The Battle of Neighborhoods

In [1]:
```
import numpy as np
import pandas as pd
from bs4 import BeautifulSoup
import requests
```

In [73]:
```
!conda install -c conda-forge folium=0.5.0 --yes
import folium
```

```
Solving environment: done

## Package Plan ##

  environment location: /opt/conda/envs/DSX-Python35

  added / updated specs:
    - folium=0.5.0


The following packages will be downloaded:

    package                    |            build
    ---------------------------|-----------------
    folium-0.5.0               |             py_0          45 KB  conda-forge
    certifi-2018.8.24          |        py35_1001         139 KB  conda-forge
    altair-2.2.2               |           py35_1         462 KB  conda-forge
    vincent-0.4.4              |             py_1          28 KB  conda-forge
    ca-certificates-2019.6.16  |       hecc5488_0         145 KB  conda-forge
    openssl-1.0.2r             |       h14c3975_0         3.1 MB  conda-forge
    branca-0.3.1               |             py_0          25 KB  conda-forge
    ------------------------------------------------------------
                                           Total:         4.0 MB

The following NEW packages will be INSTALLED:

    altair:          2.2.2-py35_1      conda-forge
    branca:          0.3.1-py_0        conda-forge
    folium:          0.5.0-py_0        conda-forge
    vincent:         0.4.4-py_1        conda-forge

The following packages will be UPDATED:

    ca-certificates: 2019.1.23-0                 --> 2019.6.16-hecc5488_0 conda-forge
    certifi:         2018.8.24-py35_1            --> 2018.8.24-py35_1001  conda-forge

The following packages will be DOWNGRADED:

    openssl:         1.0.2s-h7b6447c_0           --> 1.0.2r-h14c3975_0    conda-forge


Downloading and Extracting Packages
folium-0.5.0         | 45 KB     | ##################################### | 100%
certifi-2018.8.24    | 139 KB    | ##################################### | 100%
altair-2.2.2         | 462 KB    | ##################################### | 100%
vincent-0.4.4        | 28 KB     | ##################################### | 100%
ca-certificates-2019 | 145 KB    | ##################################### | 100%
openssl-1.0.2r       | 3.1 MB    | ##################################### | 100%
branca-0.3.1         | 25 KB     | ##################################### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
```

For this assignment, I need to explore and cluster the neighborhoods in Toronto.

For the Toronto neighborhood data, a Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto. I need to scrape the Wikipedia page and wrangle the data, clean it, and then read it into a pandas dataframe so that it is in a structured format .

Here is my code to scrape the Wikipedia page and transform the data in the table into a pandas dataframe:

In [2]:
```
url = 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
```

In [3]:
```
response = requests.get(url)
response.status_code
```

Out[3]: 200

In [4]:
```
soup = BeautifulSoup(response.content, 'html.parser')
```

In [5]:
```
stat_table = soup.find_all('table', class_ = 'wikitable sortable')

len(stat_table)
```

Out[5]: 1

In [6]:
```
stat_table = stat_table[0]
data =[]


for row in stat_table.find_all('tr'):
    col = row.find_all('td')
    if len(col) == 3:
     #   print("...")
        data.append((col[0].text.strip(), col[1].text.strip(), col[2].text.strip()))
```

```
df = pd.DataFrame(data, columns=["PostalCode", "Borough", "Neighborhood"])
df.head(10)
```

Out[6]:

	PostalCode	Borough	Neighborhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Harbourfront
5	M5A	Downtown Toronto	Regent Park
6	M6A	North York	Lawrence Heights
7	M6A	North York	Lawrence Manor
8	M7A	Queen's Park	Not assigned
9	M8A	Not assigned	Not assigned

In [7]:
```
print(df.shape)
#df.head()
```

```
(288, 3)
```

Data Preparation

The dataframe consist of three columns: PostalCode, Borough, and Neighborhood

Only process the cells that have an assigned borough. Ignore cells with a borough that is Not assigned.

If a cell has a borough but a Not assigned neighborhood, then the neighborhood will be the same as the borough. So for the 9th cell in the table on the Wikipedia page, the value of the Borough and the Neighborhood columns will be Queen's Park.

More than one neighborhood can exist in one postal code area. For example, in the table on the Wikipedia page, you will notice that M5A is listed twice and has two neighborhoods: Harbourfront and Regent Park. These two rows will be combined into one row with the neighborhoods separated with a comma.

In [7]:
```
real =[]

for row in stat_table.find_all('tr'):
    col = row.find_all('td')
    if len(col) == 3:
        string = col[1].text.strip()

        if (string != 'Not assigned'):
            neigh = col[2].text.strip()
            if (neigh == 'Not assigned'):

                real.append((col[0].text.strip(), col[1].text.strip(), col[1].text.strip()))
            else :
                real.append((col[0].text.strip(), col[1].text.strip(), col[2].text.strip()))


df = pd.DataFrame(real, columns=["PostalCode", "Borough", "Neighborhood"])
df.head(11)
```

Out[7]:

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Harbourfront
3	M5A	Downtown Toronto	Regent Park
4	M6A	North York	Lawrence Heights
5	M6A	North York	Lawrence Manor
6	M7A	Queen's Park	Queen's Park
7	M9A	Etobicoke	Islington Avenue
8	M1B	Scarborough	Rouge
9	M1B	Scarborough	Malvern
10	M3B	North York	Don Mills North

In [8]:
```
df = df.groupby(["PostalCode", "Borough"], as_index=False).agg(lambda x: ", ".join(x))
df.head()
```

Out[8]:

	PostalCode	Borough	Neighborhood
0	M1B	Scarborough	Rouge, Malvern
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

Now that I have built a dataframe of the postal code of each neighborhood along with the borough name and neighborhood name, in order to utilize the Foursquare location data, we need to get the latitude and the longitude coordinates of each neighborhood.

In [9]:
```
path="http://cocl.us/Geospatial_data"
location_coordinates=pd.read_csv(path)
location_coordinates.head()
```

Out[9]:

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353

1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

In [13]:
```
location_coordinates.shape
```

Out[13]:
```
(103, 3)
```

In [10]:
```
location_coordinates.rename(columns={"Postal Code": "PostalCode"}, inplace=True)

canada_data = df.merge(location_coordinates, on="PostalCode", how="left")
canada_data.head(10)
```

Out[10]:

	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
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848

Data Modeling

Explore the neighborhoods in Toronto. I decided to work with only boroughs that contain the word Toronto

In [12]:
```
toronto_data = canada_data[canada_data['Borough'].str.contains('Toronto',)]
toronto_data
```

Out[12]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
37	M4E	East Toronto	The Beaches	43.676357	-79.293031
41	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188
42	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572
43	M4M	East Toronto	Studio District	43.659526	-79.340923
44	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
45	M4P	Central Toronto	Davisville North	43.712751	-79.390197
46	M4R	Central Toronto	North Toronto West	43.715383	-79.405678
47	M4S	Central Toronto	Davisville	43.704324	-79.388790
48	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160
49	M4V	Central Toronto	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049
50	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
51	M4X	Downtown Toronto	Cabbagetown, St. James Town	43.667967	-79.367675
52	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
53	M5A	Downtown Toronto	Harbourfront, Regent Park	43.654260	-79.360636
54	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937
55	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
56	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
57	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
58	M5H	Downtown Toronto	Adelaide, King, Richmond	43.650571	-79.384568
59	M5J	Downtown Toronto	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752
60	M5K	Downtown Toronto	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576
61	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817
63	M5N	Central Toronto	Roselawn	43.711695	-79.416936
64	M5P	Central Toronto	Forest Hill North, Forest Hill West	43.696948	-79.411307
65	M5R	Central Toronto	The Annex, North Midtown, Yorkville	43.672710	-79.405678
66	M5S	Downtown Toronto	Harbord, University of Toronto	43.662696	-79.400049
67	M5T	Downtown Toronto	Chinatown, Grange Park, Kensington Market	43.653206	-79.400049
68	M5V	Downtown Toronto	CN Tower, Bathurst Quay, Island airport, Harbo...	43.628947	-79.394420
69	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846
70	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280
75	M6G	Downtown Toronto	Christie	43.669542	-79.422564
76	M6H	West Toronto	Dovercourt Village, Dufferin	43.669005	-79.442259
77	M6J	West Toronto	Little Portugal, Trinity	43.647927	-79.419750
78	M6K	West Toronto	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
82	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763

83	M6R	West Toronto	Parkdale, Roncesvalles	43.648960	-79.456325
84	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450
87	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558

In [43]:
```
toronto_data.shape
```

Out[43]:
```
(38, 5)
```

let's visualize the city of Toronto and their neighborhoods with some pop-up text that would get displayed when you hover over a marker

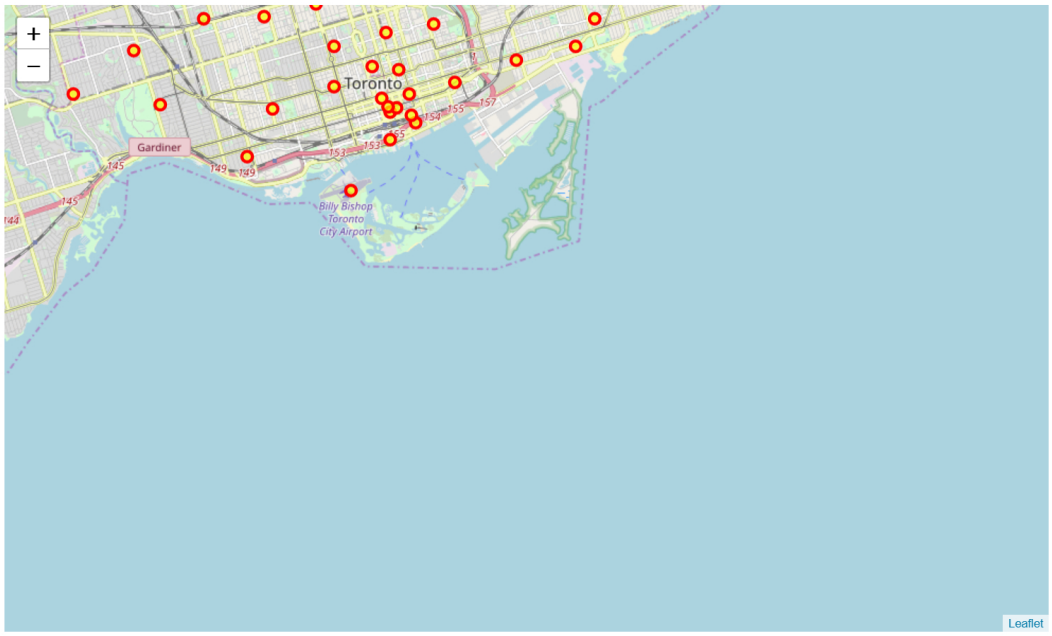
In [75]:
```
latitude = 43.6532
longitude = -79.3872

map_toronto = folium.Map(location=[latitude, longitude], zoom_start=10)

for lat, lng, borough, neighborhood in zip(toronto_data['Latitude'],toronto_data['Longitude'], toronto_data['Borough'],
                                           toronto_data['Neighborhood']):
    label = '{}, {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
       [lat, lng],   radius=5,
         popup=label,
        color='red',
        fill=True,
        fill_color='yellow',
        fill_opacity=0.7,
        parse_html=False).add_to(map_toronto)

map_toronto
```

Out[75]:



USING FOURSQUARE

Define Foursquare Credentials and Version

In [ ]:
```
```

In [13]:
```
# 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

import json
from pandas.io.json import json_normalize

CLIENT_ID = 'SEPJDVINYQAUSUGAT4I1CJUJFDMDMMOBOM1K51P4CMUGXN4D' # your Foursquare ID
CLIENT_SECRET = 'UBTN3RGQRXY4KZRK1MLAT4HNW5RTHGVXAFKGDUVXO0Q3MI1C' # your Foursquare Secret
VERSION = '20180605'
search_query = 'Indian'
radius = 1000
LIMIT = 50
```

In [14]:
```
# 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 [21]:
```
def getNearbyVenues(names, latitudes, longitudes):
```

```
    collective =pd.DataFrame(columns = ['name','categories','distance', 'lat','lng', 'neigh', 'neigh-lat', 'neigh-long'])



    for name, latitude, longitude in zip(names, latitudes, longitudes):

        list=[]
        i = 0
        print(name)

        #if name == 'Roselawn' or name == 'Forest Hill North, Forest Hill West':
         #   continue

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit=
{}'.format(CLIENT_ID, CLIENT_SECRET, latitude, longitude, VERSION, search_query, radius, LIMIT)

        # make the GET request
        results = requests.get(url).json()['response']['venues']
        temp = json_normalize(results)

        if len(results) == 0:
            continue

        if i ==  0:
            filtered_columns = ['name', 'categories'] + [col for col in temp.columns if col.startswith('location.')] + ['id']
            dataframe_filtered = temp.loc[:, filtered_columns]
            i = i+1
       # filter the category for each row

        dataframe_filtered['categories']= dataframe_filtered.apply(get_category_type, axis=1)

# clean column names by keeping only last term
        dataframe_filtered.columns = [column.split('.')[-1] for column in dataframe_filtered.columns]

        x = dataframe_filtered[['name','categories','distance', 'lat','lng']]

        counting = len(dataframe_filtered)
        print("coount =",counting)

        value=0

        while(value != counting):
            list.append((name, latitude, longitude))
            value = value+1

        frame = pd.DataFrame(list, columns=['neigh', 'neigh-lat', 'neigh-long'])
        x = x.join(frame)
        collective = x.append(collective, ignore_index=False)

        #print(collective.dtypes)
        #print(collective,"\n going on...")

    return(collective)
```

In [22]:

```
#print(toronto_data.head())
LIMIT = 100
toronto_venues = getNearbyVenues(names=toronto_data['Neighborhood'],
                                 latitudes =toronto_data['Latitude'],
                                 longitudes =toronto_data['Longitude']
                                )
#toronto_venues
```

```
The Beaches
The Danforth West, Riverdale
The Beaches West, India Bazaar
coount = 6
Studio District
coount = 1
Lawrence Park
Davisville North
coount = 3
North Toronto West
coount = 1
Davisville
coount = 5
Moore Park, Summerhill East
coount = 2
Deer Park, Forest Hill SE, Rathnelly, South Hill, Summerhill West
coount = 3
Rosedale
coount = 2
Cabbagetown, St. James Town
coount = 2
Church and Wellesley
coount = 8
Harbourfront, Regent Park
coount = 5
Ryerson, Garden District
coount = 13
St. James Town
coount = 15
Berczy Park
coount = 10
Central Bay Street
coount = 12
Adelaide, King, Richmond
coount = 21
Harbourfront East, Toronto Islands, Union Station
coount = 12
Design Exchange, Toronto Dominion Centre
coount = 19
Commerce Court, Victoria Hotel
coount = 19
Roselawn
coount = 1
Forest Hill North, Forest Hill West
coount = 1
The Annex, North Midtown, Yorkville
coount = 3
Harbord, University of Toronto
```

```
coount = 5
Chinatown, Grange Park, Kensington Market
coount = 8
CN Tower, Bathurst Quay, Island airport, Harbourfront West, King and Spadina, Railway Lands, South Niagara
Stn A PO Boxes 25 The Esplanade
coount = 15
First Canadian Place, Underground city
coount = 19
Christie
coount = 6
Dovercourt Village, Dufferin
Little Portugal, Trinity
coount = 3
Brockton, Exhibition Place, Parkdale Village
coount = 4
High Park, The Junction South
coount = 9
Parkdale, Roncesvalles
coount = 6
Runnymede, Swansea
coount = 2
Business Reply Mail Processing Centre 969 Eastern
coount = 4
```

In [60]:
```
toronto_venues
```

In [19]:
```
toronto_venues.shape
```

Out[19]: (245, 8)

In [29]:
```
toronto_venues.set_index('distance', inplace=True)
toronto_venues.reset_index(inplace=True)
toronto_venues
```

Out[29]:

	distance	name	categories	lat	lng	neigh	neigh-lat	neigh-long
0	1020	Indian Record Shop	Record Shop	43.671905	-79.321990	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558
1	1050	Indian Rasoi	Indian Restaurant	43.672086	-79.323403	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558
2	1026	Little India Neighbourhood	Neighborhood	43.671918	-79.322837	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558
3	1050	Gerrard India Bazaar	Shopping Plaza	43.672086	-79.323403	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558
4	298	Durbar Indian Cuisine	Indian Restaurant	43.648903	-79.484795	Runnymede, Swansea	43.651571	-79.484450
5	593	Bukhara indian cuisine	Indian Restaurant	43.651105	-79.477104	Runnymede, Swansea	43.651571	-79.484450
6	739	Bloor St. & Indian Rd.	Road	43.655601	-79.456300	Parkdale, Roncesvalles	43.648960	-79.456325
7	561	Indian Mound Traffic Island	Park	43.653977	-79.457034	Parkdale, Roncesvalles	43.648960	-79.456325
8	381	Indian road	Road	43.652362	-79.455718	Parkdale, Roncesvalles	43.648960	-79.456325
9	985	Indian Grove	Road	43.657400	-79.459998	Parkdale, Roncesvalles	43.648960	-79.456325
10	806	Indian Road Settlement House	Building	43.641902	-79.454078	Parkdale, Roncesvalles	43.648960	-79.456325
11	698	Scotiabank	Bank	43.654955	-79.458877	Parkdale, Roncesvalles	43.648960	-79.456325
12	426	Indian Road Crescent Public School	School	43.662491	-79.459608	High Park, The Junction South	43.661608	-79.464763
13	954	Bloor St. & Indian Rd.	Road	43.655601	-79.456300	High Park, The Junction South	43.661608	-79.464763
14	802	Leela Indian Food Bar	Indian Restaurant	43.665326	-79.473306	High Park, The Junction South	43.661608	-79.464763
15	605	Indian Grove	Road	43.657400	-79.459998	High Park, The Junction South	43.661608	-79.464763
16	391	503 Indian Grove	None	43.664610	-79.462226	High Park, The Junction South	43.661608	-79.464763
17	1053	Indian Mound Traffic Island	Park	43.653977	-79.457034	High Park, The Junction South	43.661608	-79.464763
18	545	594 Indian Grove	Building	43.666318	-79.462915	High Park, The Junction South	43.661608	-79.464763
19	1261	Indian road	Road	43.652362	-79.455718	High Park, The Junction South	43.661608	-79.464763
20	879	Scotiabank	Bank	43.654955	-79.458877	High Park, The Junction South	43.661608	-79.464763
21	758	Ali's West Indian Roti Shop	Caribbean Restaurant	43.640745	-79.435913	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
22	612	indian ink tattoo	Tattoo Parlor	43.642218	-79.426556	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
23	967	Maurya East Indian Roti Cuisine	Indian Restaurant	43.638825	-79.416494	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
24	702	A&N Canadian & West Indian Cuisine	Caribbean Restaurant	43.641275	-79.434397	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
25	153	Indian Grill	Indian Restaurant	43.646555	-79.419874	Little Portugal, Trinity	43.647927	-79.419750
26	839	indian ink tattoo	Tattoo Parlor	43.642218	-79.426556	Little Portugal, Trinity	43.647927	-79.419750
27	1046	Maurya East Indian Roti Cuisine	Indian Restaurant	43.638825	-79.416494	Little Portugal, Trinity	43.647927	-79.419750
28	739	Banjara Indian Cuisine	Indian Restaurant	43.662916	-79.421911	Christie	43.669542	-79.422564
29	865	630 Maroli Indian Kerala Restaurant	Restaurant	43.664442	-79.414456	Christie	43.669542	-79.422564
...	...	...	...	...	...	...	...	...
215	867	Joe's Indian Restaurant	Indian Restaurant	43.658150	-79.381563	Church and Wellesley	43.665860	-79.383160
216	771	Maja Indian Cuisine	Indian Restaurant	43.671986	-79.378685	Church and Wellesley	43.665860	-79.383160
217	903	Bhoj Indian Cuisine	Indian Restaurant	43.672765	-79.389063	Church and Wellesley	43.665860	-79.383160
218	1011	Mami's Indian Cuisine	Food Truck	43.656986	-79.385840	Church and Wellesley	43.665860	-79.383160

219	762	Consulate General of India in Toronto	Embassy / Consulate	43.671869	-79.378621	Church and Wellesley	43.665860	-79.383160
220	993	Maja Indian Cuisine	Indian Restaurant	43.671986	-79.378685	Cabbagetown, St. James Town	43.667967	-79.367675
221	982	Consulate General of India in Toronto	Embassy / Consulate	43.671869	-79.378621	Cabbagetown, St. James Town	43.667967	-79.367675
222	848	Maja Indian Cuisine	Indian Restaurant	43.671986	-79.378685	Rosedale	43.679563	-79.377529
223	860	Consulate General of India in Toronto	Embassy / Consulate	43.671869	-79.378621	Rosedale	43.679563	-79.377529
224	682	Indian Affairs	Capitol Building	43.688585	-79.392120	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049
225	1316	Rajput's Indian Bistro	None	43.674801	-79.396966	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049
226	525	Chef of India	Indian Restaurant	43.687391	-79.393663	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049
227	729	Indian Affairs	Capitol Building	43.688585	-79.392120	Moore Park, Summerhill East	43.689574	-79.383160
228	879	Chef of India	Indian Restaurant	43.687391	-79.393663	Moore Park, Summerhill East	43.689574	-79.383160
229	531	Banjara Indian Cuisine	Indian Restaurant	43.707810	-79.393296	Davisville	43.704324	-79.388790
230	172	Debu's Nouvelle Indian Cuisine	None	43.702893	-79.387958	Davisville	43.704324	-79.388790
231	172	Marigold Indian Bistro	Indian Restaurant	43.702881	-79.388008	Davisville	43.704324	-79.388790
232	1102	Indian Street Food Co.	Indian Restaurant	43.708032	-79.376086	Davisville	43.704324	-79.388790
233	1104	Eat Indian By Amaya	Indian Restaurant	43.705592	-79.375178	Davisville	43.704324	-79.388790
234	1305	Banjara Indian Cuisine	Indian Restaurant	43.707810	-79.393296	North Toronto West	43.715383	-79.405678
235	603	Banjara Indian Cuisine	Indian Restaurant	43.707810	-79.393296	Davisville North	43.712751	-79.390197
236	1112	Debu's Nouvelle Indian Cuisine	None	43.702893	-79.387958	Davisville North	43.712751	-79.390197
237	1112	Marigold Indian Bistro	Indian Restaurant	43.702881	-79.388008	Davisville North	43.712751	-79.390197
238	647	Siddhartha Indian	Indian Restaurant	43.659117	-79.348948	Studio District	43.659526	-79.340923
239	609	Indian Record Shop	Record Shop	43.671905	-79.321990	The Beaches West, India Bazaar	43.668999	-79.315572
240	416	Sidhartha Indian Resturaunt	Indian Restaurant	43.672695	-79.316339	The Beaches West, India Bazaar	43.668999	-79.315572
241	718	Indian Rasoi	Indian Restaurant	43.672086	-79.323403	The Beaches West, India Bazaar	43.668999	-79.315572
242	633	The Famous Indian Restaurant	Indian Restaurant	43.672339	-79.321941	The Beaches West, India Bazaar	43.668999	-79.315572
243	669	Little India Neighbourhood	Neighborhood	43.671918	-79.322837	The Beaches West, India Bazaar	43.668999	-79.315572
244	718	Gerrard India Bazaar	Shopping Plaza	43.672086	-79.323403	The Beaches West, India Bazaar	43.668999	-79.315572

245 rows × 8 columns

In [30]:
```
toronto_onehot = pd.get_dummies(toronto_venues[['categories']], prefix="", prefix_sep="")

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

# 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[30]:

	Vegetarian / Vegan Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Fast Food Restaurant	Food Truck	Indian Restaurant	Neighborhood	Office	Park	Record Shop	Re
0	0	0	0	0	0	0	0	0	0	0	Business Reply Mail Processing Centre 969 Eastern	0	0	1	0
1	0	0	0	0	0	0	0	0	0	1	Business Reply Mail Processing Centre 969 Eastern	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	Business Reply Mail Processing Centre 969 Eastern	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	Business Reply Mail Processing Centre 969 Eastern	0	0	0	0
4	0	0	0	0	0	0	0	0	0	1	Runnymede, Swansea	0	0	0	0

In [ ]:
```
```

In [31]:
```
toronto_onehot.shape
```

Out[31]: (245, 19)

In [32]:
```
toronto_grouped = toronto_onehot.groupby('Neighborhood').mean().reset_index()
```

```
toronto_grouped
```

Out[32]:

	Neighborhood	Vegetarian / Vegan Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Fast Food Restaurant	Food Truck	Indian Restaurant	Office	P
0	Adelaide, King, Richmond	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.047619	0.047619	0.809524	0.047619	0.000
1	Berczy Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.800000	0.000000	0.000
2	Brockton, Exhibition Place, Parkdale Village	0.000000	0.000000	0.000000	0.000000	0.000000	0.5	0.000	0.000000	0.000000	0.250000	0.000000	0.000
3	Business Reply Mail Processing Centre 969 Eastern	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.250000	0.000000	0.000
4	Cabbagetown, St. James Town	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.500	0.000000	0.000000	0.500000	0.000000	0.000
5	Central Bay Street	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.083333	0.833333	0.083333	0.000
6	Chinatown, Grange Park, Kensington Market	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.875000	0.125000	0.000
7	Christie	0.166667	0.166667	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.500000	0.000000	0.000
8	Church and Wellesley	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.125	0.000000	0.125000	0.750000	0.000000	0.000
9	Commerce Court, Victoria Hotel	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.052632	0.052632	0.842105	0.000000	0.000
10	Davisville	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.800000	0.000000	0.000
11	Davisville North	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.666667	0.000000	0.000
12	Deer Park, Forest Hill SE, Rathnelly, South Hi...	0.000000	0.000000	0.000000	0.000000	0.333333	0.0	0.000	0.000000	0.000000	0.333333	0.000000	0.000
13	Design Exchange, Toronto Dominion Centre	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.052632	0.052632	0.789474	0.052632	0.000
14	First Canadian Place, Underground city	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.052632	0.052632	0.789474	0.052632	0.000
15	Forest Hill North, Forest Hill West	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	1.000000	0.000000	0.000
16	Harbord, University of Toronto	0.000000	0.200000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.600000	0.200000	0.000
17	Harbourfront East, Toronto Islands, Union Station	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.083333	0.000000	0.833333	0.000000	0.000
18	Harbourfront, Regent Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.600000	0.000000	0.000
19	High Park, The Junction South	0.000000	0.000000	0.111111	0.111111	0.000000	0.0	0.000	0.000000	0.000000	0.111111	0.000000	0.111
20	Little Portugal, Trinity	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.666667	0.000000	0.000
21	Moore Park, Summerhill East	0.000000	0.000000	0.000000	0.000000	0.500000	0.0	0.000	0.000000	0.000000	0.500000	0.000000	0.000
22	North Toronto West	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	1.000000	0.000000	0.000
23	Parkdale, Roncesvalles	0.000000	0.000000	0.166667	0.166667	0.000000	0.0	0.000	0.000000	0.000000	0.000000	0.000000	0.166
24	Rosedale	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.500	0.000000	0.000000	0.500000	0.000000	0.000
25	Roselawn	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	1.000000	0.000000	0.000
26	Runnymede, Swansea	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	1.000000	0.000000	0.000
27	Ryerson, Garden District	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.076923	0.923077	0.000000	0.000
28	St. James Town	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.066667	0.733333	0.000000	0.000
	Stn A PO												

29	Boxes 25 The Esplanade	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.800000	0.000000	0.000
30	Studio District	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	1.000000	0.000000	0.000
31	The Annex, North Midtown, Yorkville	0.333333	0.333333	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.000000	0.000000	0.000
32	The Beaches West, India Bazaar	0.000000	0.000000	0.000000	0.000000	0.000000	0.0	0.000	0.000000	0.000000	0.500000	0.000000	0.000

In [ ]:
```
Top 10 venues
```

In [35]:
```
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 [36]:
```
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)

neighborhoods_venues_sorted.head()
```

Out[36]:

	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	Indian Restaurant	Food Truck	Fast Food Restaurant	Office	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
1	Berczy Park	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
2	Brockton, Exhibition Place, Parkdale Village	Caribbean Restaurant	Tattoo Parlor	Indian Restaurant	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Embassy / Consulate	Food Truck
3	Business Reply Mail Processing Centre 969 Eastern	Record Shop	Indian Restaurant	Shopping Plaza	Tattoo Parlor	Embassy / Consulate	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant
4	Cabbagetown, St. James Town	Embassy / Consulate	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Food Truck

Kmeans cluster analysis

In [ ]:
```
```

In [93]:
```
#CLUSTERING THE 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_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[20:30]
```

Out[93]:
```
array([1, 2, 1, 0, 4, 1, 1, 1, 1, 1], dtype=int32)
```

Adding cluster lables and printing

In [94]:
```
#neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

toronto_merged = toronto_data.drop([37,41,44,68,76], axis=0)

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

toronto_merged.head() # check the last columns!
```

Out[94]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Commc Venu
42	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572	2	Indian Restaurant	Record Shop	Shopping Plaza	Tattoo Parlor	Embassy / Consulate	Astrologer	Bank
43	M4M	East Toronto	Studio District	43.659526	-79.340923	1	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building

45	M4P	Central Toronto	Davisville North	43.712751	-79.390197	1	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building
46	M4R	Central Toronto	North Toronto West	43.715383	-79.405678	1	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building
47	M4S	Central Toronto	Davisville	43.704324	-79.388790	1	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building

In [ ]:

In [ ]:

In [ ]:
```
Visualising Toronto map with all the clusters
```

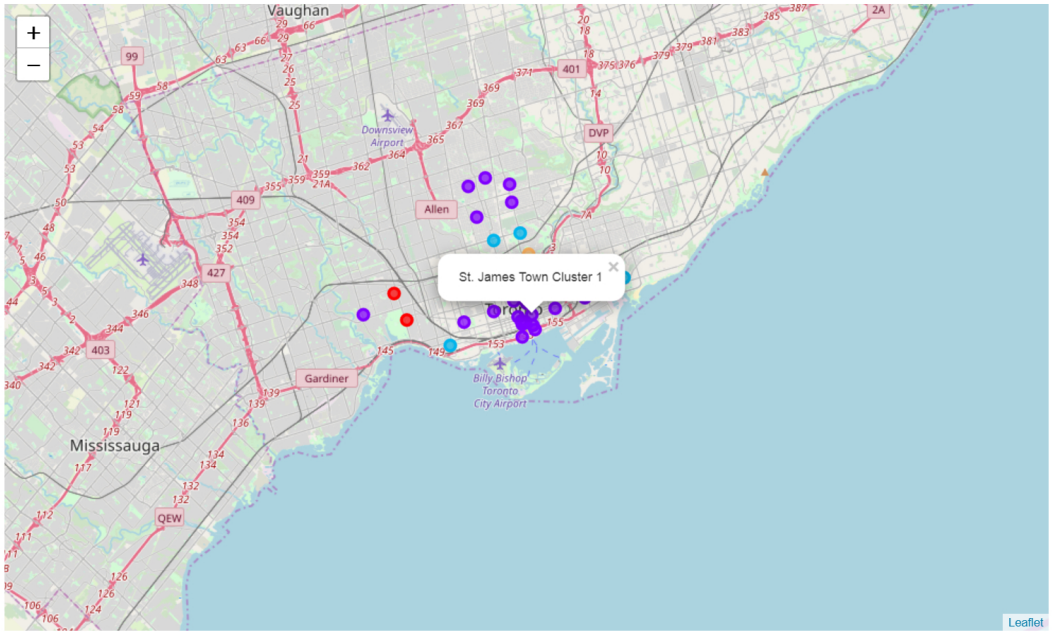
In [95]:
```
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# 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(toronto_merged['Latitude'], toronto_merged['Longitude'], toronto_merged['Neighborhood'], toron
to_merged['Cluster Labels']):
    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[95]:



Examine the clusters

Now let us examine the clusters and see how they differ from each other in terms of popular venues.

In [42]:
```
cluster_0 = toronto_merged.loc[toronto_merged['Cluster Labels'] == 0,
                               toronto_merged.columns[
                                   [2] + list(range(
                                       5, toronto_merged.shape[1]))]]
cluster_0
```

Out[42]:

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
82	High Park, The Junction South	0.0	Road	School	Bank	Park	Building	Indian Restaurant	Tattoo Parlor	Embassy / Consulate	Astrologer	Capitol Building
83	Parkdale, Roncesvalles	0.0	Road	Bank	Building	Park	Tattoo Parlor	Fast Food Restaurant	Astrologer	Capitol Building	Caribbean Restaurant	Embassy / Consulate

In [47]:
```
cluster_0.shape
```

Out[47]:
```
(2, 12)
```

In [43]:
```
cluster_1 =toronto_merged.loc[toronto_merged['Cluster Labels'] == 1,
                              toronto_merged.columns[
```

```
                        [2] + list(range(5, toronto_merged.shape[1]))]]
cluster_1
```

Out[43]:

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
43	Studio District	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
45	Davisville North	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
46	North Toronto West	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
47	Davisville	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
52	Church and Wellesley	1.0	Indian Restaurant	Food Truck	Embassy / Consulate	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Tattoo Parlor
53	Harbourfront, Regent Park	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
54	Ryerson, Garden District	1.0	Indian Restaurant	Food Truck	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Tattoo Parlor
55	St. James Town	1.0	Indian Restaurant	Food Truck	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Tattoo Parlor
56	Berczy Park	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
57	Central Bay Street	1.0	Indian Restaurant	Food Truck	Office	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
58	Adelaide, King, Richmond	1.0	Indian Restaurant	Food Truck	Fast Food Restaurant	Office	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
59	Harbourfront East, Toronto Islands, Union Station	1.0	Indian Restaurant	Fast Food Restaurant	Tattoo Parlor	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
60	Design Exchange, Toronto Dominion Centre	1.0	Indian Restaurant	Food Truck	Fast Food Restaurant	Office	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
61	Commerce Court, Victoria Hotel	1.0	Indian Restaurant	Food Truck	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Tattoo Parlor
63	Roselawn	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
64	Forest Hill North, Forest Hill West	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
66	Harbord, University of Toronto	1.0	Indian Restaurant	Astrologer	Office	Tattoo Parlor	Fast Food Restaurant	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
67	Chinatown, Grange Park, Kensington Market	1.0	Indian Restaurant	Office	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate

69	Stn A PO Boxes 25 The Esplanade	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
70	First Canadian Place, Underground city	1.0	Indian Restaurant	Food Truck	Fast Food Restaurant	Office	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
77	Little Portugal, Trinity	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
84	Runnymede, Swansea	1.0	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Food Truck

In [48]:
```
cluster_1.shape
```

Out[48]: (22, 12)

In [44]:
```
cluster_2 = toronto_merged.loc[toronto_merged['Cluster Labels'] == 2,
                  toronto_merged.columns[
                      [2] + list(range(5, toronto_merged.shape[1]))]]
cluster_2
```

Out[44]:

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
42	The Beaches West, India Bazaar	2.0	Indian Restaurant	Record Shop	Shopping Plaza	Tattoo Parlor	Embassy / Consulate	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant
48	Moore Park, Summerhill East	2.0	Capitol Building	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
49	Deer Park, Forest Hill SE, Rathnelly, South Hi...	2.0	Capitol Building	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Caribbean Restaurant	Embassy / Consulate	Food Truck
75	Christie	2.0	Indian Restaurant	Vegetarian / Vegan Restaurant	Astrologer	Restaurant	Fast Food Restaurant	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate
78	Brockton, Exhibition Place, Parkdale Village	2.0	Caribbean Restaurant	Tattoo Parlor	Indian Restaurant	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Embassy / Consulate	Food Truck
87	Business Reply Mail Processing Centre 969 Eastern	2.0	Record Shop	Indian Restaurant	Shopping Plaza	Tattoo Parlor	Embassy / Consulate	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant

In [49]:
```
cluster_2.shape
```

Out[49]: (6, 12)

In [50]:
```
cluster_3 =toronto_merged.loc[toronto_merged['Cluster Labels'] == 3,
                  toronto_merged.columns[
                      [2] + list(range(5, toronto_merged.shape[1]))]]
cluster_3
```

Out[50]:

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
65	The Annex, North Midtown, Yorkville	3.0	Vegetarian / Vegan Restaurant	Astrologer	Shopping Plaza	Bank	Building	Capitol Building	Caribbean Restaurant	Embassy / Consulate	Fast Food Restaurant	Tattoo Parlor

In [51]:
```
cluster_3.shape
```

Out[51]: (1, 12)

In [46]:
```
cluster_4 = toronto_merged.loc[toronto_merged['Cluster Labels'] == 4,
                  toronto_merged.columns[
                      [2] + list(range(5, toronto_merged.shape[1]))]]
cluster_4
```

Out[46]:

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
50	Rosedale	4.0	Embassy / Consulate	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Food Truck
51	Cabbagetown, St. James Town	4.0	Embassy / Consulate	Indian Restaurant	Tattoo Parlor	Fast Food Restaurant	Astrologer	Bank	Building	Capitol Building	Caribbean Restaurant	Food Truck

In [52]:
```
cluster_4.shape
```

Out[52]: (2, 12)

The best Community in the neighborhood

In [69]:
```
neighborhood_live = toronto_venues[toronto_venues['name'].str.contains('Indian Bazaar')]

neighborhood_live
```

Out[69]:

	distance	name	categories	lat	lng	neigh	neigh-lat	neigh-long
63	1339	Indian Bazaar	Neighborhood	43.655653	-79.364153	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846
186	1018	Indian Bazaar	Neighborhood	43.655653	-79.364153	St. James Town	43.651494	-79.375418
207	322	Indian Bazaar	Neighborhood	43.655653	-79.364153	Harbourfront, Regent Park	43.654260	-79.360636

In [70]:
```
neighborhood_live2 =toronto_venues[toronto_venues['name'].str.contains('Little India')]
neighborhood_live2
```

Out[70]:

	distance	name	categories	lat	lng	neigh	neigh-lat	neigh-long
2	1026	Little India Neighbourhood	Neighborhood	43.671918	-79.322837	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558
243	669	Little India Neighbourhood	Neighborhood	43.671918	-79.322837	The Beaches West, India Bazaar	43.668999	-79.315572

There are 22 neighborhoods in Cluster_1. It seems that Indian Restaurants, Fast Food Resturants are very popular in this neighborhood and in general, in Toronto. Also to be noted there are Corporate Buildings like Bank, Capitol Building, Embassy/Consulate office, in this Cluster_1. in Toronto.

The Best Neighborhoods To Live In Toronto

St.James Town, Harbourfront, Regent Park, The Esplanade are

the Neighborhoods where most of the Indian community lives. With help of Kmeans analysis, It is observed that they are grouped in Cluster_1 where other amenities like Indian Restaurents, Fast food Restaurant, Food Truck, Tatoo Parlor, and Corporate offices like Bank, Embassy, Capitol Buildings, Astrologer are well established for a comfortable stay in Toronto.

Less densly populated Indian commuties live in the neighborhoods are

Business Reply Mail Processing Centre 969 Eastern , The Beaches West, Indian Bazaar

They are grouped under Cluster_2 where other amenities like Indian Restaurant, Record Shop, Shopping Plaza, Tattoo Parlor, Embassy / Consulate Office, Astrologer, Bank and Capitol Buildings are available.

In [ ]:
```

```