Introduction/Business Problem

When it comes to real estate, location is everything. A business can do everything right but still fail to succeed without being in the proper location. New York City is a complex and sprawling city with a large population, which could potentially mean a lot of customers. However, New York can differ greatly borough by borough and neighborhood by neighborhood, and is already saturated by a large number of businesses. In this project, I will determine where a new restaurant should open in New York City using data on different neighborhoods.

Data

I plan to use Foursquare location data for this project, as well as data on the population per neighborhood in New York available on the state government website at this link: https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoodpop.htm. I will use these datasets to highlight areas in the city with lots of people and not many restaurants, working on the assumption that the number of people in an area is a key driver of the number of people who eat at restaurants.

Methodology

Using json data from Foursquare, I was able to determine the most popular venues for each neighborhood in New York City. Then, using the population data from New York State website, I determined which borough had the largest population. The remainder of the analysis was then focused on Manhattan only. Using kmeans to cluster the neighborhoods, I then determined which cluster would be the most apt for a new restaurant by picking a cluster that is not already saturated with restaurants.

Results

The cluster of Manhattan neighborhoods with the least amount of restaurants included Roosevelt Island, Battery Park City, and Stuyvesant Town. Both Roosevelt Island and Battery Park City primarily occupied by parks and are therefore not densely populated. Stuyvesant Town, on the other hand, has people living in it, as well as few restaurants, leaving it as the prime opportunity to have a successful new restaurant.

Discussion

One observation on the use Foursquare data is that it provides details not only whether restaurants are in these neighborhoods but also what kind of restaurants they are. The conclusion may have changed based on the type of restaurant that we want to open. There may be densely populated neighborhoods that have many Mexican restaurants, for instance, but few Italian ones. However, this was outside the scope of the exercise.

Conclusion

Stuyvesant Town is the neighborhood in New York City that is a good opportunity for a new restaurant.