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?

The optimal value for ridge is 0.01 and for lasso is 0.00001 (1e-5).

When we double the values for both ridge and lasso, there is very slight difference in the r2 score and mean square error.

	Metric	Ridge Regression	Lasso Regression	Ridge Regression R-2	Lasso Regression R-2
0	R2 Score (Train)	0.946900	0.947200	0.946000	0.947000
1	R2 Score (Test)	0.868100	0.867600	0.869400	0.868700
2	MSE (Train)	0.091104	0.091104	0.092195	0.091104
3	MSE (Test)	0.147309	0.147648	0.146629	0.146969

The top 10 variables also remain the same.

Ridge alpha = 0.01

```
In [80]: #selecting the top 10 variables
         ridge_coef.sort_values(by='Mod',ascending=
Out[80]:
```

	Features	Coefficient	Mod
0	LotFrontage	6.858800	6.858800
67	MSZoning_RM	2.969938	2.969938
68	Street_Pave	2.823709	2.823709
70	LotShape_IR3	2.695100	2.695100
72	LandContour_HLS	2.683479	2.683479
66	MSZoning_RL	2.624948	2.624948
69	LotShape_IR2	2.620848	2.620848
71	LotShape_Reg	2.513939	2.513939
58	MSSubClass_85	-1.666734	1.666734
16	BsmtUnfSF	0.737276	0.737276

Ridge alpha =0.02

	Features	Coefficient	Mod
0	LotFrontage	6.858800	6.858800
67	MSZoning_RM	2.969938	2.969938
68	Street_Pave	2.823709	2.823709
70	LotShape_IR3	2.695100	2.695100
72	LandContour_HLS	2.683479	2.683479
66	MSZoning_RL	2.624948	2.624948
69	LotShape_IR2	2.620848	2.620848
71	LotShape_Reg	2.513939	2.513939
58	MSSubClass_85	-1.666734	1.666734
16	BsmtUnfSF	0.737276	0.737276

Lasso alpha = 0.00001

	Feature	Coef	mod
0	LotFrontage	6.681246	6.681246
67	RoofMatl_Metal	3.186275	3.186275
68	RoofMatl_Roll	3.039109	3.039109
70	RoofMatl_WdShake	2.887829	2.887829
72	Exterior1st_AsphShn	2.872657	2.872657
69	RoofMatl_Tar&Grv	2.826127	2.826127
66	RoofMatl_Membran	2.813687	2.813687
71	RoofMatl_WdShngl	2.713555	2.713555
58	Condition2_RRAe	-1.701285	1.701285
13	2ndFlrSF	1.060698	1.060698

Lasso alpha = 0.00002

	Feature	Coef	mod
0	LotFrontage	6.681246	6.681246
67	RoofMatl_Metal	3.186275	3.186275
68	RoofMatl_Roll	3.039109	3.039109
70	RoofMatl_WdShake	2.887829	2.887829
72	Exterior1st_AsphShn	2.872657	2.872657
69	RoofMatl_Tar&Grv	2.826127	2.826127
66	RoofMatl_Membran	2.813687	2.813687
71	RoofMatl_WdShngl	2.713555	2.713555
58	Condition2_RRAe	-1.701285	1.701285
13	2ndFlrSF	1.060698	1.060698

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?

The optimal value of LAMBDA we got in case of Ridge and Lasso is:

- Ridge -0.01
- Lasso - 0.00001

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.946900	0.947200
1	R2 Score (Test)	0.868100	0.867600
2	MSE (Train)	0.091104	0.091104
3	MSE (Test)	0.147309	0.147648

While comparing the r2 and MSE values for ridge and lasso, both models give almost same accuracy values.
I have chosen the Lasso regression since this model helps in feature reduction making the model simple to interpret and doesn't impact the results.

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?

After dropping the top 5 predictor variables and recreating the lasso model with the same alpha 1e-5

the top 5 variables are

	Feature	Coef	mod
0	LotArea	9.172316	9.172316
57	Condition2_RRAe	-1.429154	1.429154
12	2ndFlrSF	0.995141	0.995141
66	RoofMatl_Tar&Grv	0.577244	0.577244
68	Exterior1st_BrkComm	0.513965	0.513965

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?

- The model is expected to be as simple as possible and simpler models are considered as more 'generic', though its accuracy will be decreased but it will be more robust.

- This can be understood from the Bias-Variance trade-off. The simpler the model the more the bias but less variance becoming generalizable. Whereas the complex model will have high variance and low bias.
- Sometimes underfitting and overfitting are the problems associated with the model. Hence, it is important to have balance in Bias and Variance to avoid such problems. This is possible with “Regularization”.
- Regularization helps in managing the model complexity by essentially shrinking the coefficients towards zero. This avoids the model becoming too complex, thus reducing the risk of overfitting.
- Regularization method should be used to keep the model optimum simpler. It penalizes the model if it becomes more complex.
- Regularization method helps to achieve the Bias-Variance trade off. It compromises by increasing bias to a optimum position where Total Error is minimum.
- This point also known as Optimum Model Complexity where Model is sufficient simpler to be generalisable and also complex enough to be robust.
- Making a model simple lead to Bias-Variance trade off.