# Battle of the Neighborhoods

Coursera:

Applied Data Science Capstone Project

IBM

Dimitrios Tselikas
Junior Data Scientist

# CASE AND BACKGROUND

With around half of its population born outside the country, Toronto is often referred to as ‘the most multicultural city in the world.’ Located in Canada, this city boasts 200 ethnic groups with over 140 languages spoken

At this study we will examine which is the ideal Neighborhood in the Toronto Downtown for a family to move in. We will not consider price and cost of living in the criterias, the analysis will be based mainly on the venues around each neighborhood that can provide a comfortable and safe for the family, style of living.

Downtown Toronto is the main central business district of Toronto, Ontario, Canada. Located entirely within the district of Old Toronto, it is approximately 17 square kilometers in area, bounded by Bloor Street to the northeast and Dupont Street to the northwest, Lake Ontario to the south, the Don Valley to the east, and Bathurst Street to the west. It is also the location of the City of Toronto government and the Government of Ontario. The area is made up of Canada's largest concentration of skyscrapers and businesses that form Toronto's skyline. Downtown Toronto has the third most skyscrapers in North America exceeding 200 metres (656 ft) in height, behind New York City and Chicago

# DATA DESCRIPTION

We will use the services of Foursquare API to explore the data of Toronto (as we did in the Module of the 3rd week), in terms of their neighborhoods. The data include the information about the places around each neighborhood like restaurants, hotels, coffee shops, parks, theaters, art galleries, museums and many more. We selected one Borough from Toronto to analyze its neighborhoods, the seleceted Borough was Downtown Toronto. We will use machine learning technique, “Clustering” to segment the neighborhoods with similar objects on the basis of each neighborhood data. These objects will be given priority on activity in their respective neighborhoods. This will help to locate which area/neighborhood can provide us a 'peaceful' way of living.

As we did in Week 3's assesment, we have extracted table of Toronto’s Borough from Wikipedia page. Then we arrange the data according to our requirements. In the arrangement phase, which applied multiple steps including but not limited to, eliminating “Not assigned” values, combine neighborhoods which have same geographical coordinates at each borough and sorted against the concerned borough. For data verification and further exploration, we use Foursquare API to get the coordinates of Downtown Toronto and explore its neighborhoods. The neighborhoods are further characterized as venues and venue categories.

# Data acquisition

The data we used are the following:

A)https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, in order to obtain the data that is in the table of postal codes and to transform the data into a pandas dataframe like the one shown below

B) A link to a csv file that has the geographical coordinates of each postal code: http://cocl.us/Geospatial_data.

- Extracted the PostalCodes and Toronto Neighborhoods from the page using Pythons library BeatifulSoup.

Created the map of Toronto using Folium.

| | Postcode | Borough | Neighbourhood |
|---|---|---|---|
| 0 | M3A | North York | Parkwoods |
| 1 | M4A | North York | Victoria Village |
| 2 | M5A | Downtown Toronto | Harbourfront, Regent Park |
| 3 | M6A | North York | Lawrence Heights, Lawrence Manor |
| 4 | M9A | Etobicoke | Islington Avenue |

# Connected to Foursquare and specified the analysis to Toronto Downtown Borough.

An example of the venues in Toronto Downtown is shown below.

```
toronto_sites.shape
toronto_sites.head()
```

| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---|---|---|---|---|---|---|
| 0 | Harbourfront, Regent Park | 43.65426 | -79.360636 | Roselle Desserts | 43.653447 | -79.362017 | Bakery |
| 1 | Harbourfront, Regent Park | 43.65426 | -79.360636 | Tandem Coffee | 43.653559 | -79.361809 | Coffee Shop |
| 2 | Harbourfront, Regent Park | 43.65426 | -79.360636 | Toronto Cooper Koo Family Cherry St YMCA Centre | 43.653191 | -79.357947 | Gym / Fitness Center |
| 3 | Harbourfront, Regent Park | 43.65426 | -79.360636 | Body Blitz Spa East | 43.654735 | -79.359874 | Spa |
| 4 | Harbourfront, Regent Park | 43.65426 | -79.360636 | Morning Glory Cafe | 43.653947 | -79.361149 | Breakfast Spot |

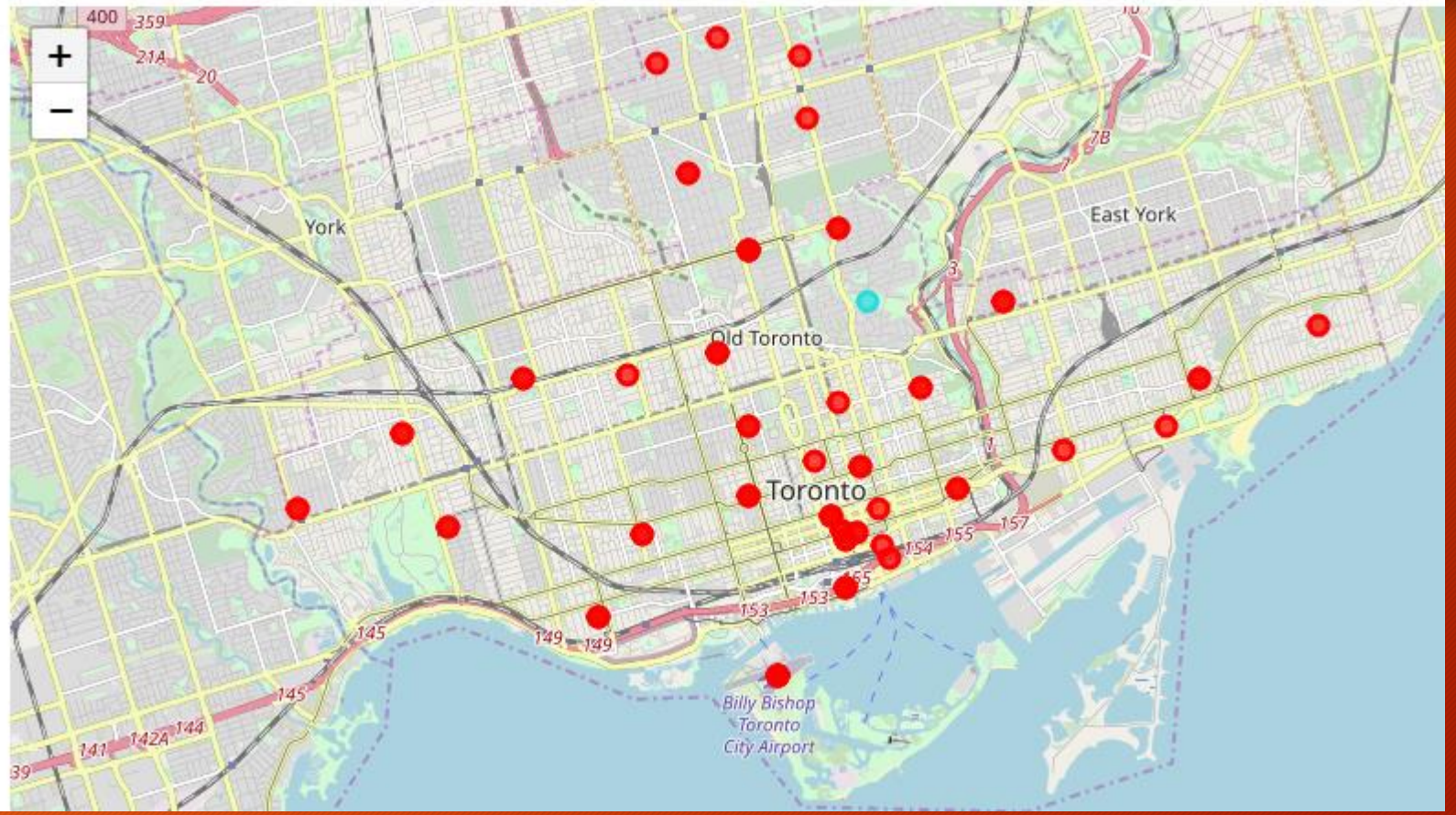
We have groupped the table by neighborhood and by taking the mean of the frequency of occurrence of each category, we can display the top 10 venues for each neighborhood.

Here is an example.

| | 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 | Bar | Thai Restaurant | Hotel | American Restaurant | Asian Restaurant | Bakery | Steakhouse | Restaurant |
| 1 | Berczy Park | Coffee Shop | Cocktail Bar | Cheese Shop | Seafood Restaurant | Steakhouse | Farmers Market | Beer Bar | Bakery | Café | Italian Restaurant |
| 2 | CN Tower, Bathurst Quay, Island airport, Harbo... | Airport Service | Airport Lounge | Airport Terminal | Boat or Ferry | Sculpture Garden | Harbor / Marina | Coffee Shop | Airport Gate | Airport Food Court | Airport |
| 3 | Cabbagetown, St. James Town | Coffee Shop | Restaurant | Italian Restaurant | Café | Pizza Place | Pub | Bakery | Japanese Restaurant | Sandwich Place | Diner |
| 4 | Central Bay Street | Coffee Shop | Italian Restaurant | Café | Ice Cream Shop | Sandwich Place | Burger Joint | Bar | Salad Place | Spa | Bubble Tea Shop |

Run *k*-means to cluster the neighborhood into 5 clusters.
Create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood and visualized the resulting clusters.

[40]:



# Through the analysis we got the 5 following clusters.

## The first cluster

| | Borough | Neighbourhood | 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 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Downtown Toronto | Harbourfront, Regent Park | 43.654260 | -79.360636 | 0 | Coffee Shop | Park | Pub | Bakery | Café | Mexican Restaurant | Theater | Restaurant | |
| 1 | Downtown Toronto | Ryerson, Garden District | 43.657162 | -79.378937 | 0 | Clothing Store | Coffee Shop | Cosmetics Shop | Bakery | Café | Middle Eastern Restaurant | Bookstore | Fast Food Restaurant | G |
| 2 | Downtown Toronto | St. James Town | 43.651494 | -79.375418 | 0 | Coffee Shop | Café | Restaurant | Hotel | Clothing Store | Bakery | Italian Restaurant | Cosmetics Shop | |
| 3 | Downtown Toronto | Berczy Park | 43.644771 | -79.373306 | 0 | Coffee Shop | Cocktail Bar | Cheese Shop | Seafood Restaurant | Steakhouse | Farmers Market | Beer Bar | Bakery | |
| 4 | Downtown Toronto | Central Bay Street | 43.657952 | -79.387383 | 0 | Coffee Shop | Italian Restaurant | Café | Ice Cream Shop | Sandwich Place | Burger Joint | Bar | Salad Place | |
| 6 | Downtown Toronto | Adelaide, King, Richmond | 43.650571 | -79.384568 | 0 | Coffee Shop | Café | Bar | Thai Restaurant | Hotel | American Restaurant | Asian Restaurant | Bakery | S |
| 7 | Downtown Toronto | Harbourfront East, Toronto Islands, Union Station | 43.640816 | -79.381752 | 0 | Coffee Shop | Hotel | Aquarium | Café | Brewery | Fried Chicken Joint | Scenic Lookout | Restaurant | |
| 8 | Downtown Toronto | Design Exchange, Toronto Dominion Centre | 43.647177 | -79.381576 | 0 | Coffee Shop | Café | Hotel | Restaurant | American Restaurant | Bar | Gastropub | Seafood Restaurant | De |
| 9 | Downtown Toronto | Commerce Court, Victoria Hotel | 43.648198 | -79.379817 | 0 | Coffee Shop | Café | Hotel | Restaurant | American Restaurant | Gym | Seafood Restaurant | Steakhouse | De |

# Rest of the Clusters

Parks and Playgrounds

```
In [44]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 1, toronto_merged.columns[[1] + list(range(2, toronto_merged.shape[1]))]]
```

Out[44]:

| | Borough | Neighbourhood | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 13 | Downtown Toronto | Rosedale | 43.679563 | -79.377529 | 1 | Park | Playground | Trail | Cosmetics Shop | Doner Restaurant | Dog Run | Discount Store | Diner | Dim Sum Restaurant |

Ethnic shops and restaurants

```
In [45]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 2, toronto_merged.columns[[1] + list(range(2, toronto_merged.shape[1]))]]
```

Out[45]:

| | Borough | Neighbourhood | 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 M Comr Ve |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | Downtown Toronto | Harbord, University of Toronto | 43.662696 | -79.400049 | 2 | Café | Restaurant | Bakery | Bar | Bookstore | Japanese Restaurant | Italian Restaurant | Dessert Shop | |
| 11 | Downtown Toronto | Chinatown, Grange Park, Kensington Market | 43.653206 | -79.400049 | 2 | Vietnamese Restaurant | Bar | Café | Vegetarian / Vegan Restaurant | Chinese Restaurant | Mexican Restaurant | Coffee Shop | Bakery | Dump Restau |

Airport and transfer services

```
In [46]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 3, toronto_merged.columns[[1] + list(range(2, toronto_merged.shape[1]))]]
```

Out[46]:

| | Borough | Neighbourhood | 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 | Con V |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | Downtown Toronto | CN Tower, Bathurst Quay, Island airport, | 43.628947 | -79.39442 | 3 | Airport Service | Airport Lounge | Airport Terminal | Boat or Ferry | Sculpture Garden | Harbor / Marina | Coffee Shop | Airport Gate | Airport Food Court | A |

Convenience stores

```
In [47]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 4, toronto_merged.columns[[1] + list(range(2, toronto_merged.shape[1]))]]
```

Out[47]:

| | Borough | Neighbourhood | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | Downtown Toronto | Christie | 43.669542 | -79.422564 | 4 | Grocery Store | Café | Park | Coffee Shop | Candy Store | Restaurant | Athletics & Sports | Italian Restaurant | Convenience Store |

# RESULTS AND ANALYSIS

By examining the results from our clusters we reach out the following conclusions.

The five clusters are separated based on the type of the venues among each neighborhood. So we have neighborhoods that are ideal for a coffee break and dinning (included in the first cluster), neighborhoods with parks playgrounds and local stores (cluster number two), neighborhoods with Ethnic shops and restaurants (third cluster), neighborhoods next to airport and transfer services (cluster number four) and finally neighborhoods with Convenience stores (last cluster).

# Rosedale, Toronto Downtown

Considering our initial goal, we are looking for neighborhoods in the Downtown of Toronto that are ideal for a family to live in, we can say that a neighborhood included in the second cluster with Parks playgrounds etc can be the perfect place. So our analysis suggests that Rosedale is the place we are looking for. By having a look at the wikipedia page for Rosedale, Toronto we can clarify that this area is a very good choice indeed(https://en.wikipedia.org/wiki/Rosedale,_Toronto).

"Rosedale is a neighbourhood in Toronto, Ontario, Canada,it is located north of Downtown Toronto and is one of its oldest suburbs. It is also one of the wealthiest and most highly priced neighbourhoods in Canada.[2] Rosedale has been ranked the best neighbourhood in Toronto to live in by Toronto Life".