**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

The optimal value for Ridge and Lasso transformation is

Ridge : 20

Lasso : 0.001

When we double the alpha the ranking of most important predictors changes

Also we tend to loose biasness and get more variance

After the change most important predictor

Will be

Ridge : Overall Quality

Lasso: GrLivArea

Top 5 Ridge

Graphical user interface, table

Description automatically generated with medium confidence

Bottom 5 Ridge

Text

Description automatically generated with low confidence

When alpha is doubled

Top 5 Ridge

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface, text, application

Description automatically generated

**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?**

Ans:

R2 scores for

Ridge

Ridge Regression r2\_train:0.9119 , r2\_test:0.8888

Lasso

Lasso Regression r2\_train:0.9059 , r2\_test:0.887

No significant change in Lasso and Ridge I would chose **Lasso** as it gives me feature reduction as well

**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?**

Ans

Table

Description automatically generated

The Five most important variables are

Condition2\_PosN, GrLivArea, Neighborhood\_Crawfor, Neighborhood\_Somerst, OverallQual

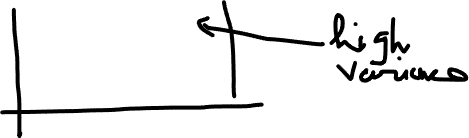
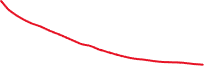
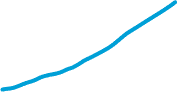
**Question 4**

**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?**

To make model simple and robust we must try to see that we are not overfitting the model

A simplistic model may be low on bias and high on variance whereas a complex model may be high on bias and low on variance . We need to balance the two

To draw an anology we need to make them meet like the demand and supply curve



|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |