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: Optimal value of alpha for ridge(10) and lasso(0.007). By Doubling the alpha of ridge some coefficient value for feature will decrease(previous value also less) and some will increase (Previous value having higher) , and cost function value will be higher than previous. In Lasso , Doubling of alpha value Feature will reduce in count(14 feature will be zero) and accuracy will also decrease.

Most Predictor: Ridge-> RoofMatl_CompShg

Lasso-> 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: Based on alpha values, we need to find out the best model performance. So by choosing alpha different values finalize the best alpha values which predict the best model performance where overfitting will be less.

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?

Answer:

1:OverallQual

2:GarageArea

3:Fireplaces

4:TotalBsmtSF

5: OverallCond

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: A **model** is considered to be **robust** if its output dependent variable is consistently accurate even if one or more of the input independent variables (features) or assumptions are drastically changed due to unforeseen circumstances.