

IBM Coursera Capstone

The Battle of Neighborhoods

By Jerry Lin

13, April 2020

Introduction

This is the report for the IBM Data Science capstone project that I put together for the IBM Data Science Certificate Course. This report contains the week 4 assignment of Applied Data Science Capstone. I have prepared a report on looking at the data of Toronto, Canada

Problem

Toronto is a capital city of Canada, with various business opportunities and venues. This city is very highly market competitive and doing business there is highly priced. Businesses also need to choose their location carefully in order to have some success themselves.

Target Audience

This research is produced by myself, along with National Finance Company. We are to locate and recommend to business on which regions of neighborhoods in Toronto will fit the best.

Data Description

In this data we would have to explore, segment, and cluster the neighborhood of Toronto. We can get the data from the Wikipedia page, since it is not available at any other resources at all. That page has all the information and therefore need to organize it. That data has coordinates for each of the neighborhoods as well as residences, and various names.

The variables that it has are: Neighborhood, Neighborhood latitude, Neighborhood longitude, Venue name, Venue Category, Venue Latitude, and Venue longitude.

	Postalcode	Borough	Neighbourhood	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

List of venues:

The Beaches
The Danforth West / Riverdale
India Bazaar / The Beaches West
Studio District
Lawrence Park
Davisville North
North Toronto West

Davisville
Moore Park / Summerhill East
Summerhill West / Rathnelly / South Hill / Forest Hill SE / Deer Park
Rosedale
St. James Town / Cabbagetown
Church and Wellesley
Regent Park / Harbourfront
Garden District, Ryerson
St. James Town
Berczy Park
Central Bay Street
Richmond / Adelaide / King
Harbourfront East / Union Station / Toronto Islands
Toronto Dominion Centre / Design Exchange
Commerce Court / Victoria Hotel
Roselawn

Venues with neighborhood:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	Glen Manor Ravine	43.676821	-79.293942	Trail
1	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health Food Store
2	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	Pub
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborhood
4	The Danforth West / Riverdale	43.679557	-79.352188	Pantheon	43.677621	-79.351434	Greek Restaurant

Venues with specific neighborhood listed with latitude and longitude:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
	Berczy Park	56	56	56	56	56	56
	Brockton / Parkdale Village / Exhibition Place	23	23	23	23	23	23
	Business reply mail Processing CentrE	16	16	16	16	16	16
	CN Tower / King and Spadina / Railway Lands / Harbourfront West / Bathurst\ n Quay / South Niagara / Island airport	14	14	14	14	14	14
	Central Bay Street	64	64	64	64	64	64
	Christie	18	18	18	18	18	18
	Church and Wellesley	77	77	77	77	77	77
	Commerce Court / Victoria Hotel	100	100	100	100	100	100
	Davisville	33	33	33	33	33	33
	Davisville North	8	8	8	8	8	8
	Dufferin / Dovercourt Village	16	16	16	16	16	16
	First Canadian Place / Underground city	100	100	100	100	100	100
	Forest Hill North & West	4	4	4	4	4	4
	Garden District, Ryerson	100	100	100	100	100	100
	Harbourfront East / Union Station / Toronto Islands	100	100	100	100	100	100

Top venues:

	Cluster Labels	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	0	Berczy Park	Coffee Shop	Beer Bar	Restaurant	Cheese Shop	Cocktail Bar	Bakery	Italian Restaurant	Seafood Restaurant	Farmers Market	Café
1	0	Brockton / Parkdale Village / Exhibition Place	Café	Nightclub	Coffee Shop	Breakfast Spot	Gym	Bakery	Performing Arts Venue	Pet Store	Climbing Gym	Restaurant
2	0	Business reply mail Processing Centre	Yoga Studio	Smoke Shop	Garden Center	Garden	Fast Food Restaurant	Farmers Market	Light Rail Station	Comic Shop	Park	Recording Studio
3	0	CN Tower / King and Spadina / Railway Lands / ...	Airport Service	Airport	Harbor / Marina	Bar	Coffee Shop	Sculpture Garden	Boat or Ferry	Boutique	Rental Car Location	Airport Terminal
4	0	Central Bay Street	Coffee Shop	Italian Restaurant	Sandwich Place	Café	Spa	Burger Joint	Japanese Restaurant	Sushi Restaurant	Bubble Tea Shop	Salad Place

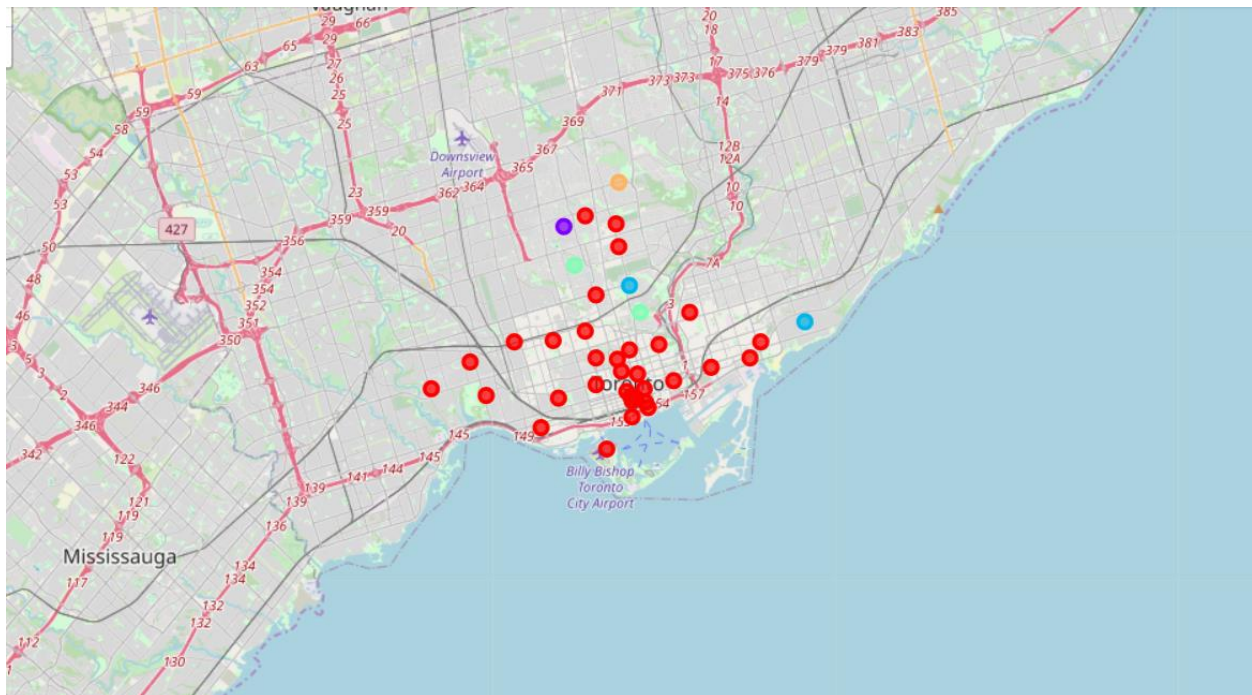
Methodology

Data scrapping from the Wikipedia page contains the update population statistics of the Toronto neighborhoods that was recorded. We need to find the population of each neighborhood as this impacts business to attract customers.

We would need to get coordinates and locations for each neighborhood that we can find. The Geocoder API has been used for this project and Foursquare.com to get the coordinates information.

Not only did we use the coordinates but we use machine learning concepts. Those concepts are K-Means to segment and cluster the neighborhoods to find the similarities. That way we can find the regions and areas that we would pick out for businesses.

Result



From the map above that we created, we can see the top 5 clusters of similar neighborhoods. They were grouped together of the nearby venues. This is really important because we can determine which cluster can offer business a large expansion. From the tables we see that coffee shops are most common while at the beaches.

Conclusion

As a result, we can see that through the first cluster, that East Toronto is the best bet to put a business there. The most common venues at some neighborhoods has mostly been coffee shop, airport service and few shops as well.