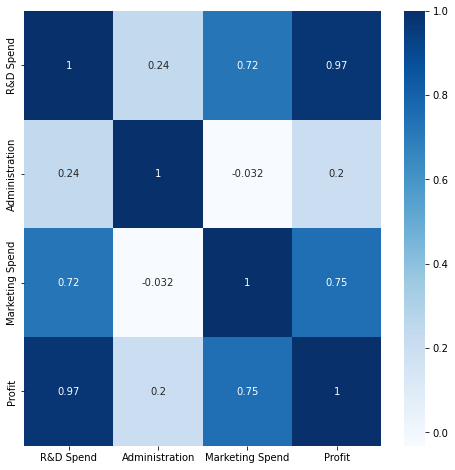
**REPORT**

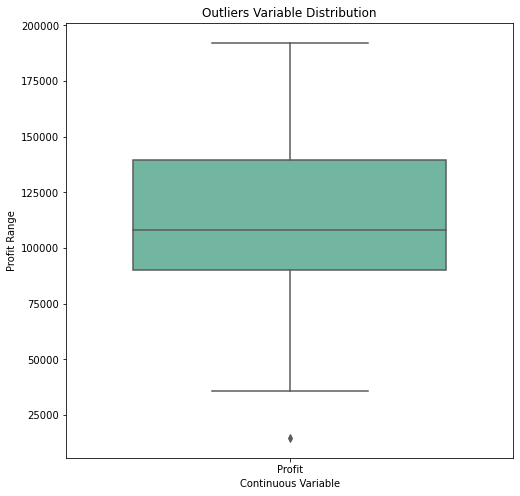
During the development a validation set was used to evaluate the model. The final architecture and hyperparameters were chosen because they performed the best among the tried combinations.

**Correlation matrix:**



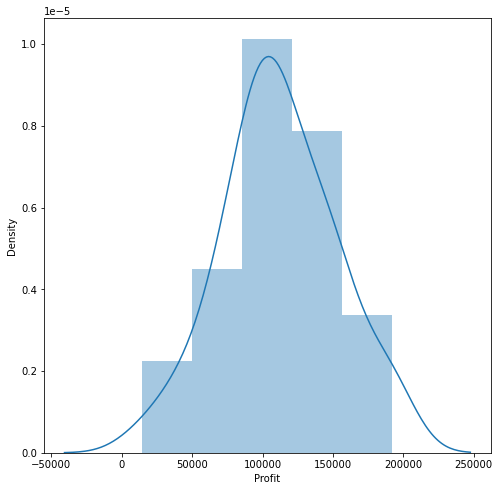
We can see the direct correlation with profit from how it is shown in the heatmap of the correlation plot.

**Outlier detection in the target variable:**

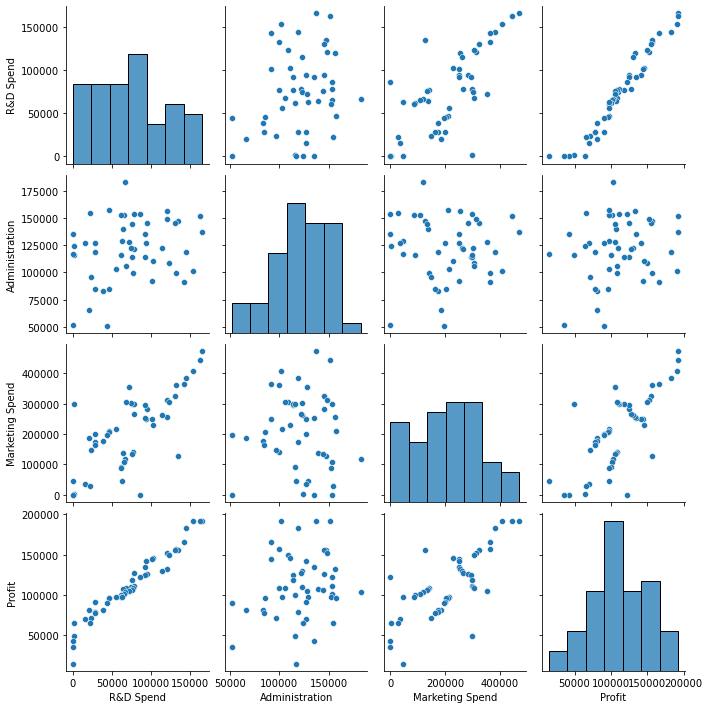


In the boxplot we can see the outliers in the profit it won’t create much negative impact as the data in not very large.

**Histogram of the Profit**



**Pair Plot:**

****

1. As we can see in the pair pot, Research and development are directly proportional to the investment that we can do.

2. The marketing spend seems to be directly proportional with the profit.

3. There is no relationship between the second column and profit that is our target column.

The results of various regression algorithms are as follows:

**Linear regression:**

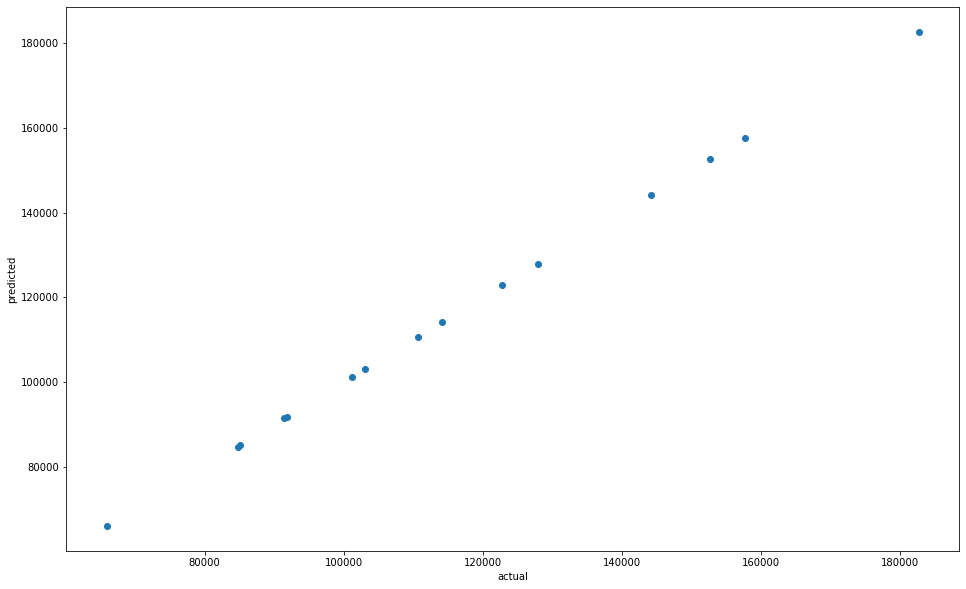
1. The model score/performance on testing the data is1.0
2. The Mean squared error and the standard deviation:

* Mean Squared Error: -5.006317648667604e-21
* Standard Deviation: 6.964140604836841e-21

3.The mean absolute error and standard deviation

* Mean Absolute Error: -3182.9153333333347
* Standard Deviation: 2693.529565014873

1. Model Score/Performance on Training data is 1.0



**Logistic regression:**

1. The model training accuracy: 0.688
2. The Mean squared error and the standard deviation

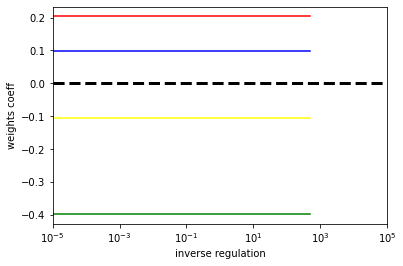
* Mean Squared Error: -0.559
* Standard Deviation: 0.08395832299420945

1. The mean absolute error and standard deviation

* Mean Absolute Error: -0.36400000000000005
* Standard Deviation: 0.04963869458396343

1. Root squared value of the model

* R squared value: 0.15700226402845158



**Ridge:**

The predicted score/performance of the tested model is 136897.800

