

Capstone: Battle of the Neighborhoods - Chinese Restaurants

Iain Drew

1. Introduction

A small group of restaurateurs are looking to open a Chinese restaurant somewhere in New York City. New York City comprises 5 boroughs with Manhattan as its epicenter, a densely populated borough and one of the world's major commercial, financial and cultural centers. There are numerous restaurants of similar genre, and this group of investors would like to understand what areas/neighborhoods would be advantageous to opening and sustaining a successful business. This report explores which neighborhoods and boroughs of New York City have the most and best Chinese restaurants and I will attempt to answer the question "Where should the restaurateurs open a Chinese restaurant?"

1.1 Problem

- What is the best area for a small group of investors to open a Chinese restaurant in New York City?
- What boroughs and neighborhoods have the most and highest ranked Chinese restaurants?

2. Data

Data for this project will be structured into data frames with the geographic data as the starting point and adding Foursquare metadata (Names, ID, rankings etc.) to additional data frames later in the workflow. These features will be used to evaluate the counts of restaurants per borough and neighborhood as well as identify the highest ranked restaurants and their locations. Throughout the workflow visualizations will be used to better understand the data as well.

2.1 Data sources

- New York City data containing neighborhoods, boroughs, latitudes and longitudes is obtained from: https://cocl.us/new_york_dataset
- All data related to locations and quality of Chinese restaurants will be obtained from the Foursquare API.

- New York City boundary data is from: <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>

2.2 Data Preparation

Initial geographic data on NYC will be gathered and structured into a data frame with features for Borough, Neighborhood, Latitude, and Longitude.

	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: Data frame of NYC geographic data

From this df we can get some preliminary information about how many neighborhoods per borough exist. Here we see Queens has the greatest number of neighborhoods and Manhattan has the least.

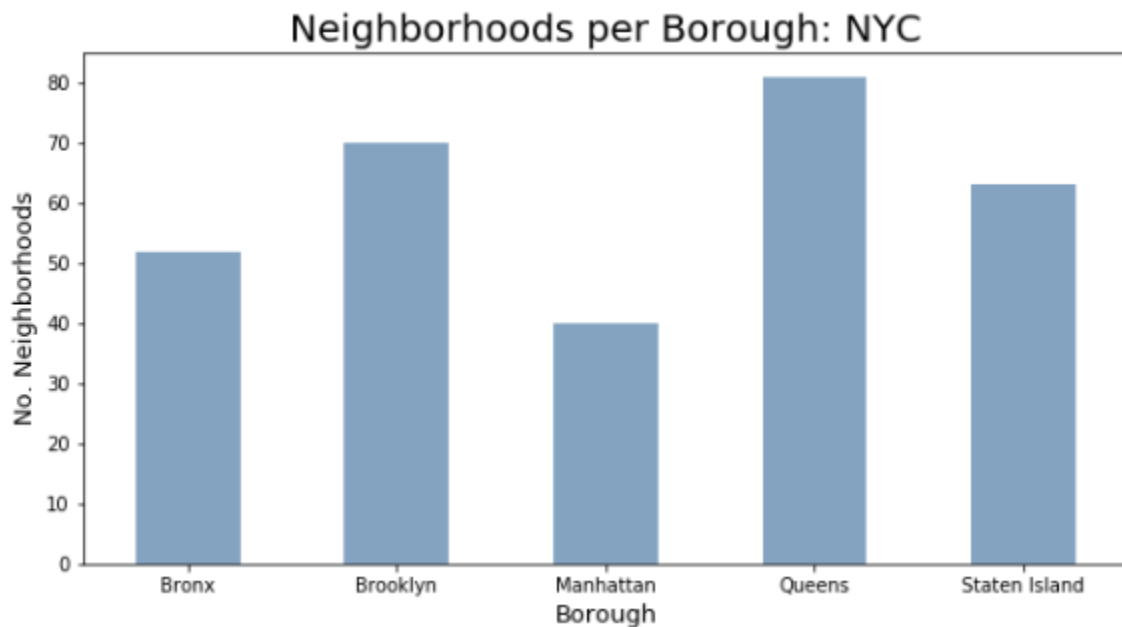


Figure 2: Bar plot of Neighborhoods per Borough in NYC

Next, a data frame was created to contain Borough, Neighborhood, ID and Name (of Chinese restaurants) from Foursquare API to compare how many Chinese restaurants exist in each borough and neighborhood.

	Borough	Neighborhood	ID	Name
257	Staten Island	Prince's Bay	4bb7d282b35776b0b83dc801	Island Taste
258	Brooklyn	Madison	4be32bd7b02ec9b61fc34ec0	win hing
259	Bronx	Bronxdale	4c3f7f2eda3dc928b8f6c5b9	Peking Kitchen
260	Bronx	Allerton	4bc11181abf495219f7dc093	Li's Kitchen
261	Queens	Hammels	4c72ad2bad69b60cb31b84b9	301 Chinese Resturant

Figure 3: Data frame of Chinese restaurant Names, ID with geographic location data.

Methodology

- Data will be collected from https://cocl.us/new_york_dataset, then cleaned and transformed into a dataframe.
- Foursquare data will be collected to locate all venues then filtered for Chinese restaurants. Metadata for likes and ratings by users will be counted and put into the dataframe.
- Data will get sorted by ranking.
- Data will be visually assessed via Python libraries.

Results:

There is a total of 269 Chinese restaurants spread throughout the 5 boroughs of NYC. The borough of Queens had the greatest number of Chinese restaurants at 83.

	ID	Name
Borough		
Bronx	46	46
Brooklyn	56	56
Manhattan	60	60
Queens	83	83
Staten Island	24	24

Figure 4: table of Chinese restaurants per borough

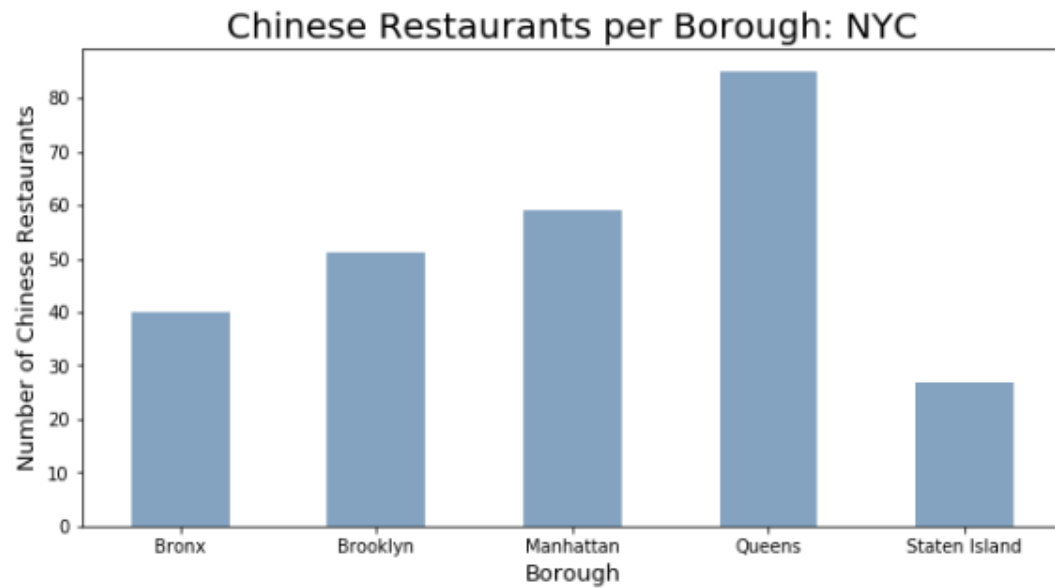


Figure 5: Bar plot of Chinese restaurants per Borough NYC

The neighborhood with the greatest number of Chinese restaurants was Chinatown in the borough of Manhattan with about 10 restaurants.

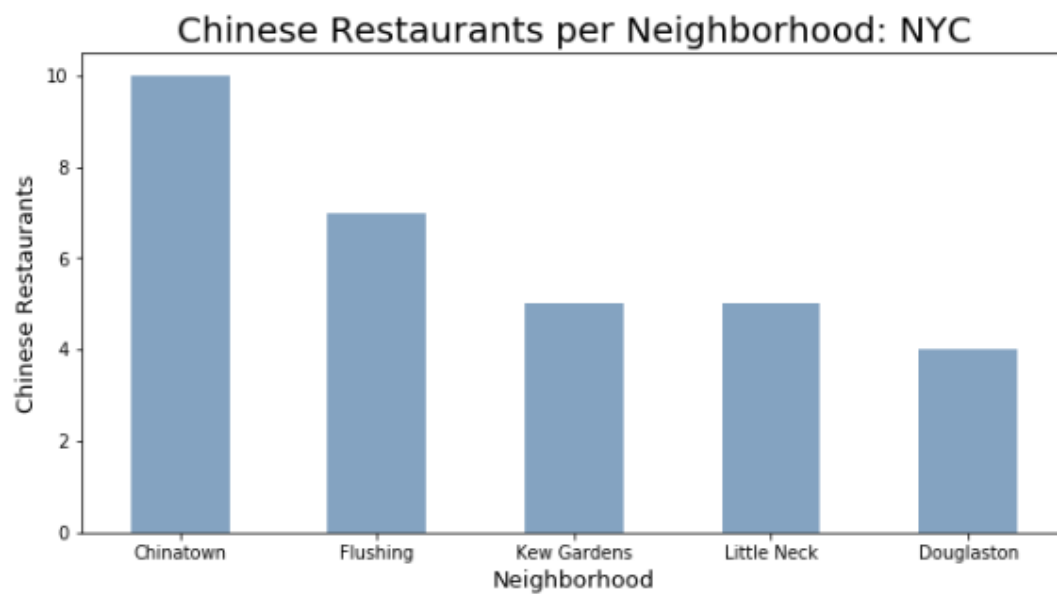


Figure 6: Chinese restaurants per Neighborhood NYC

The dataframe that was created from another Foursquare API call included the “Likes” and ratings for the Chinese restaurants. After cleaning up the dataframe to remove any restaurants that did not have any ratings, 162 restaurants remained for further analysis.

```

:

```

	Borough	Neighborhood	ID	Name	Likes	Rating
0	Bronx	Eastchester	0		0	0.0
1	Bronx	Kingsbridge	0		0	0.0
2	Bronx	Woodlawn	0		0	0.0
3	Bronx	Norwood	0		0	0.0
4	Bronx	Pelham Parkway	4b9d6b45f964a52078ab36e3	Mr. Q's Chinese Restaurant	9	7.4

```

:
#drop rows with zero ratings
indexratingzero=chi_rest_stats_nyc[chi_rest_stats_nyc['Rating']==0].index
chi_rest_stats_nyc.drop(indexratingzero, inplace=True)
chi_rest_stats_nyc.tail()

```

```

:

```

	Borough	Neighborhood	ID	Name	Likes	Rating
263	Brooklyn	Homecrest	4b46743ff964a520892126e3	Golden "Z" Restaurant	9	7.6
264	Brooklyn	Homecrest	4d28b924ebacb1f72efbf34f	Wing Hing Seafood Restaurant	25	6.7
266	Bronx	Bronxdale	4c3f7f2eda3dc928b8f6c5b9	Peking Kitchen	9	7.4
267	Bronx	Allerton	4bc11181abf495219f7dc093	Li's Kitchen	2	7.1
268	Queens	Hammels	4c72ad2bad69b60cb31b84b9	301 Chinese Resturant	9	6.3

```

:
chi_rest_stats_nyc.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 162 entries, 4 to 268
Data columns (total 6 columns):
Borough      162 non-null object
Neighborhood 162 non-null object
ID           162 non-null object
Name         162 non-null object
Likes        162 non-null int64
Rating       162 non-null float64
dtypes: float64(1), int64(1), object(4)

```

Figure 7: Dataframe with Likes and Ratings cleaned of restaurants with zero rating.

The restaurant with the greatest number of likes and highest rating was **Buddakan** in the neighborhood of Chelsea, Manhattan with 1473 Likes and a rating of 9.2.

```

#check restaurant with max likes and rating
chi_rest_stats_nyc.loc[chi_rest_stats_nyc['Rating'].idxmax()]

```

```

Borough      Manhattan
Neighborhood  Chelsea
ID           44e9b421f964a520a5371fe3
Name         Buddakan
Likes        1473
Rating       9.2
Name: 117, dtype: object

```

Figure 8: Restaurant with greatest number of likes and highest rating

Next, I looked at the average ratings of Chinese restaurants for each borough. Here we see that on average the best restaurants are in Manhattan.

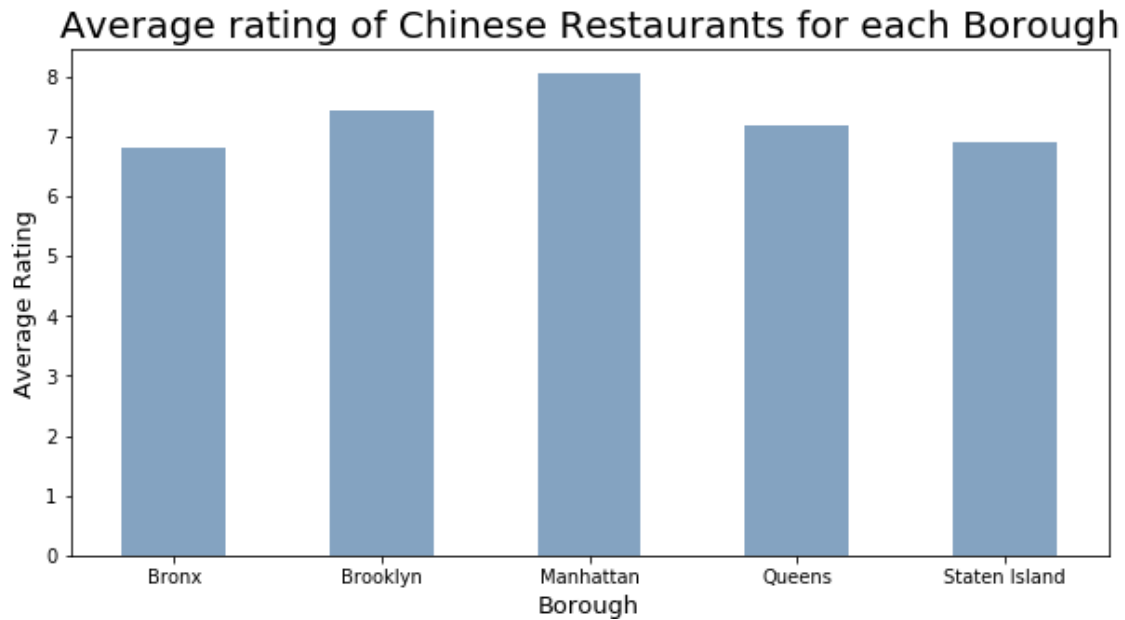


Figure 9: Bar plot of average rating of Chinese restaurants for each borough.

Finally, maps were created to visually inspect the average ratings across the boroughs as well as map the neighborhoods with an average rating greater than or equal to 6.

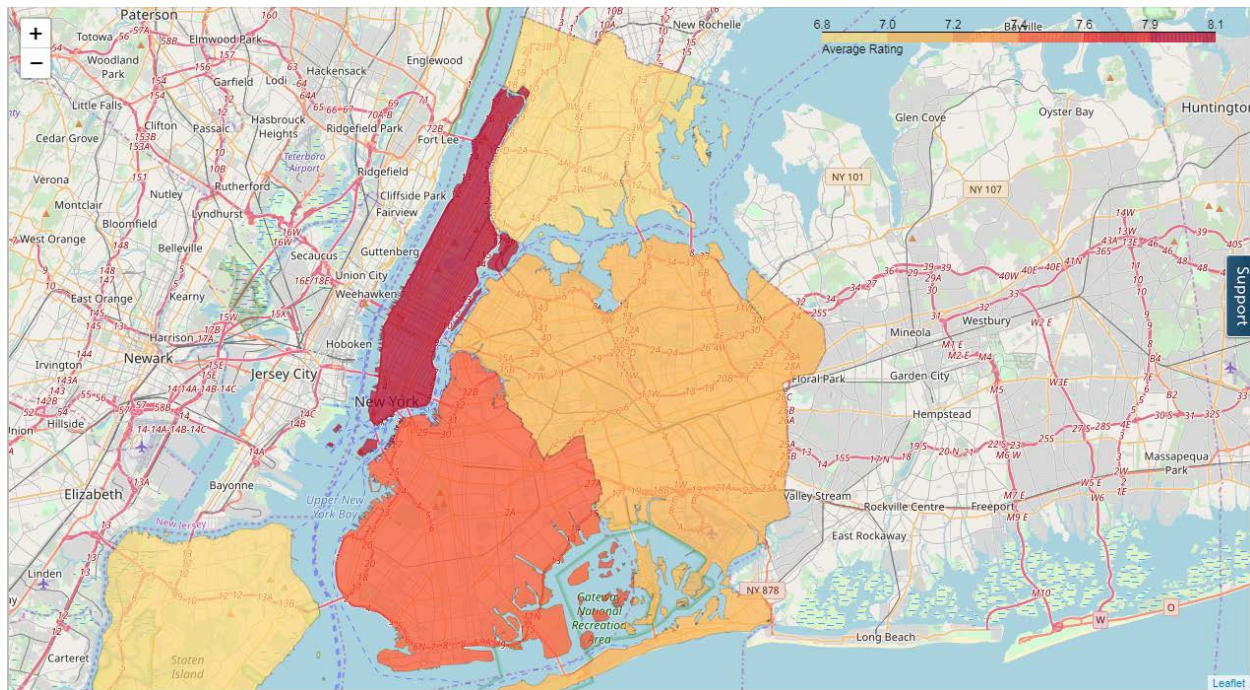


Figure 10: Choropleth map of NYC with average ratings of Chinese restaurants per borough.



Figure 11: Screen shot of marker cluster map for average rating greater than or equal to 6 for neighborhoods.

Discussion and Conclusion:

Manhattan and Queens have the greatest number of Chinese restaurants per borough with the Bronx and Staten Island having the least. Interestingly Manhattan and Brooklyn have the highest *average* ratings for Chinese restaurants, but the most liked and highest rated Chinese restaurant is in Little Neck, Queens. From this information, I would look at Brooklyn for opening a new Chinese restaurant. Brooklyn has an average rating of 7.4 and although this report didn't include real estate analysis it's common knowledge that Manhattan (average rating 8.06) has some of the most expensive real estate in the world and thus would likely be more expensive to lease or own property zoned for a restaurant. F