Suggestion on address of Chinese restaurant in Manhattan

Xiaolin Zhu

April 27, 2019

1. Introduction

1.1 Background

Manhattan is one of the five boroughs in New York City. It is Central Business District of New York City's. It is a good idea to start business in Manhattan, especially for catering. Because Manhattan is most densely populated borough of New York, it has so many appetites to feed.

As a multicultural city, Manhattan has various restaurants, such as Italy food, Suchi of Japan, Thailand food and Chinese food. Nowadays, Chinese food is getting more attention. If you have a plan to open a Chinese restaurant in Manhattan, Which community is a good choice.

There are many factors affect a restaurant’s popularity, such as population density of a venue, rivals, environment, pricing, flavor and so on. In this project, we don’t have much information, but we can still use the dataset to help to chose a good place for a new Chinese restaurant. For example, we could cluster manhattan’s communities to find a district similar to Chinatown.

1.2 Problem

The data we are going to explore in Manhattan, would help us to decide where is a good choice to open a Chinese restaurant. The dataset includes how many shops and what kind of shops on a venue, how many and what kind of restaurants on a venue. These factors give an impression of what biggest feature of a venue is. This project aims to find a good place to open a Chinese restaurant in Manhattan city.

1.3 Interest

Someone who want to open a Chinese restaurant in Manhattan and choose a address for a good business.

2. Data acquisition and cleaning

2.1 Data source

New York city has a total of 5 boroughs and 306 neighborhoods, we can get these data and the latitude, longitude coordinates of each neighborhood from the this linkage: https://geo.nyu.edu/catalog/nyu_2451_34572 . We could seperate Manhattan's data from it.

Then, we are going to utilize the Foursquare API to explore the neighborhoods and gain data of venues. we are going to get the top 100 venues of neighbors within a radius of 500 meters. Latitude and longitude coordinates of each venue will be added to the dataset, as well as categories of shops.

2.2 Data Processing

After we get categories of shops of communities in Manhattan, we are going to extract data has categories is dinner and restaurant. Then data of Chinese restaurant will be further extracted.

In this part, data cleaning is not necessary, but the original dataset don’t have any numeric data. All the numeric data would be used in this project is calculated, such as the number of total shops in each venue, the mount of restaurant and Chinese restaurant of each community and percentage of restaurant in each community.

During the process of analysis, we will get some data to help us gain shop categories of venues. For example, we need Latitude and longitude of Neighborhood (community) to obtain venues.

There will not much predictive work to do, because the dataset is not big enough to do this job. It seems no causal relationship or dependent relationship between the amount

of Chinese restaurant and the total amount of restaurant or communities.

Finally, Selected dataset would be segmented , clustered by k-mean clustering. Based on clusters and what we analyzed before, Recommendation will be given where is suitable for starting Chinese Restaurant business.

3. Exploratory Data Analysis

3.1 Calculate how many shops there are in a community

Manhattan has 40 communities (which are marked as neighborhoods in jupter notebook). With help of Foursquare API, we obtained 333 unique categories and 3320 shops. These data was gained on the condition of a radius of 500 meters of each neighborhood and maximum number returned was set to 100.

3.2 The mount of all restaurants in Manhattan

We selected data of restaurants in Manhattan of different neighborhoods. As shown below, fig1 shows 40 neighborhoods has different amount of restaurants. Some neighborhoods have similar numbers, such as Greenwich Village and Chinatown has 44 restaurants.

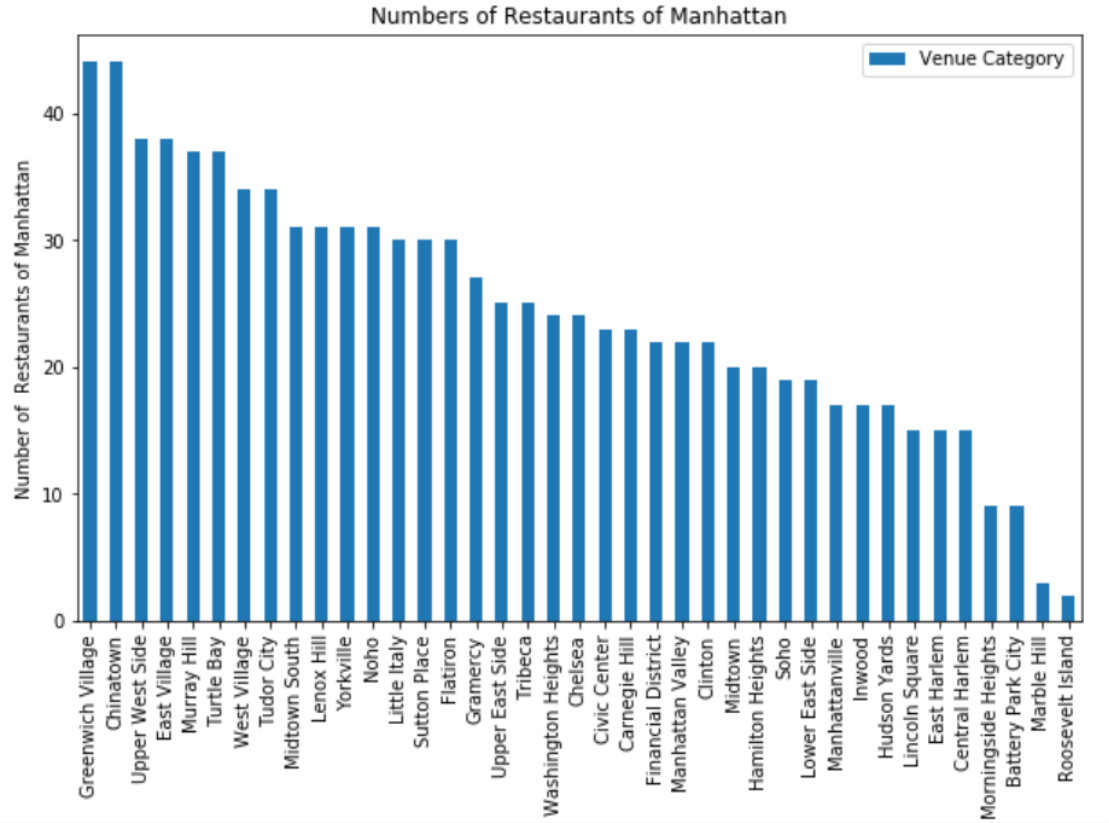

Fig.1 Mount of restaurants in 40 neighborhoods of Manhattan

3.3 Percentage of restaurants in Manhattan

We used the number of restaurants and the mount of all shop in each neighborhood to calculate the percentage of restaurants in Manhattan. The following picture shows which neighborhood is easier to find delicacy.

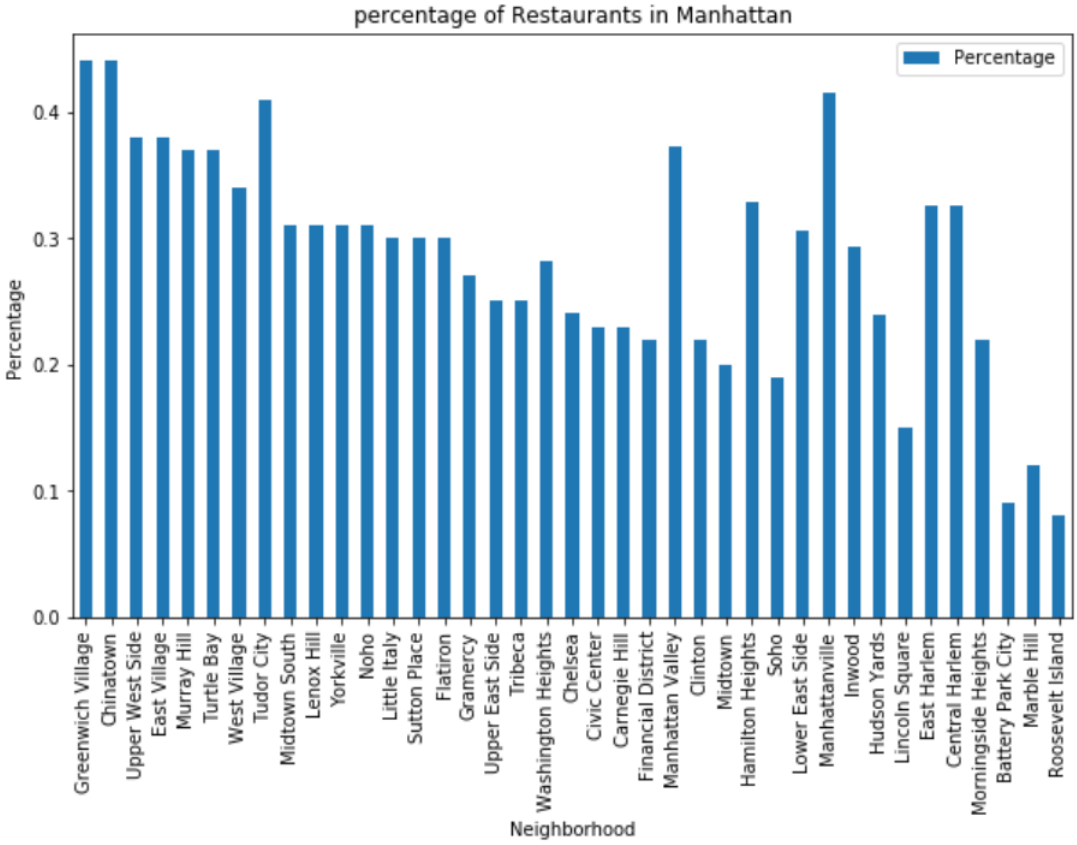

Fig 2 Percentage of restaurants in each neighborhood of Manhattan

3.4 The amount of Chinese restaurants in Manhattan

There are five different kinds of Chinese restaurants, which are marked as Chinese Restaurant, Shanghai Restaurant, Dim Sum Restaurant, Hotpot Restaurant and Dumpling Restaurant. So we extracted all of them to get a whole picture of Chinese restaurants in Manhattan. Chinatown has the most Chinese restaurant.

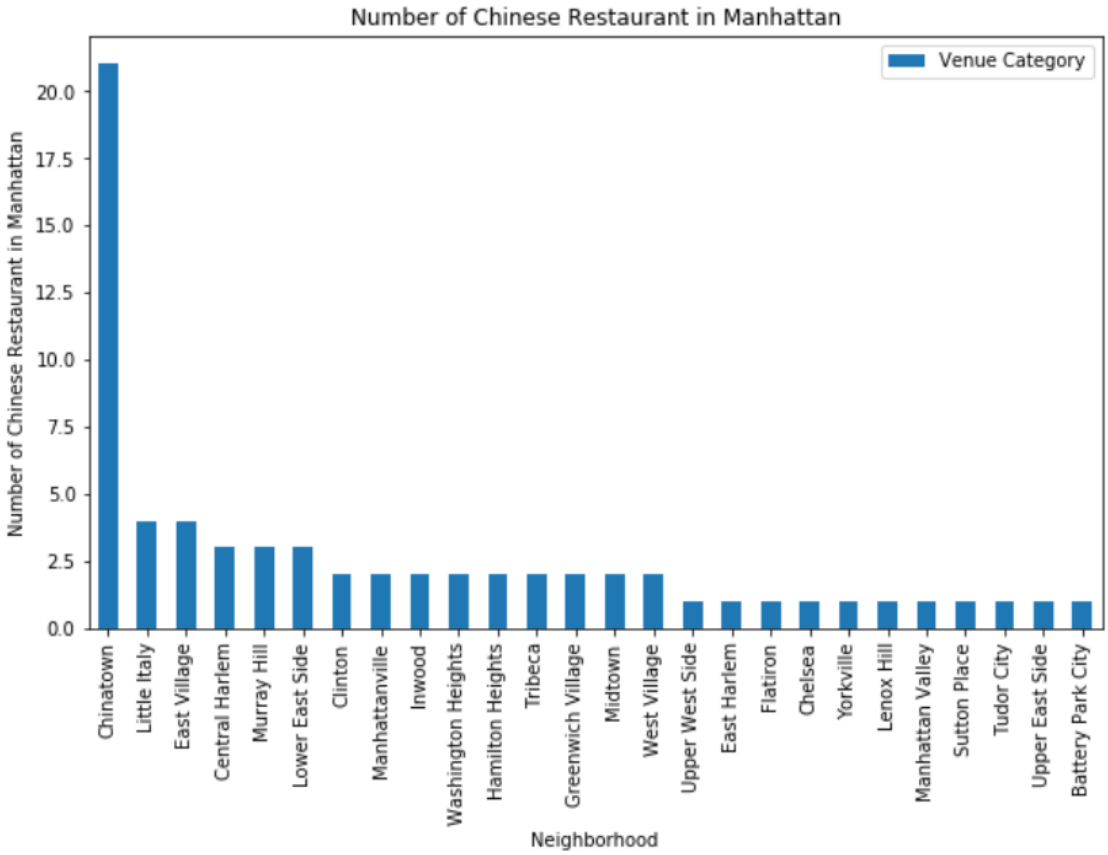

Fig.3 The mount of Chinese restaurant in 40 neighborhoods of Manhattan

3.5 Percentage of Chinese restaurant of all restaurants

We used the amount of Chinese restaurants and all restaurants in each neighborhood to calculate the percentage of restaurants in Manhattan.

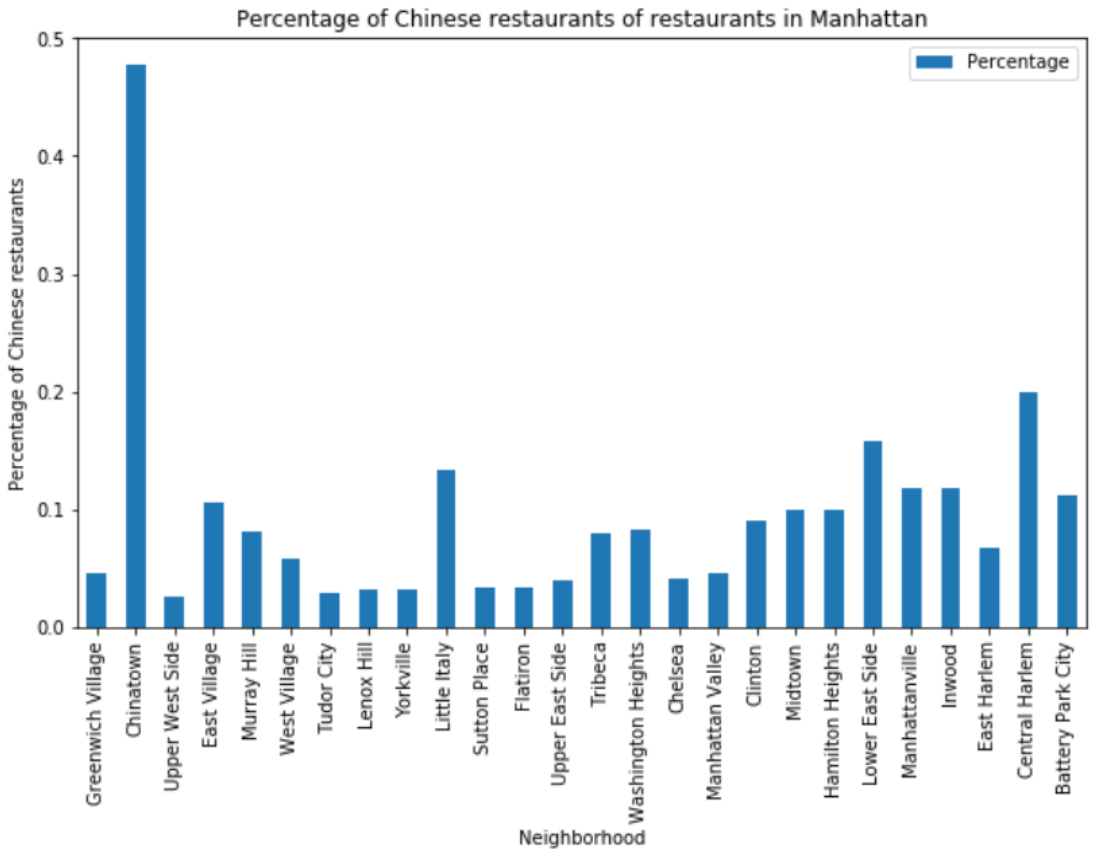

Fig.4 Percentage of Chinese restaurants of restaurant in each neighborhood of Manhattan

4. Clustering neighborhoods

We used k-means clustering to cluster 40 neighborhoods of Manhattan that have similar characteristics. This analysis based on data with 10 tops frequency of occurrence of each category. The value of k was set to be 10, because we want to know which neighborhoods are most similar to Chinatown.

5. Result

The result has 10 clusters. What we are eager to is the one has Chinatown. Seven neighborhoods is clustered with Chinatown. There are Washington Heights, Central Harlem, Midtown, Lower East Side, Little Italy, Soho and Flatiron. More details is one page 9th.

6. Conclusion

We analyzed 3320 shops of 40 neighborhoods in Manhattan, segmented and clustered 40 neighborhoods. Washington Heights, Central Harlem, Midtown, Lower East Side, Little Italy, Soho and Flatiron are similar to Chinatown.

Chinatown has 44 restaurants, 21 of which are Chinese restaurants. The percentage of Chinese restaurant in Manhatan is 47.7%, which is highest. This is a big feature of Chinatown. Surprisely, Soho has no Chinese restaurant in 10 shops, but was still clustered together with others who has Chinese restaurant.

Other six communities, which are Little Italy, Flatiron, Washington Heights, Midtown, Lower East Side and Central Harlem, are recommended to be good choices for someone who would like to open a Chinese restaurant in Manhattan.

Neighborhood	number_of_restaurant	number_of_Chinese_restaurant	Percentage
Chinatown	44	21	0.477273
Little Italy	30	4	0.133333
Flatiron	30	1	0.033333
Washington Heights	24	2	0.083333
Midtown	20	2	0.100000
Lower East Side	19	3	0.157895
Central Harlem	15	3	0.200000

Fig.5 Amount of restaurants of seven clustered neighborhoods

7. Discussion

As it has mentioned before, there are many factor affect a restaurant is going to be popular or not. In this case, a lot of work remains to do. If we have source, we'd better to investigate more information about the six communities, such as preference and income of customers, density of population and venues and competition conditions. What's more, how much you are going to invest on Chinese restaurant is also important. Rent should also be taken into consideration.

Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Manhattan	Chinatown	40.715618	-73.994279	4	Chinese Restaurant	Dim Sum Restaurant	Vietnamese Restaurant	Cocktail Bar	American Restaurant	Bar	Bakery	Noodle House	Hotpot Restaurant	Bubble Tea Shop
Manhattan	Washington Heights	40.851903	-73.936900	4	Café	Bakery	Latin American Restaurant	Mobile Phone Shop	Grocery Store	Deli / Bodega	Gym	Spanish Restaurant	Tapas Restaurant	Shoe Store
Manhattan	Central Harlem	40.815976	-73.943211	4	Chinese Restaurant	African Restaurant	American Restaurant	Cosmetics Shop	Public Art	French Restaurant	Seafood Restaurant	Gym / Fitness Center	Boutique	Bookstore
Manhattan	Midtown	40.754691	-73.981669	4	Hotel	Clothing Store	Steakhouse	Theater	Cocktail Bar	Bookstore	Coffee Shop	Sporting Goods Shop	American Restaurant	Bakery
Manhattan	Lower East Side	40.717807	-73.980890	4	Coffee Shop	Chinese Restaurant	Café	Art Gallery	Shoe Store	Ramen Restaurant	Cocktail Bar	Sandwich Place	Bakery	Japanese Restaurant
Manhattan	Little Italy	40.719324	-73.997305	4	Bakery	Café	Sandwich Place	Seafood Restaurant	Salon / Barbershop	Clothing Store	Ice Cream Shop	Chinese Restaurant	Mediterranean Restaurant	Bubble Tea Shop
Manhattan	Soho	40.722184	-74.000657	4	Clothing Store	Boutique	Women's Store	Men's Store	Shoe Store	Italian Restaurant	Furniture / Home Store	Coffee Shop	Art Gallery	Mediterranean Restaurant
Manhattan	Flatiron	40.739673	-73.990947	4	Yoga Studio	Gym	Japanese Restaurant	Gym / Fitness Center	American Restaurant	Bakery	Cosmetics Shop	Salon / Barbershop	New American Restaurant	Women's Store

Fig.6 Clustering Result of Chinatown