Rental Analysis of Bangalore City Using Various Parameters

D. Megha Sri Satya Sai

1. Introduction:

Bangalore is a megacity and the third-most populous city and fifth-most populous urban agglomeration 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 organizations ISRO, Infosys, Wipro and HAL are headquartered in the city. A demographically diverse city, Bangalore is the second fastest-growing major metropolis in India.[[20]](https://en.wikipedia.org/wiki/Bangalore#cite_note-20) Bengaluru has one of the most highly educated workforces in the world.[[21]](https://en.wikipedia.org/wiki/Bangalore#cite_note-India_Today-21) It is home to many educational and research institutions, such as Indian Institute of Science (IISc), Indian Institute of Management (Bangalore) (IIMB), International Institute of Information Technology, Bangalore (IIITB), National Institute of Fashion Technology, Bangalore.

It is Attracted by majority of population as the intake increases people find it difficult to search for new apartments that satisfy their needs. So in this project I am using the freely available dataset that represents a vast range of apartments around the Bangalore city along with their specifications. To this dataset I added latitude and longitude coordinates making it easy to plot, later I ran Foursquare API over this dataset to find the most famous places around the apartments. Making it easier for the client to choose a more suitable apartment satisfying his metrics.

1. Data Acquisition and cleaning:

2.1 Data Acquisition:

For our analysis we need a dataset that consists of a list of available apartments around the Bangalore city along with major specifications like Number of Bedrooms, Price, Size of Flat, Number of Balcony’s etc.., more the specifications better the dataset. I found such type of dataset at Kaggle titled as “[Bengaluru House Price Data Set](https://www.kaggle.com/amitabhajoy/bengaluru-house-price-data)”. The Data Set Consists of the Following Attributes:

* Availability: Shows when will the apartment will be vacant.
* area\_type: describes the area's built up type.
* location: describes the location of the apartment.
* size: denotes the size of apartment whether 3BHK or 2BHk.
* Society: shows the society of the apartment.
* Total\_sqft: denotes the measurements of the apartment on sqft's.
* bath: denotes number of bathrooms available.
* balcony: denotes the number of balcony's available.
* price: denotes the price of the apartment.
* latitude: denotes the latitude of the neighborhood
* longitude: denotes the longitude of the neighborhood
  1. Data Cleaning and Preprocessing:

Data downloaded or scraped from multiple sources were combined into one table. There were a lot of missing values, because of lack of record keeping. The only problem with the dataset is Latitudes and longitudes aren’t included in the dataset making it difficult for visualization, so we need to insert the Latitudes and longitudes values manually which is a very difficult task to fill them googling each and every one, To tackle this I used Foursquare API and filled the Latitudes and longitudes values into the dataset but this work is very time taking so I took on 600 data points out of 13321 data points.



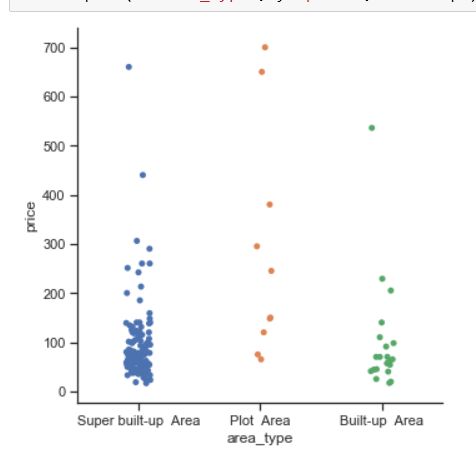
1. Exploratory Data Analysis

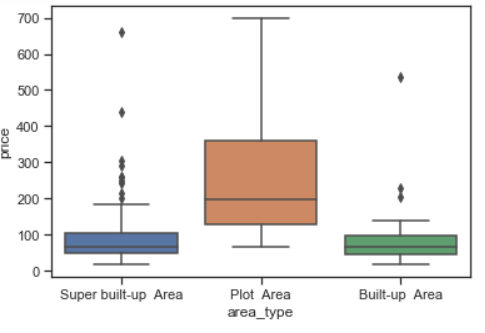
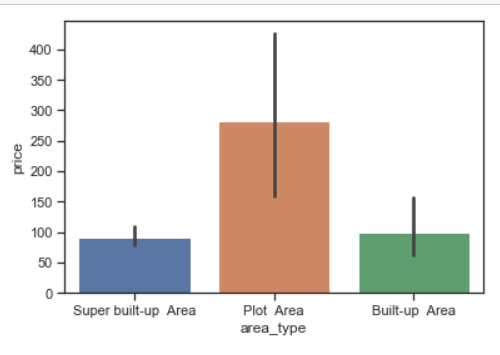
3.1 Inserting New Features:

In the Kaggle dataset is Latitudes and longitudes aren’t included in the dataset making it difficult for visualization, so we need to insert the Latitudes and longitudes values manually which is a very difficult task to fill them googling each and every one, To tackle this I used Foursquare API and filled the Latitudes and longitudes values into the dataset.

* 1. Visualization of various features:

3.2.1 Area\_type vs Price:





We can’t differentiate these categorical values.

3.2.2 Size vs Price

