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)The Optimal value of alpha is 1.0 in Ridge regression and the optimum value of alpha is 10 in Lasso regression.

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)As the r2_score of lasso is slightly higher than ridge for the test dataset, we will choose Lasso 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)The variables are:1stFlrSF,GrLivArea,Street_Pave,RoofMatl_Metal,RoofStyle_Shed

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) We can try to use the model testing error and training error to be as close as possible. Probably then we can say the model is robust.