# Advanced Regression Assignment- Subjective questions

**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?

The optimum value of lambda for Ridge and Lasso is given below:
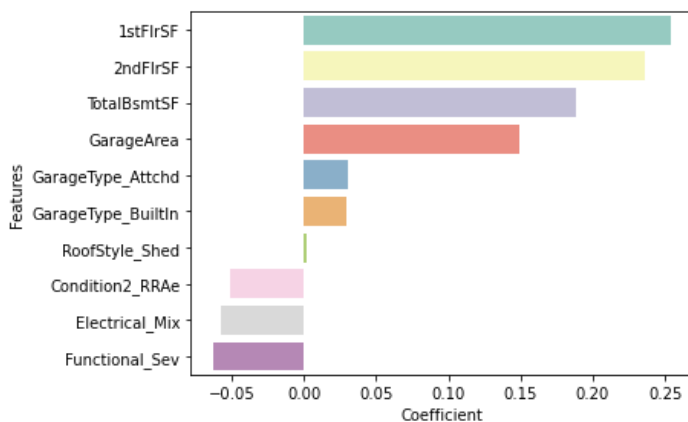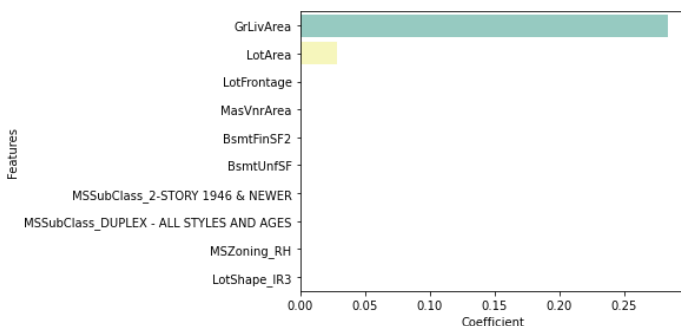
1. Ridge: 2.0
2. Lasso: 0.0001

Changes in model:

For Lasso the value did not change noticeably but for Ridge, the r2 score dropped a bit on the test set.

After change:

Ridge:



Lasso:



**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?

I choose Lasso because r2_score and mean squared error is better than Ridge and it also performs feature elimination

## 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?
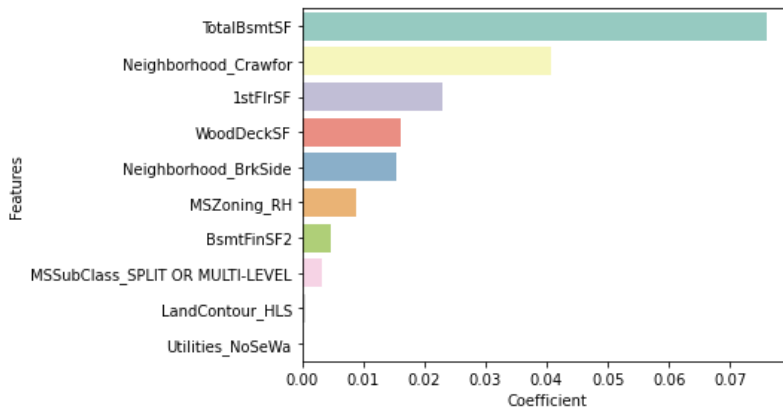


The next 5 most important features are:

1. MSZoning_RH
2. BsmtFinSF2
3. MSSubClass_SPLIT OR MULTI LEVEL
4. LandContours_HLS
5. Utilities_NoSeWa

## 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?

Linear Regression is not a generalizable algorithm because it fails when extrapolation occurs.

Also, we have to make sure the assumptions of linear regression are met.

To make a model more robust and generalizable, We can go for models which are not affected by outliers. One such example is Tree-based models.

The accuracy of the model is maintained to unseen if we go for other techniques like a random forest or boosting algorithms.