BEST LOCATION FOR A CHINESE RESTAURANT IN NEW YORK - IBM CAPSTONE PROJECT

Zhicheng Zhu

Introduction

■ 1.1 Business Problem

■ New York is a city located in New York State, USA, and is the most populous city in the United States. In this project, we going to choose this city as an example to use data analyst skill I learnt in the IBM data science course.
We intend to find the best place to open an Chinese restaurant in New York at Queens. The target audience could be the one who want to find a place to open the Chinese restaurant in New York. We will find the place without too much Chinese restaurant and also with enough people lives. To use k-means clustering and foursquare to analyze data and create maps and also use panda to import data.
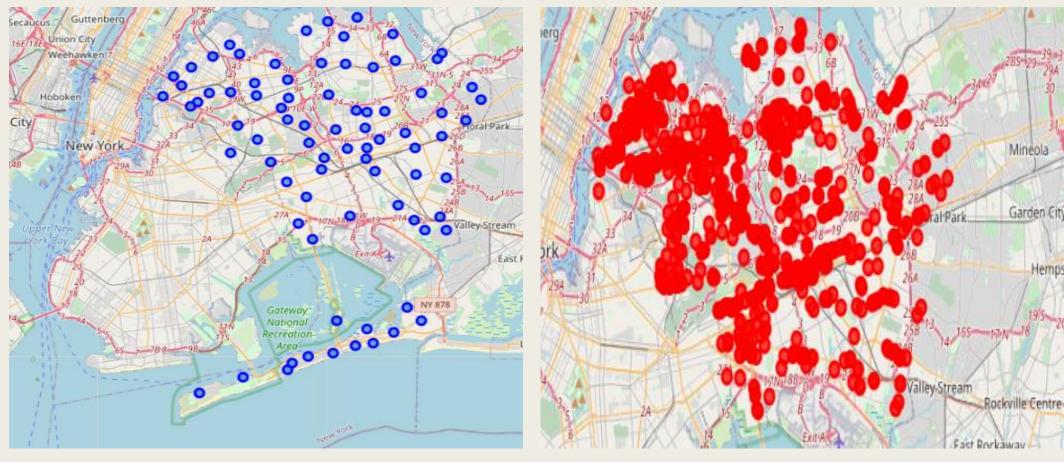
Introduction

- 1.2Methodology

- The purpose of the analysis is to highlight the best Chinese restaurant locations in order to provide entrepreneurs with valuable suggestions. To this end, we will use the machine learning K-Means algorithm to divide New York into several clusters and give a "scoring" form for developing restaurant business from the most interesting area to the least interesting area. But before that, we will first explore the downtown area of New York to get all the restaurant information (data) in the downtown area. This will help us easily highlight the most visited food districts.

Data Mining

- 2.1 Data Collection¶

- All geographic coordinates of Queens, New York will be downloaded here: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json we will Create a table in which each region is listed as a row, and the longitude and latitude are listed as columns.

- In the next step, we collect the locations of each department (listed in the table above) and look at the most common locations. In this step, we will use the Foursquare API to collect the name, category, GPS coordinates of all the data about the place. After collecting the data and sorting it into a panda data frame, we will get a table listing the top 10 locations for each department.

- The final task is to use unsupervised machine learning techniques to cluster them according to the most common places, and use the Folio library to visualize all clusters.

Data Mining

- After we use the data frame we found that, in Queens, there are currently 999 Chinese restaurants opens. So we need do next step to analyze data, and found at least top or top 10 position for our business.

Data Mining

	Borough	Neighborhood	Latitude	Longitude
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

- After we using API to search in Queens, we will find the top 5 location

Data Mining

■ Then, foursquare will going to help us to find all kind of restaurant in all around the queens

Cluster 1

■ Then use this function to group the neighborhoods into clusters, and use the K-means clustering algorithm to complete this task. In addition, the Folium library can also visualize the communities in Queens and its emerging clusters.

	Neighborhood	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
0	Astoria	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
4	Howard Beach	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
5	Corona	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
7	Kew Gardens	Chinese Restaurant	Japanese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant
10	Long Island City	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
11	Sunnyside	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
12	East Elmhurst	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
13	Maspeth	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
15	Glendale	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
17	Woodhaven	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
18	Ozone Park	Chinese Restaurant	Caribbean Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
19	South Ozone Park	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
26	Glen Oaks	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
27	Bellerose	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
31	Jamaica Center	Chinese Restaurant	Indian Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant
33	Queens Village	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
34	Hollis	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
35	South Jamaica	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
37	Rochdale	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
38	Springfield Gardens	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
39	Cambria Heights	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
40	Rosedale	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
41	Far Rockaway	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
42	Broad Channel	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
44	Steinway	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
45	Beechhurst	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
46	Bay Terrace	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
47	Edgemere	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
48	Arverne	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
49	Rockaway Beach	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
50	Neponsit	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
52	Floral Park	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
53	Holliswood	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
58	Lindenwood	Chinese Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant

Cluster 2 and 3

■ Then use this function to group the neighborhoods into clusters, and use the K-means clustering algorithm to complete this task. In addition, the Folium library can also visualize the communities in Queens and its emerging clusters.

	Neighborhood	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
2	Jackson Heights	Chinese Restaurant	Taiwanese Restaurant	Spanish Restaurant	Seafood Restaurant	Cantonese Restaurant	Indian Restaurant	Vietnamese Restaurant	Fried Chicken Joint	Bakery	Buffet
3	Elmhurst	Chinese Restaurant	Taiwanese Restaurant	Indian Restaurant	Shanghai Restaurant	Seafood Restaurant	Cantonese Restaurant	Vietnamese Restaurant	Fried Chicken Joint	Bakery	Buffet
6	Forest Hills	Chinese Restaurant	Shanghai Restaurant	Szechuan Restaurant	Kosher Restaurant	Dim Sum Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant
16	Rego Park	Chinese Restaurant	Shanghai Restaurant	Taiwanese Restaurant	Kosher Restaurant	Dim Sum Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant
20	College Point	Chinese Restaurant	Taiwanese Restaurant	Buffet	Vietnamese Restaurant	Halal Restaurant	Bakery	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant
22	Bayside	Chinese Restaurant	Taiwanese Restaurant	Shanghai Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop
28	Kew Gardens Hills	Chinese Restaurant	Asian Restaurant	Indian Chinese Restaurant	Gluten-free Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop
29	Fresh Meadows	Chinese Restaurant	Asian Restaurant	Cantonese Restaurant	Coffee Shop	Vegetarian / Vegan Restaurant	Halal Restaurant	Bakery	Buffet	Caribbean Restaurant	Dim Sum Restaurant
32	Oakland Gardens	Chinese Restaurant	Vietnamese Restaurant	Taiwanese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant
51	Murray Hill	Chinese Restaurant	Taiwanese Restaurant	Szechuan Restaurant	Asian Restaurant	Shanghai Restaurant	Korean Restaurant	Fried Chicken Joint	Bakery	Buffet	Cantonese Restaurant
67	Forest Hills Gardens	Chinese Restaurant	Szechuan Restaurant	Dim Sum Restaurant	Japanese Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant
72	Hunters Point	Chinese Restaurant	Taiwanese Restaurant	Szechuan Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop

	Neighborhood	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
9	Flushing	Chinese Restaurant	Hotpot Restaurant	Dim Sum Restaurant	Szechuan Restaurant	Cantonese Restaurant	Fried Chicken Joint	Bakery	Dongbei Restaurant	Dumpling Restaurant	Food Stand
56	Hillcrest	Chinese Restaurant	Asian Restaurant	Cantonese Restaurant	Dim Sum Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	Bakery	Buffet	Caribbean Restaurant	Coffee Shop

Cluster 4

■ Then use this function to group the neighborhoods into clusters, and use the K-means clustering algorithm to complete this task. In addition, the Folium library can also visualize the communities in Queens and its emerging clusters.

	Neighborhood	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
1	Woodside	Chinese Restaurant	Halal Restaurant	Indian Restaurant	Vietnamese Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant
8	Richmond Hill	Chinese Restaurant	Shanghai Restaurant	Japanese Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop
14	Ridgewood	Chinese Restaurant	Asian Restaurant	Buffet	Cantonese Restaurant	Halal Restaurant	Bakery	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
21	Whitestone	Chinese Restaurant	Asian Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
23	Auburndale	Chinese Restaurant	Asian Restaurant	Cantonese Restaurant	Halal Restaurant	Bakery	Buffet	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
30	Briarwood	Chinese Restaurant	Asian Restaurant	Indian Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant
36	St. Albans	Chinese Restaurant	Asian Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
54	Jamaica Estates	Chinese Restaurant	Asian Restaurant	Cantonese Restaurant	Caribbean Restaurant	Halal Restaurant	Bakery	Buffet	Coffee Shop	Dim Sum Restaurant	Dongbei Restaurant
57	Ravenswood	Chinese Restaurant	Noodle House	Latin American Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant	Coffee Shop
60	Lefrak City	Chinese Restaurant	Taiwanese Restaurant	Shanghai Restaurant	Kosher Restaurant	Vietnamese Restaurant	Gluten-free Restaurant	Bakery	Buffet	Cantonese Restaurant	Caribbean Restaurant
68	Jamaica Hills	Chinese Restaurant	Caribbean Restaurant	Indian Restaurant	Vietnamese Restaurant	Halal Restaurant	Bakery	Buffet	Cantonese Restaurant	Coffee Shop	Dim Sum Restaurant

Result

■ After the project, We got the conclusion, The cluster 3 will be the best position to have the restaurant

	Neighborhood	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
9	Flushing	Chinese Restaurant	Hotpot Restaurant	Dim Sum Restaurant	Szechuan Restaurant	Cantonese Restaurant	Fried Chicken Joint	Bakery	Dongbei Restaurant	Dumpling Restaurant	Food Stand
56	Hillcrest	Chinese Restaurant	Asian Restaurant	Cantonese Restaurant	Dim Sum Restaurant	Vegetarian / Vegan Restaurant	Halal Restaurant	Bakery	Buffet	Caribbean Restaurant	Coffee Shop

Discussion and conclusion

■ All the result could be works for limited data but not for a large and unstable one. The result reflects that it could be correct or not. Which is unstable. There are a lot of area might be good for the restaurant but the analyst result is not include. Which means it might be not reliable. It need people to double check could be the best way after the data analyse. It can only provide some of position as recommend and the truth should be explore by real people.