

**This data science project is to build a real estate price prediction website.**

Dataset link: <https://www.kaggle.com/amitabhajoy/bengaluru-house-price-data>

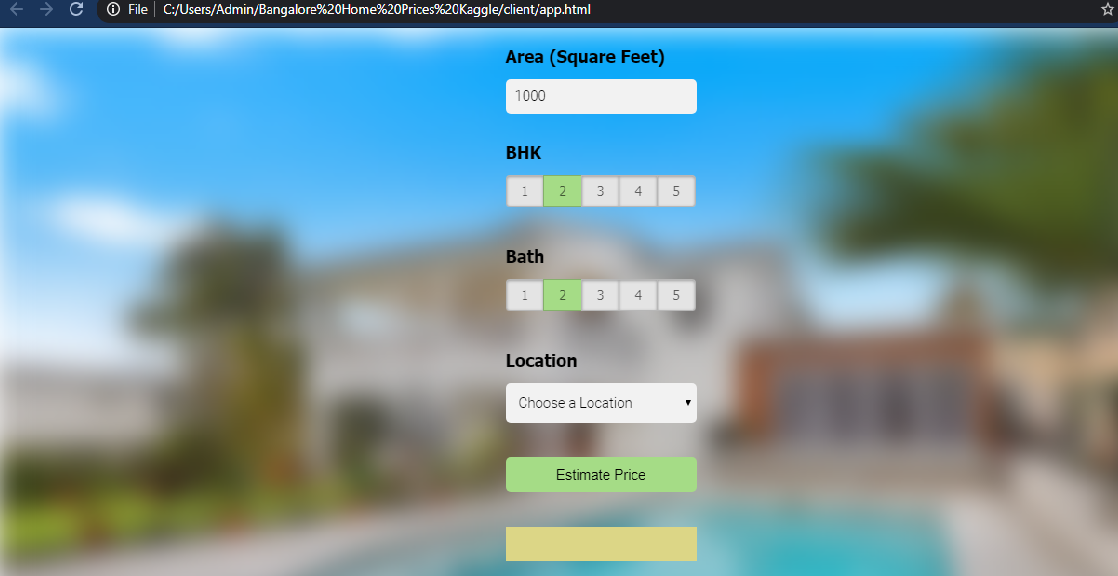
Languages and Tools used:

1. Python
2. Numpy and Pandas for data cleaning
3. Matplotlib for data visualization
4. Sklearn for model building
5. Jupyter notebook, visual studio code and pycharm as IDE
6. Python flask for http server
7. HTML/CSS/Javascript for UI

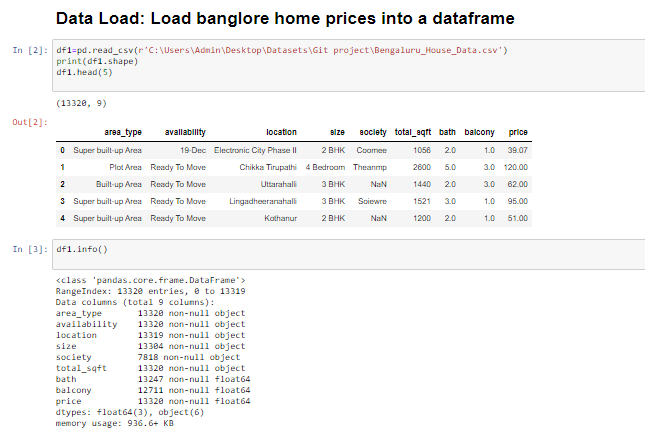
**Step by step process:**

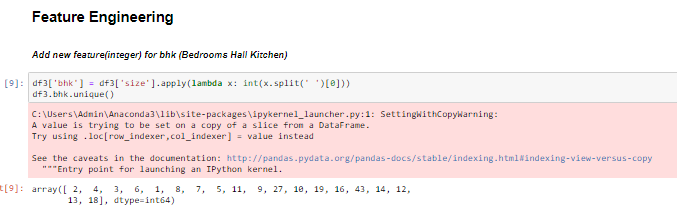
1. Used Python language on Jupyter Notebook for lading the data, Data cleaning, Outlier Detection, Feature Engineering, and Machine Learning Model building. Also used Gridsearchcv for hyperparameter tuning and K Fold Cross Validation for model evaluation
2. Wrote a Python Flask Server that uses the saved model to serve http requsts.
3. Built web page using HTML, CSS and JavaScript that allows user to enter values for estimating house price and it will call Python Flask server to retrieve the predicted price.

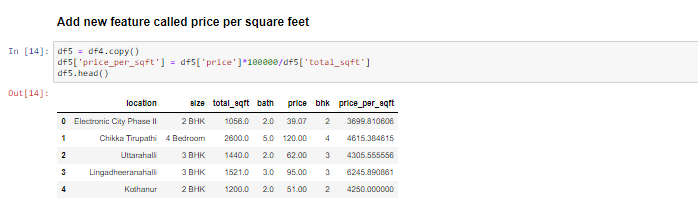
Below is the screenshot of the webpage built:

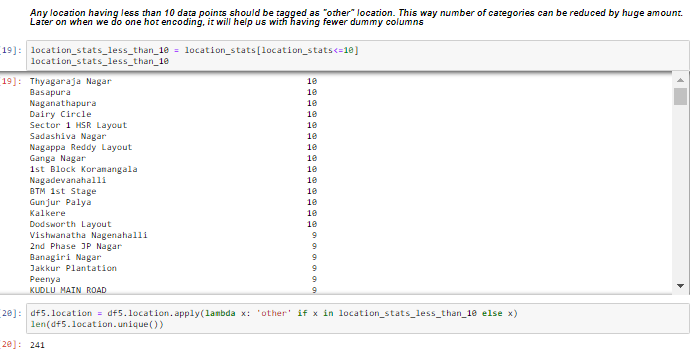


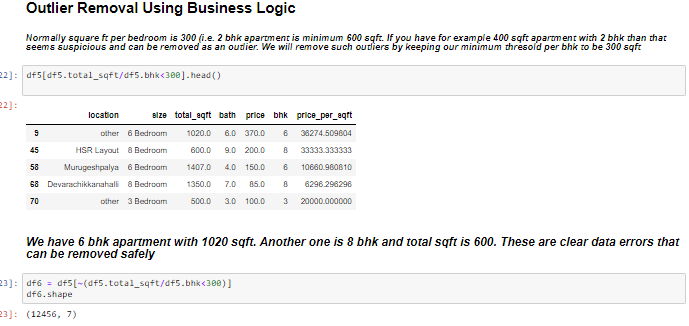
**Snippets of the work done on various stages:**

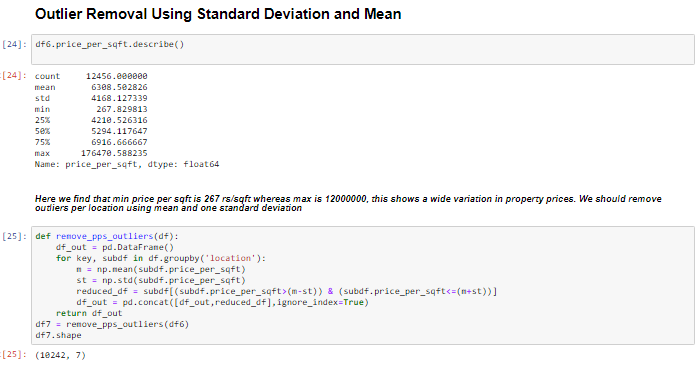
****

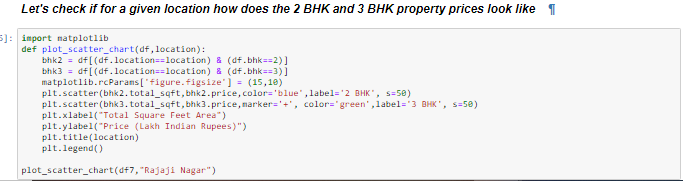
****

****

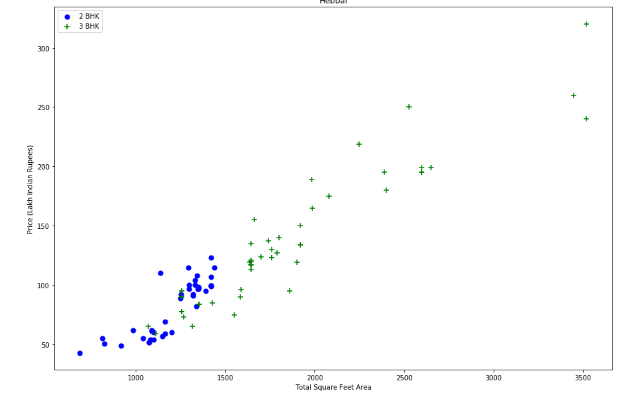
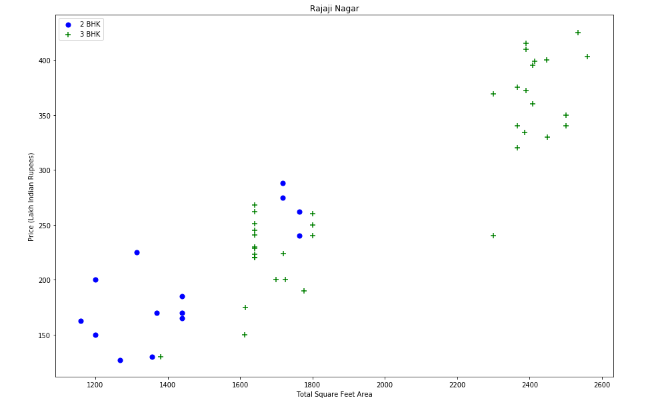
****

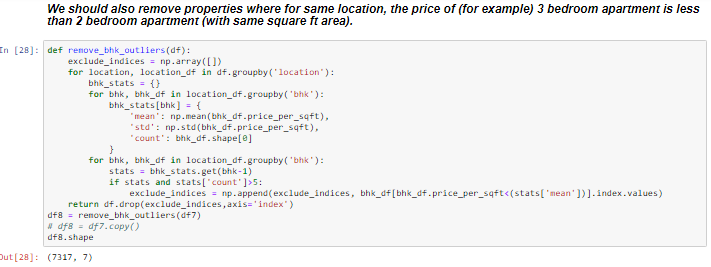
****

****

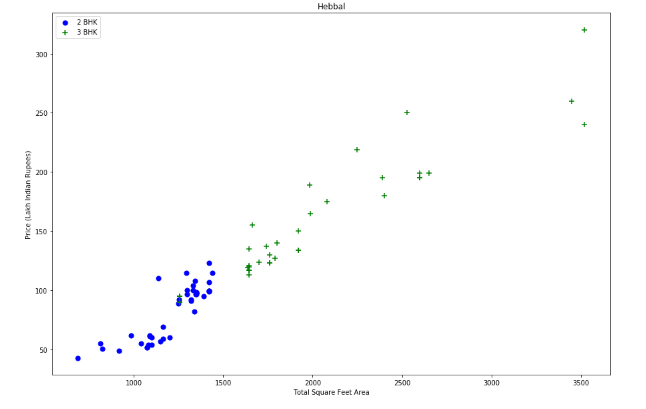
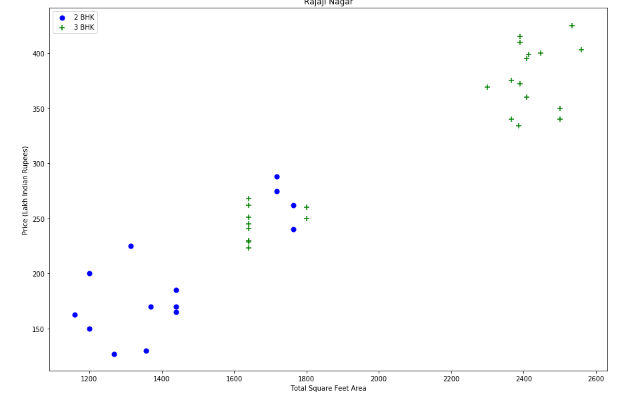
****

**Before outlier removal:**

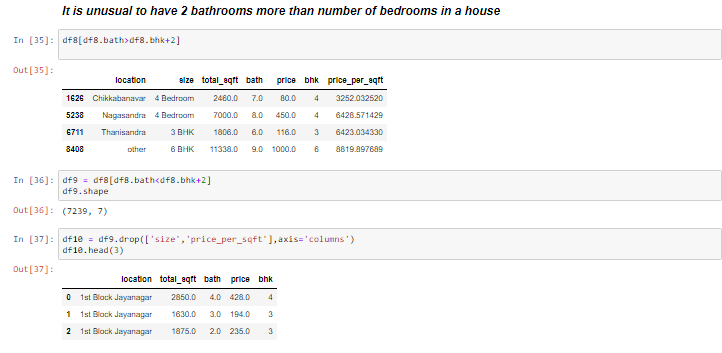
****

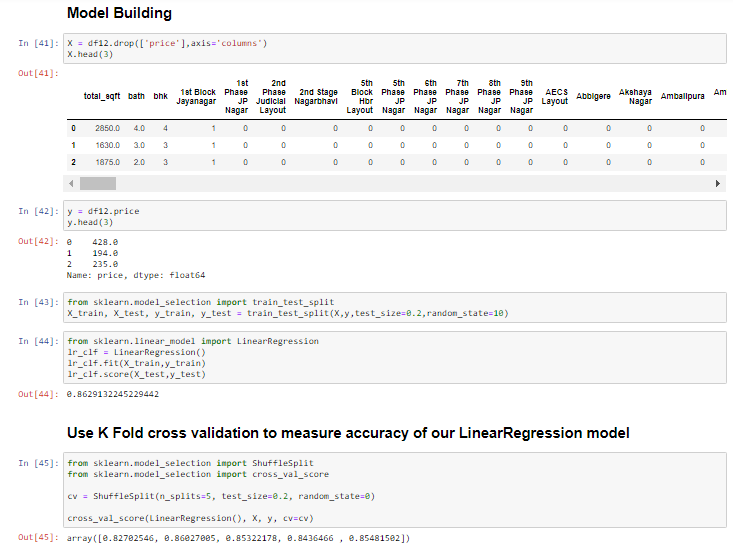
****

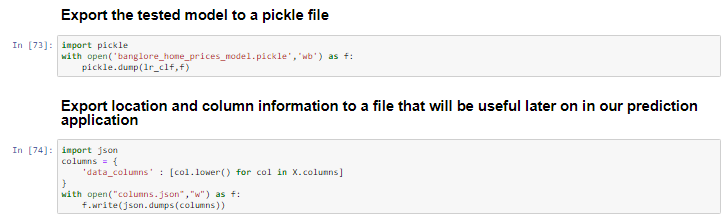
**After outlier removal:**

****

**Outlier removal using bathroom feature:**

****

****

****