Applied Data Science Capstone- The Battle of Neighborhoods

Where to Open a New East Asian Restaurant in Toronto

Wei Zhong

08/2020

Introduction

Canada is an immigrant country. As a result, there is a large number of immigrants in the metropolises of Canada who originate from different foreign countries. Toronto is one of the examples, in which about 46 percent of population are immigrants, according to the public data source. The immigrants will also bring unique traditions, cultures besides help developing the economy of the country. Food is one of the diversities that will be taken in. Hence, there is a variety of local restaurants offer special food that originate from other different countries. For example, you can come across Japanese restaurants in Toronto. As immigrant population continues to grow, new restaurants are also opening. Under such circumstances, we are interested in searching for a good region of the Toronto city to launch a new East Asian restaurant. The three main countries in East Asian are China, Korean and Japanese. This report will be specifically useful for the stakeholders who are interested to open a new East Asian restaurant in Toronto.

Data

We need the neighbor information of Toronto. We can load the table into dataframe from the following Wikipedia page: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. In order to utilize the Foursquare location data, we need to get the latitude and the longitude coordinates of each neighborhood. We will load the geographical coordinates of each postal code from the csv file via: http://cocl.us/Geospatial_data. We define a Pandas dataframe storing the neighborhood information of Toronto, shown in Figure 1. And we will couple it with the Foursquare location data to analyze the distribution of the East Asian restaurants in Toronto.

Methodology

There are 10 boroughs and 103 different regions in Toronto. Firstly, we use k-means to divide the 103 regions to 10 groups according to their coordinates. We visualize the 103 regions and 10 groups on the map of Toronto, shown in Figure 2. The 10 large dots represent the centroids of each group while the small dots are the coordinates of the 103 regions. Each of the large circles cover a radius of 4000 meters

around the centroids. We can see that the 10 large circles cover most part of Toronto City. Then we will explore the number of Chinese, Japanese and Korean restaurants in the 10 circles covering the city of Toronto, respectively, using Foursquare API.

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
...	...	...	...	...	...
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.653654	-79.506944
99	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
100	M7Y	East Toronto	Business reply mail Processing Centre, South C...	43.662744	-79.321558
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.636258	-79.498509
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.628841	-79.520999

Figure 1 A few rows of the Pandas dataframe storing the neighborhood information.

Figure 2 10 groups of neighborhoods in Toronto by K-means

Results

We explored the Korean, Japanese and Chinese restaurants around the centers of the 10 groups of neighborhoods in Toronto using Foursquare API. And the distributions are shown in Figure 3(a), (b) and (c), respectively. We can observe that the Korean restaurants are mostly found in Group 4 and 8 while the each of the other circles have no more than 5 Korean restaurants except Group 6, Figure 3(a). Japanese restaurants are more widely distributed in Toronto than Korean restaurants, Figure 3(b). And Figure 3(c) shows Chinese restaurants are distributed uniformly in Toronto compared to Korean and Japanese restaurants.

Discussion

We have visually explored the distributions of Chinese, Japanese and Korean restaurants respectively in the 10 groups of neighborhoods. Also, we add a column "East Asian Restaurant" including the total number of Chinese, Japanese and Korean restaurants in each group and thus build a table summarizing the number of East Asian restaurants in the 10 groups of neighborhoods in Toronto, shown in Table 1.

We can use bar plot the visualize the number of restaurants originated from China, Japan and Korea, as shown in Figure 4. For Chinese restaurants, the competition is only low in Group 5. On the other hand, Group 2, 5, and 9 do not have many Japanese restaurants. And Korean restaurants do not have strong competitions in most groups except Group 4, 6 and 8. If we consider all Chinese, Japanese and Korean restaurants as "East Asian Restaurant", we will see that only Group 5 and 9 are not quite crowded with East Asian Restaurants, Figure 5.

Table 1 Number of East Asian restaurants in the 10 groups of neighborhoods in Toronto

Group	Chinese Restaurant	Japanese Restaurant	Korean Restaurant	East Asian Restaurant
0	10	22	1	33
1	9	10	5	24
2	17	3	2	22
3	11	6	3	20
4	47	42	42	131
5	1	2	2	5
6	47	22	12	81
7	21	11	1	33
8	25	32	44	101
9	11	0	1	12

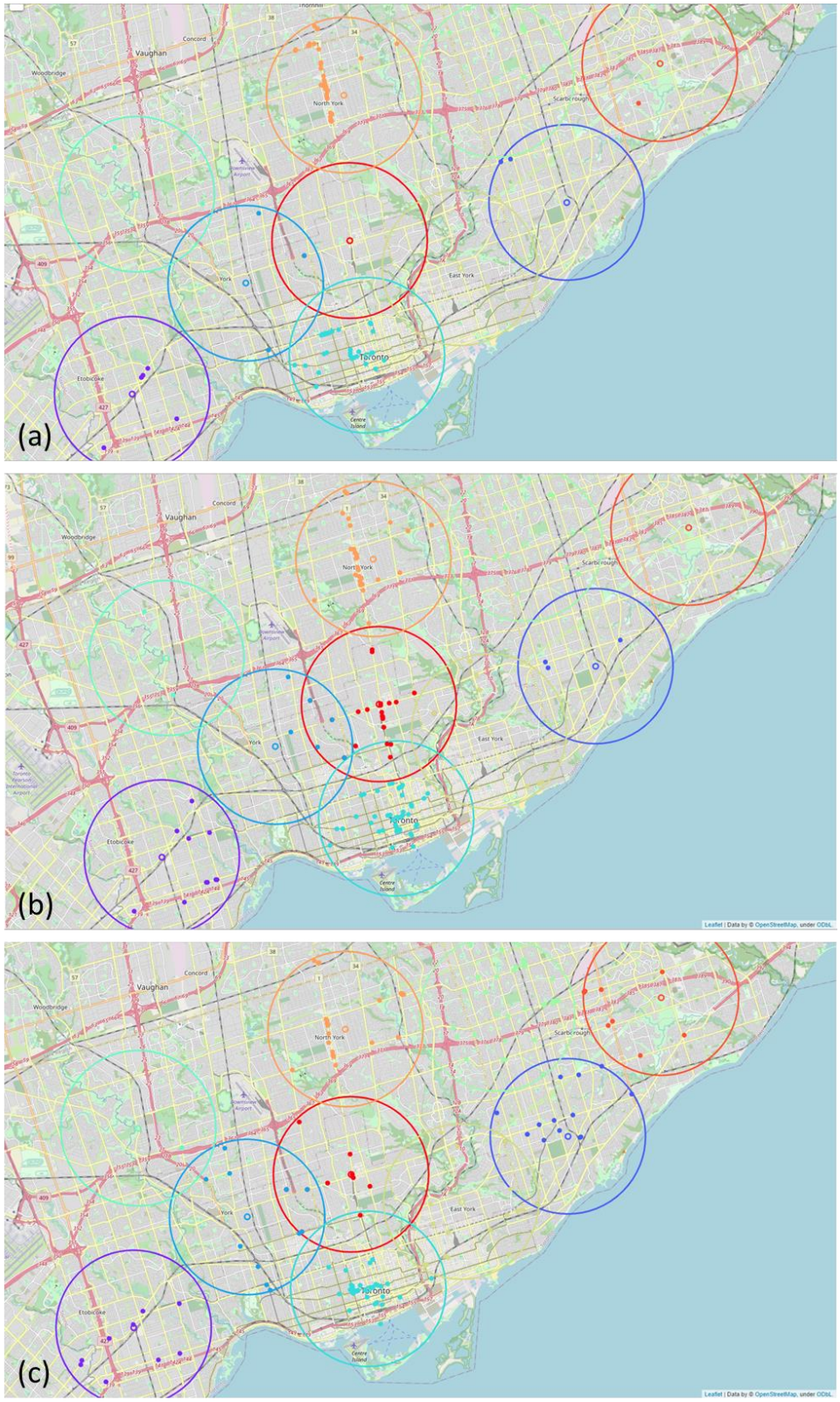

Figure 3 Restaurants distributions in Toronto: (a) Korean restaurants; (b) Japanese restaurants; (c) Chinese restaurants

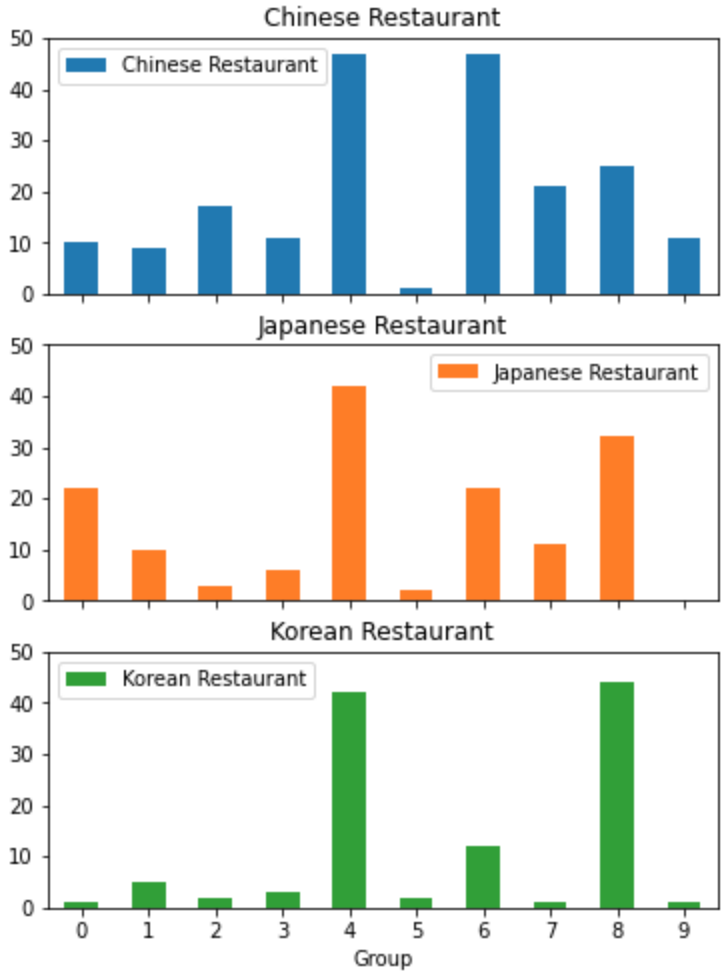

Figure 4 Bar plot of the number of Chinese, Japanese and Korean restaurants in the 10 groups of neighborhoods in Toronto

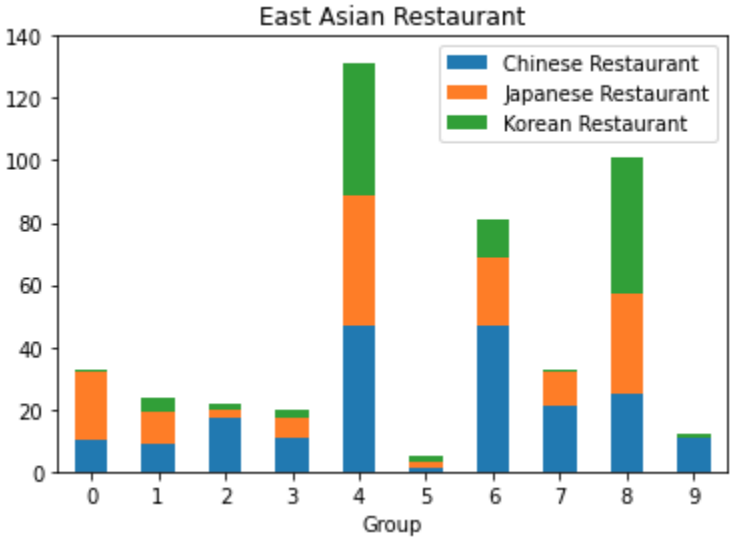

Figure 5 Bar plot of the number of East Asian Restaurants in the 10 groups of neighborhoods in Toronto

Conclusion

To sum up, we have used K-means to make 10 groups of neighborhoods from 103 regions of Toronto. And then we explored the number of Korean, Japanese, and Chinese restaurants within the 4000 meters radius of each center of those 10 groups and learned the distribution of the restaurants. We assumed there is a relation between number of restaurants and competition strength. We determined that Group 5 is weak in competition for Chinese restaurants, while Japanese restaurant density is low in Group 2, 5 and 9. Korean restaurants only have strong completions in Group 4, 6, and 8. We also treated them as a whole of East Asian Restaurants and found that there are not many of them in Group 5 and 9. Hence, we have identified the candidates of neighborhoods that are suitable to open a new Chinese, Japanese, Korean restaurants. The stakeholders can further investigate those neighborhoods to determine the final location to launch the new restaurant.