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: The optimal Value range for alpha is between 50 and 10.

There will not be any change if we double these values.

The most important predictor variables are – GrLivArea,Neighborhood_NAmes ,Neighborhood_OldTo wn,KitchenQual ,SaleCondition_Partial ,RoofStyle_Gable , TotalBsmtSF ,SaleType_New ,MSZoni ng_RM,GarageType_Attchd

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:

We will chose ridse as it is little better then the lasso.

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?

Anser:

We can consider this one:

'GrLivArea','GarageType_Attchd','MSZoning_RM','SaleType_New','TotalBsmtSF

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?

Answer:

Model is not over fit and it is as simple as robust. The accuracy will go up