



Peer-graded Assignment: Capstone Project - The Battle of Neighborhoods (Week 2)

Sharookh Ali

REPORT

Table of Content

1. Introduction
2. Data Extraction
3. Methodology
4. Observation & Discussion
5. Conclusion

1. Introduction: A description of the problem and a discussion of the background.

1. In Module 3, we explored New York City and the city of Toronto and segmented and clustered their neighborhoods. Both cities are very diverse and are the financial capitals of their respective countries.
2. One interesting idea would be to compare the neighborhoods of the two cities and determine how similar or dissimilar they are.
3. In this assignment we will compare the neighborhoods of the two Indian cities namely Hyderabad and Mumbai
4. Is Hyderabad more like Mumbai in terms of Shops, Restaurants , Cafes etc?
5. In Hyderabad or Mumbai, if someone is looking to open a restaurant, where would we recommend that they open it? Similarly, if a contractor is trying to start their own business, where would you recommend that they setup their office? and other important points will be discussed
6. If a person wants to shift from Hyderabad to Mumbai or vice versa, then how likely he will be adjusted to these cities will be discussed.

2. Data Extraction: Extracting the data of HYDERABAD & MUMBAI

The data extracted may include:

1. List of Cafes and foodstalls of Hyderabad with their latitude and longitude
2. List of Hotels, Book stores, Restaurant, Bakeries, and all relevant locations
3. All data is limited within 10km radius

3. Methodology

1. First extracting the geographical coordinates of Hyderabad & Mumbai through geolocator is to be done.
2. Then FOURSQUARE credentials are specified for the analysis of maps.
3. Then the data is gathered with FOURSQUARE API url with a limit of 10 kms radius and maximum count of 200 in both Mumbai and Hyderabad
4. After the data is gathered, the venues are sorted according to their categories and printed by including both latitude and longitude.
5. Then the maps of Mumbai and Hyderabad are represented with the help of Folium.
6. Then the value count with respect to each category is done and printed for both Hyderabad and Mumbai individually
7. The corresponding value counts are represented as bar graphs and the observations, intuitions are further discussed in "OBSERVATION & DISCUSSION" section.

```
[1] import json
import requests
import matplotlib.pyplot as plt
from pandas.io.json import json_normalize
import numpy as np
import time
import pandas as pd
pd.set_option('display.max_columns', None)
```

```

pd.set_option('display.max_rows', None)
from geopy.geocoders import Nominatim
import folium
print('Libraries imported.')

```

Libraries imported.

▼ Extracting the geographical coordinates of Hyderabad through geolocator

```

✓ [2] address = 'Hyderabad'
geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geographical coordinates of ', address, 'are {}, {}'.format(latitude, longitude))
neighborhood_latitude=latitude
neighborhood_longitude=longitude

/usr/local/lib/python3.7/dist-packages/geopy/geocoders/osm.py:143: UserWarning: Using Nominatim with the default "geopy/1.17.0" `user_agent` is str
UserWarning
The geographical coordinates of Hyderabad are 17.360589, 78.4740613.

```

```

✓ [3] CLIENT_ID = '4PDNWKYQ3G5K3PEKC00M02CUQWZGZJRTW1BTP5RE1M42VX2I'
CLIENT_SECRET = '10L5BDNLNT0YCQGDKZC5GBUDK5PVMREH4G4PIUAZPDLT0DW'
VERSION = '20180604'

```

```

✓ [4] radius = 10000
limit = 200
# importing the data with URL
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{},radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    neighborhood_latitude,
    neighborhood_longitude,
    radius,
    limit)
results = requests.get(url).json()

```

▼ Determining the categories of each location in data

```

✓ [5] 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']

```

▼ Showing the list of venues

```

✓ [6] venues = results['response'][0]['groups'][0]['items']
borough = results['response'][0]['groups'][0]['items']
HYD_venues = json_normalize(venues) # flatten JSON
HYD_borough= json_normalize(borough)

# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat', 'venue.location.lng']
HYD_venues = HYD_venues.loc[:, filtered_columns]
# filter the category for each row
HYD_venues['venue.categories'] = HYD_venues.apply(get_category_type, axis=1)
# clean columns
HYD_venues.columns = [col.split(".")[1] for col in HYD_venues.columns]
HYD_venues.head(100)

68 Taj Mahal Hotel Hotel 17.391942 78.476915
69 Starbucks Coffee Shop 17.416999 78.416344
70 Hyderabad Golf Club Golf Course 17.393581 78.404188
71 Crossword Bookstore 17.419411 78.448145
72 Golconda Fort Historic Site 17.383858 78.403270
73 Pista House Asian Restaurant 17.396746 78.424301

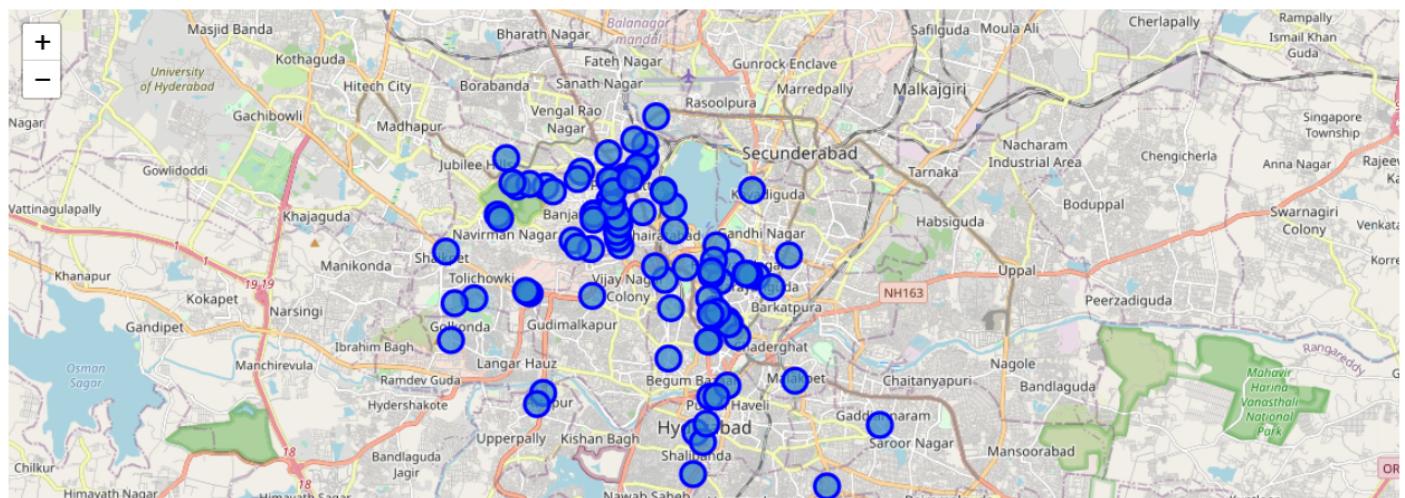
```

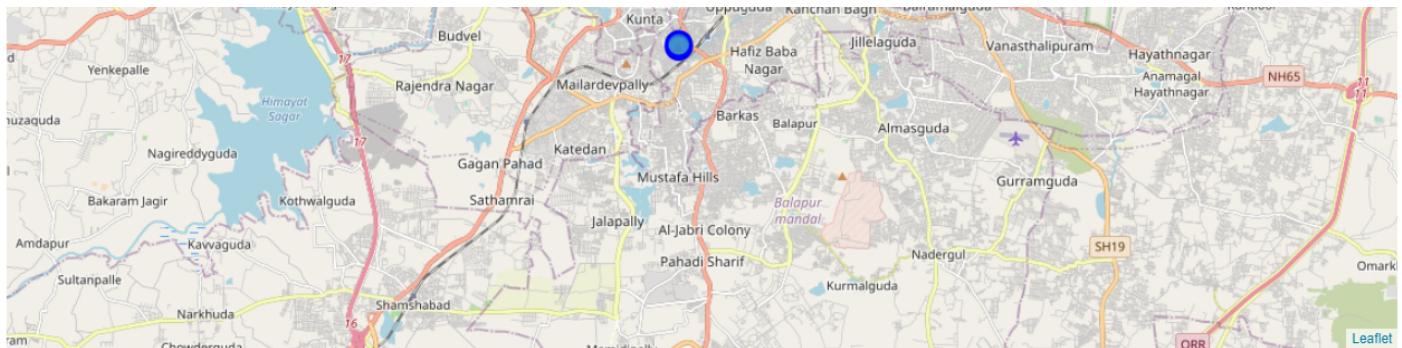
74	Ofen	Bakery	17.415806	78.443039
75	Hollywood	Shoe Store	17.389916	78.476241
76	Tank Bund	Lake	17.409019	78.477395
77	Blue Fox	Indian Restaurant	17.401401	78.485653
78	Mings Court	Chinese Restaurant	17.404872	78.476563
79	Kismet	Nightclub	17.423500	78.462598
80	Q-Mart	Department Store	17.426681	78.438616
81	Chef Inam's Steak House	Steakhouse	17.408716	78.438714
82	Karachi Bakery	Bakery	17.407400	78.401911
83	Vac's	Bakery	17.432224	78.418737
84	KBR Park	Park	17.425369	78.425063
85	Hotel Megacity	Hotel Bar	17.401133	78.475990
86	Rain Lounge	Nightclub	17.425615	78.420337
87	Himalayas	Bookstore	17.426359	78.451211
88	Bombay Bakery	Bakery	17.394869	78.475475
89	SRIKANYA COMFORT	Andhra Restaurant	17.429810	78.455347
90	Kakatiya Deluxe Mess	Diner	17.433435	78.447090
91	Southern Mirchi	South Indian Restaurant	17.426419	78.448560
92	Santosh Dabha	Diner	17.391209	78.477575
93	Chutney's	Vegetarian / Vegan Restaurant	17.426219	78.447569
94	Chinese Pavilion	Asian Restaurant	17.423456	78.448695
95	Abid's	Neighborhood	17.390543	78.475724
96	Mohini Restaurant	Indian Restaurant	17.401926	78.476153
97	PVR Cinemas	Multiplex	17.426516	78.453261
98	Vivanta by Taj	Hotel	17.443498	78.460556
99	Little Italy	Italian Restaurant	17.416178	78.416876

[6]

Creating Hyderabad map with Folium

```
[7] MAP_HYD = folium.Map(location=[latitude, longitude], zoom_start=12)
# add markers to map
for lat, lng, label in zip(HYD_venues['lat'], HYD_venues['lng'], HYD_venues['name']):
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=10,
        popup=label,
        color='blue',
        fill_color="#3186cc",
        fill_opacity=0.7,
    ).add_to(MAP_HYD)
MAP_HYD
```





```
[8] from folium import plugins
    import seaborn as sns
    import matplotlib.cm as cm
    import matplotlib.colors as colors
    from sklearn.cluster import KMeans
    print('Additional libraries imported.')
```

Additional libraries imported.

Extracting the data of Mumbai

In this case also the data extracted may include:

1. List of Cafes and foodstalls of Mumbai with their latitude and longitude
2. List of Hotels, Book stores, Restaurant, Bakaries, and all relevant locations
3. All data is limited with in 10km radius

```
[9] address = 'Mumbai'
geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geographical coordinates of ',address,'are {}, {}'.format(latitude, longitude))
neighborhood_latitude=latitude
neighborhood_longitude=longitude

/usr/local/lib/python3.7/dist-packages/geopy/geocoders/osm.py:143: UserWarning: Using Nominatim with the default "geopy/1.17.0" `user_agent` is str
UserWarning
The geographical coordinates of Mumbai are 19.0759899, 72.8773928.
```

```
[10] radius = 2000
limit = 200
# importing the data with URL
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    neighborhood_latitude,
    neighborhood_longitude,
    radius,
    limit)
results = requests.get(url).json()
```

Determining the categories of each location in data & Showing the list of venues

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

MB_venues = json_normalize(venues) # flatten JSON

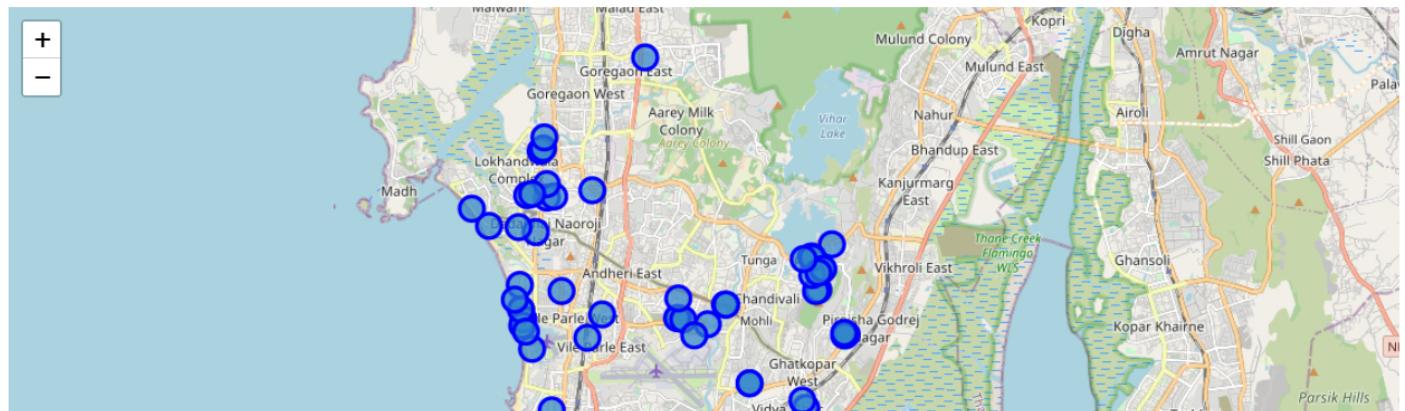
# filter columns
filtered_columns = ['venue.name', 'venue.categories', 'venue.location.lat', 'venue.location.lng']
MB_venues = MB_venues.loc[:, filtered_columns]
# filter the category for each row
MB_venues['venue.categories'] = MB_venues.apply(get_category_type, axis=1)
# clean columns
MB_venues.columns = [col.split(".")[1] for col in MB_venues.columns]
MB_venues.head(100)

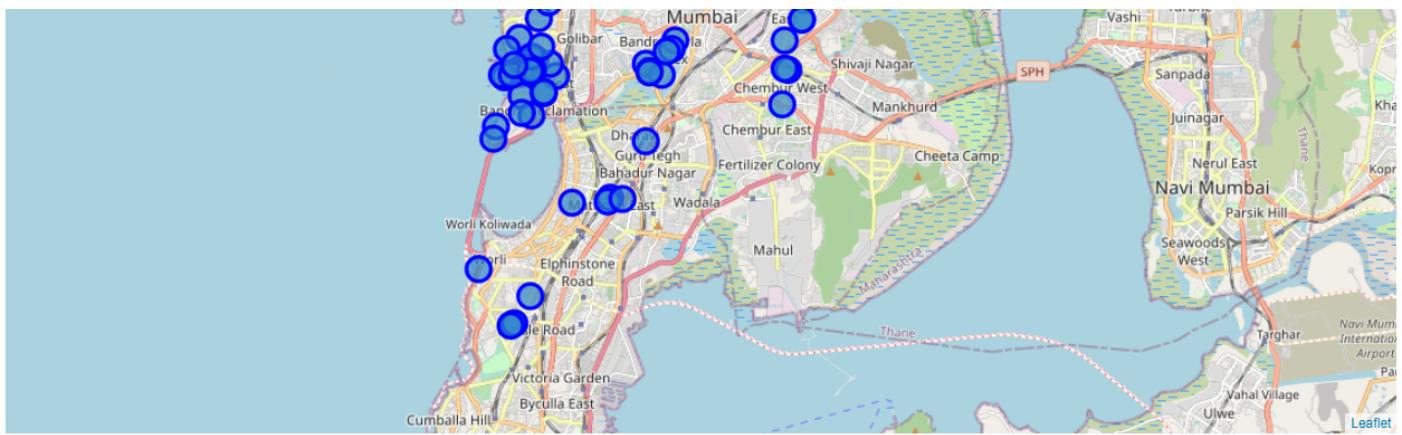
68          Urban Tadka      Indian Restaurant  19.099458  72.916215
69        Cafe Mangii      Italian Restaurant  19.120260  72.907052
--
```

70	Ram Krishna Hotel	Indian Restaurant	19.098702	72.844528
71	Elco Chat Corner	Indian Restaurant	19.055762	72.833216
72	Mia Cucina	Italian Restaurant	19.120170	72.907217
73	Starbucks	Coffee Shop	19.099942	72.916407
74	Kitchen Garden by Suzette	Salad Place	19.061545	72.829262
75	Mirchi And Mime	North Indian Restaurant	19.120020	72.907448
76	Elco Veg Restaurant	Snack Place	19.055791	72.833253
77	Hiranandani Circle	Plaza	19.117042	72.910317
78	Jai Hind Lunch Home	Seafood Restaurant	19.002183	72.829512
79	Theobroma	Bakery	19.115923	72.908921
80	Tanjore Tiffin Room	South Indian Restaurant	19.128438	72.817150
81	Doolally Taproom	Brewery	19.135917	72.833094
82	Powai Lake Promenade	Scenic Lookout	19.119665	72.904994
83	Worli Sea Face	Scenic Lookout	19.009216	72.815022
84	The Bagel Shop	Bagel Shop	19.063140	72.824902
85	Indigo Delicatessen	Mediterranean Restaurant	19.136450	72.827565
86	Hard Rock Cafe Andheri	American Restaurant	19.135995	72.835335
87	Jam Jar Diner	Diner	19.132830	72.812268
88	Jaffer Bhai's Delhi Darbar	Mughlai Restaurant	19.137714	72.845909
89	The St. Regis Mumbai	Hotel	18.994570	72.824404
90	The Westin Mumbai Garden City	Hotel	19.172654	72.860518
91	Woodside	Bar	19.148197	72.831419
92	The Little Door	Pub	19.139265	72.833180
93	Bistro 1 Cafe	Café	19.136656	72.829136
94	High Street Phoenix	Shopping Mall	18.994967	72.825032
95	The Sahib Room & Kipling Bar	Indian Restaurant	18.994333	72.823850
96	Pishu's Juice Center	Juice Bar	19.128122	72.825384
97	Pratap's The Dhaba	Indian Restaurant	19.148876	72.832097
98	Novotel Juhu	Hotel	19.108725	72.824338
99	LSD - Love Sugar Dough	Cupcake Shop	19.151546	72.832341

Creating Mumbai map with Folium

```
[12]: MAP_MB = folium.Map(location=[latitude, longitude], zoom_start=12)
# add markers to map
for lat, lng, label in zip(MB_venues['lat'], MB_venues['lng'], MB_venues['name']):
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=10,
        popup=label,
        color='blue',
        fill_color="#3186cc",
        fill_opacity=0.7,
    ).add_to(MAP_MB)
MAP_MB
```





Value count of HYDERABAD

```
[13]: HYD_count = HYD_venues['categories'].value_counts()
MB_count = MB_venues['categories'].value_counts()
print(">>>>HYDERABAD<<<<")
print(HYD_count)
```

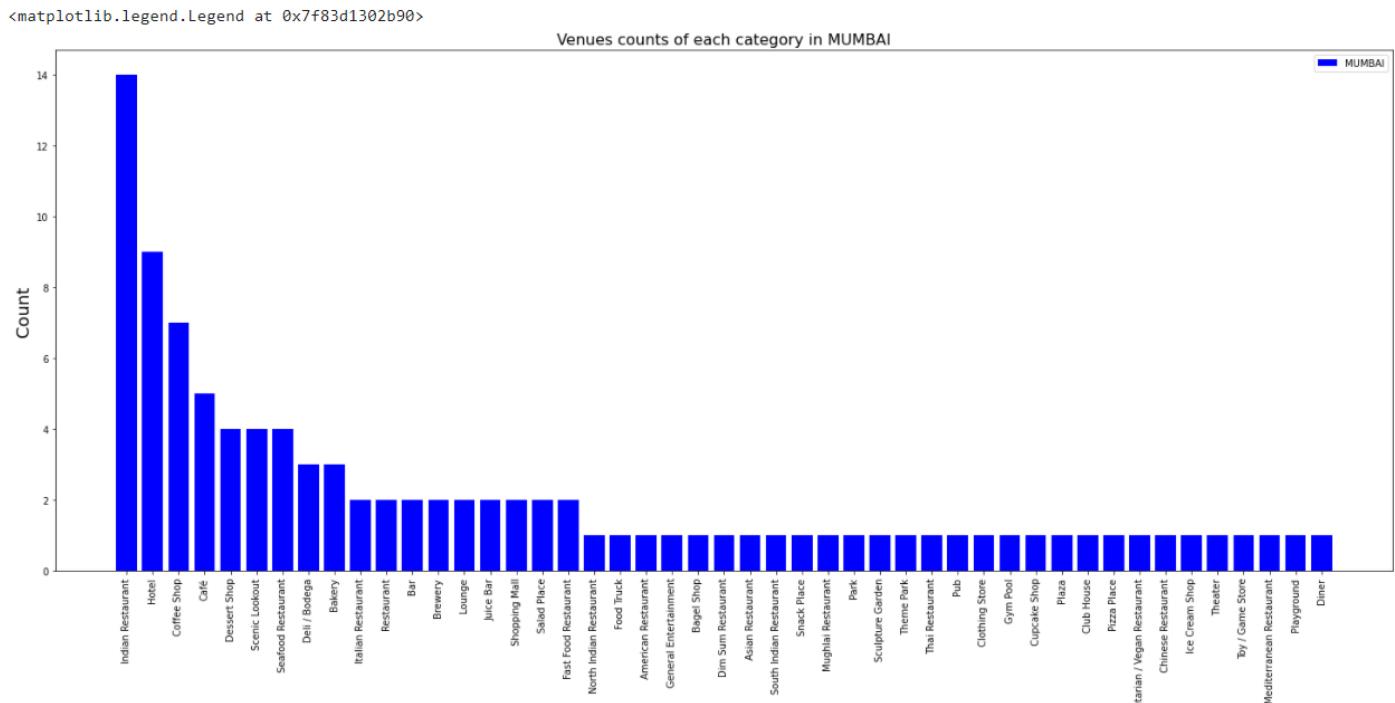
```
>>>>HYDERABAD<<<<
Indian Restaurant      12
Bakery                 7
Hotel                  6
Diner                  3
Ice Cream Shop          3
Multiplex                3
Café                   3
South Indian Restaurant  3
Asian Restaurant          3
Juice Bar                3
Shopping Mall            2
Vegetarian / Vegan Restaurant 2
Historic Site            2
Coffee Shop              2
Hotel Bar                2
BBQ Joint                2
Snack Place              2
History Museum            2
Chaat Place              2
Chinese Restaurant          2
Restaurant                2
Nightclub                 2
Bookstore                 2
Fast Food Restaurant       2
Breakfast Spot             1
Pub                      1
Dessert Shop              1
Hookah Bar                1
Shoe Store                 1
Italian Restaurant          1
Resort                     1
Scenic Lookout             1
Deli / Bodega              1
Performing Arts Venue       1
American Restaurant         1
Lounge                     1
Steakhouse                 1
Department Store            1
Burger Joint                1
Lake                      1
Golf Course                 1
Movie Theater                1
Food Truck                  1
Park                      1
Food Court                  1
Andhra Restaurant             1
Neighborhood                 1
Monument / Landmark           1
Bistro                      1
Hyderabadi Restaurant        1
Name: categories, dtype: int64
```

Value count of MUMBAI

```
[14]: print("")
print(">>>>MUMBAI<<<<")
print(MB_count)
```

```
>>>>MUMBAI<<<<
Indian Restaurant          14
Hotel                      9
Coffee Shop                 7
Café                       5
Dessert Shop                4
Scenic Lookout               4
Seafood Restaurant            4
Deli / Bodega                 3
Bakery                      3
Italian Restaurant             2
Restaurant                   2
Bar                         2
Brewery                     2
Lounge                      2
Juice Bar                    2
Shopping Mall                  2
Salad Place                  2
Fast Food Restaurant            2
North Indian Restaurant           1
Food Truck                   1
American Restaurant              1
General Entertainment            1
Bagel Shop                   1
Dim Sum Restaurant              1
Asian Restaurant                 1
South Indian Restaurant           1
Snack Place                   1
Mughlai Restaurant                1
Park                         1
Sculpture Garden                1
Theme Park                     1
Thai Restaurant                  1
Pub                          1
Clothing Store                  1
Gym Pool                      1
Cupcake Shop                  1
Plaza                        1
Club House                     1
Pizza Place                   1
Vegetarian / Vegan Restaurant           1
Chinese Restaurant                 1
Ice Cream Shop                  1
Theater                      1
Toy / Game Store                  1
Mediterranean Restaurant           1
Playground                     1
Diner                        1
Name: categories, dtype: int64
```

```
[15] MB_count = MB_venues['categories'].value_counts()
plt.figure(figsize = (25, 10))
plt.xticks(rotation = 90)
plt.xlabel("Category of place", fontsize = 18)
plt.ylabel("Count", fontsize = 18)
plt.title("Venues counts of each category in MUMBAI", fontsize = 16)
plt.bar(MB_count.index, MB_count.values,color='blue')
plt.legend([ "MUMBAI"])
```



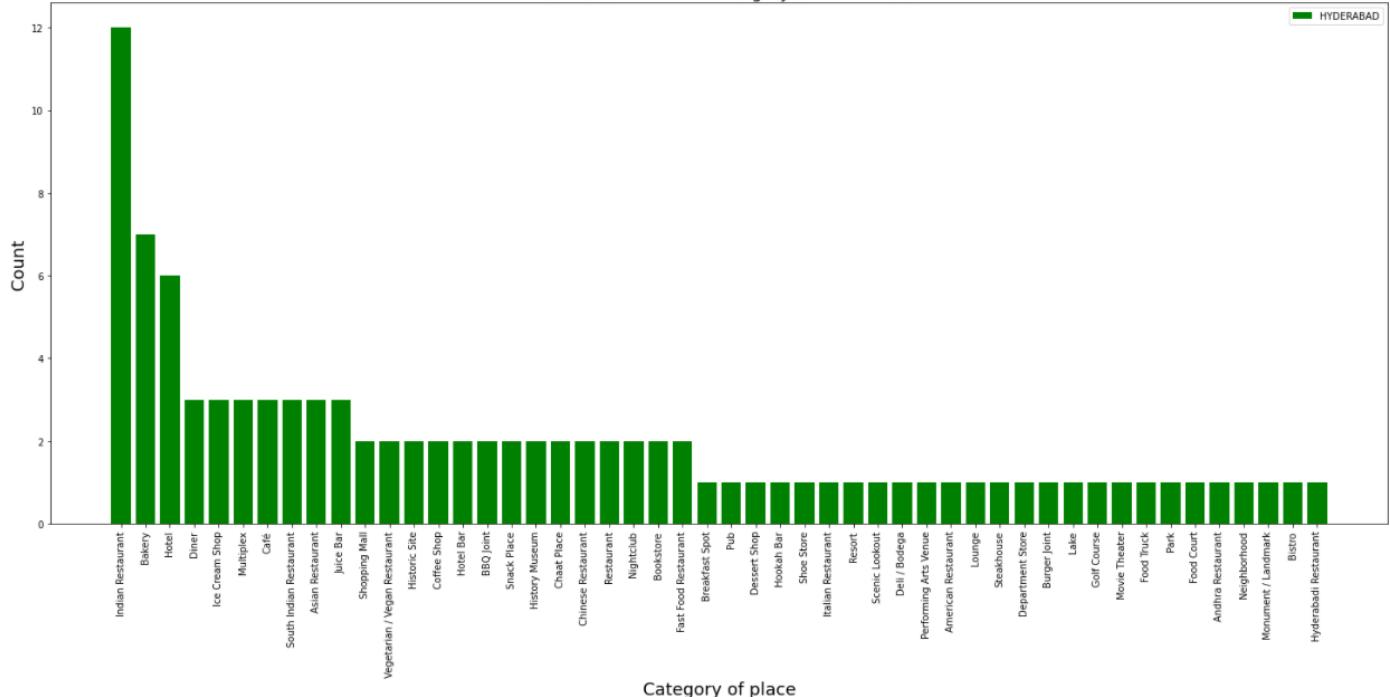
```

1s [16] plt.figure(figsize = (25, 10))
    plt.xticks(rotation = 90)
    plt.xlabel("Category of place", fontsize = 18)
    plt.ylabel("Count", fontsize = 18)
    plt.title("Venues counts of each category in HYDERABAD", fontsize = 18)
    plt.bar(HYD_count.index, HYD_count.values, color='green')
    plt.legend(["HYDERABAD"])

```

<matplotlib.legend.Legend at 0x7f83d0d0d390>

Venues counts of each category in HYDERABAD



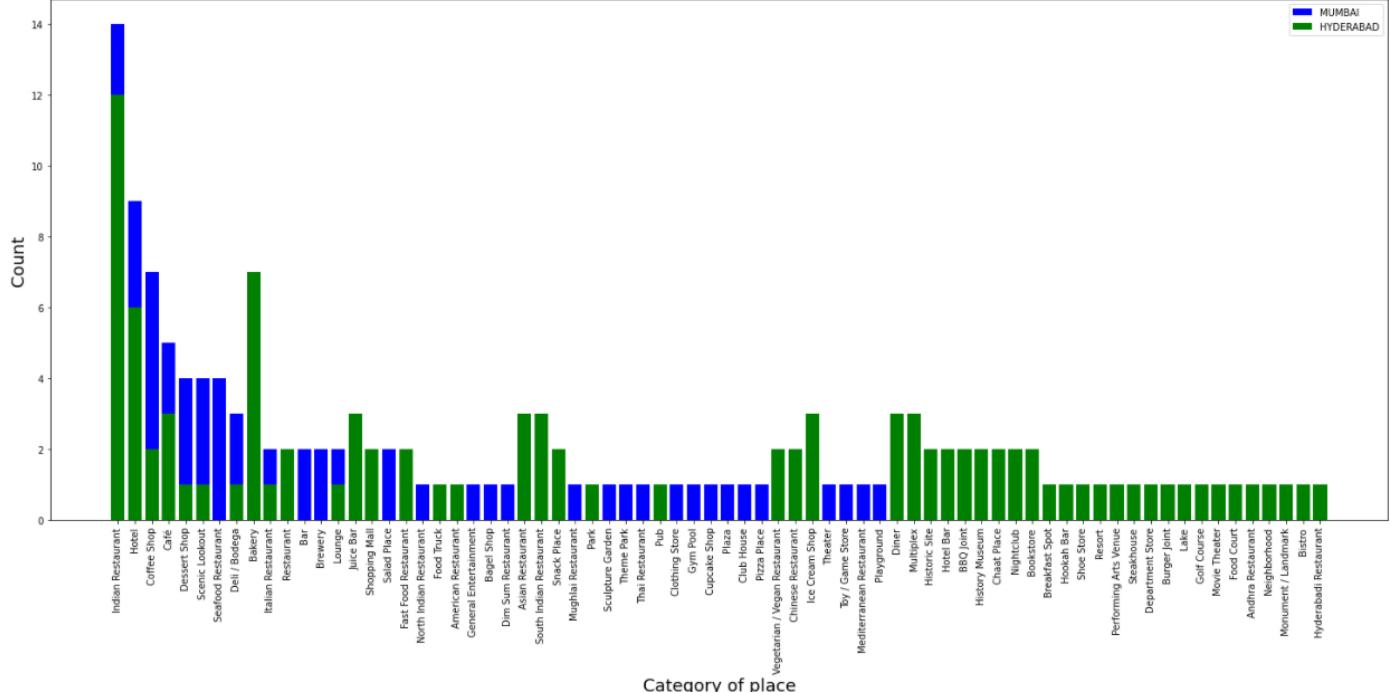
```

2s [17] plt.figure(figsize = (25, 10))
    plt.xticks(rotation = 90)
    plt.xlabel("Category of place", fontsize = 18)
    plt.ylabel("Count", fontsize = 18)
    plt.title("Comparative plot of venues counts of each category", fontsize = 16)
    plt.bar(MB_count.index, MB_count.values,color='blue')
    plt.bar(HYD_count.index, HYD_count.values, color='green')
    plt.legend(["MUMBAI","HYDERABAD"])

```

<matplotlib.legend.Legend at 0x7f83d0ca6fd0>

Comparative plot of venues counts of each category



▼ 4.Observations & Discussion

1. From above bar graph, we can observe that both HYDERABAD and MUMBAI are similar in terms of Shops, Restaurants , Cafes etc.
2. However, the Mumbai is more versatile in terms of venues compare to Hyderabad.
3. In Mumbai,when we observe the Folium map created with the help of Foursquare credentials, more venues are clustered in "Bandra West & Juhu" area and less venues can be spotted in " Vidya Vihar " area.
4. In Hyderabad,up on similar observation, more venues are clustered in "Abids, Ameerpet & Panjagutta" area and less venues can be spotted in " Golkonda " area.
5. Mumbai has significantly more number of Indian restaurants, Hotels, Coffe shops, lounges.
6. On other hand, Hyderabad has more number of Bakeries tha Mumbai
7. In terms of Shopping malls, Chinese restaurants, Asian restaurants, Ice cream shops, Juice bar both Mumbai and Hyderabad are more similar.
8. Rest miscellaneous venues are found to be single or two for both Mumbai and Hyderabad.

▼ 5.Conclusion

1. From above observations we can conclude the following things:
2. In Mumbai, if some one wants to open a restaurant, one can prefer "Bandra West or Juhu" area similarly in Hyderabad one can prefer "Abids or Ameerpet" area.
3. Opening Indian restaurant or Hotel in Hyderabad will be comparatively better than Mumbai because Mumbai already has a significantly more number of this category venues than Hyderabad.
4. One can always prefer starting up a missellenius type of venues as they were low and easy entry to the market in both Mumbai and Hyderabad
5. On overall, if a person wants to shift from Hyderabad to Mumbai or vice versa , there wouldn't be any problems as both cities are more similar in terms of venues according to Foursquare API data.

▼ Ending Remarks

Coursera has given a very good opportunity to learn Data science. Thanks to Coursera, IBM and Peers.

Happy learning!

✓ [17]
0s

✓ [17]
0s

✓ [17]
0s