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?

Ans:1
The value of Alpha or Lambda (in ridge and lasso regression) will depend on the data available. If its value is kept at zero, the model will overfit and if it keeps on increasing its value the model will become underfit. So if the value of Alpha/Lambda is double after finding its optimal value the model will underfit. In other words, bias will increase, and variance will decrease. If variance or bias is increased the summation of error will also increase and the model will fail in predictions.

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?

Ans: 2
It depends on what the business wants. If the business thinktank considers all the parameters to be important, then proceed with Ridge regression.
If there are lots of variables and the business think tank wants to know significant variables that affect the business most, then proceed with Lasso.

In Ridge regression, it considers all the variables and coefficients of some variable might be close to zero but not zero.

In Lasso Regression, it considers a subset of available variables. It considers variables that have a coefficient value above a certain threshold rest variables coefficient is assigned to zero.

Question 3

After building the model, you realized 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?

Ans: 3

In the above case, a new model should be generated by removing the most important variable from the training data. After rebuilding a new model, check for variable coefficient values not equal to zero.ro.

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?

Ans: 4

By keeping the value of Alpha/Lambda optimal. By keeping the value of alpha optimal, the model variance and bias are balanced and error summation is also minimum.

Before performing ridge and lasso regression, data transformation like logarithmic, square or exponential transformation should be performed if residual plot indicates the presence of nonlinear relations.