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?

The best alpha value for Ridge is 1, and the best alpha value for Lasso is 0.00009. With these alphas, the model's R2 was roughly 0.84.

The prediction accuracy remains at 0.84 after doubling the alpha values in the Ridge and Lasso, although there is a

The values of the coefficients changed hardly. The Jupiter notebook contains the new model's creation and demonstration. The co-efficient modifications are listed below.

	Ridge Co-Efficient
OverallQual	0.150680
Total_Bathrooms	0.077448
TotRmsAbvGrd	0.077298
LotArea	0.061657
OverallCond	0.050075
GarageArea	0.048209
GarageCars	0.045955
Total_porch_sf	0.042586
Fireplaces	0.036753
YrBltAndRemod	0.031608
BsmtQual	0.028832
KitchenQual	0.025950
ExterQual	0.022895
HeatingQC	0.019662
MasVnrArea	0.019293
CentralAir	0.016066
GarageCond	0.015683
LandSlope	0.015425
LotFrontage	0.014829
GarageQual	0.014406

	Lasso Co-Efficient
OverallQual	0.171306
TotRmsAbvGrd	0.082627
Total_Bathrooms	0.078237
LotArea	0.070260
GarageArea	0.056832
OverallCond	0.046498
Total_porch_sf	0.040009
Fireplaces	0.037228
GarageCars	0.031538
YrBltAndRemod	0.028666
BsmtQual	0.026465
KitchenQual	0.023775
LotFrontage	0.022855
HeatingQC	0.017679
ExterQual	0.017392
CentralAir	0.017140
MasVnrArea	0.014516
BsmtFinType1	0.009526
LandSlope	0.009073
BsmtExposure	0.007747

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?

▪ The optimum lambda value in case of Ridge and Lasso is as follows:-

• Ridge – 1

• Lasso – 0.00009

▪ The Mean Squared Error in case of Ridge and Lasso are:

• Ridge - 0.0014664101752543564

• Lasso - 0.0015156112834320717

▪ The Mean Squared Error of both the models are almost same. ▪ Since Lasso helps in feature reduction (as the coefficient value of some of the features become zero), Lasso has a better edge over Ridge and should be used as the final model.

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?

	Lasso Co-Efficient
LotFrontage	0.146535
Total_porch_sf	0.072445
HouseStyle_2.5Unf	0.062900
HouseStyle_2.5Fin	0.050487
Neighborhood_Veenker	0.042532

Question 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?

According to Occam's Razor, if two models exhibit similar "performance" in finite training or test data, we should choose the one that makes fewer on the test data for the reasons listed below:-

Simpler models are typically more "generic" and have a wider range of applications.

Less complex models can be effectively trained with fewer training data, making simpler models simpler to train.

More reliable models are simpler ones.

o Simple models have high bias and low variance, while complicated models have high bias and low variance when the training data set is changed.

More mistakes are made in the training set by simpler models. Overfitting occurs when complex models are used, which causes them to perform admirably on training examples but horribly on subsequent test samples.

In order to make the model more reliable and generalizable, simplify it without making it uselessly simpler.

The model can be made easier by regularisation. Regularization aids in striking the difficult balance between keeping the model straightforward and preventing it from being overly simplistic and useless. Regression regularisation entails multiplying the squares or absolute values of the model's parameters by a regularisation term, which is added to the cost.

Additionally, a model's bias-variance trade-off results from simplification.

• A complicated model is particularly unstable and highly sensitive to any changes in the training data because it must be changed for any tiny change in the dataset.

• Even if additional data points are added or subtracted, a more basic model that abstracts out any pattern revealed by the provided data points is unlikely to change drastically.

How accurate the model is expected to be on test data is quantified by bias. If there is enough training data, a complicated model can provide an accurate job forecast.

A. Models that are too naive, such as those that respond the same to every test input and make no distinction at all, have a very big bias since their predicted error for every test input is quite high.

Variance describes how much the model has changed compared to how the training set has changed.

Thus, by maintaining the balance between Bias and Variance, which reduces the overall error as illustrated in the graph below, the correctness of the model may be maintained.