

Capstone Project

Identification of Ideal Localities for Residence in
Bangalore

April 2020

By: Arohan Ajit



Introduction

The city of Bangalore is located in centre of South India in the state of Karnataka. Bangalore has a population of 8,430,000 and area of 709 sq. km. Bangalore is one of the key destination for IT sector companies and experts. City has a huge population most of which are IT sector employees. It is also a preferred location for all the aspiring Software Engineers, Data Scientists and ML Engineers in India. Bengaluru is sometimes referred to as the "Silicon Valley of India" (or "IT capital of India") because of its role as the nation's leading information technology (IT) exporter. Indian technological organisations ISRO, Infosys, Wipro and HAL are headquartered in the city. Bangalore is the second fastest-growing major metropolis in India. Bengaluru has one of the most highly educated workforces in the world. It is home to many educational and research institutions.

Business Problem

Due to its booming status, the population – one of the significant problems of the country, has been growing in this city at a rapid pace. As such, the problem of providing comfortable residence to the new and incoming people in Bangalore is becoming a bigger challenge. Uncomfortable homes leads to extra stress and thereby reduction if overall efficiency. Finding a home in the right locality is becoming a bigger challenge day by day and a resolution to this problem can be beneficial for the customers and profitable for businesses.

Target Audience

Target audience of this project would be all the entrepreneurs and companies looking to step into the domain of real estate in Bangalore. Quote an article from housing.com: 'Buyer sentiments are expected to remain positive in the upcoming quarters. New launches are expected to increase, even though sales is likely to pick up pace only gradually. Key markets for the future are locations with the availability of land parcels, proximity and connectivity to commercial hotspots due to its emergence as an office market and proximity to commercial hotspots'. It is quite apparent that realty business is here to stay and has only upward to grow. Owing to growing popularity of Bangalore as an ideal urban city to live and work in, these businesses can target incoming customers in order to provide them with a residence in a locality of their choice. The potential for this market is growing owing to rapid urbanisation and rising population of the country.

Data

- The district data was taken from Wikipedia page listing all the wards in Bangalore and their respective assembly constituency. It is used to mark and clearly define various wards or localities present in Bangalore as designated by the Government. Each wards is present under an Assembly constituency: https://en.wikipedia.org/wiki/List_of_wards_in_Bangalore
- In order to plot the location of city and its wards, we need to get the coordinates of each ward and city as a whole in terms of latitude and longitude. Coordinates of the all the wards in the city was taken from Geocoder ARCGIS API: <https://geocoder.readthedocs.io/>
- After getting details of each ward, we need to find out what are the trending places located in that area – places which people find useful and serve their purpose of comfortable residence. We also need to get the location of such places and to which category they belong. Details of Places of Interest was taken from Foursquare API: <https://developer.foursquare.com/>
- We need the geo data of the city and its wards mapping out the coordinates of each and every part of the ward. Such data is normally stored in form of geojson files. These files contain detailed information regarding shape and location of a place. The geojson of Bangalore and its wards was sourced from an Open github repository: https://github.com/datameet/Municipal_Spatial_Data/blob/master/Bangalore/BBMP.GeoJSON

Methodology

- First we need to acquire data about all the wards in Bangalore. This was sourced from Wikipedia page containing a table detailing list of all the wards present in the Bangalore as well as their Assembly Constituency. This data was scraped using Beautiful Soup Library. You'll notice the Latitude and Longitude field are empty.

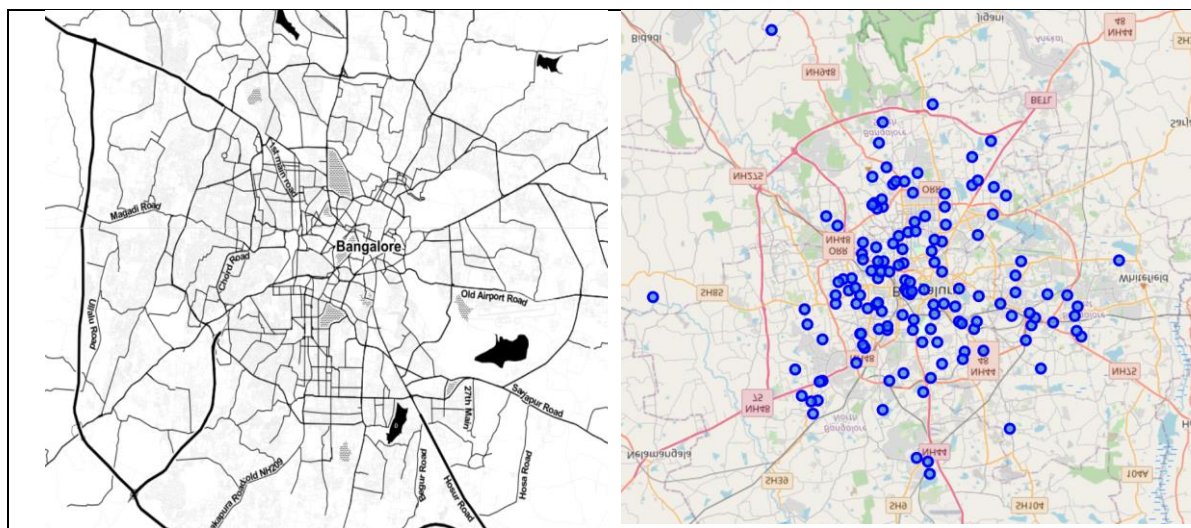
Ward number		Area	Ward	Latitude	Longitude
0	1	Yelahanka	Kempegowda		
1	2	Yelahanka	Chowdeshwari		
2	3	Yelahanka	Attur		
3	4	Yelahanka	Yelahanka Satellite Town		
4	5	Byatarayanapura	Jakkur		

- Now the location data is sourced from the geocoder API. We used the ARCGIS model which provided the necessary coordinates for each ward in Bangalore. This information is supplemented into our dataframe to complete initial data.

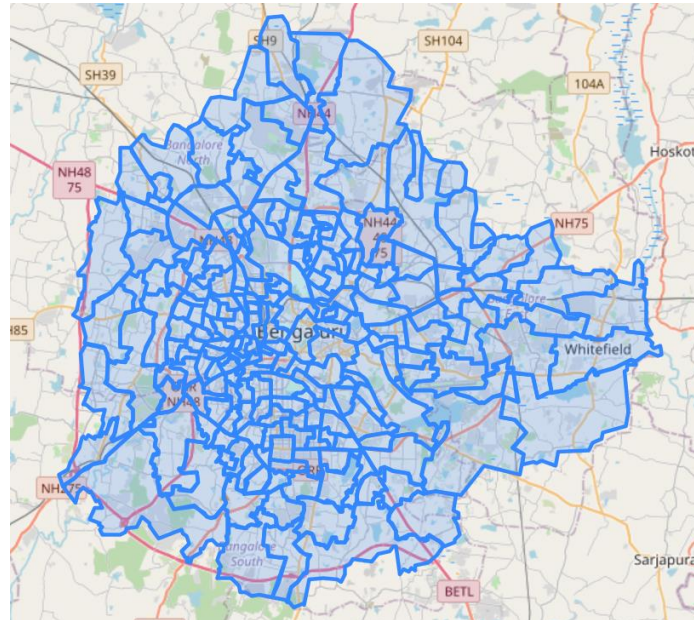
Ward number		Area	Ward	Latitude	Longitude
0	1	Yelahanka	Kempegowda	13.0993	77.5926
1	2	Yelahanka	Chowdeshwari	13.1081	77.5937
2	4	Yelahanka	Yelahanka Satellite Town	13.097	77.5841
3	5	Byatarayanapura	Jakkur	12.9519	77.5445
4	10	Byatarayanapura	Doddabommasandra	13.0625	77.5591

but: double click to hide

- Based on this data, a map of city is generated and all the wards are highlighted with a marker.



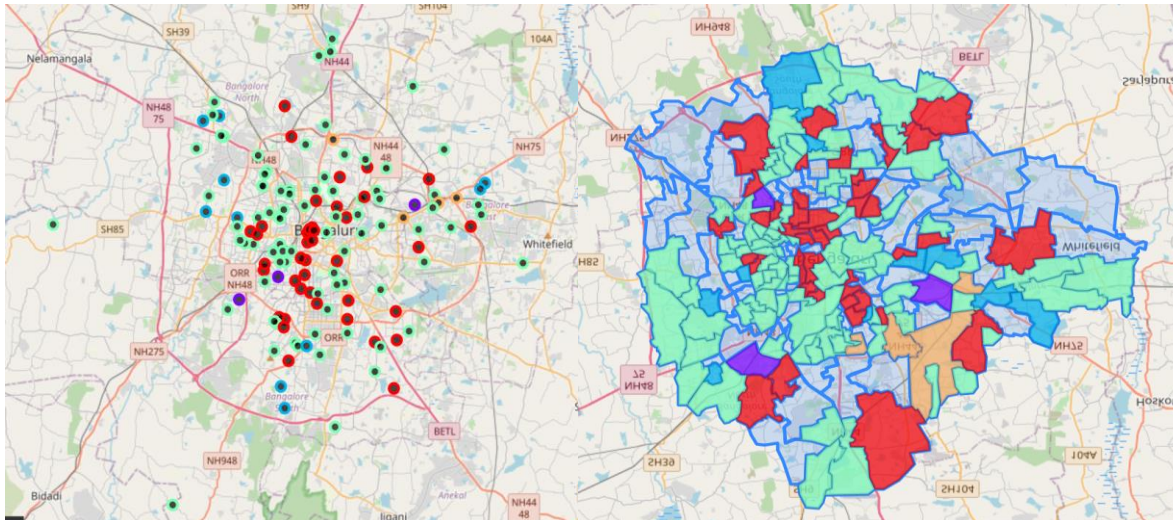
- In order to create a detailed map of Bangalore we need to source its geojson file. After getting the geojson file, we use it to create a detailed map of Bangalore highlighting its each and every district. GeoJSON file contains detailed coordinates for each ward present in Bangalore.



- Using the Foursquare API, all the trending venues in different wards of Bangalore were identified and their location coordinates were recorded. Thereafter top 10 most common venues in each ward were calculated.

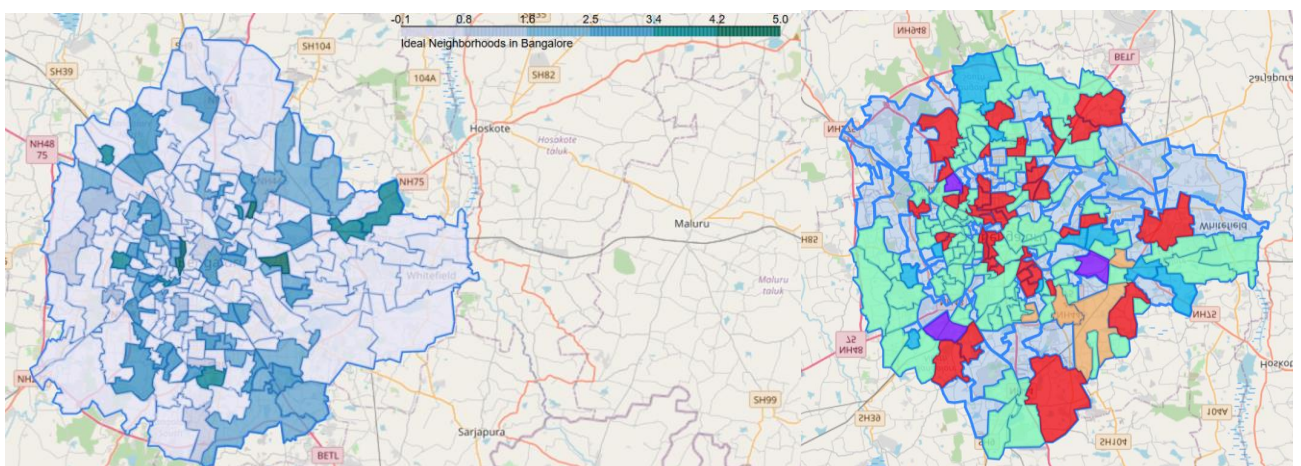
	Ward number	Area	Ward	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	1	Yelahanka	Kempegowda	Smoke Shop	American Restaurant	Ice Cream Shop	Train Station	Café
1	10	Byatarayanapura	Doddabommasandra	Lake	Indian Restaurant	Supermarket	Snack Place	Women's Store
2	100	Rajajinagar	Basaveshwara Nagar	Fast Food Restaurant	Ice Cream Shop	Indian Restaurant	Market	Snack Place
3	101	Rajajinagar	Kamakshipalya	ATM	Bus Station	Fish & Chips Shop	Field	Fast Food Restaurant
4	103	Govindarajnagara	Kaveripura	Park	Women's Store	Dog Run	Fish & Chips Shop	Field

- Based on the results, wards are clustered into different clusters using K-Means clustering based on their similarities and differences. Also a list of preferred facilities were identified and wards were ranked according to number of facilities they contained.



Results

- It was identified that wards in the red cluster was the hotspot for the residential areas. Most of the wards present in this cluster were identified as favourable for living as they contained most of the preferred features



- The green cluster was the 2nd most favoured cluster in terms of residential areas. Green clusters also contain workplaces and stores which might be the reason for its increasing popularity as a residential area
- Other clusters don't have any wards that have shown the potential to be highly demanded area of residence

Discussion

- Since the districts have been sandwiched together, there is a chance of discrepancy in the data as one popular location may come up as a place of interest in multiple districts
- Further analysis of the green cluster is required in order to pin point the causes of its growing popularity as a residential area even though it's mainly located in suburbs
- Preferred features can vary from customer to customer as each one has different view as to what is considered an ideal residential location and what accessories are preferable to them near their residence
- Further research can be done in order to find cause of less residential interest in other wards of the city.
- It is to be noted with high interest in residence, also comes a high cost living in central Bangalore which is something that needs to be taken into account while recommending a location

Conclusion

In this project, we have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, performing machine learning by clustering the data into clusters based on their similarities, and lastly providing recommendations to the relevant stakeholders i.e. realtors and real estate developers regarding the best locations to develop a residential project. Red Cluster in central part of Bangalore has been identified as the hotspot of preferred residential area which can be utilised to its full potential. The findings of this project will help the relevant business to capitalize on the opportunities on high potential locations while avoiding areas in their decisions which will generate low interest.