**HOUSE RENT ANALYSIS**

horizontal line

# 

# Introduction

This report consists of analysis of house rent . For the review of data , it is sure that the dataset is from India. It has columns such as BHK,Rent,Total floors,Furniture Status,etc.The goal of this analysis is to predict the house rent for the client.

## Background

The analysis is performed in Jupyter Notebook with the help of pandas library of python. In Addition to pandas , other required libraries are also imported such as seaborn,sklearn,matplotlib,etc.

## Data Analysis

After loading the dataset in the Jupyter notebook , data cleaning was performed. Some columns contained string values which were converted to numeric format such as Floor.Also there were some independent variables such as Point of Contact which needed to be dropped.

After cleaning the data, visualization was performed with some dependent variables.

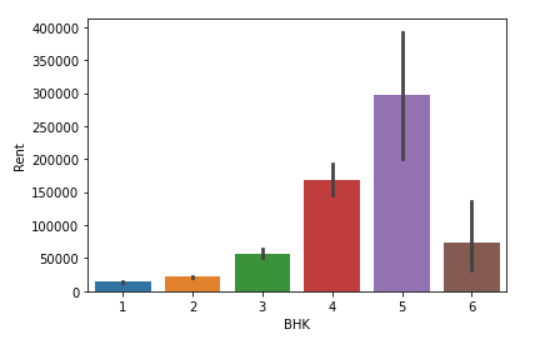


Fig : Barplot between BHK and Rent

Generally we pay more rent for more BHK, so from above bar diagram we can also say that more clearly. But house with 6 BHK have low rent price . I think that is due to lack of water or may be customer.

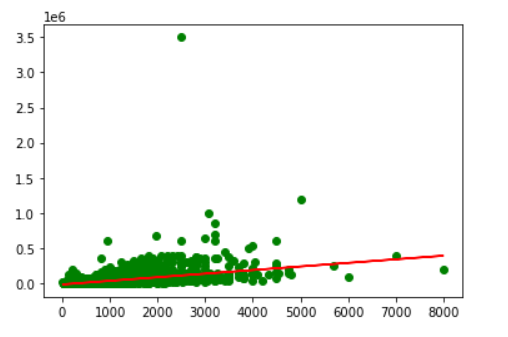


Fig:Scatter plot between size and rent

Above figure shows the scatter plot between Area and Rent. Area can be our important features because the price of rent seems to be highly correlated to Area.i.e rent increases as area increases.

## Model

In this analysis , I chose a linear regression model for prediction.

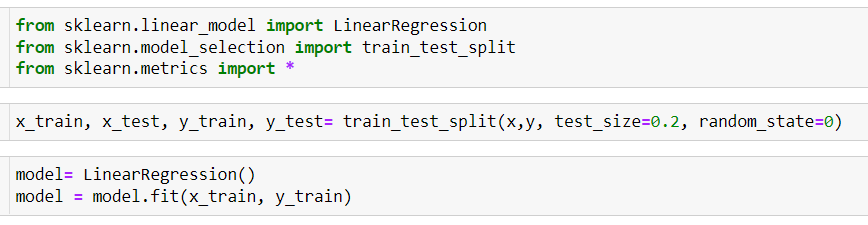


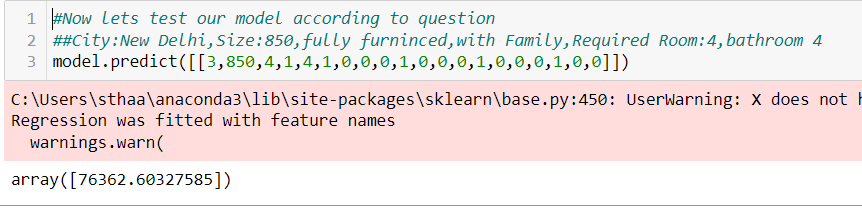
Fig: Code for linear regression

At first , data were separated for training and testing purposes. The training data contains 80 percent of data and for testing the remaining 20 percent were taken.

## Result

As per the requirement of the client ,now the model can predict the house rent for given specifications.

If the client has specifications such as City:New Delhi,Size:850,fully furnished,with Family,Required Room:4,bathroom 4, then following result was obtained



## Conclusion

Here linear regression model was used. Maybe that's why we got low precision and accuracy ,we could follow different algorithms to increase our accuracy and precision .

Root Mean Sq. Error = 43888.874627792065

R square= 0.5130891091376243

*For more detail in report kindly open AnoopShrestha.html file.*