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:
In the case of ridge regression, we observe that as alpha increases from 0, the negative mean absolute error decreases, and the train error exhibits an upward trend with higher alpha values. After analyzing the curve, we found that the test error is minimized when alpha is set to 2. Therefore, we opted to use alpha equal to 2 for our ridge regression model.

As for lasso regression, we chose a very small alpha value of 0.01. When we increase alpha, the model tends to penalize coefficients more, leading to the sparsity of the coefficient values, with many of them becoming zero. Initially, the negative mean absolute error was around 0.4 for alpha.

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:
Regularizing coefficients is crucial to enhance prediction accuracy, reduce variance, and improve model interpretability. Ridge regression employs a tuning parameter called lambda, penalizing the square of the magnitude of coefficients, which is determined through cross-validation. This penalty helps minimize the residual sum of squares, making coefficients with higher values receive more significant penalties. By increasing lambda, the model's variance decreases while the bias remains constant. Unlike Lasso Regression, Ridge regression includes all variables in the final model.

On the other hand, Lasso regression also utilizes a tuning parameter (lambda), but its penalty is the absolute value of the magnitude of coefficients, identified through cross-validation. As lambda increases, Lasso shrinks the coefficients towards zero, effectively setting some variables exactly equal to 0. This process allows Lasso to perform variable selection, eliminating variables with zero values from the model as lambda grows. When lambda is small, Lasso behaves like simple linear regression, but with increasing lambda values, shrinkage occurs, leading to the exclusion of variables with negligible impact on the model.

Question-3:
After building the model, you realized 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:
Those 5 most important predictor variables that will be excluded are :-
1. GrLivArea
2. OverallQual
3. OverallCond
4. TotalBsmtSF
5. GarageArea

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:
The objective is to keep the model as simple as possible, even if its accuracy might decrease. The benefit of simplicity lies in the model's robustness and generalizability. This principle aligns with the Bias-Variance trade-off, where a simpler model tends to have higher bias but lower variance, making it more generalizable. Consequently, a robust and generalizable model performs equally well on both training and test data, exhibiting minimal changes in accuracy between the two.

Bias refers to the model's error when it struggles to learn from the data, resulting in poor performance on both training and testing data. On the other hand, variance indicates the model's error when it overlearns from the data, leading to exceptional performance on the training data but poor performance on unseen testing data.

To strike the right balance and avoid overfitting or underfitting, it is essential to find the optimal trade-off between bias and variance in the model. This ensures that the model performs well on both the data it was trained on and new, unseen data.

~~**Question-2:**~~
~~Repeat the above procedure.~~