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 – Optimal value of alpha for ridge is '7' and for lasso is '0.0004'.On doubling the alphas the feature coefficients shifted more towards '0'. 'Overallqual' and 'GrliveArea' are most important predictors

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 lasso since the model is simplified by eliminating more than 50% of the features and yet yielding the similar results as ridge.

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 – Updated model without top five predictors gives following top predictors:

- PoolQC_Gd
- 1stFlrSF
- Condition2_PosN
- 2ndFlrSF
- SaleCondition_Alloca

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 – r2 value on training and test datasets are very close which shows that the model is robust and generalisable. RSME value is sufficiently low which shows that the model is significantly accurate.