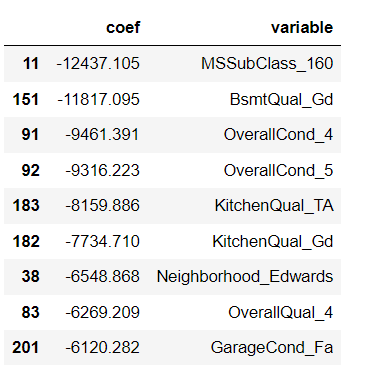
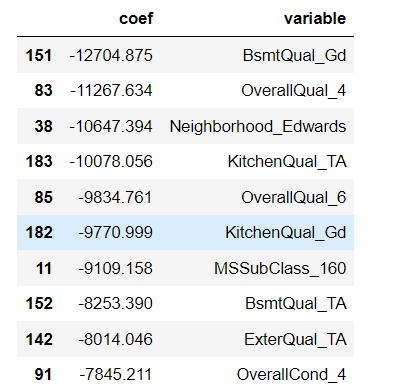
1. What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

The optimal value of alpha is around 50 for lasso and 4 for ridge regression.

Top 10 lasso variables after doubling alpha [ This is before variable selection]



Top 10 Ridge variables after doubling alpha



1. You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

Lasso Regression. As it makes coefficient of more number of variables to Zero, the model is less complex.

1. After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?
2. How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

Looking at similar test and train errors distributions[ Actuals vs predicted plots],Test vs Train R Squared, we can cnfirm that our models have low bias and low variance and are robust and generalizable. The similar rsquared across different types of the models confirm the same.