**Water**

Tools:

* Python on Jupyter notebook
* Geopandas
* Folium

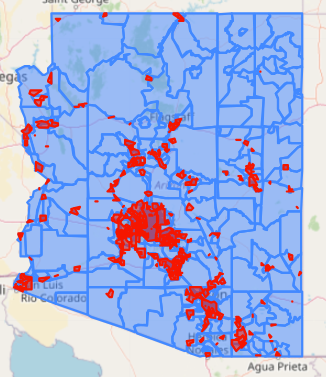
Water Bill Infrastructure Data:

* Data collected from - AZ Water and Wastewater Rates Dashboard - <https://dashboards.efc.sog.unc.edu/az>
* Water billing information was available at - <https://efc.sog.unc.edu/resource/arizona-rates-resources/>
* Primary Service Areas were scrapped from <https://dashboards.efc.sog.unc.edu/az>

Census Data

* Geographies for CDP/Towns - <https://www2.census.gov/geo/tiger/TIGER_RD18/LAYER/PLACE/> combined and pruned for Arizona
* Census Tract Info - <https://www2.census.gov/geo/tiger/TIGER_RD18/LAYER/TRACT/>

Procedure:

* Geography information for Primary Service Information was joined with the Water Rates Table
* Water Rates table for Primary Service Area were mapped to Census Tract info  
  Red – CDPs/Towns, Blue – Census Tracts  
  
* Performed spatial join on the two tables on intersection operation.
  + Max value of 10,000 Gallons was taken for Census Tracts where there are multiple utility companies were servicing
  + A value of $1000 was given to tracts were there was no service
  + Max and Min of Household incomes were added to the table based on info from the Water Rates table
* Data processing was done on the Water Rates table. Final table contains the following columns

**GEOID, Utility/Rate Structure, NAMELSAD, Median Household Income in 2020 (Min, Max), CDP, 10,000 Gallons**

**A screenshot of a map

Description automatically generated with medium confidence**

**Air Quality**

Tools:

* Python on Jupyter notebook
* Geopandas
* Folium
* K-means clustering
* Nominatim – Open Street Map

Census Data

* Geographies for CDP/Towns - <https://www2.census.gov/geo/tiger/TIGER_RD18/LAYER/PLACE/> combined and pruned for Arizona
* Census Tract Info - <https://www2.census.gov/geo/tiger/TIGER_RD18/LAYER/TRACT/>

Air Quality Data

* Ozone(44201) , PM2.5 FRM/FEM Mass (88101) and PM10 Mass (81102) - https://aqs.epa.gov/aqsweb/airdata/download\_files.html#Annual

Procedure

* Daily data was available for the year 2022. The average value for every CBSA was taken for Ozone, PM2.5 and PM10.
* Combined the 3 tables with geography information from census data and converted to geodataframe
* Mapped CBSAs to census tract

A map of arizona with red dots

Description automatically generated with low confidencered – Areas where air quality data was available, blue – census tract info

* Used K-Means clustering algorithm to find air quality values where data was not available.
* A picture containing text, screenshot, map

  Description automatically generatedFinal table contains columns - **'NAMELSAD', 'geometry', 'NAME', 'PM10\_Mean', 'PM25\_Mean', 'OZ\_Mean'**

Currently working on –

Geocoding addresses of sensors from data available to calculate sensors/sq mile ratio