Please limit your answers to less than 500 words per question.

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 Alpha for Ridge is 500 and for Lass as well it is 500.

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: Will be usging the Lasso regression lamba as the number of predictor variables is high and the Lasso regression automatically does the feature elimination.

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?

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?

Answer: If we ensure the model is not too complex and understands the date will with regularization techniques, then we can say the model is robust and generalisable. Accuracy observed on the test data is almost close to the training data we can safely assume that the model is robust and generalisable.