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 of alpha for Lasso regression, also known as the L1 regularization parameter, depends on the specific dataset and the problem at hand. In general, alpha controls the strength of the regularization penalty in Lasso regression. A larger alpha will result in more coefficients being exactly zero, which can help with feature selection and model interpretability. Conversely, a smaller alpha allows more features to be included in the model.

To determine the optimal value of alpha for Lasso regression, it is common to use techniques such as cross-validation or grid search to find the alpha value that minimizes the prediction error or optimizes a specific performance metric for the given dataset. This process is often part of model tuning and selection in machine learning workflows.
When we double these values, the model performance remains same in both the cases.

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:
We see similar performance for both the techniques.

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:
SaleType, MSZoning

Question 4
How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?
Answer:
We have to make sure model is not over fitting and is as simple as possible, we are ensuring that it is better and generalizable.