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:
The optrimal values for Ridge and Lasso regression as seen above are :

1. Ridge regression : 10
2. Lasso regression : 0.001

For ridge regression, we see that the negative mean absolute error decreases as alpha increases from 0, while the train error increases as alpha increases, so we choose alpha as 10 for our ridge regression.
For lasso regression, the model does not penalize much and keeps most of the coefficients with a very small value of 0.001. The negative mean absolute error was 0.8 for this alpha value. For ridge regression, I have used a higher value of alpha, 10, which makes the model simpler and more general by penalizing more. The graph shows that this alpha value increases the errors for both test and train. For lasso regression, increasing alpha also simplifies and generalizes the model by penalizing more and reducing more coefficients to zero, which reduces the r2 score.

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

```
MSZoning_RL                 0.084193
Neighborhood_Crawfor        0.075269
GrLivArea                   0.072783
OverallQual                 0.064186
Neighborhood_StoneBr        0.063769
Neighborhood_Somerst        0.062651
SaleCondition_Normal        0.062570
OverallCond                 0.058242
SaleCondition_Partial       0.058239
CentralAir_Y                0.054267
```

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

```
GrLivArea                 0.123538
SaleCondition_Partial     0.105583
Neighborhood_Crawfor      0.077803
OverallQual               0.077141
Neighborhood_Somerst      0.075884
MSZoning_RL               0.064558
OverallCond               0.061317
SaleCondition_Normal      0.049859
CentralAir_Y              0.048754
Condition1_Norm           0.046582
Name: Lasso, dtype: float64
```

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:

	Variable	Coeff
0	constant	11.748
31	MSZoning_RL	0.084
50	Neighborhood_Crawfor	0.075
13	GrLivArea	0.073
66	Neighborhood_StoneBr	0.064
4	OverallQual	0.064
208	SaleCondition_Normal	0.063
65	Neighborhood_Somerst	0.063
209	SaleCondition_Partial	0.058
5	OverallCond	0.058

- 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.

Looking at Ridge and Lasso together:

	Ridge	Lasso
MSSubClass	-0.011551	-0.005153
LotFrontage	0.008587	0.009616
LotArea	0.016726	0.015715
OverallQual	0.064186	0.077141
OverallCond	0.058242	0.061317
MasVnrArea	0.003663	0.004365
BsmtFinSF1	0.027161	0.031170
BsmtFinSF2	0.009169	0.003095
TotalBsmtSF	0.038708	0.036913
1stFlrSF	0.040641	0.005117

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