In [ ]:

```
## Aryan Varshney - 101903605 & Kanishk Chawla - 101903578
## ML Project
## Submitted To :- Dr.Joohi Chauhan Mam
```

In [1]:

```
#Importing all the required libraries.

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 [2]:

```
df = pd.read_csv('data/complete_dataset.csv')
```

In [3]:

```
df.head(10)
```

Out[3]:

	Unnamed: 0	Borough	Neighborhoods	Latitude	Longitude	Population	City	Avera
0	0	Central	Cantonment area	12.972442	77.580643	866377	Bangalore	189
1	1	Central	Domlur	12.960992	77.638726	743186	Bangalore	568
2	2	Central	Indiranagar	12.971891	77.641151	474289	Bangalore	419
3	3	Central	Jeevanbheemanagar	12.962900	77.659500	527874	Bangalore	66
4	4	Central	Malleswaram	13.003100	77.564300	893629	Bangalore	532
5	5	Central	Pete area	12.962700	77.575800	730999	Bangalore	507
6	6	Central	Rajajinagar	12.990100	77.552500	981362	Bangalore	609
7	7	Central	Sadashivanagar	13.006800	77.581300	662625	Bangalore	599
8	8	Central	Seshadripuram	12.993500	77.578700	396862	Bangalore	584
9	9	Central	Shivajinagar	12.985700	77.605700	77836	Bangalore	558

In [4]:

```
df.drop('Unnamed: 0',axis=1,inplace=True)
```

Creating 'Population' and 'Income' dataframe from the main dataframe

In [5]:

```
bangalore_population = pd.DataFrame(df[['Borough','Neighborhoods','Population']])
```

In [6]:

```
bangalore_population.head()
```

Out[6]:

	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 [7]:

```
bangalore_population.to_csv('data/population_dataset.csv')
```

In [8]:

```
bangalore_income = pd.DataFrame(df[['Borough','Neighborhoods','AverageIncome']])
```

In [9]:

```
bangalore_income.head()
```

Out[9]:

	Borough	Neighborhoods	AverageIncome
0	Central	Cantonment area	18944.099792
1	Central	Domlur	56837.022198
2	Central	Indiranagar	41991.817435
3	Central	Jeevanbheemanagar	6667.447632
4	Central	Malleswaram	53270.063892

In [10]:

```
bangalore_income.to_csv('data/income_dataset.csv')
```

In [11]:

```
bangalore_latitude = df['Latitude'].mean()
bangalore_longitude = df['Longitude'].mean()
print("Latitude and Longitude of Bangalore are : ",bangalore_latitude,bangalore_lon
```

```
Latitude and Longitude of Bangalore are :  12.962339620312497 77.60175
294687502
```

In [12]:

```
CLIENT_ID = 'ZK40VO3EPO1ZY5C04PNYSMQH2ZFZ3U1TPL0QME3WEY553DMH'
CLIENT_SECRET = 'RQBRNE2T3OTMYY4BBJ3YWLS0P2FK413O22MQV1QHBHAD04WG'
VERSION = '20180606'
LIMIT = 150
```

In [13]:

```
unique_boroughs_of_bangalore = df['Borough'].unique().tolist()
```

In [14]:

```
unique_boroughs_of_bangalore
```

Out[14]:

```
['Central',
 'Eastern',
 'NorthEastern',
 'Northern',
 'SouthEastern',
 'Southern',
 'SouthernSuburbs',
 'Western']
```

In [15]:

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

In [16]:

```
borough_colors
```

Out[16]:

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

In [17]:

```
bangalore_map = folium.Map(location=[bangalore_latitude,bangalore_longitude],zoom_s
```

In [18]:

```
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(bangalore_map)
```

In [19]:

```
bangalore_map
```

Out[19]:

Exploring bangalore Neighborhoods using FourSquare API

In [20]:

```
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_se
            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 ve
    nearby_venues.columns = ['Neighborhood',
                             'Borough',
                             'Neighborhood Latitude',
                             'Neighborhood Longitude',
                             'Venue',
                             'Venue Latitude',
                             'Venue Longitude',
                             'Venue Category']

    return(nearby_venues)
```

In [21]:

```
bangalore_venues = getNearbyVenues(names=df['Neighborhoods'],latitudes=df['Latitude
```

```
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 :  Vidyaranyapura
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 :  Kamakshipalya
Fetching venues for :  Kengeri
Fetching venues for :  Mahalakshmi Layout
Fetching venues for :  Nagarbhavi
Fetching venues for :  Nandini Layout
Fetching venues for :  Nayandahalli
Fetching venues for :  Rajarajeshwari Nagar
Fetching venues for :  Vijayanagar
```

In [22]:

```
print("Total number of venues found in Bangalore are : ",bangalore_venues.shape[0])
```

```
Total number of venues found in Bangalore are :  597
```

In [23]:

```
bangalore_venues.head(5)
```

Out[23]:

	Neighborhood	Borough	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Cat
0	Cantonment area	Central	12.972442	77.580643	Hotel Fishland	12.975569	77.578592	S Rest
1	Cantonment area	Central	12.972442	77.580643	Sapna Book House	12.976355	77.578461	Boo
2	Cantonment area	Central	12.972442	77.580643	Vasudev Adigas	12.973707	77.579257	Rest
3	Cantonment area	Central	12.972442	77.580643	Adigas Hotel	12.973554	77.579161	Rest
4	Cantonment area	Central	12.972442	77.580643	Kamat Yatrinivas	12.975985	77.578125	Rest

In [24]:

```
bangalore_venues.groupby('Venue Category').count()['Neighborhood'].sort_values(asce
```

Out[24]:

```
Venue Category
Indian Restaurant       114
Café                     33
Bakery                   28
Fast Food Restaurant     22
Ice Cream Shop           19
Pizza Place              17
Department Store         16
Restaurant               15
Coffee Shop              15
Chinese Restaurant       13
Name: Neighborhood, dtype: int64
```

In [25]:

```
print("Total number of unique categories in bangalore are : ",len(bangalore_venues[
```

```
Total number of unique categories in bangalore are :  132
```

Getting number of venues per neighborhood

In [26]:

```
individual_bangalore_venue_count = bangalore_venues.groupby(['Borough','Neighborhoo
```

In [27]:

```
individual_bangalore_venue_count
```

Out[27]:

Borough	Neighborhood	Borough
Central	Cantonment area	5
	Domlur	12
	Indiranagar	69
	Jeevanbheemanagar	4
	Malleswaram	6
	Pete area	4
	Rajajinagar	12
	Sadashivanagar	19
	Seshadripuram	8
	Shivajinagar	13
	Ulsoor	5
	Vasanth Nagar	25
Eastern	Bellandur	25
	CV Raman Nagar	7
	Hoodi	5
	Krishnarajapuram	1
	Mahadevapura	5
	Marathahalli	7
	Varthur	1
	Whitefield	9
NorthEastern	Banaswadi	5
	HBR Layout	5
	Horamavu	3
	Kammanahalli	10
	Lingarajapuram	2
	Ramamurthy Nagar	6
Northern	Hebbal	5
	Jalahalli	3
	Mathikere	16
	R. T. Nagar	7
	...	...
SouthEastern	Yeshwanthpur	8
	BTM Layout	33
	Bommanahalli	6

Borough	Neighborhood	Borough
Southern	Bommasandra	1
	Electronic City	5
	HSR Layout	7
	Koramangala	6
	Madiwala	22
	Banashankari	8
	Basavanagudi	11
	Girinagar	4
	J. P. Nagar	31
	Jayanagar	14
	Kumaraswamy Layout	6
	Padmanabhanagar	3
	Uttarahalli	11
SouthernSuburbs	Arekere	27
	Begur	4
	Gottigere	9
	Hulimavu	8
	Kothnur	5
Western	Basaveshwaranagar	20
	Kamakshipalya	1
	Kengeri	4
	Mahalakshmi Layout	11
	Nagarbhavi	6
	Nandini Layout	1
	Nayandahalli	4
	Rajarajeshwari Nagar	6
	Vijayanagar	4

62 rows × 1 columns

In [28]:

```
individual_bangalore_venue_count.rename(columns={'Borough':'NumberOfVenues'},inplac
```

In [29]:

```
individual_bangalore_venue_count.reset_index(inplace=True)
```

In [30]:

```
individual_bangalore_venue_count
```

Out[30]:

	Borough	Neighborhood	NumberOfVenues
0	Central	Cantonment area	5
1	Central	Domlur	12
2	Central	Indiranagar	69
3	Central	Jeevanbheemanagar	4
4	Central	Malleswaram	6
5	Central	Pete area	4
6	Central	Rajajinagar	12
7	Central	Sadashivanagar	19
8	Central	Seshadripuram	8
9	Central	Shivajinagar	13
10	Central	Ulsoor	5
11	Central	Vasanth Nagar	25
12	Eastern	Bellandur	25
13	Eastern	CV Raman Nagar	7
14	Eastern	Hoodi	5
15	Eastern	Krishnarajapuram	1
16	Eastern	Mahadevapura	5
17	Eastern	Marathahalli	7
18	Eastern	Varthur	1
19	Eastern	Whitefield	9
20	NorthEastern	Banaswadi	5
21	NorthEastern	HBR Layout	5
22	NorthEastern	Horamavu	3
23	NorthEastern	Kammanahalli	10
24	NorthEastern	Lingarajapuram	2
25	NorthEastern	Ramamurthy Nagar	6
26	Northern	Hebbal	5
27	Northern	Jalahalli	3
28	Northern	Mathikere	16
29	Northern	R. T. Nagar	7
...	...	...	...
32	Northern	Yeshwanthpur	8
33	SouthEastern	BTM Layout	33
34	SouthEastern	Bommanahalli	6
35	SouthEastern	Bommasandra	1

	Borough	Neighborhood	NumberOfVenues
36	SouthEastern	Electronic City	5
37	SouthEastern	HSR Layout	7
38	SouthEastern	Koramangala	6
39	SouthEastern	Madiwala	22
40	Southern	Banashankari	8
41	Southern	Basavanagudi	11
42	Southern	Girinagar	4
43	Southern	J. P. Nagar	31
44	Southern	Jayanagar	14
45	Southern	Kumaraswamy Layout	6
46	Southern	Padmanabhanagar	3
47	Southern	Uttarahalli	11
48	SouthernSuburbs	Arekere	27
49	SouthernSuburbs	Begur	4
50	SouthernSuburbs	Gottigere	9
51	SouthernSuburbs	Hulimavu	8
52	SouthernSuburbs	Kothnur	5
53	Western	Basaveshwaranagar	20
54	Western	Kamakshipalya	1
55	Western	Kengeri	4
56	Western	Mahalakshmi Layout	11
57	Western	Nagarbhavi	6
58	Western	Nandini Layout	1
59	Western	Nayandahalli	4
60	Western	Rajarajeshwari Nagar	6
61	Western	Vijayanagar	4

62 rows × 3 columns

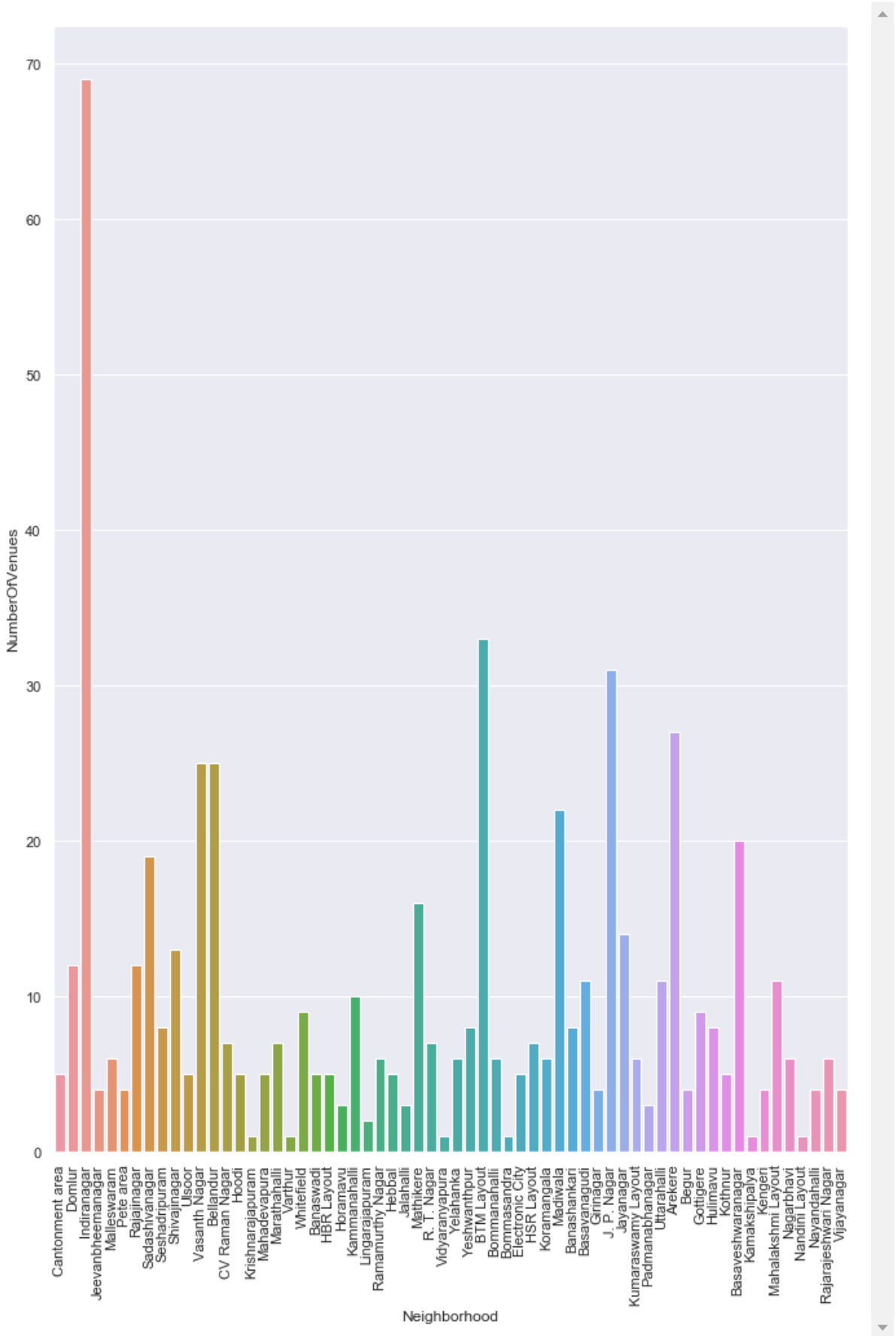
In [31]:

```
sns.set(rc={'figure.figsize':(11,16)})
plot = sns.barplot(x='Neighborhood',y='NumberOfVenues',data=individual_bangalore_ver
plot.set_xticklabels(individual_bangalore_venue_count['Neighborhood'],rotation=90)
```

Out[31]:

```
[Text(0,0,'Cantonment area'),
 Text(0,0,'Domlur'),
 Text(0,0,'Indiranagar'),
 Text(0,0,'Jeevanbheemanagar'),
 Text(0,0,'Malleswaram'),
 Text(0,0,'Pete area'),
 Text(0,0,'Rajajinagar'),
 Text(0,0,'Sadashivanagar'),
 Text(0,0,'Seshadripuram'),
 Text(0,0,'Shivajinagar'),
 Text(0,0,'Ulsoor'),
 Text(0,0,'Vasanth Nagar'),
 Text(0,0,'Bellandur'),
 Text(0,0,'CV Raman Nagar'),
 Text(0,0,'Hoodi'),
 Text(0,0,'Krishnarajapuram'),
 Text(0,0,'Mahadevapura'),
 Text(0,0,'Marathahalli'),
 Text(0,0,'Varthur'),
 Text(0,0,'Whitefield'),
 Text(0,0,'Banaswadi'),
 Text(0,0,'HBR Layout'),
 Text(0,0,'Horamavu'),
 Text(0,0,'Kammanahalli'),
 Text(0,0,'Lingarajapuram'),
 Text(0,0,'Ramamurthy Nagar'),
 Text(0,0,'Hebbal'),
 Text(0,0,'Jalahalli'),
 Text(0,0,'Mathikere'),
 Text(0,0,'R. T. Nagar'),
 Text(0,0,'Vidyaranyapura'),
 Text(0,0,'Yelahanka'),
 Text(0,0,'Yeshwanthpur'),
 Text(0,0,'BTM Layout'),
 Text(0,0,'Bommanahalli'),
 Text(0,0,'Bommasandra'),
 Text(0,0,'Electronic City'),
 Text(0,0,'HSR Layout'),
 Text(0,0,'Koramangala'),
 Text(0,0,'Madiwala'),
 Text(0,0,'Banashankari'),
 Text(0,0,'Basavanagudi'),
 Text(0,0,'Girinagar'),
 Text(0,0,'J. P. Nagar'),
 Text(0,0,'Jayanagar'),
 Text(0,0,'Kumaraswamy Layout'),
 Text(0,0,'Padmanabhanagar'),
 Text(0,0,'Uttarahalli'),
 Text(0,0,'Arekere'),
 Text(0,0,'Begur'),
 Text(0,0,'Gottigere'),
 Text(0,0,'Hulimavu'),
 Text(0,0,'Kothnur'),
```

```
Text(0,0,'Basaveshwaranagar'),
Text(0,0,'Kamakshipalya'),
Text(0,0,'Kengeri'),
Text(0,0,'Mahalakshmi Layout'),
Text(0,0,'Nagarbhavi'),
Text(0,0,'Nandini Layout'),
Text(0,0,'Nayandahalli'),
Text(0,0,'Rajarajeshwari Nagar'),
Text(0,0,'Vijayanagar')]
```



From the above graph we can see that inderanagar has most number of venues and soo on...

Exploring Whitefield venues

In [32]:

```
bangalore_venues[bangalore_venues['Neighborhood']=='Whitefield']
```

Out[32]:

	Neighborhood	Borough	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
233	Whitefield	Eastern	12.9698	77.7499	Golds Gym	12.972217	77.750529
234	Whitefield	Eastern	12.9698	77.7499	Herbs and Spices	12.968145	77.750862
235	Whitefield	Eastern	12.9698	77.7499	Chef Baker's	12.969730	77.751302
236	Whitefield	Eastern	12.9698	77.7499	Royal Orchid Suites	12.966774	77.751262
237	Whitefield	Eastern	12.9698	77.7499	The river side bar and kitchen	12.967298	77.749614
238	Whitefield	Eastern	12.9698	77.7499	Glen's Bakehouse	12.967490	77.749565
239	Whitefield	Eastern	12.9698	77.7499	refresh	12.965621	77.749234
240	Whitefield	Eastern	12.9698	77.7499	Cheenavala	12.966435	77.749368
241	Whitefield	Eastern	12.9698	77.7499	burgundy restaurant	12.966812	77.751380

One hot encoding for letting us to compare different venues based on some common scale

In [33]:

```
bangalore_venues_onehot = pd.get_dummies(bangalore_venues[['Venue Category']])
```

In [34]:

```
bangalore_venues_onehot
```

Out[34]:

	Venue Category_ATM	Venue Category_American Restaurant	Venue Category_Andhra Restaurant	Venue Category_Art Museum	Venue Category_Asian Restaurant	Ca
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	
5	0	0	0	0	0	
6	0	0	0	0	0	
7	0	0	0	0	0	
8	0	0	0	0	0	
9	0	0	0	0	0	
10	0	0	0	0	0	
11	0	0	0	0	0	
12	0	0	0	0	0	
13	0	0	0	0	0	
14	0	0	0	0	0	
15	0	0	0	0	0	
16	0	0	0	0	1	
17	0	0	0	0	0	
18	0	0	0	0	0	
19	0	0	0	0	0	
20	0	0	0	0	0	
21	0	0	0	0	0	
22	0	0	0	0	0	
23	0	0	0	0	0	
24	0	0	0	0	1	
25	0	0	0	0	0	
26	0	0	0	0	0	
27	0	0	0	0	0	
28	0	0	0	0	0	
29	0	0	0	0	0	
...	...	...	...	...	...	
567	0	0	0	0	0	
568	0	0	0	0	0	

	Venue Category_ATM	Venue Category_American Restaurant	Venue Category_Andhra Restaurant	Venue Category_Art Museum	Venue Category_Asian Restaurant	Ca
569	0	0	0	0	0	
570	0	0	0	0	0	
571	0	0	0	0	0	
572	0	0	0	0	0	
573	0	0	0	0	0	
574	0	0	0	0	0	
575	0	0	0	0	0	
576	0	0	0	0	0	
577	0	0	0	0	0	
578	0	0	0	0	0	
579	0	0	1	0	0	
580	0	0	0	0	0	
581	0	0	0	0	0	
582	0	0	0	0	0	
583	0	0	0	0	0	
584	0	0	0	0	0	
585	0	0	0	0	0	
586	0	0	0	0	0	
587	0	0	0	0	0	
588	0	0	0	0	0	
589	0	0	0	0	0	
590	0	0	0	0	0	
591	0	0	0	0	0	
592	0	0	0	0	0	
593	0	0	0	0	0	
594	0	0	0	0	0	
595	0	0	0	0	0	
596	0	0	0	0	0	

597 rows × 132 columns

In [35]:

```
bangalore_venues_onehot['Neighborhood'] = bangalore_venues['Neighborhood']
bangalore_venues_grouped = bangalore_venues_onehot.groupby('Neighborhood').mean().r
bangalore_venues_grouped
```

Out[35]:

	Neighborhood	Venue Category_ATM	Venue Category_American Restaurant	Venue Category_Andhra Restaurant	Venue Category_Art Museum	Ca
0	Arekere	0.000000	0.000000	0.000000	0.00	
1	BTM Layout	0.000000	0.000000	0.000000	0.00	
2	Banashankari	0.000000	0.000000	0.000000	0.00	
3	Banaswadi	0.000000	0.000000	0.000000	0.00	
4	Basavanagudi	0.000000	0.000000	0.000000	0.00	
5	Basaveshwaranagar	0.000000	0.000000	0.000000	0.00	
6	Begur	0.000000	0.000000	0.000000	0.00	
7	Bellandur	0.000000	0.000000	0.000000	0.00	
8	Bommanahalli	0.000000	0.000000	0.000000	0.00	
9	Bommasandra	0.000000	0.000000	0.000000	0.00	
10	CV Raman Nagar	0.000000	0.000000	0.000000	0.00	
11	Cantonment area	0.000000	0.000000	0.000000	0.00	
12	Domlur	0.000000	0.000000	0.000000	0.00	
13	Electronic City	0.000000	0.000000	0.000000	0.00	
14	Girinagar	0.000000	0.000000	0.000000	0.00	
15	Gottigere	0.000000	0.000000	0.000000	0.00	
16	HBR Layout	0.000000	0.000000	0.000000	0.00	
17	HSR Layout	0.000000	0.000000	0.000000	0.00	
18	Hebbal	0.000000	0.000000	0.000000	0.00	
19	Hoodi	0.000000	0.000000	0.000000	0.00	
20	Horamavu	0.000000	0.000000	0.000000	0.00	
21	Hulimavu	0.000000	0.000000	0.000000	0.00	
22	Indiranagar	0.000000	0.014493	0.000000	0.00	
23	J. P. Nagar	0.000000	0.000000	0.000000	0.00	
24	Jalahalli	0.000000	0.000000	0.000000	0.00	
25	Jayanagar	0.000000	0.000000	0.000000	0.00	
26	Jeevanbheemanagar	0.000000	0.000000	0.000000	0.00	
27	Kamakshipalya	0.000000	0.000000	0.000000	0.00	
28	Kammanahalli	0.000000	0.000000	0.100000	0.00	
29	Kengeri	0.000000	0.000000	0.000000	0.00	
...	...	...	...	...	...	
32	Krishnarajapuram	0.000000	0.000000	0.000000	0.00	

	Neighborhood	Venue Category_ATM	Venue Category_American Restaurant	Venue Category_Andhra Restaurant	Venue Category_Art Museum	Ca
33	Kumaraswamy Layout	0.000000	0.000000	0.000000	0.00	
34	Lingarajapuram	0.000000	0.000000	0.000000	0.00	
35	Madiwala	0.000000	0.000000	0.000000	0.00	
36	Mahadevapura	0.000000	0.000000	0.000000	0.00	
37	Mahalakshmi Layout	0.000000	0.000000	0.000000	0.00	
38	Malleswaram	0.000000	0.000000	0.000000	0.00	
39	Marathahalli	0.000000	0.000000	0.000000	0.00	
40	Mathikere	0.000000	0.000000	0.000000	0.00	
41	Nagarbhavi	0.000000	0.000000	0.166667	0.00	
42	Nandini Layout	0.000000	0.000000	0.000000	0.00	
43	Nayandahalli	0.000000	0.000000	0.000000	0.00	
44	Padmanabhanagar	0.000000	0.000000	0.000000	0.00	
45	Pete area	0.000000	0.000000	0.000000	0.00	
46	R. T. Nagar	0.000000	0.000000	0.000000	0.00	
47	Rajajinagar	0.000000	0.000000	0.000000	0.00	
48	Rajarajeshwari Nagar	0.000000	0.000000	0.000000	0.00	
49	Ramamurthy Nagar	0.166667	0.000000	0.000000	0.00	
50	Sadashivanagar	0.000000	0.000000	0.000000	0.00	
51	Seshadripuram	0.000000	0.000000	0.000000	0.00	
52	Shivajinagar	0.000000	0.000000	0.000000	0.00	
53	Ulsoor	0.000000	0.000000	0.000000	0.00	
54	Uttarahalli	0.000000	0.000000	0.000000	0.00	
55	Varthur	0.000000	0.000000	0.000000	0.00	
56	Vasanth Nagar	0.000000	0.000000	0.000000	0.04	
57	Vidyaranyapura	0.000000	0.000000	0.000000	0.00	
58	Vijayanagar	0.000000	0.000000	0.000000	0.00	
59	Whitefield	0.000000	0.000000	0.000000	0.00	
60	Yelahanka	0.000000	0.000000	0.000000	0.00	
61	Yeshwanthpur	0.000000	0.000000	0.000000	0.00	

62 rows × 133 columns

In [36]:

```
number_of_top_venues = 5
```

In [37]:

```
for hood in bangalore_venues_grouped['Neighborhood']:
    print('---------',hood,'---------')
    temp = bangalore_venues_grouped[bangalore_venues_grouped['Neighborhood'] == hood
    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
    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

In [38]:

```
def return_most_common_venues(row, number_of_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)
    return row_categories_sorted.index.values[0:number_of_top_venues]
```

In [39]:

```
number_of_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Neighborhood']
for ind in np.arange(number_of_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'] = bangalore_venues_grouped['Neighborhoo

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

neighborhoods_venues_sorted.head()
```

Out[39]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	Arekere	Venue Category_Indian Restaurant	Venue Category_Sporting Goods Shop	Venue Category_Department Store	Venue Category_Pizza Place
1	BTM Layout	Venue Category_Indian Restaurant	Venue Category_Snack Place	Venue Category_Ice Cream Shop	Venue Category_Bakery
2	Banashankari	Venue Category_Indian Restaurant	Venue Category_Café	Venue Category_Clothing Store	Venue Category_North Indian Restaurant
3	Banaswadi	Venue Category_Indian Restaurant	Venue Category_Café	Venue Category_Vegetarian / Vegan Restaurant	Venue Category_Bakery
4	Basavanagudi	Venue Category_Indian Restaurant	Venue Category_Hookah Bar	Venue Category_Metro Station	Venue Category_Mediterranean Restaurant

In [40]:

```
neighborhoods_venues_sorted.shape
```

Out[40]:

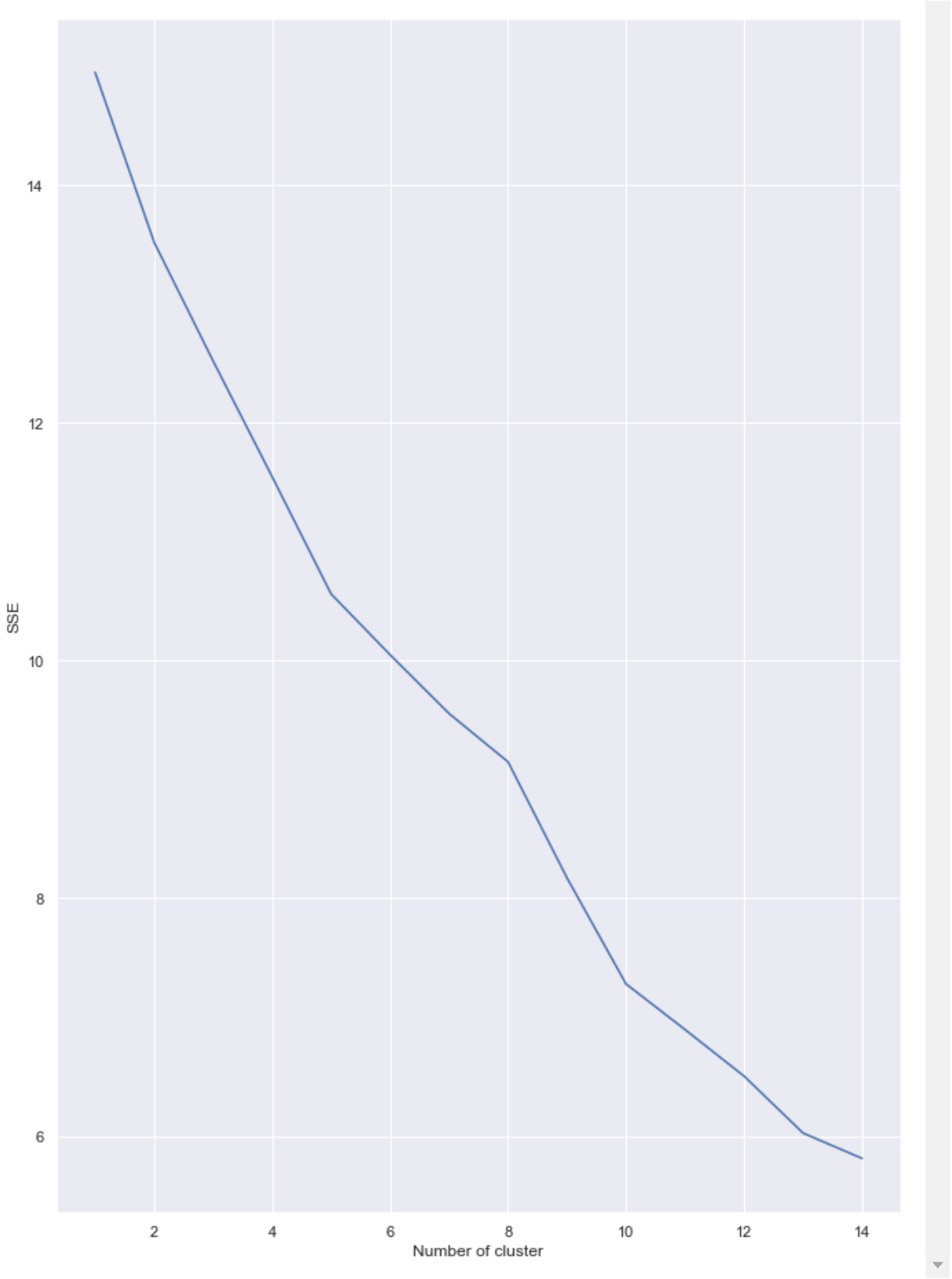
```
(62, 11)
```

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

In [41]:

```
sse = {}
for k in range(1,15):
    kmeans = KMeans(n_clusters=k,random_state=0)
    kmeans.fit(bangalore_venues_grouped.drop('Neighborhood',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()
```



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

In [42]:

```
kmeans = KMeans(n_clusters=5,random_state=0)
```

In [43]:

```
kmeans.fit(bangalore_venues_grouped.drop('Neighborhood',axis=1))
```

Out[43]:

```
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 [44]:

```
bangalore_venues_grouped['Cluster'] = kmeans.labels_
```

In [45]:

```
bangalore_venues_grouped.groupby('Cluster')['Neighborhood'].count()
```

Out[45]:

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

In [46]:

```
bangalore_venues_grouped.columns
```

Out[46]:

```
Index(['Neighborhood', 'Venue Category_ATM',
       'Venue Category_American Restaurant',
       'Venue Category_Andhra Restaurant', 'Venue Category_Art Museu
m',
       '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 Rest
aurant',
       'Venue Category_Vegetarian / Vegan Restaurant',
       'Venue Category_Vietnamese Restaurant', 'Venue Category_Wine Ba
r',
       'Venue Category_Women's Store', 'Venue Category_Yoga Studio',
       'Cluster'],
      dtype='object', length=134)
```

In [47]:

```
neighborhoods_venues_sorted = neighborhoods_venues_sorted.merge(bangalore_venues_gr
```

In [48]:

```
neighborhoods_venues_sorted.head(4)
```

Out[48]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	Co
0	Arekere	Venue Category_Indian Restaurant	Venue Category_Sporting Goods Shop	Venue Category_Department Store	Venue Category_Pizza Place	
1	BTM Layout	Venue Category_Indian Restaurant	Venue Category_Snack Place	Venue Category_Ice Cream Shop	Venue Category_Bakery	Cate
2	Banashankari	Venue Category_Indian Restaurant	Venue Category_Café	Venue Category_Clothing Store	Venue Category_North Indian Restaurant	Ca
3	Banaswadi	Venue Category_Indian Restaurant	Venue Category_Café	Venue Category_Vegetarian / Vegan Restaurant	Venue Category_Bakery	C

4 rows × 144 columns

In [49]:

```
neighborhoods_venues_sorted.columns
neighborhoods_venues_sorted = neighborhoods_venues_sorted.merge(bangalore_venues,on
```

In [50]:

```
# create map
map_clusters = folium.Map(location=[bangalore_latitude, bangalore_longitude], zoom_
```

In [51]:

```
# 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 Latitud
    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 [52]:

```
map_clusters
```

Out[52]:

Finding similar locations based on user input (recommending location)

We need to analyze factors such as population and Income to recommend.

In [53]:

```
bangalore_income.head()
```

Out[53]:

	Borough	Neighborhoods	AverageIncome
0	Central	Cantonment area	18944.099792
1	Central	Domlur	56837.022198
2	Central	Indiranagar	41991.817435
3	Central	Jeevanbheemanagar	6667.447632
4	Central	Malleswaram	53270.063892

In [54]:

```
bangalore_population.head()
```

Out[54]:

	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 [55]:

```
bangalore_population['Normalized_population'] = bangalore_population['Population']/
bangalore_population.head()
```

Out[55]:

	Borough	Neighborhoods	Population	Normalized_population
0	Central	Cantonment area	866377	0.880810
1	Central	Domlur	743186	0.755567
2	Central	Indiranagar	474289	0.482190
3	Central	Jeevanbheemanagar	527874	0.536668
4	Central	Malleswaram	893629	0.908516

In [56]:

```
bangalore_income['Normalized_income'] = bangalore_income['AverageIncome']/bangalore_
bangalore_income.head()
```

Out[56]:

	Borough	Neighborhoods	AverageIncome	Normalized_income
0	Central	Cantonment area	18944.099792	0.293051
1	Central	Domlur	56837.022198	0.879225
2	Central	Indiranagar	41991.817435	0.649581
3	Central	Jeevanbheemanagar	6667.447632	0.103140
4	Central	Malleswaram	53270.063892	0.824047

In [57]:

```
bangalore_venues_grouped.head(1)
```

Out[57]:

	Neighborhood	Venue Category_ATM	Venue Category_American Restaurant	Venue Category_Andhra Restaurant	Venue Category_Art Museum	Category Rest
0	Arekere	0.0	0.0	0.0	0.0	

1 rows × 134 columns

recommending Veg restaurants

In [58]:

```
bangalore_veg = bangalore_venues_onehot.groupby(['Neighborhood']).sum().reset_index
bangalore_veg.head()
```

Out[58]:

	Neighborhood	Venue Category_ATM	Venue Category_American Restaurant	Venue Category_Andhra Restaurant	Venue Category_Art Museum	Category Rest
0	Arekere	0	0	0	0	
1	BTM Layout	0	0	0	0	
2	Banashankari	0	0	0	0	
3	Banaswadi	0	0	0	0	
4	Basavanagudi	0	0	0	0	

5 rows × 133 columns

In [59]:

```
bangalore_veg['Venue Category_Vegetarian / Vegan Restaurant']
```

Out[59]:

```
0     0
1     2
2     0
3     1
4     0
5     0
6     0
7     0
8     0
9     0
10    0
11    0
12    0
13    0
14    0
15    0
16    0
17    0
18    0
19    0
20    0
21    0
22    0
23    1
24    0
25    0
26    0
27    0
28    0
29    0
     ..
32    0
33    0
34    0
35    0
36    0
37    0
38    0
39    0
40    0
41    0
42    1
43    0
44    0
45    0
46    0
47    0
48    0
49    0
50    0
51    0
52    0
53    0
54    0
55    0
```

```
56    1
57    0
58    0
59    0
60    0
61    0
Name: Venue Category_Vegetarian / Vegan Restaurant, Length: 62, dtyp
e: uint8
```

In [60]:

```
bangalore_veg = bangalore_veg[['Neighborhood','Venue Category_Vegetarian / Vegan Re
```

In [61]:

```
bangalore_veg.rename(columns={'Venue Category_Vegetarian / Vegan Restaurant':'Numbe
```

In [62]:

```
bangalore_veg.head()
```

Out[62]:

	Neighborhood	NumberOfVegRestaurants
0	Arekere	0
1	BTM Layout	2
2	Banashankari	0
3	Banaswadi	1
4	Basavanagudi	0

In [63]:

```
bangalore_veg['NumberOfNonVeganrestaurants'] = 1-(bangalore_veg['NumberOfVegRestaur
```

In [64]:

```
bangalore_veg.head(10)
```

Out[64]:

	Neighborhood	NumberOfVegRestaurants	NumberOfNonVeganrestaurants
0	Arekere	0	1.0
1	BTM Layout	2	0.0
2	Banashankari	0	1.0
3	Banaswadi	1	0.5
4	Basavanagudi	0	1.0
5	Basaveshwaranagar	0	1.0
6	Begur	0	1.0
7	Bellandur	0	1.0
8	Bommanahalli	0	1.0
9	Bommasandra	0	1.0

In [65]:

```
bangalore_veg.rename(columns={'Neighborhood':'Neighborhoods'},inplace=True)
```

Building a target neighborhood by providing a sample restaurant : say ' Whitefield '

In [66]:

```
target_cluster_dataframe = neighborhoods_venues_sorted.loc[neighborhoods_venues_sor
```

In [67]:

```
target_cluster_dataframe.reset_index()
```

Out[67]:

	index	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5t Cc
0	574	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
1	575	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
2	576	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
3	577	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
4	578	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
5	579	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
6	580	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
7	581	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category
8	582	Whitefield	Venue Category_Bakery	Venue Category_Gym / Fitness Center	Venue Category_Café	Venue Category_Hotel Bar	Category

9 rows × 152 columns

In [68]:

```
target_cluster = target_cluster_dataframe.iloc[0].at['Cluster']
```

In [69]:

```
target_cluster
```

Out[69]:

4

In [70]:

```
print("The target cluster is : ",target_cluster)
```

```
The target cluster is :  4
```

In [71]:

```
possible_neighborhoods = neighborhoods_venues_sorted[neighborhoods_venues_sorted['C
possible_neighborhoods.head()
```

Out[71]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	
104	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cat
105	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cat
106	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cat
107	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cat
207	Hulimavu	Venue Category_Bakery	Venue Category_Indian Restaurant	Venue Category_Badminton Court	Venue Category_South Indian Restaurant	Cat

5 rows × 151 columns

In [72]:

```
print("There are {} neighborhoods which has similar characteristics to Whitefield."
```

```
There are 51 neighborhoods which has similar characteristics to Whitef
ield.
```

In [73]:

```
possible_neighborhoods.reset_index().head()
```

Out[73]:

	index	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	104	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store
1	105	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store
2	106	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store
3	107	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store
4	207	Hulimavu	Venue Category_Bakery	Venue Category_Indian Restaurant	Venue Category_Badminton Court	Venue Category_South Indian Restaurant

5 rows × 152 columns

In [74]:

```
possible_neighborhoods.rename(columns={'Neighborhood':'Neighborhoods'},inplace=True
```

```
C:\Users\coreML\Anaconda3\lib\site-packages\pandas\core\frame.py:3781:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/indexing.html#indexing-view-versus-copy (http://pandas.pyd
ata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy)
  return super(DataFrame, self).rename(**kwargs)
```

Now we need to create ranking by combining population and income factors as well?

In [75]:

```
possible_neighborhoods = possible_neighborhoods.merge(bangalore_population[['Neighb
possible_neighborhoods = possible_neighborhoods.merge(bangalore_income[['Neighborho
possible_neighborhoods = possible_neighborhoods.merge(bangalore_veg[['Neighborhoods
```

In [76]:

```
possible_neighborhoods.head()
```

Out[76]:

	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	
0	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cate
1	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cate
2	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cate
3	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	Venue Category_Clothing Store	Cate
4	Hulimavu	Venue Category_Bakery	Venue Category_Indian Restaurant	Venue Category_Badminton Court	Venue Category_South Indian Restaurant	Cate

5 rows × 157 columns

In [77]:

```
possible_neighborhoods['Ranking'] = possible_neighborhoods['Normalized_population']
recommended_neighborhoods = possible_neighborhoods.sort_values(by='Ranking',ascendi
recommended_neighborhoods.reset_index(inplace=True, drop=True)
```

In [78]:

```
recommended_neighborhoods.head()
```

Out[78]:

	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th M Comr Ve
0	Malleswaram	Venue Category_Ice Cream Shop	Venue Category_Breakfast Spot	Venue Category_Bakery	Venue Category_Indian Restaurant	Ve Category_F C
1	Malleswaram	Venue Category_Ice Cream Shop	Venue Category_Breakfast Spot	Venue Category_Bakery	Venue Category_Indian Restaurant	Ve Category_F C
2	Malleswaram	Venue Category_Ice Cream Shop	Venue Category_Breakfast Spot	Venue Category_Bakery	Venue Category_Indian Restaurant	Ve Category_F C
3	Malleswaram	Venue Category_Ice Cream Shop	Venue Category_Breakfast Spot	Venue Category_Bakery	Venue Category_Indian Restaurant	Ve Category_F C
4	Malleswaram	Venue Category_Ice Cream Shop	Venue Category_Breakfast Spot	Venue Category_Bakery	Venue Category_Indian Restaurant	Ve Category_F C

5 rows × 158 columns

In [79]:

```
top3 = recommended_neighborhoods.groupby(['Neighborhoods','1st Most Common Venue','
```

In [80]:

```
top3_df = pd.DataFrame(top3).reset_index()
```

In [81]:

```
top3_df.head(3)
```

Out[81]:

	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]

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

Here our model will recommend these neighborneeds and top 3 common venues to visit.

Thank you for reviewing !

In [ ]:

```

```