

Final Report---Investment of new restaurant in Toronto

1. Introduction

Canada is one of the world's top immigration destinations, with over 250,000 new immigrants arriving in Canada every year to take advantage of the country's strong economy and progressive immigration system. Canada Immigration benefits include having access to a safe and clean environment, a well-developed public transport system, healthcare, education, and good working conditions. Meanwhile, immigrants are a necessary component to achieve economic growth and keep taxpayer-funded systems such as pensions and health care stable and balanced.

As the immigration growing, more immigrants and local residents have desired to establish business or investment in Canada. Since the market is highly competitive in the highly developed city like Toronto, any new business venture or expansion need to be carefully analysed. And the insight derived from analysis will give a comprehensive understanding of the business environment, which helps for strategical targeting the market.

2. Business Problem

Toronto is a famous city for its diversity culture attracting global immigrants. To start a restaurant or food catering business, the targeted customers and suppliers are crucial to success for the marketing strategy. Moreover, the types of cuisines for the potential customers are also influenced by the immigrant history.

This report will first review the immigrant history to Canada, thus exploring the potential popular cuisines based on the immigrant's background. Then provide the new investor a comprehensive understanding of the current state of restaurant market in Toronto. And deliver guidance on where to locate the restaurant using the information obtained through the Foursquare API, along with the traffic information, in order to decide an appropriate location. Additionally, with accessibility of 'tips' and 'rating' of each restaurant by Foursquare API, the further improvement advice could be addressed to enhance the success of the new restaurant investment for the investor.

3. Data Collection

The data used in this report from multiple sources:

- 1) Canada immigration history, which includes the dataset of number of immigrants in the last decades and original countries of immigrants.

https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DV0101EN/labs/Data_Files/Canada.xlsx

- 2) List of postal codes of Canada, and the areas with postal codes beginning with M are located within the city of Toronto in the province of Ontario.

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

3) Geographical coordinates with latitude and longitude of each postal code

http://cocl.us/Geospatial_data

4) Foursquare API, which is a social location service that allows users to explore the details of a venue. In this report, Foursquare API is used to build up the datasets of the restaurants within the specified areas, along with the restaurants ratings in order to evaluate the popularity.

5) Information of 8 peak hours' vehicle and pedestrian volume

https://www.toronto.ca/ext/open_data/catalog/data_set_files/8hrVeh&PedVolume_6-Mar-2018.xlsx

In this study, investigation is conducted based on above dataset. This report will help a new investor to have a comprehensive understanding of the current state of restaurant market in Toronto, as well as future market outlook.

Firstly, by analysing the Canada immigration history, the potential popular cuisines based on the immigrant's background will be found. Secondly, the current most popular restaurants with their locations would be discovered by using Foursquare API and Geographical coordinates in Toronto. Then locations of the most popular restaurants would be visualized in the Toronto map. Moreover, the restaurants will be divided into several clusters according to traffic flows, in order to easily figure out which areas have more potential for a new restaurant. Or the areas with less competitive restaurants would therefore be selected. Finally, by comparing the traffic flow of visitors, the best locations for starting a new restaurant business would be proposed in this report.

4. Methodology

In this report, K-means clustering method is used to partition the areas of restaurants based on specified conditions (ratings or heavy traffic flow). K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. the K-means algorithm in data mining starts with a first group of randomly selected centroids, which are used as the beginning points for every cluster, and then performs iterative (repetitive) calculations to optimize the positions of the centroids. The way k-means algorithm works is as follows:

- a) Specify number of clusters K.
- b) Initialize centroids by first shuffling the dataset and then randomly selecting K data points for the centroids without replacement.
- c) Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.
- d) Compute the sum of the squared distance between data points and all centroids.
- e) Assign each data point to the closest cluster (centroid).
- f) Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.

In this report, the module of K-means is imported from sklearn.cluster library. And the results of partition the restaurants into four clusters based on traffic flow is plotted in Figure 1.

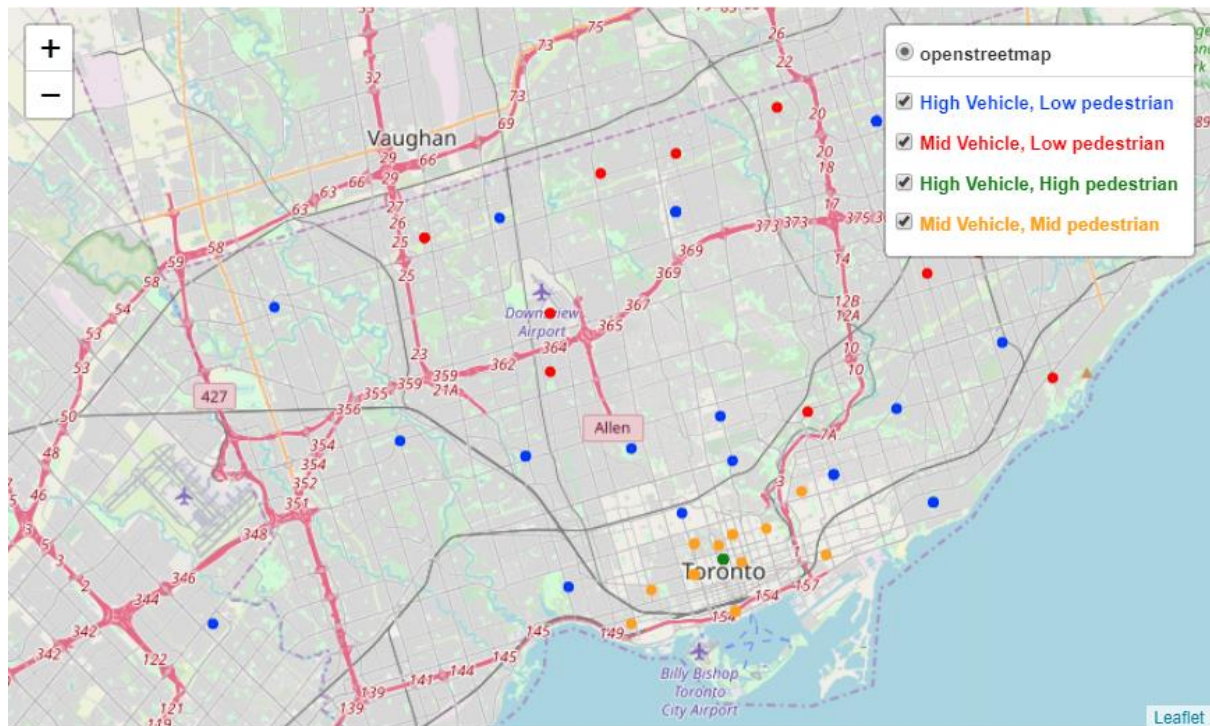


Figure 1 Four clusters of restaurants based on traffic flow by K-means

5. Results and Discussion

Canada is one of the world's top immigration destinations, attracting new immigrants with pleasant environment, safe neighbourhood, well-developed public transport system, healthcare, education, and good working conditions. As given in Figure 2, the immigrants from Asia take major proportions in the total immigrations.

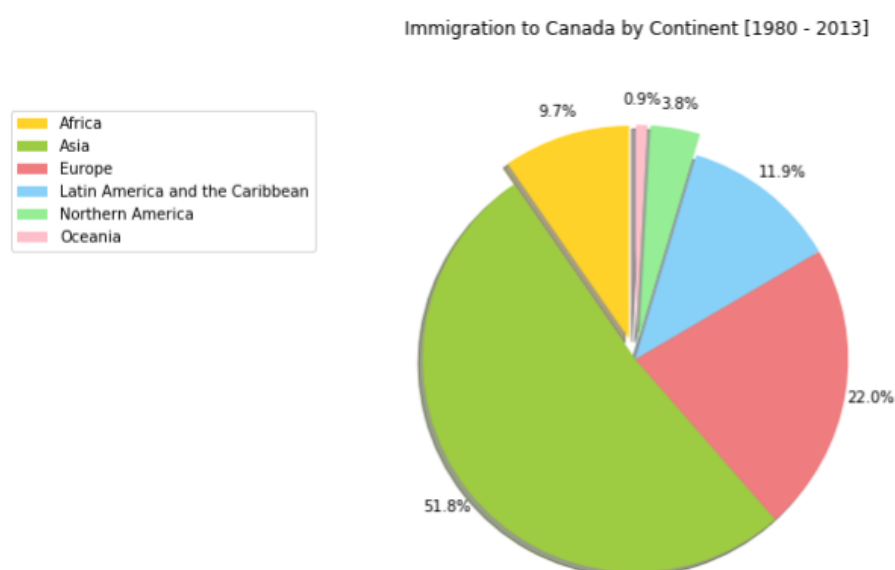


Figure 2 Immigration to Canada by continents

The above dataset contains the statistical immigration history from 195 countries over six continents. After refining and cleaning the dataset, it is found that the immigrations from India, China and United Kingdom of Great Britain and Northern Ireland are the top 3 over the total immigrations, as given in Figure 3. And the trending of immigrations from China and India to Canada is still increasing since year of 1995 by the statistics, as shown in Figure 4. Therefore, it is reasonable to infer that the demands of Indian and Chinese restaurants are high in Canada due to the immigrant's background. Moreover, the cuisines of Indian and Chinese food will be popular because of the density of populations of Indian and Chinese immigrants in Canada.

	Continent	1980s	1990s	2000s	Total
Country					
India	Asia	82154	180395	303591	658817
China	Asia	32003	161528	340385	625833
United Kingdom of Great Britain and Northern Ireland	Europe	179171	261966	83413	545673
Philippines	Asia	60764	138482	172904	481847
Pakistan	Asia	10591	65302	127598	228997
United States of America	Northern America	76824	56915	75173	232621
Iran (Islamic Republic of)	Asia	21477	54871	65794	164632
Sri Lanka	Asia	14796	70421	49678	145964
Republic of Korea	Asia	16259	38189	68183	138072
Poland	Europe	57602	64864	13629	138389

Figure 3 Top 10 countries of immigration to Canada

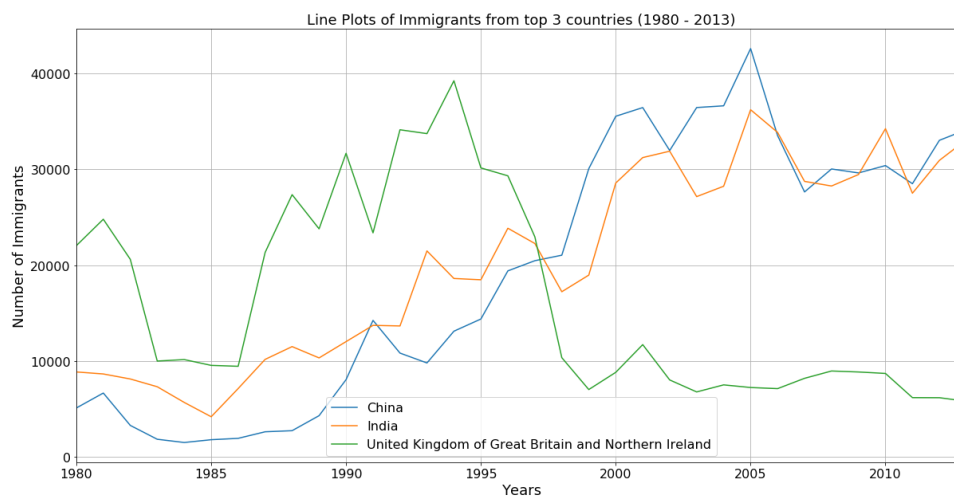


Figure 4 Number of immigrants from China, India and UK over 1980-2013

Based on above analysis, it was assumed the trending of cousins of Indian and Chinese food would be more popular in Canada, therefore a business of starting an Indian or Chinese restaurant would be profitable for an investor.

Using Foursquare API, the information of neighbourhood restaurants in Toronto could be excavated into forms of datasets. By sorting and counting, the number of different cuisines

of restaurants is summarized in Figure 5, and it is noted that the numbers of Chinese and Indian restaurants are on top, which also implies the popularity of Chinese and Indian cuisines in Toronto.

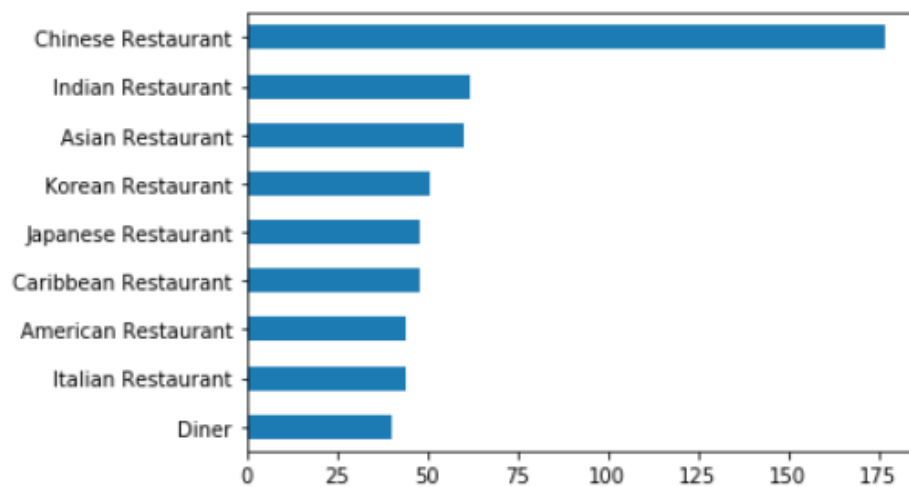


Figure 5 Number of different categories of restaurant in Toronto

Therefore, this report will focus on proposing a business prospect on starting a Chinese restaurant in Toronto. Firstly, the numbers of Chinese restaurants are statistical counted in each borough. It is clear that the areas of Downtown Toronto and Scarborough have many Chinese restaurants already, following by areas of North York. And in the other boroughs, there are less available Chinese restaurants.

Borough	Number of Chinese Restaurants
Downtown Toronto	34
Scarborough	32
North York	13
East York	5
East Toronto	4
Central Toronto	3
West Toronto	3
York	3
Etobicoke	2
Mississauga	1

Figure 6 Number of Chinese restaurants in Toronto borough

Moreover, with limited available data of the rating of each restaurant, the top 10 rated Chinese restaurants are listed in Figure 7, along with number of tips. It is worthy noted that

the borough of Scarborough has more high rated Chinese restaurants, that verifying the relative mature and developed catering industry of Chinese restaurants so far.

	Borough	name	rating	numbertips
1	Scarborough	The Royal Chinese Restaurant 避風塘小炒	7.7	11.0
2	Scarborough	Kim Kim restaurant	7.4	29.0
3	Scarborough	Super Hakka Restaurant	7.1	5.0
4	Scarborough	Phoenix Restaurant 金鳳餐廳	6.9	4.0
5	North York	Szechuan Gourmet Restaurant	6.9	8.0
6	Scarborough	Lucky Chinese (Haka) Restaurant	6.7	5.0
7	North York	Lee Town Restaurant	6.5	6.0
8	Scarborough	Perfect Chinese Restaurant 雅境海鮮酒家	6.3	42.0
9	Scarborough	Lucky Hakka	6.0	9.0
10	Scarborough	Fortune Restaurant	5.8	1.0

Figure 7 Chinese Restaurants in Toronto with top rating

With the help of Foursquare API, it is achievable to plot the restaurants map with returned information of a venue using popup menu. As visualized in Figure 8, it is easy to search the venue of a restaurant by clicking the popup menu, with returned information like restaurant's name and rating. Therefore, a quick and direct visualization is able to be provided for the investors for accurate information.

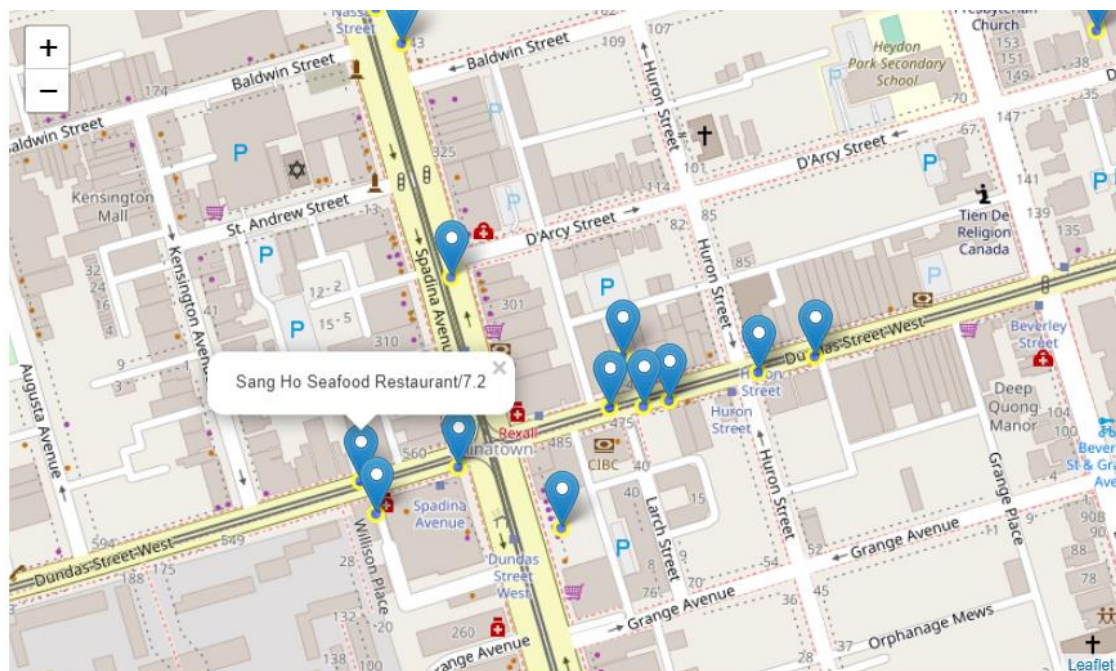


Figure 8 Map of Chinese Restaurants in Toronto with popup menu showing name and rating

Consequently, four clusters of Chinese restaurants based on traffic flow by K-means method are displayed in Figure 9. The four clusters are described as High vehicle with low pedestrian; Middle vehicle with low pedestrian; High vehicle with high pedestrian and Middle vehicle with middle pedestrian. It is found that in Downtown Toronto, more Chinese restaurants are available because of more coming visitors. Elsewhere, Chinese restaurants are scattered located, but many are accessible by vehicle.

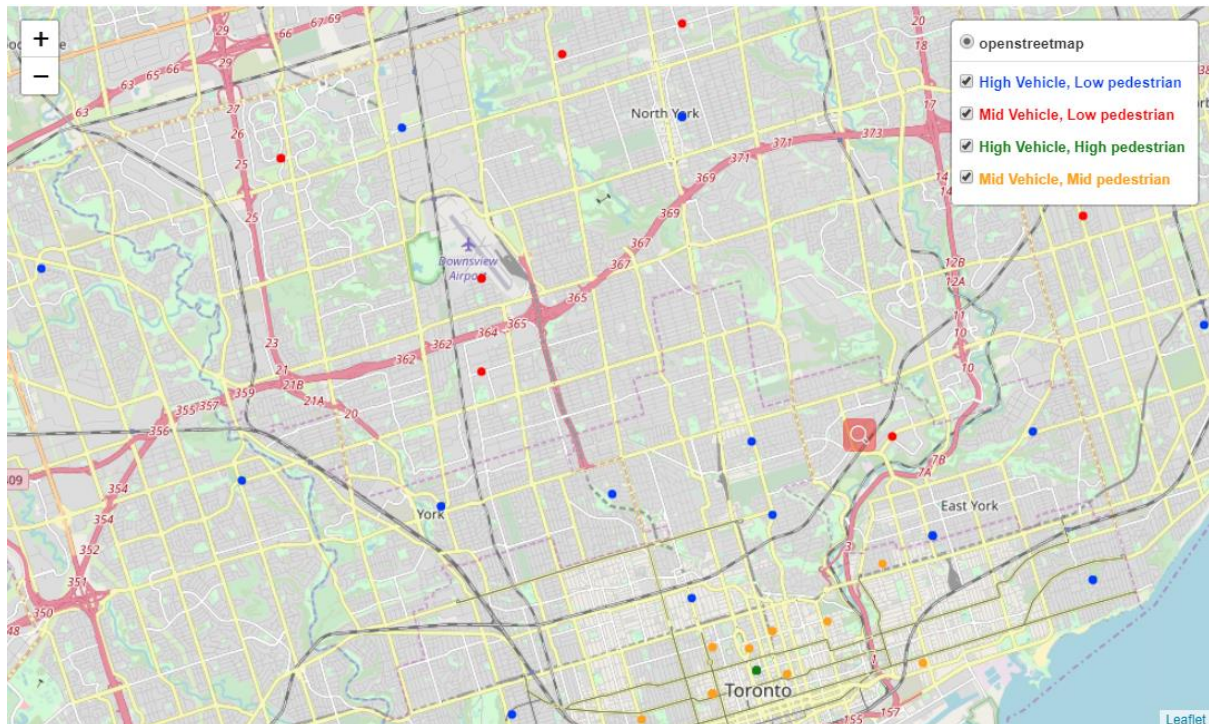


Figure 9 Four clusters of Chinese restaurants based on traffic flow by K-means method

6. Conclusion

The objective of this report is to provide a comprehensive understanding of the current restaurant business market of Toronto for a new investor.

Firstly, by analysing the Canada immigration history, the potential popular cuisines based on the immigrant's background are found to be Chinese and Indian food.

Secondly, the current most popular restaurants are tabulated using Foursquare API in different borough of Toronto, and it is found that the borough of Scarborough has more high rated Chinese restaurants, which implies the relative mature and developed catering industry of Chinese restaurants so far. Moreover, combined with analysis of traffic flow by K-means method, more Chinese restaurants are available in Downtown Toronto because of more coming visitors.

Finally, considering the market saturation and competition with current competitor, it is suggested to start a new Chinese restaurant business in the borough of East York or periphery of Downtown Toronto.