

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?

-> For Ridge optimal value of alpha is **0.01** & for Lasso it is **0.0001**.

If we double the value of optimal alpha the difference between the train and test score increases which suggests that the model is getting **overfitted**.

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?

-> For Ridge optimal value of Lambda is **0.01** & for Lasso it is **0.0001**.

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?

-> The next 5 most important predictors are

Exterior1st

HouseStyle

Neighborhood

LotConfig

Utilities

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?

-> We can make sure that the model is robust by getting the r^2 scores of test and train data and comparing them. If both are very close to each other you can conclude that the model is robust.

Using regularization you can tune the model in such a way that you can make the model more robust and also get high accuracy.