

IBM Applied Data Science Capstone

The Battle of Neighborhoods

Problem:

If someone is looking to open a Chinese restaurant in New York City, which neighbors or borough should he open it?

Background:

New York City is one of the most popular and largest cities in the United States. Many restaurants and buildings are being constructed for business. The audience is looking to find a location that is best suit for Chinese restaurant. My job is to conduct a business analysis of New York City neighborhood and boroughs to find the best location.

Data:

The data that will used for this problem is New York City neighborhood and boroughs. The dataset should also contain with latitude, longitude, and venue. The URL I used for data collection is from Wikipedia and here is the link: https://cocl.us/new_york_dataset. I included four columns in the dataset, which are borough, neighborhood, latitude, and longitude. Here is the dataset:

	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

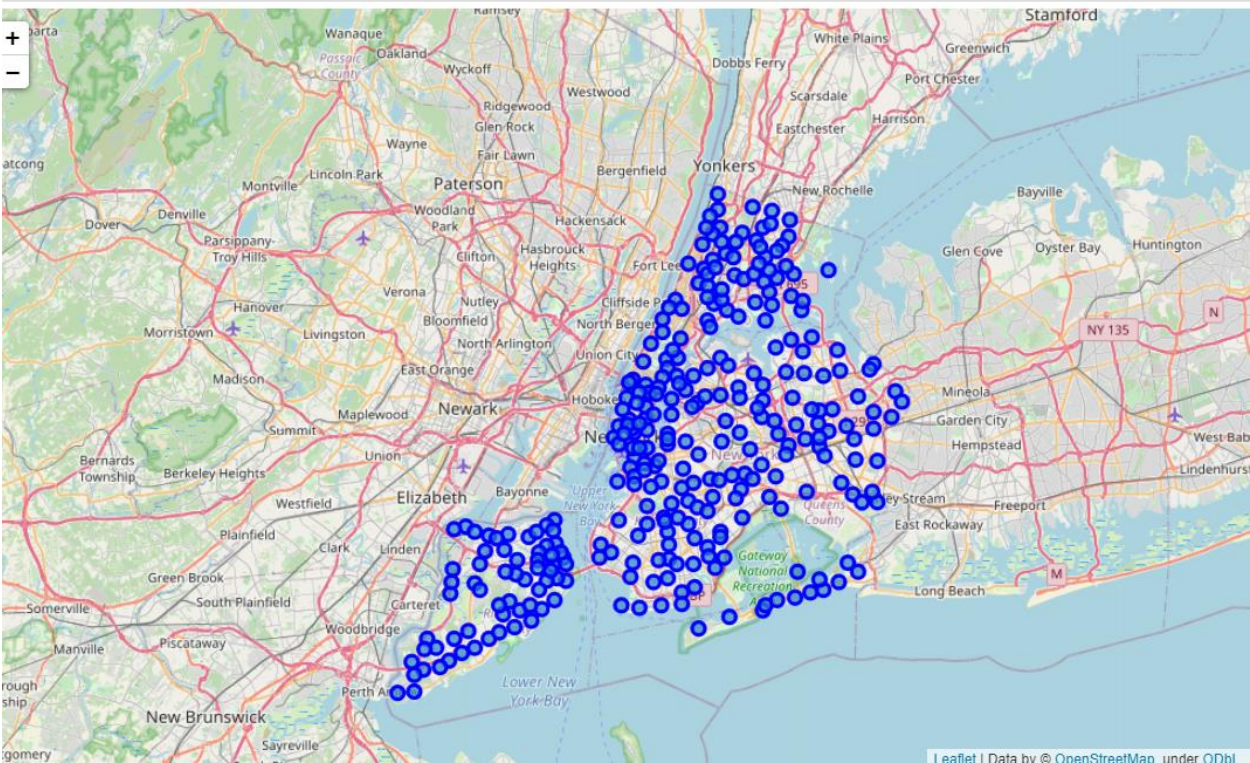
Methodology:

- Data cleaning/ data preprocessing
- Map visualization
- Four Square API

Discussion:

After data cleaning, there are total of 306 neighborhoods and 5 boroughs in the dataset. The geographical coordinate of NYC is 40.7127281, -74.0060152. Here is a map of New York

contains neighborhoods and boroughs.



In order to do Foursquare API, here is my account information:

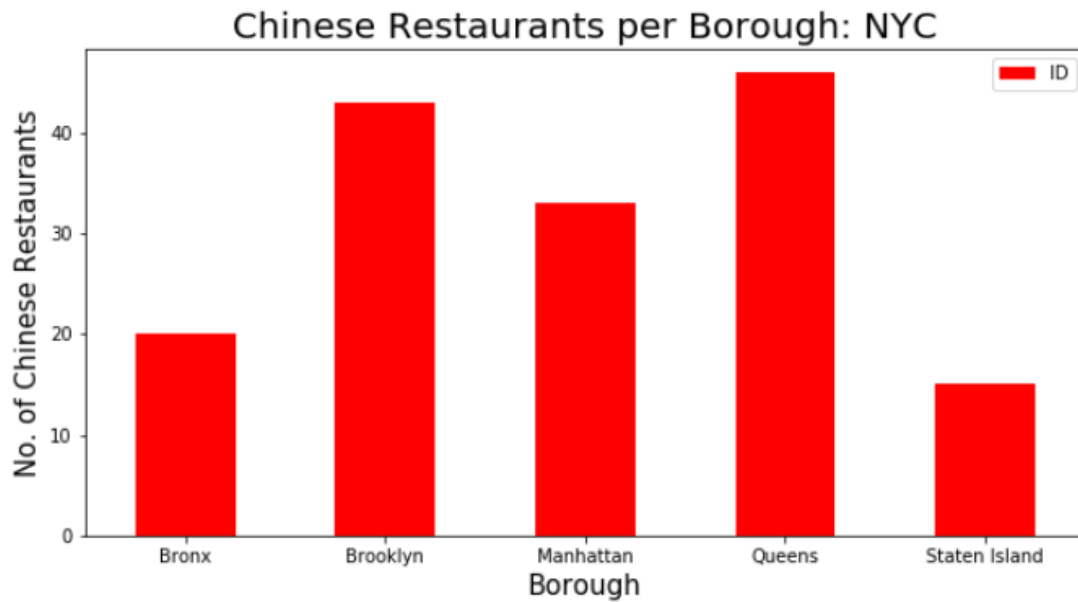
Your credentials:
CLIENT_ID: Q010I4YFZUOR3JFMPFZUWESFPAMLD4SKEQ1QQQH5SL5MDE
CLIENT_SECRET: A5L3JKC2JYCEWC4KJNUL1TFBOOT0AD4YZESDOONNEN1ZY2DG

In order to find all the Chinese restaurants, I need to find additional information from Foursquare. I added two additional columns which is name and ID. Here is the new table after adding two more columns.

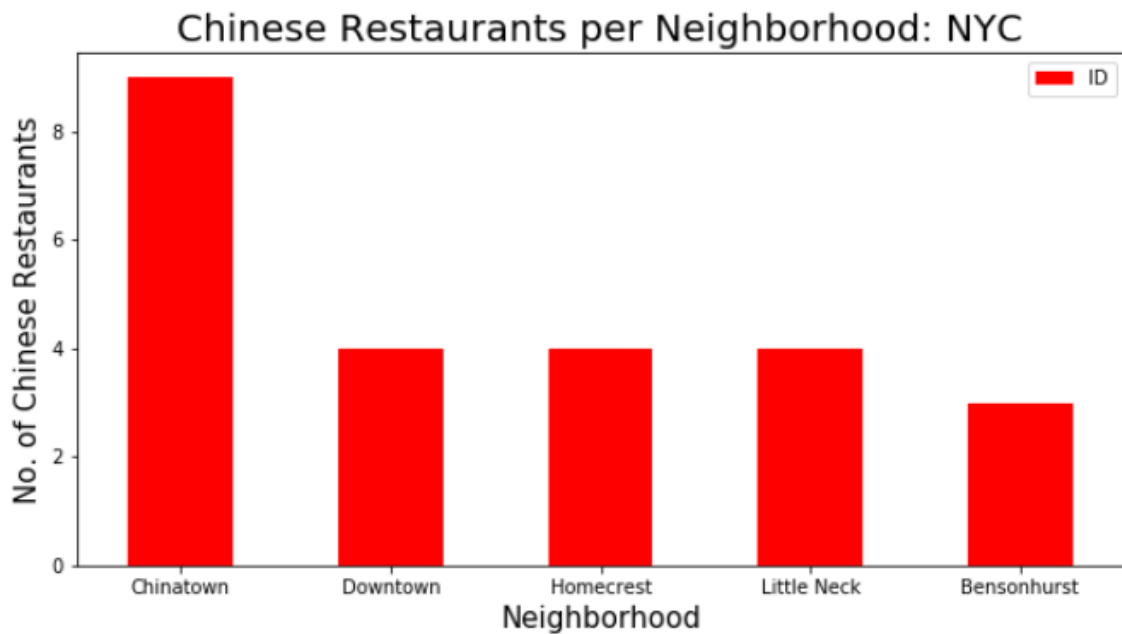
[48]:

	Unnamed: 0	Borough	Neighborhood	ID	Name
152	152	Brooklyn	Homecrest	4d064e228620224b0490b240	Kowloon Island
153	153	Staten Island	Prince's Bay	4bb7d282b35776b0b83dc801	Island Taste
154	154	Bronx	Allerton	4c7dbf7f10916dcba78e2a96	Chef King
155	155	Bronx	Kingsbridge Heights	4d965fd061a3a1cde4989a42	Lucky Star Chinese Restaurant
156	156	Brooklyn	Erasmus	4e10c87122713f7d7bcb9d1d	Kam Man Chinese Food

There are 157 restaurants total in New York.



Queen and Brooklyn are the top two boroughs in New York City that contains the most chinese restaurant.



Chinatown has the highest number of chinese restaurant in NYC.

Here are the names of chinese restaurants:

	Borough	Neighborhood	ID	Name
48	Manhattan	Chinatown	4db3374590a0843f295fb69b	Spicy Village
49	Manhattan	Chinatown	5894c9a15e56b417cf79e553	Xi'an Famous Foods
50	Manhattan	Chinatown	4a96bf8ff964a520ce2620e3	Wah Fung Number 1 Fast Food 華豐快餐店
51	Manhattan	Chinatown	5c965dad5455b2002c058659	Yi Ji Shi Mo Noodle Corp
52	Manhattan	Chinatown	4aff2b4f964a520583a22e3	Sun Hing Lung Ho Fun Tofu
53	Manhattan	Chinatown	59d828f0916bc1155fde2c04	Hwa Yuan
54	Manhattan	Chinatown	4c7ab2ac2d3ba143bb1d92d0	Sun's Organic Tea Shop
55	Manhattan	Chinatown	3fd66200f964a520ceea1ee3	Deluxe Green Bo Restaurant
56	Manhattan	Chinatown	3fd66200f964a520ede41ee3	Big Wong King 大旺

Conclusion:

In conclusion, in order to find the best location to start a chinese restaurant, Chinatown is a good option since it is the most popular neighborhood that contains the largest number of chinese restaurant. Boroughs such as Queens and brooklyn can also be good options to start a chinese restaurant. Even though Manhattan is also popular, the land price will be much more expensive to other neighborhoods.