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 Ridge is: 5.0
Optimal value of Lesso is: 0.001

If the alpha value doubled the model will go towards underfitting.

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:

Regression	Ridge	Lesso
R-square Score	0.75	0.76

R-Square value is slightly better for Lesso regression. Therefore, Lesso regression model is chosen for 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 top 5 most important feature is '1stFlrSF', '2ndFlrSF', 'OverallQual', 'OverallCond', 'BsmtFinSF1'. Una vailable of these feature, it would be good to take the second top 5 most important feature, that is

Lasso	
TotRmsAbvGrd	0.102561
GarageArea	0.067278
BsmtQual	0.064246
KitchenQual	0.059514
Fireplaces	0.054310

Question 4

How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

Ans:
Root Mean Square Error = 0.2060 on test data for the Lesso regression.