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 optimal value of alpha for ridge regression is 50 and for lasso regression is 0.001.

	R2_Train	RMSE_Train	R2_Test	RMSE_Test
Base LR Model	0.919872	21267.0	0.900467	20681.0
Base Ridge Model	0.919872	21267.0	0.900467	20681.0
Best Ridge Model	0.919078	21471.0	0.901191	20490.0
Base Laso Model	0.000000	82080.0	-0.001085	77106.0
Best Laso Model	0.919034	21471.0	0.902861	20423.0
Best Ridge Double Alpha	0.917939	21709.0	0.900580	20461.0
Best Lasso Double Alpha	0.917898	21700.0	0.903868	20193.0
Best Lasso Assessment 3	0.895719	23691.0	0.868087	23349.0

If we choose double the value of alpha for both ridge and lasso, the model would under fit the data as the regularization or penalty on the variables would increase. Therefore, we can say that as alpha increases the model performance would increase up to a certain point and then it would just start to under fit.

According to Ridge Regression Model, our top 5 predictors variables are:

```
OverallQual    0.068270
GrLivArea      0.051523
OverallCond    0.041635
HouseAge      -0.039602
1stFlrSF       0.038176
```

According to the Lasso Regression Model, our top 5 predictors variables are:

```
GrLivArea      0.106858
OverallQual    0.081386
HouseAge      -0.058611
OverallCond    0.048120
TotalBsmtSF    0.038351
```

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?

	R2_Train	RMSE_Train	R2_Test	RMSE_Test
Base LR Model	0.919872	21267.0	0.900467	20681.0
Base Ridge Model	0.919872	21267.0	0.900467	20681.0
Best Ridge Model	0.919078	21471.0	0.901191	20490.0
Base Laso Model	0.000000	82080.0	-0.001085	77106.0
Best Laso Model	0.919034	21471.0	0.902861	20423.0
Best Ridge Double Alpha	0.917939	21709.0	0.900580	20461.0
Best Lasso Double Alpha	0.917898	21700.0	0.903868	20193.0

We found the best fit Ridge and Lasso model with the optimal value of lambda for both of the models. We can see that the Lasso gives us a smaller RMSE when compared to RMSE given by Ridge on the test data and therefore, we choose Lasso over Ridge in this case. In other words when Lasso is used our predictions on an average are off by $20,423, which is lesser than that of Ridge which is $20,490.

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?

Our top 5 variables with their coefficients are:

```
2ndFlrSF      0.095193
1stFlrSF      0.090547
BsmtFinSF1    0.082529
BsmtUnfSF     0.047053
GarageCars    0.040196
```

Question 4

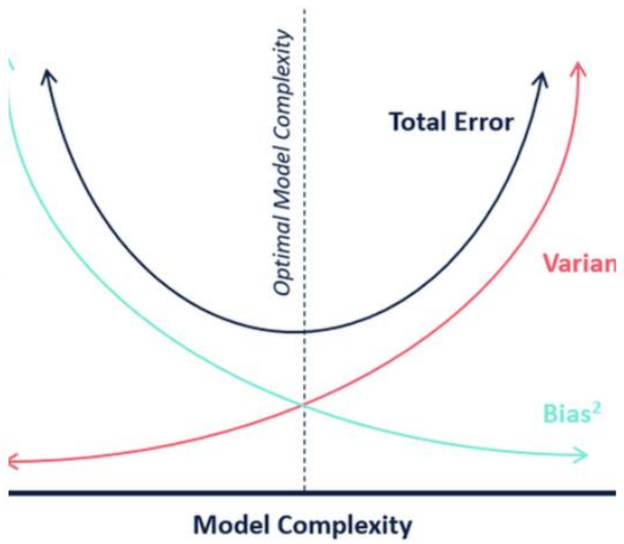
How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

Robustness of a machine learning model refers to how the model performs on unseen or new data or data it is not trained, therefore when we call a machine learning model robust, we must see very little deterioration in the performance of the model when it is used on unseen data.

Higher Robustness of a model depicts more generalisability of the model.

	R2_Train	RMSE_Train	R2_Test	RMSE_Test
Base LR Model	0.919872	21267.0	0.900467	20681.0
Base Ridge Model	0.919872	21267.0	0.900467	20681.0
Best Ridge Model	0.919078	21471.0	0.901191	20490.0
Base Laso Model	0.000000	82080.0	-0.001085	77106.0
Best Laso Model	0.919034	21471.0	0.902861	20423.0
Best Ridge Double Alpha	0.917939	21709.0	0.900580	20461.0
Best Lasso Double Alpha	0.917898	21700.0	0.903868	20193.0

We can clearly see in the above image that in the ‘Best Laso Model’ we can clearly see that the model performs (R2_Train | R2_Test) does not deviate much when tested on unseen data.



As we increase the Variance in the model to make it more robust and generalisable, we decrease the bias, in the model which in turn depicts the accuracy of the data, therefore we have to find a break even point between the bias and the variance of the data.