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?

- **Before Doubling**

Optimal value of Alpha for Ridge Regression = 1.0
Optimal value of Alpha for Lasso Regression = 0.0001

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.837178	0.834105
1	R2 Score (Test)	0.815744	0.819146
2	RSS (Train)	2.003897	2.041715
3	RSS (Test)	1.004077	0.985538
4	RMSE (Train)	0.044302	0.044718
5	RMSE (Test)	0.047825	0.047381

- **After Doubling**

After Doubling alpha for Ridge Regression = 2.0
After Doubling alpha for Lasso Regression = 0.0002

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.822501	0.823786
1	R2 Score (Test)	0.814520	0.814478
2	RSS (Train)	2.184526	2.168714
3	RSS (Test)	1.010746	1.010971
4	RMSE (Train)	0.046256	0.046088
5	RMSE (Test)	0.047983	0.047989

- R2 scores of train and test sets dropped by 1%
- RMSE increased by 0.002
- RSS increased by 0.18

Important Predictor variables after change implemented

- Above Ground Living area square feet
- Masonry Veneer Area
- Total rooms above grade
- Wood deck square feet
- Exterior first brick face
- 1st Floor Square Feet
- 2nd floor square feet
- Lot area

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 Lasso with alpha = 0.0001. Why because:

- R2 score after lasso regression is high compared to Ridge regression. Difference between r2 scores of train and test is Low compared to Ridge's train and test R2 scores difference.
- RSS and RMSE are error terms these values must be low for a good model. Lasso has less RMSE,RSS compared to ridge.
- Lasso does feature selection by equating insignificant variables coefficients to zero. But Ridge keeps every insignificant variable by reducing their coefficients closed to zero but not zero.

- Choosing best hyper parameter value gives best coefficients to variables.
- Higher alpha value leads to underfitting. Lower alpha leads to overfitting. Choosing the optimal value will overcome the problem of overfitting.

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?

After dropping top 5 variables R2 scores dropped by 4%

Five most important predictor variables are :

1. First floor square feet
2. Basement finished square feet
3. Wood deck square feet
4. Street
5. Overall quality

	Lasso
1stFlrSF	0.290667
BsmtFinSF1	0.056985
WoodDeckSF	0.041430
Street	0.040486
OverallQual	0.027293

	Metric	Lasso Regression
0	R2 Score (Train)	0.795996
1	R2 Score (Test)	0.817291
2	RSS (Train)	2.510736
3	RSS (Test)	0.995643
4	RMSE (Train)	0.049589
5	RMSE (Test)	0.047623

Question 4

How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of The model and why?

- Model is Robust and Generalizable when data has no missing values, no outliers, no errors and simple(with less number of Predictor variables) with moderate bias and variance.

- Accuracy of the model can be explained by R squared metric when r2 scores are greater than 75% and error terms RSS and RMSE are very low we can say model's accuracy is good.

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.837178	0.834105
1	R2 Score (Test)	0.815744	0.819146
2	RSS (Train)	2.003897	2.041715
3	RSS (Test)	1.004077	0.985538
4	RMSE (Train)	0.044302	0.044718
5	RMSE (Test)	0.047825	0.047381