Re-envisioning Community Safety: Group 1 (Data Collection)

...

QSIDE Datathon4Justice 2024

Saturday

- Team members got technology set up (R and R Studio, keys for US Census) and started pulling data from US Census Bureau.
- Census Bureau information includes both data from US decennial census and American Community Survey (ACS). ACS provides data in both 1 year estimates and 5 year estimates.
- QSIDE team had provided list of variables to consider as starting point , including those used in some work that has already been done.
- Brainstormed ideas for additional variables that we might want to study and sources for information.

Locations of interest

- Austin, neighborhood in Chicago, Illinois
- Washtenaw County, Michigan
- Martin County, Kentucky
- Nashville, Tennessee
- Pascua Yaqui Tribe, Arizona
- Pine Ridge, South Dakota
- Sheyenne, North Dakota

API Calls and Shapefiles - Overview

- Setting up API calls using R tidycensus package to gather data for specific geographic region
  - Scope
    - Started by collecting data on **Census Tract** level from ACS 5 year dataset – along with collecting external data sets
    - Collecting variable dictionary tables for datasets gathered
    - Creating functions that allow user to view correlation (2 vars) or susceptibility (1 var) for some given granularity level
  - Improvements
    - Limiting the number of requests per second in the API call to avoid some timeout errors we were hitting for large requests
    - Once APIs are called and shapefiles are mapped, saving local files to be used in future runs to make calls more efficient

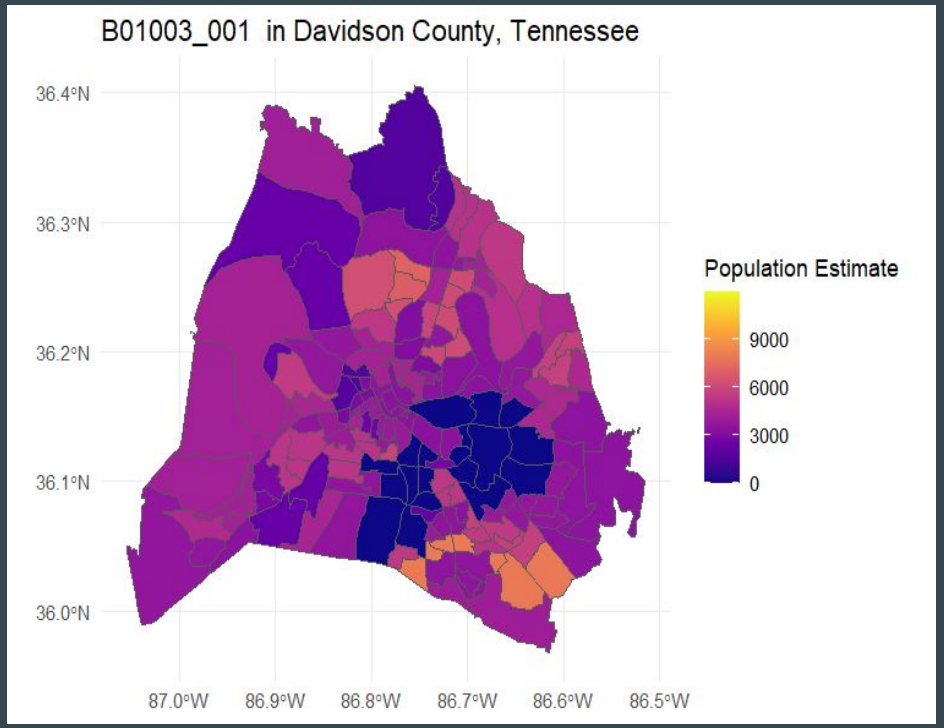
API Calls and Shapefiles - Visuals

Function to plot spatial maps for different Census Tracts within a county, including a gradient based on the selected variable. Also included correlation function that plotted against 2 variables for a certain geographical argument/s

Next Steps

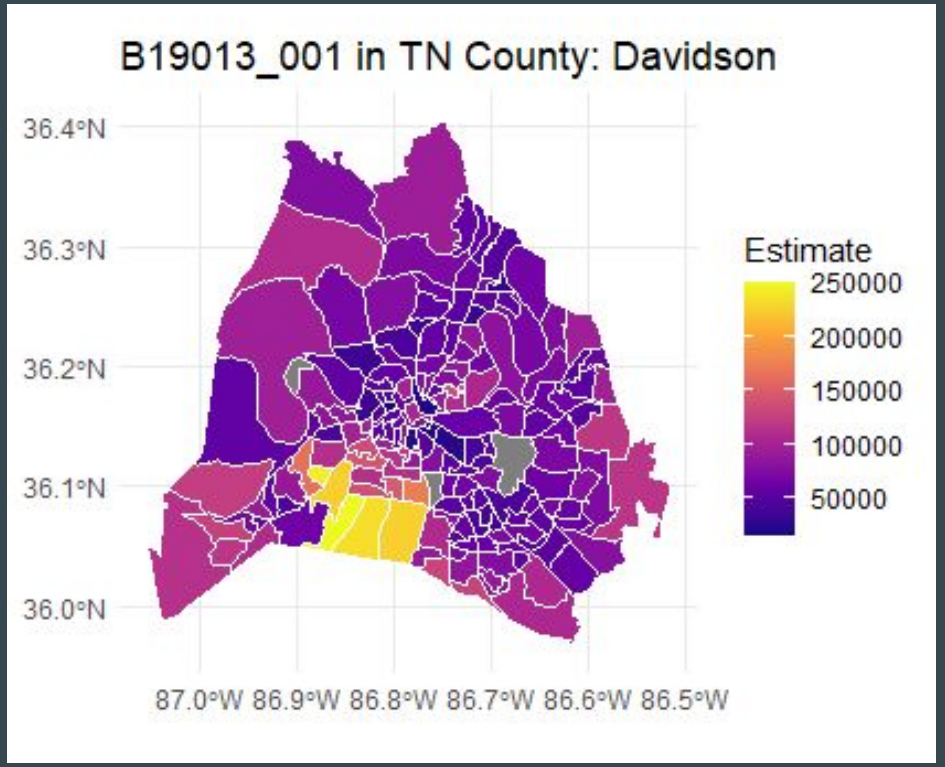
- Link certain variables not located with corresponding labels so we can plot & map important data
- Transform additional datasets for easy lookups and potential joining with Census ACS data
- Incorporate additional variables -- this will require some mapping to which variables area available for which geography levels and which datasets we need to query to find them. Others in group 1 have been working on this area in finding that information
- Utilize notes from discussions had with group regarding other external factors that are important to focus on, and use new data to produce interesting maps based on food accessibility, internet availability, and more

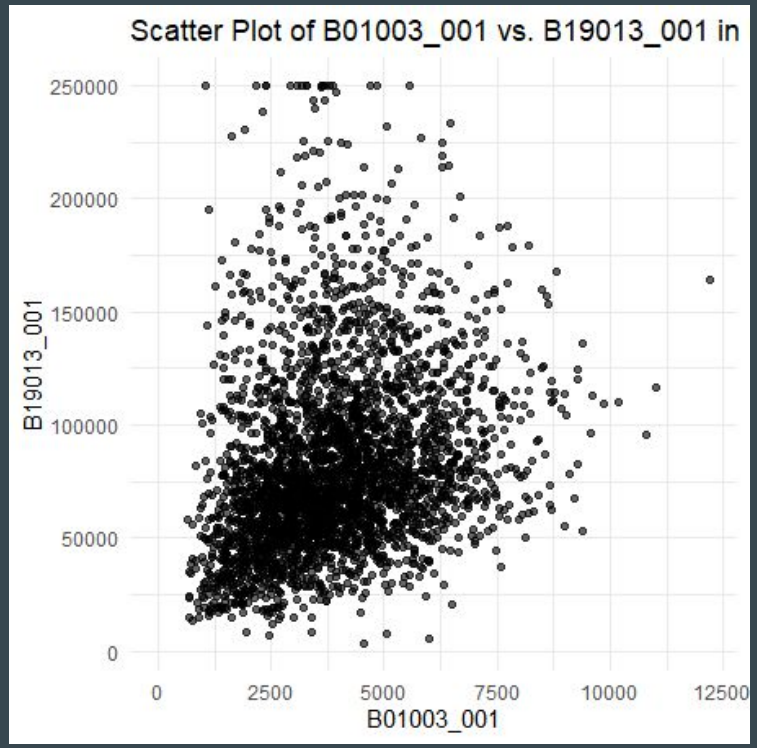
Examples for Population Size and Median Income→



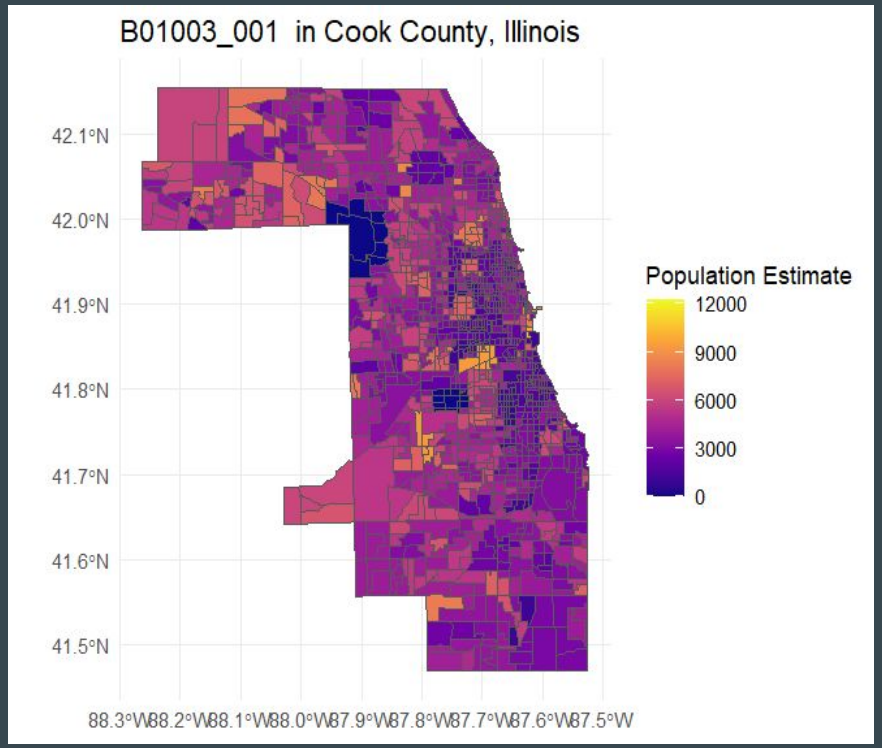
Total Population in Davidson County, TN

Median Household Income in Davidson County, TN

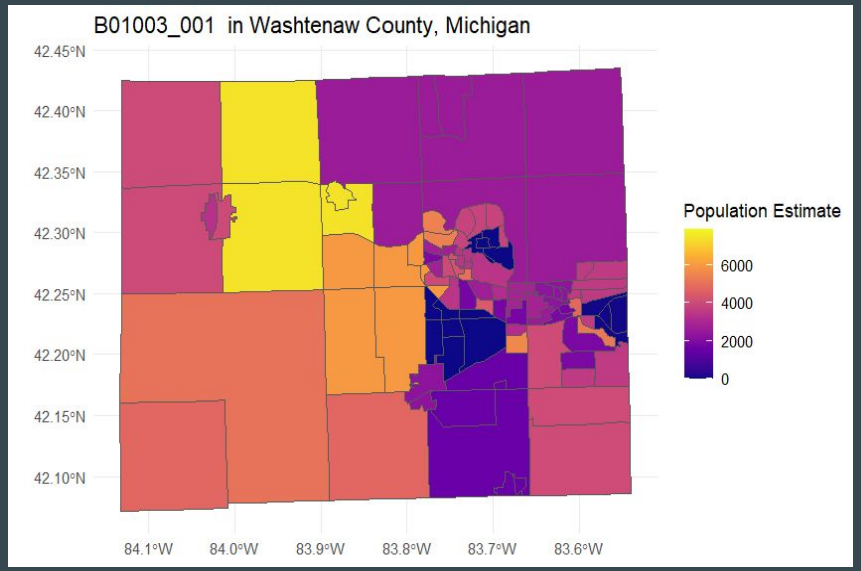


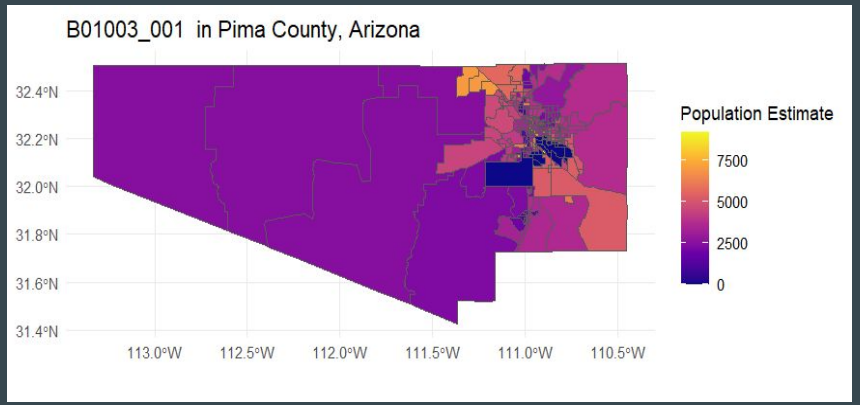


Illinois variable comparison

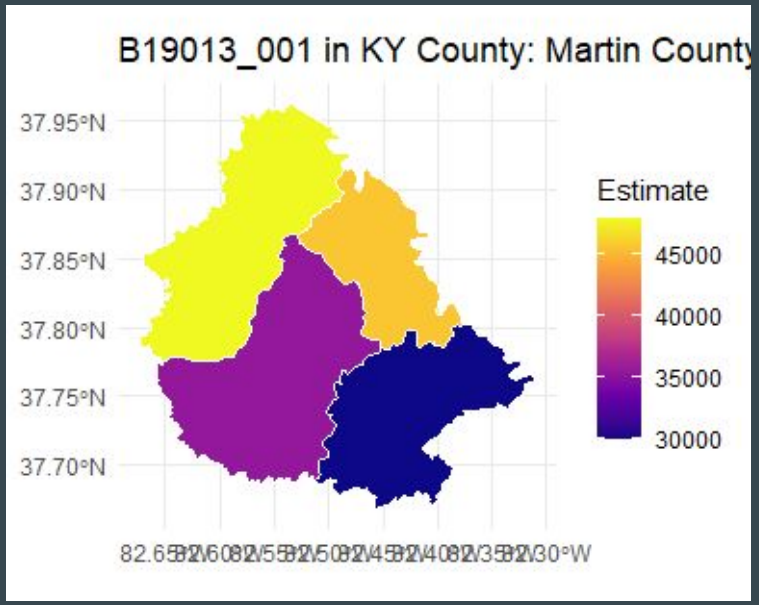


Total Population in Cook County, IL
(40% of IL resides here)





< Total Population in Washtenaw, MI (top) and Pima, AZ County (bottom)



Median Household Income in Martin County, KY

OpenStreetMap

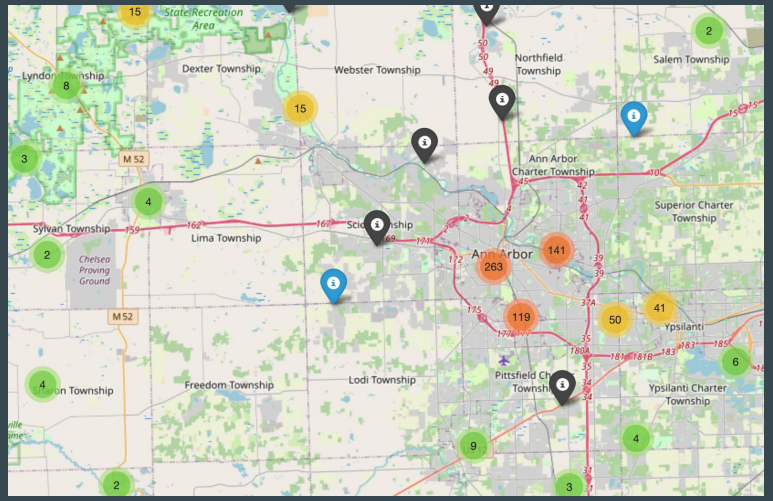
- Free open-content map that includes data on many different features (list of “map features” here: OpenStreetMap features (website) .
- Some features we think that might be interesting to this project include: public libraries, clinics, hospitals, public transportation stations, food stores.
- In Python: have code to pull data on these features.

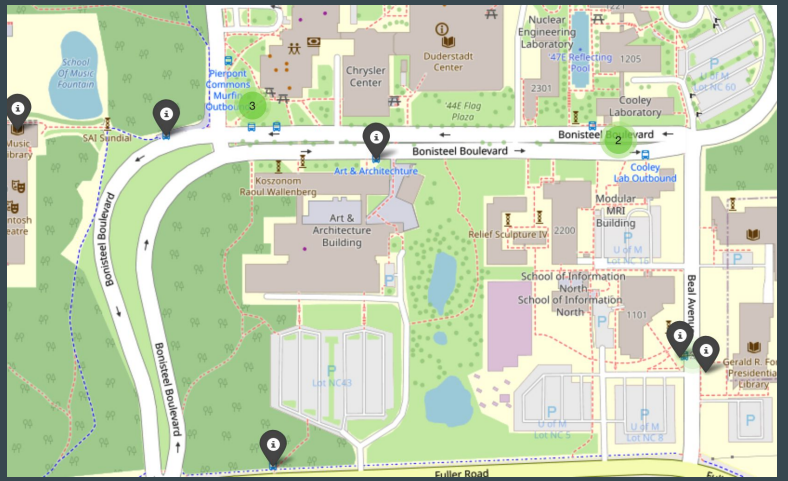
Values:

- Kindergarten, library, school, social_facility, clinic, hospital, police, prison, shower, drinking_water, toilets
- Platform, station
- Aboriginal_lands, postal_code
- Religious (buildings)
- Counseling
- Water_tap
- Food, water (shops)

OpenStreetMap

- With the data obtained from OSM, we developed (WiP) an interactive map that shows the community fabric features
- An easy resource to look for key places for community wellbeing!





Reported Crime Data (County-level)

- Uniform Crime Reporting Program
  - Crime data reported yearly by law enforcement agencies to the FBI
    - Includes juvenile records, some info about sex, age, race, count breakdown by offense type
- This iteration:
  - Summed up total reported offenses reported by each agency in 2022
  - Linked agency's county to FIPS county codes so that we can join this data with Census ACS
  - Provided documentation for future iterations
- Future iterations:
  - More detailed breakdowns/counts; yearly time-series
- Next steps for this indicator
  - May be useful for dimension reduction (selecting other indicators that may best explain)

Additional data sources

- USDA Food Access Research Atlas: https://www.ers.usda.gov/data-products/food-access-research-atlas/ (website)

  Provides food access data at census tract level.

- National Neighborhood Data Archive (NaNDA): (website) https://nanda.isr.umich.edu/

  Data at census tract level. Includes many different variables, such as polluting sites, traffic volume, parks, and weather.