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?

Optimal value of Alpha:

Ridge: 6.579332246575679
Lasso: 0.00019179102616724886

If we increase the value of alpha in both cases, the value of the coefficients will reduce. At the same time bias will increase and the variance will decrease.

If we choose double the value of Alpha, the most important predictor variable in

Ridge: GrLivArea
Lasso: BsmtFinSF1

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?

I will choose Lasso as here with a lesser number of variables we are able to get similar results. This will decrease the compute time and power that is required. The model will be more interpretable too.

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?

Columns removed were

column_name	co_efficient
GrLivArea	0.128340
OverallQual	0.111399
RoofMatl_WdShngl	0.110064
TotalBsmtSF	0.067872
Neighborhood_NoRidge	0.066517

New Columns

column_name	co_efficient
BsmtFinSF1	0.428872
BsmtUnfSF	0.149751
MasVnrArea	0.088815
LotArea	0.086470
GarageArea	0.085388

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?

We can make the model robust by ensuring that there is no overfitting. When there is overfitting, we will notice that the model performs very well on the training data and not so well on the test data. This can be done through regularization techniques like Ridge and Lasso. By reducing the features we can make sure that the model does not learn from the noise in the data. Hence it will perform well in the test data too.