

1. Introduction and business problem

Toronto is the provincial capital of Ontario and the most populous city in Canada, with a population of 2,731,571 in 2016. Current to 2016, the Toronto census metropolitan area (CMA), of which the majority is within the Greater Toronto Area (GTA), held a population of 5,928,040, making it Canada's most populous CMA. Toronto is the anchor of an urban agglomeration, known as the Golden Horseshoe in Southern Ontario, located on the northwestern shore of Lake Ontario. Toronto is an international centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.

People have travelled through and inhabited the Toronto area, situated on a broad sloping plateau interspersed with rivers, deep ravines, and urban forest, for more than 10,000 years. After the broadly disputed Toronto Purchase, when the Mississauga surrendered the area to the British Crown, the British established the town of York in 1793 and later designated it as the capital of Upper Canada. During the War of 1812, the town was the site of the Battle of York and suffered heavy damage by United States troops. York was renamed and incorporated in 1834 as the city of Toronto. It was designated as the capital of the province of Ontario in 1867 during Canadian Confederation. The city proper has since expanded past its original borders through both annexation and amalgamation to its current area of 630.2 km² (243.3 sq mi).

2. Data

To find the best location for our Coffee Shop, we will use the following sources of information:

- Toronto has a total of 6 boroughs and 140 neighbourhoods. In order to segment the neighbourhoods and explore them, we will essentially need a dataset that contains the 6 boroughs and the neighbourhoods that exist in each borough as well as the latitude and longitude coordinates of each neighbourhood. Luckily, this dataset exists for free on the web. Feel free to try to find this dataset on your own, but here is the link to the dataset: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- From Foursquare Venues Categories - <https://developer.foursquare.com/docs/resources/categories> Coffee Shops Id - 4bf58dd8d48988d1e0931735

3. Methodology

In this project, I will use the basic methodology as taught in Week 3 lab.

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M4E	East Toronto	The Beaches	43.676357	-79.293031
1	M4K	East Toronto	The Danforth West,Riverdale	43.679557	-79.352188
2	M4L	East Toronto	The Beaches West,India Bazaar	43.668999	-79.315572
3	M4M	East Toronto	Studio District	43.659526	-79.340923
4	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790

Above, I have found the corresponding latitude and longitude values of different addresses. Then I've used the Foursquare API to explore neighbourhoods in Toronto. Finally, I've applied the explore function to Coffee Shops in each neighbourhood.

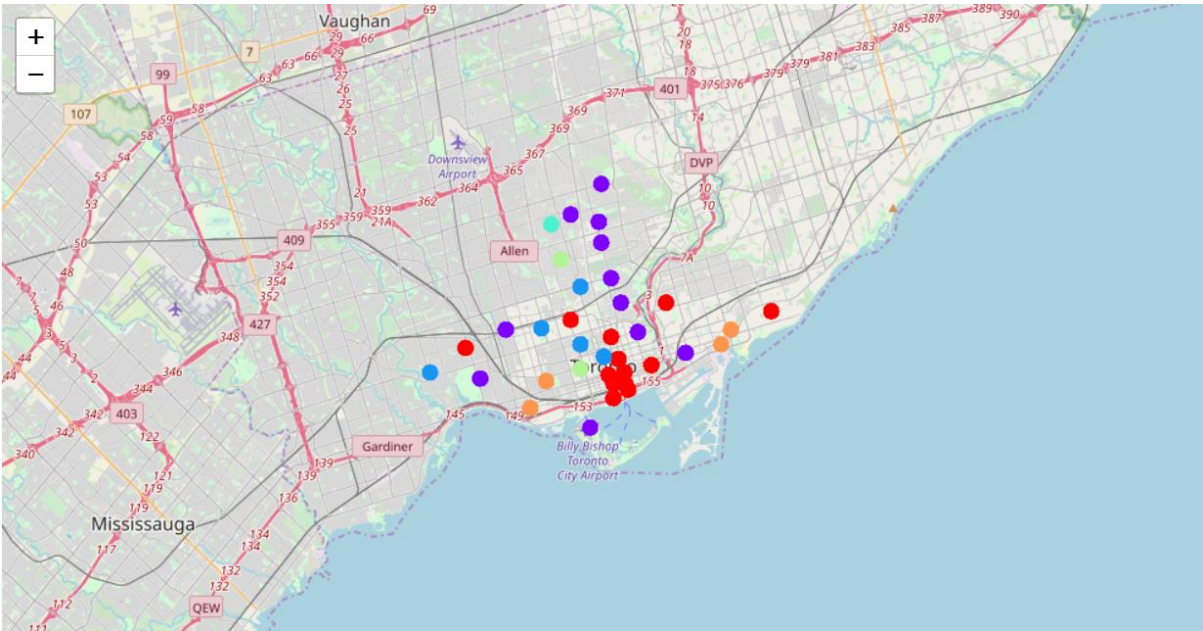
	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	Tori's Bakeshop	43.672114	-79.290331	Vegetarian / Vegan Restaurant
1	The Beaches	43.676357	-79.293031	Starbucks	43.680806	-79.285137	Coffee Shop
2	The Beaches	43.676357	-79.293031	Starbucks	43.669564	-79.301969	Coffee Shop
3	The Beaches	43.676357	-79.293031	Grinder	43.683073	-79.299875	Coffee Shop
4	The Beaches	43.676357	-79.293031	Tim Hortons	43.680799	-79.282907	Coffee Shop

Coffee Shops in Toronto

	Neighbourhood	Bakery	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café	Chinese Restaurant	Coffee Shop	...	Gaming Cafe	Gas Station	Grocery Store	Ice Cream Shop	Marijuana Dispensary
0	The Beaches	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
1	The Beaches	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0
2	The Beaches	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0
3	The Beaches	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0
4	The Beaches	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0

Then use this feature to group the neighbourhoods into clusters K-means clustering algorithm will be used to complete this task. And also, the Folium library to visualize the neighbourhoods in Toronto and its emerging clusters.

	Neighbourhood	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
0	Adelaide,King,Richmond	Coffee Shop	Café	Bar	Ice Cream Shop	Chinese Restaurant	Vegetarian / Vegan Restaurant	Dessert Shop	Bike Shop	Bookstore	Boutique
1	Berczy Park	Coffee Shop	Café	Bar	Vegetarian / Vegan Restaurant	Tea Room	Bike Shop	Bookstore	Boutique	Cafeteria	Chinese Restaurant
2	Brockton,Exhibition Place,Parkdale Village	Coffee Shop	Bakery	Bike Shop	Café	Food Truck	Tea Room	Dessert Shop	Bar	Bookstore	Boutique
3	Business Reply Mail Processing Centre 969 Eastern	Coffee Shop	Bakery	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café	Chinese Restaurant
4	CN Tower,Bathurst Quay,Island airport,Harbourf...	Coffee Shop	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café	Chinese Restaurant



4. Results

K-mean Cluster Using K-mean to clustering data area with less number of sushi bars Cluster 0

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 0, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

	Neighbourhood Latitude	Venue Longitude	Venue Category	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 Common Venue
0	43.676357	-79.290331	Vegetarian / Vegan Restaurant	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou
1	43.676357	-79.285137	Coffee Shop	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou
2	43.676357	-79.301969	Coffee Shop	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou
3	43.676357	-79.299875	Coffee Shop	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou
4	43.676357	-79.282907	Coffee Shop	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou
5	43.676357	-79.302124	Coffee Shop	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou
6	43.676357	-79.299733	Coffee Shop	0	Coffee Shop	Vegetarian / Vegan Restaurant	Café	French Restaurant	Dessert Shop	Bar	Bike Shop	Bookstore	Bou

Cluster 1

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 1, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

	Neighbourhood Latitude	Venue Longitude	Venue Category	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 Mo Common Venu
49	43.659526	-79.342565	Café	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri
50	43.659526	-79.342461	Coffee Shop	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri
51	43.659526	-79.341241	Coffee Shop	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri
52	43.659526	-79.341510	Coffee Shop	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri
53	43.659526	-79.342295	Coffee Shop	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri
54	43.659526	-79.338577	Coffee Shop	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri
55	43.659526	-79.353500	Coffee Shop	1	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteri

Cluster 2

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 2, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

	Neighbourhood Latitude	Venue Longitude	Venue Category	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 Co
175	43.686412	-79.398612	Coffee Shop	2	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cal
176	43.686412	-79.392148	Coffee Shop	2	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cal
177	43.686412	-79.395570	Coffee Shop	2	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cal
178	43.686412	-79.393281	Coffee Shop	2	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cal
179	43.686412	-79.394475	Coffee Shop	2	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cal
180	43.686412	-79.396840	Coffee Shop	2	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cal

Cluster 3

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 3, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

	Neighbourhood Latitude	Venue Longitude	Venue Category	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
749	43.711695	-79.413698	Coffee Shop	3	Coffee Shop	Tea Room	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café
750	43.711695	-79.411762	Coffee Shop	3	Coffee Shop	Tea Room	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café
751	43.711695	-79.413824	Coffee Shop	3	Coffee Shop	Tea Room	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café
752	43.711695	-79.409193	Coffee Shop	3	Coffee Shop	Tea Room	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café
753	43.711695	-79.430534	Coffee Shop	3	Coffee Shop	Tea Room	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café
754	43.711695	-79.405920	Coffee Shop	3	Coffee Shop	Tea Room	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria	Café

Cluster 4

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 4, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

	Neighbourhood Latitude	Venue Longitude	Venue Category	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
756	43.696948	-79.413698	Coffee Shop	4	Coffee Shop	Tea Room	Café	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
757	43.696948	-79.411762	Coffee Shop	4	Coffee Shop	Tea Room	Café	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
758	43.696948	-79.412803	Coffee Shop	4	Coffee Shop	Tea Room	Café	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
759	43.696948	-79.409193	Coffee Shop	4	Coffee Shop	Tea Room	Café	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
760	43.696948	-79.413824	Coffee Shop	4	Coffee Shop	Tea Room	Café	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
761	43.696948	-79.412601	Café	4	Coffee Shop	Tea Room	Café	Vegetarian / Vegan Restaurant	Bar	Bike Shop	Bookstore	Boutique	Cafeteria

Cluster 5

```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 5, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

	Neighbourhood Latitude	Venue Longitude	Venue Category	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
38	43.668999	-79.312404	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
39	43.668999	-79.308015	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
40	43.668999	-79.308204	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
41	43.668999	-79.309945	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
42	43.668999	-79.301969	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
43	43.668999	-79.302124	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
44	43.668999	-79.328110	Bakery	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria
45	43.668999	-79.306890	Coffee Shop	5	Coffee Shop	Bakery	Café	Tea Room	Bar	Bike Shop	Bookstore	Boutique	Cafeteria

Based on dataframe analysis above Cluster 3 and Cluster 5 areas are the best places to open a new Coffee Shop.

5. Discussion

In this section, I would be discussing the observations I have noted and the recommendation that I can make based on the results.

This analysis is performed on limited data. This may be right or may be wrong. But if good amount of data is available there is scope to come up with better results.

- There is high competition in The Beaches, The Danforth West and Riverdale so it is very risky to open business in these areas.
- Forest Hill North and Forest Hill West also have potential.
- It can be done more detailed analysis by adding other factors such as transportation, demographics of inhabitants.

6. Conclusion

Although all of the goals of this project were met there is definitely room for further improvement and development as noted above. However, the goals of the project were met and, with some more work, could easily be developed into a reliable application that could be used by businesses to decide on a location.