The Battle of Neighborhoods

1. Introduction/Business Problem

- Let's say a person got a job offer from a great company with great career prospects in other city or other borough of his/her current city.
- I think a person would love to shift to a location which is exactly or almost similar to his/her last location
- He/she loves the great amenities and other types of venues that exist in his/her current neighbourhood like school, gym, swimming pool, Amusement park, restaurants, coffee-shops, spencer etc.
- So I'll find out what are borough-neighbourhoods are very similar to give current location.

2. Data

New York Data

	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

Toronto Data

	Postal Code	Borough	Neighbourhood	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

3. Getting Venues Using Foursquare API

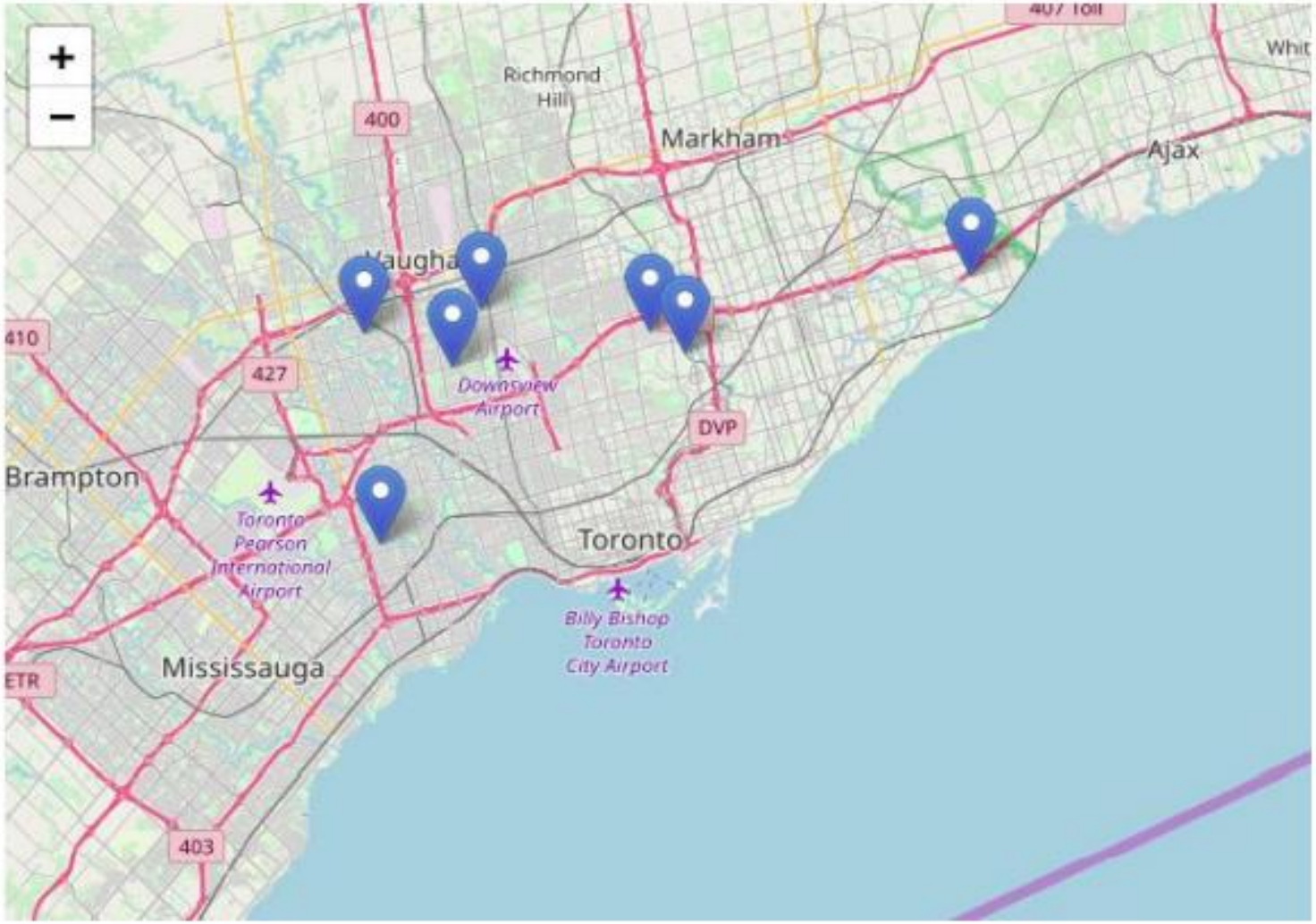
	Borough	Neighbourhood	Latitude	Longitude	ATM	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	A S
0	Staten Island	St. George	40.644982	-74.079353	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.014085	
1	Staten Island	New Brighton	40.640615	-74.087017	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.022222	
2	Staten Island	Stapleton	40.626928	-74.077902	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.027027	
3	Staten Island	Rosebank	40.615305	-74.069805	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	
4	Staten Island	West Brighton	40.631879	-74.107182	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.016129	

Figure - 3: New York Data with Venue Categories

4. Analysis – Common Venue Categories

There are some common venue categories in both data. Figure 4 is showing that there are total 304 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.

4. Analysis - Top 7 similar borough in Toronto City



- Top 7 similar borough in Toronto City Let's say current location is Bronx, Riverdale, New York. We'll get the index, latitude and longitude of this location form 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 similar to current location. We'll sort these values and get the top 7 boroughs. Then we'll visualize it using the folium.

5. Conclusion

- Cosine similarity is used here to find how similarity between two boroughs. This model can be implemented within a city also. This will help to find a suitable place for people. This will give most similar borough and neighbourhood