

# The Battle of Neighborhood - Scarborough

Capstone project

August, 2020

# Background - Scarborough

- ▶ Location: eastern part of Toronto
- ▶ Population: over 630,000 (2016 census)
  - ▶ 57% foreign-born
  - ▶ 55.9% Asian
- ▶ Business
  - ▶ Agincourt: heavily concentration of Chinese businesses and restaurants
  - ▶ Cuisines from all of world are everywhere
- ▶ The most diverse and multicultural area in the Greater Toronto

# Business Problem

- ▶ An investor plans to open a restaurant, what's the answer for the following questions:
  - ▶ Target consumers
  - ▶ Market situation
  - ▶ Competing restaurants

# Data

- ▶ Wikipedia page: pull region information
  - ▶ [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
- ▶ Geopy library
  - ▶ Latitude and longitude values
- ▶ Foursquare API
  - ▶ It is a location data provider with information about all manner of venues
  - ▶ Information includes venue names, locations, latitudes, longitudes etc.
  - ▶ Within 1000-meter radius will be selected

# Methodology - data cleansing

	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

- Dataframe: Postalcode, Borough, and Neighborhood
- Latitude and longitude will be merged with above table
- Utilize Foursquare API to explore the neighborhoods

	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

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Malvern, Rouge	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
2	Guildwood, Morningside, West Hill	43.763573	-79.188711	RBC Royal Bank	43.766790	-79.191151	Bank
3	Guildwood, Morningside, West Hill	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	Electronics Store
4	Guildwood, Morningside, West Hill	43.763573	-79.188711	Sail Sushi	43.765951	-79.191275	Restaurant
5	Guildwood, Morningside, West Hill	43.763573	-79.188711	Big Bite Burrito	43.766299	-79.190720	Mexican Restaurant

# Methodology - data exploration



# Methodology - Machine learning

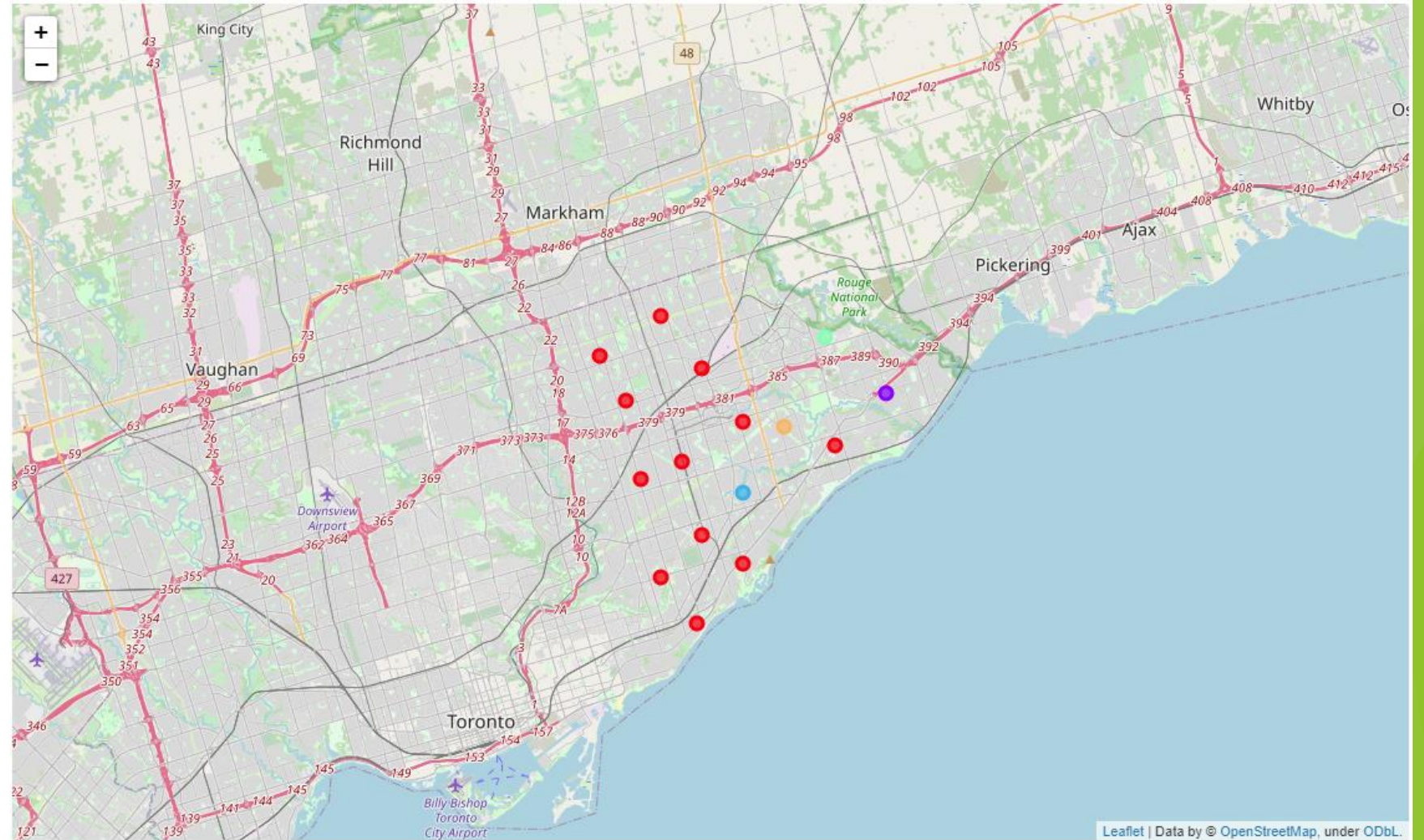
- ▶ One-hot encoding
  - ▶ Categorized data before fitting a ML model

[illegible]



# Methodology - Machine learning

- K-mean clustering
  - 5 clusters





# Methodology - data analysis

## ► 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
5	Scarborough	0.0	Spa	Playground	Vietnamese Restaurant	College Stadium	Grocery Store	General Entertainment	Gas Station	Gaming Cafe	Fried Chicken Joint	Fast Food Restaurant
14	Scarborough	0.0	Playground	Park	Vietnamese Restaurant	College Stadium	Grocery Store	General Entertainment	Gas Station	Gaming Cafe	Fried Chicken Joint	Fast Food Restaurant

# Methodology - data analysis

## ► 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
2	Scarborough	1.0	Intersection	Bank	Restaurant	Rental Car Location	Breakfast Spot	Electronics Store	Medical Center	Mexican Restaurant	Vietnamese Restaurant	Convenience Store
3	Scarborough	1.0	Coffee Shop	Soccer Field	Korean Restaurant	Indian Restaurant	Auto Garage	Department Store	Grocery Store	General Entertainment	Athletics & Sports	Gas Station
4	Scarborough	1.0	Thai Restaurant	Athletics & Sports	Hakka Restaurant	Bakery	Bank	Gas Station	Caribbean Restaurant	Lounge	Fried Chicken Joint	Discount Store
6	Scarborough	1.0	Discount Store	Chinese Restaurant	Coffee Shop	Department Store	Vietnamese Restaurant	Convenience Store	Hakka Restaurant	Grocery Store	General Entertainment	Gas Station
7	Scarborough	1.0	Bakery	Ice Cream Shop	Soccer Field	Intersection	Bus Station	Metro Station	Department Store	Grocery Store	General Entertainment	Gas Station
9	Scarborough	1.0	College Stadium	General Entertainment	Skating Rink	Café	Vietnamese Restaurant	Grocery Store	Gas Station	Gaming Cafe	Fried Chicken Joint	Fast Food Restaurant
10	Scarborough	1.0	Indian Restaurant	Pet Store	Chinese Restaurant	Gaming Cafe	Light Rail Station	Vietnamese Restaurant	Restaurant	Coffee Shop	Gas Station	Soccer Field
11	Scarborough	1.0	Middle Eastern Restaurant	Auto Garage	Bakery	Sandwich Place	Breakfast Spot	Vietnamese Restaurant	Grocery Store	General Entertainment	Gas Station	Gaming Cafe
12	Scarborough	1.0	Skating Rink	Latin American Restaurant	Breakfast Spot	Lounge	Vietnamese Restaurant	Convenience Store	General Entertainment	Gas Station	Gaming Cafe	Fried Chicken Joint
13	Scarborough	1.0	Pizza Place	Convenience Store	Thai Restaurant	Italian Restaurant	Bank	Gas Station	Fried Chicken Joint	Fast Food Restaurant	Chinese Restaurant	Noodle House
15	Scarborough	1.0	Fast Food Restaurant	Coffee Shop	Chinese Restaurant	Indian Restaurant	Breakfast Spot	Discount Store	Pharmacy	Pizza Place	Grocery Store	Sandwich Place

# Methodology - data analysis

## ► Cluster 3,4,5

	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
0	Scarborough	2.0	Fast Food Restaurant	Vietnamese Restaurant	Thai Restaurant	Hakka Restaurant	Grocery Store	General Entertainment	Gas Station	Gaming Cafe	Fried Chicken Joint	Electronics Store

	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
8	Scarborough	3.0	American Restaurant	Intersection	Motel	Convenience Store	Grocery Store	General Entertainment	Gas Station	Gaming Cafe	Fried Chicken Joint	Fast Food Restaurant

	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	4.0	Moving Target	Bar	Vietnamese Restaurant	Grocery Store	General Entertainment	Gas Station	Gaming Cafe	Fried Chicken Joint	Fast Food Restaurant	Electronics Store

# Results

- ▶ Cluster 2 has the most neighborhood
- ▶ Cluster 1 has 2 and cluster 3, 4 and 5 has one neighborhood

# Conclusions

- ▶ Utilizing different data for analysis
  - ▶ Wiki page, Foursquare API
- ▶ Machine learning used for clustering
  - ▶ One-hot encoding, K-mean clustering
- ▶ Possibility to open a new restaurant in Scarborough
  - ▶ 1<sup>st</sup> option: Golden Mile, Clairlea and Oakridge (neighborhood 7)
  - ▶ 2<sup>nd</sup> option: Birch Cliff and Cliffside West (neighborhood 9)



# Discussion

- ▶ In neighborhood 7, no completion. Neighborhood 9, only 1 Vietnamese restaurant.
- ▶ More factors could be considered for the research
  - ▶ Income, age, marital status, etc.