

# IBM NALAIYA THIRAN

## ASSIGNMENT 1

Name: INDHU PRAKASH K V

Roll No: 737319CSR064

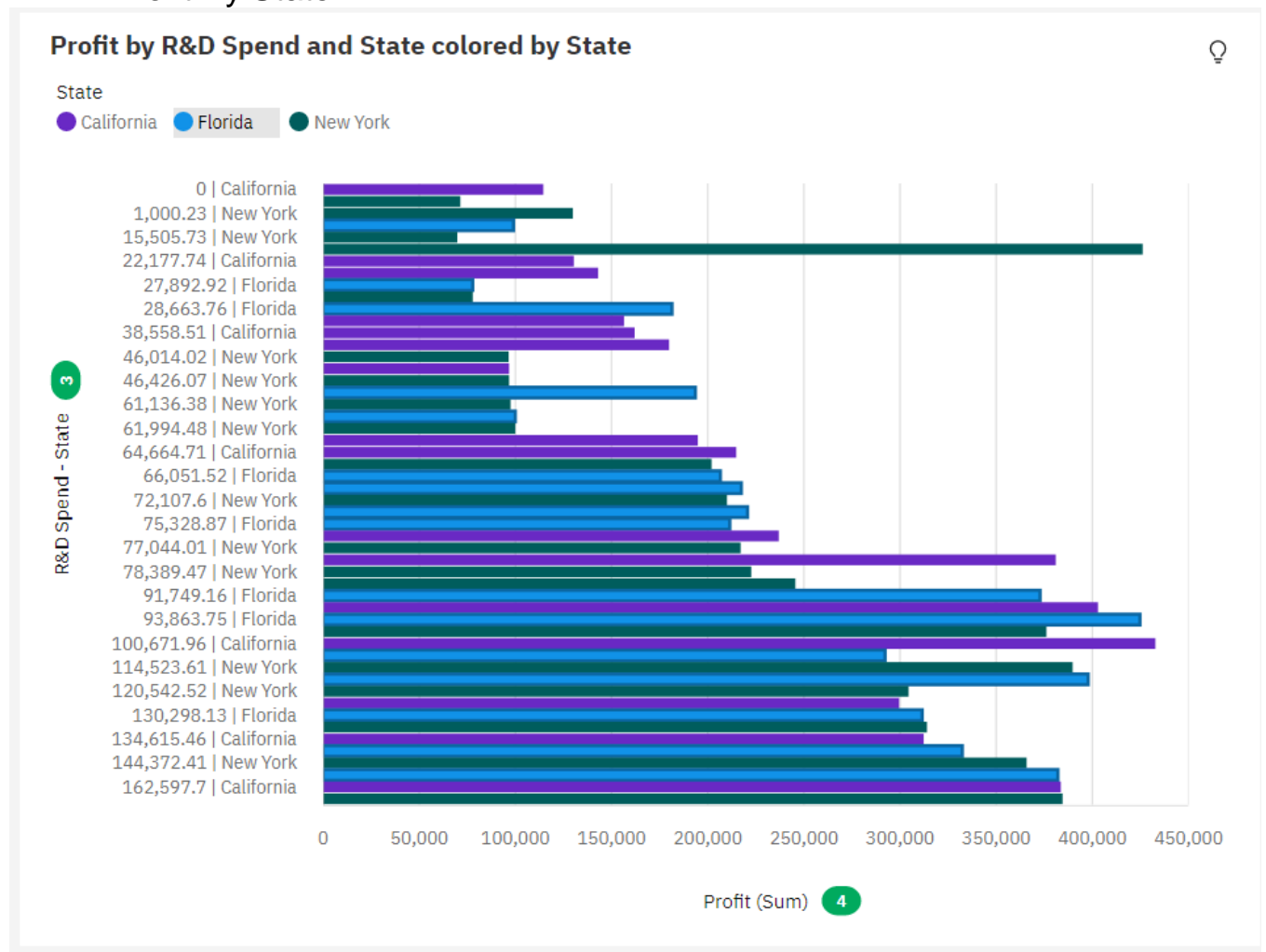
### Challenge:

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

Dataset: [50-Startup-Dataset](#)

## Exploration and Visualization:

### 1. Profit By State

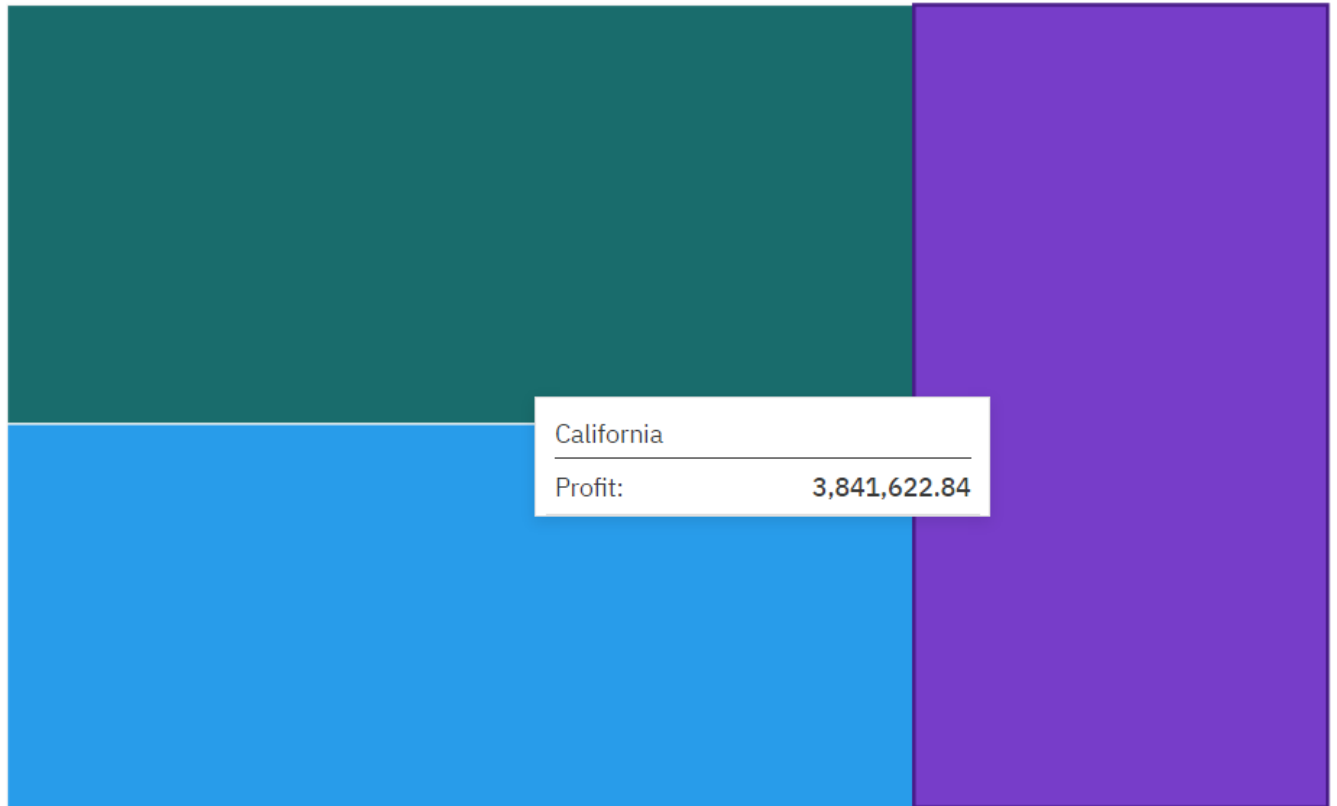
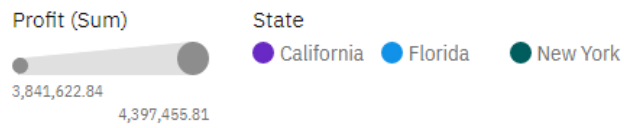


### Inference:

- For Profit, the most significant values of R&D Spend - State are 55493.95|Florida and 27892.92|Florida, whose respective Profit values add up to 6, or 11.5 % of the total.
- Over all R & D spend - states and states, the sum of Profit is 52.
- For Profit, the most significant values of State are State\_CAT77798 and State\_CAT99937, whose respective Profit values add up to 6, or 11.5 % of the total.
- The summed values of Profit range from 0 to 2.

## 2. Profit by State Colored by State

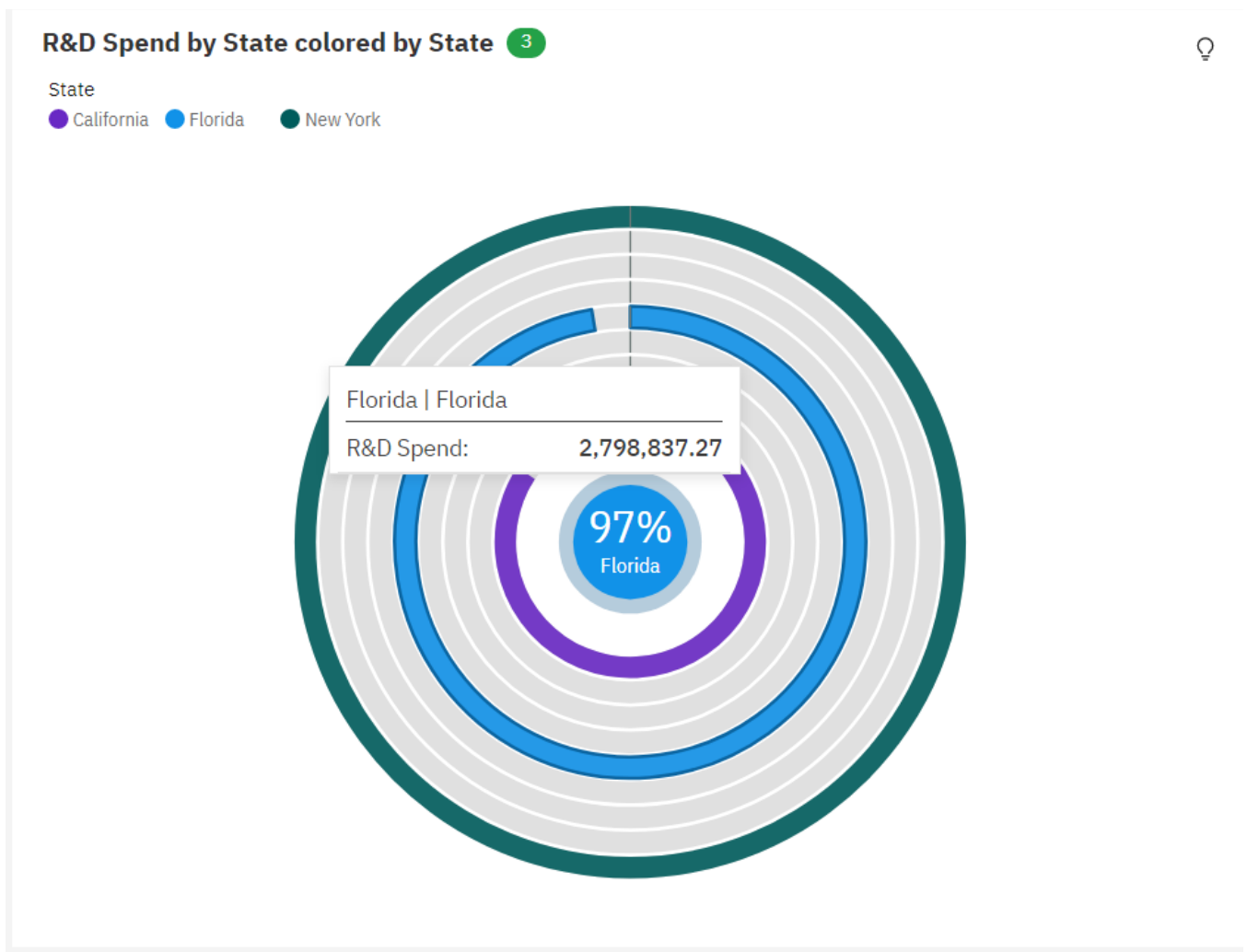
### Profit for State hierarchy 4



### Inference:

- Across all states, the sum of Profit is over 12 million.
- Profit ranges from over 3.8 million, in California, to nearly 4.4 million, in New York.

### 3. Profit and Marketing Spend by State



#### Inference:

- For R&D Spend, the most significant values of State are New York and Florida, whose respective R&D Spend values add up to nearly 5.7 million, or 70.1 % of the total.
- Over all states and states, the sum of R&D Spend is almost 8.1 million.
- The summed values of R&D Spend range from over 2.4 million to nearly 2.9 million.

#### 4. R&D Spend, Administration and Marketing Spend by State

##### State, R&D Spend, Profit, State, Profit

State - Measures

California | State

California | R&D Spend

California | Profit

Florida | State

Florida | R&D Spend

Florida | Profit

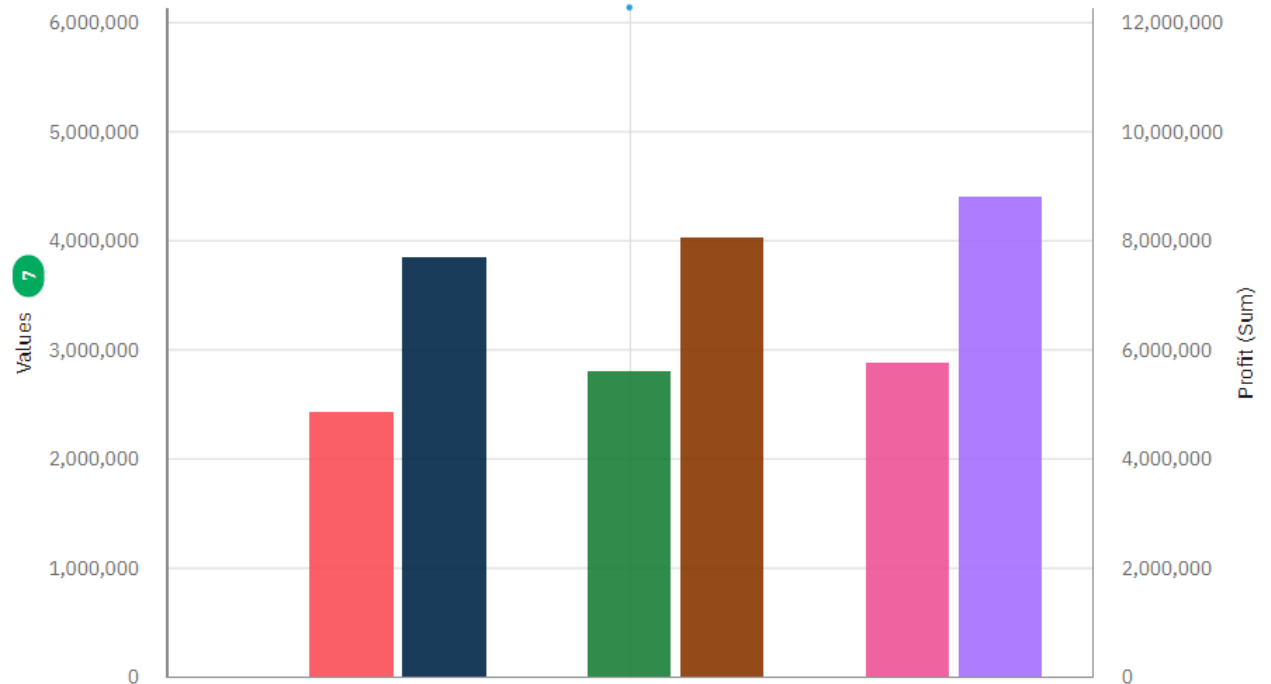
New York | State

New York | R&D Spend

New York | Profit

Line

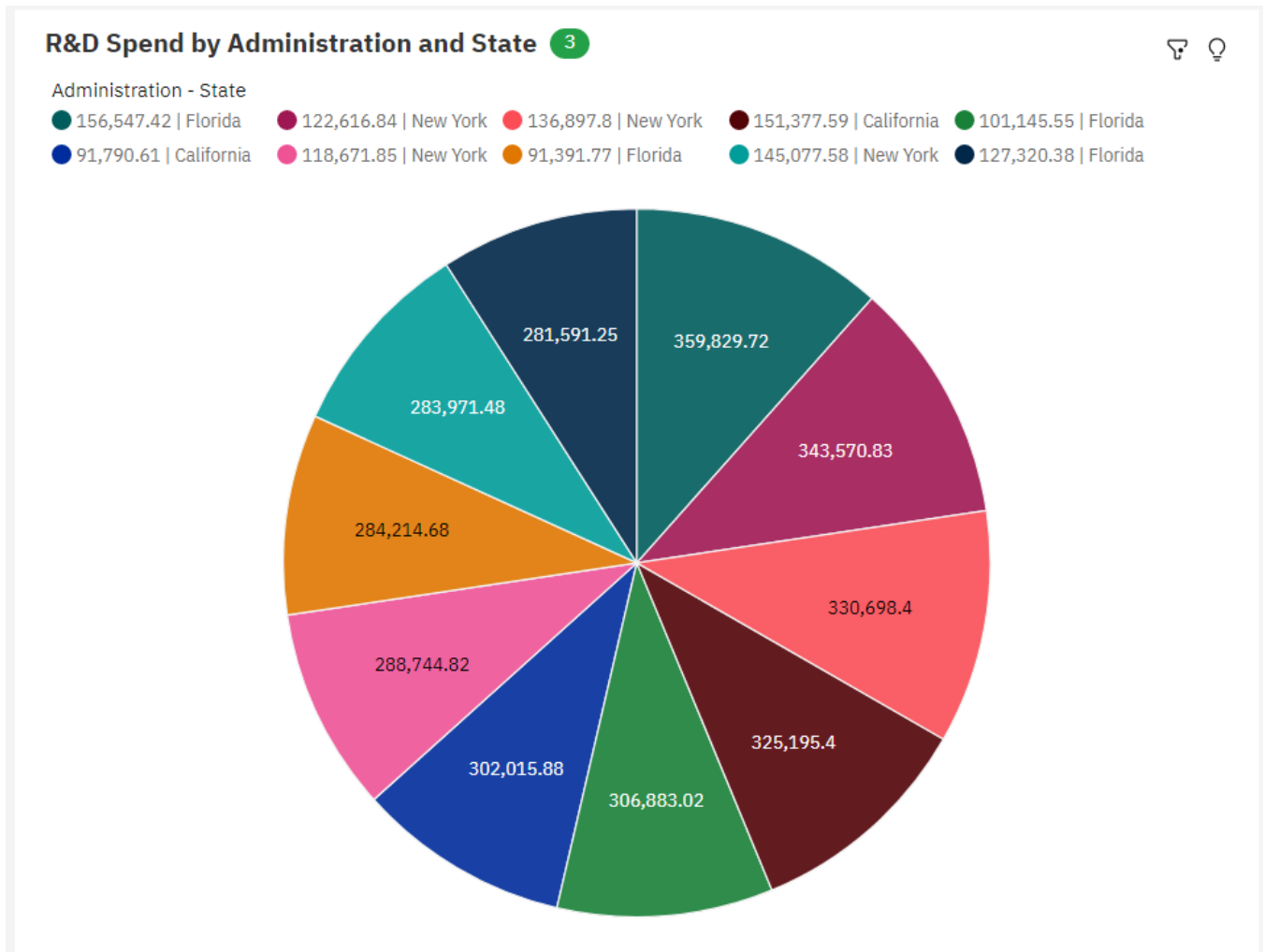
Profit (Sum)



##### Inference:

- The total number of results for State is 108.
- The total number of results for Profit is 108.
- The average of Profit is almost 114 thousand.
- The total number of results for R&D Spend is 108.
- The average of R&D Spend is almost 75 thousand.

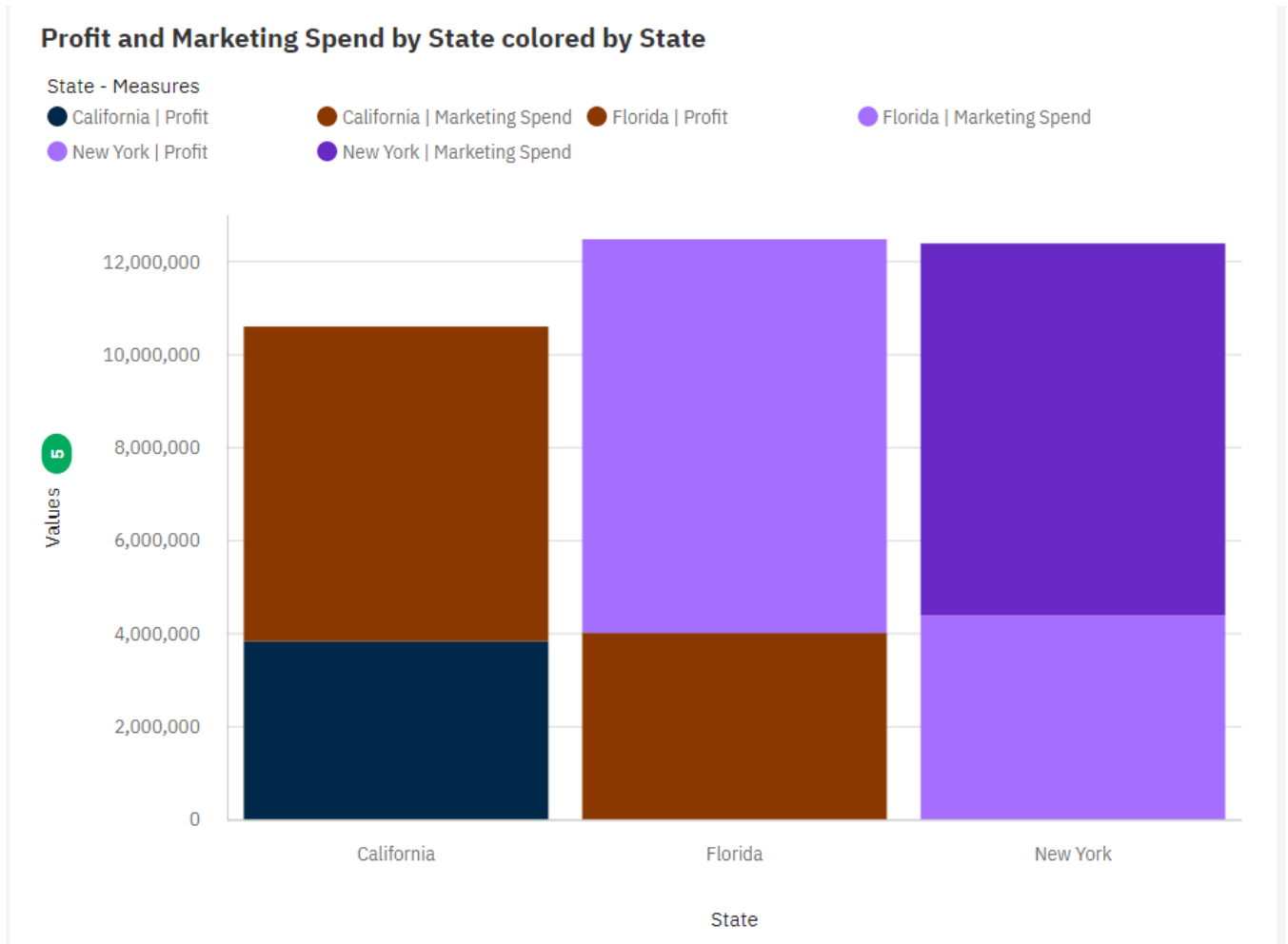
## 5. R&D Spend and Profit by State



### Inference:

- Over all administration - states, the sum of R&D Spend is 8.
- For R&D Spend, the most significant values of Administration - State are 91790.61|California and 151377.59|California, whose respective R&D Spend values add up to 4, or 50 % of the total.
- R&D Spend ranges from 0, when Administration - State is 156547.42|Florida, to 2, when Administration - State is 151377.59|California.

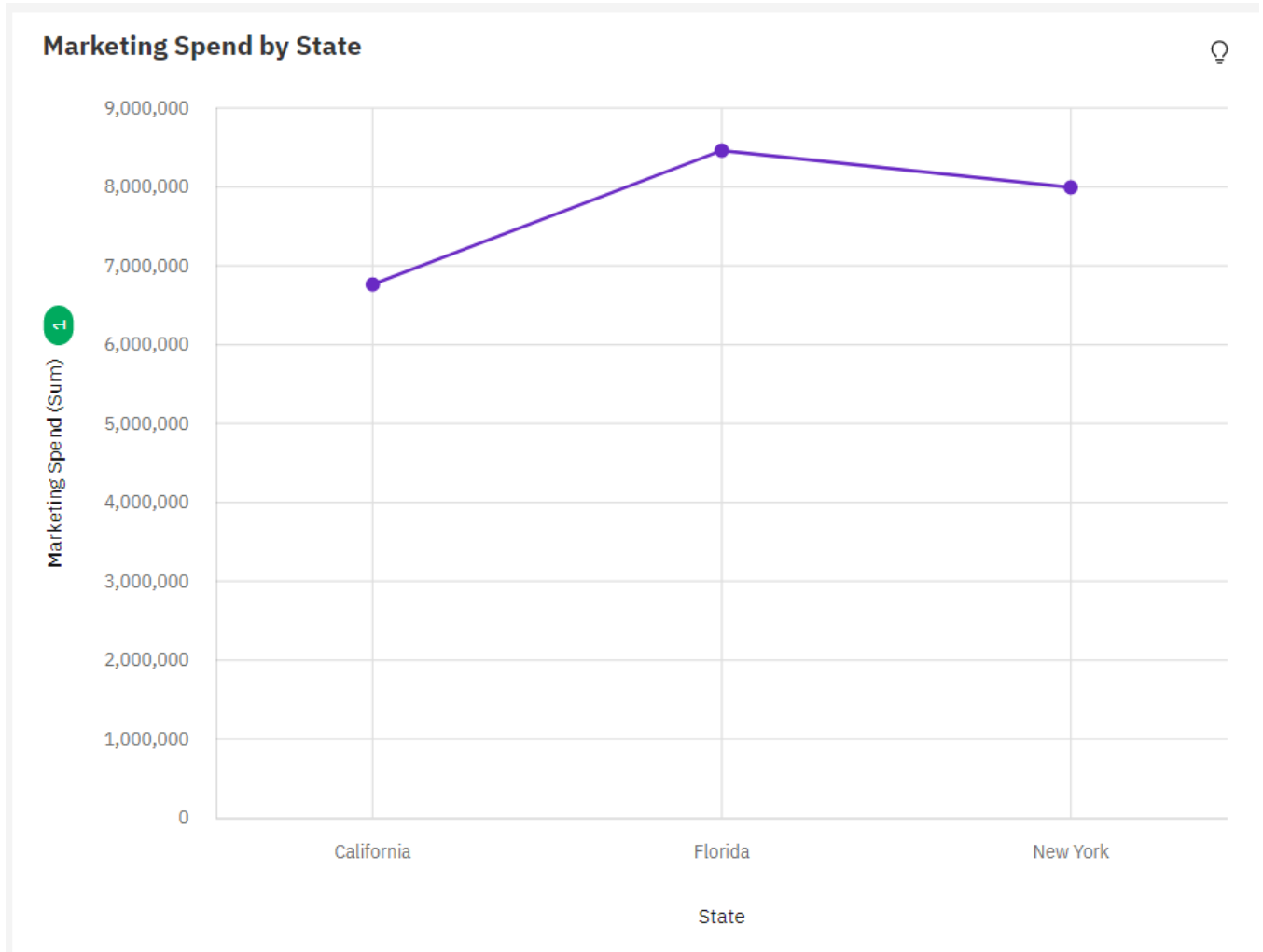
## 6. Profit Prediction Using Decision Tree:



### Inference:

- The total number of results for Profit, across all states, is 108.
- Over all states, the average of Profit is nearly 114 thousand.
- The most common values of State are New York (36.1 %) and California (33.3 %), together occurring 75 times, which is 69.4 % of the total.
- The total number of results for Marketing Spend, across all states, is 108.
- Over all states, the average of Marketing Spend is nearly 215 thousand.
- Profit ranges from over 3.8 million, in California, to nearly 4.4 million, in New York.
- Marketing Spend ranges from almost 6.8 million, in California, to almost 8.5 million, in Florida.

## 7. Profit by Administration (Group)

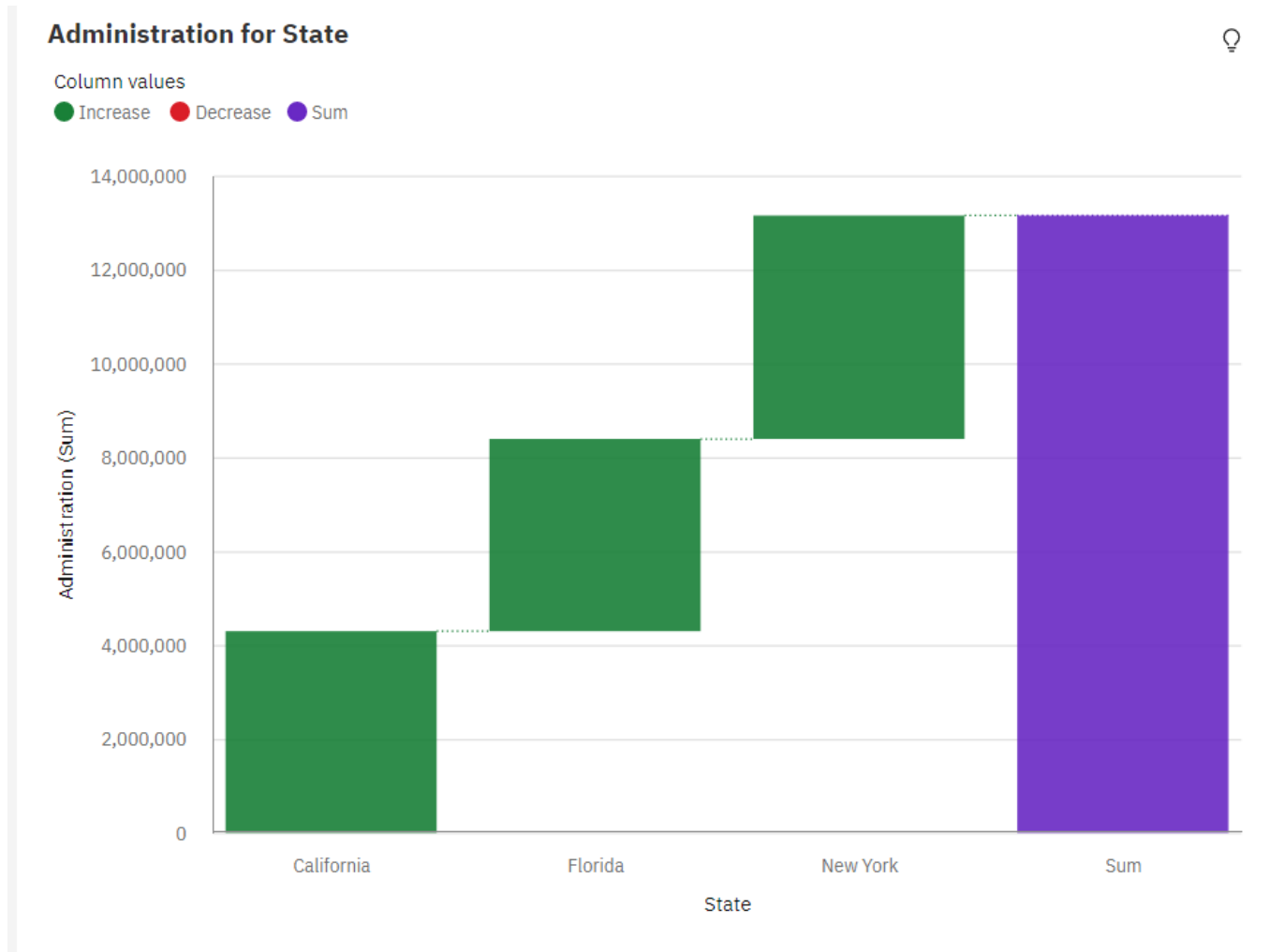


### Inference:

- Across all states, the sum of Marketing Spend is over 23 million.
- For Marketing Spend, the most significant values of State are Florida and New York, whose respective Marketing Spend values add up to over 16 million, or 70.9 % of the total.
- Marketing Spend ranges from nearly 6.8 million, in California, to almost 8.5 million, in Florida.



## 8. Data Relationship Diagram:



### Inference:

- Across all states, the sum of Administration is over 13 million.
- Administration ranges from nearly 4.1 million, in Florida, to nearly 4.8 million, in New York.

### Conclusion:

The dataset was added to Cognos Analytics and successfully explored and visualized the data.