

### Question 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?

Answer: In Lasso Regression  $\alpha=0.5$

In Ridge Regression  $\alpha=10$

If we double the value of Alpha in Lasso, the model will be more regularized and it can lead some coefficient to zero. Hence it can lead model to underfit.

In Ridge also the model will move towards underfit, but the coefficient tilt towards zero but not exactly zero.

OverallQual

### Question 2

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?

Answer: In Lasso regression when we tune with hyperparameter some of the coefficients value become exactly zero and it helps to eliminate the predictor variables that are not contributing to the predicting the output variable.

Whereas in case of Ridge regression even on tuning we cannot make coefficients equal to zero and we have to still carry the predictors that are not significant.

So we can choose Lasso over Ridge

### Question 3

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?

### Question 4

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?

Answer: We can make our model robust by applying regularization and cross validation approach. By using cross validation we generate multiple models using k-fold approach

and select the model which perform best. Then we validate our model to the unseen test data which will show better results.