Exploring the Best Place to Live in North York, Toronto

Capstone Project - The Battle of Neighborhoods

LAI Ho Ching *22 Aug 2020*

1. INTRODUCTION

Toronto, is one of the most popular residential places for both local and non-local. As the capital city of the Canadian province of Ontario, Toronto has around a hundred neighbourhoods distributed in several boroughs. In this project, I am focusing on North York, Toronto, to explore its popular shops, facilities and respective clusters of distribution. It aims to provide a list of living place recommendations for non-local, who wish to move to North York.

To understand North York, the following analysis will explore the borough's transport, cultural development, food and beverage industry, markets and recreational facilities. Cuisine and shops can give a broad picture of how people there live, thus a list of recommendations of living places is achievable.

I hope the recommendation can provide information for non-local who are considering moving to North York, and able to pick suitable neighbourhoods they like.

2. DATA SECTION

The following websites, csv file and application are used to search for data:

a) Scrape this Wikipedia page

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M for postal codes of Canada

To gather sufficient information about the borough 'North York' and its neighbourhoods.

b) Download a csv file http://cocl.us/Geospatial_data for geographical coordinates of each postal code

To get the latitude and longitude of each borough and neighbourhoods, as to prepare for getting location data.

c) Apply Foursquare API including my client ID and client secret for exploring and searching functions

To receive the location data of shops and facilities in North York, so that the result can provide information of neighbourhoods with different ways of living and culture.

d) Use JupyterLab, pandas, numpy, matplotlib, folium, geopy, sklearn, etc. to perform data analyze

To apply python in real-life situations and provide reader-friendly visualized results.

e) Utilize GitHub repository for notebook management

To perform version control and source code management effectively.

3. METHODOLOGY SECTION

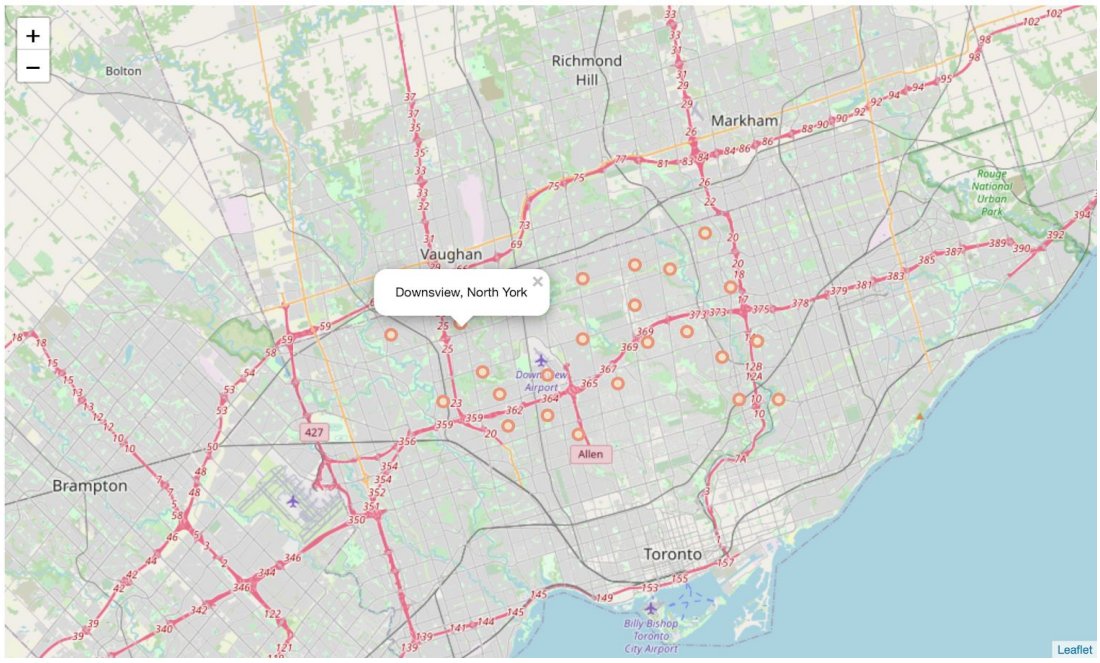
A data frame of only including neighbourhoods in North York is created, as shown below:

[68]:

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
17	M2H	North York	Hillcrest Village	43.803762	-79.363452
18	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556
19	M2K	North York	Bayview Village	43.786947	-79.385975
20	M2L	North York	York Mills, Silver Hills	43.757490	-79.374714
21	M2M	North York	Willowdale, Newtonbrook	43.789053	-79.408493
22	M2N	North York	Willowdale, Willowdale East	43.770120	-79.408493
23	M2P	North York	York Mills West	43.752758	-79.400049
24	M2R	North York	Willowdale, Willowdale West	43.782736	-79.442259
25	M3A	North York	Parkwoods	43.753259	-79.329656
26	M3B	North York	Don Mills	43.745906	-79.352188
27	M3C	North York	Don Mills	43.725900	-79.340923
28	M3H	North York	Bathurst Manor, Wilson Heights, Downsview North	43.754328	-79.442259
29	M3J	North York	Northwood Park, York University	43.767980	-79.487262
30	M3K	North York	Downsview	43.737473	-79.464763
31	M3L	North York	Downsview	43.739015	-79.506944
32	M3M	North York	Downsview	43.728496	-79.495697
33	M3N	North York	Downsview	43.761631	-79.520999
34	M4A	North York	Victoria Village	43.725882	-79.315572
62	M5M	North York	Bedford Park, Lawrence Manor East	43.733283	-79.419750
71	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
72	M6B	North York	Glencairn	43.709577	-79.445073
79	M6L	North York	North Park, Maple Leaf Park, Upwood Park	43.713756	-79.490074
96	M9L	North York	Humber Summit	43.756303	-79.565963
97	M9M	North York	Humberlea, Emery	43.724766	-79.532242

A map of North York with neighbourhoods superimposed on top is created using folium, as shown below:

[116]:



To get the venues nearby North York, Foursquare API is used to get results. The top 10 most common venues of each neighbourhood is printed as below:

[183]:

	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	Park	Bank	Sandwich Place	Middle Eastern Restaurant	Community Center	Mobile Phone Shop	Deli / Bodega	Bridal Shop	Pharmacy
1	Bayview Village	Chinese Restaurant	Café	Bank	Intersection	Japanese Restaurant	Women's Store	Dog Run	Construction & Landscaping	Convenience Store	Cosmetics Shop
2	Bedford Park, Lawrence Manor East	Coffee Shop	Italian Restaurant	Sandwich Place	Juice Bar	Liquor Store	Butcher	Pharmacy	Pizza Place	Café	Hobby Shop
3	Don Mills	Gym	Japanese Restaurant	Restaurant	Coffee Shop	Beer Store	Construction & Landscaping	Café	Italian Restaurant	Discount Store	Asian Restaurant
4	Downsview	Construction & Landscaping	Park	Gas Station	Grocery Store	Vietnamese Restaurant	Coffee Shop	Other Repair Shop	Moving Target	Latin American Restaurant	Liquor Store
5	Fairview, Henry Farm, Oriole	Clothing Store	Coffee Shop	Juice Bar	Restaurant	Bank	Pharmacy	Department Store	Cosmetics Shop	Salon / Barbershop	Electronics Store
6	Glencairn	Grocery Store	Sushi Restaurant	Pub	Japanese Restaurant	Furniture / Home Store	Fried Chicken Joint	Comfort Food Restaurant	Community Center	Construction & Landscaping	Convenience Store
7	Hillcrest Village	Golf Course	Tennis Court	Mediterranean Restaurant	Pool	Dog Run	Fried Chicken Joint	French Restaurant	Comfort Food Restaurant	Community Center	Construction & Landscaping
8	Humber Summit	Furniture / Home Store	Arts & Crafts Store	Pizza Place	Business Service	Women's Store	Discount Store	Community Center	Construction & Landscaping	Convenience Store	Cosmetics Shop
9	Humberlea, Emery	Paper / Office Supplies Store	Shoe Store	Convenience Store	Baseball Field	Furniture / Home Store	Discount Store	Comfort Food Restaurant	Community Center	Construction & Landscaping	Golf Course

Then, I run k-means to cluster the neighbourhood into 5 clusters:

Run k-means to cluster the neighborhood into 5 clusters

[206]:
```
NY_grouped_clustering = NY_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=5, random_state=0).fit(NY_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_
```

[206]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 4, 3, 0, 0, 0, 1, 0, 0, 0, 2], dtype=int32)
```

After that, a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood is created.

[221]:

	Postal Code	Borough	Neighbourhood	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
0	M2H	North York	Hillcrest Village	43.803762	-79.363452	1.0	Golf Course	Tennis Court	Mediterranean Restaurant	Pool	Dog Run	Fried Chicken Joint	French Restaurant
1	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556	0.0	Clothing Store	Coffee Shop	Juice Bar	Restaurant	Bank	Pharmacy	Department Store
2	M2K	North York	Bayview Village	43.786947	-79.385975	0.0	Chinese Restaurant	Café	Bank	Intersection	Japanese Restaurant	Women's Store	Dog Run
3	M2N	North York	Willowdale, Willowdale East	43.770120	-79.408493	0.0	Ramen Restaurant	Coffee Shop	Korean Restaurant	Café	Juice Bar	Sandwich Place	Movie Theater
4	M2P	North York	York Mills West	43.752758	-79.400049	2.0	Park	Gym Pool	Convenience Store	Women's Store	Dog Run	Comfort Food Restaurant	Community Center

To visualize the resulting clusters, folium is used to create a map. Here is the code:

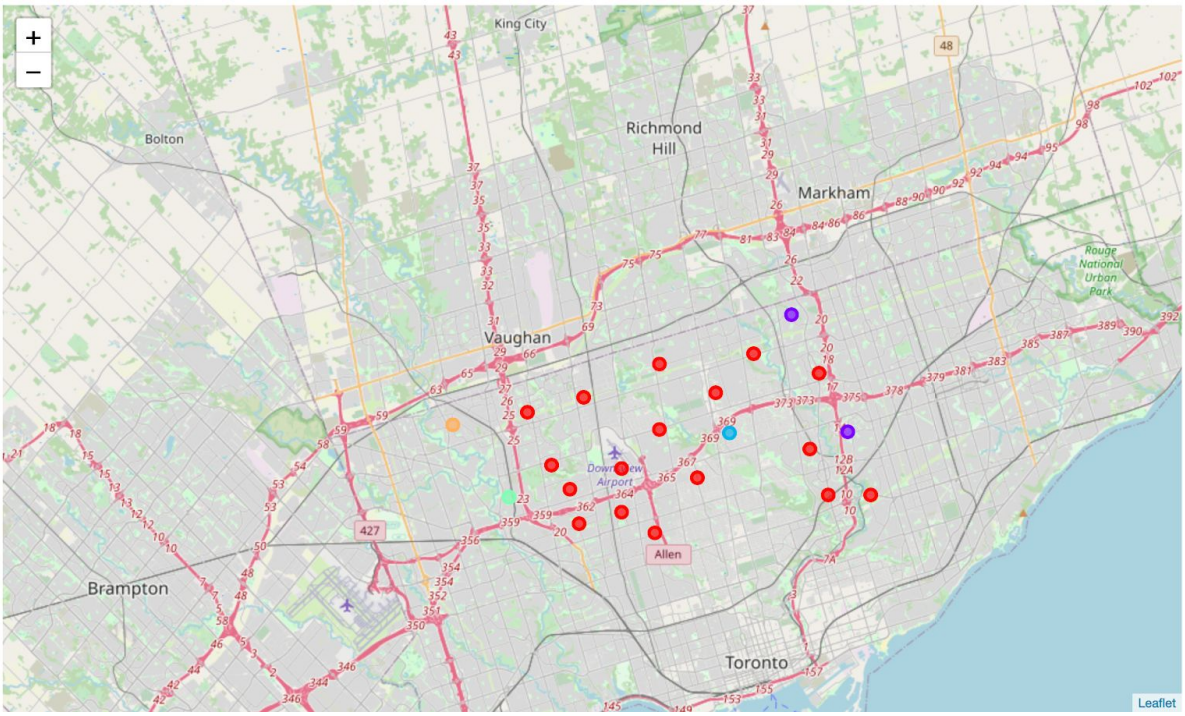
Visualize the resulting clusters

[220]:
```
# create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(NY_merged['Latitude'], NY_merged['Longitude'], NY_merged['Neighbourhood'], NY_merged['Cluster Label
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[int(cluster-1)],
        fill=True,
        fill_color=rainbow[int(cluster-1)],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

4. RESULTS SECTION

Here is the final map with clusters and neighbourhoods label.

[220]:



There are five clusters in North York:

Cluster #	Name of Neighbourhood	Colour of Dot
Cluster 0	Downsview	Red dots
Cluster 1	Hillcrest Village	Purple dots
Cluster 2	York Mills West	Blue dot
Cluster 3	Humberlea, Emery	Green dot
Cluster 4	Humber Summit	Yellow dot

For ***Cluster 0, Downsview*** is close to the airport, where construction & landscaping, park, gas station, grocery store, Vietnamese restaurant and coffee shop are mostly common. From this information, I recommend people who **fly frequently, drive their cars, cook their own meals, love Vietnamese cuisine and are coffee-addicted** to live in areas of red dots.

For ***Cluster 1, Hillcrest Village*** is close to main roads, where golf course, tennis court, Mediterranean restaurant, pool and dog run are the most common venues. Therefore, the places with purple dots are suitable for **golf, tennis or swimming lovers, dog owners and people who love Mediterranean cuisine** to live.

For **Cluster 2, York Mills** West has parks, gym pools, convenience stores, women's stores and dog run. This place marked in blue dot is great for **runners, swimmers, busy urbanites who rely on convenience stores, women who love shopping and dog owners.**

For **Cluster 3,** the neighbourhoods of **Humberlea and Emery** are surrounded by paper/office supplies store, shoe store, convenience store, baseball field and furniture/home store. **Office owners, home designers, shoe lovers, baseball players** will like this place marked in green dot.

For **Cluster 4, Humber Summit** is very close to Downsview. This place has plenty of furniture/home store, arts & crafts store, pizza place, business service and women's store. Thus, I suggest people who **love to decorate their home with artworks and crafts, pizza lovers, business owners and women who like shopping** to live in the area marked in yellow dot.

5. DISCUSSION SECTION

Except for the above mentioned clusters, several neighbourhoods also have distinct cultural and living styles. According to the grouping result:

For cuisine lovers, ***Bayview Village*** has many Chinese restaurants, cafes and Japanese restaurants. ***Willowdale and Willowdale East*** have plenty of ramen restaurants, coffee shops, Korean restaurants, cafes, juice bars and sandwich places.

People who love Japanese culture may like the place ***Glencairn*** where sushi restaurants, Japanese restaurants and grocery stores are very popular.

There is a Hockey arena in ***Victoria Village***, where French restaurants and Portuguese restaurants are common.

6. CONCLUSION

Five clusters of neighbourhoods are suggested for people with different living styles and habits. Four neighbourhoods having special features and facilities are also mentioned. I hope this analysis can help both local and non-local to pick the most suitable place to live in North York, Toronto.