**Problem statement – Part II**

**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 :**

A.Optimal value of alpha : Ridge regression = 100, Lasso regression = 100

### B. If I choose double value of alpha for ridge and lasso then more features removed from model and R-squard value also decreased.

### For ridge regression R-squard value decreased from 83% to 47% on training set and 88% to 47% on test set

### For lasso regression R- squard value decreased from 88% to 72% on training set and 88% to 75% on test set

### 

C. Top features – OverallQual, ExterQual, 1stFlrSF, LowQualFinSF, KitchenQual

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 :

I will choose lasso regression.

Lasso is better because it has removed all unwanted features.

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 :

After removing top five most important predictor nest top most important predictors are

ExterQual, BsmtFinSF1, BsmtUnfSF, FullBath, GarageCars

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 :

To make model more rebust and generalize three thing are important

1.R-Squard – It should be more than 70%. In my case 83% on training set and 88% on test set

2.P-value – P-value should be < 0.05

3.VIF – VIF should be < 5