Q1 : 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?

```
Ans: The Optimal value of alpha for Ridge is 1.0 and for Lasso it is 100. Hence
doubling it would result it in 2 and 200. In case of doubling the parameters few
selected variables are not present . New parameters are
```

FullBath,
BsmtFinSF1
BsmtUnfSF
GarageArea
HalfBath

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?

▪ The optimum lambda value in case of Ridge and Lasso are:

• Ridge Regression – 1.0

• Lasso Regression – 100

▪ The R Squared value is similar in both case , based on the RMSE value lasso model is chosen. And Lasso also helps in feature elimination. Hence Lasso has a better advantage over Ridge and should be used as the final 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?

Ans: The five most important predictor variables in the current lasso model were

1stFlrSF, OverallQual, 2ndFlrSF, OverallCond, LotArea
After dropping these variables.

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?

According to Occam's Razor, If given two models that show similar 'behaviour' in the finite training
or test data, we should pick the one that makes fewer on the test data due to following reasons:-

1. Simpler models are also more robust in nature.
2. Simpler models require very less training samples for effective training than the more complex ones and hence are easier to train.
3. Simpler models are usually more generic in nature and are more widely applicable to different types of datasets .

Regularization can be used to make the model simpler. Regularization helps to balance between keeping the model simple and not making it too naïve to be of any use. In other terms, it ensures the model doesn't under or overfit. For regression, regularization involves adding a regularization term to the cost function that adds up the absolute values or the squares of the parameters of the model.

To create a simple model leads to Bias-Variance Trade-off:

- A complex model will need to change for every negligible change made in the dataset and is
considered very unstable, unreliable and extremely sensitive to any changes in the training
data.
- A simple model on the other hand, brings out some pattern followed by the data points given therefore making it highly unlikely to change even if more points are added or removed.

'Bias' quantifies how accurate is the model likely to be on test data. A complex model can do
an accurate job prediction provided there is enough training data. Models that are too naïve,
for e.g., one that gives same answer to all test inputs and makes no discrimination whatsoever has a very large bias as its expected error across all test inputs are very high.

‘Variance’ refers to the degree of changes in the model itself with respect to changes in the training data.

Thus accuracy of the model can be maintained by keeping the balance between Bias and Variance as it minimizes the total error .