# Capstone Project: Coffee Shop in Toronto

# Introduction: Business Problem

## 1. Project Summary:

In this project I will try to examine three different neighborhoods in the city of Toronto, in order to support prospect business owner's plan to launch new Coffee shop based on the most common venue types for each neighborhood. Furthermore, after delivering the results of this project it could be used to evaluate other possible type of venues that could potentially be launched.

## 1.1 Problem Background:

Toronto is a big and beautiful city full of different sightseeing spots and various restaurants and small coffee shops. This project aims to explore different boroughs on the matter of frequency of venues per cluster, thus ensuring that opportunity to launch prospect business is existing.

## 1.2 Target audience:

This project aims to navigate entrepreneurs and help them finding the most frequent venue types so that they would be able to make appropriate decision on launching coffee shop.

# Data

## 2. Type of data

Based on the definition of the target deliverable, data that will be used:

- List of current postal codes with regard to neighborhoods in Toronto;
- Data on latitude and longitude of venues and neighborhoods;
- Foursqare API, to get actual list of venues nearby target neighborhood, that eventually will be used to create a map;

As specified earlier, three different neighborhoods will be evaluated namely: Downtown Toronto, North York and Etobicoke.

# Methodology

**3.** First off, to begin analysis on venue types we must get borough name along with postal codes and coordinates.

## 3.1 Data preparation

```
In [2]:  url = 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
         geo_data = pd.read_csv('https://cocl.us/Geospatial_data')
         dfp = pd.read_html(url)

In [4]:  df_edited = dfp[0]
         df = df_edited[['Postal code','Borough', 'Neighborhood']]
         df1 = df[df.Borough != 'Not assigned']

In [5]:  geo_data = geo_data.rename(columns = {'Postal Code':'Postal code'})
         geo_data.head(2)
```

Out[5]:

| | Postal code | Latitude | Longitude |
|---|---|---|---|
| 0 | M1B | 43.806686 | -79.194353 |
| 1 | M1C | 43.784535 | -79.160497 |

```
In [6]:  df2 = df1.merge(geo_data, on = 'Postal code', how = 'left')
         df2
```

Picture-1. Loading first part of data.

As can be seen on Picture-1, after retrieving data from first source labeled “url” that contains postal codes of Toronto’s boroughs, coordinates dataframe was created so that eventually new dataframe would be made from merger of first and second dataframes.

```
In [15]: nyork_venues = getNearbyVenues(names=nyork_data['Neighborhood'],
                                        latitudes=nyork_data['Latitude'],
                                        longitudes=nyork_data['Longitude'],)


         downtown_venues = getNearbyVenues(names=downtown_data['Neighborhood'],
                                        latitudes=downtown_data['Latitude'],
                                        longitudes=downtown_data['Longitude'],)

         etob_venues = getNearbyVenues(names=etob_data['Neighborhood'],
                                        latitudes=etob_data['Latitude'],
                                        longitudes=etob_data['Longitude'],)
```

```
Parkwoods
Victoria Village
Lawrence Manor / Lawrence Heights
Don Mills
Glencairn
Don Mills
Hillcrest Village
Bathurst Manor / Wilson Heights / Downsview North
Fairview / Henry Farm / Oriole
Northwood Park / York University
Bayview Village
Downsview
York Mills / Silver Hills
Downsview
North Park / Maple Leaf Park / Upwood Park
Humber Summit
Willowdale / Newtonbrook
Downsview
Bedford Park / Lawrence Manor East
Humberlea / Emery
```

Picture-2. Retrieving venues for selected boroughs.

After defining necessary IDs for foursquare API, the function was defined, which would retrieve nearby venues for selected boroughs (all details can be found in the attached Notebook).

```
In [23]:
num_top_venues = 5

for hood in nyork_grouped['Neighborhood']:
    print("----"+hood+"----")
    temp = nyork_grouped[nyork_grouped['Neighborhood'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_top_venues))
    print('\n')
```

```
----Bathurst Manor / Wilson Heights / Downsview North----
           venue  freq
0    Coffee Shop  0.11
1           Bank  0.11
2  Shopping Mall  0.05
3    Supermarket  0.05
4  Deli / Bodega  0.05


----Bayview Village----
                 venue  freq
0   Chinese Restaurant  0.25
1                 Bank  0.25
2                 Café  0.25
3  Japanese Restaurant  0.25
4    Accessories Store  0.00


----Bedford Park / Lawrence Manor East----
                venue  freq
```

Picture-3. Top 5 venues for North York neighborhoods.

In the preliminary results, we could see that in 4 neighborhoods in North York, Coffee shop is in top 5 venues for this borough.

```
In [71]:
# Downtown Toronto
num_top_venues = 5

for hood in downtown_grouped['Neighborhood']:
    print("----"+hood+"----")
    temp = downtown_grouped[downtown_grouped['Neighborhood'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_top_venues))
    print('\n')
```

```
----Berczy Park----
               venue  freq
0        Coffee Shop  0.05
1       Cocktail Bar  0.05
2           Beer Bar  0.04
3  Italian Restaurant  0.04
4      Farmers Market  0.04


----CN Tower / King and Spadina / Railway Lands / Harbourfront West / Bathurst Quay / South Niagara / Island airport----
               venue  freq
0    Airport Service  0.17
1     Airport Lounge  0.11
2   Airport Terminal  0.11
3        Coffee Shop  0.06
4  Airport Food Court  0.06


----Central Bay Street----
               venue  freq
```

Picture-4. Top 5 venues for Downtown Toronto neighborhoods.

Here preliminary results get interesting, in Downtown Toronto there are 15 neighborhoods that have Coffee shops in top 5 venues, which indicates that there are more Coffee shops in this area compared to North York.

```
In [72]:
# Etobicoke
num_top_venues = 5

for hood in etob_grouped['Neighborhood']:
    print("----"+hood+"----")
    temp = etob_grouped[etob_grouped['Neighborhood'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_top_venues))
    print('\n')
```

```
----Alderwood / Long Branch----
         venue  freq
0  Pizza Place   0.2
1  Coffee Shop   0.1
2          Pub   0.1
3          Gym   0.1
4     Pharmacy   0.1


----Eringate / Bloordale Gardens / Old Burnhamthorpe / Markland Wood----
         venue  freq
0   Beer Store  0.12
1    Pet Store  0.12
2  Pizza Place  0.12
3         Café  0.12
4  Coffee Shop  0.12


----Kingsview Village / St. Phillips / Martin Grove Gardens / Richview Gardens----
                venue  freq
```

Picture-5. Top 5 venues for Etobicoke neighborhoods.

Conversely, in Etobicoke the number of neighborhoods with Coffee shops listed as top 5, is 2. Quite less than in North York and Downtown Toronto.

# Results

## 4. Most common venue types

## North York

Out[74]:

| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
|---|---|---|---|---|---|---|
| 0 | Bathurst Manor / Wilson Heights / Downsview North | Bank | Coffee Shop | Supermarket | Pizza Place | Bridal Shop |
| 1 | Bayview Village | Japanese Restaurant | Chinese Restaurant | Café | Bank | Electronics Store |
| 2 | Bedford Park / Lawrence Manor East | Restaurant | Sandwich Place | Coffee Shop | Italian Restaurant | Sushi Restaurant |
| 3 | Don Mills | Restaurant | Beer Store | Asian Restaurant | Coffee Shop | Japanese Restaurant |
| 4 | Downsview | Grocery Store | Park | Liquor Store | Gym / Fitness Center | Baseball Field |
| 5 | Fairview / Henry Farm / Oriole | Clothing Store | Coffee Shop | Fast Food Restaurant | Japanese Restaurant | Bakery |
| 6 | Glencairn | Japanese Restaurant | Playground | Bakery | Pub | Discount Store |
| 7 | Hillcrest Village | Dog Run | Golf Course | Pool | Mediterranean Restaurant | Discount Store |
| 8 | Humber Summit | Pizza Place | Empanada Restaurant | Women's Store | Discount Store | Concert Hall |
| 9 | Humberlea / Emery | Food Service | Baseball Field | Women's Store | Electronics Store | Construction & Landscaping |
| 10 | Lawrence Manor / Lawrence Heights | Clothing Store | Women's Store | Shoe Store | Boutique | Coffee Shop |
| 11 | North Park / Maple Leaf Park / Upwood Park | Construction & Landscaping | Bakery | Park | Women's Store | Dog Run |
| 12 | Northwood Park / York University | Coffee Shop | Metro Station | Caribbean Restaurant | Massage Studio | Bar |
| 13 | Parkwoods | Park | Food & Drink Shop | Fireworks Store | Discount Store | Comfort Food Restaurant |
| 14 | Victoria Village | Coffee Shop | Grocery Store | Hockey Arena | Portuguese Restaurant | Nail Salon |
| 15 | Willowdale | Coffee Shop | Ramen Restaurant | Pizza Place | Restaurant | Grocery Store |
| 16 | York Mills / Silver Hills | Park | Cafeteria | Women's Store | Concert Hall | Construction & Landscaping |
| 17 | York Mills West | Convenience Store | Bank | Bar | Park | Women's Store |

Picture-6. The most common venue types for North York.

Based on the results obtained for North York, it can be spotted that for 1st most common venue type, there are 2 neighborhoods with Coffee shops, while in the third and fourth columns for 3rd and 4th most common venues, there is 1 neighborhood per each column.

## Downtown Toronto

Out[76]:

| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
|---|---|---|---|---|---|---|
| 0 | Berczy Park | Coffee Shop | Cocktail Bar | Beer Bar | Farmers Market | Bakery |
| 1 | CN Tower / King and Spadina / Railway Lands / ... | Airport Service | Airport Lounge | Airport Terminal | Boat or Ferry | Harbor / Marina |
| 2 | Central Bay Street | Coffee Shop | Italian Restaurant | Café | Sandwich Place | Salad Place |
| 3 | Christie | Grocery Store | Café | Park | Athletics & Sports | Gas Station |
| 4 | Church and Wellesley | Coffee Shop | Sushi Restaurant | Japanese Restaurant | Restaurant | Hotel |
| 5 | Commerce Court / Victoria Hotel | Coffee Shop | Restaurant | Café | Hotel | Gym |
| 6 | First Canadian Place / Underground city | Coffee Shop | Café | Restaurant | Hotel | Gym |
| 7 | Garden District, Ryerson | Coffee Shop | Clothing Store | Café | Cosmetics Shop | Japanese Restaurant |
| 8 | Harbourfront East / Union Station / Toronto Is... | Coffee Shop | Aquarium | Italian Restaurant | Restaurant | Café |
| 9 | Kensington Market / Chinatown / Grange Park | Café | Coffee Shop | Vietnamese Restaurant | Dessert Shop | Mexican Restaurant |
| 10 | Queen's Park / Ontario Provincial Government | Coffee Shop | Diner | Sushi Restaurant | Gym | Mexican Restaurant |
| 11 | Regent Park / Harbourfront | Coffee Shop | Bakery | Park | Pub | Restaurant |
| 12 | Richmond / Adelaide / King | Coffee Shop | Café | Restaurant | Gym | Thai Restaurant |
| 13 | Rosedale | Park | Trail | Playground | Cupcake Shop | Donut Shop |
| 14 | St. James Town | Café | Coffee Shop | Italian Restaurant | Gastropub | Cocktail Bar |
| 15 | St. James Town / Cabbagetown | Coffee Shop | Italian Restaurant | Market | Restaurant | Café |
| 16 | Stn A PO Boxes | Coffee Shop | Italian Restaurant | Café | Seafood Restaurant | Hotel |
| 17 | Toronto Dominion Centre / Design Exchange | Coffee Shop | Hotel | Café | Restaurant | Japanese Restaurant |
| 18 | University of Toronto / Harbord | Café | Italian Restaurant | Restaurant | Bar | Japanese Restaurant |

Picture-7. The most common venue types for Downtown Toronto.

Observing results for Downtown Toronto, it is obvious that Coffee shops are common venue type for this area based on the number of Coffee shops in the first column, which clearly indicates that this area is saturated with amount of Coffee shops.

## Etobicoke

Out[78]:

| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
|---|---|---|---|---|---|---|
| 0 | Alderwood / Long Branch | Pizza Place | Pharmacy | Athletics & Sports | Skating Rink | Gym |
| 1 | Eringate / Bloordale Gardens / Old Burnhamthor... | Pizza Place | Pet Store | Beer Store | Liquor Store | Cosmetics Shop |
| 2 | Kingsview Village / St. Phillips / Martin Grov... | Sandwich Place | Mobile Phone Shop | Pizza Place | Park | Wings Joint |
| 3 | Mimico NW / The Queensway West / South of Bloo... | Wings Joint | Kids Store | Bakery | Burger Joint | Burrito Place |
| 4 | New Toronto / Mimico South / Humber Bay Shores | Café | American Restaurant | Pizza Place | Gym | Liquor Store |
| 5 | Northwest | Drugstore | Rental Car Location | Wings Joint | Chinese Restaurant | Fried Chicken Joint |
| 6 | Old Mill South / King's Mill Park / Sunnylea /... | Baseball Field | Breakfast Spot | Locksmith | Construction & Landscaping | Wings Joint |
| 7 | South Steeles / Silverstone / Humbergate / Jam... | Grocery Store | Pizza Place | Fast Food Restaurant | Discount Store | Japanese Restaurant |
| 8 | The Kingsway / Montgomery Road / Old Mill North | River | Pool | Park | Wings Joint | Chinese Restaurant |
| 9 | Westmount | Pizza Place | Chinese Restaurant | Intersection | Sandwich Place | Discount Store |

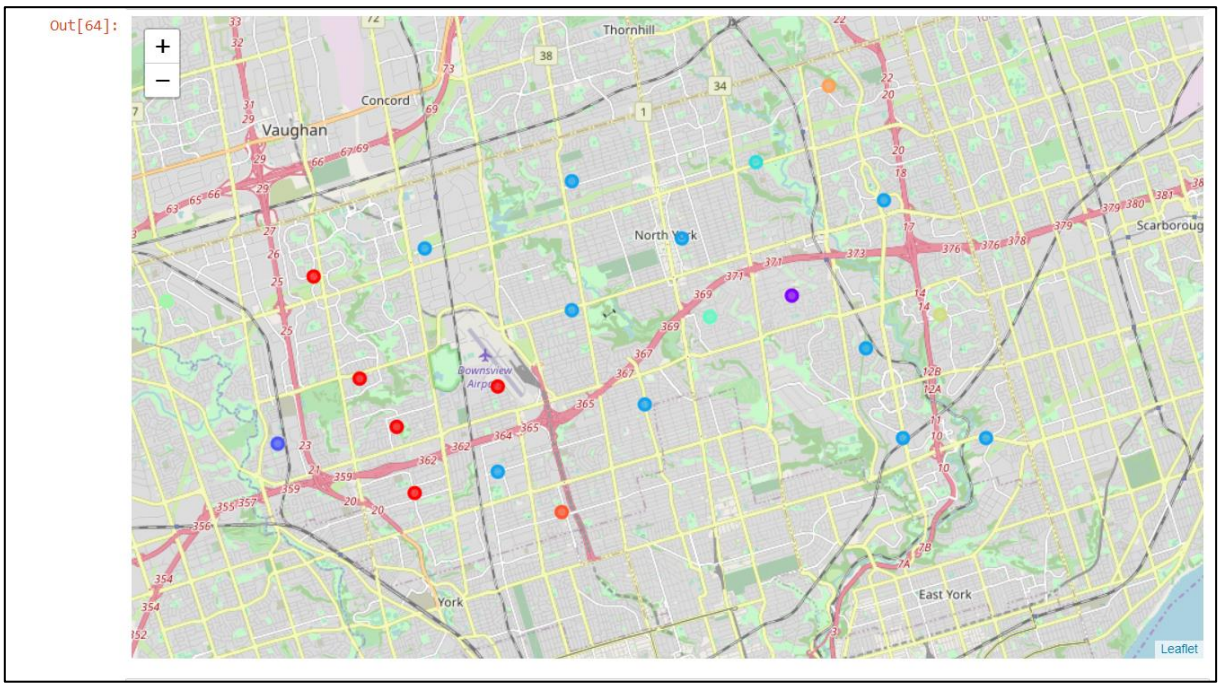
Picture-8. The most common venue types for Etobicoke.

The results are different for Etobicoke borough, where Coffee shops are not that common compared to Pizza places or restaurants for instance.

These factors, however, are not enough in my view to evaluate the possibility to open new Coffee shop. Size of each borough shall be taken in consideration as well. Hence, cluster density along with number of people that potentially live in those boroughs add weight in making final decision.

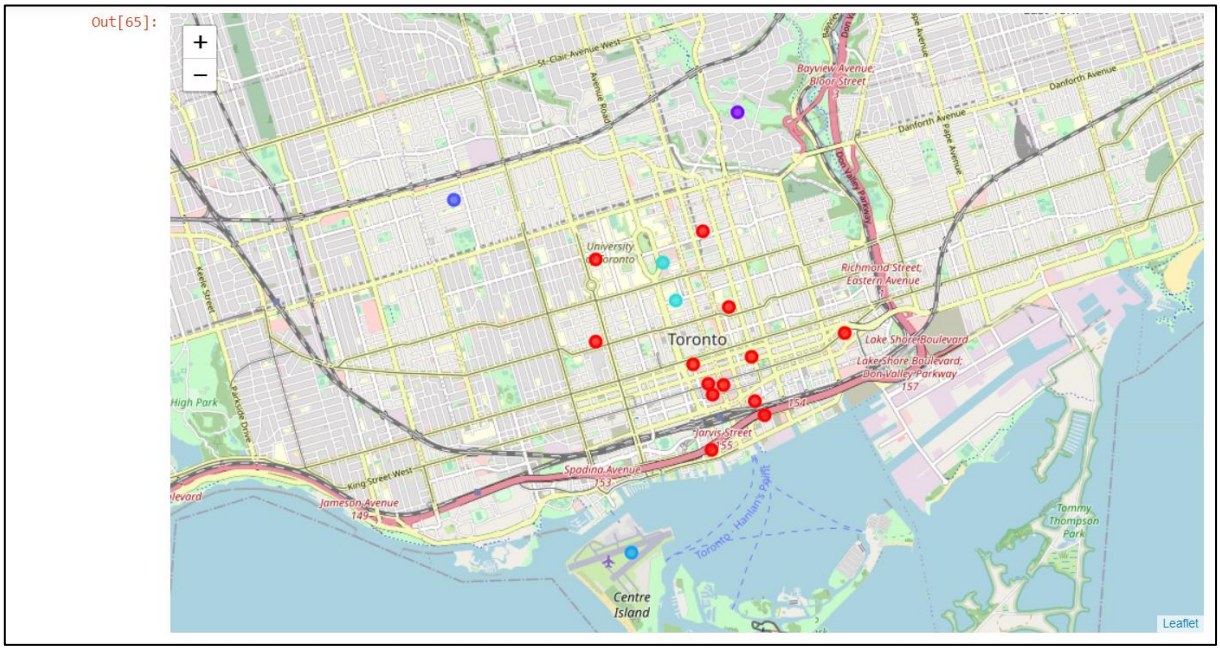
## 4.1 District Maps

## North York



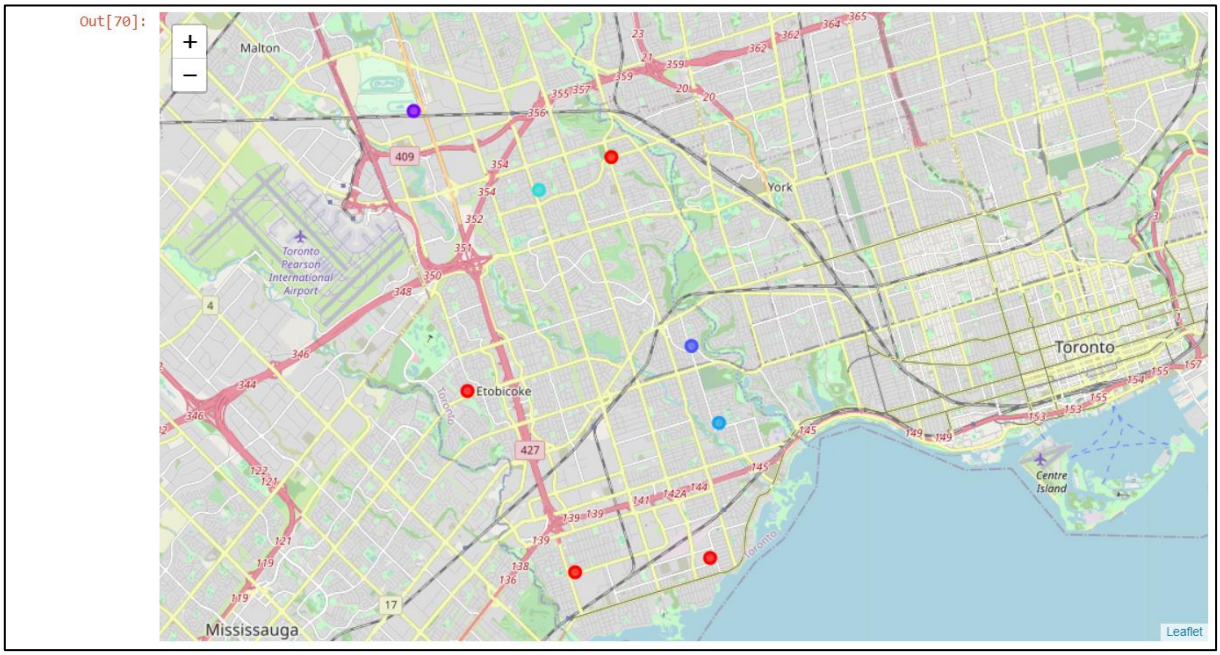

Picture-9. Map of clusters in North York.

## Downtown Toronto




Picture-10. Map of clusters in Downtown Toronto.

## Etobicoke




Picture-11. Map of clusters in Etobicoke.

# Discussion

As was mentioned earlier, factors such as frequency of most common venues would not be enough in evaluating the possibility to open new Coffee shop. Geographical size of each borough shall be taken in consideration as well. As it could be seen on maps of North York, Downtown Toronto and Etobicoke, cluster densities vary greatly. For Downtown Toronto, clusters seem closer to each other, in contrast to Etobicoke with distant clusters or North York, where clusters are at comparatively moderate distance from each other.

Observed distances of clusters in each borough show us how separated clusters are. In this case it is helpful to make assumption that in order to cover more neighborhoods clusters shall be in moderate distance vicinity from each other. Even though it could be tempting to conclude that Etobicoke has fewer Coffee shops, which is true, however cluster are located such that they are on the edge of area forming rectangle. North York on the other hand, has its clusters disseminated on moderate distance from each other, thus making it easier to cover more neighborhoods for prospect coffee shop, add the fact that there are not so many coffee shops in this area, unlike in Downtown Toronto. Those factors yield us North York as optimal place to open new coffee shop.

# Conclusion

In this study, I analyzed three different boroughs in Toronto, to find an optimal place for new coffee shop. Foursquare API was used to fetch venue types near sorted neighborhoods, thus making it easier to create new dataframes with most common venue types for selected neighborhoods. Folium library was used to generate maps, that covered clusters along with neighborhoods in those boroughs. For the scope of this project three boroughs were covered, while in other instances we are not limited to geographical area or number of boroughs for example. Eventually, all these tools helped to formulate assumption to choose from observed variables.