Battle-of-Toronto-Neighborhoods (/github/Adityakolluru7/Battle-of-Toronto-Neighborhoods/tree/master)
/
Toronto neighbourhood.ipynb (/github/Adityakolluru7/Battle-of-Toronto-Neighborhoods/tree/master/Toronto neighbourhood.ipynb)

In [2]:
```
#Importing required modules
import pandas as pd
import numpy as np
import lxml
import requests
!pip install folium
import folium
import json
from sklearn.cluster import KMeans
from pandas.io.json import json_normalize
import matplotlib.cm as cm
import matplotlib.colors as colors
```

```
Collecting folium
  Downloading https://files.pythonhosted.org/packages/a4/f0/44e69d50519880287cc41e7c8a
     |████████████████████████████████| 102kB 7.5MB/s ta 0:00:011
Requirement already satisfied: numpy in /opt/conda/envs/Python36/lib/python3.6/site-pa
Collecting branca>=0.3.0 (from folium)
  Downloading https://files.pythonhosted.org/packages/13/fb/9eacc24ba3216510c6b59a4ea1
Requirement already satisfied: requests in /opt/conda/envs/Python36/lib/python3.6/site
Requirement already satisfied: jinja2>=2.9 in /opt/conda/envs/Python36/lib/python3.6/s
Requirement already satisfied: idna<2.9,>=2.5 in /opt/conda/envs/Python36/lib/python3.
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/envs/Python36/lib/pyth
Requirement already satisfied: urllib3<1.25,>=1.21.1 in /opt/conda/envs/Python36/lib/p
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in /opt/conda/envs/Python36/lib/p
Requirement already satisfied: MarkupSafe>=0.23 in /opt/conda/envs/Python36/lib/python
Installing collected packages: branca, folium
Successfully installed branca-0.4.1 folium-0.11.0
```

In [4]:
```
#url of the html file
url = "https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"
```

In [5]:
```
#Reading just the first table from the html
tables =pd.read_html(url)
df=tables[0]
```

In [6]:
```
#Droping all the values where the Borough is Not assigned
df_tor = df[~df['Borough'].isin(['Not assigned'])]
df_tor.head()
```

Out[6]:

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

In [7]:
```
#Reading latitude and longitude
lat = pd.read_csv("https://cocl.us/Geospatial_data")
```

In [8]:
```
#Creating a DataFrame with the latitude and longitude
column_names = ['Postal Code', 'Latitude', 'Longitude']
df_lat = pd.DataFrame(data=lat,columns=column_names)
```

In [9]:
```
#merging both dataframes
df_tor = pd.merge(df_tor,df_lat, on ="Postal Code")
```

In [10]:
```
#DataFrame with all the information
df_tor.head()
```

Out[10]:

	Postal Code	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	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

In [67]:
```
#Adding Toronto Coordinates
address = 'Toronto, CA'

geolocator = Nominatim(user_agent="ca_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
```

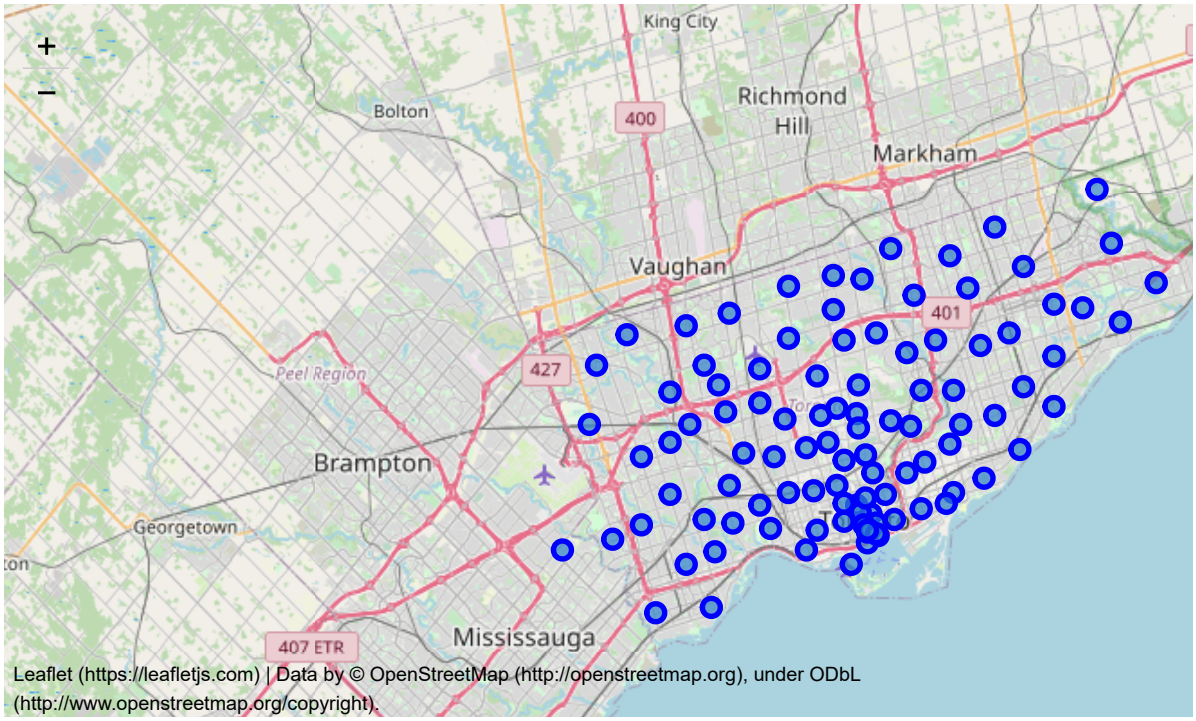
In [11]:

```
latitude = 43.6532
longitude = -79.3832
# create map of Toronto using latitude and longitude values
map_toronto = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, neighborhood in zip(df_tor['Latitude'], df_tor['Longitude'], df_
    label = '{}, {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_toronto)

map_toronto
```

Out[11]:



In [12]:
```
#Analyzing the Boroughs in Downtown Toronto
tor_data = df_tor[df_tor['Borough'] == 'Downtown Toronto'].reset_index(drop=True)
tor_data
```

Out[12]:

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
5	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
6	M6G	Downtown Toronto	Christie	43.669542	-79.422564
7	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568
8	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752
9	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576
10	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817
11	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049
12	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049
13	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420
14	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
15	M5W	Downtown Toronto	Stn A PO Boxes	43.646435	-79.374846
16	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675
17	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280
18	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160

In [73]:
```
#Adding Downtown Toronto Coordinates
address = 'Downtown Toronto, CA'

geolocator = Nominatim(user_agent="ca_explorer")
location = geolocator.geocode(address)
lat_dt = location.latitude
lon_dt = location.longitude
```

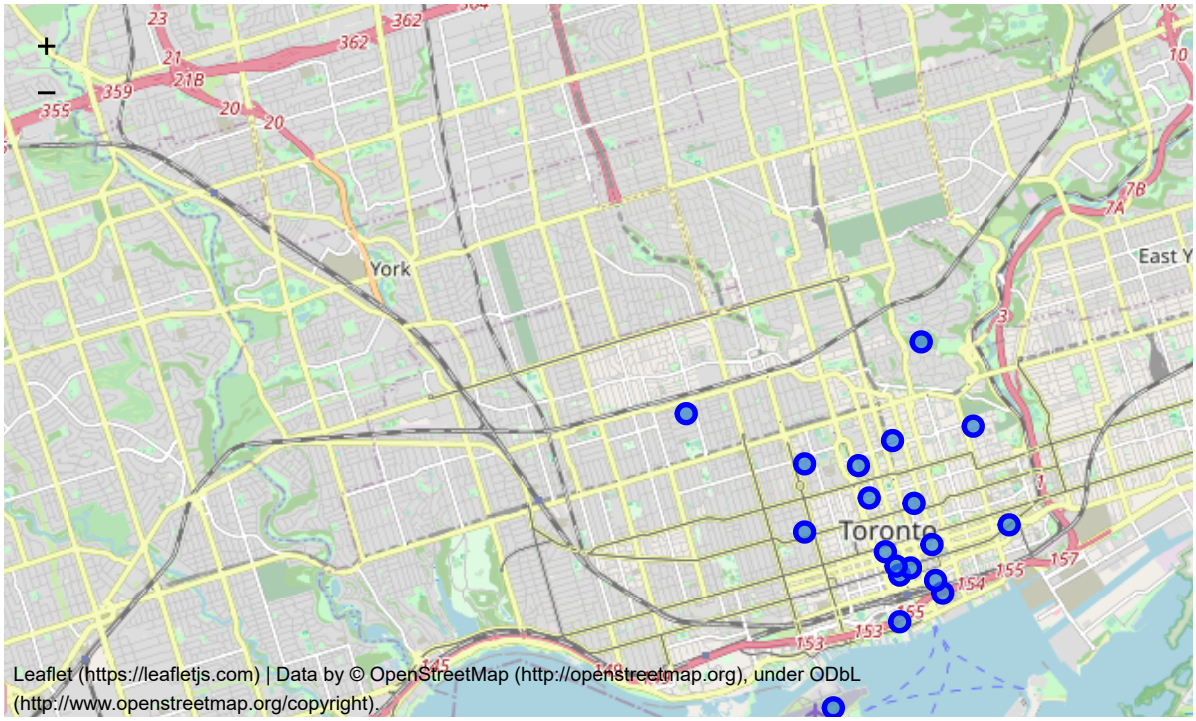
In [13]:

```
lat_dt= 43.6548
lon_dt= -79.3883
# create map of Toronto using latitude and longitude values
map_dttoronto = folium.Map(location=[lat_dt, lon_dt], zoom_start=12)

# add markers to map
for lat, lng, borough, neighborhood in zip(tor_data['Latitude'], tor_data['Longitude'],
    label = '{}, {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_dttoronto)

map_dttoronto
```

Out[13]:



In [38]:

```
#foursquare credentials
CLIENT_ID = 'EG4WQUKYXRNU1RZLTCLKYSKCROTXC41BMQX13I0B452HDPDW' # your Foursquare ID
CLIENT_SECRET = 'QKVUXUUVP3VXWVCWHRNAQKW3UKKL1EFKYBGW04RLDZXY440T' # your Foursquare Se
VERSION = '20180605'
LIMIT = 30
print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)
```

```
Your credentails:
CLIENT_ID: EG4WQUKYXRNU1RZLTCLKYSKCROTXC41BMQX13I0B452HDPDW
CLIENT_SECRET:QKVUXUUVP3VXWVCWHRNAQKW3UKKL1EFKYBGW04RLDZXY440T
```

In [15]:

```
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
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # 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 [16]:
```
Toronto_venues = getNearbyVenues(names=tor_data['Neighborhood'],
                                 latitudes=tor_data['Latitude'],
                                 longitudes=tor_data['Longitude']
                                )
```

```
Regent Park, Harbourfront
Queen's Park, Ontario Provincial Government
Garden District, Ryerson
St. James Town
Berczy Park
Central Bay Street
Christie
Richmond, Adelaide, King
Harbourfront East, Union Station, Toronto Islands
Toronto Dominion Centre, Design Exchange
Commerce Court, Victoria Hotel
University of Toronto, Harbord
Kensington Market, Chinatown, Grange Park
CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Nia
Rosedale
Stn A PO Boxes
St. James Town, Cabbagetown
First Canadian Place, Underground city
Church and Wellesley
```

In [17]:
```
print(Toronto_venues.shape)
Toronto_venues.head()
```

```
(519, 7)
```

Out[17]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery
1	Regent Park, Harbourfront	43.65426	-79.360636	Tandem Coffee	43.653559	-79.361809	Coffee Shop
2	Regent Park, Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149	Breakfast Spot
3	Regent Park, Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	43.653249	-79.358008	Distribution Center
4	Regent Park, Harbourfront	43.65426	-79.360636	Body Blitz Spa East	43.654735	-79.359874	Spa

In [18]:
```
Toronto_venues.groupby('Neighborhood').count()
```

Out[18]:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Berczy Park	30	30	30	30	30	30
CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island airport	18	18	18	18	18	18
Central Bay Street	30	30	30	30	30	30
Christie	17	17	17	17	17	17
Church and Wellesley	30	30	30	30	30	30
Commerce Court, Victoria Hotel	30	30	30	30	30	30
First Canadian Place, Underground city	30	30	30	30	30	30
Garden District, Ryerson	30	30	30	30	30	30
Harbourfront East, Union Station, Toronto Islands	30	30	30	30	30	30
Kensington Market, Chinatown, Grange Park	30	30	30	30	30	30
Queen's Park, Ontario Provincial Government	30	30	30	30	30	30
Regent Park, Harbourfront	30	30	30	30	30	30
Richmond, Adelaide, King	30	30	30	30	30	30
Rosedale	4	4	4	4	4	4
St. James Town	30	30	30	30	30	30
St. James Town, Cabbagetown	30	30	30	30	30	30
Stn A PO Boxes	30	30	30	30	30	30
Toronto Dominion Centre, Design Exchange	30	30	30	30	30	30
University of Toronto, Harbord	30	30	30	30	30	30

In [19]:

```
# one hot encoding
tor_onehot = pd.get_dummies(Toronto_venues[['Venue Category']], prefix="", prefix_sep='

# add neighborhood column back to dataframe
tor_onehot['Neighborhood'] = Toronto_venues['Neighborhood']

# move neighborhood column to the first column
fixed_columns = [tor_onehot.columns[-1]] + list(tor_onehot.columns[:-1])
tor_onehot = tor_onehot[fixed_columns]

tor_onehot.head()
```

Out[19]:

	Yoga Studio	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Aquarium	Art Gallery	...	T Restaur
0	0	0	0	0	0	0	0	0	0	0	...	
1	0	0	0	0	0	0	0	0	0	0	...	
2	0	0	0	0	0	0	0	0	0	0	...	
3	0	0	0	0	0	0	0	0	0	0	...	
4	0	0	0	0	0	0	0	0	0	0	...	

5 rows × 151 columns

In [20]:
```
tor_grouped = tor_onehot.groupby('Neighborhood').mean().reset_index()
tor_grouped
```

Out[20]:

	Neighborhood	Yoga Studio	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Aqua
0	Berczy Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
1	CN Tower, King and Spadina, Railway Lands, Har...	0.000000	0.055556	0.055556	0.055556	0.111111	0.166667	0.111111	0.000000	0.0(
2	Central Bay Street	0.033333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
3	Christie	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
4	Church and Wellesley	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
5	Commerce Court, Victoria Hotel	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033333	0.0(
6	First Canadian Place, Underground city	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033333	0.0(
7	Garden District, Ryerson	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
8	Harbourfront East, Union Station, Toronto Islands	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0:
9	Kensington Market, Chinatown, Grange Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
10	Queen's Park, Ontario Provincial Government	0.033333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
11	Regent Park, Harbourfront	0.033333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
12	Richmond, Adelaide, King	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033333	0.0(
13	Rosedale	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
14	St. James Town	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033333	0.0(
15	St. James Town, Cabbagetown	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
16	Stn A PO Boxes	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(
17	Toronto Dominion Centre, Design Exchange	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033333	0.0(
18	University of Toronto, Harbord	0.033333	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0(

19 rows × 151 columns

In [21]:

```
num_top_venues = 5

for hood in tor_grouped['Neighborhood']:
    print("----"+hood+"----")
    temp = tor_grouped[tor_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_top
    print('\n')
```

```
----Berczy Park----
                venue  freq
0         Coffee Shop  0.07
1        Cocktail Bar  0.07
2  Seafood Restaurant  0.07
3           Beer Bar  0.07
4               Park  0.03


----CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South
                 venue  freq
0      Airport Service  0.17
1       Airport Lounge  0.11
2     Airport Terminal  0.11
3  Rental Car Location  0.06
4              Airport  0.06


----Central Bay Street----
             venue  freq
0      Coffee Shop  0.23
1             Café  0.07
2      Yoga Studio  0.03
3              Spa  0.03
4  Bubble Tea Shop  0.03


----Christie----
                venue  freq
0       Grocery Store  0.24
1                Café  0.18
2                Park  0.12
3  Athletics & Sports  0.06
4          Restaurant  0.06


----Church and Wellesley----
                venue  freq
0        Burger Joint  0.07
1         Coffee Shop  0.03
2  Salon / Barbershop  0.03
3         Pizza Place  0.03
4           Bookstore  0.03


----Commerce Court, Victoria Hotel----
                venue  freq
0                Café  0.17
1           Gastropub  0.07
2  Japanese Restaurant  0.07
```

```
3        Deli / Bodega  0.07
4          Coffee Shop  0.07


----First Canadian Place, Underground city----
                venue  freq
0               Café  0.13
1         Coffee Shop  0.10
2          Restaurant  0.10
3  Seafood Restaurant  0.07
4           Tea Room  0.07


----Garden District, Ryerson----
                venue  freq
0               Café  0.10
1            Theater  0.07
2         Coffee Shop  0.07
3  Mexican Restaurant  0.03
4       Shopping Mall  0.03


----Harbourfront East, Union Station, Toronto Islands----
                  venue  freq
0                  Park  0.07
1                 Hotel  0.07
2                 Plaza  0.07
3  Performing Arts Venue  0.03
4           Supermarket  0.03


----Kensington Market, Chinatown, Grange Park----
                          venue  freq
0                          Café  0.10
1          Vietnamese Restaurant  0.07
2  Vegetarian / Vegan Restaurant  0.07
3             Mexican Restaurant  0.07
4                   Coffee Shop  0.03


----Queen's Park, Ontario Provincial Government----
                venue  freq
0         Coffee Shop  0.20
1     Sushi Restaurant  0.07
2          Yoga Studio  0.03
3  Fried Chicken Joint  0.03
4                Café  0.03


----Regent Park, Harbourfront----
            venue  freq
0     Coffee Shop  0.20
1            Park  0.10
2         Theater  0.07
3          Bakery  0.07
4  Breakfast Spot  0.07


----Richmond, Adelaide, King----
          venue  freq
0          Café  0.10
1   Coffee Shop  0.10
2  Concert Hall  0.03
3     Gastropub  0.03
```

```
4    Restaurant  0.03


----Rosedale----
         venue  freq
0         Park  0.50
1   Playground  0.25
2        Trail  0.25
3  Music Venue  0.00
4       Market  0.00


----St. James Town----
                     venue  freq
0                Gastropub  0.10
1                    Café  0.10
2              Coffee Shop  0.07
3  New American Restaurant  0.03
4                   Church  0.03


----St. James Town, Cabbagetown----
                venue  freq
0                Café  0.07
1  Italian Restaurant  0.07
2          Restaurant  0.07
3         Coffee Shop  0.07
4              Bakery  0.07


----Stn A PO Boxes----
                venue  freq
0  Seafood Restaurant  0.07
1        Cocktail Bar  0.07
2          Restaurant  0.07
3            Beer Bar  0.07
4                Café  0.07


----Toronto Dominion Centre, Design Exchange----
                venue  freq
0          Coffee Shop  0.17
1                Café  0.13
2  Japanese Restaurant  0.07
3        Deli / Bodega  0.07
4           Restaurant  0.07


----University of Toronto, Harbord----
                venue  freq
0                Café  0.13
1  Japanese Restaurant  0.07
2   Italian Restaurant  0.07
3                 Bar  0.07
4           Bookstore  0.07
```

In [22]:
```
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 [23]:
```
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_sorted1 = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted1['Neighborhood'] = tor_grouped['Neighborhood']

for ind in np.arange(tor_grouped.shape[0]):
    neighborhoods_venues_sorted1.iloc[ind, 1:] = return_most_common_venues(tor_grouped.

neighborhoods_venues_sorted1.head()
```

Out[23]:

	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
0	Berczy Park	Cocktail Bar	Coffee Shop	Seafood Restaurant	Beer Bar	Concert Hall	Creperie	Jazz Club	Bistro
1	CN Tower, King and Spadina, Railway Lands, Har...	Airport Service	Airport Lounge	Airport Terminal	Plane	Harbor / Marina	Rental Car Location	Boutique	Boat or Ferry
2	Central Bay Street	Coffee Shop	Café	Yoga Studio	Comic Shop	Bar	Seafood Restaurant	Sandwich Place	Italian Restaurant
3	Christie	Grocery Store	Café	Park	Athletics & Sports	Coffee Shop	Nightclub	Candy Store	Restaurant
4	Church and Wellesley	Burger Joint	Bubble Tea Shop	Beer Bar	Salon / Barbershop	Bookstore	Restaurant	Breakfast Spot	Ramen Restaurant

In [24]:
```
# set number of clusters
kclusters = 5

tor_grouped_clustering = tor_grouped.drop('Neighborhood', 1)

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

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

Out[24]:
```
array([4, 2, 0, 3, 4, 0, 0, 4, 4, 4], dtype=int32)
```

In [25]:
```
# add clustering labels
neighborhoods_venues_sorted1.insert(0, 'Cluster Labels', kmeans.labels_)

tor_merged = tor_data

# merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhc
tor_merged = tor_merged.join(neighborhoods_venues_sorted1.set_index('Neighborhood'), or
```

In [ ]:
```
```

In [26]:
```
tor_merged.head()
```

Out[26]:

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th M Comm Ve
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0	Coffee Shop	Park	Bakery	Breal
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	0	Coffee Shop	Sushi Restaurant	Wings Joint	D
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	4	Café	Theater	Coffee Shop	Col Ce
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	4	Café	Gastropub	Coffee Shop	
4	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	4	Cocktail Bar	Coffee Shop	Seafood Restaurant	Beer

In [27]:

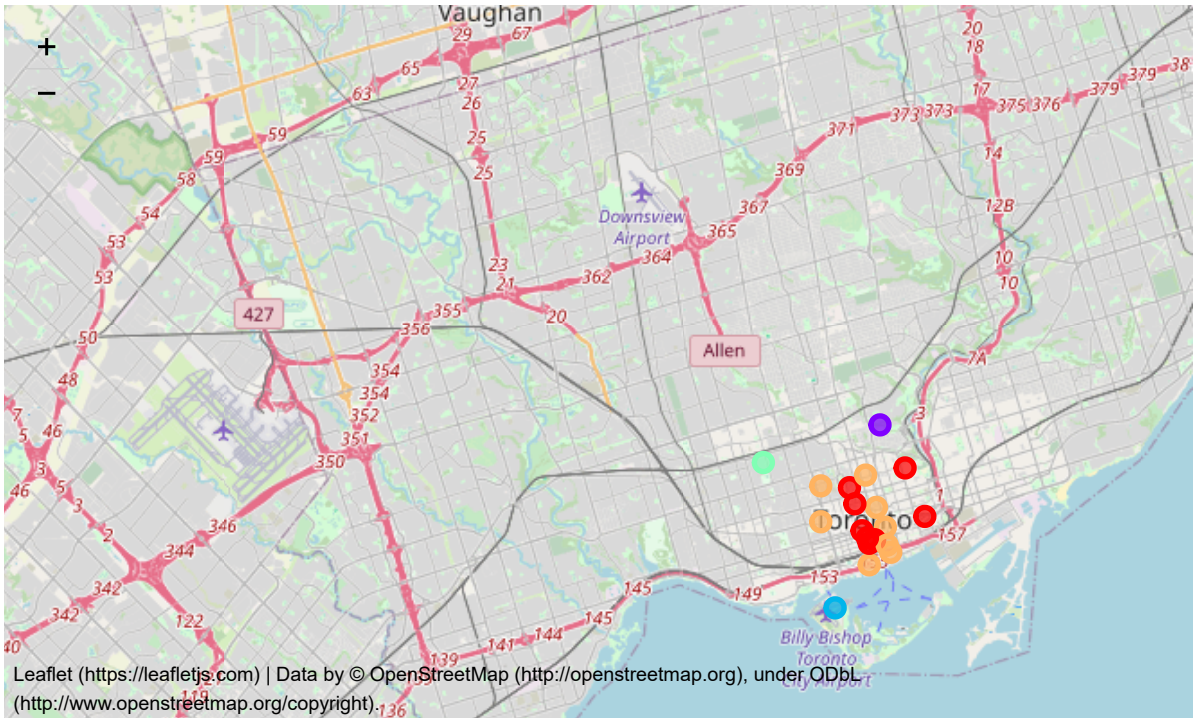
```
# create downtown Toronto map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# 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(tor_merged['Latitude'], tor_merged['Longitude'], tor_
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    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[27]:



In [32]:
```
#Cluster 0
tor_merged.loc[tor_merged['Cluster Labels'] == 0, tor_merged.columns[[2] + list(range(5
```

Out[32]:

	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
0	Regent Park, Harbourfront	0	Coffee Shop	Park	Bakery	Breakfast Spot	Theater	Yoga Studio	Historic Site
1	Queen's Park, Ontario Provincial Government	0	Coffee Shop	Sushi Restaurant	Wings Joint	Diner	Mexican Restaurant	Italian Restaurant	Hobby Shop
5	Central Bay Street	0	Coffee Shop	Café	Yoga Studio	Comic Shop	Bar	Seafood Restaurant	Sandwich Place
7	Richmond, Adelaide, King	0	Coffee Shop	Café	Seafood Restaurant	Asian Restaurant	Restaurant	Colombian Restaurant	Hotel
9	Toronto Dominion Centre, Design Exchange	0	Coffee Shop	Café	Deli / Bodega	Restaurant	Japanese Restaurant	Train Station	Gym / Fitness Center
10	Commerce Court, Victoria Hotel	0	Café	Coffee Shop	Gastropub	Restaurant	Japanese Restaurant	Deli / Bodega	Ice Cream Shop
16	St. James Town, Cabbagetown	0	Restaurant	Bakery	Café	Italian Restaurant	Coffee Shop	Market	Pub
17	First Canadian Place, Underground city	0	Café	Restaurant	Coffee Shop	Seafood Restaurant	Tea Room	Bookstore	Deli / Bodega

In [33]:
```
#cluster 1
tor_merged.loc[tor_merged['Cluster Labels'] == 1, tor_merged.columns[[2] + list(range(5
```

Out[33]:

	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 V
14	Rosedale	1	Park	Playground	Trail	Wings Joint	College Rec Center	Creperie	Cosmetics Shop	Co

In [34]:
```
#cluster 2
tor_merged.loc[tor_merged['Cluster Labels'] == 2, tor_merged.columns[[2] + list(range(5
```

Out[34]:

	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 Mo Comm Venu
13	CN Tower, King and Spadina, Railway Lands, Har...	2	Airport Service	Airport Lounge	Airport Terminal	Plane	Harbor / Marina	Rental Car Location	Boutique	Boat Fer

In [35]:
```
#cluster 3
tor_merged.loc[tor_merged['Cluster Labels'] == 3, tor_merged.columns[[2] + list(range(5
```

Out[35]:

	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 Mo Comm Venu
6	Christie	3	Grocery Store	Café	Park	Athletics & Sports	Coffee Shop	Nightclub	Candy Store	Restaura

In [36]:
```
#cluster 4
tor_merged.loc[tor_merged['Cluster Labels'] == 4, tor_merged.columns[[2] + list(range(5
```

Out[36]:

	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
2	Garden District, Ryerson	4	Café	Theater	Coffee Shop	College Rec Center	Shopping Mall	Sandwich Place	Bar
3	St. James Town	4	Café	Gastropub	Coffee Shop	Gym	Restaurant	Cosmetics Shop	Creperie
4	Berczy Park	4	Cocktail Bar	Coffee Shop	Seafood Restaurant	Beer Bar	Concert Hall	Creperie	Jazz Club
8	Harbourfront East, Union Station, Toronto Islands	4	Hotel	Park	Plaza	Roof Deck	Skating Rink	IT Services	Ice Cream Shop
11	University of Toronto, Harbord	4	Café	Italian Restaurant	Bookstore	Restaurant	Japanese Restaurant	Bar	Bakery
12	Kensington Market, Chinatown, Grange Park	4	Café	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Cheese Shop	Record Shop	Coffee Shop
15	Stn A PO Boxes	4	Cocktail Bar	Café	Beer Bar	Seafood Restaurant	Restaurant	Food Truck	Vegetarian / Vegan Restaurant
18	Church and Wellesley	4	Burger Joint	Bubble Tea Shop	Beer Bar	Salon / Barbershop	Bookstore	Restaurant	Breakfast Spot

In [ ]:

In [ ]: