Exploring Frozen Dessert Shop in Scarborough, Toronto using Foursquare

Meenu Shanmugam
May 02, 2020

1.Introduction:

Scarborough, Toronto, is the most diverse and multicultural administrative division of Toronto, Canada. In addition, this multicultural city is full of parkland, safe and clean. Steadied by strong, highly regulated banks and buoyed by an educated workforce, this city provides a lot of business opportunities and a business-friendly environment. With a low net debt-to-GDP ratio, it has attracted many different players into the market. Hence, any new business venture or expansion is going to be highly competitive and needs to be analyzed carefully. The insights derived from analysis will give a good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk and the Return on Investment will be reasonable.

Problem Definition

Scarborough is known for its excellent cuisine and its food culture includes an array of international cuisines influenced by the city's immigrant history. Frozen dessert shops have become so popular in Toronto now it seems that there's one on every corner, not only in major cities but also in smaller cities. Starting a frozen dessert shop will be a good business opportunity, but you would like to tell yourself apart from others to enjoy long-term success.

If you intend a frozen dessert shop which will demand higher prices, you need to focus in neighborhoods and outlets that already attract sophisticated clients. If you intend an inexpensive buffet restaurant, point to the masses searching for affordable high-traffic locations with large shopping centers and other local points of interest.

For this analysis, I plan to focus on the Scarborough area, as my father intends to open his business in that borough. We typically define potential neighborhoods based on the number of frozen dessert Shops which are operating right in each neighborhood. Scarborough has full potential but is also a really challenging district to open a business due to high competition. New frozen dessert Shop should be open in a vicinity that's an inadequate neighborhood, so that the shop can attract more customers. Therefore, this analysis is critical to confirm that we would attract enough customers and not very near to other frozen dessert Shops.

2.Data

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

We will use Scarborough dataset which we scrapped from wikipedia on Week 3. This dataset contains latitude and longitude, zip codes.

Foursquare API Data:

We need data about different venues in all neighborhoods of Scarborough. "Foursquare" locational information helps to gain that information. Foursquare is a location technology platform which provides information about events and venues for a given area of interest. This location data consists of data points about venue names, locations, menus and even photos. As all the required information for this study can be obtained via API, foursquare location platform will be used as the primary data source.

Once the list of neighborhoods is finalized, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meters.

The retrieved data from Foursquare contained venue information of the longitude and latitude of the zip code (within the chosen 100 meters distance). The information obtained per venue as follows:

1. Neighborhood

2. Neighborhood Latitude

3. Neighborhood Longitude

4. Venue

5. Venue Latitude

6. Venue Longitude

7. Venue Category

3.Methodology

As a first step, I retrieve the venues in Scarborough from Foursquare. I extract the location data from the Foursquare API for all venues up to a distance of 5000 meter radius from the center of Scarborough. Using this, I fetch the venue information including coordinates.

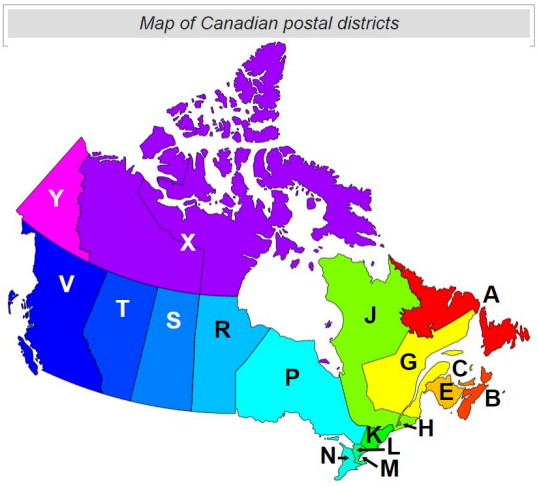
Using data cleaning, the dataset from the API will be analysed based on the Neighborhood,venue names, latitude, and longitude values. The final data will include the venue name, category, address, latitude and longitude.

Using this dataset, I begin by analyzing the top venue types that exist in Scarborough. I will then explore the venues on maps. This will allow us to better understand the location of various venues and the places where many venues co-exist and identify places worth for business. The venues will be plotted using proper color coding such that a simple glance at the map would reveal the location of the venues as well as give information about them. I aim to identify places which can be recommended based on their demand. I'll also use the "**clustering ML approach"** on the venues and see if we can draw meaningful information out of what kind of venues exist in Scarborough.

As a final step, I will analyse these plots and try to draw conclusions on what places can be recommended to start the business. I'll discuss my findings and any inferences I can draw.

3.1 Exploratory Data Analysis:

In this project, I used the basic methodology taught in Week 3 lab. The first step is to scrape the Canada postal codes data from Wikipedia (List of postal codes of Canada: M).



This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M are located within the city of Toronto in the province of Ontario. Only the first three characters are listed, corresponding to the Forward Sortation Area.

Toronto - 103 FSAs

Postalcode	Borough	Neighborhood
M1A	Not assigned	
M2A	Not assigned	
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park / Harbourfront

Using this data I built a dataframe of the postal code of each neighborhood along with the borough name and neighborhood name.Once I cleaned the data I got the latitude and the longitude coordinates of each neighborhood¶corresponding to each postalcode from Geospatial_data.The following are the neighborhoods in Toronto.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
17	M2H	North York	Hillcrest Village	43.803762	-79.363452
18	M2J	North York	Fairview , Henry Farm , Oriole	43.778517	-79.346556
19	M2K	North York	Bayview Village	43.786947	-79.385975
20	M2L	North York	York Mills , Silver Hills	43.757490	-79.374714
21	M2M	North York	Willowdale , Newtonbrook	43.789053	-79.408493
22	M2N	North York	Willowdale	43.770120	-79.408493

And the total number of neighborhoods in Toronto is as follows:

```
#Quickly examine the resulting dataframe
print('The dataframe has {} boroughs and {} neighborhoods.'.format(
        len(toronto_geo_df['Borough'].unique()),
        toronto_geo_df.shape[0]
    )
)
```

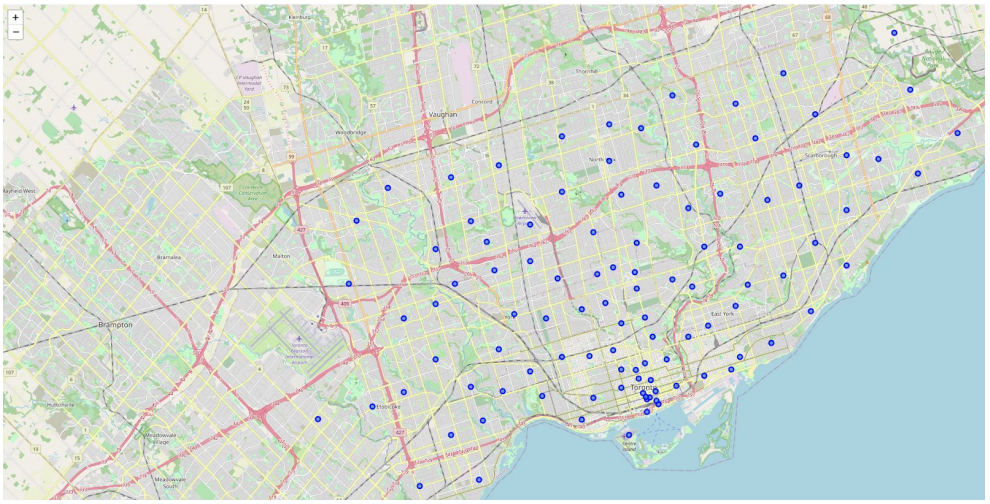
```
The dataframe has 10 boroughs and 103 neighborhoods.
```

Using the above information Wikipedia I Created a map of Toronto with neighborhoods superimposed on top.The coordinates of Toronto is shown below.

```
The geograpical coordinate of Toronto City are 43.6534817, -79.3839347.
```

The map shows Toronto with neighborhoods which are blue dots.There are nearly 10 boroughs and 103 neighborhoods.

Create a map of Toronto with neighborhoods superimposed on top



Toronto covers an area of 630 square kilometres (243 sq mi),[91] with a maximum north–south distance of 21 kilometres (13 mi) and a maximum east–west distance of 43 km (27 mi). It has a 46-kilometre (29 mi) long waterfront shoreline, on the northwestern shore of Lake Ontario.

The next step is to extract the Scarborough data from the Toronto data.There neighborhoods with their coordinates shown below.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Malvern , Rouge	43.806686	-79.194353
1	M1C	Scarborough	Rouge Hill , Port Union , Highland Creek	43.784535	-79.160497
2	M1E	Scarborough	Guildwood , Morningside , West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

There are 17 neighborhoods for Scarborough which is shown below.

Explore and cluster the neighborhoods in Scarborough

```
]: #Quickly examine the resulting dataframe
print('The dataframe has {} boroughs and {} neighborhoods.'.format(
        len(scarborough_data['Borough'].unique()),
        scarborough_data.shape[0]
    )
)
```

```
The dataframe has 1 boroughs and 17 neighborhoods.
```

The coordinates of Scarborough is shown below.

```
The geograpical coordinate of Scarborough are 54.2820009, -0.4011868.
```

Create a map of Scarborough with neighborhoods superimposed on top

The next step is to collect the Scarborough's **Frozen Dessert Shop venue** information from Foursquare. Snap-to-Place that assigns users to specific locations, real time access to places, and Geo-tag is supported by Foursquare Places API. For measurement and analysis, additionally Foursquare allows developers to build audience segments. JSON is the preferred response format.
Using the Foursquare API I retrieved the Venue information regarding **Frozen Dessert Shop**.The category id for the **Frozen Dessert Shop** is 4bf58dd8d48988d110941735.The Venue information is as follows.

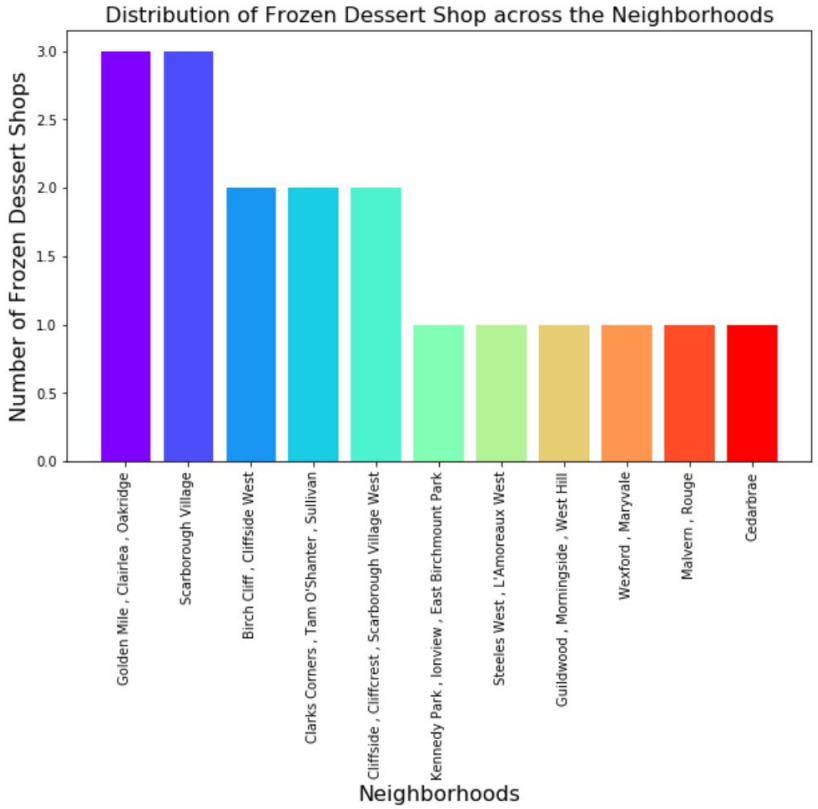
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Malvern , Rouge	43.806686	-79.194353	McDonald's	43.798261	-79.201449	Fast Food Restaurant
1	Guildwood , Morningside , West Hill	43.763573	-79.188711	McDonald's	43.768334	-79.188288	Fast Food Restaurant
2	Cedarbrae	43.773136	-79.239476	McDonald's	43.780955	-79.234213	Fast Food Restaurant
3	Scarborough Village	43.744734	-79.239476	Dairy Queen	43.739580	-79.236991	Ice Cream Shop
4	Scarborough Village	43.744734	-79.239476	Dairy Queen	43.739506	-79.236894	Ice Cream Shop
5	Scarborough Village	43.744734	-79.239476	McDonald's	43.738985	-79.239869	Fast Food Restaurant
6	Kennedy Park , Ionview , East Birchmount Park	43.727929	-79.262029	McDonald's	43.724591	-79.253462	Fast Food Restaurant
7	Golden Mile , Clairlea , Oakridge	43.711112	-79.284577	Dairy Queen	43.710378	-79.290701	Ice Cream Shop
8	Golden Mile , Clairlea , Oakridge	43.711112	-79.284577	Dairy Queen	43.710296	-79.291126	Ice Cream Shop
9	Golden Mile , Clairlea , Oakridge	43.711112	-79.284577	McDonald's	43.709284	-79.295245	Fast Food Restaurant

Foursquare returned 18 venues of Scarborough

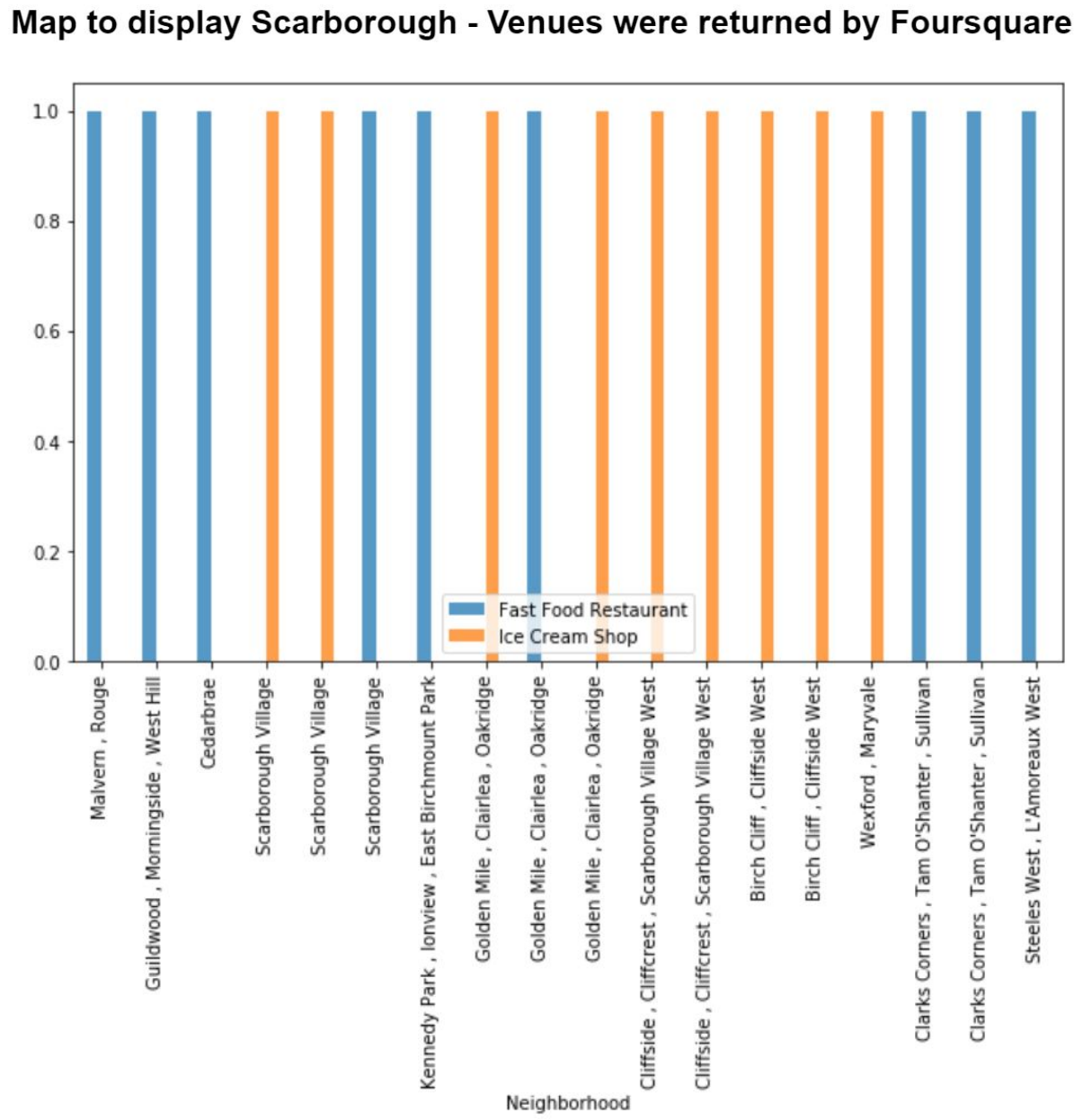
```
print('{} venues were returned by Foursquare.'.format(toronto_venues_iceshop.shape[0]))
```

```
18 venues were returned by Foursquare.
```

The following graph shows the distribution of Scarborough's **Frozen Dessert Shop** of each neighborhood .The graph shows 3 **Frozen Dessert Shop** for the first two neighborhoods and two for the next two neighborhoods and one for others.



The graph shows the Scarborough's **Frozen Dessert Shop venue** information from Foursquare. If we analyze the venue information of each neighborhood the graph shows the Fast Food Restaurant and Ice Cream Shop.The graph shows the neighborhood and the corresponding venue details.



4.Results Section

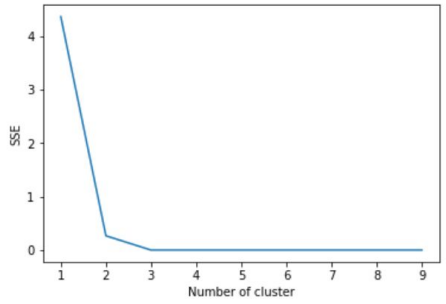
Finally, I cluster all the venues based on their location and more to identify similar venues and the relationship amongst them. I used KMeans clustering and decided to cluster the venues into two separate groups. Each neighborhood along with the top 2 most common venues is shown below. Mostly the venues are Ice Cream Shop and Fast food Restaurant.

Using K-Means Clustering Approach | Most Common Venue

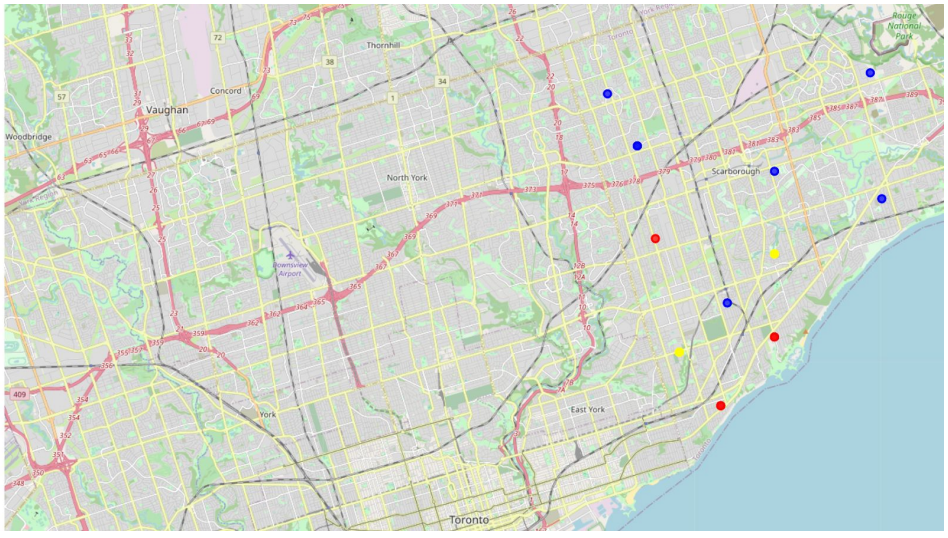
	Neighborhood	1st Most Common Venue	2nd Most Common Venue
0	Birch Cliff , Cliffside West	Ice Cream Shop	Fast Food Restaurant
1	Cedarbrae	Fast Food Restaurant	Ice Cream Shop
2	Clarks Corners , Tam O'Shanter , Sullivan	Fast Food Restaurant	Ice Cream Shop
3	Cliffside , Cliffcrest , Scarborough Village West	Ice Cream Shop	Fast Food Restaurant
4	Golden Mile , Clairlea , Oakridge	Ice Cream Shop	Fast Food Restaurant
5	Guildwood , Morningside , West Hill	Fast Food Restaurant	Ice Cream Shop
6	Kennedy Park , Ionview , East Birchmount Park	Fast Food Restaurant	Ice Cream Shop
7	Malvern , Rouge	Fast Food Restaurant	Ice Cream Shop
8	Scarborough Village	Ice Cream Shop	Fast Food Restaurant
9	Steeles West , L'Amoreaux West	Fast Food Restaurant	Ice Cream Shop
10	Wexford , Maryvale	Ice Cream Shop	Fast Food Restaurant

Using Clustering K-means ML approaches the value of K=3.Using Cluster method analysis the following are the neighborhood and the corresponding venue information.

Find optimal k for the clustering process



Most Common Venues near Neighborhood | Using Clustering



Discussion

After collecting data from the Foursquare, we got a list of 18 different venues. We also had to inspect their latitude and longitude values as well as their names.

We identified that from the total set of venues, the majority of them were Fast Food Restaurant and Ice Cream Shop. A visitor who loves Ice Cream dessert would surely benefit from coming to Scarborough.

When we plot these venues on the map, we discover that there are clusters of venues around Warden Avenue, Morningside Park and Kingston Subdivision Transit. Out of these clusters, Kingston Subdivision Transit has more number of venues.

On plotting the venues based on their location on the map, and deciding to use “Hotelling's Model of Spatial Competition” then opening in Kingston Subdivision Transit makes sense. However, if we need to avoid competition then less dense Morningside Park clusters should be tried. Warden Avenue cluster neither has a lot of competitors nor less dense.

Finally, through clusters we identified that there are many venues which are relatively located closer and some are located apart. Anybody can use this information to decide where a Frozen dessert shop can be opened.

5.Conclusion

The purpose of this project was to explore the places where a Frozen Dessert shop can be started. The venues have been identified using Foursquare API and have been plotted on the map. The map reveals that there are three major areas available to start the business: Warden Avenue, Morningside Park and Kingston Subdivision Transit. Based on the investor's preferences, he/she can choose amongst the three places.

This project only considers the density of the venues as criteria to start the business. However, this can be further extended to consider the price of the frozen dessert and user ratings.

References:

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
https://en.wikipedia.org/wiki/Postal_codes_in_Canada#Forward_sortation_areas