

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 value of lambda for ridge regression is 10

Optimal value of lambda for lasso regression is 0.01

After doubling the value,

```
GrLivArea          1.11
OverallQual_8      1.09
OverallQual_9      1.08
Functional_Typ     1.07
Neighborhood_Crawfor 1.07
TotalBsmtSF        1.05
Exterior1st_BrkFace 1.05
CentralAir_Y       1.04
YearRemodAdd       1.04
Condition1_Norm    1.03
Name: Lasso, dtype: float64
```

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?

Depends on the use-case, if we need feature selection, then we will use lasso regression, if we need smaller magnitude of coefficients then we use 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?

```
2ndFlrSF          0.10
Functional_Typ     0.07
1stFlrSF          0.07
MSSubClass_70     0.06
Neighborhood_Somerst 0.06
Name: Lasso, dtype: float64
```

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?

A model is robust when any variations in the data doesn't affect the performance much,

A model is generalisable when it is able to predict the unseen data properly.

It should always be a tradeoff between the bias and variance for which we can use lasso and ridge predictions.

