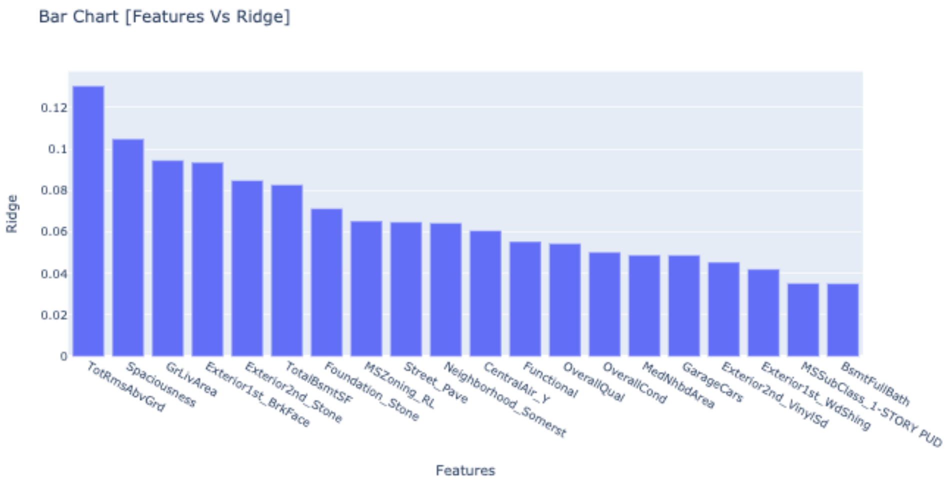
Question 1

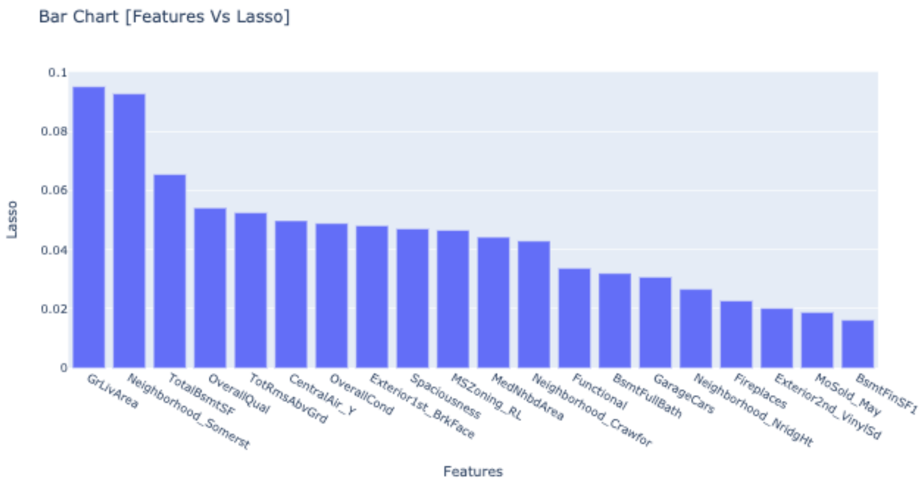
What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer: Optimal Value of alpha for lasso is 0.001 and for ridge is also 1

	Metric	Linear Regression	RFE Linear Regression	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.94	0.92	0.92	0.92
1	R2 Score (Test)	0.85	0.89	0.89	0.88
2	Adjusted R2 Score (Train)	0.92	0.91	0.91	0.91
3	Adjusted R2 Score (Test)	0.77	0.87	0.87	0.82
4	RSS (Train)	10.34	12.92	12.99	12.38
5	RSS (Test)	0.03	0.02	0.02	0.02
6	RMSE (Train)	0.10	0.11	0.11	0.11
7	RMSE (Test)	0.16	0.14	0.14	0.14

Before Changing:



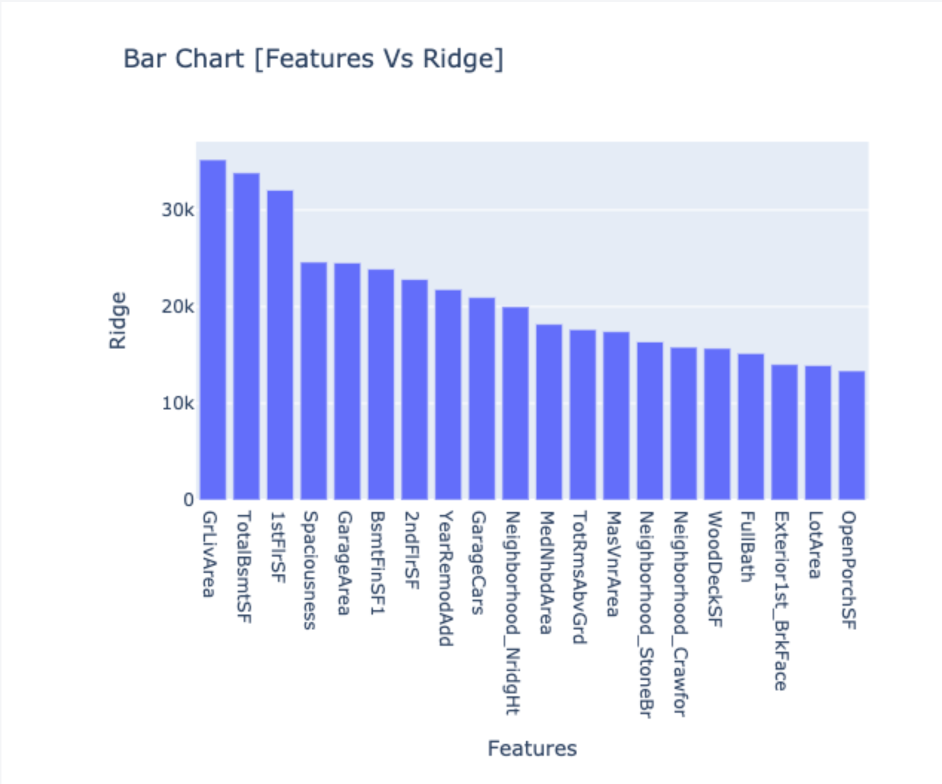


Impact of Doubling Alpha Values:

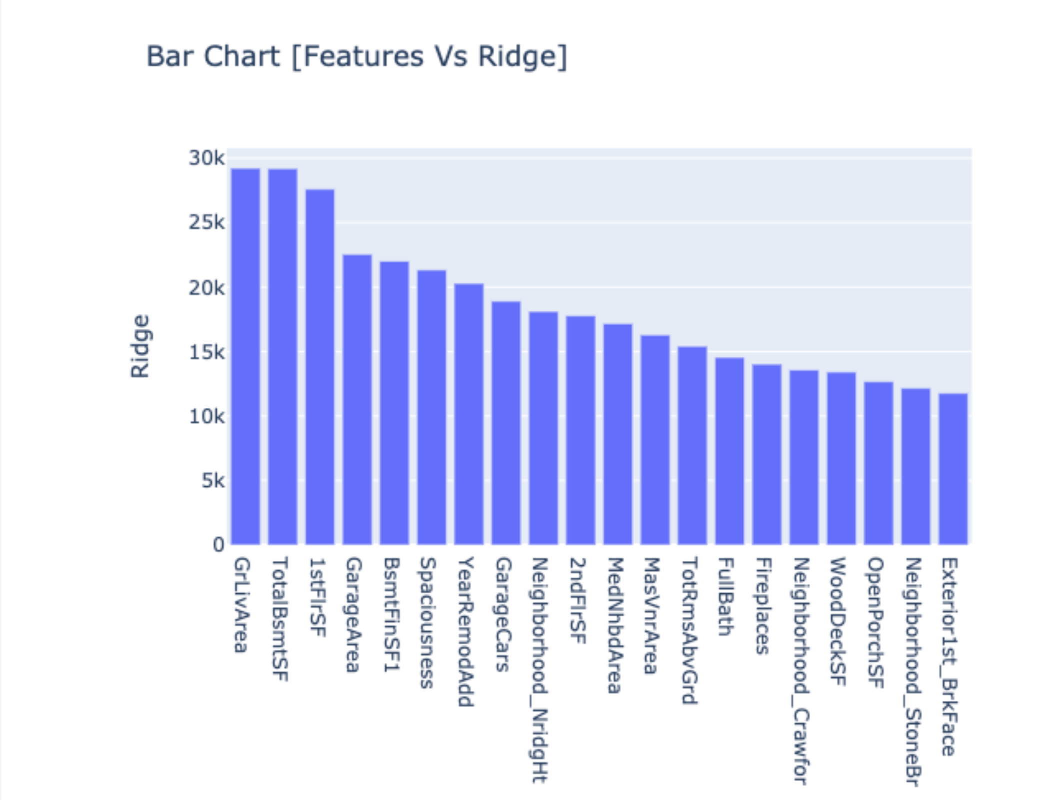
1. R-square/RMSE has changed:
    a. Ridge: Test R2 Score has remained same.
    b. Lasso – Train R2 Score has decreased from 92% to 91%

	Metric	Linear Regression	RFE Linear Regression	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.94	0.92	0.92	0.91
1	R2 Score (Test)	0.85	0.89	0.89	0.88
2	Adjusted R2 Score (Train)	0.92	0.91	0.91	0.90
3	Adjusted R2 Score (Test)	0.77	0.87	0.87	0.82
4	RSS (Train)	10.34	12.92	13.08	14.10
5	RSS (Test)	0.03	0.02	0.02	0.02
6	RMSE (Train)	0.10	0.11	0.11	0.12
7	RMSE (Test)	0.16	0.14	0.14	0.14

2. Features:
    a. Ridge:
        i. Before Change:



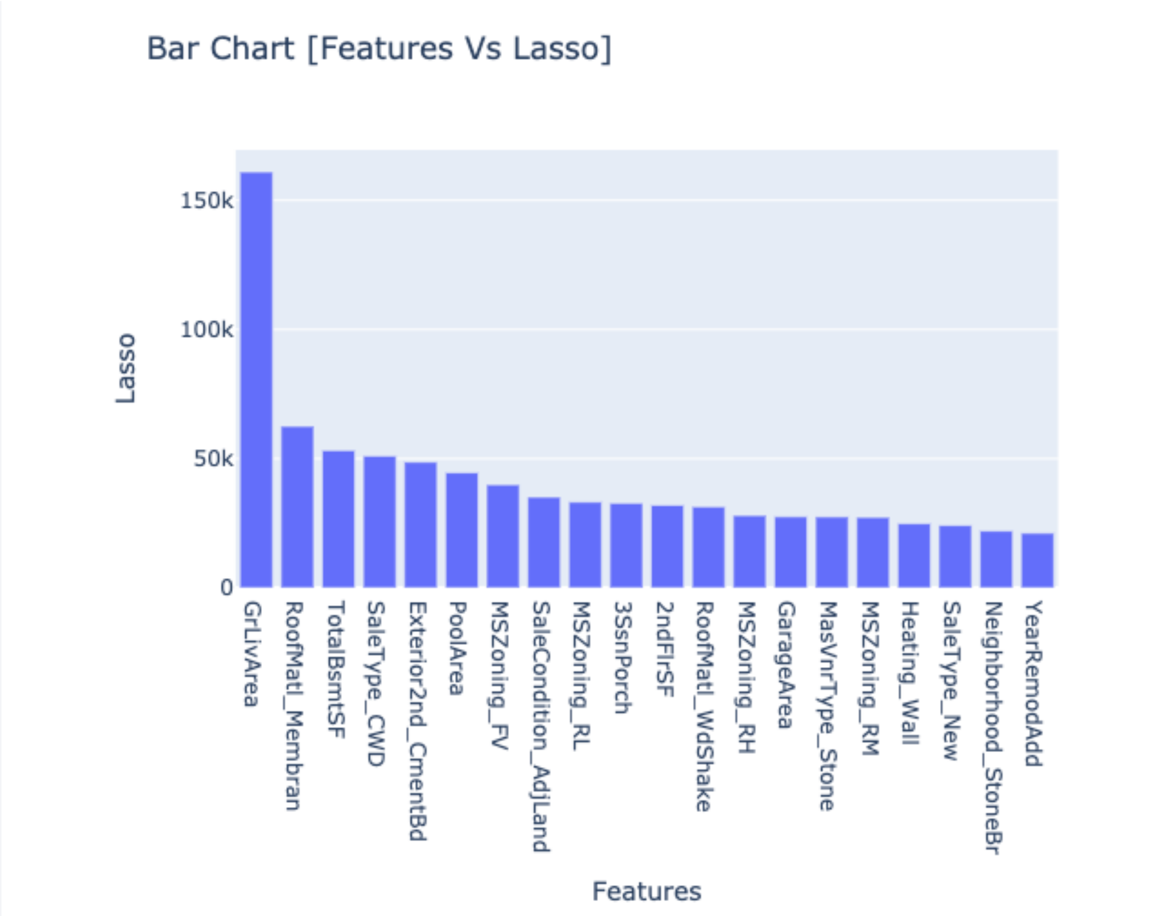
Ii After Change:



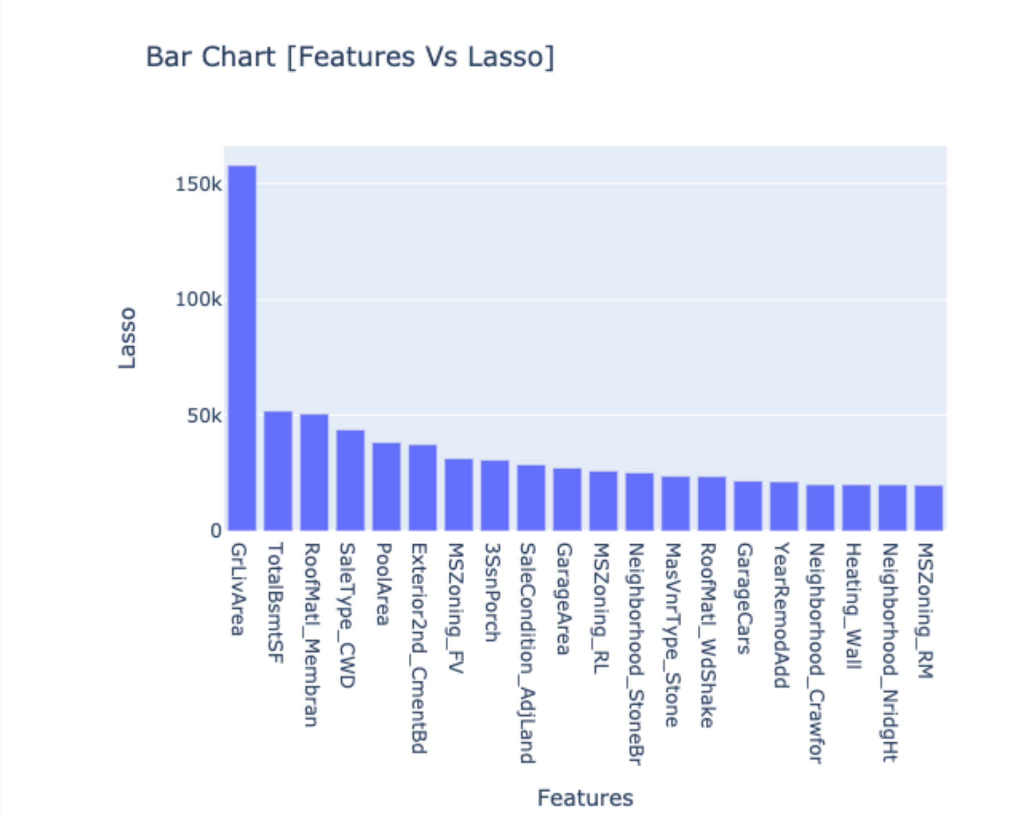
Analysis: Top features have shuffled and coefficient values also got decreased as well.

b. Lasso:

i. Before Change:



ii. After:



Analysis: Top features have shuffled and no significant change in coefficient values

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?

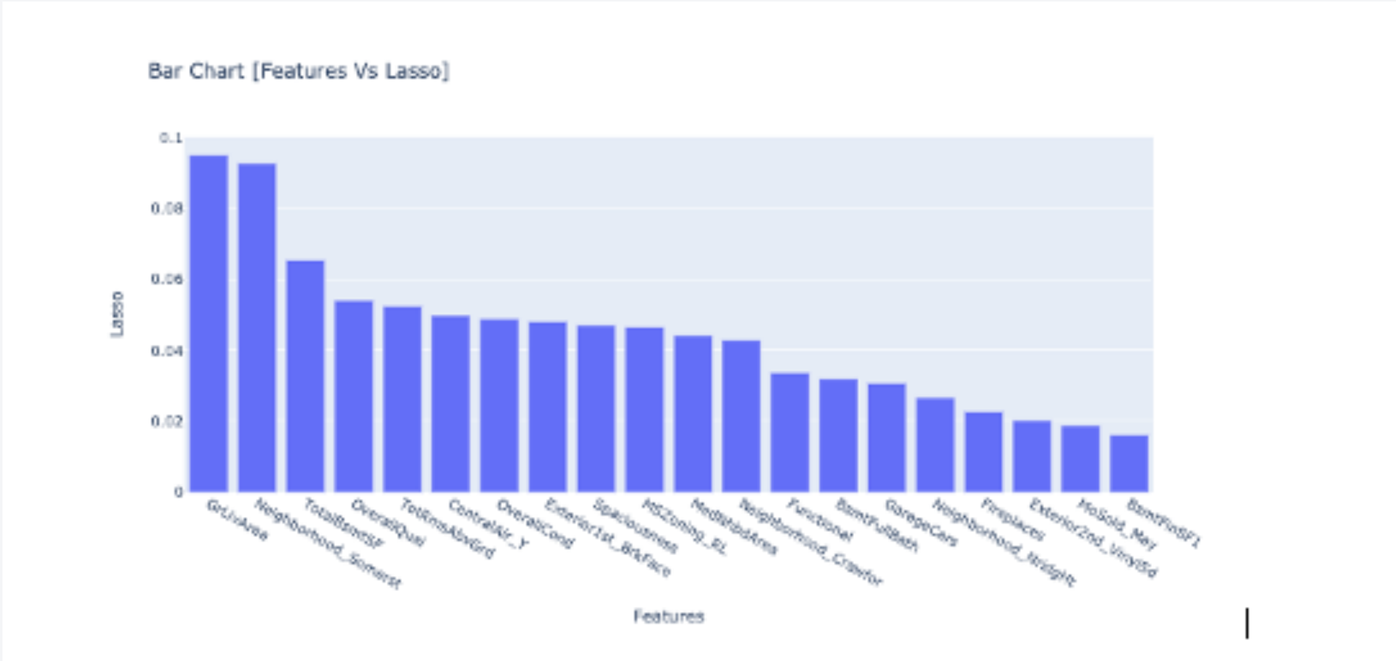
Answer: Ridge will be preferred choice as the RMSE/R2-Square are similar as of Lasso but it will be simpler as the number of features are less.

	Metric	Linear Regression	RFE Linear Regression	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.94	0.92	0.92	0.92
1	R2 Score (Test)	0.85	0.89	0.89	0.88
2	Adjusted R2 Score (Train)	0.92	0.91	0.91	0.91
3	Adjusted R2 Score (Test)	0.77	0.87	0.87	0.82
4	RSS (Train)	10.34	12.92	12.99	12.38
5	RSS (Test)	0.03	0.02	0.02	0.02
6	RMSE (Train)	0.10	0.11	0.11	0.11
7	RMSE (Test)	0.16	0.14	0.14	0.14

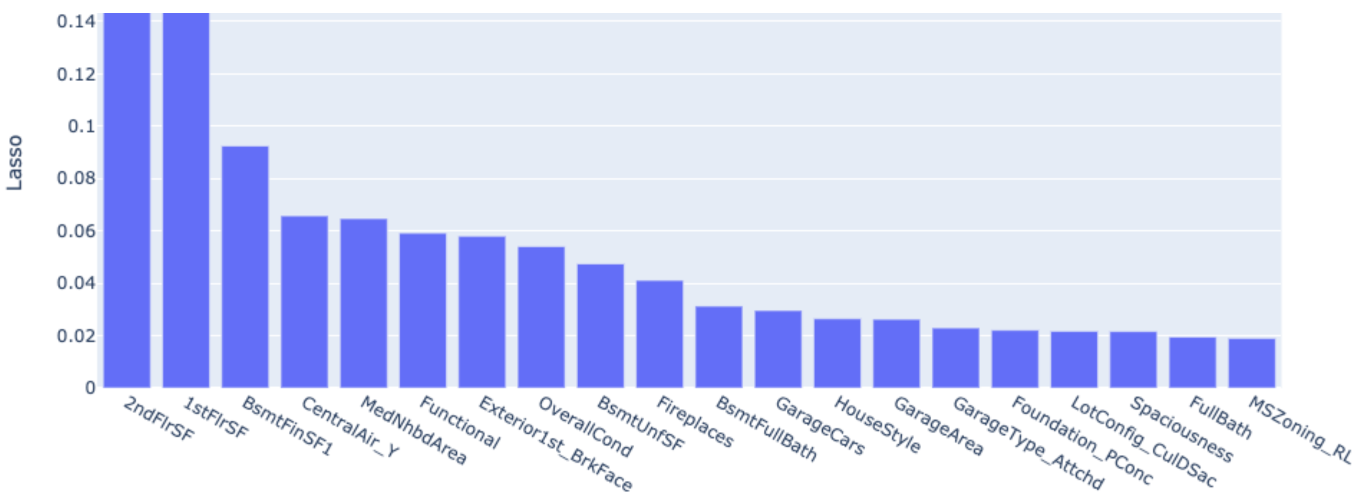
Question 3

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?

Top five features in Lasso before exclusion:



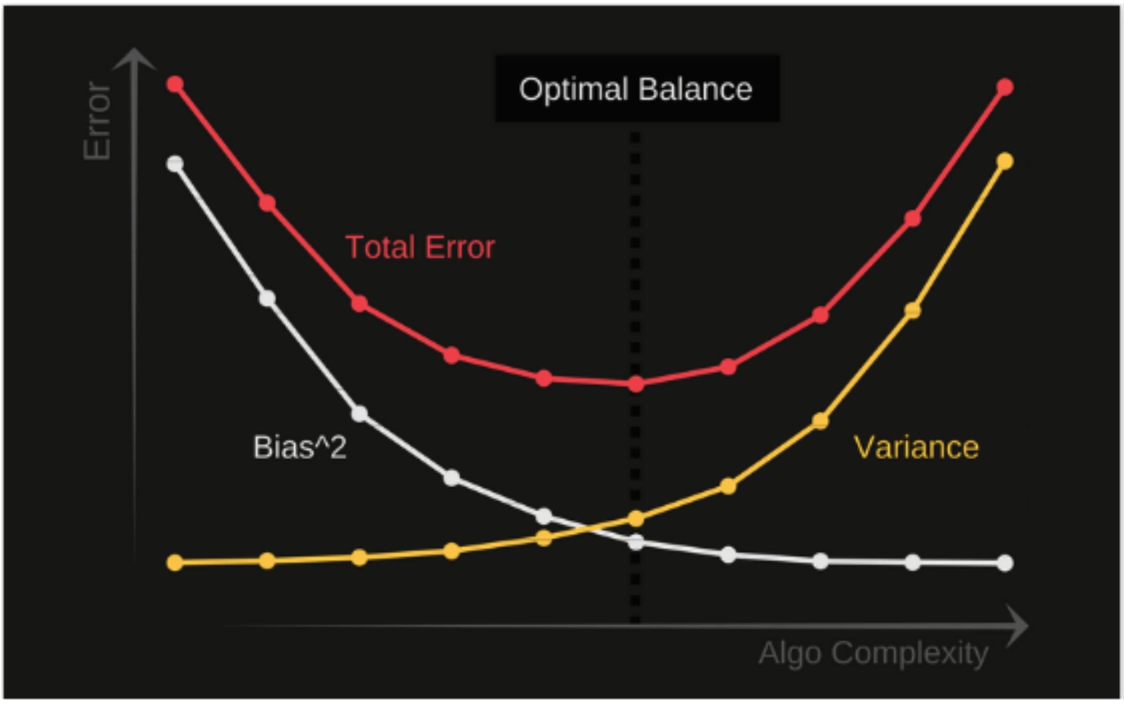
Post exclusion of top 5 features. New ones are:



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?

Answer:
To make a model robust and generalisable, we need to find a balance between bias and variance.
A simple model would usually have high bias and low variance, whereas a complex model would have low bias and high variance. In either case, the total error would be high.
To build a good model, we need to find a good balance between bias and variance such that it minimizes the total error.



An optimal balance of bias and variance would never overfit or underfit the model.
Accuracy of complex model is good on train data but it does not perform well on test data as it learns and mimic the train data, where as too simple model will not perform well on train data itself and will not be able to understand the data.
Hence, we need to do a trade-off b/w bias and variance, i.e., low bias and low variance, such that the model identifies all the patterns that it should and is also able to perform well with unseen data.