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> The optimal value for Ridge Regression is : 0.0001
2> The optimal value for Lasso Regression is : 0.0001

3> Doubled alpha values of Ridge is 0.0002 and Lasso is 0.0002

After doubling the optimal lambda values for both Ridge and Lasso Regression, we don't see any significant changes in both metrics and the features. There are few, very minor variations here and there but overall very similar.

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: The reasons why I will prefer Lasso over Ridge Regression:
The values of R2 Score, RSS and MSE for Lasso Regression are slightly better than Ridge Regression in this particular model. In Lasso Regression, we can push the model co-efficient to actual zero value. This means that the features that have co-efficient value of 0 can be removed from the model. This results in feature selection. Model complexity also reduces because we can remove features with zero co-efficient.

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:

	Lasso
LassoOverallQual	1.103737
LotArea	0.562864
BedroomAbvGr	0.532631
GarageCars	0.505633
BsmtFullBath	0.246006

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?

Answer

A robust model has low variance. This means that an unprecedented change in one or more features does not significantly alter the value of the predicted variable. Similarly, a generalisable model has reduced model complexity. As the number of features increase in the model, it becomes more complex which usually leads to low bias but high variance. A generalisable model has just enough features that it has as much low variance as possible.

This can be observed from the Bias-Variance trade-off visual shown below.
The OLS (Ordinary least squares) regression model is very sensitive to outliers and they induce high variance. To reduce this, we can go ahead with regularization (Ridge/Lasso) which include a penalty term in the cost function of the model. This penalty term will move the coefficients of the model towards 0 and thus it reduces model complexity (as feature addition is heavily discouraged). This reduces over-fitting in the model.

So regularization gets us high variance with a small trade-off in bias. Thus it helps us build a model which is robust and generalisable. A robust and generalisable model will have a good, consistent train as well as test accuracy.