

**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 to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

**Answer:**

The optimal value of alpha for ridge and lasso regression is 2 and 0.002 respectively with no substantial difference in the model performance.

The most important predictor variables for both cases are mentioned below.

Ridge regression: GrLivArea, MSZoning\_RL, OverallQual, MSZoning\_RM, TotalBsmtSF

Lasso regression: GrLivArea, OverallQual, MSZoning\_RL, MSZoning\_RM, TotalBsmtSF

**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:**

Both types of regression are providing with similar results in terms of test/train set model predictions. I would prefer ridge regression in this case because it gives a slightly better model accuracy.

**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:**

The most important variables after doing this experiment are:

1stFlrSF, FullBath, HalfBath, OverallCond, d\_BsmtQual

**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:**

A robust model can be made by adding some noise to the test data as well as treat the outliers in the dataset properly. It is also done using regularization techniques like ridge and lasso regression. This leads to better model accuracy because of hyperparameter tuning that we do.