

# **The Battle of the Neighborhoods - Report**

## **Introduction & Background:**

The City of Toronto is the most populous city in Canada. The diverse population of Toronto reflects its current and historical role as an important destination for immigrants to Canada. More than 50 percent of residents belong to a visible minority population group, and over 200 distinct ethnic origins are represented among its inhabitants.

The city is a prominent center for music, theatre, motion picture and television production. Its economy is highly diversified with strengths in technology, design, financial services, life sciences, education, arts, fashion, business services, environmental innovation, food services, and tourism.

Due to its background, market is highly competitive and cost of doing business is obviously one of the highest. Any company looking to start a business in this city should thoroughly analyze all aspects before making a decision. This will help reduce the risk of failure and to yield profit from the business.

## **Problem Description:**

Javari Food Company LLC is looking to open their first Mediterranean restaurant in the city of Toronto. The company has several locations in United States and is looking to expand in Canada. The City of Toronto is famous for its excellent cuisine. The food culture includes an array of international cuisines influenced by the city's immigrant history.

The food industry in Toronto is highly competitive, so several factors will influence for a restaurant to be successful:

- Location of the restaurant
- Location of other restaurant serving similar food
- City demographics
- Operating hours
- Other nearby venues – Entertainment zones, Parks, movie theatres that bring more people to the area

Company has secured funding to open its first Canadian location, but it is very critical for company to be successful in this location so it can replicate the success to open new locations in Canada.

## Objective:

The objective of this project is to analyze, locate and recommend the company's management which neighborhood in the city of Toronto will be best choice to start their restaurant.

## Data:

**City being analyzed:** Toronto, ON, Canada

We will be using the below datasets for analyzing Toronto city:

1. Toronto Postal codes and neighborhood information will be retrieved from the following link:

[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	Rouge,Malvern
1	M1C	Scarborough	Highland Creek ,Rouge Hill,Port Union
2	M1E	Scarborough	Guildwood,Morningside,West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

2. Toronto city geospatial coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. The geospatial information is retrieved from the link: [http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data)

	Postcode	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

3. Toronto population from the census data gathered every 5 years from 1996 to 2016 taken from the following link:

[https://en.wikipedia.org/wiki/Demographics\\_of\\_Toronto](https://en.wikipedia.org/wiki/Demographics_of_Toronto)

	2016	2011	2006	2001	1996
<b>Ethnic Origins</b>					
Canadian	746960	728745.0	651635.0	861945.0	710755.0
English	732555	777110.0	804100.0	783770.0	891735.0
Chinese	700705	594735.0	537060.0	435685.0	359450.0
Indian	643370	572250.0	484655.0	345855.0	255685.0
Irish	544380	543600.0	531865.0	487215.0	480980.0
Scottish	543760	545365.0	561050.0	517115.0	534595.0
Italian	484360	475090.0	466155.0	429385.0	414310.0
Filipino	274675	246345.0	181330.0	140405.0	102525.0
German	271815	262830.0	259015.0	220140.0	224525.0
French	247790	249375.0	241395.0	220535.0	236315.0

4. Number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**

## Methodology:

### Business Understanding:

Our main goal is to get optimal location for new restaurant business in Toronto, Canada for Javari Food Company LLC.

### Analytic Approach:

Toronto city has many neighborhoods, but we are choosing 4 neighborhoods namely – Downtown Toronto, East Toronto, West Toronto and Central Toronto. We will first extract the data and merge them with the venues data and later cluster the restaurants venue to identify the optimum location for the restaurant.

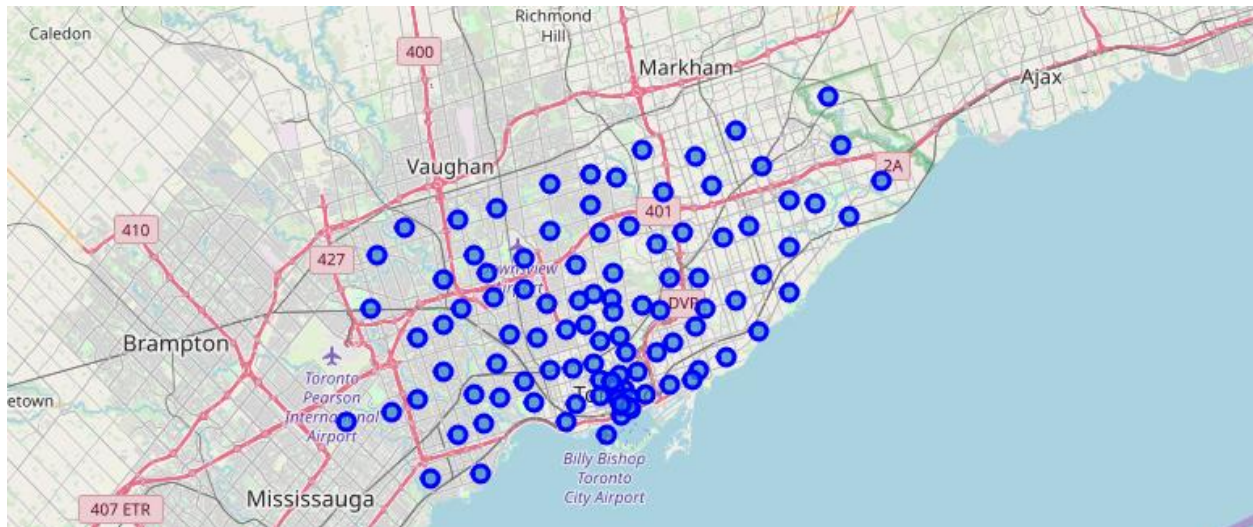
### Exploratory Data Analysis:

#### Data-1 – Toronto Geographical coordinates data

1. We load both Toronto neighborhood information and geographical coordinates data
2. The data is loaded into pandas dataframe and will be used to get venues information from Foursquare APIs./;

3. We used geopy and folium libraires to create the map of Toronto City with neighborhoods superimposed on top.

## Neighborhood Visualization of Toronto City

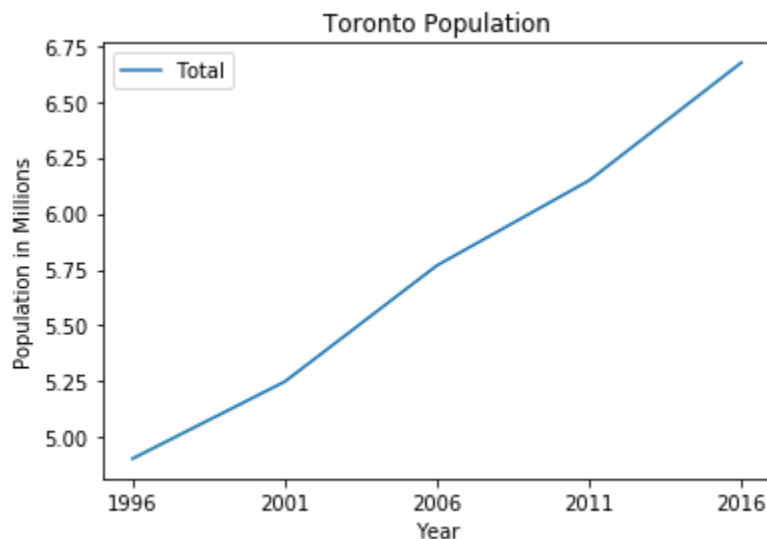


## Data 2 – Toronto population

We take the census data for the city of Toronto and load that to the data frame. We use this data to analyze population with different ethnic origins living in the city which will help us understand how different restaurants in this area will be successful.

We can see that the population in the Toronto city is steadily increasing each year which in-turn indicates that the demand for the restaurants also increases steadily. Moreover, the population is more diverse in terms of the ethnicity and culture which drives more diverse demand for different food also.

	2016	2011	2006	2001	1996
<b>Ethnic Origins</b>					
Canadian	746960	728745.0	651635.0	861945.0	710755.0
English	732555	777110.0	804100.0	783770.0	891735.0
Chinese	700705	594735.0	537060.0	435685.0	359450.0
Indian	643370	572250.0	484655.0	345855.0	255685.0
Irish	544380	543600.0	531865.0	487215.0	480980.0



### **Data 3 – Venue information for each neighborhood**

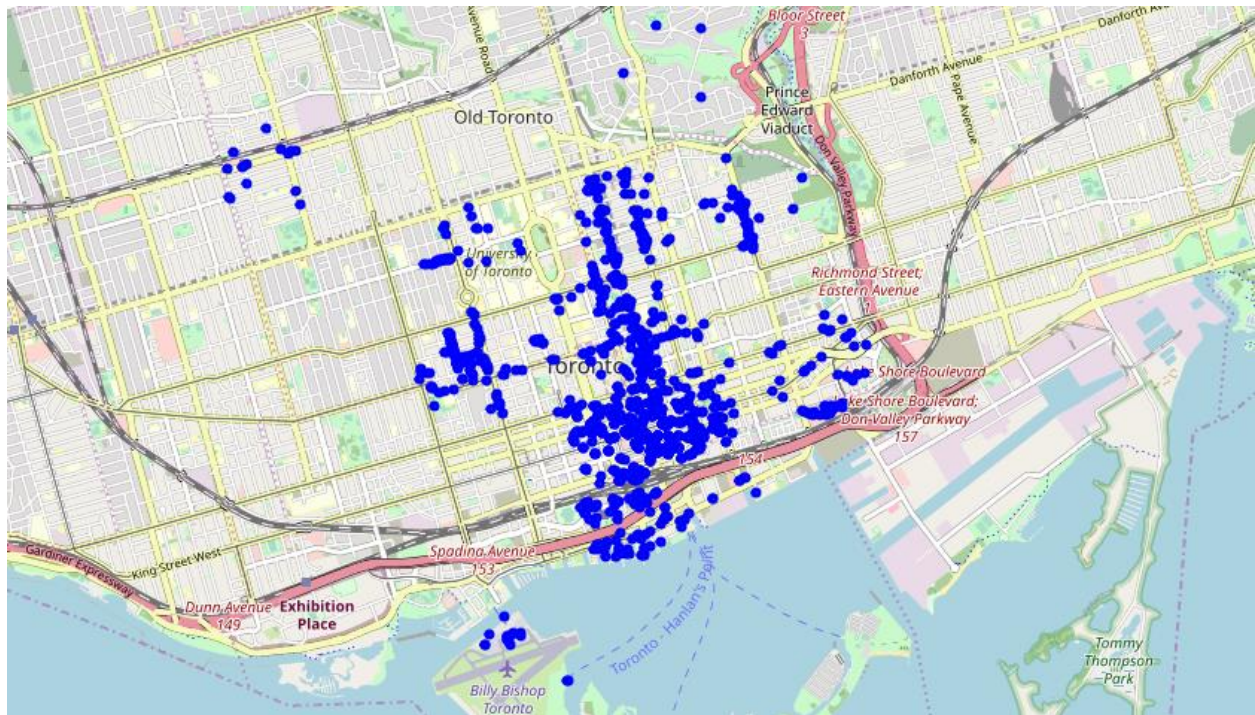
Toronto City geographical coordinates data has been utilized as an input for the Foursquare API that has been leveraged to retrieve venue details for each neighborhood. After retrieving the venue information for each neighborhood, we filtered the restaurants and grouped by each borough to identify the Top 5 restaurants in each Borough.

We used folium libraries to create venue visualization for restaurants in neighborhoods under each borough. This gives a good visualization of how crowded the neighborhoods are with different restaurants.

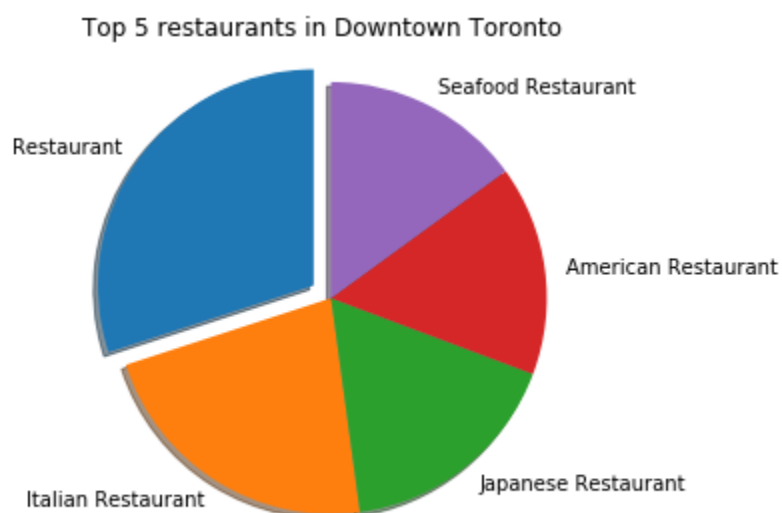
After grouping the information based on borough, we filtered only restaurant venues as we are focused on restaurant business and then identified the top 5 restaurant categories, so we can see which restaurants are more popular in these areas before making the decision.

### **Venue Visualization of Downtown Toronto:**

Downtown Toronto has different 749 venues and 204 unique venue types. The venue visualization of Downtown Toronto is generated below:



We also filtered restaurant venue types and identified top 5 restaurant venue types in Downtown Toronto. Venues data seems to have generic restaurant type that has more restaurants, but the next best one is Italian restaurant type followed by Japanese, American and Seafood.



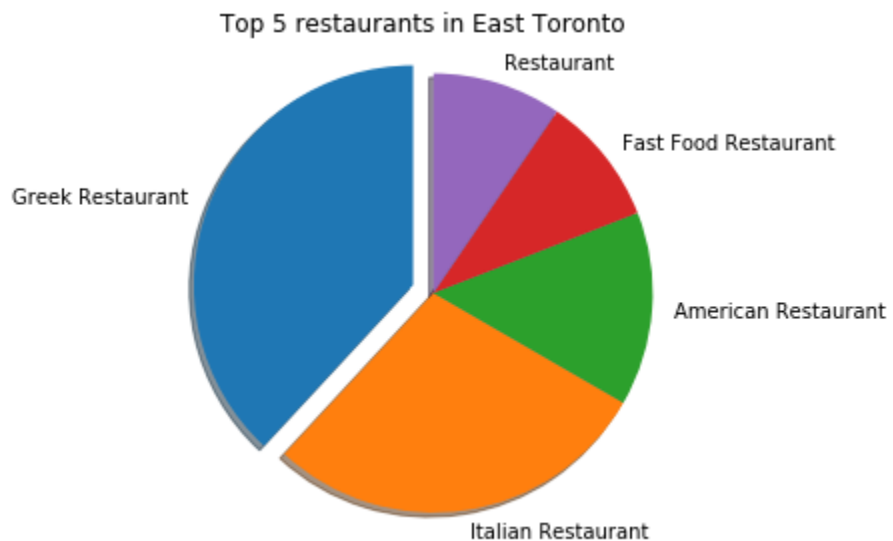


### Venue Visualization of East Toronto:

East Toronto has different 121 venues and 68 unique venue types. The venue visualization of East Toronto is generated below:

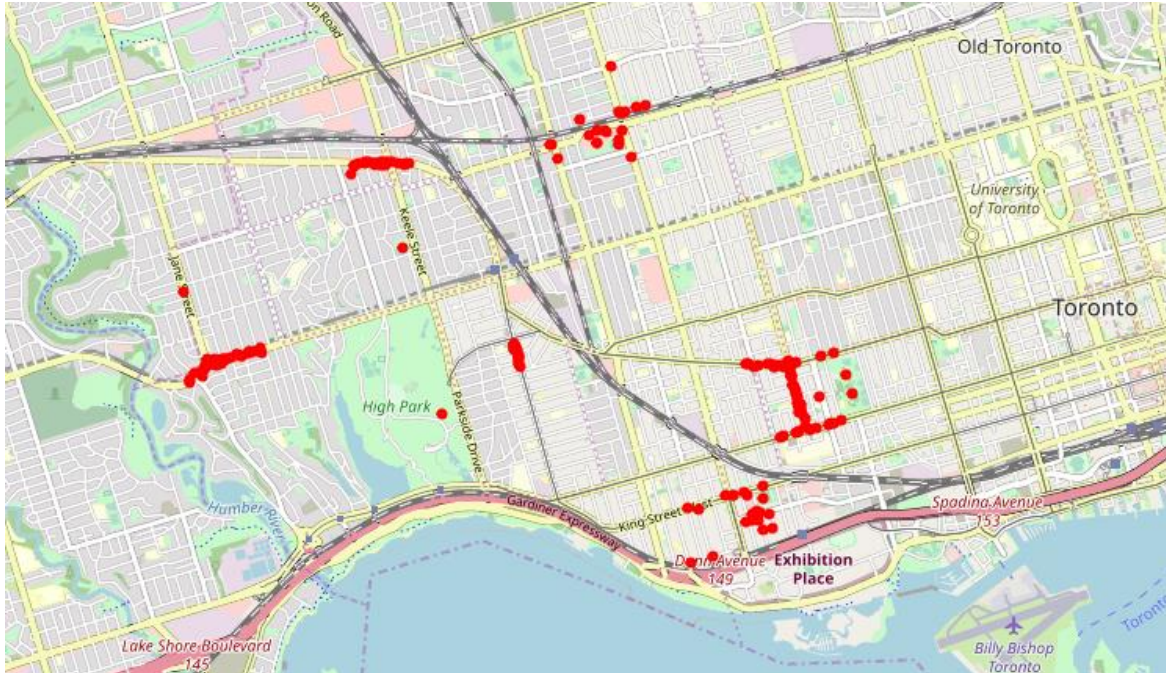


Based on the top 5 restaurants type, Greek restaurant seems to be more popular in East Toronto and has more number of restaurants than any other restaurant type.



### Venue Visualization of West Toronto:

West Toronto has different 176 unique venues and 90 unique venue types. The venue visualization of West Toronto is generated below:



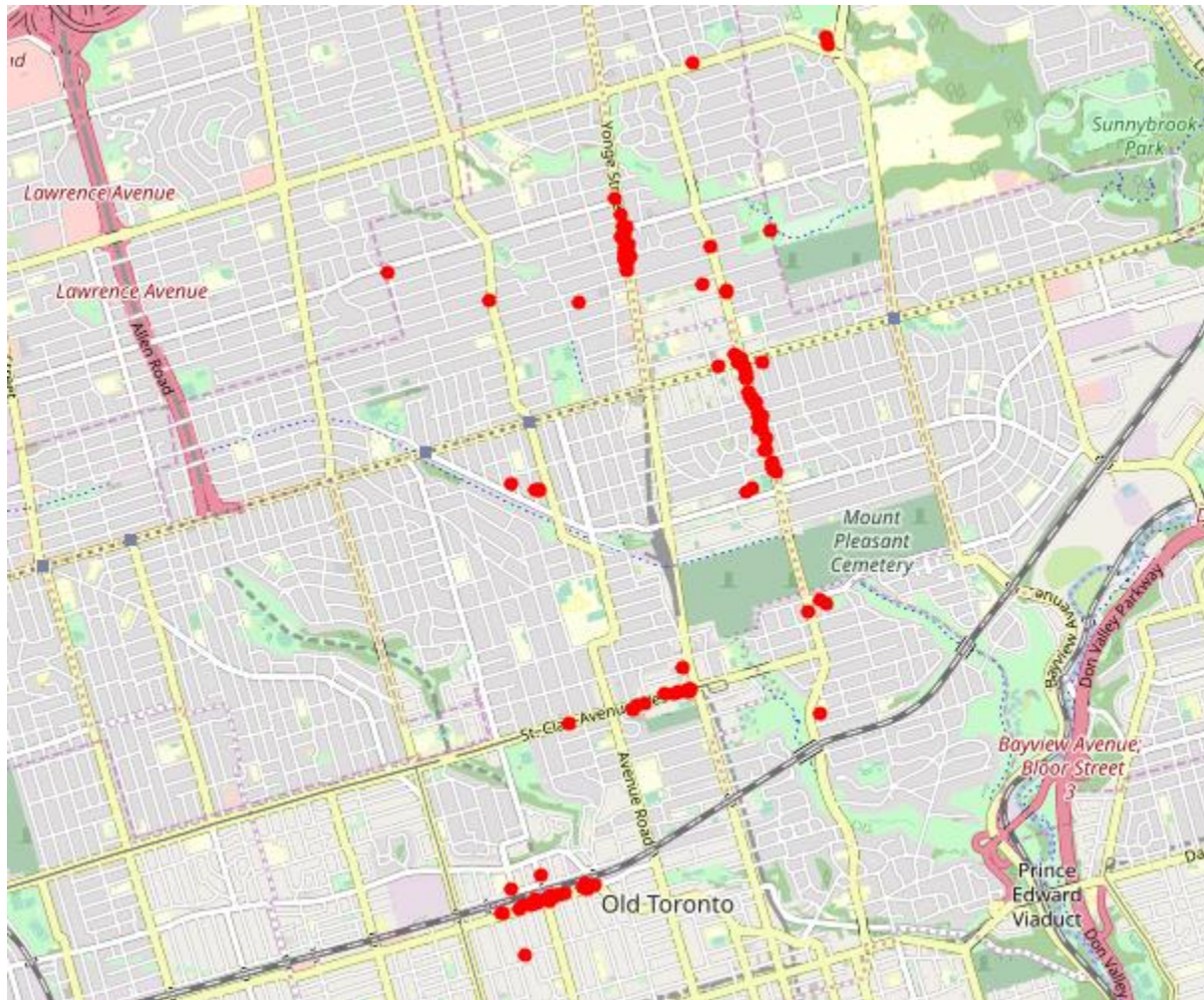
Based on the top 5 restaurants type, Italian restaurant seems to be more popular in West Toronto and has more number of restaurants than any other restaurant type.



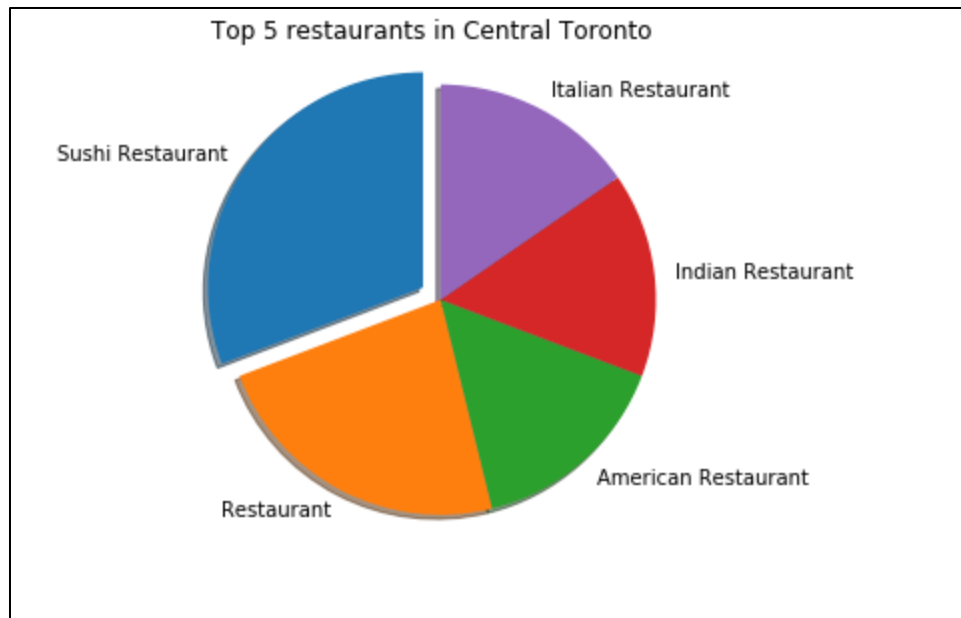


### Venue Visualization of Central Toronto:

Central Toronto has different 102 unique venues and 64 unique venue types. The venue visualization of Central Toronto is generated below:



Based on the top 5 restaurants type, Sushi restaurant seems to be more popular in Central Toronto and has more number of restaurants than any other restaurant type.



## Results:

### Neighborhood K-Means clustering based on mean occurrence of venue category:

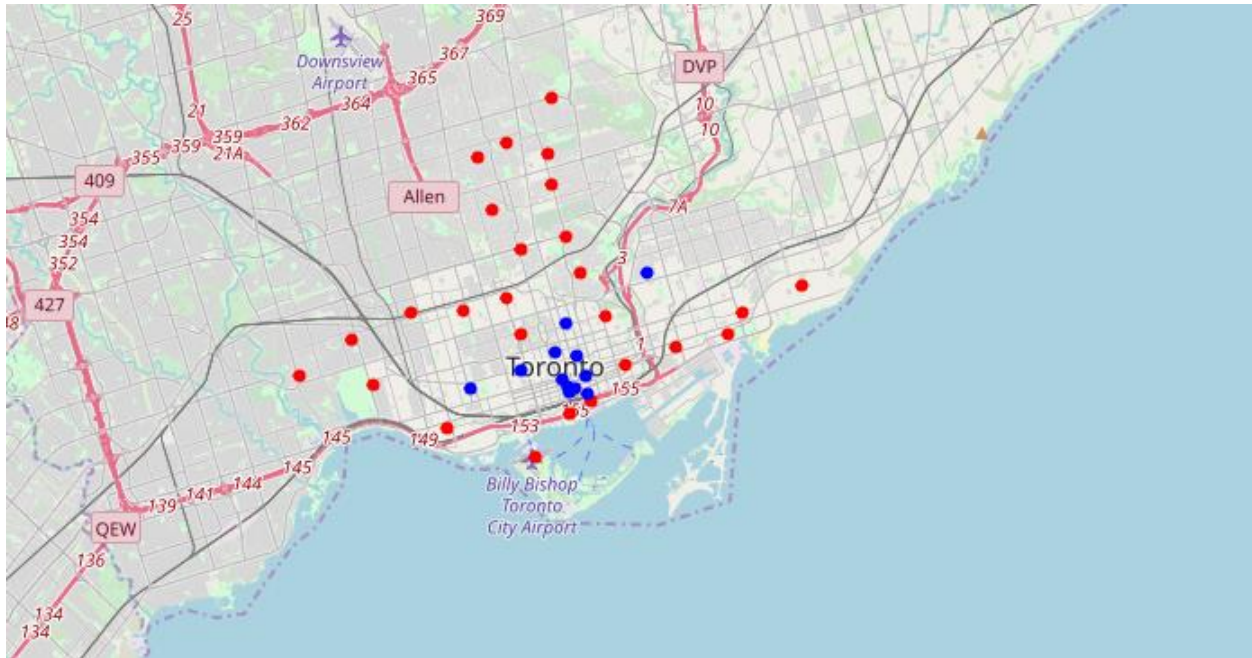
To cluster the neighborhoods, we used the K-Means clustering algorithm. K-Means clustering algorithm partitions  $n$  observations into  $K$  clusters in which each observation belongs to the cluster with the nearest mean. It uses the iterative refinement approach.

We used Silhouette coefficient method to determine the optima value of  $K$  for the dataset. A higher Silhouette Coefficient score relates to a model with better-defined clusters.

We ran the K-Means clustering model with  $K$  ranging from 2 to 9 and identified that Silhouette coefficient was higher when  $K=2$ . We used  $K=2$  and ran the K-Mean clustering and used the results to analyze the clusters.

### Clustering for all 4 Toronto boroughs:

In the below map visualization, we can see the different types of clusters created by using K-Means for Downtown, East, West and Central Toronto.



**Cluster 0:** The Total and Total Sum of Cluster0 has highest value. It shows the markets are saturated. Number of restaurants are very high in this cluster.

**Cluster 1:** The Total and Total Sum of Cluster1 has smallest value. It shows the markets are not saturated. There are some neighborhoods with no restaurants.

	Total	Total Sum
cluster0	24.500000	49.000000
cluster1	4.461538	8.923077

East, Central and Downtown Toronto all have some neighborhoods that has no restaurants.

	Borough	Neighbourhood	Latitude	Longitude	Total	Cluster_Labels
0	East Toronto	The Beaches	43.676357	-79.293031	0	1
1	Central Toronto	Lawrence Park	43.728020	-79.388790	0	1
2	Central Toronto	Davisville North	43.712751	-79.390197	0	1
3	Downtown Toronto	Rosedale	43.679563	-79.377529	0	1
4	Central Toronto	Roselawn	43.711695	-79.416936	0	1
5	Downtown Toronto	CN Tower,Bathurst Quay,Island airport,Harbourf...	43.628947	-79.394420	0	1

## **Discussion:**

Toronto is a big city with a high population density and with very diverse ethnic and cultural background. Due to diverse culture and people immigrated from multiple countries, Toronto has wide variety of restaurants. We focused on neighborhoods in 4 boroughs – Downtown, East, West and Central Toronto.

We used K-means algorithm as part of this clustering study and identified the optimum K value using Silhouette coefficient as 2. We also identified neighborhoods with no restaurants which shows that there is no nearby competition to begin with. It is possible that there may be reasons for no restaurants in that area which might make them unsuitable for new restaurants. Therefore, these locations can be used as markers/starting points for more detailed local analysis based on other factors.

## **Conclusion:**

Goal of this project is to identify neighborhoods in the City of Toronto with less or no restaurants in order to aid management in narrowing down the search for optimal location for a new Mediterranean restaurant. Using Foursquare data, we have identified all venues in each neighborhood in the city, filtered the restaurants and grouped them to identify the number of different restaurants in the area. Clustering of those locations was then performed to create zones of interest and later identified the neighborhoods that has no restaurants which will help management to explore possible location for the new restaurant.

Final decision on optimal restaurant location will be made by management based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location such as proximity to major attractions, levels of noise, proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.