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. The R2 test score on the Lasso Regression Model is slightly better than that of Ridge Regression Model. Moreover, the training accuracy is slightly reduced; hence, making the model an optimal choice as it seems to perform better on the unseen data.

The Lasso Regression test is slightly lower than that of the Ridge Regression Model which implies Lasso Regression performs better on the unseen test datas. Also, Lasso can help in feature selection (the coefficient values of some of the insignificant predictor variables became 0), which implies Lasso Regression has a better edge over Ridge Regression. And, the variables which have been predicted by Lasso Regression can be applied in order to choose variables for predicting the price of any particular house in this data set.

Moreover, while choosing a type of regression in the real world, we have to deal with the confounding dangers of outliers, non-normality of errors and overfitting especially in sparse datasets among others. In such cases Lasso Regression will be very much beneficial.

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. Robustness of a model implies, either the testing error of the model is consistent with the training error, the model performs well with enough stability even after adding some noise to the dataset. Thus, the robustness (or generalizability) of a model is a measure of its successful application to data sets other than the one used for training and testing.

By implementing regularization, we can control the trade-off between model complexity and the bias which is directly connected to the robustness of the model. Regularization will help in dealing with the coefficients for making the model too complex; thereby allowing only the optimal amount of complexity to the model. It helps in controlling the robustness of the model by making the model optimally simpler.