

Coursera Capstone Project

The Battle of Neighbourhoods in Chicago

Brian Wu

September 4, 2021

1. Introduction/Business Problem

A sushi franchise owner is seeking perfect locations to open branches where he can introduce the finest sushi to the residents of the city. However, he is new to the city and couldn't decide where to set root for the business to grow. The three rules for starting a business are 1)location, 2)location, and 3)location! Therefore, he seeks help from data scientists and engineers to solve the problem that could be the deciding factor to this expansion.

2. Data

Based on definition of our problem, factors that will influence the decision are:

- Postal and geolocation data of the neighbourhoods in Chicago.
- Number of restaurants and the ratio of restaurants venues to other venues.
- Number of sushi restaurants in the neighbourhoods.

3. Methodology

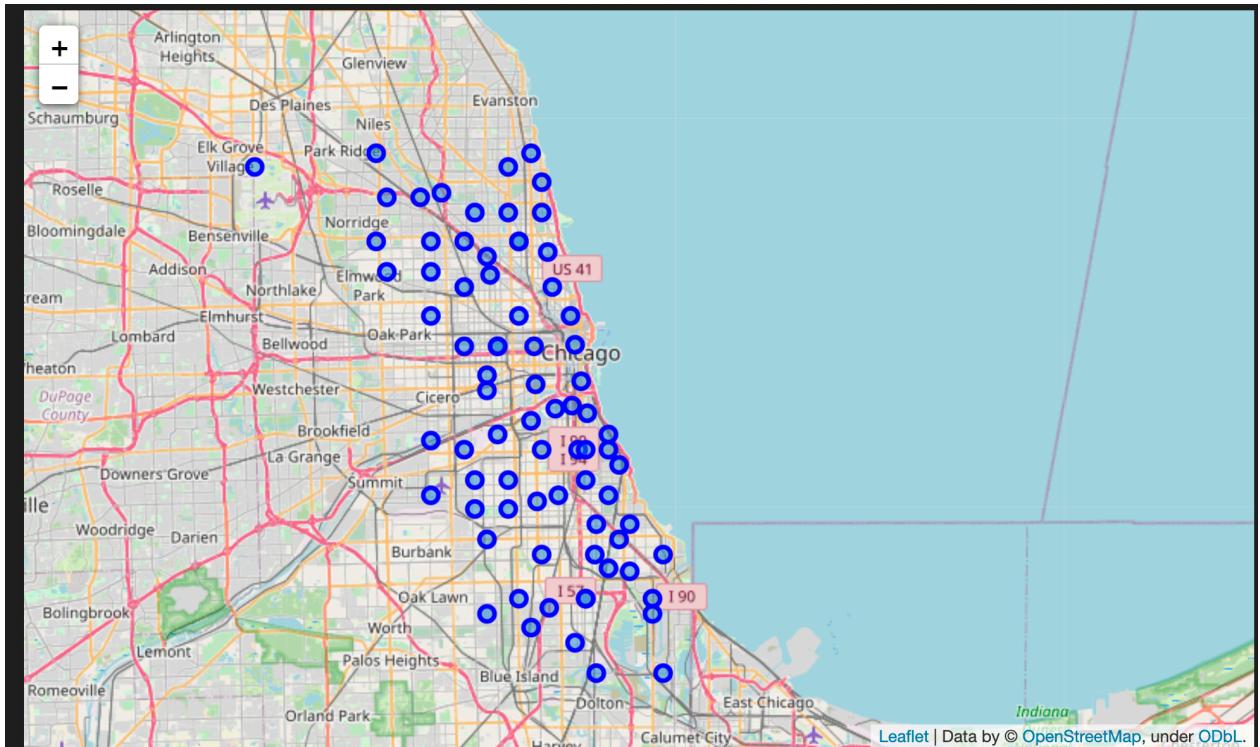
We will first identify and map the neighbourhoods on a folium map to provide a taste of where and how the neighbourhoods in Chicago are located. Next, we will start to explore each neighbourhoods, using Foursquare API to help us find restaurants venues and all other types of venues, and make exploratory data analyses to see what the neighbourhoods are made up of. Then, we will use Foursquare API again to find all the sushi restaurants in each neighbourhood to identify potential competitors, and count them as one of the important factors. Finally, do a K-means clustering analysis to categorize them into 5 clusters, which could be served as indicators to help us determine whether a neighbourhood has the potential to start a successful business.

3.1 Find the coordinates of the neighbourhoods

Out[108]:

	No.	Name	Latitude	Longitude
0	01	Rogers Park	42.010000	-87.670000
1	02	West Ridge	42.000000	-87.690000
2	03	Uptown	41.970000	-87.660000
3	04	Lincoln Square	41.970000	-87.690000
4	05	North Center	41.950000	-87.680000
...
72	73	Washington Heights	41.703833	-87.653667
73	74	Mount Greenwood	41.700000	-87.710000
74	75	Morgan Park	41.690000	-87.670000
75	76	O'Hare	42.000000	-87.920000
76	77	Edgewater	41.990000	-87.660000

77 rows × 4 columns



3.2 A brief summary of venues and the frequency of restaurants of each neighbourhoods

Restaurant Frequency in ascending order

No.		Name	Latitude	Longitude	Restaurant	Freq.
0	60	Bridgeport	41.837500	-87.647500		0.00
1	56	Garfield Ridge	41.816667	-87.760000		0.00
2	28	Near West Side	41.880000	-87.666667		0.00
3	46	South Chicago	41.740000	-87.550000		0.04
4	57	Archer Heights	41.810000	-87.730000		0.08

Restaurant Frequency in descending order

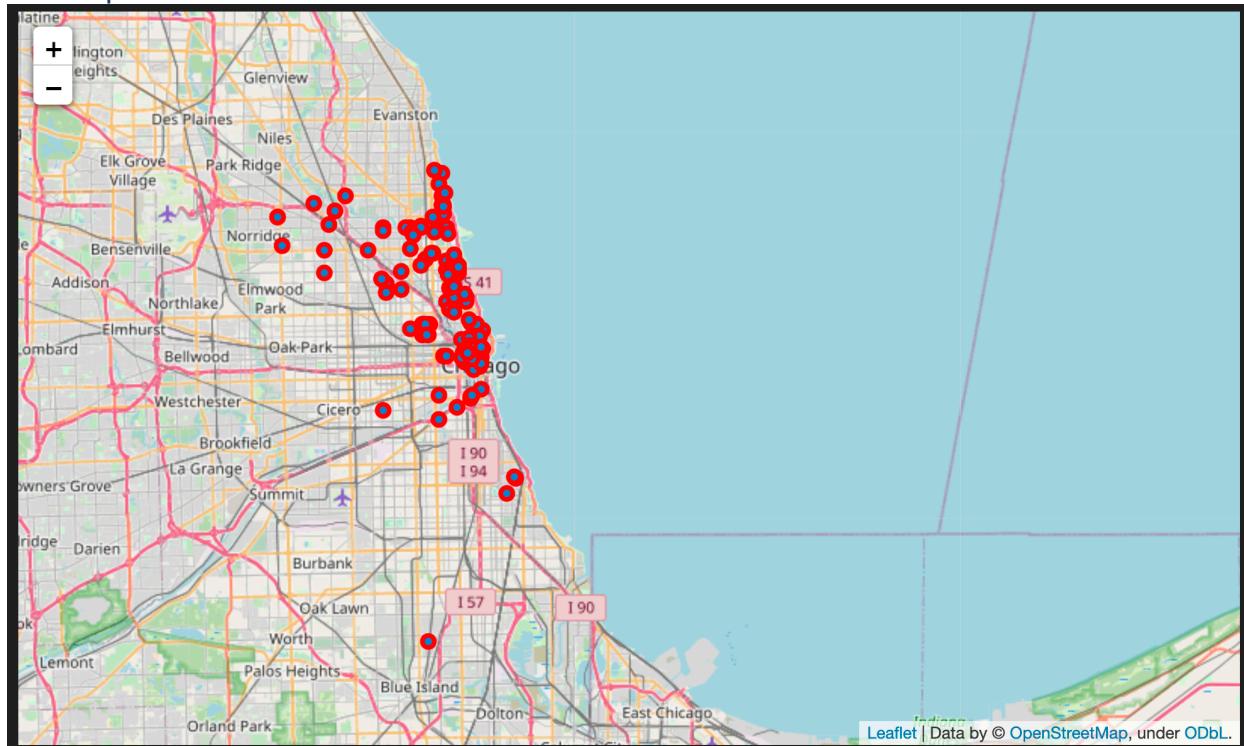
No.		Name	Latitude	Longitude	Restaurant	Freq.
0	75	Morgan Park	41.690000	-87.670000		0.54
1	21	Avondale	41.940000	-87.710000		0.44
2	65	West Lawn	41.770000	-87.720000		0.44
3	67	West Englewood	41.775833	-87.664167		0.42
4	73	Washington Heights	41.703833	-87.653667		0.37

3.3 Now we know how each neighbourhood is made of, we will start to find all the sushi restaurants in the neighbourhoods.

Out[117]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Restaurant Name	Restaurant Latitude	Restaurant Longitude
0	Rogers Park	42.01	-87.67	Asahi Roll	42.005543	-87.660996
1	Rogers Park	42.01	-87.67	Hana	42.005825	-87.660520
2	Rogers Park	42.01	-87.67	Hira's Cafe	42.007936	-87.666718
3	Uptown	41.97	-87.66	Agami Contemporary Sushi	41.967519	-87.658831
4	Uptown	41.97	-87.66	Dib Sushi Bar & Thai Cuisine	41.969042	-87.655973
5	Uptown	41.97	-87.66	Taketei Sushi	41.978093	-87.658353
6	Uptown	41.97	-87.66	Ora	41.975715	-87.668389
7	Uptown	41.97	-87.66	Wabi Sabi Rotary	41.964322	-87.654553
8	Uptown	41.97	-87.66	Gorilla Sushi Bar	41.965832	-87.666872
9	Lincoln Square	41.97	-87.69	Sushi Tokoro	41.968376	-87.688964

3.4 Map out the sushi restaurants

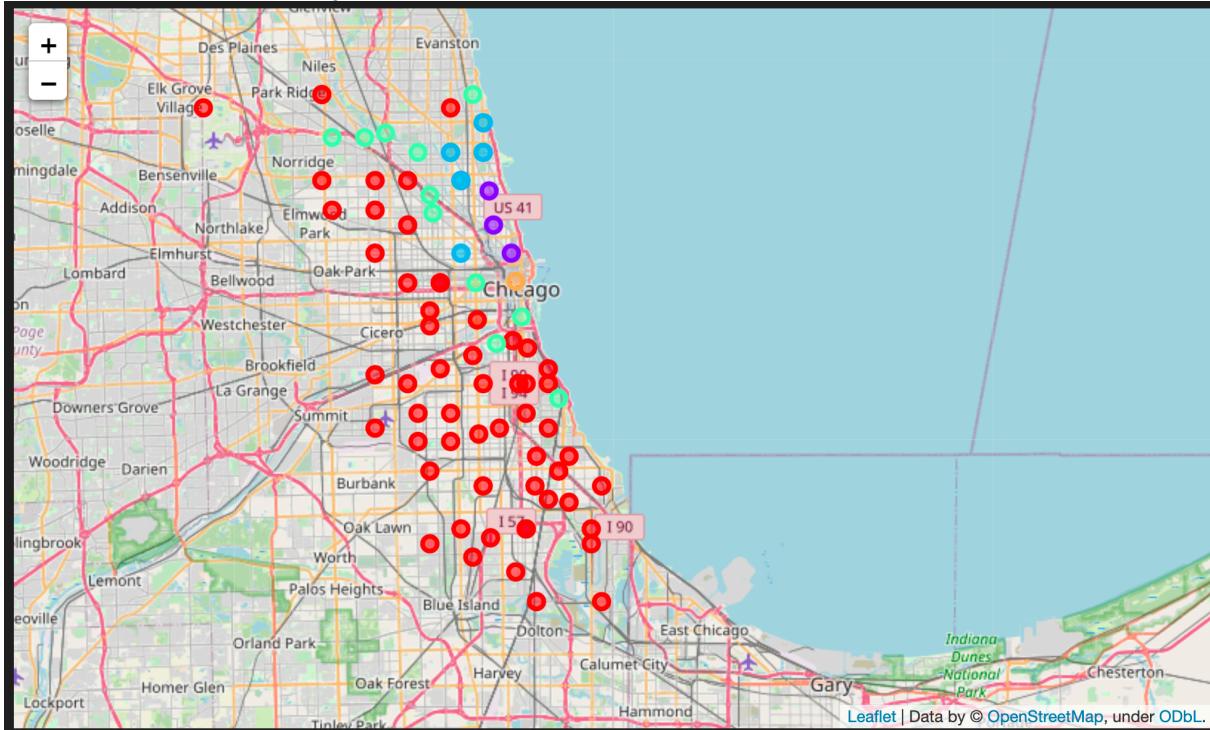


3.5 Sushi restaurants count of each neighbourhood

Out[119]:

	Name	count
0	Albany Park	2
1	Avondale	3
2	Belmont Cragin	1
3	Bridgeport	2
4	Dunning	1
5	Edgewater	7
6	Forest Glen	2
7	Hyde Park	2
8	Irving Park	1
9	Jefferson Park	3
10	Lake View	16
11	Lincoln Park	13
12	Lincoln Square	6
13	Logan Square	4
14	Loop	24
15	Lower West Side	1

3.6 KMean cluster analysis on the data



4. Results and Discussion

The analysis shows that neighbourhoods of cluster 0(red) shows a high potential of opening a successful sushi restaurant, since there is little competition and moderate rate of restaurants occupied the area. Cluster 3(green) has second less in competition, followings are cluster 2(blue), cluster 1(purple), and cluster 4(orange), which is highly crowded with sushi restaurant. Therefore, we can start our location search by the neighbourhoods with red or orange labels; neighbourhoods with blue labels are not recommended; and neighbourhoods with purple or green labels are places we strongly don't recommend.

5. Conclusion

The purpose of this project is to identify the spots in Chicago where there are less competition and show high potential for sushi restaurant businesses to grow. Our analysis could give the stakeholders a better understanding of the ecosystem of the Chicago's restaurant business and the areas around. Location matters the most when starting a business from scratch. Based on our analysis on the neighbourhoods, we enable the stakeholders and give them a head start to spot and secure their seats in this highly competitive food market.