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=>

The optimal value **of Ridge is 7.0 and Lasso is 0.0001.**

If we double the values of Ridge and Lasso we get below results

I) Ridge

When Ridge is doubled 14 =>R2_Train=>0.9449722923341373
R2_Test=>.8958823659193424
RSS=>7.012480634181597
MSE=>0.006868247437983935

When Ridge value is 7=> R2_Train=>0.9499272425912836
R2_Test=>0.8956990590558629
RSS=>6.381044323358727
MSE=>0.006249798553730389

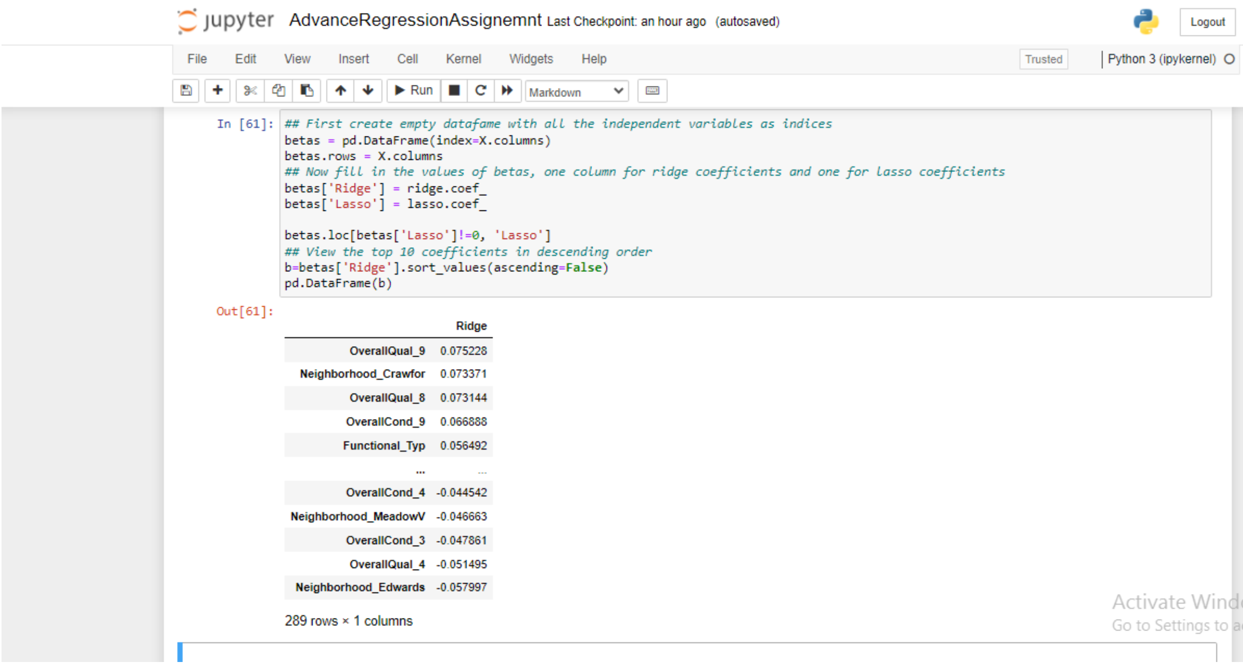
II) Lasso

When Lasso value is doubled 0.002=>R2_Train 0.9549356100064951
R2_Test=>0.869813747408074
RSS=>5.742800773013187
MSE=>0.0056246824417367156

When Lasso value is 0.001=>R2_Train=> 0.9588379295275048
R2_Test=0.8513472198556492
RSS=>5.245506932687617
MSE=>0.005137616976187675

Below is the list of significant predictors

OverallQual_9	0.075228
Neighborhood_Crawfor	0.073371
OverallQual_8	0.073144
OverallCond_9	0.066888
Functional_Typ	0.056492
Exterior1st_BrkFace	0.048813
Neighborhood_Somerst	0.048021
MSSubClass_70	0.047183
OverallCond_7	0.040676
Condition1_Norm	0.038663
BsmtCond_Gd	0.038453



In [61]:

```
## First create empty datafame with all the independent variables as indices
betas = pd.DataFrame(index=X.columns)
betas.rows = X.columns
## Now fill in the values of betas, one column for ridge coefficients and one for lasso coefficients
betas['Ridge'] = ridge.coef_
betas['Lasso'] = lasso.coef_

betas.loc[betas['Lasso']!=0, 'Lasso']
## View the top 10 coefficients in descending order
b=betas['Ridge'].sort_values(ascending=False)
pd.DataFrame(b)
```

Out[61]:

	Ridge
OverallQual_9	0.075228
Neighborhood_Crawfor	0.073371
OverallQual_8	0.073144
OverallCond_9	0.066888
Functional_Typ	0.056492
...	...
OverallCond_4	-0.044542
Neighborhood_MeadowV	-0.046663
OverallCond_3	-0.047861
OverallQual_4	-0.051495
Neighborhood_Edwards	-0.057997

289 rows × 1 columns

As per my observation there is no any significant change when we double the value of Ridge and Lasso.

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 2=>I will use Ridge when I have many correlated predictors and want to avoid multicollinearity and I still want to keep of option to self-eliminate the variables. I will use Lasso when feature selection is crucial, Which ever model helps me achieve simple model I will use that Regression, I personally like lasso regression but it depends on usecase.

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 3=> In my case below are the top predictors

['Condition2_PosA','OverallCond_9','OverallQual_9','SaleType_ConLD','Neighborhood_Crawfor'] after deleting these variables and rebuilding the model. I get below new variables

SaleCondition_Alloca	0.093274
Exterior1st_BrkFace	0.083785
MSSubClass_70	0.075167
SaleCondition_AdjLand	0.070221
MSZoning_FV	0.068726

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 4=>

1) A model is robust when any variation in the data does not affect its performance much.
2) A generalizable model is able to adapt properly to new, previously unseen data, drawn from the same distribution as the one used to create the model.
3) To make sure a model is robust and generalizable, we have to take care it doesn't overfit.
4) This is because an overfitting model has very high variance and a smallest change in data affects the model prediction heavily.
5) Such a model will identify all the patterns of a training data, but fail to pick up the patterns in unseen test data.
6) In other words, the model should not be too complex in order to be robust and generalizable. If we look at it from the perspective of Accuracy, 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. Addition of bias means that accuracy will decrease. In general, we have to find strike some balance between model accuracy and complexity. This can be achieved by Regularization techniques like Ridge Regression and Lasso.