

**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:

Optimal alpha value for Ridge = 2

Optimal alpha value for Lasso = 9e-05

- ➔ For Ridge regression, if we double the alpha, it's increasing the R-square value of the model
- ➔ For Lasso Regression, if we double the alpha, it's decreasing the R-square value of the model
- ➔ LotArea and Overall condition features are the predictor variables and it's coefficients are dropped slightly

**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:

For Ridge its 2 and for Lasso it's 9e-05, I choose Lasso regression as it has Lambda value near zero but not zero and it explains the Sales Price better than Ridge regression model

**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?

Answer:

I'll take remaining top 5 features after removing the previous top 5 features.

LotArea, MasVnrArea, **Neighborhood\_Crawfor (Crawford)**, Overall Condition – Fair & WoodDeckSF

**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?

**Answer:**

We can remove outliers, cleaning the data properly and tuning the hyperparameter to optimal value and not hard coding the model will make model robust.