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?

Values for ridge and lasso are 6 and 500 respectively. If we double the value of alpha,
then the results are similar,but there is a minor decrease in r2 score ie. not a major
change. The most important predictor vars are:
For positive impact:
1. Roof Material
2. GrLivArea
3. YearBuilt
4. GarageCars
For negative impact:
1. Condition2_PosN
2. Kitchen Quality(TA)
3. Edwards Neighbourhood

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 apply the optimal value for lambda because it results in better r2 scores.

Question 3

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?

Initially here are the 5 most important vars:
1. RoofMatl
2. GrLivArea
3. OverallQual
4. YearBuilt
5. GarageCars

Now after removing these vars and building the model again:
1. 1stFlrSF
2. 2ndFlrSF
3. KitchenQual
4. GarageArea
5. Neighborhood

## 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?

To ensure that the model is robust, we need to ensure that:
1. It should be less complex
2. Standardization is used
3. Regularization is implemented