**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?

**Answer:**

For Ridge regression, while we plot curve between negative mean absolute error and alpha we observe that as the alpha value increases from 0 the error term decreases and the train error is increasing pattern when the alpha increases. When the value of alpha is 2 the test error is minimum so, we decided to go with value of alpha equal to 2 for our ridge regression.

For lasso regression I have decided to keep very small value that is 0.01, when we increase the value of alpha the model try to penalize more and try to make most of the coefficient value zero. Initially it came as 0.4 in negative mean absolute error and alpha.

**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?

**Answer:**

It is important to regularize coefficients and improve the prediction accuracy with decrease in variance, and making the model interpretably. Ridge regression, uses a tuning parameter called lambda as the penalty is square of magnitude of coefficients which is identified by cross validation. Residual sum or squares should be small by using the penalty. The penalty is lambda times sum of squares of the coefficients, hence the coefficients that have greater values gets penalized. As we increase the value of lambda the variance in model is dropped and bias remains constant. Ridge regression includes all variables in final model unlike Lasso Regression.

Lasso regression, uses a tuning parameter called lambda as the penalty is absolute value of magnitude of coefficients which is identified by cross validation. As the lambda value increases Lasso shrinks the coefficient towards zero and it make the variables exactly equal to 0. Lasso also does variable selection. When lambda value is small it performs simple linear regression and as lambda value increases, shrinkage takes place and variables with 0 values are neglected by the model.

**Question-3:**

After building the model, you realized that the five most important predictor variables in the lasso model are not available in the incoming data. You will have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?

**Answer:**

The five most important Predictor variables that will be excluded are:

1. GrLivArea
2. OverallQual
3. OverallCond
4. TotalBsmtSF
5. GarageArea

**Question-4:**

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

**Answer:**

The model should be simple, although its accuracy will decrease it will be more robust and generalisable. The model should be accurate for datasets other than the ones which were used during training. It can be realized using the Bias-Variance trade-off.

If the model is simple the bias is more and variance will be less and more generalisable. A complex model will need to change for every little change in the dataset and hence is very unstable and extremely sensitive to any changes in the training data. A simpler model that abstracts out some pattern followed by the data points given is unlikely to change wildly even if more points are added or removed.