

## 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?

**Answer :**

#### *Ridge regression*

For Ridge Regression the Alpha value predicted is 8.0 and when u double the alpha the R-squared value have changed from 8.21 To 8.18 which is not big but when the alpha is considered 20 and when it is doubled it is 40 in that case there is a considerable difference in the R- Squared value.

#### *Lasso regression*

For Lasso Regression the Alpha value predicted is 9.0 and when it is doubled it is 18 and the R-squared value is not varied the difference in both the cases is almost negligible but when alpha is taken 200 the R-Squared is .82 and when it is doubled the R-squared changed drastically.

‘OverallQual’ variable is the most important variable in this model.

### 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 :**

The optimal solution for this problem is to check and analyse the model in both regressions and get to a conclusion as how the R-Squared values is changing according to the alpha values. However for this model according to the alpha and change in the R-squared value I will chose Lasso Regression we can slightly make changes in alpha and keep the model as simple as possible and need not try to overfit.

### **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 variable. Which are the five most important variables now.

#### **Answer:**

1. 'KitchenQual'
2. 'Squarefoot'
3. 'MSSubclass'
4. 'GrLivArea'
5. 'Neighbourhood'

These 5 variables are the next important predictor variables as we have predicted the R-Squared value by changing this variables has changed significantly.

### **Question 4:**

How can u make sure the model is robust and genneralisable? What are the implications of the same for accuracy of the model and why?

#### **Answer:**

We can tell the accuracy of a model by plotting the Train data vs test data values of a model. The model needs to be simple and should not be overfitting.