

# **REPORT**

## **Introduction**

### ***Background***

The last five years have seen a surge in the number of coffee drinkers and people who visit coffee houses on a regular basis. Following this trend, the number of coffee houses that are being established grows by the day. In such a case, a person who is interested in establishing a similar coffee house, needs to conduct a thorough research.

Research, here, implies to have a thorough understanding of the number of coffee houses that are already established in various neighbourhoods, which of these coffee houses are successful (in terms of profits earned and the number of footfalls each day) and what factors contribute to the success of that particular establishment. Keeping these points in mind, and simultaneously taking steps in the correct direction will allow a person who is interested in setting up a coffee house, to be successful.

### ***Problem***

The dataset of Toronto has a list of all the boroughs and neighbourhood. This project allows us to explore the neighbourhoods of all the boroughs, to find the ten most commonly visited in each neighbourhood and cluster these neighbourhoods using k-means clustering. This will give an idea of which venues the public generally visits and will allow the interested party to determine, whether the establishment of a coffee house in a particular neighbourhood would be profitable or not.

### ***Interest***

This project is targeted to those people who are interested in setting up a coffee house. Using this project, they can decide where to establish their coffee house and what they can offer to the public, that would be different from the offerings of their competitors.

# Data

## Data Acquisition

The data containing the postal codes of Canada was obtained from [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M). This dataframe consisted of three columns - Postal Code, Borough and Neighbourhood. The next set of data containing the latitudes and longitudes of each neighbourhood was obtained from Google Maps Geocoding API. As Google was charging for the same, the Geocoder Python package was used instead (<https://geocoder.readthedocs.io/index.html>). The one drawback with this package is that it can be unreliable; to bypass this unreliability, alternatively, the data could also be collected from [http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data), where the data was available in the csv format. The dataset that was finally created, consists of the following columns:

- Neighbourhood – Name of the neighbourhood in the borough
- Borough – Name of the borough
- Latitude – Latitude of the borough
- Longitude – Longitude of the borough

## Data Cleaning

From the first data set, only the cells that have an assigned borough are selected. The others are ignored. However, if a borough does not have an assigned neighbourhood, then it will be considered the same as the borough.

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

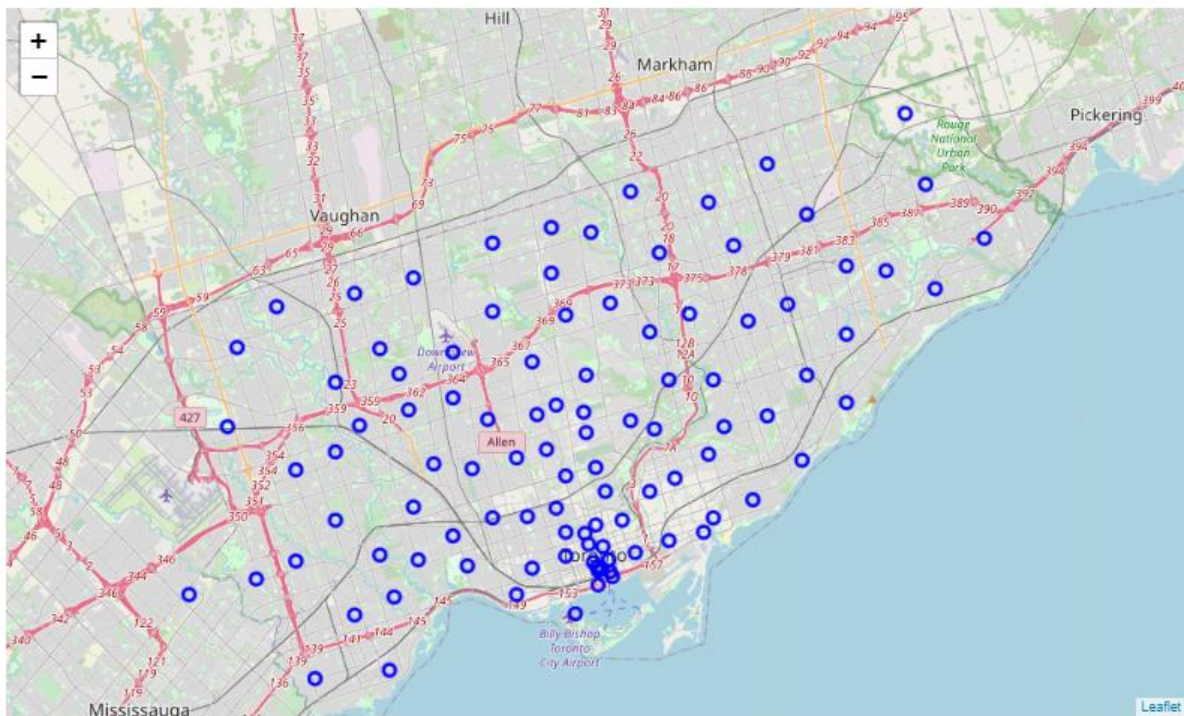
The second dataset containing the co-ordinates of all the boroughs is merged with the above dataset at the postal code.

	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

This merged dataframe is used to explore the boroughs of Toronto using Foursquare.

## Methodology

There are about 102 boroughs in Toronto.



## Results

With the help of k-means clustering, these boroughs are segregated into five clusters. These clusters were further examined to determine the 10 most visited venues with these clusters.

### 1. Cluster 1:

	Borough	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
93	Etobicoke	0	Pizza Place	Dance Studio	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Drugstore	Donut Shop	Doner Restaurant	Dog Run	Distribution Center
96	North York	0	Pizza Place	Dance Studio	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Drugstore	Donut Shop	Doner Restaurant	Dog Run	Distribution Center

### 2. Cluster 2:

	Borough	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Common V
	Scarborough	1	Park	Playground	Women's Store	Distribution Center	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Doner Resta
	North York	1	Park	Women's Store	Dog Run	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Doner Resta
	North York	1	Park	Convenience Store	Women's Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Dog Run	Doner Resta
	North York	1	Construction & Landscaping	Park	Food & Drink Shop	Dog Run	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Women's Store
	East York	1	Intersection	Convenience Store	Park	Dog Run	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Women's Store
	Central Toronto	1	Park	Swim School	Bus Line	Women's Store	Distribution Center	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Dog Run
	Central Toronto	1	Park	Playground	Tennis Court	Women's Store	Discount Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner
	Downtown Toronto	1	Park	Trail	Playground	Distribution Center	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store
	York	1	Park	Women's Store	Pool	Distribution Center	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Doner Resta
	York	1	Convenience Store	Park	Women's Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Dog Run	Doner Resta

### 3. Cluster 3:

Borough	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
Scarborough	2	Fast Food Restaurant	Women's Store	Deli / Bodega	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Drugstore	Donut Shop	Doner Restaurant
Scarborough	2	Electronics Store	Moving Target	Breakfast Spot	Medical Center	Intersection	Mexican Restaurant	Bank	Rental Car Location	Restaurant
Scarborough	2	Coffee Shop	Korean Restaurant	Women's Store	Dog Run	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center
Scarborough	2	Fried Chicken Joint	Gas Station	Bank	Hakka Restaurant	Athletics & Sports	Caribbean Restaurant	Thai Restaurant	Bakery	Department Store
Scarborough	2	Grocery Store	Playground	Distribution Center	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store
Scarborough	2	Convenience Store	Bus Station	Coffee Shop	Discount Store	Department Store	Dog Run	Dim Sum Restaurant	Diner	Distribution Center
Scarborough	2	Bus Line	Bakery	Soccer Field	Bus Station	Intersection	Metro Station	Park	Ice Cream Shop	Doner Restaurant
Scarborough	2	Motel	American Restaurant	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Dog Run
Scarborough	2	College Stadium	Café	Skating Rink	General Entertainment	Women's Store	Distribution Center	Dessert Shop	Dim Sum Restaurant	Diner
Scarborough	2	Indian Restaurant	Light Rail Station	Vietnamese Restaurant	Chinese Restaurant	Brewery	Pet Store	German Restaurant	Curling Ice	Drugstore
North York	2	Gym	Beer Store	Restaurant	Restaurant	Shop	Restaurant	Restaurant	Goods Shop	Supermarket
North York	2	Coffee Shop	Bank	Gift Shop	Fried Chicken Joint	Bridal Shop	Sandwich Place	Diner	Deli / Bodega	Restaurant
North York	2	Metro Station	Massage Studio	Coffee Shop	Furniture / Home Store	Bar	Caribbean Restaurant	Distribution Center	Dim Sum Restaurant	Diner
North York	2	Park	Grocery Store	Home Service	Discount Store	Food Truck	Baseball Field	Bank	Liquor Store	Hotel
North York	2	Park	Grocery Store	Home Service	Discount Store	Food Truck	Baseball Field	Bank	Liquor Store	Hotel
North York	2	Park	Grocery Store	Home Service	Discount Store	Food Truck	Baseball Field	Bank	Liquor Store	Hotel
North York	2	Park	Grocery Store	Home Service	Discount Store	Food Truck	Baseball Field	Bank	Liquor Store	Hotel
North York	2	French Restaurant	Pizza Place	Coffee Shop	Portuguese Restaurant	Hockey Arena	Discount Store	Dance Studio	Deli / Bodega	Department Store
East York	2	Pizza Place	Pharmacy	Gastropub	Café	Intersection	Athletics & Sports	Bank	Gym / Fitness Center	Pet Store
East York	2	Skating Rink	Park	Curling Ice	Bus Stop	Beer Store	Dumpling Restaurant	Drugstore	Donut Shop	Department Store
...	...	...	...	...	...	...	...	...	...	...
Downtown Toronto	2	Café	Coffee Shop	Vegetarian / Vegan Restaurant	Bar	Mexican Restaurant	Vietnamese Restaurant	Bakery	Grocery Store	Park
Downtown Toronto	2	Airport Service	Airport Lounge	Airport Terminal	Bar	Rental Car Location	Airport	Airport Food Court	Sculpture Garden	Harbor / Marina

Downtown Toronto	2	Coffee Shop	Italian Restaurant	Café	Gym	Seafood Restaurant	Japanese Restaurant	Hotel	Beer Bar	Restauran
Downtown Toronto	2	Coffee Shop	Café	Hotel	Gym	Japanese Restaurant	Restaurant	Salad Place	Steakhouse	Asian Restauran
North York	2	Clothing Store	Accessories Store	Women's Store	Event Space	Shoe Store	Miscellaneous Shop	Furniture / Home Store	Boutique	Vietnames Restauran
North York	2	Japanese Restaurant	Sushi Restaurant	Pizza Place	Pub	Discount Store	Department Store	Dessert Shop	Dim Sum Restaurant	Diner
York	2	Trail	Playground	Field	Dog Run	Hockey Arena	Discount Store	Deli / Bodega	Department Store	Dessert S
Downtown Toronto	2	Grocery Store	Café	Park	Restaurant	Candy Store	Baby Store	Italian Restaurant	Coffee Shop	Nightclub
West Toronto	2	Pharmacy	Bakery	Pet Store	Coffee Shop	Café	Middle Eastern Restaurant	Bar	Supermarket	Bank
West Toronto	2	Bar	Coffee Shop	Restaurant	Vietnamese Restaurant	Café	Asian Restaurant	Men's Store	Mexican Restaurant	Brewery
West Toronto	2	Café	Coffee Shop	Breakfast Spot	Nightclub	Bakery	Gym	Stadium	Burrito Place	Restauran
North York	2	Basketball Court	Park	Bakery	Construction & Landscaping	Women's Store	Dim Sum Restaurant	Diner	Discount Store	Distributio Center
York	2	Skating Rink	Sandwich Place	Bar	Discount Store	Women's Store	Distribution Center	Dessert Shop	Dim Sum Restaurant	Diner
York	2	Bus Line	Breakfast Spot	Convenience Store	Women's Store	Dog Run	Dim Sum Restaurant	Diner	Discount Store	Distributio Center
West Toronto	2	Thai Restaurant	Café	Mexican Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Bar	Speakeasy	Bakery	Italian Restauran
West Toronto	2	Breakfast Spot	Gift Shop	Restaurant	Dessert Shop	Bookstore	Bar	Italian Restaurant	Movie Theater	Cuban Restauran
West Toronto	2	Café	Coffee Shop	Pub	Sushi Restaurant	Italian Restaurant	Pizza Place	Indie Movie Theater	Fish & Chips Shop	Bar
Downtown Toronto	2	Coffee Shop	Diner	Gym	Park	Mexican Restaurant	Japanese Restaurant	Italian Restaurant	Hobby Shop	General Entertainn
Mississauga	2	Hotel	Coffee Shop	Gym	American Restaurant	Middle Eastern Restaurant	Sandwich Place	Gas Station	Burrito Place	Mediterrar Restauran
East Toronto	2	Light Rail Station	Yoga Studio	Auto Workshop	Skate Park	Burrito Place	Spa	Farmers Market	Fast Food Restaurant	Restauran
Etobicoke	2	Gym	Bakery	Hobby Shop	Fried Chicken Joint	Fast Food Restaurant	Liquor Store	Mexican Restaurant	Pharmacy	Pizza Plac
Etobicoke	2	Pizza Place	Gym	Dance Studio	Sandwich Place	Pub	Pool	Pharmacy	Skating Rink	Coffee Sh
Etobicoke	2	River	Pool	Distribution Center	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store
Etobicoke	2	Gym	Fast Food Restaurant	Flower Shop	Tanning Salon	Discount Store	Burger Joint	Sandwich Place	Bakery	Social Clu



Etobicoke	2	Filipino Restaurant	Women's Store	Deli / Bodega	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Drugstore	Donut Shop	Doner Restaurant
Etobicoke	2	Pet Store	Beer Store	Pharmacy	Pizza Place	Coffee Shop	Convenience Store	Café	Shopping Plaza	Liquor Store
Etobicoke	2	Pizza Place	Chinese Restaurant	Coffee Shop	Intersection	Sandwich Place	Middle Eastern Restaurant	Discount Store	Women's Store	Dessert Shop
Etobicoke	2	Pizza Place	Mobile Phone Shop	Bus Line	Sandwich Place	Women's Store	Discount Store	Dessert Shop	Dim Sum Restaurant	Diner
Etobicoke	2	Grocery Store	Beer Store	Fried Chicken Joint	Fast Food Restaurant	Pizza Place	Sandwich Place	Pharmacy	Airport Service	Airport Terminal
Etobicoke	2	Rental Car Location	Drugstore	Women's Store	Discount Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner

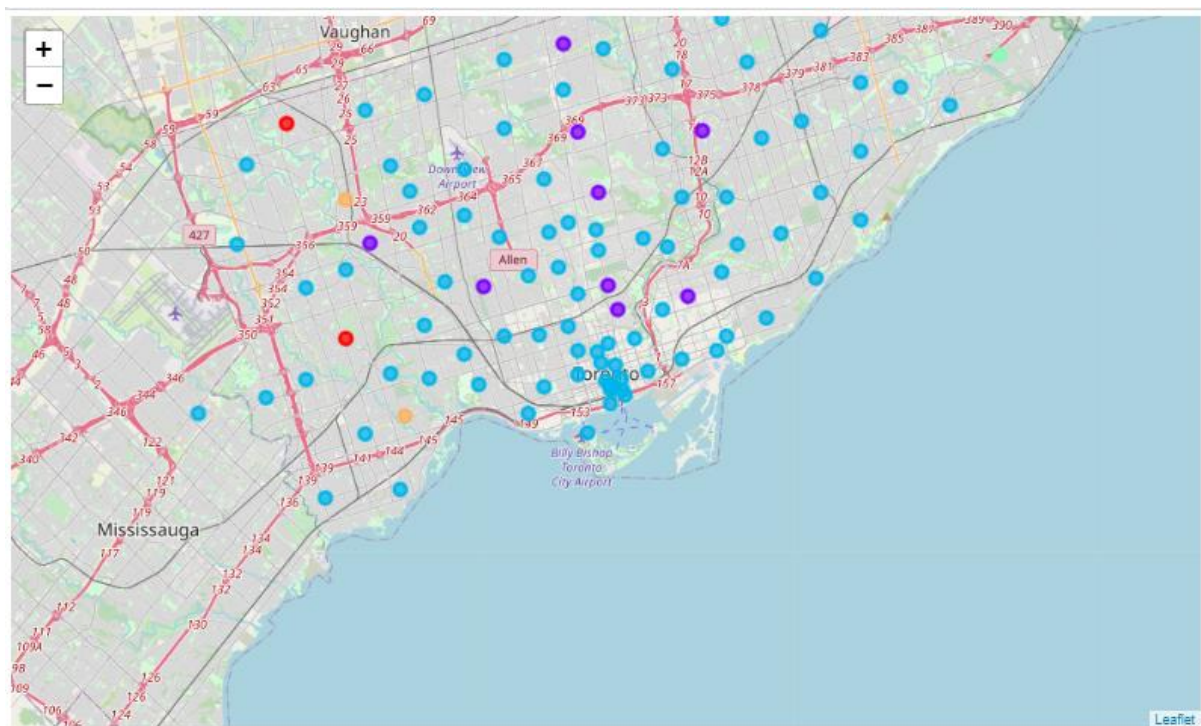
#### 4. Cluster 4:

	Borough	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Scarborough	3	Bar	Women's Store	Dog Run	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Doner Restaurant	Falafel Restaurant

#### 5. Cluster 5:

	Borough	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
91	Etobicoke	4	Business Service	Baseball Field	Dog Run	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Women's Store	Falafel Restaurant
97	North York	4	Baseball Field	Women's Store	Dog Run	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Doner Restaurant	Falafel Restaurant

These clusters can be visualized on a map using the folium library.



Each cluster is coded with different colours – blue, yellow, red, purple and green – for the ease of presentation.

## **Discussion**

The aim of this project was to help someone who is interested in setting up their own café or coffee house. They could refer to the data regarding the most commonly visited venues in different clusters and accordingly decide where they would want to establish their venture.

After analysing the customer behaviour in the above five clusters, it can be concluded that the third cluster is an extremely active clusters with a lot of establishments that are popular among the masses. Conversely, it can be observed that the remaining four clusters have lesser establishments, especially there are very few cafés or coffee houses. Establishing one here would be profitable as the people living here would not have to travel far to access a popular one.

## **Conclusion**

This project helps anyone who is interested in establishing a coffee house in Toronto. It is always a good option to use technology to stay ahead in the game and not act on the basis of trial and error, as setting up any establishment would require large number of investments.

In this project, I have taken coffee house as an example, but it can be used for other establishments also. One simply has to analyse which are the commonly visited venues in each cluster and make a decision accordingly.