Applied Data Science Capstone

Introduction

In [52]:

```
from IPython.display import Image
Image("Blr.jpg")
```

Out[52]:

History of Bengaluru

Bengaluru is the capital of the Indian state of Karnataka. It is a historic city noted for its many monuments, temples, mosques and bazaars. A multitude of influences has shaped the character of the city in the last 400 years.

The city is forming its role and outlook as part of the booming service industry revolution, and is trying to preserve and popularize its history.

The History of this city, has deeply affected the culture, language, and cuisine of the people living here, and the areas once part of Bengaluru state.

Bengaluru is the capital of southern India's Karnataka state. A major center for the technology industry, it's home to many upscale restaurants and shops. Its historic sites include Golconda Fort, a former diamond-trading center that was once the Qutb Shahi dynastic capital. The Charminar, a 16th-century mosque whose 4 arches support towering minarets, is an old city landmark near the long-standing Laad Bazaar.

Importing Libraries

In [2]:

```
from bs4 import BeautifulSoup
import requests
import numpy as np
import pandas as pd
pd.set_option('display.max_columns', None)
```

```
pd.set_option('display.max_rows', None)

import json # library to handle JSON files

import urllib
!conda install -c conda-forge geocoder --yes
import geocoder

!conda install -c conda-forge geopy --yes
from geopy.geocoders import Nominatim # convert an address into latitude and longitude
values import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas datafram

# Matplotlib and associated plotting
modules import matplotlib.cm as cm
import matplotlib.colors as colors

# import k-means from clustering stage
from sklearn.cluster import KMeans

!conda install -c conda-forge folium=0.5.0 --
yes import folium # map rendering library
print('Libraries imported.')
```

```
Solving environment: done

## Package Plan ##

  environment location: /home/jupyterlab/conda

  added / updated specs:
    - geocoder


The following packages will be downloaded:

    package                    |            build
                               |
    orderedset-2.0             |  ----------py36_0         231 KB conda-forge
    geocoder-1.38.1            |            py_0           52 KB conda-forge
    ratelim-0.1.6              |          py36_0            5 KB conda-forge
      ------------------------------------------------------------
                                           Total:         288 KB

The following NEW packages will be INSTALLED:

    geocoder:   1.38.1-py_0  conda-forge
    orderedset: 2.0-py36_0   conda-forge
    ratelim:    0.1.6-py36_0 conda-forge


Downloading and Extracting Packages
orderedset-2.0       | 231 KB    | ##################################### | 100%
geocoder-1.38.1      | 52 KB     | ##################################### | 100%
ratelim-0.1.6        | 5 KB      | ##################################### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
Solving environment: done

# All requested packages already

installed. Solving environment: done

# All requested packages already

installed. Libraries imported.
```

Data Section

In [3]:

```
url = "Blrdata.csv"
```

```
df = pd.read_csv(url)
df.head()
```

Out[3]:

	officename	pincode	officeType	Deliverystatus	divisionname	regionname	circlename	Taluk	Districtnam
0	A.Gs Office S.O	500004	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Khairatabad	Bengaluru
1	A.Gs. Staff Quarters S.O	500045	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Khairatabad	Bengaluru
2	Anandnagar S.O (Bengaluru)	500004	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Khairatabad	Bengaluru
3	AP Police Academy PO	500091	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Rajendra Nagar	Bengaluru
4	Appa Himayathsagar B.O	500008	B.O	Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Rajendranagar	K.V.Rangared

Focus is only on Blrerbad City

In [4]:

```
Blr_data = df.rename(columns={'divisionname' : 'Borough', 'officename' : 'Neighbourhood'})
Blr_data.head()
```

Out[4]:

	Neighbourhood	pincode	officeType	Deliverystatus	Borough	regionname	circlename	Taluk	Districtname
0	A.Gs Office S.O	500004	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Khairatabad	Bengaluru
1	A.Gs. Staff Quarters S.O	500045	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Khairatabad	Bengaluru
2	Anandnagar S.O (Bengaluru)	500004	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Khairatabad	Bengaluru
3	AP Police Academy PO	500091	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Rajendra Nagar	Bengaluru
4	Appa Himayathsagar B.O	500008	B.O	Delivery	Bengaluru City	Bengaluru City	Andhra Pradesh	Rajendranagar	K.V.Rangareddy

Focus is only on Blrerbad City

In [5]:

```
Blr_data = Blr_data[Blr_data.Borough == "Bengaluru
City"] Blr_data.tail()
```

Out[5]:

	Neighbourhood	pincode	officeType	Deliverystatus	Borough	regionname	circlename	Taluk	Districtname
72	Vidhan Sabha S.O (Bengaluru)	500004	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Karnataka	Nampally	Bengaluru

	Neighbourhood	pincode	officeType	Deliverystatus	Borough	regionname	circlename	Taluk	Districtname
73	Vijay Nagar Colony S.O (Bengaluru)	500057	S.O	Delivery	Bengaluru City	Bengaluru City	Karnataka	Asifnagar	Bengaluru
74	Vivekananda Nagar Colony S.O	500018	S.O	Non-Delivery	Bengaluru City	Bengaluru City	Karnataka	Balanagar	Bengaluru
75	Wattinagulapalli B.O	500075	B.O	Delivery	Bengaluru City	Bengaluru City	Karnataka	Rajendra Nagar	K.V.Rangareddy
76	Yousufguda S.O	500045	S.O	Delivery	Bengaluru City	Bengaluru City	Karnataka	Khairatabad	Bengaluru

Join all by Pincode

In [6]:

```
def neighbourhood_list(grouped):
    return ', '.join(sorted(grouped['Neighbourhood'].tolist()))

grp = Blr_data.groupby(['pincode', 'Borough'])
Blrcity_Blr_city = grp.apply(neighbourhood_list).reset_index(name = 'Neighbourhood')
```

Displaying Bengalurucity data

In [7]:

```
Blrcity_Blr_city.head()
```

Out[7]:

	pincode	Borough	Neighbourhood
0	500001	Bengaluru City	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...
1	500004	Bengaluru City	A.Gs Office S.O, Anandnagar S.O (Bengaluru), B...
2	500006	Bengaluru City	Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O
3	500008	Bengaluru City	Appa Himayathsagar B.O, Dargah Hussain Shahwal...
4	500018	Bengaluru City	Bharat Nagar Colony S.O, Erragadda S.O, Fathen...

Getting Coordinates as per pincode

In [8]:

```
def get_latlng(postal_code):
    # initialize your variable to
    None lat_lng_coords = None
    # loop until you get the coordinates
    while(lat_lng_coords is None):
        g = geocoder.arcgis('{}, Bengaluru,
        KARNATAKA'.format(postal_code)) lat_lng_coords = g.latlng
    return lat_lng_coords

print('done')
```

```
done
```

In [9]:

```
get_latlng('500001')
```

Out[9]:

```
[17.390585000000044, 78.47038817100008]
```

Now appling this functions for all pincode of areas of South Mumbai

In [10]:

```
latitude = []
longitude = []

for row in Blrcity_Blr_city['pincode']:
    coordinate = get_latlng(row)
    latitude.append(coordinate[0])
    longitude.append(coordinate[1])

Blrcity_Blr_city['latitude'] = latitude
Blrcity_Blr_city['longitude'] = longitude
Blrcity_Blr_city.head()
```

Out[10]:

	pincode	Borough	Neighbourhood	latitude	longitude
0	500001	Bengaluru City	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388
1	500004	Bengaluru City	A.Gs Office S.O, Anandnagar S.O (Bengaluru), B...	17.403781	78.462525
2	500006	Bengaluru City	Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O	17.371224	78.454180
3	500008	Bengaluru City	Appa Himayathsagar B.O, Dargah Hussain Shahwal...	17.396335	78.406792
4	500018	Bengaluru City	Bharat Nagar Colony S.O, Erragadda S.O, Fathen...	17.457435	78.445780

Next, we are going to start utilizing the Foursquare API to explore the South Mumbabi for data collection

In [11]:

```
CLIENT_ID = 'Q5U3IIIC5I51HWBVM3D1S5IG0GNAIPRLPE4R2CKMJGS2NBMD' # your Foursquare ID
CLIENT_SECRET = 'AONPJJBL5CACYQBOHBA3JKVMSSYFEA41VBSXAZHMKPXRX5ST' # your Foursquare
Secret VERSION = '20180605' # Foursquare API version

print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)
```

```
Your credentails:
CLIENT_ID: Q5U3IIIC5I51HWBVM3D1S5IG0GNAIPRLPE4R2CKMJGS2NBMD
CLIENT_SECRET:AONPJJBL5CACYQBOHBA3JKVMSSYFEA41VBSXAZHMKPXRX5ST
```

In [12]:

```
first_lat = Blrcity_Blr_city['latitude'][0]
first_long = Blrcity_Blr_city['longitude'][0]
radius = 500
LIMIT = 100

# create URL
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&r
adius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    first_lat,
    first_long,
    radius,
    LIMIT)
url # display URL
```

Out[12]:

```
'https://api.foursquare.com/v2/venues/explore?
&client_id=Q5U3IIIC5I51HWBVM3D1S5IG0GNAIPRLPE4R2CKMJGS2NBMD&client_secret=AONPJJBL5CACYQBOHBA3JKVMS
A41VBSXAZHMKPXRX5ST&v=20180605&ll=17.390585000000044,78.47038817100008&radius=500&limit=100'
```

Getting the results

In [13]:

```
results = requests.get(url).json()
```

In [14]:

```
# function that extracts the category of the
venue def get_category_type(row):
    try:
        categories_list =
    row['categories'] except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']
```

We are ready to clean the json and structure it into *pandas* data frame

In [15]:

```
venues = results['response']['groups'][0]['items']

nearby_venues = json_normalize(venues) # flatten JSON

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat',
'venue.location.lng'] nearby_venues =nearby_venues.loc[:, filtered_columns]
# filter the category for each row
nearby_venues['venue.categories'] = nearby_venues.apply(get_category_type, axis=1)

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

nearby_venues.columns] nearby_venues.head()
```

Out[15]:

	name	categories	lat	lng
0	Jagdish Market	Mobile Phone Shop	17.391815	78.474880
1	Al-Yousufain Fast Food Center	Fried Chicken Joint	17.391338	78.470147
2	Hotel Royalton	Hotel	17.393211	78.473504
3	Bengaluru Deccan Railway Station	Platform	17.392863	78.467555
4	Hollywood Foot Wear	Shoe Store	17.391627	78.474748

In [16]:

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

```
7 venues were returned by Foursquare.
```

Methodlogy Selection

Creating a Map

In [17]:

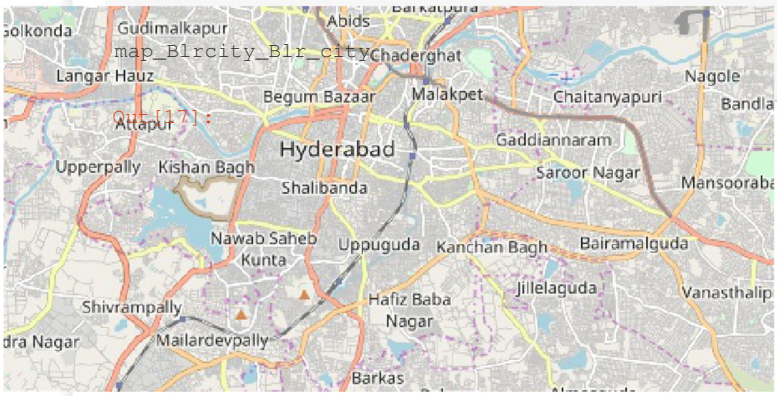
```
address = 'Bengaluru, KARNATAKA'

geolocator = Nominatim(user_agent="userid")
```

```
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude

# create map of South Mumbai using latitude and longitude values
map_Blrcity_Blr_city = folium.Map(location=[latitude, longitude], zoom_start=11.5)

# add markers to map
for lat, lng, borough, neighborhood in
zip(Blrcity_Blr_city['latitude'],Blrcity_Blr_city['longitude '],
Blrcity_Blr_city['Borough'], Blrcity_Blr_city['Neighbourhood']):
    label = '{}, {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='red',
        fill = True,
        fill_color='#cc3139',
        fill_opacity=0.7,
        parse_html=False).add_to(map_Blrcity_Blr_city)
```



Out[17]:

Explore Bengaluru City

In [18]:

```
def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?
&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
```

```
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby
        venue venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])
    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in
    venue_list]) nearby_venues.columns = ['Neighbourhood',
                  'Neighbourhood Latitude',
                  'Neighboruhood Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

In [19]:

```
Blrcity_venues = getNearbyVenues(names=Blrcity_Blr_city['Neighbourhood'],
                                 latitudes=Blrcity_Blr_city['latitude'],
                                 longitudes=Blrcity_Blr_city['longitude']
                                )
```

```
Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O, Seetharampet S.O, State Bank Of Bengaluru S.O
A.Gs Office S.O, Anandnagar S.O (Bengaluru), Bazarghat S.O (Bengaluru), Khairatabad H.O, Parishram
Bhawan S.O, Vidhan Sabha S.O (Bengaluru)
Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O
Appa Himayathsagar B.O, Dargah Hussain Shahwali B.O, Golconda S.O, Blrer Shah Kote B.O, Kakatiya
N agar S.O, Lunger House S.O, Nanakramguda B.O, Sakkubai Nagar S.O, Toli Chowki S.O
Bharat Nagar Colony S.O, Erragadda S.O, Fathenagar Colony S.O, Moosapet B.O, Sanath Nagar Colony S
.O, Sanathnagar I E S.O, Swarajyanagar S.O, Vivekananda Nagar Colony
S.O Lingampalli S.O
Central Secretariat S.O
Humayunnagar S.O, Murad Nagar S.O (Bengaluru), Shantinagar S.O (Bengaluru)
Gagan Mahal S.O, Himayathnagar S.O, Narayanguda S.O, Ramakrishna Mutt S.O
Ibrahim Bagh Lines S.O
Gachibowli S.O, Manuu S.O
Dr.B R Ambedkar O.U  S.O, Jubilee Hills S.O
Banjara Hills S.O
Sanjeev Reddy Nagar S.O, Vengal Rao Nagar S.O
Raj Bhawan S.O (Bengaluru)
A.Gs. Staff Quarters S.O, Yousufguda S.O
CUC S.O
Vijay Nagar Colony S.O (Bengaluru)
LIC Division S.O, New Mla Quarters S.O
Srinagar Colony S.O
Aziz Nagar B.O, Bhaskar Nagar S.O, C.B.I.T S.O, Gandipet B.O, Himayathnagar B.O, Janwada
B.O, Kokapet B.O, Mancherevula B.O, Narsingi B.O, Wattinagulapalli B.O Cyberabad S.O,
Madhapur B.O
I.M.Colony S.O, Somajiguda S.O
Kondapur B.O, Kothaguda S.O (K.V.Rangareddy)
Manikonda S.O
AP Police Academy PO, Blrershahkote S.O
Putlibowli S.O, State Bank Of India S.O
Film Nagar S.O
```

'Top 5 regions of Bengaluru City with Highest Number of Restaurants'

In [21]:

```
print(Blrcity_venues.shape)
```

```
Blrcity_venues.head()
```

```
(243, 7)
```

Out[21]:

	Neighbourhood	Neighbourhood Latitude	Neighboruhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388	Jagdish Market	17.391815	78.474880	Mobile Phone Shop
1	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388	Al-Yousufain Fast Food Center	17.391338	78.470147	Fried Chicken Joint
2	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388	Hotel Royalton	17.393211	78.473504	Hotel
3	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388	Bengaluru Deccan Railway Station	17.392863	78.467555	Platform
4	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388	Hollywood Foot Wear	17.391627	78.474748	Shoe Store

In [22]:

```
Blrcity_venues.groupby('Neighbourhood').count()
```

Out[22]:

	Neighbourhood Latitude	Neighboruhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighbourhood						
A.Gs Office S.O, Anandnagar S.O (Bengaluru), Bazarghat S.O (Bengaluru), Khairatabad H.O, Parishram Bhawan S.O, Vidhan Sabha S.O (Bengaluru)	15	15	15	15	15	15
A.Gs. Staff Quarters S.O, Yousufguda S.O	2	2	2	2	2	2
Appa Himayathsagar B.O, Dargah Hussain Shahwali B.O, Golconda S.O, Blrer Shah Kote B.O, Kakatiya Nagar S.O, Lunger House S.O, Nanakramguda B.O, Sakkubai Nagar S.O, Toli Chowki S.O	6	6	6	6	6	6
Aziz Nagar B.O, Bhaskar Nagar S.O, C.B.I.T S.O, Gandipet B.O, Himayathnagar B.O, Janwada B.O, Kokapet B.O, Mancherevula B.O, Narsingi B.O, Wattinagulapalli B.O	16	16	16	16	16	16
Banjara Hills S.O	9	9	9	9	9	9
Bharat Nagar Colony S.O, Erragadda S.O, Fathenagar Colony S.O, Moosapet B.O, Sanath Nagar Colony S.O, Sanathnagar I E S.O, Swarajyanagar S.O, Vivekananda Nagar Colony S.O	4	4	4	4	4	4
CUC S.O	2	2	2	2	2	2
Central Secretariat S.O	2	2	2	2	2	2
Cyberabad S.O, Madhapur B.O	36	36	36	36	36	36
Dr.B R Ambedkar O.U S.O, Jubilee Hills S.O	4	4	4	4	4	4

Film Nagar S.O	Neighbourhood 9 Latitude	Neighboruhood 9 Longitude	Venue 9	Venue 9 Latitude	Venue 9 Longitude	Venue 9 Category
Gachibowli S.O, Manuu S.O Neighbourhood	8	8	8	8	8	8
Gagan Mahal S.O, Himayathnagar S.O, Narayanguda S.O, Ramakrishna Mutt S.O	22	22	22	22	22	22
Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O, Seetharampet S.O, State Bank Of Bengaluru S.O	7	7	7	7	7	7
Humayunnagar S.O, Murad Nagar S.O (Bengaluru), Shantinagar S.O (Bengaluru)	13	13	13	13	13	13
I.M.Colony S.O, Somajiguda S.O	16	16	16	16	16	16
Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O	2	2	2	2	2	2
Kondapur B.O, Kothaguda S.O (K.V.Rangareddy)	20	20	20	20	20	20
LIC Division S.O, New Mla Quarters S.O	12	12	12	12	12	12
Lingampalli S.O	4	4	4	4	4	4
Manikonda S.O	3	3	3	3	3	3
Putlibowli S.O, State Bank Of India S.O	6	6	6	6	6	6
Raj Bhawan S.O (Bengaluru)	8	8	8	8	8	8
Sanjeev Reddy Nagar S.O, Vengal Rao Nagar S.O	4	4	4	4	4	4
Srinagar Colony S.O	8	8	8	8	8	8
Vijay Nagar Colony S.O (Bengaluru)	5	5	5	5	5	5

Let's find out how many unique categories can be curated from all the returned venues

In [23]:

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

```
There are 83 uniques categories.
```

Analyze Each Neighborhood

In [24]:

```
# one hot encoding
Blrcity_onehot = pd.get_dummies(Blrcity_venues[['Venue Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
Blrcity_onehot['Neighbourhood'] = Blrcity_venues['Neighbourhood']

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

Blrcity_onehot.head()
```

Out[24]:

	Neighbourhood	ATM	Andhra Restaurant	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Bakery	Beer Garden	Boutique	Breakfast Spot	Burger Joint	Bus Station	C
0	Gandhi Bhawan S.O (Bengaluru), Moazzampura	0	0	0	0	0	0	0	0	0	0	0	0

	Neighbourhood	ATM	Andhra Restaurant	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Bakery	Beer Garden	Boutique	Breakfast Spot	Burger Joint	Bus Station	C
1	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	0	0	0	0	0	0	0	0	0	0	0	0
2	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	0	0	0	0	0	0	0	0	0	0	0	0
3	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	0	0	0	0	0	0	0	0	0	0	0	0
4	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	0	0	0	0	0	0	0	0	0	0	0	0

And let's examine the new dataframe size.

In [25]:

```
Blrcity_onehot.shape
```

Out[25]:

```
(243, 84)
```

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

In [26]:

```
Blrcity_grouped = Blrcity_onehot.groupby('Neighbourhood').mean().reset_index()
Blrcity_grouped
```

Out[26]:

	Neighbourhood	ATM	Andhra Restaurant	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Bakery	Beer Garden	Boutique	Breakfast Spot	Burger Joint
0	A.Gs Office S.O, Anandnagar S.O (Bengaluru), B...	0.0	0.000000	0.000000	0.000000	0.000000	0.066667	0.000000	0.000000	0.00	0.000000
1	A.Gs. Staff Quarters S.O, Yousufguda S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
2	Appa Himayathsagar B.O, Dargah Hussain Shahwal...	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
3	Aziz Nagar B.O, Bhaskar Nagar S.O, C.B.I.T	0.0	0.000000	0.000000	0.062500	0.000000	0.000000	0.000000	0.000000	0.00	0.000000

	S.O...										
4	Neighbourhood	ATM	Andhra Restaurant	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Bakery	Beer Garden	Boutique	Breakfast Spot	Burger Joint
	Bharat Nagar										
5	Colony S.O, Erragadda S.O, Fathen...	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
6	CUC S.O	0.5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
7	Central Secretariat S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
8	Cyberabad S.O, Madhapur B.O	0.0	0.027778	0.000000	0.027778	0.000000	0.055556	0.000000	0.000000	0.00	0.000000
9	Dr.B R Ambedkar O.U S.O, Jubilee Hills S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
10	Film Nagar S.O	0.0	0.000000	0.000000	0.111111	0.111111	0.000000	0.111111	0.000000	0.00	0.000000
11	Gachibowli S.O, Manuu S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
12	Gagan Mahal S.O, Himayathnagar S.O, Narayangud...	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.045455	0.00	0.045455
13	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
14	Humayunnagar S.O, Murad Nagar S.O (Bengaluru),...	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
15	I.M.Colony S.O, Somajiguda S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
16	Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
17	Kondapur B.O, Kothaguda S.O (K.V.Rangareddy)	0.0	0.050000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.05	0.000000
18	LIC Division S.O, New Mla Quarters S.O	0.0	0.000000	0.083333	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
19	Lingampalli S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
20	Manikonda S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
21	Putlibowli S.O, State Bank Of India S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
22	Raj Bhawan S.O (Bengaluru)	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000
23	Sanjeev Reddy Nagar S.O, Vengal Rao Nagar S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.250000	0.000000	0.000000	0.25	0.000000
24	Srinagar Colony S.O	0.0	0.000000	0.000000	0.000000	0.000000	0.125000	0.000000	0.000000	0.00	0.000000

25	Vijay Nagar Colony S.O Neighbourhood (Bengaluru)	0.0 ATM	0.000000 Andhra Restaurant	Arts & Crafts Store	0.000000 Asian Restaurant	0.000000 Athletics & Sports	0.000000 Bakery	0.000000 Beer Garden	0.000000 Boutique	0.20 Breakfast Spot	0.000000 Burger Joint

Let's confirm the new size

In [28]:

```
Blrcity_grouped.shape
```

Out[28]:

```
(26, 84)
```

Let's print each neighborhood along with the top 5 most common venues

In [29]:

```
num_top_venues = 5

for hood in Blrcity_grouped['Neighbourhood']:
    print("----"+hood+"----")
    temp = Blrcity_grouped[Blrcity_grouped['Neighbourhood'] == hood].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(num_top_venues))
    print('\n')
```

```
----A.Gs Office S.O, Anandnagar S.O (Bengaluru), Bazarghat S.O (Bengaluru), Khairatabad H.O,
Parishram Bhawan S.O, Vidhan Sabha S.O (Bengaluru)----
                          venue  freq
0              Indian Restaurant  0.13

2          Bengalurui Restaurant  0.13
3Performing Arts Venue  0.07
4  Paper / Office Supplies Store  0.07


----A.Gs. Staff Quarters S.O,  Yousufguda S.O----
                       venue  freq
0       Fast Food Restaurant   0.5
1            Sandwich  Place   0.5
2                        ATM   0.0
3   Middle Eastern Restaurant   0.0
4      Performing Arts  Venue   0.0


----Appa Himayathsagar  B.O, Dargah Hussain Shahwali B.O, Golconda S.O, Blrer Shah Kote B.O,
Kakatiya Nagar S.O, Lunger House S.O, Nanakramguda B.O, Sakkubai Nagar S.O, Toli Chowki S.O----
                venue  freq
0      Women's Store  0.17

2      Historic Site  0.17
3Golf Course  0.17
4  Indian Restaurant  0.17


----Aziz Nagar B.O, Bhaskar Nagar S.O, C.B.I.T S.O, Gandipet B.O, Himayathnagar B.O, Janwada B.O,
Kokapet B.O, Mancherevula B.O, Narsingi B.O, Wattinagulapalli B.O----
         venue  freq
0        Hotel  0.19
1  Coffee  Shop  0.12
2 Hotel Pool 0.06
3Cafeteria  0.06
4Hotel Bar  0.06


----Banjara Hills S.O----
            venue  freq
```

```
0     Coffee Shop  0.44
1            Café  0.11
2  Sandwich Place  0.11
3Bakery  0.11
4   Deli / Bodega  0.11
```

```
----Bharat Nagar Colony S.O, Erragadda S.O, Fathenagar Colony S.O, Moosapet B.O, Sanath Nagar
Colony S.O, Sanathnagar I E S.O, Swarajyanagar S.O, Vivekananda Nagar Colony S.O----
                      venue  freq
0              Train Station  0.50
1Department Store  0.25
 2               Bus Station 0.25
 3                       ATM 0.00
4  Middle Eastern Restaurant  0.00
```

```
----CUC S.O----
                    venue freq
0                     ATM  0.5
1             Pizza Place  0.5
2                Pharmacy  0.0
3   Performing Arts Venue  0.0
4                    Park  0.0
```

```
----Central Secretariat S.O----
                  venue freq
 0  Chinese Restaurant  0.5
 1              Garden  0.5
 2                 ATM  0.0
 3   Mobile Phone Shop  0.0
 4            Pharmacy  0.0
```

```
----Cyberabad S.O, Madhapur B.O----
                venue freq
 0               Café 0.11
1        Coffee Shop  0.08
2  Indian Restaurant  0.08
3              Hotel  0.08
4Jewelry Store  0.06
```

```
----Dr.B R Ambedkar O.U  S.O, Jubilee Hills S.O----
                      venue  freq
0              Women's Store  0.25
 1                       Park 0.25
 2                  Food Court 0.25
 3                       Lake 0.25
4  Middle Eastern Restaurant  0.00
```

```
----Film Nagar S.O----
                venue  freq
0           Irish Pub  0.11
1Gift Shop  0.11
2  Italian Restaurant  0.11
3Asian Restaurant  0.11
4  Athletics & Sports  0.11
```

```
----Gachibowli S.O, Manuu S.O----
                venue  freq
0      Sandwich Place  0.25
1Coffee Shop  0.25
3  College Rec Center  0.12
4                 Gym  0.12
```

```
----Gagan Mahal S.O, Himayathnagar S.O, Narayanguda S.O, Ramakrishna Mutt S.O----
                           venue  freq
0              Indian Restaurant  0.14
 1                     Restaurant 0.09
 2                          Café 0.09
```

```
3  Vegetarian / Vegan Restaurant  0.09
4Fast Food Restaurant  0.09


----Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O, Seetharampet S.O, State Bank Of Bengaluru S.O-
---
                 venue  freq
0           Shoe Store  0.14
1Train Station  0.14
2             Platform  0.14
4  Fried Chicken Joint  0.14


----Humayunnagar S.O, Murad Nagar S.O (Bengaluru), Shantinagar S.O (Bengaluru)----
                  venue  freq
0  Fast Food Restaurant  0.31
1Indian Restaurant  0.15
2Farmers Market  0.08
3Pizza Place  0.08
4             Juice Bar  0.08


----I.M.Colony  S.O, Somajiguda S.O----
                 venue freq
0   Indian Restaurant 0.12
1               Hotel 0.12
2        Pizza  Place 0.12
3     Sandwich  Place 0.12
4         Coffee Shop 0.12


----Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O----
                      venue   freq
0             Women's Store    0.5
1    South Indian Restaurant   0.5
2  Mediterranean Restaurant   0.0
3                   Pharmacy   0.0
4     Performing Arts Venue    0.0


----Kondapur B.O, Kothaguda S.O (K.V.Rangareddy)----
                 venue freq
0  Indian Restaurant  0.25
1Grocery Store  0.15
2   Department Store  0.10
3Breakfast Spot  0.05
4               Café  0.05


 ----LIC Division S.O, New Mla Quarters S.O----
                        venue  freq
0   Chinese Restaurant  0.17
1Restaurant  0.17
2Indian Restaurant  0.08
3  Arts & Crafts Store  0.08
4                 Park 0.08


----Lingampalli S.O----
                 venue freq
0          Snack Place  0.25
2             Pharmacy  0.25
3Indian Restaurant  0.25
4  Indie Movie Theater  0.00


----Manikonda S.O----
                venue  freq
0           Pharmacy  0.33
1Coffee Shop  0.33
2Food Court  0.33
 3                ATM 0.00
 4  Mobile Phone Shop 0.00
```

```
----Putlibowli S.O, State Bank Of India S.O----
                venue  freq
0          Chaat Place  0.17
1  Indie Movie Theater  0.17
2Restaurant  0.17
3Coffee Shop  0.17
4Bus Station  0.17


----Raj Bhawan S.O (Bengaluru)----
               venue  freq
0  Indian Restaurant  0.12
1Scenic Lookout  0.12
2Pizza Place  0.12
3              Plaza  0.12
4Food Court  0.12


----Sanjeev Reddy Nagar S.O, Vengal Rao Nagar S.O----
              venue  freq
0  Department Store  0.25
1Pharmacy  0.25
2            Bakery  0.25
3Breakfast Spot  0.25

4               ATM  0.00


----Srinagar Colony S.O----
                    venue  freq
0  Rajasthani Restaurant  0.12
1             Hookah Bar  0.12
2Dessert Shop  0.12
3Department Store  0.12
4            Coffee Shop  0.12


----Vijay Nagar Colony S.O (Bengaluru)----
               venue freq
 0              Café  0.2
 1  Department Store  0.2
 2    Breakfast Spot  0.2
 3      Soccer Field  0.2
4     Shopping Mall   0.2
```

Let's put that into a pandas dataframe

First, let's write a function to sort the venues in descending order.

In [30]:

```
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

Now let's create the new dataframe and display the top 10 venues for each neighborhood.

In [31]:

```
num_top_venues = 10

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

# create columns according to number of top
venues columns = ['Neighbourhood']
for ind in
    np.arange(num_top_venues): try:
```

```
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighbourhoods_venues_sorted = pd.DataFrame(columns=columns)
neighbourhoods_venues_sorted['Neighbourhood'] = Blrcity_grouped['Neighbourhood']

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

neighbourhoods_venues_sorted
```

Out[31]:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	A.Gs Office S.O, Anandnagar S.O (Bengaluru), B...	Bengalurui Restaurant	Hotel	Indian Restaurant	Ice Cream Shop	Café	Paper / Office Supplies Store	Middle Eastern Restaurant	Performing Arts Venue
1	A.Gs. Staff Quarters S.O, Yousufguda S.O	Fast Food Restaurant	Sandwich Place	Women's Store	Furniture / Home Store	Donut Shop	Electronics Store	Farmers Market	Food
2	Appa Himayathsagar B.O, Dargah Hussain Shahwal...	Women's Store	Café	Bengalurui Restaurant	Indian Restaurant	Historic Site	Golf Course	Electronics Store	Farmers Market
3	Aziz Nagar B.O, Bhaskar Nagar S.O, C.B.I.T S.O...	Hotel	Coffee Shop	Hotel Pool	Indian Restaurant	Restaurant	Sandwich Place	Cafeteria	Lounge
4	Banjara Hills S.O	Coffee Shop	Deli / Bodega	Café	Hookah Bar	Bakery	Sandwich Place	Electronics Store	Farmers Market
5	Bharat Nagar Colony S.O, Erragadda S.O, Fathen...	Train Station	Bus Station	Department Store	Women's Store	Donut Shop	Electronics Store	Farmers Market	Fast Food Restaurant
6	CUC S.O	ATM	Pizza Place	Hot Dog Joint	Fried Chicken Joint	Dessert Shop	Donut Shop	Electronics Store	Farmers Market
7	Central Secretariat S.O	Garden	Chinese Restaurant	Furniture / Home Store	Donut Shop	Electronics Store	Farmers Market	Fast Food Restaurant	Food
8	Cyberabad S.O, Madhapur B.O	Café	Coffee Shop	Indian Restaurant	Hotel	Bakery	Mediterranean Restaurant	Jewelry Store	Pizza Place
9	Dr.B R Ambedkar O.U S.O, Jubilee Hills S.O	Women's Store	Food Court	Park	Lake	Gym	Historic Site	Donut Shop	Hookah Bar
10	Film Nagar S.O	Italian Restaurant	Asian Restaurant	Athletics & Sports	Beer Garden	Coffee Shop	Irish Pub	Mediterranean Restaurant	Gift Shop
11	Gachibowli S.O, Manuu S.O	Coffee Shop	Sandwich Place	Cafeteria	College Rec Center	Gym	Café	Women's Store	Farmers Market
12	Gagan Mahal S.O, Himayathnagar S.O,	Indian Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Café	Restaurant	Hookah Bar	Department Store	Coffee Shop

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
	Narayangud... Gandhi Bhawan S.O (Bengaluru),		Shoe	Train	Electronics		Mobile Phone	Fried Chicken	Donut
13	Moazzampura S.O...	Hotel	Store	Station	Store	Platform	Shop	Joint	Shop
14	Humayunnagar S.O, Murad Nagar S.O (Bengaluru),...	Fast Food Restaurant	Indian Restaurant	Hookah Bar	Juice Bar	Pizza Place	Farmers Market	Restaurant	Department Store
15	I.M.Colony S.O, Somajiguda S.O	Sandwich Place	Hotel	Pizza Place	Coffee Shop	Indian Restaurant	Convenience Store	Donut Shop	Fast Food Restaurant
16	Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O	Women's Store	South Indian Restaurant	Garden	Donut Shop	Electronics Store	Farmers Market	Fast Food Restaurant	Food
17	Kondapur B.O, Kothaguda S.O (K.V.Rangareddy)	Indian Restaurant	Grocery Store	Department Store	Ice Cream Shop	Pizza Place	Café	Breakfast Spot	Smoke Shop
18	LIC Division S.O, New Mla Quarters S.O	Chinese Restaurant	Restaurant	Park	Sandwich Place	Hotel	Arts & Crafts Store	Ice Cream Shop	Indian Restaurant
19	Lingampalli S.O	Gym	Indian Restaurant	Snack Place	Pharmacy	Women's Store	Fried Chicken Joint	Dessert Shop	Donut Shop
20	Manikonda S.O	Coffee Shop	Food Court	Pharmacy	Women's Store	Furniture / Home Store	Donut Shop	Electronics Store	Farmers Market
21	Putlibowli S.O, State Bank Of India S.O	Bus Station	Multiplex	Coffee Shop	Indie Movie Theater	Restaurant	Chaat Place	Furniture / Home Store	Farmers Market
22	Raj Bhawan S.O (Bengaluru)	Food Court	Café	Pizza Place	Indian Restaurant	Plaza	Fast Food Restaurant	Restaurant	Scenic Lookout
23	Sanjeev Reddy Nagar S.O, Vengal Rao Nagar S.O	Department Store	Bakery	Breakfast Spot	Pharmacy	Furniture / Home Store	Donut Shop	Electronics Store	Farmers Market
24	Srinagar Colony S.O	Department Store	Hookah Bar	Ice Cream Shop	Bakery	Coffee Shop	Rajasthani Restaurant	Dessert Shop	Food
25	Vijay Nagar Colony S.O (Bengaluru)	Shopping Mall	Soccer Field	Café	Breakfast Spot	Department Store	Furniture / Home Store	Donut Shop	Electronics Store

Cluster Neighborhoods

Run k-means to cluster the neighborhood into 5 clusters.

In [35]:

```
# set number of clusters
kclusters = 5

Blrcity_grouped_clustering = Blrcity_grouped.drop('Neighbourhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Blrcity_grouped_clustering)

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

Out[35]:

```
array([3, 0, 3, 3, 1, 3, 3, 4, 3, 3], dtype=int32)
```

Let's create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood

In [37]:

```
Blrcity_merged = Blrcity_Blr_city

# add clustering labels
Blrcity_merged['Cluster Labels'] = kmeans.labels_[1]

Blrcity_merged = Blrcity_merged.join(neighbourhoods_venues_sorted.set_index('Neighbourhood'),
on=' Neighbourhood')

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

Out[37]:

	pincode	Borough	Neighbourhood	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Com V
0	500001	Bengaluru City	Gandhi Bhawan S.O (Bengaluru), Moazzampura S.O...	17.390585	78.470388	0	Hotel	Shoe Store	Train Station	Electronics Store	Platfo
1	500004	Bengaluru City	A.Gs Office S.O, Anandnagar S.O (Bengaluru), B...	17.403781	78.462525	0	Bengalurui Restaurant	Hotel	Indian Restaurant	Ice Cream Shop	Café
2	500006	Bengaluru City	Karwan Sahu S.O, Kulsumpura S.O, Mangalhat S.O	17.371224	78.454180	0	Women's Store	South Indian Restaurant	Garden	Donut Shop	Electr Store
3	500008	Bengaluru City	Appa Himayathsagar B.O, Dargah Hussain Shahwal...	17.396335	78.406792	0	Women's Store	Café	Bengalurui Restaurant	Indian Restaurant	Histo Site
4	500018	Bengaluru City	Bharat Nagar Colony S.O, Erragadda S.O, Fathen...	17.457435	78.445780	0	Train Station	Bus Station	Department Store	Women's Store	Donu Shop

Finally, let's visualize the resulting clusters

In [38]:

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

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

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(Blrcity_merged['latitude'], Blrcity_merged['longitude'],
Blrcity_ merged['Neighbourhood'], Blrcity_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
```

```
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

Out[38]:

Examine Clusters

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

Cluster 1

In [43]:

```
Blrcity_merged.loc[Blrcity_merged['Cluster Labels'] == 0, Blrcity_merged.columns[[1] + list(range(5
, Blrcity_merged.shape[1]))]].head( )
```

Out[43]:

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Mo Comm Ven
0	Bengaluru City	0	Hotel	Shoe Store	Train Station	Electronics Store	Platform	Mobile Phone Shop	Fried Chicken Joint	Donut Shop	Farmers Market
1	Bengaluru City	0	Bengalurui Restaurant	Hotel	Indian Restaurant	Ice Cream Shop	Café	Paper / Office Supplies	Middle Eastern	Performing Arts	South Indian

								Store	Restaurant	Venue	Restaur
2	Bengaluru	0	Women's	South Indian	Garden	Donut	Electronics	Farmers	Fast Food	Food	Food

	City Borough	Cluster Labels	Store 1st Most Common Venue	Restauran2ndMost Common Café Venue	3rd Most Common Bengalurui Venue	Shop 4th Most Common Indian Venue	Store 5th Most Common Historic Venue	Market 6th Most Common Golf Venue	Restaurant 7th Most Common Electronics Venue	8th Most Common Farmers Venue	Court 9th Mo Comm Fast Foo Ven
3	Bengaluru City	0	Women's Store		Restaurant	Restaurant	Site	Course	Store	Market	Restaur
4	Bengaluru City	0	Train Station	Bus Station	Department Store	Women's Store	Donut Shop	Electronics Store	Farmers Market	Fast Food Restaurant	Food

Conclusion

In majority of part of Bengaluru City "Women's Store" & "Food Court" is the most popular type of service.