## 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 to double the value of
alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

**Answer:**
Ridge Alpha: .01

lasso Alpha: 20

R2 score on training and testing data got reduced

(Experiments in notebook)

## 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?

**Answer:**

R2 score of lasso is high compared to ridge, also see less differance between the test and train data, so we choose lasso over ridge

## 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?

Answer:
- GrLivArea
- TotalBsmtSF
- MSSubClass
- LotFrontage

(Experiments in notebook)

## 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?

**Answer:**
Techniques uses
1) Regularisation techniques such as Ridge and Lasso
2) Cross-validation: This is a technique used to evaluate the performance of a model by dividing the dataset into training and testing sets.
3) Feature selection: By selecting the most relevant features for the model

Model can be made more robust and generalizable by using techniques such as cross-validation, regularization,
ensemble methods, and feature selection. While these techniques can help to improve the robustness and generalizability
it will reduce the accurecy as its modelled for generic data set.

In summary, a model that is robust and generalizable may have a slightly lower accuracy on a specific dataset, but it will perform better on new unseen data,