

# Capstone Project - The Battle of Neighborhoods (Week 2)

## Introduction & Problem Statement

The city of Toronto supports one of the large urban settings in the world. With the likes of Financial District - Toronto Stock Exchange (3<sup>rd</sup> largest in North America), Canadian Film & TV Industry, The Digital Corridor and more, the need for coffee is boundless around the city. We will look to explore which neighborhoods in Toronto are closest to coffee shops to residents. This can influence the decision to stay/rent/buy housing in those neighborhoods for coffee lovers.

1. Find neighborhoods cluster(s) with the most coffee shops near them
2. Identify a few desirable neighborhoods to stay that have access to most coffee shops
3. Find the neighborhoods cluster with the lowest coffee shops near them
4. This presents a business opportunity for entrepreneurs to invest in or open new coffee shops near those neighborhoods.

## Data Section

1. We will source Toronto neighborhood data from Wikipedia  
[https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
2. Get geolocation data from cocl.us (used in previous labs)  
[http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data)
3. Using FourSquare API, get the venues data for the neighborhood latitude/longitudes  
[https://api.foursquare.com/v2/venues/explore?&client\\_id={}&client\\_secret={}&v={}&ll={},{}&radius={}&limit={}'](https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}')  
Format -> CLIENT\_ID, CLIENT\_SECRET, VERSION, lat, lng, radius, LIMIT

## Methodology Section

1. Source Toronto neighborhood data from Wikipedia

|   | Postal Code | Borough          | Neighborhood              |
|---|-------------|------------------|---------------------------|
| 0 | M1A         | NaN              | NaN                       |
| 1 | M2A         | NaN              | NaN                       |
| 2 | M3A         | North York       | Parkwoods                 |
| 3 | M4A         | North York       | Victoria Village          |
| 4 | M5A         | Downtown Toronto | Regent Park, Harbourfront |

2. Cleanup empty and NaN cells

|   | Postal Code | 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                                       |
| 5 | M1J         | Scarborough | Scarborough Village                             |
| 6 | M1K         | Scarborough | Kennedy Park, Ionview, East Birchmount Park     |
| 7 | M1L         | Scarborough | Golden Mile, Clairlea, Oakridge                 |
| 8 | M1M         | Scarborough | Cliffside, Cliffcrest, Scarborough Village West |
| 9 | M1N         | Scarborough | Birch Cliff, Cliffside West                     |

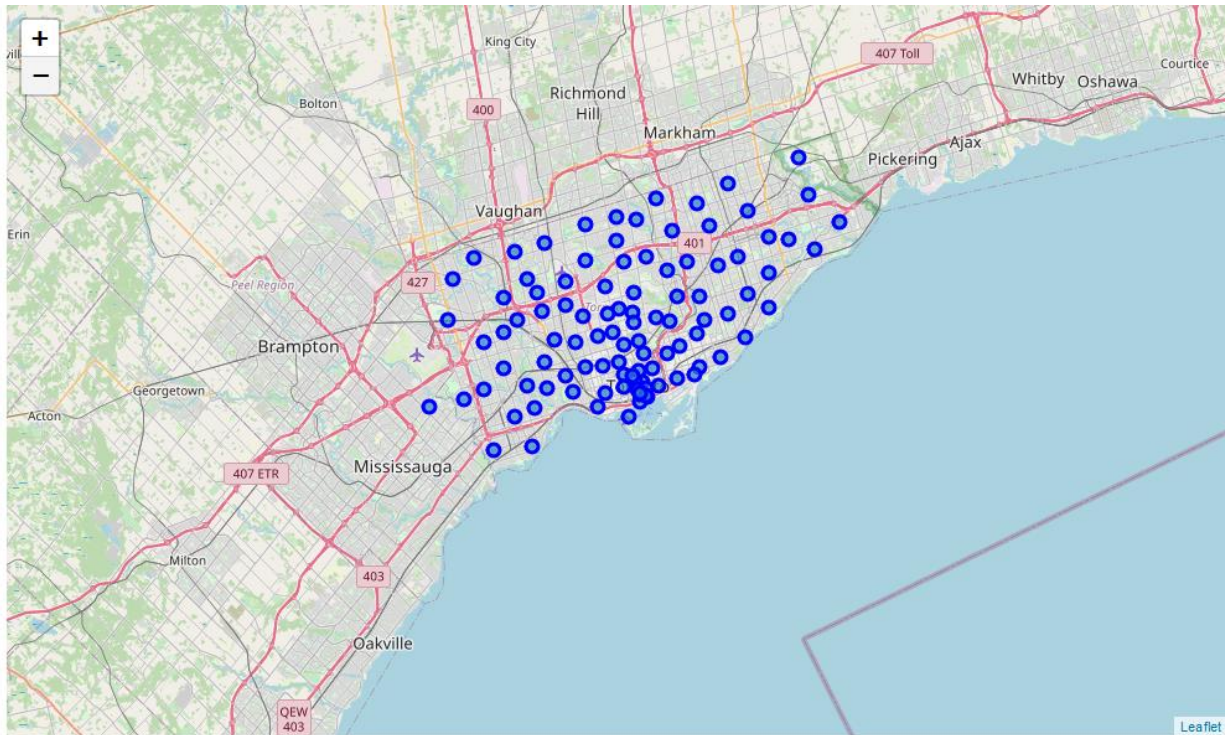
3. Get geolocation data from cocl.us (used in previous labs)

|   | Postal Code | Latitude  | Longitude  |
|---|-------------|-----------|------------|
| 0 | M1B         | 43.806686 | -79.194353 |
| 1 | M1C         | 43.784535 | -79.160497 |
| 2 | M1E         | 43.763573 | -79.188711 |
| 3 | M1G         | 43.770992 | -79.216917 |
| 4 | M1H         | 43.773136 | -79.239476 |

4. Merge the two data to obtain latitude and longitude for the neighborhoods

|   | Postal Code | 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 |

5. Plot the data on a map



6. Using FourSquare API, get the venues data for the neighborhood latitude/longitudes

|   | 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 | Malvern, Rouge                         | 43.806686             | -79.194353             | Interprovincial Group | 43.805630      | -79.200378      | Print Shop           |
| 2 | Rouge Hill, Port Union, Highland Creek | 43.784535             | -79.160497             | Royal Canadian Legion | 43.782533      | -79.163085      | Bar                  |
| 3 | Guildwood, Morningside, West Hill      | 43.763573             | -79.188711             | RBC Royal Bank        | 43.766790      | -79.191151      | Bank                 |
| 4 | Guildwood, Morningside, West Hill      | 43.763573             | -79.188711             | G & G Electronics     | 43.765309      | -79.191537      | Electronics Store    |

7. Filter venues to use data for Coffee Shops

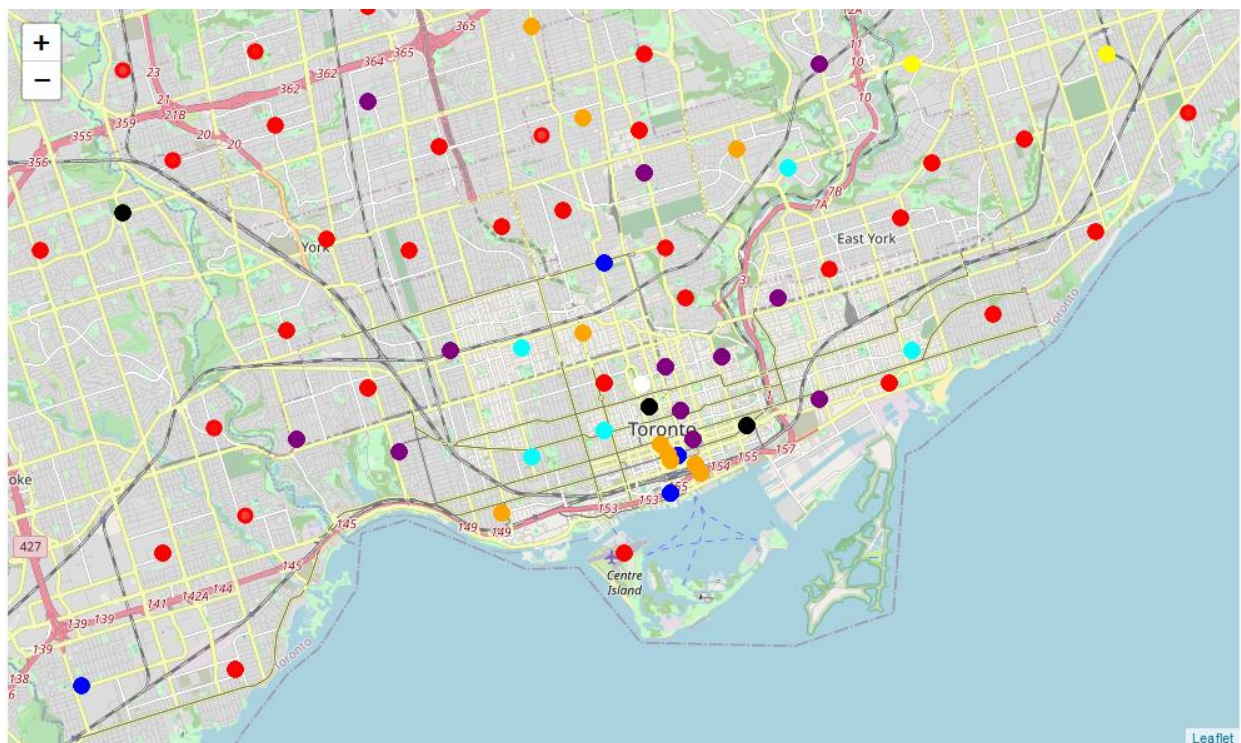
|   | Neighborhood                                    | Coffee Shop |
|---|---|-------------|
| 0 | Agincourt                                       | 0.000000    |
| 1 | Alderwood, Long Branch                          | 0.142857    |
| 2 | Bathurst Manor, Wilson Heights, Downsview North | 0.200000    |
| 3 | Bayview Village                                 | 0.000000    |
| 4 | Bedford Park, Lawrence Manor East               | 0.100000    |

8. Run KMeans clustering

|   | Neighborhood                                    | Coffee Shop | Cluster Labels |
|---|---|-------------|----------------|
| 0 | Agincourt                                       | 0.000000    | 1              |
| 1 | Alderwood, Long Branch                          | 0.142857    | 5              |
| 2 | Bathurst Manor, Wilson Heights, Downsview North | 0.200000    | 4              |
| 3 | Bayview Village                                 | 0.000000    | 1              |
| 4 | Bedford Park, Lawrence Manor East               | 0.100000    | 0              |

## Results Section

9. Plot the cluster data on a map



10. Analyze and deduce best clusters/neighborhoods for coffee lovers

```

Cluster 0 shape : (0, 9)
Cluster 1 shape : (29, 9)
Cluster 2 shape : (2, 9)
Cluster 3 shape : (3, 9)
Cluster 4 shape : (43, 9)
Cluster 5 shape : (22, 9)
Cluster 6 shape : (10, 9)
Cluster 7 shape : (57, 9)
Cluster 8 shape : (8, 9)

```

## Discussion Section

11. Analyze and deduce neighborhoods where there are low or now coffee shops

Lets view cluster 7 as it has the most coffee shops

Out[34]:

|    | Neighborhood                                    | Coffee Shop | Cluster Labels | Neighborhood Latitude | Neighborhood Longitude | Venue                             | Venue Latitude | Venue Longitude | Venue Category |
|----|---|-------------|----------------|-----------------------|------------------------|-----------------------------------|----------------|-----------------|----------------|
| 17 | Bathurst Manor, Wilson Heights, Downsview North | 0.100000    | 7              | 43.754328             | -79.442259             | Tim Hortons                       | 43.754767      | -79.443250      | Coffee Shop    |
| 19 | Bathurst Manor, Wilson Heights, Downsview North | 0.100000    | 7              | 43.754328             | -79.442259             | Starbucks                         | 43.755797      | -79.440471      | Coffee Shop    |
| 43 | Bedford Park, Lawrence Manor East               | 0.086957    | 7              | 43.733283             | -79.419750             | Tim Hortons                       | 43.735356      | -79.419605      | Coffee Shop    |
| 47 | Bedford Park, Lawrence Manor East               | 0.086957    | 7              | 43.733283             | -79.419750             | Starbucks                         | 43.732604      | -79.419136      | Coffee Shop    |
| 72 | Berczy Park                                     | 0.089286    | 7              | 43.644771             | -79.373306             | Starbucks                         | 43.644489      | -79.368639      | Coffee Shop    |
| 87 | Berczy Park                                     | 0.089286    | 7              | 43.644771             | -79.373306             | Everyday Gourmet (Teas & Coffees) | 43.648757      | -79.371645      | Coffee Shop    |
| 95 | Berczy Park                                     | 0.089286    | 7              | 43.644771             | -79.373306             | Moe's Moe's                       | 43.641640      | -79.377552      | Coffee         |

Lets view cluster 4 Neighborhoods as this has the next most coffee shops

|     | Neighborhood                 | Coffee Shop | Cluster Labels | Neighborhood Latitude | Neighborhood Longitude | Venue                | Venue Latitude | Venue Longitude | Venue Category |
|-----|------------------------------|-------------|----------------|-----------------------|------------------------|----------------------|----------------|-----------------|----------------|
| 298 | Church and Wellesley         | 0.066667    | 4              | 43.665860             | -79.383160             | Starbucks            | 43.664980      | -79.380510      | Coffee Shop    |
| 319 | Church and Wellesley         | 0.066667    | 4              | 43.665860             | -79.383160             | Piedmont Coffee Bar  | 43.668286      | -79.382520      | Coffee Shop    |
| 326 | Church and Wellesley         | 0.066667    | 4              | 43.665860             | -79.383160             | Rooster Coffee House | 43.669654      | -79.379871      | Coffee Shop    |
| 352 | Church and Wellesley         | 0.066667    | 4              | 43.665860             | -79.383160             | The Drink            | 43.664167      | -79.380149      | Coffee Shop    |
| 356 | Church and Wellesley         | 0.066667    | 4              | 43.665860             | -79.383160             | Croissant Tree       | 43.669575      | -79.382331      | Coffee Shop    |
| 482 | Davisville                   | 0.060606    | 4              | 43.704324             | -79.388790             | Starbucks            | 43.705923      | -79.389548      | Coffee Shop    |
| 496 | Davisville                   | 0.060606    | 4              | 43.704324             | -79.388790             | Second Cup           | 43.704344      | -79.388659      | Coffee Shop    |
| 534 | Don Mills                    | 0.076923    | 4              | 43.725900             | -79.340923             | Tim Hortons          | 43.722897      | -79.339117      | Coffee Shop    |
| 546 | Don Mills                    | 0.076923    | 4              | 43.725900             | -79.340923             | Delimark Cafe        | 43.727536      | -79.339547      | Coffee Shop    |
| 586 | Dufferin, Dovercourt Village | 0.062500    | 4              | 43.669005             | -79.442259             | Coffee Time          | 43.666837      | -79.446197      | Coffee Shop    |
| 610 | Fairview, Henry Farm, Oriole | 0.075758    | 4              | 43.778517             | -79.346556             | Aroma Espresso Bar   | 43.777700      | -79.344652      | Coffee Shop    |
| 614 | Fairview, Henry Farm, Oriole | 0.075758    | 4              | 43.778517             | -79.346556             | Starbucks            | 43.777990      | -79.344091      | Coffee Shop    |
| 622 | Fairview, Henry Farm, Oriole | 0.075758    | 4              | 43.778517             | -79.346556             | Tim Hortons          | 43.777964      | -79.344715      | Coffee Shop    |
| 623 | Fairview, Henry Farm, Oriole | 0.075758    | 4              | 43.778517             | -79.346556             | Tim Hortons          | 43.774993      | -79.346303      | Coffee Shop    |



Lets view cluster 0 Neighborhoods as this has no Coffee Shops

Out[38]:

|    | Neighborhood    | Coffee Shop | Cluster Labels | Neighborhood Latitude | Neighborhood Longitude | Venue                            | Venue Latitude | Venue Longitude | Venue Category            |
|----|-----------------|-------------|----------------|-----------------------|------------------------|----------------------------------|----------------|-----------------|---------------------------|
| 0  | Agincourt       | 0.0         | 0              | 43.794200             | -79.262029             | Panagio's Breakfast & Lunch      | 43.792370      | -79.260203      | Breakfast Spot            |
| 1  | Agincourt       | 0.0         | 0              | 43.794200             | -79.262029             | Twilight                         | 43.791999      | -79.258584      | Lounge                    |
| 2  | Agincourt       | 0.0         | 0              | 43.794200             | -79.262029             | El Pulgarcito                    | 43.792648      | -79.259208      | Latin American Restaurant |
| 3  | Agincourt       | 0.0         | 0              | 43.794200             | -79.262029             | Mark's                           | 43.791179      | -79.259714      | Clothing Store            |
| 4  | Agincourt       | 0.0         | 0              | 43.794200             | -79.262029             | Commander Arena                  | 43.794867      | -79.267989      | Skating Rink              |
| 33 | Bayview Village | 0.0         | 0              | 43.786947             | -79.385975             | Sun Star Chinese Cuisine<br>華星小炒 | 43.787914      | -79.381234      | Chinese Restaurant        |
| 34 | Bayview Village | 0.0         | 0              | 43.786947             | -79.385975             | TD Canada Trust                  | 43.788074      | -79.380367      | Bank                      |
| 35 | Bayview Village | 0.0         | 0              | 43.786947             | -79.385975             | Maxim's Cafe and Patisserie      | 43.787863      | -79.380751      | Café                      |

## Conclusion Section

### Conclusion

1. Cluster 7 Neighborhoods has the maximum coffee shops to live around, in Neighborhoods like Berczy Park, First Canadian Place, Underground city and Richmond, Adelaide, King.
2. Cluster 4 Neighborhood has the second most coffee shops to live around, in Neighborhoods like Church and Wellesley, Garden District, Ryerson & St. James Town.
3. Cluster 0 can do with some coffee shops, around Neighborhoods like Agincourt & Bayview Village.