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?

Ans:

- Optimal value of alpha in Ridge regression model is **10**
- Optimal value of alpha in Lasso regression model is **0.001**

Most important predictor after changing the alpha value are:

- GrLivArea
- OverallQual_8
- OverallQual_9
- Functional_Typ
- Neighborhood_Crawfor
- Exterior1st_BrkFace

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?

Ans:

There are more features available, Lasso can reduce it by feature selection. So Lasso will be better than ridge regression.

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:

After dropping top 5 lasso predictor, following are the new top predictor

- Condition2_PosA
- OverallCond_9
- SaleType_ConLD
- 2ndFlrSF
- OverallCond_8

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?

Ans:

- A good model should fit well for unseen data and the used data for model building
- Model should not overfit. When overfitting happens test data will not fit properly.
- Model should be simple.
- High accuracy models are complex in nature. So we have to decrease variance