



Recommendation for The Location and Style of a New Chinese Restaurant in New York City Neighborhoods

Xiaozhou (John) Zhang

The business problem and solution

- The NYC area is home to the largest and most prominent ethnic Chinese population outside of Asia, hosting Chinese populations representing all 34 provincial-level administrative units of China.
- Chinese cuisines are an important part of Chinese culture, which includes cuisines originating from the diverse regions of China. There are a variety of styles of cooking in China and these styles are distinctive from one another due to factors such as availability of resources, climate, geography, history, cooking techniques and lifestyle.
- With the use of Python and its vast amount of Data Science libraries, Through the location data analysis we can accurately predict where to open a new Chinese restaurant and decide its style to fit the taste of the favor of the neighborhoods well.
- This data will allow us to report on and provide feedback to interested parties looking at locations and styles to open a new Chinese restaurant in a neighborhood.

The Data

- Following the Business Problem, the following data will be used to tailored to retrieve the data for the systematic analysis:
 - Neighborhood data of the NYC which has a total of 5 boroughs and 306 neighborhoods
 - The list of Chinese restaurants category IDs that were retrieved from Foursquare
- We are going to use the NYC neighborhood data to get the geographical coordinates of all its neighborhoods. The geographical coordinates will be used as the center locations to retrieve the information for recommended Chinese restaurants around it with Foursquare's Endpoint "explore". The preferred Chinese restaurants around certain neighborhoods will be determined with the clustering analysis and visualized on the map.

Methodology

- Libraries used in this study
 - pandas library for data analysis
 - numpy library for handling data in a vectorized manner
 - json library for handling JSON files
 - matplotlib library and associated plotting modules for data visualization
 - KMeans module from sklearn.cluster library for clustering analysis
 - folium library for map rendering
- The coordinates for the 306 neighborhoods of NYC for retrieving nearby recommended Chinese restaurants within the radius of 500 meters of all NYC neighborhoods.
- The retrieved data was analyzed with KMeans clustering with the determined optimal k value and the clusters were visualized by map rendering with folium library.

Glimpse of recommended Chinese restaurants in NYC neighborhoods

Partial view of IDs of Chinese restaurants

	Restaurant Category	Category ID
0	Anhui Restaurant	52af3a5e3cf9994f4e043bea
1	Beijing Restaurant	52af3a723cf9994f4e043bec
2	Cantonese Restaurant	52af3a7c3cf9994f4e043bed
3	Cha Chaan Teng	58daa1558bbb0b01f18ec1d3
4	Chinese Aristocrat Restaurant	52af3a673cf9994f4e043beb
5	Chinese Breakfast Place	52af3a903cf9994f4e043bee
6	Dim Sum Restaurant	4bf58dd8d48988d1f5931735
7	Dongbei Restaurant	52af3a9f3cf9994f4e043bef
8	Fujian Restaurant	52af3aaa3cf9994f4e043bf0
9	Guizhou Restaurant	52af3ab53cf9994f4e043bf1

Count of Chinese restaurants from neighborhoods

Neighborhood		
Bath Beach	4	Jackson Heights
Battery Park City	2	Lenox Hill
Bay Ridge	1	Lincoln Square
Bayside	1	Little Italy
Bensonhurst	2	Little Neck
Chelsea	1	Manhattan Valley
Chinatown	65	Midtown
Clinton	1	Midtown South
College Point	1	Murray Hill
Downtown	2	New Dorp
Dyker Heights	1	Noho
East Village	8	North Side
East Williamsburg	2	Oakland Gardens
Elmhurst	13	Queensboro Hill
Financial District	1	Rego Park
Flushing	25	Sutton Place
Gramercy	4	Tudor City
Greenpoint	2	Turtle Bay
Greenwich Village	1	University Heights
Hillcrest	1	Upper East Side
Homecrest	1	Vinegar Hill
Hunters Point	2	West Village
		Yorkville

Partial view of the retrieved Chinese restaurants

NYC_CNRestaurants.head()

(241, 7)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Restaurant	Restaurant Latitude	Restaurant Longitude	Restaurant Category
0	University Heights	40.855727	-73.910416	Merry Land Buffet	40.853287	-73.907745	Dim Sum Restaurant
1	Bay Ridge	40.625801	-74.030621	Kathy's Dumplings	40.627894	-74.029275	Dim Sum Restaurant
2	Bensonhurst	40.611009	-73.995180	大華小炒皇	40.614777	-73.994774	Cantonese Restaurant
3	Bensonhurst	40.611009	-73.995180	Double Ben Cafe 大家樂	40.612389	-74.000717	Cha Chaan Teng
4	Greenpoint	40.730201	-73.954241	Chiko	40.731871	-73.954369	Szechuan Restaurant

Distribution of Recommended Chinese Restaurants in NYC

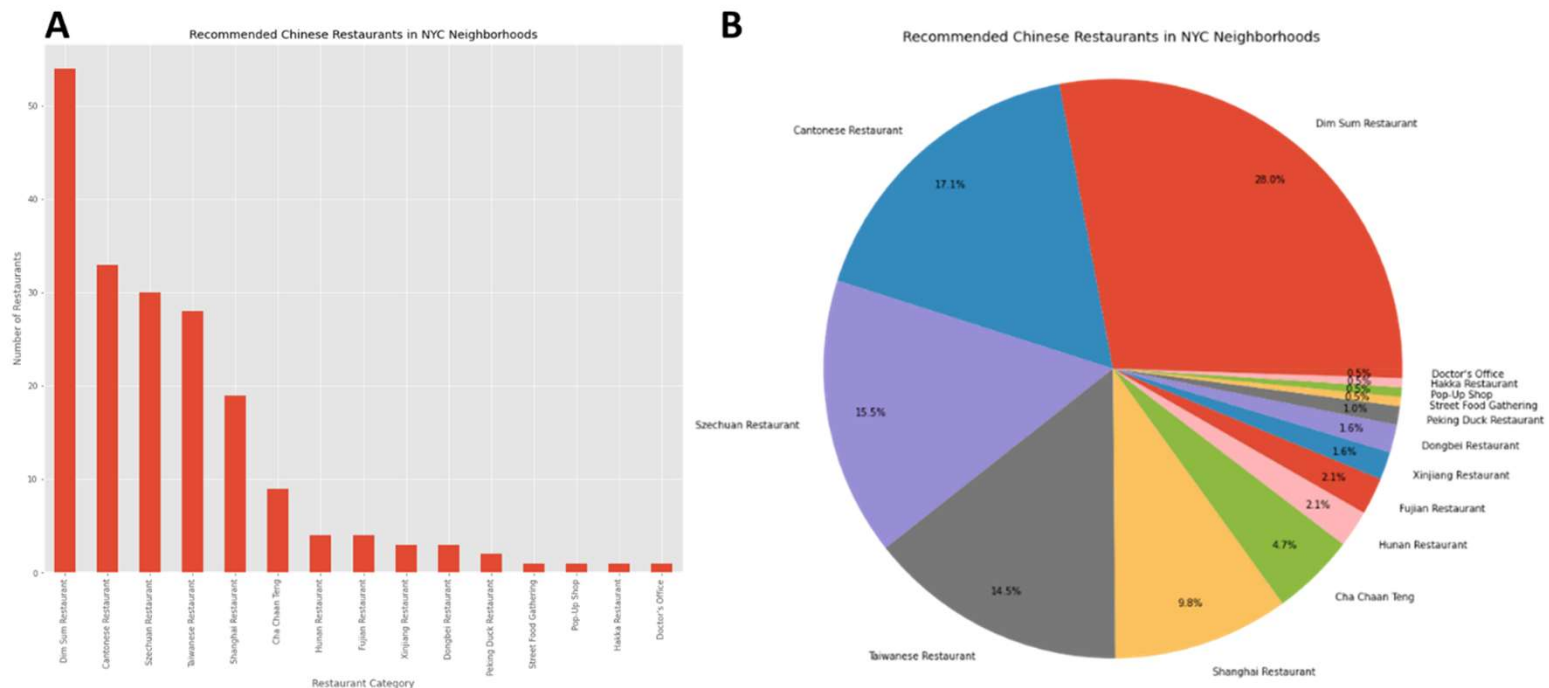
Total number of restaurants
1 65



Hot spots that have the high numbers of recommended Chinese restaurants

- The recommended Chinese restaurants are mostly located in the boroughs of Manhattan and Northern Queens, which centralize at the Chinatown and Flushing neighborhoods, which is not surprising since these two neighborhoods have the largest Chinese American population and they are famous and attractive places for international tourists who want to experience Chinese culture and have a taste of Chinese dishes.

The most recommend Chinese restaurants in NYC neighborhoods



- Dim Sum is the mostly recommended restaurant (54 counts, 28% of the total recommended restaurants).
- The following four highly recommended restaurants after Dim Sum are Cantonese (17.1%), Szechuan (15.5%), Taiwanese (14.5%), and Shanghai (9.8%) restaurants.

The top 5 most commonly recommended Chinese restaurants from different neighborhoods

```
----Chinatown----
      restaurant  freq
0   Dim Sum Restaurant  0.28
1   Cantonese Restaurant  0.25
2   Taiwanese Restaurant  0.11
3     Cha Chaan Teng  0.09
4   Shanghai Restaurant  0.08

----Flushing----
      restaurant  freq
0   Dim Sum Restaurant  0.28
1   Szechuan Restaurant  0.24
2   Taiwanese Restaurant  0.16
3   Shanghai Restaurant  0.12
4   Cantonese Restaurant  0.08

----Clinton----
      restaurant  freq
0   Dim Sum Restaurant  1.0
1   Cantonese Restaurant  0.0
2     Cha Chaan Teng  0.0
3   Doctor's Office  0.0
4   Dongbei Restaurant  0.0

----Gramercy----
      restaurant  freq
0   Dim Sum Restaurant  0.50
1   Szechuan Restaurant  0.25
2   Taiwanese Restaurant  0.25
3   Cantonese Restaurant  0.00
4     Cha Chaan Teng  0.00

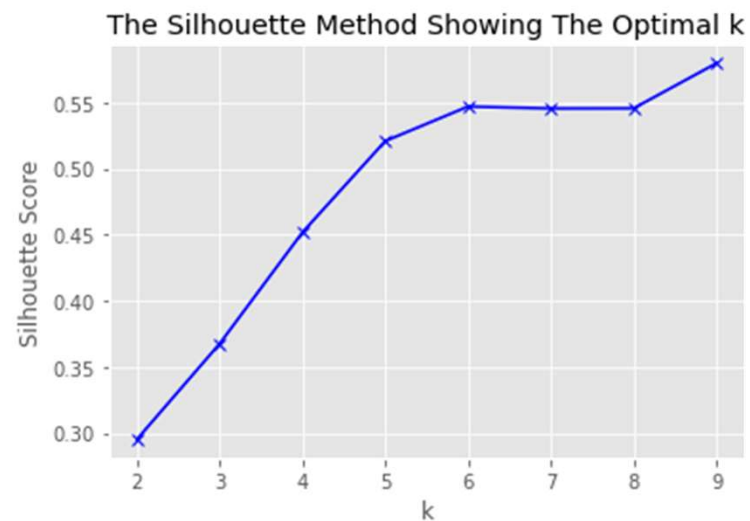
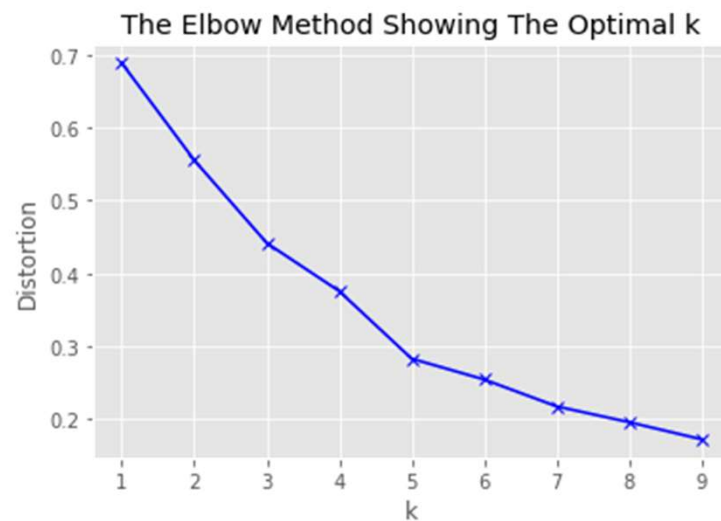
----College Point----
      restaurant  freq
0   Taiwanese Restaurant  1.0
1   Cantonese Restaurant  0.0
2     Cha Chaan Teng  0.0
3   Dim Sum Restaurant  0.0
4   Doctor's Office  0.0

----Greenpoint----
      restaurant  freq
0   Shanghai Restaurant  0.5
1   Szechuan Restaurant  0.5
2   Cantonese Restaurant  0.0
3     Cha Chaan Teng  0.0
4   Dim Sum Restaurant  0.0

----Downtown----
      restaurant  freq
0   Shanghai Restaurant  1.0
1   Cantonese Restaurant  0.0
2     Cha Chaan Teng  0.0
3   Dim Sum Restaurant  0.0
4   Doctor's Office  0.0

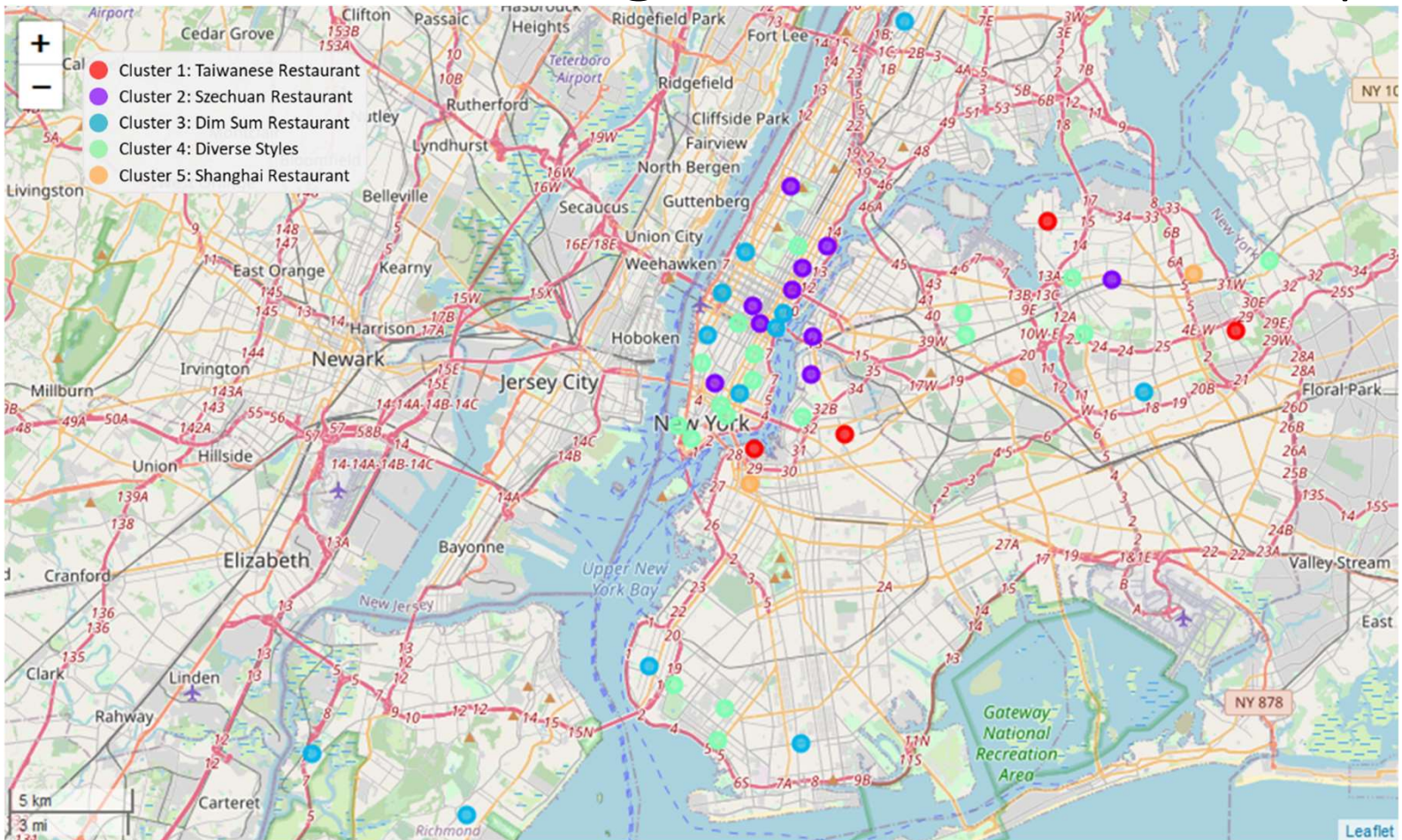
----Greenwich Village----
      restaurant  freq
0   Szechuan Restaurant  1.0
1   Cantonese Restaurant  0.0
2     Cha Chaan Teng  0.0
3   Dim Sum Restaurant  0.0
4   Doctor's Office  0.0
```


Determine the optimal k for clustering analysis



- Both methods showed the turning point at $k=5$ and then 5 was selected for the further clustering analysis

Visualize the resulting clusters on NYC map



Conclusion

- Based on the analysis of the retrieved data from Foursquare and the neighborhoods clustering results, I would recommend one to open the new Chinese restaurant in the area of Chinatown or Flushing neighborhoods from where there are highest population of Chinese Americans and with high traffic for tourists who want to experience Chinese culture and food.
- To brace the competition, I will recommend the style of the restaurant to be Dim Sum, Cantonese, and Szechuan restaurants.
- Keep it in mind that some neighborhoods in Manhattan and Northern Queen are quite different which have the very diverse choice on their first most commonly recommended restaurants (including styles of Cantonese, Dim Sum, Szechuan, Cha Chuan, Street Food Gathering, and Shanghai). I guess this opens the door for the restaurant owner to try some new Chinese cuisines.