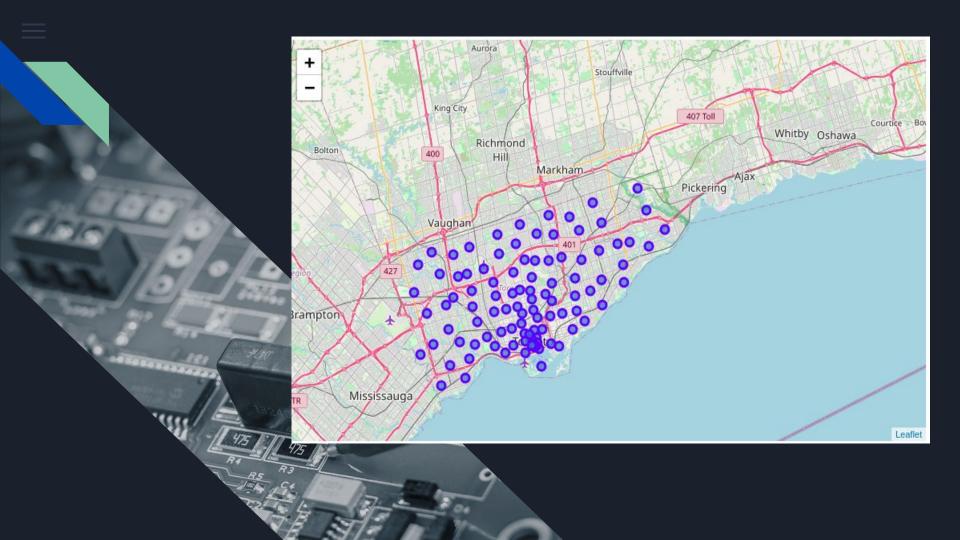


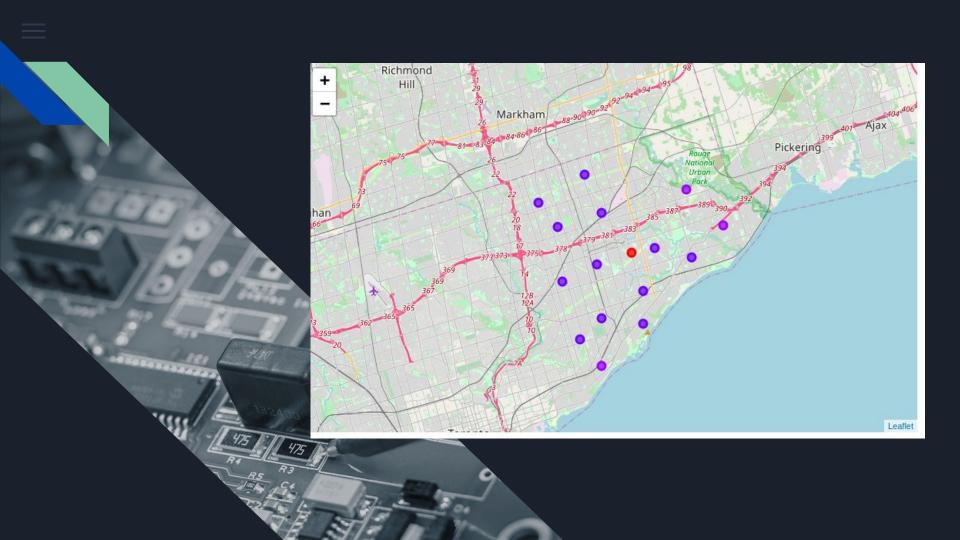
Capstone Project

## Methodology

- Machine Learning(k-means clustering)
  - 2. Exploratory Data Analysis

3. Inferential Analysis





## Datasets

	name	categories	lat	lng
0	Disney Store	Toy / Game Store	43.775537	-79.256833
1	SEPHORA	Cosmetics Shop	43.775017	-79.258109
2	American Eagle Outfitters	Clothing Store	43.776012	-79.258334
3	St. Andrews Fish & Chips	Fish & Chips Shop	43.771865	-79.252645
4	Tommy Hilfiger	Clothing Store	43.776015	-79.257369

$\Box$		Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
	Neighborhood						
	Agincourt	31	31	31	31	31	31
	Alderwood / Long Branch	10	10	10	10	10	10
	Bathurst Manor / Wilson Heights / Downsview North	24	24	24	24	24	24
	Bayview Village	5	5	5	5	5	5
	Bedford Park / Lawrence Manor East	25	25	25	25	25	25

	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	Chinese Restaurant	Shopping Mall	Pool	Latin American Restaurant	Supermarket	Sushi Restaurant	Motorcycle Shop	Breakfast Spot	Sandwich Place	Print Shop
1	Alderwood / Long Branch	Convenience Store	Pizza Place	Gas Station	Pub	Pharmacy	Skating Rink	Coffee Shop	Sandwich Place	Athletics & Sports	Gym
2	Bathurst Manor / Wilson Heights / Downsview North	Bank	Coffee Shop	Bridal Shop	Trail	Shopping Mall	Middle Eastern Restaurant	Sandwich Place	Fried Chicken Joint	Supermarket	Sushi Restaurant
3	Bayview Village	Park	Construction & Landscaping	Trail	Farm	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Falafel Restaurant
4	Bedford Park / Lawrence Manor East	Restaurant	Coffee Shop	Sandwich Place	Italian Restaurant	Intersection	Greek Restaurant	Liquor Store	Sports Club	Juice Bar	Boutique

### Clustering

E>

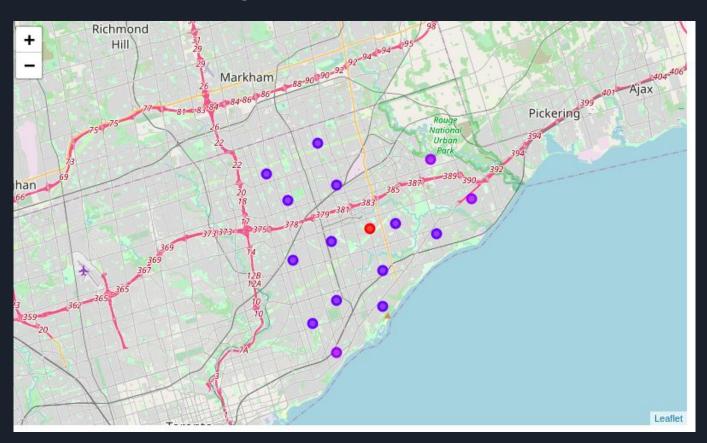
```
[ ] # Using K-Means to cluster neighborhood into 3 clusters
    Scarborough_grouped_clustering = Scarborough_grouped.drop('Neighborhood', 1)
    kmeans = KMeans(n_clusters=3, random_state=0).fit(Scarborough_grouped_clustering)
    neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

Scarborough_merged =df_2.iloc[:16,:]

# merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhood
Scarborough_merged = Scarborough_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')
Scarborough_merged.head()
```

}	Postalco	de 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 N	1B Scarborough	Malvern / Rouge	43.808626	-79.189913	1	Park	Trail	Women's Store	Farm	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Falafel Restaurant
	1 M	1C Scarborough	Rouge Hill / Port Union / Highland Creek	43.785779	-79.157368	1	Bar	Park	Women's Store	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Falafel Restaurant	Farm
	2 N	1E Scarborough	Guildwood / Morningside / West Hill		-79.185284	2	Pizza Place	Fast Food Restaurant	Bank	Park	Coffee Shop	Medical Center	Supermarket	Laundromat	Fried Chicken Joint	Sports Bar
	3 N	1G Scarborough	Woburn	43.771545	-79.218135	2	Coffee Shop	Park	Business Service	Women's Store	Farmers Market	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Falafel Restaurant
	4 N	1H Scarborough	Cedarbrae	43.768791	-79.238813	0	Thai Restaurant	Athletics & Sports	Bank	Hakka Restaurant	Gas Station	Fish Market	Fish & Chips Shop	Dumpling Restaurant	Food	Eastern European Restaurant

#### Result (selected neighbourhoods)



# THANK YOU