

A NEW HINT TO TRANSPORTATION - ANALYSIS OF THE NYC BIKE SHARE SYSTEM

TASK:

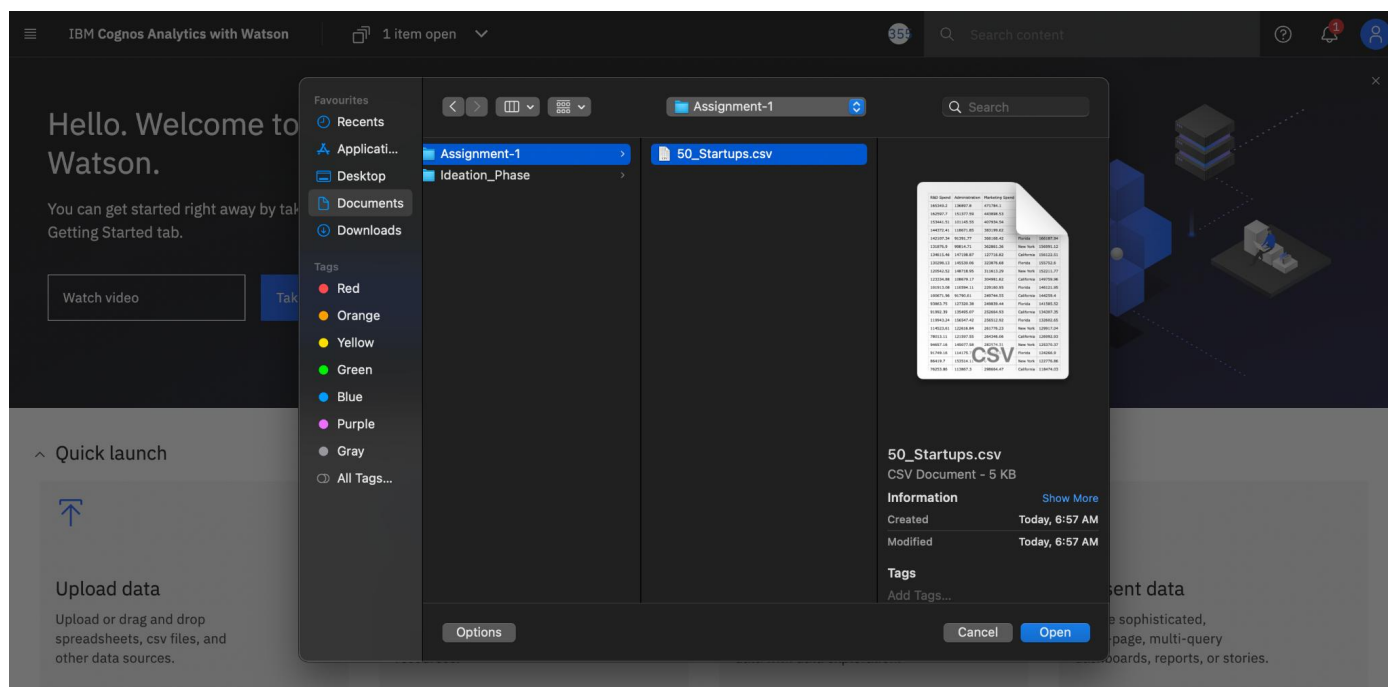
Upload the dataset to Cognos Analytics, Explore and Visualize the 50-startups dataset.

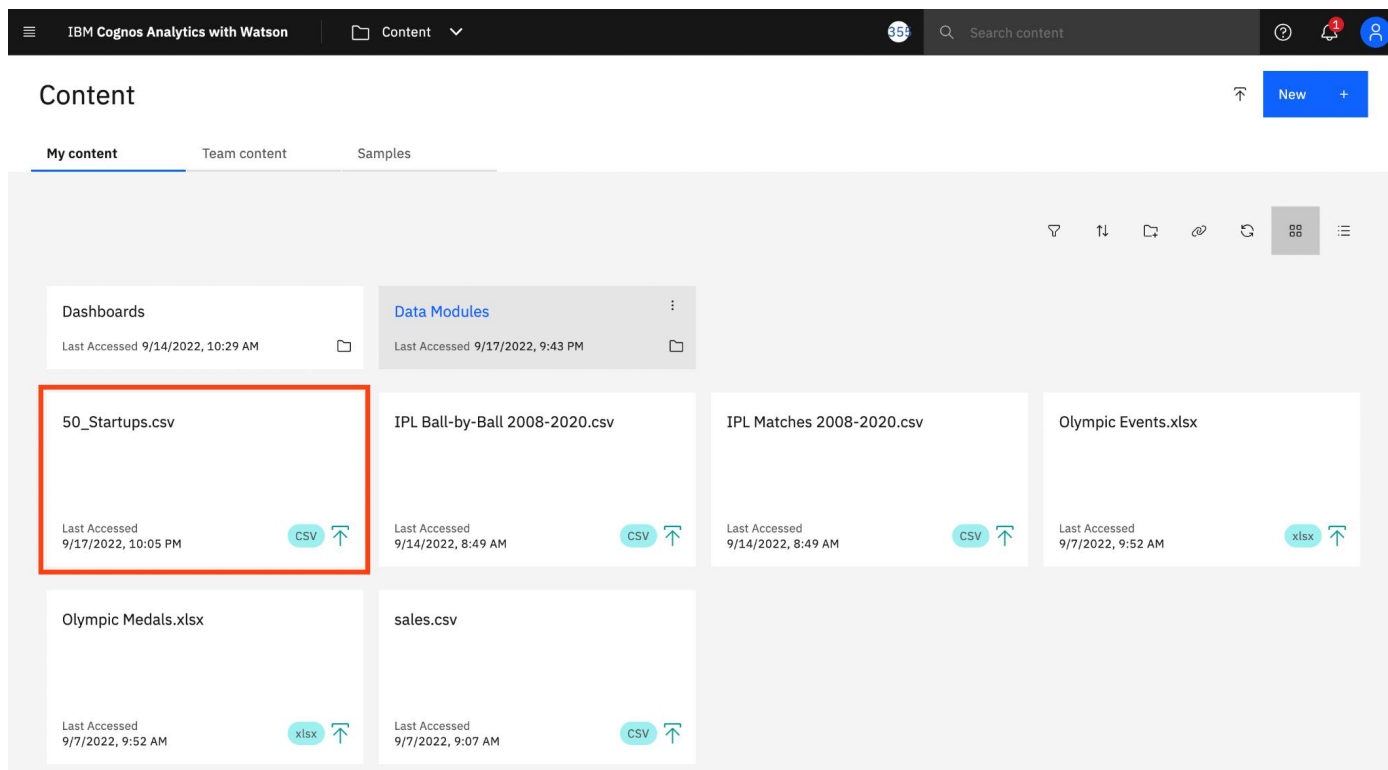
DATASET USED:

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.

SOLUTION:



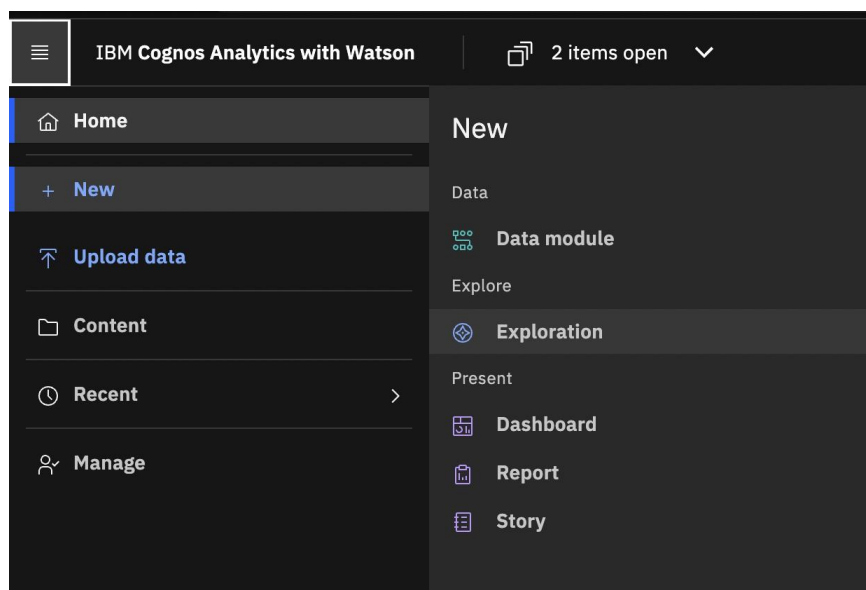


The screenshot shows the 'Content' page in IBM Cognos Analytics. The top navigation bar includes the IBM Cognos Analytics with Watson logo, a 'Content' dropdown, a search bar, and user profile icons. The main section is titled 'Content' and has tabs for 'My content', 'Team content', and 'Samples'. A 'New +' button is in the top right. The content area displays a grid of items: 'Dashboards', 'Data Modules', and several CSV and XLSX files. The file '50_Startups.csv' is highlighted with a red box. Other files include 'IPL Ball-by-Ball 2008-2020.csv', 'IPL Matches 2008-2020.csv', 'Olympic Events.xlsx', 'Olympic Medals.xlsx', and 'sales.csv'. Each file card shows the last accessed time and a small icon indicating the file type (CSV or XLSX).

As highlighted by the red box, the dataset *50_Startups.csv* was successfully **uploaded** onto the IBM Cognos platform.

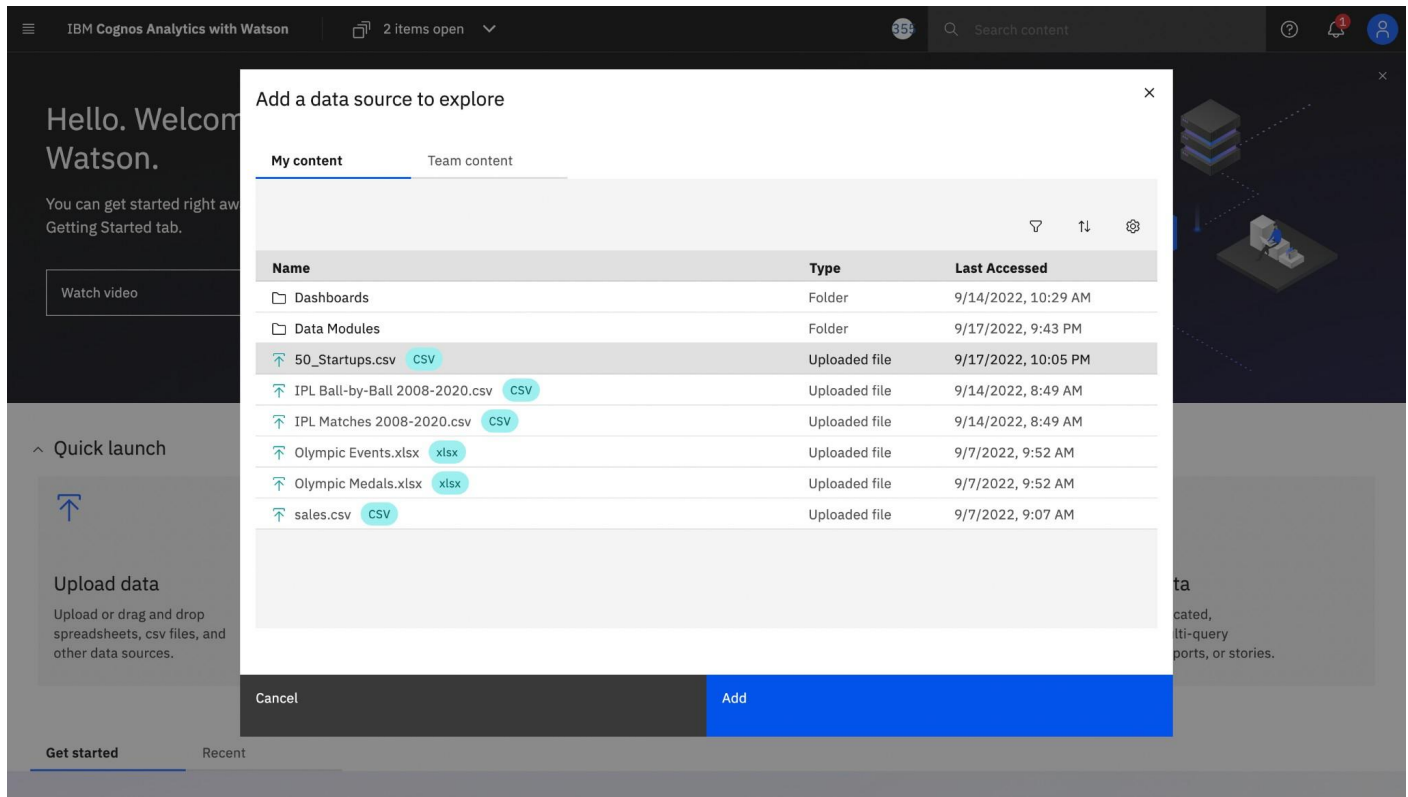
EXPLORATION AND VISUALIZATION OF THE DATASET:

As a standard step for exploration and visualization, A new exploration was created so that the dataset can be analyzed and visualized to provide further insights.



The screenshot shows the 'New' menu in IBM Cognos Analytics. The left sidebar contains navigation options: Home, New, Upload data, Content, Recent, and Manage. The main area displays the 'New' menu with options: Data, Data module, Explore, Exploration, Present, Dashboard, Report, and Story. The 'Exploration' option is highlighted.

The data source was then added to carry out the exploration.



In order to decide which visualizations to create, the data was first analyzed and the features understood. The grid data was accessed using the **Prepare Data** option.

+

Q

Search

New data module

Navigation paths

50_Startups.csv

#

Row Id

R&D Spend

Administration

Marketing Spend

State

Profit

Grid

Relationships

Custom tables

Row Id

R&D Spend

Administration

Marketing Spend

State

Profit

1

165349.2

136897.8

471784.1

New York

192261.83

2

162597.7

151377.59

443898.53

California

191792.06

3

153441.51

101145.55

407934.54

Florida

191050.39

4

144372.41

118671.85

383199.62

New York

182901.99

5

142107.34

91391.77

366168.42

Florida

166187.94

6

131876.9

99814.71

362861.36

New York

156991.12

7

134615.46

147198.87

127716.82

California

156122.51

8

130298.13

145530.06

323876.68

Florida

155752.6

9

120542.52

148718.95

311613.29

New York

152211.77

10

123334.88

108679.17

304981.62

California

149759.96

11

101913.08

110594.11

229160.95

Florida

146121.95

12

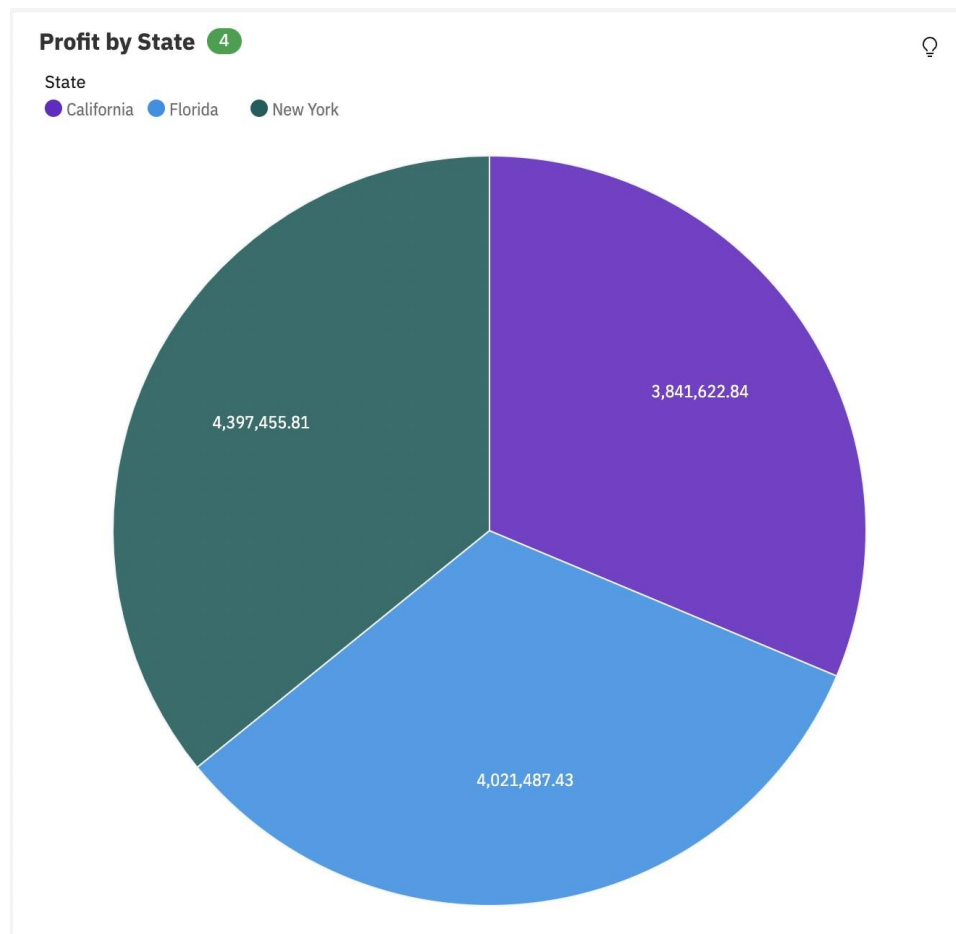
100671.96

91790.61

249744.55

California

144259.4

Visualization #1:

This Pie chart visualization depicts the total profit earned by startups belonging to each state present in the Dataset.

Fields

Segments* Required field

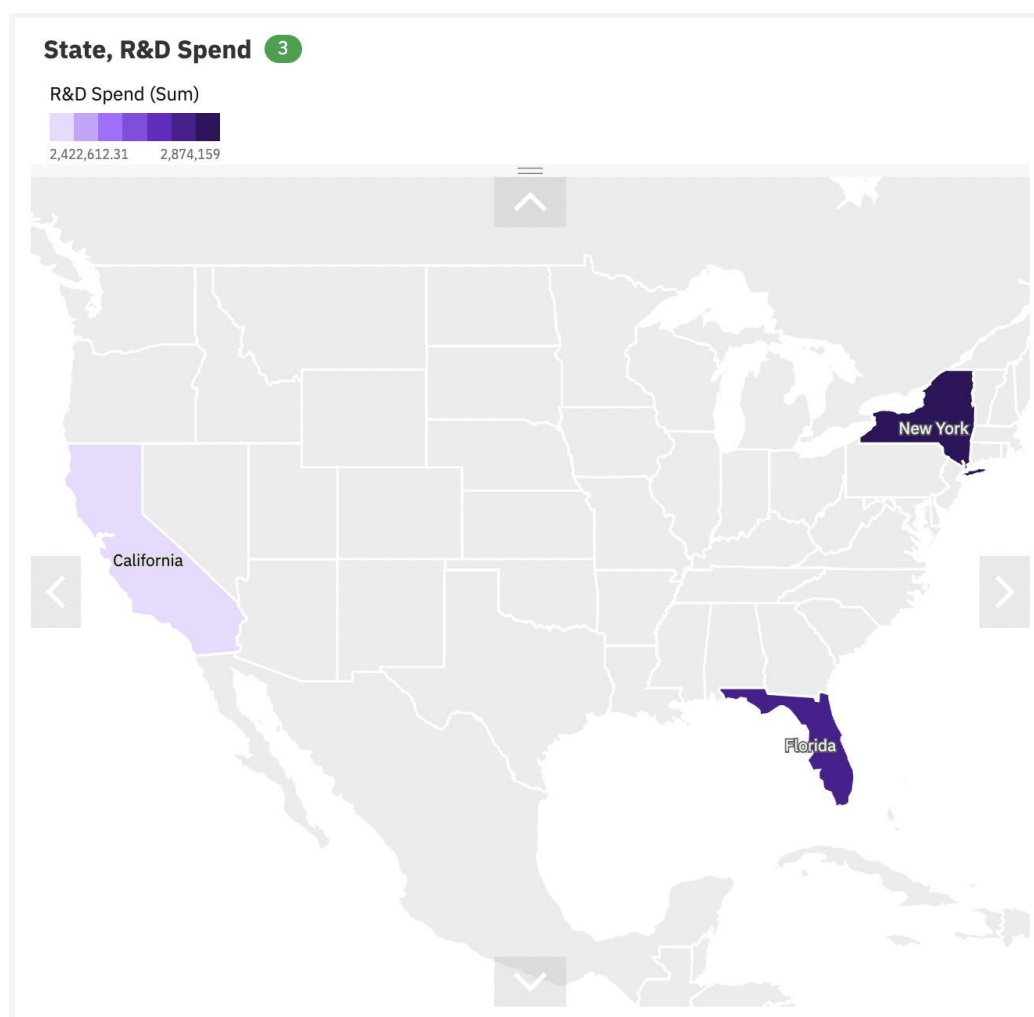
State

Click or drag data here

Size* Required field

Profit

Visualization #2:



This Geospatial Map (Geolocation) visualization shows the State-wise startup expenditure towards Research and Development field.

The color gradient indicates the amount startups in the states spent, with the startups in New York spending the most (indicated by **dark purple**), while California spending the least in R&D (indicated by **lavender**).

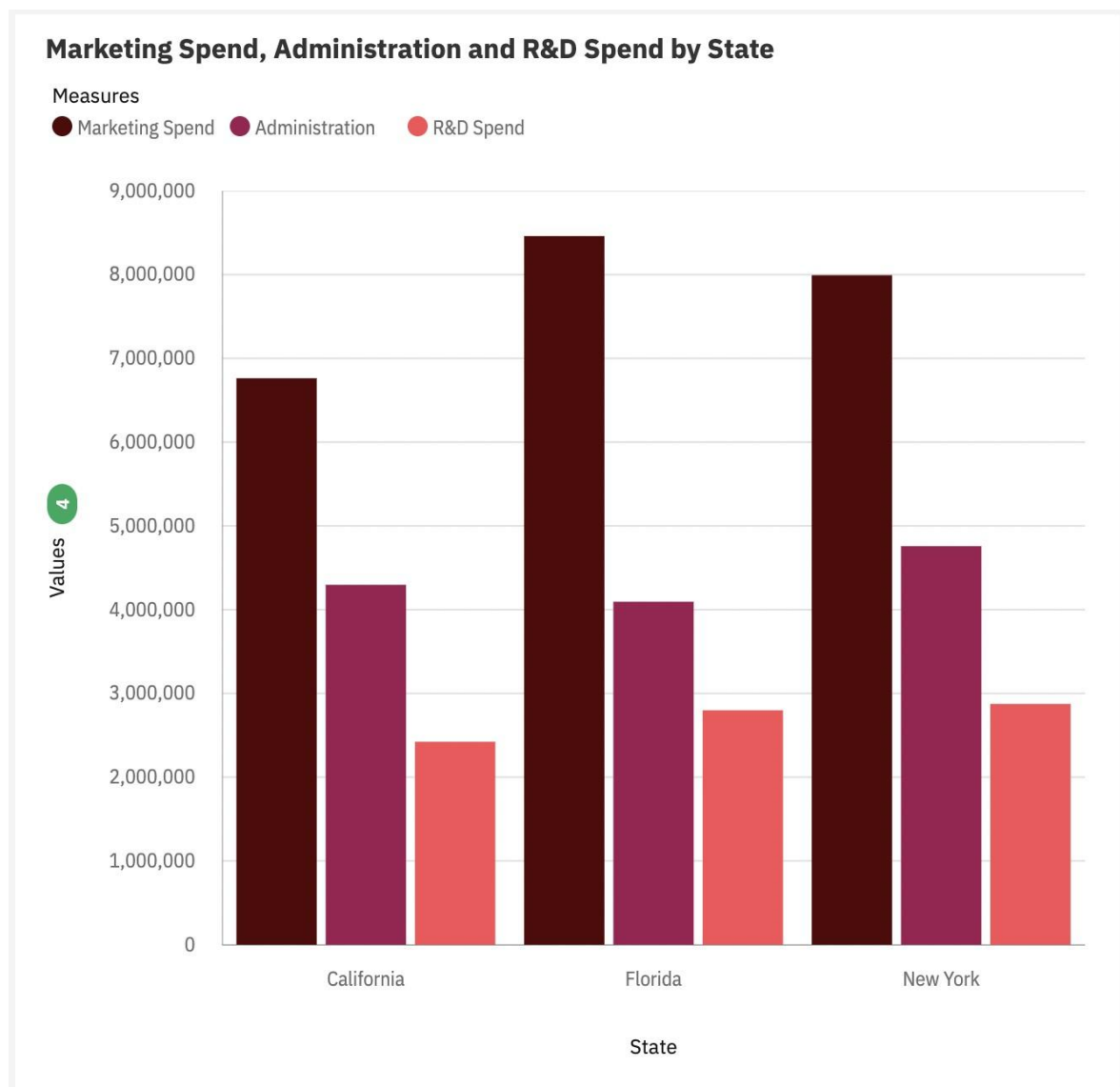
Fields

📍 **Regions*** Required field

⋮ State ⋮

■ ■ **Regions heat**

⋮ R&D Spend ⋮

Visualization #3:

This visualization shows the category-wise expenditure by startups belonging to the states of California, Florida, and New York. This will instantly provide insight with regard to the amount of money spent for each category, bearing in mind the location of the startups.

Fields

Bars

State

Click or drag data here

Length*

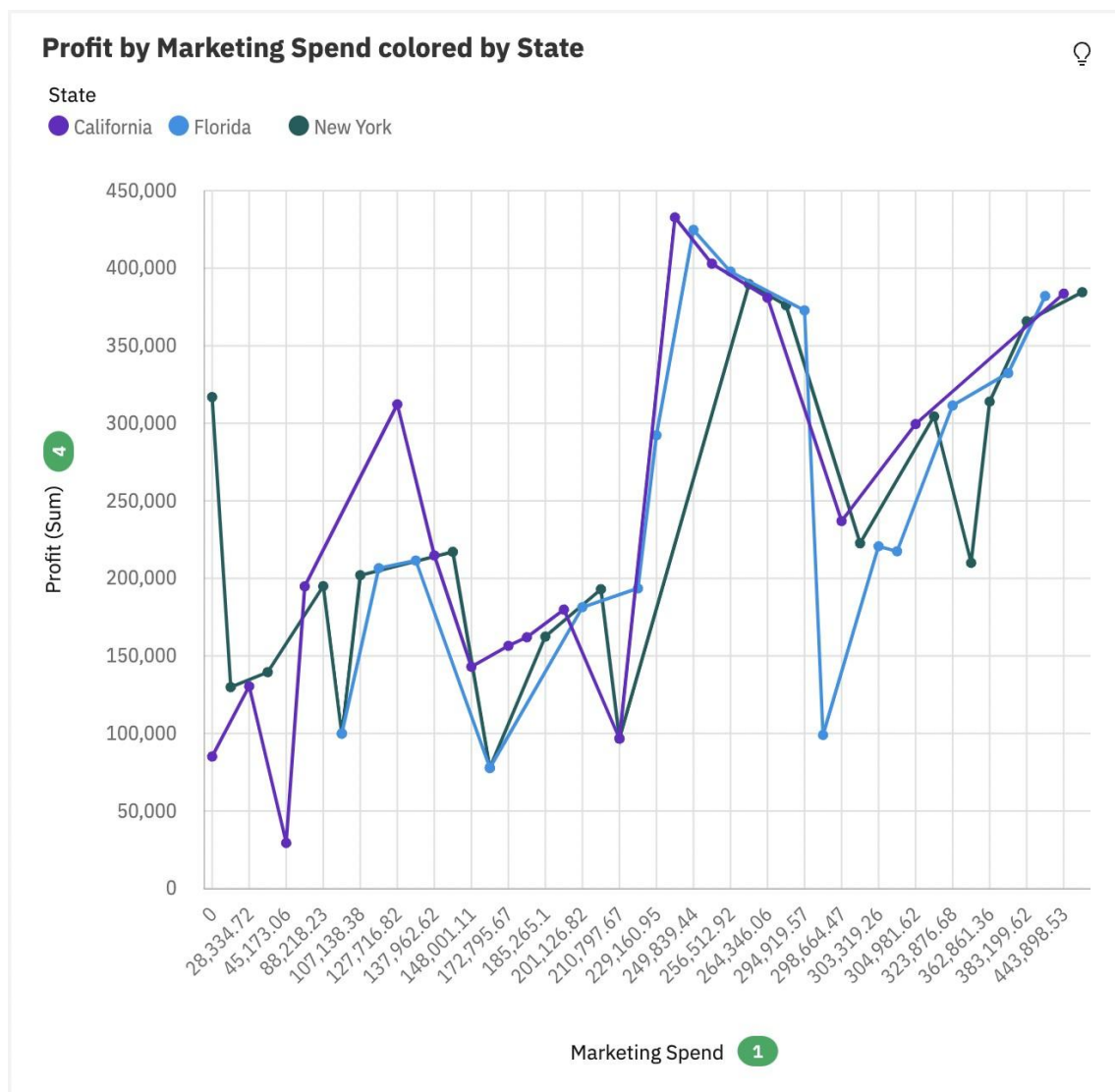
Required field

Marketing Spend

Administration

R&D Spend

Click or drag data here

Visualization #4:

This visualization provides insight regarding the profit made by comparing it with the Marketing Spend - for the startups in all three states as indicated by different colored lines.

Fields

↗ **x-axis*** Required field

Marketing Spend

⋮

Click or drag data here

Color

State

⋮

Click or drag data here

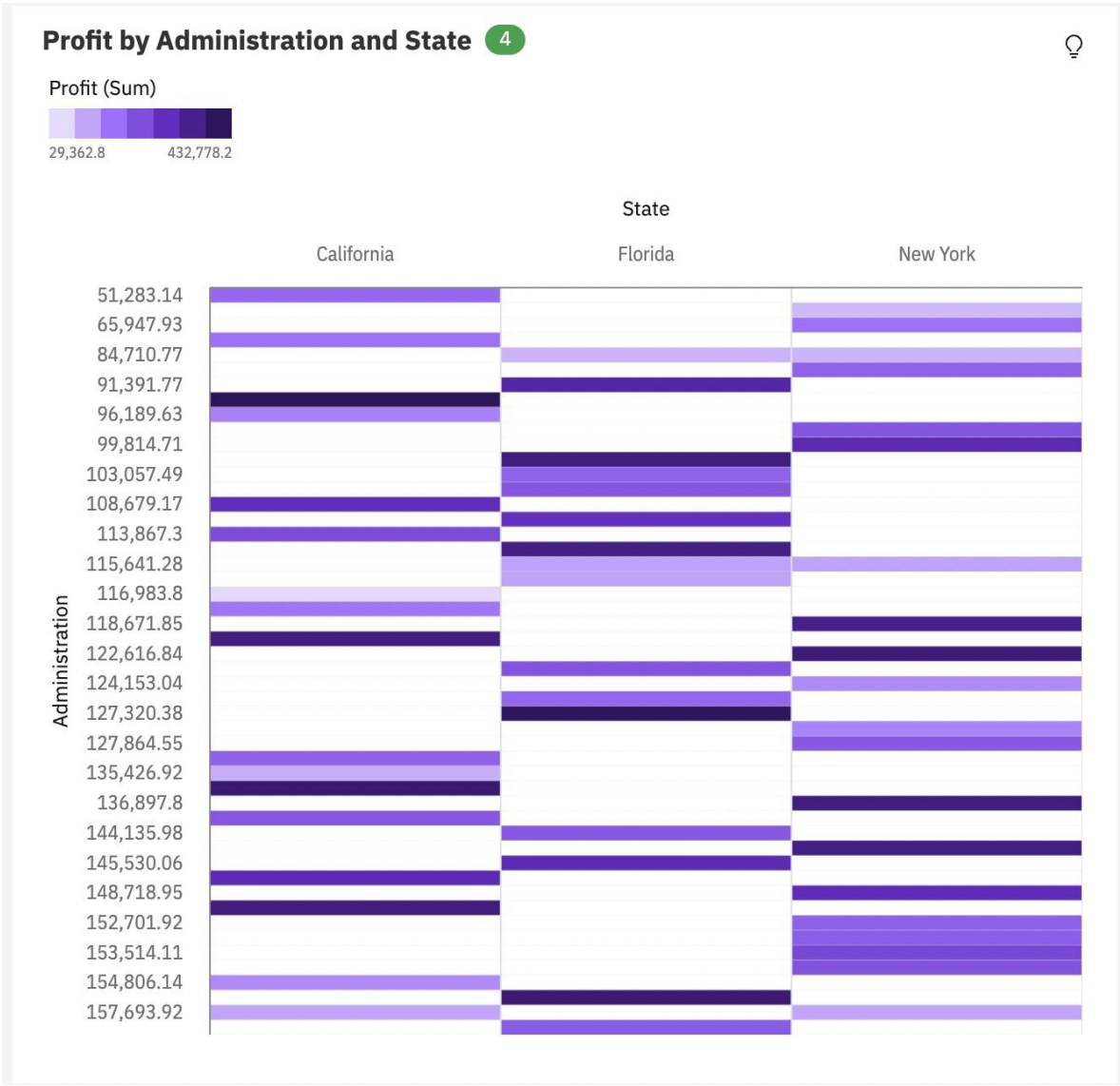
y-axis* Required field

Profit

⋮

Click or drag data here

Visualization #5:



This visualization represents a heatmap that bases its gradient color on the profit earned - Statewise with regards to Administration spending of companies in each state.

Fields

Rows

Administration

Click or drag data here

Columns

State

Click or drag data here

Heat*

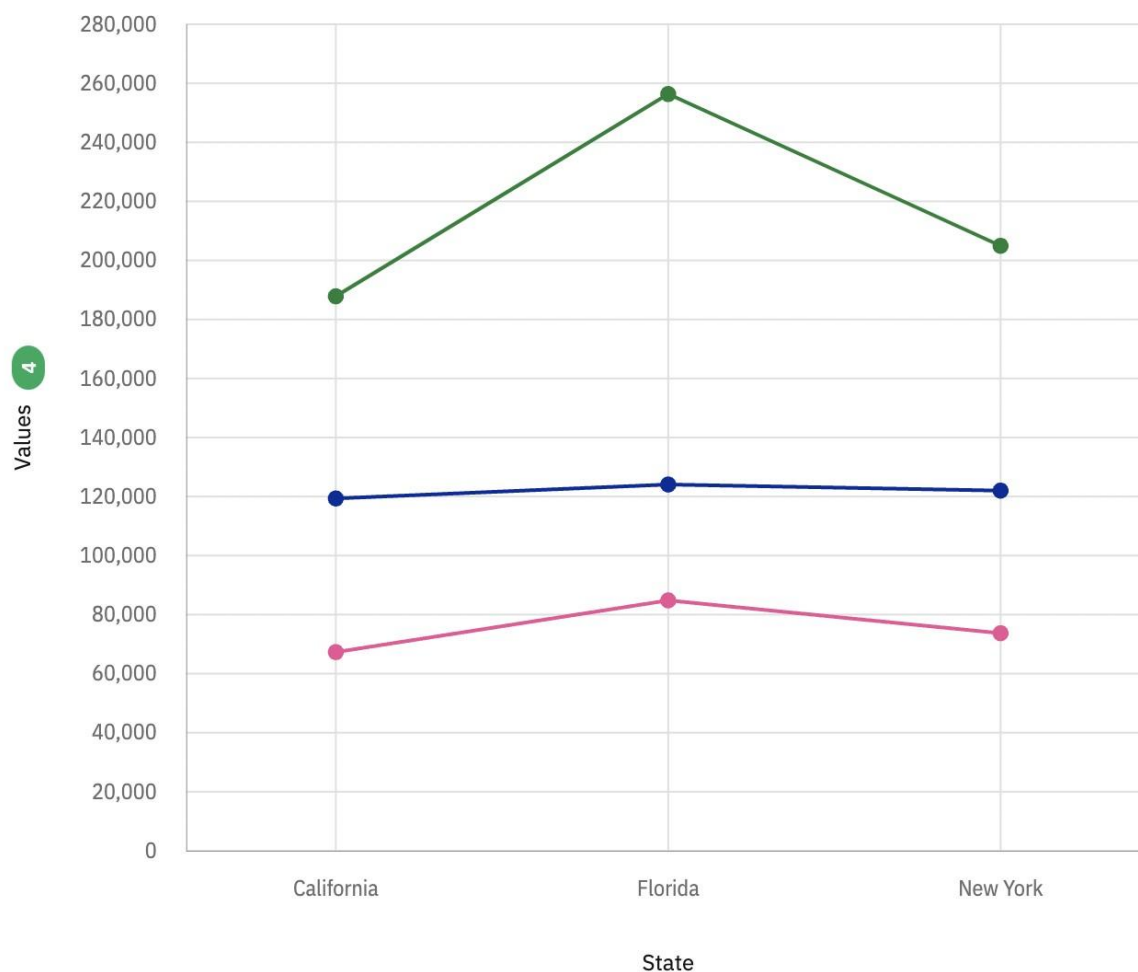
Profit

Required field

Visualization #6:**Marketing Spend, Administration and R&D Spend by State**

Measures

● Marketing Spend ● Administration ● R&D Spend



This visualization represents the average count of each expenditure for a startup - Marketing spend, Administration, and R&D Spend - mapped state-wise.

Fields

↶ x-axis*

Required field

:: State

⋮

Color

:: Measures group (3)

Click or drag data here

y-axis*

Required field

:: Marketing Spend

⋮

:: Administration

⋮

:: R&D Spend

⋮

Click or drag data here