Capstone Project- Battle of Neighborhood

Table of Contents

- Business Problem

- Data Selection & Source

- Methodology

- Results & Decision

- Conclusion

Business Problem

The aim of this project is to find a safe and secure location for opening of Italian restaurants in the city of New York. New York is the most populous city in the United States with an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km2) which makes it the most densely populated major city in the United States(Wikipedia). New York City has the largest population of Italian Americans in the United States of America as well as North America. Over 1.3 million Italians and Italian-Americans live in the greater New York City metro area, with about 800,000 living within one of the five New York City boroughs. For this big population the need to find and enjoy Italian cuisine is on the rise. We will make use of our data science tools to analyse data and focus on the safest borough and explore its neighborhoods to find the best italian restaurants in the city of New York. Also we are going to cluster neighborhoods in order to recommend the best locations of opening italian restaurant and the best places to enjoy italian cuisine.

Data

Based on definition of our problem, factors that will influence our decision are:

- Finding boroughs to include boundaries, latitude, longitude, restaurants, and restaurant ratings

- Finding the most common venues for easy access

- Selecting the right neighborhood within the borough

- Data source: https://cocl.us/new_york_dataset & https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm
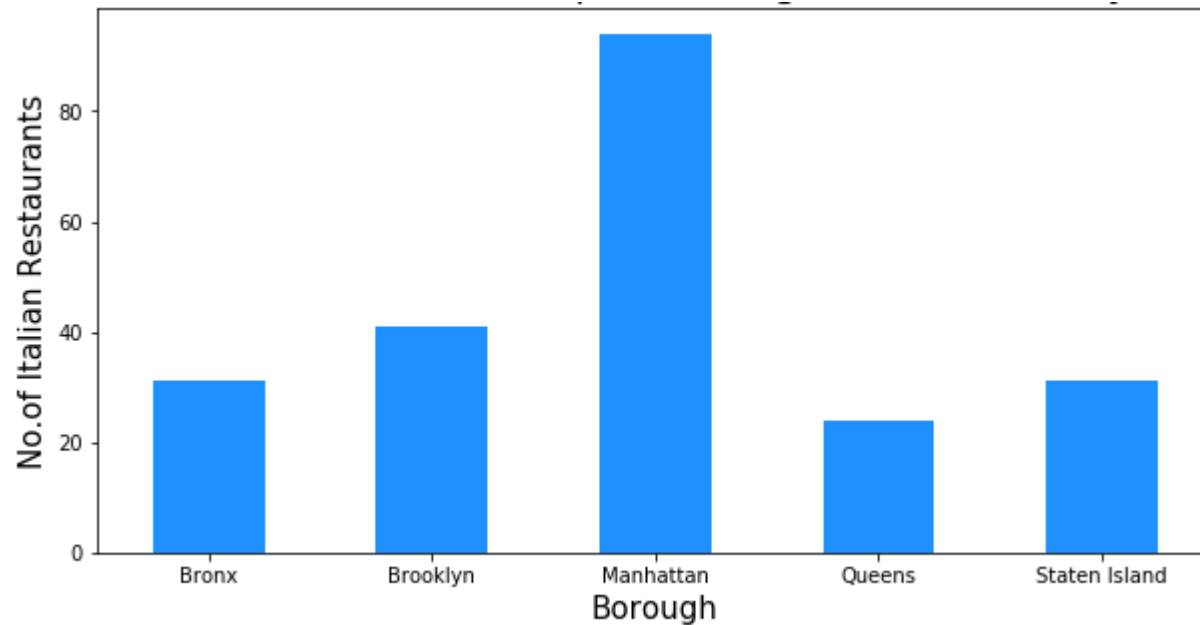
Methodology

- Collect Inspection Data

- Using Foursquare locating all venues filtered by Italian restaurants. Ratings, tips, and likes by users

- Explore and Understand Data

- Data preparation and preprocessing

- Modeling
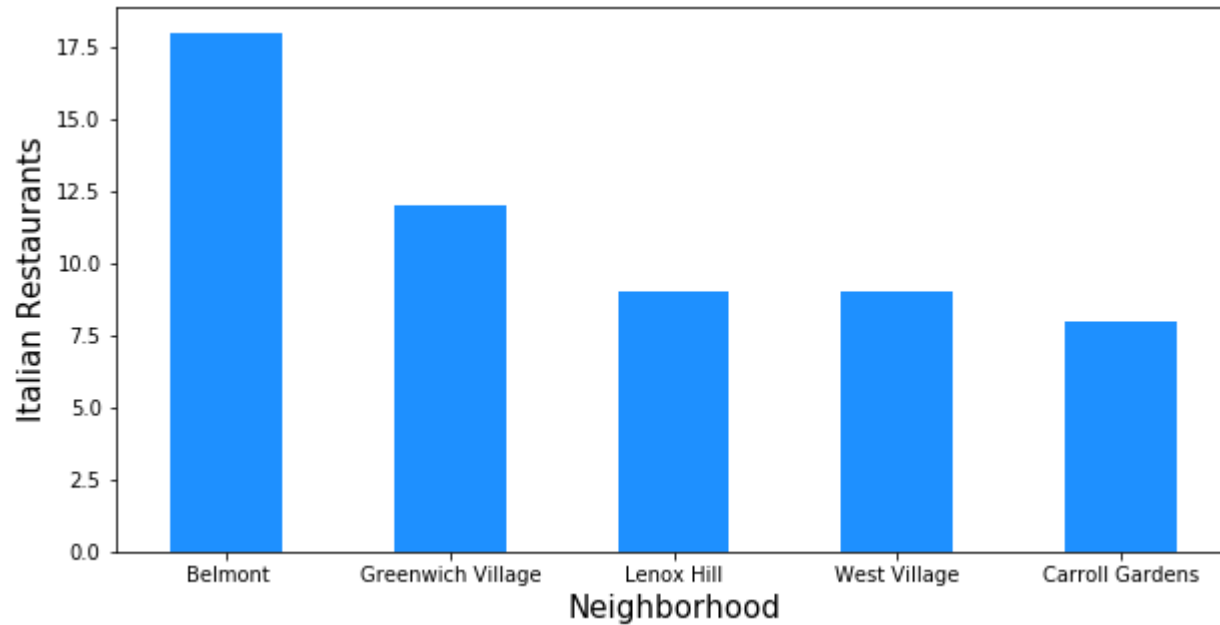
- Data Visualization using various Python libraries.

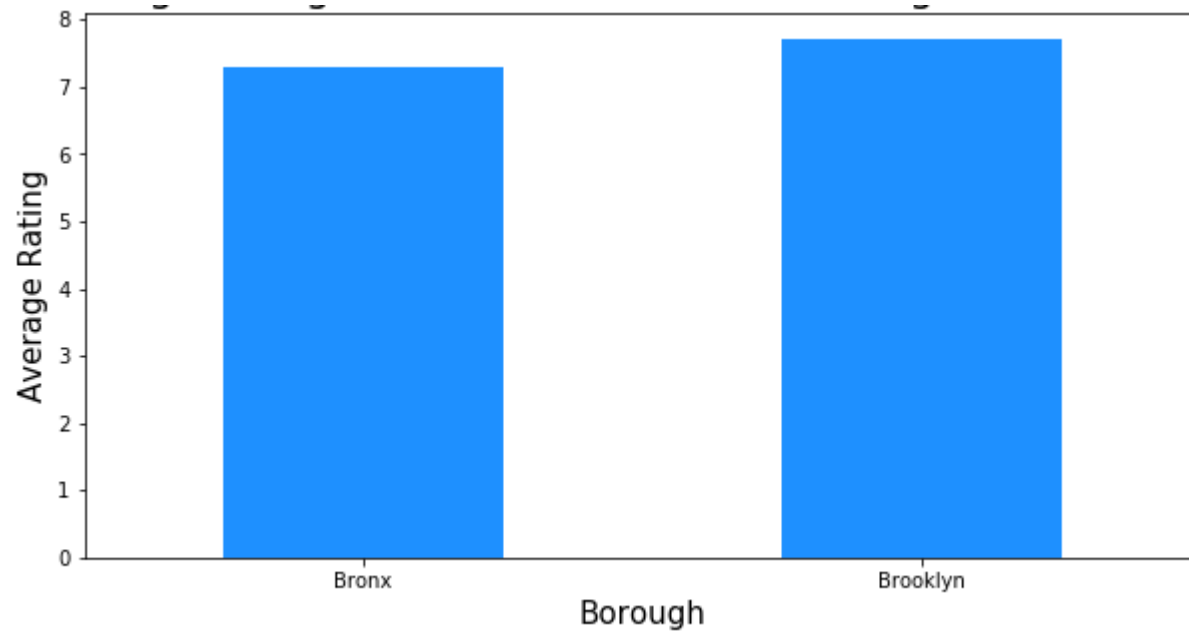# Neighborhoods per Borough: New York City
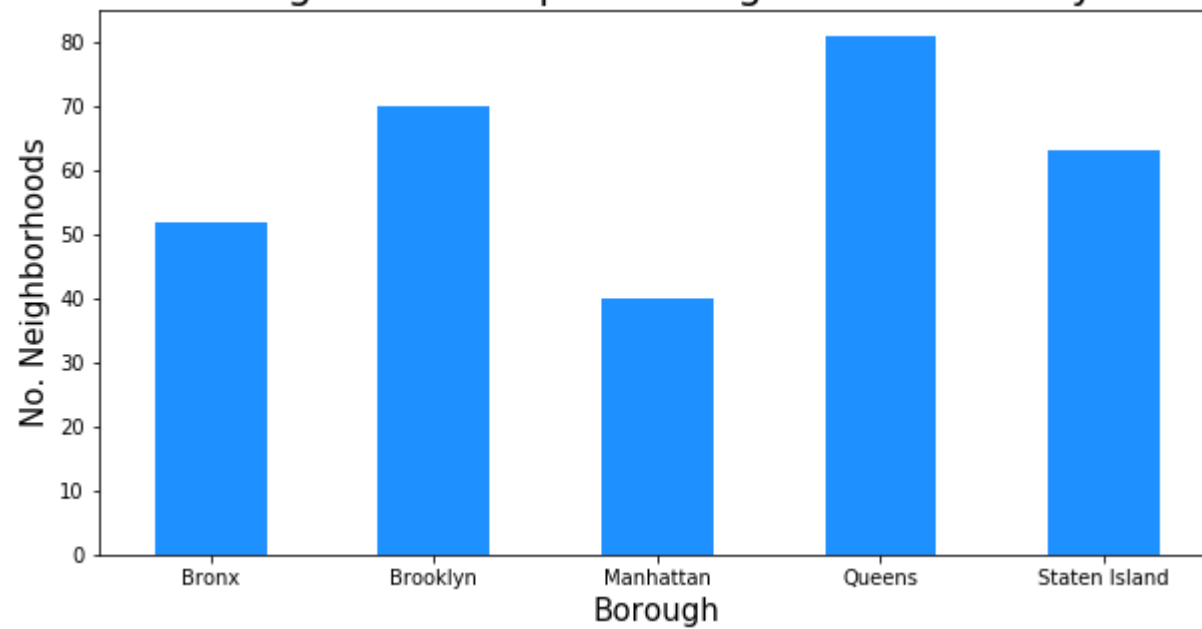
Italian Restaurants per Borough: New York City

Italian Restaurants per Neighborhood: New York city

Most Average Rating in Borough New York

Most Average Rating per Neighborhood

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Brooklyn	Bushwick	40.698116	-73.925258	9.4
1	Bronx	City Island	40.847247	-73.786488	8.2
2	Brooklyn	Cobble Hill	40.687920	-73.998561	8.5
3	Bronx	Morris Park	40.847549	-73.850402	8.5
4	Bronx	Pelham Gardens	40.862966	-73.841612	8.7
5	Bronx	Woodlawn	40.898273	-73.867315	8.6

Results & Decision

- Here we can conclude that borough wise Manhattan has most Italian restaurant. On the other hand Bronx has as twice as many in half the number of restaurants in Manhattan. After analyzing the data we found that Bronx's Belmont neighborhood has more restaurants than Manhattan neighborhood Lenox Hill. Considering Likes and Tips, Belmont is much ahead out of other neighborhoods. Most importantly, Bushwick neighborhood has highest average rating than other neighborhood in Brooklyn. In Bronx, Woodlawn and Pelham Garden's restaurant have highest average ratings than other neighborhood.

- Finally, we drew the conclusion that Bronx is the best option to open an Italian restaurants. In Bronx, most restaurants in neighborhood have on average 8.5 rating out of 10. Finally, I would go to Belmont in Bronx for the best Italian food based on 145 likes.

Conclusion

- To sum up, finding best places to open Italian restaurant throughout the city in New York we included boundaries, latitude, longitude, restaurants, and restaurant ratings and sorted the most common venues for easy access and finally selected the right neighborhood within the borough .Although in this report, it demonstrates the relations between location and ratings, but ratings might not reflect the operation status of the restaurant. A restaurant with a high rating could still be unprofitable, which is unsuccessful from a business perspective As a final note, all of the above analysis is depended on the adequacy and accuracy of Four Square data. To suggest more practical and profitable ideas, the relationship between customer reactions and financial performance should be evaluated.