Finding the Top Neighborhoods To Open a Restaurant in NYC

Dhananjayan P N

Introduction

- New York City is the most populous city in the United States.
- With an estimated 2019 population of 8,336,817 distributed over about 302.6 square mile, New York is also the most densely populated major city in the United States.
- With such a densely populated city, it is found that New York is filled with restaurants in almost all neighborhoods. In such a city, it is often difficult for someone to find the best place to open their own restaurant.

In this project, we will be leveraging data from various sources on the internet, to help us find the best neighborhoods for a stakeholder to open their restaurant, in order to get guaranteed business and footfall.

Data Acquisition

- In order to carry forward this project we will need to gather the required data from the internet. The larger the dataset, the better our model will perform in finding the best neighborhoods in New York.
- One of the main datasets we will be using is the neighborhood dataset for New York City. We will use this NYC dataset to get all the data we need for NYC Neighborhoods and Boroughs.
- We will also be leveraging the Foursquare API in this project, in order to get data related to all the restaurants in a neighborhood.

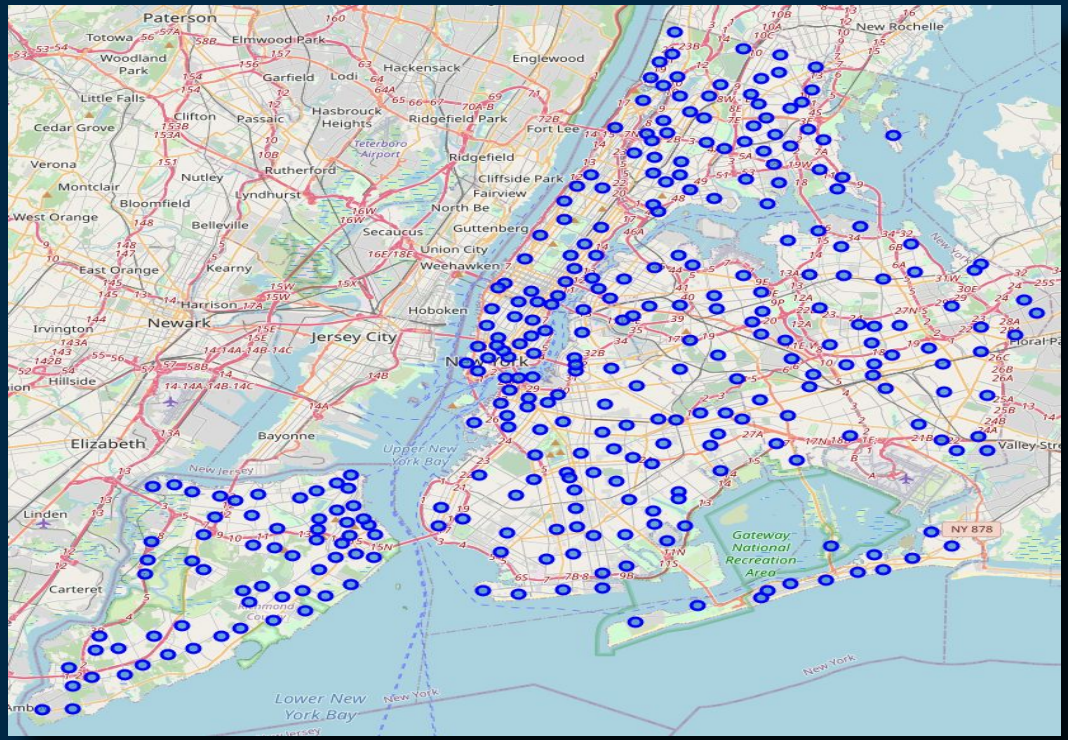
Data Cleaning

The NYC dataset that we are using in this project is extremely accurate, clean and structured, and hence it doesn't require any cleaning (like removing missing values, etc). The dataset can be used as is due to its high quality.

The Foursquare API is an API that has been created to ensure that the user is able to retrieve clean, accurate and a large variety of data for personal and commercial purposes. Hence, the Foursquare data does not require any cleaning.

Methodology

We will first visualize the data we had acquired from the NYC dataset on a real world map using Folium.

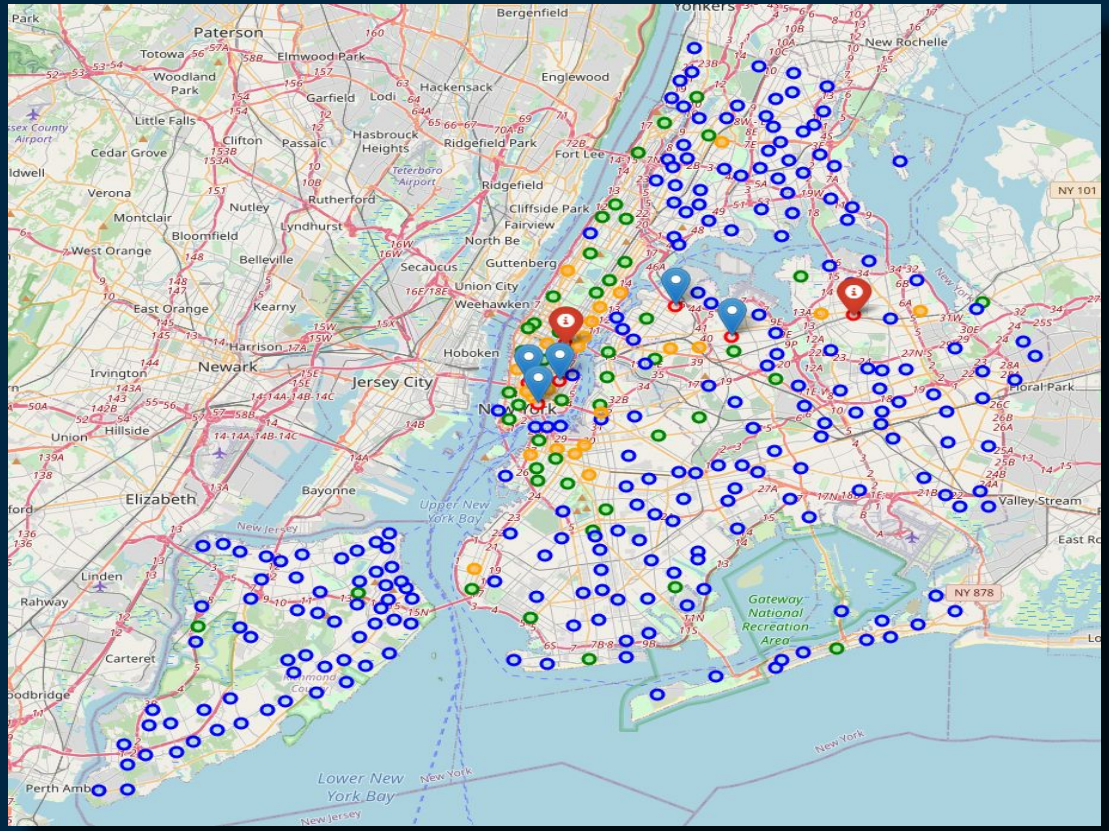


Now that we have visualized it we will do the following to use this data for our project:

- use the Foursquare API data to get the number of restaurants in every neighborhood.
- Merge this restaurants data with our NYC dataset to get a complete dataset with neighborhoods and the number of restaurants they have.
- Use this full-fledged dataset to find the top neighborhoods for a stakeholder to open a restaurant

Once we have found out the top neighborhoods with the most restaurants, we will visualize this merged dataset on a real NYC map to get a sense of where they are located.

NYC Map with Restaurant Density



Results

Our analysis shows that although there are restaurants in almost every neighborhood in NYC, a majority of them are pockets of low restaurant density, with a few neighborhoods having average and medium density.
Highest concentration of restaurants was detected in the following neighborhoods in decreasing order of the number of restaurants:

1. Murray Hill, Manhattan, 47 restaurants
2. Murray Hill, Queens, 47 restaurants
3. Jackson Heights, Queens, 39 restaurants
4. Astoria, Queens, 37 restaurants
5. East Village, Manhattan, 37 restaurants
6. Greenwich Village, Manhattan, 37 restaurants

Discussion

- From the results we can state that all the neighborhoods of high restaurant density were located in the Manhattan and Queens boroughs. It is therefore also implied that a stakeholder will experience high footfall if they open a new restaurant in any of these neighborhoods.
- Manhattan and Queens also have numerous neighborhoods that have average to medium restaurant density.
- Low restaurant density was found in a majority of neighborhoods in NYC. This also shows that the results of this project will enable stakeholders to ensure that they do not open their restaurants in areas with low density.

Conclusion

- The purpose of this project was to identify neighborhoods in NYC with high restaurant density in order to aid stakeholders in narrowing down the search for optimal location for a new restaurant. By using the NYC neighborhoods and boroughs dataset along with the data from the FourSquare API we were able to successfully analyze the neighborhoods and find the neighborhoods with the most restaurants.
- A future feature implementation and analysis of this project can include factors like attractiveness of each location (proximity to park or water), levels of noise, proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc, to help stakeholders decide on the optimal location for their restaurant.