**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 1

The optimal value of Alpha for ridge and lasso are:

• Ridge Alpha value = 20

• Lasso Alpha Value = 0.001

So new alpha value for Ridge = 20 following are thepredictors –

Lasso:

Variable Coeff

0 constant 11.828

13 GrLivArea 0.127

210 SaleCondition\_Partial 0.116

50 Neighborhood\_Crawfor 0.095

4 OverallQual 0.084

209 SaleCondition\_Normal 0.060

5 OverallCond 0.057

31 MSZoning\_RL 0.050

29 MSZoning\_FV 0.046

70 Condition1\_Norm 0.044

Ridge:

Variable Coeff

0 constant 11.816

50 Neighborhood\_Crawfor 0.091

13 GrLivArea 0.074

4 OverallQual 0.072

209 SaleCondition\_Normal 0.063

210 SaleCondition\_Partial 0.062

66 Neighborhood\_StoneBr 0.058

31 MSZoning\_RL 0.056

29 MSZoning\_FV 0.056

5 OverallCond 0.054

**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 2

Lasso Regression is more preferable as Lasso helps prevent overfitting by penalizing the absolute values of the regression coefficients, leading to feature selection and a simpler, more interpretable model.

**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 3

Five most predictor variables are –

LotFrontage

LotArea

OverallCond

MasVnrArea

BsmtFinSF1

**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 4

The importance of finding a balance between model accuracy and complexity to ensure robustness and generalizability. Regularization techniques like Ridge and Lasso Regression play a crucial role in achieving this balance by controlling the model's complexity and addressing the bias-variance trade-off. This approach is fundamental for building models that perform well on both training and new, unseen datasets.