Exploring Restaurants in New York City using Foursquare API

Jon Doyle

August 6, 2020

1. Introduction

1.1 Background

New York City, simply called New York and abbreviated as NYC, 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, New York City is also the most densely populated major city in the United States.

Situated on one of the world's largest natural harbors, New York City is composed of five boroughs, each of which is a county of the State of New York. The five boroughs—Brooklyn, Queens, Manhattan, the Bronx, and Staten Island—were consolidated into a single city in 1898. The city and its metropolitan area constitute the premier gateway for legal immigration to the United States. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. New York is home to more than 3.2 million residents born outside the United States, the largest foreign-born population of any city in the world as of 2016.

The Restaurants industry in New York has experienced steady growth over the five years to 2018, as convenient and affordable food remains popular with consumers. While the low price point of the industry's products typically places restaurants and coffee shops with a competitive advantage over other segments of the foodservices sector, rising consumer sentiment has increased competition overall. Furthermore, with the rise of fast-casual concepts rapidly gaining market share, major industry operators have had to alter their offerings to effectively compete.

Given the history mentioned this project will attempt to answer the following questions in the Problem section.

1.2 Problem

To find the answers to the following questions:

1) What is the best location in New York City for Chinese Food?
2) Which areas have potential for additional Chinese Restaurants?
3) Which is the best place to stay if you prefer Chinese Food?
4) List and visualize all major parts of New York City that has high ranked Chinese restaurants

1.3 Interested audience

The target audience for such a project is twofold. Firstly, any person who is visiting New York City can use the plots and maps from this project with an interested in Chinese food to select areas to travel to. Secondly, an investor can use this information to determine if they want to invest in a Chinese restaurant.

2. Data

2.1 Data Sources

For this project we need the following data:

New York City data that contains list Boroughs, Neighbourhoods along with their latitude and longitude:
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

I retrieved the following for each Borough

- **Borough:** Names of NYC Boroughs
- **Neighbourhood:** Neighbourhoods within the Borough
- **Longitude:** The longitude value of the Borough.
- **Latitude:** The latitude value of the Borough

Chinese restaurants in each neighbourhood of New York City:

Data Source: Foursquare API (https://api.foursquare.com/v2/venues/)

Description: By using this API we will get all the venues in each neighbourhood. We can filter these venues to get only Chinese restaurants.

I retrieved the following for each venue:

- **Name:** The name of the venue.
- **Category:** The category type as defined by the API.
- **Venue ID:** The latitude value of the venue.
- **Like Count**: Number of Likes
- **Rating:** Rating of the Venue
- **Tip Counts:** Count of Tips

GeoSpace data:
Data Source: https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm

Description: By using this geo space data we will get the New York Borough boundaries that will help us visualize choropleth map.

2.2 Data Cleaning

As part of using the Foursquare API the data was retrieved and filtered down to just Chinese Restaurants. Additionally, when retrieving venue details if ratings information was missing, I needed to set the values to 0 with the assumption that it was missing data. Since the Developer account is limited in how much information is retrieved, I did not drop data from the sample.

3. Methodology and Exploratory Data Analysis

1. We begin by collecting the New York city data from "https://cocl.us/new_york_dataset".
2. We will find all venues for each neighborhood using FourSquare API Developer Account.
3. We will then filter out all Chinese Restaurant venues.
4. Next using FourSquare API, we will find the Ratings, Tips, and Like count for all the Indian Restaurants.
5. Next, we will sort the data keeping Ratings as the constraint.
6. Finally, we will visualize the Ranking of neighborhoods using python's Folium library.

3.1 Location information for Borough's

I began by collecting the New York city data from the following link https://cocl.us/new_york_dataset and breaking down the Borough Name, Neighborhood, Latitude and Longitude that will be needed to feed the Foursquare API. Below is a sample output. The analysis produced 306 different Neighborhoods in New York City and that Queens has the most neighborhoods.

	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

Figure 1- New City Borough, Neighbour and Location Information

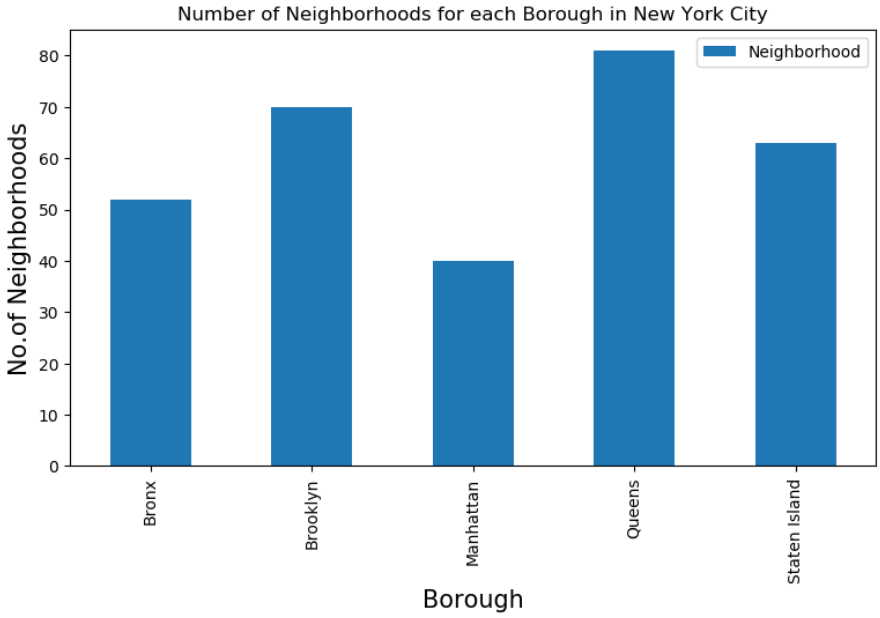


Figure 2- Diagram of Number of Neighbourhoods by Borough

3.2 Venue Information

Next, I will explore the venue information from the Foursquare API. In the figures below after filtering down to Chinese Restaurants. We can see that Queens has the largest population of Chinese restaurants and that Chinatown is the largest concentration for a neighborhood in Manhattan. Additionally, we see that there are 215 Chinese Restaurants across New York City.

.

	Borough	Neighborhood	ID	Name
0	Bronx	Kingsbridge	4da39def540ea1cdfb3b95de	Yeung Hing Chinese Restsurant
1	Bronx	Norwood	4e6aa73eaeb7c31e43294be9	Sing Fei Chinese Restaurant
2	Bronx	Norwood	4ce313fc438b224bbc6c80a3	Happy Dragon
3	Bronx	Norwood	4c3a5b280a71c9b6762844c9	Wok Wok
4	Bronx	Pelham Parkway	4b9d6b45f964a52078ab36e3	Mr. Q's Chinese Restaurant

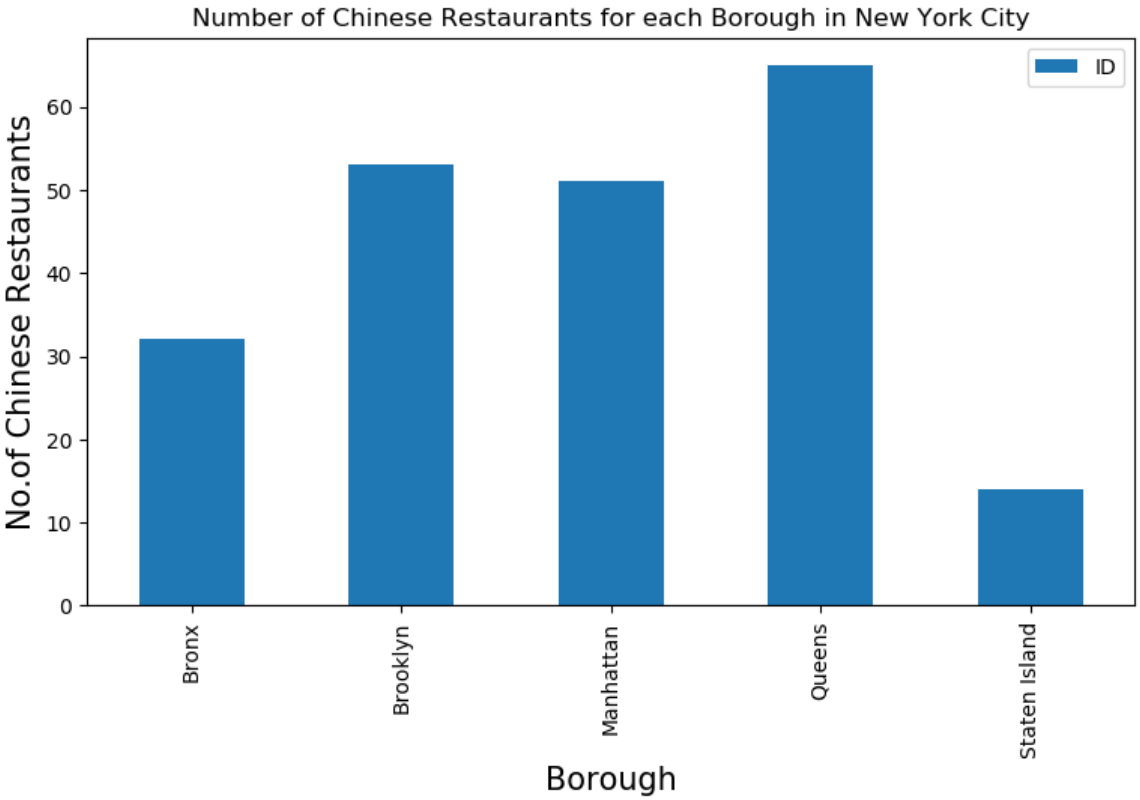


Figure 3 Ranking of Chinese Restaurants by Borough

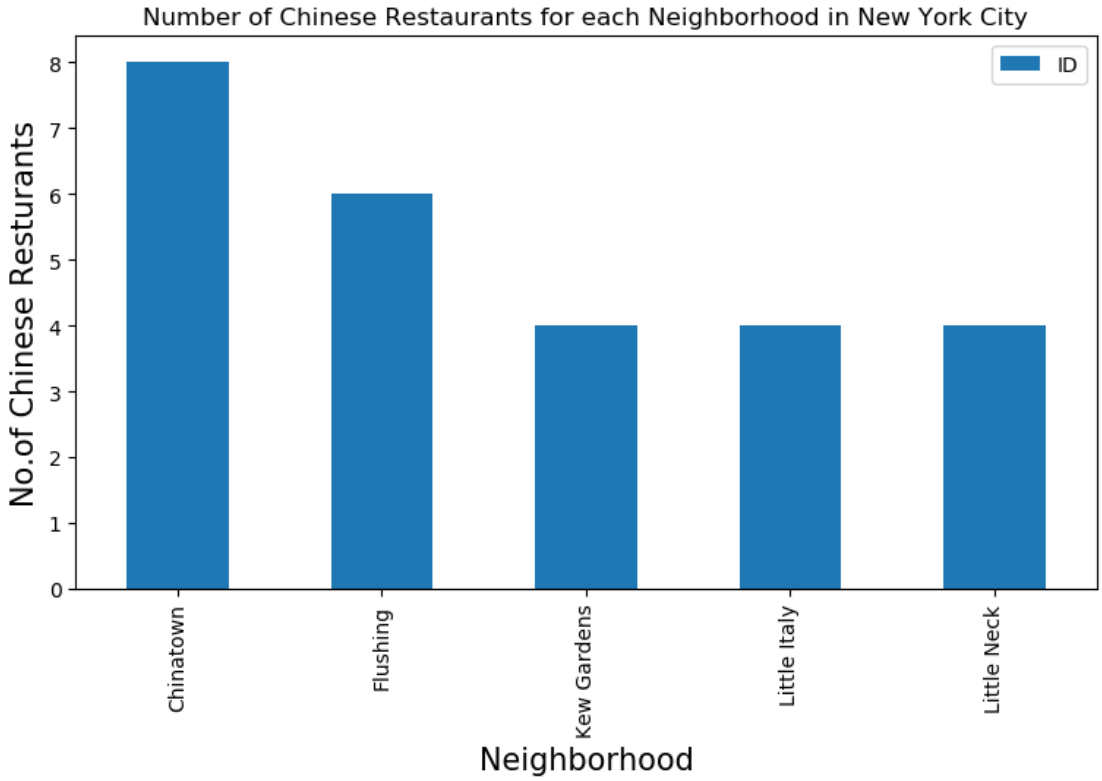


Figure 4- Top 5 Neighbourhood Rankings

3.3 Detailed Venue Information

Below is detailed venue information that we can develop after being pulled out of FourSquare

3.3.1 Restaurant with maximum Likes

Borough	Manhattan
Neighborhood	Chelsea
Name	Buddakan
Likes	1489
Rating	8.8
Tips	521

3.3.2 Restaurant with maximum Rating

Borough	Manhattan
Neighborhood	Lower East Side
Name	Kings County Imperial
Likes	67
Rating	9
Tips	10

3.3.3 Restaurant with maximum Tips

Borough	Manhattan
Neighborhood	Chelsea
Name	Buddakan
Likes	1489
Rating	8.8
Tips	521

3.3.4 Ranking of Boroughs

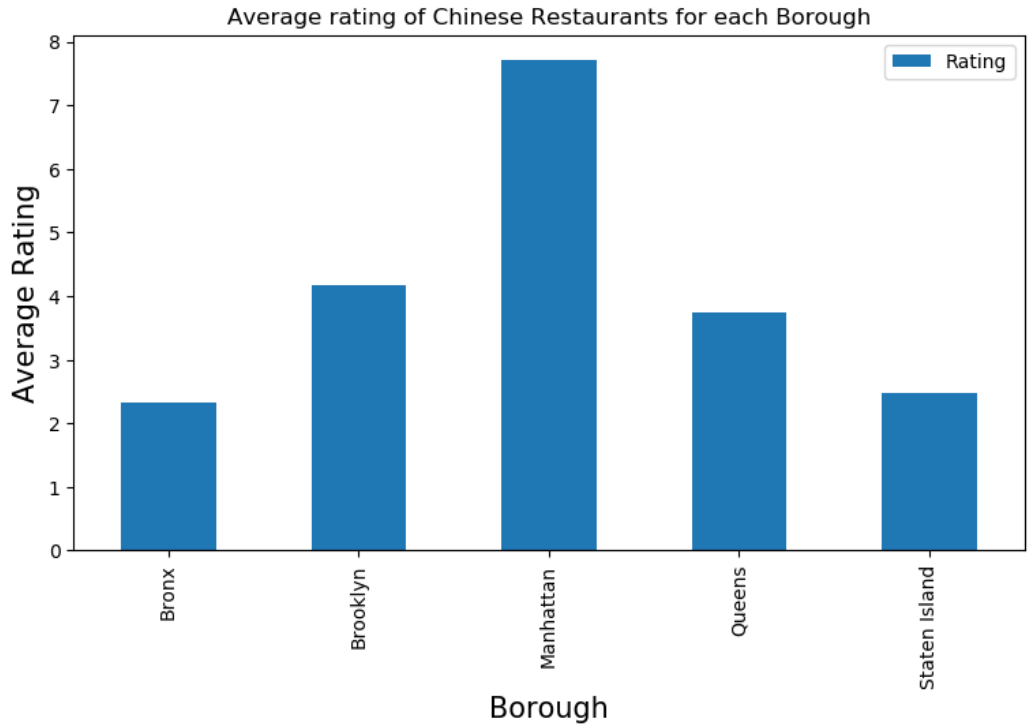


Figure 5 Rating of Each Borough

3.3.5 Neighborhoods with average rating greater or equal 8.0 to visualize on map.

	Neighborhood	Average Rating
12	Civic Center	9.100000
69	Tribeca	9.100000
0	Astoria	9.000000
5	Blissville	9.000000
75	West Village	8.800000
44	Midtown South	8.800000
43	Midtown	8.800000
29	Gramercy	8.733333
25	Fort Greene	8.700000
11	Chelsea	8.700000

Figure 6 Top 10 Restaurants Based on Average Rating

3.4 Visualize Top Rated Neighborhoods on a map.

In order to visual the information on the map we needed to join the data back the initial extract we used which then allows us to map it out using Folium.

3.4.1 Initial Step needed to bring back Location Information as shown below:

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Manhattan	Chelsea	40.744035	-74.003116	8.800000
1	Staten Island	Chelsea	40.594726	-74.189560	8.800000
2	Manhattan	Chinatown	40.715618	-73.994279	8.425000
3	Brooklyn	Downtown	40.690844	-73.983463	8.233333
4	Manhattan	East Village	40.727847	-73.982226	8.500000
5	Queens	Elmhurst	40.744049	-73.881656	8.200000
6	Manhattan	Greenwich Village	40.726933	-73.999914	8.300000
7	Manhattan	Little Italy	40.719324	-73.997305	8.550000
8	Manhattan	Midtown	40.754691	-73.981669	8.200000
9	Manhattan	Midtown South	40.748510	-73.988713	8.700000
10	Manhattan	Murray Hill	40.748303	-73.978332	8.650000
11	Queens	Murray Hill	40.764126	-73.812763	8.650000
12	Brooklyn	North Side	40.714823	-73.958809	8.900000
13	Manhattan	Soho	40.722184	-74.000657	8.600000
14	Brooklyn	South Side	40.710861	-73.958001	8.550000
15	Manhattan	Tribeca	40.721522	-74.010683	8.600000
16	Manhattan	Upper West Side	40.787658	-73.977059	8.700000
17	Manhattan	West Village	40.734434	-74.006180	8.850000
18	Brooklyn	Windsor Terrace	40.656946	-73.980073	8.900000
19	Queens	Woodside	40.746349	-73.901842	8.400000

Figure 7Data for Mapping Rankings

3.4.2 Creating a Map Using Folium to Visualize the Top Neighborhoods

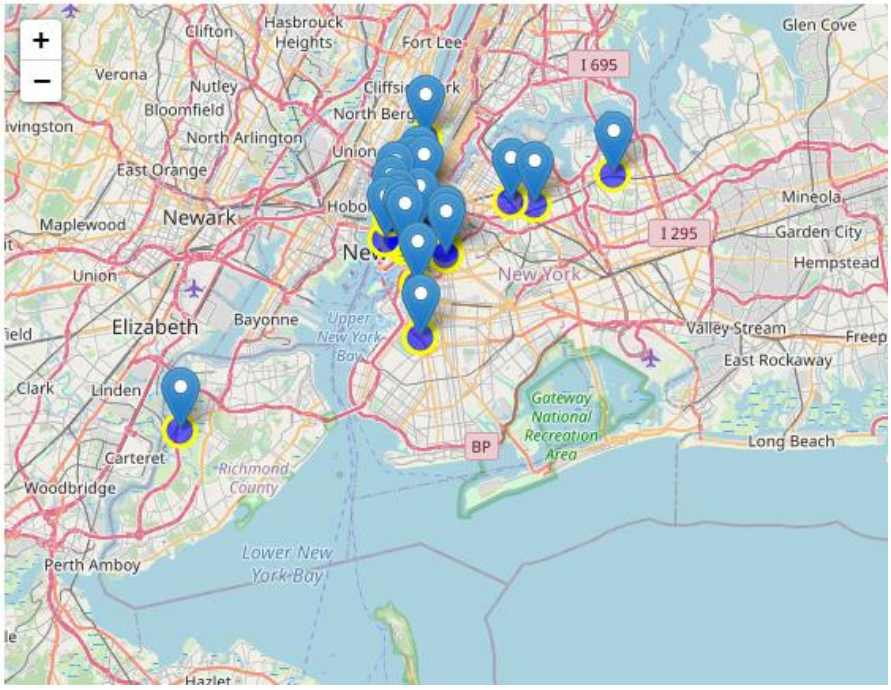


Figure 8 Map of Top-Rated Restaurants in each Borough

4. Results and Discussion

After collecting data from the Foursquare, filtering it for Chinese Restaurants along with retrieving specific venue details we can answer the questions we initially started with.

1) What is best location in New York City for Chinese Food?

In the borough of Manhattan, the following locations are tops based on:

Ratings: Kings County Imperial in the Lower East Side
Tips: Buddakan in Chelsea
Likes: Buddakan in Chelsea

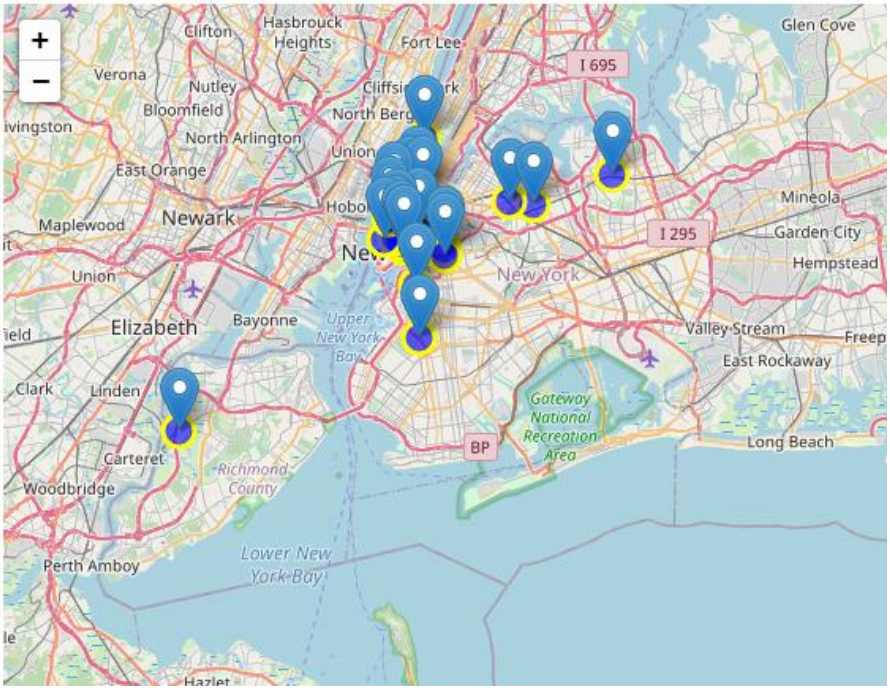
2) Which areas have potential for additional Chinese Restaurants?

Staten Island has the fewest Chinese Restaurants however, the Bronx has the lowest rated Chinese Restaurants so depending on the type of Chinese Restaurant could further research these to areas to which one is a better fit.

3) Which is the best place to stay if you prefer Chinese Food?

Manhattan is the best place to stay to find the best rated Chinese Food.

4) List and visualize all major parts of New York City that has high ranked Chinese restaurants.



Any person who is visiting New York City can use the plots and maps from this project with an interested in Chinese food to select areas to travel to. Secondly, an investor can use this information to determine if they want to invest in a Chinese restaurant.

5. Conclusion

The purpose of this project was to answer questions specific to Chinese restaurants for either a visitor, resident or investor. The venues where identified using Foursquare API. Although Queens has the largest number of restaurants the overall quality of restaurants in Manhattan came out on top for where to go eat. While Staten Island could be looked at for an investment opportunity due to number of restaurants a high-quality venue could provide more value in Bronx due it its lower ratings.

It is worth noting that using the Foursquare Developer API (personal account) does have its limits and results using a paid account could lead to deeper analysis. Additionally, joining data in from a service like Zomato for pricing could also expand the how the data can be looked at for a subsequent project.