**Machine Learning : Linear Regression With Multiple Variables**

PROJECT REPORT

AIM OF THE PROJECT:

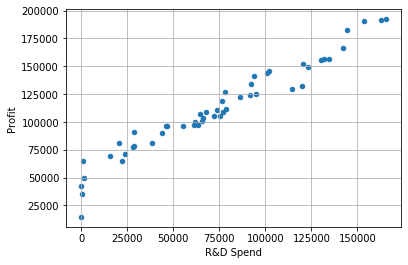
Create a machine learning model using linear regression that can predict PROFIT for startup based on various factors like R&D Spend, Adminstration , Marketing Spend and State.

ROLE OF EACH COLUMN:

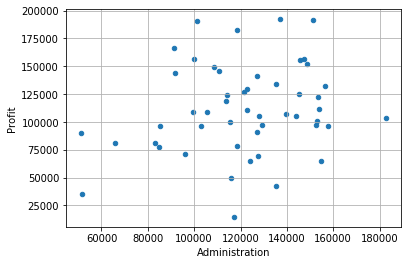
It can be observed the numerical variables are: R&D Spend, Adminstration , Marketing Spend, while the categorical variable is state.

Since Profit is the target variable , we will be drawing graph showing the relation between numerical variables and the target variable

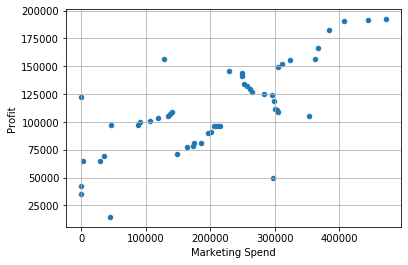
1)RELATION BETWEEN R&D Spend and profit



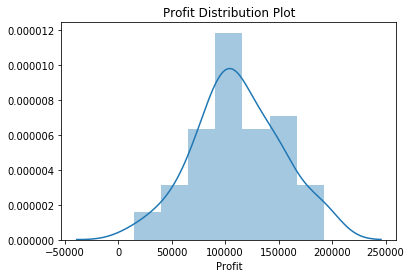
2)RELATION BETWEEN ADMINSTRATION AND PROFIT



3)RELATION BETWEEN MARKETING SPEND AND PROFIT



4)Distribution of Target Variable

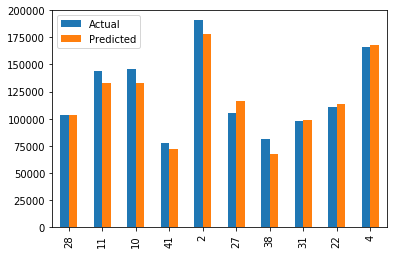


Since the target variable is linearly distributed .so linear regression will hold true.

Predicting output

Using linear regression output (here profit) was predicted for the test data and graph was plotted between actual and predicted values

Actual v/s Predicted values



CONCLUSION

A machine learning model was created using linear regression which predicted the output(here profit) with an accuracy of 93.4%

Link to the dataset: https://www.kaggle.com/karthickveerakumar/startup-logistic-regression