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 : The optimal value of alpha for ridge is 2 and that for lasso regression is 0.001. The accuracy of the model remains same almost if we choose double the value of alpha for both ridge and lasso with a small change in coefficient values. This has been demonstrated in the attached Jupiter notebook. With that, we could expect not much change due to doubling the coefficient in the prediction model.

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 : From the analysis, we found that the optimum Lambda for Lasso = 0.001 and that for Ridge = 2. However, the mean squared error came out to be almost same for both of the model i.e. approx. 0.0018. since, in this case Lasso helps through its very property of "feature selection" along with penalising the coefficient, we could use Lasso over Ridge in the model for this use case.

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 : As per the solution in the notebook, the five top predictors are as following :

Total_sqr_footage	0.204642
GarageArea	0.103822
TotRmsAbvGrd	0.064902
OverallCond	0.042168
CentralAir_Y	0.033113

After removing the above predictors, when the Lasso model is built, the following are the new important variables :

LotFrontage
Total_porch_sf
HouseStyle_2.5Unf
HouseStyle_2.5Fin
Neighborhood_Veenker