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?

The optimal value of alpha in Ridge regression is 3.0 The optimal value for Lasso regression is 0.0001.

When the alpha is doubled there is a difference in coefficients

The important predictors after the change are not changed

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?

I will choose ridge regression as the Rsquare value is Mean square error is less in Lasso

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?

All the variables are derived from the incoming data

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?

The A model is robust and generalizable if it does not overfit. This is because an overfitting model has very high variance. Any change in data affects the model prediction. An overfitting model will try to fit the training data more accurately but may fail in testing data.