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 values of alpha for ridge and lasso regression are **100** and **0.0001 r**espectively**.** As the alpha value increases/ doubles the variance becomes less and the bias becomes more, which leads to underfitting of the model. The most Important predictor is **OverallQual**, which rates the overall material and finish of the house.

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: The Mean Squared Error (**MSE**) for ridge and lasso regression are 0.**0184** and 0.**0174** respectively. The lower the MSE value, the better the model. Hence choosing Lasso Regression Model.

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: **GrLivArea**, **OverallQual, MSZoning_RL, MSZoning_FV and GarageCars** would be the five most important predictor variables.

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: A model needs to be made robust and generalizable, so that there is no implication of the outliers in the training data. The model should also be generalisable, so that the model can perform better on the unseen data. The model should be accurate in prediction and we should expect to have a model which can predict the target variable optimally.