THE BATTLE OF NEIGHBORHOODS

Comparison between
North York, Toronto and Queen, New York

INTRODUCTION

Problem

- The objective is to compare North York, Toronto and Queens, New York City to determine how they are similar or dissimilar.

Interest

- A person who is considering moving will be interested since it will allow them to compare the most common venues and strategically choose based on their needs or likes.

Toronto-Data acquisition and cleaning

1. The website https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M will be scrapped using Beautiful Soup retrieving data such as Postal Code, Boroughs, Neighborhood.
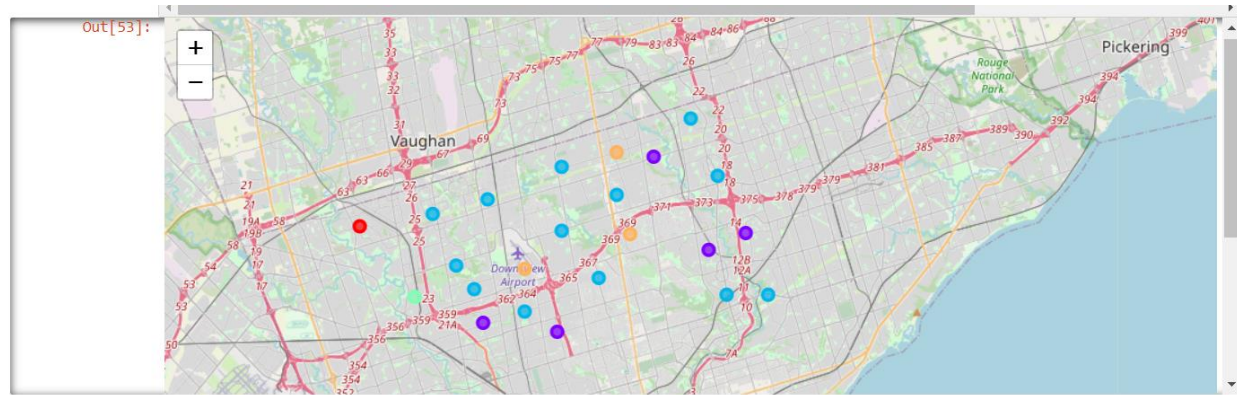
2. Geocoder Package, https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv, will provide Latitude and Longitude Coordinates of neighbourhoods in Toronto with their Postal Code.

3. Merge both dataframes and filter neighborhoods in North York.

Out[10]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
3	M3B	North York	Don Mills North	43.745906	-79.352188
4	M6B	North York	Glencairn	43.709577	-79.445073

4. Use Foursquare API, https://foursquare.com/developers/apps, to get venue data facilitating the 10 most common venues.

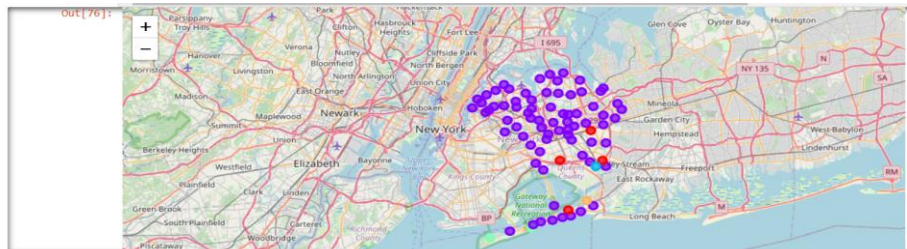5. Segment data into cluster using k-means clustering.

6. Merge neighborhood dataframe, venue dataframe and clusters.

New York-Data acquisition and cleaning

1. Use the url='https://cocl.us/new_york_dataset' to retrieve the Borough, Neighborhood, Latitude and Longitude of New York and filter neighborhoods in Queens.

Out[17]:

	Borough	Neighborhood	Latitude	Longitude
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

2. Use Foursquare API, https://foursquare.com/developers/apps, to get venue data facilitating the 10 most common venues.

3. Segment data into cluster using k-means clustering

4.Merge neighborhood dataframe, venue dataframe and clusters.

Results and Discussion

Cluster 1

Out[62]:

	Borough	Neighborhood	Cluster Labels	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	North York	Humber Summit	0	Gym	Pizza Place	Construction & Landscaping	Women's Store	Discount Store	Comfort Food Restaurant	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store

Cluster 2

Out[64]:

Borough	Neighborhood	Cluster Labels	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
North York	Parkwoods	1	Food & Drink Shop	Fast Food Restaurant	Park	Women's Store	Dim Sum Restaurant	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop
North York	Don Mills North	1	Gym	Caribbean Restaurant	Café	Dessert Shop	Japanese Restaurant	Women's Store	Discount Store	Construction & Landscaping	Convenience Store	Cosmetics Shop
North York	Glencairn	1	Park	Japanese Restaurant	Asian Restaurant	Bakery	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store
North York	Bayview Village	1	Chinese Restaurant	Café	Bank	Japanese Restaurant	Women's Store	Discount Store	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega
North York	North Park, Maple Leaf Park, Upwood Park	1	Park	Construction & Landscaping	Massage Studio	Bakery	Discount Store	Comfort Food Restaurant	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store

```
Construction & Landscaping    5
Convenience Store             5
Cosmetics Shop                5
Park                          3
Women's Store                 3
Comfort Food Restaurant       3
Japanese Restaurant           3
Discount Store                3
Deli / Bodega                 3
Café                          2
dtype: int64
```

Cluster 3

`67]:`

	Borough	Neighborhood	Cluster Labels	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	North York	Victoria Village	2	Coffee Shop	Hockey Arena	Portuguese Restaurant	Intersection	Diner	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega
2	North York	Lawrence Manor, Lawrence Heights	2	Clothing Store	Furniture / Home Store	Women's Store	Vietnamese Restaurant	Arts & Crafts Store	Athletics & Sports	Boutique	Coffee Shop	Event Space	Accessories Store
5	North York	Don Mills South	2	Coffee Shop	Gym	Restaurant	Sandwich Place	Clothing Store	Chinese Restaurant	Dim Sum Restaurant	Italian Restaurant	Discount Store	Bike Shop
6	North York	Hillcrest Village	2	Golf Course	Pool	Fast Food Restaurant	Mediterranean Restaurant	Dog Run	Dim Sum Restaurant	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	Convenience Store
		Bathurst Manor, Wilson				Mobile		Middle					

```
Coffee Shop                  11
Convenience Store             8
Construction & Landscaping    7
Comfort Food Restaurant       5
Deli / Bodega                 4
Dim Sum Restaurant            4
Women's Store                 4
Pizza Place                   4
Discount Store                4
Grocery Store                 4
dtype: int64
```

Cluster 4

`Out[69]:`

	Borough	Neighborhood	Cluster Labels	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
19	North York	Humberlea, Emery	3	Baseball Field	Women's Store	Distribution Center	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop

Cluster 5

`Out[71]:`

	Borough	Neighborhood	Cluster Labels	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	North York	Downsview East	4	Airport	Park	Women's Store	Diner	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega
6	North York	Willowdale, Newtonbrook	4	Park	Women's Store	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop
2	North York	York Mills West	4	Park	Convenience Store	Women's Store	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop

```
Park                        3
Women's Store               3
Convenience Store           3
Deli / Bodega               3
Construction & Landscaping  3
Cosmetics Shop              3
Comfort Food Restaurant     3
Coffee Shop                 3
Department Store            2
Dessert Shop                2
dtype: int64
```

Toronto

- In cluster 2 the most common venues are Construction & Landscaping, Convenience Store and Cosmetics Shop.

- In cluster 3 the most common venue is a Coffee Shop.

- In cluster 5 the most common venues are Park, Women's Store, Convenience Store, Deli / Bodega, Construction & Landscaping, Cosmetics Shop, Comfort Food Restaurant and Coffee Shop.

Cluster 1

Out[77]:

	Borough	Neighborhood	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
19	Queens	South Ozone Park	Park	Deli / Bodega	Bar	Food Truck	Grocery Store	Hotel	Donut Shop	Fast Food Restaurant	Sandwich Place	Event Space
34	Queens	Hollis	Park	Baseball Field	Shopping Mall	Mobile Phone Shop	Chinese Restaurant	Lounge	Sandwich Place	Discount Store	Asian Restaurant	Bakery
59	Queens	Laurelton	Caribbean Restaurant	Cosmetics Shop	Train Station	Park	Sculpture Garden	Dance Studio	Creperie	Flower Shop	Fish Market	Fish & Chips Shop
63	Queens	Somerville	Park	Surf Spot	Yoga Studio	Farmers Market	Electronics Store	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Fast Food Restaurant

```
Park                    4
Fast Food Restaurant    2
Event Space             2
Sandwich Place          2
Chinese Restaurant      1
Train Station           1
Hotel                   1
Bar                     1
Creperie                1
Caribbean Restaurant    1
dtype: int64
```

Cluster 2

Out[79]:

	Borough	Neighborhood	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Queens	Astoria	Bar	Middle Eastern Restaurant	Greek Restaurant	Hookah Bar	Café	Seafood Restaurant	Bakery	Indian Restaurant	Pizza Place	Italian Restaurant
1	Queens	Woodside	Grocery Store	Thai Restaurant	Filipino Restaurant	Latin American Restaurant	Bakery	Donut Shop	American Restaurant	Pub	Bar	Sandwich Place
2	Queens	Jackson Heights	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Bakery	Mobile Phone Shop	Thai Restaurant	Supermarket	Diner	Pizza Place	Grocery Store
3	Queens	Elmhurst	Mexican Restaurant	Thai Restaurant	South American Restaurant	Bubble Tea Shop	Vietnamese Restaurant	Chinese Restaurant	Bank	Snack Place	Fast Food Restaurant	Seafood Restaurant
4	Queens	Howard Beach	Pharmacy	Italian Restaurant	Fast Food Restaurant	Sandwich Place	Breakfast Spot	Supermarket	Jewelry Store	Tapas Restaurant	Mexican Restaurant	Donut Shop

```
Pizza Place           34
Deli / Bodega         28
Donut Shop            25
Chinese Restaurant    22
Bakery                22
Bank                  21
Sandwich Place        19
Supermarket           18
Grocery Store         17
Pharmacy              17
dtype: int64
```

Cluster 3

Out[81]:

	Borough	Neighborhood	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
64	Queens	Brookville	Deli / Bodega	Yoga Studio	Filipino Restaurant	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Fish & Chips Shop

Cluster 4

	Borough	Neighborhood	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
43	Queens	Breezy Point	Beach	Monument / Landmark	Trail	Bus Stop	Fast Food Restaurant	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market
50	Queens	Neponsit	Beach	Lounge	Yoga Studio	Food Truck	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

```
Empanada Restaurant      2
Fast Food Restaurant     2
Event Space              2
Beach                    2
Farm                     2
Falafel Restaurant       2
Farmers Market           2
Trail                    1
Yoga Studio              1
Bus Stop                 1
dtype: int64
```

Cluster 5

	Borough	Neighborhood	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 Most Common Venue	9th Most Common Venue	10th Most Common Venue
79	Queens	Bayswater	Playground	Yoga Studio	Fast Food Restaurant	Electronics Store	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market	Filipino Restaurant

New York

- In cluster 1 the most common venue is the Park.

- In cluster 2 the most common venue is the Pizza Place.

- In cluster 4 the most common venues are Empanada Restaurant, Fast Food Restaurant, Event Space, Beach, Farm, Falafel Restaurant and Farmers Market.

- North York and Queens have Park as a most common venue.

Conclusion

This data can be of great used to persons who are interested in moving since they can view the data and based on their preference/needs they can select a neighborhood based on an informed decision to live according to the 10 most common venues. As well as we can see the most common venue in each cluster, this can narrow their interest.

A woman will most likely to move to North York, Toronto since Women's Store and Cosmetics Shop are common.

However, a family will most likely move to Queens, New York since Parks and Beaches are common.

Different types of clustering algorithms can be applied to cluster the neighborhoods and more than 100 venues in a neighborhood can be considered for analysis using Foursquare API.

In the same manner, data can also be processed for other cities.