

# Capstone Project Report

## 1.Introduction

The client has just moved to from New York, US to Toronto, Canada. She would like to get some insights on the neighborhoods of Toronto, to find a place that is similar to her previous neighborhood, i.e. Flatiron in New York, to ensure smooth transition and her lifestyle.

## 2.Data

### Neighborhoods of Toronto

For this project, we first make a list of neighborhoods of Toronto and their geographical coordinates. The neighborhood names for Toronto will be scraped from Wikipedia website: [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) and geographical coordinates are obtained from: [https://cocl.us/Geospatial\\_data](https://cocl.us/Geospatial_data)

The processed list of neighborhoods is shown in the below figure.

	Postcode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	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
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848

## Venue data of the neighborhoods

The venue data of the neighborhoods in Toronto and of Flatiron, New York will be collected using Foursquare API. 'Explore' each neighborhood to get a list of up to 100 venues in the proximity.

For all neighborhoods, there are 336 venues categories returned, including: yoga studio, accessories store, farmer markets, electronics store, etc. The venue categories, converted into 'frequency' will define the characteristics of each neighborhood, for example, the profile of Flatiron neighborhood in New York below.

```
----Flatiron----
```

	venue	freq
0	Gym / Fitness Center	0.06
1	American Restaurant	0.06
2	New American Restaurant	0.05
3	Cycle Studio	0.04
4	Japanese Restaurant	0.04
5	Hotel	0.03
6	Café	0.03
7	Wine Shop	0.03
8	Mediterranean Restaurant	0.02
9	Bookstore	0.02

## 3. Methodology

In order to find neighborhoods that resemble the Flatiron neighborhood in New York, a Flatiron-index will be calculated for each neighborhood. The Flatiron-index for a neighborhood is the sum of the product of its venue category frequencies and those of Flatiron's. The higher the index means there are more venues of the popular venue categories of Flatiron in that particular neighborhood.

## 4. Results

The Flatiron-index for all the neighborhoods are calculated, and the top 15 neighborhoods are shown in the below table. The neighborhoods with less than 20 venues returned are filtered out.

	Flatiron	Venue
Neighborhood		
Stn A PO Boxes 25 The Esplanade	0.012500	100
Commerce Court, Victoria Hotel	0.012400	100
First Canadian Place, Underground city	0.012300	100
Mimico NW, The Queensway West, Royal York South West, South of Bloor	0.012143	56
Design Exchange, Toronto Dominion Centre	0.011900	100
St. James Town	0.011900	100
Forest Hill North, Forest Hill West	0.011333	45
Don Mills North	0.011290	31
Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West	0.010741	81
Flemingdon Park, Don Mills South	0.010698	43
Downsview Northwest	0.010606	33
Adelaide, King, Richmond	0.010300	100
Berczy Park	0.010000	100
Brockton, Exhibition Place, Parkdale Village	0.009900	100
Church and Wellesley	0.009900	100

The profile of the top 2 neighborhoods, Stn A PO and Commerce Court are shown below.

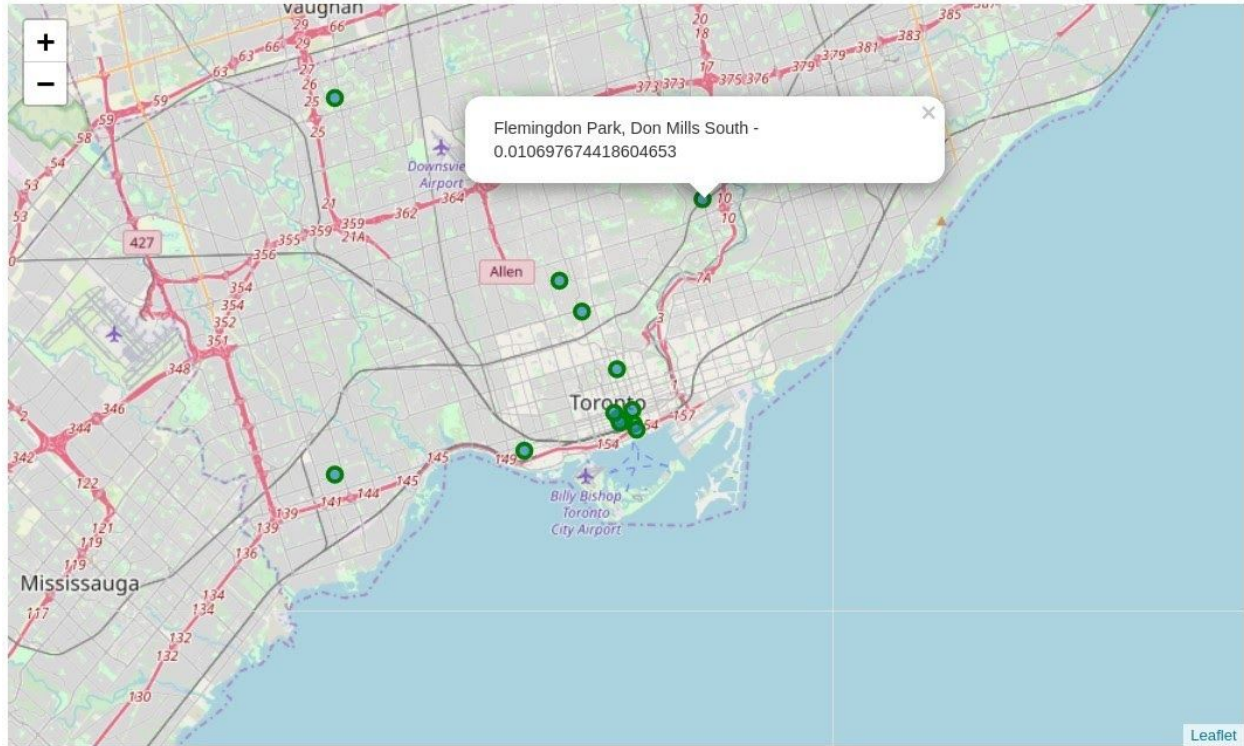
----Stn A P0 Boxes 25 The Esplanade----

	venue	freq
0	Coffee Shop	0.12
1	Café	0.07
2	Hotel	0.05
3	Restaurant	0.04
4	Beer Bar	0.04
5	Japanese Restaurant	0.04
6	Breakfast Spot	0.03
7	Seafood Restaurant	0.03
8	Gastropub	0.03
9	Art Gallery	0.02

----Commerce Court, Victoria Hotel----

	venue	freq
0	Coffee Shop	0.10
1	Café	0.07
2	Hotel	0.05
3	Restaurant	0.04
4	Steakhouse	0.03
5	Gastropub	0.03
6	Concert Hall	0.03
7	Beer Bar	0.03
8	Japanese Restaurant	0.03
9	Vegetarian / Vegan Restaurant	0.02

And finally, a map of Toronto is created with the top 15 neighborhoods marked, and their Flatiron-index as labels.



## 5. Discussion

The similarity index method is effective in showing the neighborhoods with characteristics matching the Flatiron neighborhood in New York - having coffee shops, restaurant, hotels, etc.

However, some categories such as 'hotel' are not relevant to residents. It would be more effective to drop those venue categories, and keep only the categories that the client would use for consideration, in order to better filter the neighborhoods of the client's interest.

Furthermore, the index does not take into account the count of venues, so neighborhoods that have few venues (less than 20) were dropped from the data to avoid misleading results.

## 6. Conclusion

In this project, the venue data of neighborhoods in Toronto were collected, and an index was calculated for each neighborhood that reflects their similarity to the Flatiron neighborhood in New York. A list and map of the top 15 neighborhoods were made for the client's consideration.