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?

Ans: Optimal Value of Alpha for ridge = 0.05 and for lasso = 0.0001.

Ridge regression- It enhanced Test R2 score

Actual Optimal alpha = 0.05, Train R2 = 0.86, Test R2 = 0.80

Twice Optimal alpha = 0.1 , Train R2 = 0.86, Test R2 = 0.82

Lasso regression- It made model underfit

Actual Optimal alpha = 0.0001, Train R2 = 0.84, Test R2 = 0.836

Twice Optimal alpha = 0.0002 , Train R2 = 0.83, Test R2 = 0.84 (Underfit)

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?

Ans: Lasso since lasso regression is giving better R2 score on test data compared to Ridge Regression.

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?

Ans: 'GrLivArea', 'OverallQual', 'MasVnrArea', 'YearBuilt', 'Fireplaces'

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?

Ans: Model is robust and generalisable if it is performing well on unseen data and not too complex. When the model is robust and generalisable it gives good accuracy on unseen data since it has captured most of patterns during training of model.