BEST LOCATION TO OPEN A HALAL RESTAURANT IN TORONTO

1. Introduction

Background

Islam is the second-largest religion in Canada. The most Muslim populated areas in Canada are in the Greater Toronto Area (GTA). The number of Muslim is increasing geometrically and according to statistics, the size of the Muslim population in Canada may cross 15 per cent of the total population by 2030. This also means that the demand for halal food by Muslim will increase as well. Thus, offering halal options in GTA is a pragmatic business decision for anyone in the food industry.

Business Problem

A client from Turkey wants to open his chain restaurant in Toronto. The restaurant will focus on serving halal food option to cater to Muslims around a neighbourhood that lacks this service. Setting up a business can be challenging at a location that is foreign to a business owner. Therefore, the client requires assistance to find the best borough/neighbourhood in Toronto to open a Halal Restaurant.

The challenge to find a strategic neighbourhood is to ensure the following criteria are taken into consideration as well:

1. Where the targeted customer is most clustered at. (in this case, we will choose Mosque venue as the key factor to assume the target market cluster for a halal restaurant).
2. Little to no competition would help to reduce risk and gain a reasonable return on investment.
3. The availability of suppliers such as farmers market, wholesale market to purchase fresh ingredients to ensure quality and cost-effectiveness.

Target Audience

The success criteria of this project will be the best recommendation of the Borough/Neighborhood choice. The insight from this project analysis will be beneficial for audiences such as follows:

1. Business owners focusing on the halal industry.
2. Muslim who plans to migrate to Toronto.
3. Muslim students from oversea who chooses to continue their study in Toronto.
4. Muslim traveller who looks for accommodation around a Muslim friendly neighbourhood.
5. Anyone who wish to eat halal food option.

2. Data

For this project we need the following data:

1. Toronto dataset that contains list of Boroughs, Neighborhoods along with their latitude and longitude.

    - Data source:https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&oldid=945633050
    - Data usage: To explore various Boroughs and Neighborhoods of Toronto.

2. Mosques in each Boroughs of Toronto.

    - Data source: Foursquare API
    - Data usage: Identifying Mosques location can help justify where most Muslims are clustered at, as Muslims frequently gather at or visit a mosque.

3. Halal restaurants nearby the mosques.

    - Data source: Foursquare API
    - Data usage: Identifying competitions and which Borough/Neighborhood lacking halal service option.

4. Farmers Markets, Supermarkets, Fish Markets, etc

    - Data source: Foursquare API
    - Data usage: To identify which location will be the strategically near to get fresh supplies and other stocks.

3. Methodology

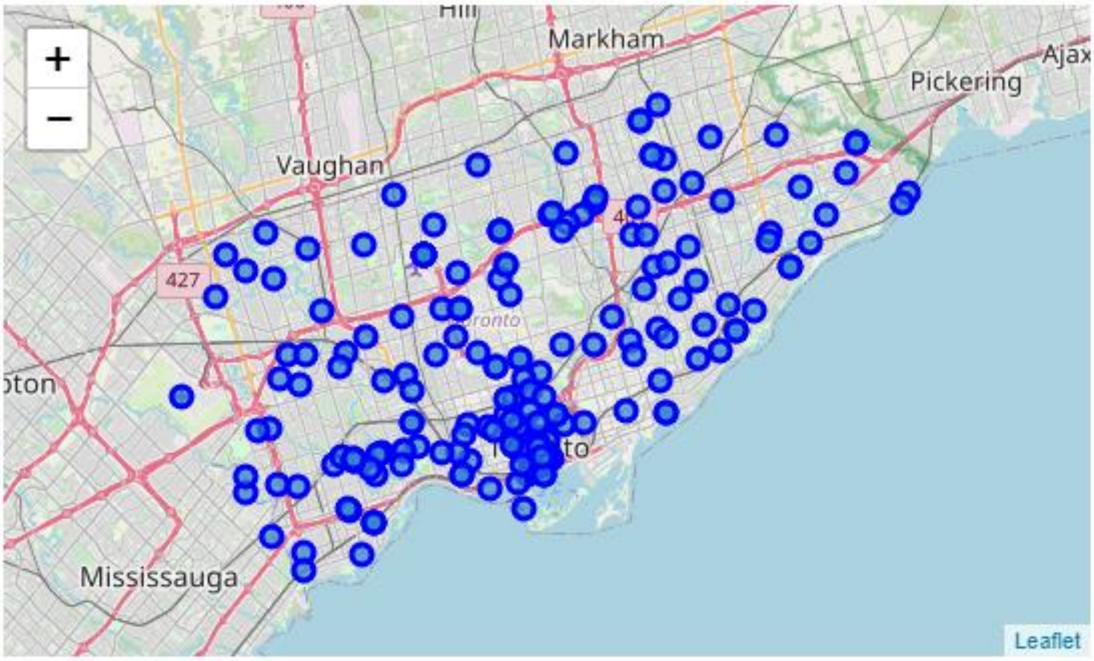
The methods used to get the data required for this project are as follows:

1. Toronto Data

   Firstly, scraped the list of Toronto postal code from Wikipedia using **BeautifulSoup**. Then Clean the data by dropping rows where Borough columns is 'Not assigned'. Lastly use **geopy** to get the latitude and longitude of each Neighbourhood. This data will be used to explore the venues like mosque, halal restaurant and markets around each neighbourhood leveraging their latitude and longitude data. Through analysis, there are a total of **9 Boroughs** and **201 Neighbourhoods** in the Toronto dataframe (**toronto_data**).

	Borough	Neighborhood	Latitude	Longitude
0	North York	Parkwoods	43.758800	-79.320197
1	North York	Victoria Village	43.732658	-79.311189
2	Downtown Toronto	Harbourfront	43.640080	-79.380150
3	North York	Lawrence Heights	43.722778	-79.450933
4	North York	Lawrence Manor	43.722079	-79.437507

   The curated dataframe is then used to visualise the map of Toronto with its neighbourhoods superimposed on top using python **folium** library. Below is the map generated using folium library.



2. Nearby Mosque, Halal Restaurant and Market Data (Nearby Venue Data)

The **Foursquare API** is used to explore nearby mosque, halal restaurant and market around the 201 neighbourhoods and segment them. A '**getNearbyVenues**' function is created. This function loops through all 201 neighbourhood of Toronto and create an API GET request URL to search each location within a radius = 1000 meters and return nearby venues with a limit = 500 venues maximum. Specific category ID is also defined in the GET request URL as the following:

- 4bf58dd8d48988d138941735 for Mosque.
- 52e81612bcbc57f1066b79ff for Halal Restaurant.
- 4bf58dd8d48988d1fa941735 for Farmers Market.
- 4bf58dd8d48988d10e951735 for Fish Market.
- 52f2ab2ebcbc57f1066b8b1c for Fruit & Vegetable Store.
- 52f2ab2ebcbc57f1066b8b46 for Supermarket.

Next, the GET request is made to Foursquare API and only relevant information for each nearby venue is extracted from it. The data is then appended to a python 'list'. Lastly the python 'list' is unfolded to append it to dataframe being returned by the function. Below is the returned dataframe with **1616 venues** returned from Foursquare API:

```
print(tor_venues.shape)
tor_venues.head(10)
```

```
(1616, 7)
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.758800	-79.320197	Food Basics	43.760549	-79.326045	Supermarket
1	Parkwoods	43.758800	-79.320197	Metro	43.756643	-79.313639	Supermarket
2	Victoria Village	43.732658	-79.311189	Al Ansar İslamic Centre	43.738526	-79.320284	Mosque
3	Victoria Village	43.732658	-79.311189	Adam's Halal Pub	43.736584	-79.308075	Halal Restaurant
4	Harbourfront	43.640080	-79.380150	Warerfront Artisan Market	43.638320	-79.388133	Farmers Market
5	Harbourfront	43.640080	-79.380150	Family Foods	43.646726	-79.372442	Farmers Market
6	Harbourfront	43.640080	-79.380150	Autum Market - Indoor Farmer Market	43.647636	-79.384113	Farmers Market
7	Harbourfront	43.640080	-79.380150	Farmers MRKT	43.647038	-79.379798	Farmers Market
8	Harbourfront	43.640080	-79.380150	Metro Hall Farmer's Market	43.645920	-79.387507	Farmers Market
9	Harbourfront	43.640080	-79.380150	Longo's Maple Leaf Square	43.642517	-79.381393	Supermarket

3.1. Exploratory Data Analysis

Out of 1616 venues returned, is it found that there are **24 unique categories**. The following is a list of the unique categories and the number of venues for each category:

```
print('There are {} uniques categories.'.format(len(tor_venues['Venue Category'].unique())))
tor_venues.groupby('Venue Category')['Venue Category'].count().sort_values(ascending=False)
```

```
There are 24 uniques categories.
Venue Category
Farmers Market               456
Grocery Store                329
Supermarket                  243
Mosque                       142
Fish Market                  122
Fruit & Vegetable Store      112
Halal Restaurant             101
Middle Eastern Restaurant     31
Comfort Food Restaurant       19
Market                        17
Big Box Store                 15
Afghan Restaurant              6
Pizza Place                    6
Food Court                     5
Seafood Restaurant             2
Chinese Restaurant             2
Kebab Restaurant               1
Pakistani Restaurant           1
Flea Market                    1
Park                           1
Dessert Shop                   1
Burger Joint                   1
Steakhouse                     1
Turkish Restaurant             1
Name: Venue Category, dtype: int64
```

Data Cleaning

The aim of the project is to identify the most strategic neighbourhood area to start up a halal restaurant by clustering the neighbourhood using the venues' categories. So, it is important to know the key factor that helps in choosing the strategic area. For this project the aim to focus on the location of:

- **Mosque** - Identifying where most mosque are located is the key factor to identify which neighbourhood area has the most demand for halal restaurant location as typically most Muslims will be clustered nearby the mosque.
- **Halal Restaurant** - this data helps to identify the best place to set up a halal restaurant by showing which area has been saturated with competition. Knowing the competition can be a great help in reducing business risk.
- **Markets** - another important factor for setting up a restaurant business is to identify how convenient is the location in terms of distance for getting fresh supply of menu ingredient stocks. Are there any nearby farm markets, wet markets, grocery store, etc? This data will help in answering just that.We will specifically use 'Supermarket', 'Fish Market', 'Farmers Market' and 'Fruit & Vegetable Store' venue categories to be explored around each neighbourhood.

Thus, it is important to remove all the venues from the Toronto venues (**tor_venues**) dataframe which have generalised categories. Generalised categories here mean any venues category that is not 'Mosque', 'Halal Restaurant', 'Supermarket', 'Fish Market', 'Farmers Market' and 'Fruit &

Vegetable Store’. To remove the generalised categories venues, firstly all the unique categories are fed into a python ‘list’.

Next, manually create the general categories python ‘list’ by assigning all the unique categories except the specific categories mentioned above. The following is the categories listed as ‘general’:

```
# manually create a list of generalized categories
general_categories = ['Grocery Store', 'Middle Eastern Restaurant',
                      'Comfort Food Restaurant', 'Market', 'Big Box Store',
                      'Pizza Place', 'Chinese Restaurant', 'Burger Joint',
                      'Park', 'Afghan Restaurant', 'Dessert Shop',
                      'Kebab Restaurant', 'Turkish Restaurant',
                      'Pakistani Restaurant','Seafood Restaurant',
                      'Flea Market', 'Food Court', 'Steakhouse']
```

Now, in order to get the ‘list’ of all the categories which are required for further analysis, just simply do a subtraction of the two python ‘list’ i.e. subtract general categories list set from unique categories list set. The curated python ‘list’ of required categories will be used to remove all the venues with generalised categories.

After the subtraction only 6 unique categories are left, as compared to 24 earlier. This means, 75% of the data was a noise for the analysis. This essential step helped to capture the true data points of interest.

Feature Engineering

Now, to understand the most common venues within 1 Kilometre of an area, each neighbourhood needs to be analysed individually. This process is done using the ‘**one hot encoding**’ function of python pandas library. One hot encoding converts categorial variables, in this case the ‘Venue Category’, into a form that could be fitted to Machine Learning algorithms to do prediction.

```
# one hot encoding
tor_onehot = pd.get_dummies(tor_venues[['Venue Category']], prefix="", prefix_sep="")
tor_onehot.head()
```

	Farmers Market	Fish Market	Fruit & Vegetable Store	Halal Restaurant	Mosque	Supermarket
0	0	0	0	0	0	1
1	0	0	0	0	0	1
2	0	0	0	0	1	0
3	0	0	0	1	0	0
4	1	0	0	0	0	0

With this data, count the number of venues of each category in each neighbourhood.

```
venue_counts = tor_onehot.groupby('Neighborhood').sum()
venue_counts.head(5)
```

	Farmers Market	Fish Market	Fruit & Vegetable Store	Halal Restaurant	Mosque	Supermarket
Neighborhood						
Adelaide\n	18	3	2	3	4	5
Agincourt	0	0	1	0	0	1
Agincourt North	0	0	0	0	1	1
Bathurst Manor	0	0	0	0	1	0
Bathurst Quay\n	2	0	1	0	0	1

The ‘Venue Category’ top chart can also be found by counting their occurrences around Toronto.The analysis shows that ‘Farmers Market’, ‘Mosque’ and ‘Supermarket’ are among the top 3.

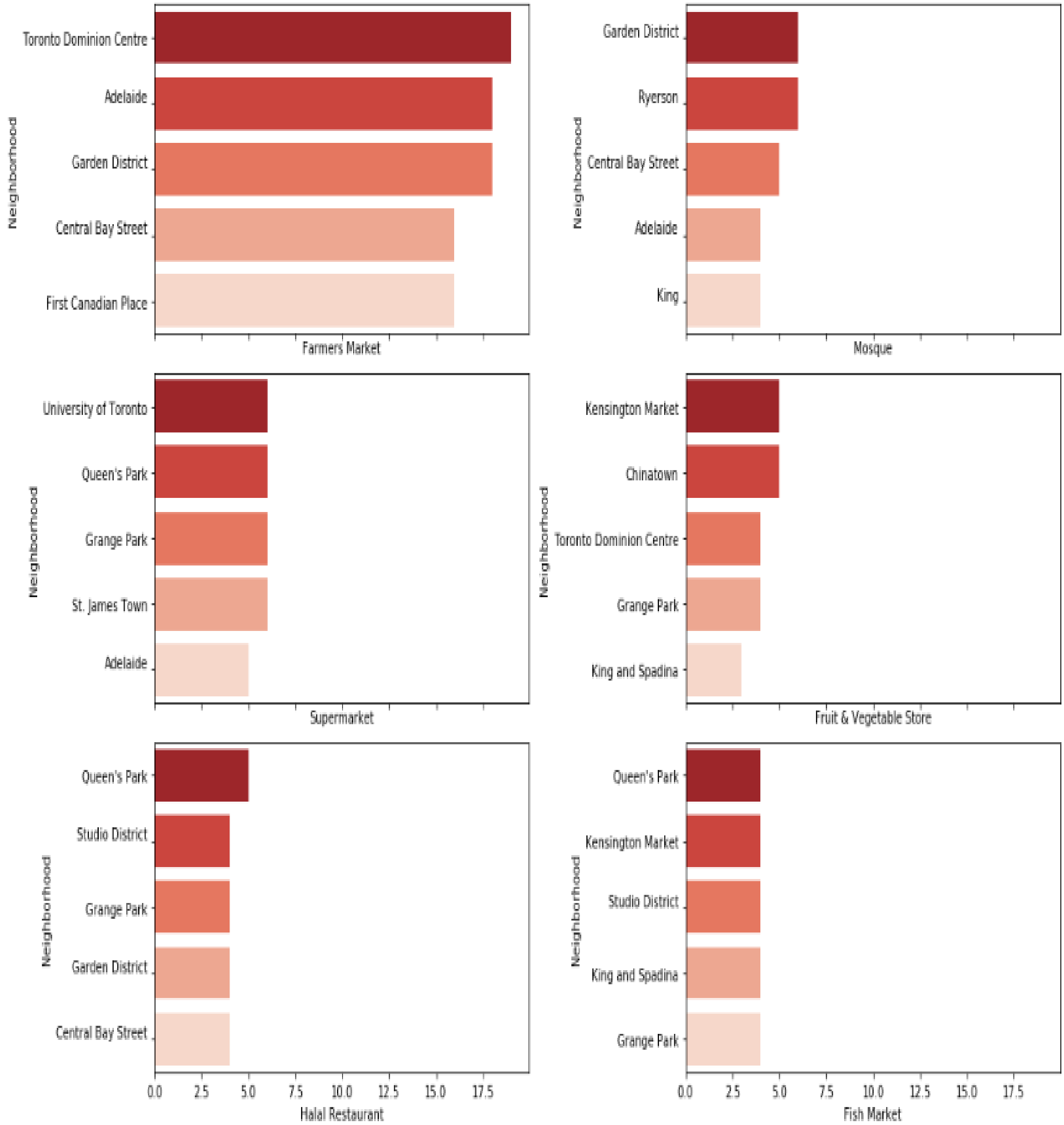
```
venue_counts_described = venue_counts.describe().transpose()
```

```
venue_top_chart = venue_counts_described.sort_values('max', ascending=False)[0:6]
venue_top_chart
```

	count	mean	std	min	25%	50%	75%	max
Farmers Market	170.0	2.682353	4.732393	0.0	0.0	1.0	2.0	19.0
Mosque	170.0	0.835294	1.322066	0.0	0.0	0.0	1.0	6.0
Supermarket	170.0	1.429412	1.652778	0.0	0.0	1.0	2.0	6.0
Fruit & Vegetable Store	170.0	0.658824	1.099251	0.0	0.0	0.0	1.0	5.0
Halal Restaurant	170.0	0.594118	1.174165	0.0	0.0	0.0	1.0	5.0
Fish Market	170.0	0.717647	1.136920	0.0	0.0	0.0	1.0	4.0

Data Visualisation

These categories are further plotted individually on bar graph using python ‘**seaborn**’ library. The graph below plots the top 5 neighbourhoods for each category.



Next, group the rows by neighborhood and take the mean of the frequency of occurrences of each category.

	Neighborhood	Farmers Market	Fish Market	Fruit & Vegetable Store	Halal Restaurant	Mosque	Supermarket
0	Adelaide\n	0.514286	0.085714	0.057143	0.085714	0.114286	0.142857
1	Agincourt	0.000000	0.000000	0.500000	0.000000	0.000000	0.500000
2	Agincourt North	0.000000	0.000000	0.000000	0.000000	0.500000	0.500000
3	Bathurst Manor	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000
4	Bathurst Quay\n	0.500000	0.000000	0.250000	0.000000	0.000000	0.250000

From here, let's create a dataframe called '**neighborhood_venues_sorted'** that define the top 5 most common venues within 1 Kilometres of each neighborhood. This dataframe can be used to define the strategic location for setting up a halal restaurant business.

```
for ind in np.arange(tor_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(tor_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Adelaide\n	Farmers Market	Supermarket	Mosque	Halal Restaurant	Fish Market
1	Agincourt	Supermarket	Fruit & Vegetable Store	Mosque	Halal Restaurant	Fish Market
2	Agincourt North	Supermarket	Mosque	Halal Restaurant	Fruit & Vegetable Store	Fish Market
3	Bathurst Manor	Mosque	Supermarket	Halal Restaurant	Fruit & Vegetable Store	Fish Market
4	Bathurst Quay\n	Farmers Market	Supermarket	Fruit & Vegetable Store	Mosque	Halal Restaurant

3.2. Machine Learning

For this project, k-means will be used to clusters the neighbourhood into several groups. K-means is one of the simplest unsupervised machine learning technique. It finds groupsin the data points, with the number of groups represented by the variable K. The algorithm works iteratively to cluster each data point to one of K groups based on certain similar feature.

To implement this algorithm, it is best to first find the optimal number of clusters, which is K. There are several techniques to determine the optimal value of K such as ‘The Silhouette Coefficient Method’, ‘The Elbow Method’, etc. In general, there is no one method to determine the exact value of K,but an accurate estimation can be obtained from those techniques.Let’s use the silhouette method to find the optimal value of K.

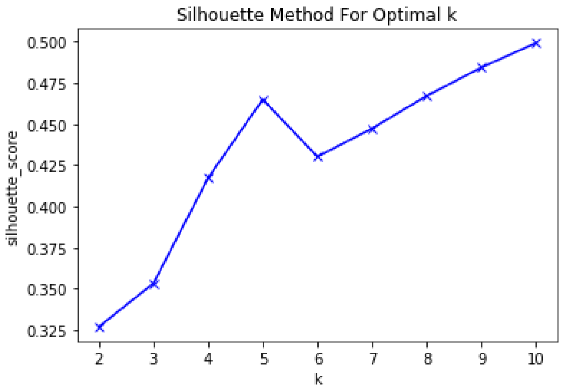
The SilhouetteCoefficient Method

Silhouette method measure how close or similar each data point in one cluster is to points in the neighbouring clusters. For this method, it requires a minimum of 2 clusters to define dissimilarity number of cluster, thus range of value K should start at 2 and end with own preference.

The following is the implementation of this method. Here range of K is set to be between 2 to 11. shown in the graph below, the point of peak is a good indication that the underlying model fits best at that point. Therefore, the best k (number of cluster) is **k = 5**:

```
sil = []
K_sil = range(2,11)
# minimum 2 clusters required, to define dissimilarity
for k in K_sil:
    kmeans = KMeans(n_clusters = k).fit(tor_grouped_clustering)
    labels = kmeans.labels_
    sil.append(silhouette_score(tor_grouped_clustering, labels, metric = 'euclidean'))
    #print("For n_clusters={}, The Silhouette Coefficient is {}".format(k, sil[k-2]))
```

```
plt.plot(K_sil, sil, 'bx-')
plt.xlabel('k')
plt.ylabel('silhouette_score')
plt.title('Silhouette Method For Optimal k')
plt.show()
```



k-Means

Now that the optimal value of K have been chosen, next is to run the k-Means algorithm with the number of clusters = 5. Then print the counts of neighbourhood of each cluster:

```
# set number of clusters
kclusters = 5

# run k-means clustering
kmeans = KMeans(init="k-means++", n_clusters=kclusters, n_init=10).fit(tor_grouped_clustering)

print(Counter(kmeans.labels_))
```

```
Counter({1: 82, 2: 30, 0: 26, 3: 19, 4: 13})
```

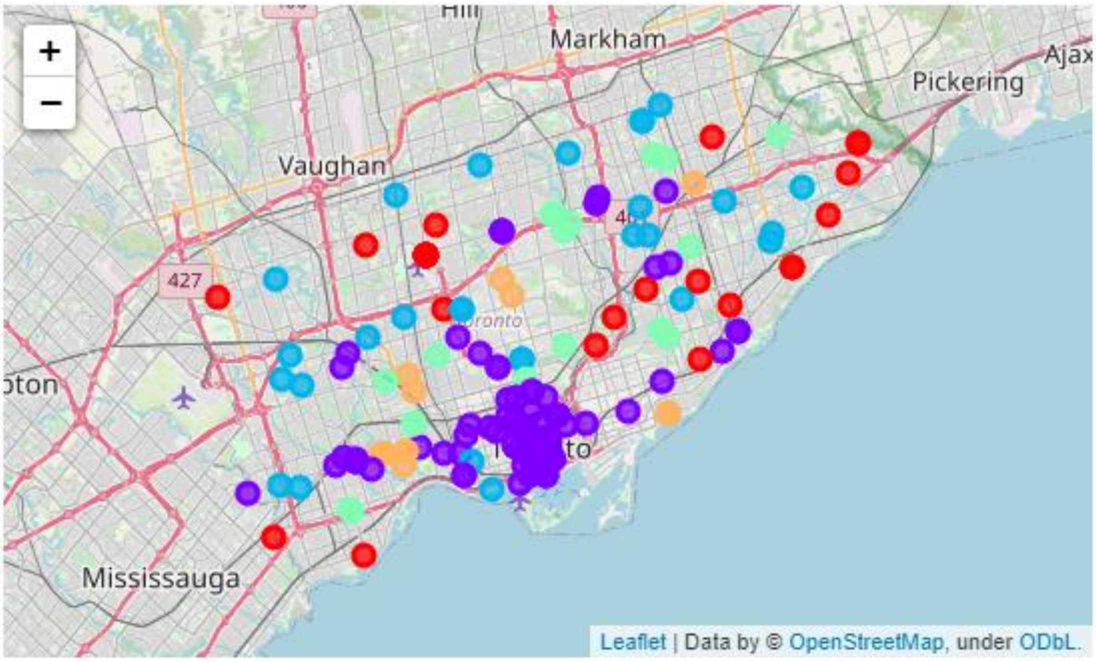
Next is to add the cluster labels to the 'neighborhood_venues_sorted' dataframe. Then, merged 'neighborhood_venues_sorted' with 'toronto_data' dataframe to add the Borough, Latitude and Longitude of each neighbourhood. The final dataframe looked like shown below:

```
# add clustering labels
try:
    neighborhoods_venues_sorted.drop('Cluster Labels', axis=1)
except:
    neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)
```

```
# merge neighborhoods_venues_sorted withtor_data to add latitude/longitude for each neighborhood
tor_merged = neighborhoods_venues_sorted.join(toronto_data.set_index('Neighborhood'), on='Neighborhood')
tor_merged.head()
```

	Cluster Labels	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	Borough	Latitude	Longitude
0	1	Adelaide	Farmers Market	Supermarket	Mosque	Halal Restaurant	Fish Market	Downtown Toronto	43.650486	-79.379498
1	4	Agincourt	Supermarket	Fruit & Vegetable Store	Mosque	Halal Restaurant	Fish Market	Scarborough	43.785353	-79.278549
2	0	Agincourt North	Supermarket	Mosque	Halal Restaurant	Fruit & Vegetable Store	Fish Market	Scarborough	43.808038	-79.266439
3	0	Bathurst Manor	Mosque	Supermarket	Halal Restaurant	Fruit & Vegetable Store	Fish Market	North York	43.763893	-79.456367
4	1	Bathurst Quay	Farmers Market	Supermarket	Fruit & Vegetable Store	Mosque	Halal Restaurant	Downtown Toronto	43.635790	-79.398329

With the above data, let's visualise the Toronto's neigborhoods utilizing the python 'folium' library. This time the generated map will show the desired segmentation of the Toronto's neighbourhoods:



4. Results

Following are the results of the clusters analysis:

Cluster – 0 (*Red Cluster*)

```
for col in required_column:
    print(cluster_0[col].value_counts(ascending = False))
    print("---------------------------------------------")
```

```
Mosque          23
Supermarket      3
Name: 1st Most Common Venue, dtype: int64
---------------------------------------------
Supermarket          21
Mosque                3
Halal Restaurant      2
Name: 2nd Most Common Venue, dtype: int64
---------------------------------------------
North York         11
Scarborough        10
Etobicoke           3
Central Toronto     1
East York           1
Name: Borough, dtype: int64
---------------------------------------------
```

There are **26 neighborhoods** in cluster 0. As appear on the analysis above, '**Mosque**' shows the highest occurrence with 23 occurrences in the '1st Most Common Venue' across different neighbourhoods followed by '**Supermarket**' with 21 occurrences in the '2nd Most Common Venue'. The majority of these neighbourhoods are in **North York** and **Scarborough** borough.

So, Cluster – 0 can be termed as 'Mosque' and 'Supermarket' dominant cluster.

Cluster – 1 (*Purple Cluster*)

```
for col in required_column:
    print(cluster_1[col].value_counts(ascending = False))
    print("---------------------------------------------")
```

```
Farmers Market            54
Supermarket               16
Mosque                     5
Halal Restaurant           4
Fish Market                3
Fruit & Vegetable Store    1
Name: 1st Most Common Venue, dtype: int64
---------------------------------------------
Supermarket               47
Farmers Market            15
Fish Market                7
Fruit & Vegetable Store    6
Mosque                     5
Halal Restaurant           3
Name: 2nd Most Common Venue, dtype: int64
---------------------------------------------
Downtown Toronto    32
Etobicoke           14
Central Toronto     10
West Toronto         8
North York           7
Scarborough          6
East Toronto         4
York                 1
East York            1
Name: Borough, dtype: int64
---------------------------------------------
```

There are **82 neighborhoods** in cluster 1. As appear on the analysis above, obviously this is a '**Farmers Market**' and '**Supermarket**' dominant cluster. The majority of these neighbourhoods are in **Downtown Toronto**.

So, Cluster – 1 can be termed as 'Farmers Market' and 'Supermarket' dominant cluster.

Cluster – 2 (*Blue Cluster*)

```
for col in required_column:
    print(cluster_2[col].value_counts(ascending = False))
    print("---------------------------------------------")
```

```
Supermarket    30
Name: 1st Most Common Venue, dtype: int64
---------------------------------------------
Mosque                        20
Farmers Market                 4
Fruit & Vegetable Store        3
Halal Restaurant               3
Name: 2nd Most Common Venue, dtype: int64
---------------------------------------------
Etobicoke          9
Scarborough        9
North York         8
West Toronto       2
Central Toronto    2
Name: Borough, dtype: int64
---------------------------------------------
```

There are **30 neighborhoods** in cluster 2. As appear on the analysis above, '**Supermarket**' and '**Mosque**' are also the top two most common venues for this cluster. The majority of these neighbourhoods are in **Etobicoke**, **Scarborough** and **North York** borough.

So, Cluster – 2 can be termed as 'Supermarket' and 'Mosque' dominant cluster.

Cluster – 3 (*Green Cluster*)

```
for col in required_column:
    print(cluster_3[col].value_counts(ascending = False))
    print("---------------------------------------------")
```

```
Fish Market    11
Supermarket     8
Name: 1st Most Common Venue, dtype: int64
---------------------------------------------
Supermarket    10
Fish Market     7
Mosque          2
Name: 2nd Most Common Venue, dtype: int64
---------------------------------------------
Scarborough         6
North York          3
York                2
Etobicoke           2
Central Toronto     2
East York           2
Downtown Toronto    1
West Toronto        1
Name: Borough, dtype: int64
---------------------------------------------
```

There are **19 neighborhoods** in cluster 3. As appear on the analysis above, '**Fish Market**' shows to be the '1st Most Common Venue' across the different neighbourhoods in this cluster and followed by '**Supermarket**' with 10 occurrences in the '2nd Most Common Venue'. The majority of these neighbourhoods are in **Scarborough** borough.

So, Cluster – 3 can be termed as 'Fish Market' and 'Supermarket' dominant cluster.

Cluster – 4 (*Orange Cluster*)

```
for col in required_column:
    print(cluster_4[col].value_counts(ascending = False))
    print("---------------------------------------------")
```

```
Fruit & Vegetable Store    12
Supermarket                 1
Name: 1st Most Common Venue, dtype: int64
---------------------------------------------
Supermarket                6
Fish Market                4
Mosque                     2
Fruit & Vegetable Store    1
Name: 2nd Most Common Venue, dtype: int64
---------------------------------------------
York               3
Etobicoke          3
East Toronto       2
West Toronto       2
Scarborough        1
Central Toronto    1
North York         1
Name: Borough, dtype: int64
---------------------------------------------
```

There are **13 neighborhoods** in cluster 4. As appear on the analysis above, '**Fruit & Vegetable Store**' shows to be the '1st Most Common Venue' across the different neighbourhoods in this cluster and followed by '**Supermarket**' with 6 occurrences in the '2nd Most Common Venue'. The majority of these neighbourhoods are in **York** and **Etobicoke** borough.

So, Cluster – 4 can be termed as 'Fruit & Vegetable Store' and 'Supermarket' dominant cluster.

5. Discussion

To understand the clusters segmentation, three analysis were done on every cluster as follows:

1. Count of 1st most common venue
2. Count of 2nd most common venue
3. Count of Borough

These data interpret the property of clusters based on the similarities between the neighbourhoods within its cluster.

Table below summarise the k-Mean unsupervised machine learning technique results:

	Count of occurrences within the cluster		
Cluster	**1st Most Common Venue**	**2nd Most Common Venue**	**Borough**
0	Mosque	Supermarket	North York, Scarborough
1	Farmers Market	Supermarket	Downtown Toronto
2	Supermarket	Mosque	Scarborough, Etobicoke, North York
3	Fish Market	Supermarket	Scarborough
4	Fruit & Vegetable Store	Supermarket	York, Etobicoke

From the analysis, it is obvious that Supermarket is the most common venue across all the clusters which also means across all neighbourhoods. From all 6 unique venue categories that were focused on for this project, Halal Restaurant appears to be an unpopular or less common venue around the neighbourhoods.

Following defined the clusters segmented by k-Means unsupervised machine learning technique:

- Cluster 0 – Mosque
- Cluster 1 – Farmers Market
- Cluster 2 – Supermarket
- Cluster 3 – Fish Market
- Cluster 4 – Fruit & Vegetable Store

6. Conclusion

In conclusion, Toronto neighbourhoods were explored to analyse the best location for setting up a new restaurant that strictly focus on catering halal menu option. Mosque, Halal Restaurant, Fish Market, Farmers Market, Fruit & Vegetable Store and Supermarket venue categories were used to study the environment of each neighbourhood. These specific venue categories are some of the main keys in helping business owners in a number of ways. For example, where targeted audience are most saturated at, area with less competition and plenty of venues to get business supplies easily.

As a result, the best location to set up a new halal restaurant in general would be in Scarborough, Etobicoke and North York borough. It is highly recommended to select neighbourhoods from the cluster 0 and/or 2 of these boroughs. The mosque and the different types of market venues met the condition set in deciding the best location for this case. And for competition from other halal restaurants, it is not much of a concern as from the analysis done shows that halal restaurant is scarce.

Last but not least, the analysis performed in this project has some limitation. The accuracy of data depends on the data provided by Foursquare. The venues explored were only within 1km of the neighbourhoods and changing it can affect the clusters. The analysis scope can be expanded further to distinctly point out which neighbourhood would be the best location for the client to set up a halal restaurant.