**Dataset**:-StartUP

**Problem statement**:- A prediction model for profit of 50\_startups data.

Dependent variable:- Profit

Independent Variable:- R.DSpend, Administration ,Marketting.Spend ,State

Now,lets make EDA of the given data:-

**Business Moment 1:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **RDSPEND** | **ADMINISTRATION** | **MARKETINGSPEED** | **STATE** | **PROFIT** |
| MEAN | **73722** | **121345** | **211025** | **-** | **112013** |
| MEDIAN | **73051** | **122700** | **212716** | **-** | **107978** |

**Business Moment 2:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **VARIANCE** | **SD** | **RANGE** |
| RD SPEND | 2107017150 | 45902.26 | (0-165349.20) |
| ADMINISTRATION | 784997271 | 28017.80 | (51283.14-182645.56) |
| MARKETTING SPEED | 14954920097 | 122290.31 | (0-471784.10) |
| STATE | NA | NA | NA |
| PROFIT | 1624588173 | 40306.18 | (14681.40-192261.83) |

**Business moment 3 :**

SKEWNESS VALUES:-

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | RD SPEND | 0.1590405 | | ADMINISTRATION | -0.4742301 | | MARKETING SPEED | -0.04506632 | | STATE | NA | | PROFIT | 0.02258638 | |
| OBSERVATION:-   * R.D.Spend, Profit are right skewed * Adminstration , Marketing.spend are left skewed |

**Business Moment 4:-(kurtiosis)**

|  |  |
| --- | --- |
| RD SPEND | 0.1590405 |
| ADMINISTRATION | -0.4742301 |
| MARKETING SPEED | -0.04506632 |
| STATE | NA |
| PROFIT | 2.824704 |

THERE ARE NO HEAVY PEAKS!!

**Linear Model:-**

R-squared value is 0.712

Probability-value is less than 0.05

**Logarithmic transformation:-**

Multiple R-squared: 0.669,

Adjusted R-squared: 0.6621

p-value: 4.16e-13

**Exponential transformation:-**

Multiple R-squared: 0.7617

Adjusted R-squared: 0.7462

p-value: 2.275e-14