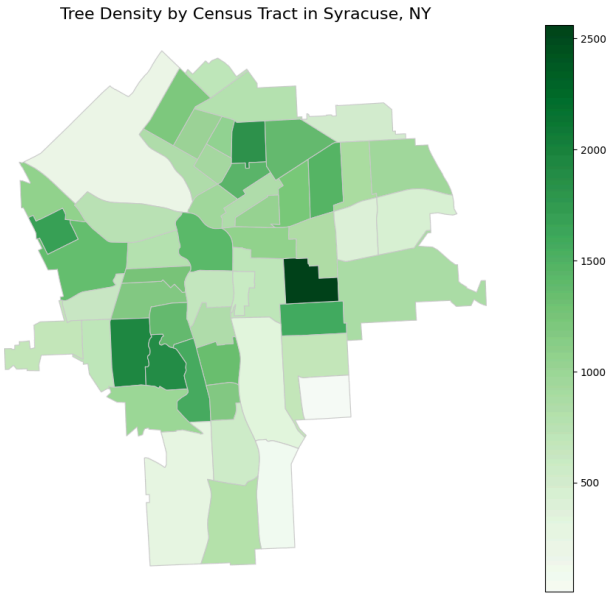
Jena Klimaszewski
Python Programming
Yanan Wu
May 5, 2025

Trees and Income in Syracuse, NY Report

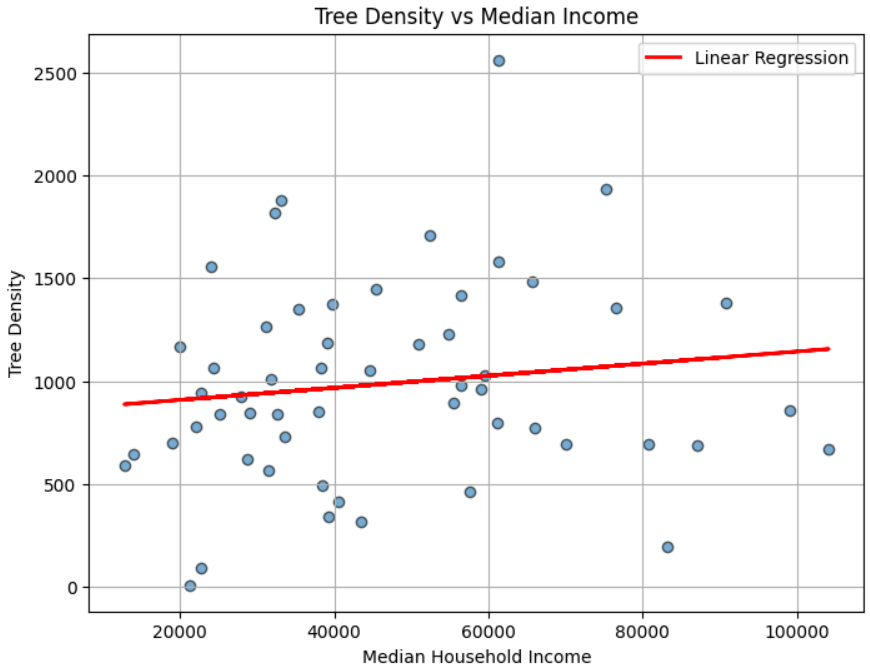
This project analyzes the relationship between he number of trees and income per census tract in Syracuse, NY. Syracuse is known for being a relatively green city with significant income disparities. For the project, I used Python and its various packages to determine if high-income areas have more trees than low-income areas of the city.

To start, I gathered data from Syracuse's open data portal. I read the data using the geopandas and path libraries. To retrieve the census income data, I utilized the censusdata package to read in the 5-year American Community Survey data for Onondaga County. For accuracy, I projected both datasets to NAD 83 (EPSG:26918). To make each visualization, I used matplotlib to keep things simple. To clean the census data, I used the re package to match the different formats of the census tract codes from each dataset.

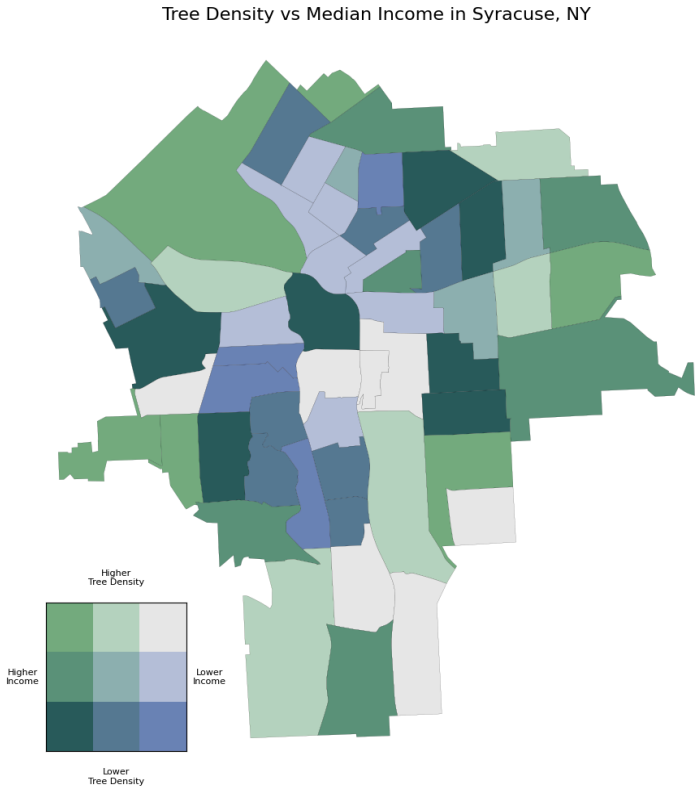
To measure tree density, I used a spatial join to match each tree point to the census tract that it is in. Following this, I grouped the trees by tract and calculated the number of trees per tract. For the tree density map, I simply plotted this dataset. This map showed that the southwest and eastern parts of the city tend to have more trees, while the southern, northeastern, and central parts of the city have fewer trees.



To demonstrate the relationship between income and tree density, I first created a scatterplot comparing tree density with median household income. Each dot represented a census tract, with income on the x-axis and tree density on the y-axis. I added a linear regression line using the numpy package to show the general trend of the data. While there was a slight upward slope, it still indicates a weak trend between income and tree density. This means that other factors, such as zoning and parks, may have a greater influence on tree presence in neighborhoods.



The final visualization I created was a bivariate colorpleth map showing the relationship between income and tree density. This map classifies each variable into three quantile-based categories (low, medium, high). These combinations create nine possible classes, each assigned a color. Tracts labeled with darker green-blue tones represented areas with both high income and high tree density, while pale or grayish areas tended to have low values for both. This map makes it easier to see how the two variables interact better than two individual maps would.



The results of the maps suggest that higher-income neighborhoods tend to have more trees, but the correlation is not strong enough to say that definitively. Some lower-income areas still had a decent number of trees, and some high-income tracts have less. This shows that income could play a role, but it's not the only thing that affects where trees are located.

With additional time, the analysis could be further improved by incorporating more data, such as park locations and zoning information. These layers would provide a better understanding of greenery in the city. Interactive maps created with libraries like Plotly could also make the data more digestible.