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. For Ridge the alpha value is 0.1 and For Lasso alpha value we got is 0.0001. After taking the double of alpha value we have seen that the r2 score for both have been increased . Initially the r2 score for Ridge was icreased from 74% to 79% ,but test r2 score for Lasso was 83% which we can use . The important variables after the change is implemented are :MSSubClass,LotFrontage, LotArea, OverallQuality,Overall Cond etc.

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. I'll go with Lasso Regression as here coefficients are descent and it is pretty descent on test data.

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. R2 score has decreased after dropping the values we got
BsmtFinSF2,BsmtUnfSF,1stFlrSF, 2ndFlrSF,GrLivArea

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?

Here we can see that there is Overfirring in Linear Regression and we is Overcome this to some extent in Ridge and Lasso Regression . Here we cannot confirm that Linear Regression model is a generalized model. So we have used Ridge and lasso to cater this problem