


Branch: master ▾

Find file Copy path

Capstone_Project / Case study_continued.ipynb


 **DeepentiA** Add files via upload

be34d6f now

1 contributor

<>

RawBlameHistory



2374 lines (2374 sloc) 180 KB

```
In [58]: url= "https://raw.githubusercontent.com/DeepentiA/Capstone_Project/master/best%20food.csv"
df5 = pd.read_csv(url)
df5.head()
```

```
In [59]: import requests
url ="https://raw.githubusercontent.com/DeepentiA/Capstone_Project/master/Gurgaon_dataset.csv"
df = pd.read_csv(url)
print(df)
df.head(20)
```

	Unnamed: 0	Borough	Neighborhoods	Latitude	Longitude	\
0	0	Central	Cantonment area	12.972442	77.580643	
1	1	Central	Domlur	12.960992	77.638726	
2	2	Central	Indiranagar	12.971891	77.641151	
3	3	Central	Jeevanbheemanagar	12.962900	77.659500	
4	4	Central	Malleswaram	13.003100	77.564300	
5	5	Central	Pete area	12.962700	77.575800	
6	6	Central	Rajajinagar	12.990100	77.552500	
7	7	Central	Sadashivanagar	13.006800	77.581300	
8	8	Central	Seshadripuram	12.993500	77.578700	
9	9	Central	Shivajinagar	12.985700	77.605700	
10	10	Central	Ulsoor	12.981700	77.628600	
11	11	Central	Vasanth Nagar	12.991100	77.592000	
12	12	Eastern	Bellandur	12.926000	77.676200	
13	13	Eastern	CV Raman Nagar	12.985500	77.663900	
14	14	Eastern	Hoodi	12.992200	77.715900	
15	15	Eastern	Krishnarajapuram	13.004000	77.687800	
16	16	Eastern	Mahadevapura	12.991300	77.687400	
17	17	Eastern	Marathahalli	12.959200	77.697400	
18	18	Eastern	Varthur	12.938900	77.741200	
19	19	Eastern	Whitefield	12.969800	77.749900	
20	20	NorthEastern	Banaswadi	13.012000	77.647100	
21	21	NorthEastern	HBR Layout	13.037700	77.628800	
22	22	NorthEastern	Horamavu	13.025100	77.659700	
23	23	NorthEastern	Kammanahalli	13.015900	77.637900	
24	24	NorthEastern	Lingarajapuram	13.008400	77.630600	
25	25	NorthEastern	Ramamurthy Nagar	13.018400	77.678100	
26	26	Northern	Hebbal	13.035800	77.597000	
27	27	Northern	Jalahalli	13.052800	77.541900	
28	28	Northern	Mathikere	13.033400	77.564000	
29	29	Northern	Peenya	13.028500	77.519700	
30	30	Northern	R. T. Nagar	13.019600	77.596800	
31	31	Northern	Vidyaranyapura	13.085800	77.556100	
32	32	Northern	Yelahanka	13.100500	77.594000	
33	33	Northern	Yeshwanthpur	13.028000	77.540900	
34	34	SouthEastern	Bommanahalli	12.898400	77.617900	
35	35	SouthEastern	Bommasandra	12.816800	77.698900	
36	36	SouthEastern	BTM Layout	12.916600	77.610100	
37	37	SouthEastern	Electronic City	12.840711	77.676369	
38	38	SouthEastern	HSR Layout	12.908100	77.647600	
39	39	SouthEastern	Koramangala	12.927900	77.627100	
40	40	SouthEastern	Madiwala	12.922600	77.617400	
41	41	Southern	Banashankari	12.925500	77.546800	
42	42	Southern	Basavanagudi	12.942100	77.575400	
43	43	Southern	Girinagar	12.938600	77.544000	
44	44	Southern	J. P. Nagar	12.910500	77.585700	
45	45	Southern	Jayanagar	12.925000	77.593800	
46	46	Southern	Kumaraswamy Layout	12.903800	77.561800	
47	47	Southern	Padmanabhanagar	12.915600	77.556800	
48	48	Southern	Uttarahalli	12.907000	77.552100	
49	49	SouthernSuburbs	Anjanapura	12.860400	77.561200	
50	50	SouthernSuburbs	Arekere	12.887500	77.597000	
51	51	SouthernSuburbs	Begur	12.878800	77.637700	
52	52	SouthernSuburbs	Gottigere	12.855200	77.586900	
53	53	SouthernSuburbs	Hulimavu	12.878900	77.609000	
54	54	SouthernSuburbs	Kothnur	12.873400	77.582000	
55	55	Western	Basaveshwaranagar	12.988600	77.538300	
56	56	Western	Kamakshipalya	12.988700	77.527100	
57	57	Western	Kengeri	12.899600	77.482700	
58	58	Western	Mahalakshmi Layout	13.014600	77.551400	
59	59	Western	Nagarbhavi	12.959900	77.508300	
60	60	Western	Nandini Layout	13.016000	77.533800	
61	61	Western	Nayandahalli	12.941100	77.524800	
62	62	Western	Rajarajeshwari Nagar	12.924200	77.519100	
63	63	Western	Vijayanagar	12.971900	77.529900	

	Population	City	AverageIncome
0	866377	Gurgaon	18944.099790
1	743186	Gurgaon	56837.022200
2	474289	Gurgaon	41991.817440
3	527874	Gurgaon	6667.447632
4	893629	Gurgaon	53270.063890
5	730999	Gurgaon	50712.430220
6	981362	Gurgaon	60967.535870
7	662625	Gurgaon	59943.541560

```

8      396862 Gurgaon  58407.090340
9      77836  Gurgaon  55850.962100
10     656726 Gurgaon  41007.219540
11     942711 Gurgaon  26168.448090
12     208094 Gurgaon   7227.731930
13     122714 Gurgaon  54335.368710
14     330409 Gurgaon  22591.063480
15     351936 Gurgaon  36934.737730
16     905568 Gurgaon  35915.973330
17     249182 Gurgaon  58448.658520
18     546186 Gurgaon  36433.267300
19     83029  Gurgaon  44637.984600
20     632031 Gurgaon  53349.701180
21     822101 Gurgaon  43625.383680
22     679059 Gurgaon  13930.195750
23     730628 Gurgaon  23146.347760
24     223119 Gurgaon  55404.539140
25     468662 Gurgaon  56428.329770
26     926976 Gurgaon  39022.694140
27     569724 Gurgaon   9836.954154
28     757626 Gurgaon  61046.854720
29     661647 Gurgaon  24706.335040
30     241187 Gurgaon  64644.480990
31     704566 Gurgaon  31889.365790
32     15846  Gurgaon   5269.688407
33     333688 Gurgaon  27803.161570
34     737166 Gurgaon  11422.449970
35     723807 Gurgaon  24735.995810
36     835743 Gurgaon  49310.618880
37     781260 Gurgaon   8350.934536
38     534598 Gurgaon  45739.216290
39     288089 Gurgaon  40110.283710
40     901659 Gurgaon  15025.903390
41     810407 Gurgaon  57524.209530
42     426903 Gurgaon  63161.962220
43     783535 Gurgaon  41146.060250
44     255412 Gurgaon  61115.381150
45     103930 Gurgaon   8891.317323
46     983614 Gurgaon  10941.420380
47     561382 Gurgaon  47292.334640
48     722264 Gurgaon  63166.190380
49     940039 Gurgaon  44218.922550
50     138760 Gurgaon  29378.716630
51     594887 Gurgaon  61640.098300
52     381292 Gurgaon  25803.820490
53     949207 Gurgaon  33486.513520
54     65249  Gurgaon  55508.825050
55     628910 Gurgaon  41177.478490
56     968497 Gurgaon  39132.402990
57     14705  Gurgaon  19167.834100
58     306403 Gurgaon   8416.293069
59     623843 Gurgaon  38627.411760
60     638555 Gurgaon  32490.969170
61     474920 Gurgaon  46826.803890
62     516920 Gurgaon  12533.785280
63     791549 Gurgaon  51966.782270

```

Out[59]:

	Unnamed: 0	Borough	Neighborhoods	Latitude	Longitude	Population	City	AverageIncome
0	0	Central	Cantonment area	12.972442	77.580643	866377	Gurgaon	18944.099790
1	1	Central	Domlur	12.960992	77.638726	743186	Gurgaon	56837.022200
2	2	Central	Indiranagar	12.971891	77.641151	474289	Gurgaon	41991.817440
3	3	Central	Jeevanbheemanagar	12.962900	77.659500	527874	Gurgaon	6667.447632
4	4	Central	Malleswaram	13.003100	77.564300	893629	Gurgaon	53270.063890
5	5	Central	Pete area	12.962700	77.575800	730999	Gurgaon	50712.430220
6	6	Central	Rajajinagar	12.990100	77.552500	981362	Gurgaon	60967.535870
7	7	Central	Sadashivanagar	13.006800	77.581300	662625	Gurgaon	59943.541560
8	8	Central	Seshadripuram	12.993500	77.578700	396862	Gurgaon	58407.090340
9	9	Central	Shivajinagar	12.985700	77.605700	77836	Gurgaon	55850.962100
10	10	Central	Ulsoor	12.981700	77.628600	656726	Gurgaon	41007.219540
11	11	Central	Vasanth Nagar	12.991100	77.592000	942711	Gurgaon	26168.448090
12	12	Eastern	Bellandur	12.926000	77.676200	208094	Gurgaon	7227.731930
13	13	Eastern	CV Raman Nagar	12.985500	77.663900	122714	Gurgaon	54335.368710
14	14	Eastern	Hoodi	12.992200	77.715900	330409	Gurgaon	22591.063480
15	15	Eastern	Krishnarajapuram	13.004000	77.687800	351936	Gurgaon	36934.737730
16	16	Eastern	Mahadevapura	12.991300	77.687400	905568	Gurgaon	35915.973330

17	17	Eastern	Marathahalli	12.959200	77.697400	249182	Gurgaon	58448.658520
18	18	Eastern	Varthur	12.938900	77.741200	546186	Gurgaon	36433.267300
19	19	Eastern	Whitefield	12.969800	77.749900	83029	Gurgaon	44637.984600

In [60]: `import os`

In [61]: `df.drop('Unnamed: 0',axis=1,inplace=True)`
`df.head()`

Out[61]:

	Borough	Neighborhoods	Latitude	Longitude	Population	City	AverageIncome
0	Central	Cantonment area	12.972442	77.580643	866377	Gurgaon	18944.099790
1	Central	Domlur	12.960992	77.638726	743186	Gurgaon	56837.022200
2	Central	Indiranagar	12.971891	77.641151	474289	Gurgaon	41991.817440
3	Central	Jeevanbheemanagar	12.962900	77.659500	527874	Gurgaon	6667.447632
4	Central	Malleswaram	13.003100	77.564300	893629	Gurgaon	53270.063890

In []:

In [62]: `url = "https://raw.githubusercontent.com/Deepentia/Capstone_Project/master/Gurgaon_population.csv"`
`df1 = pd.read_csv(url)`
`df1.head()`

Out[62]:

	Borough	Neighborhoods	Population
0	Central	Cantonment area	866377
1	Central	Domlur	743186
2	Central	Indiranagar	474289
3	Central	Jeevanbheemanagar	527874
4	Central	Malleswaram	893629

In [63]: `url = "https://raw.githubusercontent.com/Deepentia/Capstone_Project/master/Gurgaon_income.csv"`
`df2 = pd.read_csv(url)`
`df2.head()`

Out[63]:

	Borough	Neighborhoods	AverageIncome
0	Central	Cantonment area	18944.099790
1	Central	Domlur	56837.022200
2	Central	Indiranagar	41991.817440
3	Central	Jeevanbheemanagar	6667.447632
4	Central	Malleswaram	53270.063890

In [64]: `!pip install geopy`
`!pip install beautifulsoup4`

Requirement already satisfied: geopy in /home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (1.20.0)
Requirement already satisfied: geographiclib<2,>=1.49 in /home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geopy) (1.50)
Requirement already satisfied: beautifulsoup4 in /home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (4.8.1)
Requirement already satisfied: soupsieve>=1.2 in /home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from beautifulsoup4) (1.9.5)

In [65]: `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`

`from geopy.geocoders import Nominatim # convert an address into Latitude and Longitude values`

`import requests # Library to handle requests`
`from bs4 import BeautifulSoup # Library to parse HTML and XML documents`

`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

import folium # map rendering library

print("Libraries are imported.")

Libraries are imported.
```

```
In [66]: address = 'Gurgaon'

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

The geograpical coordinates of Gurgaon are 28.4646148, 77.0299194.
```

```
In [67]: map_GGN = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for latitude,longitude, Borough, Neighborhoods in zip(df['Latitude'],df['Longitude'], df['Borough'], df['N
eighborhoods']):
    label = '{} {}'.format(Neighborhoods, 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_GGN)

map_GGN
```

```
Out[67]:
```

```
In [68]: CLIENT_ID = 'AJ5GIVG4RTCVSPB1CQBINDQIL435MQJ5X3203DX2C2SL2NZM'
CLIENT_SECRET = 'HT0QEMRESM552MVJNVZEX44HOYVPHNLYM0HZDRQJTH3MFIOQ'
VERSION = '20180606'
LIMIT = 150
```

```
In [69]: gurgaon_latitude = df['Latitude'].mean()
gurgaon_longitude = df['Longitude'].mean()
print("Latitude and Longitude of Gurgaon are : ",gurgaon_latitude,gurgaon_longitude)

Latitude and Longitude of Gurgaon are : 12.9623396203125 77.601752946875
```

```
In [70]: unique_boroughs_of_gurgaon = df['Borough'].unique().tolist
```

```
In [71]: unique_boroughs_of_gurgaon
```

```
Out[71]: <function ndarray.tolist>
```

```
In [72]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import requests
import folium
import seaborn as sns
import matplotlib.cm as cm
import matplotlib.colors as colors
from sklearn.cluster import KMeans
```

```
In [73]: borough_colors = {}
for i in unique_boroughs_of_gurgaon:
    borough_colors[i] = '#%02X%02X%02X' % tuple(np.random.choice(range(256),size=3))
```

```
-----
TypeError                                Traceback (most recent call last)
<ipython-input-73-447cd1c30259> in <module>
      1 borough_colors = {}
----> 2 for i in unique_boroughs_of_gurgaon:
      3     borough_colors[i] = '#%02X%02X%02X' % tuple(np.random.choice(range(256),size=3))

TypeError: 'builtin_function_or_method' object is not iterable
```

```
{'Central': '#2ECA6E', 'Eastern': '#0E1B15', 'NorthEastern': '#B12985', 'Northern': '#AF4FF5', 'SouthEastern': '#68B1E1', 'Southern': '#EACF24',
'SouthernSuburbs': '#F29B23', 'Western': '#BD51B8'}
```

```
In [ ]: gurgaon_map = folium.Map(location=[gurgaon_latitude,gurgaon_longitude],zoom_start=12,control_scale=True)
```

```
In [ ]: for lat,lng,boro,nei in zip(df['Latitude'],
                                   df['Longitude'],
                                   df['Borough'],
                                   df['Neighborhoods']):
    label_text = boro + ' - ' + nei
    label = folium.Popup(label_text,parse_html=True)
    folium.CircleMarker(
        [lat,lng],
        tooltip = label_text,
        radius = 4,
        popup = label,
        color=borough_colors[boro],
        fill=True,
        fill_color = borough_colors[boro],
        fill_opacity=0.7).add_to(gurgon_map)
```

Exploring Gurgaon Neighborhoods using FourSquare API

```
In [ ]: def getNearbyVenues(names, boro, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, boro, lat, lng in zip(names, boro, latitudes, longitudes):
        print("Fetching venues for : ",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']

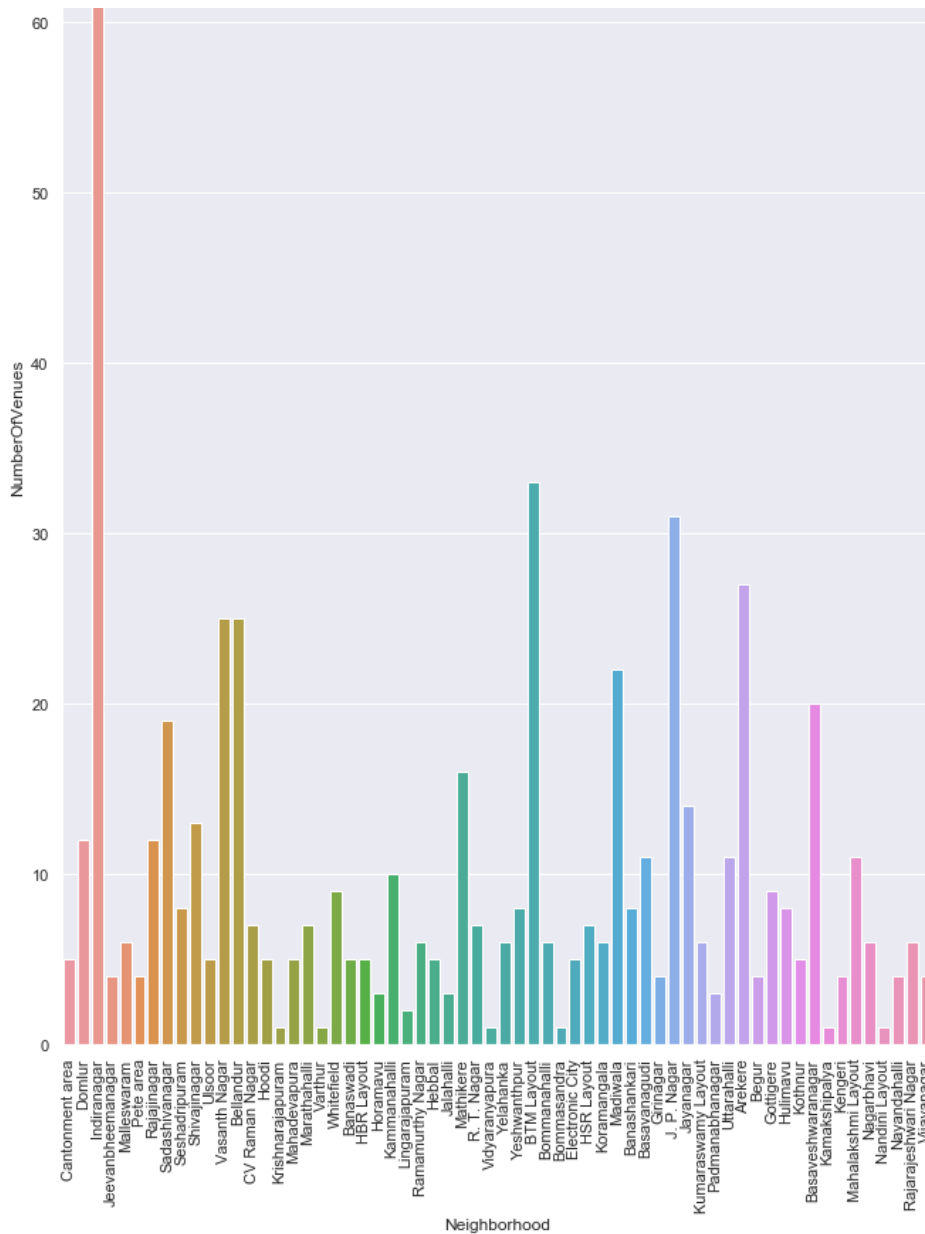
        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            boro,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
    nearby_venues.columns = ['Neighborhoods',
                            'Borough',
                            'Neighborhood Latitude',
                            'Neighborhood Longitude',
                            'Venue',
                            'Venue Latitude',
                            'Venue Longitude',
                            'Venue Category']

    return(nearby_venues)
```

```
In [ ]: gurgaon_venues = getNearbyVenues(names=df['Neighborhoods'],latitudes=df['Latitude'],longitudes=df['Longitude'],boro=df['Borough'])
```

Fetching venues for : Cantonment area Fetching venues for : Domlur Fetching venues for : Indiranagar Fetching venues for : Jeevanbheemanagar
 Fetching venues for : Malleswaram Fetching venues for : Pete area Fetching venues for : Rajajinagar Fetching venues for : Sadashivanagar
 Fetching venues for : Seshadripuram Fetching venues for : Shivajinagar Fetching venues for : Ulsoor Fetching venues for : Vasanth Nagar Fetching
 venues for : Bellandur Fetching venues for : CV Raman Nagar Fetching venues for : Hoodi Fetching venues for : Krishnarajapuram Fetching
 venues for : Mahadevapura Fetching venues for : Marathahalli Fetching venues for : Varthur Fetching venues for : Whitefield Fetching venues for :
 Banaswadi Fetching venues for : HBR Layout Fetching venues for : Horamavu Fetching venues for : Kammanahalli Fetching venues for :
 Lingarajapuram Fetching venues for : Ramamurthy Nagar Fetching venues for : Hebbal Fetching venues for : Jalahalli Fetching venues for :
 Mathikere Fetching venues for : Peenya Fetching venues for : R. T. Nagar Fetching venues for : Vidyananyapura Fetching venues for : Yelahanka
 Fetching venues for : Yeshwanthpur Fetching venues for : Bommanahalli Fetching venues for : Bommasandra Fetching venues for : BTM Layout
 Fetching venues for : Electronic City Fetching venues for : HSR Layout Fetching venues for : Koramangala Fetching venues for : Madiwala
 Fetching venues for : Banashankari Fetching venues for : Basavanagudi Fetching venues for : Girinagar Fetching venues for : J. P. Nagar Fetching
 venues for : Jayanagar Fetching venues for : Kumaraswamy Layout Fetching venues for : Padmanabhanagar Fetching venues for : Uttarahalli
 Fetching venues for : Anjanapura Fetching venues for : Arekere Fetching venues for : Begur Fetching venues for : Gottigere Fetching venues for :
 Hulimavu Fetching venues for : Kothnur Fetching venues for : Basaveshwaranagar Fetching venues for : Kamakshinhalva Fetching venues for :



```
In [ ]: number_of_top_venues = 5
```

```
In [ ]: for hood in bangalore_venues_grouped['Neighborhood']:
    print('-----',hood,'-----')
    temp = bangalore_venues_grouped[bangalore_venues_grouped['Neighborhood'] == hood].T.reset_index()
    temp.columns = ['Venue','Frequency']
    temp = temp.iloc[1:]
    temp['Frequency'] = temp['Frequency'].astype(float)
    temp = temp.round({'Frequency': 2})
    print(temp.sort_values('Frequency', ascending=False).reset_index(drop=True).head(number_of_top_venues))
    print('\n')
```

```
----- Arekere ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.19 1 Venue Category_Sporting Goods Shop 0.15 2 Venue
Category_Department Store 0.11 3 Venue Category_Pizza Place 0.07 4 Venue Category_Bus Line 0.04 ----- BTM Layout ----- Venue Frequency
0 Venue Category_Indian Restaurant 0.24 1 Venue Category_Ice Cream Shop 0.09 2 Venue Category_Snack Place 0.09 3 Venue
Category_Vegetarian / Vegan Restaurant 0.06 4 Venue Category_Pizza Place 0.06 ----- Banashankari ----- Venue Frequency 0 Venue
Category_Indian Restaurant 0.25 1 Venue Category_Café 0.25 2 Venue Category_Clothing Store 0.25 3 Venue Category_Pizza Place 0.12 4 Venue
Category_North Indian Restaurant 0.12 ----- Banaswadi ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.4 1 Venue
Category_Bakery 0.2 2 Venue Category_Vegetarian / Vegan Restaurant 0.2 3 Venue Category_Café 0.2 4 Venue Category_Outlet Store 0.0 -----
Basavanagudi ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.36 1 Venue Category_Mediterranean Restaurant 0.09 2 Venue
Category_Café 0.09 3 Venue Category_Indian Sweet Shop 0.09 4 Venue Category_Restaurant 0.09 ----- Basaveshwaranagar ----- Venue
Frequency 0 Venue Category_Indian Restaurant 0.15 1 Venue Category_Fast Food Restaurant 0.10 2 Venue Category_Ice Cream Shop 0.10 3
Venue Category_Juice Bar 0.05 4 Venue Category_Supermarket 0.05 ----- Begur ----- Venue Frequency 0 Venue Category_Bakery 0.25 1
Venue Category_Indian Sweet Shop 0.25 2 Venue Category_Stadium 0.25 3 Venue Category_Clothing Store 0.25 4 Venue Category_Outlet Store
0.00 ----- Bellandur ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.20 1 Venue Category_Fast Food Restaurant 0.12 2 Venue
Category_Kerala Restaurant 0.08 3 Venue Category_Café 0.08 4 Venue Category_Coffee Shop 0.04 ----- Bommanahalli ----- Venue Frequency
0 Venue Category_Furniture / Home Store 0.17 1 Venue Category_Department Store 0.17 2 Venue Category_Gym / Fitness Center 0.17 3 Venue
Category_Athletics & Sports 0.17 4 Venue Category_Auto Garage 0.17 ----- Bommasandra ----- Venue Frequency 0 Venue Category_Indian
Restaurant 1.0 1 Venue Category_ATM 0.0 2 Venue Category_Outlet Store 0.0 3 Venue Category_Multicuisine Indian Restaurant 0.0 4 Venue
```


Category_Multiplex 0.0 ----- CV Raman Nagar ----- Venue Frequency 0 Venue Category_Pizza Place 0.29 1 Venue Category_Indian Restaurant 0.29 2 Venue Category_Department Store 0.14 3 Venue Category_Park 0.14 4 Venue Category_Shop & Service 0.14 ----- Cantonment area ----- - Venue Frequency 0 Venue Category_Indian Restaurant 0.4 1 Venue Category_Seafood Restaurant 0.2 2 Venue Category_Bookstore 0.2 3 Venue Category_Restaurant 0.2 4 Venue Category_Outlet Store 0.0 ----- Domlur ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.33 1 Venue Category_Café 0.17 2 Venue Category_Rajasthani Restaurant 0.08 3 Venue Category_Asian Restaurant 0.08 4 Venue Category_Pizza Place 0.08 ----- Electronic City ----- Venue Frequency 0 Venue Category_Outlet Store 0.2 1 Venue Category_Auto Garage 0.2 2 Venue Category_Bus Stop 0.2 3 Venue Category_Toll Plaza 0.2 4 Venue Category_Furniture / Home Store 0.2 ----- Girinagar ----- Venue Frequency 0 Venue Category_Fast Food Restaurant 0.25 1 Venue Category_Soccer Field 0.25 2 Venue Category_Park 0.25 3 Venue Category_Ice Cream Shop 0.25 4 Venue Category_North Indian Restaurant 0.00 ----- Gottigere ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.67 1 Venue Category_Department Store 0.11 2 Venue Category_Grocery Store 0.11 3 Venue Category_Pharmacy 0.11 4 Venue Category_North Indian Restaurant 0.00 ----- HBR Layout ----- Venue Frequency 0 Venue Category_Restaurant 0.2 1 Venue Category_Road 0.2 2 Venue Category_Coffee Shop 0.2 3 Venue Category_Café 0.2 4 Venue Category_North Indian Restaurant 0.2 ----- HSR Layout ----- Venue Frequency 0 Venue Category_Punjabi Restaurant 0.14 1 Venue Category_Chinese Restaurant 0.14 2 Venue Category_Pizza Place 0.14 3 Venue Category_Indian Restaurant 0.14 4 Venue Category_Café 0.14 ----- Hebbal ----- Venue Frequency 0 Venue Category_Department Store 0.2 1 Venue Category_Indian Restaurant 0.2 2 Venue Category_Market 0.2 3 Venue Category_Coffee Shop 0.2 4 Venue Category_Park 0.2 ----- Hoodi - ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.4 1 Venue Category_Yoga Studio 0.2 2 Venue Category_Breakfast Spot 0.2 3 Venue Category_Bus Station 0.2 4 Venue Category_North Indian Restaurant 0.0 ----- Horamavu ----- Venue Frequency 0 Venue Category_Bakery 0.33 1 Venue Category_Indian Restaurant 0.33 2 Venue Category_Cosmetics Shop 0.33 3 Venue Category_Paintball Field 0.00 4 Venue Category_Multiplex 0.00 ----- Hulimavu ----- Venue Frequency 0 Venue Category_Bakery 0.25 1 Venue Category_South Indian Restaurant 0.12 2 Venue Category_Juice Bar 0.12 3 Venue Category_Gym / Fitness Center 0.12 4 Venue Category_Indian Restaurant 0.12 ----- Indiranagar ----- - Venue Frequency 0 Venue Category_Café 0.10 1 Venue Category_Pub 0.09 2 Venue Category_Lounge 0.07 3 Venue Category_Indian Restaurant 0.06 4 Venue Category_Ice Cream Shop 0.06 ----- J. P. Nagar ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.29 1 Venue Category_Snack Place 0.06 2 Venue Category_Chinese Restaurant 0.06 3 Venue Category_Diner 0.03 4 Venue Category_Hyderabadi Restaurant 0.03 ----- Jalahalli ----- Venue Frequency 0 Venue Category_Indie Movie Theater 0.33 1 Venue Category_Playground 0.33 2 Venue Category_Convenience Store 0.33 3 Venue Category_ATM 0.00 4 Venue Category_Outlet Store 0.00 ----- Jayanagar ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.36 1 Venue Category_Flea Market 0.07 2 Venue Category_Multiplex 0.07 3 Venue Category_Fast Food Restaurant 0.07 4 Venue Category_Asian Restaurant 0.07 ----- Jeevanbheemanagar ----- Venue Frequency 0 Venue Category_Fast Food Restaurant 0.50 1 Venue Category_Pizza Place 0.25 2 Venue Category_Gym 0.25 3 Venue Category_ATM 0.00 4 Venue Category_Nightclub 0.00 ----- Kamakshipalya ----- Venue Frequency 0 Venue Category_South Indian Restaurant 1.0 1 Venue Category_ATM 0.0 2 Venue Category_Movie Theater 0.0 3 Venue Category_Pizza Place 0.0 4 Venue Category_Pharmacy 0.0 ----- Kammanahalli ----- Venue Frequency 0 Venue Category_Kerala Restaurant 0.1 1 Venue Category_Department Store 0.1 2 Venue Category_Snack Place 0.1 3 Venue Category_Chinese Restaurant 0.1 4 Venue Category_Fast Food Restaurant 0.1 ----- Kengeri ----- Venue Frequency 0 Venue Category_Italian Restaurant 0.25 1 Venue Category_Food Court 0.25 2 Venue Category_Indian Restaurant 0.25 3 Venue Category_Restaurant 0.25 4 Venue Category_North Indian Restaurant 0.00 ----- Koramangala ----- Venue Frequency 0 Venue Category_Hotel Bar 0.17 1 Venue Category_Restaurant 0.17 2 Venue Category_Indian Restaurant 0.17 3 Venue Category_Breakfast Spot 0.17 4 Venue Category_Bakery 0.17 ----- Kothnur ----- Venue Frequency 0 Venue Category_Department Store 0.2 1 Venue Category_Women's Store 0.2 2 Venue Category_Athletics & Sports 0.2 3 Venue Category_Restaurant 0.2 4 Venue Category_Performing Arts Venue 0.2 ----- Krishnarajapuram ----- Venue Frequency 0 Venue Category_Clothing Store 1.0 1 Venue Category_ATM 0.0 2 Venue Category_Mughlai Restaurant 0.0 3 Venue Category_Playground 0.0 4 Venue Category_Pizza Place 0.0 ----- Kumaraswamy Layout ----- Venue Frequency 0 Venue Category_Breakfast Spot 0.17 1 Venue Category_Indian Restaurant 0.17 2 Venue Category_Fast Food Restaurant 0.17 3 Venue Category_Sandwich Place 0.17 4 Venue Category_Café 0.17 ----- Lingarajapuram ----- Venue Frequency 0 Venue Category_Train Station 0.5 1 Venue Category_Electronics Store 0.5 2 Venue Category_Outlet Store 0.0 3 Venue Category_Multicuisine Indian Restaurant 0.0 4 Venue Category_Multiplex 0.0 ----- Madiwala ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.32 1 Venue Category_Fast Food Restaurant 0.09 2 Venue Category_Diner 0.05 3 Venue Category_Neighborhood 0.05 4 Venue Category_Halal Restaurant 0.05 ----- Mahadevapura ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.4 1 Venue Category_Bus Station 0.2 2 Venue Category_Convenience Store 0.2 3 Venue Category_Shopping Mall 0.2 4 Venue Category_ATM 0.0 ----- Mahalakshmi Layout ----- Venue Frequency 0 Venue Category_Coffee Shop 0.18 1 Venue Category_Hotel 0.18 2 Venue Category_Convenience Store 0.09 3 Venue Category_Grocery Store 0.09 4 Venue Category_Department Store 0.09 ----- Malleswaram ----- Venue Frequency 0 Venue Category_Ice Cream Shop 0.50 1 Venue Category_Bakery 0.17 2 Venue Category_Breakfast Spot 0.17 3 Venue Category_Indian Restaurant 0.17 4 Venue Category_Outlet Store 0.00 ----- Marathahalli ----- Venue Frequency 0 Venue Category_Clothing Store 0.43 1 Venue Category_Indian Restaurant 0.29 2 Venue Category_Movie Theater 0.14 3 Venue Category_Bakery 0.14 4 Venue Category_Pizza Place 0.00 ----- Mathikere ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.38 1 Venue Category_Kerala Restaurant 0.06 2 Venue Category_Ice Cream Shop 0.06 3 Venue Category_Shop & Service 0.06 4 Venue Category_Shoe Store 0.06 ----- Nagarbhavi ----- Venue Frequency 0 Venue Category_Gym 0.17 1 Venue Category_Andhra Restaurant 0.17 2 Venue Category_Café 0.17 3 Venue Category_Breakfast Spot 0.17 4 Venue Category_Indian Restaurant 0.17 ----- Nandini Layout ----- Venue Frequency 0 Venue Category_Vegetarian / Vegan Restaurant 1.0 1 Venue Category_ATM 0.0 2 Venue Category_North Indian Restaurant 0.0 3 Venue Category_Mughlai Restaurant 0.0 4 Venue Category_Multicuisine Indian Restaurant 0.0 ----- Nayandahalli ----- Venue Frequency 0 Venue Category_Electronics Store 0.25 1 Venue Category_Road 0.25 2 Venue Category_Fast Food Restaurant 0.25 3 Venue Category_Toll Plaza 0.25 4 Venue Category_ATM 0.00 ----- Padmanabhanagar ----- Venue Frequency 0 Venue Category_Café 0.67 1 Venue Category_Snack Place 0.33 2 Venue Category_ATM 0.00 3 Venue Category_North Indian Restaurant 0.00 4 Venue Category_Multiplex 0.00 ----- Pete area ----- Venue Frequency 0 Venue Category_Market 0.25 1 Venue Category_Bus Station 0.25 2 Venue Category_Park 0.25 3 Venue Category_Historic Site 0.25 4 Venue Category_ATM 0.00 ----- R. T. Nagar ----- Venue Frequency 0 Venue Category_Juice Bar 0.14 1 Venue Category_Gym 0.14 2 Venue Category_Resort 0.14 3 Venue Category_Fast Food Restaurant 0.14 4 Venue Category_Park 0.14 ----- Rajajinagar ----- Venue Frequency 0 Venue Category_Bakery 0.25 1 Venue Category_Indian Restaurant 0.25 2 Venue Category_Snack Place 0.25 3 Venue Category_Park 0.17 4 Venue Category_Pharmacy 0.08 ----- Rajarajeshwari Nagar ----- Venue Frequency 0 Venue Category_Pizza Place 0.17 1 Venue Category_Food Court 0.17 2 Venue Category_Café 0.17 3 Venue Category_Indian Chinese Restaurant 0.17 4 Venue Category_Ice Cream Shop 0.17 ----- Ramamurthy Nagar ----- Venue Frequency 0 Venue Category_ATM 0.17 1 Venue Category_Bakery 0.17 2 Venue Category_Multicuisine Indian Restaurant 0.17 3 Venue Category_South Indian Restaurant 0.17 4 Venue Category_Supermarket 0.17 ----- Sadashivanagar ----- Venue Frequency 0 Venue Category_Coffee Shop 0.16 1 Venue Category_Café 0.11 2 Venue Category_Department Store 0.11 3 Venue Category_Ice Cream Shop 0.11 4 Venue Category_Women's Store 0.05 ----- Seshadripuram ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.25 1 Venue Category_Hotel 0.12 2 Venue Category_Asian Restaurant 0.12 3 Venue Category_Chaat Place 0.12 4 Venue Category_Chinese Restaurant 0.12 ----- Shivajinagar ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.38 1 Venue Category_Fast Food Restaurant 0.15 2 Venue Category_Donut Shop 0.08 3 Venue Category_South Indian Restaurant 0.08 4 Venue Category_Furniture / Home Store 0.08 ----- Ulsoor ----- Venue Frequency 0 Venue Category_Café 0.4 1 Venue Category_Bridal Shop 0.2 2 Venue Category_Burger Joint 0.2 3 Venue Category_Bakery 0.2 4 Venue Category_Pizza Place 0.0 ----- Uttarahalli ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.27 1 Venue Category_Bakery 0.18 2 Venue Category_Gym 0.09 3 Venue Category_Paintball Field 0.09 4

Venue Category_Fast Food Restaurant 0.09 ----- Varthur ----- Venue Frequency 0 Venue Category_Supermarket 1.0 1 Venue Category_ATM 0.0 2 Venue Category_Nightclub 0.0 3 Venue Category_Mughlai Restaurant 0.0 4 Venue Category_Multicuisine Indian Restaurant 0.0 -----
 Vasanth Nagar ----- Venue Frequency 0 Venue Category_Indian Restaurant 0.20 1 Venue Category_Coffee Shop 0.16 2 Venue Category_Italian Restaurant 0.08 3 Venue Category_Hotel 0.08 4 Venue Category_Chinese Restaurant 0.08 ----- Vidyaranyapura ----- Venue Frequency 0
 Venue Category_Bus Station 1.0 1 Venue Category_ATM 0.0 2 Venue Category_Mughlai Restaurant 0.0 3 Venue Category_Playground 0.0 4 Venue Category_Pizza Place 0.0 ----- Vijayanagar ----- Venue Frequency 0 Venue Category_Bakery 0.50 1 Venue Category_Indian Restaurant 0.25 2 Venue Category_North Indian Restaurant 0.25 3 Venue Category_Outlet Store 0.00 4 Venue Category_Multicuisine Indian Restaurant 0.00 -----
 Whitefield ----- Venue Frequency 0 Venue Category_Bakery 0.22 1 Venue Category_Swiss Restaurant 0.11 2 Venue Category_Kerala Restaurant 0.11 3 Venue Category_Eastern European Restaurant 0.11 4 Venue Category_Café 0.11 ----- Yelahanka ----- Venue Frequency 0 Venue Category_Jewelry Store 0.17 1 Venue Category_Smoke Shop 0.17 2 Venue Category_Indian Restaurant 0.17 3 Venue Category_Restaurant 0.17 4 Venue Category_Train Station 0.17 ----- Yeshwanthpur ----- Venue Frequency 0 Venue Category_Bar 0.12 1 Venue Category_Restaurant 0.12 2 Venue Category_Chinese Restaurant 0.12 3 Venue Category_Mediterranean Restaurant 0.12 4 Venue Category_Hotel 0.12

Frequency of each neighborhood and its top 5 venues can be known. This is very important is we can analyze top neighborhoods with most busy restaurants

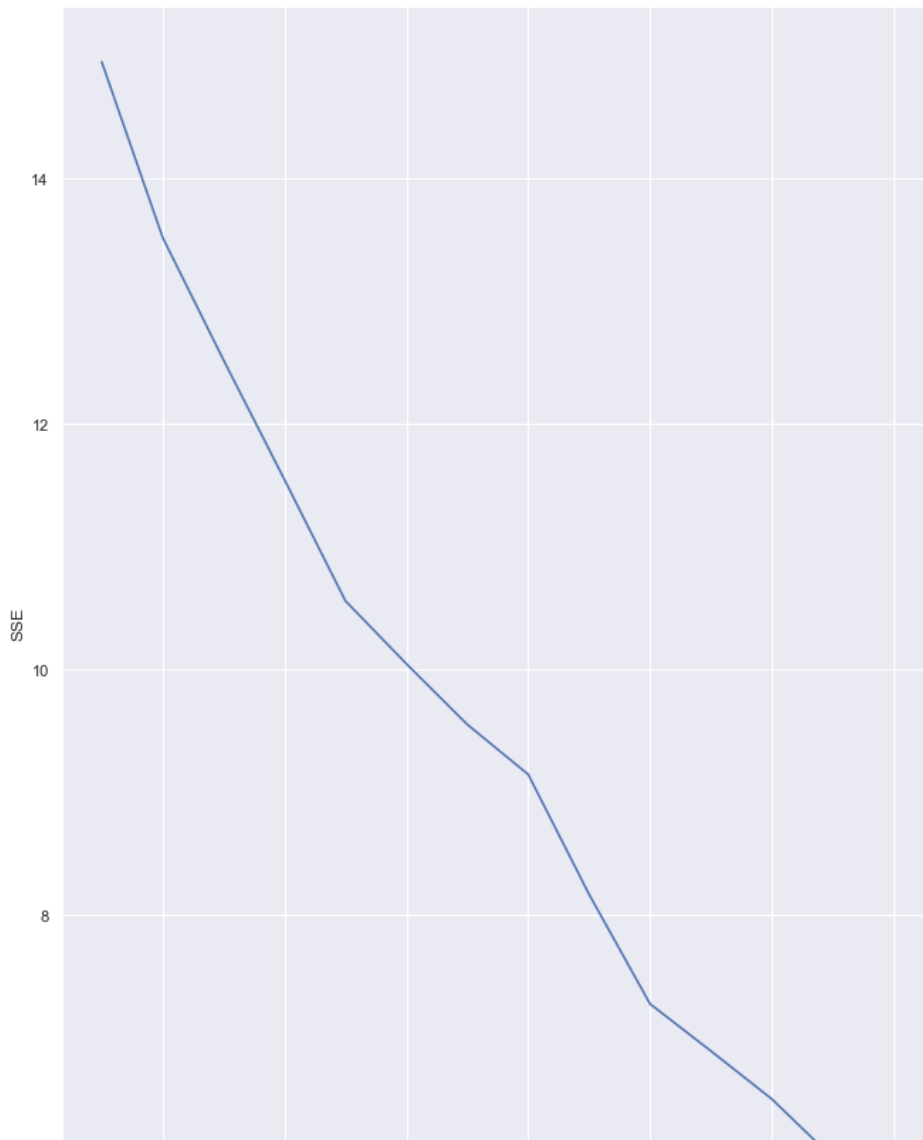
Next challenge is to find the optimal k value for clustering and we do it using the elbow method

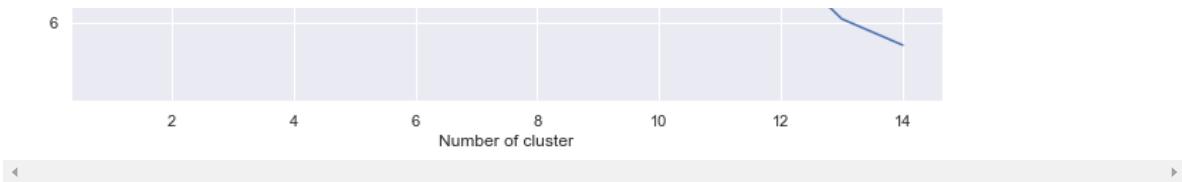
```
In [ ]: sse = {}
for k in range(1,15):
    kmeans = KMeans(n_clusters=k,random_state=0)
    kmeans.fit(gurgaon_venues_grouped.drop('Neighborhoods',axis=1))
    bangalore_venues_grouped['Cluster'] = kmeans.labels_
    sse[k] = kmeans.inertia_

plt.figure()
plt.plot(list(sse.keys()), list(sse.values()))
plt.xlabel("Number of cluster")
plt.ylabel("SSE")
plt.show()
```

```
In [78]: from IPython.display import Image
from IPython.core.display import HTML
Image(url= "https://github.com/DeepentiA/Capstone_Project/blob/master/download.png?raw=true")
```

Out[78]:





From the above graph, we can see the optimal value for cluster is 5

```
In [ ]: kmeans = KMeans(n_clusters=5, random_state=0)
```

```
In [ ]: kmeans.fit(gurgaon_venues_grouped.drop('Neighborhoods', axis=1))
```

KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300, n_clusters=5, n_init=10, n_jobs=None, precompute_distances='auto', random_state=0, tol=0.0001, verbose=0)

```
In [ ]: gurgaon_venues_grouped.groupby('Cluster')['Neighborhoods'].count()
```

Cluster 0 22 1 4 2 10 3 17 4 9 Name: Neighborhood, dtype: int64

```
In [ ]: gurgaon_venues_grouped.columns
```

Index(['Neighborhood', 'Venue Category_ATM', 'Venue Category_American Restaurant', 'Venue Category_Andhra Restaurant', 'Venue Category_Art Museum', 'Venue Category_Asian Restaurant', 'Venue Category_Athletics & Sports', 'Venue Category_Auto Garage', 'Venue Category_BBQ Joint', 'Venue Category_Badminton Court', ... 'Venue Category_Toll Plaza', 'Venue Category_Train Station', 'Venue Category_Travel & Transport', 'Venue Category_Udupi Restaurant', 'Venue Category_Vegetarian / Vegan Restaurant', 'Venue Category_Vietnamese Restaurant', 'Venue Category_Wine Bar', 'Venue Category_Women's Store', 'Venue Category_Yoga Studio', 'Cluster'], dtype='object', length=134)

```
In [ ]: # create map
map_clusters = folium.Map(location=[gurgaon_latitude, gurgaon_longitude], zoom_start=11)
```

```
In [ ]: # set color scheme for the clusters
x = np.arange(6)
ys = [i + x + (i*x)**2 for i in range(6)]
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(neighborhoods_venues_sorted['Neighborhood Latitude'], neighborhoods_venues_sorted['Neighborhood Longitude'], neighborhoods_venues_sorted['Neighborhood'], neighborhoods_venues_sorted['Cluster']):
    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)
```

```
In [ ]: top3 = recommended_neighborhoods.groupby(['Neighborhoods', '1st Most Common Venue', '2nd Most Common Venue', '3rd Most Common Venue'])['Ranking'].unique()
```

```
In [ ]: top3_df = pd.DataFrame(top3).reset_index()
```

```
In [86]: url= "https://raw.githubusercontent.com/DeepentiA/Capstone_Project/master/best%20food.csv"
df5 = pd.read_csv(url)
df5.head()
```

Out[86]:

	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Ranking
0	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	[0.7361321887351776]
1	Hulimavu	Venue Category_Bakery	Venue Category_Indian Restaurant	Venue Category_Badminton Court	[0.7638135476902764]
2	Kamakshipalya	Venue Category_South Indian Restaurant	Venue Category_Yoga Studio	Venue Category_Food Truck	[0.80418735993893]

Conclusion :

Here, according to the data, we see that 'Begur, Hulimavu and kamakshipalya' are top 3 neighborhoods to find similar food to whitefield