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 : Optimal value of Alpha for Ridge and Lasso regression came as 2 and 0.0001

When we increased it to 4 for Ridge we can see the magnitude of coefficients were decreased plus the variance decreased.

Most important predictor variables is **Street**

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 will apply Lasso as it allows to shrinks the coefficients and helps to reduce the model complexity and multi-collinearity.It makes model more robust

This is impossible in the ridge regression model values can be shrunk close to zero, but never equal to zero.

Thus lasso provides ways for automatic feature selection how ever Ridge doesn't

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?

Based on the list of missing predictor variables most important variables are

1.) Sqft area of house
2.) Ageof house
3.) Neighborhood
4.) SaleType

5.) Street

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?

1.) By removing outliers
2.) Test accuracy is not having too much of a difference from Training set

If model is not ribust it will not help in doing correct predictions