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?

We found that the best lambda values for Ridge and Lasso are different: 2.0 for Ridge and 0.0001 for Lasso.

If we were to double the alpha values for both Ridge and Lasso:

For Ridge:

- **There's a small increase in the mean squared error.**
- **The R-squared values for both training and testing data stay almost the same.**

For Lasso:

- **There's a slight increase in the mean squared error.**
- **The R-squared value for the training data decreases slightly.**
- **The R-squared value for the testing data decreases significantly, indicating poorer predictions.**
- **This higher alpha penalizes the model more, causing more coefficients to shrink towards zero.**

Ridge:

1. **Total_sqr_footage**

2. **OverallQual**

3. **GrLivArea**

4. **Neighborhood_StoneBr**

5. **OverallCond**

6. TotalBsmtSF

7. LotArea

Lasso:

1.Total_sqr_footage

2. OverallQual

3. YearBuilt

4. GrLivArea

5. Neighborhood_StoneBr

6. OverallCond

7. LotArea

8. Neighborhood_Crawfor

9. Neighborhood_NridgHt

10. GarageCars

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?

We found the best lambda values for Ridge and Lasso:

- Ridge: 2.0
- Lasso: 0.0001

The R-squared values we obtained are:

- Ridge: Train = 0.930, Test = 0.896 (Difference = 0.046)
- Lasso: Train = 0.927, Test = 0.902 (Difference = 0.025)

The Mean Squared Error for Ridge and Lasso is:

- Ridge: 0.00297
- Lasso: 0.00280

We noticed that Lasso has slightly lower Mean Squared Error compared to Ridge. Additionally, the difference between the R-squared values of the training and testing data is smaller for Lasso than for Ridge.

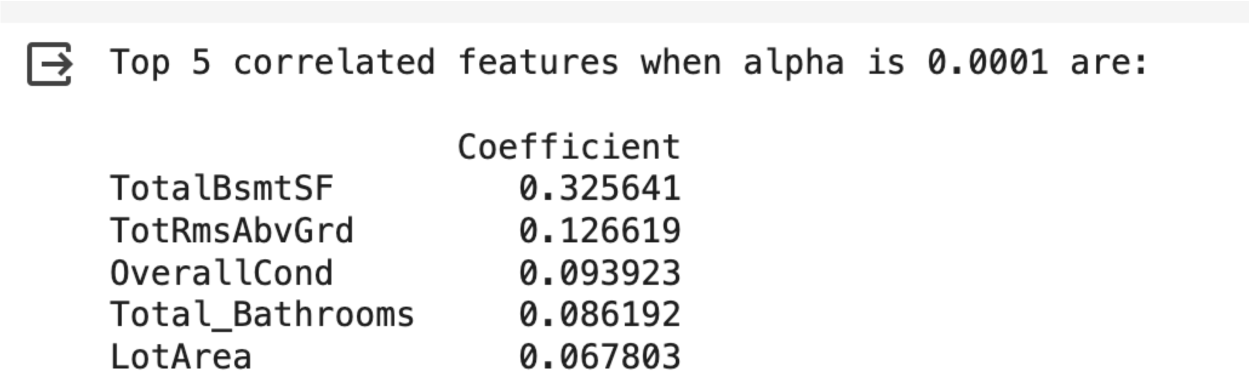
Moreover, Lasso's feature reduction capability and its enhancement of model interpretability due to coefficient magnitude consideration give it an edge over Ridge.

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?

After removing the top 5 most influential predictor variables in the Lasso model, we reanalyzed and identified a new set of top 5 predictors:

1. TotalBsmtSF
2. TotRmsAbvGrd
3. OverallCond
4. Total_Bathrooms
5. LotArea



```
Top 5 correlated features when alpha is 0.0001 are:

                 Coefficient
TotalBsmtSF         0.325641
TotRmsAbvGrd        0.126619
OverallCond         0.093923
Total_Bathrooms     0.086192
LotArea             0.067803
```

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?

To make sure a model works well in different situations, we need to:

1. **Split the data into two parts: one for training the model and one for testing it.**
2. **Test the model on different parts of the data multiple times to see if it works consistently.**
3. **Choose only the important information for the model and use techniques to prevent it from becoming too focused on specific details.**
4. **Adjust the settings of the model to make sure it's not too focused on the training data.**
5. **Finally, check if the model works well on completely new data.**

This might mean the model doesn't perform as well on the training data, but it's more likely to work better in real-life situations where we use it on new information.