IBM Applied Data Science Capstone

Capstone Project - The Battle of Neighborhoods

Evaluating Brazilian restaurants in NYC

Rafael Andrello Rubo

November 2020

Introduction

- The Brazilian immigration wave to the US hit its peak in the 1990s and early 2000s. During the 17 years between 1990 and 2007 over 275,000 Brazilians migrated to the United States. In total, around 7% of the entire US Brazilian population lives in New York City.

- According to the United States Census Bureau, in a survey from 2012, there were 72,635 brazilians living in NYC. It is the main metropolitan area of brazilian population, followed by Boston (63,930) and Miami (43,930).

- So, where exactly do Brazilians hang out in the city? Where can you find feijoada (a stew of beans with beef and pork) and other delicious Brazilian foods?

- This project will attempt to answer the questions "Where should an investor open a Brazilian Restaurant in NYC?" and "Where should I go If I want great Brazilian food?"

Data

- In order to answer the above questions, data on New York City neighborhoods, boroughs to include boundaries, latitude, longitude, restaurants, and restaurant ratings and tips are required.

- 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

- All data related to locations and quality of Brazilian restaurants will be obtained via the FourSquare API utilized via the Request library in Python.

Methods

- Data will be collected from https://cocl.us/new_york_dataset and cleaned and processed into a dataframe.

- FourSquare be used to locate all venues and then filtered by Brazilian restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.

- Data will be sorted based on rankings.

- Finally, the data be will be visually assessed using graphing from Python libraries.

Results

- After retrieving data from neighborhoods and boroughs of New York City, the loaded data was visualized:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585
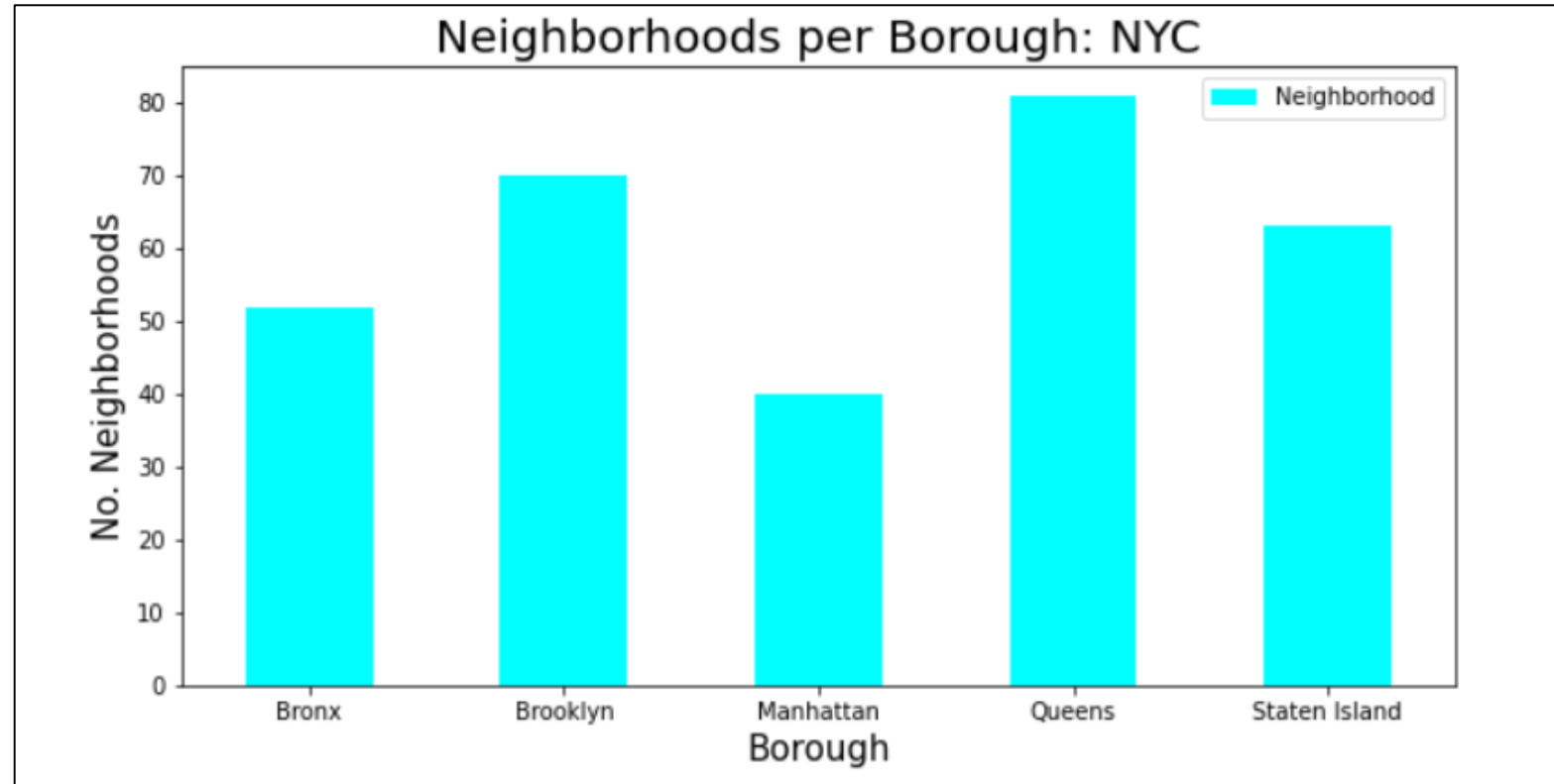
Results

- Using the FourSquare API, Brazilian restaurants were retrieved from New York City.

	Borough	Neighborhood	ID	Name
0	Manhattan	West Village	3fd66200f964a52008e71ee3	Casa
1	Queens	Astoria	4bdf502a89ca76b062b75d5e	Favela Grill

- There are only two brazilian restaurants in NYC, even though the brazilian population reaches its tens of thousands in the City.

- At least, that is the information retrieved. It is possible that there are many other restaurants that are not classified as "brazilian restaurant".

Results

- These are commom issues in data science. We will work with these two instances of data and try to get the most of them.
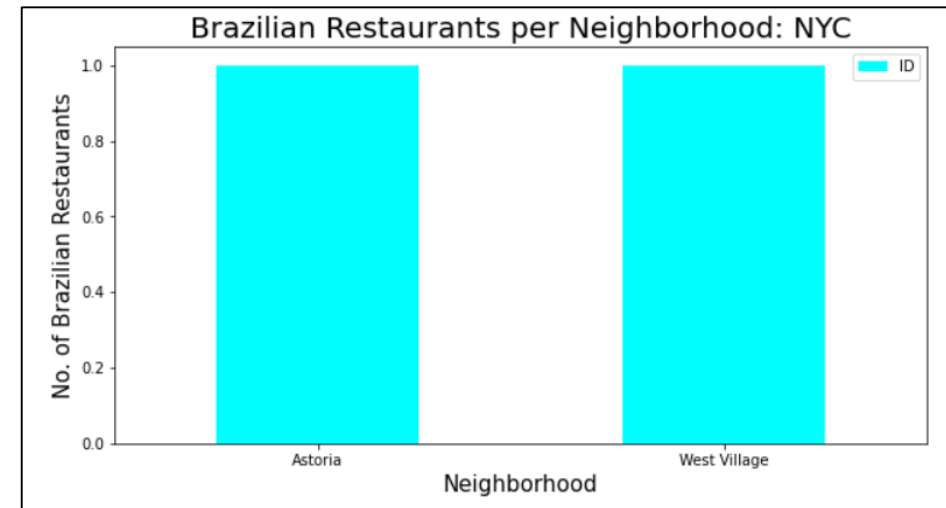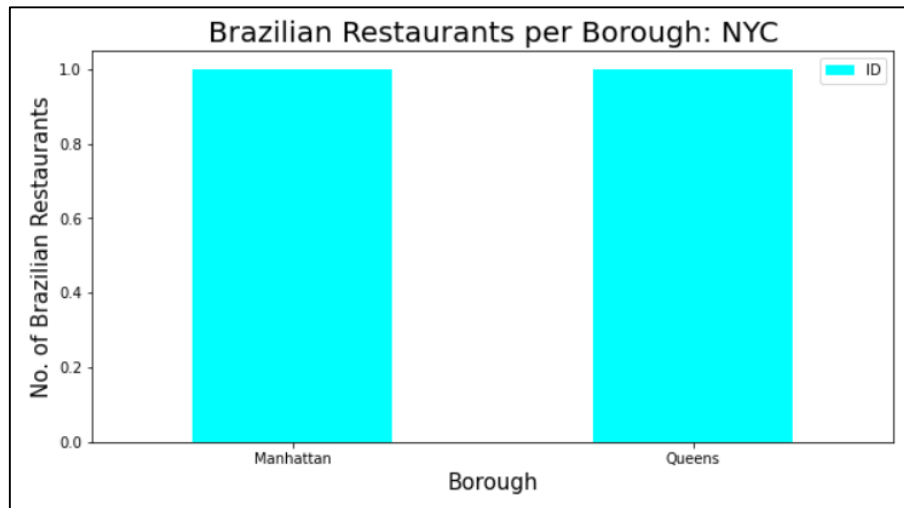
- While many Brazilians have dispersed throughout the city, Astoria seems to be the main home of Brazilians in NYC. It is expected that we would find a brazilian restaurant there.

- West Village's option may be more related to touristic activities, considering it is on the Manhattan Borough.

Results

- The stats for both restaurants were obtained via the FourSquare API.
- Average from one observation! Statistics with no data...
- Even though it makes no sense to get too much statistics from these two observations, some were obtained in order to create a code that would be useful with a different dataset.


Average rating of Brazilian Restaurants for each Borough
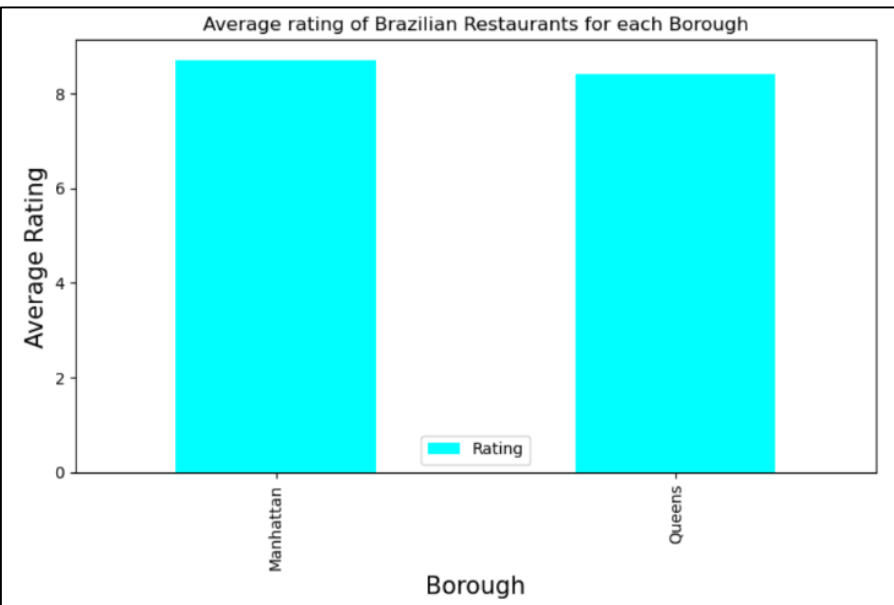
	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Manhattan	West Village	3fd66200f964a52008e71ee3	Casa	194	8.7	66
1	Queens	Astoria	4bdf502a89ca76b062b75d5e	Favela Grill	86	8.4	57

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Queens	Astoria	40.768509	-73.915654	8.4
1	Manhattan	West Village	40.734434	-74.006180	8.7

	Likes	Rating	Tips
count	2.000000	2.000000	2.000000
mean	140.000000	8.550000	61.500000
std	76.367532	0.212132	6.363961
min	86.000000	8.400000	57.000000
25%	113.000000	8.475000	59.250000
50%	140.000000	8.550000	61.500000
75%	167.000000	8.625000	63.750000
max	194.000000	8.700000	66.000000

Results

- Finally, the data was plotted in a map for better visualization.
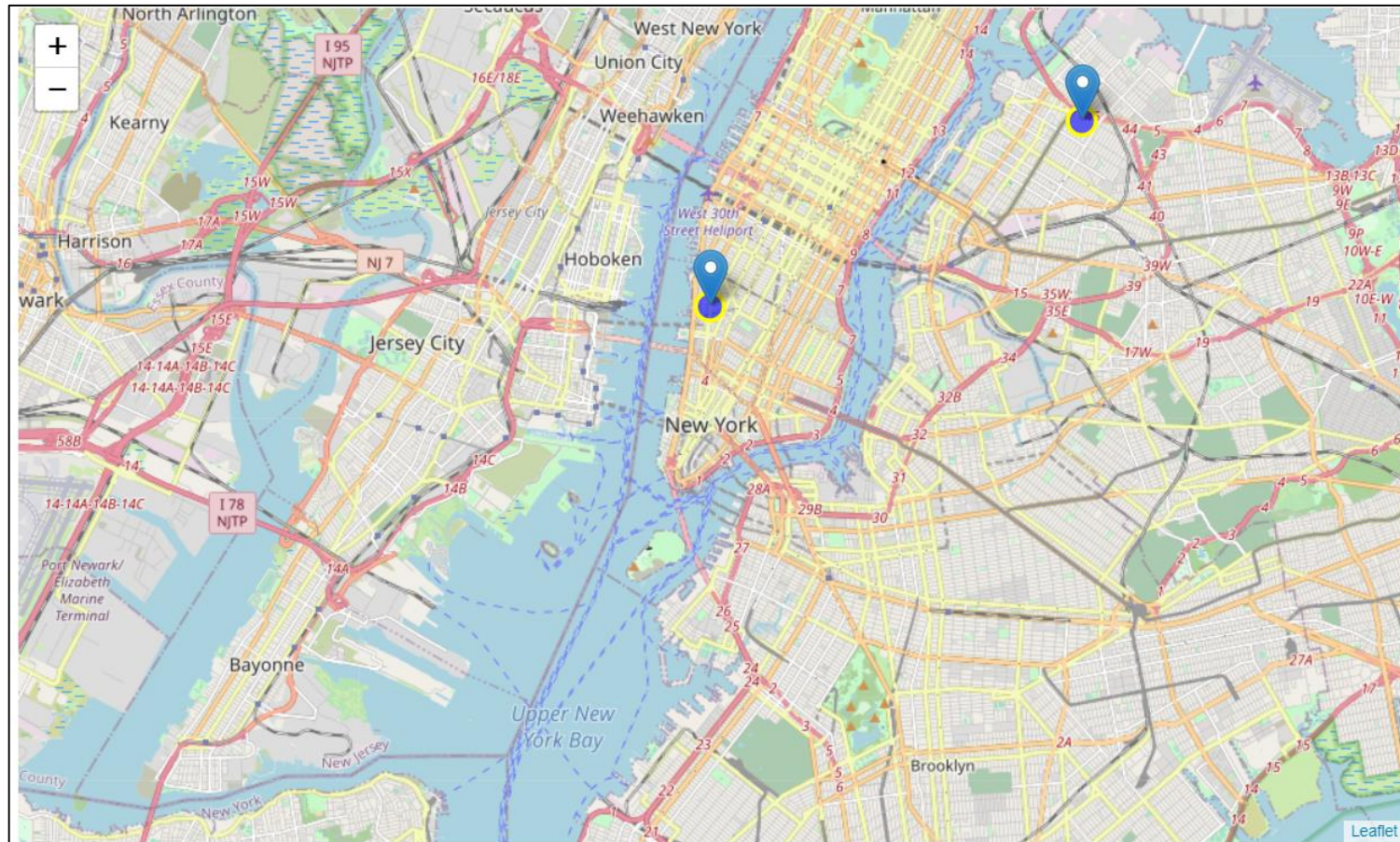
Conclusions

- Both West Village (Manhattan) and Astoria (Queens) have high rates for their brazilian restaurants.

- During the course of this project, there was this main issue related to the amount of data. Even though there is an elevated brazilian population in NYC and there probably are other brazilian restaurant options, these alternatives are not registered in FourSquare, or they are classified as something other than "brazilian restaurant".

- Even with that issue, it was possible to create a project that used most of the tools and methods that were studied during the course.

- The West Village restaurant seems to be the best brazilian restaurant option according to its statistics in FourSquare. However, it is closely followed by the Astoria restaurant. Other factors, such as accessibility might be added to the evaluation in order to choose between them. Anyway, the first choice would be the West Village restaurant, but it would be interesting to try both of them.

- If it were to open a new Brazilian restaurant, Manhattan or Queens would be interesting choices of borough. Manhattan seems to be able to have a greater public, considering its higher number of "Likes".