

KC_HOUSING

Presented by:
Nduta Tabitha Elsie

PROBLEM STATEMENT

- The main purpose of the project is to develop a pricing algorithm to help set/predict a target price for the houses in the agency. The goal is to save the company some time and to help ensure consistency in pricing with consideration of the attributes the house has to offer.
- Also identify the variable that impact the price of the house the most.

LEAD QUESTIONS

- What's the most common prices ?
- Does the number of bathrooms and price correlate ?
- Does the grade affect the price ?
- What are the correlations of price and the other attributes ?
- Are there any categorical data?

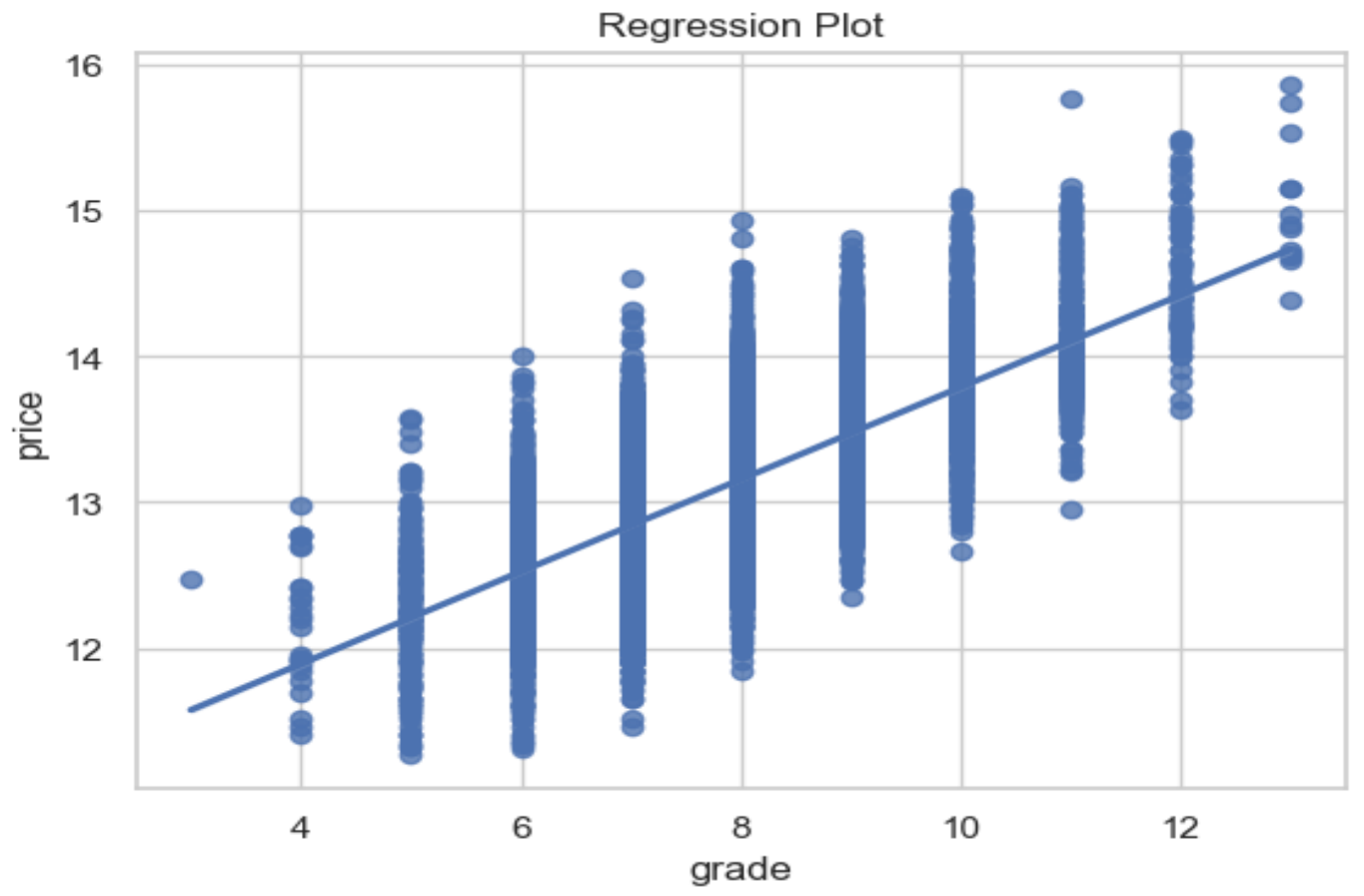
PRICES OF MOST HOUSES

- Most houses prices seems to be ranging at 13.



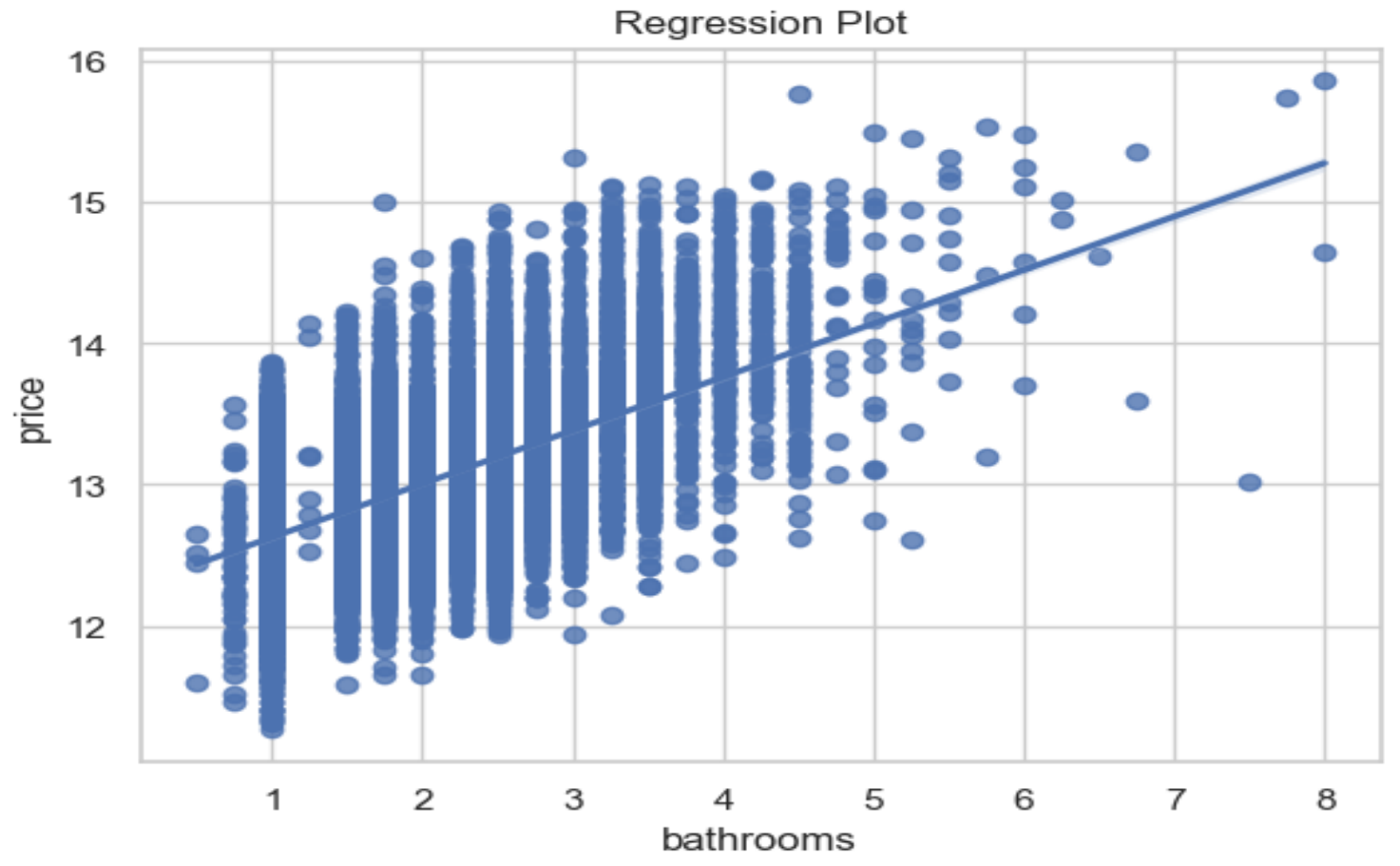
EFFECT OF GRADE ON PRICE

- High grade impacts positively on the price of a house.



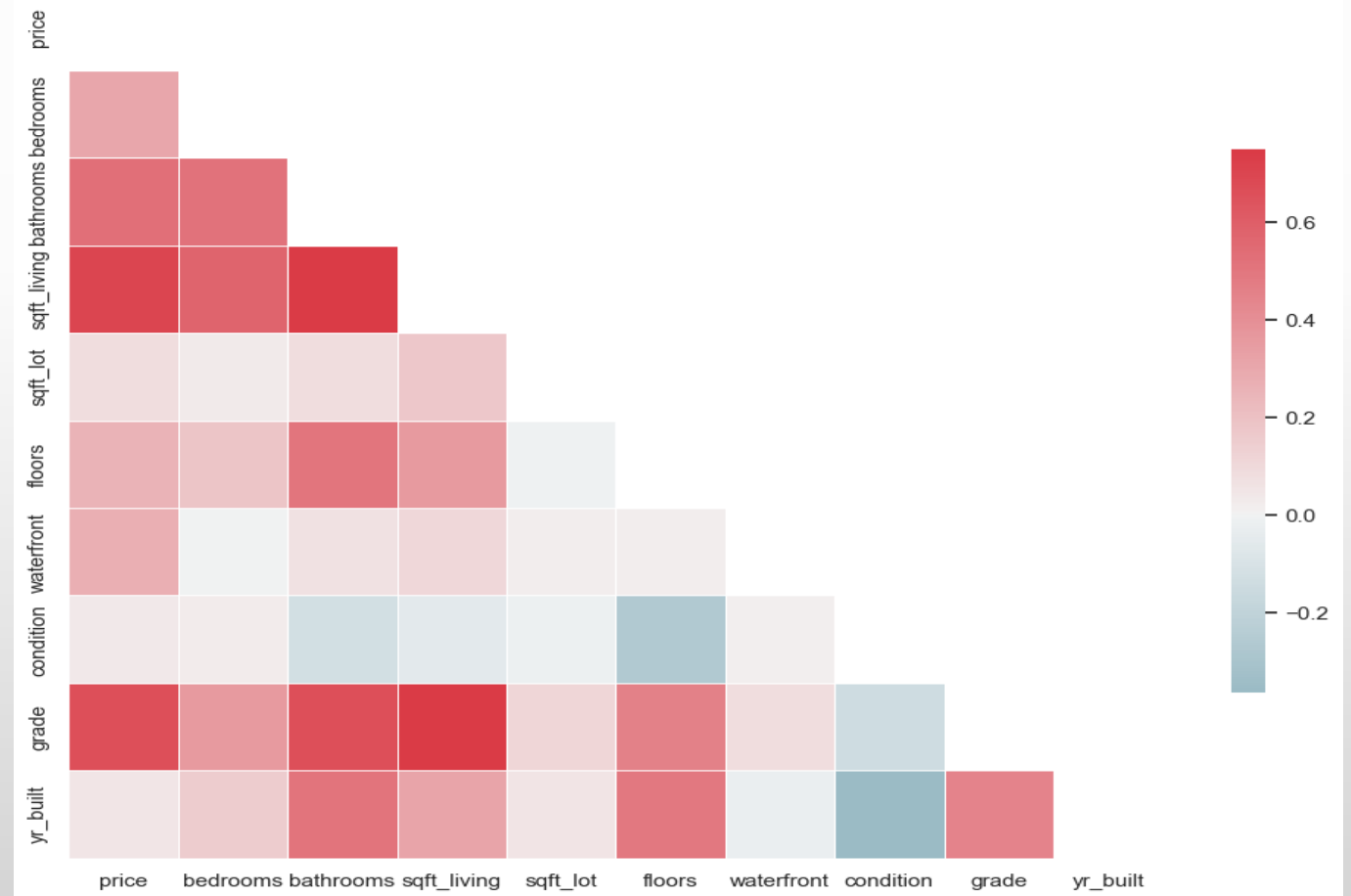
EFFECT OF BATHROOMS ON PRICE

- Increase of number of bathrooms impacts positively on the price of a house.



CORRELATION OF THE VARIABLES

- The dark red shades show high correlation.
- Sqft living vs price
- Sqft living vs bathroom
- Price vs grade



MODEL FITTING

The model is a good fit since the training data and test data do not over fit or underfit



INTERPRETING THE DATA

Train R-squared: 0.6430815412512239 Test R-squared: 0.644954224982312

- The Rsqd of both the training data and testing data seem to be at close range this is verified by the graph of the overfitting of both sets having a good fit. Meaning the model is efficient. To increase the efficiency of the pricing the should be iterated as seen from the Rsqd which is higher ,meaning the relation of the attributed to the pricing higher ,up to 3 iterations.