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?

- Optimal value of alpha for Ridge = 2
- Optimal value of alpha for Lasso = 0.0001

- When doubled alpha for Ridge, there is no significant change in the metrics.
- When double alpha for Lasso, R2 score specially for testing dataset dropped slightly.

- There is no change in the most important variables when the change is implemented. Alpha was 0.0001 when double to 0.0002 No change. When 0.0004 there is change.

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?

I must take alpha as “0.0001”. the metrics associated with alpha is 0.0001 is better. R2 score was bit higher for train and test datasets. Also, residual sum of squares, mean square error and root mean square error were low.

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?

1. **1stFlrSF**
2. **2ndFlrSF**
3. **BsmtFinSF1**
4. **GarageArea**
5. **Neighborhood_Crawfor**

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?

Robust model is a model that it performs well on unseen data. That means it can explain most of the variance in the training set and it can generalize what it had learned from the training set on the test set.

There is a trade-off between bias and variance, variance is high when the model had learned extensively on the training set that can't generalise on the test set. Bias on the other hand, is a very simple model that it didn't learn anything from the training set. A robust model is somewhere between the two.