In [2]:
```python
import numpy as np # library to handle data in a vectorized manner

import pandas as pd # library for data analsysis
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json # library to handle JSON files

#!conda install -c conda-forge geopy --yes # uncomment this line if
you haven't completed the Foursquare API lab
from geopy.geocoders import Nominatim # convert an address into lat
itude and longitude values

import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into
a pandas dataframe

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors

# import k-means from clustering stage
from sklearn.cluster import KMeans

#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this
line if you haven't completed the Foursquare API lab
import folium # map rendering library

print('Libraries imported.')
```

Libraries imported.

Use geopy library to get the latitude and longitude values of Toronto.

In [3]:
```python
address = 'Toronto'

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

The geograpical coordinate of Toronto are 43.653963, -79.387207.

Create a map of Toronto with neighborhoods superimposed on top.

In [29]: *#Load Toronto neighbourhoods with coordinates prepared in part 2 of the assignment*
toronto_df = pd.read_csv('toronto_coords_acquisition.csv')
toronto_df.head()

Out[29]:

	PostalCode	Borough	Neighborhood	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

Limit boroughs to those containing Toronto in their name

In [26]: toronto_df_sel = toronto_df[toronto_df['Borough'].str.contains('Toronto')]
toronto_df_sel.shape

Out[26]: (38, 5)

In [31]:
```python
# create map of New York 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(toronto_df_sel['Latitude'], toronto_df_sel['Longitude'], toronto_df_sel['Borough'], toronto_df_sel['Neighborhood']):
    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[31]:
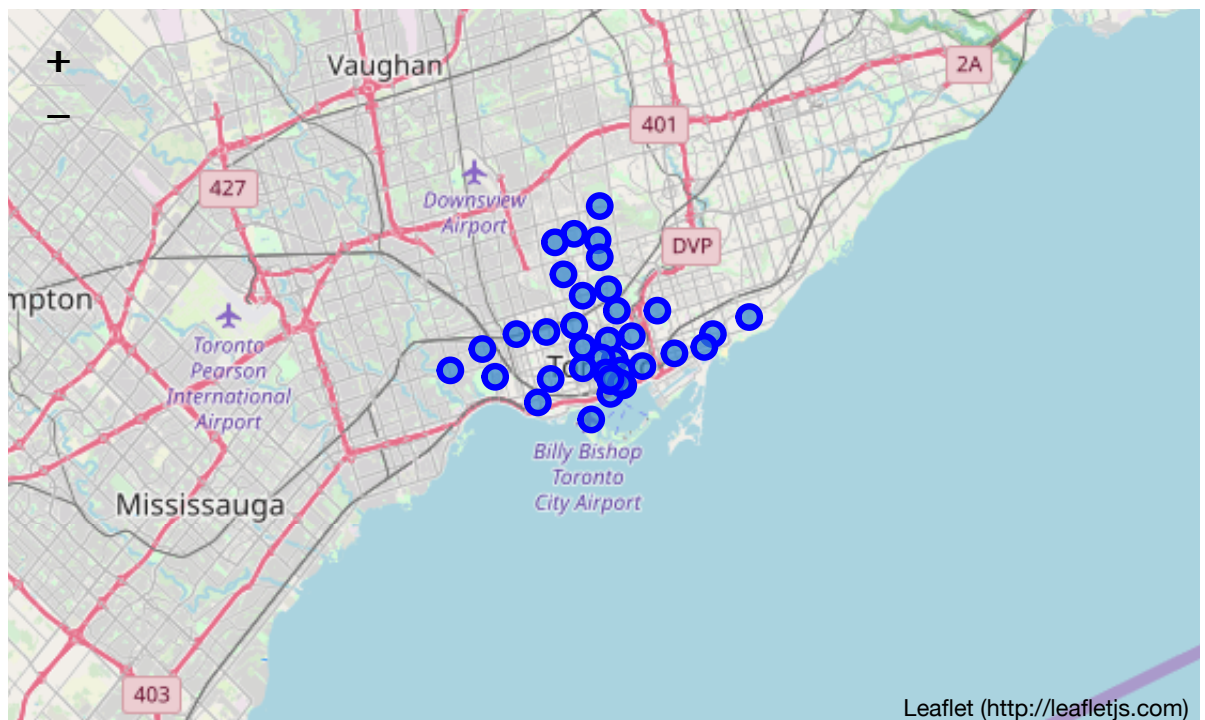
Four square credentials

I deleted my credentials for security reasons after executing the code

In [139]:
```python
CLIENT_ID = 'Your client ID' # your Foursquare ID
CLIENT_SECRET = 'Your client secret' # your Foursquare Secret
VERSION = '20180605' # Foursquare API version

print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)
```

```
Your credentails:
CLIENT_ID: Your client ID
CLIENT_SECRET:Your client secret
```

Let's explore the first neighborhood in our dataframe

In [33]:
```python
neighb = toronto_df.loc[0, 'Neighborhood'].split(',')[0]
neighb
```

Out[33]: `'Rouge'`

In [34]:
```python
neighborhood_latitude = toronto_df.loc[0, 'Latitude'] # neighborhood latitude value
neighborhood_longitude = toronto_df.loc[0, 'Longitude'] # neighborhood longitude value

neighborhood_name = neighb # neighborhood name

print('Latitude and longitude values of {} are {}, {}.'.format(neighborhood_name,
                                                               neighborhood_latitude,
                                                               neighborhood_longitude))
```

```
Latitude and longitude values of Rouge are 43.8066863, -79.1943534
0000003.
```

Now, let's get the top 100 venues that are in Rouge within a radius of 500 meters.

In [140]:
```python
# type your answer here
LIMIT = 100 # limit of number of venues returned by Foursquare API
radius = 500 # define radius
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    neighborhood_latitude,
    neighborhood_longitude,
    radius,
    LIMIT)
url # display URL
```

Out[140]:
```
'https://api.foursquare.com/v2/venues/explore?&client_id=Your clie
nt ID&client_secret=Your client secret&v=20180605&ll=43.8066863,-7
9.19435340000003&radius=500&limit=100'
```

In [36]:
```python
results = requests.get(url).json()
results
```

Out[36]: {'meta': {'code': 200, 'requestId': '5cd7b6bddb04f559d656b450'},
 'response': {'warning': {'text': "There aren't a lot of results n
ear you. Try something more general, reset your filters, or expand
the search area."},
  'headerLocation': 'Malvern',
  'headerFullLocation': 'Malvern, Toronto',
  'headerLocationGranularity': 'neighborhood',
  'totalResults': 1,
  'suggestedBounds': {'ne': {'lat': 43.811186304500005,
    'lng': -79.18812958073043},
   'sw': {'lat': 43.8021862955, 'lng': -79.20057721926962}},
  'groups': [{'type': 'Recommended Places',
    'name': 'recommended',
    'items': [{'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
      'venue': {'id': '4bb6b9446edc76b0d771311c',
       'name': "Wendy's",
       'location': {'crossStreet': 'Morningside & Sheppard',
        'lat': 43.80744841934756,
        'lng': -79.19905558052072,
        'labeledLatLngs': [{'label': 'display',
          'lat': 43.80744841934756,
          'lng': -79.19905558052072}],
        'distance': 387,
        'cc': 'CA',
        'city': 'Toronto',
        'state': 'ON',
        'country': 'Canada',
        'formattedAddress': ['Toronto ON', 'Canada']},
       'categories': [{'id': '4bf58dd8d48988d16e941735',
         'name': 'Fast Food Restaurant',
         'pluralName': 'Fast Food Restaurants',
         'shortName': 'Fast Food',
         'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v
2/food/fastfood_',
          'suffix': '.png'},
         'primary': True}],
       'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4bb6b9446edc76b0d771311c-0'}]}]}}

```
In [37]:  # function that extracts the category of the venue
          def get_category_type(row):
              try:
                  categories_list = row['categories']
              except:
                  categories_list = row['venue.categories']

              if len(categories_list) == 0:
                  return None
              else:
                  return categories_list[0]['name']
```

```
In [38]:  venues = results['response']['groups'][0]['items']

          nearby_venues = json_normalize(venues) # flatten JSON

          # filter columns
          filtered_columns = ['venue.name', 'venue.categories', 'venue.locati
          on.lat', 'venue.location.lng']
          nearby_venues =nearby_venues.loc[:, filtered_columns]

          # filter the category for each row
          nearby_venues['venue.categories'] = nearby_venues.apply(get_categor
          y_type, axis=1)

          # clean columns
          nearby_venues.columns = [col.split(".")[-1] for col in nearby_venue
          s.columns]

          nearby_venues.head()
```

Out[38]:

	name	categories	lat	lng
0	Wendy's	Fast Food Restaurant	43.807448	-79.199056

Only one venue was returned for Rouge. Let us explore more neighborhoods.

Explore Neighborhoods in Toronto

If neighbourhoods have the same postal code, I will keep them grouped as it was done in parts 1 and 2 of this assignment.

```
In [40]:  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={}&v={}&ll={},{}&radius={}&limit={}'.format(
                      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_lis
          t for item in venue_list])
              nearby_venues.columns = ['Neighborhood',
                            'Neighborhood Latitude',
                            'Neighborhood Longitude',
                            'Venue',
                            'Venue Latitude',
                            'Venue Longitude',
                            'Venue Category']

              return(nearby_venues)
```

In [43]:
```python
toronto_venues = getNearbyVenues(names=toronto_df_sel['Neighborhood'],
                                 latitudes=toronto_df_sel['Latitude'],
                                 longitudes=toronto_df_sel['Longitude']
                                )
```

```
The Beaches
The Danforth West, Riverdale
The Beaches West, India Bazaar
Studio District
Lawrence Park
Davisville North
North Toronto West
Davisville
Moore Park, Summerhill East
Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West
Rosedale
Cabbagetown, St. James Town
Church and Wellesley
Harbourfront, Regent Park
Ryerson, Garden District
St. James Town
Berczy Park
Central Bay Street
Adelaide, King, Richmond
Harbourfront East, Toronto Islands, Union Station
Design Exchange, Toronto Dominion Centre
Commerce Court, Victoria Hotel
Roselawn
Forest Hill North, Forest Hill West
The Annex, North Midtown, Yorkville
Harbord, University of Toronto
Chinatown, Grange Park, Kensington Market
CN Tower, Bathurst Quay, Island airport, Harbourfront West, King a
nd Spadina, Railway Lands, South Niagara
Stn A PO Boxes 25 The Esplanade
First Canadian Place, Underground city
Christie
Dovercourt Village, Dufferin
Little Portugal, Trinity
Brockton, Exhibition Place, Parkdale Village
High Park, The Junction South
Parkdale, Roncesvalles
Runnymede, Swansea
Business Reply Mail Processing Centre 969 Eastern
```

```
In [66]: print(toronto_venues.shape)
         toronto_venues.head()
```

```
(1700, 7)
```

Out[66]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Ve Categ
0	The Beaches	43.676357	-79.293031	The Big Carrot Natural Food Market	43.678879	-79.297734	Health F S
1	The Beaches	43.676357	-79.293031	Grover Pub and Grub	43.679181	-79.297215	
2	The Beaches	43.676357	-79.293031	St-Denis Studios Inc.	43.675031	-79.288022	Music Ve
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neighborh
4	The Danforth West, Riverdale	43.679557	-79.352188	Pantheon	43.677621	-79.351434	G Restau

There is a Venue Category called Neighborhood

```
In [67]: toronto_venues[toronto_venues['Venue Category']=='Neighborhood']
```

Out[67]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	
3	The Beaches	43.676357	-79.293031	Upper Beaches	43.680563	-79.292869	Neig
76	Studio District	43.659526	-79.340923	Leslieville	43.662070	-79.337856	Neig
730	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	43.653232	-79.385296	Neig
814	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	Harbourfront	43.639526	-79.380688	Neig

This value may interfere with the column name 'Neighborhood' when creating one-hot encoding later. Besides, the meaning of this Category is unclear. For this reason, I decided to remove the rows corresponding to this category.

In [68]:
```python
print('Toronto venues shape before removal of Neighborhood Venue Ca
tegory: ', toronto_venues.shape)
toronto_venues = toronto_venues[toronto_venues['Venue Category'] !=
'Neighborhood']
print('Toronto venues shape after removal of Neighborhood Venue Cat
egory: ', toronto_venues.shape)
```

Toronto venues shape before removal of Neighborhood Venue Category
: (1700, 7)
Toronto venues shape after removal of Neighborhood Venue Category:
(1696, 7)

Four rows have been removed

Number of venues per neighborhood (or grouped neighbourhoods if they have the same postal code)

In [69]:
```python
toronto_venues.groupby('Neighborhood').count()
```

Out[69]:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Adelaide, King, Richmond	99	99	99	99	99	99
Berczy Park	57	57	57	57	57	57
Brockton, Exhibition Place, Parkdale Village	19	19	19	19	19	19
Business Reply Mail Processing Centre 969 Eastern	17	17	17	17	17	17
CN Tower, Bathurst Quay, Island airport, Harbourfront West, King and Spadina, Railway Lands, South Niagara	15	15	15	15	15	15
Cabbagetown, St. James Town	44	44	44	44	44	44
Central Bay Street	88	88	88	88	88	88
Chinatown, Grange Park, Kensington Market	100	100	100	100	100	100
Christie	16	16	16	16	16	16

Church and Wellesley	88	88	88	88	88	88
Commerce Court, Victoria Hotel	100	100	100	100	100	100
Davisville	32	32	32	32	32	32
Davisville North	7	7	7	7	7	7
Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West	14	14	14	14	14	14
Design Exchange, Toronto Dominion Centre	100	100	100	100	100	100
Dovercourt Village, Dufferin	20	20	20	20	20	20
First Canadian Place, Underground city	100	100	100	100	100	100
Forest Hill North, Forest Hill West	4	4	4	4	4	4
Harbord, University of Toronto	34	34	34	34	34	34
Harbourfront East, Toronto Islands, Union Station	99	99	99	99	99	99
Harbourfront, Regent Park	46	46	46	46	46	46
High Park, The Junction South	23	23	23	23	23	23
Lawrence Park	4	4	4	4	4	4
Little Portugal, Trinity	63	63	63	63	63	63
Moore Park, Summerhill East	3	3	3	3	3	3
North Toronto West	20	20	20	20	20	20
Parkdale, Roncesvalles	15	15	15	15	15	15
Rosedale	4	4	4	4	4	4
Roselawn	1	1	1	1	1	1
Runnymede, Swansea	40	40	40	40	40	40
Ryerson, Garden District	100	100	100	100	100	100

St. James Town	100	100	100	100	100	100
Stn A PO Boxes 25 The Esplanade	95	95	95	95	95	95
Studio District	36	36	36	36	36	36
The Annex, North Midtown, Yorkville	24	24	24	24	24	24
The Beaches	3	3	3	3	3	3
The Beaches West, India Bazaar	22	22	22	22	22	22
The Danforth West, Riverdale	44	44	44	44	44	44

Number of unique venue categories

In [70]:
```python
print('There are {} uniques categories.'.format(len(toronto_venues[
'Venue Category'].unique())))
```

There are 235 uniques categories.

Analyze each neighborhood

In [74]:
```python
# one hot encoding
toronto_onehot = pd.get_dummies(toronto_venues[['Venue Category']],
prefix="", prefix_sep="")

# add neighborhood column back to dataframe
toronto_onehot['Neighborhood'] = toronto_venues['Neighborhood']

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

toronto_onehot.head()
```

Out[74]:

	Neighborhood	Adult Boutique	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Termina
0	The Beaches	0	0	0	0	0	0	0	0
1	The Beaches	0	0	0	0	0	0	0	0
2	The Beaches	0	0	0	0	0	0	0	0
4	The Danforth West, Riverdale	0	0	0	0	0	0	0	0
5	The Danforth West, Riverdale	0	0	0	0	0	0	0	0

In [75]: `toronto_onehot.shape`

Out[75]: `(1696, 236)`

Group rows by neighborhood and by taking the mean of the frequency of occurrence of each category

In [76]:
```python
toronto_grouped = toronto_onehot.groupby('Neighborhood').mean().res
et_index()
toronto_grouped
```

Out[76]:

	Neighborhood	Adult Boutique	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service
0	Adelaide, King, Richmond	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
1	Berczy Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0

2	Brockton, Exhibition Place, Parkdale Village	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
3	Business Reply Mail Processing Centre 969 Eastern	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
4	CN Tower, Bathurst Quay, Island airport, Harbo...	0.000000	0.000000	0.066667	0.066667	0.066667	0.133333	0.2
5	Cabbagetown, St. James Town	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
6	Central Bay Street	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
7	Chinatown, Grange Park, Kensington Market	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
8	Christie	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
9	Church and Wellesley	0.011364	0.011364	0.000000	0.000000	0.000000	0.000000	0.0
10	Commerce Court, Victoria Hotel	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
11	Davisville	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
12	Davisville North	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
13	Deer Park, Forest Hill SE, Rathnelly, South Hi...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
14	Design Exchange, Toronto Dominion Centre	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
15	Dovercourt Village, Dufferin	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
16	First Canadian Place, Underground city	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
17	Forest Hill North, Forest Hill West	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0

18	Harbord, University of Toronto	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
19	Harbourfront East, Toronto Islands, Union Station	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
20	Harbourfront, Regent Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
21	High Park, The Junction South	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
22	Lawrence Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
23	Little Portugal, Trinity	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
24	Moore Park, Summerhill East	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
25	North Toronto West	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
26	Parkdale, Roncesvalles	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
27	Rosedale	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
28	Roselawn	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
29	Runnymede, Swansea	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
30	Ryerson, Garden District	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
31	St. James Town	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
32	Stn A PO Boxes 25 The Esplanade	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
33	Studio District	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
34	The Annex, North Midtown, Yorkville	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
35	The Beaches	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
36	The Beaches West, India Bazaar	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
37	The Danforth West, Riverdale	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0

In [77]: `toronto_grouped.shape`

Out[77]: `(38, 236)`

5 most common venues per neighborhood

In [79]:
```python
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_top_venues))
    print('\n')
```

```
----Adelaide, King, Richmond----
                venue  freq
0          Coffee Shop  0.06
1                 Café  0.05
2           Steakhouse  0.04
3      Thai Restaurant  0.04
4  American Restaurant  0.04


----Berczy Park----
            venue  freq
0     Coffee Shop  0.09
1    Cocktail Bar  0.05
2      Restaurant  0.04
3            Café  0.04
4        Beer Bar  0.04


----Brockton, Exhibition Place, Parkdale Village----
            venue  freq
0  Breakfast Spot  0.11
1     Coffee Shop  0.11
2            Café  0.11
3    Climbing Gym  0.05
4   Grocery Store  0.05


----Business Reply Mail Processing Centre 969 Eastern----
                venue  freq
0  Light Rail Station  0.12
1         Yoga Studio  0.06
```

	venue	freq
2	Auto Workshop	0.06
3	Garden Center	0.06
4	Garden	0.06

----CN Tower, Bathurst Quay, Island airport, Harbourfront West, King and Spadina, Railway Lands, South Niagara----

	venue	freq
0	Airport Service	0.20
1	Airport Lounge	0.13
2	Airport Terminal	0.13
3	Boat or Ferry	0.07
4	Airport	0.07

----Cabbagetown, St. James Town----

	venue	freq
0	Coffee Shop	0.09
1	Bakery	0.05
2	Pizza Place	0.05
3	Market	0.05
4	Italian Restaurant	0.05

----Central Bay Street----

	venue	freq
0	Coffee Shop	0.15
1	Café	0.06
2	Italian Restaurant	0.05
3	Burger Joint	0.03
4	Chinese Restaurant	0.02

----Chinatown, Grange Park, Kensington Market----

	venue	freq
0	Café	0.08
1	Vegetarian / Vegan Restaurant	0.06
2	Bar	0.05
3	Mexican Restaurant	0.04
4	Coffee Shop	0.04

----Christie----

	venue	freq
0	Café	0.19
1	Grocery Store	0.19
2	Park	0.12
3	Coffee Shop	0.06
4	Baby Store	0.06

----Church and Wellesley----

	venue	freq
0	Coffee Shop	0.07

```
1    Japanese Restaurant  0.07
2               Gay Bar  0.06
3     Sushi Restaurant  0.05
4            Restaurant  0.03
```

----Commerce Court, Victoria Hotel----
```
                venue  freq
0          Coffee Shop  0.11
1                 Café  0.07
2                Hotel  0.06
3           Restaurant  0.05
4  American Restaurant  0.04
```

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

----Davisville North----
```
               venue  freq
0  Food & Drink Shop  0.14
1     Breakfast Spot  0.14
2                Gym  0.14
3              Hotel  0.14
4               Park  0.14
```

----Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West----
```
                venue  freq
0                  Pub  0.14
1          Coffee Shop  0.14
2  American Restaurant  0.07
3     Sushi Restaurant  0.07
4          Supermarket  0.07
```

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

----Dovercourt Village, Dufferin----
```
                 venue   freq
```

```
0              Bakery  0.10
1            Pharmacy  0.10
2         Supermarket  0.10
3  Brazilian Restaurant  0.05
4                 Bar  0.05
```

```
----First Canadian Place, Underground city----
          venue  freq
0          Café  0.09
1   Coffee Shop  0.09
2         Hotel  0.04
3    Restaurant  0.03
4     Gastropub  0.03
```

```
----Forest Hill North, Forest Hill West----
               venue  freq
0              Trail  0.25
1   Sushi Restaurant  0.25
2           Bus Line  0.25
3      Jewelry Store  0.25
4     Adult Boutique  0.00
```

```
----Harbord, University of Toronto----
                venue  freq
0               Café  0.15
1                Bar  0.06
2         Restaurant  0.06
3           Bookstore  0.06
4  Japanese Restaurant  0.06
```

```
----Harbourfront East, Toronto Islands, Union Station----
                venue  freq
0         Coffee Shop  0.12
1               Hotel  0.05
2            Aquarium  0.05
3   Italian Restaurant  0.04
4                Café  0.04
```

```
----Harbourfront, Regent Park----
          venue  freq
0   Coffee Shop  0.15
1           Pub  0.07
2          Park  0.07
3        Bakery  0.07
4          Café  0.07
```

```
----High Park, The Junction South----
                venue  freq
```

```
0     Mexican Restaurant  0.09
1                   Café  0.09
2                    Bar  0.09
3    Arts & Crafts Store  0.04
4            Flea Market  0.04
```

----Lawrence Park----

```
                         venue  freq
0    Construction & Landscaping  0.25
1                          Park  0.25
2                   Swim School  0.25
3                      Bus Line  0.25
4                Adult Boutique  0.00
```

----Little Portugal, Trinity----

```
                      venue  freq
0                       Bar  0.13
1           Asian Restaurant  0.05
2                Coffee Shop  0.05
3    New American Restaurant  0.03
4                 Restaurant  0.03
```

----Moore Park, Summerhill East----

```
                       venue  freq
0                  Playground  0.33
1                       Trail  0.33
2                Tennis Court  0.33
3                 Music Venue  0.00
4    Mediterranean Restaurant  0.00
```

----North Toronto West----

```
             venue  freq
0    Clothing Store  0.15
1       Coffee Shop  0.10
2       Yoga Studio  0.05
3     Metro Station  0.05
4               Spa  0.05
```

----Parkdale, Roncesvalles----

```
                 venue  freq
0        Breakfast Spot  0.13
1             Gift Shop  0.13
2              Bookstore  0.07
3     Italian Restaurant  0.07
4                   Bank  0.07
```

----Rosedale----

```
                    venue   freq
```

```
0                         Park  0.50
1                   Playground  0.25
2                        Trail  0.25
3                  Music Venue  0.00
4      Mediterranean Restaurant  0.00
```

----Roselawn----

```
                          venue  freq
0                        Garden   1.0
1                Adult Boutique   0.0
2                        Market   0.0
3      Mediterranean Restaurant   0.0
4                   Men's Store   0.0
```

----Runnymede, Swansea----

```
                 venue  freq
0                 Café  0.10
1           Pizza Place  0.08
2           Coffee Shop  0.08
3      Sushi Restaurant  0.05
4                   Gym  0.05
```

----Ryerson, Garden District----

```
                         venue  freq
0                   Coffee Shop  0.08
1                Clothing Store  0.07
2                Cosmetics Shop  0.04
3                          Café  0.04
4      Middle Eastern Restaurant  0.03
```

----St. James Town----

```
               venue  freq
0        Coffee Shop  0.08
1               Café  0.05
2         Restaurant  0.05
3              Hotel  0.05
4      Cosmetics Shop  0.04
```

----Stn A PO Boxes 25 The Esplanade----

```
                 venue  freq
0           Coffee Shop  0.11
1                  Café  0.04
2            Restaurant  0.04
3      Seafood Restaurant  0.03
4              Beer Bar  0.03
```

----Studio District----

```
                   venue  freq
```

```
0                   Café   0.11
1           Coffee Shop    0.08
2     Italian Restaurant   0.06
3                 Bakery   0.06
4   American Restaurant    0.06
```

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

```
----The Beaches----
                    venue  freq
0       Health Food Store   0.33
1                     Pub   0.33
2             Music Venue   0.33
3  New American Restaurant   0.00
4             Men's Store   0.00
```

```
----The Beaches West, India Bazaar----
                 venue  freq
0       Sandwich Place   0.09
1                 Park   0.09
2  Fast Food Restaurant   0.05
3    Italian Restaurant   0.05
4      Sushi Restaurant   0.05
```

```
----The Danforth West, Riverdale----
                   venue  freq
0      Greek Restaurant   0.18
1           Coffee Shop   0.09
2        Ice Cream Shop   0.07
3  Furniture / Home Store   0.05
4     Italian Restaurant   0.05
```

Put results into Pandas DataFrame

In [104]:
```python
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 [122]:
```python
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_grouped.iloc[ind, :], num_top_venues)

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

```
(38, 11)
```

Out[122]:

	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 Com... V...
0	Adelaide, King, Richmond	Coffee Shop	Café	Steakhouse	American Restaurant	Thai Restaurant	Hotel	E
1	Berczy Park	Coffee Shop	Cocktail Bar	Bakery	Beer Bar	Steakhouse	Seafood Restaurant	C
2	Brockton, Exhibition Place, Parkdale Village	Café	Breakfast Spot	Coffee Shop	Grocery Store	Bar	Caribbean Restaurant	Resta
3	Business Reply Mail Processing Centre 969 Eastern	Light Rail Station	Yoga Studio	Auto Workshop	Garden Center	Garden	Fast Food Restaurant	Fa N
4	CN Tower, Bathurst Quay, Island airport, Harbo...	Airport Service	Airport Terminal	Airport Lounge	Plane	Sculpture Garden	Boutique	E

Cluster Neighborhoods

In [126]:
```python
# set number of clusters
kclusters = 5

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

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

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

Out[126]: array([4, 4, 4, 4, 0, 4, 4, 4, 4, 4], dtype=int32)

Create a dataframe

In [131]:
```python
# add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labe
ls_)
#neighborhoods_venues_sorted.head()
toronto_merged = toronto_df_sel

# merge toronto_grouped with toronto_data to add latitude/longitude
for each neighborhood
toronto_merged = toronto_merged.join(neighborhoods_venues_sorted.se
t_index('Neighborhood'), on='Neighborhood')

toronto_merged # check the last columns!
```

Out[131]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2 C
37	M4E	East Toronto	The Beaches	43.676357	-79.293031	3	Music Venue	Hea
41	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188	4	Greek Restaurant	Cofi
42	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572	4	Sandwich Place	
43	M4M	East Toronto	Studio District	43.659526	-79.340923	4	Café	Cofi
44	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790	0	Park	Con Lanc
45	M4P	Central Toronto	Davisville North	43.712751	-79.390197	4	Gym	
46	M4R	Central Toronto	North Toronto West	43.715383	-79.405678	4	Clothing Store	Cofi
47	M4S	Central Toronto	Davisville	43.704324	-79.388790	4	Sandwich Place	
48	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160	2	Playground	Tenr
49	M4V	Central Toronto	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049	4	Coffee Shop	
50	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529	2	Park	Pla
51	M4X	Downtown	Cabbagetown, St. James	43.667967	-79.367675	4	Coffee	

		Toronto	Town				Shop	
52	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	4	Coffee Shop	J Re
53	M5A	Downtown Toronto	Harbourfront, Regent Park	43.654260	-79.360636	4	Coffee Shop	
54	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937	4	Coffee Shop	
55	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	4	Coffee Shop	
56	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	4	Coffee Shop	Coc
57	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	4	Coffee Shop	
58	M5H	Downtown Toronto	Adelaide, King, Richmond	43.650571	-79.384568	4	Coffee Shop	
59	M5J	Downtown Toronto	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	4	Coffee Shop	
60	M5K	Downtown Toronto	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	4	Coffee Shop	
61	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	4	Coffee Shop	
63	M5N	Central Toronto	Roselawn	43.711695	-79.416936	1	Garden	
64	M5P	Central Toronto	Forest Hill North, Forest Hill West	43.696948	-79.411307	0	Trail	
65	M5R	Central Toronto	The Annex, North Midtown, Yorkville	43.672710	-79.405678	4	Coffee Shop	S
66	M5S	Downtown Toronto	Harbord, University of Toronto	43.662696	-79.400049	4	Café	B
67	M5T	Downtown Toronto	Chinatown, Grange Park, Kensington Market	43.653206	-79.400049	4	Café	Veg Re
68	M5V	Downtown Toronto	CN Tower, Bathurst Quay, Island airport, Harbo...	43.628947	-79.394420	0	Airport Service	

69	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846	4	Coffee Shop	Re
70	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	4	Café	Cof
75	M6G	Downtown Toronto	Christie	43.669542	-79.422564	4	Grocery Store	
76	M6H	West Toronto	Dovercourt Village, Dufferin	43.669005	-79.442259	4	Pharmacy	Sup
77	M6J	West Toronto	Little Portugal, Trinity	43.647927	-79.419750	4	Bar	Cof
78	M6K	West Toronto	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191	4	Café	E
82	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763	4	Mexican Restaurant	
83	M6R	West Toronto	Parkdale, Roncesvalles	43.648960	-79.456325	4	Breakfast Spot	C
84	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450	4	Café	Piz
87	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558	4	Light Rail Station	Yog

Visualize clusters

In [132]:
```python
# create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_star
t=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'], toron
to_merged['Longitude'], toronto_merged['Neighborhood'], toronto_mer
ged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), par
se_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[132]:

Leaflet (http://leafletjs.com)
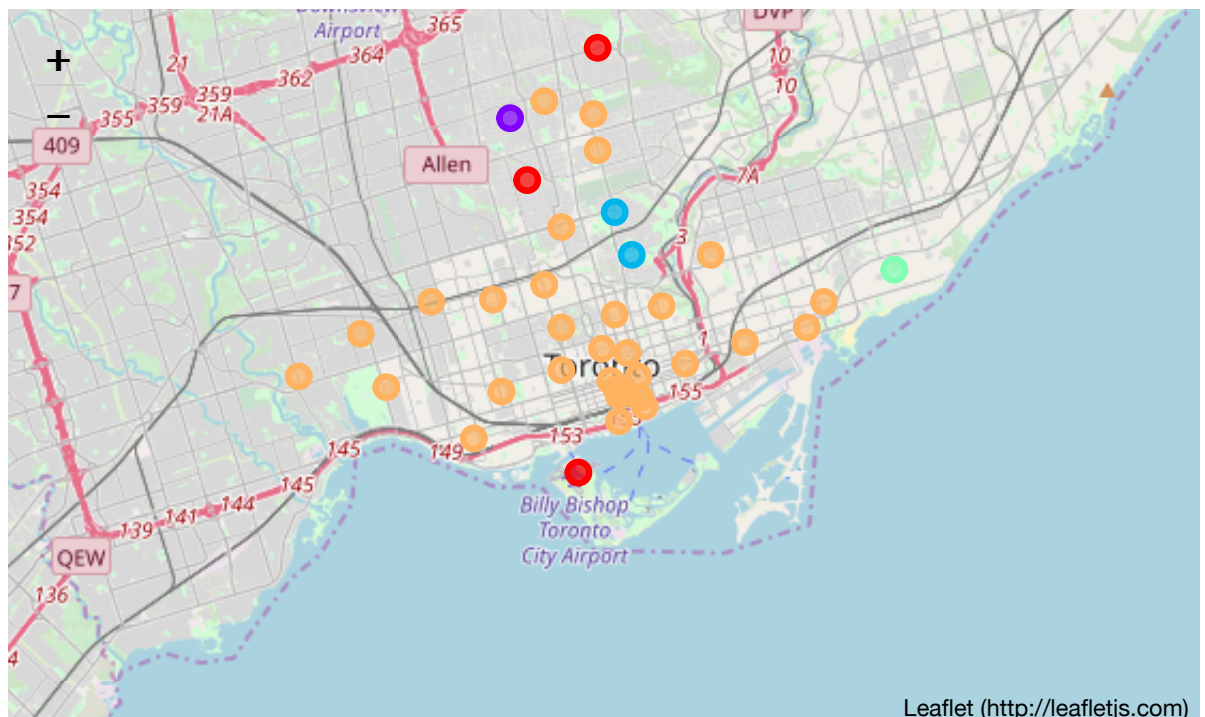
Examine Clusters

Find out the characteristics (venues) which distinguish a neighborhood from other neighborhoods

Cluster1

```
In [133]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 0, toronto_m
          erged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Out[133]:

	Borough	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
44	Central Toronto	0	Park	Construction & Landscaping	Swim School	Bus Line	Yoga Studio	Dim Sum Restaurant
64	Central Toronto	0	Trail	Jewelry Store	Sushi Restaurant	Bus Line	Yoga Studio	Diner
68	Downtown Toronto	0	Airport Service	Airport Terminal	Airport Lounge	Plane	Sculpture Garden	Boutique

Cluster 2</h2>

```
In [134]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 1, toronto_m
          erged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Out[134]:

	Borough	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	7 C
63	Central Toronto	1	Garden	Dessert Shop	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	E Re

Cluster 3

In [136]:
```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 2, toronto_m
erged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Out[136]:

	Borough	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
48	Central Toronto	2	Playground	Tennis Court	Trail	Yoga Studio	Doner Restaurant	Dim Sum Restaurant
50	Downtown Toronto	2	Park	Playground	Trail	Yoga Studio	Dessert Shop	Event Space

Cluster 4

In [137]:
```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 3, toronto_m
erged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Out[137]:

	Borough	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	7 C
37	East Toronto	3	Music Venue	Health Food Store	Pub	Department Store	Event Space	Ethiopian Restaurant	Ele

Cluster 5

In [138]:
```
toronto_merged.loc[toronto_merged['Cluster Labels'] == 4, toronto_m
erged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

Out[138]:

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th M Comm Ve
41	East Toronto	4	Greek Restaurant	Coffee Shop	Ice Cream Shop	Bookstore	Furniture / Home Store	Ita Restau
42	East Toronto	4	Sandwich Place	Park	Burger Joint	Italian Restaurant	Steakhouse	Fis Cl S
43	East Toronto	4	Café	Coffee Shop	Bakery	Italian Restaurant	American Restaurant	Sandv P
45	Central	4	Gym	Hotel	Sandwich	Park	Breakfast	Clotl

	Toronto					Place		Spot	S
46	Central Toronto	4	Clothing Store	Coffee Shop	Yoga Studio	Fast Food Restaurant	Pet Store	Chir Restau
47	Central Toronto	4	Sandwich Place	Dessert Shop	Pizza Place	Coffee Shop	Italian Restaurant	(
49	Central Toronto	4	Coffee Shop	Pub	Pizza Place	Sushi Restaurant	Bagel Shop	F Chic J
51	Downtown Toronto	4	Coffee Shop	Pub	Pizza Place	Restaurant	Café	Ma
52	Downtown Toronto	4	Coffee Shop	Japanese Restaurant	Gay Bar	Sushi Restaurant	Restaurant	Night
53	Downtown Toronto	4	Coffee Shop	Café	Park	Pub	Bakery	The
54	Downtown Toronto	4	Coffee Shop	Clothing Store	Cosmetics Shop	Café	Middle Eastern Restaurant	Restau
55	Downtown Toronto	4	Coffee Shop	Hotel	Café	Restaurant	Cosmetics Shop	Ita Restau
56	Downtown Toronto	4	Coffee Shop	Cocktail Bar	Bakery	Beer Bar	Steakhouse	Seaf Restau
57	Downtown Toronto	4	Coffee Shop	Café	Italian Restaurant	Burger Joint	Chinese Restaurant	Sandv Pl
58	Downtown Toronto	4	Coffee Shop	Café	Steakhouse	American Restaurant	Thai Restaurant	H
59	Downtown Toronto	4	Coffee Shop	Hotel	Aquarium	Café	Italian Restaurant	F Chic J
60	Downtown Toronto	4	Coffee Shop	Hotel	Café	Restaurant	Gastropub	Seaf Restau
61	Downtown Toronto	4	Coffee Shop	Café	Hotel	Restaurant	American Restaurant	Bal
65	Central Toronto	4	Coffee Shop	Sandwich Place	Café	Pizza Place	Pharmacy	Lic S
66	Downtown Toronto	4	Café	Bookstore	Restaurant	Bar	Japanese Restaurant	Bal
67	Downtown Toronto	4	Café	Vegetarian / Vegan Restaurant	Bar	Coffee Shop	Mexican Restaurant	Bal
69	Downtown Toronto	4	Coffee Shop	Restaurant	Café	Cocktail Bar	Seafood Restaurant	Beer
70	Downtown Toronto	4	Café	Coffee Shop	Hotel	Deli / Bodega	Bar	Restau
75	Downtown Toronto	4	Grocery Store	Café	Park	Baby Store	Athletics & Sports	Restau
	West					Music			Mic

76	Toronto	4	Pharmacy	Supermarket	Bakery	Venue	Pool	Eas Restau
77	West Toronto	4	Bar	Coffee Shop	Asian Restaurant	Bakery	Restaurant	C
78	West Toronto	4	Café	Breakfast Spot	Coffee Shop	Grocery Store	Bar	Caribb Restau
82	West Toronto	4	Mexican Restaurant	Café	Bar	Park	Bakery	D
83	West Toronto	4	Breakfast Spot	Gift Shop	Restaurant	Dessert Shop	Eastern European Restaurant	
84	West Toronto	4	Café	Pizza Place	Coffee Shop	Sushi Restaurant	Gym	Ita Restau
87	East Toronto	4	Light Rail Station	Yoga Studio	Auto Workshop	Garden Center	Garden	Fast F Restau

Conclusion

From the above results, wee see that the majority of popular venues are located in cluster 5. This cluster may be analyzed further in future.

In [ ]: