Ames Iowa Housing Prices

By: David Adoni

Problem Statement

I have been tasked with creating a machine learning model based on the Ames Housing Dataset. This model will predict the price of a property at sale.

The Ames Housing Dataset is contains over 70 columns of different features relating to houses, and many of variables needed to be modified for the data to be usable.

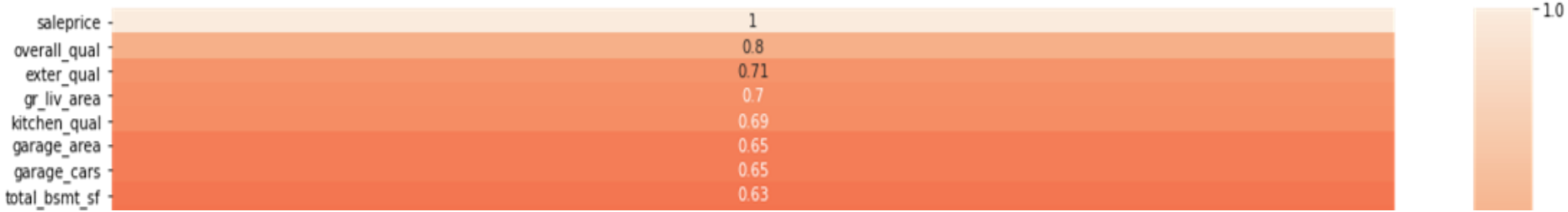
Steps Taken

1. Get the data
2. Clean the data
3. Run EDA
4. Create models

Variables with Correlation to Price

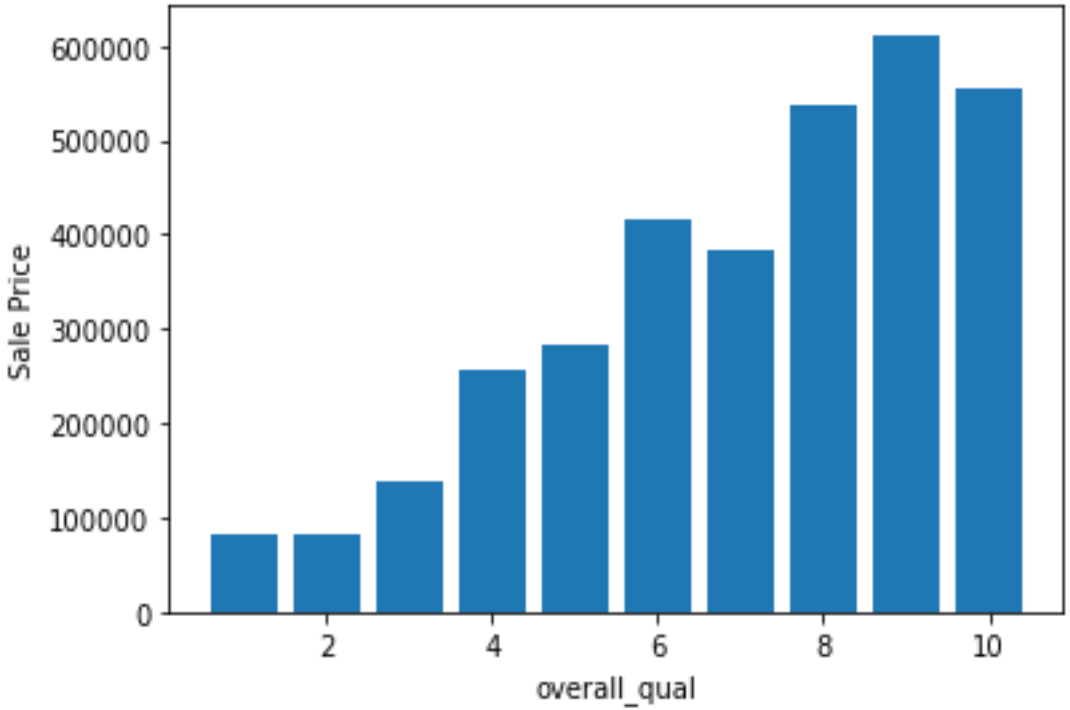
With respect to the Sale Price the top 5 factors that increase price are:

1. Overall Quality
2. External Quality
3. Garage Living area
4. Kitchen Quality
5. Garage Area/Cars



Overall Quality and Sale Price

Overall Quality clearly indicates the better the quality, the more homebuyers are willing to pay for the home



Model Results

Model Type	Training Score	Testing Score
Linear Regression	0.8143	0.8689
Lasso	0.813	0.8711
Ridge	0.8102	0.8697

Conclusions

Although the model that was created is not perfect, it can use new data specific to the Ames Housing set and find prices with a chance of being $32,910 off of the actual price.

The variables that correlated the most are key components to these predictions