

## 1. Introduction/Business Problem

In this new age of globalization & the opening of worldwide opportunities for any talented individual generates a new reality that is: that one person might work in different cities or even countries in their lifetime (For this example we will be using New York & Toronto). Let's say a person got a job offer from an international company with great growth opportunities in another city or other borough/neighborhood of his/her current city. That person would probably like to shift a location that resembles their current lifestyle and that implies that he/she will like to have coffee shops, gyms, libraries, malls or specialized stores. My goal is to find out what are borough/neighborhoods have similar services, amenities or store so that a person shifts within the city or from one city to another city her/his current way of life can be maintained as much as possible (If that person wishes it of course). The idea is that is that this will help them to know where they can live as similar as possible as where they currently live.

## 2 Data

### 2.1 New York Data

I'll be using the borough/neighborhood data of two cities. The first city would be New York and data of this can be downloaded from this [link](#) . This data would be *json* format and it can be transformed into *pandas dataframe*. I would take only those are important for me like borough, neighbourhood, latitude and longitude. For more information of this data you can visit this [site](#).

|   | Borough | Neighbourhood | Latitude  | Longitude  |
|---|---------|---------------|-----------|------------|
| 0 | Bronx   | Wakefield     | 40.894705 | -73.847201 |
| 1 | Bronx   | Co-op City    | 40.874294 | -73.829939 |
| 2 | Bronx   | Eastchester   | 40.887556 | -73.827806 |
| 3 | Bronx   | Fieldston     | 40.895437 | -73.905643 |
| 4 | Bronx   | Riverdale     | 40.890834 | -73.912585 |

### 2.2 Toronto Data

The second city is Toronto which Postal Code, borough and neighborhood will be scrapped from [wikipedia page](#) There are some rows where "Not assigned" is written. the rows where borough is "Not assigned" (only) will be dropped. If neighborhood is 'Not assigned' but Borough is assigned, then I will make the corresponding borough as neighborhood. Merge the rows if Postal Code and Borough of two or more rows same and merged neighborhood will be separated by comma ",". Latitude and longitude information can be downloaded from [here](#). Finally this data and Toronto data can be merged together. Once both data are ready, we'll use Foursquare API to get the venues near each neighborhood. To get the nearby venues data we must have Foursquare account and app which will give the credentials.

|   | Postal Code | 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 |
| 5 | M1J         | Scarborough | Scarborough Village                             | 43.744734 | -79.239476 |
| 6 | M1K         | Scarborough | Kennedy Park, Ionview, East Birchmount Park     | 43.727929 | -79.262029 |
| 7 | M1L         | Scarborough | Golden Mile, Clairlea, Oakridge                 | 43.711112 | -79.284577 |
| 8 | M1M         | Scarborough | Cliffside, Cliffcrest, Scarborough Village West | 43.716316 | -79.239476 |
| 9 | M1N         | Scarborough | Birch Cliff, Cliffside West                     | 43.692657 | -79.264848 |

### 3. Methodology

After getting above data, I'll find the nearby venues to each borough and neighborhood & pair them in both cities. Foursquare API will be used to get the any nearby venue. To use the Foursquare API I'll need the developer account and an app which will give the *client id* and *client secrets*. Both cities may yield in different number of venue categories, but I'll take only common venue categories. Now we can find a borough and neighborhood of a city and see how much similar to the boroughs of another city it is.

In other words, I can find top similar boroughs and neighborhoods in another city.

### 4. Getting Venues Using Foursquare API

Foursquare API will be used to get the nearby venues. Now I will be going to find the venues in 1km range from a borough and neighborhood pair.

#### 4.1 For New York City

The shape of this data is (306, 474). 302 columns are venue categories column.

|   | Borough       | Neighbourhood | Latitude  | Longitude  | Accessories Store | Adult Boutique | Alghan Restaurant | African Restaurant | Airport Lounge | Airport Service | Airport Terminal | American Restaurant | Amphitheater | Animal Shelter | Antique Shop | Aquarium | Arcade | Arepa Restaurant | Argentinian Restaurant | Art Gallery | Art Museum | Art & Crafts Store | Entertai |
|---|---------------|---------------|-----------|------------|-------------------|----------------|-------------------|--------------------|----------------|-----------------|------------------|---------------------|--------------|----------------|--------------|----------|--------|------------------|------------------------|-------------|------------|--------------------|----------|
| 0 | Staten Island | St. George    | 40.644982 | -74.079353 | 0.0               | 0.0            | 0.0               | 0.0                | 0.0            | 0.0             | 0.0              | 0.013699            | 0.0          | 0.0            | 0.0          | 0.0      | 0.0    | 0.0              | 0.0                    | 0.013699    | 0.000000   | 0.0                | 0.0      |
| 1 | Staten Island | New Brighton  | 40.640615 | -74.087017 | 0.0               | 0.0            | 0.0               | 0.0                | 0.0            | 0.0             | 0.0              | 0.026316            | 0.0          | 0.0            | 0.0          | 0.0      | 0.0    | 0.0              | 0.0                    | 0.000000    | 0.000000   | 0.0                | 0.0      |
| 2 | Staten Island | Stapleton     | 40.626928 | -74.077902 | 0.0               | 0.0            | 0.0               | 0.0                | 0.0            | 0.0             | 0.0              | 0.000000            | 0.0          | 0.0            | 0.0          | 0.0      | 0.0    | 0.0              | 0.0                    | 0.000000    | 0.000000   | 0.0                | 0.0      |
| 3 | Staten Island | Rosebank      | 40.615305 | -74.069805 | 0.0               | 0.0            | 0.0               | 0.0                | 0.0            | 0.0             | 0.0              | 0.000000            | 0.0          | 0.0            | 0.0          | 0.0      | 0.0    | 0.0              | 0.0                    | 0.000000    | 0.022222   | 0.0                | 0.0      |
| 4 | Staten Island | West Brighton | 40.631879 | -74.107182 | 0.0               | 0.0            | 0.0               | 0.0                | 0.0            | 0.0             | 0.0              | 0.018182            | 0.0          | 0.0            | 0.0          | 0.0      | 0.0    | 0.0              | 0.0                    | 0.000000    | 0.000000   | 0.0                | 0.0      |

#### 4.2 For Toronto City

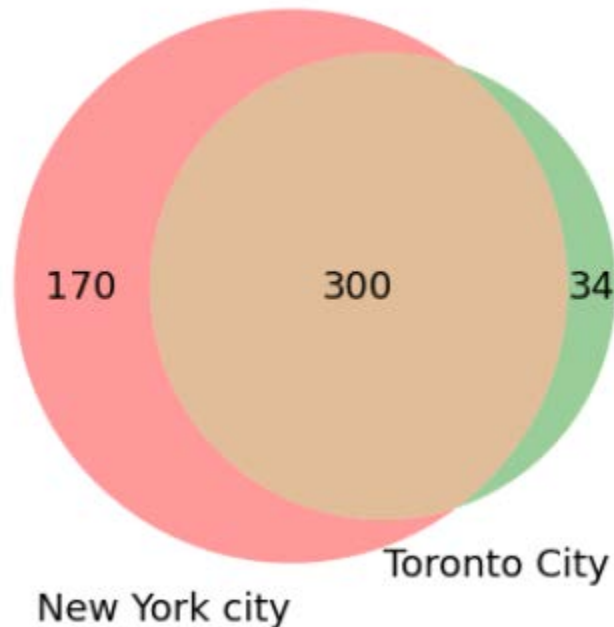
The shape of this data is (103, 339). 334 columns are venue categories columns.

| Postal Code | Borough | Neighbourhood | Latitude   | Longitude | Accessories Store | Afghan Restaurant | African Restaurant | Airport | American Restaurant | Amphitheater | Animal Shelter | Antique Shop | Aquarium | Art Gallery | Art Museum | Arts & Crafts Store | Asian Restaurant | Athletics & Sports | Auto Dealership | Auto Workshop | Automotive Shop |
|-------------|---------|---------------|--|-----------|-------------------|-------------------|--------------------|---------|---------------------|--------------|----------------|--------------|----------|-------------|------------|---------------------|------------------|--------------------|-----------------|---------------|-----------------|
| 0           | MBV     | Etobicoke     | New Toronto, Mimico South, Humber Bay Shores       | 43.605647 | -79.501321        | 0.0               | 0.0                | 0.0     | 0.0                 | 0.055556     | 0.0            | 0.0          | 0.0      | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                | 0.0             | 0.0           | 0.000000        |
| 1           | MBW     | Etobicoke     | Alderswood, Long Branch                            | 43.602414 | -79.543484        | 0.0               | 0.0                | 0.0     | 0.0                 | 0.000000     | 0.0            | 0.0          | 0.0      | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                | 0.0             | 0.0           | 0.000000        |
| 2           | MBX     | Etobicoke     | The Kingsway, Montgomery Road, Old Mill North      | 43.663654 | -79.506944        | 0.0               | 0.0                | 0.0     | 0.0                 | 0.000000     | 0.0            | 0.0          | 0.0      | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                | 0.0             | 0.0           | 0.000000        |
| 3           | MBY     | Etobicoke     | Old Mill South, King's Mill Park, Sunnyside, Hu... | 43.636258 | -79.498509        | 0.0               | 0.0                | 0.0     | 0.0                 | 0.000000     | 0.0            | 0.0          | 0.0      | 0.0         | 0.0        | 0.0                 | 0.000000         | 0.0                | 0.0             | 0.0           | 0.000000        |
| 4           | MBZ     | Etobicoke     | Mimico NW, The Queensway West, South of River      | 43.628841 | -79.520999        | 0.0               | 0.0                | 0.0     | 0.0                 | 0.000000     | 0.0            | 0.0          | 0.0      | 0.0         | 0.0        | 0.0                 | 0.017241         | 0.0                | 0.0             | 0.0           | 0.017241        |

## 5. Analysis

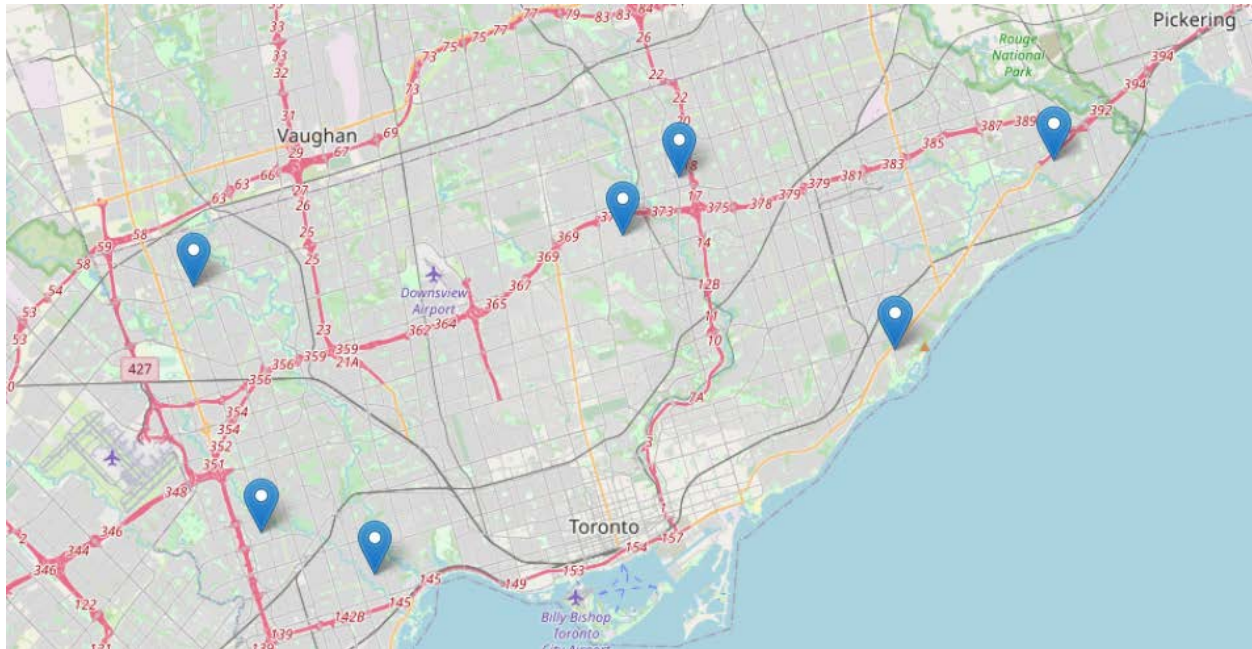
### 5.1 Common Venue Categories

There are some common venue categories in both data. My data is showing is showing that there are total 300 common venue categories in both cities. So, we'll take only these common categories of data. Other than common categories of the data we'll be dropped. This is represented in the figure below:



### 5.2 Top 7 similar borough in Toronto City

Let's say current location is Staten Island, St. George, New York. We'll get the index, latitude and longitude of this location from the New York city data. We can get the all common venue categories data of corresponding index and then we'll multiply with Toronto City data. This will give how much each borough is like your current location. We'll sort these values and get the top 7 boroughs. Then we'll visualize it using the *folium*.



## 6. Conclusion

This model can be implemented within any city. This will help to find a suitable place for people. This will give most similar borough and neighborhoods. Saving time, with real data & providing the chance to anyone to make a thoroughly analysis when they want to change cities where they want to work. In other simplifies something based on your lifestyle. Once again this can be done with other data sets such as prizes, population etc...