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?**

   → The optimal value of alpha for ridge is 2 and for lasso is 0.0006. With these values the R2 of ridge is 0.84 and lasso is 0.71. After doubling the alpha values the R2 value of ridge decreased from 0.84 to 0.71 and the R2 value of Lasso decreased from 0.71 to 0.69. There is no significant change is seen in the coefficients of each model as seen in the ridge coefficients below. The new model is created and demonstrated in the Jupyter notebook.

	Ridge Co-Efficient
Neighborhood_NridgHt	0.273916
Total_porch_sf	0.237259
Neighborhood_NoRidge	0.221662
Neighborhood_Crawfor	0.201495
Neighborhood_StoneBr	0.198474
BsmtExposure_Gd	0.136391
OpenPorchSF	0.126886
Exterior1st_BrkFace	0.121172
Condition2_PosA	0.101061
RoofStyle_Hip	0.098515
HouseStyle_2Story	0.089006
Exterior1st_CemntBd	0.080674
Neighborhood_NWAmes	0.078999
SaleCondition_Alloca	0.078835
HouseStyle_2.5Fin	0.076579
RoofMatl_WdShngl	0.074949
SaleCondition_Partial	0.074412
LotConfig_CulDSac	0.073452
Exterior1st_Stucco	0.068428
Neighborhood_BrkSide	0.066988

	Ridge Doubled Alpha Co-Efficient
Neighborhood_NridgHt	0.284417
Total_porch_sf	0.240801
Neighborhood_NoRidge	0.233808
Neighborhood_StoneBr	0.213485
Neighborhood_Crawfor	0.210620
BsmtExposure_Gd	0.137013
OpenPorchSF	0.128831
Exterior1st_BrkFace	0.124436
Condition2_PosA	0.120391
RoofStyle_Hip	0.097365
HouseStyle_2Story	0.087783
SaleCondition_Alloca	0.087734
HouseStyle_2.5Fin	0.087545
Exterior1st_CemntBd	0.086992
RoofMatl_WdShngl	0.084805
Neighborhood_NWAmes	0.080865
LotConfig_CulDSac	0.074998
SaleCondition_Partial	0.073052
Neighborhood_Somerst	0.072816
Exterior1st_Stucco	0.072246

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?**

   ➔ The optimum value of lambda in case of ridge is 2 and for lasso is 0.0006. We can see the R2 value of ridge is 0.84 and R2 value of lasso is 0.71 and also the Mean Square Error for ridge is 0.025 and for lasso is 0.048. Thus considering all of the factors above I will be choosing the ridge model as it is the most significant among the two.

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 the lasso model are:
   - Neighborhood_NridgHt
   - Total_porch_sf
   - Neighborhood_NoRidge
   - Neighborhood_Crawfor
   - Neighborhood_StoneBr

We build a Lasso model without the top 5 predictors after removing them from the dataset.

The R2 of the new model is 0.65 and Mean Squared Error is 0.057 which is quite high.

The new top 5 predictors are:

- Condition2_PosA
- OpenPorchSF
- Electrical_Mix
- RoofMatl_WdShngl
- HousStyle_2.5Fin

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?**

➔ We know that simple model are more robust and generalizable,but choosing just simple model is not enough. We should test the accuracy as well on train and test dataset, and if there are two model which perform equally well on test dataset then we should always go for the simple model,as simpler model is more robust and generalizable.
Also,Making a model simply leads to Bias-Variance Trade-off:

- A complex model will need to change for every little change in the dataset and hence is very unstable and extremely sensitive to any changes in the training data.
- A simpler model that abstracts out some pattern followed by the data points given is unlikely to change wildly even if more points are added or removed.

Bias quantifies how accurate is the model likely to be on test data. A complex model can do an

Accurate job prediction provided there is enough training data. Models that are too naïve,

For e.g, one that gives same answer to all test inputs and makes no discrimination whatsoever

Has a very large bias as its expected error across all test inputs are very high.

Variance refers to the degree of changes in the model itself with respect to changes in the

Training data.

Thus accuracy of the model can be maintained by keeping the balance between Bias and

Variance as it minimizes the total error as shown in the below graph.