Coursera Capstone IBM Applied Data Science

Opening an Italian Restaurant in New York

Business Problem

- The objective of this Capstone project is to analyze and select the best locations in the city of New York to open a new Italian restaurant. Using Data Science methodology and instruments such as Data Analysis and Visualization, this project aims to provide solutions to answer the business question: Where in the city of New York, should the investor open an Italian Restaurant?

Data

- Data required:
1. List of borough and neighborhoods of New York
2. Latitude and longitude coordinates of those neighborhoods
3. Venue data, particularly data related to restaurants
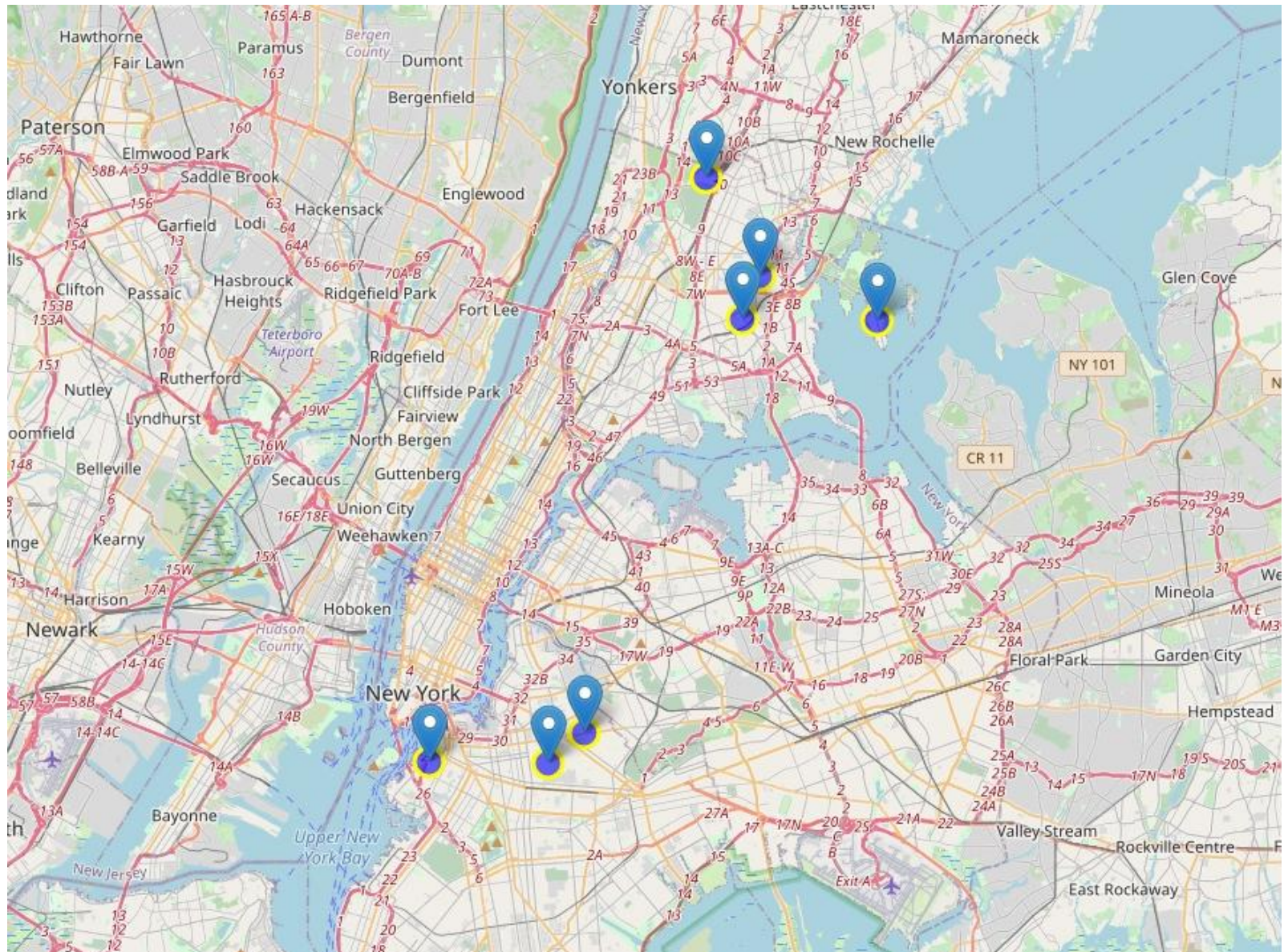
- Sources of data:
1. Open source dataset of New York: https://cocl.us/new_york_dataset
2. Geocoder package for latitude and longitude coordinates
3. Foursquare API for venue data

Methodology

- Collect the data from https://cocl.us/new_york_dataset.

- Cleaneand process the data into a dataframe.

- Using FourSquare, locate all venues and then filter by Japanese restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.

- Sort the data based on rankings.

- Visually assess the data be will be using graphing from Python libraries.

Results

- Manhattan has the highest number of Italian restaurants comparing to the other boroughs.

- Manhattan has the highest average rating for Italian Restaurants.

- Two neighborhoods with the highest average rating for Italian Restaurants are Cobble Hill and North Side (both located in Brooklyn).

Conclusion

- Based on the results of our analysis, Bronx and Brooklyn are the best locations for Italian cuisine in NYC.

- Brooklyn is a more promising place to open a new Italian restaurant.

- Neighborhoods Bedford Stuyvesant or Bushwick are the best places in Brooklyn to open a new Italian restaurant.