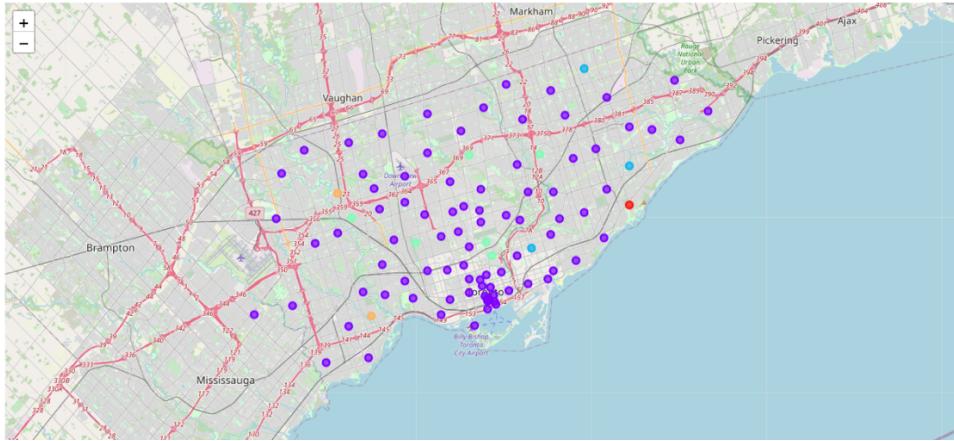


# Capstone Project

Identify Suitable Neighbourhoods for New Residents and Franchise Owners



## Introduction:

- This project will focus on the problems people face when choosing a neighbourhood to live in Toronto based on their preference of public/private lifestyle and what amenities and venues (if any) are closest and most visited. This will focus on new residents and people who wish to relocate within the city. While locations can be scouted or mapped accordingly beforehand, there is no way of knowing the traffic of amenities and venues near your location (i.e. parks, restaurants) and whether they are commonly visited. This project will list the most common venues for each cluster of neighbourhoods. This project will also target urban planners and franchise owners who want to expand their businesses to new locations, based on frequency of traffic in local venues. Decisions can therefore be made according to the preferred location for residence or business operation.

## **Data:**

- Publicly provided lists of Toronto neighbourhoods will be used to create clusters from Wikipedia - [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- Longitude and Latitude Data will also be used and mapped accordingly - [http://cocl.us/Gespatial\\_data](http://cocl.us/Gespatial_data)
- Also venue information and location data will be pulled from Foursquare using my account.
  - I will import venue addresses and their categories using API information. Then link these to all listed neighbourhood

## Methodology

- First, I collected all nearby venues using Foursquare's live database and returned the relevant information for each nearby venue. Then I scraped all neighbourhood locations from the publicly listed page using Beautiful Soup in Python.

- Then checked the data frame results by displaying top (head) venues:

Checking Dataframe Result

```
In [30]: toronto_venues.head(10)
Out[30]:
```

	Neighborhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.753259	-79.399656	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
2	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena
3	Victoria Village	43.725882	-79.315572	Tim Hortons	43.725517	-79.313103	Coffee Shop
4	Victoria Village	43.725882	-79.315572	Portugil	43.725819	-79.312785	Portuguese Restaurant
5	Victoria Village	43.725882	-79.315572	The Frig	43.727051	-79.317418	French Restaurant
6	Victoria Village	43.725882	-79.315572	Eglington Ave E & Sloane Ave/Bermondsey Rd	43.726086	-79.313620	Intersection
7	Victoria Village	43.725882	-79.315572	Pizza Nova	43.725824	-79.312860	Pizza Place
8	Regent Park, Harbourfront	43.654260	-79.360636	Rosette Desserts	43.653447	-79.362017	Bakery
9	Regent Park, Harbourfront	43.654260	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop

- I verified how many venues returned:

```
In [32]: toronto_venues.groupby('Neighborhood').count()
Out[32]:
```

Neighborhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Agincourt	5	5	5	5	5	5
Alderwood, Long Branch	10	10	10	10	10	10
Bathurst Manor, Wilson Heights, Downsview North	19	19	19	19	19	19
Bayview Village	4	4	4	4	4	4
Bedford Park, Lawrence Manor East	24	24	24	24	24	24
Berczy Park	57	57	57	57	57	57
Birch Cliff, Cliffside West	4	4	4	4	4	4
Brockton, Parkdale Village, Exhibition Place	23	23	23	23	23	23
Business reply mail Processing CentrE	18	18	18	18	18	18
CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island Airport	17	17	17	17	17	17

- Next, I analyzed the neighborhoods using “one hot encoding” method.
- Then added the neighbourhood column back to the data frame and then moved it to the first column:

	Accessories Store	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop	Aquarium	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Auto Garage	Auto Workshop	BBQ Joint	Baby Store	Bagel Shop	Bakery	Bank	Bar	Baseball Field	Baseball Stadi
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- Next I grouped the neighbourhoods using the .groupby function:

Grouping	
In [37]:	
toronto_grouped = toronto_onehot.groupby('Neighborhood').mean().reset_index()	toronto_grouped
4	Lawrence Manor East
5	Berczy Park
6	Birch Cliff, Cliffside West
7	Brockton, Parkdale Village, Exhibition Place
8	Business reply mail Processing Centre
9	CN Tower, King and Spadina, Railway Lands, Har...
10	Caledonia-Fairbanks
	Canada Post

- I printed the neighborhoods according to Top Venue: (List continues on)

----Agincourt----	
	venue freq
0	Latin American Restaurant 0.2
1	Breakfast Spot 0.2
2	Lounge 0.2
3	Skating Rink 0.2
4	Clothing Store 0.2
----Alderwood, Long Branch----	
	venue freq
0	Pizza Place 0.2
1	Dance Studio 0.1
2	Skating Rink 0.1
3	Sandwich Place 0.1
4	Athletics & Sports 0.1
----Bathurst Manor, Wilson Heights, Downsview North----	
	venue freq

- Then sorted the data in alphabetical order according to neighborhood name:

	Neighborhood	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	Agincourt	Lounge	Skating Rink	Latin American Restaurant	Breakfast Spot	Clothing Store	Yoga Studio	Drugstore	Distribution Center	Dog Run	Doner Restaurant
1	Alderwood, Long Branch	Pizza Place	Sandwich Place	Skating Rink	Pharmacy	Gym	Dance Studio	Coffee Shop	Athletics & Sports	Pub	Dim Sum Restaurant
2	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Fried Chicken Joint	Bridal Shop	Sandwich Place	Diner	Restaurant	Middle Eastern Restaurant	Supermarket	Sushi Restaurant
3	Bayview Village	Café	Bank	Chinese Restaurant	Japanese Restaurant	Yoga Studio	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop
4	Bedford Park, Lawrence Manor East	Sandwich Place	Italian Restaurant	Restaurant	Coffee Shop	Sushi Restaurant	Greek Restaurant	Thai Restaurant	Liquor Store	Comfort Food Restaurant	Juice Bar

- This portion of research involves my cluttering methods used to begin categorizing/ordering the neighborhoods.
- Using clustering by Run K means and creating an Array to organize the process:

```
array([1, 1, 1, 1, 1, 1, 1, 1, 1, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1,
       1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 4, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 3, 1, 1, 1, 1, 1, 1, 1, 4, 1, 1, 1, 3, 1, 1, 1, 1, 1, 1, 3,
       1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 3,
       1, 1, 1, 1, 1, 3], dtype=int32)
```

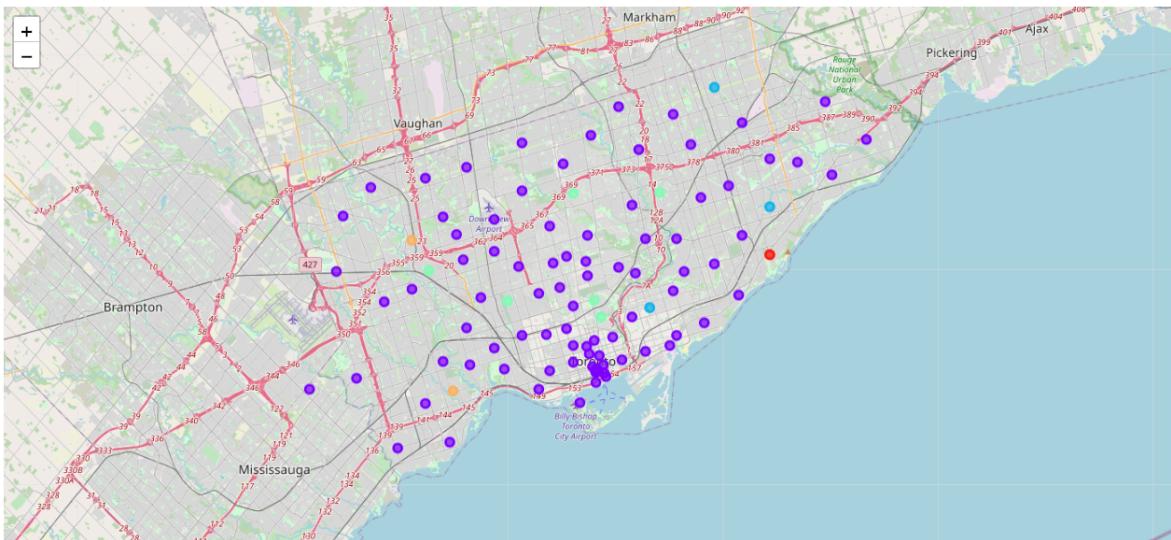
- Now I created a new data frame to include clusters AND the top venues for each neighbourhood:

	Postal code	Borough	Neighborhood	Latitude	Longitude	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
0	M3A	North York	Parkwoods	43.753259	-79.329656	3.0	Food & Drink Shop	Park	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Yoga Studio
1	M4A	North York	Victoria Village	43.725882	-79.315572	1.0	Pizza Place	Coffee Shop	Intersection	French Restaurant	Portuguese Restaurant	Hockey Arena	Eastern European Restaurant	Electronics Store	Drugstore	Dessert Shop
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	1.0	Coffee Shop	Bakery	Pub	Park	Breakfast Spot	Café	Theater	Yoga Studio	Cosmetics Shop	Shoe Store
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	1.0	Clothing Store	Furniture / Home Store	Coffee Shop	Event Space	Shoe Store	Sporting Goods Shop	Miscellaneous Shop	Arts & Crafts Store	Accessories Store	Boutique
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	1.0	Coffee Shop	Sushi Restaurant	Diner	Burger Joint	Burrito Place	Juice Bar	Café	Japanese Restaurant	Italian Restaurant	Beer Bar

- Then dropped unnecessary Data using .dropna and created a new variable to merge the labeled clusters and defined colors, radius, size etc.:

```
toronto_merged=toronto_merged.dropna()
toronto_merged['Cluster_Labels'] = toronto_merged.Cluster_Labels.astype(int)
```

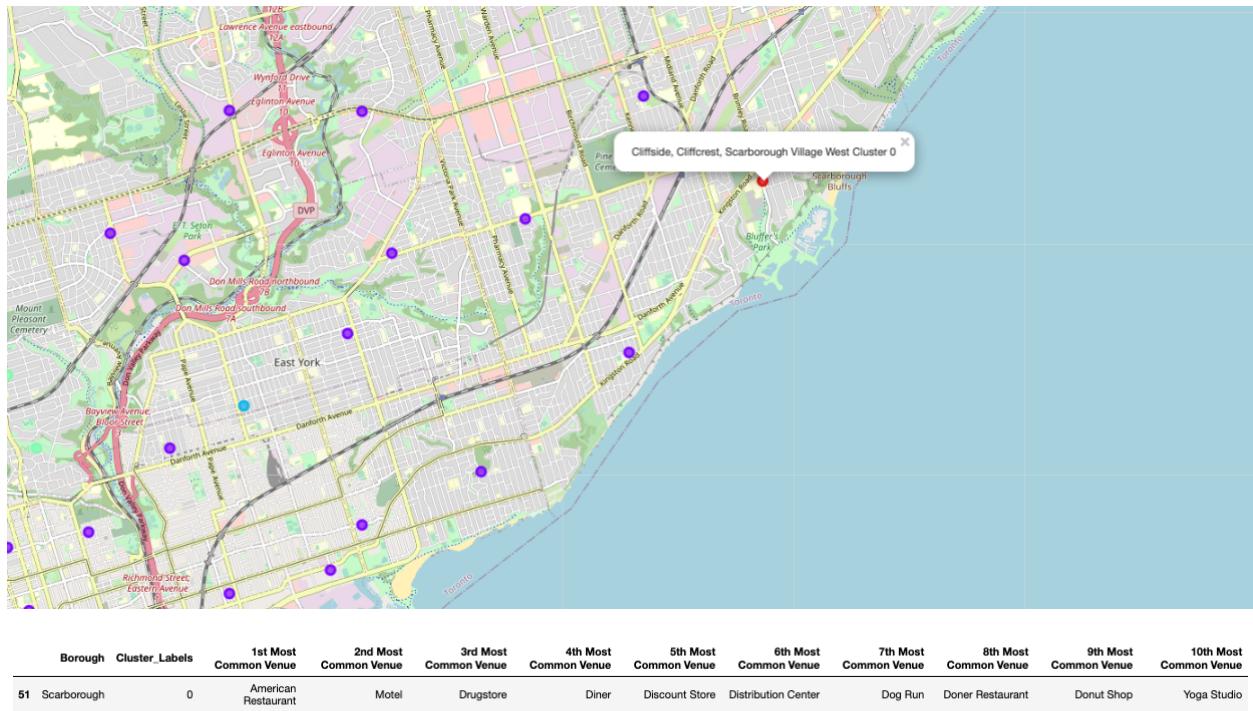
- With this data, I created a map, set color scheme and added all the markers.



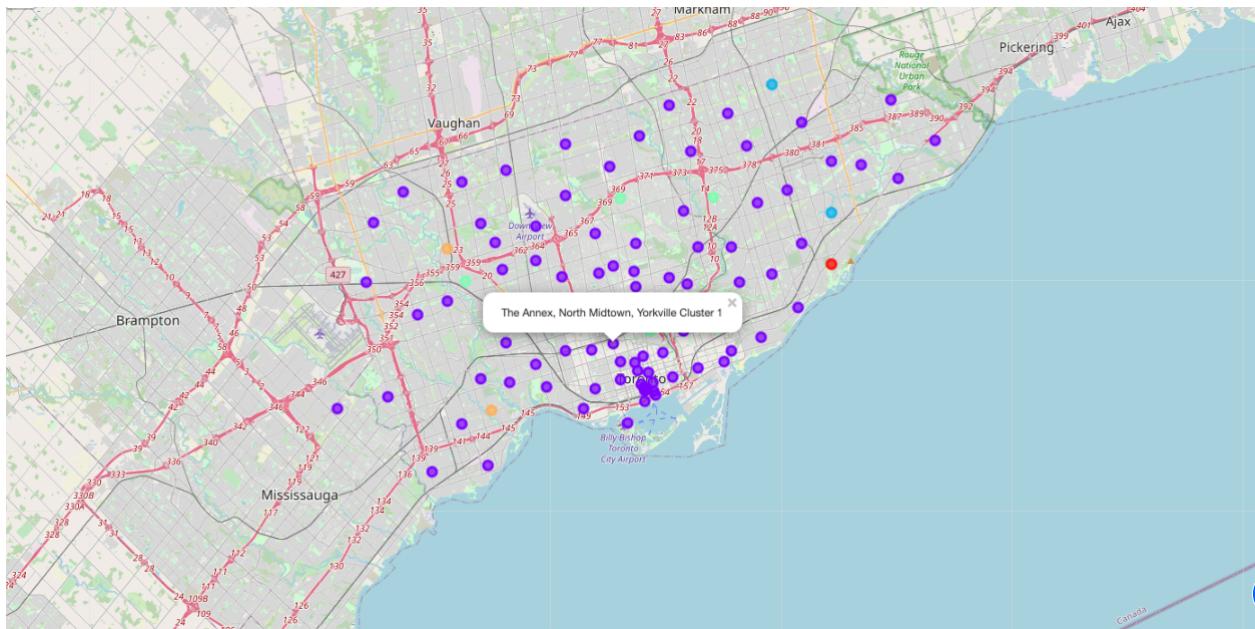
## Results:

- These new clusters created a visual for new residents and franchise owners while plotting their preferred neighborhood based on the according clusters. Each option lists the Most Common Venues from 1<sup>st</sup> to 10<sup>th</sup>:

- Cluster 0 (labelled zero according to variable name) :



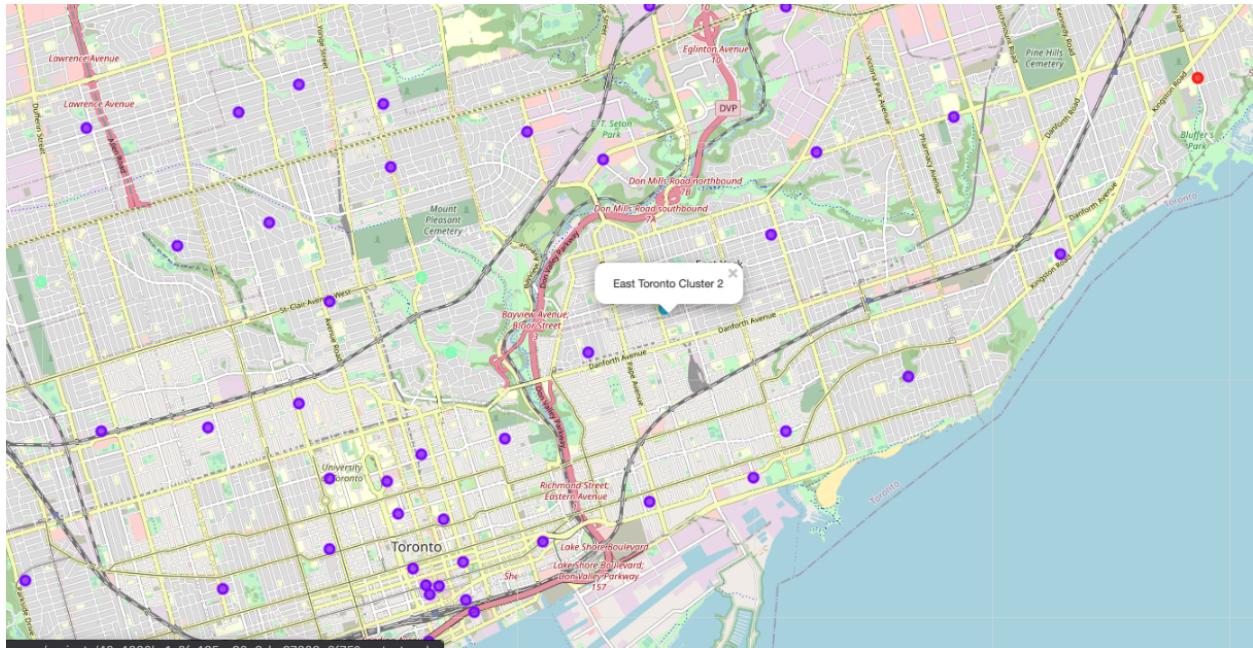
- Cluster 1:



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
1	North York	1	Pizza Place	Coffee Shop	Intersection	French Restaurant	Portuguese Restaurant	Hockey Arena	Eastern European Restaurant	Electronics Store	Drugstore
2	Downtown Toronto	1	Coffee Shop	Bakery	Pub	Park	Breakfast Spot	Café	Theater	Yoga Studio	Cosmetics Shop
3	North York	1	Clothing Store	Furniture / Home Store	Coffee Shop	Event Space	Shoe Store	Sporting Goods Shop	Miscellaneous Shop	Arts & Crafts Store	Accessories Store
4	Downtown Toronto	1	Coffee Shop	Sushi Restaurant	Diner	Burger Joint	Burrito Place	Juice Bar	Café	Japanese Restaurant	Italian Restaurant
6	Scarborough	1	Fast Food Restaurant	Print Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Yoga Studio
7	North York	1	Gym	Beer Store	Restaurant	Japanese Restaurant	Coffee Shop	Asian Restaurant	Dim Sum Restaurant	Sandwich Place	Bike Shop
8	East York	1	Pizza Place	Gastropub	Pharmacy	Gym / Fitness Center	Breakfast Spot	Fast Food Restaurant	Intersection	Bank	Athletics & Sports
Downtown											
Midtown											

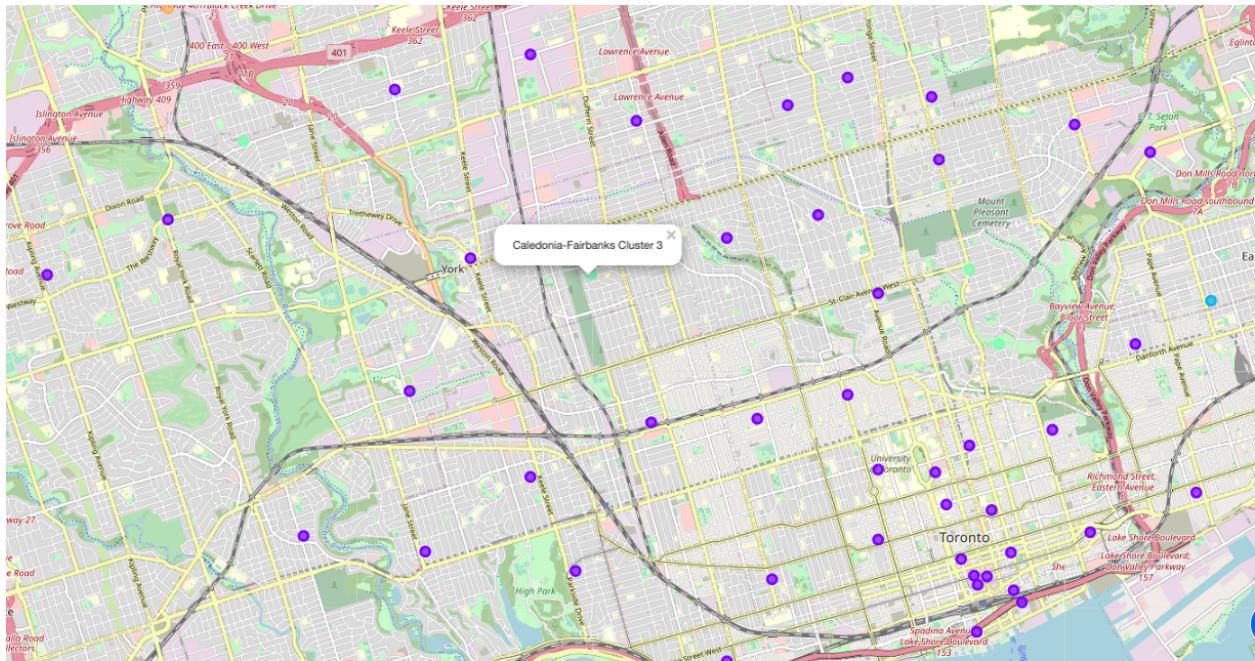
(List continues, see notebook)

- Cluster 2



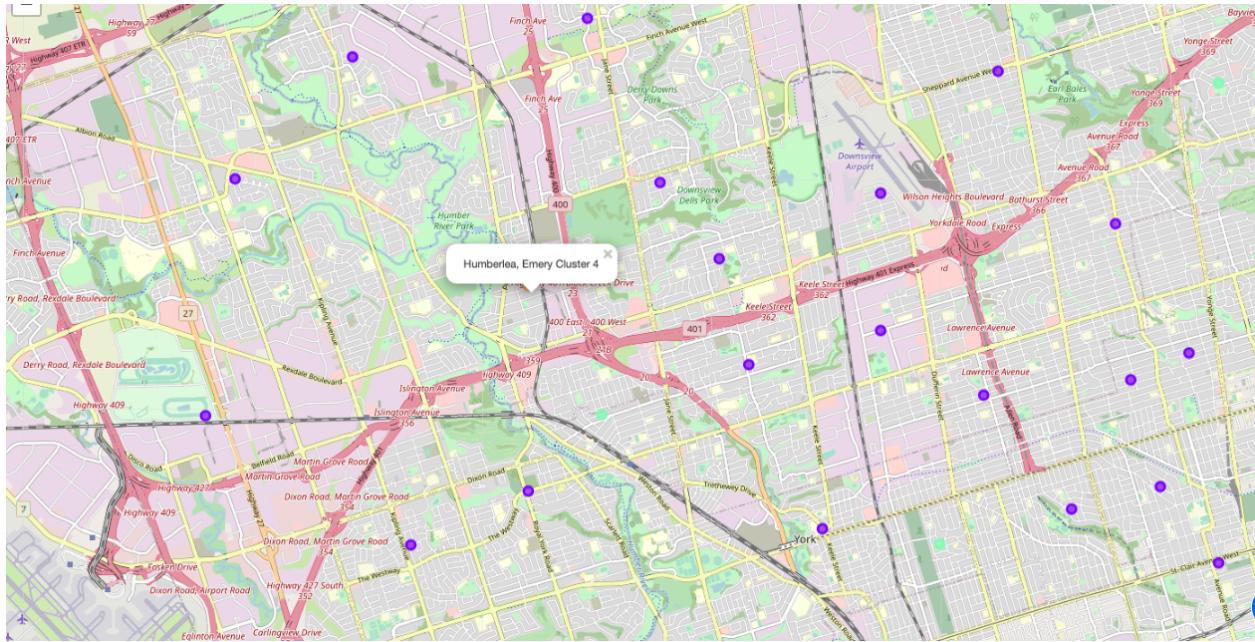
Borough	Cluster_Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	
32	Scarborough	2	Playground	Yoga Studio	Donut Shop	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Drugstore
35	East York	2	Park	Coffee Shop	Convenience Store	Yoga Studio	Donut Shop	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant
85	Scarborough	2	Playground	Park	Coffee Shop	Yoga Studio	Donut Shop	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant

- 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	10th Most Common Venue
0	North York	3 Food & Drink Shop	Park	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Yoga Studio
21	York	3 Park	Women's Store	Spa	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Electronics Store
64	York	3 Park	Yoga Studio	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Eastern European Restaurant
66	North York	3 Park	Convenience Store	Bank	Bar	Yoga Studio	Drugstore	Distribution Center	Dog Run	Doner Restaurant	Donut Shop
83	Central Toronto	3 Park	Yoga Studio	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Eastern European Restaurant
91	Downtown Toronto	3 Park	Playground	Trail	Yoga Studio	Doner Restaurant	Dim Sum Restaurant	Diner	Discount Store	Distribution Center	Dog Run

- 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	
57	North York	4	Paper / Office Supplies Store	Baseball Field	Yoga Studio	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Eastern European Restaurant
101	Etobicoke	4	Baseball Field	Yoga Studio	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Eastern European Restaurant	Filipino Restaurant

## **Discussion:**

- Based on the labelled clusters, users have a graphic understanding of commonly visited venues according to neighbourhood, according to their relationship with nearby clusters. Important note - the selected locations for each cluster are approximate and chosen to highlight the surrounding area.
- However it is clear that certain clusters (specifically 1 and 2) involve predominant retail and restaurant venues as the most commonly visited according to Foursquare. However the outer clusters of Toronto's main city area, list parks and recreation more commonly in their top venues. This provides a unique recommendation for new resident who yearn for fresh living space within a city centre (or out if preferred). Noting the commonly visited outdoor venue (on the contrary) will signify higher tourist traffic, which may not be desirable. Regardless it provides the potential resident with abundant information regarding public venues.
- Recommendations for franchise owners would involve the avoidance of clustered areas that already contain similar businesses. However usage of parks and recreation spaces (especially those commonly visited in downtown clusters) may prove useful if they acquire permits for sales. This is dependent on the owner's preference, however the results illustrate a perspective on the difference of commonly visited venues away from the city centre. Specifically useful to note the remains of common food venues regardless of the neighborhood or its distance from the city centre.

## Conclusion

- The details in this report only remotely reflect multiple moving constants. Therefore it may be difficult to apply the findings while many other factors should be considered. It is also limited to the Foursquare inputs, which may be subject to change. There is also the important weather factors which may change the visitation of most venues (i.e. indoor vs outdoor, according to the season). Thus I believe including a recent seasonal report may better provide the customer with substantial knowledge when seeking franchise locations. However in terms of social traffic and trending venues, I believe this report will provide residents and expanding franchisers with useful geographical and social information.