**IBM FINAL CAPSTONE**

**Battle of the Neighborhoods – Part 2**

**Italian Restaurants**

Xavier Lowe

**May 4, 2020**

**-------------------------------------------------------------------------------------------------------**

**Introduction**

This final project explores the best locations for Italian restaurants throughout the city of New York. Food Business News stated that worldwide pasta sales were up for the second year in a row with the United Sates holding the largest market (Donley, 2018). New York is a major metropolitan area with more than 8.4 million (Quick Facts, 2018) people living within city limits. Most of the Italian immigration into the United States occurred during the late 19th and early 20th century with over two million immigrants between 1900 and 1910.

Italian families first settled in Little Italy’s neighborhood around Mulberry Street as has continued to thrive ever since. With almost 700,000 Manhattan inhabitants reporting Italian ancestry, the need to find and enjoy Italian cuisine is on the rise. This report explores which neighborhoods and boroughs of New York City have the most as well as the best Italian restaurants. Additionally, I will attempt to answer the questions “Where should I open an Italian Restaurant?” and “Where should I stay If I want great Italian food?”

Therefore, there are a couple different questions that this analysis project will try to answer.

1. What is / are the best location(s) for Italian cuisine in New York City?
2. In what Neighborhood and/or borough should I open an Italian restaurant to have the best chance of being successful?
3. Where would I go in New York City to have the best Italian 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. There are a couple of data sources that will be used as listed below.

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>

New York City data containing neighborhood boundaries will be obtained from the data source: <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>

All data related to locations and quality of Italian restaurants will be obtained via the Four Square API utilized via the Request library in Python. The different data points and services will be used to create an overall data frame that can be used to analyze the population density, restaurant location density and other variables as it pertains to the analysis.

**Methodology**

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 Italian restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.

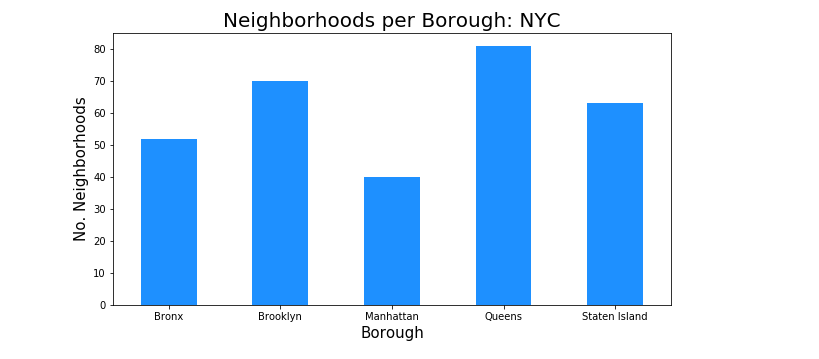
Data will be sorted based on rankings

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

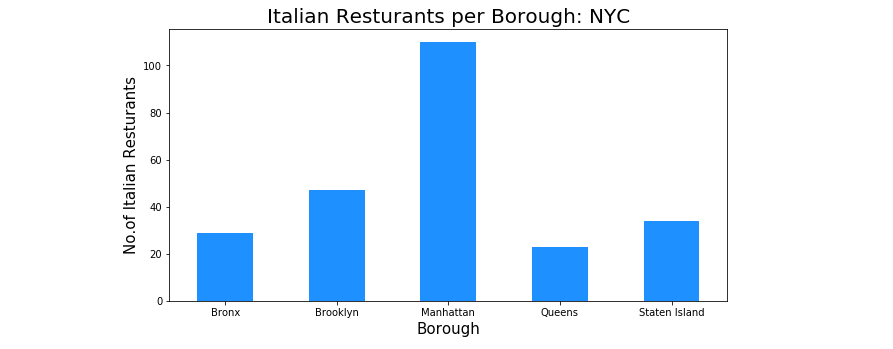
**Neighborhood population calculation**

While the data for total population and number of neighborhoods within any given borough is available within the dataset, I decided that it would be easier to visual the overall top cities with the highest amount of neighborhoods first. This is depicted by the first graph (figure 1) where the data clearly illustrates that Queens and Brooklyn have the highest amount of neighborhoods within their respective borough’s, with Manhattan having the lowest.

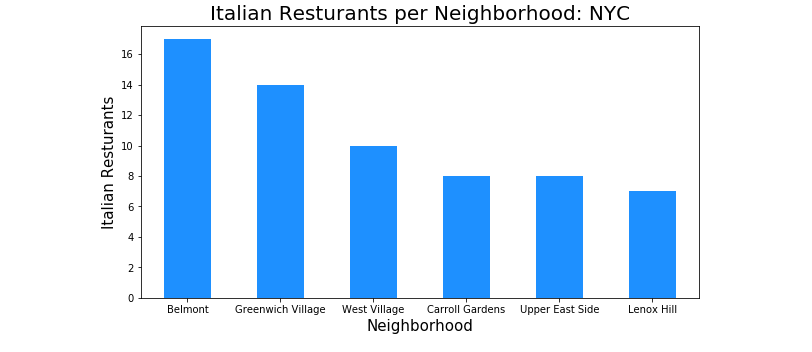
urSquare be used to locate all venues and then filtered by Italian restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.

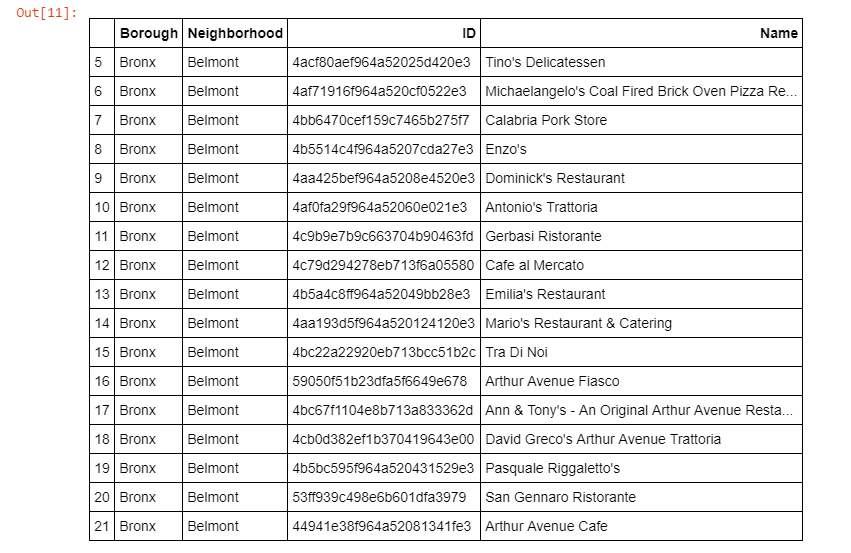
Figure 1

Next, I wanted to utilize the same established boroughs but focus on the concentration of Italian Restaurants within the boroughs of the New York City area. As you can see from Figure 2 below, the highest concentration of Italian Restaurants are within Manhattan, while the lowest concentration of restaurants are located in Queens. As I continued my analysis, I saw below that although Manhattan had the least number of neighborhoods, it does have the highest number if Italian restaurants. Additionally, we see how many restaurants the top 6 neighborhoods have. The neighborhood of Belmont has the highest number of Italian restaurants in all of NYC and is actually located in the borough of Bronx vice Manhattan (Figure 3 & 4)

 Figure 2

As I continued my analysis, I saw below that although Manhattan had the least number of neighborhoods, it does have the highest number if Italian restaurants. Additionally, we see how many restaurants the top 6 neighborhoods have. The neighborhood of Belmont has the highest number of Italian restaurants in all of NYC and is actually located in the borough of Bronx vice Manhattan (Figure 3 & 4)

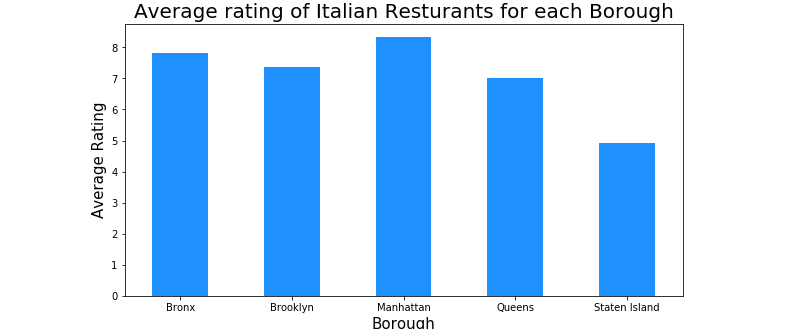
Figure 3

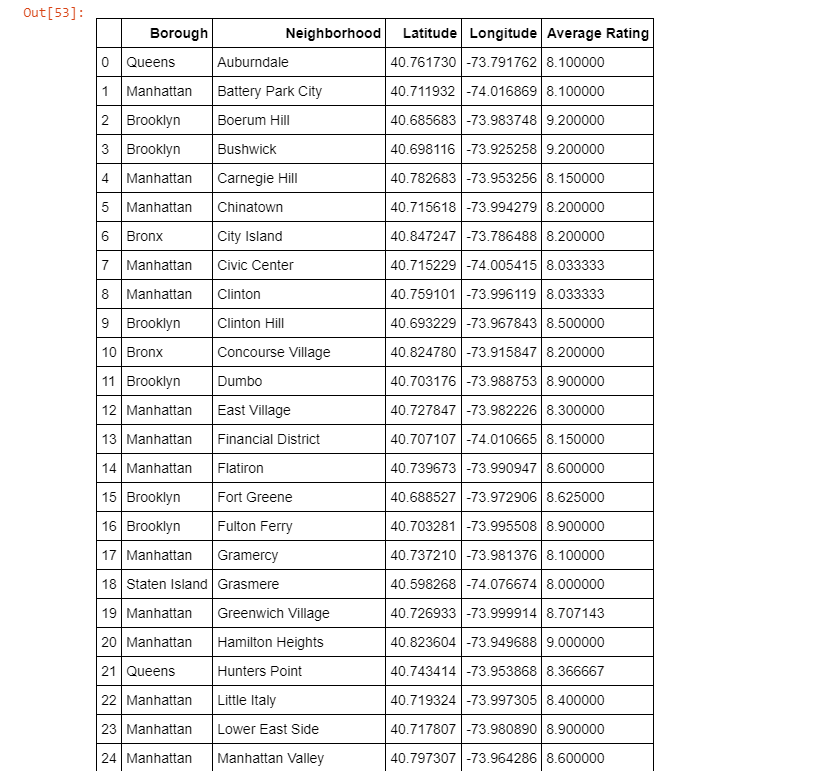
Figure 4

All the above figures describe various aspects of the distributions of both neighborhoods and restaurants within the given boroughs of New York City. These graphs/ data frames along with a host of others to look at distribution data in a number of different ways. These data points directly correlate to the questions and overall purpose of the report/analysis and supplies us with the information about location of current restaurants to inspire future locations when possible.

**Results/Discussion**

Based upon the analysis, Manhattan and Queens have the best rated Italian restaurants on average. Queens and The Bronx have the least amount of Italian restaurants per borough. However, of note, 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, I would state that Manhattan and Queens are the best locations for Italian cuisine in NYC. This can be seen and illustrated by the different data frames and graphs below with have the average rating calculated for each neighborhood as well as location and disparity

Figure 5

Figure 6

To have the best shot of success, I would open an Italian restaurants in Queens. Queens has multiple neighborhoods with average ratings exceeding 8.0 of a scale of 1.0 to 10.0 and has the least number of Italian restaurants making competition easier than in other boroughs. Finally, I would go to Gramercy in Manhattan for the best Italian food based on 131 likes. As a final note, all of the above analysis is depended on the adequacy and accuracy of Four Square data. A more comprehensive analysis and future work would need to incorporate data from other external databases.

**Conclusion**

This final project explores the best locations for Italian restaurants throughout the city of New York and therefore where there would be the best opportunity to create a new establishment. It is to be noted that this analysis is not to be used as the only data points determined to help produce a qualifying decision as there are other variables that are not accounted for in the analysis. By using information regarding most populated boroughs, current competition within boroughs in the NYC area and ratings received by patrons within those same surrounding areas. I have come to the conclusion that to have the best shot of success, I would open an Italian restaurants in Queens. Queens has multiple neighborhoods with average ratings exceeding 8.0 of a scale of 1.0 to 10.0 and has the least number of Italian restaurants making competition easier than in other boroughs. However, if the goal is just to taste Italian food, then I would go to Gramercy in Manhattan for the best Italian food.