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?

For Ridge, for alpha = 10 and for Lasso alpha = 20 we are getting similar results on training and test data.

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?

Based on the results, any one out of ridge and lasso can be selected. However we will be going ahead with Lasso.

L1 Lasso Regression : It is a Regularization Method to reduce Overfitting.

It is similar to RIDGE REGRESSION except to a very important difference: the Penalty Function now is: lambda*|slope|.

The result of the Lasso Regression is very similar to the Result given by the Ridge Regression. Both can be used in Logistic Regression, Regression with discrete values and Regression with interaction. The big difference between Ridge and Lasso start to be clear when we Increase the value on Lambda. In fact, Ridge can only shrink the slope asymptotically close to zero, while Lasso can shrink the slope all the way to zero. The advantage of this is clear when we have lots of parameters in the model.

In Ridge, when we increase the value of Lambda, the most important parameters might shrink a little bit and the less important parameter stay at high value.

In contrast, with Lasso when we increase the value of Lambda the most important parameters shrink a little bit and the less important parameters goes closed to zero. So, Lasso is able to exclude silly parameters from the 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?

GarageArea
GrLivArea
OverallQual
OverallCond
TotalBsmtSF

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?

For model to be robust and generalised, model should be as simple as possible. In case of simple model, performance on training data might not be very accurate however it will perform well on test data. Trade off curve on variance and bias can also be looked at while choosing the complexity of the model. Need to make sure, model isn't either overfitting or underfitting on the test data.