The Battle of Neighbourhoods

1.Introduction

1.1 Background

Bengaluru is a capital city of Karnataka. The centre of India's high-tech industry, the city is also known for its parks and nightlife. As a resident of city I decided to use Bengaluru in my project. In recent decades, the city has witnessed rapid growth in population and urbanized area. Many businessmen arrives here with a dream of investment.

1.2 Problem

When we think as an investor, we expect from them to prefer locality where the type of business they want to install is less intense. However, it is difficult for investors to get these information easily.

So, we can create a map and information where each neighbourhood of the city is clustered according to the venue density.

2.Data

To solve the above problem, that is to create a map and information on all the neighbourhoods of Bengaluru city, we need data of the city.

There are many sources on internet from where we can get list of neighbourhoods of the Bengaluru City, e.g. Wikipedia.

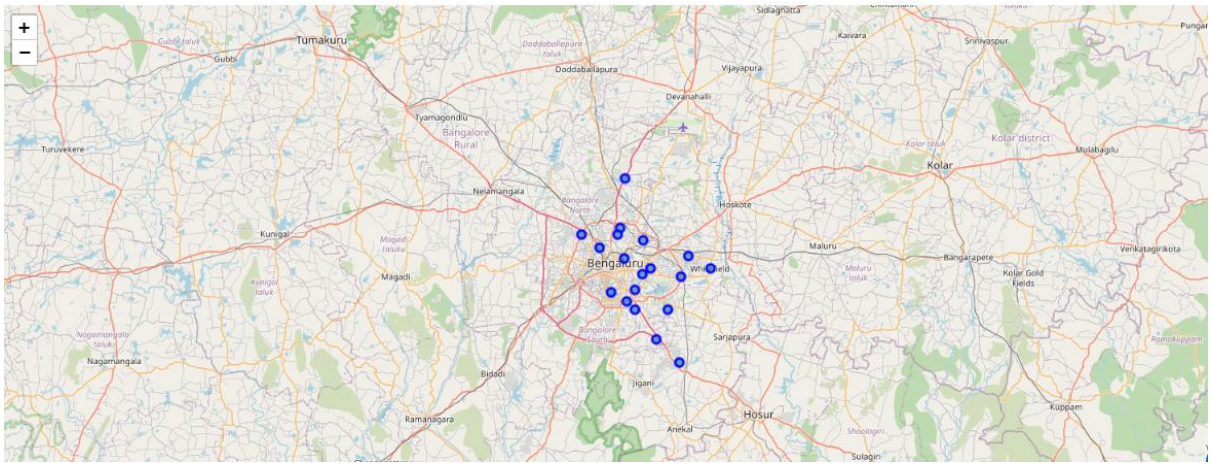
I have organized a data set in csv file with a list of all neighbourhoods along with their latitude and longitude.

Zipcode	Name	Latitude	Longitude
560071	Domlur	12.9608	77.6361
560038	Indiranagar	12.9699	77.6499
560003	Malleswaram	13.0031	77.5643
560001	Shivajinagar	12.9857	77.6057
560103	Bellandur	12.9034	77.6784
560048	Hoodi	12.9896	77.7127
560037	Marathalli	12.9569	77.7011
560066	Whitefield	12.9698	77.75
560043	Kammanahalli	13.0159	77.6379
560024	Hebbal	13.0354	77.5988
560022	Yeshwanthpur	13.025	77.534
560032	RT Nagar	13.0247	77.5948
560063	Yelahanka	13.1155	77.607
560068	Bommanahalli	12.903	77.6242
560099	Bommasandra	12.8167	77.6974
560076	BTM Layout	12.9166	77.6101
560100	Electronic City	12.8542	77.6602
560034	Koramangala	12.9352	77.6245
560041	Jaya Nagar	12.9308	77.5838

3.Methodology

The data which I have organized is clean and can be used as it for the project.

I used the **folium** library to visualize the geographic details of the Bengaluru city and its neighbourhoods. I created a map of the city with neighbourhoods superimposed on top. To create the map I used **geopy** library to get the latitude and longitude values of the city.



I utilized the **Foursquare API** to explore the venues of each neighbourhoods of the city.

I designed the limit of 100 and radius of 1000mtrs for each neighbourhoods from their given latitude and longitude present in my data set. The result gave me list of all the venues,along with

their respective geographical locations, for all the neighbourhoods. As the result was a json file, so I cleaned the file and put the data in a dataframe. Below is the head of the dataframe.

21]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Domlur	12.9608	77.6361	League of Extraordinary Gamers	12.967099	77.636919	Gaming Cafe
1	Domlur	12.9608	77.6361	Lavonne	12.963909	77.638579	Café
2	Domlur	12.9608	77.6361	Smoke House Deli	12.965584	77.641498	Deli / Bodega
3	Domlur	12.9608	77.6361	Barbeque Nation	12.962684	77.641599	BBQ Joint
4	Domlur	12.9608	77.6361	Murphy's	12.953659	77.639397	Irish Pub
5	Domlur	12.9608	77.6361	California Burrito	12.967528	77.636722	Burrito Place
6	Domlur	12.9608	77.6361	Starbucks	12.965649	77.641718	Coffee Shop
7	Domlur	12.9608	77.6361	Puma Social Club	12.967254	77.641212	Nightclub
8	Domlur	12.9608	77.6361	Anand Sweets	12.960166	77.645168	Indian Restaurant
9	Domlur	12.9608	77.6361	Drops Total Spirits	12.964527	77.641636	Liquor Store

From the above data set I retrieved the unique categories of the venue. There were total of 127 unique categories.

I created a table which shows list of top 10 venue categories for each neighbourhood according to the frequency of each venue category.

	Neighborhood	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 Most Common Venue	10th Most Common Venue
0	BTM Layout	Indian Restaurant	Ice Cream Shop	Vegetarian / Vegan Restaurant	Coffee Shop	Café	Bakery	Pizza Place	Snack Place	Chinese Restaurant	Park
1	Bellandur	Sports Bar	Badminton Court	Department Store	Indian Restaurant	Gym / Fitness Center	Other Great Outdoors	General Entertainment	Gastropub	Café	Farmers Market
2	Bommanahalli	Hotel Bar	Tea Room	Hotel Pool	Department Store	Cafeteria	Furniture / Home Store	Pizza Place	Indian Restaurant	Grocery Store	Farmers Market
3	Bommasandra	South Indian Restaurant	Sporting Goods Shop	Coffee Shop	Print Shop	Gym / Fitness Center	Gym	Cosmetics Shop	Cupcake Shop	Deli / Bodega	Department Store
4	Domlur	Indian Restaurant	Café	Coffee Shop	Hotel	Italian Restaurant	Sports Bar	Food Truck	Burger Joint	Burrito Place	Pub

We observed from the above table, that there are some common venue categories in neighbourhoods. So, I decided to segment and cluster the neighbourhood.

For, this reason I used unsupervised learning K-means algorithm of ML to cluster the neighbourhoods. K- Means algorithm is the most common cluster algorithm and fits apt for my purpose. I created 3 clusters.

```
bengaluru_merged
```

84]:

	Zipcode	Neighborhood	Latitude	Longitude	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 Most Common Venue	10th Most Common Venue
0	560071	Domlur	12.9608	77.6361	0	Indian Restaurant	Café	Coffee Shop	Hotel	Italian Restaurant	Sports Bar	Food Truck	Burger Joint	Burrito Place	Pub
1	560038	Indiranagar	12.9699	77.6499	0	Indian Restaurant	Ice Cream Shop	Pub	Bakery	Lounge	Chinese Restaurant	Café	Burger Joint	Restaurant	Cocktail Bar
2	560003	Malleswaram	13.0031	77.5643	0	Indian Restaurant	Ice Cream Shop	Vegetarian / Vegan Restaurant	South Indian Restaurant	Coffee Shop	Café	Breakfast Spot	Light Rail Station	Park	Pizza Place
3	560001	Shivajinagar	12.9857	77.6057	0	Indian Restaurant	Clothing Store	Hotel	Fast Food Restaurant	Women's Store	Bar	Shopping Mall	Café	Candy Store	Men's Store
4	560103	Bellandur	12.9034	77.6784	0	Sports Bar	Badminton Court	Department Store	Indian Restaurant	Gym / Fitness Center	Other Great Outdoors	General Entertainment	Gastropub	Café	Farmers Market
5	560048	Hoodi	12.9896	77.7127	0	Indian Restaurant	Paper / Office Supplies Store	Supermarket	Salad Place	Bus Station	Cafeteria	Soccer Field	Bed & Breakfast	Gastropub	Hotel
6	560037	Marathalli	12.9569	77.7011	0	Indian	Clothing Store	Vegetarian / Vegan	Dessert Shop	Sports Bar	Department	Hotel	Gastropub	Andhra	Chinese

Now, I will label each cluster as below:

Cluster 1: "Indian Restaurant"

Cluster 2: "Sports/Gym Center"

Cluster 3: "Market

4.Results

Let’s create 3 tables for each cluster with 5 most common venue categories.

Cluster 1: Indian Restaurant

]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Domlur	Indian Restaurant	Café	Coffee Shop	Hotel	Italian Restaurant
1	Indiranagar	Indian Restaurant	Ice Cream Shop	Pub	Bakery	Lounge
2	Malleswaram	Indian Restaurant	Ice Cream Shop	Vegetarian / Vegan Restaurant	South Indian Restaurant	Coffee Shop
3	Shivajinagar	Indian Restaurant	Clothing Store	Hotel	Fast Food Restaurant	Women's Store
4	Bellandur	Sports Bar	Badminton Court	Department Store	Indian Restaurant	Gym / Fitness Center
5	Hoodi	Indian Restaurant	Paper / Office Supplies Store	Supermarket	Salad Place	Bus Station
6	Marathalli	Indian Restaurant	Clothing Store	Vegetarian / Vegan Restaurant	Dessert Shop	Sports Bar

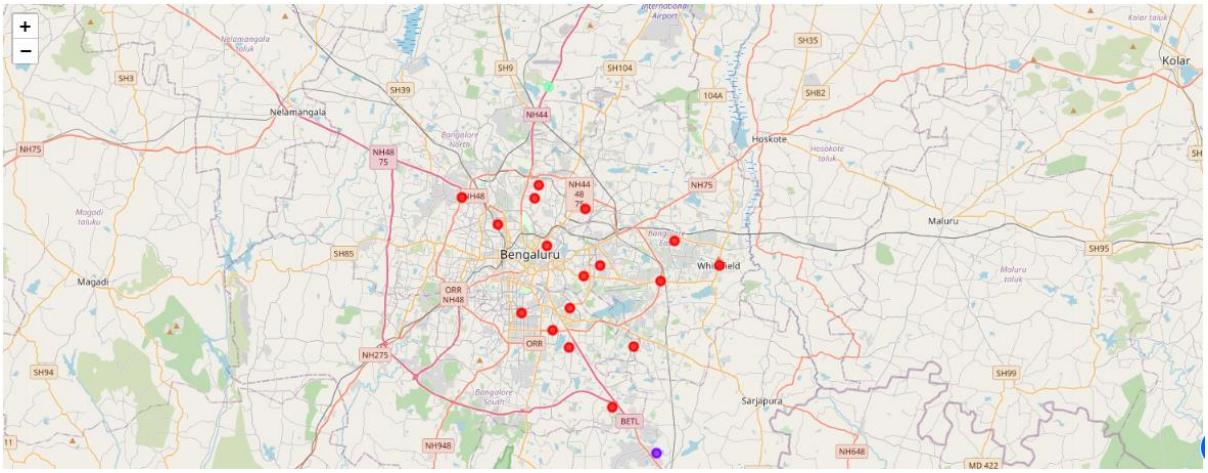
Cluster 2: Sports/Gym Center

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Bommasandra	South Indian Restaurant	Sporting Goods Shop	Coffee Shop	Print Shop	Gym / Fitness Center

Cluster 3: Market

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Yelahanka	ATM	Indian Restaurant	Restaurant	Farmers Market	Flea Market

Finally, we can see the map with the clusters



Looking at the clusters a person can make decision for investing on type or place for business.
For example if a person wants to invest on openening a restaurant then he can look on cluster 3 where there are very few restaurants.

5.Discussion

As I mentioned before, Bengaluru is a major city of India and densely populated. There are far more neighbourhoods which can be included in the dataset for the data analysis which will give more detailed information for every business type. Using Foursquare API, we can get more information about every venue and we can find out which venue is trending in which area. I used K-means clustering for clustering the neighbourhoods which gave more clear idea for business owners.
At the end, I created a map using folium library which gives a better visualization of the data analysis performed.

6.Conclusion

As more people are turning to big cities like Bengaluru for their business set up, they can achieve better outcomes using the above analysis.