

South to North Bengaluru

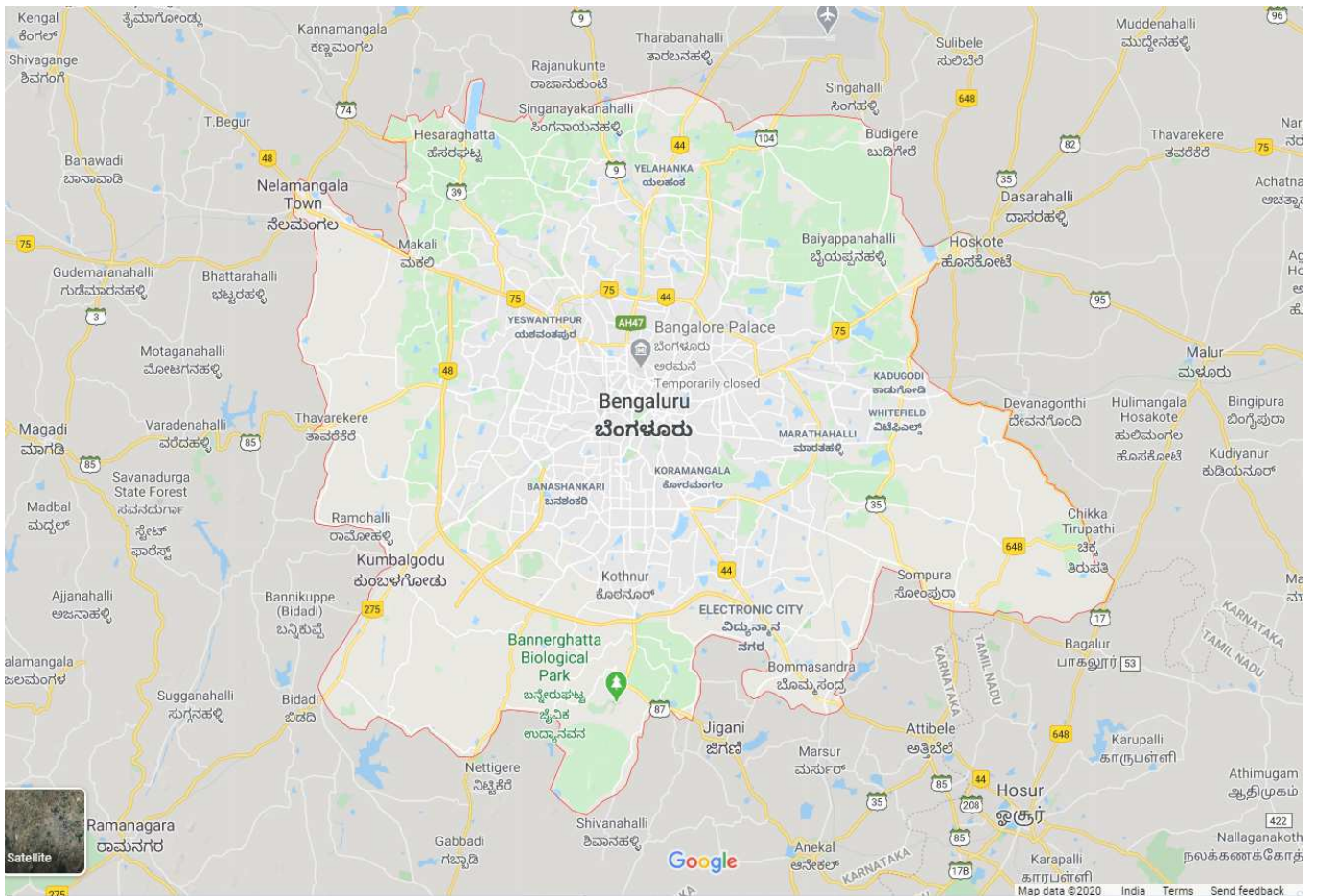
Sadhana Rajendran

17th April, 2020

1. Introduction

1.1 Bengaluru city

Bengaluru is a large metropolitan city in South India. It is a major technology and start-up hub of the country and, as a result, has seen rapid growth and expansion over the past decade.



Bengaluru attracts several techies from across the country who are employed by global Tech giants like Microsoft, Google, Uber, etc. that have their Indian offices in the city and by start-ups that are causing waves in the domestic market. These companies are usually nested within large Tech Parks with sprawling campuses scattered all across the city.

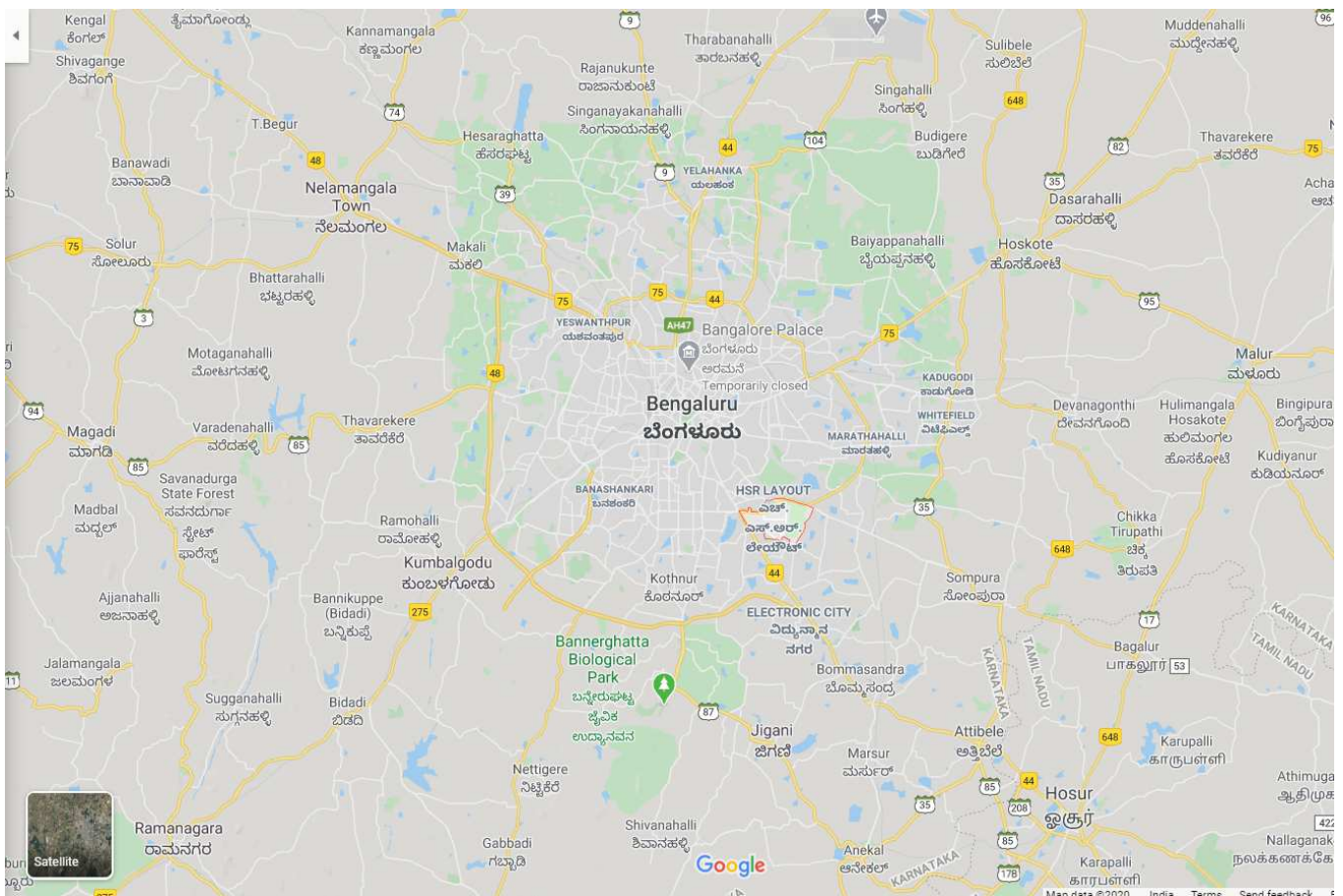
Apart from this, Bengaluru is also well known for its happening nightlife, thanks to the young working populace. The city is home to a large number of microbreweries and other eateries, catering to every palette, which attract large crowds even on weekdays.

On the downside, due to the sudden, uncontrolled expansion witnessed by the city, the traffic is a nightmare, especially at peak hours in the morning and evening. A lack of good intracity public transport system results in a lot of private vehicles on the road causing congestions. It is not uncommon for an average commuter to spend upwards of 3 hours on the road on bad days.

Working professionals tend to consider renting out flats that are in close proximity to their workplaces in neighbourhoods that have an affordable rental rates. Additionally, a neighbourhood that has easy access basic amenities, restaurants, entertainment centers is much sought after.

1.2 HSR Layout neighbourhood

HSR Layout is an area in south Bengaluru that is one such locality. It is connected to some of the largest Tech Parks in surrounding areas like Bellandur, Sarjapur, Electronic City apart from sprouting several start-ups. It is also a hub for food and shopping and is only a short distance away from Koramangala that is a popular pub hopping destination.



However, if a resident of HSR Layout lands a new job in Manyata Tech Park or World Trade Center which is in North Bengaluru, then they will have to rent out a new flat closer to their workplace to avoid long hours commuting. Hence, a neighbourhood needs to be identified for residing in, which is similar to HSR Layout but is in North Bengaluru.

This can also be extended to identify other neighbourhoods in Bengaluru that are similar to HSR Layout or to identify clusters of similar neighbourhoods. This would help those looking to move from one part of Bengaluru that they are familiar with to another, new one to shortlist localities to start their house hunting.

1.3 Problem

Grouping neighbourhoods of Bengaluru to help house hunters who have to move from one part of the city to another or small business owners who want to set up an outlet in a different part of the city, to identify potential neighbourhoods that suits their purpose with similar:

- Rental rates
- Access to amenities

With this information identify a neighbourhood in North Bengaluru that is similar to HSR Layout in South Bengaluru.

2. Data

2.1 Data description and sources

For the purpose of this project, the required data inputs and their sources are as given below:

Data input required	Data source	Example
Localities in Bengaluru	Makaan.com Bengaluru rental rates page	HSR Layout, Malleshwaram, Kalyan Nagar
Rental rates of localities	Makaan.com Bengaluru rental rates page	Average rental rates for 2BHK in th localities
Coordinates of localities	Geocoder and Geopy	Latitude and longitudes of the localities
Popular spots in a locality	Foursquare API	Restaurants, entertainment centers, bars

The details of the localities and the average property rental rates in Bengaluru to be considered will be fetched from this Makaan.com page: <https://www.makaan.com/price-trends/property-rates-for-rent-in-bangalore?page=1> (<https://www.makaan.com/price-trends/property-rates-for-rent-in-bangalore?page=1>).

A sample page of data is presented below.

Locality Name	Rental Rates					
<input type="text" value="search locality in bangalore"/>	1 BHK		2 BHK		3 BHK	
	Rent range	Avg rent	Rent range	Avg rent	Rent range	Avg rent
HSR Layout	₹ 5,750 - 40,000	₹ 16,256.02	₹ 8,000 - 1.5 L	₹ 26,750	₹ 32,000 - 1 L	₹ 54,818.18
Koramangala	₹ 7,500 - 45,000	₹ 18,732.43	₹ 16,500 - 60,000	₹ 32,327.27	₹ 30,000 - 1.3 L	₹ 57,31.25
Jigani	₹ 6,500 - 35,000	₹ 20,750	₹ 7,000	₹ 7,000	₹ 12,000 - 17,000	₹ 15,312.5
Mahadevapura	₹ 6,500 - 18,000	₹ 10,328.57	₹ 9,800 - 15,000	₹ 13,525	₹ 35,000	₹ 35,000
Whitefield Hope Farm Junction	₹ 5,000	₹ 5,000	₹ 13,500	₹ 13,500	₹ 42,000 - 2 L	₹ 96,400
Krishnarajapura	₹ 8,000 - 95,000	₹ 19,222.22	₹ 10,800 - 17,500	₹ 14,238.33	₹ 60,000 - 61,000	₹ 60,500
Whitefield	₹ 5,499 - 1.6 L	₹ 15,697.38	₹ 6,500 - 49,000	₹ 18,845.15	₹ 21,000 - 50,000	₹ 37,900
Yelahanka	₹ 4,000 - 8,000	₹ 6,625	₹ 10,000	₹ 10,000	₹ 8,000 - 70,000	₹ 32,459.78
Devanahalli	₹ 12,000 - 18,000	₹ 15,000	₹ 11,000 - 25,000	₹ 16,000	₹ 45,000 - 60,000	₹ 51,666.67
Jakkur	₹ 8,000 - 19,000	₹ 14,214.29	₹ 15,000 - 28,000	₹ 19,142.86	₹ 40,000 - 47,000	₹ 43,000
Chandapura	₹ 6,000 - 13,000	₹ 9,000	₹ 8,000 - 20,000	₹ 14,000	₹ 11,000 - 15,000	₹ 13,000
Horamavu	₹ 8,000 - 13,000	₹ 11,200	₹ 8,000 - 32,000	₹ 17,165	₹ 15,000 - 35,000	₹ 22,333.33
Chandapura Anekal Road	-	-	-	-	-	-
Varthur	₹ 9,500 - 13,000	₹ 10,357.14	₹ 11,000	₹ 11,000	₹ 22,000	₹ 22,000

2.2 Bengaluru neighbourhoods data collect and processing

Data is imported from the Makaan.com pages by scraping the site. The details of the neighbourhoods are stored along with the average rental rates of 2BHKs in the neighbourhoods. Only the 2BHK rates are considered for the purpose of this project. Also, only data from the first 10 pages of results is be considered as the data quality was observed to drop after that.

The following actions are performed on the data scraped from Makaan.com:

- Records which have poor data quality (eg.: '-' values for rental rates)
- Modify data type of the average rental rate of the neighbourhoods to float so that mathematical operations can be performed
- Remove records where the neighbourhood names were garbage (eg.: Neighbourhood value '2nd cross road')

It was observed that there were several cases where the same neighbourhood had been repeated twice albeit with slightly different spellings. In such cases, the multiple records for the same neighbourhood were combined to form one row with the value of the average rental rate taken to be the average of the records.

Next, the geographical coordinates of the neighbourhoods are fetched using Geocoder. It is to be noted that not all neighbourhoods fetch results from Geocoder. Hence, such neighbourhoods are skipped while traversing the list. These rows have latitudes and longitudes as NaN post processing which are then dropped.

2.3 Venues data collect and cleaning

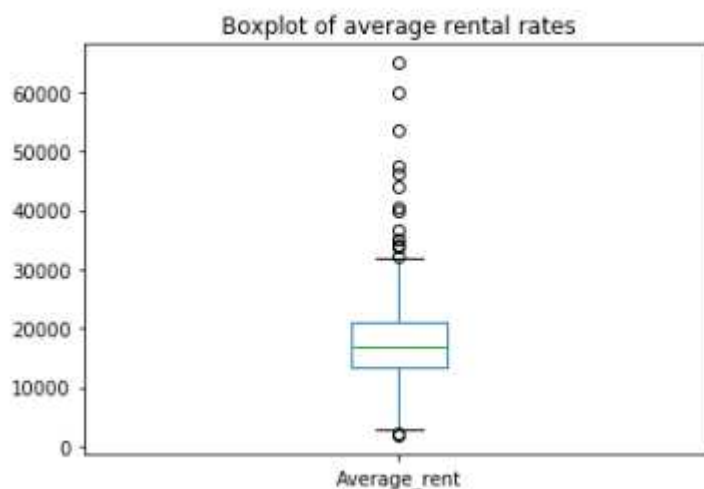
The top 150 popular venues within a 1.75km radius of a neighbourhood's coordinates are fetched using the Foursquare API and consolidated in the same table. Columns are added for each venue category encountered. Each column gives the number of venues belonging to that particular category in the area of the neighbourhood.

A new column is added to store the count of popular venues in a neighbourhood.

3. Methodology

3.1 Preliminary data analysis

A box plot is plotted to visualize the rental rate data.

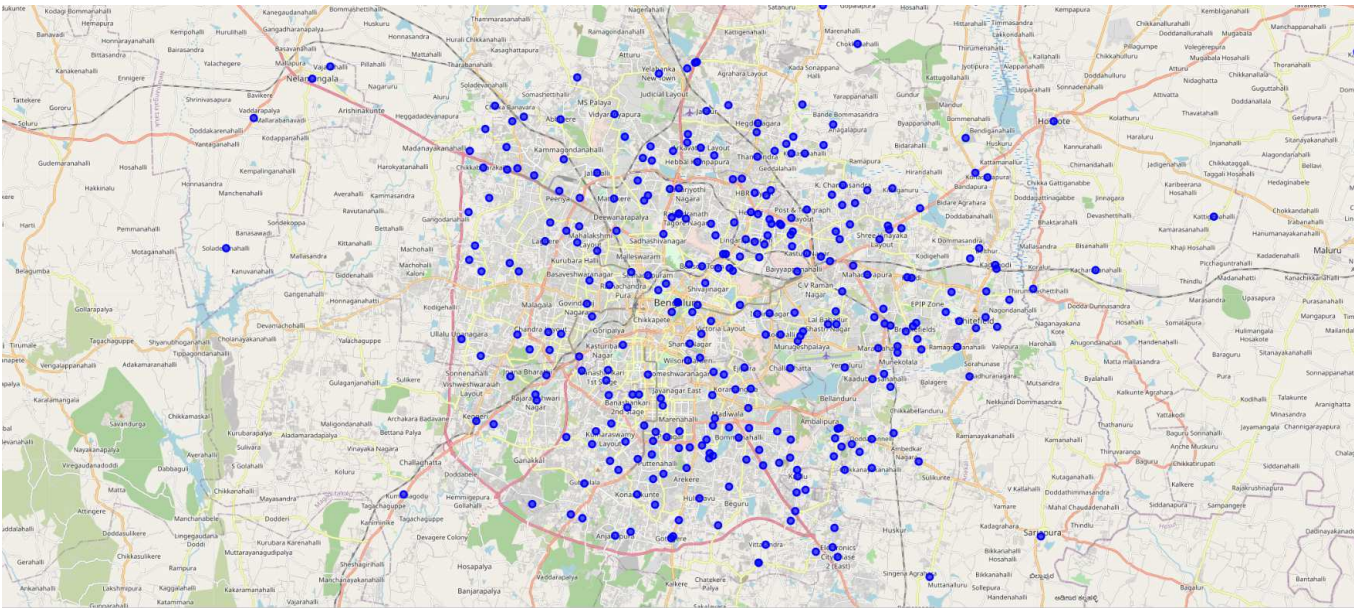


The average rental rates were fetched for 340 neighbourhoods of Bengaluru.

It can be observed from the initial analysis of the average rental rates in the neighbourhoods of Bengaluru that:

- the average rental rate across Bengaluru city is Rs. 18000 (approx) with a standard deviation of Rs. 7800 (approx)
- half the neighbourhoods have an average rental rate between Rs. 13500-21000
- with an average rental rate of Rs. 26750, HSR layout is on the higher end (4th quartile) of the spectrum
- there are several neighbourhoods in Bengaluru that can be considered to be outliers due to extremely high average rental rates

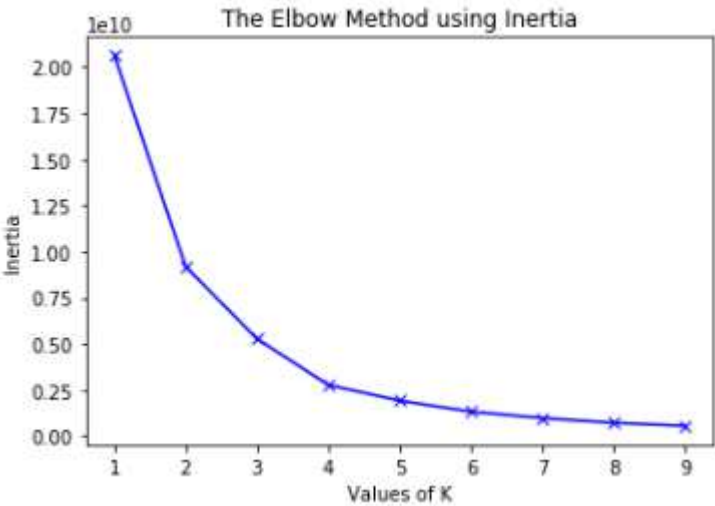
The neighbourhoods are then plotted on a map to visualize their geographic spread.



3.2 K-means clustering

K-means clustering is the chosen approach to group the neighbourhoods of Bengaluru.

As a first step, the optimal value of K to be used is determined using the elbow point method as shown below.



The value of K in the K-means clustering algorithm is chosen to be 6 as the curve shown above becomes linear from this point onwards.

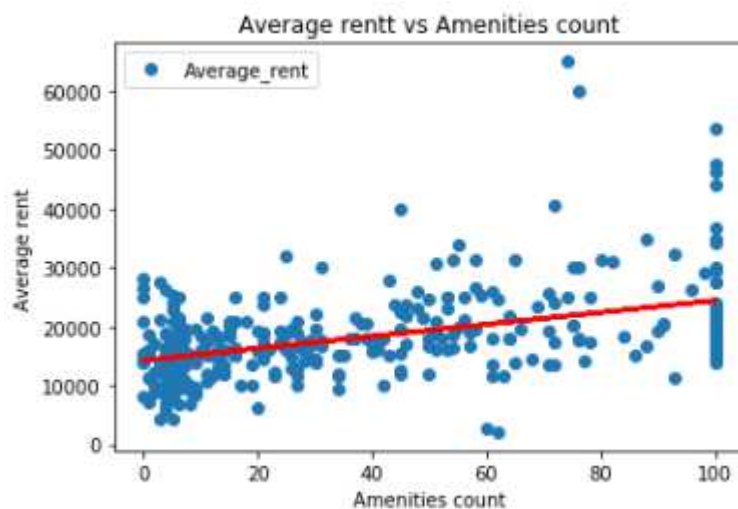
The results of the K-means clustering algorithm are stored in the same table. The Cluster labels calculated are added as a column with a value ranging between 0-5 for each neighbourhood.

The mean rental rate and the number of amenities per neighbourhood was calculated for each cluster so formed.

	Average_rent	Amenities_count
Cluster_labels		
0	9767.630769	15.369231
1	32628.382857	73.714286
2	19283.056061	44.434343
3	52727.611667	91.666667
4	15006.175514	27.710280
5	25002.187857	48.666667

3.3 Regression

Since amenities count and average rental rates of the clusters appear to be related, a linear regression is performed on the entire dataset to view the relationship.



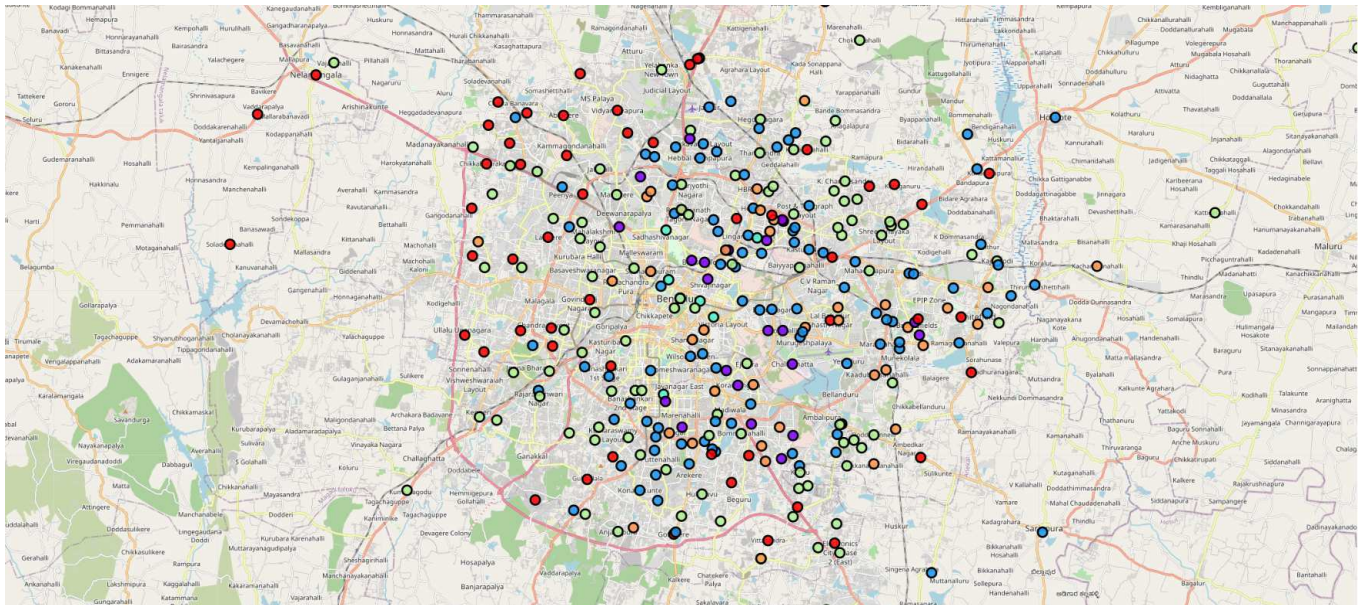
Thus, it can be seen that there exists a linear relationship between the average rental rate of a neighborhood and the amenities available in it.

4. Results

The Bengaluru city has been divided into 6 clusters based on rental rates and amenities present in the neighbourhoods.

4.1 Clustered Bengaluru neighbourhoods

The neighbourhoods are plotted on a Map again, this time with the results of the clustering. The markers are encoded with colours indicating the cluster.



4.2 HSR equivalent in North Bengaluru

Based on the clustering and the visualization of the clusters, the neighbourhoods in North Bengaluru that is similar to HSR Layout in South Bengaluru are Sanjaynagar and RMV, both of which are located pretty close together and, hence, to simply the result, we finalize Sanjaynagar to be the location of choice.

Both HSR Layout and Sanjaynagar belong to Cluster 5. A comparison of the two neighbourhoods based upon their rental rate and their amenities is given below.

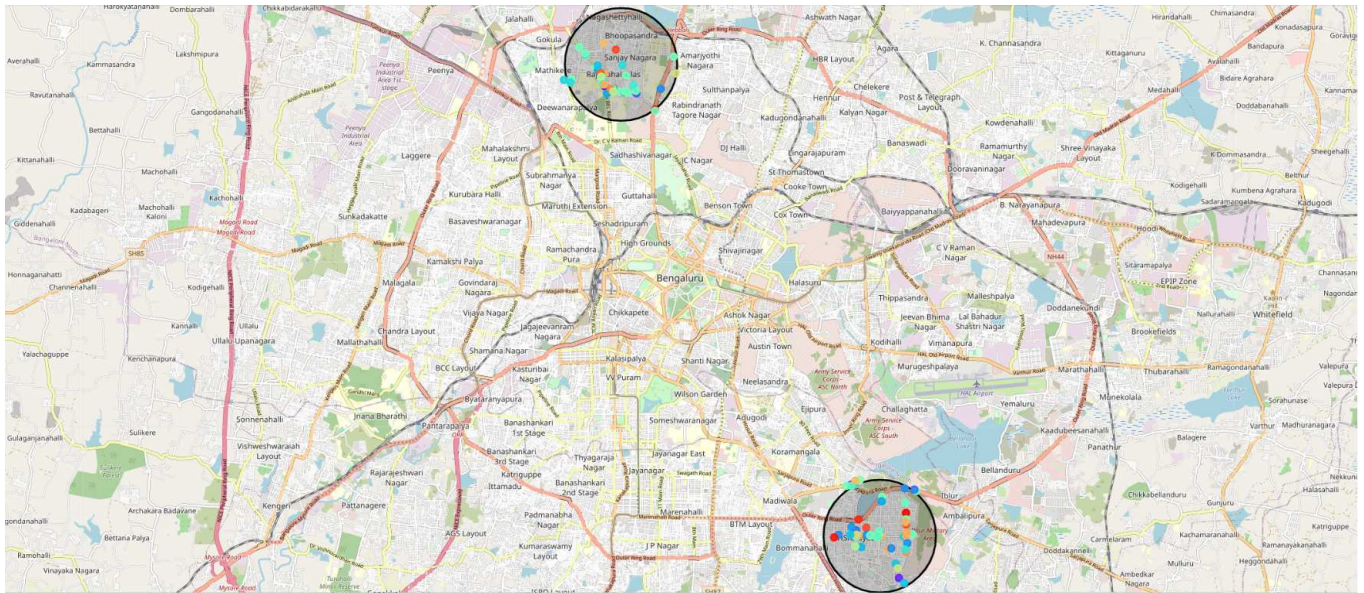
Cluster_labels	Neighbourhood	Average_rent	Amenities_count
5	HSR Layout	26750.0	58
5	Sanjaynagar	25250.0	59

Out[41]:

		Count
Venue Category	Neighbourhood	
American Restaurant	HSR Layout	1
Asian Restaurant	Sanjaynagar	1
BBQ Joint	HSR Layout	1
Bakery	HSR Layout	1
Bar	Sanjaynagar	1
Bike Shop	HSR Layout	1
Breakfast Spot	Sanjaynagar	1
Burger Joint	HSR Layout	1
	Sanjaynagar	1
Business Service	HSR Layout	1
Café	HSR Layout	7
	Sanjaynagar	2
Chettinad Restaurant	HSR Layout	1
Chinese Restaurant	HSR Layout	2
	Sanjaynagar	5
Coffee Shop	HSR Layout	1
	Sanjaynagar	1
Department Store	HSR Layout	4
	Sanjaynagar	2
Diner	Sanjaynagar	1
Fast Food Restaurant	HSR Layout	1
	Sanjaynagar	6
Fish & Chips Shop	HSR Layout	1
Frozen Yogurt Shop	Sanjaynagar	1
German Restaurant	Sanjaynagar	1
Gym	HSR Layout	2
	Sanjaynagar	1
Ice Cream Shop	HSR Layout	5
	Sanjaynagar	8
Indian Restaurant	HSR Layout	5
	Sanjaynagar	15
Italian Restaurant	HSR Layout	3

		Count
Venue Category	Neighbourhood	
	Sanjaynagar	2
Kebab Restaurant	HSR Layout	1
Lake	HSR Layout	1
Maharashtrian Restaurant	HSR Layout	1
Mediterranean Restaurant	HSR Layout	1
	Sanjaynagar	1
Mexican Restaurant	Sanjaynagar	1
Middle Eastern Restaurant	HSR Layout	1
Motorcycle Shop	Sanjaynagar	1
North Indian Restaurant	HSR Layout	1
Paper / Office Supplies Store	Sanjaynagar	1
Park	HSR Layout	1
	Sanjaynagar	1
Pizza Place	HSR Layout	5
	Sanjaynagar	4
Pub	HSR Layout	1
Punjabi Restaurant	HSR Layout	2
Restaurant	Sanjaynagar	1
Sandwich Place	Sanjaynagar	1
Smoke Shop	Sanjaynagar	1
Snack Place	HSR Layout	2
	Sanjaynagar	1
Steakhouse	HSR Layout	1
Trail	HSR Layout	1
	Sanjaynagar	1
Vegetarian / Vegan Restaurant	HSR Layout	1
	Sanjaynagar	1

Plotting the two neighbourhoods along with their most popular venues in the same Map to see their similarity.



5. Discussion ¶

- There is a close relationship between the amenities available in a neighbourhood and the average rental rates in the area - the more the amenities, the higher the rental rates of the neighbourhood
- Neighbourhoods that are cheaper but also have fewer amenities than the rest tend to be found outside the outer ring road of Bengaluru
- As one moves towards the interiors, the neighbourhoods become more desirable as is reflected in the increasing rental rates and amenities count
- The outliers found earlier which had extremely high average rental rates compared to the rest of the city are all grouped pretty close to each other in central Bengaluru
- Clusters 2 and 5 are similar in terms of amenities count and yet the former has an average rental rate that is 25% lesser - these are largely located in North East Bengaluru can be viewed as the developing areas of the city which are likely to yield high RoI in the future

6. Conclusion

By clustering, similar neighbourhoods, in terms of rental rates and amenities, have been identified across Bengaluru. Along the way, the relationship between the rental rates, amenities and location of the neighbourhoods also emerged organically.

Using this, Sanjaynagar has been identified as the best location in North Bengaluru that closely resembles HSR Layout in South Bengaluru. A comparison of the two localities has also been presented.