

# **COURSERA - IBM APPLIED DATA SCIENCE CAPSTONE**

## **JAPANESE RESTAURANT CLUSTERING IN NEW YORK CITY**

*SNEHA SRUTHI V*

*August 2020*

### **INTRODUCTION**

Konnichiwa!!

This final project explores New York City for the best locations of Japanese restaurants throughout the city. New York is a major metropolitan area with more than 8.8 million (Macrotrends, 2020) people living within city limits. The earliest Japanese immigrants to settle in NYC came in 1876, mostly merchants who set up businesses to import silk. By 1900, there were approximately 1,000 first generation immigrants in New York State, almost all of them in New York City.

As of 2019, 30,000 Japanese nationals and a total of 45,000 reside in the Greater New York City area. Many of the Japanese are from transient groups such as university students, artists, and business workers. Over 30,000 Japanese nationals live throughout the five boroughs of New York City. However, New York's "Japantown" usually referred to as Little Tokyo, is located on the Lower East Side, between 2nd and 3rd Avenues and 8th (St Mark's Place) and 10th Streets, nothing comparable to New York's famous Chinatown.

This report explores which neighborhood and boroughs of New York City have the most as well as the best Japanese restaurants. Additionally, there will be an attempt to answer the questions “Where should I open a Japanese Restaurant?” and “Where should I stay If I want great Japanese food?” to meet the needs of the working professionals and students.

## DATA

1. In order to answer the above questions, data on New York City neighborhoods, boroughs, including boundaries, latitude, longitude, restaurants, and restaurant ratings and tips are required.
2. Above information will be obtained from New York City data containing the neighborhoods and boroughs, latitudes, and longitudes. The data source used will be: [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)
3. The neighborhood boundaries of New York City will be obtained from the data source: <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>
4. Foursquare API will be utilized via the Request library in Python to procure all data related to locations and user ratings of Japanese restaurants

## PROBLEM STATEMENT

1. What is / are the best location(s) for Japanese cuisine in New York City?  
(*Borough & Neighborhoods with maximum Japanese restaurants*)
2. In what Neighborhood and/or borough should I open a Japanese restaurant to have the best chance of being successful? (*Borough & Neighborhoods with maximum & minimum Japanese restaurants*)
3. Where would I go in New York City to have the best Japanese food?  
(*Best average rating restaurants*)

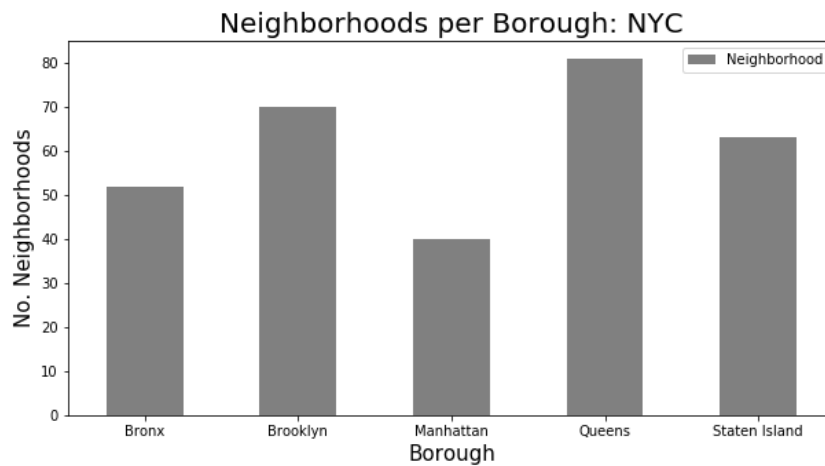
## METHODOLOGY

Data will be collected from [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset) cleaned and processed into a dataframe.

- FourSquare API will be used to locate all venues which are then filtered by category of Japanese restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.
- Data will be sorted based on average ratings
- Finally, the data be will be visually assessed using graphing from Python library like Folium

⇒ From the New York data set, we observe there are 306 neighborhoods in New York City from the dataset

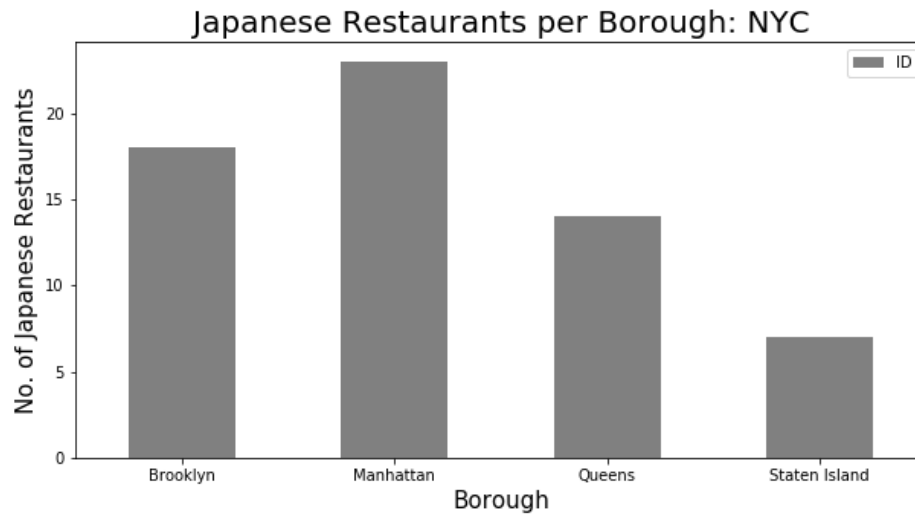
⇒ We clean and process it to obtain top 5 neighborhoods



- Queens has the highest number of neighborhoods and Manhattan the least.

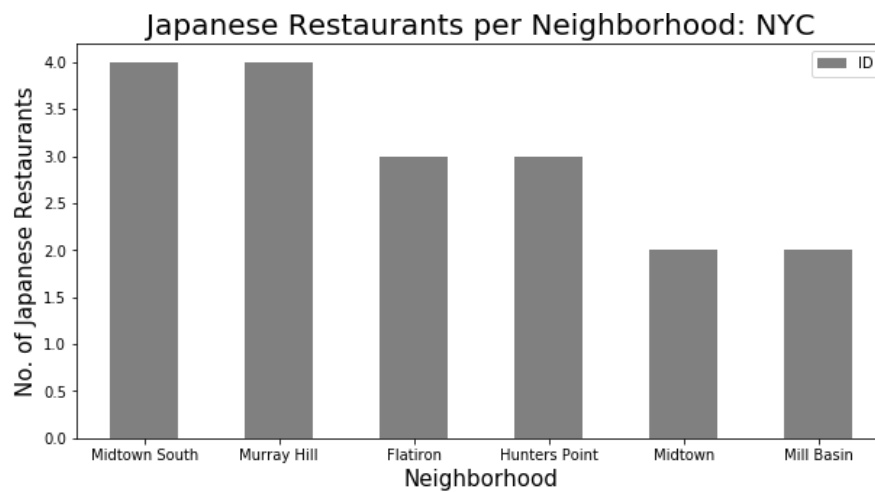
⇒ These Top 5 neighborhoods in New York City are selected and the Foursquare API will be called to analyze total number of Japanese restaurants across these neighborhoods

⇒ Let's visualize by plotting the restaurants based on borough



- We can see that there are a total of 63 Japanese restaurants in NYC
- Manhattan has the most number of restaurants, followed by Queens and Staten Island, the least

⇒ Let's plot which are the top 6 neighborhoods with Japanese restaurants

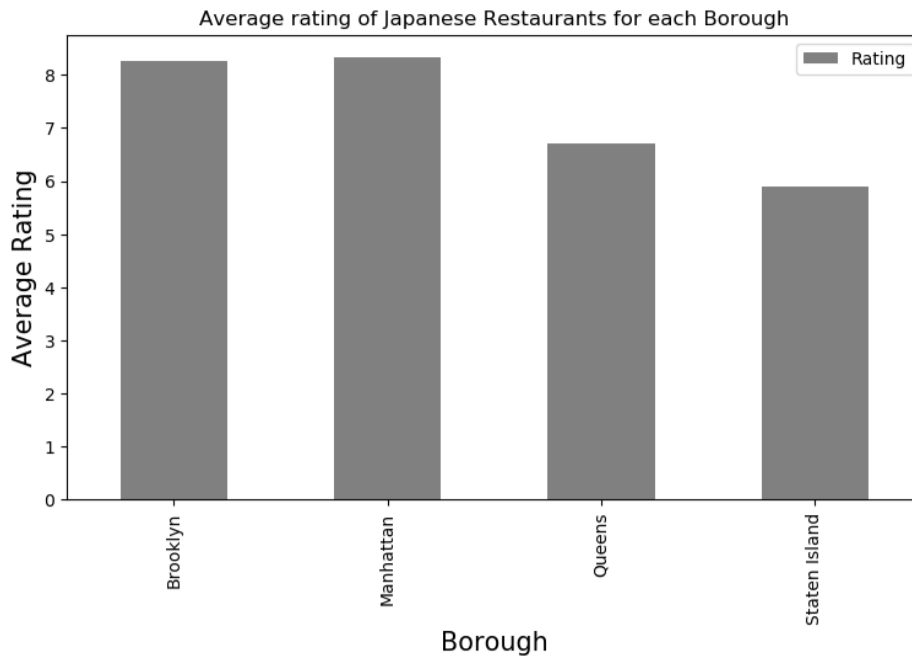


- Murray Hill and Midtown South have 4 restaurants each followed by the rest.

- ⇒ These 63 restaurants are selected and the Foursquare API will be called to analyze the average ratings, likes and tips of each, across these neighborhoods
- ⇒ Due to the limitation on number of premium calls that can be made to Foursquare API, we have restricted the number to first 50 restaurants.
- ⇒ The table below lists restaurants in Manhattan with the ratings, likes and tips out of the details collected for 50 of them. Bottom 5 are tabulated here:

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
45	Manhattan	Midtown South	51e8ce40498ea1fb0180efb2	Izakaya MEW	1055	8.7	266
46	Manhattan	Midtown South	57ab615d498eef084faa348b	Abiko Curry	132	8.3	28
47	Manhattan	Midtown South	5d7a9d59a369d40008dfcc1a	Konbini	16	8.2	2
48	Manhattan	Midtown South	58b8ed467220e661ddf34104	KazuNori: The Original Hand Roll Bar	520	9.1	125
49	Brooklyn	Mill Basin	4d7fc91f95c1a143260eecf2	Yoshi	7	7.4	7

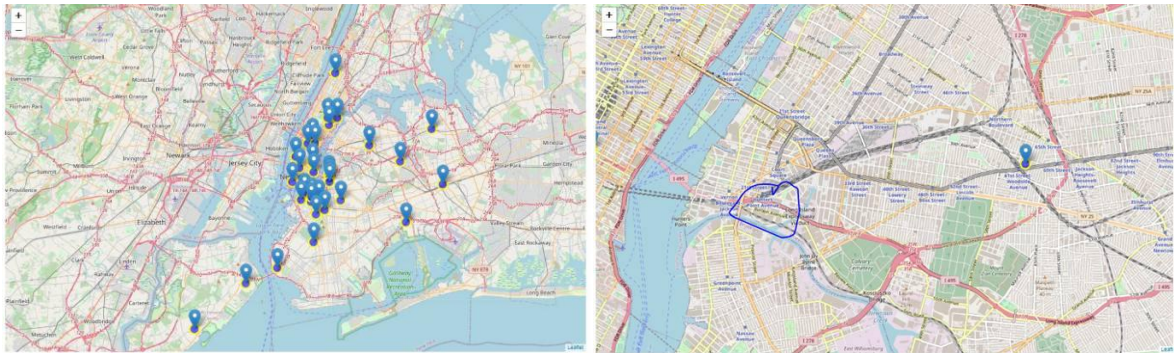
- ⇒ Let's visualize how the Boroughs fare with Average ratings



- Manhattan and Brooklyn house Japanese restaurants with Average rating  $\geq 8$  and are the most popular and visited

## RESULTS & DISCUSSION

1. Murray Hill in Manhattan and Queens is the best location for Japanese Cuisine
2. Based on the analysis, the neighborhoods in Flatiron, Hunters Point, Mill Basin can be considered for a new restaurant to be successful as there are only 1 or 2 restaurants.
3. Based on the average rating and user reviews, restaurants in Brooklyn & Manhattan have the best Japanese cuisines. These could be the best hangout places for students or working professionals to enjoy their home cuisine and satiate their taste buds at a central location.



## CONCLUSION

This project allowed me to work with a huge dataset, process and clean the dataset to suit my needs and problem statement of finding and/or opening a Japanese restaurant in New York City. Using data analysis techniques, I could cluster the different neighborhoods and identify top neighborhoods with highest number as well as most popular Japanese restaurants, by interacting with a third-party API. This enabled me to visualize these restaurants across neighborhoods which serve as a quick guide for students, working professionals or tourists to locate nearby ones. Also, this could be a good opportunity for new businessmen who require less competition, yet, most popular neighborhood to open a new restaurant.