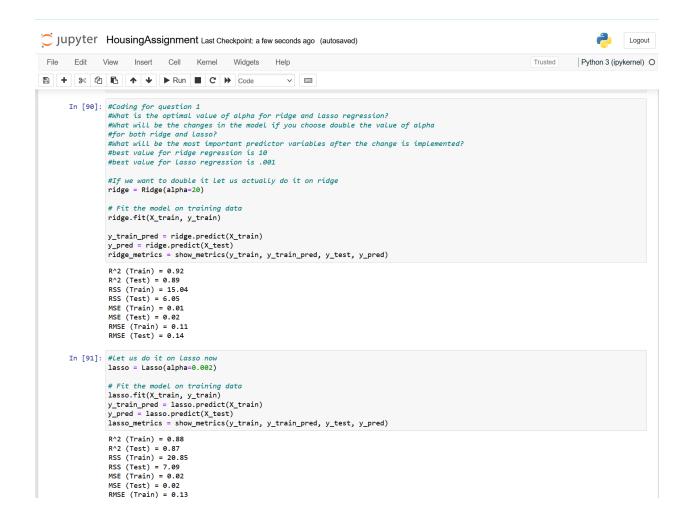
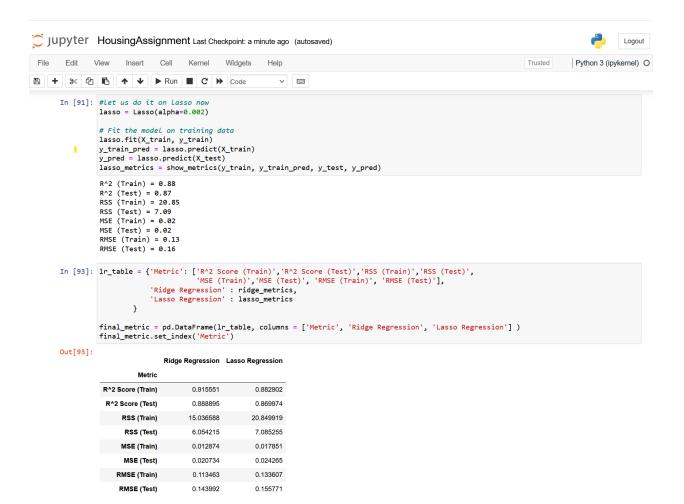
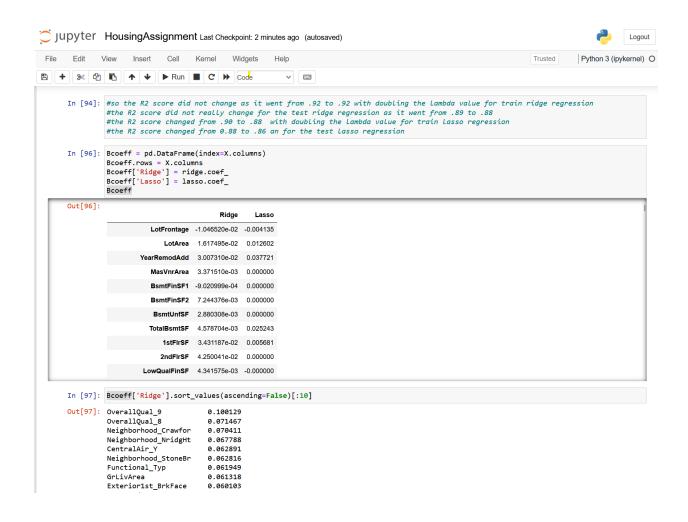
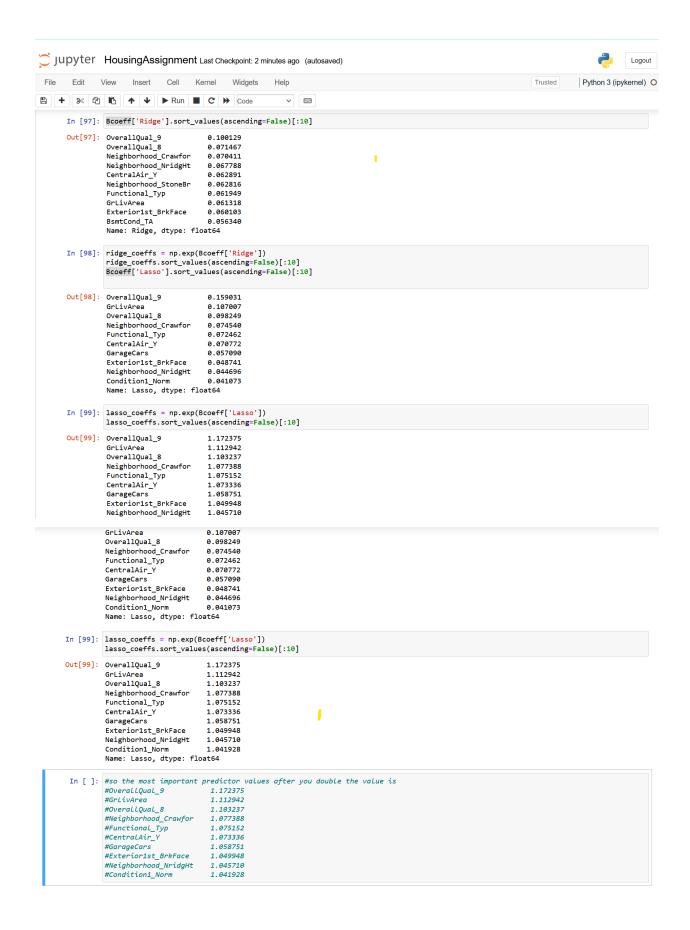
Coding for Question 1









Q1 Answer: So the R2 score did not change as it went from 0.92 to 0.92 with doubling the lambda value for train ridge regression. The R2 score did not really change for the test ridge regression as it went from 0.89 to 0.88. The R2 score changed from 0.90 to 0.88 with doubling the lambda value for train lasso regression. The R2 score changed from 0.88 to 0.86 an for the test lasso regression. The most important predictor variables now are

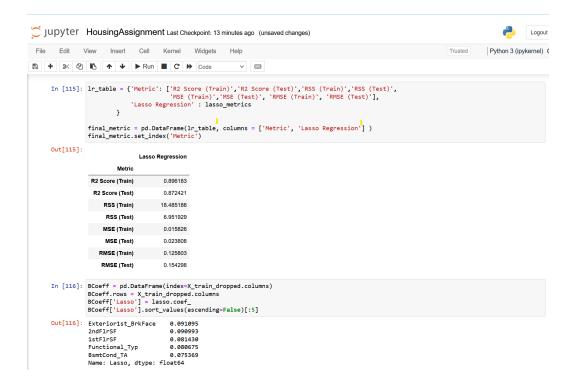
OverallQual 9 1.172375 GrLivArea 1.112942 OverallQual 8 1.103237 Neighborhood_Crawfor 1.077388 Functional Typ 1.075152 CentralAir Y 1.073336 GarageCars 1.058751 Exterior1st BrkFace 1.049948 Neighborhood NridgHt 1.045710 Condition1 Norm 1.041928

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?

I would choose to apply lasso regression as it finds the important predictor variables which is what this case was asking. Additionally, it reduces the amount of important predictor variables by taking out the non important ones.

Question 3



So the most important predictor variables after taking out the top 5 are:

Exterior1st_BrkFace 0.091095

 2ndFIrSF
 0.090993

 1stFIrSF
 0.081430

 Functional_Typ
 0.080675

 BsmtCond_TA
 0.075369

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 can be ensured that it is robust and generalisable by making sure there is thorough data preparation. Making sure there are no NA values or missing values that could affect the analysis is very important. Additionally, finding the right lambda value for lasso and ridge regression is important to control regularization and to stop overfitting. Finally, cross validation is important to make sure that the training data split into folds and evaluated to prevent overfitting.

To ensure accuracy, make sure the model is not overfit as it will not work on new test data. Additionally, not making the model too complex is helpful.