Question 1

What is the optimal value of alpha for ridge and lasso regression?

The optimal value of Lambda Values:

Ridge value: 4.0

Lasso value: 0.0002

What will be the changes in the model if you choose double the value of alph a for both ridge and lasso?

When we choose to double the value of the alpha then the normalized features in ridge regression, the column whose values are doubled gets the coefficients halved while oth ers remains unchanged.

In Lasso regression the coefficients becomes 0 while others remains unchanged.

What will be the most important predictor variables after the change is implemented?

The predictor variable with the largest standardized regression coefficient as the most important variable.

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?

The optimal value of Lambda Values:

Ridge value: 4.0

Lasso value: 0.0002

We have choose The Lasso value as its value is 0.0002 which is close to the coefficient value of one feature to 0 .

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?

The five most important predictor variables are below

1. OverallCond
2. BsmtFinSF1
3. TotalBsmtSF
4. GrLivArea
5. GarageCars

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?

When a model is said to be robust if its dependent variable (output label) is consistent & accurate even if one or more of the independent variables (input features) or assumptions are drastically changed due to unknown circumstances.

The accuracy of the model can be implemented using below ways,

1. Add more data
2. Treat missing and Outlier values
3. Feature Engineering
4. Feature Selection
5. Multiple algorithms
6. Algorithm Tuning
7. Ensemble methods
8. Cross Validation