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?

Answer:

1) As per the model built, the optimal values of alpha are as following:
   Ridge -100
   Lasso - 0.001

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.89	0.89
1	R2 Score (Test)	0.87	0.87

index	Ridge	Lasso
OverallQual	0.08343210604609204	0.10381494890417257
GarageCars	0.04600035939309428	0.0586321595079514 2
GrLivArea	0.04346135975818809	0.08394082107689524
1stFlrSF	0.03481995015160574	0.023515112751481444
OverallCond	0.034367112778828465	0.038876198959961236
YearBuilt	0.03226740713722118	0.04707735359672827
BsmtFullBath	0.025232555317118464	0.027760294784426504
SaleCondition	0.025225485692667124	0.027762954449588814
CentralAir	0.02139086521853214	0.020284278822243672
Fireplaces	0.021287907107799982	0.017644846384822468

By doubling the alpha values , we get the following scores:
   Ridge -200
   Lasso - 0.002

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.88	0.89
1	R2 Score (Test)	0.88	0.88

index	Ridge	Lasso
OverallQual	0.08343210604609204	0.10381494890417257
GarageCars	0.04600035939309428	0.0586321595079514 2
GrLivArea	0.04346135975818809	0.08394082107689524
1stFlrSF	0.03481995015160574	0.023515112751481444
OverallCond	0.034367112778828465	0.038876198959961236
YearBuilt	0.03226740713722118	0.04707735359672827
BsmtFullBath	0.025232555317118464	0.027760294784426504
SaleCondition	0.025225485692667124	0.027762954449588814
CentralAir	0.02139086521853214	0.020284278822243672
Fireplaces	0.021287907107799982	0.017644846384822468

2) Changes in the new model

Ridge Train score - decreased
Ridge Test score - increased

Lasso Train score - remained same
Lasso Test score - increased
Overall the test scores slightly increased on doubling the alpha values. This suggests the initial model had some underfitting.

3) The important predictors remained the same however, in the lasso model with double the alpha, the number of predictors penalised(made 0) were higher.

With original alpha values-

```
## View the features removed by Lasso
betas[betas['Lasso']==0]
```

	Ridge	Lasso
Condition1	0.00	-0.00
ExterQual	-0.01	-0.00
BsmtFinSF1	0.00	-0.00
BsmtFinType2	0.00	0.00
2ndFlrSF	0.02	0.00
GarageArea	0.01	-0.00
GarageCond	0.00	0.00

With Doubled alpha values-

```
## View the features removed by Lasso
betas2[betas2['Lasso']==0]
```

	Ridge	Lasso
Condition1	0.00	-0.00
Exterior1st	-0.00	-0.00
ExterQual	-0.01	-0.00
BsmtFinSF1	0.01	0.00
BsmtFinType2	0.00	0.00
BsmtUnfSF	0.00	0.00
2ndFlrSF	0.02	0.00
LowQualFinSF	0.00	0.00
FireplaceQu	-0.00	-0.00
GarageArea	0.02	0.00
GarageQual	-0.00	-0.00
GarageCond	0.00	0.00
MiscVal	-0.01	-0.00

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?

Answer:

The model to choose will depend on the use case.

If feature selection is the goal then I will choose the Lasso model with alpha=0.002

If the goal is to reduce the magnitude of the coefficients , I will choose Ridge regression with alpha =100

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?

Answer:

```
top5_lasso = ['OverallQual', 'GrLivArea', 'GarageCars', 'YearBuilt',
'OverallCond']
```

After removing the above predictors, the top predictors for the new model are

index
1stFlrSF
GarageArea
2ndFlrSF
YearRemodAdd
Fireplaces

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?

Answer :
For the model to be robust and generalised, we need to adopt strategies such as cross validation, regularisation , hyper-parameter tuning and some amount of data transformation such as log transformation etc
Generalisation of the model involves overcoming the overfitting problem and achieving the optimal bias-variance trade off.
In a generalised model the Initial accuracy for the model may be slightly low but the goal is to create a model that not only performs well on the data it was trained on but also maintains its accuracy when faced with new, unseen data.