

Battle of the Neighborhoods - Coursera Capstone Project

Introduction

This project will analyze neighborhoods between Toronto, Canada and New York City, New York. A Fortune 500 company is looking to move its headquarters to either Toronto or New York City. The company wants insight into the neighborhoods and local businesses in the cities so that its employees may have the optimum living standards and quality of life. This project will explore the similarities and dissimilarities between certain neighborhoods in the two cities, and determine which neighborhoods best fit the culture of the Fortune 500 company's employees.

Segmenting and Clustering Neighborhoods in Toronto

Use data from Toronto Neighborhood Wikipedia page to segment, cluster and explore neighborhoods in Toronto

```
In [1]: import numpy as np # Library to handle data in a vectorized manner

import pandas as pd # Library for data analysis
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 latitude and longitude values

import requests # Library to handle requests
from pandas.io.json import json_normalize # transform 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.')
```

Fetching package metadata

Solving package specifications: .

All requested packages already installed.

packages in environment at /opt/conda/envs/DSX-Python35:

#

geopy	1.18.1	py_0	conda-forge
-------	--------	------	-------------

Fetching package metadata

Solving package specifications: .

All requested packages already installed.

packages in environment at /opt/conda/envs/DSX-Python35:

#

folium	0.5.0	py_0	conda-forge
--------	-------	------	-------------

Libraries imported.

Download and Explore Toronto Dataset

The dataset being used is found at https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M).

The dataset is a list of Toronto's zipcodes which includes the boroughs and neighborhood names.

```
In [2]: #Obtain Postal Code, Borough, and Neighborhood information from Wikipedia
table = pd.read_html('https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M', header = 0)

#Obtain the first table
df_toronto = table[0]
df_toronto.head()
```

Out[2]:

	Postcode	Borough	Neighbourhood
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

Transform the data

```
In [3]: df_toronto.rename(columns = {"Postcode": "PostalCode", "Neighbourhood": "Neighborhood"}, inplace = True)

#Only process the cells that have an assigned borough. Ignore cells with a borough that is Not assigned.
df_toronto.drop(df_toronto[df_toronto.Borough == 'Not assigned'].index, inplace=True)
#df.head()

#Combine the neighborhoods that exists in one postal code
df_toronto = df_toronto.groupby(['PostalCode', 'Borough'])['Neighborhood'].apply(lambda x: ','.join(x)).reset_index()
#df.head()

#Change unassigned Neighborhood to its Borough's name
df_toronto.loc[85, 'Neighborhood'] = 'Queen\'s Park'

print (df_toronto.shape)

df_toronto.head()
```

(103, 3)

Out[3]:

	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

Join neighborhood table with latitude and longitude information

```
In [4]: #Create a dataframe of the Latitude and Longitudes of the Toronto Neighborhoods
latlong = pd.read_csv("http://cocl.us/Geospatial_data")
latlong.head()
```

Out[4]:

	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 [5]: latlong.rename(columns = {"Postal Code": "PostalCode"}, inplace = True)
        latlong.head()
```

Out[5]:

	PostalCode	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 [6]: latlong.shape
```

Out[6]: (103, 3)

```
In [7]: latlong.tail()
```

Out[7]:

	PostalCode	Latitude	Longitude
98	M9N	43.706876	-79.518188
99	M9P	43.696319	-79.532242
100	M9R	43.688905	-79.554724
101	M9V	43.739416	-79.588437
102	M9W	43.706748	-79.594054

Join latitude and longitude dataframe with neighborhood dataframe

```
In [8]: #Join the Lat and Long dataframe to Neighborhoods dataframe
        df_toronto.set_index("PostalCode")
        latlong.set_index("PostalCode")
        neighbor=pd.merge(df_toronto, latlong)
        neighbor.head()
```

Out[8]:

	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

```
In [9]: print('Toronto has {} boroughs and {} neighborhoods.'.format(
        len(neighbor['Borough'].unique()),
        neighbor.shape[0]
    )
)
```

Toronto has 11 boroughs and 103 neighborhoods.

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

```
In [10]: address = 'Toronto, CA'

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

/opt/conda/envs/DSX-Python35/lib/python3.5/site-packages/ipykernel/__main__.py:3: DeprecationWarning: Using Nominatim with the default "geopy/1.18.1" `user_agent` is strongly discouraged, as it violates Nominatim's ToS <https://operations.osmfoundation.org/policies/nominatim/> and may possibly cause 403 and 429 HTTP errors. Please specify a custom `user_agent` with `Nominatim(user_agent="my-application")` or by overriding the default `user_agent`: `geopy.geocoders.options.default_user_agent = "my-application"`. In geopy 2.0 this will become an exception.

```
app.launch_new_instance()
```

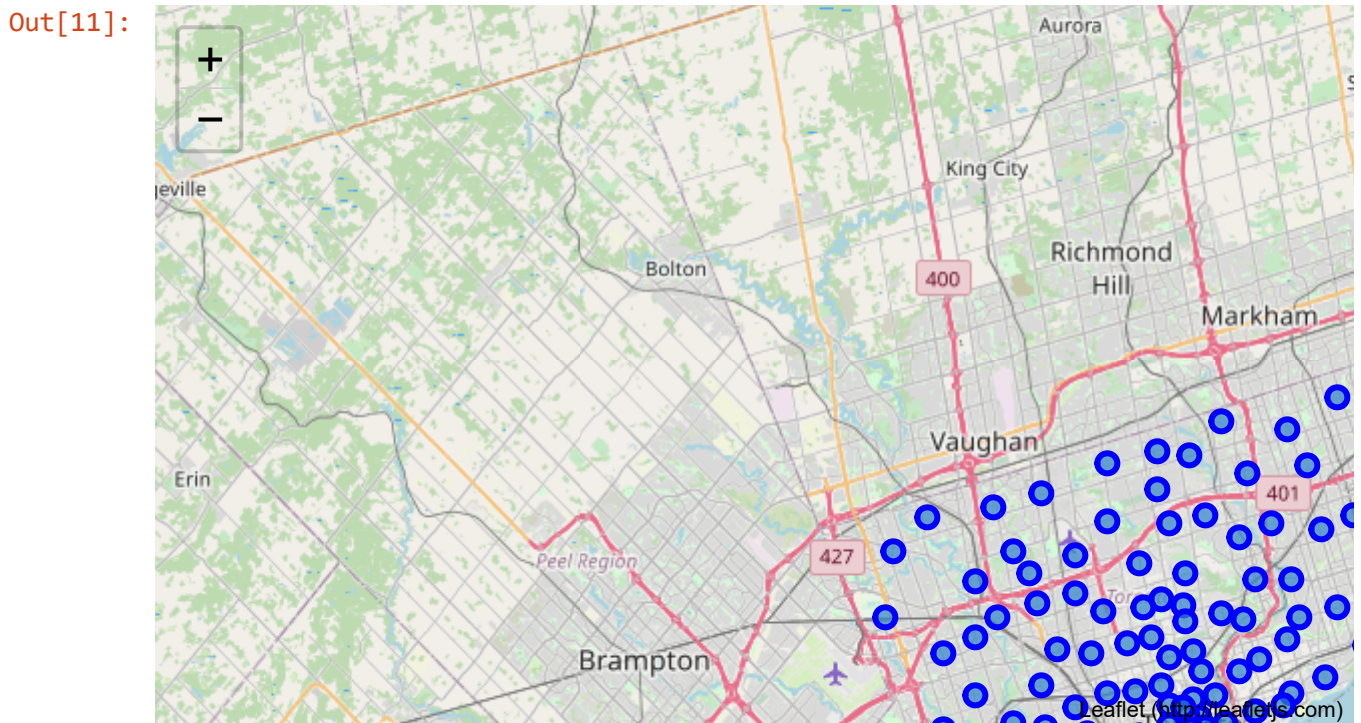
The geograpical coordinate of Toronto, Canada are 43.7170226, -79.4197830350134.

Create a map of Toronto with neighborhoods superimposed on top.

```
In [11]: # 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(neighbor['Latitude'], neighbor['Longitude'], neighbor['Borough'], neighbor['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
```



Folium is a great visualization library. Feel free to zoom into the above map, and click on each circle mark to reveal the name of the neighborhood and its respective borough.

However, for illustration purposes, let's simplify the above map and segment and cluster only the neighborhoods in Toronto. So let's slice the original dataframe and create a new dataframe of the **Scarborough** Neighborhood data.

```
In [12]: scarborough_data = neighbor[neighbor['Borough'] == 'Scarborough'].reset_index(drop=True)
scarborough_data.head()
```

```
Out[12]:
```

	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

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

```
Out[13]: (17, 5)
```

Lets get the geographical coordinates of Scarborough

```
In [14]: address = 'Scarborough, Toronto'

geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Scarborough, CA are {}, {}'.format(latitude, longitude))
```

```
/opt/conda/envs/DSX-Python35/lib/python3.5/site-packages/ipykernel/__main__.py:3: DeprecationWarning: Using Nominatim with the default "geopy/1.18.1" `user_agent` is strongly discouraged, as it violates Nominatim's ToS https://operations.osmfoundation.org/policies/nominatim/ and may possibly cause 403 and 429 HTTP errors. Please specify a custom `user_agent` with `Nominatim(user_agent="my-application")` or by overriding the default `user_agent`: `geopy.geocoders.options.default_user_agent = "my-application"`. In geopy 2.0 this will become an exception.
```

```
app.launch_new_instance()
```

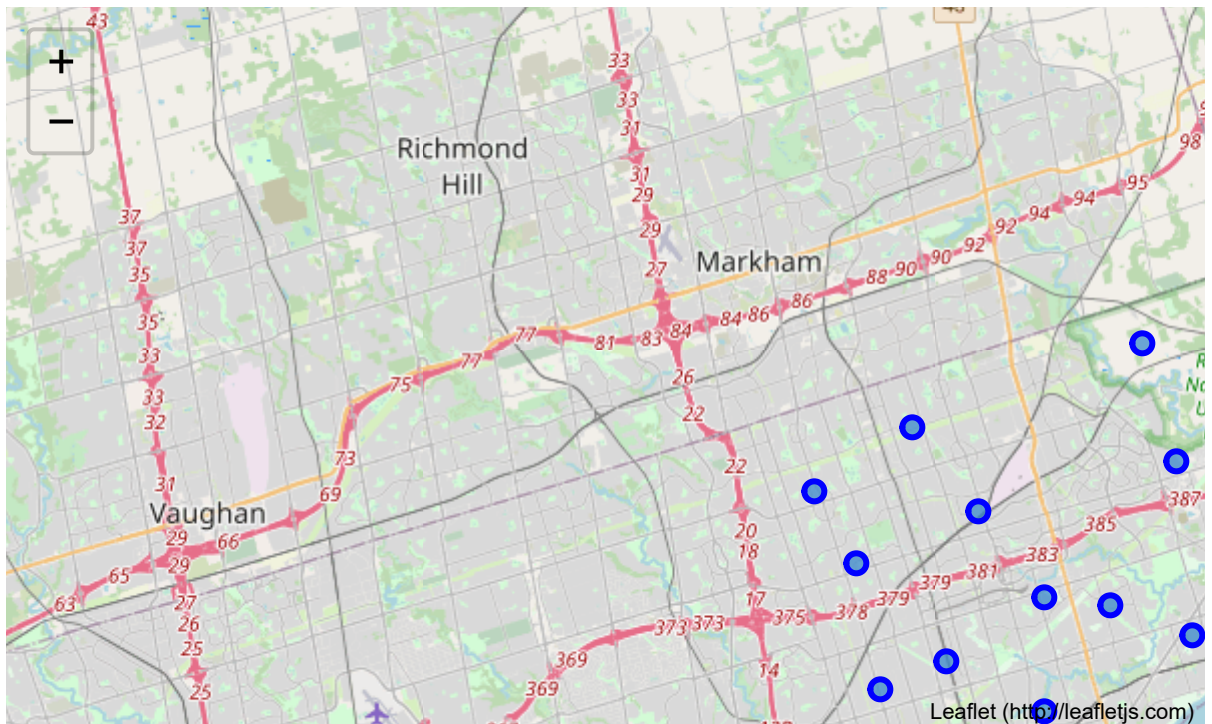
The geograpical coordinate of Scarborough, CA are 43.773077, -79.257774.


```
In [15]: # create map of Scarborough using Latitude and Longitude values
map_scarborough = folium.Map(location=[latitude, longitude], zoom_start=11)

# add markers to map
for lat, lng, label in zip(scarborough_data['Latitude'], scarborough_data['Longitude'], scarborough_data['Neighborhood']):
    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_scarborough)

map_scarborough
```

Out[15]:



Explore Scarborough neighborhood in Toronto with Foursquare API

Define Foursquare credentials and version

```
In [77]: CLIENT_ID = '*****' # your Foursquare ID
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: *****
CLIENT_SECRET:*****
```

Function to explore neighborhoods

```

In [17]: 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,
            100)

        # 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_list])
    nearby_venues.columns = ['Neighborhood',
                            'Neighborhood Latitude',
                            'Neighborhood Longitude',
                            'Venue',
                            'Venue Latitude',
                            'Venue Longitude',
                            'Venue Category']

    print('Found {} venues in {} neighborhoods.'.format(nearby_venues.shape[0], len(venues_list)))

    return(nearby_venues)

```

```

In [18]: scarborough_venues = getNearbyVenues(names=scarborough_data['Neighborhood'],
                                                latitudes=scarborough_data['Latitude'],
                                                longitudes=scarborough_data['Longitude']
                                                )

```

Found 84 venues in 17 neighborhoods.

```
In [19]: print(scarborough_venues.shape)
scarborough_venues.head()
```

(84, 7)

Out[19]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	(
0	Rouge,Malvern	43.806686	-79.194353	Wendy's	43.807448	-79.199056	F R
1	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	
2	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497	Affordable Toronto Movers	43.787919	-79.162977	
3	Guildwood,Morningside,West Hill	43.763573	-79.188711	Swiss Chalet Rotisserie & Grill	43.767697	-79.189914	
4	Guildwood,Morningside,West Hill	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	El

```
In [20]: #Venues per Neighborhood
scarborough_venues.groupby('Neighborhood').count()
```

Out[20]:

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Agincourt	4	4	4	4	4	
Agincourt North,L'Amoreaux East,Milliken,Steeles East	2	2	2	2	2	
Birch Cliff,Cliffside West	4	4	4	4	4	
Cedarbrae	8	8	8	8	8	
Clairlea,Golden Mile,Oakridge	9	9	9	9	9	
Clarks Corners,Sullivan,Tam O'Shanter	9	9	9	9	9	
Cliffcrest,Cliffside,Scarborough Village West	3	3	3	3	3	
Dorset Park,Scarborough Town Centre,Wexford Heights	6	6	6	6	6	
East Birchmount Park,Ionview,Kennedy Park	8	8	8	8	8	
Guildwood,Morningside,West Hill	6	6	6	6	6	
Highland Creek,Rouge Hill,Port Union	2	2	2	2	2	
L'Amoreaux West,Steeles West	13	13	13	13	13	
Maryvale,Wexford	4	4	4	4	4	
Rouge,Malvern	1	1	1	1	1	
Scarborough Village	1	1	1	1	1	
Woburn	4	4	4	4	4	

```
In [21]: print('There are {} distinct venues in {} categories.'.format(
        len(scarborough_venues['Venue'].unique()),len(scarborough_venues['Venue Ca
        tegory'].unique()))

        #print('There are {} uniques categories.'.format(len(scarborough_venues['Venue
        Category'].unique()))
```

There are 79 distinct venues in 52 categories.

Analyze each Neighborhood

```
In [22]: # one hot encoding
scarborough_onehot = pd.get_dummies(scarborough_venues[['Venue Category']], pr
efix="", prefix_sep="")

# add neighborhood column back to dataframe
scarborough_onehot['Neighborhood'] = scarborough_venues['Neighborhood']

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

scarborough_onehot.head()
```

Out[22]:

	Neighborhood	American Restaurant	Athletics & Sports	Bakery	Bank	Bar	Breakfast Spot	Bus Line	Bus Station
0	Rouge,Malvern	0	0	0	0	0	0	0	0
1	Highland Creek,Rouge Hill,Port Union	0	0	0	0	1	0	0	0
2	Highland Creek,Rouge Hill,Port Union	0	0	0	0	0	0	0	0
3	Guildwood,Morningside,West Hill	0	0	0	0	0	0	0	0
4	Guildwood,Morningside,West Hill	0	0	0	0	0	0	0	0

Next, let's group rows by neighborhood and by taking the mean of the frequency of occurrence of each category

```
In [23]: scarborough_grouped = scarborough_onehot.groupby('Neighborhood').mean().reset_index()
scarborough_grouped
```

Out[23]:

	Neighborhood	American Restaurant	Athletics & Sports	Bakery	Bank	Bar	Breakfast Spot	Bus Line
0	Agincourt	0.000000	0.000	0.000000	0.000	0.0	0.250000	0.000000
1	Agincourt North,L'Amoreaux East,Milliken,Steel...	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
2	Birch Cliff,Cliffside West	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
3	Cedarbrae	0.000000	0.125	0.125000	0.125	0.0	0.000000	0.000000
4	Clairlea,Golden Mile,Oakridge	0.000000	0.000	0.222222	0.000	0.0	0.000000	0.222222
5	Clarks Corners,Sullivan,Tam O'Shanter	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
6	Cliffcrest,Cliffside,Scarborough Village West	0.333333	0.000	0.000000	0.000	0.0	0.000000	0.000000
7	Dorset Park,Scarborough Town Centre,Wexford He...	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
8	East Birchmount Park,Ionview,Kennedy Park	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
9	Guildwood,Morningside,West Hill	0.000000	0.000	0.000000	0.000	0.0	0.166667	0.000000
10	Highland Creek,Rouge Hill,Port Union	0.000000	0.000	0.000000	0.000	0.5	0.000000	0.000000
11	L'Amoreaux West,Steeles West	0.000000	0.000	0.000000	0.000	0.0	0.076923	0.000000
12	Maryvale,Wexford	0.000000	0.000	0.250000	0.000	0.0	0.250000	0.000000
13	Rouge,Malvern	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
14	Scarborough Village	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000
15	Woburn	0.000000	0.000	0.000000	0.000	0.0	0.000000	0.000000

Each neighborhood with the top 5 venues

```
In [24]: num_top_venues = 10

for hood in scarborough_grouped['Neighborhood']:
    print("-----"+hood+"-----")
    temp = scarborough_grouped[scarborough_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')
```


----Agincourt----

	venue	freq
0	Lounge	0.25
1	Breakfast Spot	0.25
2	Skating Rink	0.25
3	Clothing Store	0.25
4	American Restaurant	0.00
5	Park	0.00
6	Latin American Restaurant	0.00
7	Medical Center	0.00
8	Mexican Restaurant	0.00
9	Middle Eastern Restaurant	0.00

----Agincourt North,L'Amoreaux East,Milliken,Steeles East----

	venue	freq
0	Park	0.5
1	Playground	0.5
2	Noodle House	0.0
3	Korean Restaurant	0.0
4	Latin American Restaurant	0.0
5	Lounge	0.0
6	Medical Center	0.0
7	Mexican Restaurant	0.0
8	Middle Eastern Restaurant	0.0
9	Motel	0.0

----Birch Cliff,Cliffside West----

	venue	freq
0	College Stadium	0.25
1	Skating Rink	0.25
2	Café	0.25
3	General Entertainment	0.25
4	Park	0.00
5	Lounge	0.00
6	Medical Center	0.00
7	Mexican Restaurant	0.00
8	Middle Eastern Restaurant	0.00
9	Motel	0.00

----Cedarbrae----

	venue	freq
0	Lounge	0.12
1	Bakery	0.12
2	Bank	0.12
3	Thai Restaurant	0.12
4	Athletics & Sports	0.12
5	Caribbean Restaurant	0.12
6	Hakka Restaurant	0.12
7	Fried Chicken Joint	0.12
8	Noodle House	0.00
9	Medical Center	0.00

----Clairlea,Golden Mile,Oakridge----

	venue	freq
0	Bakery	0.22
1	Bus Line	0.22
2	Park	0.11
3	Intersection	0.11
4	Soccer Field	0.11
5	Bus Station	0.11
6	Fast Food Restaurant	0.11
7	American Restaurant	0.00
8	Medical Center	0.00
9	Mexican Restaurant	0.00

----Clarks Corners,Sullivan,Tam O'Shanter----

	venue	freq
0	Pizza Place	0.22
1	Italian Restaurant	0.11
2	Thai Restaurant	0.11
3	Fried Chicken Joint	0.11
4	Chinese Restaurant	0.11
5	Pharmacy	0.11
6	Fast Food Restaurant	0.11
7	Noodle House	0.11
8	Nail Salon	0.00
9	Lounge	0.00

----Cliffcrest,Cliffside,Scarborough Village West----

	venue	freq
0	American Restaurant	0.33
1	Motel	0.33
2	Movie Theater	0.33
3	Athletics & Sports	0.00
4	Korean Restaurant	0.00
5	Latin American Restaurant	0.00
6	Lounge	0.00
7	Medical Center	0.00
8	Mexican Restaurant	0.00
9	Middle Eastern Restaurant	0.00

----Dorset Park,Scarborough Town Centre,Wexford Heights----

	venue	freq
0	Indian Restaurant	0.33
1	Vietnamese Restaurant	0.17
2	Chinese Restaurant	0.17
3	Latin American Restaurant	0.17
4	Pet Store	0.17
5	Smoke Shop	0.00
6	Skating Rink	0.00
7	Train Station	0.00
8	Thai Restaurant	0.00
9	Lounge	0.00

----East Birchmount Park,Ionview,Kennedy Park----

	venue	freq
--	-------	------

0	Discount Store	0.25
1	Bus Station	0.12
2	Department Store	0.12
3	Coffee Shop	0.12
4	Convenience Store	0.12
5	Train Station	0.12
6	Hobby Shop	0.12
7	Middle Eastern Restaurant	0.00
8	Motel	0.00
9	Mexican Restaurant	0.00

----Guildwood,Morningside,West Hill----

	venue	freq
0	Medical Center	0.17
1	Breakfast Spot	0.17
2	Mexican Restaurant	0.17
3	Rental Car Location	0.17
4	Pizza Place	0.17
5	Electronics Store	0.17
6	American Restaurant	0.00
7	Noodle House	0.00
8	Latin American Restaurant	0.00
9	Lounge	0.00

----Highland Creek,Rouge Hill,Port Union----

	venue	freq
0	Bar	0.5
1	Moving Target	0.5
2	American Restaurant	0.0
3	Park	0.0
4	Latin American Restaurant	0.0
5	Lounge	0.0
6	Medical Center	0.0
7	Mexican Restaurant	0.0
8	Middle Eastern Restaurant	0.0
9	Motel	0.0

----L'Amoreaux West,Steeles West----

	venue	freq
0	Fast Food Restaurant	0.23
1	Chinese Restaurant	0.15
2	Nail Salon	0.08
3	Pizza Place	0.08
4	Coffee Shop	0.08
5	Grocery Store	0.08
6	Pharmacy	0.08
7	Breakfast Spot	0.08
8	Japanese Restaurant	0.08
9	Sandwich Place	0.08

----Maryvale,Wexford----

	venue	freq
0	Bakery	0.25

1	Breakfast Spot	0.25
2	Smoke Shop	0.25
3	Middle Eastern Restaurant	0.25
4	American Restaurant	0.00
5	Park	0.00
6	Latin American Restaurant	0.00
7	Lounge	0.00
8	Medical Center	0.00
9	Mexican Restaurant	0.00

----Rouge,Malvern----

	venue	freq
0	Fast Food Restaurant	1.0
1	American Restaurant	0.0
2	Japanese Restaurant	0.0
3	Latin American Restaurant	0.0
4	Lounge	0.0
5	Medical Center	0.0
6	Mexican Restaurant	0.0
7	Middle Eastern Restaurant	0.0
8	Motel	0.0
9	Movie Theater	0.0

----Scarborough Village----

	venue	freq
0	Playground	1.0
1	American Restaurant	0.0
2	Noodle House	0.0
3	Korean Restaurant	0.0
4	Latin American Restaurant	0.0
5	Lounge	0.0
6	Medical Center	0.0
7	Mexican Restaurant	0.0
8	Middle Eastern Restaurant	0.0
9	Motel	0.0

----Woburn----

	venue	freq
0	Coffee Shop	0.50
1	Korean Restaurant	0.25
2	Convenience Store	0.25
3	Latin American Restaurant	0.00
4	Lounge	0.00
5	Medical Center	0.00
6	Mexican Restaurant	0.00
7	Middle Eastern Restaurant	0.00
8	Motel	0.00
9	Movie Theater	0.00

Let's put that into a *pandas* dataframe

Let's write a function to sort the venues in descending order

```
In [25]: 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]
```

Create a new dataframe and display the top ten venues for each neighborhood

```
In [26]: 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'] = scarborough_grouped['Neighborhood']

for ind in np.arange(scarborough_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(scar
borough_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted
```

Out[26]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Agincourt	Clothing Store	Skating Rink	Breakfast Spot	Lounge	College Stadium
1	Agincourt North,L'Amoreaux East,Milliken,Steel...	Playground	Park	Vietnamese Restaurant	Coffee Shop	Hakka Restaurant (
2	Birch Cliff,Cliffside West	General Entertainment	Skating Rink	Café	College Stadium	Vietnamese Restaurant
3	Cedarbrae	Thai Restaurant	Athletics & Sports	Bakery	Bank	Hakka Restaurant
4	Clairlea,Golden Mile,Oakridge	Bakery	Bus Line	Intersection	Soccer Field	Bus Station
5	Clarks Corners,Sullivan,Tam O'Shanter	Pizza Place	Thai Restaurant	Italian Restaurant	Fried Chicken Joint	Fast Food Restaurant
6	Cliffcrest,Cliffside,Scarborough Village West	American Restaurant	Motel	Movie Theater	College Stadium	Hakka Restaurant (
7	Dorset Park,Scarborough Town Centre,Wexford He...	Indian Restaurant	Vietnamese Restaurant	Latin American Restaurant	Pet Store	Chinese Restaurant
8	East Birchmount Park,Ionview,Kennedy Park	Discount Store	Coffee Shop	Hobby Shop	Bus Station	Department Store
9	Guildwood,Morningside,West Hill	Breakfast Spot	Rental Car Location	Electronics Store	Pizza Place	Medical Center
10	Highland Creek,Rouge Hill,Port Union	Bar	Moving Target	Vietnamese Restaurant	College Stadium	Hakka Restaurant (
11	L'Amoreaux West,Steeles West	Fast Food Restaurant	Chinese Restaurant	Nail Salon	Pizza Place	Japanese Restaurant
12	Maryvale,Wexford	Bakery	Smoke Shop	Breakfast Spot	Middle Eastern Restaurant	Vietnamese Restaurant
13	Rouge,Malvern	Fast Food Restaurant	Vietnamese Restaurant	Train Station	Hobby Shop	Hakka Restaurant (
14	Scarborough Village	Playground	Vietnamese Restaurant	Coffee Shop	Hakka Restaurant	Grocery Store
15	Woburn	Coffee Shop	Korean Restaurant	Convenience Store	Indian Restaurant	Hakka Restaurant (

```
In [27]: neighborhoods_venues_sorted.iloc[11,]
```

```
Out[27]: Neighborhood      L'Amoreaux West,Steeles West
1st Most Common Venue      Fast Food Restaurant
2nd Most Common Venue      Chinese Restaurant
3rd Most Common Venue      Nail Salon
4th Most Common Venue      Pizza Place
5th Most Common Venue      Japanese Restaurant
6th Most Common Venue      Coffee Shop
7th Most Common Venue      Pharmacy
8th Most Common Venue      Grocery Store
9th Most Common Venue      Sandwich Place
10th Most Common Venue     Breakfast Spot
Name: 11, dtype: object
```

4. Cluster the Scarborough Neighborhood using k-means

Run K-means to cluster neighborhood into three clusters

```
In [28]: # set number of clusters
kclusters = 3

scarborough_grouped_clustering = scarborough_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=2).fit(scarborough_grouped_
clustering)

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

```
Out[28]: array([1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 0, 1], dtype=int32)
```

Lets create a new dataframe that includes the cluster as well as the top ten venues for each neighborhood


```
In [29]: #Note that the neighborhood Upper Rouge does not have any venues, so I will drop from dataset
scarborough_data.drop(scarborough_data[scarborough_data.Neighborhood == 'Upper Rouge'].index, inplace = True)
#df_toronto.drop(df_toronto[df_toronto.Borough == 'Not assigned'].index, inplace=True)

scarborough_merged = scarborough_data

# add clustering labels
scarborough_merged['Cluster Labels'] = kmeans.labels_

# merge toronto_grouped with toronto_data to add Latitude/Longitude for each neighborhood
scarborough_merged = scarborough_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

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

Out[29]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue
0	M1B	Scarborough	Rouge,Malvern	43.806686	-79.194353	1	Fast Food Restaurant
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497	0	Bar
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711	1	Breakfast Spot
3	M1G	Scarborough	Woburn	43.770992	-79.216917	1	Coffee Shop
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	1	Thai Restaurant

```

In [30]: # create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

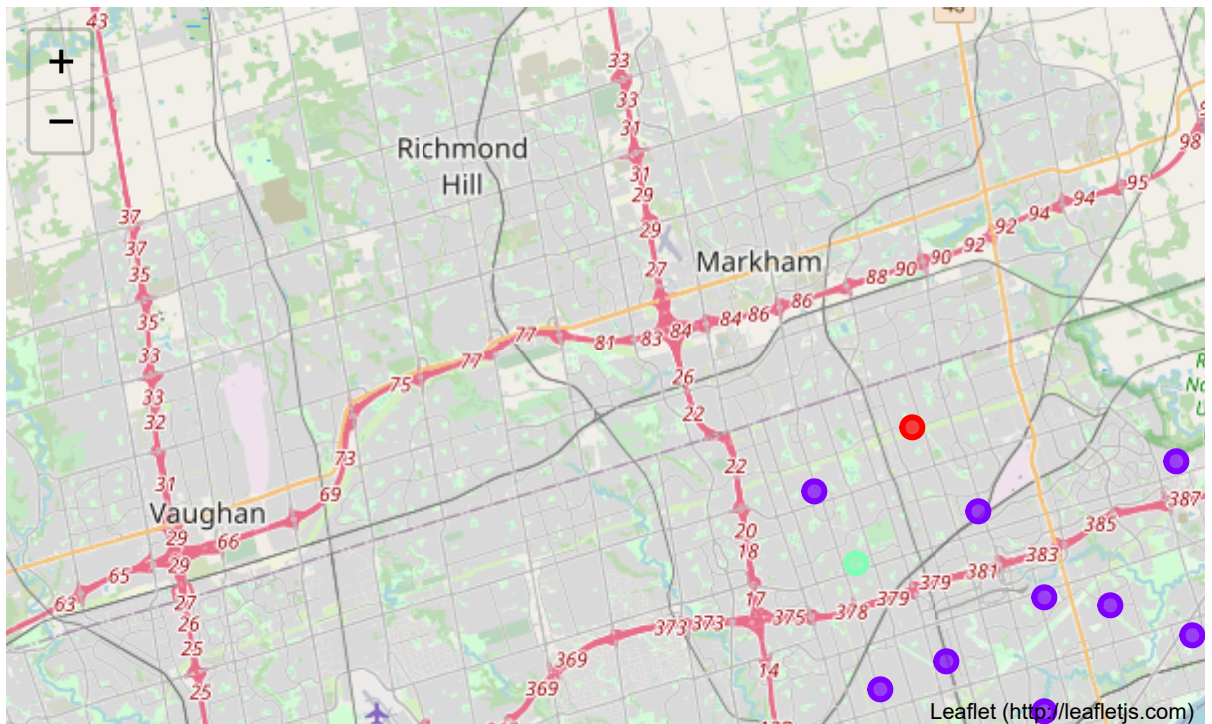
# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i+x+(i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(scarborough_merged['Latitude'], scarborough_
merged['Longitude'], scarborough_merged['Neighborhood'], scarborough_merged['C
luster 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[30]:



5. Examine Scarborough Neighborhood Cluster

Now, you can examine each cluster and determine the discriminating venue categories that distinguish each cluster. Based on the defining categories, you can then assign a name to each cluster.

Scarborough Clusters 0, 1, 2

```
In [31]: scarborough_cluster_0 = scarborough_merged.loc[scarborough_merged['Cluster Labels'] == 0, scarborough_merged.columns[[1] + list(range(4, scarborough_merged.shape[1]))]]

scarborough_cluster_1 = scarborough_merged.loc[scarborough_merged['Cluster Labels'] == 1, scarborough_merged.columns[[1] + list(range(4, scarborough_merged.shape[1]))]]

scarborough_cluster_2 = scarborough_merged.loc[scarborough_merged['Cluster Labels'] == 2, scarborough_merged.columns[[1] + list(range(4, scarborough_merged.shape[1]))]]
```

```
In [32]: scarborough_cluster_0
```

Out[32]:

	Borough	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Co
1	Scarborough	-79.160497	0	Bar	Moving Target	Vietnamese Restaurant	College Stadium	Hakka Restaurant	G
14	Scarborough	-79.284577	0	Playground	Park	Vietnamese Restaurant	Coffee Shop	Hakka Restaurant	G

In [33]: `scarborough_cluster_1`

Out[33]:

	Borough	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Scarborough	-79.194353	1	Fast Food Restaurant	Vietnamese Restaurant	Train Station	Hobby Shop	Halal Restaurant
2	Scarborough	-79.188711	1	Breakfast Spot	Rental Car Location	Electronics Store	Pizza Place	Medicine Center
3	Scarborough	-79.216917	1	Coffee Shop	Korean Restaurant	Convenience Store	Indian Restaurant	Halal Restaurant
4	Scarborough	-79.239476	1	Thai Restaurant	Athletics & Sports	Bakery	Bank	Halal Restaurant
5	Scarborough	-79.239476	1	Playground	Vietnamese Restaurant	Coffee Shop	Hakka Restaurant	Grocery Store
6	Scarborough	-79.262029	1	Discount Store	Coffee Shop	Hobby Shop	Bus Station	Department Store
7	Scarborough	-79.284577	1	Bakery	Bus Line	Intersection	Soccer Field	Bus Station
8	Scarborough	-79.239476	1	American Restaurant	Motel	Movie Theater	College Stadium	Halal Restaurant
9	Scarborough	-79.264848	1	General Entertainment	Skating Rink	Café	College Stadium	Vietnamese Restaurant
10	Scarborough	-79.273304	1	Indian Restaurant	Vietnamese Restaurant	Latin American Restaurant	Pet Store	Chinese Restaurant
11	Scarborough	-79.295849	1	Bakery	Smoke Shop	Breakfast Spot	Middle Eastern Restaurant	Vietnamese Restaurant
12	Scarborough	-79.262029	1	Clothing Store	Skating Rink	Breakfast Spot	Lounge	College Stadium
15	Scarborough	-79.318389	1	Fast Food Restaurant	Chinese Restaurant	Nail Salon	Pizza Place	Japanese Restaurant

In [34]: `scarborough_cluster_2`

Out[34]:

	Borough	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
13	Scarborough	-79.304302	2	Pizza Place	Thai Restaurant	Italian Restaurant	Fried Chicken Joint	Fast Food Restaurant	Pharmaceutical Store

6. Explore New York City Neighborhoods

New York City has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.

Link to the dataset: https://geo.nyu.edu/catalog/nyu_2451_34572 (https://geo.nyu.edu/catalog/nyu_2451_34572)

```
In [35]: !wget -q -O 'newyork_data.json' https://ibm.box.com/shared/static/fbpwbovar7lf8p5sgddm06cgipa2rxpe.json
print('Data downloaded!')
```

Data downloaded!

Load and explore dataset

```
In [36]: with open('newyork_data.json') as json_data:
        newyork_data = json.load(json_data)
```

Notice that all relevant data is in the *features* key, which is essentially a list of the neighborhoods. With that in mind, let's define a new variable that includes this data.

```
In [37]: neighborhoods_data = newyork_data['features']
```

```
In [38]: neighborhoods_data[0]
```

```
Out[38]: {'geometry': {'coordinates': [-73.84720052054902, 40.89470517661],
  'type': 'Point'},
  'geometry_name': 'geom',
  'id': 'nyu_2451_34572.1',
  'properties': {'annoangle': 0.0,
  'annoline1': 'Wakefield',
  'annoline2': None,
  'annoline3': None,
  'bbox': [-73.84720052054902,
  40.89470517661,
  -73.84720052054902,
  40.89470517661],
  'borough': 'Bronx',
  'name': 'Wakefield',
  'stacked': 1},
  'type': 'Feature'}
```

Transform the data into a pandas dataframe

```
In [39]: # define the dataframe columns
column_names = ['Borough', 'Neighborhood', 'Latitude', 'Longitude']

# instantiate the dataframe
neighborhoods = pd.DataFrame(columns=column_names)
```

```
In [40]: for data in neighborhoods_data:
    borough = neighborhood_name = data['properties']['borough']
    neighborhood_name = data['properties']['name']

    neighborhood_latlon = data['geometry']['coordinates']
    neighborhood_lat = neighborhood_latlon[1]
    neighborhood_lon = neighborhood_latlon[0]

    neighborhoods = neighborhoods.append({'Borough': borough,
                                          'Neighborhood': neighborhood_name,
                                          'Latitude': neighborhood_lat,
                                          'Longitude': neighborhood_lon}, ignore_index=True)
```

```
In [41]: neighborhoods.head()
```

```
Out[41]:
```

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

```
In [42]: neighborhoods.shape
```

```
Out[42]: (306, 4)
```

```
In [43]: print('The dataframe has {} boroughs and {} neighborhoods.'.format(
    len(neighborhoods['Borough'].unique()),
    neighborhoods.shape[0]
))
```

The dataframe has 5 boroughs and 306 neighborhoods.

Use Geolibrary to get the latitude and longitude of New York City


```
In [44]: address = 'New York City, NY'

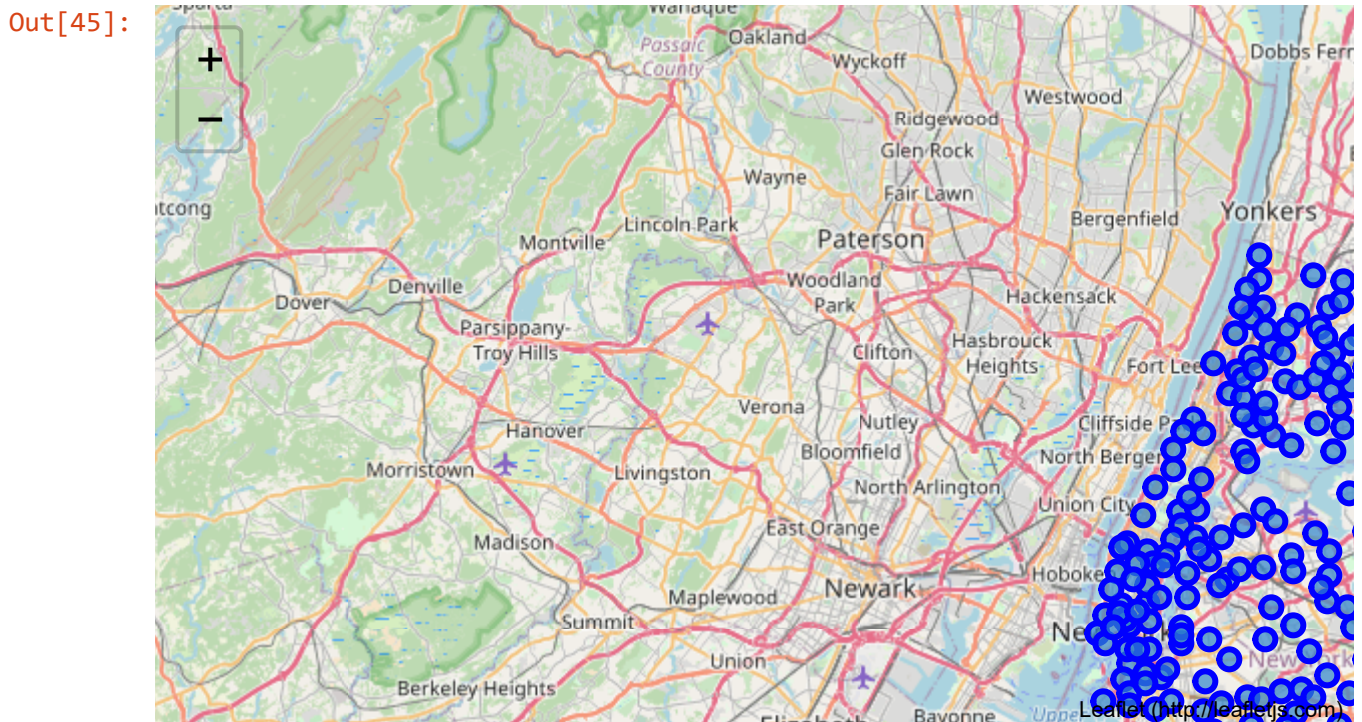
geolocator = Nominatim(user_agent = 'my-application')
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of New York City are {}, {}'.format(latitude, longitude))
```

The geographical coordinate of New York City are 40.7308619, -73.9871558.

```
In [45]: # create map of New York using Latitude and Longitude values
map_newyork = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, neighborhood in zip(neighborhoods['Latitude'], neighborhoods['Longitude'], neighborhoods['Borough'], neighborhoods['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_newyork)

map_newyork
```



Lets simplify the above map and segment and cluster only the neighborhoods in Manhattan. So let's slice the original dataframe and create a new dataframe of the Queens neighborhood data.

```
In [46]: queens_data = neighborhoods[neighborhoods['Borough'] == 'Queens'].reset_index(drop=True)
queens_data.head()
```

```
Out[46]:
```

	Borough	Neighborhood	Latitude	Longitude
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

Lets get the geographical location of Queens, NY

```
In [47]: address = 'Queens, NY'

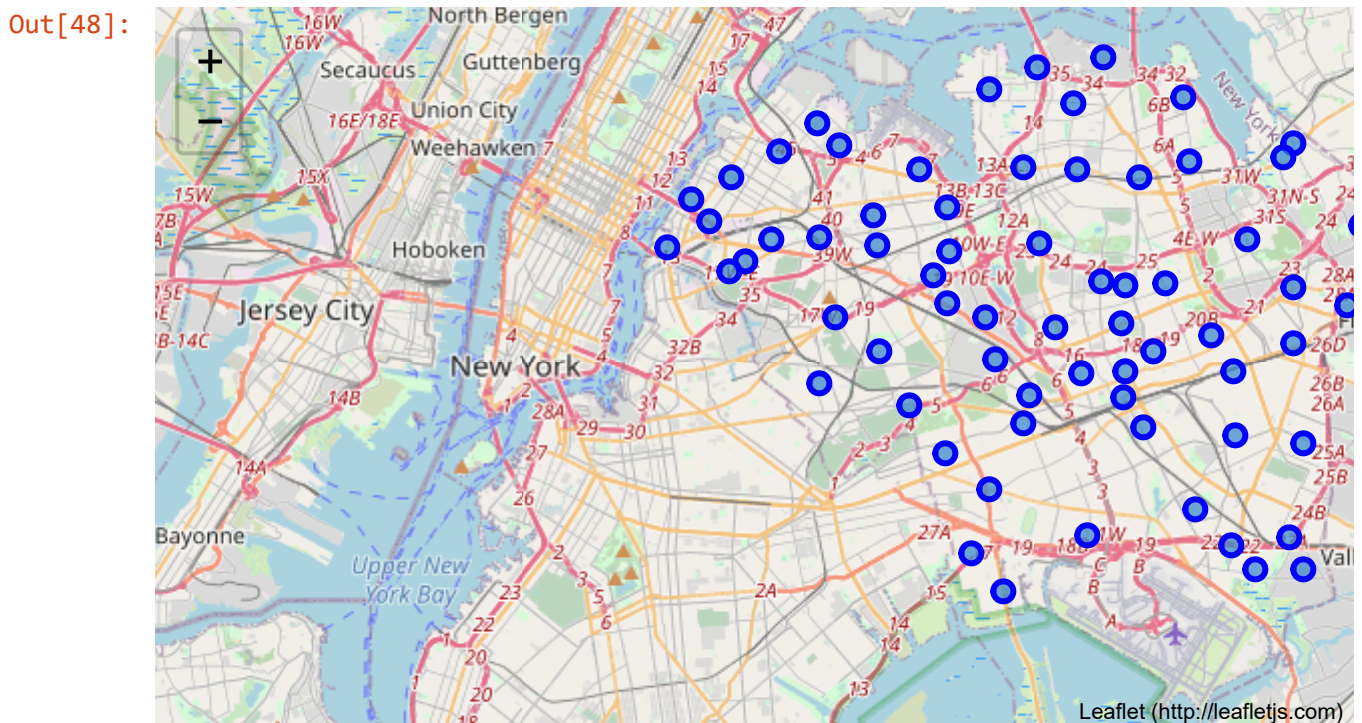
geolocator = Nominatim(user_agent = 'my-application')
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of Queens are {}, {}'.format(latitude, longitude))
```

The geograpical coordinate of Queens are 40.6524927, -73.7914214158161.


```
In [48]: # create map of Manhattan using Latitude and Longitude values
map_queens = folium.Map(location=[latitude, longitude], zoom_start=11)

# add markers to map
for lat, lng, label in zip(queens_data['Latitude'], queens_data['Longitude'],
queens_data['Neighborhood']):
    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_queens)

map_queens
```



Lets explore the Long Island City neighborhood in Queens, NY

```
In [49]: queens_data.loc[10, 'Neighborhood']
```

Out[49]: 'Long Island City'

```
In [50]: #Long Island City Latitude and Longitude values

neighborhood_latitude = queens_data.loc[10, 'Latitude'] # neighborhood Latitude value
neighborhood_longitude = queens_data.loc[10, 'Longitude'] # neighborhood Longitude value

neighborhood_name = queens_data.loc[10, 'Neighborhood'] # neighborhood name

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

Latitude and longitude values of Long Island City are 40.75021734610528, -73.93920223915505.

Top 100 venues in Long Island City neighborhood within a radius of 500 meters

First, let's create the GET request URL named `url`.

```
In [78]: # type your answer here
LIMIT = 100
radius = 500
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
```

```
Out[78]: 'https://api.foursquare.com/v2/venues/explore?&client_id=*****&client_secret=*****&v=20180605&ll=40.75021734610528,-73.93920223915505&radius=500&limit=100'
```

```
In [52]: #Send the GET request
results = requests.get(url).json()
```

```
In [53]: # 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 [54]: venues = results['response']['groups'][0]['items']

nearby_venues = json_normalize(venues) # flatten JSON

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.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_category_type, axis=1)

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

nearby_venues.head()
```

Out[54]:

	name	categories	lat	lng
0	Hilton Garden Inn New York Long Island City/Ma...	Hotel	40.750216	-73.936886
1	Clever Blend Lic	Coffee Shop	40.750228	-73.939608
2	The Beast Next Door Cafe & Bar	Bar	40.748888	-73.940876
3	Baker House Market	Convenience Store	40.752137	-73.939235
4	Etto Espresso Bar	Coffee Shop	40.748703	-73.940689

```
In [55]: print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))

65 venues were returned by Foursquare.
```

7. Analyze Each Neighborhood in Queens

```

In [56]: 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_list for item
in venue_list])
    nearby_venues.columns = ['Neighborhood',
                            'Neighborhood Latitude',
                            'Neighborhood Longitude',
                            'Venue',
                            'Venue Latitude',
                            'Venue Longitude',
                            'Venue Category']

    print('Found {} venues in {} neighborhoods.'.format(nearby_venues.shape[0],
len(venues_list)))

    return(nearby_venues)

```

```

In [57]: queens_venues = getNearbyVenues(names=queens_data['Neighborhood'],
                                         latitudes=queens_data['Latitude'],
                                         longitudes=queens_data['Longitude'])

```

Found 2108 venues in 81 neighborhoods.

```
In [58]: queens_venues.head()
```

Out[58]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Astoria	40.768509	-73.915654	Favela Grill	40.767348	-73.917897	Brazilian Restaurant
1	Astoria	40.768509	-73.915654	Orange Blossom	40.769856	-73.917012	Gourmet Shop
2	Astoria	40.768509	-73.915654	Titan Foods Inc.	40.769198	-73.919253	Gourmet Shop
3	Astoria	40.768509	-73.915654	CrossFit Queens	40.769404	-73.918977	Gym
4	Astoria	40.768509	-73.915654	Off The Hook	40.767200	-73.918104	Seafood Restaurant

```
In [59]: print(queens_venues.shape)
queens_venues.tail()
```

(2108, 7)

Out[59]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2103	Queensbridge	40.756091	-73.945631	Queensbridge Park Softball Fields	40.756055	-73.948407	Baseball Field
2104	Queensbridge	40.756091	-73.945631	SummerStage - Queensbridge Park	40.755514	-73.949125	Performing Arts Venue
2105	Queensbridge	40.756091	-73.945631	The Ravel Hotel Gym	40.753787	-73.948815	Athletic Sports
2106	Queensbridge	40.756091	-73.945631	Track 114	40.753008	-73.947833	Platform
2107	Queensbridge	40.756091	-73.945631	Ravel Rooftop	40.753971	-73.949456	Roof Deck

```
In [60]: #Venues per Neighborhood  
queens_venues.groupby('Neighborhood').count()
```

Out[60]:

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Arverne	15	15	15	15	15	15
Astoria	100	100	100	100	100	100
Astoria Heights	12	12	12	12	12	12
Auburndale	20	20	20	20	20	20
Bay Terrace	40	40	40	40	40	40
Bayside	68	68	68	68	68	68
Bayswater	2	2	2	2	2	2
Beechhurst	17	17	17	17	17	17
Bellaire	13	13	13	13	13	13
Belle Harbor	17	17	17	17	17	17
Bellerose	22	22	22	22	22	22
Blissville	18	18	18	18	18	18
Breezy Point	7	7	7	7	7	7
Briarwood	12	12	12	12	12	12
Broad Channel	5	5	5	5	5	5
Brookville	1	1	1	1	1	1
Cambria Heights	14	14	14	14	14	14
College Point	35	35	35	35	35	35
Corona	15	15	15	15	15	15
Douglaston	22	22	22	22	22	22
East Elmhurst	17	17	17	17	17	17
Edgemere	14	14	14	14	14	14
Elmhurst	35	35	35	35	35	35
Far Rockaway	31	31	31	31	31	31
Floral Park	6	6	6	6	6	6
Flushing	66	66	66	66	66	66
Forest Hills	39	39	39	39	39	39
Forest Hills Gardens	26	26	26	26	26	26
Fresh Meadows	15	15	15	15	15	15
Glen Oaks	22	22	22	22	22	22
Glendale	6	6	6	6	6	6
Hammels	18	18	18	18	18	18

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Hillcrest	18	18	18	18	18	18
Hollis	18	18	18	18	18	18
Holliswood	4	4	4	4	4	4
Howard Beach	40	40	40	40	40	40
Hunters Point	74	74	74	74	74	74
Jackson Heights	82	82	82	82	82	82
Jamaica Center	40	40	40	40	40	40
Jamaica Estates	4	4	4	4	4	4
Jamaica Hills	32	32	32	32	32	32
Kew Gardens	46	46	46	46	46	46
Kew Gardens Hills	17	17	17	17	17	17
Laurelton	6	6	6	6	6	6
Lefrak City	27	27	27	27	27	27
Lindenwood	11	11	11	11	11	11
Little Neck	47	47	47	47	47	47
Long Island City	65	65	65	65	65	65
Malba	3	3	3	3	3	3
Maspeth	30	30	30	30	30	30
Middle Village	15	15	15	15	15	15
Murray Hill	44	44	44	44	44	44
Neponsit	7	7	7	7	7	7
North Corona	22	22	22	22	22	22
Oakland Gardens	24	24	24	24	24	24
Ozone Park	32	32	32	32	32	32
Pomonok	9	9	9	9	9	9
Queens Village	16	16	16	16	16	16
Queensboro Hill	27	27	27	27	27	27
Queensbridge	17	17	17	17	17	17
Ravenswood	26	26	26	26	26	26
Rego Park	42	42	42	42	42	42
Richmond Hill	22	22	22	22	22	22

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Ridgewood	40	40	40	40	40	40
Rochdale	15	15	15	15	15	15
Rockaway Beach	46	46	46	46	46	46
Rockaway Park	25	25	25	25	25	25
Rosedale	14	14	14	14	14	14
Roxbury	9	9	9	9	9	9
Somerville	1	1	1	1	1	1
South Jamaica	10	10	10	10	10	10
South Ozone Park	10	10	10	10	10	10
Springfield Gardens	13	13	13	13	13	13
St. Albans	14	14	14	14	14	14
Steinway	28	28	28	28	28	28
Sunnyside	47	47	47	47	47	47
Sunnyside Gardens	100	100	100	100	100	100
Utopia	18	18	18	18	18	18
Whitestone	4	4	4	4	4	4
Woodhaven	25	25	25	25	25	25
Woodside	72	72	72	72	72	72

```
In [61]: print('There are {} distinct venues in {} categories.'.format(
        len(queens_venues['Venue'].unique()), len(queens_venues['Venue Category'].unique())))
```

There are 1760 distinct venues in 269 categories.

```
In [62]: # one hot encoding
queens_onehot = pd.get_dummies(queens_venues[['Venue Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
queens_onehot['Neighborhood'] = queens_venues['Neighborhood']

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

neighbor = queens_onehot['Neighborhood']
queens_onehot.drop(labels=['Neighborhood'], axis=1, inplace = True)
queens_onehot.insert(0, 'Neighborhood', neighbor)

queens_onehot.head()
```

Out[62]:

	Neighborhood	Accessories Store	Afghan Restaurant	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Ar Museum
0	Astoria	0	0	0	0	0	0	(
1	Astoria	0	0	0	0	0	0	(
2	Astoria	0	0	0	0	0	0	(
3	Astoria	0	0	0	0	0	0	(
4	Astoria	0	0	0	0	0	0	(

Group by Neighborhood, and examine the frequency of the occurrence of venue

```
In [63]: queens_grouped = queens_onehot.groupby('Neighborhood').mean().reset_index()  
         queens_grouped
```

Out[63]:

[illegible]

[illegible]

	Neighborhood	Accessories Store	Afghan Restaurant	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Musi
63	Ridgewood	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
64	Rochdale	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
65	Rockaway Beach	0.000000	0.000000	0.000000	0.043478	0.000000	0.000000	0.000
66	Rockaway Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
67	Rosedale	0.071429	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
68	Roxbury	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
69	Somerville	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
70	South Jamaica	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
71	South Ozone Park	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
72	Springfield Gardens	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
73	St. Albans	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
74	Steinway	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
75	Sunnyside	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
76	Sunnyside Gardens	0.000000	0.000000	0.040000	0.000000	0.000000	0.000000	0.000
77	Utopia	0.000000	0.055556	0.000000	0.000000	0.000000	0.000000	0.000
78	Whitestone	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
79	Woodhaven	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000
80	Woodside	0.000000	0.000000	0.041667	0.013889	0.000000	0.000000	0.000

Each Neighborhood with the top 5 venues

```
In [64]: num_top_venues = 5

for hood in queens_grouped['Neighborhood']:
    print("-----"+hood+"-----")
    temp = queens_grouped[queens_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')
```

----Arverne----

	venue	freq
0	Surf Spot	0.27
1	Metro Station	0.13
2	Bus Stop	0.07
3	Pizza Place	0.07
4	Thai Restaurant	0.07

----Astoria----

	venue	freq
0	Middle Eastern Restaurant	0.07
1	Bar	0.06
2	Hookah Bar	0.06
3	Bakery	0.05
4	Greek Restaurant	0.05

----Astoria Heights----

	venue	freq
0	Deli / Bodega	0.17
1	Supermarket	0.08
2	Bakery	0.08
3	Italian Restaurant	0.08
4	Burger Joint	0.08

----Auburndale----

	venue	freq
0	Deli / Bodega	0.05
1	Train Station	0.05
2	Italian Restaurant	0.05
3	Donut Shop	0.05
4	Pharmacy	0.05

----Bay Terrace----

	venue	freq
0	Clothing Store	0.10
1	Women's Store	0.08
2	Cosmetics Shop	0.05
3	Kids Store	0.05
4	Lingerie Store	0.05

----Bayside----

	venue	freq
0	Bar	0.09
1	Indian Restaurant	0.04
2	Donut Shop	0.04
3	Pizza Place	0.04
4	Pub	0.04

----Bayswater----

	venue	freq
0	Playground	0.5

1	Park	0.5
2	Accessories Store	0.0
3	North Indian Restaurant	0.0
4	Outdoors & Recreation	0.0

----Beechhurst----

	venue	freq
0	Gym	0.12
1	Yoga Studio	0.06
2	Boutique	0.06
3	Gym / Fitness Center	0.06
4	Frozen Yogurt Shop	0.06

----Bellaire----

	venue	freq
0	Deli / Bodega	0.08
1	Coffee Shop	0.08
2	Grocery Store	0.08
3	Greek Restaurant	0.08
4	Bus Station	0.08

----Belle Harbor----

	venue	freq
0	Beach	0.18
1	Deli / Bodega	0.12
2	Spa	0.12
3	Boutique	0.12
4	Bagel Shop	0.06

----Bellerose----

	venue	freq
0	Pizza Place	0.14
1	Italian Restaurant	0.09
2	Flower Shop	0.09
3	Chinese Restaurant	0.09
4	Deli / Bodega	0.05

----Blissville----

	venue	freq
0	Hotel	0.17
1	Deli / Bodega	0.11
2	Donut Shop	0.11
3	Hostel	0.06
4	Rental Service	0.06

----Breezy Point----

	venue	freq
0	Beach	0.29
1	Supermarket	0.14
2	Lighthouse	0.14
3	Trail	0.14

4 Board Shop 0.14

----Briarwood----

	venue	freq
0	Deli / Bodega	0.17
1	Indian Restaurant	0.08
2	Convenience Store	0.08
3	Gym	0.08
4	Pet Store	0.08

----Broad Channel----

	venue	freq
0	Deli / Bodega	0.2
1	Other Nightlife	0.2
2	Bus Station	0.2
3	Dive Bar	0.2
4	Pizza Place	0.2

----Brookville----

	venue	freq
0	Deli / Bodega	1.0
1	Accessories Store	0.0
2	North Indian Restaurant	0.0
3	Paper / Office Supplies Store	0.0
4	Outdoors & Recreation	0.0

----Cambria Heights----

	venue	freq
0	Caribbean Restaurant	0.29
1	Cosmetics Shop	0.14
2	Moving Target	0.07
3	Restaurant	0.07
4	Liquor Store	0.07

----College Point----

	venue	freq
0	Deli / Bodega	0.11
1	Pizza Place	0.06
2	Karaoke Bar	0.06
3	Latin American Restaurant	0.06
4	Pharmacy	0.06

----Corona----

	venue	freq
0	Mexican Restaurant	0.13
1	Restaurant	0.07
2	Empanada Restaurant	0.07
3	Bakery	0.07
4	Supermarket	0.07

----Douglaston----

	venue	freq
0	Deli / Bodega	0.14
1	Italian Restaurant	0.09
2	Chinese Restaurant	0.09
3	Bank	0.05
4	Diner	0.05

----East Elmhurst----

	venue	freq
0	Donut Shop	0.18
1	Bus Station	0.06
2	Lake	0.06
3	Coffee Shop	0.06
4	School	0.06

----Edgemere----

	venue	freq
0	Metro Station	0.14
1	Pizza Place	0.14
2	Seafood Restaurant	0.07
3	Scenic Lookout	0.07
4	Gift Shop	0.07

----Elmhurst----

	venue	freq
0	Thai Restaurant	0.17
1	Mexican Restaurant	0.14
2	Chinese Restaurant	0.09
3	Bubble Tea Shop	0.09
4	South American Restaurant	0.06

----Far Rockaway----

	venue	freq
0	Grocery Store	0.10
1	Chinese Restaurant	0.10
2	Pizza Place	0.10
3	Breakfast Spot	0.06
4	Fast Food Restaurant	0.06

----Floral Park----

	venue	freq
0	Indian Restaurant	0.33
1	Basketball Court	0.17
2	Grocery Store	0.17
3	Hobby Shop	0.17
4	Pizza Place	0.17

----Flushing----

	venue	freq
0	Chinese Restaurant	0.09

1	Bubble Tea Shop	0.09
2	Korean Restaurant	0.05
3	Hotpot Restaurant	0.05
4	Asian Restaurant	0.05

----Forest Hills----

	venue	freq
0	Gym	0.08
1	Yoga Studio	0.05
2	Convenience Store	0.05
3	Thai Restaurant	0.05
4	Park	0.05

----Forest Hills Gardens----

	venue	freq
0	Bakery	0.12
1	Grocery Store	0.08
2	Food	0.08
3	New American Restaurant	0.04
4	Sushi Restaurant	0.04

----Fresh Meadows----

	venue	freq
0	Bus Station	0.20
1	Chinese Restaurant	0.13
2	Moving Target	0.07
3	Hotel	0.07
4	Grocery Store	0.07

----Glen Oaks----

	venue	freq
0	Pharmacy	0.09
1	Indian Restaurant	0.05
2	Bagel Shop	0.05
3	Grocery Store	0.05
4	Park	0.05

----Glendale----

	venue	freq
0	Dance Studio	0.17
1	Brewery	0.17
2	Pizza Place	0.17
3	Arts & Crafts Store	0.17
4	Food & Drink Shop	0.17

----Hammels----

	venue	freq
0	Beach	0.28
1	Deli / Bodega	0.06
2	Fast Food Restaurant	0.06
3	Diner	0.06

----Hillcrest----

	venue	freq
0	Fast Food Restaurant	0.06
1	College Basketball Court	0.06
2	Harbor / Marina	0.06
3	College Stadium	0.06
4	College Academic Building	0.06

----Hollis----

	venue	freq
0	Shopping Mall	0.11
1	Baseball Field	0.11
2	Sandwich Place	0.11
3	Park	0.11
4	Fried Chicken Joint	0.06

----Holliswood----

	venue	freq
0	Donut Shop	0.25
1	Mobile Phone Shop	0.25
2	Supermarket	0.25
3	Playground	0.25
4	Noodle House	0.00

----Howard Beach----

	venue	freq
0	Italian Restaurant	0.08
1	Bagel Shop	0.08
2	Chinese Restaurant	0.05
3	Clothing Store	0.05
4	Deli / Bodega	0.05

----Hunters Point----

	venue	freq
0	Italian Restaurant	0.05
1	Café	0.05
2	Brewery	0.04
3	Japanese Restaurant	0.04
4	Comedy Club	0.03

----Jackson Heights----

	venue	freq
0	Latin American Restaurant	0.11
1	Mobile Phone Shop	0.06
2	South American Restaurant	0.06
3	Peruvian Restaurant	0.06
4	Bakery	0.05

----Jamaica Center----

	venue	freq
0	Mobile Phone Shop	0.10
1	Donut Shop	0.05
2	Pizza Place	0.05
3	Sandwich Place	0.05
4	Performing Arts Venue	0.05

----Jamaica Estates----

	venue	freq
0	Indian Restaurant	0.25
1	Grocery Store	0.25
2	Dog Run	0.25
3	Lounge	0.25
4	Museum	0.00

----Jamaica Hills----

	venue	freq
0	Pharmacy	0.09
1	Indian Restaurant	0.06
2	Fast Food Restaurant	0.06
3	Donut Shop	0.06
4	Fried Chicken Joint	0.06

----Kew Gardens----

	venue	freq
0	Chinese Restaurant	0.09
1	Indian Restaurant	0.04
2	Bakery	0.04
3	Pharmacy	0.04
4	Pizza Place	0.04

----Kew Gardens Hills----

	venue	freq
0	Bank	0.12
1	Bus Station	0.12
2	Playground	0.12
3	Pizza Place	0.12
4	Dance Studio	0.06

----Laurelton----

	venue	freq
0	Caribbean Restaurant	0.50
1	Cosmetics Shop	0.17
2	Train Station	0.17
3	Park	0.17
4	Accessories Store	0.00

----Lefrak City----

	venue	freq
0	Department Store	0.11

1	Cosmetics Shop	0.07
2	Bakery	0.07
3	Bank	0.04
4	Spa	0.04

----Lindenwood----

	venue	freq
0	Deli / Bodega	0.09
1	Hotel	0.09
2	Gym	0.09
3	Bank	0.09
4	Bakery	0.09

----Little Neck----

	venue	freq
0	Coffee Shop	0.06
1	Italian Restaurant	0.06
2	Spa	0.06
3	Deli / Bodega	0.06
4	Chinese Restaurant	0.06

----Long Island City----

	venue	freq
0	Coffee Shop	0.09
1	Hotel	0.09
2	Pizza Place	0.08
3	Café	0.06
4	Mexican Restaurant	0.06

----Malba----

	venue	freq
0	Tennis Court	0.33
1	Scenic Lookout	0.33
2	Rest Area	0.33
3	Accessories Store	0.00
4	North Indian Restaurant	0.00

----Maspeth----

	venue	freq
0	Pizza Place	0.10
1	Deli / Bodega	0.07
2	Bank	0.07
3	Mobile Phone Shop	0.07
4	Grocery Store	0.07

----Middle Village----

	venue	freq
0	Italian Restaurant	0.07
1	Bakery	0.07
2	Discount Store	0.07
3	Cosmetics Shop	0.07

4 Playground 0.07

----Murray Hill----

	venue	freq
0	Korean Restaurant	0.48
1	Coffee Shop	0.05
2	Supermarket	0.05
3	Bar	0.05
4	Pub	0.02

----Neponsit----

	venue	freq
0	Beach	0.57
1	Bar	0.14
2	Beach Bar	0.14
3	Scenic Lookout	0.14
4	Accessories Store	0.00

----North Corona----

	venue	freq
0	Deli / Bodega	0.18
1	Bakery	0.14
2	Pizza Place	0.09
3	Gym / Fitness Center	0.09
4	Mexican Restaurant	0.05

----Oakland Gardens----

	venue	freq
0	Chinese Restaurant	0.12
1	Korean Restaurant	0.12
2	Yoga Studio	0.08
3	Donut Shop	0.08
4	Bagel Shop	0.04

----Ozone Park----

	venue	freq
0	Pharmacy	0.12
1	Diner	0.06
2	Gym	0.06
3	Pizza Place	0.06
4	Bank	0.06

----Pomonok----

	venue	freq
0	Bus Station	0.11
1	Supermarket	0.11
2	Bar	0.11
3	Scenic Lookout	0.11
4	Bowling Alley	0.11

----Queens Village----

	venue	freq
0	Bank	0.12
1	Bus Stop	0.06
2	Pizza Place	0.06
3	Fast Food Restaurant	0.06
4	Martial Arts Dojo	0.06

----Queensboro Hill----

	venue	freq
0	Chinese Restaurant	0.15
1	Bakery	0.07
2	Bus Station	0.07
3	Bank	0.07
4	Indian Restaurant	0.04

----Queensbridge----

	venue	freq
0	Hotel	0.29
1	Sandwich Place	0.12
2	Baseball Field	0.06
3	Platform	0.06
4	Performing Arts Venue	0.06

----Ravenswood----

	venue	freq
0	Deli / Bodega	0.12
1	Grocery Store	0.08
2	Indian Restaurant	0.04
3	Fried Chicken Joint	0.04
4	Brazilian Restaurant	0.04

----Rego Park----

	venue	freq
0	Bakery	0.12
1	Sandwich Place	0.07
2	Grocery Store	0.07
3	Chinese Restaurant	0.07
4	Donut Shop	0.05

----Richmond Hill----

	venue	freq
0	Latin American Restaurant	0.09
1	Lounge	0.09
2	Pizza Place	0.09
3	Discount Store	0.09
4	Caribbean Restaurant	0.09

----Ridgewood----

	venue	freq
0	Italian Restaurant	0.08

1	Pet Store	0.05
2	Bakery	0.05
3	Grocery Store	0.05
4	Greek Restaurant	0.05

----Rochdale----

	venue	freq
0	Market	0.13
1	Sandwich Place	0.07
2	Candy Store	0.07
3	Southern / Soul Food Restaurant	0.07
4	Breakfast Spot	0.07

----Rockaway Beach----

	venue	freq
0	Beach	0.17
1	Ice Cream Shop	0.07
2	Arepa Restaurant	0.04
3	Pizza Place	0.04
4	Latin American Restaurant	0.04

----Rockaway Park----

	venue	freq
0	Beach	0.12
1	Pizza Place	0.08
2	Pharmacy	0.08
3	Bank	0.08
4	Sandwich Place	0.04

----Rosedale----

	venue	freq
0	Bus Station	0.14
1	Fried Chicken Joint	0.14
2	Pharmacy	0.07
3	Supermarket	0.07
4	Caribbean Restaurant	0.07

----Roxbury----

	venue	freq
0	Beach	0.22
1	Pizza Place	0.11
2	Fast Food Restaurant	0.11
3	Trail	0.11
4	Pub	0.11

----Somerville----

	venue	freq
0	Park	1.0
1	Accessories Store	0.0
2	Moving Target	0.0
3	Multiplex	0.0

4 Museum 0.0

----South Jamaica----

	venue	freq
0	Bus Station	0.2
1	Supermarket	0.1
2	Grocery Store	0.1
3	Bakery	0.1
4	Sandwich Place	0.1

----South Ozone Park----

	venue	freq
0	Park	0.2
1	Bar	0.2
2	Hotel	0.1
3	Fast Food Restaurant	0.1
4	Sandwich Place	0.1

----Springfield Gardens----

	venue	freq
0	Fried Chicken Joint	0.15
1	Donut Shop	0.15
2	Fast Food Restaurant	0.08
3	Park	0.08
4	Gym	0.08

----St. Albans----

	venue	freq
0	Caribbean Restaurant	0.21
1	Deli / Bodega	0.14
2	Chinese Restaurant	0.07
3	Discount Store	0.07
4	Shopping Mall	0.07

----Steinway----

	venue	freq
0	Rental Car Location	0.11
1	Italian Restaurant	0.07
2	Cosmetics Shop	0.07
3	Deli / Bodega	0.07
4	Sushi Restaurant	0.07

----Sunnyside----

	venue	freq
0	Pizza Place	0.09
1	South American Restaurant	0.06
2	Italian Restaurant	0.06
3	Chinese Restaurant	0.06
4	Discount Store	0.06

```

----Sunnyside Gardens----
      venue  freq
0          Bar  0.05
1  American Restaurant  0.04
2      Grocery Store  0.04
3          Pharmacy  0.03
4  Filipino Restaurant  0.03

```

```

----Utopia----
      venue  freq
0  Deli / Bodega  0.11
1  Basketball Court  0.11
2  South American Restaurant  0.06
3          Bus Stop  0.06
4      History Museum  0.06

```

```

----Whitestone----
      venue  freq
0  Deli / Bodega  0.25
1  Bubble Tea Shop  0.25
2  Moving Target  0.25
3          Train  0.25
4  North Indian Restaurant  0.00

```

```

----Woodhaven----
      venue  freq
0  Deli / Bodega  0.12
1          Park  0.08
2          Bank  0.08
3      Pharmacy  0.08
4      Nail Salon  0.04

```

```

----Woodside----
      venue  freq
0  Thai Restaurant  0.06
1          Pub  0.06
2      Grocery Store  0.06
3      Pizza Place  0.04
4  American Restaurant  0.04

```

```

In [65]: #Function to sort venues in descending order
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]

```

Top venues for each neighborhood

```

In [66]: 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'] = queens_grouped['Neighborhood']

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

neighborhoods_venues_sorted

```

Out[66]:

	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
0	Arverne	Surf Spot	Metro Station	Playground	Donut Shop	Bus Stop	Board Shop
1	Astoria	Middle Eastern Restaurant	Hookah Bar	Bar	Bakery	Greek Restaurant	Seafood Restaurant
2	Astoria Heights	Deli / Bodega	Plaza	Bus Station	Burger Joint	Bakery	Supermarket
3	Auburndale	Ice Cream Shop	Train Station	Mobile Phone Shop	Deli / Bodega	Fast Food Restaurant	Furniture / Home Store
4	Bay Terrace	Clothing Store	Women's Store	Lingerie Store	Cosmetics Shop	Donut Shop	Mobile Phone Shop
5	Bayside	Bar	Indian Restaurant	Donut Shop	Pizza Place	Pub	American Restaurant
6	Bayswater	Park	Playground	Yoga Studio	Filipino Restaurant	Fabric Shop	Falafel Restaurant
7	Beechhurst	Gym	Yoga Studio	Gym / Fitness Center	Pizza Place	Convenience Store	Dessert Shop
8	Bellaire	IT Services	Laundromat	Convenience Store	Construction & Landscaping	Coffee Shop	Chinese Restaurant
9	Belle Harbor	Beach	Spa	Deli / Bodega	Boutique	Donut Shop	Mexican Restaurant
10	Bellerose	Pizza Place	Flower Shop	Italian Restaurant	Chinese Restaurant	Motel	Fast Food Restaurant
11	Blissville	Hotel	Donut Shop	Deli / Bodega	Hostel	Rental Service	Movie Theater
12	Breezy Point	Beach	Board Shop	Monument / Landmark	Lighthouse	Trail	Supermarket
13	Briarwood	Deli / Bodega	Gym	Convenience Store	Fast Food Restaurant	Pet Store	Arts & Crafts Store
14	Broad Channel	Dive Bar	Pizza Place	Bus Station	Other Nightlife	Deli / Bodega	Food
15	Brookville	Deli / Bodega	Yoga Studio	Fish & Chips Shop	Fabric Shop	Falafel Restaurant	Farm
16	Cambria Heights	Caribbean Restaurant	Cosmetics Shop	Flower Shop	Pharmacy	Liquor Store	Gym / Fitness Center
17	College Point	Deli / Bodega	Pharmacy	Pizza Place	Latin American Restaurant	Karaoke Bar	Bakery
18	Corona	Mexican Restaurant	Bakery	Deli / Bodega	Pizza Place	Convenience Store	Restaurant
19	Douglaston	Deli / Bodega	Chinese Restaurant	Italian Restaurant	Bank	Pharmacy	Diner

	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
20	East Elmhurst	Donut Shop	Caribbean Restaurant	Coffee Shop	Flower Shop	Chinese Restaurant	Latin American Restaurant
21	Edgemere	Pizza Place	Metro Station	Skate Park	Asian Restaurant	Seafood Restaurant	Scenic Lookout
22	Elmhurst	Thai Restaurant	Mexican Restaurant	Bubble Tea Shop	Chinese Restaurant	South American Restaurant	Vietnamese Restaurant
23	Far Rockaway	Grocery Store	Pizza Place	Chinese Restaurant	Deli / Bodega	Fast Food Restaurant	Pharmacy
24	Floral Park	Indian Restaurant	Grocery Store	Pizza Place	Hobby Shop	Basketball Court	Fast Food Restaurant
25	Flushing	Bubble Tea Shop	Chinese Restaurant	Karaoke Bar	Asian Restaurant	Bakery	Hotpot Restaurant
26	Forest Hills	Gym	Yoga Studio	Thai Restaurant	Gym / Fitness Center	Park	Pharmacy
27	Forest Hills Gardens	Bakery	Food	Grocery Store	Ice Cream Shop	Bagel Shop	Pizza Place
28	Fresh Meadows	Bus Station	Chinese Restaurant	Hotel	Vegetarian / Vegan Restaurant	Food Truck	Moving Target
29	Glen Oaks	Pharmacy	Department Store	Bagel Shop	Middle Eastern Restaurant	Mexican Restaurant	Fast Food Restaurant
30	Glendale	Dance Studio	Brewery	Food & Drink Shop	Dim Sum Restaurant	Pizza Place	Arts & Crafts Store
31	Hammels	Beach	Deli / Bodega	Gym / Fitness Center	Food Truck	Southern / Soul Food Restaurant	Bus Station
32	Hillcrest	Ice Cream Shop	Grocery Store	Dim Sum Restaurant	Convenience Store	Donut Shop	College Stadium
33	Hollis	Baseball Field	Park	Shopping Mall	Sandwich Place	Fried Chicken Joint	Fast Food Restaurant
34	Holliswood	Mobile Phone Shop	Playground	Supermarket	Donut Shop	Filipino Restaurant	Fabric Shop
35	Howard Beach	Bagel Shop	Italian Restaurant	Deli / Bodega	Clothing Store	Fast Food Restaurant	Chinese Restaurant
36	Hunters Point	Italian Restaurant	Café	Brewery	Japanese Restaurant	Deli / Bodega	Comedy Club
37	Jackson Heights	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Mobile Phone Shop	Bakery	Mexican Restaurant

	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
38	Jamaica Center	Mobile Phone Shop	Coffee Shop	Caribbean Restaurant	Pizza Place	Performing Arts Venue	Mexican Restaurant
39	Jamaica Estates	Dog Run	Indian Restaurant	Lounge	Grocery Store	Yoga Studio	Falafel Restaurant
40	Jamaica Hills	Pharmacy	Fried Chicken Joint	Indian Restaurant	Donut Shop	Fast Food Restaurant	Chinese Restaurant
41	Kew Gardens	Chinese Restaurant	Deli / Bodega	Indian Restaurant	Bar	Bank	Bakery
42	Kew Gardens Hills	Playground	Pizza Place	Bus Station	Bank	Bagel Shop	Dance Studio
43	Laurelton	Caribbean Restaurant	Cosmetics Shop	Park	Train Station	Yoga Studio	Fast Food Restaurant
44	Lefrak City	Department Store	Bakery	Cosmetics Shop	Ice Cream Shop	Spa	Fried Chicken Joint
45	Lindenwood	Deli / Bodega	Japanese Restaurant	Fast Food Restaurant	Chinese Restaurant	Gym	Bank
46	Little Neck	Spa	Chinese Restaurant	Coffee Shop	Deli / Bodega	Italian Restaurant	Bus Station
47	Long Island City	Coffee Shop	Hotel	Pizza Place	Café	Mexican Restaurant	Bar
48	Malba	Tennis Court	Scenic Lookout	Rest Area	Farmers Market	Empanada Restaurant	Event Space
49	Maspeth	Pizza Place	Mobile Phone Shop	Bank	Diner	Grocery Store	Deli / Bodega
50	Middle Village	Cosmetics Shop	Bakery	Pizza Place	Dessert Shop	Diner	Discount Store
51	Murray Hill	Korean Restaurant	Supermarket	Coffee Shop	Bar	Dessert Shop	BBQ Joint
52	Neponsit	Beach	Bar	Scenic Lookout	Beach Bar	Event Space	Falafel Restaurant
53	North Corona	Deli / Bodega	Bakery	Pizza Place	Gym / Fitness Center	Hotel Bar	Supermarket
54	Oakland Gardens	Chinese Restaurant	Korean Restaurant	Yoga Studio	Donut Shop	Taiwanese Restaurant	Bagel Shop
55	Ozone Park	Pharmacy	Bank	Diner	Pizza Place	Gym	Ice Cream Shop
56	Pomonok	Bowling Alley	Playground	American Restaurant	Scenic Lookout	Pizza Place	Bus Station
57	Queens Village	Bank	Mexican Restaurant	Bakery	Mobile Phone Shop	Bus Station	Bus Stop

	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
58	Queensboro Hill	Chinese Restaurant	Bus Station	Bank	Bakery	Park	Frozen Yogurt Shop
59	Queensbridge	Hotel	Sandwich Place	Athletics & Sports	Ramen Restaurant	Performing Arts Venue	Park
60	Ravenswood	Deli / Bodega	Grocery Store	BBQ Joint	Café	Chinese Restaurant	Fried Chicken Joint
61	Rego Park	Bakery	Chinese Restaurant	Sandwich Place	Grocery Store	Pharmacy	Restaurant
62	Richmond Hill	Lounge	Pizza Place	Latin American Restaurant	Caribbean Restaurant	Discount Store	Deli / Bodega
63	Ridgewood	Italian Restaurant	Pharmacy	Grocery Store	Pizza Place	Pet Store	Bank
64	Rochdale	Market	Food Court	Candy Store	Sandwich Place	Bank	Liquor Store
65	Rockaway Beach	Beach	Ice Cream Shop	BBQ Joint	Food Truck	Brazilian Restaurant	Latin American Restaurant
66	Rockaway Park	Beach	Pizza Place	Pharmacy	Bank	Ice Cream Shop	Gift Shop
67	Rosedale	Bus Station	Fried Chicken Joint	Accessories Store	Baseball Field	Pharmacy	Sandwich Place
68	Roxbury	Beach	Pub	Deli / Bodega	Pizza Place	Irish Pub	Trail
69	Somerville	Park	Electronics Store	Food & Drink Shop	Food	Flower Shop	Flea Market
70	South Jamaica	Bus Station	Grocery Store	Dessert Shop	Pizza Place	Bakery	Sandwich Place
71	South Ozone Park	Bar	Park	Food Truck	Donut Shop	Deli / Bodega	Hotel
72	Springfield Gardens	Donut Shop	Fried Chicken Joint	Laundromat	Liquor Store	Gym	Chinese Restaurant
73	St. Albans	Caribbean Restaurant	Deli / Bodega	Donut Shop	Fried Chicken Joint	Café	Shopping Mall
74	Steinway	Rental Car Location	Cosmetics Shop	Deli / Bodega	Sushi Restaurant	Italian Restaurant	Bagel Shop
75	Sunnyside	Pizza Place	Discount Store	South American Restaurant	Chinese Restaurant	Bakery	Italian Restaurant
76	Sunnyside Gardens	Bar	Grocery Store	American Restaurant	Thai Restaurant	Filipino Restaurant	Korean Restaurant

	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
77	Utopia	Deli / Bodega	Basketball Court	Automotive Shop	Pizza Place	Donut Shop	Paper / Office Supplies Store
78	Whitestone	Bubble Tea Shop	Deli / Bodega	Moving Target	Train	Fish & Chips Shop	Falafel Restaurant
79	Woodhaven	Deli / Bodega	Pharmacy	Park	Bank	Metro Station	Restaurant
80	Woodside	Grocery Store	Pub	Thai Restaurant	Latin American Restaurant	American Restaurant	Pizza Place

```
In [67]: neighborhoods_venues_sorted.iloc[47,]
```

```
Out[67]: Neighborhood      Long Island City
1st Most Common Venue      Coffee Shop
2nd Most Common Venue      Hotel
3rd Most Common Venue      Pizza Place
4th Most Common Venue      Café
5th Most Common Venue      Mexican Restaurant
6th Most Common Venue      Bar
7th Most Common Venue      Donut Shop
8th Most Common Venue      Gym / Fitness Center
9th Most Common Venue      Bus Station
10th Most Common Venue     Chinese Restaurant
Name: 47, dtype: object
```

Cluster the Queens Borough using K-Means

```
In [68]: # set number of clusters
kclusters = 5

queens_grouped_clustering = queens_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=2).fit(queens_grouped_clustering)

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

```
Out[68]: array([0, 0, 0, 0, 0, 0, 4, 0, 0, 0, 0, 0, 2, 0, 0, 1, 3, 0, 0, 0, 0, 0, 0,
                0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0,
                0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 2,
                4, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], dtype=int32)
```

Dataframe that includes the cluster of each neighborhood

```
In [69]: queens_merged = queens_data

# add clustering labels
queens_merged['Cluster Labels'] = kmeans.labels_

# merge toronto_grouped with toronto_data to add Latitude/Longitude for each neighborhood
queens_merged = queens_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

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

Out[69]:

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4
0	Queens	Astoria	40.768509	-73.915654	0	Middle Eastern Restaurant	Hookah Bar	Bar	
1	Queens	Woodside	40.746349	-73.901842	0	Grocery Store	Pub	Thai Restaurant	A Re
2	Queens	Jackson Heights	40.751981	-73.882821	0	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	
3	Queens	Elmhurst	40.744049	-73.881656	0	Thai Restaurant	Mexican Restaurant	Bubble Tea Shop	Re
4	Queens	Howard Beach	40.654225	-73.838138	0	Bagel Shop	Italian Restaurant	Deli / Bodega	

```

In [70]: # create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

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

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(queens_merged['Latitude'], queens_merged['Longitude'], queens_merged['Neighborhood'], queens_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[70]:



Examine Queens Cluster Neighborhood

Now, you can examine each cluster and determine the discriminating venue categories that distinguish each cluster. Based on the defining categories, you can then assign a name to each cluster.

```
In [71]: queens_cluster_0 = queens_merged.loc[queens_merged['Cluster Labels'] == 0, queens_merged.columns[[1] + list(range(4, queens_merged.shape[1]))]]

queens_cluster_1 = queens_merged.loc[queens_merged['Cluster Labels'] == 1, queens_merged.columns[[1] + list(range(4, queens_merged.shape[1]))]]

queens_cluster_2 = queens_merged.loc[queens_merged['Cluster Labels'] == 2, queens_merged.columns[[1] + list(range(4, queens_merged.shape[1]))]]

queens_cluster_3 = queens_merged.loc[queens_merged['Cluster Labels'] == 3, queens_merged.columns[[1] + list(range(4, queens_merged.shape[1]))]]

queens_cluster_4 = queens_merged.loc[queens_merged['Cluster Labels'] == 4, queens_merged.columns[[1] + list(range(4, queens_merged.shape[1]))]]
```

In [72]: `queens_cluster_0`

Out[72]:

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	
0	Astoria	0	Middle Eastern Restaurant	Hookah Bar	Bar	Bakery	Greek Restaurant	R
1	Woodside	0	Grocery Store	Pub	Thai Restaurant	Latin American Restaurant	American Restaurant	Piz
2	Jackson Heights	0	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Mobile Phone Shop	Bakery	R
3	Elmhurst	0	Thai Restaurant	Mexican Restaurant	Bubble Tea Shop	Chinese Restaurant	South American Restaurant	Vie R
4	Howard Beach	0	Bagel Shop	Italian Restaurant	Deli / Bodega	Clothing Store	Fast Food Restaurant	R
5	Corona	0	Mexican Restaurant	Bakery	Deli / Bodega	Pizza Place	Convenience Store	R
7	Kew Gardens	0	Chinese Restaurant	Deli / Bodega	Indian Restaurant	Bar	Bank	
8	Richmond Hill	0	Lounge	Pizza Place	Latin American Restaurant	Caribbean Restaurant	Discount Store	
9	Flushing	0	Bubble Tea Shop	Chinese Restaurant	Karaoke Bar	Asian Restaurant	Bakery	R
10	Long Island City	0	Coffee Shop	Hotel	Pizza Place	Café	Mexican Restaurant	
11	Sunnyside	0	Pizza Place	Discount Store	South American Restaurant	Chinese Restaurant	Bakery	R
13	Maspeth	0	Pizza Place	Mobile Phone Shop	Bank	Diner	Grocery Store	
14	Ridgewood	0	Italian Restaurant	Pharmacy	Grocery Store	Pizza Place	Pet Store	
17	Woodhaven	0	Deli / Bodega	Pharmacy	Park	Bank	Metro Station	R
18	Ozone Park	0	Pharmacy	Bank	Diner	Pizza Place	Gym	lc
19	South Ozone Park	0	Bar	Park	Food Truck	Donut Shop	Deli / Bodega	
20	College Point	0	Deli / Bodega	Pharmacy	Pizza Place	Latin American Restaurant	Karaoke Bar	
21	Whitestone	0	Bubble Tea Shop	Deli / Bodega	Moving Target	Train	Fish & Chips Shop	R
22	Bayside	0	Bar	Indian Restaurant	Donut Shop	Pizza Place	Pub	' R

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	
23	Auburndale	0	Ice Cream Shop	Train Station	Mobile Phone Shop	Deli / Bodega	Fast Food Restaurant	F Ho
24	Little Neck	0	Spa	Chinese Restaurant	Coffee Shop	Deli / Bodega	Italian Restaurant	Bu
25	Douglaston	0	Deli / Bodega	Chinese Restaurant	Italian Restaurant	Bank	Pharmacy	
26	Glen Oaks	0	Pharmacy	Department Store	Bagel Shop	Middle Eastern Restaurant	Mexican Restaurant	F R
27	Bellerose	0	Pizza Place	Flower Shop	Italian Restaurant	Chinese Restaurant	Motel	F R
28	Kew Gardens Hills	0	Playground	Pizza Place	Bus Station	Bank	Bagel Shop	
29	Fresh Meadows	0	Bus Station	Chinese Restaurant	Hotel	Vegetarian / Vegan Restaurant	Food Truck	
30	Briarwood	0	Deli / Bodega	Gym	Convenience Store	Fast Food Restaurant	Pet Store	Arts
32	Oakland Gardens	0	Chinese Restaurant	Korean Restaurant	Yoga Studio	Donut Shop	Taiwanese Restaurant	Ba
33	Queens Village	0	Bank	Mexican Restaurant	Bakery	Mobile Phone Shop	Bus Station	
34	Hollis	0	Baseball Field	Park	Shopping Mall	Sandwich Place	Fried Chicken Joint	F R
35	South Jamaica	0	Bus Station	Grocery Store	Dessert Shop	Pizza Place	Bakery	\$
36	St. Albans	0	Caribbean Restaurant	Deli / Bodega	Donut Shop	Fried Chicken Joint	Café	:
37	Rochdale	0	Market	Food Court	Candy Store	Sandwich Place	Bank	Liq
38	Springfield Gardens	0	Donut Shop	Fried Chicken Joint	Laundromat	Liquor Store	Gym	R
39	Cambria Heights	0	Caribbean Restaurant	Cosmetics Shop	Flower Shop	Pharmacy	Liquor Store	
40	Rosedale	0	Bus Station	Fried Chicken Joint	Accessories Store	Baseball Field	Pharmacy	\$
41	Far Rockaway	0	Grocery Store	Pizza Place	Chinese Restaurant	Deli / Bodega	Fast Food Restaurant	F
42	Broad Channel	0	Dive Bar	Pizza Place	Bus Station	Other Nightlife	Deli / Bodega	

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	
44	Steinway	0	Rental Car Location	Cosmetics Shop	Deli / Bodega	Sushi Restaurant	Italian Restaurant	Bar
45	Beechhurst	0	Gym	Yoga Studio	Gym / Fitness Center	Pizza Place	Convenience Store	
46	Bay Terrace	0	Clothing Store	Women's Store	Lingerie Store	Cosmetics Shop	Donut Shop	Pharmacy
47	Edgemere	0	Pizza Place	Metro Station	Skate Park	Asian Restaurant	Seafood Restaurant	
48	Arverne	0	Surf Spot	Metro Station	Playground	Donut Shop	Bus Stop	Bar
49	Rockaway Beach	0	Beach	Ice Cream Shop	BBQ Joint	Food Truck	Brazilian Restaurant	Bar / Restaurant
50	Neponsit	0	Beach	Bar	Scenic Lookout	Beach Bar	Event Space	Restaurant
51	Murray Hill	0	Korean Restaurant	Supermarket	Coffee Shop	Bar	Dessert Shop	Eatery
53	Holliswood	0	Mobile Phone Shop	Playground	Supermarket	Donut Shop	Filipino Restaurant	Farmers Market
54	Jamaica Estates	0	Dog Run	Indian Restaurant	Lounge	Grocery Store	Yoga Studio	Restaurant
55	Queensboro Hill	0	Chinese Restaurant	Bus Station	Bank	Bakery	Park	Yoga Studio
56	Hillcrest	0	Ice Cream Shop	Grocery Store	Dim Sum Restaurant	Convenience Store	Donut Shop	
57	Ravenswood	0	Deli / Bodega	Grocery Store	BBQ Joint	Café	Chinese Restaurant	
58	Lindenwood	0	Deli / Bodega	Japanese Restaurant	Fast Food Restaurant	Chinese Restaurant	Gym	
59	Laurelton	0	Caribbean Restaurant	Cosmetics Shop	Park	Train Station	Yoga Studio	Farmers Market / Restaurant
60	Lefrak City	0	Department Store	Bakery	Cosmetics Shop	Ice Cream Shop	Spa	
61	Belle Harbor	0	Beach	Spa	Deli / Bodega	Boutique	Donut Shop	Restaurant
62	Rockaway Park	0	Beach	Pizza Place	Pharmacy	Bank	Ice Cream Shop	
63	Somerville	0	Park	Electronics Store	Food & Drink Shop	Food	Flower Shop	Farmers Market
64	Brookville	0	Deli / Bodega	Yoga Studio	Fish & Chips Shop	Fabric Shop	Falafel Restaurant	

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	
66	North Corona	0	Deli / Bodega	Bakery	Pizza Place	Gym / Fitness Center	Hotel Bar	Sup
67	Forest Hills Gardens	0	Bakery	Food	Grocery Store	Ice Cream Shop	Bagel Shop	Piz
70	Pomonok	0	Bowling Alley	Playground	American Restaurant	Scenic Lookout	Pizza Place	Bu
71	Astoria Heights	0	Deli / Bodega	Plaza	Bus Station	Burger Joint	Bakery	Sup
72	Hunters Point	0	Italian Restaurant	Café	Brewery	Japanese Restaurant	Deli / Bodega	
73	Sunnyside Gardens	0	Bar	Grocery Store	American Restaurant	Thai Restaurant	Filipino Restaurant	R
74	Blissville	0	Hotel	Donut Shop	Deli / Bodega	Hostel	Rental Service	
75	Roxbury	0	Beach	Pub	Deli / Bodega	Pizza Place	Irish Pub	
76	Middle Village	0	Cosmetics Shop	Bakery	Pizza Place	Dessert Shop	Diner	
77	Malba	0	Tennis Court	Scenic Lookout	Rest Area	Farmers Market	Empanada Restaurant	Eve
78	Hammels	0	Beach	Deli / Bodega	Gym / Fitness Center	Food Truck	Southern / Soul Food Restaurant	Bu
79	Bayswater	0	Park	Playground	Yoga Studio	Filipino Restaurant	Fabric Shop	R
80	Queensbridge	0	Hotel	Sandwich Place	Athletics & Sports	Ramen Restaurant	Performing Arts Venue	

In [73]: queens_cluster_1

Out[73]:

	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
15	Glendale	1	Dance Studio	Brewery	Food & Drink Shop	Dim Sum Restaurant	Pizza Place	Arts & Crafts Store	Fis Cr Sl

In [74]: queens_cluster_2

Out[74]:

	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
12	East Elmhurst	2	Donut Shop	Caribbean Restaurant	Coffee Shop	Flower Shop	Chinese Restaurant	Latin American Restaurant
31	Jamaica Center	2	Mobile Phone Shop	Coffee Shop	Caribbean Restaurant	Pizza Place	Performing Arts Venue	Mexican Restaurant
52	Floral Park	2	Indian Restaurant	Grocery Store	Pizza Place	Hobby Shop	Basketball Court	Fast Food Restaurant
65	Bellaire	2	IT Services	Laundromat	Convenience Store	Construction & Landscaping	Coffee Shop	Chinese Restaurant
68	Jamaica Hills	2	Pharmacy	Fried Chicken Joint	Indian Restaurant	Donut Shop	Fast Food Restaurant	Chinese Restaurant

In [75]: queens_cluster_3

Out[75]:

	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
16	Rego Park	3	Bakery	Chinese Restaurant	Sandwich Place	Grocery Store	Pharmacy	Restaurant
43	Breezy Point	3	Beach	Board Shop	Monument / Landmark	Lighthouse	Trail	Supermarket

In [76]: queens_cluster_4

Out[76]:

	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
6	Forest Hills	4	Gym	Yoga Studio	Thai Restaurant	Gym / Fitness Center	Park	Pharmacy	
69	Utopia	4	Deli / Bodega	Basketball Court	Automotive Shop	Pizza Place	Donut Shop	Paper / Office Supplies Store	Coffee Shop