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 changes made in the model if we choose to double the value of alpha for both ridge and lasso are as follows

We have alpha values for Ridge and Lasso:

1. Rigde : 2

2. Lasso : 0.01

If we double the alpha values:

For Ridge,

There is a slight increase in the mean square error. R2 for train and test remains same.

For Lasso,

There is a slight increase in the mean square error. R2 value of train slightly decreases. There is a huge fall in R2 values for test. And worsening in model prediction.

It also penalizes the model even more & a greater number of coefficients of a variable shrink towards zero.

The most important predictor variables after the changes are implemented are as follows:

For Ridge :

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 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

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 optimal value of LAMBDA we have in case of Ridge and Lasso is:

1. Ridge - 2.0

2. Lasso - 0.01

The r2 value we have in case of Ridge and Lasso is:

Ridge - Train = 0.936, Test = 0.907, difference – 0.029

Lasso - Train = 0.885, Test = 0.889, difference –0.004

The RMSE in case of Ridge and Lasso is:

Ridge – 0.114

Lasso - 0.125

Also difference of r2 between train & test is less in lasso when as compared to ridge.

Also, since lasso helps in feature reduction (as the coefficient value of one of the lasso's features to be shrunk toward 0) and helps to increase model interpretation by taking the magnitude of the coefficients thus lasso has a better edge over ridge.

3. After building the model, you realized 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?

We will exclude the following 5 predictor variables (explanation in the end section of jupyter notebook):

1. GrLivArea
2. OverallQual
3. OverallCond
4. TotalBsmtSF
5. GarageArea

We will have the following new variables (explanation in the end section of jupyter notebook):

	Coefficient
1stFlrSF	0.163463
2ndFlrSF	0.134380
BsmtFinSF1	0.052595
Fireplaces	0.027523
Foundation_PConc	0.026834

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?

We can have a model with simplicity as much as possible.But accuracy can decrease slightly, but it is generalizable model. Also it can be understood using the Bias-Variance trade-off. The simpler the model the more the bias but less variance and more generalizable. Its implication in terms of accuracy is that a robust and generalizable model will perform equally well on both training and test data i.e. the accuracy does not change much for training and test data.

Bias: Bias is that, error in model, when the model is weak to learn from the data. High bias means model is unable to learn details in the data. Model performs poor on training and testing data.

Variance: Variance is error in model, when model tries to over learn from the data. High variance means model performs exceptionally well on training data as it has very well trained on this of data but performs very poor on testing data as it was unseen data for the model.

It is important to note that we have to have balance in Bias and Variance to avoid overfitting and under-fitting of data.