Subjective 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 to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

   Answer:
   - Alpha Value for Ridge is 10 and
   - Alpha Value for Lasso is 0.001

   When we double the value of Ridge alpha i.e., 10 * 2, this implies more penalty on the curve. The model becomes more generalized means not fitting any of the data from the data set.

   When we double the value of Lasso alpha i.e., 0.001 * 2, this implies more penalization. Most of the co-efficient values become 0. R2 square also decreases.

   The most important predictor variables after the change (Predictors are the same but there is a change in the co-efficient value)

   - FireplaceQu
   - OverallQual
   - OverallCond
   - FullBath
   - GrLivArea
   - GarageCars

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 r2 square is slightly higher for the Ridge. So, I will go with Ridge regression.

3. After building the model, you realize 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: Most important predictor variables for Lasso Regression
   - 2ndFlrSF,
   - TotalBsmtSF
   - KitchenQual
   - ExterQual
   - BsmtQual

   (Detail code solution added in the Python book)

4. How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

   Answer: The following considerations help to build the robust and generalizable model:
   - Proper EDA and Data cleaning
   - Outliers' treatment well before splitting train and test data.

- You should adjust the parameters and hyperparameters to find the optional values that have minimized error and maximized performance.
- Regularization techniques such as lasso, ridge, etc. to reduce the complexity of your model and prevent overfitting.
- Test accuracy should not be lesser than the training accuracy score.

These all considerations help to build a robust and general model