**Problem Statement - Part II**

**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 1**

Optimal value for Lasso : 0.001

Optimal value for Ridge : 50

As we increase the value of the alpha the values of the coeff in Ridge move towards 0 and in Lasso they become 0.

**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 2**

I have chosen Lasso as the difference between the R2 score of the train and test data isn’t much and Lasso has reduced the no. of features resulting in simpler 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 3**

Below are the 5 most important predictor variables:

OverallQual,

GrLivArea,

GarageArea,

Fireplaces,

YearRemodAdd

**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 4**

A Simple model (reducing the features) makes the model robust and generalisable. This results in Bias-Variance trade-off.

This reduces the accuracy of the model but accuracy of train & test set should be similar.