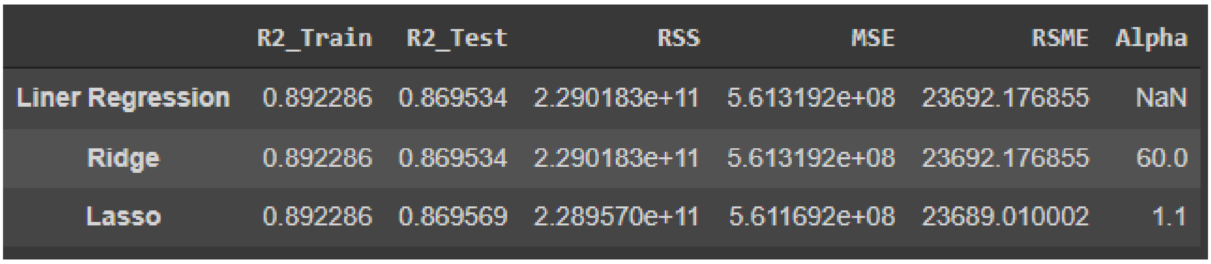
Question 1

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?

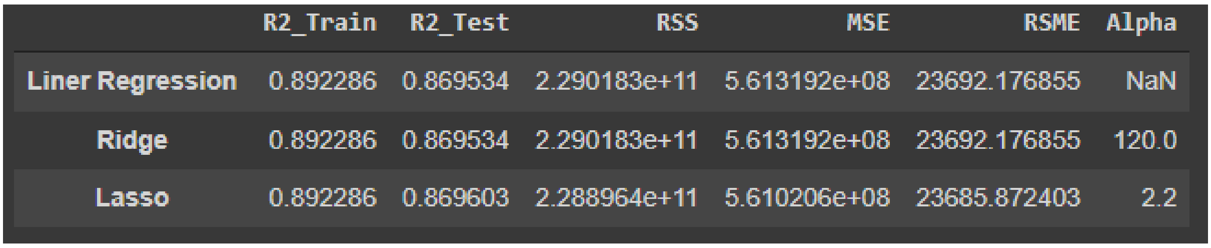
Ans:

The Optimal value of alpha for ridge is **60.0** ,
and Optimal Value of alpha in Lasso is **1.1**



	R2_Train	R2_Test	RSS	MSE	RSME	Alpha
Liner Regression	0.892286	0.869534	2.290183e+11	5.613192e+08	23692.176855	NaN
Ridge	0.892286	0.869534	2.290183e+11	5.613192e+08	23692.176855	60.0
Lasso	0.892286	0.869569	2.289570e+11	5.611692e+08	23689.010002	1.1

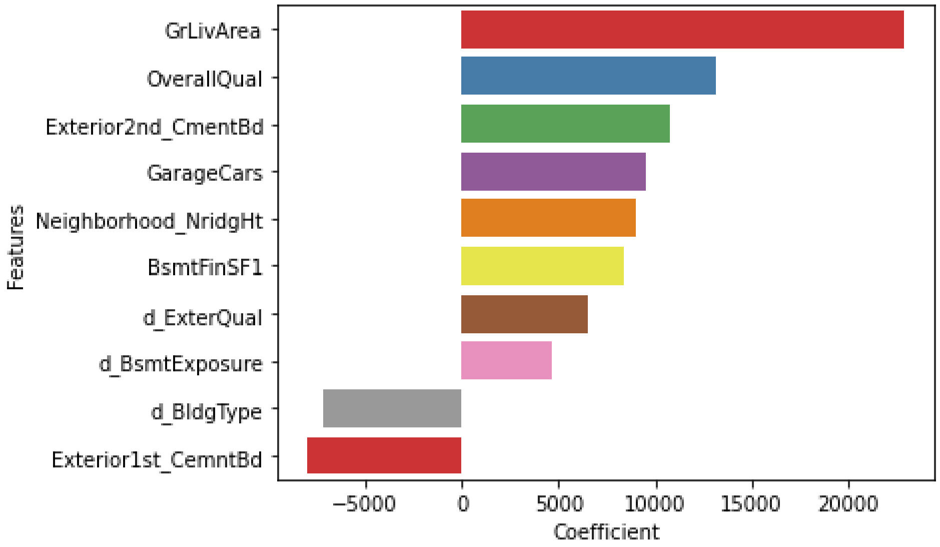
If we double the value of ridge i.e **120** and Optimal Value of Lasso is **2.2,** We can see slight drop in RSME for Lasso Regression

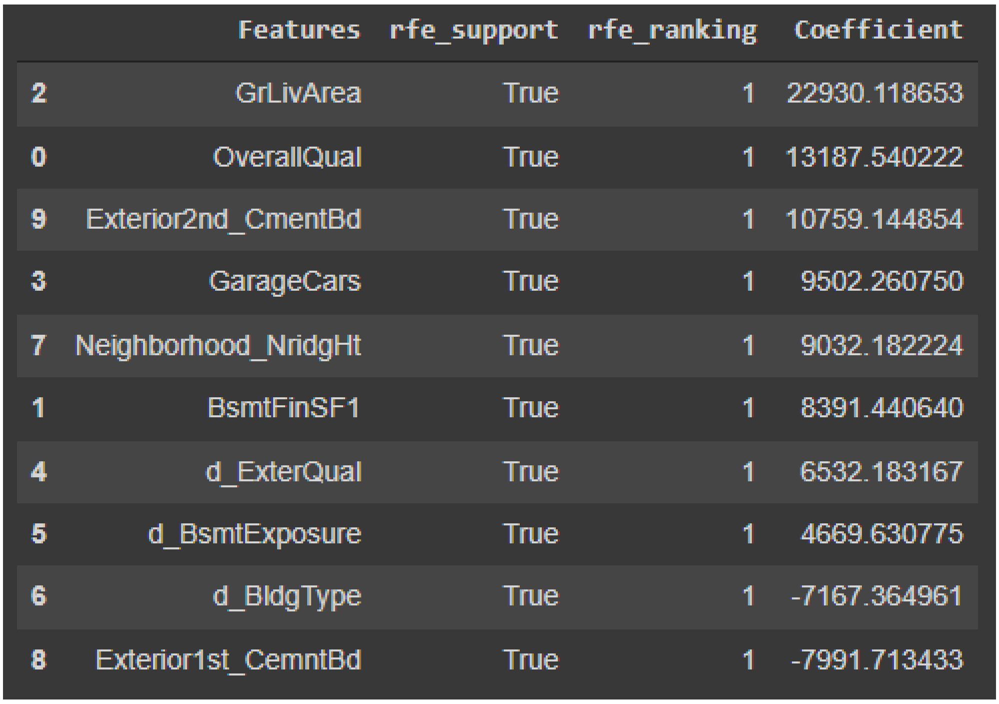


	R2_Train	R2_Test	RSS	MSE	RSME	Alpha
Liner Regression	0.892286	0.869534	2.290183e+11	5.613192e+08	23692.176855	NaN
Ridge	0.892286	0.869534	2.290183e+11	5.613192e+08	23692.176855	120.0
Lasso	0.892286	0.869603	2.288964e+11	5.610206e+08	23685.872403	2.2

Also, Most important predictor variables will be
'GrLivArea', 'OverallQual', 'Exterior2nd_CmentBd', 'GarageCars', 'Neighborhood_NridgHt', 'BsmtFinSF1', 'd_ExterQual', 'd_BsmtExposure', 'd_BldgType', 'Exterior1st_CemntBd'

We can see feature importance and its coefficient in this graph and Table below:





	Features	rfe_support	rfe_ranking	Coefficient
2	GrLivArea	True	1	22930.118653
0	OverallQual	True	1	13187.540222
9	Exterior2nd_CmentBd	True	1	10759.144854
3	GarageCars	True	1	9502.260750
7	Neighborhood_NridgHt	True	1	9032.182224
1	BsmtFinSF1	True	1	8391.440640
4	d_ExterQual	True	1	6532.183167
5	d_BsmtExposure	True	1	4669.630775
6	d_BldgType	True	1	-7167.364961
8	Exterior1st_CemntBd	True	1	-7991.713433

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?

Ans:
After Looking at model metrics, we can see that there is no major difference in model metrics, however, the general practice of keeping models simple and robust is one key principle while building a model.

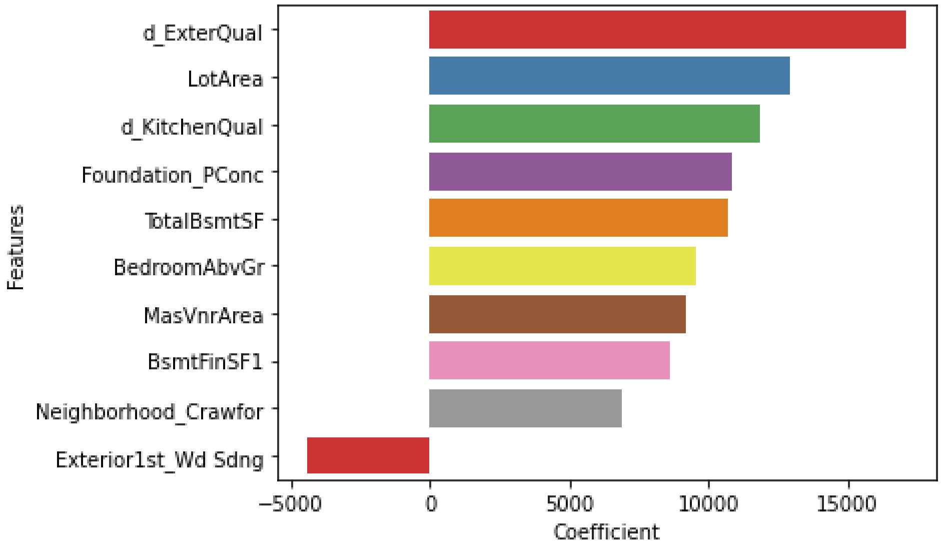
From the above Models, I select Lasso Regression, for the reasons listed below:

1. It has a lower number of features which in turn suggests that the model is not complex while giving similar or better metrics.
2. Lasso regularisation helps with feature selection
3. A Regularisation technique is always preferred over unregularized models since it avoids overfitting of train data and gives a robust model.

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?

Ans:
After creating a new Lasso Model, with excluding top 5 important predictor, i.e.
'GrLivArea', 'OverallQual', 'Exterior2nd_CmentBd', 'GarageCars', 'Neighborhood_NridgHt'

New, top 5 Important predictor variables are :
'd_ExterQual', 'LotArea', 'd_KitchenQual', 'Foundation_PConc', 'TotalBsmtSF'



From the above experiments, we have found that there is a significant decrease in R2 score of modified lasso model with removal of top 5 predictors.

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?

Ans:
In order to keep the model robust and generalisable, we can take into consideration a few pointers of a good model.

1. A model should not be overfitted to the train dataset. We can check for this by looking at accuracy metrics of model on train and test datasets. If the model performs very well on the train dataset and very poorly on the test dataset, It indicates that Model is overfitted.
2. A model should not be underfitted. We can check this by looking at the train and test metrics which would be very poor.
3. There should not be High Variance and High Bias in the Model Prediction; we can reduce this by regularisation techniques.
4. Model with lower number of features is always preferred over Model with higher number of features with similar performance metrics
5. There should be the least amount of collinearity in the model.