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?

- Alpha for ridge found out is 0.05

- Alpha for lasso found out is 0

- Top Features: LotFrontage, LotArea, MasVnrArea, BsmtFinSF1, BsmtFinSF2

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?

- We will use ridge as it is giving greater normalised residuals distribution

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?

- 'BsmtUnfSF', ''2ndFlrSF','LowQualFinSF','FullBath''KitchenAbvGr'

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?

- To make model robust and generalisable 3 features are required:

    - Model accuracy should be $> 70\text{-}75\%$:

    - P-value of all the features is < 0.05

    - VIF of all the features are < 5