Project Scope Update

Since its inception, the scope of the project has remained unchanged, continuing to analyze the spatial relationship between Airbnb listing prices in New York City and the distribution of subway stations, as well as neighborhoods' median household income. The project's goal is to develop a spatial regression model that quantifies the influence of these factors on listing prices.

Data Sources

Airbnb listing data: CSV file downloaded from InsideAirbnb, containing over 4,000 active listings (including prices, geographic coordinates, property types, and guest ratings).

New York subway station data: CSV file obtained from NY Open Data, containing geographic locations and line information for 473 subway stations.

Census socioeconomic data: Retrieved via Python code calling the U.S. Census Bureau's ACS API, covering median household income and population density data for 195 New York City neighborhoods.

API Used

Census Bureau American Community Survey (ACS) 5-Year API which is used to obtain official socioeconomic data at the community level. Implemented using Python and the requests library, this solution employs valid API key authentication to retrieve target variables, applicable to New York State and its associated counties and districts.

Issues / Difficulties

Spatial visualization has not been fully implemented. When overlaying subway station data layers onto the folium map, data formats do not match. Additionally, the spatial regression model and Moran's I spatial autocorrelation test have not yet been implemented.