

CAPSTONE PROJECT

THE BATTLE OF NEIGHBORHOODS

Comparison between North York, Toronto and Queen, New York

INTRODUCTION

Background

North York, Toronto is 176.87 km² with a multicultural and diverse population of 869,401 recorded in 2016 that has a lot of history. Two of the largest shopping centers are Yorkdale Shopping Centre and Fairview Mall. The Don River rushes through North York. Queens, New York is 460 km² with the most ethnically diverse population of 2,230,722 recorded in 2010. Food lovers crave the Greek cuisine of Astoria. Queens Botanical Garden has a spectacular flora. They are both financial capitals that are diverse.

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.

DATA

Toronto

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

New York

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

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

METHODOLOGY

Toronto

- Collect from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. Clean the data such that if a borough is Not assigned ignore, merge neighborhoods with same postal code and if a neighborhood is Not assigned but has a borough then the neighborhood will be the same as the borough. Then process data into a dataframe.

Out[5]:

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Queen's Park	Ontario Provincial Government

- Get postal code, latitude and longitude from https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv into a dataframe.

Out[7]:

	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

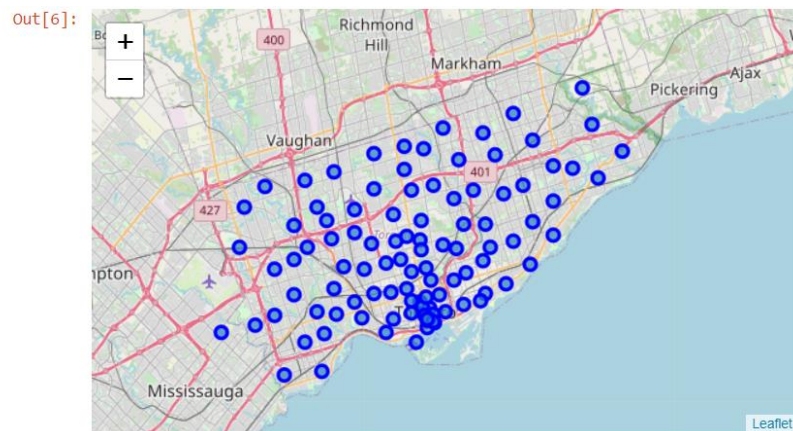
- Merge both dataframes.

Out[9]:

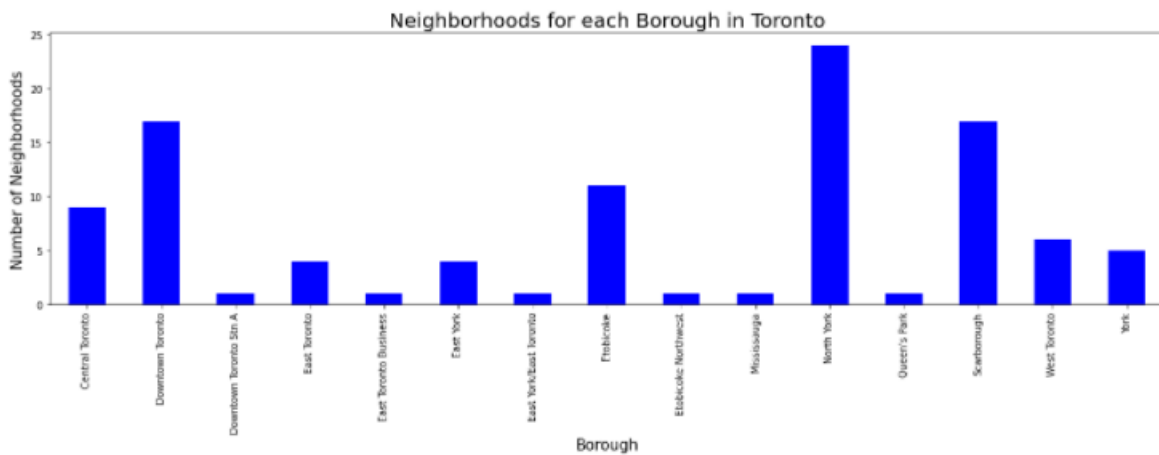
	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494

- Explore the above dataframe

The below map shows all the neighborhoods in Toronto.



Toronto has 15 boroughs and 103 neighborhoods.



- Filter all the data in respect to neighborhoods that belong to North York, Toronto.

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

- Use FourSquare to get all venues and explore then filtered the 10 most common venues.

Out[48]:

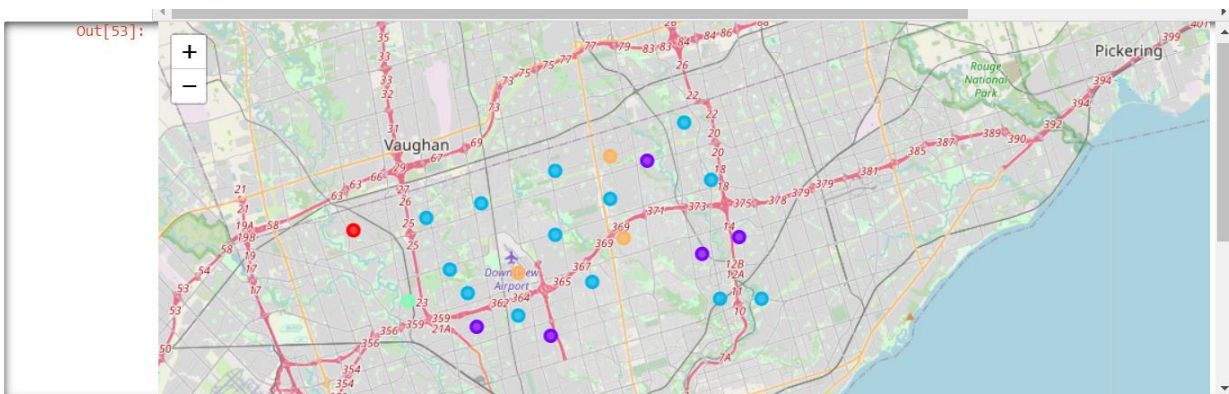
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Mobile Phone Shop	Diner	Middle Eastern Restaurant	Intersection	Pharmacy	Pizza Place	Deli / Bodega	Bridal Shop
1	Bayview Village	Chinese Restaurant	Café	Bank	Japanese Restaurant	Women's Store	Discount Store	Construction & Landscaping	Convenience Store	Cosmetics Shop	Deli / Bodega
2	Bedford Park, Lawrence Manor East	Italian Restaurant	Coffee Shop	Pizza Place	Sandwich Place	Greek Restaurant	Restaurant	Fast Food Restaurant	Grocery Store	Indian Restaurant	Juice Bar
3	Don Mills North	Gym	Caribbean Restaurant	Café	Dessert Shop	Japanese Restaurant	Women's Store	Discount Store	Construction & Landscaping	Convenience Store	Cosmetics Shop
4	Don Mills South	Coffee Shop	Gym	Restaurant	Sandwich Place	Clothing Store	Chinese Restaurant	Dim Sum Restaurant	Italian Restaurant	Discount Store	Bike Shop

- The data will be segmented into clusters using unsupervised k-means clustering.
- Merge dataframes with neighborhood data, venue data and cluster data.

(23, 16)

	PostalCode	Borough	Neighborhood	Latitude	Longitude	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	
0	M3A	North York	Parkwoods	43.753259	-79.329656	1	Food & Drink Shop	Fast Food Restaurant	Park	Women's Store	Dim Sum Restaurant	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	Co
1	M4A	North York	Victoria Village	43.725882	-79.315572	2	Coffee Shop	Hockey Arena	Portuguese Restaurant	Intersection	Diner	Comfort Food Restaurant	Construction & Landscaping	Convenience Store	
2	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	2	Clothing Store	Furniture / Home Store	Women's Store	Vietnamese Restaurant	Arts & Crafts Store	Athletics & Sports	Boutique	Coffee Shop	Ev
3	M3B	North York	Don Mills North	43.745906	-79.352188	1	Gym	Caribbean Restaurant	Café	Dessert Shop	Japanese Restaurant	Women's Store	Discount Store	Construction & Landscaping	Co

- Visualization of clusters.



New York

- Collect data from https://cocl.us/new_york_dataset such as borough, neighborhood, latitude and longitude process it into a dataframe.

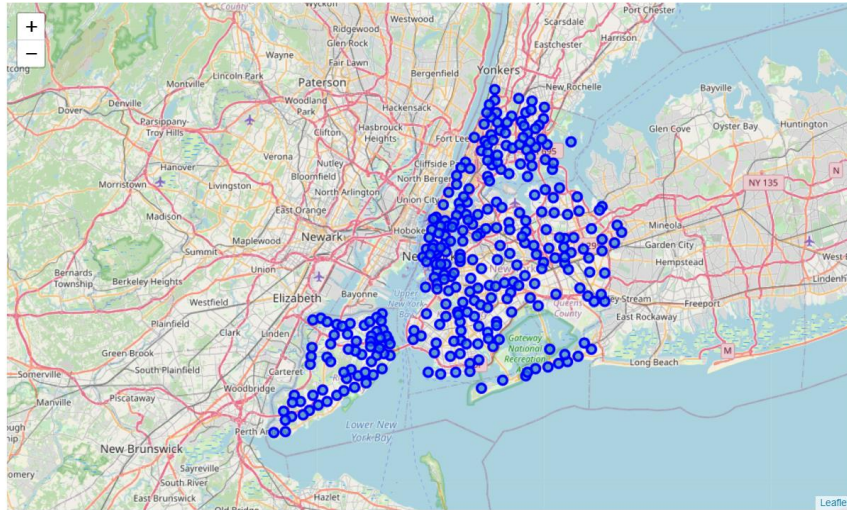
Out[12]:

	Borough	Neighborhood	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

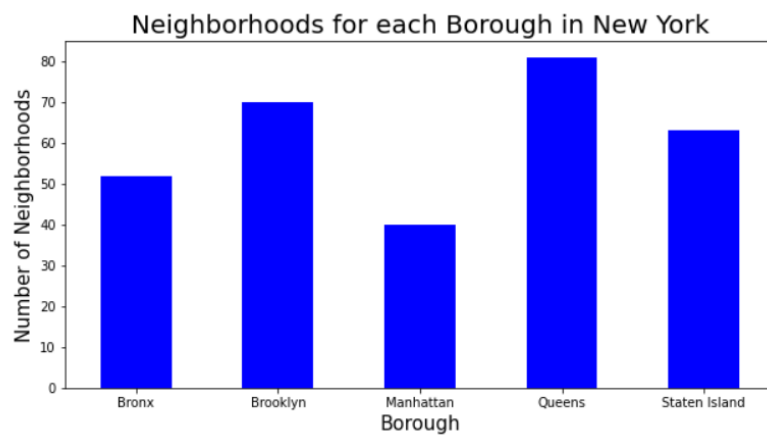
- Explore the data in the dataframe.

The below map show all the neighborhoods in New York City.

Out[13]:



There are 5 boroughs and 306 neighborhoods.



- Filter all the data in respect to neighborhoods that belong to Queens, New York.

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

- Use FourSquare to get all venues and explore then filtered the 10 most common venues.

Out[49]:

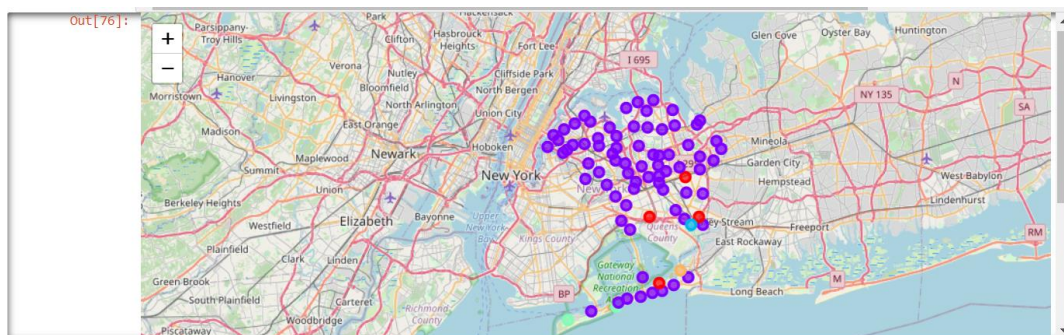
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Arverne	Surf Spot	Metro Station	Sandwich Place	Wine Shop	Coffee Shop	Playground	Bus Stop	French Restaurant	Café	Pizza Place
1	Astoria	Bar	Middle Eastern Restaurant	Greek Restaurant	Hookah Bar	Café	Seafood Restaurant	Bakery	Indian Restaurant	Pizza Place	Italian Restaurant
2	Astoria Heights	Burger Joint	Bakery	Plaza	Playground	Hostel	Pizza Place	Bus Station	Italian Restaurant	Supermarket	Bowling Alley
3	Auburndale	Italian Restaurant	Mattress Store	Train Station	Mobile Phone Shop	Fast Food Restaurant	Supermarket	Noodle House	Bar	Discount Store	Gymnastics Gym
4	Bay Terrace	Clothing Store	American Restaurant	Kids Store	Shoe Store	Mobile Phone Shop	Cosmetics Shop	Donut Shop	Women's Store	Lingerie Store	Gluten-free Restaurant

- The data will be segmented into clusters using unsupervised k-means clustering.
- Merge dataframe with neighborhood dataframe, venue dataframe and cluster data.

(81, 15)

	Borough	Neighborhood	Latitude	Longitude	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
0	Queens	Astoria	40.768509	-73.915654	1	Bar	Middle Eastern Restaurant	Greek Restaurant	Hookah Bar	Café	Seafood Restaurant	Bakery	Indian Restaurant	Pizza Place	Rest
1	Queens	Woodside	40.746349	-73.901842	1	Grocery Store	Thai Restaurant	Filipino Restaurant	Latin American Restaurant	Bakery	Donut Shop	American Restaurant	Pub	Bar	Sar
2	Queens	Jackson Heights	40.751981	-73.882821	1	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Bakery	Mobile Phone Shop	Thai Restaurant	Supermarket	Diner	Pizza Place	G
3	Queens	Elmhurst	40.744049	-73.881656	1	Mexican Restaurant	Thai Restaurant	South American Restaurant	Bubble Tea Shop	Vietnamese Restaurant	Chinese Restaurant	Bank	Snack Place	Fast Food Restaurant	Se Rest

- Visualization of clusters.



Both final dataframes will be segmented according to clusters. The clusters will be explored to see which is the most common venue in each cluster.

Toronto

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, Willowdale				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

New York

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 Shop	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

Out[84]:

	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

Out[88]:

	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

Results and Discussion

There are a great number of different types of venue categories. There are also different neighborhoods in the same row with the same most common venues.

Toronto

North York has 24 neighborhoods with 98 unique venues.

Clusters in descending order with the amount of neighborhoods to each cluster are 2, 1, 4 and 0 and 3.

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.

New York

Queens has 81 neighborhoods with 279 unique venues.

Clusters in descending order with the amount of neighborhoods to each cluster are 1, 0, 3 and 1 and 4.

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

The purpose of this project was to compare North York, Toronto and Queens, New York City.

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.