Capstone Project Report

Ideal location for Opening an Asian Supermarket in the Toronto

1. Introduction

1.1 Background

One of my clients is the owner of an Asian product wholesale company and would like to expand his business and increase diversity by opening up a retail supermarket in Toronto. This is his first supermarket and is planning to cover various Asian products including frozen foods, grocery, meat, produce, seafood, snacks, liquor, dairy, etc. This Asian supermarket will be his first step to involve in the retail industry of Canada and thus he wants it to be successful.

1.2 Business Problem

Since my client has been running an Asian food wholesale company for more than 10 years and owns a distribution center located in Hamilton, there is no concern on the good supplying and logistics. However, my client did not have business in the Toronto in the past, so basically he knows very less information about this area. As we both agreed, an ideal location is one of the most important factor for a consumer-oriented retail business to be success. My client needs us to provide the analysis helps find out the best location for his first supermarket and avoiding too many competitors surrounding.

1.3 Interest

The Chairman (Business Owner), CEO (Chief Executive Officer), and COO (Chief Operation Officer) from the Asian food wholesale company would be very interested in the reasons for recommending the location as a best candidate and would like to know how the surroundings of this location look like.

2. Data and Data Acquisition

For this analysis, there are two datasets are being used:

- Toronto neighborhood data
- Foursquare location Data

The neighborhood data in Toronto is obtained from Wikipedia page “List of postal codes of Canada: M”. Foursquare location data is calling by using the coordinates which are explored from the dataset of Toronto neighborhood in Python. The Foursquare API returns venues with necessary information (i.e. information regarding the existing food supermarkets in Toronto) for analysis and supports the final recommendation.

3. Methodology

First of all, it's necessary to make sure having the clean data before doing any analysis. After loading the list from Wikipedia page regarding the Toronto, there are several issues found from its neighborhood dataset, such as below rows with Borough column of "Not assigned", too many data are not belonging to the target city – Toronto, and duplicated data may exist in the dataset.

In order to have accurate dataset, multiple tasks need to be completed. Rows with no Borough are dropped, only keep the target city Toronto, and duplicates are removed based on the name of neighborhood. After the above steps are performed, the geocoder is used to capture the coordinates for neighborhoods accordingly.

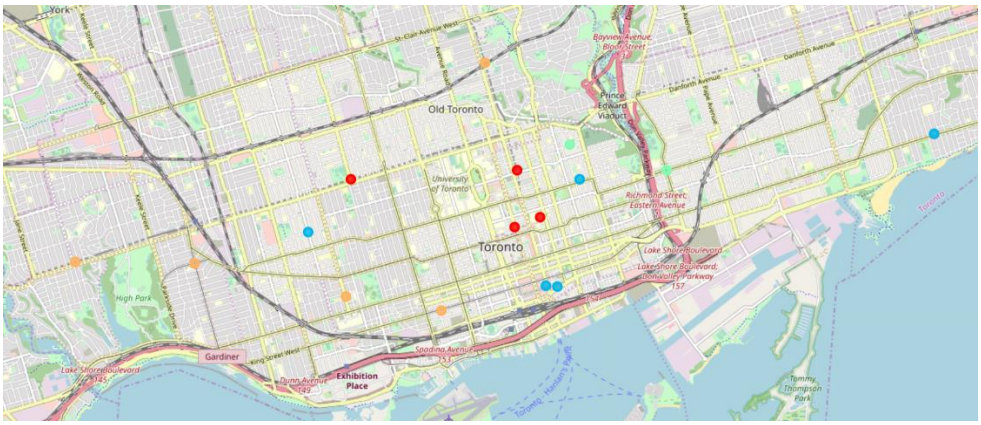
By calling the venue and venue category from the Foursquare API based on the coordinates of the neighborhoods from the above list, it returns a data frame which now has been captured all the necessary information for analysis. Venue category is the key indicator and now the data frame totally includes 223 unique categories. Since the supermarket is planning to have various products, based on the names of venues categories, filter them to the relevant categories only by string the key words, such as "Market", "Supermarket", "Liquor", "Bakery", "Japanese", "Thai", "Asian", "Chinese", etc.

	Neighborhood	Asian Restaurant	Bakery	Bus Line	Bus Stop	Chinese Restaurant	Deli / Bodega	Dessert Shop	Farmers Market	Fish Market	Food & Drink Shop	Gourmet Shop	Grocery Store	Japanese Restaurant	Liquor Store	Malay Restaurant	Market	Pastry Shop	Ramen Restaurant	Supermarket	Thai Restaurant	Vietnamese Restaurant
0	Berczy Park	0.000000	0.230769	0.000000	0.000000	0.000000	0.076923	0.000000	0.153846	0.076923	0.000000	0.076923	0.076923	0.230769	0.000000	0.0	0.000000	0.0	0.000000	0.000000	0.076923	0.000000
1	Cabbagetown, St. James Town	0.000000	0.200000	0.000000	0.000000	0.100000	0.100000	0.000000	0.000000	0.000000	0.000000	0.000000	0.100000	0.200000	0.100000	0.0	0.100000	0.0	0.000000	0.000000	0.100000	0.000000
2	Central Bay Street	0.000000	0.250000	0.000000	0.000000	0.250000	0.000000	0.125000	0.000000	0.000000	0.000000	0.000000	0.000000	0.125000	0.000000	0.0	0.000000	0.0	0.125000	0.000000	0.125000	0.000000
3	Christie	0.000000	0.000000	0.000000	0.000000	0.111111	0.000000	0.222222	0.000000	0.000000	0.000000	0.000000	0.222222	0.222222	0.000000	0.0	0.000000	0.0	0.111111	0.000000	0.000000	0.111111
4	Church and Wellesley	0.000000	0.000000	0.000000	0.000000	0.090909	0.000000	0.090909	0.000000	0.000000	0.090909	0.000000	0.090909	0.363636	0.000000	0.0	0.000000	0.0	0.181818	0.000000	0.090909	0.000000

Constructing data frame to present the mean of frequency of occurrence of each relevant categories based on each Neighborhoods. Then, applying K means clustering method according to the values of top venues and plotting the map accordingly.

4. Results

The clusters are showed in below image:



Cluster 0: Red Points – Areas with restaurants that serving many Asian element foods (ex: Japanese, Chinese, Thai, etc.) and has bus lines/stops surrounding. But some small businesses like bakery and liquor stores nearby.

Cluster 1: Purple Points – Areas with dessert shops, deli and farmers markets and with a few of Asian food restaurants. But since the location is far from the main business district, it's not included in the above snapshot and thus will not be considered further.

Cluster 2: Blue Points – Areas are consist of many food related businesses such as bakery, pastry shop, and various makers. However, many Japanese restaurants are located in these areas, and they has convenient transportation.

Cluster 3: Green Points – Areas with some grocery stores and Vietnamese restaurants. Also, there are bakery stores are around there. Bus stop is located at St. James Town.

Cluster 4: Orange Points – Areas with many gourmet shops and Asian restaurants. Bakery stores are popular and a few retail shops are opening in these areas.

5. Discussion

An ideal location for a retail business should avoid too many competitors nearby, and it should close to the properties that may benefit to the business. Therefore, the location should far from other competitors such as markets. If there is no such spot available, look for a place with fewer competitors are surrounding. Also, the location should be close to Asian food restaurants as they may become potential buyers in the future. Bus lines/stops may be another element that needs to be considered. Older people may choose to take the bus instead of driving to a destination. So having more convenient transportation, more customers may come to the supermarket. Therefore, after considering the above mentioned factors and gathering the results, it's recommend to pick Central Bay Street for opening up the first supermarket as there are many Asian restaurants but only has few markets and shops as competitor. Bus lines/stops are around the street.

6. Conclusion

This report provides the analysis on the neighborhoods in Toronto and recommends a location for opening up an Asian supermarket based on that. It provides proof for the recommendation which should address the most questions or concerns from the management team of the Asian food wholesale company regarding the location selection, and thus build a bridge toward success on their first retail supermarket.