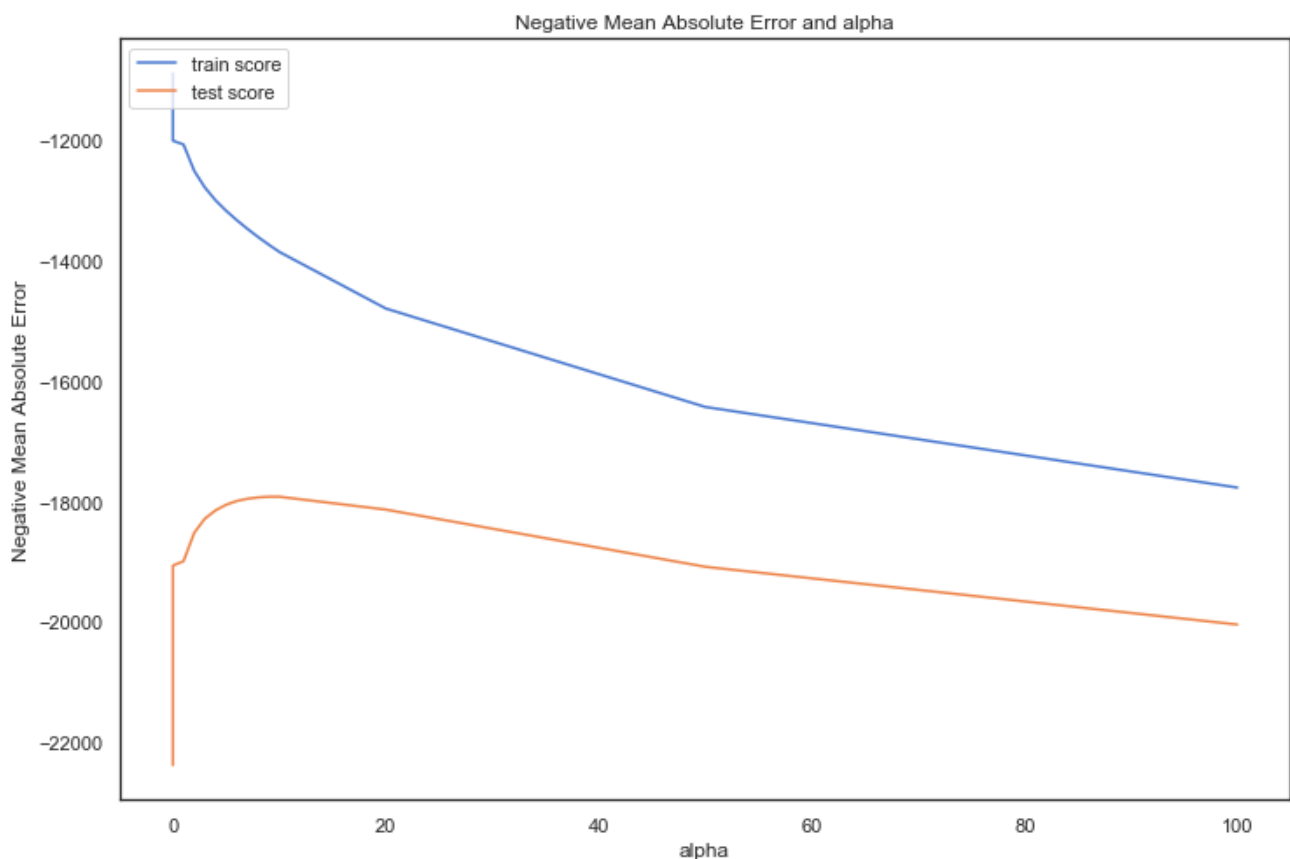


Subjective Questions

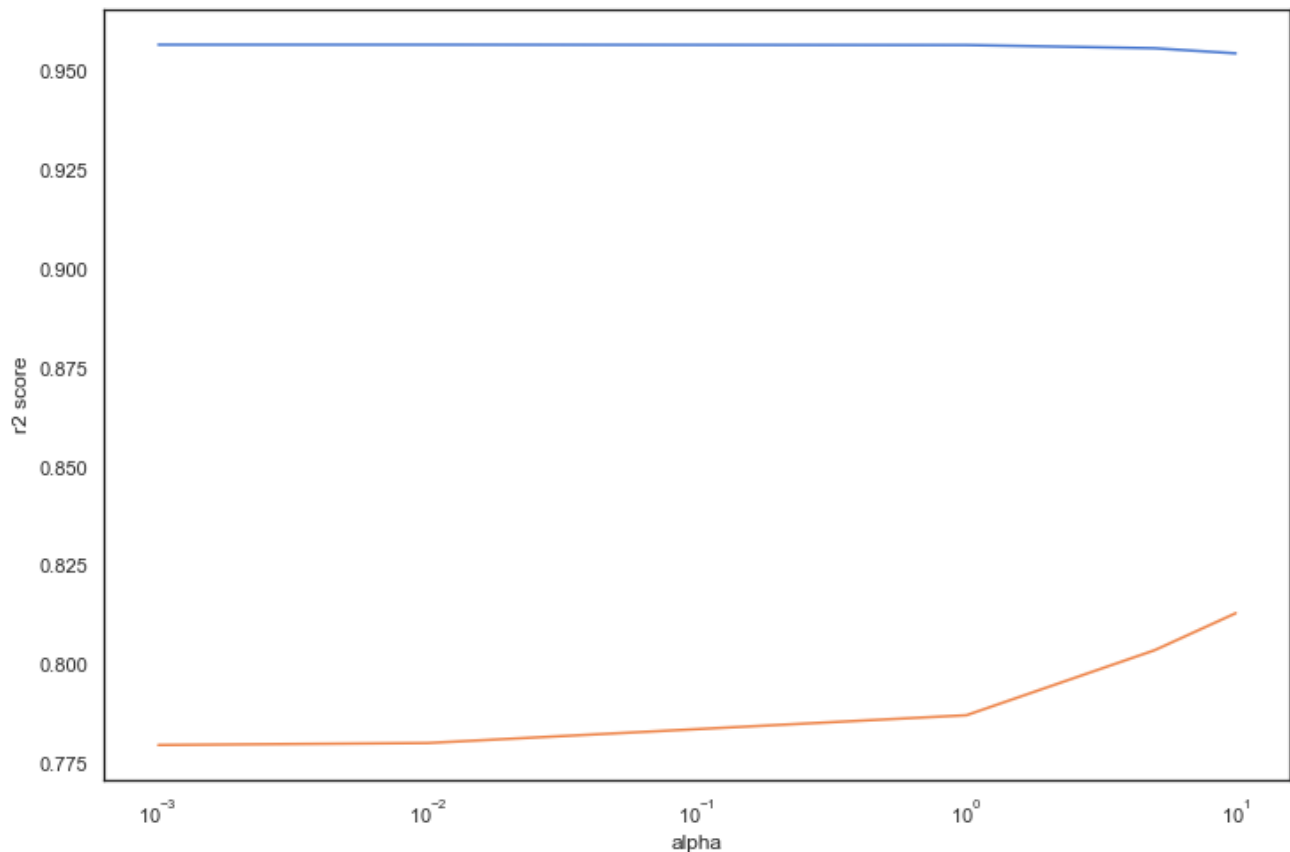
- What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?**

We would decide that on the basis of plots and chose a value of alpha where we have good training as well as the test score.

Based on the below plot, we choose 4 as the value for lambda for Ridge Regression as it has the best train as well as the score.



Based on the below plot, we choose 50 as the value for lambda for Lasso Regression as it has the best train as well as the score.



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?

Lasso Regression model is simpler, hence we would consider that, since most coefficients are zero. Optimum value of lambda for Lasso Regression is 50

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 five most important predictor variables in the lasso model are not available in the income data are:

1. Building type
2. Exterior
3. Fireplace
4. Functional
5. Housing Style

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?

A model is considered to be robust if the model is stable and does not change drastically upon changing the training set. The model is considered generalizable if it does not overfits the training data and works well with new data. Its implication in terms of accuracy is that a robust and generalizable model will perform equally well on both training and test data, the accuracy does not change much for training and test data.