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 value of Ridge and Lasso Regressions are :-

Regression Model	Alpha Values	Train R2 score	Test R2 score
Ridge	1	0.9325	0.8807
Lasso	0.001	0.9044	0.8808

Top 5 Predictors in Ridge and Lasso Regressions - Before

	Ridge
RoofMatl_WdShngl	0.2542
RoofMatl_CompShg	0.1889
Neighborhood_StoneBr	0.1733
RoofMatl_Membran	0.1374
Neighborhood_NridgHt	0.1303

	Lasso
Neighborhood_Crawfor	0.1274
Neighborhood_StoneBr	0.1125
Neighborhood_NridgHt	0.1074
OverallQual	0.0665
GarageCars	0.0654

After Doubling the alpha value of Ridge and Lasso Regressions the values are :-

Regression Model	Alpha Values	Train R2 score	Test R2 score
Ridge	2	0.9295	0.8800
Lasso	0.002	0.8903	0.8801

Top 5 Predictors in Ridge and Lasso Regressions - After

	Ridge
RoofMatl_WdShngl	0.1618
Neighborhood_StoneBr	0.1613
Neighborhood_NridgHt	0.1257
RoofMatl_CompShg	0.1134
Neighborhood_Crawfor	0.1130

	Lasso
Neighborhood_Crawfor	0.0920
OverallQual	0.0738
Neighborhood_NridgHt	0.0657
GarageCars	0.0594
Condition1_Norm	0.0430

If we double the alpha values in Ridge Regression then the L2 regularization penalty increases which will shrink the coefficients towards zero but not to zero. The model may become more biased towards the mean which can result in lower variance and overfitting reduction

If we double the alpha values in Lasso Regression then the L1 regularization penalty increases which can set the coefficients to zero. This leads to more feature selection and sparsity in the model. The model may become more biased which can result in higher variance and underfitting.

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 Optimal Values of Ridge and Lasso regressions are :-

Model	Optimal Alpha
Ridge	1
Lasso	0.001

The Training and testing scores are also n nearly same

Regression Model	Train R2 score	Test R2 score	MSE Train	MSE Test
Ridge	0.9325	0.8807	0.099	0.150
Lasso	0.9044	0.8808	0.118	0.150

I will choose Lasso regression because the Train and Test Scores are nearer than Ridge regression. Also Lasso regression will not consider all the variables.

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 :-
The top most variables after dropping the five most important predictor variables are

	Lasso
Functional_Typ	0.0847
LandContour_HLS	0.0786
SaleType_New	0.0662
Condition1_Norm	0.0600
Exterior1st_BrkFace	0.0594

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 implications of ensuring that a model is robust and generalizable are that it will perform well on new, unseen data that was not used during training. This is important because

models are typically deployed in real-world scenarios where the input data can change over time or come from different sources. A model that is robust and generalizable will be better equipped to handle such situations and make accurate predictions or classifications. However, achieving high levels of robustness and generalization often comes at the expense of accuracy, as the model may be less complex or constrained than a more specialized model that overfits the training data. It is therefore important to strike a balance between accuracy and robustness/generalization depending on the specific use case and requirements.

In our case the difference between training and testing R2 score is small, So we can consider that our model is Robust and generalisable.