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 : Optimal value of alpha:

For Ridge regression :1.0

For Lasso Regression :0.0001

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 Ridge regression because its optimal value of alpha is high.

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:

1. LotArea
2. FullBath
3. ExterCond
4. 1stFlrSF
5. MSZoning_RH

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:

1. In order to make sure the model is robust, we need to select simple models. Linear models are always simplest models.

2. To apply Linear models, the data needs to be linear. So we need to check non-linearity in data.

3. If we discover that there is non-linearity in data, we can handle them to keep the data in Linear frame work. We can either perform Polynomial regression, or data transformation, or non-linear regression.

4. To bring accuracy in data, we need to Regularise the data. While doing so, we must check if the model is overfit or not. This can be done by testing the model with unseen data. A model is said to be perfectly trained when there is **low variance and low bias**.

Thank You