Question1: 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?

Answer1: In the case of ridge regression:- When we plot the curve between negative mean absolute error and alpha we see that as the value of alpha increase from 0 the error term decrease and the train error is showing increasing trend when value of alpha increases .when the value of alpha is 2 the test error is minimum so we decided to go with value of alpha equal to 2 for our ridge regression.

For lasso regression I have decided to keep very small value that is 0.01, when we increase the value of alpha the model try to penalize more and try to make most of the coefficient value zero. Initially it came as 0.4 in negative mean absolute error and alpha.

When we double the value of alpha for our ridge regression no we will take the value of alpha equal to 10 the model will apply more penalty on the curve and try to make the model more generalized that is making model more simpler and no thinking to fit every data of the data set .from the graph we can see that when alpha is 10 we get more error for both test and train.
Similarly when we increase the value of alpha for lasso we try to penalize more our model and more coefficient of the variable will reduced to zero, when we increase the value of our r2 square also decreases.
The most important variable after the changes has been implemented for ridge regression are as follows:-

1. MSZoning_FV
2. MSZoning_RL
3. Neighborhood_Crawfor
4. MSZoning_RH
5. MSZoning_RM
6. SaleCondition_Partial
7. Neighborhood_StoneBr
8. GrLivArea
9. SaleCondition_Normal
10. Exterior1st_BrkFace

The most important variable after the changes has been implemented for lasso regression are as follows:-

1. GrLivArea
2. OverallQual
3. OverallCond
4. TotalBsmtSF
5. BsmtFinSF1
6. GarageArea
7. Fireplaces
8. LotArea
9. LotArea
10. LotFrontage

Question2: 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?

Answer2: It is important to regularize coefficients and improve the prediction accuracy also with the decrease in variance, and making the model interpretably.

Ridge regression, uses a tuning parameter called lambda as the penalty is square of magnitude of coefficients which is identified by cross validation. Residual sum or squares should be small by using the penalty. The penalty is lambda times sum of squares of the coefficients, hence the coefficients that have greater values gets penalized. As we increase the value of lambda the variance in model is dropped and bias remains constant. Ridge regression includes all variables in final model unlike Lasso Regression.

Lasso regression, uses a tuning parameter called lambda as the penalty is absolute value of magnitude of coefficients which is identified by cross validation. As the lambda value increases Lasso shrinks the coefficient towards zero and it make the variables exactly equal to 0. Lasso also does variable selection. When lambda value is small it performs simple linear regression and as lambda value increases, shrinkage takes place and variables with 0 value are neglected by the model.

Question3 : 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?

Answer3:

Metric	R2 Score (Train)	R2 Score (Test)	RSS (Train)	RSS (Test)	MSE (Train)	MSE (Test)
Linear Regression	0.9	0.84	102.25	72.88	0.32	0.41
Ridge Regression	0.89	0.85	110.0	67.41	0.33	0.39
Lasso Regression	0.89	0.85	109.62	67.73	0.33	0.39
Ridge Regression Double Lambda	0.89	0.85	117.35	67.29	0.34	0.39
Lasso Regression Double Lambda	0.88	0.85	119.47	67.5	0.34	0.39
Ridge Regression Drop Top 5	0.87	0.84	135.18	74.23	0.36	0.41
Lasso Regression Drop Top 5	0.87	0.84	132.13	72.51	0.36	0.41

- There is a drop in r2 values for train and test data
- There is increase in RSS and MSE values for train and test data The Top 5 predictors after changes.

Using Ridge :

- 'OverallQual_5'
- 'OverallQual_6',
- 'OverallQual_4'
- 'KitchenQual_TA',

- ‘KitchenQual_Fa'

Using Lasso :

- 'OverallQual_5’
- 'OverallQual_4'
- 'OverallQual_6'
- 'OverallQual_3'
- 'Fireplaces_3'

Question4 : 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?

Answer4:
To make sure a Model can learn well, we take a few important steps:

Using L1 and L2 to Improve Models:
We can use special techniques (L1 and L2 regularization) to make our models strong and accurate.

- What Makes a Good Model:

A good model stays accurate even if some things change in the input (robust) and it works well for a wide range of situations (generalizable).

- Simpler is Better:

Simpler models, with fewer complicated parts, are usually better and work well in many situations.

- Measuring Model Complexity:

We can measure how complicated a model is by looking at how many pieces it has, how much math it uses (like polynomials), or how big it is (like decision trees).

- Balancing Accuracy and Complexity:

We want our model to be just right- not too complicated (low complexity) and also not too simple, finding a good balance for accurate predictions.

- Trade-off:

There's a balance between how well our model fits our training data (accuracy on train set) and how well it works on new data (accuracy on test set).

- What Happens When We Simplify:

When we make our model simpler, it might not be as accurate on the training data, but it could get better at predicting on new, unseen data.

Implications:

- The accuracy of the model on train set might go down
- The accuracy of the model on test set might go up