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?

   GridSearchCV algorithm was used to find best alpha for both the regularization techniques.
   Alpha was chosen at 20 for ridge regression and at 0.001 for lasso regression.
   If we double the values of alpha, it will mean more penalty or reduce coefficients further.
   This will also cause underfitting, reducing the test score further.

   Ridge , Lasso with Alpha = 20 , Alpha =0.001 respectively

	Metric	Linear Regression	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.878723	0.870999	0.870918
1	R2 Score (Test)	0.839932	0.835069	0.835446

   We observe here that R2 score on test is at 83 % for both regularization

   Ridge , Lasso with Alpha = 40 , Alpha =0.002 respectively

	Metric	Linear Regression	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.880501	0.872657	0.872399
1	R2 Score (Test)	0.788291	0.817897	0.814299

   We observe here that R2 score on test is at 81 % for both regularization , reduced after we increased or doubled alpha, causing underfitting. The predictor variables after doubling the alpha remain the same only ,the coefficients magnitude reduce further as we increase alpha. Hence, we should see more coefficients moving towards zero for ridge and lasso reducing to zero as we increase alpha.

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 the lasso regression with lambda value of 0.01 which is show a good Train and Test score. The gap between test and train R2 score is reduced, implies reduced overfitting. Also the MSE is lesser compared to ridge.
   Although R2 test values of Lasso are very close to ridge regression, but lasso is chosen because it was able to make few coefficients towards zero for the model compared to ridge thus reducing multicollinearity and being more robust compared to ridge.

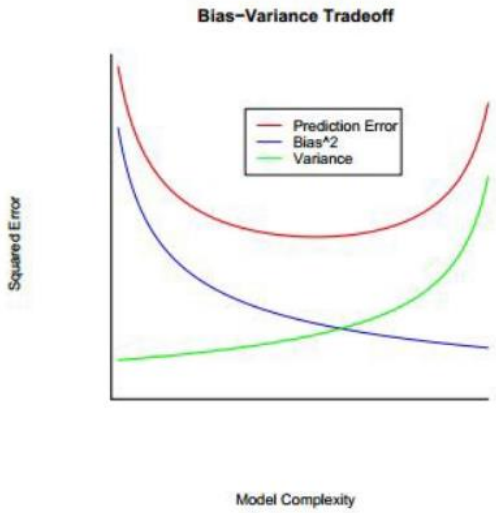
3. After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?

   The coefficient values received from ridge and lasso are sorted to find the significant variables ,from these if we drop the top 5 ,and build the model , then these ones will become significant .Example for lasso we would have :

   1stFlrSF
   2ndFlrSF
   GarageArea
   KitchenQual
   BsmtUnfSF

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?

   Regularization can be used to make the model robust and simpler.
   Regularization constrains/ regularizes or shrinks the coefficient estimates towards zero, this technique discourages learning a more complex model, so as to avoid the risk of overfitting. Regularization also reduces the variance of the model, without increase in its bias. The tuning parameter λ, used in the regularization techniques, controls the impact on bias and variance. As the value of λ rises, it reduces the value of coefficients and thus reducing the variance. Till a point, this increase in λ is beneficial as it is only reducing the variance (hence avoiding overfitting), without losing any important properties in the data. But after certain value, the model starts losing important properties, giving rise to bias in the model and thus underfitting. Therefore, the value of λ should be carefully selected.
   The accuracy of the model can be maintained by keeping the balance between Bias and Variance as it minimizes the total error as shown in the below graph.


Bias−Variance Tradeoff
Prediction Error
Bias^2
Variance
Squared Error
Model Complexity