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?

Optimal value of alpha for ridge is 20 and lasso is 0.001

There is a slight change in variance and r2 value decreases, if we double the alpha value.

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?

We need to regularize coefficients and increase prediction accuracy by decreasing the variance.

As we increase the value of lambda, variance decreases in ridge and in lasso when lambda is small it does simple linear regression, if lambda increases variables with 0 value are neglected.

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?

The simpler the model, more the bias and lesser the variance where the model is robust and generalizable. Accuracy doesn’t change much for train and test data.