Q.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. Usually the Optimal value is 10, If we double the value R2 score decreases for both train and test set.

Q.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: The model we will choose to apply will totally depend on the case. For example:

1.If we don't want too large coefficients and reduction of coeff. magnitude is one of our primary goals, **Ridge Regression** should be the preference.

2. Too many variables and one of our primary goals is feature selection, then we will use **Lasso**.

Q.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: We drop those features and will build the model again. According to coeff. of lasso sorted in the descending order, the top 5 we get are the five most important predictor variables now.

Q.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: Robust and Generalisable model means variation of data, doesn't affect the model too much and is prepared for new data which is not seen in training and test. Overfitting and underfitting is taken care of in such models.

The Accuracy depends on complexity, taking care of factors which would cause error, or more variance in data.