Question 1

What is the optimal value of alpha for ridge and lasso regression?
For ridge : 100
For lasso : 500

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?

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?

Lasso. Lasso model has a higher R2 score (test) than Ridge and a lower RSS (test) and MSE (Test) than Ridge.

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?
The model should not be overfitting. It should have low bias and low variance, such that the model identifies all the patterns that it should and is also able to perform well with unseen data.

What are the implications of the same for the accuracy of the model and why?
If the model identifies all necessary underlying patterns in the data without overfitting , it should perform well on unseen data as well so the accuracy will be good.