

IBM NALIYA THIRAN

TEAM MEMBERS

JESWIN W – 917719IT040

JEYA GANESH A V – 917719IT041

RAHUL HARIESH B – 917719IT074

SANJAYKUMAR S – 917719IT082

ASSIGNMENT – 1

Dataset Link:

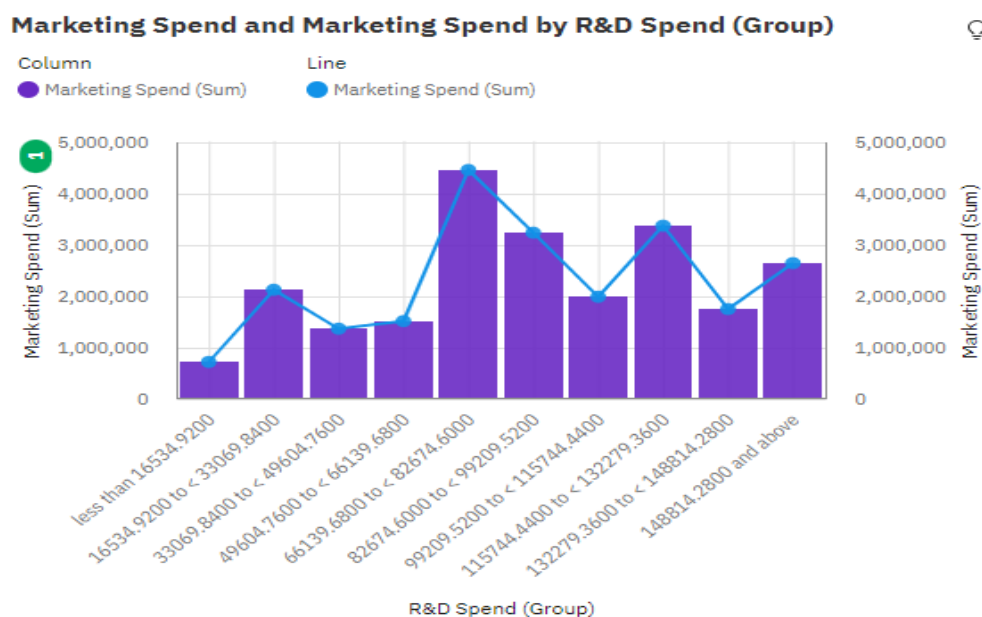
https://drive.google.com/file/d/100OUpJFG8GQjhFztfG5JP_jvzRgoihr/view

Challenge:

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

Exploration of Dataset:

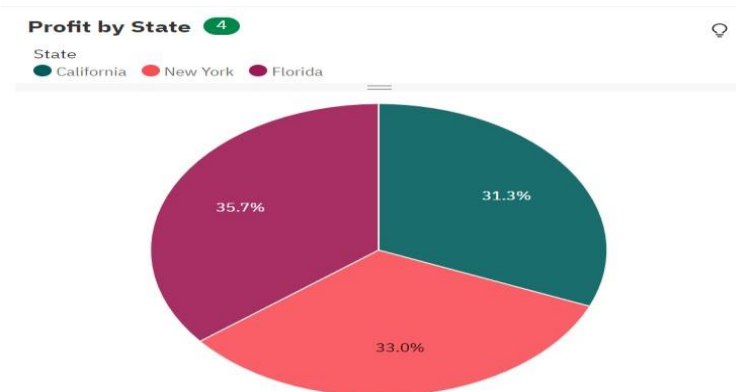
Visualization the Dataset:



Inference:

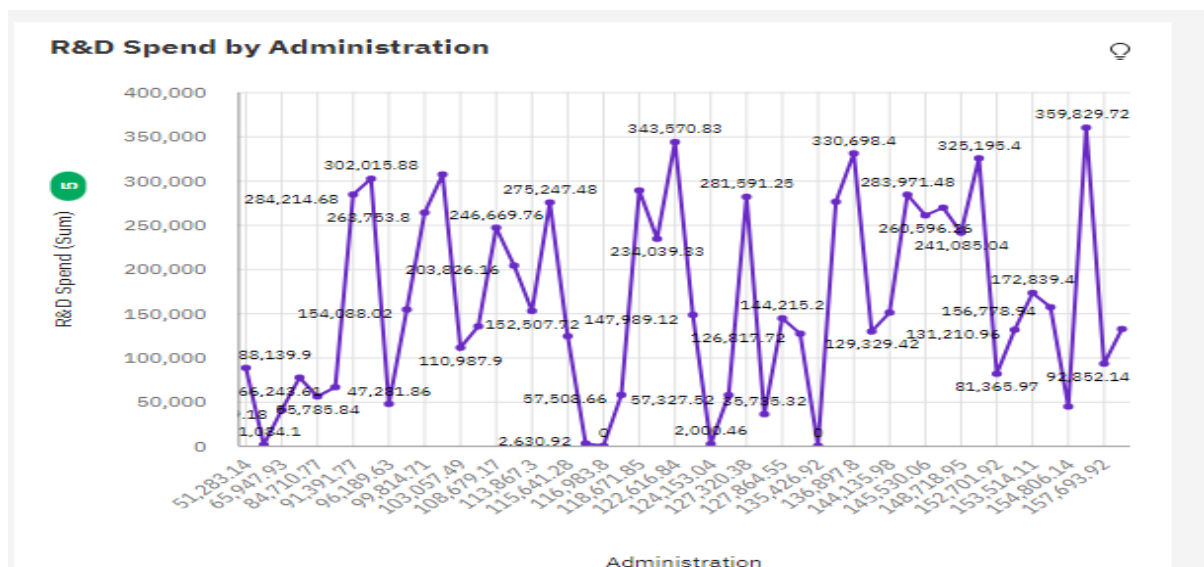
Marketing Spend ranges from almost 724 thousand, when R&D Spend (Group) is less than 16534.9200, to almost 4.5 million, when R&D Spend (Group) is 66139.6800 to < 82674.6000.. cross all values of R&D Spend (Group), the sum of Marketing Spend is over 23 million.

Profit and Marketing spend by State:



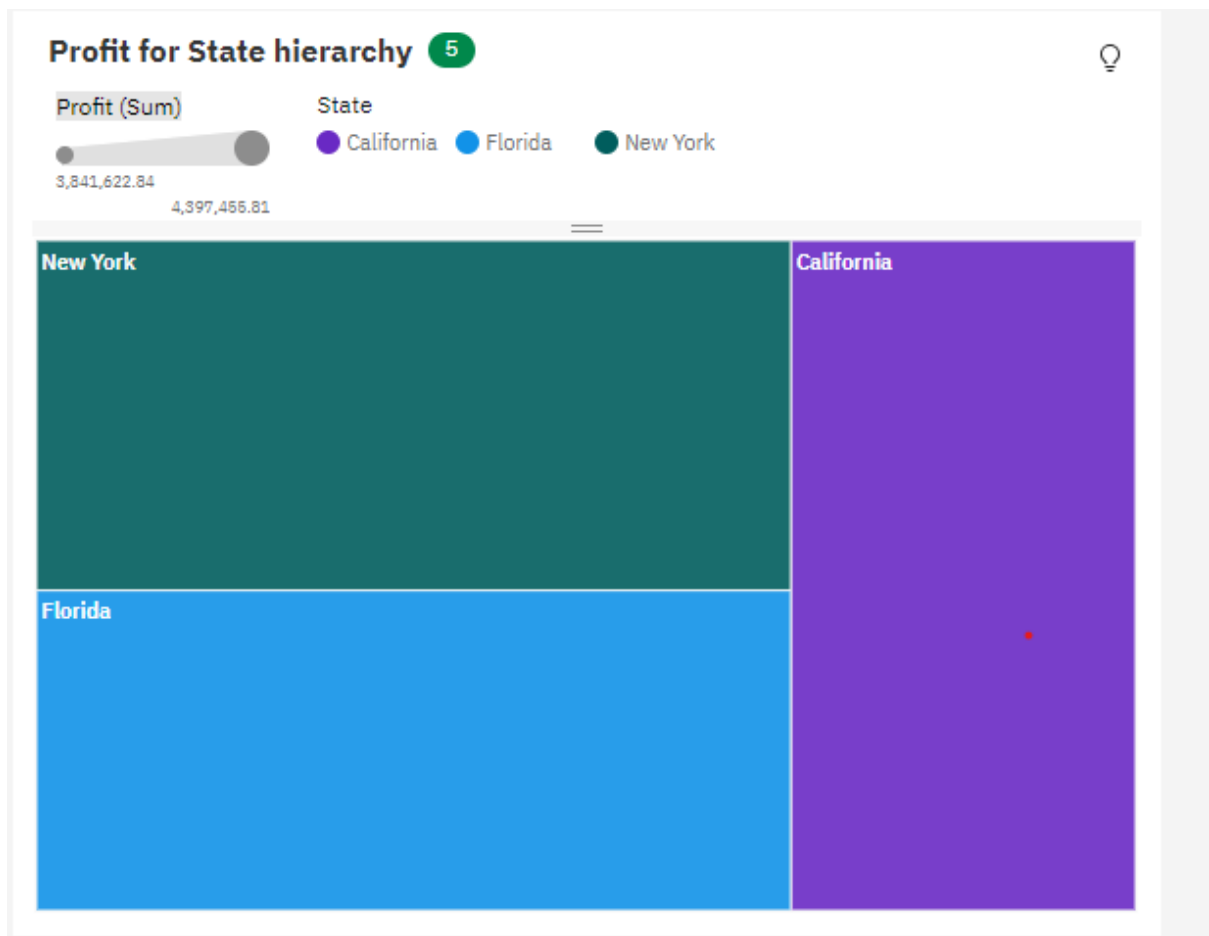
Inference:

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.



Inference:

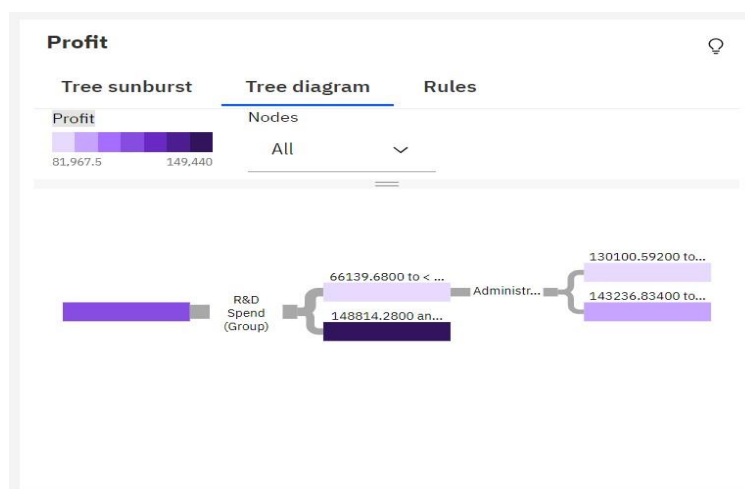
Over all administrations, the sum of R&D Spend is nearly 8.1 million. R&D Spend ranges from 0, when Administration is 116983.8, to nearly 360 thousand, when Administration is 156547.42.



Inference:

The average profit is high in Florida with more than 121,000 profit. Then followed by Florida and then by California.

Decision Tree:



Inference:

R&D Spend (Group) and Administration (Group) predict Profit with a strength of 62.7%.

R&D Spend (Group) is the most significant predictor of Profit being ten times better than any other field.

Marketing Spend by Profit with points for Marketing Spend**Inference:**

The marketing spend and profit are related in a medium level.

Profit and Marketing Spend have a medium positive linear association, being $\text{Profit} = 64558 + 0.1002 * \text{Marketing Spend}$.