

Clustering neighborhoods in New York City: Recipient groups of healthcare service

Capstone Project - The Battle of
Neighborhoods



Introduction

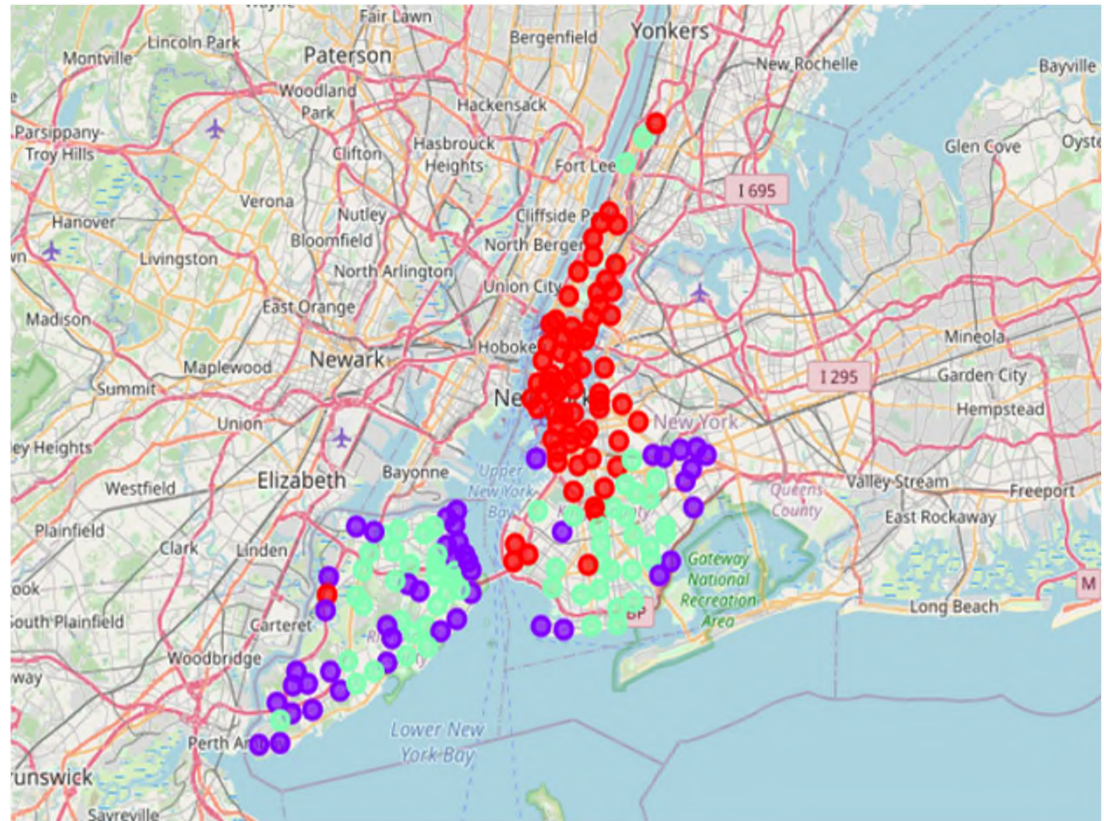
- A living environment has a significant influence on individual health status
- The government not only focuses on improving medical services to citizens but also promotes citizen's health
- it is advantageous for the government to accurately understand the physical environment of its neighborhoods including accessibility of health-related facilities and financial conditions so that they can have a strategic guideline of healthcare policies.

Data acquisition and cleaning

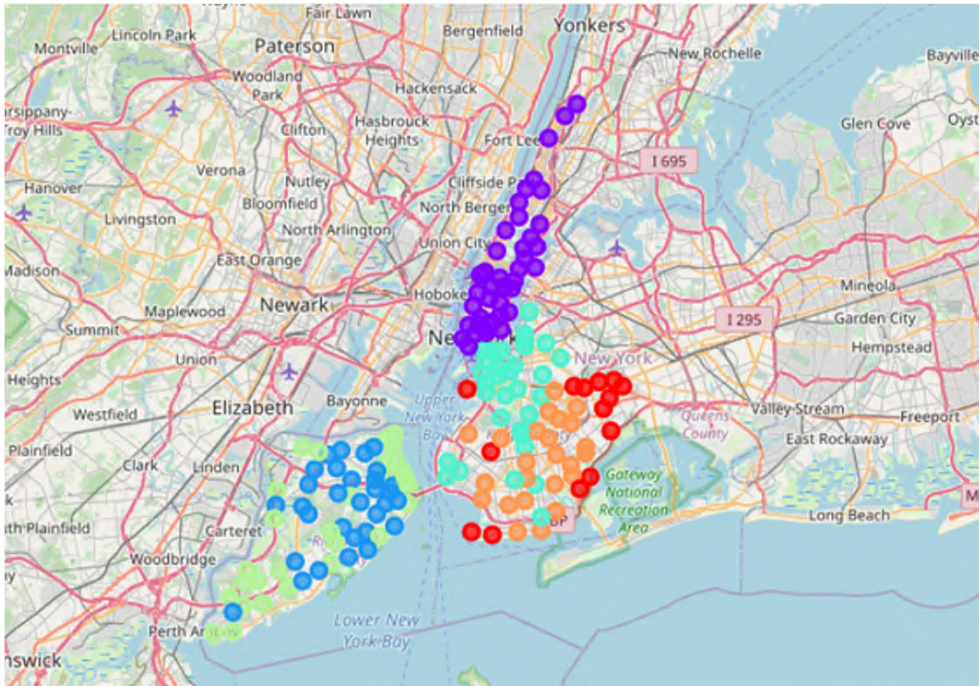
- Given that we access the living environment directly related to health,
- there are four factors to be measured:
 - number of medical centers in the neighborhood (e.g. hospitals, doctor's office, etc)
 - number of sports facilities in the neighborhood (e.g. gym, yoga studio, etc)
 - GDP, median income and poverty rate of the borough to which the neighborhood belongs
 - density rate of the borough to which the neighborhood belongs
- Data sources:
 - number of medical centers and sports facilities in every neighborhood will be obtained using Foursquare API
 - GDP, median income, poverty rate, and density rate of the borough will be obtained using Wikipedia which specifies the original sources of the data

Clustering neighborhoods A

- The neighborhoods are divided into clusters only in terms of accessibility of medical and fitness-related facilities.
- While all the neighborhoods in Manhattan have the highest accessibility of both medical services and sports facilities, the other two boroughs share similar patterns: the neighborhoods situated toward inland relatively have lower opportunities to have many options in that regard.



Clustering neighborhoods B



- When all features were taken into consideration, neighborhoods will be divided into several clusters within its borough. It indicates the predictors such as the population density, GDP, the median income, and the poverty rate shows the distinct characteristic of the boroughs
- Manhattan with the highest accessibility and income has a very strong position among others because it remains the same cluster in both clusterings.



Conclusion

- The objective of this project is to help government health agencies that provide healthcare services get a better understanding of the citizens of New York City, especially the Manhattan data, the Brooklyn, and the Staten Island. As the neighborhoods were categorized into several groups based on their physical living environment, the agencies and companies can establish a public campaign or marketing strategy with different approaches depending on neighborhoods.