The Battle of Neighborhood

New York Vs Toronto

Suruthi Vinoth Kannan
4/20/2020

The Battle of Neighborhood

NYC Vs Toronto

1

Introduction

1.1 Background

In this project we are going to explore the neighborhoods of the most popular cities, New York and Toronto. They are the international centers of business, finance, arts, culture, and are recognized as the most multicultural cosmopolitans in the world. They are diverse in many ways. One interesting idea would be to compare the neighborhoods of the two cities and determine how similar or dissimilar they are.

1.2 Problem Description

Let me explain the idea of this project through a scenario. Say you live in the New York City of USA. You love your neighborhood, mainly because of all the great amenities and other types of venues that exist in the neighborhood, such as gourmet fast food joints, pharmacies, parks, and so on. Now say you receive a job offer from a great company which is in Toronto, Canada with great career prospects. However, given the far distance from your current place you unfortunately must move if you decide to accept the offer.

Wouldn't it be great if you are able to determine a neighborhood that are the same as your current neighborhood, and if not, perhaps a similar neighborhood that are at least closer to your new job?

1.3 Target Audience

We will study and analyze the neighborhoods of both the cities and group them into similar clusters and, analyze those clusters to gather meaningful information. That information can be used to find out neighborhoods that are same as your current neighborhood or at least similar.

The information provided by this project would be useful for people who are interested in relocating to different places and are interested in finding new neighborhoods that are highly similar to their existing neighborhood.

The Battle of Neighborhood

NYC Vs Toronto

2

Data Description

2.1 Data

The NYC neighborhood data exists for free on the web. Here is the link to the dataset:
https://geo.nyu.edu/catalog/nyu_2451_34572.
For Toronto neighborhood data, a Wikipedia page:
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M,
exists that has all the information we needed to explore. We will scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas data frame.

Note: There are different website scraping libraries and packages in Python. We can simply use pandas to read the table into a pandas data frame. Another way, for more complicated cases of web scraping is using the Beautiful Soup package. Here is the package's main documentation page:

http://beautiful-soup-4.readthedocs.io/en/latest/

3

Methodology

3.1 Method

For this problem, we will get the services of Foursquare API to explore the data of two cities, 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. In order to utilize the Foursquare location data, we need to get the latitude and the longitude coordinates of each neighborhood. Here is a link to a csv file that has the geographical coordinates of each postal code of Toronto: http://cocl.us/Geospatial_data.

We will use the machine learning technique, "Clustering" to segment the neighborhoods with similar objects on the basis of each neighborhood data. Finally, you will use the Folium library to visualize the neighborhoods in New York City and their emerging clusters.

The Battle of Neighborhood

NYC Vs Toronto

4

Results/ Visualizations

Data of NYC and Toronto neighborhoods are created after web scraping and some preprocessing process. The neighborhood data is visualized in map with the help of folium.

```
ny_neighborhoods.head()
```

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

```
citymap('New York','USA',ny_neighborhoods)
```
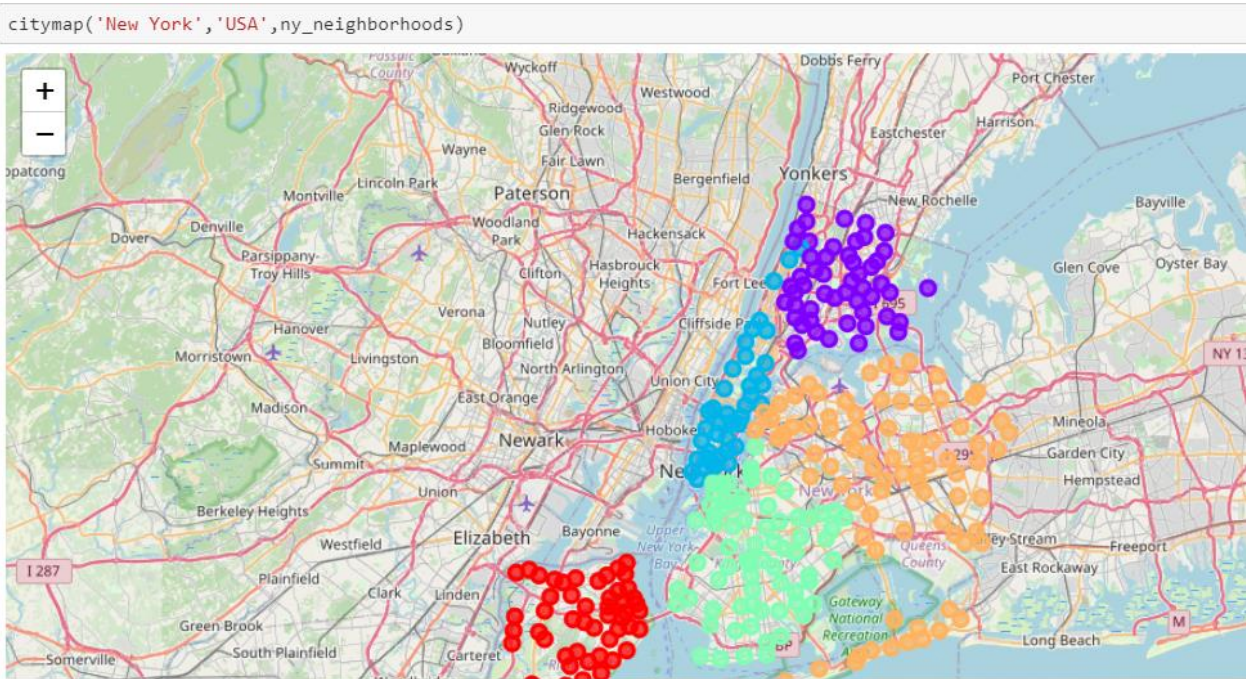
```
toronto_neighborhoods.head()
```

	City	Borough	Neighborhood	Latitude	Longitude
0	Toronto	North York	Parkwoods	43.758800	-79.320197
1	Toronto	North York	Victoria Village	43.732658	-79.311189
2	Toronto	Etobicoke	Islington Avenue	43.622575	-79.514215
3	Toronto	Scarborough	Malvern / Rouge	43.809196	-79.221701
4	Toronto	North York	Don Mills	43.775347	-79.345944

```
citymap('Toronto','Canada',toronto_neighborhoods)
```
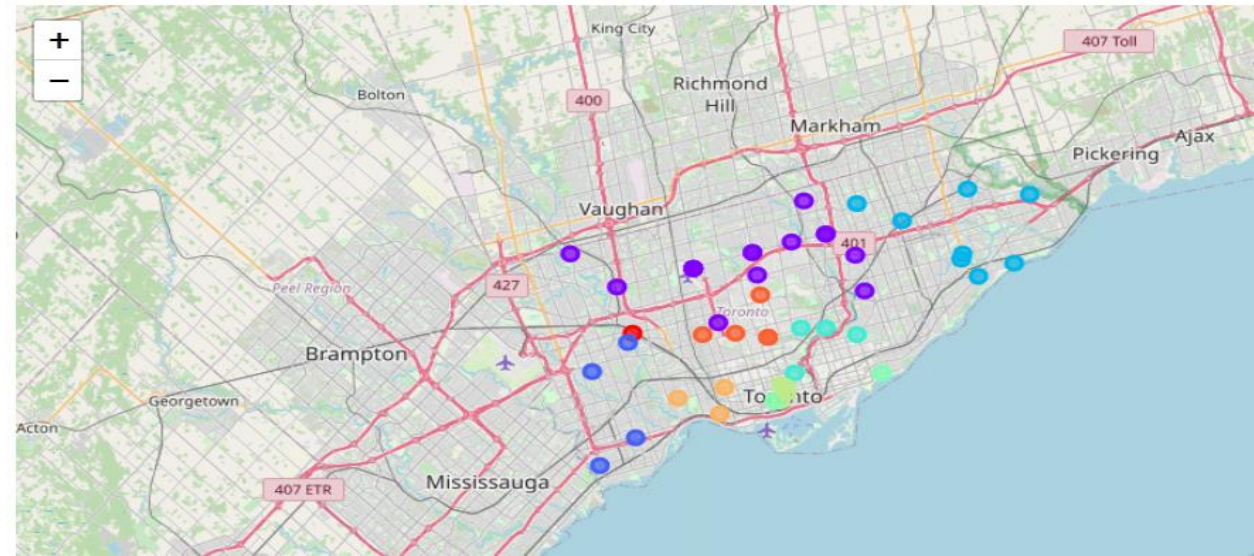

The NYC and Toronto neighborhoods are combined together into a single data frame, for further analysis.

```
allneighborhoods.head()
```

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

The venues around NYC and Toronto neighborhoods are collected
using foursquare API and put it together in a single data frame, for
further analysis.

```
# obtain the nearby venues of each neighborhood in New York
ny_venues = getNearbyVenues(ny_neighborhoods)
ny_venues.head()
```

	City	Borough	Neighborhood	VenueName	VenueCategory
0	New York	Bronx	Wakefield	Lollipops Gelato	Dessert Shop
1	New York	Bronx	Wakefield	Ripe Kitchen & Bar	Caribbean Restaurant
2	New York	Bronx	Wakefield	Ali's Roti Shop	Caribbean Restaurant
3	New York	Bronx	Wakefield	Carvel Ice Cream	Ice Cream Shop
4	New York	Bronx	Wakefield	Jimbo's	Burger Joint

```
# obtain the nearby venues of each neighborhood in toronto
toronto_venues = getNearbyVenues(toronto_neighborhoods)
toronto_venues.head()
```

	City	Borough	Neighborhood	VenueName	VenueCategory
0	Toronto	North York	Parkwoods	Allwyn's Bakery	Caribbean Restaurant
1	Toronto	North York	Parkwoods	LCBO	Liquor Store
2	Toronto	North York	Parkwoods	Tim Hortons	Café
3	Toronto	North York	Parkwoods	Dollarama	Discount Store
4	Toronto	North York	Parkwoods	A&W	Fast Food Restaurant

```
#check the new all venues df
allvenues.head()
```

	City	Borough	NeighborhoodName	VenueName	VenueCategory
0	New York	Bronx	Wakefield	Lollipops Gelato	Dessert Shop
1	New York	Bronx	Wakefield	Ripe Kitchen & Bar	Caribbean Restaurant
2	New York	Bronx	Wakefield	Ali's Roti Shop	Caribbean Restaurant
3	New York	Bronx	Wakefield	Carvel Ice Cream	Ice Cream Shop
4	New York	Bronx	Wakefield	Jimbo's	Burger Joint

The one hot encoded venue and the neighborhoods data are joined
and sorted with 10 most common venues on each neighborhood. The
encoded data are then fitted to the cluster model for prediction.

```
# join all the venues and neighborhood data
allvenues_grouped = allneighborhoods.join(allvenues_grouped, on='NeighborhoodName')
allvenues_grouped.head(2)
```

|	City	Borough	NeighborhoodName	Latitude	Longitude	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	Airport Lounge	Airport Service	American Restaurant	Anim Shelte
0	New York	Bronx	Wakefield	40.894705	-73.847201	0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	New York	Bronx	Co-op City	40.874294	-73.829939	0.02	0.0	0.0	0.0	0.0	0.0	0.0	0.0

```
allvenues_sorted.head()
```

	City	Borough	NeighborhoodName	Latitude	Longitude	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	8 C
0	New York	Bronx	Wakefield	40.8947	-73.8472	Pharmacy	Supermarket	Caribbean Restaurant	Donut Shop	Gas Station	Chinese Restaurant	Fast Food Restaurant	Fri Ch Joi
1	New York	Bronx	Co-op City	40.8743	-73.8299	Shopping Mall	Kids Store	Mattress Store	Mexican Restaurant	Mobile Phone Shop	Furniture / Home Store	Supermarket	Fa Re
2	New York	Bronx	Eastchester	40.8876	-73.8278	Caribbean Restaurant	Shopping Mall	Fast Food Restaurant	Diner	Discount Store	Donut Shop	Mobile Phone Shop	Bu Joi
3	New York	Bronx	Fieldston	40.8954	-73.9056	Pizza Place	Bus Station	Bar	Sandwich Place	Bank	Plaza	Deli / Bodega	Pa
4	New York	Bronx	Riverdale	40.8908	-73.9126	Pizza Place	Bank	Bar	Park	Plaza	Playground	Dog Run	Ba Sh

```
allvenues_sorted.tail()
```

	City	Borough	NeighborhoodName	Latitude	Longitude	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 Com Ve
348	Toronto	Downtown Toronto	Rosedale	43.6563	-79.3809	Furniture / Home Store	Caribbean Restaurant	Coffee Shop	Sandwich Place	Fast Food Restaurant	Cosmetics Shop	Japane Restau
349	Toronto	Etobicoke	Alderwood / Long Branch	43.6017	-79.5452	Discount Store	Pharmacy	Convenience Store	Pizza Place	Coffee Shop	Sandwich Place	Park
350	Toronto	Etobicoke	Northwest	43.6715	-79.5525	Liquor Store	Coffee Shop	Café	Pharmacy	Grocery Store	Pet Store	Sushi Restau
351	Toronto	Scarborough	Upper Rouge	43.8049	-79.1658	Park	Trail	Zoo Exhibit	Bus Station	Campground	Gas Station	Shoppi Mall
352	Toronto	Downtown Toronto	Church and Wellesley	43.6612	-79.3821	Coffee Shop	Japanese Restaurant	Gastropub	Park	Plaza	Mexican Restaurant	Salon / Barber

The predictions of the clustering model are visualized as dendrogram given below.

Hierarchical Clustering Dendrogram
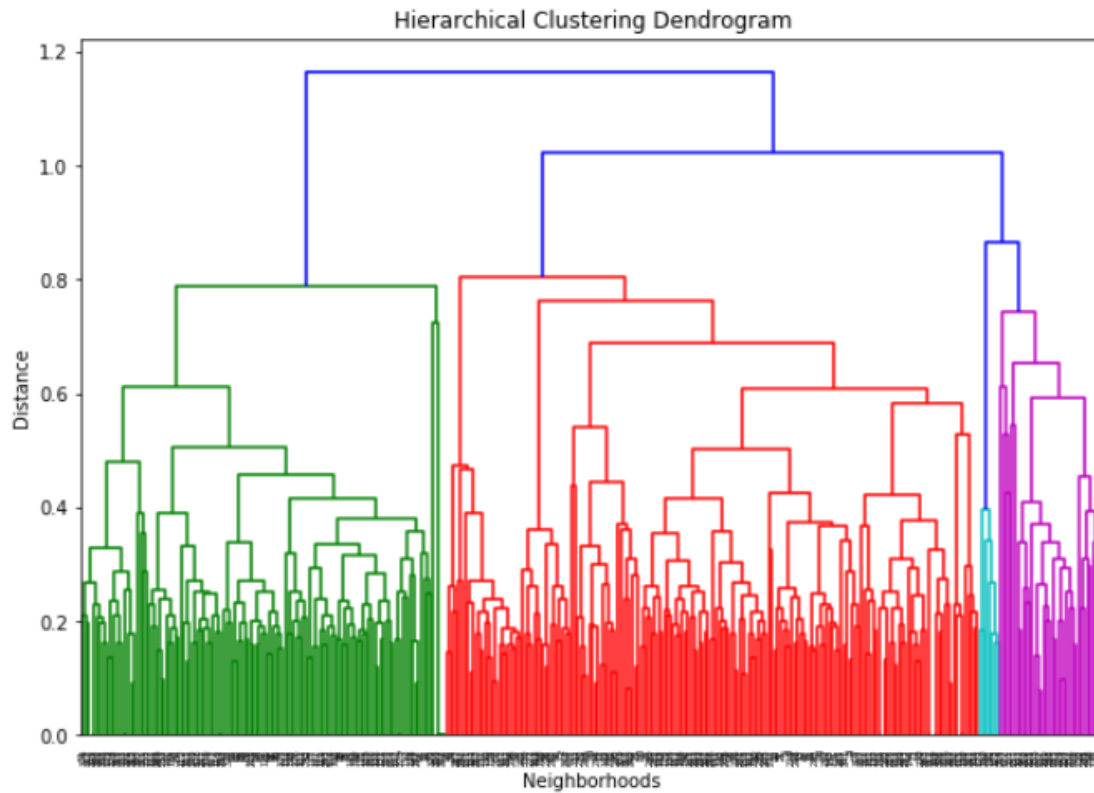
Total number of neighborhoods in each cluster

```
for i in range(7):
    num = allvenues_sorted[allvenues_sorted['Cluster_Labels'] == i].shape[0]
    print('Total number of neighborhoods in cluster {} is {}'.format(i, num))
```

```
Total number of neighborhoods in cluster 0 is 36
Total number of neighborhoods in cluster 1 is 5
Total number of neighborhoods in cluster 2 is 141
Total number of neighborhoods in cluster 3 is 7
Total number of neighborhoods in cluster 4 is 26
Total number of neighborhoods in cluster 5 is 121
Total number of neighborhoods in cluster 6 is 17
```

Let's take a look on each cluster, to understand it better.

Cluster 1

	City	Borough	NeighborhoodName	Latitude	Longitude	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 Cor V
327	Toronto	North York	Downsview	43.7493	-79.4622	Athletics & Sports	Coffee Shop	Swim School	Sandwich Place	Soccer Field	Sporting Goods Shop	French Restaurant	Gas Stat
328	Toronto	North York	Downsview	43.7493	-79.4622	Athletics & Sports	Coffee Shop	Swim School	Sandwich Place	Soccer Field	Sporting Goods Shop	French Restaurant	Gas Stat
329	Toronto	North York	Humber Summit	43.7601	-79.5718	Sports Bar	Mexican Restaurant	Coffee Shop	Filipino Restaurant	Ethiopian Restaurant	Event Service	Event Space	Exh
330	Toronto	North York	Downsview	43.7493	-79.4622	Athletics & Sports	Coffee Shop	Swim School	Sandwich Place	Soccer Field	Sporting Goods Shop	French Restaurant	Gas Stat
334	Toronto	North York	Downsview	43.7493	-79.4622	Athletics & Sports	Coffee Shop	Swim School	Sandwich Place	Soccer Field	Sporting Goods Shop	French Restaurant	Gas Stat

Cluster 1 is not an outlier, it's a neighborhood with the sports complex.

Cluster 3

	City	Borough	NeighborhoodName	Latitude	Longitude	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue	Common Venue	C
85	New York	Brooklyn	Sea Gate	40.5764	-74.0079	Beach	Supermarket	Spa	Basketball Court	Chinese Restaurant	Park	Pharmacy	D S
178	New York	Queens	Rockaway Beach	40.5828	-73.8224	Beach	Bar	Beach Bar	Supermarket	Bagel Shop	Deli / Bodega	Eastern European Restaurant	W
179	New York	Queens	Neponsit	40.572	-73.8575	Beach	Spa	Deli / Bodega	Pub	Harbor / Marina	Boutique	Bakery	P
190	New York	Queens	Belle Harbor	40.5762	-73.854	Beach	Spa	Pub	Mexican Restaurant	Smoke Shop	Bakery	Bagel Shop	E S
191	New York	Queens	Rockaway Park	40.5803	-73.8415	Beach	Pizza Place	Donut Shop	Pub	Bank	Bar	Bagel Shop	S
204	New York	Staten Island	South Beach	40.5802	-74.0796	Beach	Pier	Food	Playground	Bus Stop	Athletics & Sports	American Restaurant	B F
302	New York	Queens	Hammels	40.5873	-73.8055	Beach	Surf Spot	Board Shop	Bar	Supermarket	Donut Shop	Wine Bar	T P

Cluster 3 is not also an oulier, it's a neighborhood with beach as the most common venue

Cluster 6

223	New York	Brooklyn	Rugby	40.6556	-73.9269	Caribbean Restaurant	Pizza Place	Bank	Mobile Phone Shop	Donut Shop	Pharmacy	Grocery Store	S P
259	New York	Brooklyn	Remsen Village	40.6521	-73.9167	Caribbean Restaurant	Donut Shop	Fast Food Restaurant	Deli / Bodega	Fried Chicken Joint	Discount Store	Liquor Store	S P
261	New York	Brooklyn	Paerdegat Basin	40.6313	-73.9023	Caribbean Restaurant	Harbor / Marina	Grocery Store	Pizza Place	Gym	Skate Park	Fast Food Restaurant	A R
281	New York	Brooklyn	Weeksville	40.675	-73.9305	Caribbean Restaurant	Southern / Soul Food Restaurant	Discount Store	Café	Park	Pizza Place	Juice Bar	D S
300	New York	Brooklyn	Erasmus	40.6469	-73.9482	Caribbean Restaurant	Mobile Phone Shop	Bank	Discount Store	Pharmacy	Lounge	Gym / Fitness Center	P P

Cluster 6 contains the neighborhoods in New York with a lot of nearby Caribbean restaurants, pizza places and fast food restaurants.

Cluster 4

217	New York	Staten Island	Tottenville	40.5053	-74.2466	Italian Restaurant	Deli / Bodega	Pizza Place	Bakery	Sandwich Place	Beach	Bagel Shop	T
230	New York	Staten Island	Old Town	40.5963	-74.0875	Italian Restaurant	Pizza Place	Bank	Bus Stop	Deli / Bodega	Gas Station	Pharmacy	P P
231	New York	Staten Island	Dongan Hills	40.5887	-74.0964	Italian Restaurant	Pizza Place	Bank	Sandwich Place	Ice Cream Shop	Donut Shop	Baseball Field	P
236	New York	Staten Island	Huguenot	40.5319	-74.1917	Bank	Park	Train Station	Sandwich Place	Moving Target	Ice Cream Shop	Donut Shop	N
251	New York	Staten Island	Richmond Town	40.5696	-74.1341	Italian Restaurant	Donut Shop	Japanese Restaurant	Art Museum	Thrift / Vintage Store	Grocery Store	Bank	E
287	New York	Staten Island	Egbertville	40.5791	-74.1273	Clothing Store	Bagel Shop	Trail	Health & Beauty Service	Flower Shop	Park	Massage Studio	C P
292	New York	Staten Island	Lighthouse Hill	40.5765	-74.1379	Golf Course	Italian Restaurant	Bagel Shop	Cosmetics Shop	Art Museum	Café	History Museum	I S

Cluster 4 contains the neighborhoods in New York with a lot of nearby Italian restaurants and pizza places.

Cluster 0

256	New York	Staten Island	Randall Manor	40.6356	-74.0981	Bus Stop	Park	Bar	Coffee Shop	Fast Food Restaurant	Ice Cream Shop
257	New York	Staten Island	Howland Hook	40.6384	-74.1862	Harbor / Marina	Peruvian Restaurant	Border Crossing	Lounge	Waterfront	Park
284	New York	Staten Island	Manor Heights	40.6018	-74.1206	Chinese Restaurant	Donut Shop	Bus Stop	Bagel Shop	Liquor Store	Pharmacy
285	New York	Staten Island	Willowbrook	40.6037	-74.1321	Bus Stop	Chinese Restaurant	Deli / Bodega	Bagel Shop	Liquor Store	Convenience Store
288	New York	Queens	Roxbury	40.5674	-73.8921	Baseball Field	Beach	Theater	Art Gallery	Food Truck	Sports Club
305	New York	Staten Island	Fox Hills	40.6173	-74.0817	Bus Stop	Spanish Restaurant	Scenic Lookout	Hardware Store	Sandwich Place	Discount Store
314	Toronto	Scarborough	Guildwood / Morningside / West Hill	43.7535	-79.18	Restaurant	Baseball Field	Convenience Store	Park	Filipino Restaurant	Ethiopian Restaurant
351	Toronto	Scarborough	Upper Rouge	43.8049	-79.1658	Park	Trail	Zoo Exhibit	Bus Station	Campground	Gas Station

Cluster 0 contains the neighborhoods a lot of recreation activities like theme park, baseball field, park, surf spot, ferry, zoo, campground and many.

Cluster 5

	City	Borough	NeighborhoodName	Latitude	Longitude							
340	Toronto	Central Toronto	Forest Hill North & West	43.701	-79.4256	Coffee Shop	Sushi Restaurant	Pizza Place	Italian Restaurant	Trail	Skating Rink	Dry Cleaner
342	Toronto	North York	Willowdale	43.7615	-79.4109	Coffee Shop	Grocery Store	Japanese Restaurant	Fried Chicken Joint	Ramen Restaurant	Burger Joint	Café
343	Toronto	West Toronto	Parkdale / Roncesvalles	43.6399	-79.4397	Bakery	Tibetan Restaurant	Park	Indian Restaurant	Café	Gift Shop	Diner
345	Toronto	Central Toronto	Davisville	43.6979	-79.3973	Sushi Restaurant	Italian Restaurant	Coffee Shop	Restaurant	Indian Restaurant	Park	Gaming Cafe
346	Toronto	West Toronto	Runnymede / Swansea	43.6518	-79.4759	Coffee Shop	Café	Pub	Italian Restaurant	Bakery	Bank	Sushi Restaurant
350	Toronto	Etobicoke	Northwest	43.6715	-79.5525	Liquor Store	Coffee Shop	Café	Pharmacy	Grocery Store	Pet Store	Sushi Restaurant
352	Toronto	Downtown Toronto	Church and Wellesley	43.6612	-79.3821	Coffee Shop	Japanese Restaurant	Gastropub	Park	Plaza	Mexican Restaurant	Salon / Barbershop

Cluster 5 contains the neighborhoods with a lot of Bar, Pub, Lounge, japanese restaurant and sushi places.

Cluster 2

```
cluster2.head(10)
```

|	City	Borough	NeighborhoodName	Latitude	Longitude	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
0	New York	Bronx	Wakefield	40.8947	-73.8472	Pharmacy	Supermarket	Caribbean Restaurant	Donut Shop	Gas Station	Chinese Restaurant	Fast Food Restaurant
1	New York	Bronx	Co-op City	40.8743	-73.8299	Shopping Mall	Kids Store	Mattress Store	Mexican Restaurant	Mobile Phone Shop	Furniture / Home Store	Supermarket
3	New York	Bronx	Fieldston	40.8954	-73.9056	Pizza Place	Bus Station	Bar	Sandwich Place	Bank	Plaza	Deli / Bodega
4	New York	Bronx	Riverdale	40.8908	-73.9126	Pizza Place	Bank	Bar	Park	Plaza	Playground	Dog Run
5	New York	Bronx	Kingsbridge	40.8817	-73.9028	Pizza Place	Mexican Restaurant	Donut Shop	Bar	Burger Joint	Bakery	Spanish Restaurant
6	New York	Manhattan	Marble Hill	40.8766	-73.9107	Mexican Restaurant	Sandwich Place	Café	Pizza Place	Spanish Restaurant	Coffee Shop	Donut Shop
337	Toronto	York	Weston	43.7002	-79.5162	Park	Train Station	Coffee Shop	Breakfast Spot	Laundromat	Skating Rink	Bar
341	Toronto	Etobicoke	Westmount	43.6936	-79.521	Pizza Place	Park	Bank	Train Station	Gas Station	Convenience Store	Lau
344	Toronto	Scarborough	Agincourt	43.7854	-79.2785	Chinese Restaurant	Vietnamese Restaurant	Shopping Mall	Cantonese Restaurant	Restaurant	Korean Restaurant	Cal
347	Toronto	Scarborough	Steeles West / L'Amoreaux West	43.7982	-79.318	Chinese Restaurant	Pizza Place	Grocery Store	Bank	Bakery	Coffee Shop	Fas Re:
348	Toronto	Downtown Toronto	Rosedale	43.6563	-79.3809	Furniture / Home Store	Caribbean Restaurant	Coffee Shop	Sandwich Place	Fast Food Restaurant	Cosmetics Shop	Jap Re:
349	Toronto	Etobicoke	Alderwood / Long Branch	43.6017	-79.5452	Discount Store	Pharmacy	Convenience Store	Pizza Place	Coffee Shop	Sandwich Place	Pal

Cluster 2 contains the neighborhoods in Malls, Supermarkets, Pharmacy, furniture and stores.

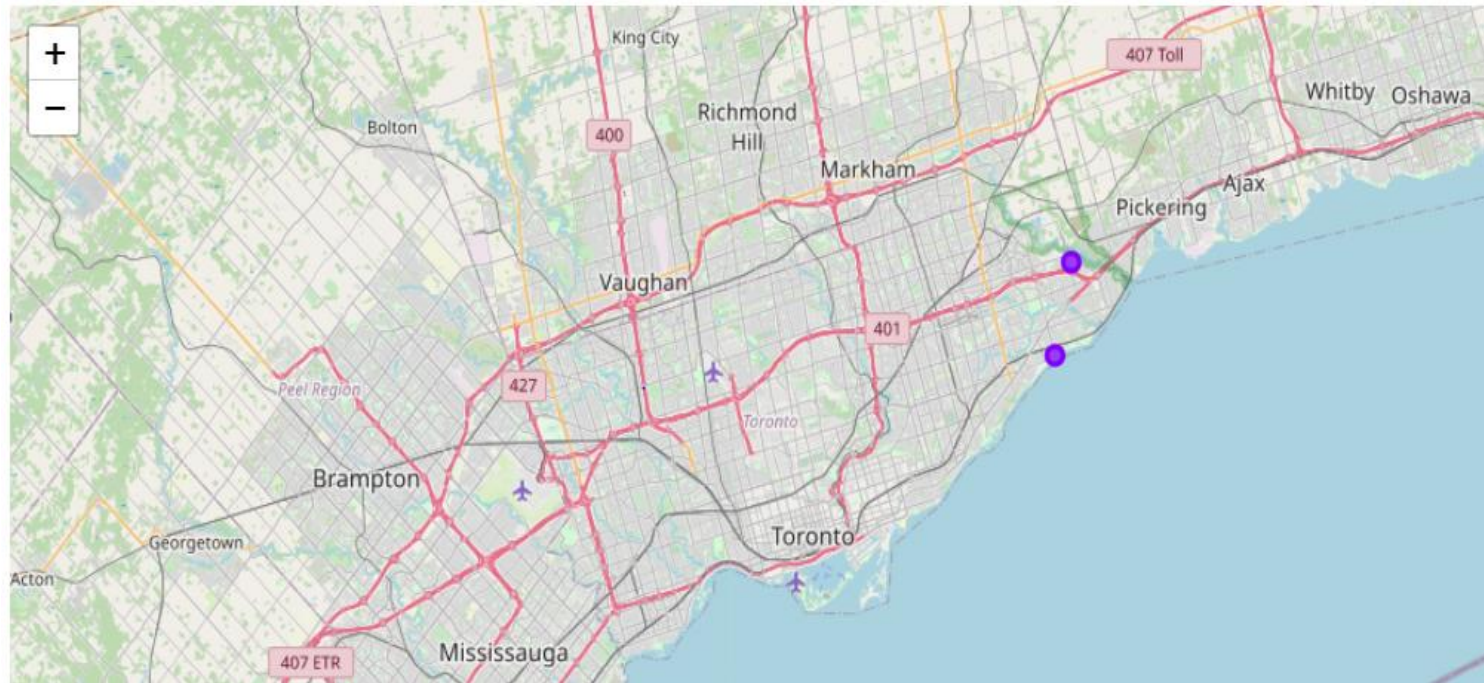# Explore the neighborhoods with similar clusters using map

Cluster 0 - Neighborhoods with Recreational places in New York

```
clustermap('New York','USA',cluster0)
```
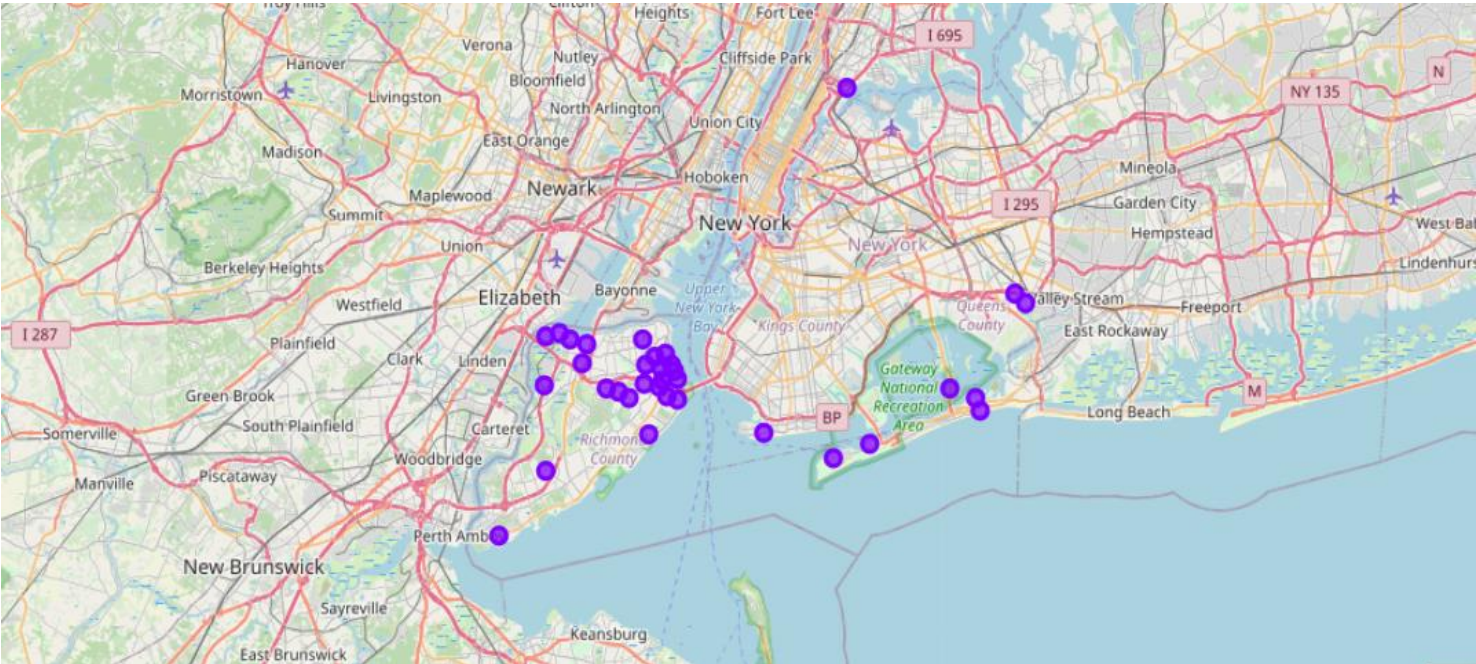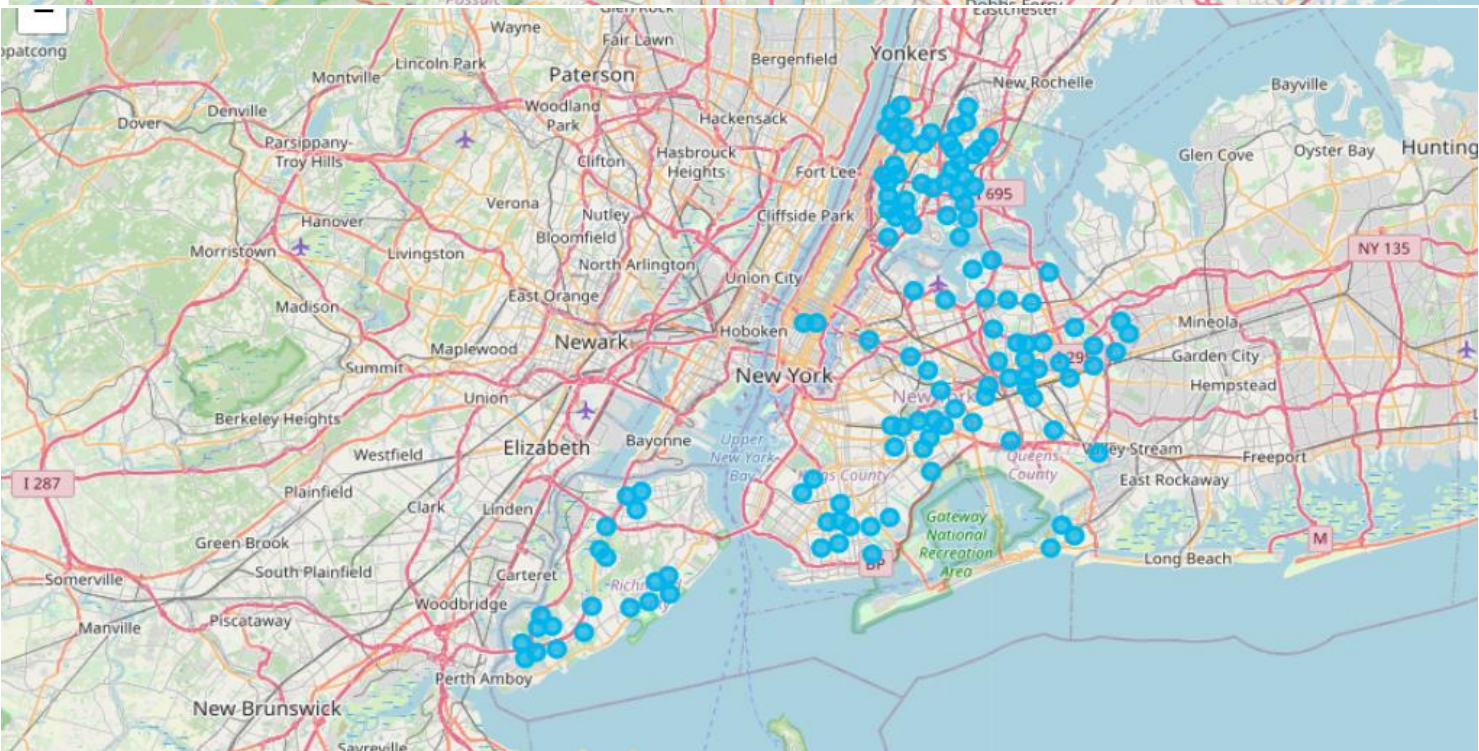

Cluster 0 - Neighborhoods with Recreational places in Toronto

```
clustermap('Toronto','Canada',cluster0)
```
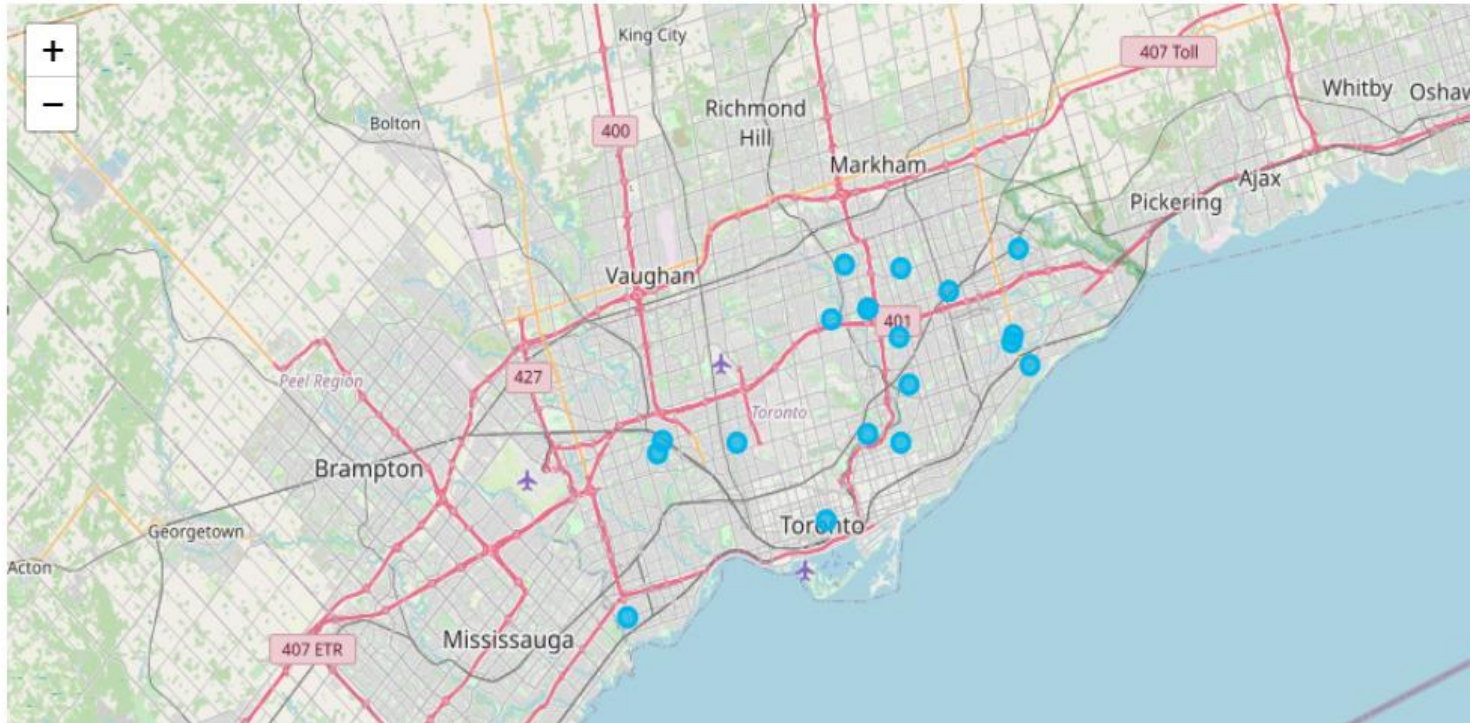
Cluster2 - Neighborhood with Suprmarkets, Malls, Pharmacies, Furniture and other stores

```
clustermap('New York','USA',cluster2[cluster2['City'] == 'New York'])
```



```
clustermap('Toronto','Canada',cluster2[cluster2['City'] == 'Toronto'])
```

The Battle of Neighborhood
NYC Vs Toronto

6

Conclusion

6.1 Conclusion

With all the analysis we have done, we can conclude that,

Cluster 0 - Neighborhood with all type of Recreational centers

Cluster 1 - Neighborhood with Sports Complex

Cluster 2 - Neighborhood with Super markets, malls - (Residential)

Cluster 3 - Neighborhood with Beach

Cluster 4 - Neighborhood with more no. of Italian restaurants

Cluster 5 - Neighborhood with Pub, Bar and Sushi places

Cluster 6 - Neighborhood with more no. of Caribbean restaurants