House Price Prediction Model For King County

Business overview.

- Real estate industry is a complex business that is pushed by various factors that affect the prices.
- Due to these factors, giving a valuation to a property is a challenging task.
- The project aims to explore King county house data set and factor in some of the factors provided, which will help to derive a model that will help to predict the price of a house.
- Overall the model will help buyers and sellers to identify factors that will increase the price of the house.

Challanges Facing Real Estate in Kings County

- Lack of affordable housing due to the rapid growth in population due to more people moving to the area. Demand for houses is going higher.
- Property owners or developers might give false information about the house grade to drive the prices up. The rating may also not be accurate due to other factors.
- Scarcity of available units for residence within Kings County has driven up the cost of house units between the various grades ie: average, good, excellent and luxurious.
- The disparities in the housing market play out on a sub regional basis within King County Properties located in desirable areas, such as waterfront or downtown areas are limited and highly valued.
- Limited supply of land in the region, particularly in desirable areas close to job centers and transportation. The scarcity of land can limit the number of available units and drive up prices.

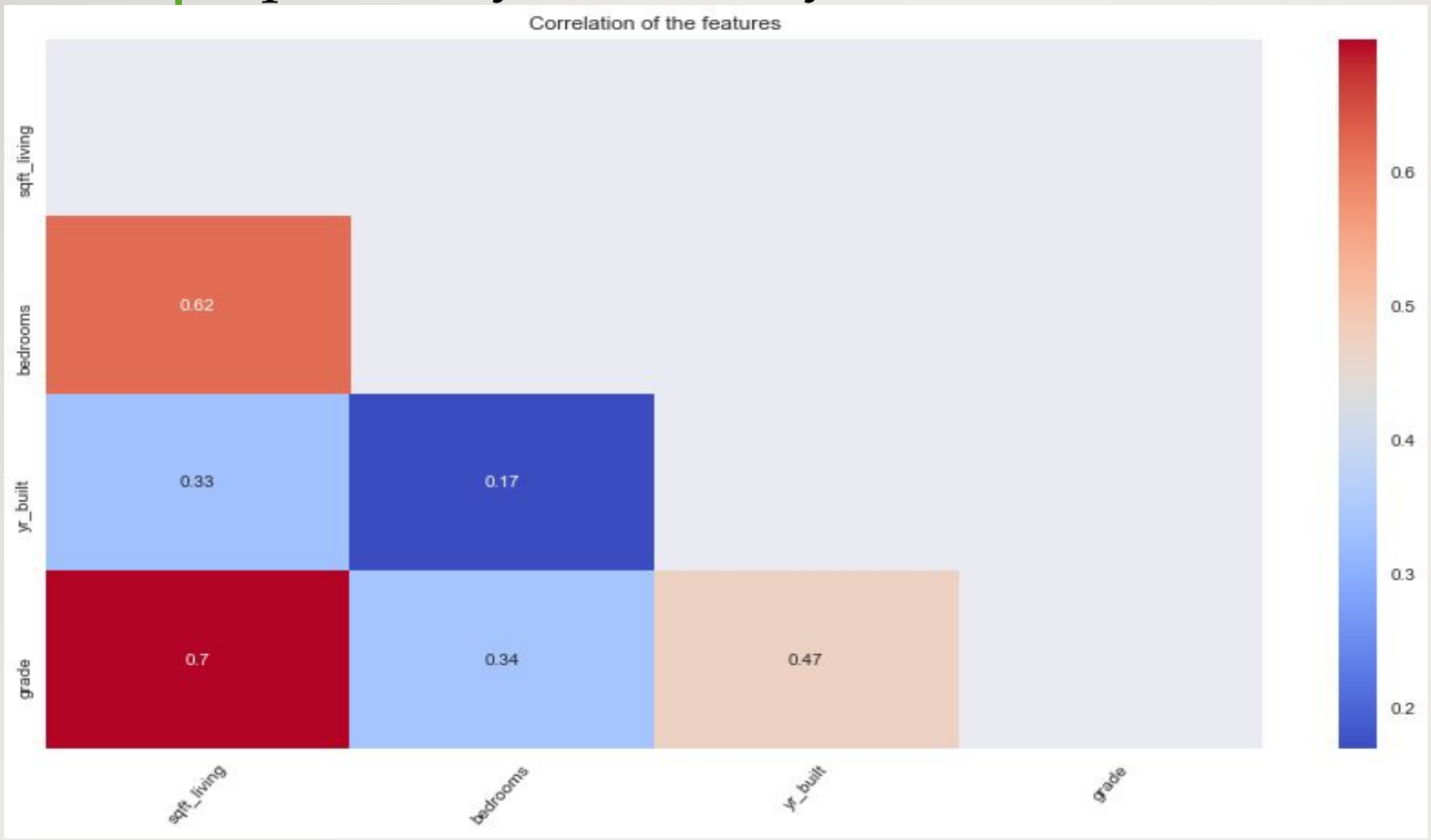
Objectives

- Determining the relationship between the square footage of the house and the sale price of the houses.
- Examine the relationship between the overall grade of the house and the sale price of the houses.
- Explore the relationship between the year built and the sale price of the houses.
- Investigate the relationship between the number of bedrooms and the sale price of the houses.

Source of data set.

- The data came from house sale prices in King County, Washington between the year (2014 to 2015.)
- The data had 18,678 showing the number of house sold and 21 features.

Exploratory data analysis



- Relationship between key features.

Models

First Model

Baseline

- Analysed how the size of living space affected price.
- Living space size could only predict the price with a 36.8 % accuracy.
- Made updates to the data but there was no improvement in the accuracy.

Second Model

Living space and Bedrooms

- We checked how related the two features are to price.
- This still did not give us a good value for prediction.
- It could only account for 38.1% accuracy.

Third Model

Grades

- We checked if grade could have any influence.
- It had a small effect on the price.
- It had a 40% accuracy thus a little better than the previous models.

Final Model

Combined features.

- We combined this features to identify their effect.
- This proved to be better as it has a 56.8% accuracy.
- However, it did not meet the target.

CONCLUSION

- We can conclude that there is a strong linear relationship between price and square foot living and square foot above.
- We can conclude that as the grade of the house increases, its price also increases and this will determine the selling price of the house.
- We can conclude that the number of bedrooms has a weak positive correlation with the house price

Recommendations

- Build houses that have a high grade rating.
- Target houses with a big living square footage.
- The bedrooms are also a factor in the price so they can look for buildings with at least 4 bedrooms

Authors

- Peter Onsomu
- James Wainaina
- Kelvin Muriuki
- Paul Machau
- Robert Mbugua

https://trello.com/b/VXmOMtwL/project