**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 10 and for ridge is also 10

Table

Description automatically generated

 Impact of Doubling Alpha Values:

1. R-square/RMSE has changed:
   1. Ridge: Test R2 Score has decreased from 89 to 87%
   2. Lasso – Test R2 Score hs increased from 89% to 90%

Table

Description automatically generated

1. Features:
   1. Ridge:
      1. Before Change:

Chart, bar chart, histogram

Description automatically generated

Ii After Change:

Chart, bar chart

Description automatically generated

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

* 1. Lasso:
     1. Before Change:

Chart, bar chart, histogram

Description automatically generated

* + 1. After:

Chart, bar chart, histogram

Description automatically generated

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(209 against 214 in Lasso)

Table

Description automatically generated

**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:

Chart, bar chart, histogram

Description automatically generated

Post exclusion of top 5 features. New ones are:

Chart

Description automatically generated

**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?