Jim's Real Estate House Price Model

Ricky Wong
June 10, 2022

Summary

- Linear regression model of house prices to advise clients if renovating improves the price of their property

- Find the best predictors that affect the house price

Outline

- Business Understanding
- Data & Methods
- Results
- Assumption checks
- Accuracy
- Conclusion

Business Understanding

- Real Estate agents to provide advice about house prices and renovations

- Use house price data to create linear regression model

Data & Methods

- Check for multicollinearity to remove correlated predictors

- Normalise continuous data for linear model

- Categorical data transformed using one-hot encoding

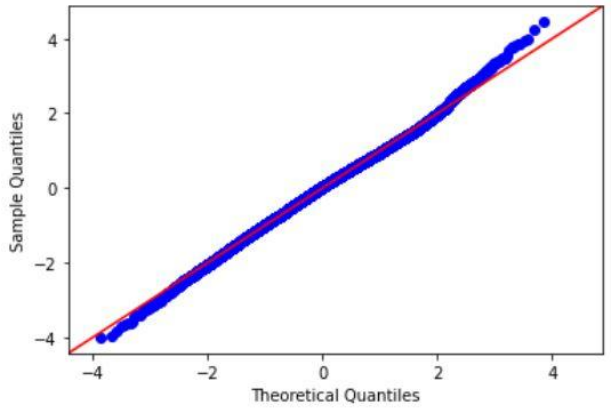
Results - Model 1 & Model 2

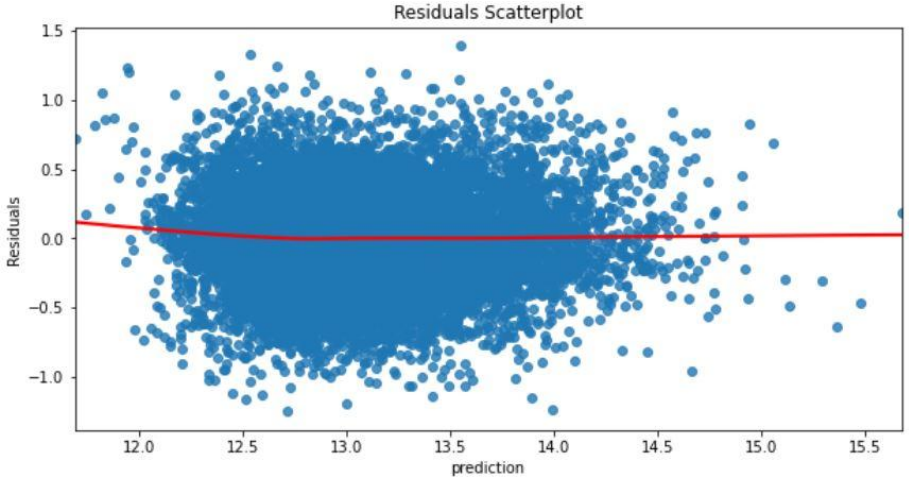
- Used all the available predictors
- R-Squared value of 0.633
- std_above and conditions had p-values greater than 0.05
- Skewness & Kurtosis values high

- Removing the two predictors have lowered our R-Squared score slightly
- Skew and Kurtosis is still quite high
- Doing a QQ-plot shows it is not normal

Results - Model 3

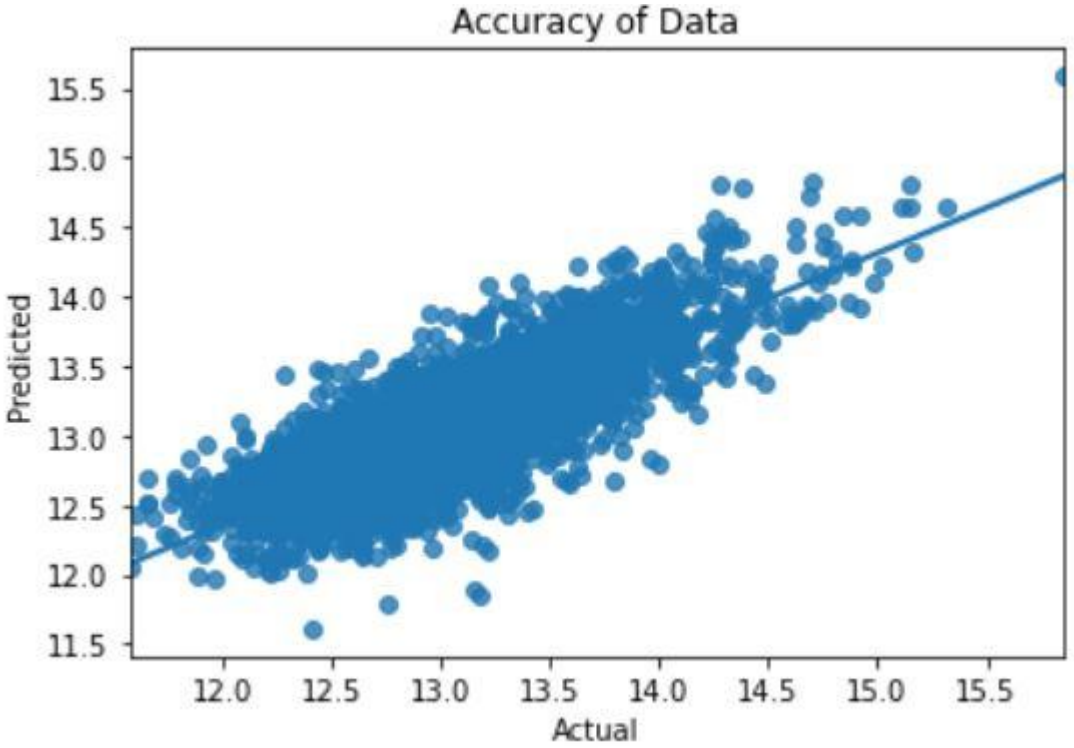
- Log transformation on price improved the distribution
- Homoscedasticity check
- R-Squared value increased. 64.6% of the variance is explained by the model.





Accuracy

- Train and test MSE very similar
- Accuracy of the model is 63.42%
- Good correlation between actual and predicted results



Conclusion

- 64.6% of the variance is explained by the model

- Living area has strongest relationship with price

- Renovation data may be insufficient for model

Thank you!

Email: wong_ricky@hotmail.com

GitHub: @wong-ricky