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

- Optimal value of alpha for Rigde : 7.0
- Optimal value of alpha for Lasso : 0.0001

Below is complete result of Linear, Ridge and lasso Model

	Metric	Linear Regression	Ridge Regression	Lasso Regression	Ridge Regression Alpha*2	Lasso Regression Alpha*2
0	R2 Score (Train)	0.901222	0.900626	0.898730	0.899767	0.898730
1	R2 Score (Test)	0.879666	0.879683	0.881204	0.879219	0.881204
2	RSS (Train)	0.093280	0.093843	0.095634	0.094655	0.095634
3	RSS (Test)	0.051236	0.051229	0.050581	0.051426	0.050581
4	MSE (Train)	0.009558	0.009587	0.009678	0.009628	0.009678
5	MSE (Test)	0.010816	0.010815	0.010746	0.010836	0.010746

Changes in Ridge metrics:

R2 score decrease if increase the value of alpha for train set.
R2 score almost same on test data.

Changes in Lasso metrics:

R2 score same for both train and test.

Important predictor variables after the change is implemented: -

Ridge

	Features	Coefficient
0	LotFrontage	-0.0015
1	LotArea	0.0011
2	OverallQual	0.0056
3	OverallCond	0.0039
4	BsmtQual	0.0018
5	BsmtExposure	0.0019
6	BsmtFinType1	0.0015
7	HeatingQC	0.0012
8	GrLivArea	0.0074
9	BsmtFullBath	0.0019

Lasso

	Features	Coefficient
0	LotFrontage	-0.0010
1	LotArea	0.0010
2	OverallQual	0.0064
3	OverallCond	0.0038
4	BsmtQual	0.0017
5	BsmtExposure	0.0016
6	BsmtFinType1	0.0014
7	HeatingQC	0.0012
8	GrLivArea	0.0077
9	BsmtFullBath	0.0018

After changing the alpha predictors remains the same and lasso will push the model coefficient towards 0. R2 score get decreases.

To handle the variances with slight compromise of bias. For both Ridge and lasso having same predictor variables.

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-

I will choose lasso there is less variance in train and test set. Over all model accuracy is good. RSE value for lasso is lower as compared to Ridge. Lasso also does variable selection. When lambda value is small it performs simple linear regression and as lambda value increases, shrinkage takes place and variables with 0 value are neglected by the model.

In Ridge over all accuracy is good but we can see variance in train and test data set. Ridge regression includes all variables in

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:-

```
Top 5 features in original lasso model:  ['GrLivArea', 'MSZoning_RL', 'OverallQual', 'AgeOfHouse', 'MSZoning_RM']
```

Shape after removal of top5 features

```
Shape of DF Before dropping top 5 variables (1460, 131)
Shape of DF after dropping top 5 variables (1460, 126)
```

Below are the top 5 features after removal of important predictor.

	Features	Coefficient	Abs_Coefficient_Lasso(Desc_Sort)
0	FullBath	0.0050	0.0050
1	GarageArea	0.0045	0.0045
2	OverallCond	0.0041	0.0041
3	FireplaceQu	0.0041	0.0041
4	TotalBsmtSF	0.0039	0.0039

```
R2Score_Train:  0.8725459178951727
R2Score_Test:  0.8425818197871623
RSS_Train:  0.12036010794045635
RSS_Test:  0.0670256281058195
MSE_Train:  0.00011788453275265068
MSE_Test:  0.0001530265481868025
```

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?

- A Model is robust if there is any variation in data does not a affect much performance. The model should be simple but at same time robust as well.
- A simple model is more robust and does not change significantly if the training data points undergo small changes.
- A generalisable models are bound to perform better on unseen data sets.
- If model is robust and generalizable, we have to make sure that, it doesn't overfit.
- Overfitted model has very high variance means small change in data affects the model prediction. Over fitted model doesn’t perform well on unseen data.
- Such a model mugs up pattern of training dataset and fail to make prediction on unseen data.
- We can understand this with Bias Variance Trade-Off
- Simpler the model more the bias but less variance and more generalizable.
- Its implication in terms of accuracy is that a robust and generalisable model will perform equally well on both training and test data i.e. the accuracy does not change much for training and test data.
- A too complex model will have a very high accuracy . So, to make our model more robust and generalizable, we will have to decrease variance which will lead to some bias penalty. Addition of bias means that accuracy will decrease.