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?**
   I chose alpha=0.1 for the lasso regression and alpha=10 for ridge regression.
   After doubling the values, model performance remains the same for ridge regression. But, for Lasso Regression model, the r2 value get decreased a lot.
   After the changes,
   For lasso regression, the most important predictor variables are
   - GrLivArea
   - TotalBsmtSF
   - LotArea
   - HalfBath
   - KitchenAbvGr
   - GarageType
   - FireplaceQu
   - Fireplaces
   - Functional
   - TotRmsAbvGrd

   For ridge regression, the most important predictor variables are
   - PoolQC
   - OverallQual
   - BsmtFullBath
   - GarageCars
   - OverallCond
   - FullBath
   - Fireplaces
   - KitchenQual
   - BldgType
   - SaleCondition

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?**
   In the process of model selection between Ridge and Lasso regression, both approaches demonstrated comparable performance on the test dataset. However, the decision leans toward favoring the Lasso regression model due to its inherent feature elimination capability. Lasso's effectiveness in shrinking certain coefficients to zero is particularly advantageous for enhancing model interpretability. Therefore, the final choice is Lasso regression, selected for its ability to strike a balance between predictive accuracy and feature selection without specifying numerical metrics. This decision aims to streamline the model and improve its interpretability without relying on specific performance scores.

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?**

The five most important predictor variables in our Lasso model are:

GrLivArea

TotalBsmtSF

LotArea

HalfBath

KitchenAbvGr

After creating another model excluding these top 5 predictor variables, the five most important variables now are:

SaleType_New

GarageType_Attchd

Neighborhood_StoneBr

TotRmsAbvGrd

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?**

Ensuring the robustness and generalizability of a model involves preventing overfitting and maintaining simplicity. If a model is overly complex, it may perform exceptionally well on the training data but struggle to generalize to new, unseen data. The accuracy of a model tends to increase when it is overfitted to the training set; however, this compromises its ability to generalize. A truly robust model should exhibit strong accuracy not only on the training dataset but also on a separate testing dataset. By striking a balance between complexity and performance on both training and testing data, we ensure that the model is robust, capable of providing accurate predictions across various scenarios, and not simply memorizing the training data.