QGIS: Table Joins

Emily Barrett

	REGION	KYGEONET	SHAPE_area	SHAPE_len	YR_S2201_with_an	YR_S2201_with_an	2201_with_ann_H	2201_with_ann_HC	2201_with_ann_H
1	EAST REGION	7B6407538C-23...	13052913832.50...	668411.6762070...	0500000US21095	21095	11249	246	(X)
2	SOUTH CENTR...	7B6407538C-23...	5727897454.779...	373059.6876180...	0500000US21053	21053	4025	212	(X)
3	SOUTH CENTR...	7B6407538C-23...	9811755114.790...	557142.3594599...	0500000US21003	21003	7705	257	(X)
4	EAST REGION	7B6407538C-23...	8617286923.620...	577231.0987319...	0500000US21153	21153	5057	167	(X)
5	BLUEGRASS RE...	7B6407538C-23...	12355002668.60...	676211.2533729...	0500000US21151	21151	33036	423	(X)
6	SOUTH CENTR...	7B6407538C-23...	8395339385.989...	531239.9975929...	0500000US21229	21229	4507	163	(X)

Description: Not every dataset you want to use comes in a spatial format (shapefile, geopackage, KLM etc.). Often the data you are most interested in is located in a table or a spreadsheet and you need to link it with your existing spatial data for use in your mapping analysis. The operation to link two tables using a common attribute is known as a **table join**.

Example: We want to map the number of McDonald's per 100,00 people in America. We know that we can access a shapefile of American states that has an attribute table that lists information such as the name of the state and its population. We also know there are additional data sources that in tabular form (spreadsheets) list the number of McDonald's per state. BUT there are no shapefiles that contain the geographic information we need to map out the states AND have information about the number of McDonald's in each state. This is where we use a table join to connect the two tables together using a common attribute.

What might this common attribute be?
- A name or abbreviation (such as Tennessee or TN)
- An alphanumeric code (zip code, geographic location code, GEOID)
- Anything else where the columns in the two tables perfectly match

By joining the tables together, we can now make a map that shows the number of McDonald's per 100,00 people in each state.

Part 1: Perform a table join in QGIS

1. Follow this tutorial to learn how to perform a table join in QGIS. **Note:** You can directly download the data needed to follow along with this tutorial from this webpage. You do not need to get the census data for this step. Remember to unzip your downloaded files.
2. Tutorial link: https://www.qgistutorials.com/en/docs/3/performing_table_joins.html
3. How to know if your table joined worked? You can check by opening the attribute table for the shapefile: if you see all of your values, you successfully completed the join! If you get NULL values or cannot see anything then try it again.

Part 2: Download tabular data from the US Census Bureau, data.census.gov

1. **Note:** The tutorial you followed above referenced American FactFinder. As of March 2020, the Census Bureau is transitioning away from this site to data.census.gov. We will use this site as the primary source for tabular census data.
2. One of the most common datasets that we perform table joins on is census data. Following this tutorial, download a tabular dataset for a specific geographic region and scale (states, counties, census tracts etc.) of your choice. Example: Income by census tracts, for Davidson County, TN.
3. Tutorial link: https://www.census.gov/data/academy/data-gems/2019/navigate-datacensus-gov.html

Part 3: Download a spatial dataset from TIGER LINE shapefiles

1. Using page 11 from this tutorial download a shapefile that corresponds to the geographic region and scale that you used above. For example, census tracts in KY.
2. Link to TIGER: https://www.census.gov/geographies/mapping-files/timeseries/geo/tiger-line-file.html
3. Open your downloaded shapefile in QGIS and explore the attribute table. Compare it to the tabular data that you downloaded in part 2. What common attribute could you use to join these files together? Use this attribute to join together your dataset using the same steps from above.