

## **ASSIGNMENT – 1**

Assignment Date	10 September 2022
Student Name	Aiswarya A
Student Roll Number	95071915003
Maximum Marks	2 Marks

## **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:**

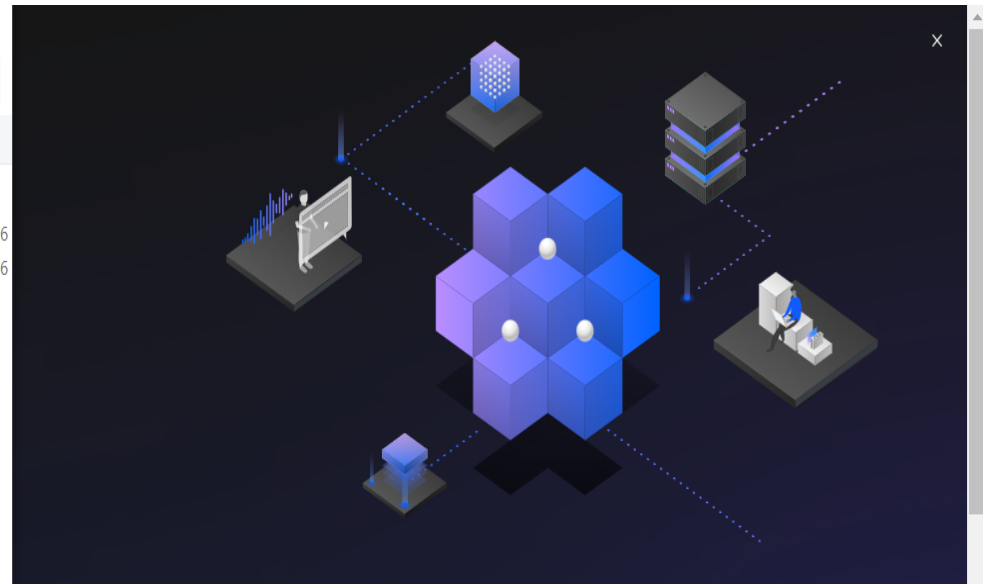
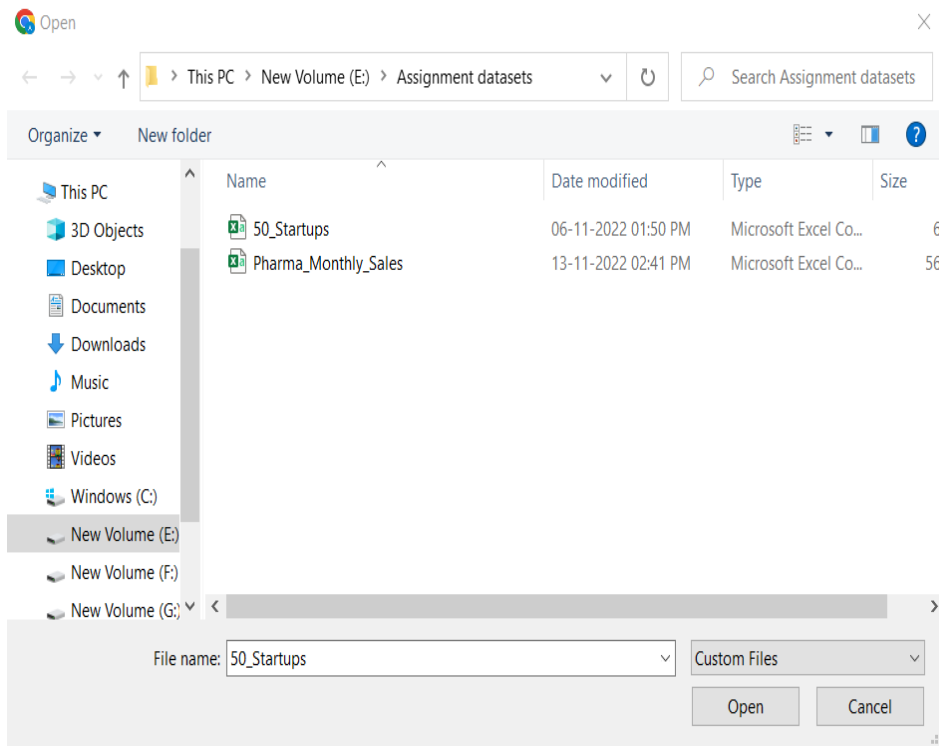
Upload the dataset to Cognos Analytics, explore and visualize the dataset

## Uploading the Dataset:

### Step 1 :

The screenshot displays the IBM Cognos Analytics with Watson user interface. At the top, a dark blue header bar contains the text "IBM Cognos Analytics with Watson" on the left, a notification badge with the number "354" in the center, and a search bar with the placeholder text "Search content" on the right. To the right of the search bar are icons for help, notifications, and user profile. Below the header, a light blue banner message states: "completed. Click More Info for details and to subscribe to future events", with "Dismiss" and "More info" links on the right. A left-hand navigation menu is visible, featuring options: "Home", "+ New", "Upload data" (highlighted with a blue bar), "Content", "Recent", and "Manage". The main content area features a large dark blue hero section with the title "Cognos Analytics with Watson." and a sub-header "Taking a look at our introduction video, product tour and Getting". A blue button labeled "Take a product tour" is positioned below the sub-header. To the right of the text is an isometric illustration of a data cube, server racks, a person at a computer, and a line graph. Below the hero section, three light gray cards are displayed in a row: "Prepare data" (with a data cube icon and description: "Use data modules to clean and connect data from multiple resources."), "Exploration" (with a magnifying glass icon and description: "Quickly find unbiased answers by identifying trends in your data with data exploration."), and "Present data" (with a dashboard icon and description: "Create sophisticated, multi-page, multi-query dashboards, reports, or stories.").

## Step 2 :



### Upload data

Upload or drag and drop spreadsheets, csv files, and other data sources.

### Prepare data

Use data modules to clean and connect data from multiple resources.

### Exploration

Quickly find unbiased answers by identifying trends in your data with data exploration.

### Present data

Create sophisticated, multi-page, multi-query dashboards, reports, or stories.

## Analyzing 50\_Startups.csv

Cancel Details

### Select sources

🔍

My content

🔍

📁

👤

🕒

📅

⬆️

🔍 Type any text to filter items in this folder

📁 Data Modules  
11/13/2022, 10:21 PM

📄 50\_Startups.csv  
11/15/2022, 2:00 PM

📄 IPL Ball-by-Ball 2008-2020.csv  
11/9/2022, 8:56 PM

📄 IPL dashboard  
11/13/2022, 11:37 AM

📄 IPL Matches 2008-2020.csv  
11/9/2022, 8:54 PM

📄 Olympic Events.xlsx  
11/3/2022, 12:06 PM

📄 Olympic Medals.xlsx  
11/3/2022, 12:07 PM

📄 Pharma exploration  
11/14/2022, 12:31 AM

📄 Pharma Sales dashboard  
11/15/2022, 12:25 PM

📄 Pharma\_Monthly\_Sales.csv

**Filter by**

[Reset filters](#)

Type

☐ Folders

☐ Packages

☐ Files

☐ Data sets

☐ Modules

Modified

☒ All

☐ Today

☐ Yesterday

☐ Past week

☐ Past month

Cancel

OK

IBM Cognos Analytics with Watson

New data module

35

Search content

?

🔔

👤

📄

🔗

↶

↷

📄

📄

Properties

Data module

+

🔍

🔍 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
	13	93863.75	127320.38	249839.44	Florida	141585.52
	14	91992.39	135495.07	252664.93	California	134307.35

## Add a data source to explore



My content

Team content



Name	Type	Last Accessed
Data Modules	Folder	11/13/2022, 10:21 PM
50_Startups.csv <span>CSV</span>	Uploaded file	11/15/2022, 2:00 PM
IPL Ball-by-Ball 2008-2020.csv <span>CSV</span>	Uploaded file	11/9/2022, 8:56 PM
IPL Matches 2008-2020.csv <span>CSV</span>	Uploaded file	11/9/2022, 8:54 PM
Olympic Events.xlsx <span>xlsx</span>	Uploaded file	11/3/2022, 12:06 PM
Olympic Medals.xlsx <span>xlsx</span>	Uploaded file	11/3/2022, 12:07 PM
Pharma_Monthly_Sales.csv <span>CSV</span>	Uploaded file	11/13/2022, 6:22 PM
sales.csv <span>CSV</span>	Uploaded file	11/2/2022, 12:26 PM
US Superstore data.xls <span>xls</span>	Uploaded file	11/6/2022, 10:13 AM

Cancel

Add



## Data Exploration :

Create +

Cards

Data relationships

Profit

Marketing Spend

Administration

R&D Spend

### Select a starting point

Every exploration includes a data relationships card.

Start with any column. You can always change it later.

Enter data column. Not sure? Try Profit, Marketing Spend

Try starting with one of these...

## Explore data relationships

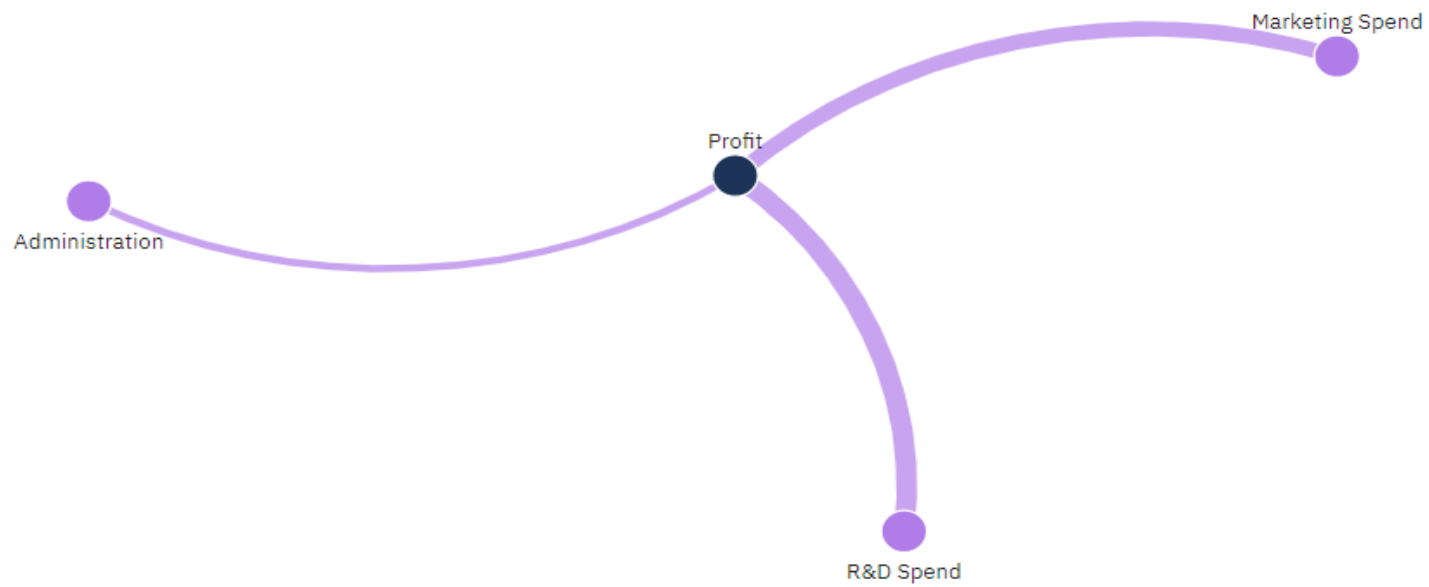
50\_Startups.csv

[Reset to original](#)

🔍 Profit

×

Edit diagram ▾

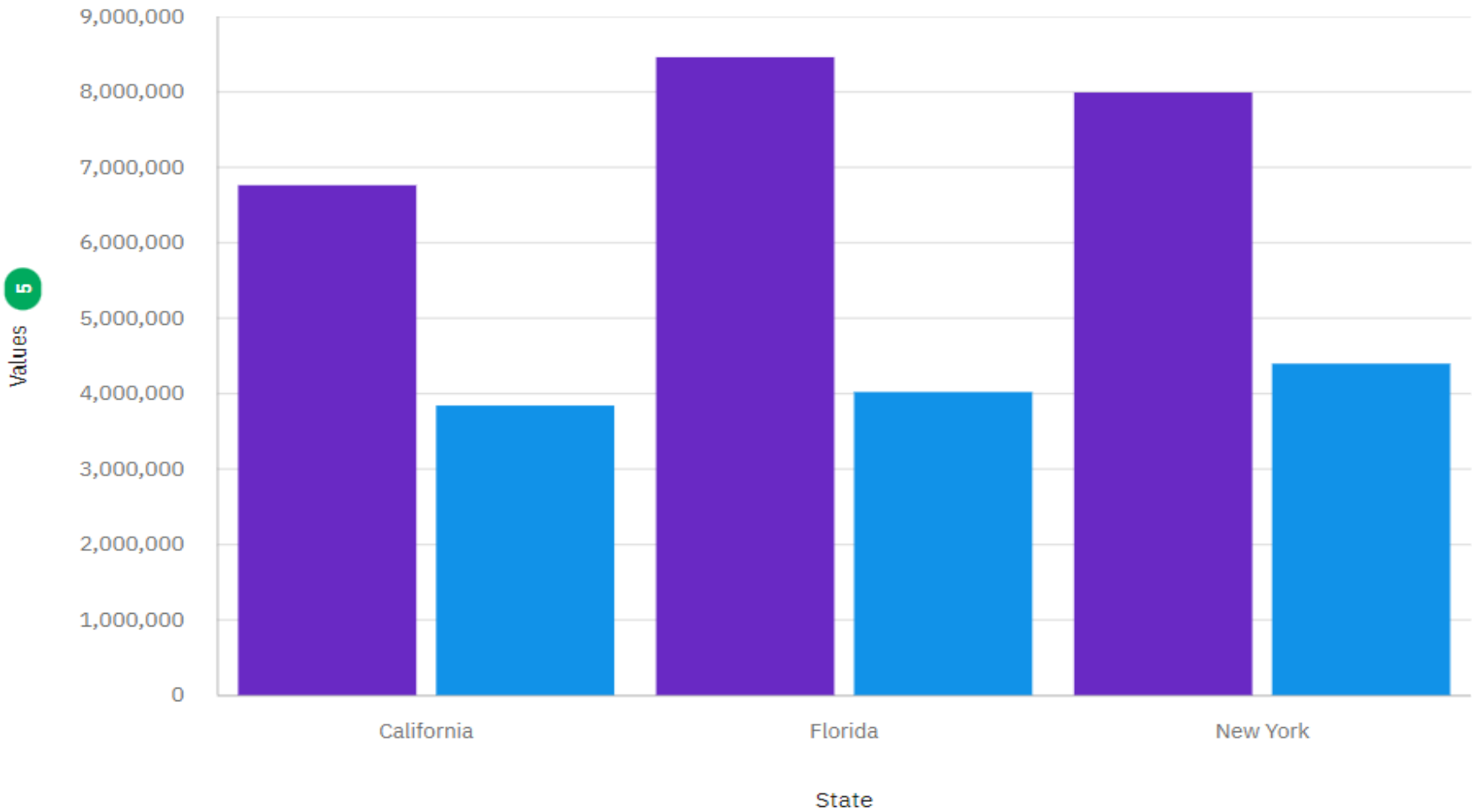


Relationship diagram ⓘ

10% —●—————●———— 100%

Marketing Spend and Profit by State

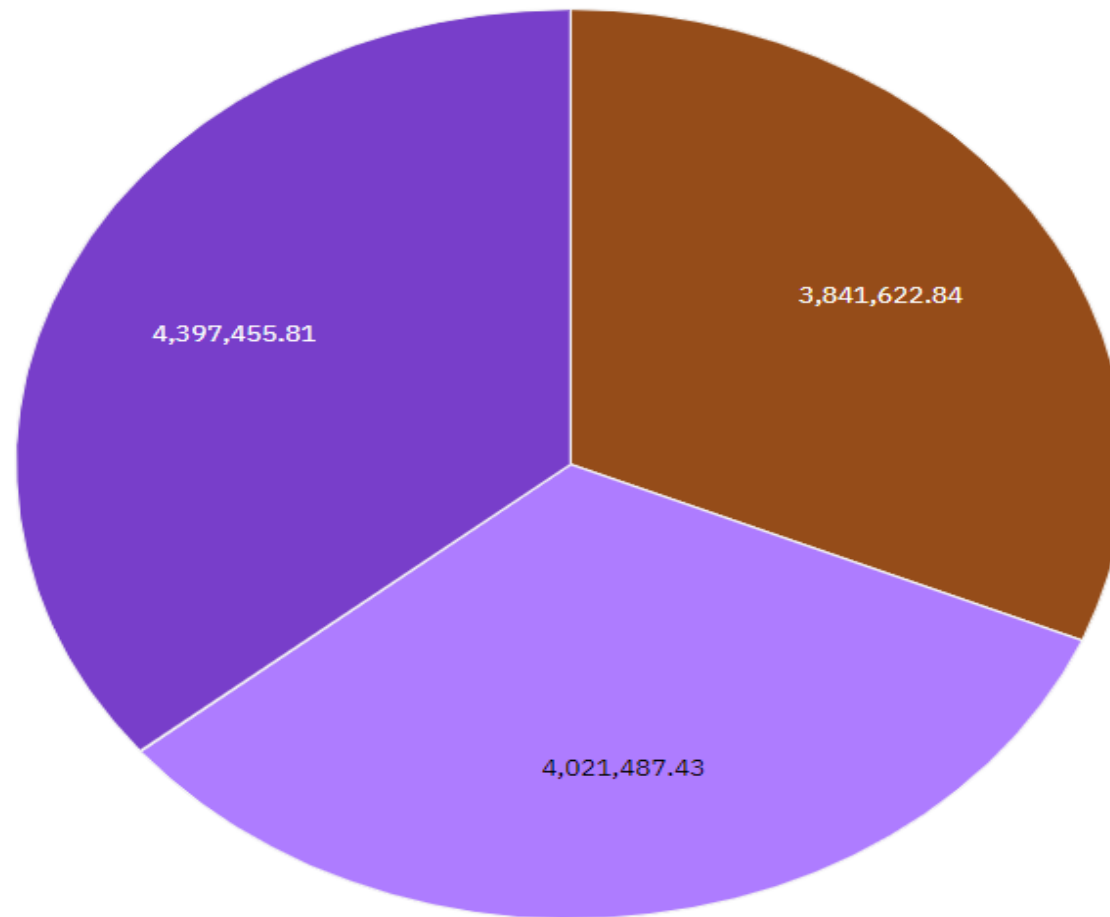
Measures  
Marketing Spend Profit



## Profit by State 4

State

California Florida New York

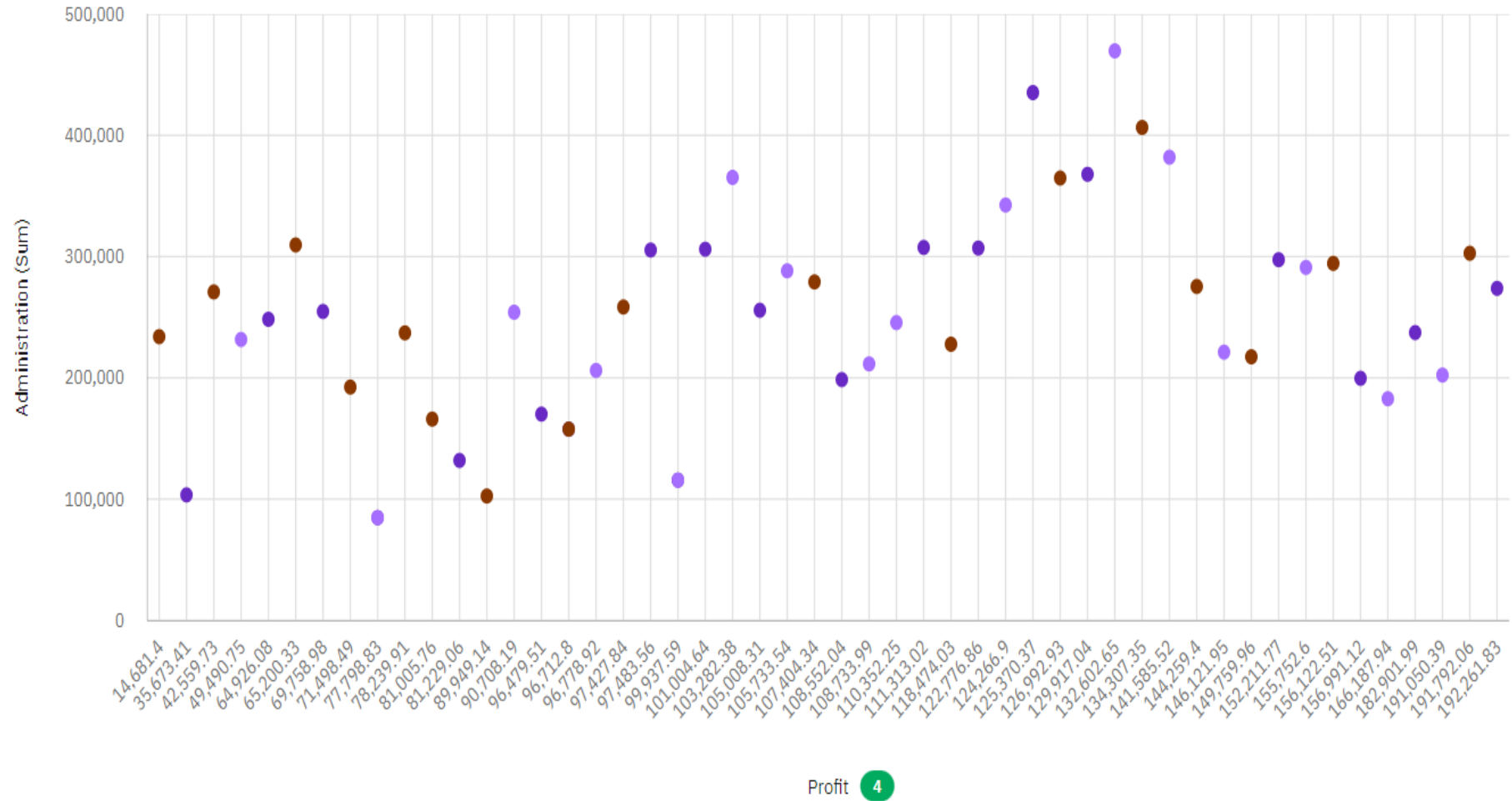


## Administration by Profit colored by State



State

California Florida New York

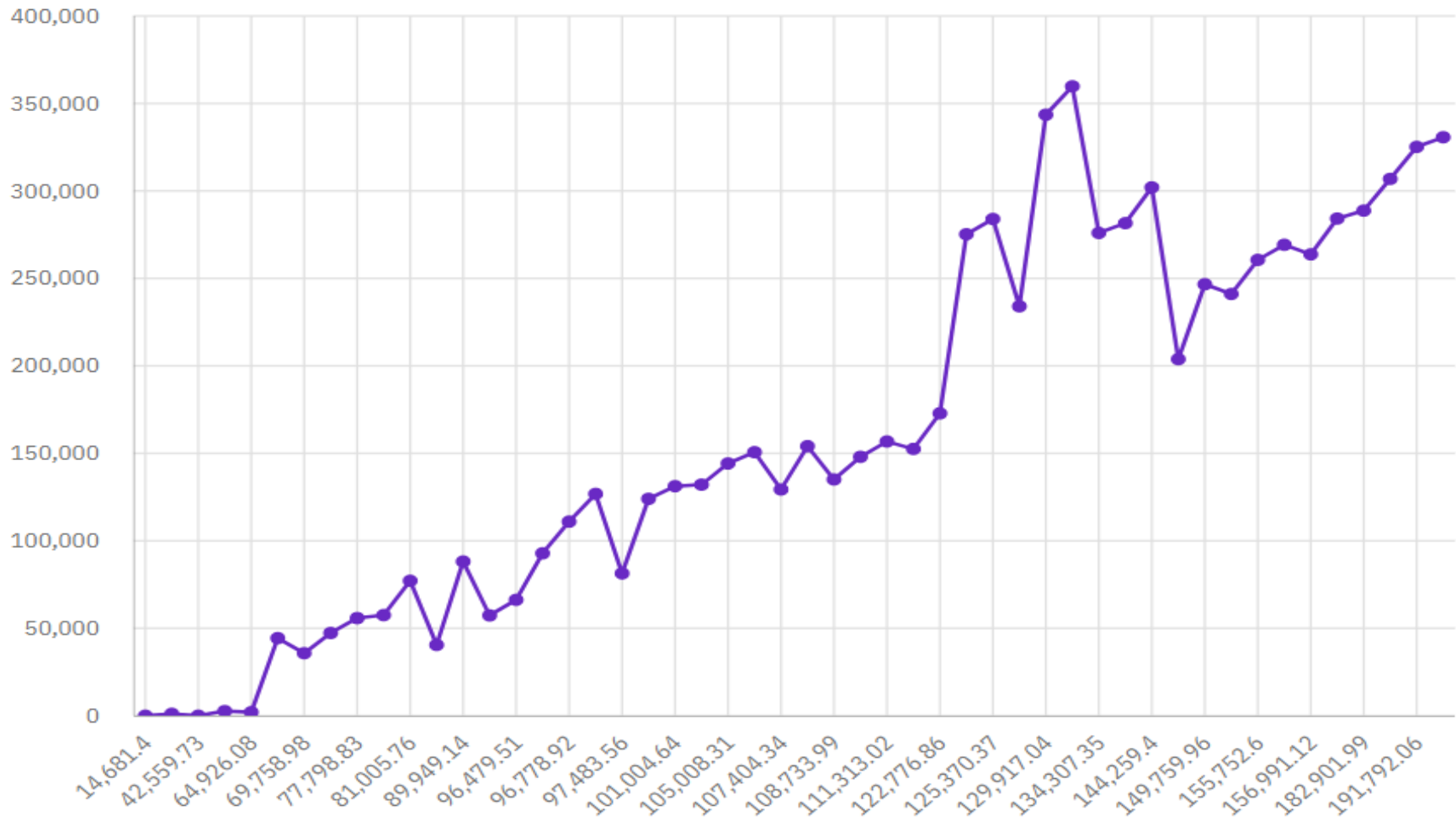


## R&D Spend by Profit



3

R&D Spend (Sum)



Profit

4

State, Marketing Spend, Profit, R&D Spend and Administration 8

State	R&D Spend	Administration	Marketing Spend	Profit
California	2,422,612.31	4,296,570.55	6,762,349.47	3,841,622.84
Florida	2,798,837.27	4,094,296.26	8,460,023.62	4,021,487.43
New York	2,874,159	4,758,218.39	7,992,514.7	4,397,455.81
Summary	8,095,608.58	13,149,085.2	23,214,887.79	12,260,566.08

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

Measures

R&D Spend Administration Marketing Spend

