

Segmenting and Clustering Toronto

February 3, 2019

0.1 Segmenting and Clustering Neighborhoods in Toronto

0.1.1 Project Objectives:

- Scrap Toronto Neighbourhood Data from Wikipedia
- wrangle the data, clean it, and then read it into a pandas dataframe so that it is in a structured format like the New York dataset
- explore and cluster the neighborhoods in the city of Toronto

0.1.2 Import libraries needed for the project

```
In [1]: # -*- coding: utf-8 -*-
import os,sys
import urllib
import requests
from urllib.request import urlopen
from bs4 import BeautifulSoup
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
from geopy.geocoders import Nominatim # convert an address into latitude and longitude v
import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe
# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors
# import k-means from clustering stage
from sklearn.cluster import KMeans
#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you haven't c
import folium # map rendering library
print('Libraries imported.')
```

Libraries imported.

0.1.3 Scrap Toronto Neighbourhood data from Wikipedia using postal codes

```
In [2]: URL = "https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"
        page = urlopen(URL)
        soup = BeautifulSoup(page)
        page.close()

        fp = open("data.csv", "w")
        tables = soup.findAll('table')
        tab = tables[0]
        for tr in tab.tbody.findAll('tr'):
            #print(tr.findAll('th'))
            for th in tr.findAll('th'):
                text = th.getText().strip()+','
                fp.write(text)
            for td in tr.findAll('td'):
                text = td.getText().strip()+','
                fp.write(text)
            fp.write('\n')
        fp.close()
```

0.1.4 Represent data in a Dataframe

```
In [3]: dfs = pd.read_csv('data.csv')
        dfs.drop('Unnamed: 3', axis=1, inplace = True)
        dfs.head()
```

```
Out[3]:
```

	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

```
In [4]: dfs.shape
```

```
Out[4]: (289, 3)
```

```
In [5]: dfs1 = dfs[ ~ dfs['Borough'].str.contains('Not assigned')]
        dfs1.shape
```

```
Out[5]: (212, 3)
```

```
In [6]: grouped = dfs1.groupby(['Postcode', 'Borough'], as_index=False)
        dfs2 = pd.DataFrame(grouped.sum())
        dfs2.head()
```

```
Out[6]:
```

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	RougeMalvern
1	M1C	Scarborough	Highland CreekRouge HillPort Union

2	M1E	Scarborough	GuildwoodMorningsideWest Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

```
In [7]: dfs2.shape
```

```
Out[7]: (103, 3)
```

```
In [9]: for i in range(len(dfs2)):
        #if dfs2.iloc[i,2] == "Not assigned"
        line_data=dfs2.iloc[i,:]
        if line_data['Neighbourhood'] == 'Not assigned':
            line_data['Neighbourhood'] = line_data['Borough']
dfs2.head()
```

```
Out[9]:
```

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	RougeMalvern
1	M1C	Scarborough	Highland CreekRouge HillPort Union
2	M1E	Scarborough	GuildwoodMorningsideWest Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

0.1.5 Use Geospatial_Coordinates.csv to trace dataframe downloaded and save in the path

```
In [10]: geo_data = pd.read_csv("Geospatial_Coordinates.csv")
        geo_data.rename(columns={'Postal Code':'Postcode'}, inplace = True)
        geo_data.head()
```

```
Out[10]:
```

	Postcode	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

0.1.6 Merge dfs2 and geo_data by column "Postcode"

```
In [12]: dfs3 = pd.merge(dfs2, geo_data, on='Postcode')
        dfs3.head()
```

```
Out[12]:
```

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	RougeMalvern	43.806686	-79.194353
1	M1C	Scarborough	Highland CreekRouge HillPort Union	43.784535	-79.160497
2	M1E	Scarborough	GuildwoodMorningsideWest Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

```

1 -79.160497
2 -79.188711
3 -79.216917
4 -79.239476

```

```
In [13]: print('The dataframe has {} boroughs and {} neighborhoods.'.format(len(dfs3['Borough'])).
```

The dataframe has 11 boroughs and 103 neighborhoods.

0.1.7 Find the latitude and longitude of Toronto of Canada

```
In [14]: address = 'Toronto, CA'
         geolocator = Nominatim()
         location = geolocator.geocode(address)
         latitude = location.latitude
         longitude = location.longitude
         print('The geographical coordinate of Toronto are {}, {}'.format(latitude, longitude))
```

The geographical coordinate of Toronto are 43.653963, -79.387207.

0.1.8 Create map of Toronto using latitude and longitude values

```
In [15]: map_Toronto = folium.Map(location=[latitude, longitude], zoom_start=10)
         # add markers to map
         for lat, lng, borough, neighborhood in zip(dfs3['Latitude'], dfs3['Longitude'], dfs3['Borough'], dfs3['Neighbourhood']):
             label = '{} {}'.format(neighborhood, borough)
             popup = folium.Popup(label, parse_html=True)
             folium.CircleMarker(
                 [lat, lng],
                 radius=5,
                 popup=popup,
                 color='blue',
                 fill=True,
                 fill_color='#3186cc',
                 fill_opacity=0.7).add_to(map_Toronto)
         map_Toronto
```

```
Out[15]: <folium.folium.Map at 0x7fd989f86b00>
```

0.1.9 Scarborough will be used for the analyses

```
In [17]: Scarborough_data = dfs3[dfs3['Borough'] == 'Scarborough'].reset_index(drop=True)
         Scarborough_data.head()
```

```
Out[17]:
```

	Postcode	Borough	Neighbourhood	Latitude	\
0	M1B	Scarborough	RougeMalvern	43.806686	
1	M1C	Scarborough	Highland CreekRouge HillPort Union	43.784535	

2	M1E	Scarborough	GuildwoodMorningsideWest Hill	43.763573
3	M1G	Scarborough	Woburn	43.770992
4	M1H	Scarborough	Cedarbrae	43.773136

	Longitude
0	-79.194353
1	-79.160497
2	-79.188711
3	-79.216917
4	-79.239476

0.1.10 Get the coordinates of 'Scarborough'

```
In [18]: address = 'Scarborough, CA'
         geolocator = Nominatim()
         location = geolocator.geocode(address)
         latitude = location.latitude
         longitude = location.longitude
         print('The geograpical coordinate of Scarborough are {}, {}'.format(latitude, longitude))
```

The geograpical coordinate of Scarborough are 43.773077, -79.257774.

0.1.11 Create the map of 'Scarborough' using latitude and longitude values

```
In [19]: 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['Label']):
             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).add_to(map_Scarborough)
         map_Scarborough
```

```
Out[19]: <folium.folium.Map at 0x7fd9893fa7b8>
```

0.1.12 My foursquare ID

```
In [20]: CLIENT_ID = 'IZJNDNMGN2UGOQC4VTSC00X3FIJCWLEOMDME03AZXXKSFK1' # your Foursquare ID
         CLIENT_SECRET = '2VRK4XTUNZGWEC5FU3ITX45S3KRWYECKYHE1OVKOEJLMJHGJ' # your Foursquare Secret
         VERSION = '20180605' # Foursquare API version
         print('My foursquare credentails:')
         print('CLIENT_ID: ' + CLIENT_ID)
         print('CLIENT_SECRET: ' + CLIENT_SECRET)
```



```

Out[25]: {'meta': {'code': 200, 'requestId': '5c5761036a60717aef4cb3ea'},
  'response': {'warning': {'text': "There aren't a lot of results near you. Try something else.", 'code': 404},
    'headerLocation': 'Malvern',
    'headerFullLocation': 'Malvern, Toronto',
    'headerLocationGranularity': 'neighborhood',
    'totalResults': 2,
    'suggestedBounds': {'ne': {'lat': 43.8111863045, 'lng': -79.18812958073042},
      'sw': {'lat': 43.80218629549999, 'lng': -79.2005772192696}},
    'groups': [{'type': 'Recommended Places',
      'name': 'recommended',
      'items': [{'reasons': {'count': 0,
        'items': [{'summary': 'This spot is popular',
          'type': 'general',
          'reasonName': 'globalInteractionReason'}]}],
      'venue': {'id': '4bb6b9446edc76b0d771311c',
        'name': "Wendy's",
        'location': {'crossStreet': 'Morningside & Sheppard',
          'lat': 43.80744841934756,
          'lng': -79.19905558052072,
          'labeledLatLngs': [{'label': 'display',
            'lat': 43.80744841934756,
            'lng': -79.19905558052072}]},
        'distance': 387,
        'cc': 'CA',
        'city': 'Toronto',
        'state': 'ON',
        'country': 'Canada',
        'formattedAddress': ['Toronto ON', 'Canada']},
        'categories': [{'id': '4bf58dd8d48988d16e941735',
          'name': 'Fast Food Restaurant',
          'pluralName': 'Fast Food Restaurants',
          'shortName': 'Fast Food',
          'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/food/fastfood_',
            'suffix': '.png'},
          'primary': True}],
        'photos': {'count': 0, 'groups': []}},
      'referralId': 'e-0-4bb6b9446edc76b0d771311c-0'},
    {'reasons': {'count': 0,
      'items': [{'summary': 'This spot is popular',
        'type': 'general',
        'reasonName': 'globalInteractionReason'}]}],
      'venue': {'id': '5539e7d2498edaf4b02673ca',
        'name': 'Interprovincial Group',
        'location': {'address': '1315 Morningside Avenue',
          'lat': 43.8056297,
          'lng': -79.2003784,
          'labeledLatLngs': [{'label': 'display',
            'lat': 43.8056297,

```

```

        'lng': -79.2003784}],
    'distance': 498,
    'postalCode': 'M1B 3C5',
    'cc': 'CA',
    'city': 'Scarborough',
    'state': 'ON',
    'country': 'Canada',
    'formattedAddress': ['1315 Morningside Avenue',
        'Scarborough ON M1B 3C5',
        'Canada']},
    'categories': [{'id': '52f2ab2ebcbc57f1066b8b28',
        'name': 'Print Shop',
        'pluralName': 'Print Shops',
        'shortName': 'Print Shop',
        'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v2/shops/default_',
            'suffix': '.png'},
        'primary': True}],
    'photos': {'count': 0, 'groups': []}},
    'referralId': 'e-0-5539e7d2498edaf4b02673ca-1'}}]]}}

```

0.1.17 Function to get the category's type

In [26]: *# 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']

```

0.1.18 Check how many venues from the url_json

```

In [27]: 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[27]:

```

	name	categories	lat	lng
--	------	------------	-----	-----


```

0           Wendy's Fast Food Restaurant 43.807448 -79.199056
1 Interprovincial Group Print Shop 43.805630 -79.200378

```

```
In [28]: print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))
```

2 venues were returned by Foursquare.

0.1.19 Prepare a function to find the nearby venues

```
In [29]: 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={}&lat={}&lng={}&radius={}&limit={}'
        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']

    return(nearby_venues)

```

```
In [30]: Scarborough_venues = getNearbyVenues(names=Scarborough_data['Neighbourhood'],
                                              latitudes=Scarborough_data['Latitude'],
                                              longitudes=Scarborough_data['Longitude']
                                              )
```

```
RougeMalvern
Highland CreekRouge HillPort Union
GuildwoodMorningsideWest Hill
Woburn
Cedarbrae
Scarborough Village
East Birchmount ParkIonviewKennedy Park
ClairleaGolden MileOakridge
CliffcrestCliffsideScarborough Village West
Birch CliffCliffside West
Dorset ParkScarborough Town CentreWexford Heights
MaryvaleWexford
Agincourt
Clarks CornersSullivanTam O'Shanter
Agincourt NorthL'Amoreaux EastMillikenSteeles East
L'Amoreaux WestSteeles West
Upper Rouge
```

```
In [31]: Scarborough_venues
```

```
Out[31]:
```

	Neighborhood	Neighborhood Latitude \
0	RougeMalvern	43.806686
1	RougeMalvern	43.806686
2	Highland CreekRouge HillPort Union	43.784535
3	Highland CreekRouge HillPort Union	43.784535
4	GuildwoodMorningsideWest Hill	43.763573
5	GuildwoodMorningsideWest Hill	43.763573
6	GuildwoodMorningsideWest Hill	43.763573
7	GuildwoodMorningsideWest Hill	43.763573
8	GuildwoodMorningsideWest Hill	43.763573
9	GuildwoodMorningsideWest Hill	43.763573
10	GuildwoodMorningsideWest Hill	43.763573
11	Woburn	43.770992
12	Woburn	43.770992
13	Woburn	43.770992
14	Woburn	43.770992
15	Cedarbrae	43.773136
16	Cedarbrae	43.773136
17	Cedarbrae	43.773136
18	Cedarbrae	43.773136
19	Cedarbrae	43.773136
20	Cedarbrae	43.773136

21	Cedarbrae	43.773136
22	Scarborough Village	43.744734
23	Scarborough Village	43.744734
24	Scarborough Village	43.744734
25	East Birchmount ParkIonviewKennedy Park	43.727929
26	East Birchmount ParkIonviewKennedy Park	43.727929
27	East Birchmount ParkIonviewKennedy Park	43.727929
28	East Birchmount ParkIonviewKennedy Park	43.727929
29	East Birchmount ParkIonviewKennedy Park	43.727929
30	East Birchmount ParkIonviewKennedy Park	43.727929
31	East Birchmount ParkIonviewKennedy Park	43.727929
32	ClairleaGolden MileOakridge	43.711112
33	ClairleaGolden MileOakridge	43.711112
34	ClairleaGolden MileOakridge	43.711112
35	ClairleaGolden MileOakridge	43.711112
36	ClairleaGolden MileOakridge	43.711112
37	ClairleaGolden MileOakridge	43.711112
38	ClairleaGolden MileOakridge	43.711112
39	ClairleaGolden MileOakridge	43.711112
40	ClairleaGolden MileOakridge	43.711112
41	CliffcrestCliffsideScarborough Village West	43.716316
42	CliffcrestCliffsideScarborough Village West	43.716316
43	Birch CliffCliffside West	43.692657
44	Birch CliffCliffside West	43.692657
45	Birch CliffCliffside West	43.692657
46	Birch CliffCliffside West	43.692657
47	Dorset ParkScarborough Town CentreWexford Heights	43.757410
48	Dorset ParkScarborough Town CentreWexford Heights	43.757410
49	Dorset ParkScarborough Town CentreWexford Heights	43.757410
50	Dorset ParkScarborough Town CentreWexford Heights	43.757410
51	Dorset ParkScarborough Town CentreWexford Heights	43.757410
52	Dorset ParkScarborough Town CentreWexford Heights	43.757410
53	Dorset ParkScarborough Town CentreWexford Heights	43.757410
54	Dorset ParkScarborough Town CentreWexford Heights	43.757410
55	Dorset ParkScarborough Town CentreWexford Heights	43.757410
56	MaryvaleWexford	43.750072
57	MaryvaleWexford	43.750072
58	MaryvaleWexford	43.750072
59	MaryvaleWexford	43.750072
60	MaryvaleWexford	43.750072
61	MaryvaleWexford	43.750072
62	Agincourt	43.794200
63	Agincourt	43.794200
64	Agincourt	43.794200
65	Agincourt	43.794200
66	Agincourt	43.794200
67	Clarks CornersSullivanTam O'Shanter	43.781638
68	Clarks CornersSullivanTam O'Shanter	43.781638

69	Clarks CornersSullivanTam O'Shanter	43.781638
70	Clarks CornersSullivanTam O'Shanter	43.781638
71	Clarks CornersSullivanTam O'Shanter	43.781638
72	Clarks CornersSullivanTam O'Shanter	43.781638
73	Clarks CornersSullivanTam O'Shanter	43.781638
74	Clarks CornersSullivanTam O'Shanter	43.781638
75	Clarks CornersSullivanTam O'Shanter	43.781638
76	Agincourt NorthL'Amoreaux EastMillikenSteeles ...	43.815252
77	Agincourt NorthL'Amoreaux EastMillikenSteeles ...	43.815252
78	L'Amoreaux WestSteeles West	43.799525
79	L'Amoreaux WestSteeles West	43.799525
80	L'Amoreaux WestSteeles West	43.799525
81	L'Amoreaux WestSteeles West	43.799525
82	L'Amoreaux WestSteeles West	43.799525
83	L'Amoreaux WestSteeles West	43.799525
84	L'Amoreaux WestSteeles West	43.799525
85	L'Amoreaux WestSteeles West	43.799525
86	L'Amoreaux WestSteeles West	43.799525
87	L'Amoreaux WestSteeles West	43.799525
88	L'Amoreaux WestSteeles West	43.799525
89	L'Amoreaux WestSteeles West	43.799525
90	L'Amoreaux WestSteeles West	43.799525
91	L'Amoreaux WestSteeles West	43.799525

	Neighborhood Longitude	Venue \
0	-79.194353	Wendy's
1	-79.194353	Interprovincial Group
2	-79.160497	Royal Canadian Legion
3	-79.160497	Affordable Toronto Movers
4	-79.188711	Swiss Chalet Rotisserie & Grill
5	-79.188711	G & G Electronics
6	-79.188711	Marina Spa
7	-79.188711	Big Bite Burrito
8	-79.188711	Enterprise Rent-A-Car
9	-79.188711	Woburn Medical Centre
10	-79.188711	Eggsmart
11	-79.216917	Starbucks
12	-79.216917	Tim Hortons
13	-79.216917	Korean Grill House
14	-79.216917	Al-Hamd Biryani & Pizza
15	-79.239476	Federick Restaurant
16	-79.239476	Drupati's Roti & Doubles
17	-79.239476	Thai One On
18	-79.239476	Centennial Recreation Centre
19	-79.239476	TD Canada Trust
20	-79.239476	B&A Bakery
21	-79.239476	Popeyes Louisiana Kitchen
22	-79.239476	McCowan Park

23	-79.239476	Tumbe Cafe
24	-79.239476	Elane Barbers
25	-79.262029	Giant Tiger
26	-79.262029	Tim Hortons
27	-79.262029	Hakka No.1
28	-79.262029	Dollarama
29	-79.262029	Kennedy GO Station
30	-79.262029	Dollarama
31	-79.262029	Kennedy Station Passenger Pickup
32	-79.284577	Dairy Queen
33	-79.284577	Warden Ave & St. Clair Ave E
34	-79.284577	TTC Bus #68 Warden
35	-79.284577	Warden Subway Station
36	-79.284577	TTC Bus 102 Markham Road
37	-79.284577	Warden Station Bus Loop
38	-79.284577	Bakery On The Go
39	-79.284577	Cafe on the go
40	-79.284577	Clairlea Futbol Centre
41	-79.239476	Have A Nap Motel
42	-79.239476	Vincent's Spot
43	-79.264848	The Birchcliff
44	-79.264848	Birchmount Community Centre
45	-79.264848	Scarborough Gardens
46	-79.264848	Birchmount Stadium
47	-79.273304	Kairali
48	-79.273304	Kim Kim restaurant
49	-79.273304	El Pulgarcito
50	-79.273304	Karaikudi Chettinad South Indian Restaurant
51	-79.273304	Big Al's Pet Supercentre
52	-79.273304	Pho Vietnam
53	-79.273304	Mill St. Brewery
54	-79.273304	Salvation Army Thrift Store
55	-79.273304	Blinds To Go
56	-79.295849	Wexford Restaurant
57	-79.295849	Subway
58	-79.295849	Lebanese bakery
59	-79.295849	Wexford Heights Plaza
60	-79.295849	Scarborough Garage Door Repair
61	-79.295849	Sequoia Lounge
62	-79.262029	Panagio's Breakfast & Lunch
63	-79.262029	Twilight
64	-79.262029	Subway
65	-79.262029	Mark's
66	-79.262029	Royal Chinese Seafood Restaurant
67	-79.304302	Remezzo Italian Bistro
68	-79.304302	The Royal Chinese Restaurant
69	-79.304302	Eight Noodles
70	-79.304302	Kub Khao

71	-79.304302	Little Caesars
72	-79.304302	Popeyes
73	-79.304302	Eaton Centre USA
74	-79.304302	KFC
75	-79.304302	Gusto Pizza
76	-79.284577	Port Royal Park
77	-79.284577	Milliken Public School Playground
78	-79.318389	Mr Congee Chinese Cuisine
79	-79.318389	Phoenix Restaurant
80	-79.318389	Price Chopper
81	-79.318389	Subway
82	-79.318389	Tim Hortons
83	-79.318389	Shoppers Drug Mart
84	-79.318389	KFC
85	-79.318389	Yamamoto Japanese Cuisine
86	-79.318389	McDonald's
87	-79.318389	Pizza Pizza
88	-79.318389	Coffee Time
89	-79.318389	Eggsmart
90	-79.318389	Black's Photography
91	-79.318389	Nantha's Bakery

	Venue Latitude	Venue Longitude	Venue Category
0	43.807448	-79.199056	Fast Food Restaurant
1	43.805630	-79.200378	Print Shop
2	43.782533	-79.163085	Bar
3	43.787919	-79.162977	Moving Target
4	43.767697	-79.189914	Pizza Place
5	43.765309	-79.191537	Electronics Store
6	43.766000	-79.191000	Spa
7	43.766299	-79.190720	Mexican Restaurant
8	43.764076	-79.193406	Rental Car Location
9	43.766631	-79.192286	Medical Center
10	43.767800	-79.190466	Breakfast Spot
11	43.770037	-79.221156	Coffee Shop
12	43.770827	-79.223078	Coffee Shop
13	43.770812	-79.214502	Korean Restaurant
14	43.767585	-79.219570	Indian Restaurant
15	43.774697	-79.241142	Hakka Restaurant
16	43.775222	-79.241678	Caribbean Restaurant
17	43.774468	-79.241268	Thai Restaurant
18	43.774593	-79.236500	Athletics & Sports
19	43.774952	-79.241343	Bank
20	43.774391	-79.243877	Bakery
21	43.775930	-79.235328	Fried Chicken Joint
22	43.745089	-79.239336	Playground
23	43.744058	-79.244021	Grocery Store
24	43.743706	-79.244457	Cosmetics Shop

25	43.727447	-79.266240	Department Store
26	43.726895	-79.266157	Coffee Shop
27	43.727688	-79.266057	Chinese Restaurant
28	43.726904	-79.265886	Discount Store
29	43.732275	-79.262418	Train Station
30	43.726596	-79.266830	Discount Store
31	43.732009	-79.264537	Bus Station
32	43.710378	-79.290701	Fast Food Restaurant
33	43.712057	-79.281005	Intersection
34	43.711778	-79.279714	Bus Line
35	43.711229	-79.279602	Metro Station
36	43.711381	-79.279588	Bus Line
37	43.711241	-79.279576	Bus Station
38	43.711271	-79.279506	Bakery
39	43.711151	-79.279469	Bakery
40	43.715234	-79.286506	Soccer Field
41	43.718256	-79.240135	Motel
42	43.717002	-79.242353	American Restaurant
43	43.691666	-79.264532	Café
44	43.695175	-79.262161	General Entertainment
45	43.694647	-79.262230	Skating Rink
46	43.695323	-79.261293	College Stadium
47	43.754768	-79.277199	Indian Restaurant
48	43.753833	-79.276611	Chinese Restaurant
49	43.754790	-79.277064	Latin American Restaurant
50	43.756042	-79.276276	Indian Restaurant
51	43.759279	-79.278325	Pet Store
52	43.757770	-79.278572	Vietnamese Restaurant
53	43.759472	-79.271121	Brewery
54	43.755782	-79.276208	Thrift / Vintage Store
55	43.757147	-79.278073	Furniture / Home Store
56	43.746030	-79.293843	Breakfast Spot
57	43.746008	-79.293238	Sandwich Place
58	43.746701	-79.292896	Middle Eastern Restaurant
59	43.746136	-79.293782	Shopping Mall
60	43.751288	-79.301508	Auto Garage
61	43.745645	-79.295737	Middle Eastern Restaurant
62	43.792370	-79.260203	Breakfast Spot
63	43.791999	-79.258584	Lounge
64	43.792823	-79.259681	Sandwich Place
65	43.791179	-79.259714	Clothing Store
66	43.798496	-79.262196	Chinese Restaurant
67	43.778649	-79.308264	Italian Restaurant
68	43.780505	-79.298844	Chinese Restaurant
69	43.778234	-79.308299	Noodle House
70	43.780438	-79.299837	Thai Restaurant
71	43.780563	-79.298624	Pizza Place
72	43.780476	-79.298460	Fried Chicken Joint

73	43.783572	-79.304916	Shopping Mall
74	43.779440	-79.303371	Fast Food Restaurant
75	43.783607	-79.298983	Pizza Place
76	43.815477	-79.289773	Park
77	43.815383	-79.289867	Playground
78	43.798879	-79.318335	Chinese Restaurant
79	43.798198	-79.318432	Chinese Restaurant
80	43.799445	-79.318563	Grocery Store
81	43.798983	-79.318838	Sandwich Place
82	43.798281	-79.318317	Coffee Shop
83	43.799670	-79.319315	Pharmacy
84	43.798938	-79.318854	Fast Food Restaurant
85	43.798589	-79.318558	Japanese Restaurant
86	43.798880	-79.318724	Fast Food Restaurant
87	43.797909	-79.318113	Pizza Place
88	43.797952	-79.318678	Coffee Shop
89	43.796375	-79.318681	Breakfast Spot
90	43.797292	-79.319332	Camera Store
91	43.796430	-79.319151	Indian Restaurant

In [32]: Scarborough_data.tail()

```
Out[32]:
```

	Postcode	Borough	Neighbourhood \
12	M1S	Scarborough	Agincourt
13	M1T	Scarborough	Clarks CornersSullivanTam O'Shanter
14	M1V	Scarborough	Agincourt NorthL'Amoreaux EastMillikenSteeles ...
15	M1W	Scarborough	L'Amoreaux WestSteeles West
16	M1X	Scarborough	Upper Rouge

	Latitude	Longitude
12	43.794200	-79.262029
13	43.781638	-79.304302
14	43.815252	-79.284577
15	43.799525	-79.318389
16	43.836125	-79.205636

In [33]: Scarborough_data.drop(index=16,axis=0,inplace=True)
Scarborough_data.tail()

```
Out[33]:
```

	Postcode	Borough	Neighbourhood \
11	M1R	Scarborough	MaryvaleWexford
12	M1S	Scarborough	Agincourt
13	M1T	Scarborough	Clarks CornersSullivanTam O'Shanter
14	M1V	Scarborough	Agincourt NorthL'Amoreaux EastMillikenSteeles ...
15	M1W	Scarborough	L'Amoreaux WestSteeles West

	Latitude	Longitude
11	43.750072	-79.295849
12	43.794200	-79.262029


```

13  43.781638 -79.304302
14  43.815252 -79.284577
15  43.799525 -79.318389

```

Drop Neighbourhoods with no venues

```

In [34]: print(Scarborough_venues.shape)
         Scarborough_venues.head()

```

```

(92, 7)

```

```

Out[34]:
      Neighborhood Neighborhood Latitude \
0      RougeMalvern      43.806686
1      RougeMalvern      43.806686
2 Highland CreekRoug HillPort Union    43.784535
3 Highland CreekRoug HillPort Union    43.784535
4      GuildwoodMorningsideWest Hill    43.763573

      Neighborhood Longitude      Venue Venue Latitude \
0      -79.194353      Wendy's      43.807448
1      -79.194353      Interprovincial Group    43.805630
2      -79.160497      Royal Canadian Legion    43.782533
3      -79.160497      Affordable Toronto Movers    43.787919
4      -79.188711      Swiss Chalet Rotisserie & Grill    43.767697

      Venue Longitude      Venue Category
0      -79.199056      Fast Food Restaurant
1      -79.200378      Print Shop
2      -79.163085      Bar
3      -79.162977      Moving Target
4      -79.189914      Pizza Place

```

0.2 Groupby the venues by Neighbourhood

```

In [36]: Scarborough_venues.groupby('Neighborhood').head()

```

```

Out[36]:
      Neighborhood Neighborhood Latitude \
0      RougeMalvern      43.806686
1      RougeMalvern      43.806686
2 Highland CreekRoug HillPort Union    43.784535
3 Highland CreekRoug HillPort Union    43.784535
4      GuildwoodMorningsideWest Hill    43.763573
5      GuildwoodMorningsideWest Hill    43.763573
6      GuildwoodMorningsideWest Hill    43.763573
7      GuildwoodMorningsideWest Hill    43.763573
8      GuildwoodMorningsideWest Hill    43.763573
11      Woburn      43.770992
12      Woburn      43.770992

```

13		Woburn	43.770992
14		Woburn	43.770992
15		Cedarbrae	43.773136
16		Cedarbrae	43.773136
17		Cedarbrae	43.773136
18		Cedarbrae	43.773136
19		Cedarbrae	43.773136
22		Scarborough Village	43.744734
23		Scarborough Village	43.744734
24		Scarborough Village	43.744734
25	East Birchmount Park	IonviewKennedy Park	43.727929
26	East Birchmount Park	IonviewKennedy Park	43.727929
27	East Birchmount Park	IonviewKennedy Park	43.727929
28	East Birchmount Park	IonviewKennedy Park	43.727929
29	East Birchmount Park	IonviewKennedy Park	43.727929
32		ClairleaGolden MileOakridge	43.711112
33		ClairleaGolden MileOakridge	43.711112
34		ClairleaGolden MileOakridge	43.711112
35		ClairleaGolden MileOakridge	43.711112
36		ClairleaGolden MileOakridge	43.711112
41	Cliffcrest	CliffsideScarborough Village West	43.716316
42	Cliffcrest	CliffsideScarborough Village West	43.716316
43		Birch CliffCliffside West	43.692657
44		Birch CliffCliffside West	43.692657
45		Birch CliffCliffside West	43.692657
46		Birch CliffCliffside West	43.692657
47	Dorset Park	Scarborough Town CentreWexford Heights	43.757410
48	Dorset Park	Scarborough Town CentreWexford Heights	43.757410
49	Dorset Park	Scarborough Town CentreWexford Heights	43.757410
50	Dorset Park	Scarborough Town CentreWexford Heights	43.757410
51	Dorset Park	Scarborough Town CentreWexford Heights	43.757410
56		MaryvaleWexford	43.750072
57		MaryvaleWexford	43.750072
58		MaryvaleWexford	43.750072
59		MaryvaleWexford	43.750072
60		MaryvaleWexford	43.750072
62		Agincourt	43.794200
63		Agincourt	43.794200
64		Agincourt	43.794200
65		Agincourt	43.794200
66		Agincourt	43.794200
67	Clarks Corners	SullivanTam O'Shanter	43.781638
68	Clarks Corners	SullivanTam O'Shanter	43.781638
69	Clarks Corners	SullivanTam O'Shanter	43.781638
70	Clarks Corners	SullivanTam O'Shanter	43.781638
71	Clarks Corners	SullivanTam O'Shanter	43.781638
76	Agincourt North	L'Amoreaux EastMillikenSteeles ...	43.815252
77	Agincourt North	L'Amoreaux EastMillikenSteeles ...	43.815252

78	L'Amoreaux WestSteeles West	43.799525
79	L'Amoreaux WestSteeles West	43.799525
80	L'Amoreaux WestSteeles West	43.799525
81	L'Amoreaux WestSteeles West	43.799525
82	L'Amoreaux WestSteeles West	43.799525

	Neighborhood Longitude	Venue \
0	-79.194353	Wendy's
1	-79.194353	Interprovincial Group
2	-79.160497	Royal Canadian Legion
3	-79.160497	Affordable Toronto Movers
4	-79.188711	Swiss Chalet Rotisserie & Grill
5	-79.188711	G & G Electronics
6	-79.188711	Marina Spa
7	-79.188711	Big Bite Burrito
8	-79.188711	Enterprise Rent-A-Car
11	-79.216917	Starbucks
12	-79.216917	Tim Hortons
13	-79.216917	Korean Grill House
14	-79.216917	Al-Hamd Biryani & Pizza
15	-79.239476	Federick Restaurant
16	-79.239476	Drupati's Roti & Doubles
17	-79.239476	Thai One On
18	-79.239476	Centennial Recreation Centre
19	-79.239476	TD Canada Trust
22	-79.239476	McCowan Park
23	-79.239476	Tumbe Cafe
24	-79.239476	Elane Barbers
25	-79.262029	Giant Tiger
26	-79.262029	Tim Hortons
27	-79.262029	Hakka No.1
28	-79.262029	Dollarama
29	-79.262029	Kennedy GO Station
32	-79.284577	Dairy Queen
33	-79.284577	Warden Ave & St. Clair Ave E
34	-79.284577	TTC Bus #68 Warden
35	-79.284577	Warden Subway Station
36	-79.284577	TTC Bus 102 Markham Road
41	-79.239476	Have A Nap Motel
42	-79.239476	Vincent's Spot
43	-79.264848	The Birchcliff
44	-79.264848	Birchmount Community Centre
45	-79.264848	Scarborough Gardens
46	-79.264848	Birchmount Stadium
47	-79.273304	Kairali
48	-79.273304	Kim Kim restaurant
49	-79.273304	El Pulgarcito
50	-79.273304	Karaikudi Chettinad South Indian Restaurant

51	-79.273304	Big Al's Pet Supercentre
56	-79.295849	Wexford Restaurant
57	-79.295849	Subway
58	-79.295849	Lebanese bakery
59	-79.295849	Wexford Heights Plaza
60	-79.295849	Scarborough Garage Door Repair
62	-79.262029	Panagio's Breakfast & Lunch
63	-79.262029	Twilight
64	-79.262029	Subway
65	-79.262029	Mark's
66	-79.262029	Royal Chinese Seafood Restaurant
67	-79.304302	Remezzo Italian Bistro
68	-79.304302	The Royal Chinese Restaurant
69	-79.304302	Eight Noodles
70	-79.304302	Kub Khao
71	-79.304302	Little Caesars
76	-79.284577	Port Royal Park
77	-79.284577	Milliken Public School Playground
78	-79.318389	Mr Congee Chinese Cuisine
79	-79.318389	Phoenix Restaurant
80	-79.318389	Price Chopper
81	-79.318389	Subway
82	-79.318389	Tim Hortons

	Venue Latitude	Venue Longitude	Venue Category
0	43.807448	-79.199056	Fast Food Restaurant
1	43.805630	-79.200378	Print Shop
2	43.782533	-79.163085	Bar
3	43.787919	-79.162977	Moving Target
4	43.767697	-79.189914	Pizza Place
5	43.765309	-79.191537	Electronics Store
6	43.766000	-79.191000	Spa
7	43.766299	-79.190720	Mexican Restaurant
8	43.764076	-79.193406	Rental Car Location
11	43.770037	-79.221156	Coffee Shop
12	43.770827	-79.223078	Coffee Shop
13	43.770812	-79.214502	Korean Restaurant
14	43.767585	-79.219570	Indian Restaurant
15	43.774697	-79.241142	Hakka Restaurant
16	43.775222	-79.241678	Caribbean Restaurant
17	43.774468	-79.241268	Thai Restaurant
18	43.774593	-79.236500	Athletics & Sports
19	43.774952	-79.241343	Bank
22	43.745089	-79.239336	Playground
23	43.744058	-79.244021	Grocery Store
24	43.743706	-79.244457	Cosmetics Shop
25	43.727447	-79.266240	Department Store
26	43.726895	-79.266157	Coffee Shop

27	43.727688	-79.266057	Chinese Restaurant
28	43.726904	-79.265886	Discount Store
29	43.732275	-79.262418	Train Station
32	43.710378	-79.290701	Fast Food Restaurant
33	43.712057	-79.281005	Intersection
34	43.711778	-79.279714	Bus Line
35	43.711229	-79.279602	Metro Station
36	43.711381	-79.279588	Bus Line
41	43.718256	-79.240135	Motel
42	43.717002	-79.242353	American Restaurant
43	43.691666	-79.264532	Café
44	43.695175	-79.262161	General Entertainment
45	43.694647	-79.262230	Skating Rink
46	43.695323	-79.261293	College Stadium
47	43.754768	-79.277199	Indian Restaurant
48	43.753833	-79.276611	Chinese Restaurant
49	43.754790	-79.277064	Latin American Restaurant
50	43.756042	-79.276276	Indian Restaurant
51	43.759279	-79.278325	Pet Store
56	43.746030	-79.293843	Breakfast Spot
57	43.746008	-79.293238	Sandwich Place
58	43.746701	-79.292896	Middle Eastern Restaurant
59	43.746136	-79.293782	Shopping Mall
60	43.751288	-79.301508	Auto Garage
62	43.792370	-79.260203	Breakfast Spot
63	43.791999	-79.258584	Lounge
64	43.792823	-79.259681	Sandwich Place
65	43.791179	-79.259714	Clothing Store
66	43.798496	-79.262196	Chinese Restaurant
67	43.778649	-79.308264	Italian Restaurant
68	43.780505	-79.298844	Chinese Restaurant
69	43.778234	-79.308299	Noodle House
70	43.780438	-79.299837	Thai Restaurant
71	43.780563	-79.298624	Pizza Place
76	43.815477	-79.289773	Park
77	43.815383	-79.289867	Playground
78	43.798879	-79.318335	Chinese Restaurant
79	43.798198	-79.318432	Chinese Restaurant
80	43.799445	-79.318563	Grocery Store
81	43.798983	-79.318838	Sandwich Place
82	43.798281	-79.318317	Coffee Shop

```
In [37]: print('There are {} unique categories.'.format(len(Scarborough_venues['Venue Category'])))
```

```
There are 57 unique categories.
```

0.3 Make one hot to to mechine learning

In [38]: *# one hot encoding*

```
Scarborough_onehot = pd.get_dummies(Scarborough_venues[['Venue Category']], prefix="",
# 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.columns[:-1])
Scarborough_onehot = Scarborough_onehot[fixed_columns]
Scarborough_onehot.head()
```

Out [38]:

	Neighborhood	American Restaurant	\
0	RougeMalvern	0	
1	RougeMalvern	0	
2	Highland CreekRouge HillPort Union	0	
3	Highland CreekRouge HillPort Union	0	
4	GuildwoodMorningsideWest Hill	0	

	Athletics & Sports	Auto Garage	Bakery	Bank	Bar	Breakfast Spot	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	1	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	

	Brewery	Bus Line	Bus Station	Café	Camera Store	Caribbean Restaurant	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	0	0	

	Chinese Restaurant	Clothing Store	Coffee Shop	College Stadium	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Cosmetics Shop	Department Store	Discount Store	Electronics Store	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Fast Food Restaurant	Fried Chicken Joint	Furniture / Home Store	\
0	1	0	0	
1	0	0	0	

2	0	0	0
3	0	0	0
4	0	0	0

	General Entertainment	Grocery Store	Hakka Restaurant	Indian Restaurant	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Intersection	Italian Restaurant	Japanese Restaurant	Korean Restaurant	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Latin American Restaurant	Lounge	Medical Center	Metro Station	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Mexican Restaurant	Middle Eastern Restaurant	Motel	Moving Target	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	1	
4	0	0	0	0	

	Noodle House	Park	Pet Store	Pharmacy	Pizza Place	Playground	\
0	0	0	0	0	0	0	
1	0	0	0	0	0	0	
2	0	0	0	0	0	0	
3	0	0	0	0	0	0	
4	0	0	0	0	1	0	

	Print Shop	Rental Car Location	Sandwich Place	Shopping Mall	\
0	0	0	0	0	
1	1	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Skating Rink	Soccer Field	Spa	Thai Restaurant	Thrift / Vintage Store	\
0	0	0	0	0	0	

1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0

	Train Station	Vietnamese Restaurant
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0

In [39]: Scarborough_onehot.shape

Out[39]: (92, 58)

0.3.1 Group by Neighbourhood

In [40]: Scarborough_grouped = Scarborough_onehot.groupby('Neighborhood').mean().reset_index()
Scarborough_grouped

Out[40]:

	Neighborhood	American Restaurant	\
0	Agincourt	0.0	
1	Agincourt NorthL'Amoreaux EastMillikenSteeles ...	0.0	
2	Birch CliffCliffside West	0.0	
3	Cedarbrae	0.0	
4	ClairleaGolden MileOakridge	0.0	
5	Clarks CornersSullivanTam O'Shanter	0.0	
6	CliffcrestCliffsideScarborough Village West	0.5	
7	Dorset ParkScarborough Town CentreWexford Heights	0.0	
8	East Birchmount ParkIonviewKennedy Park	0.0	
9	GuildwoodMorningsideWest Hill	0.0	
10	Highland CreekRouge HillPort Union	0.0	
11	L'Amoreaux WestSteeles West	0.0	
12	MaryvaleWexford	0.0	
13	RougeMalvern	0.0	
14	Scarborough Village	0.0	
15	Woburn	0.0	

	Athletics & Sports	Auto Garage	Bakery	Bank	Bar	Breakfast Spot	\
0	0.000000	0.000000	0.000000	0.000000	0.0	0.200000	
1	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	
2	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	
3	0.142857	0.000000	0.142857	0.142857	0.0	0.000000	
4	0.000000	0.000000	0.222222	0.000000	0.0	0.000000	
5	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	
6	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	
7	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	
8	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	

9	0.000000	0.000000	0.000000	0.000000	0.0	0.142857
10	0.000000	0.000000	0.000000	0.000000	0.5	0.000000
11	0.000000	0.000000	0.000000	0.000000	0.0	0.071429
12	0.000000	0.166667	0.000000	0.000000	0.0	0.166667
13	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
14	0.000000	0.000000	0.000000	0.000000	0.0	0.000000
15	0.000000	0.000000	0.000000	0.000000	0.0	0.000000

	Brewery	Bus Line	Bus Station	Café	Camera Store	Caribbean Restaurant	\
0	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
1	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
2	0.000000	0.000000	0.000000	0.25	0.000000	0.000000	
3	0.000000	0.000000	0.000000	0.00	0.000000	0.142857	
4	0.000000	0.222222	0.111111	0.00	0.000000	0.000000	
5	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
6	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
7	0.111111	0.000000	0.000000	0.00	0.000000	0.000000	
8	0.000000	0.000000	0.142857	0.00	0.000000	0.000000	
9	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
10	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
11	0.000000	0.000000	0.000000	0.00	0.071429	0.000000	
12	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
13	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
14	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	
15	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	

	Chinese Restaurant	Clothing Store	Coffee Shop	College Stadium	\
0	0.200000	0.2	0.000000	0.00	
1	0.000000	0.0	0.000000	0.00	
2	0.000000	0.0	0.000000	0.25	
3	0.000000	0.0	0.000000	0.00	
4	0.000000	0.0	0.000000	0.00	
5	0.111111	0.0	0.000000	0.00	
6	0.000000	0.0	0.000000	0.00	
7	0.111111	0.0	0.000000	0.00	
8	0.142857	0.0	0.142857	0.00	
9	0.000000	0.0	0.000000	0.00	
10	0.000000	0.0	0.000000	0.00	
11	0.142857	0.0	0.142857	0.00	
12	0.000000	0.0	0.000000	0.00	
13	0.000000	0.0	0.000000	0.00	
14	0.000000	0.0	0.000000	0.00	
15	0.000000	0.0	0.500000	0.00	

	Cosmetics Shop	Department Store	Discount Store	Electronics Store	\
0	0.000000	0.000000	0.000000	0.000000	
1	0.000000	0.000000	0.000000	0.000000	
2	0.000000	0.000000	0.000000	0.000000	

3	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.000000
7	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.142857	0.285714	0.000000
9	0.000000	0.000000	0.000000	0.142857
10	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000
14	0.333333	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000

	Fast Food Restaurant	Fried Chicken Joint	Furniture / Home Store	\
0	0.000000	0.000000	0.000000	
1	0.000000	0.000000	0.000000	
2	0.000000	0.000000	0.000000	
3	0.000000	0.142857	0.000000	
4	0.111111	0.000000	0.000000	
5	0.111111	0.111111	0.000000	
6	0.000000	0.000000	0.000000	
7	0.000000	0.000000	0.111111	
8	0.000000	0.000000	0.000000	
9	0.000000	0.000000	0.000000	
10	0.000000	0.000000	0.000000	
11	0.142857	0.000000	0.000000	
12	0.000000	0.000000	0.000000	
13	0.500000	0.000000	0.000000	
14	0.000000	0.000000	0.000000	
15	0.000000	0.000000	0.000000	

	General Entertainment	Grocery Store	Hakka Restaurant	Indian Restaurant	\
0	0.00	0.000000	0.000000	0.000000	
1	0.00	0.000000	0.000000	0.000000	
2	0.25	0.000000	0.000000	0.000000	
3	0.00	0.000000	0.142857	0.000000	
4	0.00	0.000000	0.000000	0.000000	
5	0.00	0.000000	0.000000	0.000000	
6	0.00	0.000000	0.000000	0.000000	
7	0.00	0.000000	0.000000	0.222222	
8	0.00	0.000000	0.000000	0.000000	
9	0.00	0.000000	0.000000	0.000000	
10	0.00	0.000000	0.000000	0.000000	
11	0.00	0.071429	0.000000	0.071429	
12	0.00	0.000000	0.000000	0.000000	
13	0.00	0.000000	0.000000	0.000000	
14	0.00	0.333333	0.000000	0.000000	

15	0.00	0.000000	0.000000	0.250000
----	------	----------	----------	----------

	Intersection	Italian Restaurant	Japanese Restaurant	Korean Restaurant	\
0	0.000000	0.000000	0.000000	0.00	
1	0.000000	0.000000	0.000000	0.00	
2	0.000000	0.000000	0.000000	0.00	
3	0.000000	0.000000	0.000000	0.00	
4	0.111111	0.000000	0.000000	0.00	
5	0.000000	0.111111	0.000000	0.00	
6	0.000000	0.000000	0.000000	0.00	
7	0.000000	0.000000	0.000000	0.00	
8	0.000000	0.000000	0.000000	0.00	
9	0.000000	0.000000	0.000000	0.00	
10	0.000000	0.000000	0.000000	0.00	
11	0.000000	0.000000	0.071429	0.00	
12	0.000000	0.000000	0.000000	0.00	
13	0.000000	0.000000	0.000000	0.00	
14	0.000000	0.000000	0.000000	0.00	
15	0.000000	0.000000	0.000000	0.25	

	Latin American Restaurant	Lounge	Medical Center	Metro Station	\
0	0.000000	0.2	0.000000	0.000000	
1	0.000000	0.0	0.000000	0.000000	
2	0.000000	0.0	0.000000	0.000000	
3	0.000000	0.0	0.000000	0.000000	
4	0.000000	0.0	0.000000	0.111111	
5	0.000000	0.0	0.000000	0.000000	
6	0.000000	0.0	0.000000	0.000000	
7	0.111111	0.0	0.000000	0.000000	
8	0.000000	0.0	0.000000	0.000000	
9	0.000000	0.0	0.142857	0.000000	
10	0.000000	0.0	0.000000	0.000000	
11	0.000000	0.0	0.000000	0.000000	
12	0.000000	0.0	0.000000	0.000000	
13	0.000000	0.0	0.000000	0.000000	
14	0.000000	0.0	0.000000	0.000000	
15	0.000000	0.0	0.000000	0.000000	

	Mexican Restaurant	Middle Eastern Restaurant	Motel	Moving Target	\
0	0.000000	0.000000	0.0	0.0	
1	0.000000	0.000000	0.0	0.0	
2	0.000000	0.000000	0.0	0.0	
3	0.000000	0.000000	0.0	0.0	
4	0.000000	0.000000	0.0	0.0	
5	0.000000	0.000000	0.0	0.0	
6	0.000000	0.000000	0.5	0.0	
7	0.000000	0.000000	0.0	0.0	
8	0.000000	0.000000	0.0	0.0	

9	0.142857	0.000000	0.0	0.0
10	0.000000	0.000000	0.0	0.5
11	0.000000	0.000000	0.0	0.0
12	0.000000	0.333333	0.0	0.0
13	0.000000	0.000000	0.0	0.0
14	0.000000	0.000000	0.0	0.0
15	0.000000	0.000000	0.0	0.0

	Noodle House	Park	Pet Store	Pharmacy	Pizza Place	Playground \
0	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.5	0.000000	0.000000	0.000000	0.500000
2	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
5	0.111111	0.0	0.000000	0.000000	0.222222	0.000000
6	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
7	0.000000	0.0	0.111111	0.000000	0.000000	0.000000
8	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.0	0.000000	0.000000	0.142857	0.000000
10	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.0	0.000000	0.071429	0.071429	0.000000
12	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.0	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.0	0.000000	0.000000	0.000000	0.333333
15	0.000000	0.0	0.000000	0.000000	0.000000	0.000000

	Print Shop	Rental Car	Location	Sandwich Place	Shopping Mall \
0	0.0		0.000000	0.200000	0.000000
1	0.0		0.000000	0.000000	0.000000
2	0.0		0.000000	0.000000	0.000000
3	0.0		0.000000	0.000000	0.000000
4	0.0		0.000000	0.000000	0.000000
5	0.0		0.000000	0.000000	0.111111
6	0.0		0.000000	0.000000	0.000000
7	0.0		0.000000	0.000000	0.000000
8	0.0		0.000000	0.000000	0.000000
9	0.0		0.142857	0.000000	0.000000
10	0.0		0.000000	0.000000	0.000000
11	0.0		0.000000	0.071429	0.000000
12	0.0		0.000000	0.166667	0.166667
13	0.5		0.000000	0.000000	0.000000
14	0.0		0.000000	0.000000	0.000000
15	0.0		0.000000	0.000000	0.000000

	Skating Rink	Soccer Field	Spa	Thai Restaurant \
0	0.00	0.000000	0.000000	0.000000
1	0.00	0.000000	0.000000	0.000000
2	0.25	0.000000	0.000000	0.000000

3	0.00	0.000000	0.000000	0.142857
4	0.00	0.111111	0.000000	0.000000
5	0.00	0.000000	0.000000	0.111111
6	0.00	0.000000	0.000000	0.000000
7	0.00	0.000000	0.000000	0.000000
8	0.00	0.000000	0.000000	0.000000
9	0.00	0.000000	0.142857	0.000000
10	0.00	0.000000	0.000000	0.000000
11	0.00	0.000000	0.000000	0.000000
12	0.00	0.000000	0.000000	0.000000
13	0.00	0.000000	0.000000	0.000000
14	0.00	0.000000	0.000000	0.000000
15	0.00	0.000000	0.000000	0.000000

	Thrift / Vintage Store	Train Station	Vietnamese Restaurant
0	0.000000	0.000000	0.000000
1	0.000000	0.000000	0.000000
2	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000
7	0.111111	0.000000	0.111111
8	0.000000	0.142857	0.000000
9	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000

```
In [41]: Scarborough_grouped.shape
```

```
Out[41]: (16, 58)
```

0.3.2 Review the top 5 venues

```
In [42]: num_top_venues = 5
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')
```

----Aginccourt----

	venue	freq
0	Sandwich Place	0.2
1	Lounge	0.2
2	Clothing Store	0.2
3	Chinese Restaurant	0.2
4	Breakfast Spot	0.2

----Aginccourt NorthL'Amoreaux EastMillikenSteeles East----

	venue	freq
0	Playground	0.5
1	Park	0.5
2	American Restaurant	0.0
3	Pet Store	0.0
4	Korean Restaurant	0.0

----Birch CliffCliffside West----

	venue	freq
0	College Stadium	0.25
1	General Entertainment	0.25
2	Skating Rink	0.25
3	Café	0.25
4	American Restaurant	0.00

----Cedarbrae----

	venue	freq
0	Hakka Restaurant	0.14
1	Bakery	0.14
2	Bank	0.14
3	Athletics & Sports	0.14
4	Thai Restaurant	0.14

----ClairleaGolden MileOakridge----

	venue	freq
0	Bus Line	0.22
1	Bakery	0.22
2	Soccer Field	0.11
3	Fast Food Restaurant	0.11
4	Bus Station	0.11

----Clarks CornersSullivanTam O'Shanter----

	venue	freq
0	Pizza Place	0.22

1	Italian Restaurant	0.11
2	Fried Chicken Joint	0.11
3	Chinese Restaurant	0.11
4	Fast Food Restaurant	0.11

----CliffcrestCliffsideScarborough Village West----

	venue	freq
0	American Restaurant	0.5
1	Motel	0.5
2	Italian Restaurant	0.0
3	Korean Restaurant	0.0
4	Latin American Restaurant	0.0

----Dorset ParkScarborough Town CentreWexford Heights----

	venue	freq
0	Indian Restaurant	0.22
1	Vietnamese Restaurant	0.11
2	Brewery	0.11
3	Furniture / Home Store	0.11
4	Pet Store	0.11

----East Birchmount ParkIonviewKennedy Park----

	venue	freq
0	Discount Store	0.29
1	Coffee Shop	0.14
2	Bus Station	0.14
3	Chinese Restaurant	0.14
4	Department Store	0.14

----GuildwoodMorningsideWest Hill----

	venue	freq
0	Pizza Place	0.14
1	Medical Center	0.14
2	Spa	0.14
3	Breakfast Spot	0.14
4	Mexican Restaurant	0.14

----Highland CreekRouge HillPort Union----

	venue	freq
0	Bar	0.5
1	Moving Target	0.5
2	American Restaurant	0.0
3	Pharmacy	0.0

4 Korean Restaurant 0.0

----L'Amoreaux WestSteeles West----

	venue	freq
0	Chinese Restaurant	0.14
1	Fast Food Restaurant	0.14
2	Coffee Shop	0.14
3	Pharmacy	0.07
4	Japanese Restaurant	0.07

----MaryvaleWexford----

	venue	freq
0	Middle Eastern Restaurant	0.33
1	Auto Garage	0.17
2	Breakfast Spot	0.17
3	Shopping Mall	0.17
4	Sandwich Place	0.17

----RougeMalvern----

	venue	freq
0	Fast Food Restaurant	0.5
1	Print Shop	0.5
2	American Restaurant	0.0
3	Pet Store	0.0
4	Korean Restaurant	0.0

----Scarborough Village----

	venue	freq
0	Grocery Store	0.33
1	Playground	0.33
2	Cosmetics Shop	0.33
3	American Restaurant	0.00
4	Pet Store	0.00

----Woburn----

	venue	freq
0	Coffee Shop	0.50
1	Korean Restaurant	0.25
2	Indian Restaurant	0.25
3	Latin American Restaurant	0.00
4	Lounge	0.00

0.3.3 Function to sort the most common venues

```
In [43]: 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 [44]: 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(Scarborough_g

neighborhoods_venues_sorted
```

```
Out[44]:
```

	Neighborhood \
0	Agincourt
1	Agincourt NorthL'Amoreaux EastMillikenSteeles ...
2	Birch CliffCliffside West
3	Cedarbrae
4	ClairleaGolden MileOakridge
5	Clarks CornersSullivanTam O'Shanter
6	CliffcrestCliffsideScarborough Village West
7	Dorset ParkScarborough Town CentreWexford Heights
8	East Birchmount ParkIonviewKennedy Park
9	GuildwoodMorningsideWest Hill
10	Highland CreekRouge HillPort Union
11	L'Amoreaux WestSteeles West
12	MaryvaleWexford
13	RougeMalvern
14	Scarborough Village
15	Woburn

	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue \
0	Clothing Store	Lounge	Breakfast Spot

1	Park	Playground	Vietnamese Restaurant
2	General Entertainment	Skating Rink	Café
3	Caribbean Restaurant	Bakery	Fried Chicken Joint
4	Bakery	Bus Line	Intersection
5	Pizza Place	Noodle House	Thai Restaurant
6	American Restaurant	Motel	Clothing Store
7	Indian Restaurant	Brewery	Furniture / Home Store
8	Discount Store	Chinese Restaurant	Bus Station
9	Breakfast Spot	Mexican Restaurant	Spa
10	Bar	Moving Target	Vietnamese Restaurant
11	Chinese Restaurant	Coffee Shop	Fast Food Restaurant
12	Middle Eastern Restaurant	Sandwich Place	Auto Garage
13	Fast Food Restaurant	Print Shop	Vietnamese Restaurant
14	Grocery Store	Cosmetics Shop	Playground
15	Coffee Shop	Korean Restaurant	Indian Restaurant
	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue \
0	Sandwich Place	Chinese Restaurant	Vietnamese Restaurant
1	General Entertainment	Furniture / Home Store	Fried Chicken Joint
2	College Stadium	Vietnamese Restaurant	Clothing Store
3	Bank	Hakka Restaurant	Thai Restaurant
4	Soccer Field	Fast Food Restaurant	Bus Station
5	Fried Chicken Joint	Italian Restaurant	Fast Food Restaurant
6	Grocery Store	General Entertainment	Furniture / Home Store
7	Latin American Restaurant	Chinese Restaurant	Pet Store
8	Department Store	Coffee Shop	Train Station
9	Medical Center	Electronics Store	Pizza Place
10	Clothing Store	Grocery Store	General Entertainment
11	Sandwich Place	Grocery Store	Pharmacy
12	Shopping Mall	Breakfast Spot	Electronics Store
13	Chinese Restaurant	Grocery Store	General Entertainment
14	Vietnamese Restaurant	Chinese Restaurant	General Entertainment
15	Vietnamese Restaurant	Clothing Store	Grocery Store
	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue \
0	General Entertainment	Furniture / Home Store	Fried Chicken Joint
1	Fast Food Restaurant	Electronics Store	Discount Store
2	Grocery Store	Furniture / Home Store	Fried Chicken Joint
3	Athletics & Sports	Auto Garage	Grocery Store
4	Metro Station	Coffee Shop	General Entertainment
5	Shopping Mall	Chinese Restaurant	Clothing Store
6	Fried Chicken Joint	Fast Food Restaurant	Electronics Store
7	Vietnamese Restaurant	Thrift / Vintage Store	Bar
8	Bar	Cosmetics Shop	Hakka Restaurant
9	Rental Car Location	Department Store	College Stadium
10	Furniture / Home Store	Fried Chicken Joint	Fast Food Restaurant
11	Pizza Place	Camera Store	Breakfast Spot
12	College Stadium	Cosmetics Shop	Department Store

13	Furniture / Home Store	Fried Chicken Joint	Electronics Store
14	Furniture / Home Store	Fried Chicken Joint	Fast Food Restaurant
15	General Entertainment	Furniture / Home Store	Fried Chicken Joint

```

10th Most Common Venue
0    Fast Food Restaurant
1      Department Store
2    Fast Food Restaurant
3    General Entertainment
4    Furniture / Home Store
5        Coffee Shop
6      Discount Store
7    College Stadium
8      Grocery Store
9      Cosmetics Shop
10     Electronics Store
11   Japanese Restaurant
12      Discount Store
13      Discount Store
14     Electronics Store
15    Fast Food Restaurant

```

0.3.4 Cluster the venues using K means

```

In [45]: # 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=0).fit(Scarborough_grouped_clustering)
# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]

```

```

Out[45]: array([0, 0, 0, 0, 0, 0, 2, 0, 0, 0], dtype=int32)

```

0.3.5 Make labels

```

In [46]: Scarborough_merged = Scarborough_data

# add clustering labels
Scarborough_merged['Cluster Labels'] = kmeans.labels_
# make the column name the same
Scarborough_merged.rename(columns={'Neighbourhood': 'Neighborhood'}, inplace = True)

```

0.3.6 Merge neighborhoods_venues_sorted to Scarborough_merged by column "Neighborhood"

```

In [47]: # merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhood
Scarborough_merged = Scarborough_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'))

```

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

```
Out[47]:
```

	Postcode	Borough	Neighborhood	Latitude	\
0	M1B	Scarborough	RougeMalvern	43.806686	
1	M1C	Scarborough	Highland CreekRouge HillPort Union	43.784535	
2	M1E	Scarborough	GuildwoodMorningsideWest Hill	43.763573	
3	M1G	Scarborough	Woburn	43.770992	
4	M1H	Scarborough	Cedarbrae	43.773136	

	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	\
0	-79.194353	0	Fast Food Restaurant	Print Shop	
1	-79.160497	0	Bar	Moving Target	
2	-79.188711	0	Breakfast Spot	Mexican Restaurant	
3	-79.216917	0	Coffee Shop	Korean Restaurant	
4	-79.239476	0	Caribbean Restaurant	Bakery	

	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	\
0	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	
1	Vietnamese Restaurant	Clothing Store	Grocery Store	
2	Spa	Medical Center	Electronics Store	
3	Indian Restaurant	Vietnamese Restaurant	Clothing Store	
4	Fried Chicken Joint	Bank	Hakka Restaurant	

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
0	General Entertainment	Furniture / Home Store	Fried Chicken Joint	
1	General Entertainment	Furniture / Home Store	Fried Chicken Joint	
2	Pizza Place	Rental Car Location	Department Store	
3	Grocery Store	General Entertainment	Furniture / Home Store	
4	Thai Restaurant	Athletics & Sports	Auto Garage	

	9th Most Common Venue	10th Most Common Venue
0	Electronics Store	Discount Store
1	Fast Food Restaurant	Electronics Store
2	College Stadium	Cosmetics Shop
3	Fried Chicken Joint	Fast Food Restaurant
4	Grocery Store	General Entertainment

0.3.7 Map the Clusters

```
In [48]: # 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['L
    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[48]: <folium.folium.Map at 0x7fd988332358>

0.3.8 verify the cluster1

In [49]: Scarborough_merged.loc[Scarborough_merged['Cluster Labels'] == 0, Scarborough_merged.co

```

Out[49]:
   Borough  Latitude  Longitude  Cluster Labels  \
0  Scarborough  43.806686  -79.194353          0
1  Scarborough  43.784535  -79.160497          0
2  Scarborough  43.763573  -79.188711          0
3  Scarborough  43.770992  -79.216917          0
4  Scarborough  43.773136  -79.239476          0
5  Scarborough  43.744734  -79.239476          0
7  Scarborough  43.711112  -79.284577          0
8  Scarborough  43.716316  -79.239476          0
9  Scarborough  43.692657  -79.264848          0
11 Scarborough  43.750072  -79.295849          0
12 Scarborough  43.794200  -79.262029          0
13 Scarborough  43.781638  -79.304302          0
14 Scarborough  43.815252  -79.284577          0
15 Scarborough  43.799525  -79.318389          0

   1st Most Common Venue  2nd Most Common Venue  3rd Most Common Venue  \
0  Fast Food Restaurant  Print Shop  Vietnamese Restaurant
1  Bar  Moving Target  Vietnamese Restaurant
2  Breakfast Spot  Mexican Restaurant  Spa
3  Coffee Shop  Korean Restaurant  Indian Restaurant
4  Caribbean Restaurant  Bakery  Fried Chicken Joint
5  Grocery Store  Cosmetics Shop  Playground
7  Bakery  Bus Line  Intersection
8  American Restaurant  Motel  Clothing Store
9  General Entertainment  Skating Rink  Café

```

11	Middle Eastern Restaurant	Sandwich Place	Auto Garage
12	Clothing Store	Lounge	Breakfast Spot
13	Pizza Place	Noodle House	Thai Restaurant
14	Park	Playground	Vietnamese Restaurant
15	Chinese Restaurant	Coffee Shop	Fast Food Restaurant

	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue \
0	Chinese Restaurant	Grocery Store	General Entertainment
1	Clothing Store	Grocery Store	General Entertainment
2	Medical Center	Electronics Store	Pizza Place
3	Vietnamese Restaurant	Clothing Store	Grocery Store
4	Bank	Hakka Restaurant	Thai Restaurant
5	Vietnamese Restaurant	Chinese Restaurant	General Entertainment
7	Soccer Field	Fast Food Restaurant	Bus Station
8	Grocery Store	General Entertainment	Furniture / Home Store
9	College Stadium	Vietnamese Restaurant	Clothing Store
11	Shopping Mall	Breakfast Spot	Electronics Store
12	Sandwich Place	Chinese Restaurant	Vietnamese Restaurant
13	Fried Chicken Joint	Italian Restaurant	Fast Food Restaurant
14	General Entertainment	Furniture / Home Store	Fried Chicken Joint
15	Sandwich Place	Grocery Store	Pharmacy

	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue \
0	Furniture / Home Store	Fried Chicken Joint	Electronics Store
1	Furniture / Home Store	Fried Chicken Joint	Fast Food Restaurant
2	Rental Car Location	Department Store	College Stadium
3	General Entertainment	Furniture / Home Store	Fried Chicken Joint
4	Athletics & Sports	Auto Garage	Grocery Store
5	Furniture / Home Store	Fried Chicken Joint	Fast Food Restaurant
7	Metro Station	Coffee Shop	General Entertainment
8	Fried Chicken Joint	Fast Food Restaurant	Electronics Store
9	Grocery Store	Furniture / Home Store	Fried Chicken Joint
11	College Stadium	Cosmetics Shop	Department Store
12	General Entertainment	Furniture / Home Store	Fried Chicken Joint
13	Shopping Mall	Chinese Restaurant	Clothing Store
14	Fast Food Restaurant	Electronics Store	Discount Store
15	Pizza Place	Camera Store	Breakfast Spot

	10th Most Common Venue
0	Discount Store
1	Electronics Store
2	Cosmetics Shop
3	Fast Food Restaurant
4	General Entertainment
5	Electronics Store
7	Furniture / Home Store
8	Discount Store
9	Fast Food Restaurant

```

11         Discount Store
12     Fast Food Restaurant
13         Coffee Shop
14         Department Store
15     Japanese Restaurant

```

0.3.9 Verify the cluster2

```
In [50]: Scarborough_merged.loc[Scarborough_merged['Cluster Labels'] == 1, Scarborough_merged.co
```

```

Out[50]:      Borough  Latitude  Longitude  Cluster Labels  1st Most Common Venue  \
10  Scarborough  43.75741  -79.273304             1      Indian Restaurant

      2nd Most Common Venue  3rd Most Common Venue  4th Most Common Venue  \
10              Brewery  Furniture / Home Store  Latin American Restaurant

      5th Most Common Venue  6th Most Common Venue  7th Most Common Venue  \
10      Chinese Restaurant              Pet Store  Vietnamese Restaurant

      8th Most Common Venue  9th Most Common Venue  10th Most Common Venue
10  Thrift / Vintage Store              Bar      College Stadium

```

0.3.10 verify the cluster3

```
In [51]: Scarborough_merged.loc[Scarborough_merged['Cluster Labels'] == 2, Scarborough_merged.co
```

```

Out[51]:      Borough  Latitude  Longitude  Cluster Labels  1st Most Common Venue  \
6  Scarborough  43.727929  -79.262029             2      Discount Store

      2nd Most Common Venue  3rd Most Common Venue  4th Most Common Venue  \
6      Chinese Restaurant              Bus Station      Department Store

      5th Most Common Venue  6th Most Common Venue  7th Most Common Venue  \
6              Coffee Shop      Train Station              Bar

      8th Most Common Venue  9th Most Common Venue  10th Most Common Venue
6      Cosmetics Shop      Hakka Restaurant      Grocery Store

```

```
In [ ]:
```