Best Places to Start an Italian Restaurant in New York

1. Introduction

1.1. Background:

This project searches for the best locations to open Italian restaurants in New York. US Census Bureau states that New York has a population of around 8.4 million (US Census Bureau, 2019). Apart from being a financial hub of the United States, New York boasts on being a multicultural city. A great share comes from the Italian immigrants who form a major part of the population, 8.2% (2006–2008 American Community Survey). Italian cusine has become so famous in New York that it is quite hard finding a reservation for an Italian restaurant.

1.2. Problem description:

New York is famous for its diverse cuisine. It is home of the finest and most diverse restaurants in the world. In addition to Italian cuisine, food culture includes also an array of international cuisines influenced by diversity of the town.

To survive in competition, it is important to perform a background analysis. Various factors need to be studied to decide on the location such as:

- Number of restaurants
- Ratings
- Location
- Reviews
- Tips

Objective of this study is to locate and recommend to any entrepreneur who wants to open a new restaurant innew York, the neighborhood.

2. Data:

This project will analyze neighborhoods and boroughs of New York.

Following data would be used for this analysis.

2.1 Data 1: Boroughs and neighborhoods in New York:

New York has in total 5 boroughs and multiple neighbourhoods. To explore, analyze and segment neighborhoods, longitude and latitude of each neighbor and borough will be added.

Links to the dataset are:

New York City data containing the neighborhoods and boroughs, latitudes, and longitudes will be obtained from the data source: https://cocl.us/new_york_dataset

2.3 Data 3: Restaurants and markets in New York:

Foursquare API. **Foursquare API** will provide the required data that permits to provide venues information for each neighbourhood.

3. Methodology:
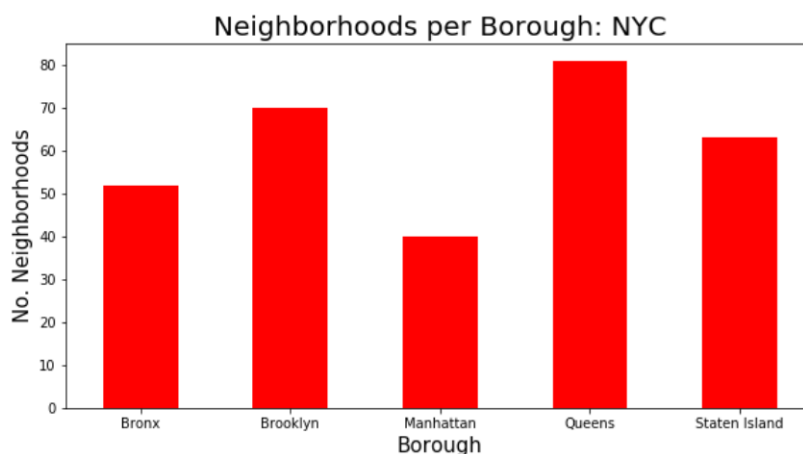
3.1 Business understanding:

Main goal of this project is to get locations that will be suitable for a business entrepreneur to open a restaurant in New York.

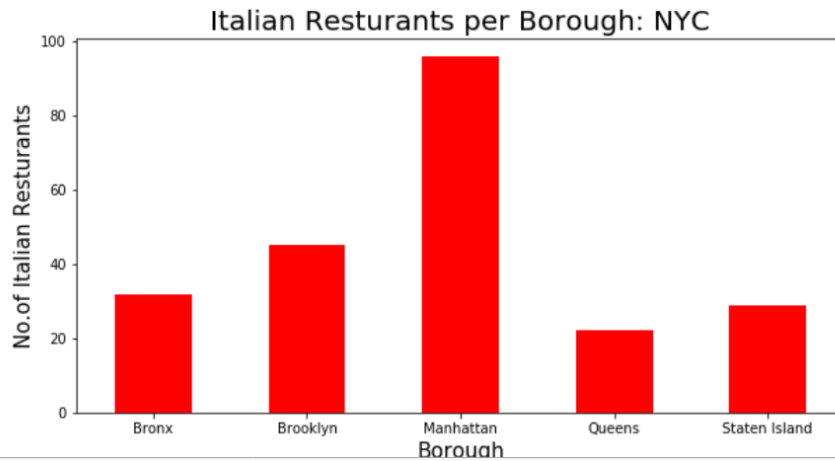## 3.3 Exploratory Data Analysis:

3.3.1. Data :

- Transform the data of nested python dictionaries into a data frame.
- This data frame contains geographical localization of famer market and information as opening time, company that is managed the market and type of products that are bought in the market.
- Geopy and Folium is two libraries that will be used to create a map of NY with marks for each market where name of market and number of borough where it is located are superimposed on top.
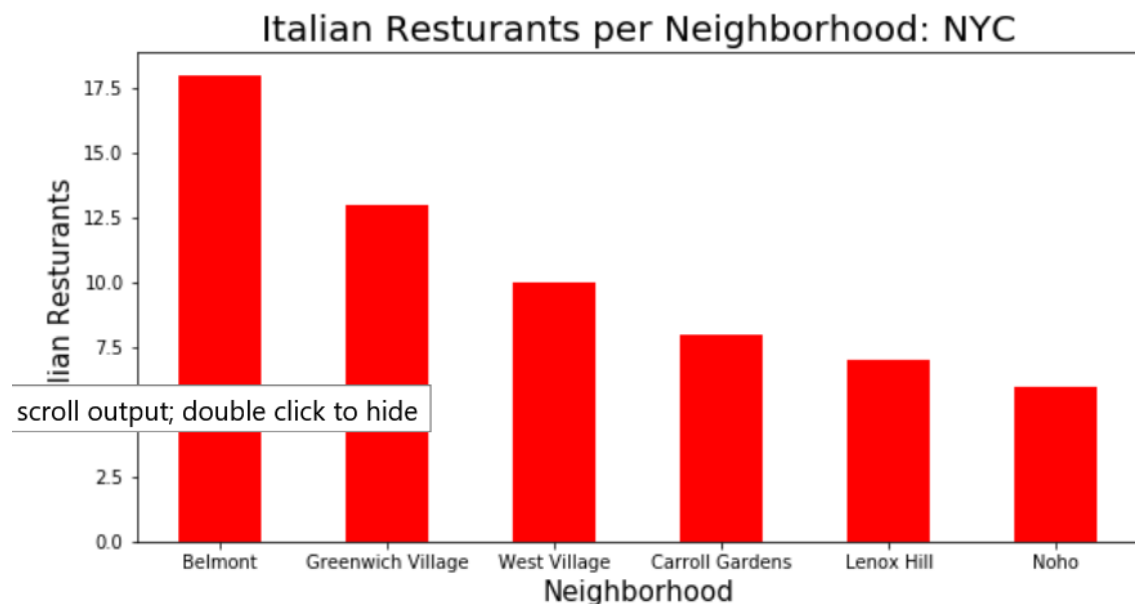
4. **Visualization**

The above bar chart shows the division of boroughs and neighborhoods in new York.
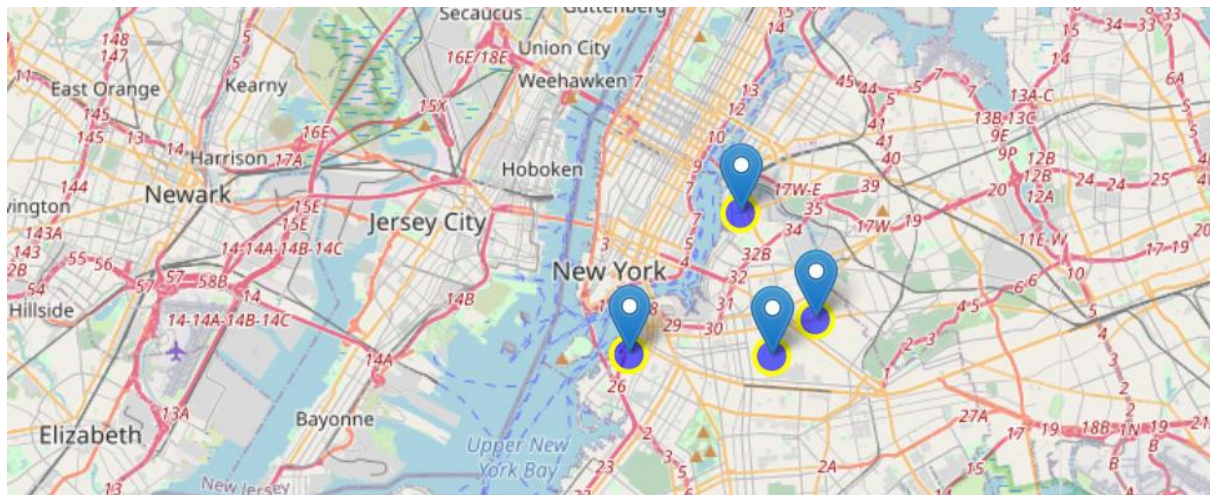
Italian Resturants per Borough: NYC

As can be seen from the above graph, Manhattan followed by Brooklyn and Bronx had the highest number of Italian restaurants. On the other hand, the graph below shows that Belmont neighbourhood had the highest number of restaurants. Noho, had the lowest restaurants,

Italian Resturants per Neighborhood: NYC

scroll output; double click to hide

3.4 Machine learning uses:

For segmenting neighborhood in New York based on existed restaurant, and finding recommendation for a good start of restaurant business, **clustering** will be used **K-mean method**. This method is used as unsupervised algorithm. It does not need previous recommendations to build a model. **K-mean method** is good for segmentation. It divides the data into clusters without any cluster-internal structures or label.

4. Results:

From venues data, dataset is filtered and used only data where venues are restaurant as we focused only on restaurant business.

4.1 Neighborhood K-Means clustering based on mean occurrence of venue category:

To cluster the neighborhoods into two clusters we used K-Means clustering Algorithm. K-Mean clustering aims to partition n observation into k clusters in which each observation belongs to the cluster with the nearest mean. It uses iterative refinement approach. In this project, 2 clusters are chosen. The map below, that is constructed using Geopy and Folium libraries, shows clustering of neighborhoods in NY into two clusters

Cluster 1 is marked by red purple markers. It is about neighborhoods where restauration market is saturated and it is not recommended here to start a restaurant business.

5. Conclusion:

Brooklyn and Bronx have the best rated Italian restaurants overall. Belmont of The Bronx is the neighborhood in all of NYC with the most Italian Restaurants. Despite Manhattan having the least number of neighborhoods in all five boroughs, it has the most Italian restaurants. Based on this information, Woodland and Pehlam Gardens are the best locations for Italian cuisine in NYC if our criteria is looking at least average ratings. The ratings are less and so the competition would be lower.

6. Future directions:

At the end, multiple data sources such as median income, rental values, population, race should be taken into account to further build on this analysis.