

TULASI MALINI V K
211719104154
IBM NALAIYA THIRAN
ASSIGNMENT - 1

The case of 50 Startups

Overview

In this article, we will be exploring the profit from the startup's dataset with the features available to us. We're using the 50-startups dataset for this problem statement

How do startups work?

Well, we can say that startups pipeline operates on the same principles which are similar to other MNCs the major difference between both of them is that on the one hand startups work to make products that are beneficial for the customers on a small scale while other established companies do that work on a large scale by re-doing something which is already being done. How startups are being funded?

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**.

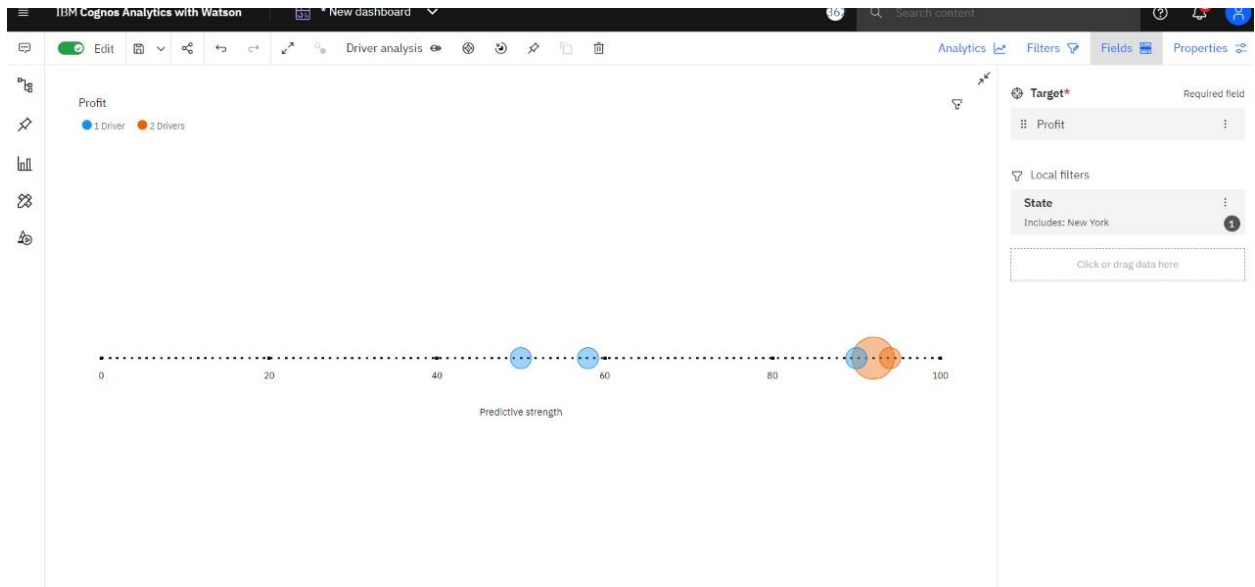
1. **R&D spending:** The amount which startups are spending on Research and development.
2. **Administration spending:** The amount which startups are spending on the Admin panel.
3. **Marketing spending:** The amount which startups are spending on marketing strategies.
4. **State:** To which state that particular startup belongs.
5. **Profit:** How much profit that particular startup is making.

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

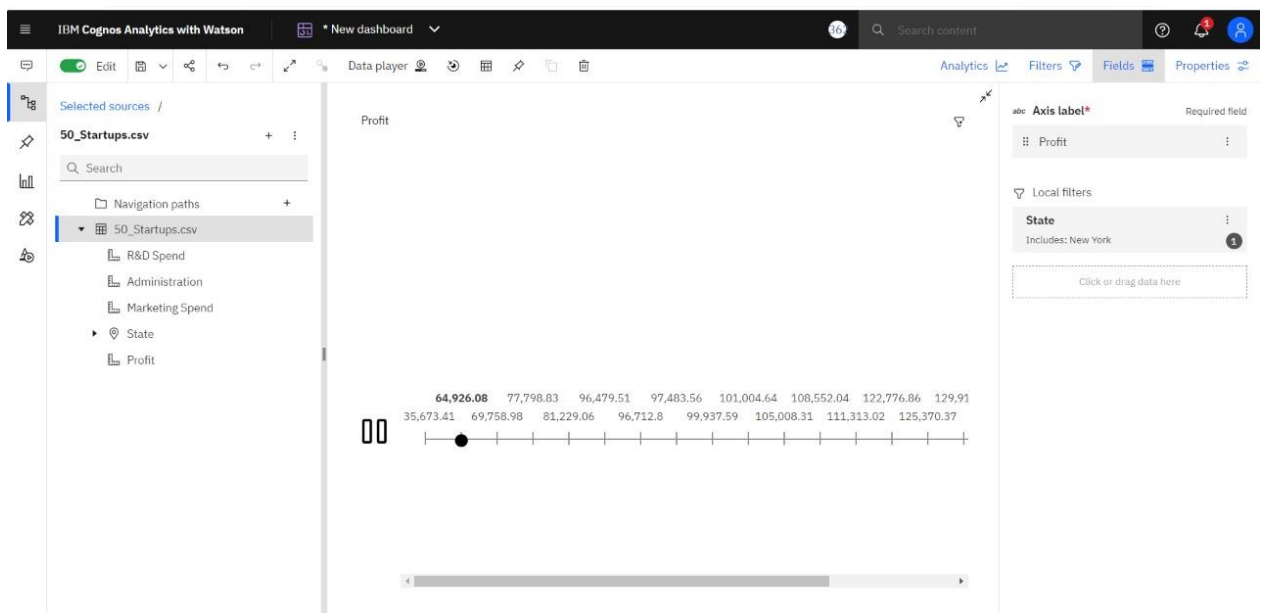
Challenge:

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

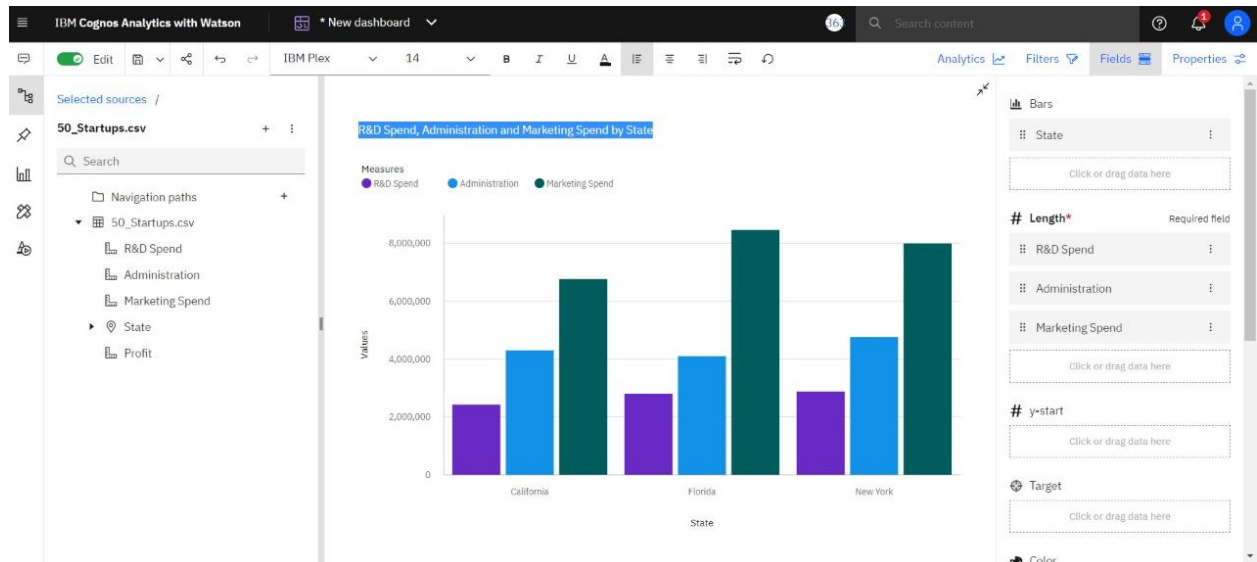
Visualizations:



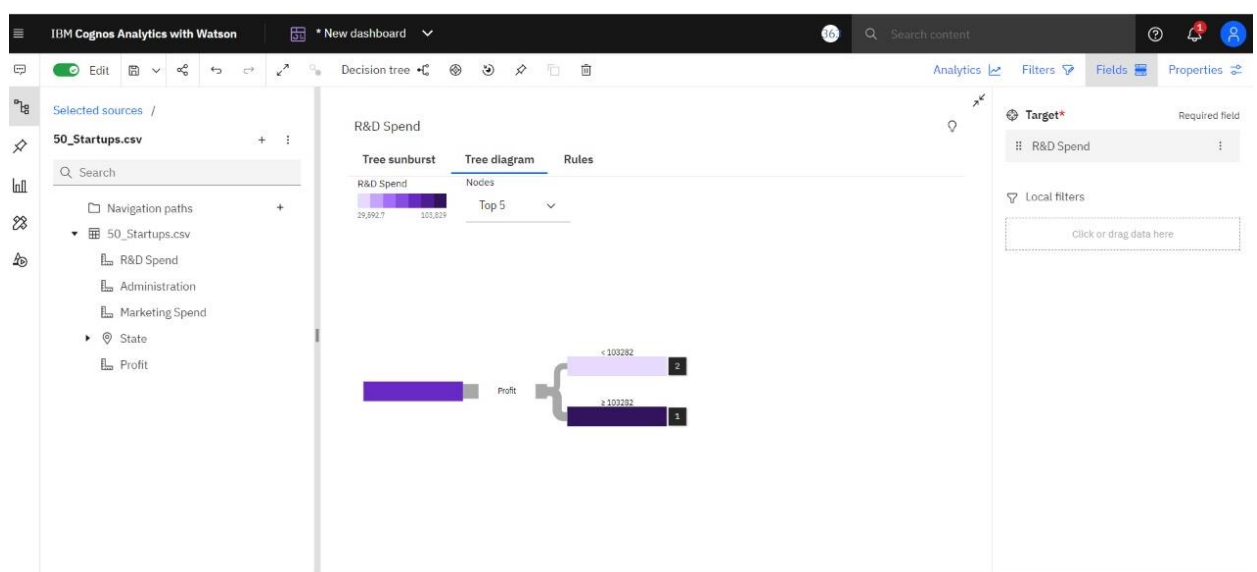
In the above Driver analysis graph the **Profit** is represented as target and it is filtered by **state (New York)**.



The above data player graph helps the user to understand the profit of New York by using the axis label as profit and the local filter as State (New York).



The following representation helps to understand the total spending of the Start-ups and analyse the spending on each sector state wise. **States** are plotted in X-axis and **R&D Spend, Administration** and **Marketing Spend** are plotted in Y- axis.



The following decision tree helps to evaluate the options. In our case **R&D Spend** contributes more (61%) to the profit