

# SETTING UP THE LOCATION OF A NEW CHINESE RESTAURANT

Jose Romero Porrino

February 2021

IBM Data Science Certificate Coursera Capstone Project

## 1. INTRODUCTION

### 1.1 Background

New York, capital of globalization and multiculturalism. This city has always been seen as the epicentre of the world, and not without a reason: New York is said to be the only place in the world where you can find a representation of all the cultures in the world. Whether it is true or not, it is undeniable its attraction to millions of emigrants and tourists, creating a unique environment where the only rule that seems to apply is “*welcoming everybody at all moments*”.

One of the motors of this environment is the food, a pillar for many cultures. Japanese, Korean, French, Indian, Chinese, Spanish, Mexican, Italian... these cuisines have gained millions of supporters through the years, as well as a few detractors. So is the popularity they are gaining, that hundreds of restaurants are starting to pop in the city, a part of a boom never seen before. Is there anything more New Yorker than an Italian guy attending a Peruvian restaurant with his South African date?

Restaurants are, without a doubt, the essence of New York.

### 1.2 Problem

However, New York is quite a big city, and even though you can use public transportation to visit any place in the city, it is true that people just stay in their boroughs in their day-to-day life. So whenever a person wants to open a thematic restaurant, the question arises: “*where should I start my business?*” Well. It will depend on the type of service you are going to provide. In this paper, my friend Wang Wei and his wife Li Xiu Ying want to open a new restaurant in the city “*The House of China*”. However, they do not know how to proceed in the analysis, and are

overwhelmed by the classic questions: “*what borough to choose?*”, or “*what exact neighbourhood is best?*”.

The main goal of this Capstone project is to try to provide the most accurate answer as possible.

### **1.3 Interest**

This project may interest any person trying to open a restaurant or any business in a city taking into consideration population and density of the city, as well as its structure, and combine it with the ethnicity and location of potential competitors in order to have an insight of how to proceed.

## **2. DATA**

The data used for this project is related mainly to the demographics of the city. We will use data related to the structure of New York City: its 5 boroughs, as well as the neighbourhoods they are divided into, taking into consideration size, population and density. Also, demographics related to ethnicity and race distribution are taken into consideration.

All this information is subtracted from the Wikipedia page:

[https://en.wikipedia.org/wiki/New\\_York\\_City](https://en.wikipedia.org/wiki/New_York_City)

[https://en.wikipedia.org/wiki/Demographics\\_of\\_New\\_York\\_City](https://en.wikipedia.org/wiki/Demographics_of_New_York_City)

However, that information is not enough. A dataset of the city is needed in the form of a geojson archive ([https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)), and finally the combination of Foursquare API in order to gain access to location of Chinese restaurants as well as their names.

## **3. METHODOLOGY AND RESULTS**

In order to proceed with the project, it was divided in two main parts. The first part was related to the demographics and data about population of New York City, while the second one was related to Chinese restaurants and identifying where they are located in order to select a place to grow the business.

### **First part: Demographics of New York**

In this part, information about the structure of New York was the main objective. New York is formed by 5 main boroughs. Some of them are notorious enough to be known worldwide, like Manhattan or Bronx. Others remain unknown to the general public, such as Staten Island.

A table with further information was created:

	borough	county	population	gdp_billions	land_sqm	land_sqkm	persons_sqm	persons_sqkm
0	The Bronx	Bronx	1418207.0	42.695	42.10	109.04	33867.0	13006.0
1	Brooklyn	Kings	2559903.0	91.559	70.82	183.42	36147.0	13957.0
2	Manhattan	New York	1628706.0	600.244	22.83	59.13	71341.0	27544.0
3	Queens	Queens	2253858.0	93.310	108.53	281.09	20767.0	8018.0
4	Staten Island	Richmond	476143.0	14.514	58.37	151.18	8157.0	3150.0

Figure 1: Demographics of the 5 NY Boroughs

As a result, it can be seen that Brooklyn is the most populous borough, but when it comes to the density, Manhattan is the borough this the highest one. In both cases, Staten Island remains as the least populous and least dense borough of them all. This affects the result in order to select whether the business will be located in an overcrowded place in order to attract as many clients as possible, or in a quiet place so the restaurant is able to supply with the highest possible quality the lower demand.

After that, ethnicity and race were taken into account. Even though this is a globalized and multicultural world and Chinese food attracts people from different cultural, social and economic backgrounds, it is true that Chinese food is consumed the most by Chinese people. Wikipedia will let us see distribution by races, and we will take Asian ethnicity as a reference even though it would be pretty inaccurate.

	Jurisdiction	% White	% Black or African American	% Asian	% Other	% Mixed Race	% Hispanic latino of other race
0	Brooklyn	41.2	36.4	7.5	10.6	4.3	19.8
1	Queens	44.1	20.0	17.6	12.3	6.1	25.0
2	Manhattan	54.4	17.4	9.4	14.7	4.1	27.2
3	Bronx	29.9	35.6	3.0	25.7	5.8	48.4
4	Staten Island	77.6	9.7	5.7	4.3	2.7	12.1
5	NYC Total	44.7	26.6	9.8	14.0	4.9	27.0

Figure 2: Races and ethnicities by NY boroughs

As a conclusion, we can see that Asian population has a higher Asian population (in % of total population) in Queens. They represent 17.6% of total population. Still, once again it is important to remind that Asian race includes many other origins and cultures than just the Chinese one. On the contrary, we can see that The Bronx and Staten Island are not home for a high percentage of Asian population.

The next thing to do, as it can be seen in the Notebook, was to establish and figure out how many neighborhoods exist in New York, as well as their location and the borough they are located in. Using a dataset and geojson map obtained from the Internet (see Part 2: Data), obtaining the 306 neighbourhoods the city of New York is divided into.

```
In [6]: new_york_data=get_new_york_data()

In [7]: new_york_data.shape

Out[7]: (306, 4)
```

Figure 3: Total number of neighborhoods in NY

Some of those neighbourhoods are more known and notorious than others: Chinatown, Upper East Side, Tribeca, Nolita... Out of curiosity, since this data will not affect the final result, a horizontal bar chart was created to know what borough had the highest number of neighbourhoods: Queens.

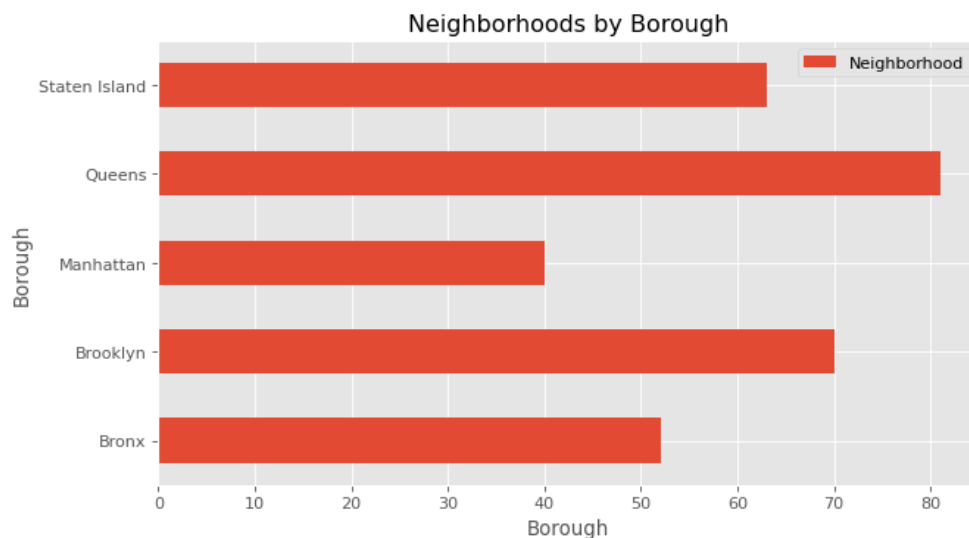


Figure 4: Number of neighborhoods by borough

## Second Part: Restaurants

Now it is time to use Foursquare APIs. They will tell us about how many restaurants are located in the city, as well as their neighbourhood and the name of the restaurants. This is needed in order to obtain a clearer vision of the map of the competition.

```
In [11]: chinese_rest_ny.head()
```

```
Out[11]:
```

	Borough	Neighborhood	ID	Name
0	Bronx	Wakefield	4e2710b7a809510950dd6b32	Double Dragon
1	Bronx	Co-op City	4c66dcfaebea593955a74d0	Chinese Buffet
2	Bronx	Eastchester	4dabc3dc93a04642f09ccabd	Xing Lung Chinese Restaurant
3	Manhattan	Marble Hill	4d965fd061a3a1cde4989a42	Lucky Star Chinese Restaurant
4	Manhattan	Marble Hill	4bcbae5d68f976b01da46183	Ming Moon Buffet

```
In [12]: chinese_rest_ny.shape #just to confirm the number of Chinese restaurants
```

```
Out[12]: (424, 4)
```

Figure 5: example of the findings – Name of the restaurant, ID, borough and neighborhood.

We can see that Foursquare database has 424 Chinese restaurants registered, as of February 9<sup>th</sup> 2021. This number may be larger, due to maybe some restaurants not being registered or detected by the Foursquare APIs.

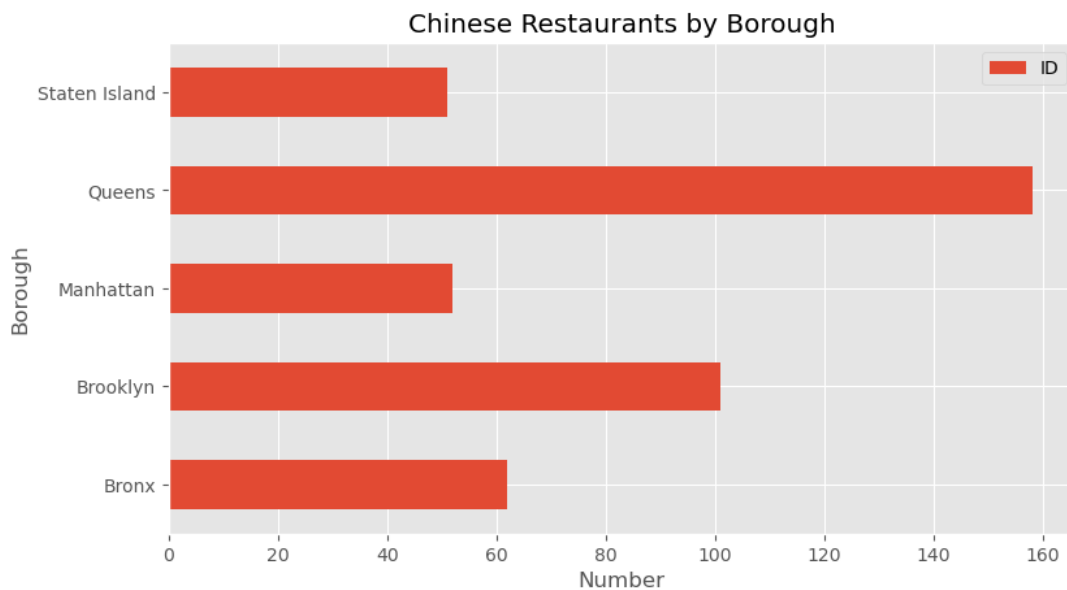


Figure 6: Bar chart of number of Chinese restaurants by borough

Once again, Queens is the absolute leader. When it comes to the total amount of NY Chinese restaurants, 158 of them (almost 38% of the total) are located in the Queens borough.

But what about the neighborhoods?

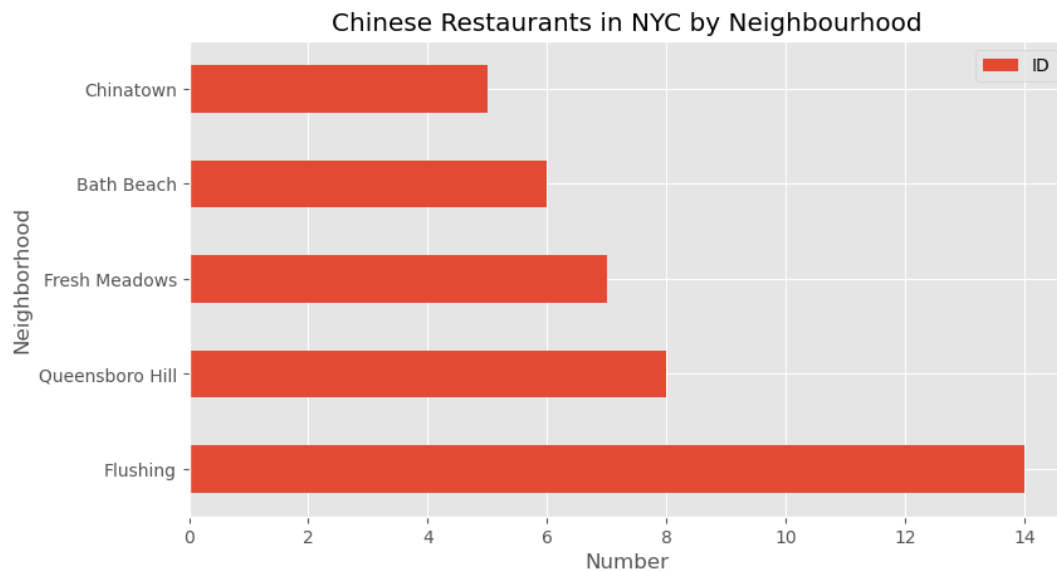


Figure 7: Bar chart of number of Chinese restaurants by neighborhood

```
In [20]: chinese_rest_ny.groupby('Neighborhood')['ID'].count().nlargest(10)

Out[20]: Neighborhood
Flushing          14
Queensboro Hill   8
Fresh Meadows     7
Bath Beach        6
Chinatown         5
Little Neck       5
Sunset Park       5
Utopia            5
Weeksville        5
Bellaire          4
Name: ID, dtype: int64
```

Figure 8: Top 10 Neighborhoods with the highest number of Chinese restaurants

From the Top 10 results, **Flushing neighborhood** is home to the highest number of Chinese restaurants. The top 3 actually is composed by Queens's neighborhoods.

## 4. RESULTS DISCUSSION

As a result of the project carried out, many observations can be inferred:

- Flushing, Queensboro Hills and Fresh Meadows neighborhoods (all located in Queens) are some of the best neighborhoods for Chinese cuisine.
- When it comes to boroughs, Queens is the absolute leader of Chinese restaurants.
- At the bottom, we can check that Manhattan and Staten Island are not very appealing for the Chinese cuisine.
- As for curiosity, the famous and notorious Chinatown Manhattan is not the neighborhood with the highest number of Chinese restaurants.

The final recommendation depends on what strategy the restaurant wants to follow.

1. If the restaurant wants to simply attract the highest number of clients and competition is not a problem for them: then they should start developing the business in Queens, as it is the second most populous borough, has the highest percentage of Asian population, and the highest number of Chinese restaurants. Flushing, Queensboro Hill or Fresh Meadows are great options.
2. However, if the restaurant wants to gain a competitive strategy by being the first-mover or not facing a high rivalry level, Manhattan could be considered as a more than acceptable option: it is the densest borough, the second with the highest Asian population (%), but at the same time one of the boroughs with the lowest number of Chinese restaurants. The market and demand is there, and competition is low.
3. We would not recommend the restaurant to open in Staten Island, as it is the least populated and least dense borough, has a very low percentage of Asian population, and is the borough with the lowest number of Chinese restaurants for a reason.

## 5. CONCLUSIONS

This project was carried out with the main intention of establishing a protocol about how to proceed in order to open a Chinese (or any other type of cuisine) restaurant in New York. New York is known by its multiculturalism and Chinese population is a very present one in the city. Chinatown can be an example of that, while Queens is known as the most diverse place in the world.

The aim of this project was to give recommendations about where to grow the business, but at the end it depends on the strategy that the owners want to follow. What is crucial is to gain information about two of the most important attributes: population and location. Population will show where a higher demand may be faced, depending not only on absolute population, but also relative (density) as well as dispersion of ethnicities around the city since we are talking about “cuisine of origin”.

As for future directions, this project may need a more accurate list of Chinese restaurants, since Foursquare may not have registered all of them, or does not identify a determined restaurant as Chinese.