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?

Alpha for Ridge is 0.05
Alpha for Lasso is 0.0001
After doubling the value of the Ridge and Lasso , R2 value did not change much
The important predictor variables remained same.

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 to apply Lasso , since it will predict the important variables and make the coefficient for others zero

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?

1) **TotalSquareFeet - Total Square Feet of the property**

   2) **Condition2 - Near to positve off site deature like park , greenbelt**

   3) **RoofMatl - Roof material is wood shingles**

   4) **KitchenQuality - Kitchen Quality is Average or Fair or Good**

   5) **2nd floor Square Feet - Second Floow Square feet area**

2) **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?

By checking the R2 score on the test data we can make sure that the model is robust , It should give good value for the test data also thus making sure that model is not overfitting or underfitting.

Applying Regularization is the best way to achieve this
We can apply Ridge or Lasso Regression which applies regularization to make sure that a model is not performing overfitting or underfitting.