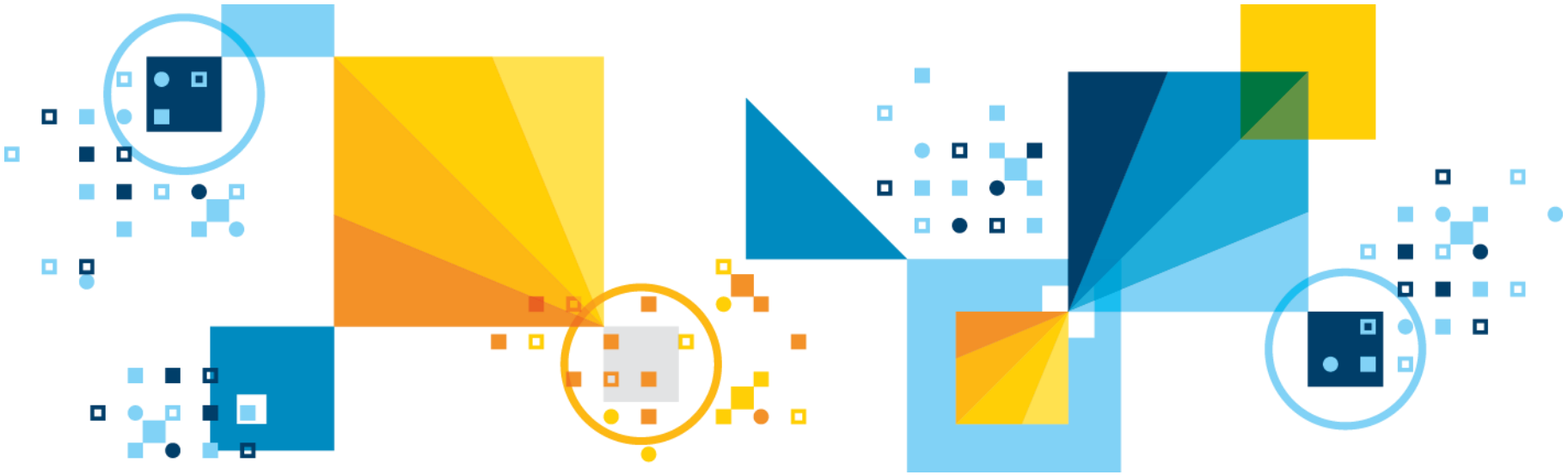
Battle of Neighborhoods
Final Presentation

by Chris Yu



Introduction

- The initial intention is to design this project as a consulting investigation for the catering industry investors by data exploration technologies.
- A Chinese chain restaurant (SWJ) wish to expand its business to North America. SWJ is a very popular brand in China. We select Queens, New York City as the first landing location for SWJ in North America.
- SWJ's target customers are the tourists visiting New York City from China. And the location of the new-open restaurant would be near tourism attractions, shopping stores and Chinese-like cuisine restaurants.

Our Mission:
To find a good location to open a Chinese chain restaurant in Queens, New York City.

Data Sources

- **Neighborhood Data:** Queens, which is the target boroughs of this investigation within New York City, has 80 neighborhoods. We obtain a list of all the locations of the neighborhoods in Queens. From: ttps://geo.nyu.edu/catalog/nyu_2451_34572

- **Geographical Coordinates:** Geographical coordinates for each neighborhood will be obtained with the aid of GEOPY Library. Each neighborhood will be assigned a latitude and longitude coordinate

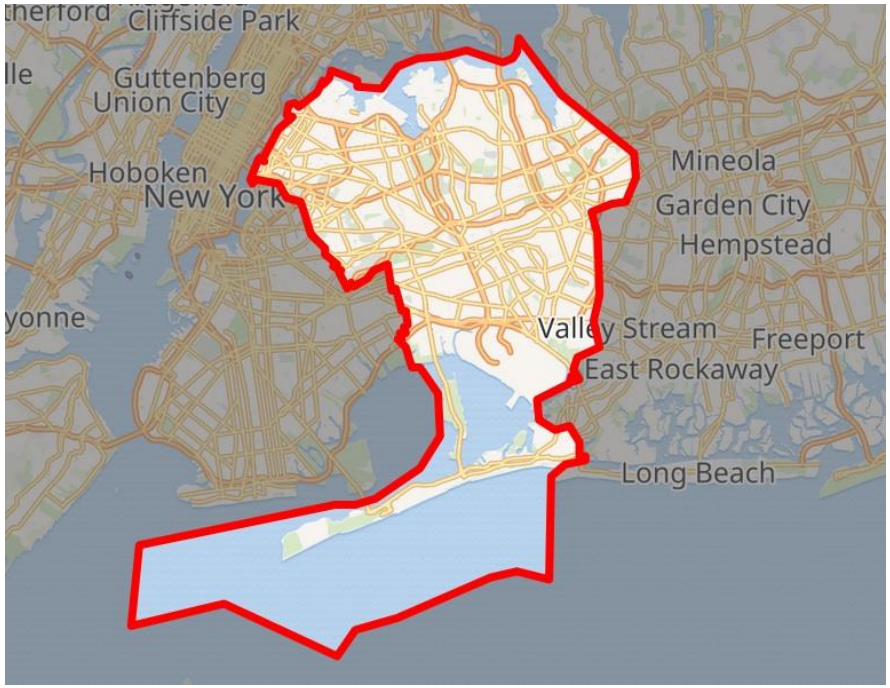
.

- **Venue categories:** We will use the Foursquare API to retrieve venues, using the coordinates obtained in Data Source 2 above. We shall further obtain a list using Foursquare API for related venues such as attractions, shops and restaurants in Queens.

Methodology and Approach

- Based on the business problem and Data Sources described above, I decide to leverage data exploration, data preparation, data visualization and machine learning technologies and tools to investigate the neighborhood in Queens, New York City. Our target is to find a good location to open a Chinese chain restaurant.

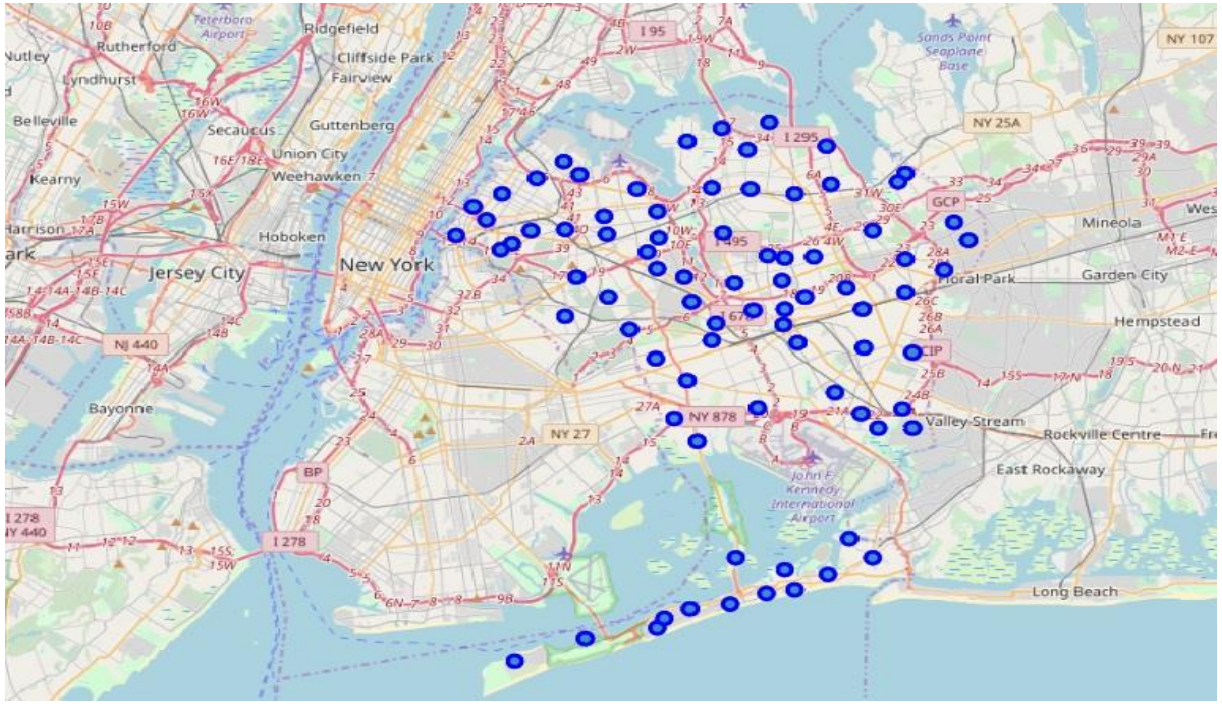
About Queens, New York City



From https://en.wikipedia.org/wiki/Queens

- Queens is the easternmost of the five boroughs of New York City. It is the largest borough geographically. the borough of Queens is the second largest in population (after Brooklyn), with an estimated 2,358,582 residents in 2017. Queens County also is the second most populous county in the U.S. state of New York, behind Brooklyn

The neighborhoods in Queens



- There’re totally 80 neighborhoods in Queens. We can visualize the distribution of the neighborhoods in the map, as well get the geographical data from public data source.

Select related venue categories

- We leverage Foursquare API to retrieve venue data in Queens and transfer the data into a dataframe, which include neighborhood, latitude, longitude, venue, venue category
- Since the new-open restaurant mainly focus on tourists from China, We selected venue categories related to tourism, attractions, shopping and restaurants. The following is the venue category list:

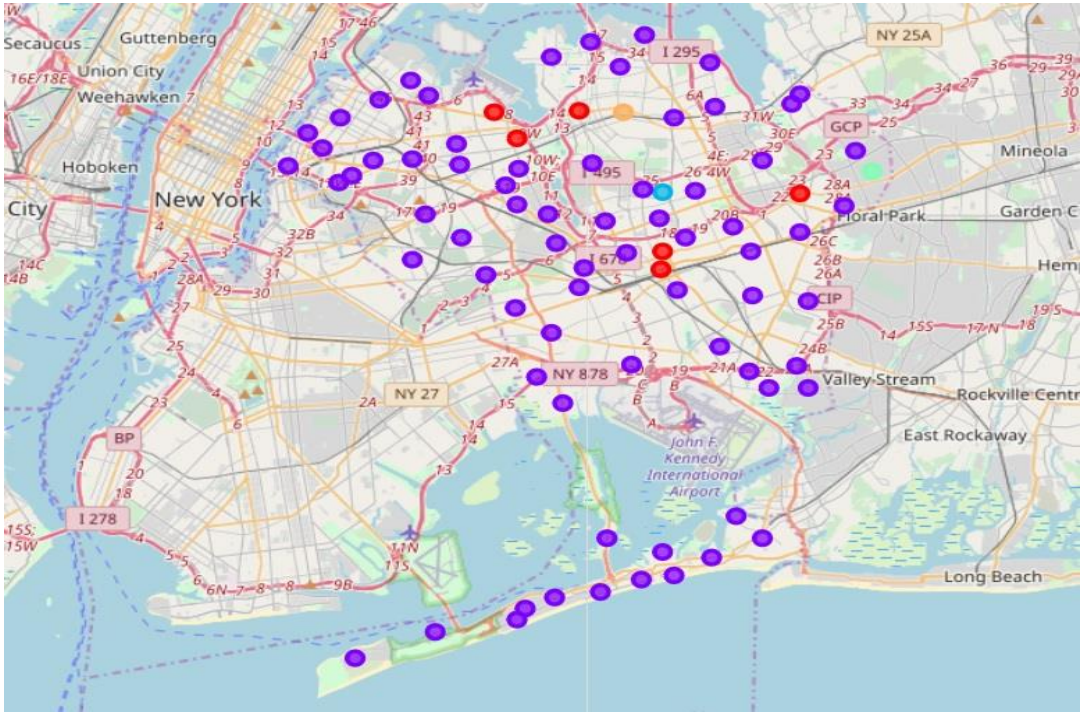
'Art Museum', 'Arts & Entertainment', 'Asian Restaurant', 'Beach', 'Cantonese Restaurant', 'Chinese Restaurant', 'Clothing Store', 'Concert Hall', 'Department Store', 'Dim Sum Restaurant', 'Diner', 'Fast Food Restaurant', 'General Entertainment', 'Gift Shop', 'Hotel', 'Japanese Restaurant', 'Korean Restaurant', "Men's Store", 'Monument / Landmark', 'Museum', 'Park', 'Restaurant', 'Seafood Restaurant', 'Shanghai Restaurant', 'Shoe Store', 'Shopping Mall', 'Sushi Restaurant', 'Taiwanese Restaurant', 'Theater'

Venue category analysis by neighborhoods

	Neighborhood	Art Museum	Arts & Entertainment	Asian Restaurant	Beach	Cantonese Restaurant	Chinese Restaurant	Clothing Store	Concert Hall	Department Store	Dim Sum Restaurant	Diner	Fast Food Restaurant	General Entertainment	Gift Shop	Hotel
0	Arverne	0.000000	0.000000	0.000000	0.058824	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
1	Astoria	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.010000	0.000000	0.000000	0.000000	0.000000
2	Astoria Heights	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
3	Auburndale	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.052632	0.000000	0.000000	0.000000
4	Bay Terrace	0.000000	0.000000	0.024390	0.000000	0.000000	0.000000	0.121951	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.024390	0.000000
5	Bayside	0.000000	0.000000	0.014706	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
6	Bayswater	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
7	Beechhurst	0.000000	0.000000	0.000000	0.000000	0.000000	0.058824	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000
8	Bellaire	0.000000	0.000000	0.000000	0.000000	0.000000	0.166667	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000

- We summarized and normalized the venue quantity grouped by neighborhoods with the important venue category list.

Clustring the neighborhoods



- In the machine learning phase, we leveraged clustering(K-means) to cluster the neighborhoods into 5 clusters based on the attributes of venues. The chart is the visualization of the 5 clusters

Results

- As a result, it is found that cluster 1 is the cluster in which the neighborhoods are good locations to open a Chinese chain restaurant.

- This cluster contains 6 neighborhoods: **Flushing, East Elmhurst, Jamaica Center, Bellaire, North Corona, Jamaica Hills.**

- We can find that all of these 6 neighborhoods have very similar popular venues like Chinese Restaurants, Clothing Store, Theater, Museum, Asian Restaurants.

Conclusion

- I investigated the neighborhoods data in Queens, New York City, with Foursquare venue categories. By using data preparation, exploration, visualization and machine learning, I finally find a group of neighborhoods as the good location to open a Chinese chain restaurant (SWJ). That's a fulfillment and practice of Data Science course.

- I suggest to open a Chinese chain restaurant in following neighborhood:

Flushing, East Elmhurst, Jamaica Center, Bellaire, North Corona, Jamaica Hills.

धन्यवाद
Hindi

多謝
Traditional Chinese

Grazie
Italian

ขอบคุณ
Thai

Gracias
Spanish

Merci
French

Спасибо
Russian

شكرا
Arabic

Obrigado
Brazilian Portuguese

Danke
German

多谢
Simplified Chinese

நன்றி
Tamil

ありがとうございました
Japanese

감사합니다
Korean