ASSIGNMENT – 1

**About the 50 startups dataset**

This particular dataset holds data from 50 startups in New York, California, and Florida. The

features in this dataset are R&D spending, Administration Spending, Marketing Spending, and

location features, while the target variable is: Profit.

1. R&D spending: The amount which startups are spending on Research and development.

2. Administration spending: The amount which startups are spending on the Admin panel.

3. Marketing spending: The amount which startups are spending on marketing strategies.

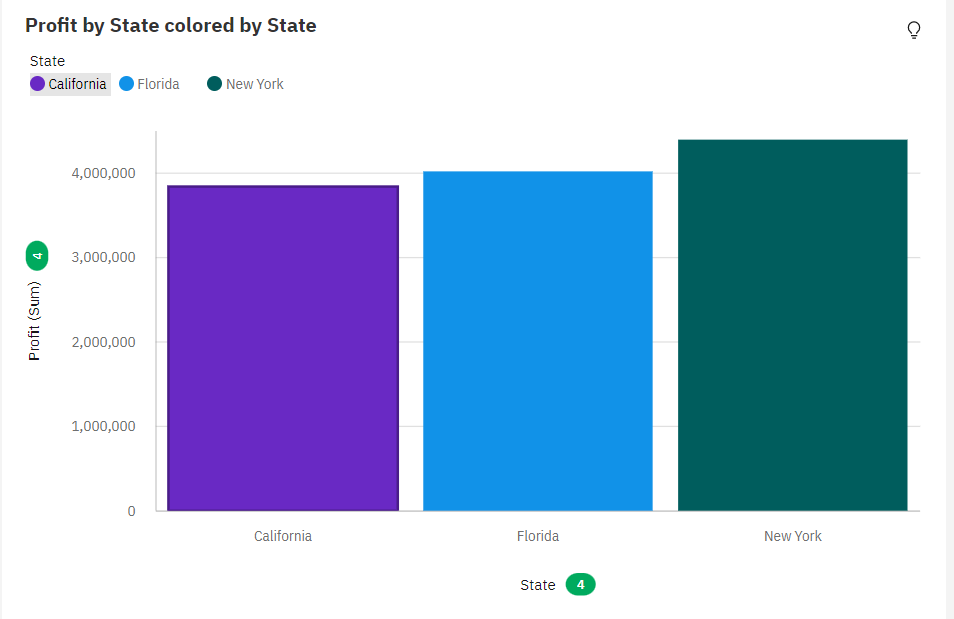
4. State: To which state that particular startup belongs.

5. Profit: How much profit that particular startup is making.

**Challenge:**

Upload the dataset to Cognos Analytics, explore and visualize the dataset

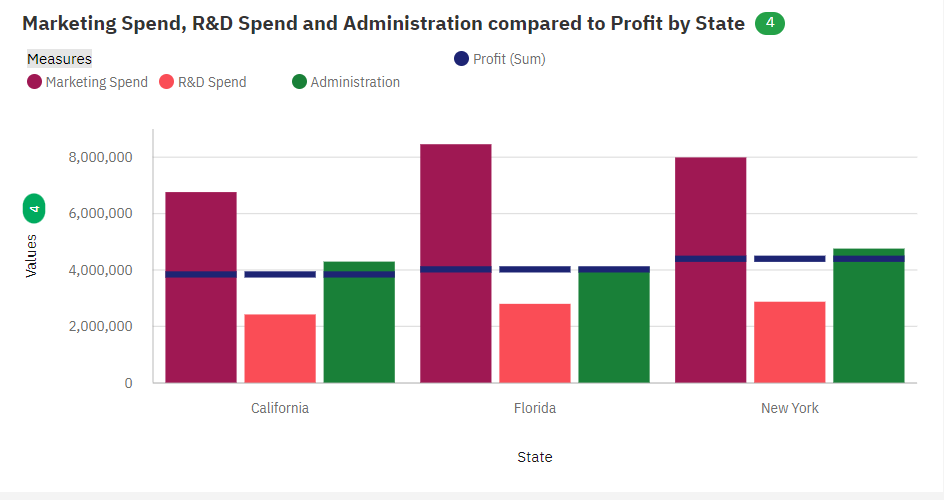
**Graph 1**



Analysed Data

* Across all **states** and **states**, the sum of **Profit** is over 12 million.
* For **Profit**, the most significant value of **State** is New York, whose respective **Profit** values add up to almost 4.4 million, or 35.9 % of the total.
* The summed values of **Profit** range from over 3.8 million to almost 4.4 million.

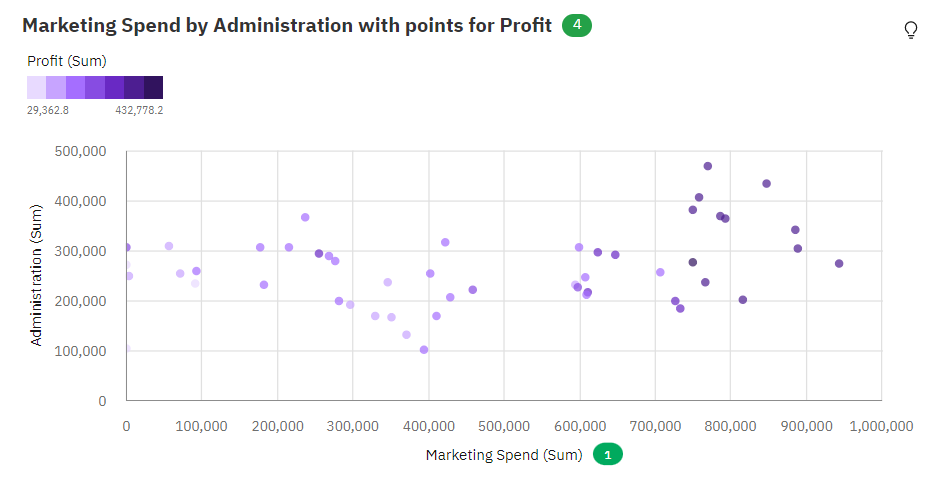
**Graph 2**

****

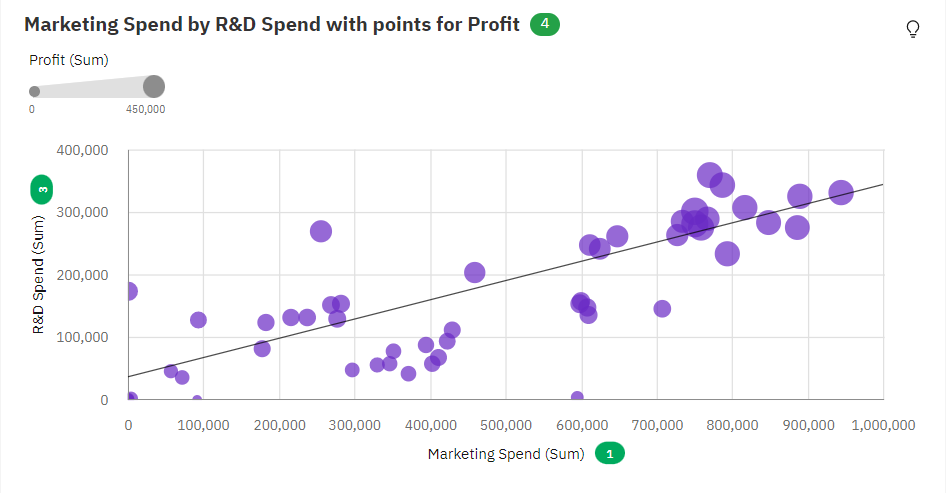
Analysed Data

* **Marketing Spend** ranges from almost 6.8 million, in California, to almost 8.5 million, in Florida.
* **R&D Spend** ranges from over 2.4 million, in California, to almost 2.9 million, in New York.
* **Administration** ranges from almost 4.1 million, in Florida, to nearly 4.8 million, in New York.

**Graph 3**

****

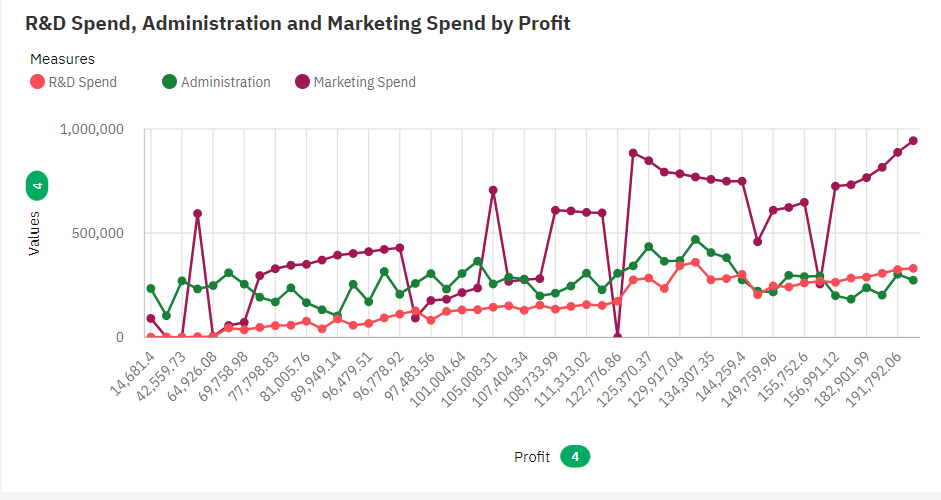
**Graph 4**

****

Analysed Data

* R&D Spend and Marketing Spend have a medium positive linear association, being R&D Spend = 36050 + 0.3064 \* Marketing Spend

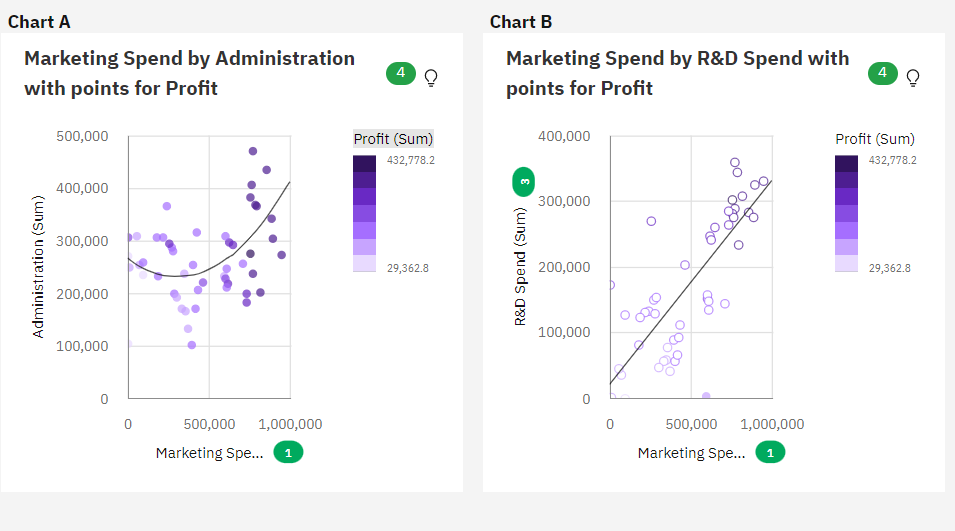
**Graph 5**

****

Analysed Data

* Administration ranges from nearly 103 thousand, when Profit is 89949.14, to nearly 470 thousand, when Profit is 132602.65.
* Marketing Spend ranges from 0, when Profit is 35673.41, to almost 944 thousand, when Profit is 192261.83.
* Overall profits, the average of Marketing Spend is nearly 215 thousand.
* Over all profits, the average of R&D Spend is almost 75 thousand.
* Over all profits, the average of Administration is nearly 122 thousand.

**Graph 6**

****

Analysed Data

* There is a weak quadratic relationship between Administration and Marketing Spend, being Administration = 265437 - 0.2329 \* Marketing Spend + 3.779E-7 \* Marketing Spend^2.
* R&D Spend and Marketing Spend have a medium positive linear association, being R&D Spend = 19744 + 0.3111 \* Marketing Spend