

Capstone Project - The Battle of the Neighborhoods

Applied Data Science Capstone by IBM/Coursera



Opening a Japanese Restaurant in New York City

Abdurrahman Öztürk
June 29, 2021



1. Introduction: Business Problem

New York City (NYC), often simply called 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), New York City is also the most densely populated major city in the United States. Located at the southern tip of the State of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban area. With almost 20 million people in its metropolitan statistical area and approximately 23 million in its combined statistical area, it is one of the world's most populous megacities. New York City has been described as the cultural, financial, and media capital of the world, significantly influencing commerce, entertainment, research, technology, education, politics, tourism, art, fashion, and sports, and is the most photographed city in the world. Home to the headquarters of the United Nations, New York is an important center for international diplomacy, and has sometimes been called the capital of the world.

A restaurant is a business which prepares and serves food and drink to customers in return for money, either paid before the meal, after the meal, or with an open account. Overall, New York 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.. So, it is evident that to survive in such competitive market it is very important to strategically plan.

Various factors need to be studied in order to decide on the location such as: New York Population, New York City Demographics, Are there any Farmers Markets, Wholesale markets etc nearby so that the ingredients can be purchased fresh to maintain quality and cost?, Are there any venues like Gyms, Entertainment zones, Parks etc nearby where floating population is high etc Who are the competitors in that location?, Cuisine served / Menu of the competitors, Segmentation of the Borough even though well-funded, XYZ Company Ltd. need to choose the correct location to start its first venture. If this is successful, they can replicate the same in other locations. This would interest anyone who wants to start a new restaurant in NYC.

This final project explores the best locations for Japanese restaurants throughout the city of New York. Potentially the owner of the new Japanese restaurant can have great success and consistent profit. Now when the idea of a healthy lifestyle conquered the minds of people all over the country, Japanese restaurants became extremely popular, as they offer a healthy alternative to regular American eating habits. However, as with any business, opening a new restaurant requires serious considerations and 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. So our project will attempt to answer the questions "Where should the investor open a Japanese Restaurant?" and "Where should I go If I want great Japanese food?"

2. Data Section

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 restaurants will be obtained via the FourSquare API utilized via the Request library in Python.

At the end of this project, questions that will be answers:

1. What are the best location(s) for Japanese cuisine in New York City?
2. Which areas have potential Japanese Restaurant Market?
3. In what neighborhood and/or borough should the investor open a Japanese restaurant to have the best chance of being successful?
4. Which is the best place to stay if you prefer Japanese Cuisine?

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). We will be using the below datasets for analysing NYC.

Data 1: Neighbourhood has a total of 5 boroughs and 306 neighbourhoods. In order to segment the neighbourhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighbourhoods that exist in each borough as well as the latitude and longitude coordinates of each neighbourhood. Data source: https://cocl.us/new_york_dataset Description: This data set contains the required information. And we will use this data set to explore various neighbourhoods of New York city.

Date 2: Japanese restaurants in each neighbourhood of New York city. Data source: Foursquare API Description: By using this api we will get all the venues in each neighbourhood. We can filter these venues to get only Indian restaurants.

Data 3: GeoSpace data Data source: <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmi-j8zm> Description: By using this geo space data we will get the New York Borough boundaries that will help us visualize choropleth map.

3. Methodology

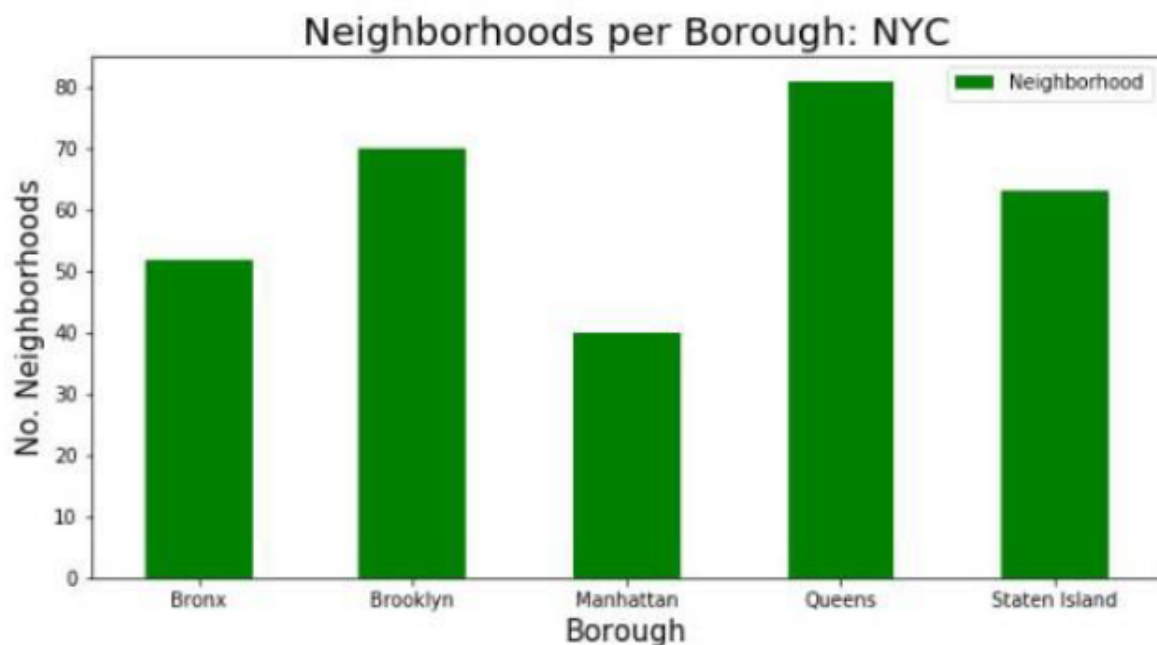
The required libraries are imported for python pandas and numpy for handling data, request module for using FourSquare API, geopy to get co-ordinates of City of New York, folium to visualize the results on a map.

- 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 Japanese 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.

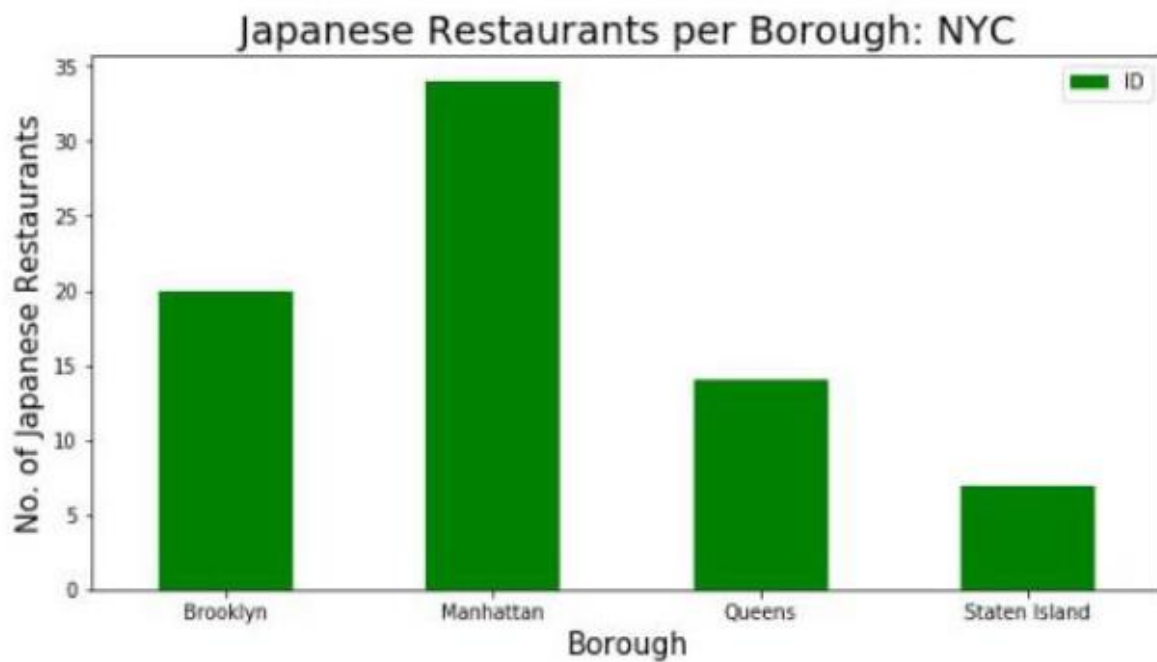
4. Results

The results of our analysis showed below:

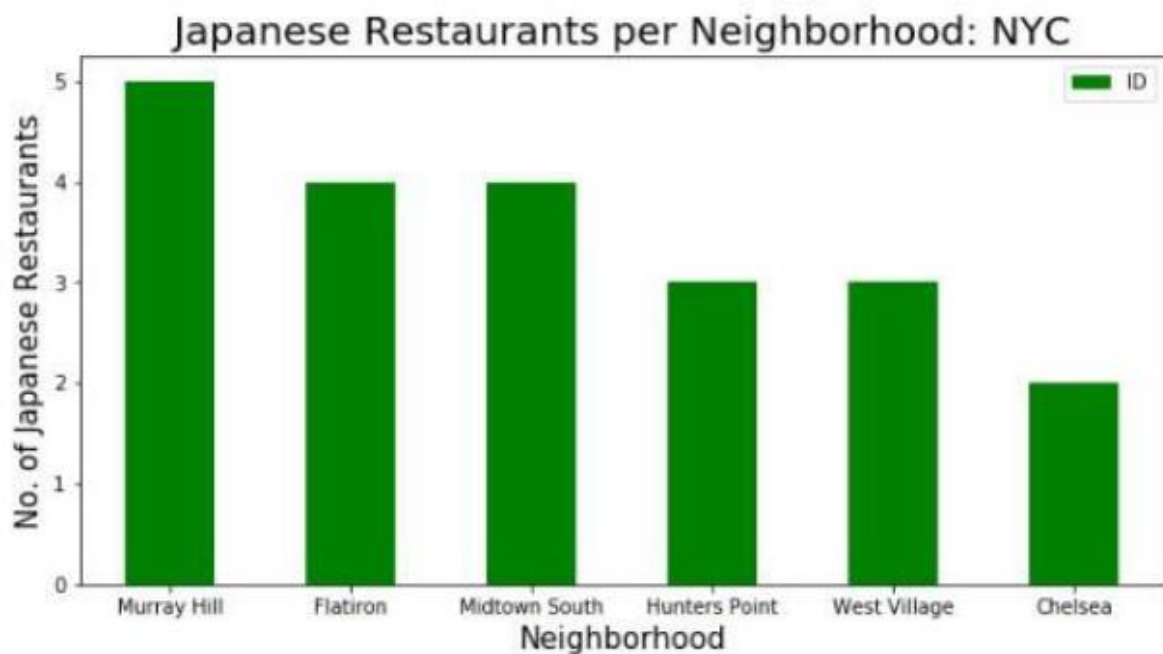
1. We see that Queens has the highest number of Neighborhoods.



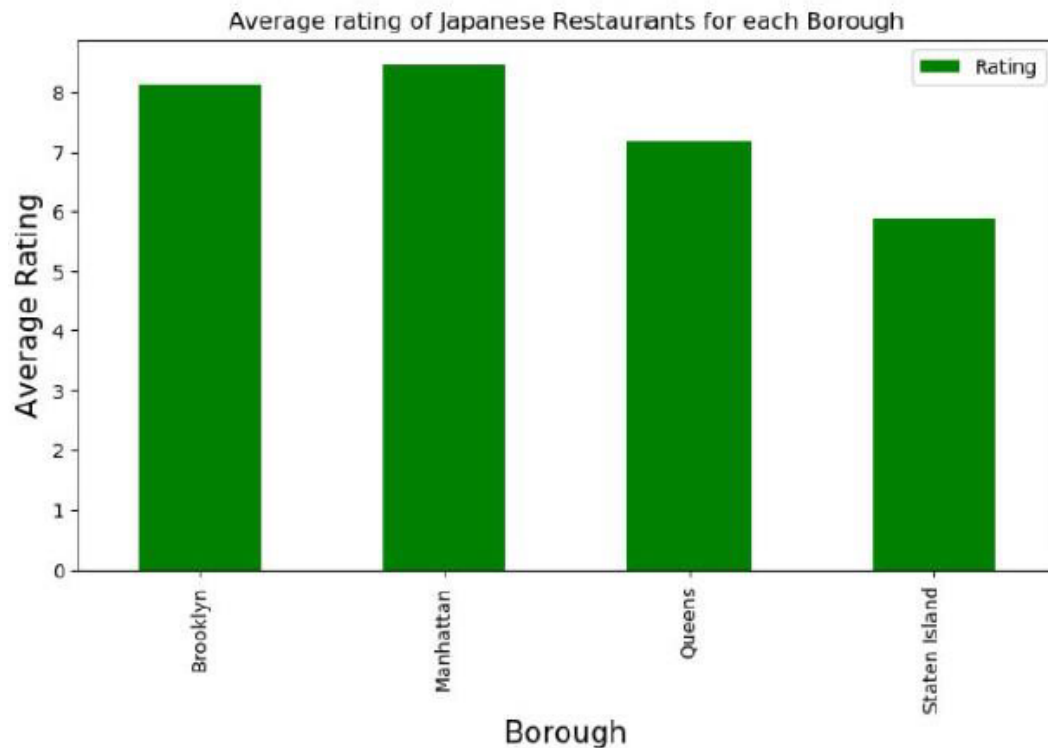
2. Although Manhattan had the least number of neighborhoods, it does have the highest number of Japanese restaurants.



3. The Murray Hill in Manhattan has the highest number of Japanese Restaurants with a total count of 5.



4. Manhattan has the highest average rating for Japanese Restaurants. Brooklyn goes right behind the Manhattan with the average rating of its Japanese restaurants only a little lower than Manhattan. Two neighborhoods with the highest average rating for Japanese Restaurants are Hunters Point and East Village.



5. The top neighborhoods with the highest average ratings of Japanese restaurants.

	Neighborhood	Average Rating
12	East Village	9.250000
20	Hunters Point	8.933333
8	Cobble Hill	8.800000
10	Downtown	8.800000
36	Soho	8.700000
40	Upper East Side	8.600000
31	North Side	8.600000
27	Midtown South	8.433333
24	Lindenwood	8.400000
23	Lenox Hill	8.400000

5. Discussion

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 a Japanese restaurant in Brooklyn. Brooklyn has multiple neighborhoods with average ratings exceeding 8.0 on a scale of 1.0 to 10.0 and has less amount of Japanese restaurants than Manhattan, making competition easier. In addition, we should keep in mind, that real estate prices in Brooklyn are much cheaper than in Manhattan. Particularly, I would recommend considering opening a Japanese Restaurant either in East Village or in Hunters Point, because both of these neighborhoods have the highest rating for Japanese restaurants. Finally, I would go to OOTOYA in Manhattan for the best Japanese food based on 1213 likes.

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.



6. Conclusion

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 investors/developers. 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 investor open a Japanese 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 a Japanese restaurant.