



CAPSTONE PROJECT

The Battle of Neighborhoods



Pedro Wong

2020/2/24



Table of Content

Instruction2

Description.....3

 Background3

 Problems3

 Data.....3

Calculation Result.....5

Conclusion8

Instruction

For this project, you will be required to submit the following:

- A description of the problem and a discussion of the background.
- A description of the data and how it will be used to solve the problem.
- A link to your Notebook on your Github repository, showing your code.
- A full report consisting of all of the following components:
 - Introduction where you discuss the business problem and who would be interested in this project.
 - Data where you describe the data that will be used to solve the problem and the source of the data.
 - Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.
 - Results section where you discuss the results.
 - Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.
 - Conclusion section where you conclude the report.
- A presentation or blogpost.

Description

Background

There are large number of people especially for widespread consumers, travelers and local small business investors are desire to find the most appropriate restaurants for themselves in big city. However, the miscellaneous info on the web contain different noise and interfere the people's choice. These kinds of info not only based on the gourmet's review and the numbers of order, but also contain the malicious comments from other competitors. Consequently, in this project, we will identify the several different Chinese restaurants in New York based on location.

Problems

1. What is the top ranked Chinese restaurants in terms of neighborhoods and boroughs?
2. How can we identify the location for small business investors to start and maintain their restaurant business in the city?

Data

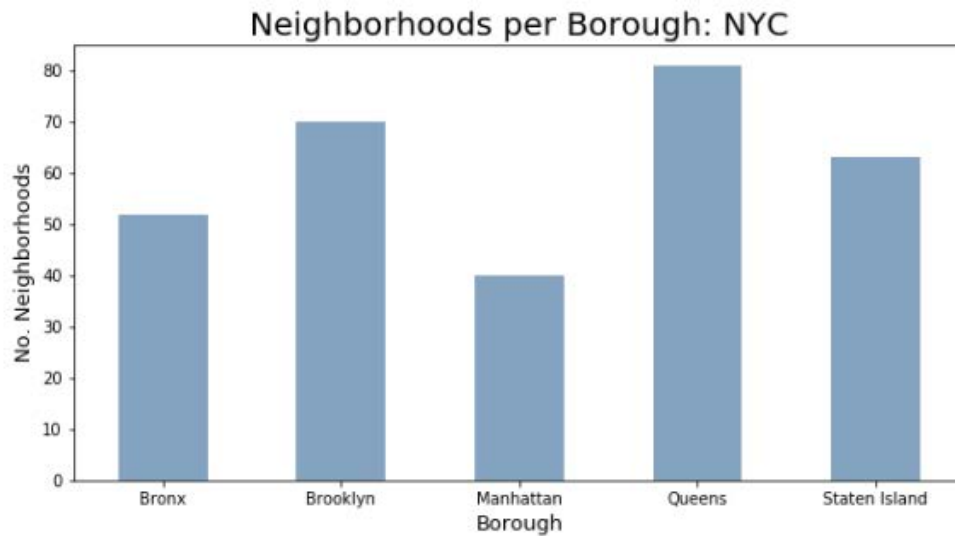
In this project, we will use the Foursquare data including ID, name and ranking etc. to data frames as the additional one in the later workflow, also we will address the geographic data to the data frames. These efficient elements will be taken into account when we count and analyze the Chinese restaurants, also play a significant role for the later process such as recognize the top ranked restaurants and location info.

Moreover, data visualization in the later process will bring a comprehensive and understandable way to observe the information. The geographic data comes from the official website of cityofnewyork.us and cocl.us. Other ratable sources come from the Foursquare.

The following picture demonstrate the initial data frame related with geographic information:

| | 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 |

The following bar chart illustrated the preliminary information involved with how many neighborhoods per borough. Apparently, there are great number of neighborhoods in Queen and the least number of neighborhoods in Manhattan.



The next step is to generate the data frame which contains Names, ID and other information. This information are both comes from Foursquare. Through this graph, we can easily compare the total count of Chinese restaurants in different borough.

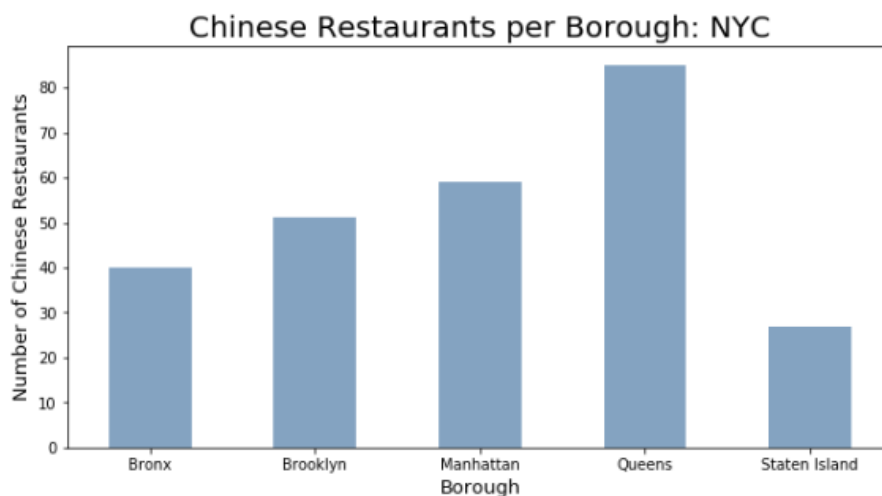
| | Borough | Neighborhood | ID | Name |
|-----|---------------|--------------|--------------------------|-----------------------|
| 257 | Staten Island | Prince's Bay | 4bb7d282b35776b0b83dc801 | Island Taste |
| 258 | Brooklyn | Madison | 4be32bd7b02ec9b61fc34ec0 | win hing |
| 259 | Bronx | Bronxdale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

Calculation Result

From the table below we can get the information that it has 269 Chinese restaurants in total throughout the 5 different area in New York, and most of them are distributed in Queens borough.

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

As you can see from the below bar chart, Chinatown has the largest number of Chinese restaurants per neighborhood (around 10 restaurants).



The following picture shows the another Foursquare, it reveals the overall rating of the restaurants and the tendency of people's choice. In the data cleaning process, we clean the unnecessary data and NAN data (162 remain) in terms of the deep 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)

```

The top popular restaurant which contains the most people's choice and top rating in the Chelsea neighborhood is Buddakan, as you can see from the following picture.

```

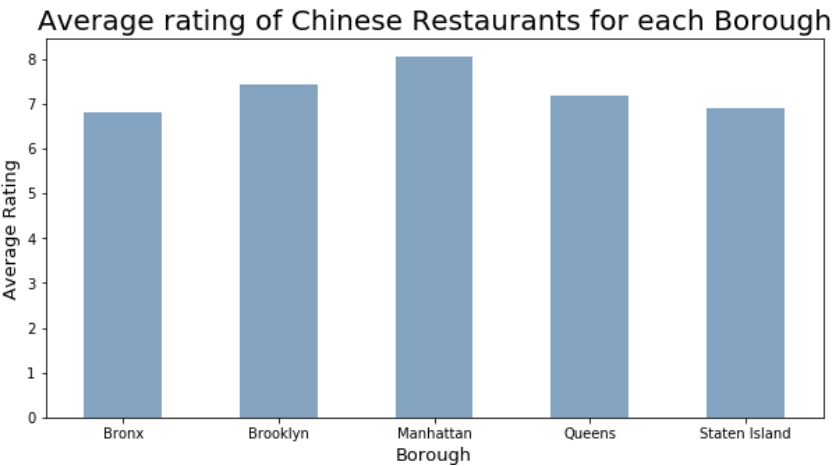
#check restaurant with max Likes and rating
chi_rest_stats_nyc.loc[chi_rest_stats_nyc['Rating'].idxmax()]

```

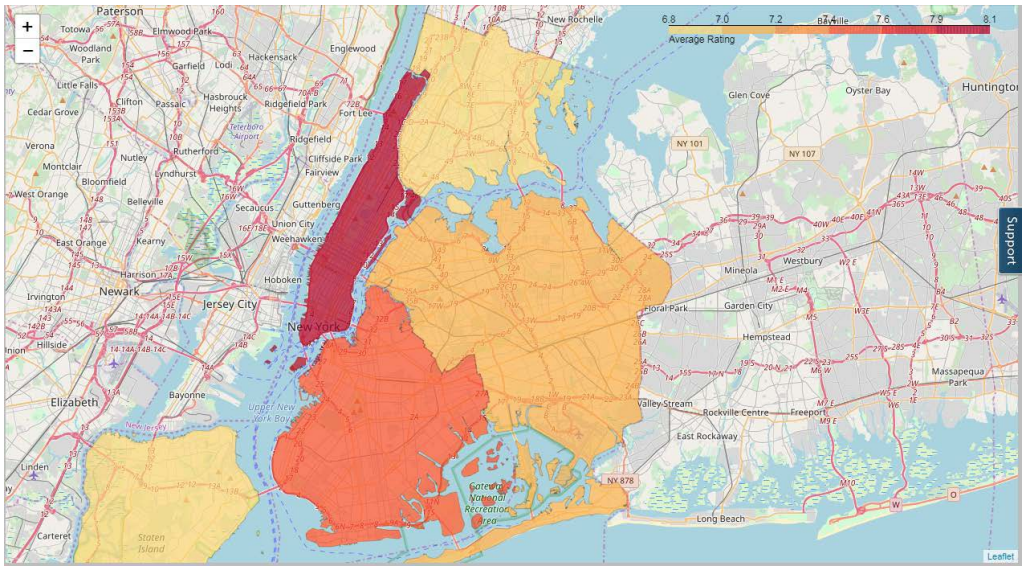
| | |
|--------------|--------------------------|
| Borough | Manhattan |
| Neighborhood | Chelsea |
| ID | 44e9b421f964a520a5371fe3 |
| Name | Buddakan |
| Likes | 1473 |
| Rating | 9.2 |

Name: 117, dtype: object

Furthermore, there is another information displays the average rating in terms of different area. From the bar chart, we can easily get the conclusion that Manhattan has the largest points which means the restaurants in there are the most popular for the people. On the opposite side, the Bronx has the least popular Chinese restaurants compare with other borough.



For the last step, the data visualization will bring the data, tables and charts out together as the untreatable map. The following maps (cluster and choropleth) show the aberage rating of Chinese restaurants.



Conclusion

Overall, the boroughs which contain the largest number of Chinese restaurants are Queens and Manhattan, while the Bronx has the least number of restaurants. For the average rating, the top-ranked area is Brooklyn. However, the best rated restaurants are more likely distributed in Neck. The recommendation and most appropriate suggestion for the small restaurant business investor is that starting a new restaurant in and Brooklyn will be a good choice. For the fresh off the boat and travelers. It is better to consider about Neck. Queen when they want to have a delicious experience and explore the Chinese cuisine culture.