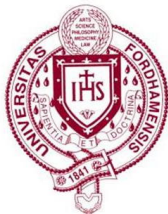# Downloading Map Data and American Community Survey Data

SDGB 7844

Prof. Nagaraja

Fordham University

Shape Files for the US (i.e., Map Data)

TIGER/Lines website: link:

- Choose the year which corresponds to the final year of your data set. For example, since we are using 5-year American Community Survey data for 2011-2015, then choose: 2015 ⟶ Download ⟶ Web interface

- Select 2015 (the webpage may have reverted back to 2016), Choose "Layer type" based on the smallest unit of geography you need. For this assignment, you want census tracts, because you want to make a map of New York County made up of census tracts; therefore, census tract is the smallest geographical unit you are interested in.

- After choosing census tract, you will select New York; you will get a map containing all of the census tracts in the state of New York. You will have remove the census tracts that do not pertain to New York County (like we removed Alaska and Hawaii from the US state map in Lecture 6).

- A folder of files will be downloaded. Don't change the name of the folder or the files inside. If you do, R may not be able to read the shapefiles correctly.

More technical stuff on shape files: link

The American Community Survey

- American Community Survey (ACS) ([link](#)) administered by U.S. Census Bureau
- for statistical purposes ONLY, individual survey data not available to other parts of the government
- demographic, economic, housing, and social characteristics
- used by governments (apportioning funds), businesses, researchers, and us!
- multi-year estimates combine data across several years $\longrightarrow$ interpret as an estimate for entire period
- ** note: only compare multi-year estimates with non-overlapping years
- multi-year estimates ([link](#)):
  - 1-year $\longrightarrow$ geographies with $\geq$ 65,000 people
  - 3-year $\longrightarrow$ geographies with $\geq$ 20,000 people
  - 5-year $\longrightarrow$ all geographies
- geographies ([link](#)): e.g., states, metropolitan areas, counties, census tracts

Data Source: American FactFinder

- American FactFinder website hosts the U.S. Census Bureau data
- Go to http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml
- Click on the link "get data" next to the American Community Survey in the "What We Provide" section. This should take you to a screen that looks similar to this:

ACS Data

- For your homework assignment, you want to extract information corresponding to:

  - New York County census tracts
  - 5-year ACS estimates for the period 2011-2015

- To get this data:

  - Use blue tabs on the left side of the screen to help you choose the table containing the information in which you are interested.
  - Note the website can be a little finicky so sometimes changing the order of how you select the settings may make a difference.
  - You will have to repeat these steps twice: once to get Table DP03 and once to get Table DP04.
  - Do not change the names of the files and folders! We will not be able to grade your assignment if you do.

ACS Data: Select Geography Parameter

ACS Data: Select Geography Parameter Part 2

ACS Data: Select Time Parameter

ACS Data: Selecting a Table

- At the top, type DP03 (or DP04) and click GO.

- The table should appear like the image below:

- Click on the blue table name, which will take you to the table. Then click Download.

Downloading the Files

Make sure these settings are selected before clicking OK:
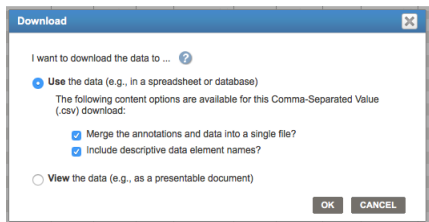
Downloading the Files

- a folder called "ACS_15_5YR_DP03" will be downloaded:
  - "ACS_15_5YR_DP03" is called the Table ID an uniquely identifies the table
  - "ACS" $\longrightarrow$ American Community Survey data
  - "15" $\longrightarrow$ data ending with the year 2015
  - "5YR" $\longrightarrow$ five-year aggregated data (in this case 2011-2015; as opposed to 1- or 3-year data)
  - "DP03" $\longrightarrow$ table number
- inside this folder, there will be four files:
  - "aff_download_readme_ann.txt" $\longrightarrow$ information on the content of the files in the folder
  - "ACS_15_5YR_DP03_metadata.csv" $\longrightarrow$ file with column codes and column descriptions
  - "ACS_15_5YR_DP03_with_ann.csv" $\longrightarrow$ data file
  - "ACS_15_5YR_DP03.txt" $\longrightarrow$ information on how to interpret the data file (important!)
- Your data is always in the file called "[Table ID]_with_ann.csv" and can be uploaded to R using read.csv()

Understanding "[Table ID]_with_ann.csv"

In addition to the information contained in "ACS_15_5YR_DP03.txt",

- `GEO.id` $\longrightarrow$ unique geography code among all U.S. Census Bureau geography codes

- `GEO.id2` $\longrightarrow$ unique geography code among all geography codes in data set (note that `GEO.id2` is a shortened version of `GEO.id`)

    - each row of the data set represents a different geography (e.g., all counties in a state, all states in a country, etc.) and `GEO.id2` uniquely identifies them

        - in this example, we only want data for one geography, the New York metropolitan area, so we only have one row in the data set

- `GEO.display-label` $\longrightarrow$ name of geography in English

- variable columns of the form `HD01_VD**` $\longrightarrow$ an estimate (e.g., mean, median, total) of variable VD** (or VC), its margin of error is given in column `HD02_VD**`

- variable columns of the form `HD03_VD**` $\longrightarrow$ a percentage estimate of variable VD** (or VC), its margin of error is given in column `HD04_VD**` (not applicable for our arriving at work example)

FIPS Codes

- FIPS = Federal Information Processing Standard
- code for every geographic area in the US and its associated territories
- different from USPS codes (U.S. Postal Service) state abbreviations
- USPS and FIPS codes for states and territories (link):

state	USPS code	FIPS code
California	CA	06
District of Columbia	DC	11
Illinois	IL	17
New York	NY	36
Ohio	OH	39
Pennsylvania	PA	42

- <u>need</u> the leading "0" in the codes (i.e., treat codes like character strings, not numbers)
- codes exist for counties, census tracts, etc. $\longrightarrow$ ALL geographies
  - this is what the GEO.id/GEO.id2 codes are in the ACS data