoods
https://popfactfinder.planning.nyc.gov/

TIGERweb
Reference map for US census geography
https://tigerweb.geo.census.gov/tigerweb/

Missouri Census Data Center
Point and click to create profiles for any US geography
http://mcdc.missouri.edu/
Today’s Topics

Introduction to the Census

Data Profiles

Comparison Tables

Review - What to Use and When to Use it
Comparsion Tables

Provide one or a few variables for many places. Census data is packaged in tables arranged in subjects. Different prefixes indicate the breadth and depth of the table.

S1903 Median Income in the Past 12 Months vs B19013 Median Household Income in the Past 12 Months

<table>
<thead>
<tr>
<th></th>
<th>ACS</th>
<th>Decennial Census</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Profiles</td>
<td>DP 02 through 05</td>
<td>DP 1</td>
</tr>
<tr>
<td>Comparative Profiles</td>
<td>CP 02 through 05</td>
<td>not available</td>
</tr>
<tr>
<td>Summary Tables</td>
<td>S subject table</td>
<td>QT quick table</td>
</tr>
<tr>
<td>Detailed Tables</td>
<td>B base table</td>
<td>P population subject</td>
</tr>
<tr>
<td></td>
<td>C collapsed table</td>
<td>H housing subject</td>
</tr>
<tr>
<td><strong>American Factfinder</strong></td>
<td>Census Bureau’s existing portal for accessing all its data</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td><a href="https://data.census.gov/">https://data.census.gov/</a></td>
<td></td>
</tr>
<tr>
<td><strong>data.census.gov</strong></td>
<td>Census Bureau’s new portal for accessing all its data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>It will be formally released summer 2019</td>
<td></td>
</tr>
<tr>
<td></td>
<td><a href="https://data.census.gov/">https://data.census.gov/</a></td>
<td></td>
</tr>
<tr>
<td><strong>Census Reporter</strong></td>
<td>Alternative source for recent ACS data</td>
<td></td>
</tr>
<tr>
<td></td>
<td><a href="https://censusreporter.org/">https://censusreporter.org/</a></td>
<td></td>
</tr>
</tbody>
</table>
Importing CSVs in Excel

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can preview below.

Delimiters:
- [ ] Tab
- [ ] Semicolon
- [x] Comma
- [ ] Space
- [ ] Other:

Text qualifier: 

Data preview:

<table>
<thead>
<tr>
<th>GEO.id</th>
<th>GEO.id2</th>
<th>GEO.display-label</th>
</tr>
</thead>
<tbody>
<tr>
<td>0500000US36001</td>
<td>36001</td>
<td>5001</td>
</tr>
<tr>
<td>0500000US36003</td>
<td>36003</td>
<td>5003</td>
</tr>
<tr>
<td>0500000US36002</td>
<td>36002</td>
<td>5002</td>
</tr>
</tbody>
</table>

Text Import Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

Column data format:
- [ ] General
- [x] Text
- [ ] Date: MDY

Text qualifier: 

Data preview:

<table>
<thead>
<tr>
<th>Text</th>
<th>Text</th>
<th>General</th>
<th>General</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEO.id</td>
<td>GEO.id</td>
<td>GEO.display-label</td>
<td>5001_VD01</td>
</tr>
</tbody>
</table>
Today’s Topics

Introduction to the Census

Data Profiles

Comparison Tables

Review - What to Use and When to Use it
Profiles vs Comparison Tables

Use profiles:
- Lots of data for one place
- Broad overview of a place
- Basic fact finding
- Comparing just a couple places

Use comparison tables:
- A few variables for many places
- Looking for specific variables
- Understanding a specific topic
- Research and analysis
Datasets

Use decennial census:
- Need actual counts
- Data for smallest geographies
- Basic demographic variables
- Long-term historical comparisons

Use ACS:
- Recent data
- For detailed social & economic characteristics
- Use 1-year period only for largest areas / groups
- Use 5-year period for everything else
A 3rd alternative for population data. Good for studying annual trends, simple to work with. Access via data.census.gov or MCDC

- Annual estimates created using calculations factoring in births, deaths, and migration with 10-year census as baseline
- Geography is limited to nation, states, counties, cities, and metro areas
- Data limited to basic variables: total population, gender, age, race, housing units
- Data includes components of change (births, deaths, migration)

https://www.census.gov/programs-surveys/popest.html
Sources

- **NYC Population Factfinder**: best for neighborhood-level data
- **TIGERweb**: visualize all census geography
- **MCDC**: Easy to access profiles for the entire US
- **American Factfinder**: full range of census datasets, tables, and geographies for the entire US
- **data.census.gov**: scheduled for release summer 2019, to replace the American Factfinder
- **Census Reporter**: profiles and tables from the most recent ACS
Timeline:

- Will be conducted Apr 1, 2020
- Basic count to be published Dec 31, 2020
- Redistricting data files to be published Mar 31, 2021
- Summary files published spring and summer 2021

With few exceptions, as a dataset the 2020 Census will be similar to 2010 in composition. Some changes:

- Same-sex marriages will be explicitly tabulated
- Respondents who choose black or white as a race can also indicate an ethnicity or national origin
- Members of military temporarily deployed overseas to be counted at their US military base
- Citizenship question is pending court challenges
Exploring the US Census: Your Guide to America’s Data
by Frank Donnelly published by SAGE Publications
Slated for late 2019
Frank Donnelly, Geospatial Data Librarian, Baruch CUNY
francis.donnelly@baruch.cuny.edu — 646-312-1657

https://www.baruch.cuny.edu/confluence/display/geoportal/GIS+Lab