**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 we plot the curve between negative mean absolute error and alpha we see that as the value of alpha increase from 0 the error term decrease and the train error is showing increasing trend when value of 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.
In case of lasso regression I assumed 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.
When we double the value of alpha for our ridge regression no we will take the value of alpha equal to 10 the model will apply more penalty on the curve and try to make the model more generalized that is making model more simpler and no thinking to fit every data of the data set .from the graph we can see that when alpha is 10 we get more error for both test and train.
Similarly when we increase the value of alpha for lasso we try to penalize more our model and more coefficient of the variable will reduced to zero, when we increase the value of our r2 square also decreases.
The most important variable after the changes has been implemented :

1. Lasso regression:
    - GrLivArea
    - OverallQual
    - OverallCond
    - TotalBsmtSF
    - BsmtFinSF1
    - GarageArea
    - Fireplaces
    - LotArea
    - LotFrontage

2. Ridge Regression:
    - MSZoning_FV
    - MSZoning_RL
    - Neighborhood_Crawfr
    - MSZoning_RH
    - MSZoning_RM
    - SaleCondition_Partial
    - Neighborhood_StoneB
    - GrLivArea
    - SaleCondition_Normal
    - Exterior1st_BrkFace

**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 the 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 value are neglected by 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?

Answer:

The 5 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 same for the accuracy of the model and why?

Answer:

The model should be as simple as possible, though its accuracy will decrease but it will be more robust and generalisable. Robustness of a model implies, either the testing error of the model is consistent with the training error, the model performs well with enough stability even after adding some noise to the dataset. Also bias-variance trade-off implies that simpler the model more the bias and less variance thus it is more generalizable .It also states that such models will perform more accurately with both training and test dataset. Complex model will need to change for every little change in dataset whereas simpler model hides some pattern followed by data points which is unlikely to change if a data is removed or added. Simpler models do not follow the patterns very strictly so they are generalizable.