IBM NALAIYA THIRAN

ASSIGNMENT 1

Name: KARAN N S
Roll No: 737819CSR076
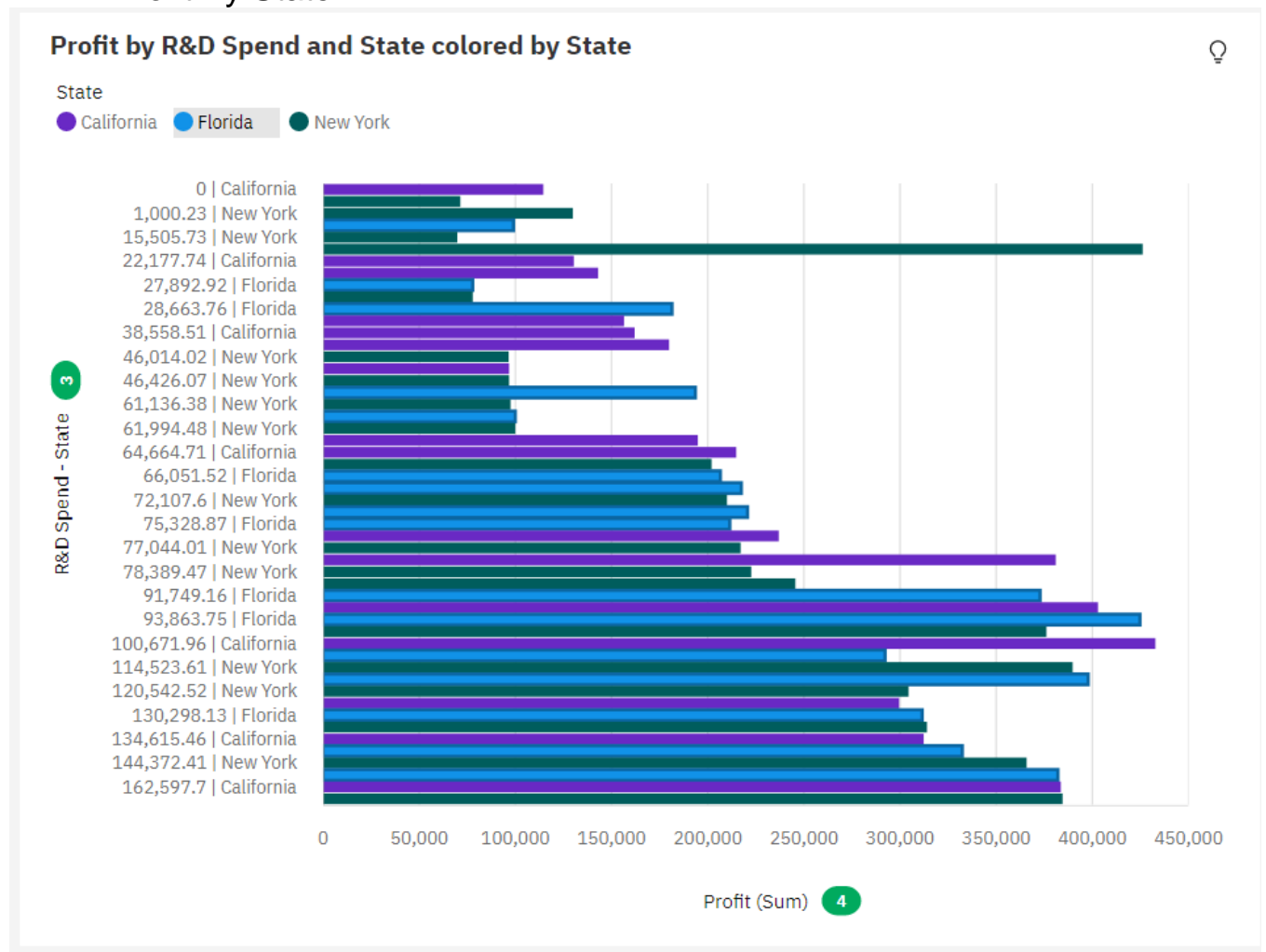
Challenge:

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

Dataset: 50-Startup-Dataset

Exploration and Visualization:

1. Profit By State

Profit by R&D Spend and State colored by State

State
● California ● Florida ● New York

Y-axis (R&D Spend - State):
- 0 | California
- 1,000.23 | New York
- 15,505.73 | New York
- 22,177.74 | California
- 27,892.92 | Florida
- 28,663.76 | Florida
- 38,558.51 | California
- 46,014.02 | New York
- 46,426.07 | New York
- 61,136.38 | New York
- 61,994.48 | New York
- 64,664.71 | California
- 66,051.52 | Florida
- 72,107.6 | New York
- 75,328.87 | Florida
- 77,044.01 | New York
- 78,389.47 | New York
- 91,749.16 | Florida
- 93,863.75 | Florida
- 100,671.96 | California
- 114,523.61 | New York
- 120,542.52 | New York
- 130,298.13 | Florida
- 134,615.46 | California
- 144,372.41 | New York
- 162,597.7 | California

X-axis: 0, 50,000, 100,000, 150,000, 200,000, 250,000, 300,000, 350,000, 400,000, 450,000
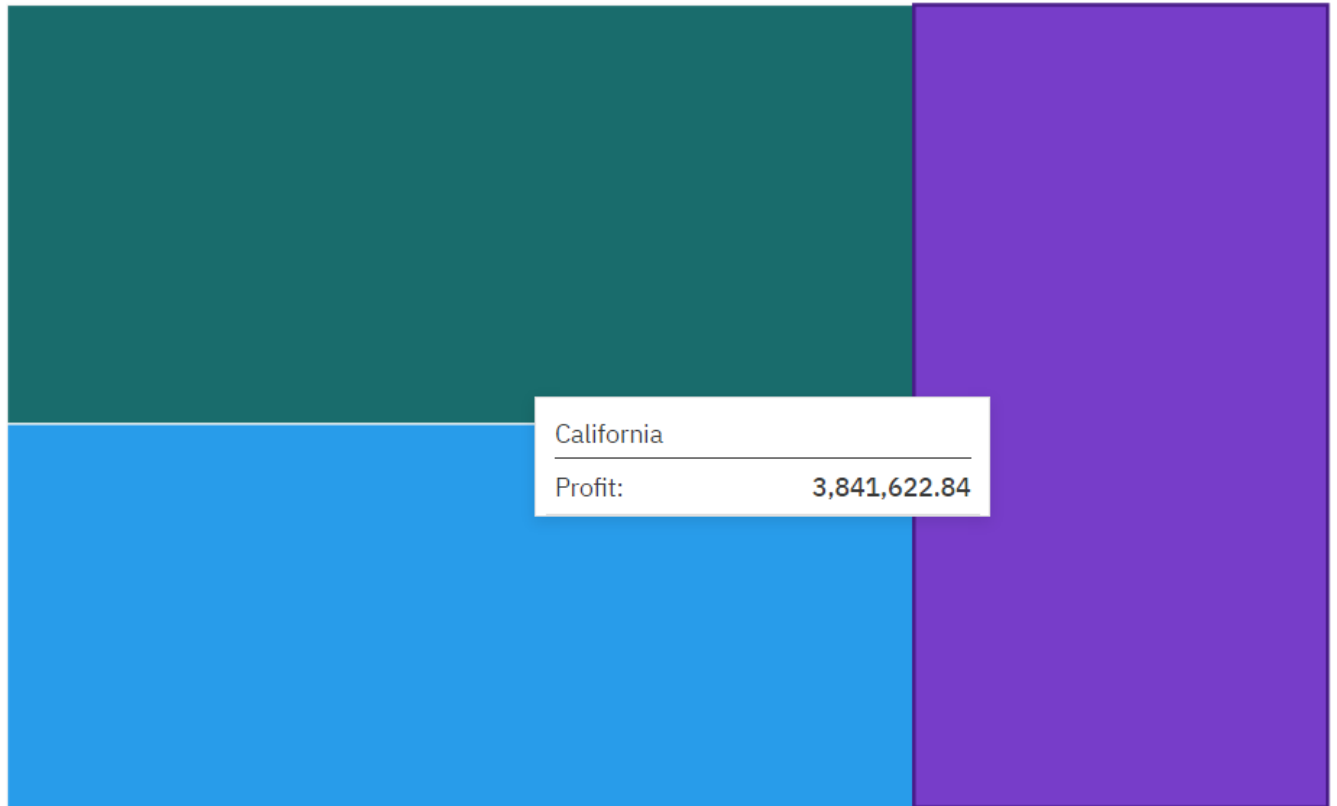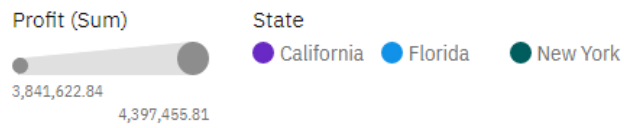
Profit (Sum) 4

Inference:

➢ For Profit, the most significant values of R&D Spend - State are 55493.95|Florida and 27892.92Florida, 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.
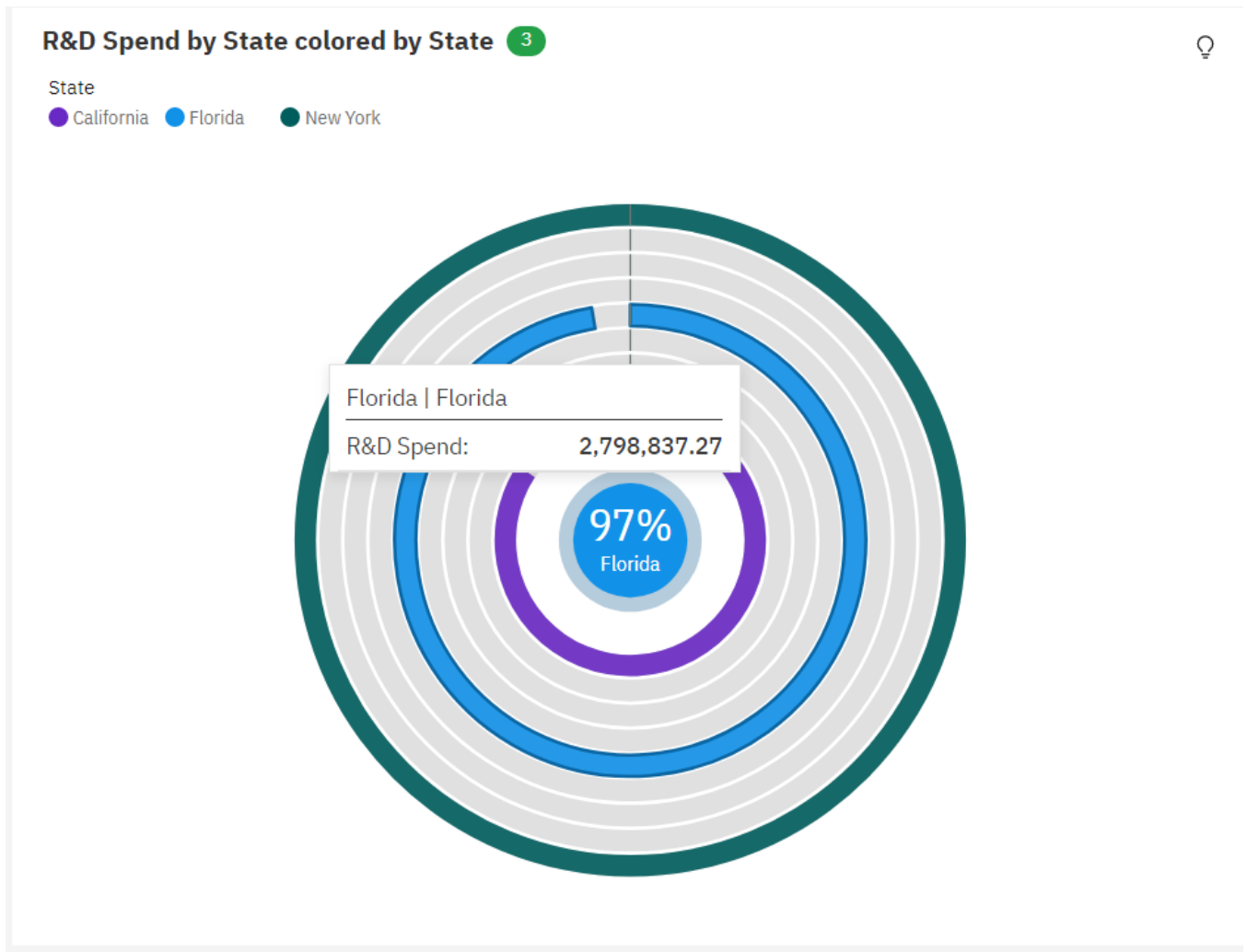
2. Profit by State Colored by State

Profit for State hierarchy 4

Profit (Sum)

3,841,622.84
      4,397,455.81

State
● California   ● Florida   ● New York

California	
Profit:	3,841,622.84

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

R&D Spend by State colored by State ③

State
● California ● Florida ● New York

Florida | Florida

R&D Spend: 2,798,837.27

97%
Florida

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
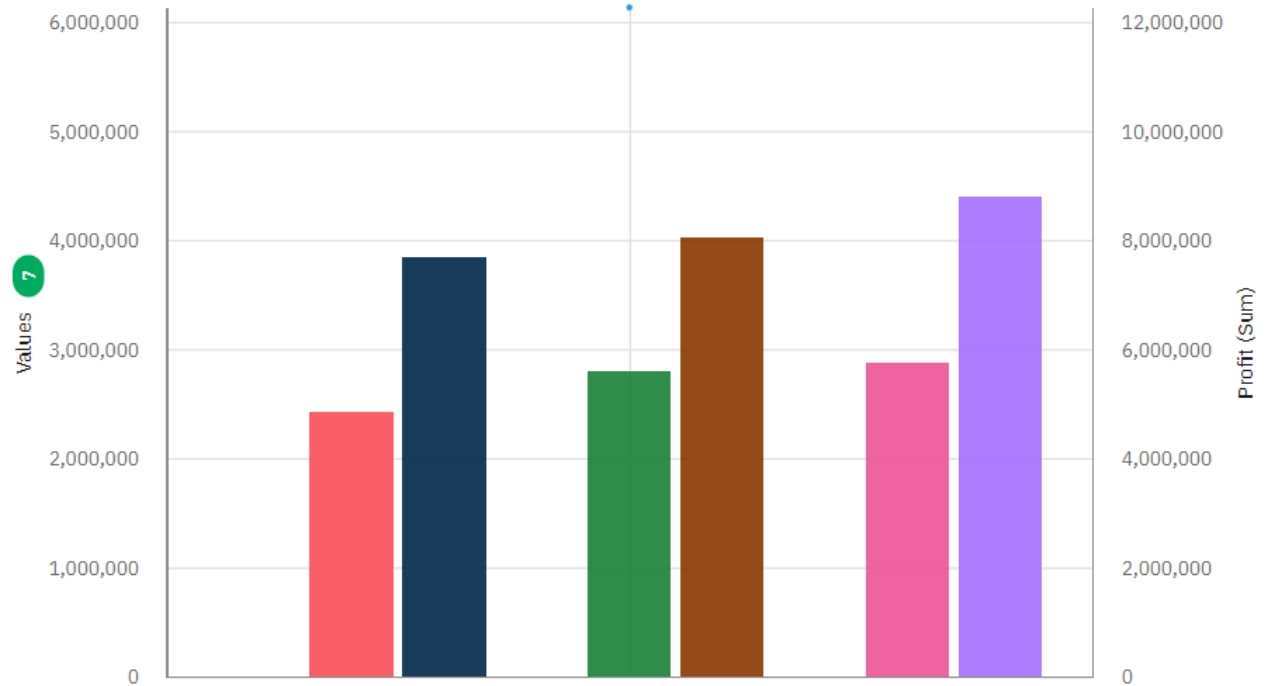- Florida | R&D Spend
- Florida | Profit
- New York | State
- New York | R&D Spend
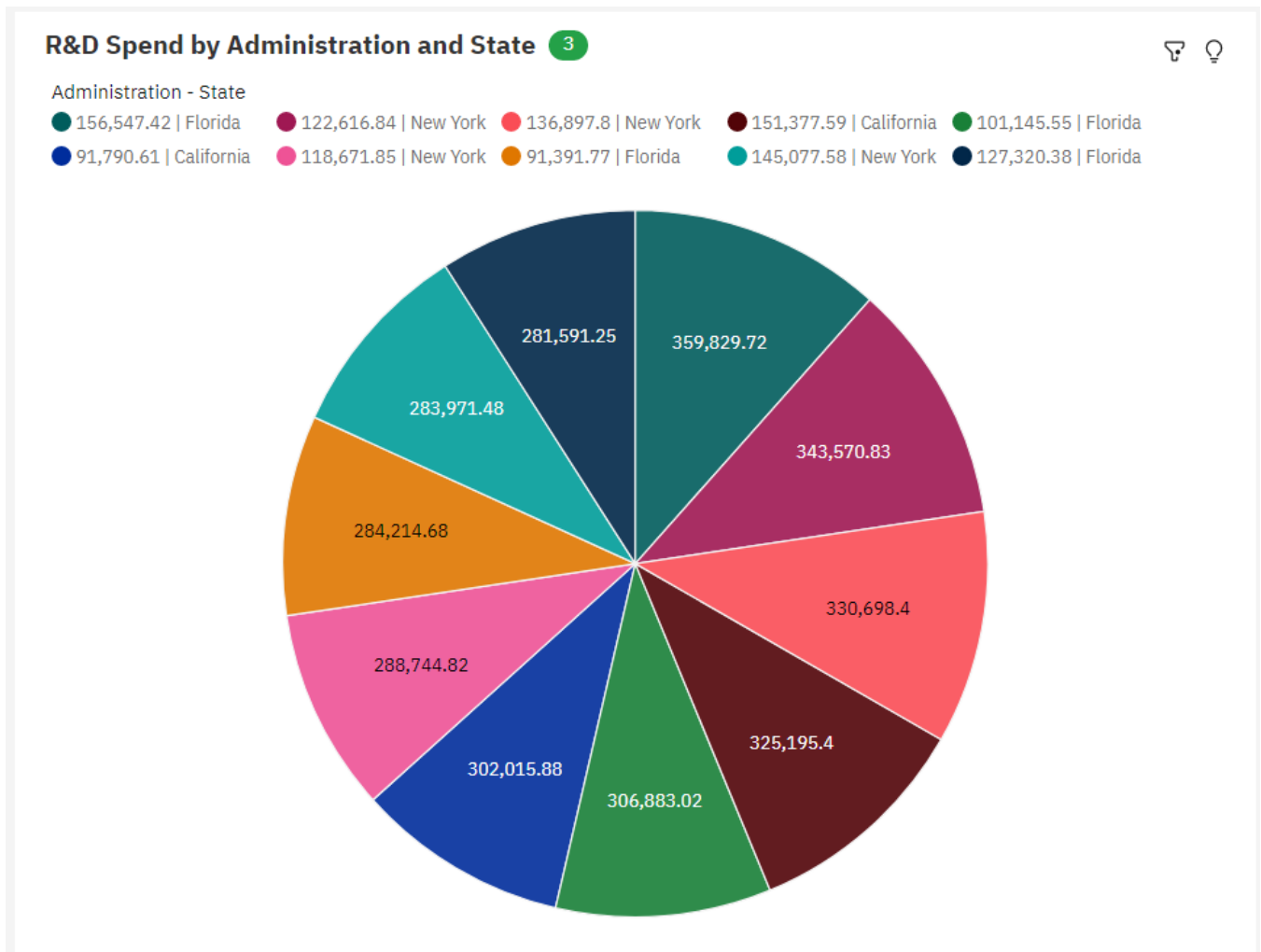- New York | Profit

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.
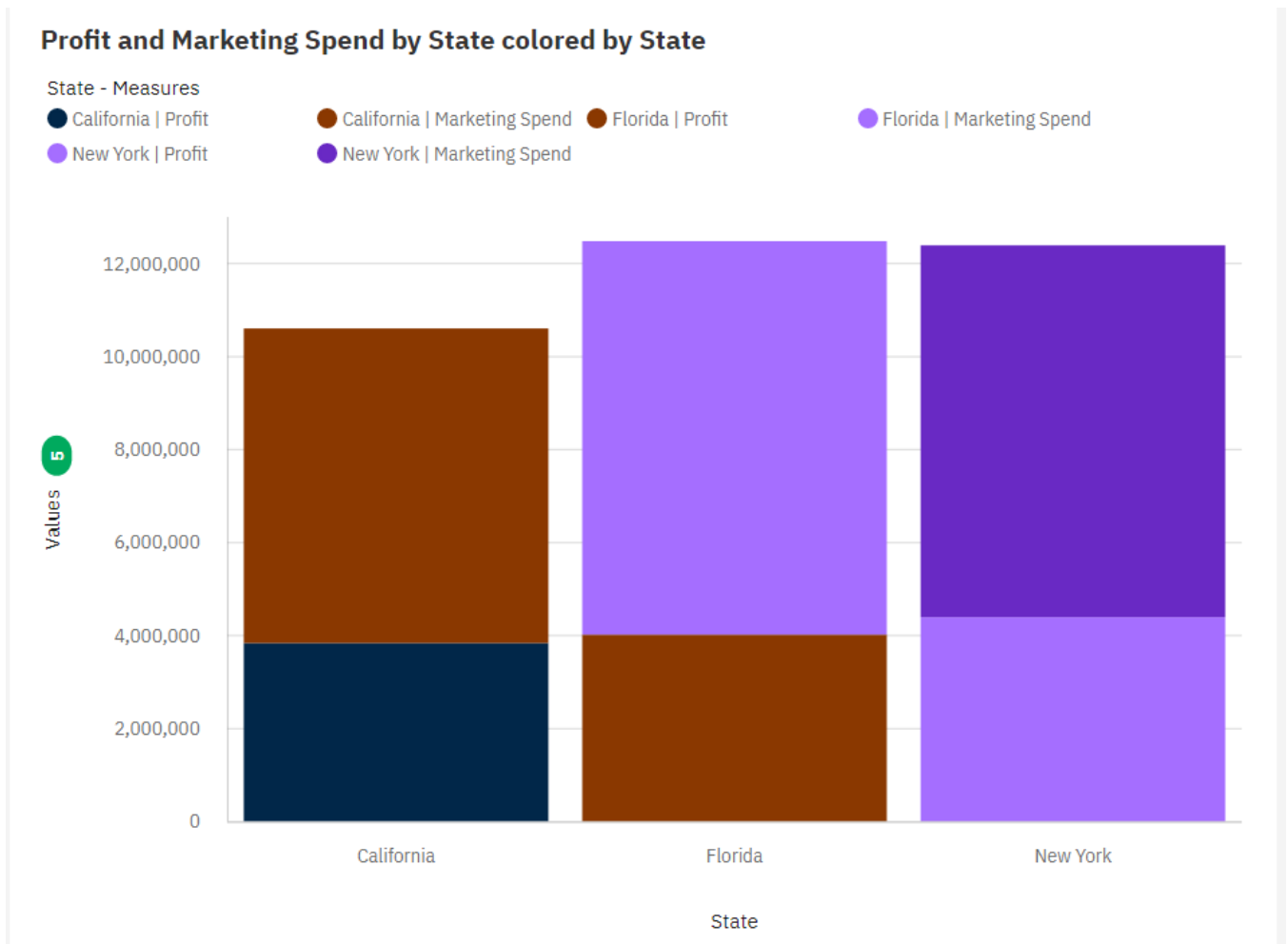
➢ The average of R&D Spend is almost 75 thousand.

5. R&D Spend and Profit by State

R&D Spend by Administration and State ③

Administration - State
- 156,547.42 | Florida
- 122,616.84 | New York
- 136,897.8 | New York
- 151,377.59 | California
- 101,145.55 | Florida
- 91,790.61 | California
- 118,671.85 | New York
- 91,391.77 | Florida
- 145,077.58 | New York
- 127,320.38 | Florida

Pie chart values: 359,829.72; 343,570.83; 330,698.4; 325,195.4; 306,883.02; 302,015.88; 288,744.82; 284,214.68; 283,971.48; 281,591.25

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.
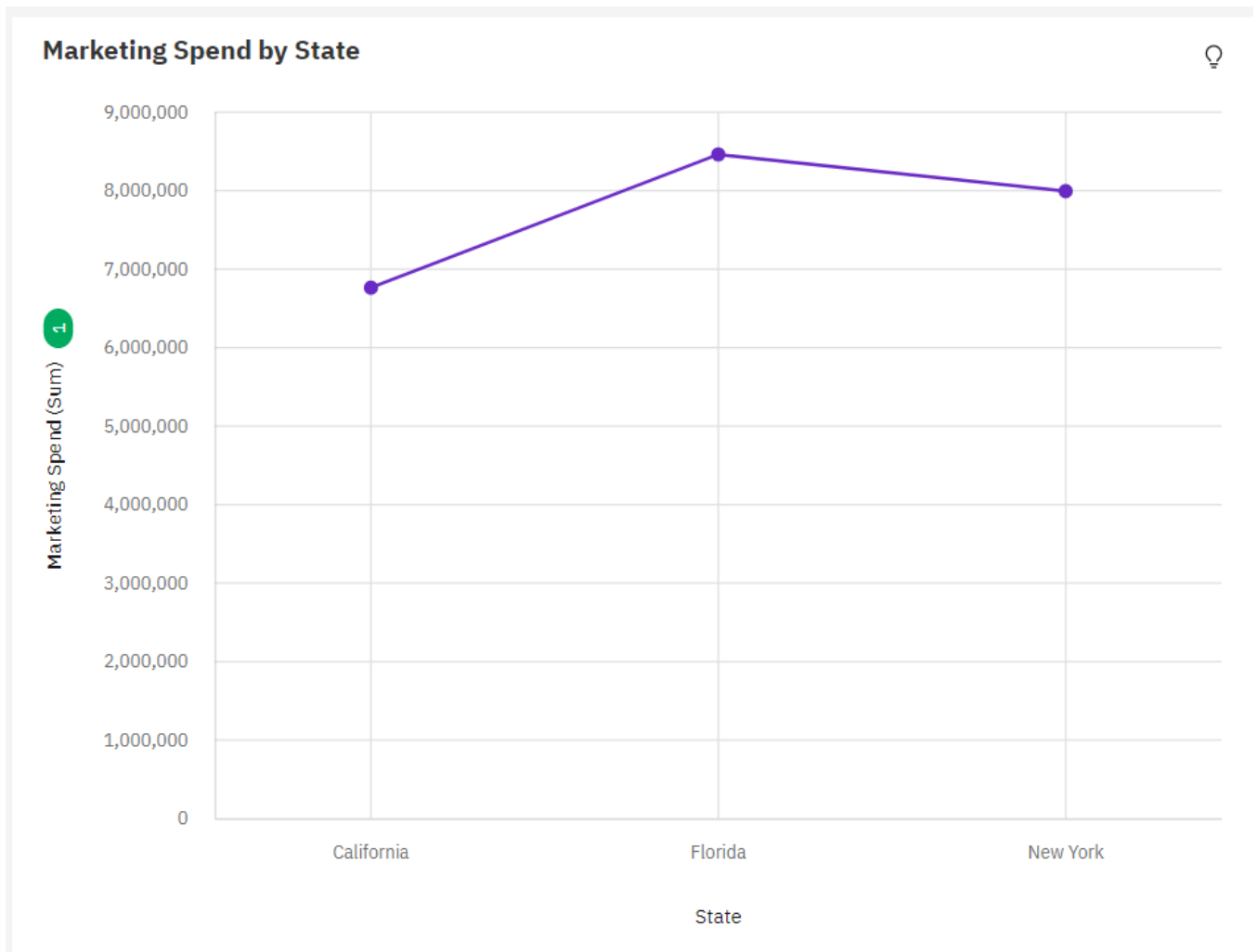
6. Profit Prediction Using Decision Tree:

Profit and Marketing Spend by State colored by State

State - Measures
- California | Profit
- California | Marketing Spend
- Florida | Profit
- Florida | Marketing Spend
- New York | Profit
- New York | Marketing Spend

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.
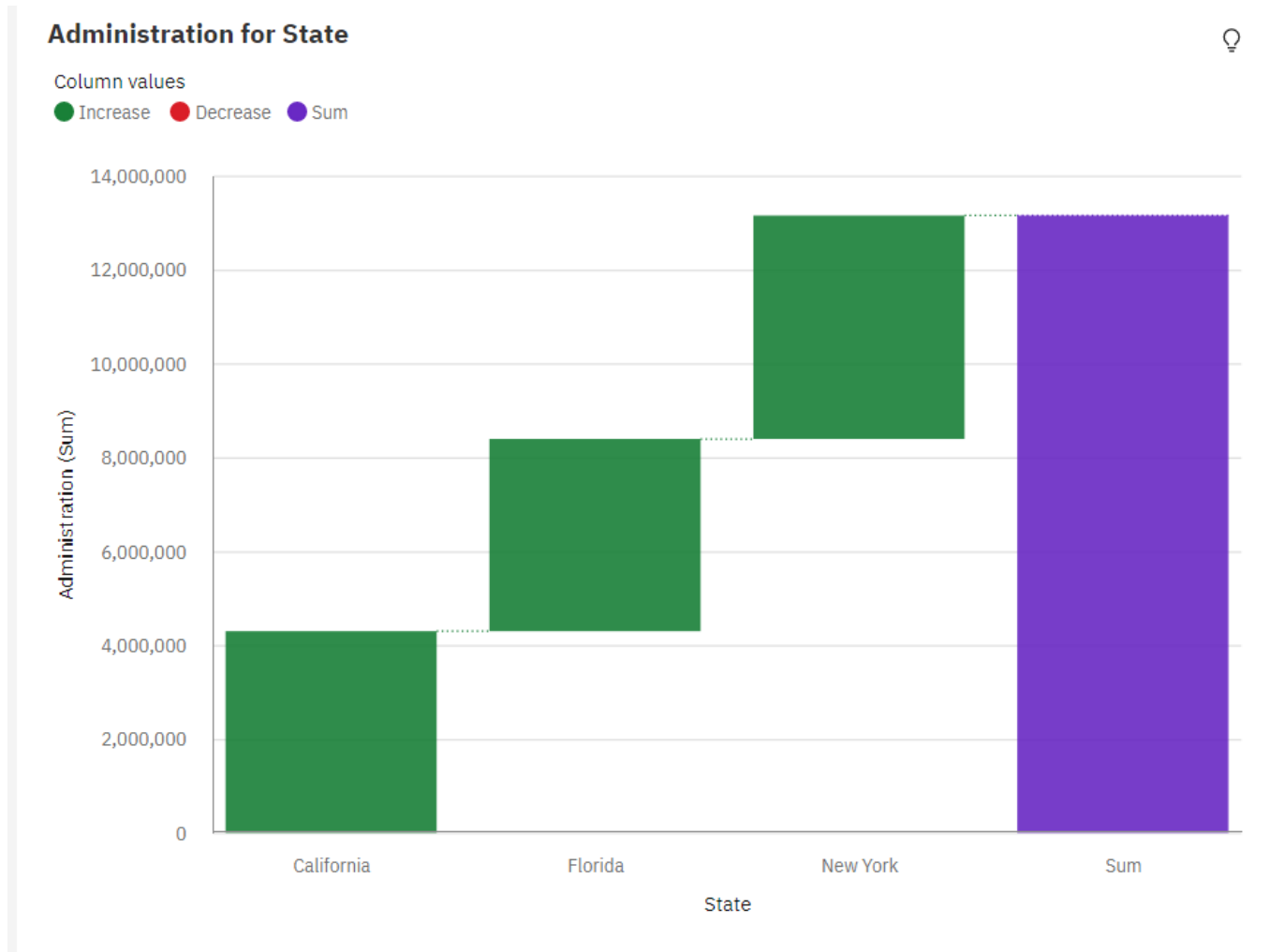
7. Profit by Administration (Group)

Marketing Spend by State

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:

Administration for State

Column values
● Increase ● Decrease ● Sum

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.