Bangalore Multiplex Data Analysis

CAPSTONE PROJECT

Data Analysis in Bangalore Multiplexes

- Problem Definition
- Data Gathering and Exploration.
- Proposed Solution
- Result and Benefits.

Problem Definition

Bangalore is a very huge city in South India and Capital of IT industry in India. This is a busy city and people lifestyle is luxurious and spends lot money in shopping movies and outings etc.

Due to this, there are lot of Multiplexes in Bangalore and still there is lot of business scope in this field. Hence, many builders are investing in constructing new malls and Multiplexes.

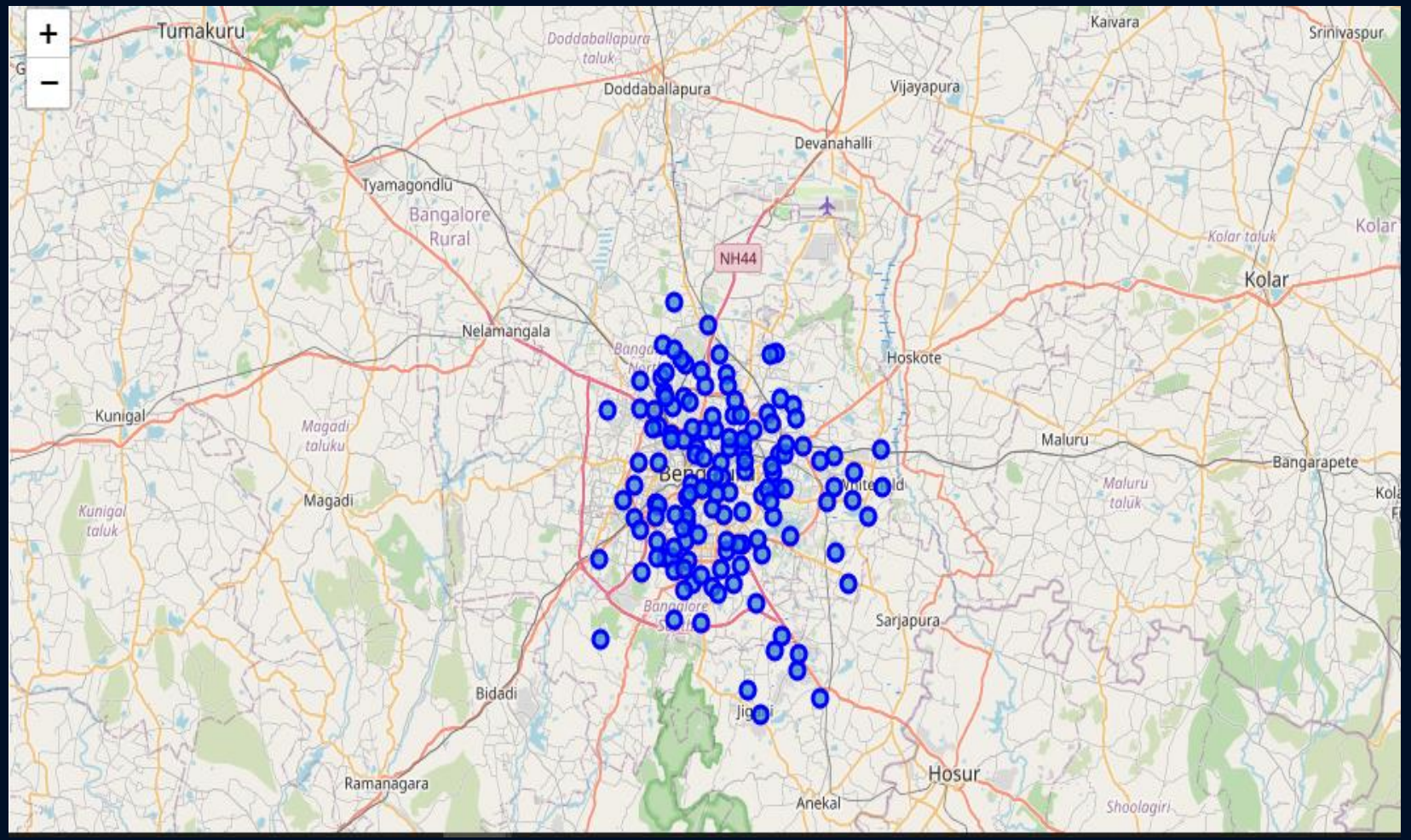
However, there is challenge in finding the in the right place to invest so as to run the business with profits. To address this, it is very much helpful if we provide some insights to the builders to make decisions on their business plans.

Data Gathering and Exploring.

- In order to perform the Analysis and provide the correct inputs, it is very important to gather the correct and sufficient data. The Data should also be enhanced and cleaned as needed.
- As part of this project the data is gathered from Wikipedia and Geocoder python packages for Bangalore Locations and Latitude and Longitudes respectively.
- As there are several regions in Bangalore we should chose only the required zones/Neighborhoods in Bangalore.

Data Gathering and Exploring - Contd

- Bangalore is segregated into several regions like South,North,East,South East,Nort east etc. However we are required to segment the neighborhoods data extract the all neighborhood.
- Added the additional parameters latitude and longitude as they are required for visualizing the data in maps.
- Used Geocoder package to get the latitude and longitude and visualize the map in the data points.
- Generated map in the next slide.


Tumakuru
Doddaballapura taluk
Doddaballapura
Kaivara
Srinivaspur
Vijayapura
Devanahalli
Tyamagondlu
Bangalore Rural
NH44
Kolar taluk
Kolar
Nelamangala
Hoskote
Kunigal
Magadi taluku
Maluru
Bangarapete
Magadi
Maluru taluk
Kunigal taluk
Sarjapura
Bidadi
Ramanagara
Anekal
Hosur
Shoolagiri

Data Gathering and Exploring - Contd

- Once the Neighborhoods are gathered and refined, it is required to get the Multiplexes data on which we should perform the analysis.
- I have explored the data using the Foursquare APIs with radius 3000 which provides me the popular Venues.
- API returned total 9584 venues and 209 unique categories.
- Further filtering the data, it is identified there are total 59 Multiplexes found in the data set which is almost same as the number of multiplexes when searched in Foursquare website.

Proposed Solution

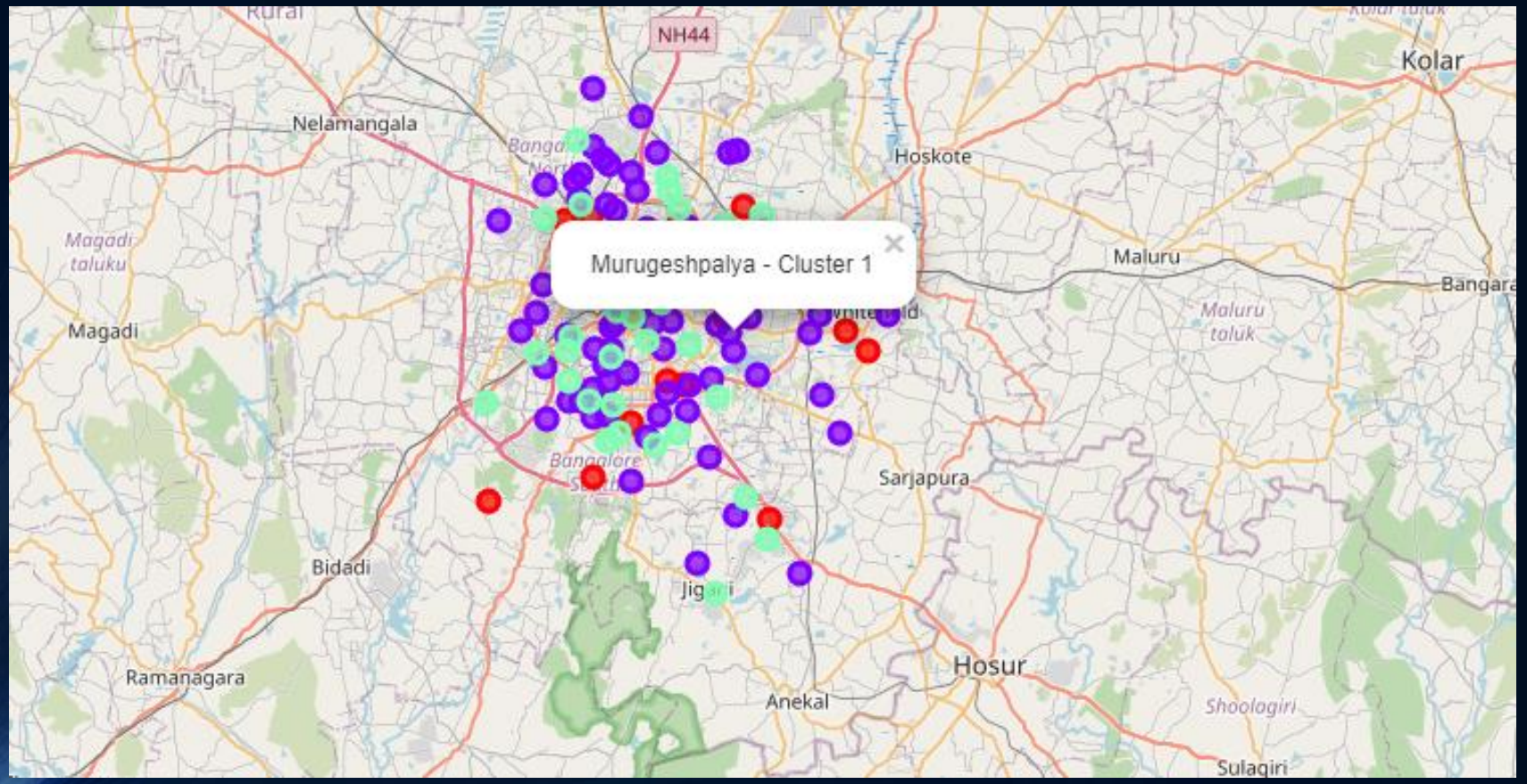
- By the detailed study of the data, it is understood that the data is unsupervised and K-means segmentation and clustering would be best model to get the appropriate results.
- Used K value 3(3 clusters) and trained model and calculated the number of multiplexes falling under each neighborhood which are part of the 3 clusters.

Results and Benefits

- Once the model is trained we received the multiplexes present in the clusters.
- Cluster 1 has 80 data points(Neighborhoods) with 0 multiplexes.
- Cluster 2 has 28 data points(Neighborhoods) with 1 multiplexes and 9 datapoints(Neighborhoods) with 2 Multiplexes
- Cluster 0 has 16 data points(Neighborhoods) with 3 multiplexes and 6 Data points with 4 Multiplexes.

Hence, it can be concluded that the Cluster 1 is good place to start new multiplex.

Map – 3 Clusters with Neighborhoods.



Results and Benefits - Contd

- To conclude, there is lot of scope for business in Bangalore with respect to Multiplexes. This Analysis can be a primary and provide most of the criteria. However, we should also consider the additional parameters in the specific neighborhood. For example, though there are 80 neighborhoods with no multiplex, all the 80 neighborhoods may not be suitable considering the space, lifestyle of the people living there etc.
- With this analysis, not only the real estate but also the govt can make use for study like to find out if there are any specific reasons for not building Multiplexes in those areas.

Thank You