Downtown Toronto Analysis – Alvaro Marrugo

IMB Data Science Professional Certificate Specialization

1. Introduction

As a customer retention analytic I aim to continuously better understand y company's customers, what is it that drives them? Why do they prefer our service over our competitors? What is it they find better in our competitors over us? So on and so forth.

These questions unfortunately are not so easily answered as there are too many variables, demographically speaking, to take into account such as age, gender, household income, household composition, education, etc.

Specifically for this report I will try to better understand Toronto's Downtown's neighborhoods (the reason I choose Downtown Toronto will be explained later on this report) demographically in order to translate information into actionable insights which will help me create a way better and robust model for my company.

For learning purposes I will not be putting different variables in the development of this model. I will be using mainly Foursquare data to categorize neighborhoods based on Foursquare users' preferences.

Note that this report is mainly informative and picturing the data itself, the code will be shared in another link to Github in order to better understand this project.

2. Data where you describe the data that will be used to solve the problem and the source of the data

The data being used for this project is coming mainly from 2 sources, one being Wikipedia to create a data frame of neighborhoods, boroughs and postal codes of the city of Toronto and nearby settlements such as Etobicoke , North York and Mississauga.

The second and most important source is Foursquare, as specified in the development of the capstone project.

3. Methodology

The development of the project is essentially divided in three different steps:

1. Gathering and cleansing of the data

2. Processing of the data

3. Presentation of the data (essentially this report and a PowerPoint presentation which will be linked for grading purposes.)

As previously stated, the information is coming mainly from foursquare data merged with a data frame created by pulling information from a Wikipedia page.

All the different coding will be shown on a different pdf file on my GitHub repertory in order to make a shorter report.

The very first step I took for the development of this project was creating a dataframe in order to identify Neighborhoods and classify them in Boroughs based on their postal codes, once this dataframe was created I merged a it with its respective latitude and longitude.

	PostalCode	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

Example of the dataframe used as identifiers of neighborhoods in order to be merged with foursquare data.

After having gathered the information I wanted to focus my research on an specific Borough instead of doing a whole city analysis.

For this, I simply grouped Toronto's borough's importance based on the amount of neighborhoods in them.

11 boroughs were found with 4 of them having a larger number of neighborhoods.

Borough	PostalCode	Neighborhood	Latitude	Longitude
Central Toronto	9	9	9	9
Downtown Toronto	18	18	18	18
East Toronto	5	5	5	5
East York	5	5	5	5
Etobicoke	12	12	12	12
Mississauga	1	1	1	1
North York	24	24	24	24
Queen's Park	1	1	1	1
Scarborough	17	17	17	17
West Toronto	6	6	6	6
York	5	5	5	5

The outstanding boroughs are North York with 24 neighborhoods, Downtown Toronto with 18, Scarborough with 17 and Etobicoke with 12.

For this specific project I decided to just analyze Downtown Toronto because of its size, economical importance for the city itself, the daily flow of people and finally a personal love to the area as I used to live there.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
1	M4X	Downtown Toronto	Cabbagetown,St. James Town	43.667967	-79.367675
2	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
3	M5A	Downtown Toronto	Harbourfront,Regent Park	43.654260	-79.360636
4	M5B	Downtown Toronto	Ryerson,Garden District	43.657162	-79.378937
5	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
6	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
7	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
8	M5H	Downtown Toronto	Adelaide,King,Richmond	43.650571	-79.384568
9	M5J	Downtown Toronto	Harbourfront East,Toronto Islands,Union Station	43.640816	-79.381752
10	M5K	Downtown Toronto	Design Exchange,Toronto Dominion Centre	43.647177	-79.381576
11	M5L	Downtown Toronto	Commerce Court,Victoria Hotel	43.648198	-79.379817
12	M5S	Downtown Toronto	Harbord,University of Toronto	43.662696	-79.400049
13	M5T	Downtown Toronto	Chinatown,Grange Park,Kensington Market	43.653206	-79.400049
14	M5V	Downtown Toronto	CN Tower,Bathurst Quay,Island airport,Harbourf...	43.628947	-79.394420
15	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846
16	M5X	Downtown Toronto	First Canadian Place,Underground city	43.648429	-79.382280
17	M6G	Downtown Toronto	Christie	43.669542	-79.422564

Note: keep in mind that there are actually more than 18 neighborhoods in downtown Toronto, as they are grouped based on their postal codes some of them share the same and are therefore grouped under the same group (I.E Church & Wellesley).

Once connected to the Foursquare API I call all the most visited venues in Downtown Toronto and apply one-hot encoding to be able to then process the information as numeric values for k-means clustering.

	Neighborhood	Yoga Studio	Adult Boutique	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop
0	Adelaide,King,Richmond	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.040000	0.000000
1	Berczy Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	CN Tower,Bathurst Quay,Island airport,Harbourf…	0.000000	0.000000	0.000000	0.071429	0.071429	0.071429	0.142857	0.142857	0.142857	0.000000	0.000000
3	Cabbagetown,St. James Town	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	Central Bay Street	0.012195	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.012195	0.000000
5	Chinatown,Grange Park,Kensington Market	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	Christie	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7	Church and Wellesley	0.011494	0.011494	0.011494	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.011494	0.000000
8	Commerce Court,Victoria Hotel	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.040000	0.000000

	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	Adelaide,King,Richmond	Coffee Shop	Café	Steakhouse	American Restaurant	Thai Restaurant	Hotel	Asian Restaurant	Clothing Store	Bakery	Bar
1	Berczy Park	Coffee Shop	Restaurant	Cocktail Bar	Beer Bar	Cheese Shop	Seafood Restaurant	Steakhouse	Italian Restaurant	Farmers Market	Café
2	CN Tower,Bathurst Quay,Island airport,Harbourf…	Airport Lounge	Airport Service	Airport Terminal	Boat or Ferry	Plane	Sculpture Garden	Boutique	Airport	Airport Food Court	Airport Gate
3	Cabbagetown,St. James Town	Restaurant	Coffee Shop	Pub	Café	Bakery	Pizza Place	Market	Italian Restaurant	Playground	Bank
4	Central Bay Street	Coffee Shop	Café	Italian Restaurant	Burger Joint	Bar	Indian Restaurant	Chinese Restaurant	Spa	Bubble Tea Shop	Salad Place
5	Chinatown,Grange Park,Kensington Market	Bar	Café	Vegetarian / Vegan Restaurant	Dumpling Restaurant	Vietnamese Restaurant	Coffee Shop	Bakery	Mexican Restaurant	Chinese Restaurant	Dessert Shop
6	Christie	Grocery Store	Café	Park	Italian Restaurant	Baby Store	Diner	Nightclub	Restaurant	Convenience Store	Coffee Shop
7	Church and Wellesley	Japanese Restaurant	Sushi Restaurant	Coffee Shop	Gay Bar	Burger Joint	Restaurant	Gastropub	Café	Mediterranean Restaurant	Fast Food Restaurant

Snippet of the categorization of Neighborhoods.

The final part of the categorization is creating 5 clusters and classifying the neighborhoods in them as follows:

	PostalCode	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
0	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529	4	Park	Playground	Trail	Women's Store	Dessert Shop	Ethiopian Restaurant
1	M4X	Downtown Toronto	Cabbagetown,St. James Town	43.667967	-79.367675	4	Restaurant	Coffee Shop	Pub	Café	Bakery	Pizza Place
2	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	3	Japanese Restaurant	Sushi Restaurant	Coffee Shop	Gay Bar	Burger Joint	Restaurant
3	M5A	Downtown Toronto	Harbourfront,Regent Park	43.654260	-79.360636	4	Coffee Shop	Bakery	Café	Park	Pub	Breakfast Spot
4	M5B	Downtown Toronto	Ryerson,Garden District	43.657162	-79.378937	0	Coffee Shop	Clothing Store	Café	Middle Eastern Restaurant	Cosmetics Shop	Japanese Restaurant
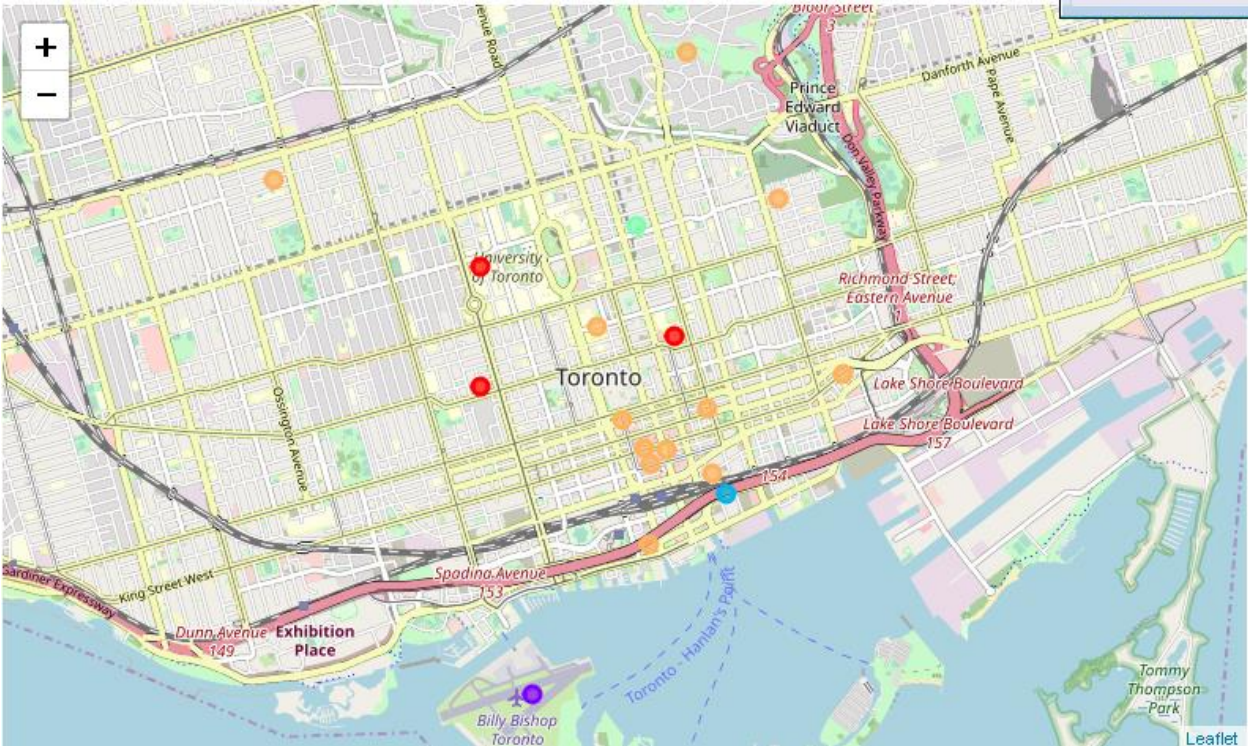
4. Results

As seen on the previous section of the report, I divided downtown Toronto into 5 different clusters in order to properly classify demographically.

Cluster 0 (red dots)

	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
4	Downtown Toronto	0	Coffee Shop	Clothing Store	Café	Middle Eastern Restaurant	Cosmetics Shop	Japanese Restaurant	Pizza Place	Bar	Ramen Restaurant	Restaurant
12	Downtown Toronto	0	Café	Japanese Restaurant	Bar	Bakery	Restaurant	Coffee Shop	Bookstore	Jazz Club	Italian Restaurant	Beer Bar
13	Downtown Toronto	0	Bar	Café	Vegetarian / Vegan Restaurant	Dumpling Restaurant	Vietnamese Restaurant	Coffee Shop	Bakery	Mexican Restaurant	Chinese Restaurant	Dessert Shop

Second largest cluster mostly restaurants. People who constantly use foursquare in the area are usually doing it for eating purposes with different kinds of restaurants (ethnically rich).

Cluster 1 (purple dot)

	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
14	Downtown Toronto	1	Airport Lounge	Airport Service	Airport Terminal	Boat or Ferry	Plane	Sculpture Garden	Boutique	Airport	Airport Food Court	Airport Gate

Nothing major to note besides the fact that, by the nature of the place, people who use foursquare here are tagging themselves on airport services or boat/ferry to go into the island.

Demographically, there is not much to say besides that people who are here simply have the need to travel a lot or like to travel.

Cluster 2 (blue dot)

	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
6	Downtown Toronto	2	Coffee Shop	Restaurant	Cocktail Bar	Beer Bar	Cheese Shop	Seafood Restaurant	Steakhouse	Italian Restaurant	Farmers Market	Café

Another neighborhood/borough mostly for eating, just as cluster 1, it would safe to say that nightlife in the area is quite active.

Cluster 3 (light green)

	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	Downtown Toronto	3	Japanese Restaurant	Sushi Restaurant	Coffee Shop	Gay Bar	Burger Joint	Restaurant	Gastropub	Café	Mediterranean Restaurant	Fast Food Restaurant

There is an interesting demographic aspect that can be noted from this cluster. It is the only one bringing gay bars in their top 10 most visited venues; it is safe to say that this is the "gay village" in downtown Toronto with also a lot of restaurant activity.

Cluster 4 (yellow dots)

	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	Downtown Toronto	4	Park	Playground	Trail	Women's Store	Dessert Shop	Ethiopian Restaurant	Electronics Store	Dumpling Restaurant	Donut Shop	Doner Restaurant
1	Downtown Toronto	4	Restaurant	Coffee Shop	Pub	Café	Bakery	Pizza Place	Market	Italian Restaurant	Playground	Bank
3	Downtown Toronto	4	Coffee Shop	Bakery	Café	Park	Pub	Breakfast Spot	Restaurant	Theater	Mexican Restaurant	Electronics Store
5	Downtown Toronto	4	Coffee Shop	Restaurant	Café	Hotel	Clothing Store	Italian Restaurant	Bakery	Breakfast Spot	Park	Gastropub
7	Downtown Toronto	4	Coffee Shop	Café	Italian Restaurant	Burger Joint	Bar	Indian Restaurant	Chinese Restaurant	Spa	Bubble Tea Shop	Salad Place
8	Downtown Toronto	4	Coffee Shop	Café	Steakhouse	American Restaurant	Thai Restaurant	Hotel	Asian Restaurant	Clothing Store	Bakery	Bar
9	Downtown Toronto	4	Coffee Shop	Hotel	Aquarium	Pizza Place	Café	Bakery	Scenic Lookout	Restaurant	Italian Restaurant	Brewery
10	Downtown Toronto	4	Coffee Shop	Café	Hotel	Restaurant	American Restaurant	Gastropub	Deli / Bodega	Gym	Italian Restaurant	Burger Joint
11	Downtown Toronto	4	Coffee Shop	Café	Restaurant	Hotel	American Restaurant	Gastropub	Deli / Bodega	Seafood Restaurant	Bakery	Steakhouse
15	Downtown Toronto	4	Coffee Shop	Restaurant	Café	Hotel	Cocktail Bar	Pub	Seafood Restaurant	Beer Bar	Italian Restaurant	Art Gallery
16	Downtown Toronto	4	Coffee Shop	Café	Hotel	Restaurant	American Restaurant	Steakhouse	Asian Restaurant	Deli / Bodega	Bakery	Seafood Restaurant

The final and largest cluster of downtown Toronto and possibly the one we can make the most of in terms of demographic data.

    a. In most neighborhoods the most visited venue is a Coffee shop followed by Cafes
    b. Most hotels in downtown Toronto seem to be in this area
    c. All activity is mostly related to food and drinks (Coffee and restaurants), hotels and spirits drinking.

5. **Discussion and observations**

Personally I believe that the information provided was in somewhat a little vague in order to do a proper demographic analysis. The very big observation that we I can tell that is 100% sure is that people in downtown Toronto like coffee, a lot, but then this does not necessarily mean that the whole analysis was bad, it was what the information that came out from the analysis.

6. **Conclusion.**

Major part of the conclusion is already written in the discussion/observation part of this report. In order to have been able to get better and more actionable insights in terms of demographics more information is needed to run this report. As many courses before this very final capstone taught us, one of the most important steps when analyzing something is firstly finding the right source a data analysts.