IBM DataScience - Capstone - wk3 Part 3 - Sven De Smit

August 25, 2019

1 IBM DataScience - Capstone - wk3 Part 3 - Sven De Smit

2 Segmenting and Clustering Neighborhoods in Toronto

In [1]:
```
#!conda install -c conda-forge folium=0.5.0 --yes
```

In [2]:
```
import pandas as pd

import requests

import folium

import matplotlib
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

2.1 Read neighborhood dataset from the file created by the previous exercise

In [3]:
```
df_postal_loc = pd.read_csv('toronto_postal_codes_with_location.csv')
print(df_postal_loc.shape)
df_postal_loc.head()
```

```
(103, 5)
```

Out[3]:

	PostalCode	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

2.2 Show neighborhoods on a map

2.2.1 Find geo location of the city to make sure that the map is centered correctly

```
In [4]: from geopy.geocoders import Nominatim

        address = 'Canada, Toronto'

        geolocator = Nominatim(user_agent="toronto_explorer")
        location = geolocator.geocode(address)
        latitude = location.latitude
        longitude = location.longitude
        print('The geograpical coordinate of Toronto City are {}, {}.'.format(latitude, longitu
```

```
The geograpical coordinate of Toronto City are 43.653963, -79.387207.
```

2.2.2 Show the neighborhoods on a city map

```
In [35]: from IPython.core.display import HTML
         map_toronto = folium.Map(location=[latitude +0.05, longitude], zoom_start=11)

         # add markers to map
         for lat, lng, borough, neighborhood in zip(df_postal_loc['Latitude'], df_postal_loc['
             label = '{}, {}'.format(neighborhood, borough)
             label = folium.Popup(label, parse_html=True)
             folium.CircleMarker(
                 [lat, lng],
                 radius=7,
                 popup=label,
                 color='red',
                 fill=True,
                 fill_color='orange',
                 fill_opacity=0.7,
                 parse_html=False).add_to(map_toronto)

         map_toronto
```

```
Out[35]: <folium.folium.Map at 0x20c46fe8748>
```

2.3 Create venue dataset with Foursquare API's

2.3.1 Read Foursquare credentials from CSV file

```
In [6]: df_credentials = pd.read_csv('credentials.txt',header=None)
        CLIENT_ID = df_credentials[1][0]
        CLIENT_SECRET = df_credentials[1][1]
```

2.3.2 Get nearby venues for all rows in the neighborhood dataset

```
In [7]: VERSION = '20180605' # Foursquare API version
        def getNearbyVenues(names, latitudes, longitudes, radius=500):

            venues_list=[]
            for name, lat, lng in zip(names, latitudes, longitudes):
                #print(name)

                # create the API request URL
                url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret
                    CLIENT_ID,
                    CLIENT_SECRET,
                    VERSION,
                    lat,
                    lng,
                    radius,
                    LIMIT)

                # make the GET request
                #print(requests.get(url).json())
                results = requests.get(url).json()["response"]['groups'][0]['items']
                #print(results)
                #print(' ')
                # return only relevant information for each nearby venue
                venues_list.append([(
                    name,
                    lat,
                    lng,
                    v['venue']['name'],
                    v['venue']['location']['lat'],
                    v['venue']['location']['lng'],
                    v['venue']['categories'][0]['name']) for v in results])

            nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_
            nearby_venues.columns = ['Neighborhood',
                          'Neighborhood Latitude',
                          'Neighborhood Longitude',
                          'Venue',
                          'Venue Latitude',
                          'Venue Longitude',
                          'Venue Category']

            return(nearby_venues)
```

```
In [8]: LIMIT = 100
        toronto_venues = getNearbyVenues(names=df_postal_loc['Neighbourhood'],
                                         latitudes=df_postal_loc['Latitude'],
```

```
                                        longitudes=df_postal_loc['Longitude']
                                       )
```

```
In [9]: print('There are {} venues in this dataset.'.format(toronto_venues.shape[0]))
        toronto_venues.head()
```

```
There are 2244 venues in this dataset.
```

Out[9]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Rouge, Malvern	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
2	Guildwood, Morningside, West Hill	43.763573	-79.188711	Swiss Chalet Rotisserie & Grill	43.767697	-79.189914	Pizza Place
3	Guildwood, Morningside, West Hill	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	Electronics Store
4	Guildwood, Morningside, West Hill	43.763573	-79.188711	Big Bite Burrito	43.766299	-79.190720	Mexican Restaurant

2.3.3 Get an idea how much venues we have per neighborhoud

```
In [10]: res = toronto_venues.groupby('Neighborhood')['Venue'].count()
         #print(type(res))
         #print(res.index)
         for n,vc in zip(res.index,res):
             print(vc,'\t',n)
```

```
100          Adelaide, King, Richmond
5           Agincourt
2           Agincourt North, L'Amoreaux East, Milliken, Steeles East
9           Albion Gardens, Beaumond Heights, Humbergate, Jamestown, Mount Olive, Silverstone, S
9           Alderwood, Long Branch
22           Bathurst Manor, Downsview North, Wilson Heights
4           Bayview Village
22           Bedford Park, Lawrence Manor East
57           Berczy Park
5           Birch Cliff, Cliffside West
9           Bloordale Gardens, Eringate, Markland Wood, Old Burnhamthorpe
```

21	Brockton, Exhibition Place, Parkdale Village
19	Business Reply Mail Processing Centre 969 Eastern
2	CFB Toronto, Downsview East
14	CN Tower, Bathurst Quay, Island airport, Harbourfront West, King and Spadina, Railw
45	Cabbagetown, St. James Town
5	Caledonia-Fairbanks
11	Canada Post Gateway Processing Centre
7	Cedarbrae
84	Central Bay Street
100	Chinatown, Grange Park, Kensington Market
15	Christie
84	Church and Wellesley
9	Clairlea, Golden Mile, Oakridge
10	Clarks Corners, Sullivan, Tam O'Shanter
3	Cliffcrest, Cliffside, Scarborough Village West
1	Cloverdale, Islington, Martin Grove, Princess Gardens, West Deane Park
100	Commerce Court, Victoria Hotel
35	Davisville
9	Davisville North
15	Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West
6	Del Ray, Keelesdale, Mount Dennis, Silverthorn
100	Design Exchange, Toronto Dominion Centre
5	Don Mills North
6	Dorset Park, Scarborough Town Centre, Wexford Heights
16	Dovercourt Village, Dufferin
3	Downsview Central
4	Downsview Northwest
5	Downsview West
4	Downsview, North Park, Upwood Park
6	East Birchmount Park, Ionview, Kennedy Park
5	East Toronto
1	Emery, Humberlea
59	Fairview, Henry Farm, Oriole
100	First Canadian Place, Underground city
20	Flemingdon Park, Don Mills South
4	Forest Hill North, Forest Hill West
4	Glencairn
7	Guildwood, Morningside, West Hill
35	Harbord, University of Toronto
100	Harbourfront East, Toronto Islands, Union Station
49	Harbourfront, Regent Park
23	High Park, The Junction South
1	Highland Creek, Rouge Hill, Port Union
6	Hillcrest Village
14	Humber Bay Shores, Mimico South, New Toronto
4	Humber Bay, King's Mill Park, Kingsway Park South East, Mimico NE, Old Mill South, T
2	Humber Summit
4	Humewood-Cedarvale

```
2          Kingsview Village, Martin Grove Gardens, Richview Gardens, St. Phillips
15          Kingsway Park South West, Mimico NW, The Queensway West, Royal York South West, Sou
12          L'Amoreaux West
14          Lawrence Heights, Lawrence Manor
4          Lawrence Park
34          Leaside
65          Little Portugal, Trinity
6          Maryvale, Wexford
4          Moore Park, Summerhill East
1          Newtonbrook, Willowdale
20          North Toronto West
3          Northwest
6          Northwood Park, York University
15          Parkdale, Roncesvalles
3          Parkwoods
40          Queen's Park
5          Rosedale
1          Roselawn
1          Rouge, Malvern
36          Runnymede, Swansea
100          Ryerson, Garden District
3          Scarborough Village
1          Silver Hills, York Mills
100          St. James Town
96          Stn A PO Boxes 25 The Esplanade
39          Studio District
24          The Annex, North Midtown, Yorkville
5          The Beaches
18          The Beaches West, India Bazaar
42          The Danforth West, Riverdale
5          The Junction North, Runnymede
2          The Kingsway, Montgomery Road, Old Mill North
16          Thorncliffe Park
4          Victoria Village
8          Westmount
2          Weston
35          Willowdale South
5          Willowdale West
3          Woburn
11          Woodbine Gardens, Parkview Hill
9          Woodbine Heights
3          York Mills West
```

2.3.4 Get an overview of the venue categories

```
In [11]: print('There are {} uniques categories.'.format(len(toronto_venues['Venue Category'].u
         toronto_venues['Venue Category'].unique()
```

```
There are 280 uniques categories.
```

```
Out[11]: array(['Fast Food Restaurant', 'Bar', 'Pizza Place', 'Electronics Store',
                'Mexican Restaurant', 'Rental Car Location', 'Medical Center',
                'Intersection', 'Breakfast Spot', 'Coffee Shop',
                'Korean Restaurant', 'Hakka Restaurant', 'Caribbean Restaurant',
                'Thai Restaurant', 'Athletics & Sports', 'Bank', 'Bakery',
                'Fried Chicken Joint', 'Playground', 'Spa', 'Convenience Store',
                'Department Store', 'Discount Store', 'Chinese Restaurant',
                'Bus Station', 'Bus Line', 'Metro Station', 'Soccer Field',
                'Motel', 'Movie Theater', 'American Restaurant', 'Café',
                'General Entertainment', 'Farm', 'Skating Rink', 'College Stadium',
                'Indian Restaurant', 'Vietnamese Restaurant', 'Pet Store',
                'Latin American Restaurant', 'Sandwich Place',
                'Middle Eastern Restaurant', 'Shopping Mall', 'Auto Garage',
                'Lounge', 'Italian Restaurant', 'Noodle House', 'Pharmacy', 'Park',
                'Grocery Store', 'Thrift / Vintage Store', 'Nail Salon',
                'Golf Course', 'Pool', 'Mediterranean Restaurant', 'Dog Run',
                'Toy / Game Store', 'Burger Joint', 'Tea Room', 'Candy Store',
                'Juice Bar', 'Salon / Barbershop', 'Smoothie Shop',
                'Clothing Store', 'Theater', 'Japanese Restaurant', 'Food Court',
                'Restaurant', 'Cosmetics Shop', 'Liquor Store', 'Video Game Store',
                'Asian Restaurant', 'Wings Joint', 'Sporting Goods Shop',
                'Burrito Place', 'Shoe Store', 'Deli / Bodega', 'Gift Shop',
                'Boutique', 'Supplement Shop', "Women's Store", 'Luggage Store',
                'Baseball Field', 'Cafeteria', 'Ramen Restaurant', 'Steakhouse',
                'Indonesian Restaurant', 'Arts & Crafts Store', 'Plaza',
                'Ice Cream Shop', 'Sushi Restaurant', 'Bubble Tea Shop', 'Hotel',
                'Butcher', 'Food & Drink Shop', 'Gym / Fitness Center',
                'Basketball Court', 'Gym', 'Bike Shop', 'Beer Store',
                'Dim Sum Restaurant', 'Supermarket', 'Bridal Shop', 'Diner',
                'Frozen Yogurt Shop', 'Video Store', 'Massage Studio',
                'Falafel Restaurant', 'Airport', 'Home Service', 'Food Truck',
                'Hockey Arena', 'Portuguese Restaurant',
                'Financial or Legal Service', 'Gastropub', 'Curling Ice',
                'Bus Stop', 'Trail', 'Health Food Store', 'Pub', 'Neighborhood',
                'Sports Bar', 'Fish & Chips Shop', 'Dessert Shop', 'Brewery',
                'Furniture / Home Store', 'Bagel Shop', 'Warehouse Store',
                'Yoga Studio', 'Housing Development', 'Greek Restaurant',
                'Fruit & Vegetable Store', 'Bookstore', 'Fish Market',
                'Cheese Shop', 'Comfort Food Restaurant', 'Seafood Restaurant',
                'Stationery Store', 'Coworking Space', 'Music Store', 'Gay Bar',
                'Swim School', 'Dance Studio', 'Gourmet Shop', 'Farmers Market',
                'Costume Shop', 'Light Rail Station', 'Building', 'Jewelry Store',
                'Taiwanese Restaurant', 'Market', 'Dive Bar', 'Outdoor Sculpture',
                'Snack Place', 'Theme Restaurant', 'Creperie', 'Hobby Shop',
                'Ethiopian Restaurant', "Men's Store", 'Smoke Shop', 'Sake Bar',
```

```
'Afghan Restaurant', 'Persian Restaurant',
'Health & Beauty Service', 'Strip Club', 'Sculpture Garden',
'Polish Restaurant', 'Historic Site', 'Chocolate Shop',
'Performing Arts Venue', 'French Restaurant', 'Event Space',
'Art Gallery', 'Antique Shop', 'Comic Shop', 'Taco Place',
'Music Venue', 'Beer Bar', 'Vegetarian / Vegan Restaurant',
'College Rec Center', 'Tanning Salon',
'Modern European Restaurant', 'Miscellaneous Shop', 'Hookah Bar',
'Wine Bar', 'Lake', 'Lingerie Store', 'Other Great Outdoors',
'Poutine Place', 'Office', 'BBQ Joint', 'Church', 'Speakeasy',
'Poke Place', 'New American Restaurant', 'Hostel', 'Cocktail Bar',
'Jazz Club', 'Camera Store', 'Tailor Shop', 'Fountain',
'German Restaurant', 'Concert Hall', 'Museum', 'Bistro',
'Basketball Stadium', 'Beach', 'Eastern European Restaurant',
'Irish Pub', 'Nightclub', 'Art Museum', 'Salad Place',
'Donut Shop', 'Opera House', 'Monument / Landmark',
'General Travel', 'Colombian Restaurant', 'Record Shop',
'Brazilian Restaurant', 'Gluten-free Restaurant', 'Roof Deck',
'Aquarium', 'Train Station', 'History Museum', 'Scenic Lookout',
'Baseball Stadium', 'Hotel Bar', 'Soup Place', 'Garden',
'Jewish Restaurant', 'College Gym', 'College Arts Building',
'Flower Shop', 'Organic Grocery', 'Dumpling Restaurant',
'Gaming Cafe', 'Martial Arts Dojo', 'Doner Restaurant',
'Filipino Restaurant', 'Hotpot Restaurant', 'Hospital',
'Airport Lounge', 'Harbor / Marina', 'Airport Food Court',
'Airport Terminal', 'Airport Gate', 'Plane', 'Airport Service',
'Boat or Ferry', 'Molecular Gastronomy Restaurant', 'Optical Shop',
'Accessories Store', 'Fraternity House', 'Field', 'Baby Store',
'Cuban Restaurant', 'Mac & Cheese Joint', 'Malay Restaurant',
'Cupcake Shop', 'Tapas Restaurant',
'Southern / Soul Food Restaurant', 'Climbing Gym', 'Stadium',
'Construction & Landscaping', 'Flea Market',
'Cajun / Creole Restaurant', 'Food', 'Indie Movie Theater',
'College Auditorium', 'College Cafeteria', 'Skate Park',
'Garden Center', 'Auto Workshop', 'Recording Studio', 'River',
'Locksmith', 'Hardware Store', 'Shopping Plaza',
'Empanada Restaurant', 'Drugstore'], dtype=object)
```

2.3.5 Create a one hot encoding matrix, with venue categories as columns and venues from the venues in the venue dataframe as rows

```
In [12]: # one hot encoding
         toronto_onehot = pd.get_dummies(toronto_venues[['Venue Category']], prefix="", prefix_
         print(toronto_onehot.shape)
         toronto_onehot.head()
```

```
(2244, 280)
```

Out[12]:

	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop	...	Train Station	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Warehouse Store	Wine Bar	Wings Joint	Women's Store	Yoga Studio
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

[5 rows x 280 columns]

2.3.6 Remove 'Neighborhood' column in the onehot dataset

For some reason there seems to be a 'Neighborhood' column in this dataset, where we would only expect venue categories

In [13]:
```
#toronto_onehot['Neighborhood']
```

In [14]:
```
toronto_onehot.drop(columns=['Neighborhood'],inplace=True)
```

In [15]:
```
#toronto_onehot['Neighborhood']
```

In [16]:
```
### Add Neighboorhood column to the onehot dataset as a first column
```

In [17]:
```
# add neighborhood column back to dataframe
toronto_onehot['Neighborhood'] = toronto_venues['Neighborhood']
fixed_columns = [toronto_onehot.columns[-1]] + list(toronto_onehot.columns[:-1])
toronto_onehot = toronto_onehot[fixed_columns]
toronto_onehot.head()
```

Out[17]:

	Neighborhood	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	...	Train Station	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Warehouse Store
0	Rouge, Malvern	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
1	Highland Creek, Rouge Hill, Port Union	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
2	Guildwood, Morningside, West Hill	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
3	Guildwood, Morningside, West Hill	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
4	Guildwood, Morningside, West Hill	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0

```
   Wine Bar  Wings Joint  Women's Store  Yoga Studio
0         0            0              0            0
1         0            0              0            0
2         0            0              0            0
3         0            0              0            0
4         0            0              0            0

[5 rows x 280 columns]
```

In [18]:
```
### For each Neighborhood, calculate the average % for each venue category (% of tota
```

In [19]:
```
toronto_grouped = toronto_onehot.groupby('Neighborhood').mean()
# Neighborhood from index => column
toronto_grouped.reset_index(inplace=True)
toronto_grouped.head()
```

Out[19]:
```
                                     Neighborhood  Accessories Store  \
0                          Adelaide, King, Richmond                0.0
1                                       Agincourt                0.0
2  Agincourt North, L'Amoreaux East, Milliken, St...                0.0
3  Albion Gardens, Beaumond Heights, Humbergate, ...                0.0
4                           Alderwood, Long Branch                0.0

   Afghan Restaurant  Airport  Airport Food Court  Airport Gate  \
0                0.0      0.0                 0.0           0.0
1                0.0      0.0                 0.0           0.0
2                0.0      0.0                 0.0           0.0
3                0.0      0.0                 0.0           0.0
4                0.0      0.0                 0.0           0.0

   Airport Lounge  Airport Service  Airport Terminal  American Restaurant  \
0             0.0             0.0               0.0                 0.03
1             0.0             0.0               0.0                 0.00
2             0.0             0.0               0.0                 0.00
3             0.0             0.0               0.0                 0.00
4             0.0             0.0               0.0                 0.00

      ...       Train Station  Vegetarian / Vegan Restaurant  \
0     ...                 0.0                           0.01
1     ...                 0.0                           0.00
2     ...                 0.0                           0.00
3     ...                 0.0                           0.00
4     ...                 0.0                           0.00

   Video Game Store  Video Store  Vietnamese Restaurant  Warehouse Store  \
0               0.0          0.0                    0.0              0.0
1               0.0          0.0                    0.0              0.0
2               0.0          0.0                    0.0              0.0
```

```
3               0.0          0.0                    0.0              0.0
4               0.0          0.0                    0.0              0.0

   Wine Bar  Wings Joint  Women's Store  Yoga Studio
0      0.01          0.0           0.01          0.0
1      0.00          0.0           0.00          0.0
2      0.00          0.0           0.00          0.0
3      0.00          0.0           0.00          0.0
4      0.00          0.0           0.00          0.0

[5 rows x 280 columns]
```

2.3.7 For each neighborhood, show the top 5 venue categories

```
In [20]: num_top_venues = 5

         for hood in toronto_grouped['Neighborhood']:
             print("----"+hood+"----")
             temp = toronto_grouped[toronto_grouped['Neighborhood'] == hood].T.reset_index()
             temp.columns = ['venue','freq']
             temp = temp.iloc[1:]
             temp['freq'] = temp['freq'].astype(float)
             temp = temp.round({'freq': 2})
             print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_to
             print('\n')
```

```
----Adelaide, King, Richmond----
             venue  freq
0      Coffee Shop  0.08
1             Café  0.05
2              Bar  0.04
3       Steakhouse  0.04
4  Thai Restaurant  0.04


----Agincourt----
                venue  freq
0      Breakfast Spot   0.2
1  Chinese Restaurant   0.2
2              Lounge   0.2
3        Skating Rink   0.2
4      Sandwich Place   0.2


----Agincourt North, L'Amoreaux East, Milliken, Steeles East----
                 venue  freq
0           Playground   0.5
1                 Park   0.5
```

```
2    Accessories Store   0.0
3        Metro Station   0.0
4  Monument / Landmark   0.0


----Albion Gardens, Beaumond Heights, Humbergate, Jamestown, Mount Olive, Silverstone, South St
                  venue  freq
0         Grocery Store  0.22
1           Pizza Place  0.11
2           Coffee Shop  0.11
3        Sandwich Place  0.11
4  Fast Food Restaurant  0.11


----Alderwood, Long Branch----
            venue  freq
0     Pizza Place  0.22
1    Skating Rink  0.11
2  Sandwich Place  0.11
3             Pub  0.11
4     Coffee Shop  0.11


----Bathurst Manor, Downsview North, Wilson Heights----
                  venue  freq
0           Coffee Shop  0.09
1             Gift Shop  0.05
2         Shopping Mall  0.05
3  Fast Food Restaurant  0.05
4        Sandwich Place  0.05


----Bayview Village----
                        venue  freq
0         Japanese Restaurant  0.25
1          Chinese Restaurant  0.25
2                        Bank  0.25
3                        Café  0.25
4  Modern European Restaurant  0.00


----Bedford Park, Lawrence Manor East----
                 venue  freq
0          Coffee Shop  0.09
1   Italian Restaurant  0.09
2  Japanese Restaurant  0.05
3     Greek Restaurant  0.05
4            Juice Bar  0.05
```

```
----Berczy Park----
                venue  freq
0         Coffee Shop  0.09
1        Cocktail Bar  0.05
2            Beer Bar  0.04
3  Seafood Restaurant  0.04
4      Farmers Market  0.04


----Birch Cliff, Cliffside West----
                   venue  freq
0           Skating Rink   0.2
1        College Stadium   0.2
2                   Farm   0.2
3  General Entertainment   0.2
4                   Café   0.2


----Bloordale Gardens, Eringate, Markland Wood, Old Burnhamthorpe----
            venue  freq
0     Pizza Place  0.11
1  Shopping Plaza  0.11
2           Café  0.11
3     Coffee Shop  0.11
4    Liquor Store  0.11


----Brockton, Exhibition Place, Parkdale Village----
                  venue  freq
0           Coffee Shop  0.10
1                 Café  0.10
2        Breakfast Spot  0.10
3          Intersection  0.05
4  Caribbean Restaurant  0.05


----Business Reply Mail Processing Centre 969 Eastern----
                venue  freq
0  Light Rail Station  0.11
1        Yoga Studio  0.05
2       Auto Workshop  0.05
3        Pizza Place  0.05
4         Comic Shop  0.05


----CFB Toronto, Downsview East----
```

```
                venue  freq
0             Airport   0.5
1                Park   0.5
2   Accessories Store   0.0
3       Metro Station   0.0
4  Monument / Landmark   0.0


----CN Tower, Bathurst Quay, Island airport, Harbourfront West, King and Spadina, Railway Lands
             venue  freq
0  Airport Terminal  0.14
1    Airport Lounge  0.14
2   Airport Service  0.14
3             Plane  0.07
4       Coffee Shop  0.07


----Cabbagetown, St. James Town----
               venue  freq
0        Coffee Shop  0.07
1               Park  0.07
2  Convenience Store  0.04
3                Pub  0.04
4             Bakery  0.04


----Caledonia-Fairbanks----
                  venue  freq
0                  Park   0.4
1         Women's Store   0.2
2                Market   0.2
3  Fast Food Restaurant   0.2
4     Accessories Store   0.0


----Canada Post Gateway Processing Centre----
                      venue  freq
0                     Hotel  0.18
1               Coffee Shop  0.18
2  Mediterranean Restaurant  0.09
3             Burrito Place  0.09
4      Gym / Fitness Center  0.09


----Cedarbrae----
                  venue  freq
0                  Bank  0.14
1       Thai Restaurant  0.14
```

```
2      Hakka Restaurant  0.14
3  Caribbean Restaurant  0.14
4    Athletics & Sports  0.14


----Central Bay Street----
                venue  freq
0         Coffee Shop  0.14
1      Ice Cream Shop  0.05
2  Italian Restaurant  0.05
3      Sandwich Place  0.04
4        Burger Joint  0.04


----Chinatown, Grange Park, Kensington Market----
                           venue  freq
0                           Café  0.07
1  Vegetarian / Vegan Restaurant  0.06
2             Chinese Restaurant  0.05
3             Mexican Restaurant  0.04
4            Dumpling Restaurant  0.04


----Christie----
           venue  freq
0  Grocery Store  0.20
1           Café  0.20
2           Park  0.13
3     Baby Store  0.07
4    Coffee Shop  0.07


----Church and Wellesley----
                 venue  freq
0         Coffee Shop  0.07
1  Japanese Restaurant  0.06
2    Sushi Restaurant  0.05
3             Gay Bar  0.04
4          Restaurant  0.04


----Clairlea, Golden Mile, Oakridge----
          venue  freq
0      Bus Line  0.22
1        Bakery  0.22
2  Soccer Field  0.11
3  Intersection  0.11
4   Bus Station  0.11
```

```
----Clarks Corners, Sullivan, Tam O'Shanter----
                  venue  freq
0           Pizza Place   0.2
1              Pharmacy   0.1
2          Noodle House   0.1
3  Fast Food Restaurant   0.1
4   Fried Chicken Joint   0.1


----Cliffcrest, Cliffside, Scarborough Village West----
                  venue  freq
0        Movie Theater  0.33
1                Motel  0.33
2  American Restaurant  0.33
3   Mexican Restaurant  0.00
4  Monument / Landmark  0.00


----Cloverdale, Islington, Martin Grove, Princess Gardens, West Deane Park----
                             venue  freq
0                             Bank   1.0
1                Accessories Store   0.0
2               Mexican Restaurant   0.0
3              Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----Commerce Court, Victoria Hotel----
                  venue  freq
0          Coffee Shop  0.09
1                Hotel  0.06
2                 Café  0.06
3           Restaurant  0.04
4  American Restaurant  0.04


----Davisville----
             venue  freq
0     Pizza Place  0.09
1    Dessert Shop  0.09
2  Sandwich Place  0.09
3      Restaurant  0.06
4            Café  0.06


----Davisville North----
```

```
              venue  freq
0           Dog Run  0.11
1               Gym  0.11
2  Food & Drink Shop  0.11
3    Sandwich Place  0.11
4    Clothing Store  0.11


----Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West----
              venue  freq
0       Coffee Shop  0.13
1               Pub  0.13
2        Restaurant  0.07
3      Liquor Store  0.07
4  Sushi Restaurant  0.07


----Del Ray, Keelesdale, Mount Dennis, Silverthorn----
                  venue  freq
0  Fast Food Restaurant  0.33
1            Restaurant  0.17
2        Sandwich Place  0.17
3   Fried Chicken Joint  0.17
4           Coffee Shop  0.17


----Design Exchange, Toronto Dominion Centre----
         venue  freq
0  Coffee Shop  0.12
1         Café  0.08
2        Hotel  0.06
3   Restaurant  0.05
4       Bakery  0.03


----Don Mills North----
                  venue  freq
0   Japanese Restaurant   0.2
1  Caribbean Restaurant   0.2
2  Gym / Fitness Center   0.2
3                 Café   0.2
4     Basketball Court   0.2


----Dorset Park, Scarborough Town Centre, Wexford Heights----
                      venue  freq
0          Indian Restaurant  0.33
1                  Pet Store  0.17
```

```
2      Vietnamese Restaurant  0.17
3  Latin American Restaurant  0.17
4         Chinese Restaurant  0.17


----Dovercourt Village, Dufferin----
                       venue  freq
0                   Pharmacy  0.12
1                     Bakery  0.12
2                Supermarket  0.12
3  Middle Eastern Restaurant  0.06
4                       Bank  0.06


----Downsview Central----
                  venue  freq
0           Food Truck  0.33
1         Home Service  0.33
2       Baseball Field  0.33
3   Mexican Restaurant  0.00
4  Monument / Landmark  0.00


----Downsview Northwest----
                   venue  freq
0          Liquor Store  0.25
1  Gym / Fitness Center  0.25
2         Grocery Store  0.25
3    Athletics & Sports  0.25
4     Accessories Store  0.00


----Downsview West----
                venue  freq
0      Grocery Store   0.4
1               Bank   0.2
2               Park   0.2
3      Shopping Mall   0.2
4  Accessories Store   0.0


----Downsview, North Park, Upwood Park----
                         venue  freq
0  Construction & Landscaping  0.25
1                      Bakery  0.25
2                        Park  0.25
3           Basketball Court  0.25
4          Accessories Store  0.00
```

```
----East Birchmount Park, Ionview, Kennedy Park----
                venue  freq
0      Discount Store  0.33
1    Department Store  0.17
2         Coffee Shop  0.17
3  Chinese Restaurant  0.17
4         Bus Station  0.17


----East Toronto----
               venue  freq
0               Park   0.4
1        Pizza Place   0.2
2        Coffee Shop   0.2
3  Convenience Store   0.2
4  Martial Arts Dojo   0.0


----Emery, Humberlea----
                             venue  freq
0                  Baseball Field   1.0
1               Accessories Store   0.0
2              Mexican Restaurant   0.0
3             Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----Fairview, Henry Farm, Oriole----
                  venue  freq
0        Clothing Store  0.14
1           Coffee Shop  0.08
2  Fast Food Restaurant  0.08
3   Japanese Restaurant  0.03
4            Shoe Store  0.03


----First Canadian Place, Underground city----
         venue  freq
0  Coffee Shop  0.09
1         Café  0.07
2   Steakhouse  0.04
3        Hotel  0.04
4   Restaurant  0.04


----Flemingdon Park, Don Mills South----
```

```
                venue  freq
0         Coffee Shop  0.10
1          Beer Store  0.10
2                 Gym  0.10
3    Asian Restaurant  0.05
4  Chinese Restaurant  0.05


----Forest Hill North, Forest Hill West----
               venue  freq
0               Park  0.25
1   Sushi Restaurant  0.25
2      Jewelry Store  0.25
3              Trail  0.25
4  Accessories Store  0.00


----Glencairn----
                 venue  freq
0   Italian Restaurant  0.25
1                  Pub  0.25
2               Bakery  0.25
3  Japanese Restaurant  0.25
4               Motel  0.00


----Guildwood, Morningside, West Hill----
                 venue  freq
0          Pizza Place  0.14
1  Rental Car Location  0.14
2       Medical Center  0.14
3   Mexican Restaurant  0.14
4    Electronics Store  0.14


----Harbord, University of Toronto----
                 venue  freq
0                Café  0.11
1               Bakery  0.06
2  Japanese Restaurant  0.06
3            Bookstore  0.06
4           Restaurant  0.06


----Harbourfront East, Toronto Islands, Union Station----
                venue  freq
0         Coffee Shop  0.11
1               Hotel  0.05
```

```
2            Aquarium  0.05
3                Café  0.04
4  Italian Restaurant  0.04


----Harbourfront, Regent Park----
         venue  freq
0  Coffee Shop  0.16
1         Café  0.06
2       Bakery  0.06
3         Park  0.06
4          Pub  0.06


----High Park, The Junction South----
                       venue  freq
0                       Café  0.09
1         Mexican Restaurant  0.09
2               Music Venue  0.04
3             Grocery Store  0.04
4  Cajun / Creole Restaurant  0.04


----Highland Creek, Rouge Hill, Port Union----
                             venue  freq
0                              Bar   1.0
1               Accessories Store   0.0
2              Mexican Restaurant   0.0
3             Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----Hillcrest Village----
                      venue  freq
0     Fast Food Restaurant  0.17
1  Mediterranean Restaurant  0.17
2       Athletics & Sports  0.17
3                     Pool  0.17
4                  Dog Run  0.17


----Humber Bay Shores, Mimico South, New Toronto----
                  venue  freq
0            Pet Store  0.07
1          Coffee Shop  0.07
2  Fast Food Restaurant  0.07
3   Seafood Restaurant  0.07
4         Liquor Store  0.07
```

```
----Humber Bay, King's Mill Park, Kingsway Park South East, Mimico NE, Old Mill South, The Quee
                        venue  freq
0  Construction & Landscaping  0.25
1                   Locksmith  0.25
2              Baseball Field  0.25
3                        Pool  0.25
4   Middle Eastern Restaurant  0.00


----Humber Summit----
                            venue  freq
0                     Pizza Place   0.5
1              Empanada Restaurant   0.5
2              Monument / Landmark   0.0
3  Molecular Gastronomy Restaurant   0.0
4       Modern European Restaurant   0.0


----Humewood-Cedarvale----
          venue  freq
0       Dog Run  0.25
1         Trail  0.25
2         Field  0.25
3  Hockey Arena  0.25
4         Motel  0.00


----Kingsview Village, Martin Grove Gardens, Richview Gardens, St. Phillips----
                            venue  freq
0                        Bus Line   0.5
1                            Park   0.5
2                Accessories Store   0.0
3              Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----Kingsway Park South West, Mimico NW, The Queensway West, Royal York South West, South of Bl
                  venue  freq
0        Hardware Store  0.07
1         Grocery Store  0.07
2           Flower Shop  0.07
3  Fast Food Restaurant  0.07
4        Sandwich Place  0.07


----L'Amoreaux West----
```

```
                   venue  freq
0      Chinese Restaurant  0.17
1    Fast Food Restaurant  0.17
2          Breakfast Spot  0.08
3  Thrift / Vintage Store  0.08
4             Pizza Place  0.08


----Lawrence Heights, Lawrence Manor----
                   venue  freq
0          Clothing Store  0.21
1       Accessories Store  0.07
2             Shoe Store  0.07
3  Furniture / Home Store  0.07
4             Coffee Shop  0.07


----Lawrence Park----
                venue  freq
0           Bus Line  0.25
1               Park  0.25
2        Swim School  0.25
3  Dim Sum Restaurant  0.25
4  Mexican Restaurant  0.00


----Leaside----
                   venue  freq
0            Coffee Shop  0.12
1    Sporting Goods Shop  0.09
2       Sushi Restaurant  0.06
3  Furniture / Home Store  0.06
4           Burger Joint  0.06


----Little Portugal, Trinity----
                    venue  freq
0                     Bar  0.12
1             Coffee Shop  0.06
2        Asian Restaurant  0.05
3  New American Restaurant  0.03
4       French Restaurant  0.03


----Maryvale, Wexford----
                       venue  freq
0            Breakfast Spot  0.17
1             Shopping Mall  0.17
```

```
2                Auto Garage  0.17
3             Sandwich Place  0.17
4  Middle Eastern Restaurant  0.17


----Moore Park, Summerhill East----
               venue  freq
0               Park  0.25
1         Playground  0.25
2                Gym  0.25
3         Restaurant  0.25
4  Accessories Store  0.00


----Newtonbrook, Willowdale----
                             venue  freq
0                             Park   1.0
1               Accessories Store   0.0
2                   Metro Station   0.0
3             Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----North Toronto West----
                 venue  freq
0  Sporting Goods Shop  0.10
1       Clothing Store  0.10
2          Coffee Shop  0.10
3          Yoga Studio  0.05
4            Gift Shop  0.05


----Northwest----
                 venue  freq
0            Drugstore  0.33
1  Rental Car Location  0.33
2                  Bar  0.33
3    Accessories Store  0.00
4        Movie Theater  0.00


----Northwood Park, York University----
                venue  freq
0      Massage Studio  0.17
1         Coffee Shop  0.17
2                 Bar  0.17
3       Metro Station  0.17
4  Falafel Restaurant  0.17
```

```
----Parkdale, Roncesvalles----
                        venue  freq
0               Breakfast Spot  0.13
1                    Gift Shop  0.13
2  Eastern European Restaurant  0.07
3                         Bank  0.07
4                Movie Theater  0.07


----Parkwoods----
                   venue  freq
0  Fast Food Restaurant  0.33
1                  Park  0.33
2     Food & Drink Shop  0.33
3     Accessories Store  0.00
4    Mexican Restaurant  0.00


----Queen's Park----
         venue  freq
0  Coffee Shop  0.22
1         Park  0.05
2        Diner  0.05
3          Gym  0.05
4  Yoga Studio  0.02


----Rosedale----
               venue  freq
0               Park   0.4
1             Trail   0.2
2         Playground   0.2
3           Building   0.2
4  Accessories Store   0.0


----Roselawn----
                 venue  freq
0               Garden   1.0
1    Accessories Store   0.0
2   Mexican Restaurant   0.0
3                Motel   0.0
4  Monument / Landmark   0.0


----Rouge, Malvern----
```

```
                           venue  freq
0             Fast Food Restaurant   1.0
1                Accessories Store   0.0
2                    Metro Station   0.0
3              Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----Runnymede, Swansea----
              venue  freq
0       Coffee Shop  0.08
1              Café  0.08
2       Pizza Place  0.06
3             Diner  0.06
4  Sushi Restaurant  0.06


----Ryerson, Garden District----
                      venue  freq
0                Coffee Shop  0.10
1             Clothing Store  0.06
2             Cosmetics Shop  0.04
3                      Café  0.03
4  Middle Eastern Restaurant  0.03


----Scarborough Village----
               venue  freq
0  Convenience Store  0.33
1                Spa  0.33
2         Playground  0.33
3  Accessories Store  0.00
4      Metro Station  0.00


----Silver Hills, York Mills----
                           venue  freq
0                      Cafeteria   1.0
1                Accessories Store   0.0
2               Mexican Restaurant   0.0
3              Monument / Landmark   0.0
4  Molecular Gastronomy Restaurant   0.0


----St. James Town----
                venue  freq
0               Hotel  0.05
1          Restaurant  0.05
```

```
2         Coffee Shop  0.05
3                Café  0.05
4  Italian Restaurant  0.05


----Stn A PO Boxes 25 The Esplanade----
                venue  freq
0         Coffee Shop  0.11
1                Café  0.04
2          Restaurant  0.04
3  Italian Restaurant  0.03
4        Cocktail Bar  0.03


----Studio District----
                  venue  freq
0                 Café  0.10
1          Coffee Shop  0.08
2  American Restaurant  0.05
3              Bakery  0.05
4   Italian Restaurant  0.05


----The Annex, North Midtown, Yorkville----
               venue  freq
0        Coffee Shop  0.12
1               Café  0.12
2     Sandwich Place  0.12
3        Pizza Place  0.08
4  Indian Restaurant  0.04


----The Beaches----
                  venue  freq
0    Health Food Store   0.2
1          Coffee Shop   0.2
2                  Pub   0.2
3                Trail   0.2
4  Monument / Landmark   0.0


----The Beaches West, India Bazaar----
               venue  freq
0        Pizza Place  0.11
1          Pet Store  0.06
2                Pub  0.06
3  Fish & Chips Shop  0.06
4           Brewery  0.06
```

```
----The Danforth West, Riverdale----
                   venue  freq
0        Greek Restaurant  0.21
1             Coffee Shop  0.10
2      Italian Restaurant  0.07
3  Furniture / Home Store  0.05
4          Ice Cream Shop  0.05


----The Junction North, Runnymede----
                  venue  freq
0     Convenience Store   0.2
1              Bus Line   0.2
2           Pizza Place   0.2
3  Caribbean Restaurant   0.2
4              Brewery   0.2


----The Kingsway, Montgomery Road, Old Mill North----
                 venue  freq
0                River   0.5
1                 Park   0.5
2   Accessories Store   0.0
3   Mexican Restaurant   0.0
4  Monument / Landmark   0.0


----Thorncliffe Park----
               venue  freq
0       Burger Joint  0.12
1  Indian Restaurant  0.12
2      Grocery Store  0.06
3     Sandwich Place  0.06
4        Supermarket  0.06


----Victoria Village----
                        venue  freq
0  Financial or Legal Service  0.25
1                Hockey Arena  0.25
2                 Coffee Shop  0.25
3       Portuguese Restaurant  0.25
4              Massage Studio  0.00


----Westmount----
```

```
                      venue  freq
0                Pizza Place  0.25
1                Coffee Shop  0.12
2  Middle Eastern Restaurant  0.12
3             Sandwich Place  0.12
4               Intersection  0.12


----Weston----
                 venue  freq
0    Convenience Store   0.5
1                 Park   0.5
2    Accessories Store   0.0
3        Metro Station   0.0
4  Monument / Landmark   0.0


----Willowdale South----
              venue  freq
0  Ramen Restaurant  0.09
1       Coffee Shop  0.09
2              Café  0.06
3  Sushi Restaurant  0.06
4       Pizza Place  0.06


----Willowdale West----
            venue  freq
0     Pizza Place   0.2
1     Coffee Shop   0.2
2  Discount Store   0.2
3         Butcher   0.2
4        Pharmacy   0.2


----Woburn----
                 venue  freq
0          Coffee Shop  0.67
1    Korean Restaurant  0.33
2               Museum  0.00
3                Motel  0.00
4  Monument / Landmark  0.00


----Woodbine Gardens, Parkview Hill----
                  venue  freq
0           Pizza Place  0.18
1  Fast Food Restaurant  0.18
```

```
2    Athletics & Sports  0.09
3               Pharmacy  0.09
4                   Bank  0.09


----Woodbine Heights----
            venue  freq
0    Skating Rink  0.22
1        Pharmacy  0.11
2        Bus Stop  0.11
3  Cosmetics Shop  0.11
4            Park  0.11


----York Mills West----
                venue  freq
0   Convenience Store  0.33
1                Bank  0.33
2                Park  0.33
3   Accessories Store  0.00
4  Mexican Restaurant  0.00
```

2.3.8 Create a dataframe containing the 10 most common venue categories for each neigborhood

```
In [21]: def return_most_common_venues(row, num_top_venues):
             row_categories = row.iloc[1:]
             row_categories_sorted = row_categories.sort_values(ascending=False)

             return row_categories_sorted.index.values[0:num_top_venues]
```

```
In [22]: import numpy as np
         num_top_venues = 10

         indicators = ['st', 'nd', 'rd']

         # create columns according to number of top venues
         columns = ['Neighborhood']
         for ind in np.arange(num_top_venues):
             try:
                 columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
             except:
                 columns.append('{}th Most Common Venue'.format(ind+1))

         # create a new dataframe
```

```
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = toronto_grouped['Neighborhood']

for ind in np.arange(toronto_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(toronto_grou

print(neighborhoods_venues_sorted.shape)
neighborhoods_venues_sorted.head()
```

```
(101, 11)
```

Out[22]:

	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	Adelaide, King, Richmond	Coffee Shop	Café	Steakhouse	Bar	Thai Restaurant	Gym	Breakfast Spot	Hotel	Asian Restaurant	Restaurant
1	Agincourt	Lounge	Breakfast Spot	Skating Rink	Chinese Restaurant	Sandwich Place	Eastern European Restaurant	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant
2	Agincourt North, L'Amoreaux East, Milliken, St...	Park	Playground	Yoga Studio	Eastern European Restaurant	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant
3	Albion Gardens, Beaumond Heights, Humbergate, ...	Grocery Store	Fast Food Restaurant	Pizza Place	Sandwich Place	Coffee Shop	Beer Store	Pharmacy	Fried Chicken Joint	Empanada Restaurant	Electronics Store
4	Alderwood, Long Branch	Pizza Place	Pool	Skating Rink	Gym	Pharmacy	Coffee Shop	Pub	Sandwich Place	Yoga Studio	Diner

2.3.9 Create neighborhood clusters based on the top 10 venue categories

```
In [23]: from sklearn.cluster import KMeans

kclusters = 5

toronto_grouped_clustering = toronto_grouped.drop('Neighborhood', 1)
```

```
        # run k-means clustering
        kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(toronto_grouped_clustering)

        print(len(kmeans.labels_))
        print(kmeans.labels_)
        # check cluster labels generated for each row in the dataframe
        #kmeans.labels_[0:10]
```

```
101
[1 1 2 1 1 1 1 1 1 1 1 1 2 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 1 1 1 1 2 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 2 1 1 1 2 2 1 1 1 1 2
 1 2 4 3 1 1 1 0 1 1 1 1 1 1 1 1 2 1 1 1 2 1 1 1 1 1 2]
```

2.3.10 Add the cluster labels to the dataset with top 10 venue categories per neighborhood and merge that dataset with the location dataset for the toronto neighborhoods

In [24]:
```
        # add clustering labels
        neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

        # merge toronto_grouped with toronto_data to add latitude/longitude for each neighbori
        toronto_merged = df_postal_loc.join(neighborhoods_venues_sorted.set_index('Neighborhoo

        print(toronto_merged.shape)
        toronto_merged.head()
```

```
(103, 16)
```

Out[24]:

	PostalCode	Borough	Neighbourhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353	3.0	Fast Food Restaurant	Eastern European Restaurant	Dive Bar
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497	1.0	Bar	Yoga Studio	Eastern European Restaurant
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711	1.0	Breakfast Spot	Electronics Store	Rental Car Location
3	M1G	Scarborough	Woburn	43.770992	-79.216917	1.0	Coffee Shop	Korean Restaurant	Yoga Studio
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	1.0	Hakka Restaurant	Thai Restaurant	Fried Chicken Joint

```
  4th Most Common Venue 5th Most Common Venue 6th Most Common Venue  \
0              Dog Run      Doner Restaurant            Donut Shop
1              Dog Run      Doner Restaurant            Donut Shop
2         Intersection    Mexican Restaurant        Medical Center
3    Electronics Store               Dog Run      Doner Restaurant
4                 Bank                Bakery    Athletics & Sports

  7th Most Common Venue 8th Most Common Venue 9th Most Common Venue  \
0             Drugstore   Dumpling Restaurant     Electronics Store
1             Drugstore   Dumpling Restaurant     Electronics Store
2           Pizza Place   Empanada Restaurant  Ethiopian Restaurant
3             Donut Shop             Drugstore   Dumpling Restaurant
4  Caribbean Restaurant      Cuban Restaurant          Costume Shop

       10th Most Common Venue
0                       Field
1               Discount Store
2  Eastern European Restaurant
3  Eastern European Restaurant
4               Farmers Market
```

2.3.11 Remove rows where the cluster label could not be calculated

```
In [25]: print(df_postal_loc.shape)
         print(neighborhoods_venues_sorted.shape)
         df1 = toronto_merged[toronto_merged.isna().any(axis=1)]
         df1
```

```
(103, 5)
(101, 12)
```

```
Out[25]:    PostalCode      Borough     Neighbourhood   Latitude  Longitude  \
         16        M1X  Scarborough       Upper Rouge  43.836125 -79.205636
         93        M9A    Etobicoke  Islington Avenue  43.667856 -79.532242

             Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
         16             NaN                  NaN                  NaN
         93             NaN                  NaN                  NaN

            3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
         16                   NaN                  NaN                  NaN
         93                   NaN                  NaN                  NaN

            6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
         16                   NaN                  NaN                  NaN
         93                   NaN                  NaN                  NaN
```

```
   9th Most Common Venue 10th Most Common Venue
16                   NaN                    NaN
93                   NaN                    NaN
```

In [26]:
```
toronto_merged.dropna(inplace=True)
print(toronto_merged.shape)
```

```
(101, 16)
```

In [27]:
```
df1 = toronto_merged[toronto_merged.isna().any(axis=1)]
df1
```

Out[27]:
```
Empty DataFrame
Columns: [PostalCode, Borough, Neighbourhood, Latitude, Longitude, Cluster Labels, 1st
Index: []
```

2.3.12 Generate a map of Toronto with the neighborhoods and how they cluster together (colour indicated)

In [31]:
```
import matplotlib.cm as cm
import matplotlib.colors as colors

# create map
map_clusters = folium.Map(location=[latitude + 0.05, 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(toronto_merged['Latitude'], toronto_merged['Longitud
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    cluster = int(cluster)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

Out[31]:
```
<folium.folium.Map at 0x20c46e68978>
```

2.4 Conclusion

I currently live in the Humber Summit neighborhood and I want to move to the Guildwood, Morningside, West Hill neighborhood.

The map shows that both neighborhoods are in the same cluster and are thus pretty similar.

Below we also see that both are neighborhoods with plenty of restaurants.

```
In [29]: f = (toronto_merged['PostalCode'] == 'M9L') | (toronto_merged['PostalCode'] == 'M1E')
         toronto_merged[f]
```

```
Out[29]:    PostalCode      Borough                       Neighbourhood   Latitude  \
         2         M1E  Scarborough  Guildwood, Morningside, West Hill  43.763573
         96        M9L   North York                       Humber Summit  43.756303

             Longitude  Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
         2  -79.188711             1.0        Breakfast Spot     Electronics Store
         96 -79.565963             1.0           Pizza Place   Empanada Restaurant

            3rd Most Common Venue        4th Most Common Venue 5th Most Common Venue  \
         2    Rental Car Location                Intersection    Mexican Restaurant
         96           Yoga Studio  Eastern European Restaurant              Dive Bar

            6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
         2         Medical Center           Pizza Place   Empanada Restaurant
         96               Dog Run      Doner Restaurant            Donut Shop

            9th Most Common Venue       10th Most Common Venue
         2   Ethiopian Restaurant  Eastern European Restaurant
         96             Drugstore          Dumpling Restaurant
```