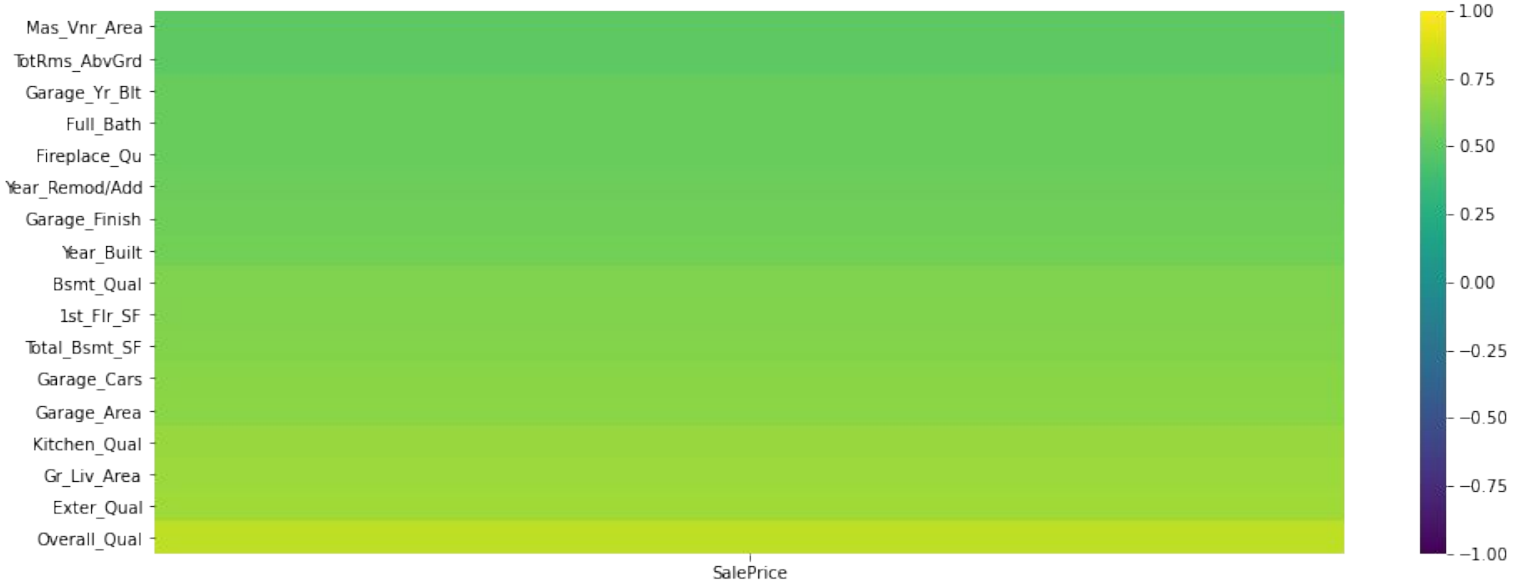
Inferring or Predicting? What's More Valuable For a Housing Model?

By: Adam Klesc

Date: 11/4/21

Why is it important? And why not both?

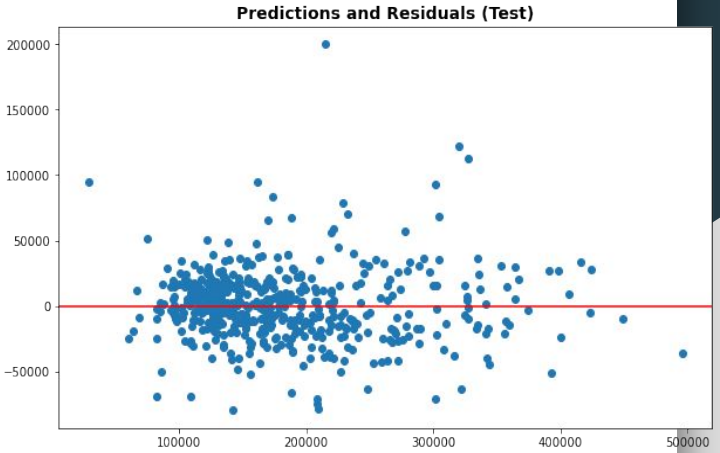
Often, our features make it impossible for both to be true. When that happens, what should we prioritize? This can be either due to collinearity, autocorrelation, or a number of different factors related to the violation of Linear Regression assumptions.

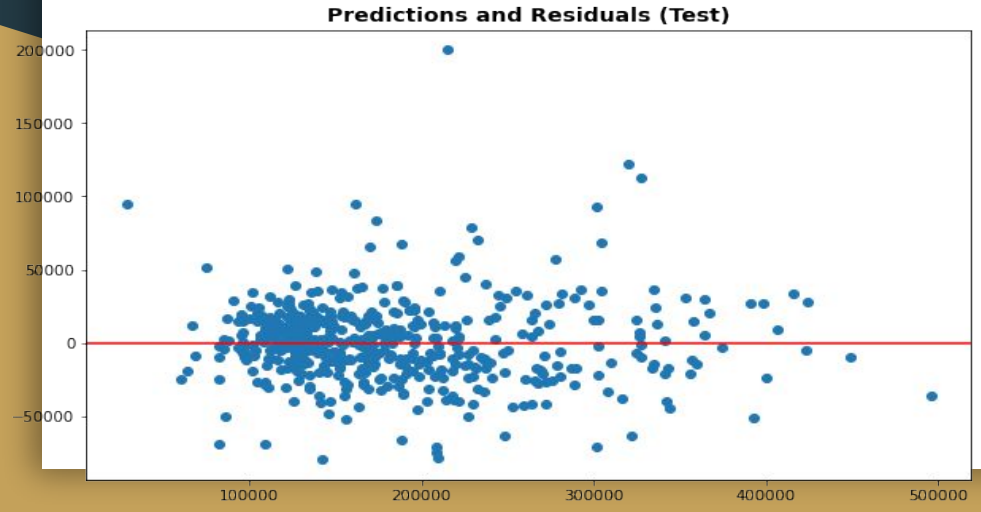


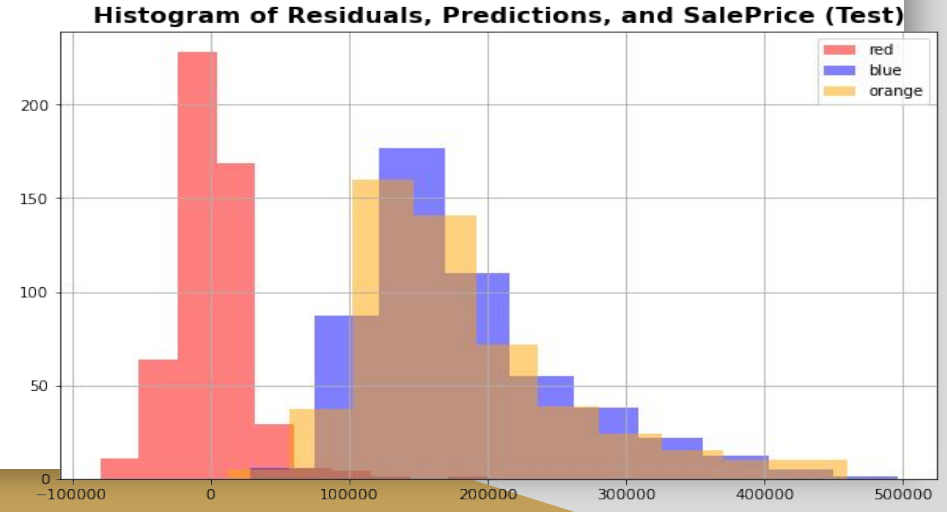
What are the 5 Assumptions of a Linear Regression Model?

- Linear relationship
- Multivariate normality
- No or little multicollinearity
- No autocorrelation
- Homoscedasticity

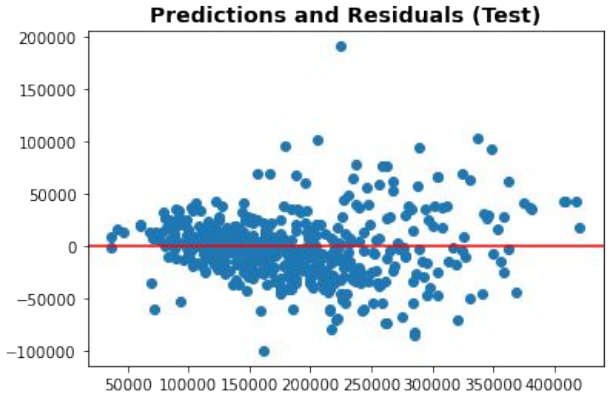
A Model That Predicts Well

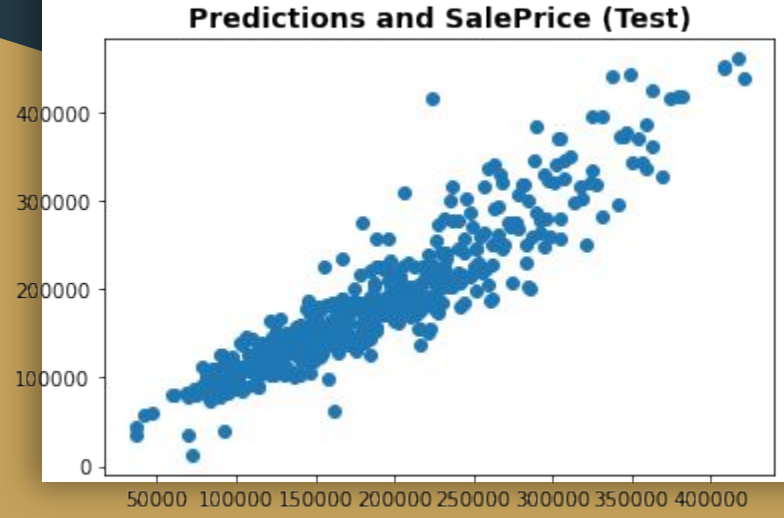


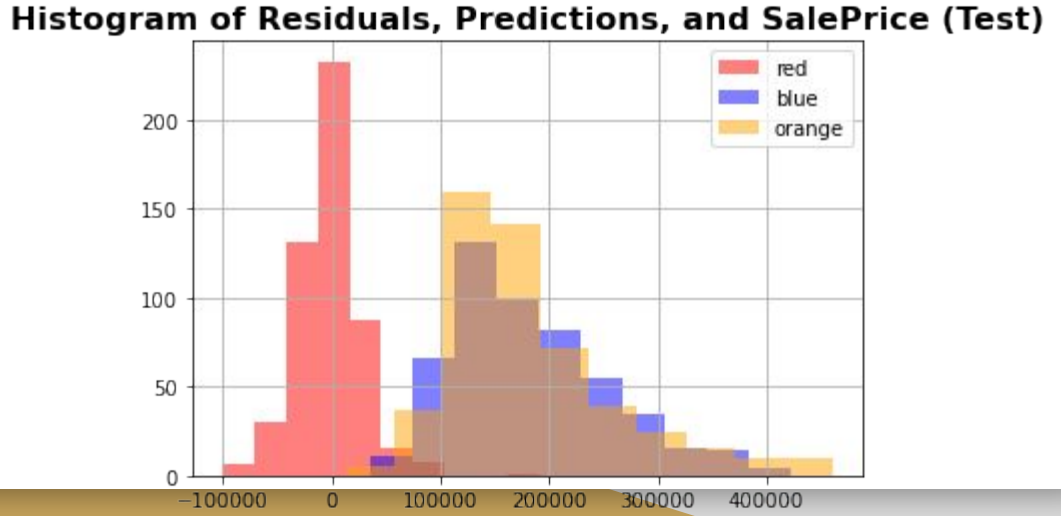




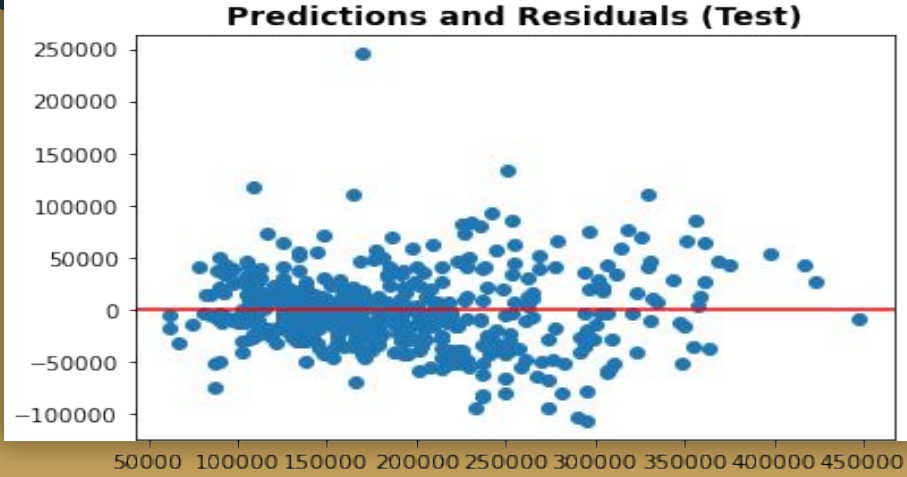
A Model That Balances Both

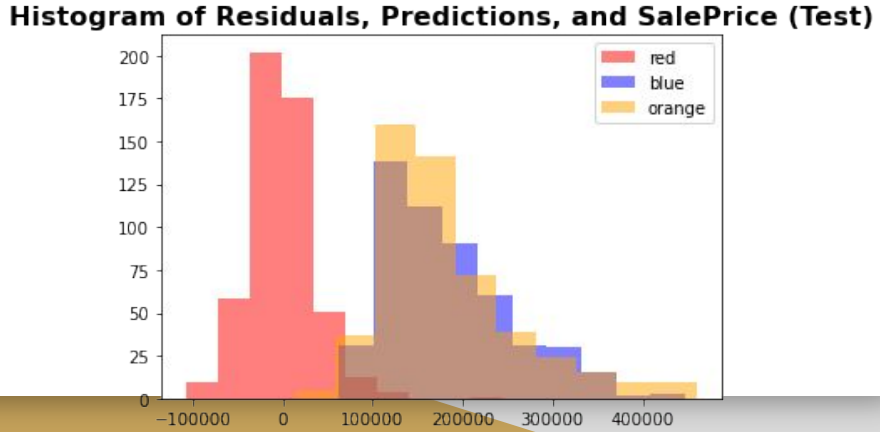






A Model That Diligently Adheres to The Assumptions





In Conclusion

- Do not be dogmatic. Be flexible
- Do not feign ignorance to the assumptions
- Do not be afraid to deviate from the assumptions to achieve a balance
- Inferences > Predictive value
- “All models are wrong, but some are useful”