

Coursera Capstone
IBM Applied Data Science Capstone
Opening an Italian Restaurant in New York
By: Ariel Romero
December 2020



Introduction

New York City's demographics show that it is one of the largest and most diverse cities in the entire world. It is the largest city in the United States with a long history of international immigration. Over the last decade, the city has been growing faster than the region. The New York region continues to be by far the leading metropolitan gateway for legal immigrants admitted into the United States. This final project explores the best locations for Italian restaurants throughout the city of New York. Potentially the owner of the new Italian restaurant can have great success and consistent profit. However, as with any business, opening a new restaurant requires a lot of preparation and it is more complicated than it seems from the first glance. In particular, the location of the restaurant is one of the most important factors that will affect whether it will have success or a failure. Location is important because you do not want to be in a location where competition is at your disadvantage. So our project will attempt to answer the questions “Where should a chef open an Italian Restaurant?” and “Where should I go if I want Italian food?”

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 chef open an Italian Restaurant? Target Audience of this project and some demographic facts: This project is particularly useful to developers and investors looking to open or invest in an Italian restaurant in the city of New York. Overall, New York City is a great place to open a restaurant with an ethnical cuisine. As New York is the most diverse city in the world (800 languages are spoken in New York). With its diverse culture, comes diversity in the food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, French, etc. Why did we decide to focus on Italian cuisine in our project? Italian restaurants have been growing throughout the years recently as dining out is becoming more and more popular.

Data

To solve the problem, we will need the following data:

- New York City data containing the neighborhoods and boroughs.
- Latitude and longitude coordinates of those neighborhoods. This is required to plot the map and get the venue data.
- Venue data, particularly data related to restaurants. We are going to use this data to perform further analysis of the neighborhoods. This project will require using of many data science skills, from web scrapping (open source dataset), working with API (Foursquare), data cleaning, data wrangling, to map visualization (Folium).

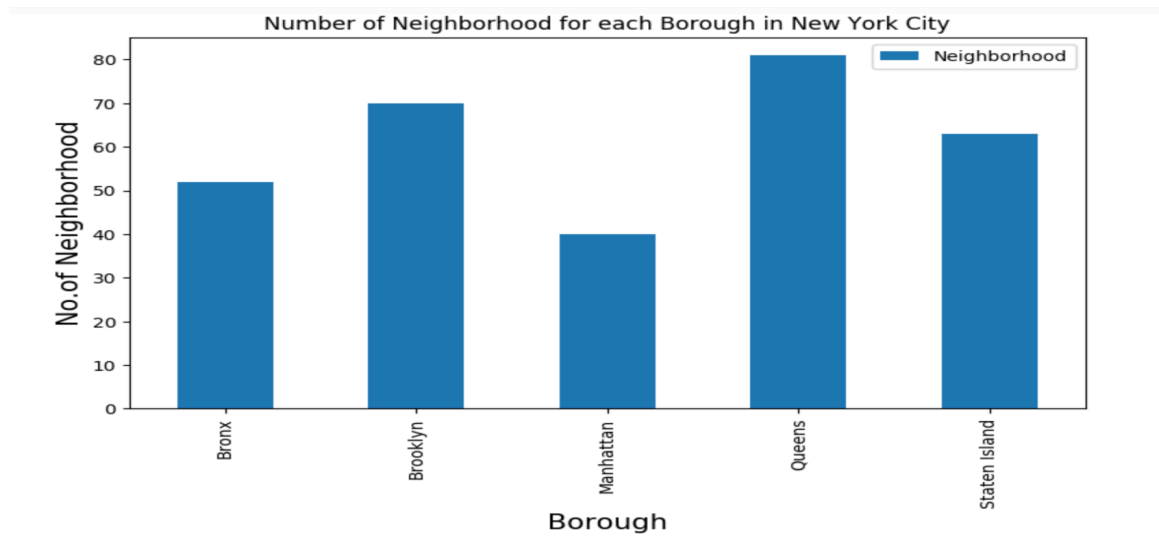
In the next Methodology section, we will discuss and describe any exploratory data analysis that we did, any inferential statistical testing that we performed, and what machine learning techniques were used.

Methodology

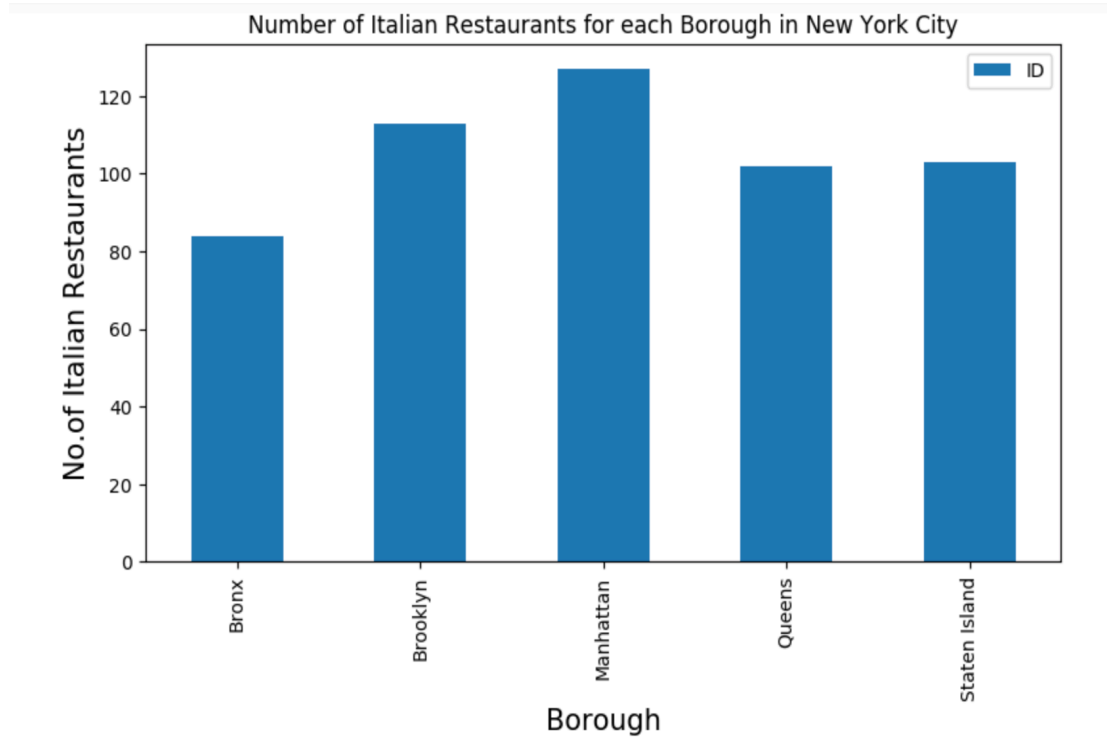
- Data will be collected from https://cocl.us/new_york_dataset and cleaned and processed into a dataframe.
- FourSquare be used to locate all venues and then filtered by Italian restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.
- Data will be sorted based on rankings.
- Finally, the data be will be visually assessed using graphing from Python libraries.

Results

1. We see that Queens has the highest number of Neighborhoods.



2. Although Manhattan had the least number of neighborhoods, it does have the highest number of Italian restaurants.



3. The neighborhood with the highest rated Italian restaurant is located in Belmont, Bronx with a rating of 8.7

	Neighborhood	Average Rating
2	Belmont	8.685714
16	Woodlawn	8.500000
9	Morrisania	8.200000
5	Fordham	8.050000
0	Baychester	8.050000
14	Van Nest	8.000000
7	Melrose	7.800000
11	Pelham Parkway	7.766667
1	Bedford Park	7.700000
8	Morris Park	7.500000

Top Ten neighborhoods with Italian Restaurants with the best average rating

Discussion Section

Based on the results of our analysis, I would state that Manhattan and Brooklyn are the best locations for Japanese cuisine in NYC. To have the best shot of success, I would open an Italian restaurant in Bronx. I believe due to relatively low amount of Italian restaurants in comparison to the rest of the Boroughs provides opportunity to expand in an untapped market. Manhattan is oversaturated with Italian restaurants which can cause difficulty to set themselves apart from competition.

Limitations and suggestions for future researches

All of the above analysis is depended on the accuracy of Four Square data. Besides, during this project, we used a free Sandbox Tier Account of Foursquare API that goes with limitations as to the number of API calls and results returned. To get better results, future research work and more comprehensive analysis could consider using a paid account to bypass these limitations as well as incorporating data from other external databases.

Conclusions

In the project we have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, performing data analysis, and lastly providing recommendations to the chefs/investors. During the project, we applied different data science methods and instruments to get the answer to our main question: "Where in the City of New York, should the chef open an Italian Restaurant?" The findings of this project will help the relevant investor better understand the

advantages and disadvantages of different New York neighborhoods/boroughs in terms of opening an Italian restaurant.