



Housing Price Prediction Project



Valary Thairu



Overview

Housing has been the long-term focus of urban development not only to the medium level economies but to the high economies too. House price prediction is a very tedious task as there are many things to consider when buying a house, things like floors, number of bedrooms, number of bathrooms, square footage living of a home space etc.

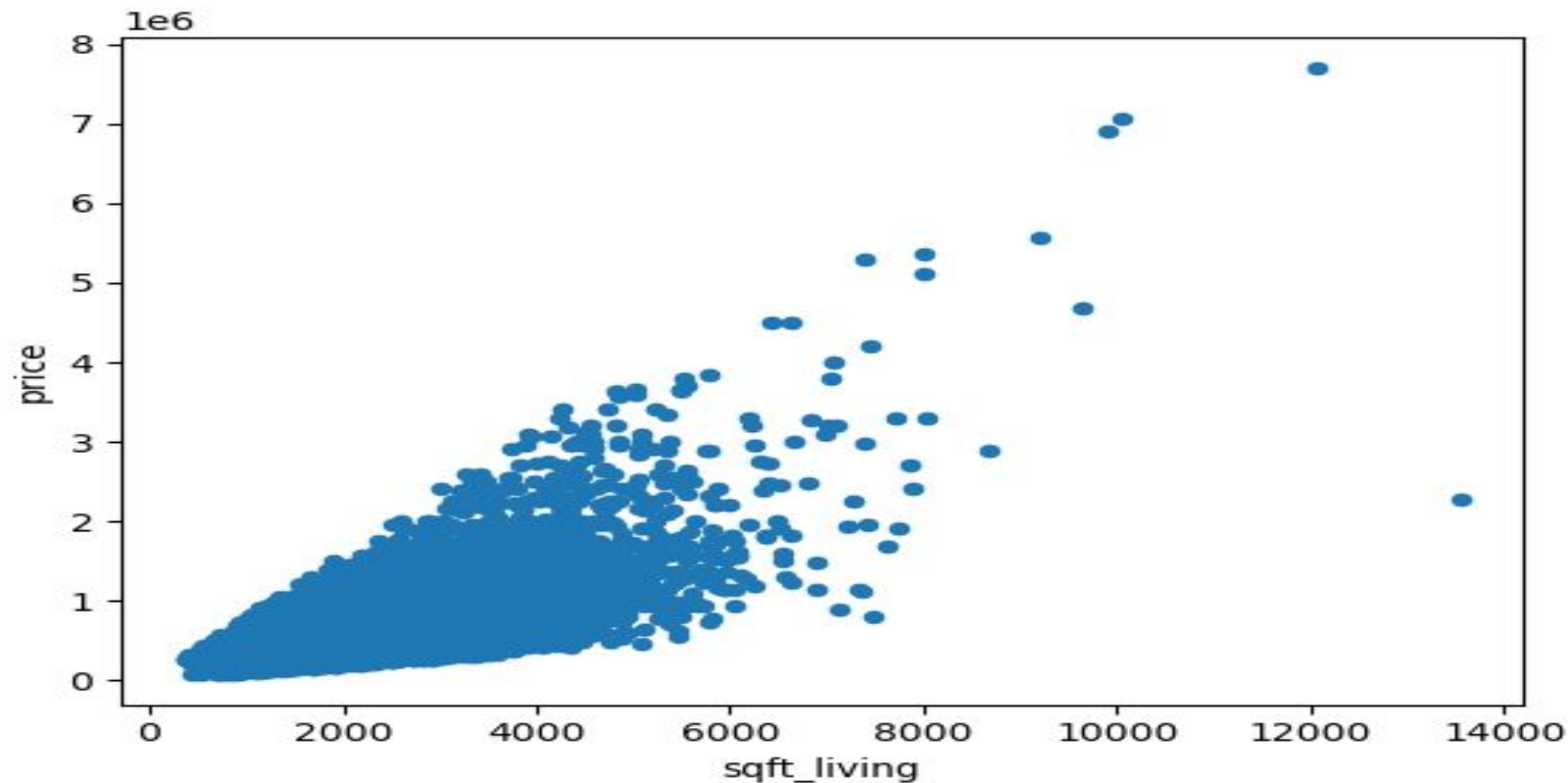
A real estate agency has hired me as a data scientist to find out the factors that affect the sale price of houses. I have the Kings County Home Sales dataset that contains data about features of houses. I used the data from the Kings County Home Sales dataset to investigate the factors considered when buying a house and make recommendations according to my findings.

Business and Data Understanding

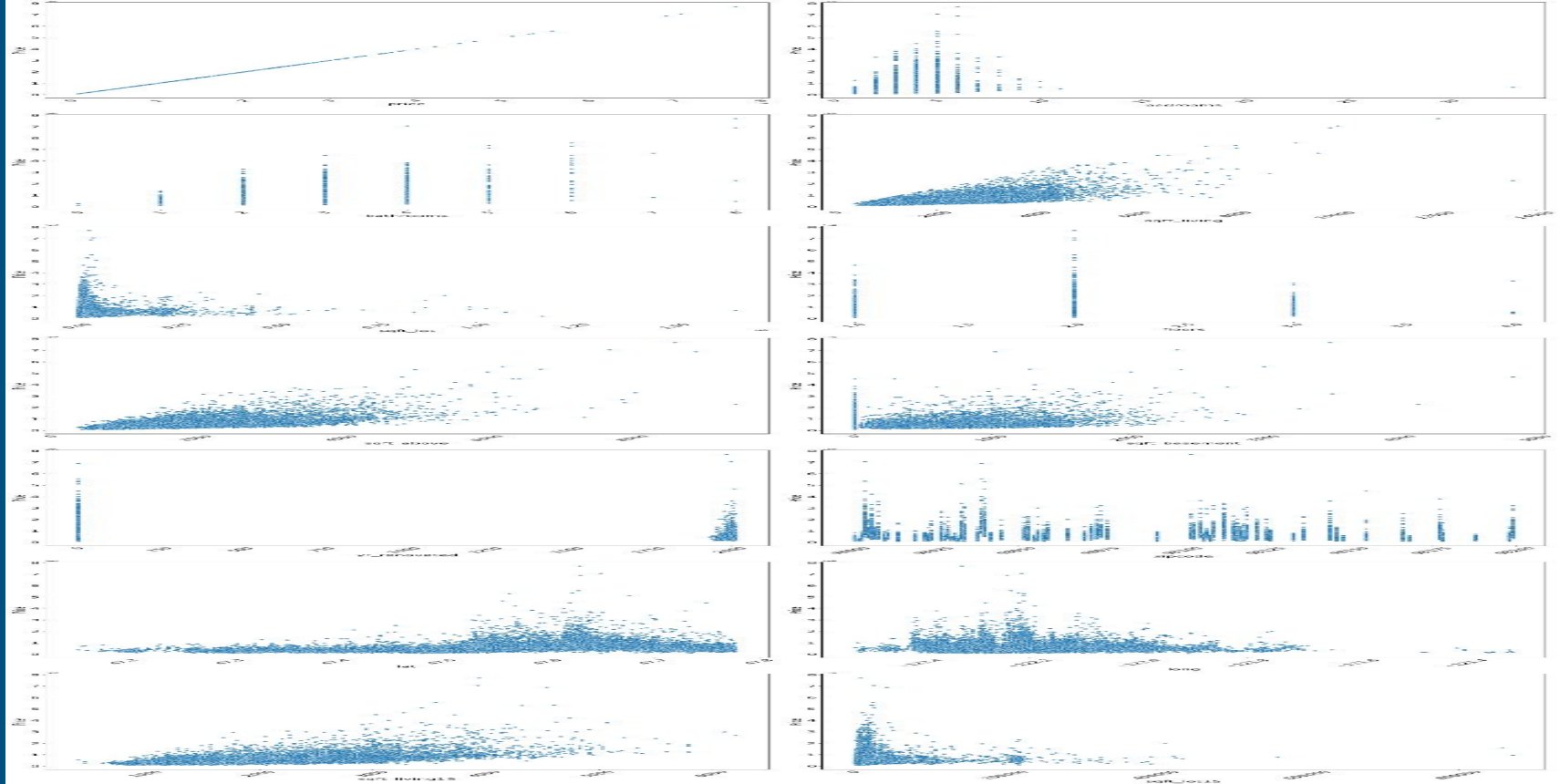
- To identify that factors that affect the sales prices of house in North-Western county.
- The data i used was from the Kings County Homes Sales dataset, it contains a range of values that provide details of different factors of a house including its price.
- Analysis of this dataset contributes to success of house pricing. In this analysis I found the relationship between the different factors of their houses and their sale prices.

Modelling

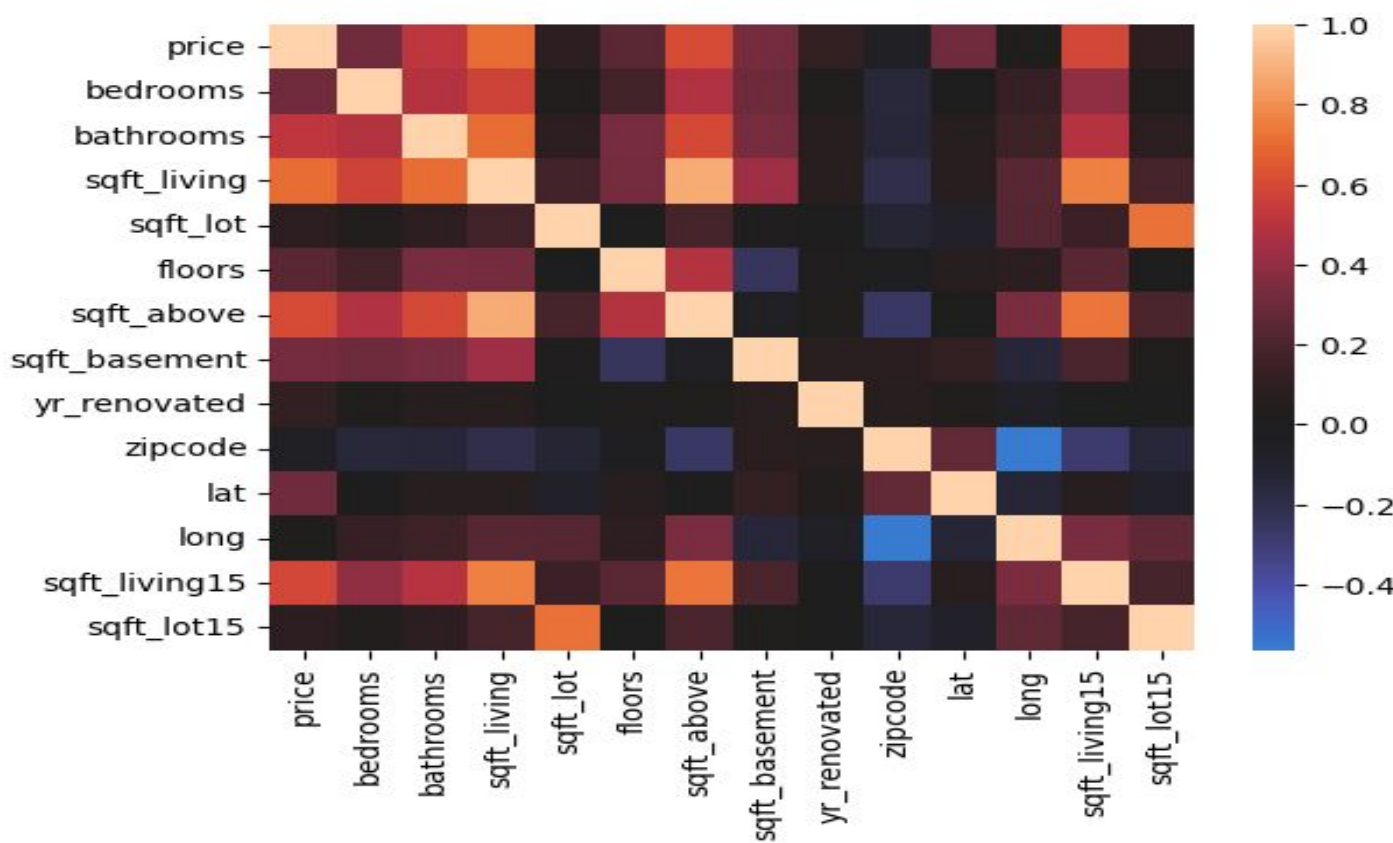
- I constructed a multiple linear regression model, to be able to find the relationships between the factors affecting house's sales price and house's sales price.
- Creating this model helped identify that the factor square foot living space of a home('sqft_living') has the highest positive effect on house's sale prices in North-Western County.
- I was also able to find the relationships between all the variables in the Kings County Home Sales dataset and house's sale price which is also in the dataset.



This displays the relationship between 'sqft_living' and house's sale price



This displays the relationships between each variable and house's sale price



This displays the correlation between variables of the dataset.

Regression results

- The model I created had some interesting results, it explained 59.1% of the variation in house's sale price. This means that these factors do have an effect on house's sale price in North-Western County.
- The model proved to be statistically significant which proves that these variables in the dataset do affect the house's sale price in North-Western County.
- Square foot living space in the home and square footage of the basement which are two factors in the dataset have proved to have the highest positive effect on house's sale price in North-Western County.

Recommendations

1. The real estate agency should increase square foot living space in home as it increases sales prices of the homes, they should also focus on increasing square foot of the basement. These two factors have the highest positive effect on houses sale price hence are very important.
2. As according to my findings, homeowners should also focus on the number of bathrooms, square-foot living space in the home, Square footage of the house apart from the basement, Square footage of the basement, and number of floors when they want to do home renovations as these factors affect the price of houses positively.

Thank you!



This project was a delight!