**Interpretations for the Analysis of Startup 50 data**

**Simple linear regression:**

* In this analysis the relationship between the independent variable **‘R&D Spend’** and the dependent variable **‘Profit’**.
* 80% of the data is used for training and 20% of data is used for testing using **train – test split.**
* When there is no amount spent for R&D, then the profit would be **Rs.49337**. This is calculated using the intercept value.
* The r2\_ score is 0.93 which means that **93%** of the variation in Profit is attributed to R&D spend and only the remaining 7% is attributed to other variables like administration expense and marketing expenditures.

**Multiple linear regression:**

* In this analysis the relationship between the independent variables **‘R&D Spend, administration expense, marketing expenditures, State of purchase’** and the dependent variable **‘Profit’**.
* The state variable is in object type. Hence it is converted into int type using **‘Label encoder’**.
* 80% of the data is used for training and 20% of data is used for testing using **train – test split.**
* When there is no amount spent on any of the independent variables, then the profit would be **Rs.54081**. This is calculated using the intercept value.
* The r2\_ score is 0.9001 which means that **90%** of the variation in Profit is attributed to R&D spend, administration expense, marketing expenditures, State of purchase and only the remaining 10% is attributed to other factors not included in the model.
* The mean squared error is **80929465** which is the prediction error in the model.