

NAME: VINOODHINI D
REG NO: 917719IT118

NAALAIYA THIRAN -ASSIGNMENT 1

QUESTION:

<https://drive.google.com/file/d/1e-YAVbSeAfITj0AS5kBufwnxITlLwrEr/view>

CSV FILE:



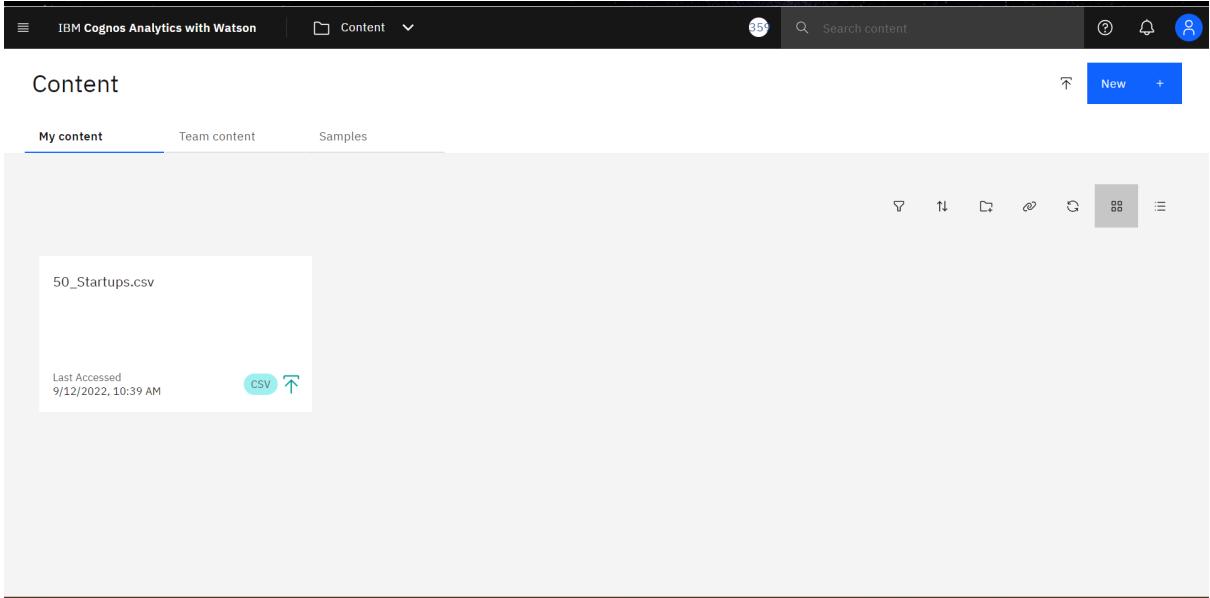
UPLOADING DATA:

A screenshot of the IBM Cognos Analytics with Watson interface. At the top, a progress bar shows 'Initializing file upload...'. A central modal window titled 'Update data' asks if the file '50_Startups.csv' already exists, with options to 'Replace data' (selected) or 'Append data'. There is also a checkbox for 'Apply to all files'. Below the modal, a list shows '50_Startups.csv' under the 'Recent' tab. To the right, there are sections for 'Present data' and other features like 'Upload data', 'Prepare data', and 'Exploration'.

A screenshot of the IBM Cognos Analytics with Watson interface. At the top, a progress bar shows '3 items open'. A central modal window titled 'Update data' asks if the file '50_Startups.csv' already exists, with options to 'Replace data' (selected) or 'Append data'. There is also a checkbox for 'Apply to all files'. Below the modal, a list shows '50_Startups.csv' under the 'Recent' tab. To the right, there are sections for 'Upload data', 'Prepare data', 'Exploration', and 'Present data'.

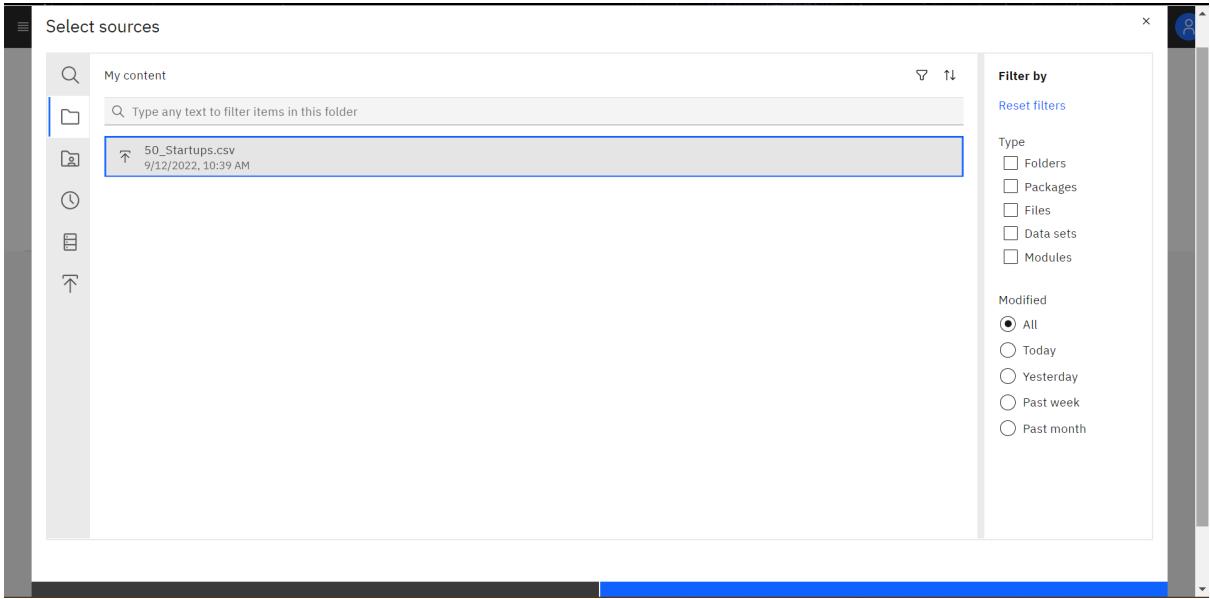
NAME: VINOODHINI D
REG NO: 917719IT118

CONTENT UPLOADED IS IN MY CONTENT:



The screenshot shows the 'Content' section of the IBM Cognos Analytics with Watson interface. At the top, there's a navigation bar with 'IBM Cognos Analytics with Watson', a 'Content' dropdown, a user icon with '359', a search bar, and notification icons. Below the navigation is a toolbar with 'Content', 'New', and other icons. The main area is titled 'Content' and has tabs for 'My content' (which is selected), 'Team content', and 'Samples'. A single item, '50_Startups.csv', is listed. It shows the file name, a 'Last Accessed' timestamp (9/12/2022, 10:39 AM), and a 'CSV' download button with a downward arrow icon.

PREPARING DATASET:
go to menu -> data module



The screenshot shows the 'Select sources' dialog box. On the left, there's a sidebar with icons for search, folder, file, and refresh. The main area is titled 'My content' and contains a search bar and a list of files. One file, '50_Startups.csv', is highlighted with a blue border. To the right of the list is a 'Filter by' panel with sections for 'Type' (Folders, Packages, Files, Data sets, Modules) and 'Modified' (All, Today, Yesterday, Past week, Past month). The 'All' option is selected under 'Modified'.

NAME: VINOODHINI D
REG NO: 917719IT118

The screenshot shows the IBM Cognos Analytics with Watson interface. On the left, there's a sidebar titled 'Data module' with a search bar. Below it, under 'Navigation paths', is a folder named 'New data module' which contains a file named '50_Startups.csv'. This file has several columns: '# Row Id', 'R&D Spend', 'Administration', 'Marketing Spend', 'State', and 'Profit'. To the right of the sidebar, there's a large preview icon with a checkmark, and below it, a section titled 'Preview data' with the instruction: 'To preview data, select a table, a column in a table, or a folder that contains columns.'

- Unique values (Identifiers)

Abc- Text values

Scale – Measures (numerical Value)

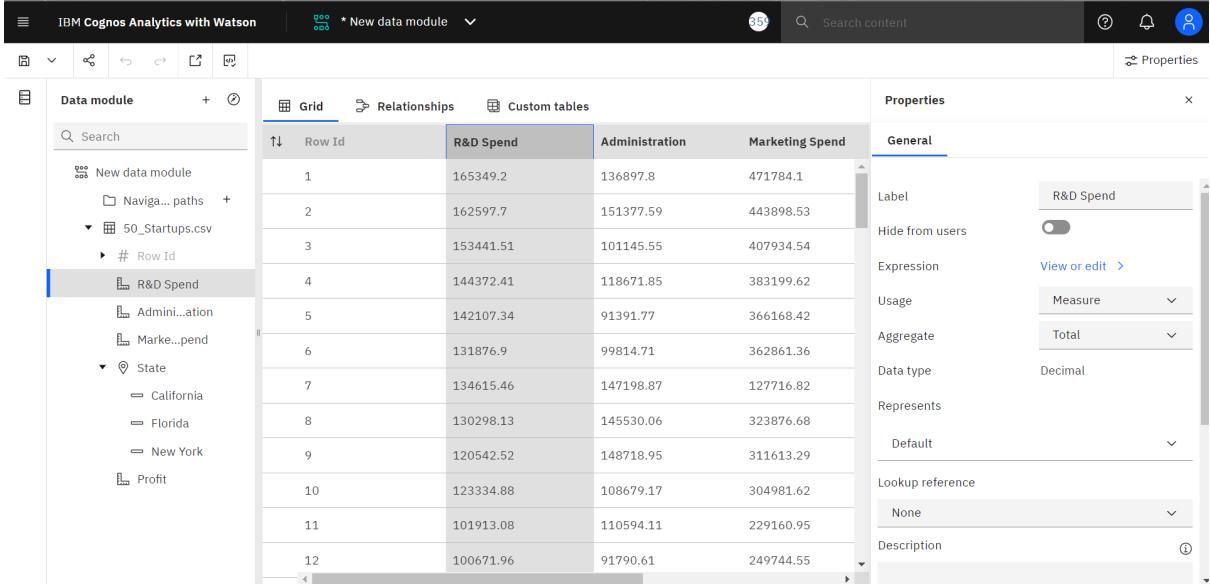
Map mark- Map value

PREPARING DATA:

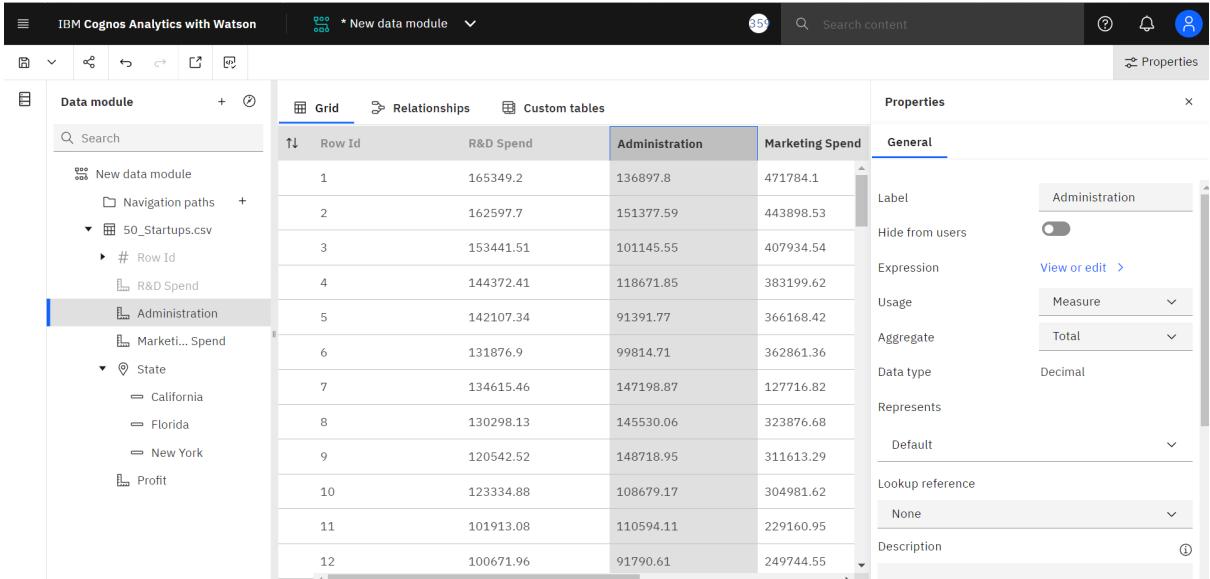
This screenshot shows the 'Grid' view of the '50_Startups.csv' data in IBM Cognos Analytics. The grid displays 12 rows of data with columns: Row Id, R&D Spend, Administration, and Marketing Spend. The 'Properties' panel on the right is open for the 'Row Id' column, showing settings for General, Navigation paths, and Data type (Identifier, Count, Integer). The 'General' tab shows 'Label' as 'Row Id' and 'Usage' as 'Identifier'. The 'Navigation paths' tab shows 'Hide from users' as 'On'. The 'Data type' is set to 'Integer'.

Row Id	R&D Spend	Administration	Marketing Spend
1	165349.2	136897.8	471784.1
2	162597.7	151377.59	443898.53
3	153441.51	101145.55	407934.54
4	144372.41	118671.85	383199.62
5	142107.34	91391.77	366168.42
6	131876.9	99814.71	362861.36
7	134615.46	147198.87	127716.82
8	130298.13	145530.06	323876.68
9	120542.52	148718.95	311613.29
10	123334.88	108679.17	304981.62
11	101913.08	110594.11	229160.95
12	100671.96	91790.61	249744.55

NAME: VINOODHINI D
REG NO: 917719IT118



The screenshot shows the IBM Cognos Analytics with Watson interface. On the left, there's a navigation pane with a search bar and links for 'New data module', 'Navigation paths', and a file named '50_Startups.csv'. The main area displays a grid with four columns: Row Id, R&D Spend, Administration, and Marketing Spend. The 'R&D Spend' column is currently selected. The grid contains 12 rows of data. To the right of the grid is a 'Properties' panel for the selected column, specifically for the 'R&D Spend' measure. The 'General' tab is selected, showing details like Label (R&D Spend), Hide from users (disabled), Expression (View or edit), Usage (Measure), Aggregate (Total), Data type (Decimal), and Represents (Default). Other tabs in the properties panel include 'Relationships' and 'Custom tables'.



This screenshot is nearly identical to the one above, but the 'Administration' column is now selected in the grid. The data in the grid remains the same, with 12 rows of data across the four columns. The 'Properties' panel on the right is also identical, showing the 'General' tab for the 'Administration' measure with its properties: Label (Administration), Hide from users (disabled), Expression (View or edit), Usage (Measure), Aggregate (Total), Data type (Decimal), and Represents (Default).

NAME: VINOODHINI D
REG NO: 917719IT118

IBM Cognos Analytics with Watson * New data module

Grid Relationships Custom tables

Row Id	R&D Spend	Administration	Marketing Spend
1	165349.2	136897.8	471784.1
2	162597.7	151377.59	443898.53
3	153441.51	101145.55	407934.54
4	144372.41	118671.85	383199.62
5	142107.34	91391.77	366168.42
6	131876.9	99814.71	362861.36
7	134615.46	147198.87	127716.82
8	130298.13	145530.06	323876.68
9	120542.52	148718.95	311613.29
10	123334.88	108679.17	304981.62
11	101913.08	110594.11	229160.95
12	100671.96	91790.61	249744.55

Properties

General

Label: Marketing Spend
 Hide from users
[View or edit >](#)
 Usage: Measure
 Aggregate: Total
 Data type: Decimal
 Represents:
 Default:
 Lookup reference: None
 Description:

IBM Cognos Analytics with Watson * New data module

Grid Relationships Custom tables

Administration	Marketing Spend	State	Profit
97.8	471784.1	New York	192261.83
77.59	443898.53	California	191792.06
45.55	407934.54	Florida	191050.39
71.85	383199.62	New York	182901.99
1.77	366168.42	Florida	166187.94
4.71	362861.36	New York	156991.12
98.87	127716.82	California	156122.51
30.06	323876.68	Florida	155752.6
18.95	311613.29	New York	152211.77
79.17	304981.62	California	149759.96
94.11	229160.95	Florida	146121.95
0.61	249744.55	California	144259.4

Properties

General Navigation paths

Label: State
 Hide from users
[View or edit >](#)
 Usage: Identifier
 Aggregate: Count Distinct
 Data type: Text
 Represents:
 Geographic location:
 City:
 Description:

IBM Cognos Analytics with Watson * New data module

Grid Relationships Custom tables

Administration	Marketing Spend	State	Profit
97.8	471784.1	New York	192261.83
77.59	443898.53	California	191792.06
45.55	407934.54	Florida	191050.39
71.85	383199.62	New York	182901.99
1.77	366168.42	Florida	166187.94
4.71	362861.36	New York	156991.12
98.87	127716.82	California	156122.51
30.06	323876.68	Florida	155752.6
18.95	311613.29	New York	152211.77
79.17	304981.62	California	149759.96
94.11	229160.95	Florida	146121.95
0.61	249744.55	California	144259.4

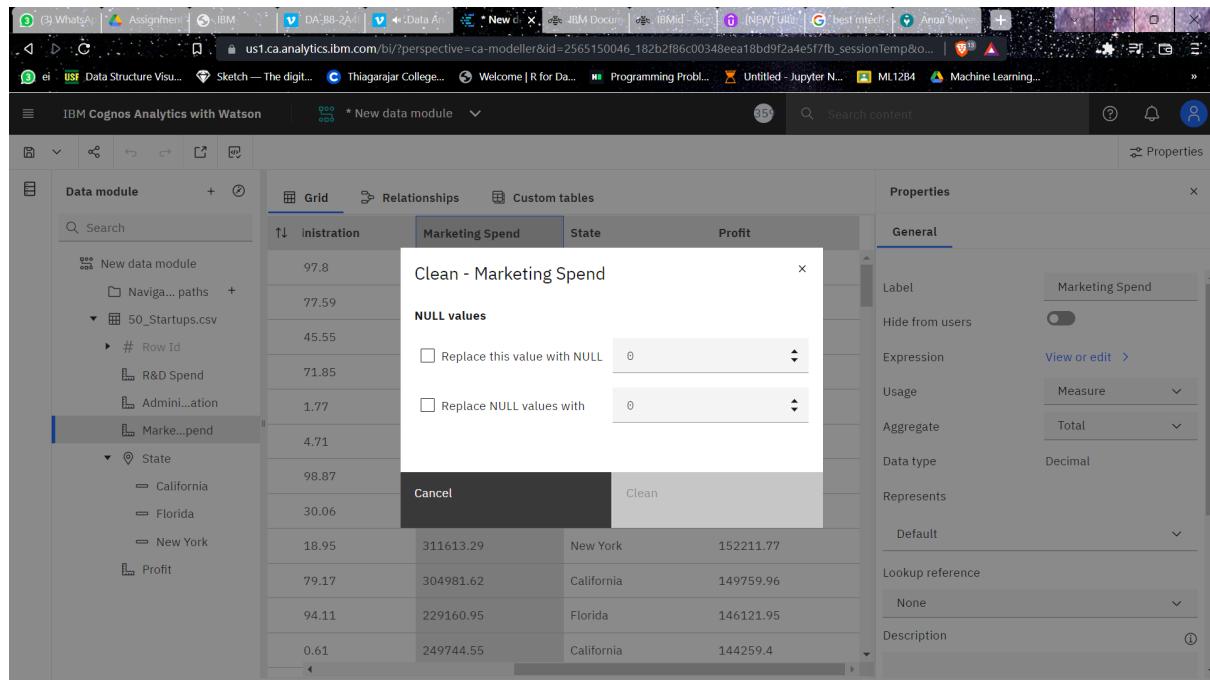
Properties

General

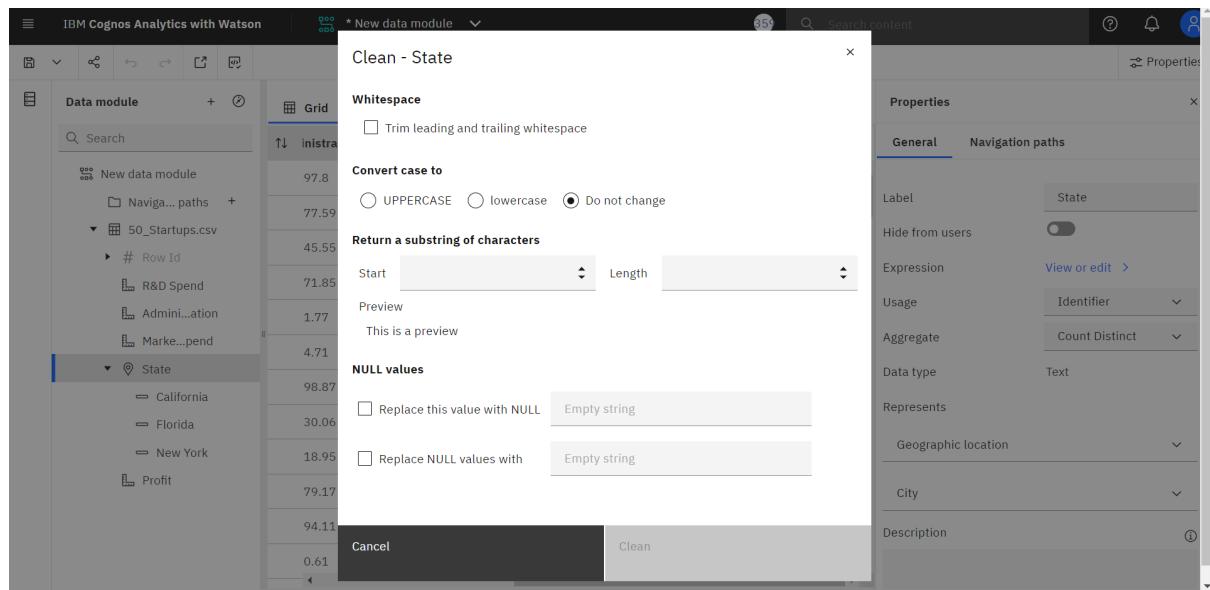
Label: Profit
 Hide from users
[View or edit >](#)
 Usage: Measure
 Aggregate: Total
 Data type: Decimal
 Represents:
 Default:
 Lookup reference: None
 Description:

NAME: VINOODHINI D
REG NO: 917719IT118

CLEANING THE DATASET:



The screenshot shows the IBM Cognos Analytics interface with a data module named '50_Startups.csv'. A context menu is open over the 'Marketing Spend' column, displaying options like 'Grid', 'Relationships', and 'Custom tables'. A modal dialog titled 'Clean - Marketing Spend' is centered, specifically targeting the 'NULL values' section. It contains two input fields: 'Replace this value with NULL' set to '0' and 'Replace NULL values with' also set to '0'. The dialog has 'Cancel' and 'Clean' buttons at the bottom.



The screenshot shows the same IBM Cognos Analytics interface with the '50_Startups.csv' data module. A context menu is open over the 'State' column, showing options like 'Grid', 'Relationships', and 'Custom tables'. A modal dialog titled 'Clean - State' is open, focusing on the 'NULL values' section. It includes two input fields: 'Replace this value with NULL' set to 'Empty string' and 'Replace NULL values with' also set to 'Empty string'. The dialog includes sections for 'Whitespace', 'Convert case to', and 'Return a substring of characters'. The dialog has 'Cancel' and 'Clean' buttons at the bottom.

NAME: VINOODHINI D
REG NO: 917719IT118

The screenshot shows the 'Clean - State' dialog box in IBM Cognos Analytics. The 'NULL values' section is active, with the checkbox 'Replace this value with NULL' checked and the value 'Florida' entered into the text field. The 'Clean' button at the bottom is highlighted in blue.

Thus Florida values are returned with null:

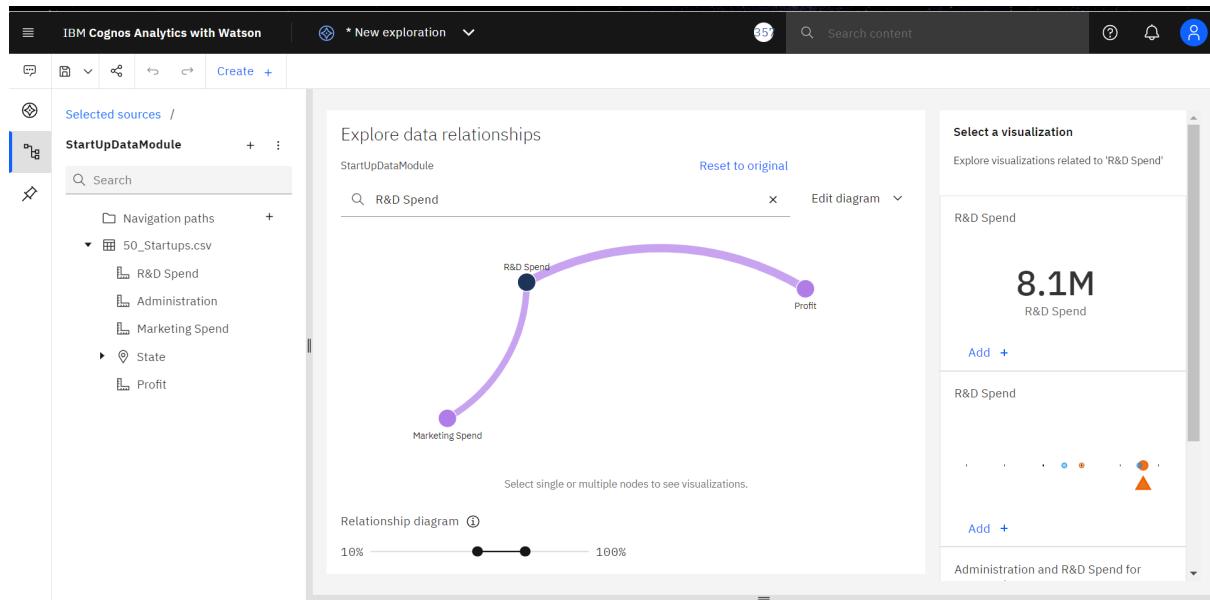
Row Id	R&D Spend	Administration	Marketing Spend	State	Profit
1	165349.2	136897.8	471784.1	New York	192261.83
2	162597.7	151377.59	443898.53	California	191792.06
3	153441.51	101145.55	407934.54	Null	191050.39
4	144372.41	118671.85	383199.62	New York	182901.99
5	142107.34	91391.77	366168.42	Null	166187.94
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	Null	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	Null	146121.95
12	100671.96	91790.61	249744.55	California	144259.4

DATA VISUALISATION:

1. R&D SPEND:

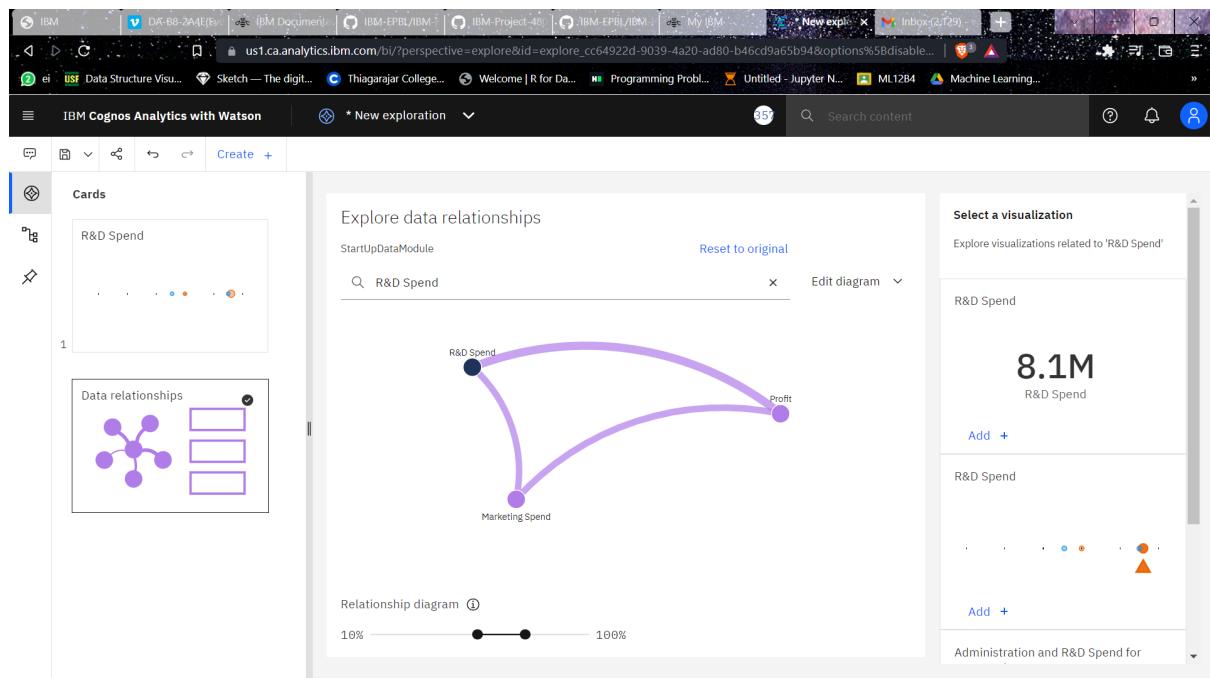
PRIMARY RELATIONSHIP:

NAME: VINOODHINI D
REG NO: 917719IT118



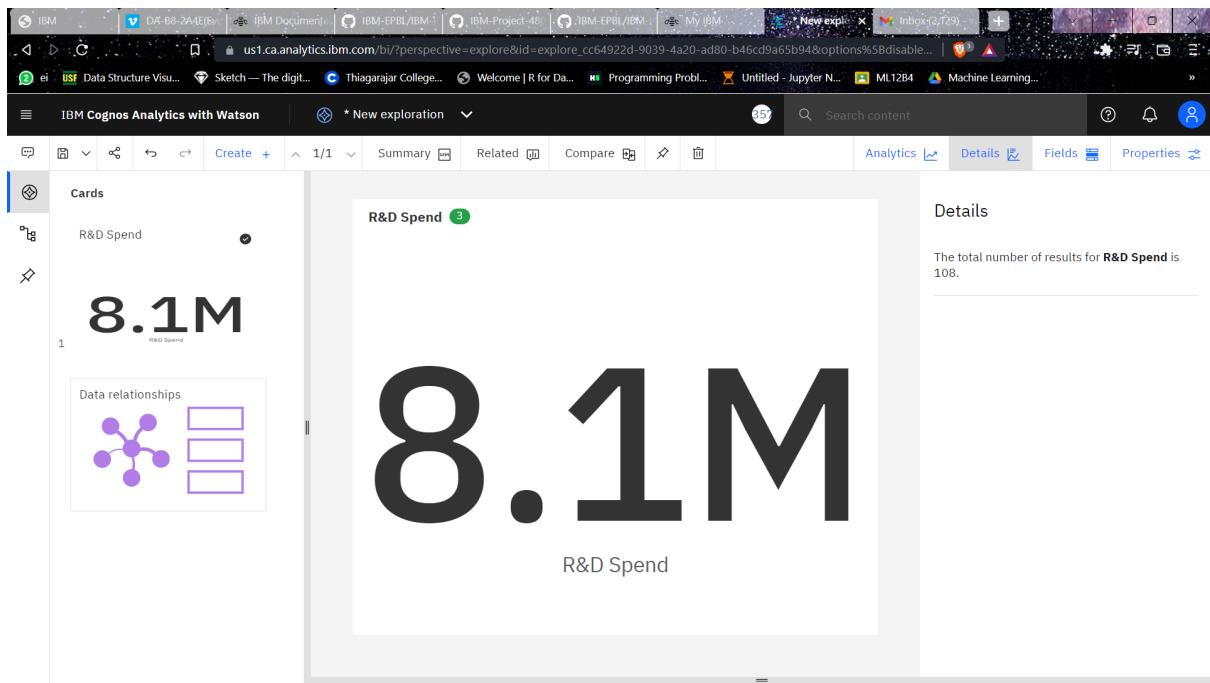
The screenshot shows the IBM Cognos Analytics with Watson interface. On the left, there's a sidebar with a tree view of 'Selected sources / StartUpDataModule'. The tree includes '50_Startups.csv' with nodes like 'R&D Spend', 'Administration', 'Marketing Spend', 'State', and 'Profit'. The main area is titled 'Explore data relationships' and shows a relationship diagram for 'R&D Spend'. Nodes include 'R&D Spend', 'Marketing Spend', and 'Profit'. Curved purple arrows connect 'R&D Spend' to 'Profit' and 'Marketing Spend' to 'Profit'. A slider at the bottom indicates a relationship strength of 10% to 100%. To the right, a panel titled 'Select a visualization' shows a summary card for 'R&D Spend' with the value '8.1M'.

SECONDARY RELATIONSHIP:



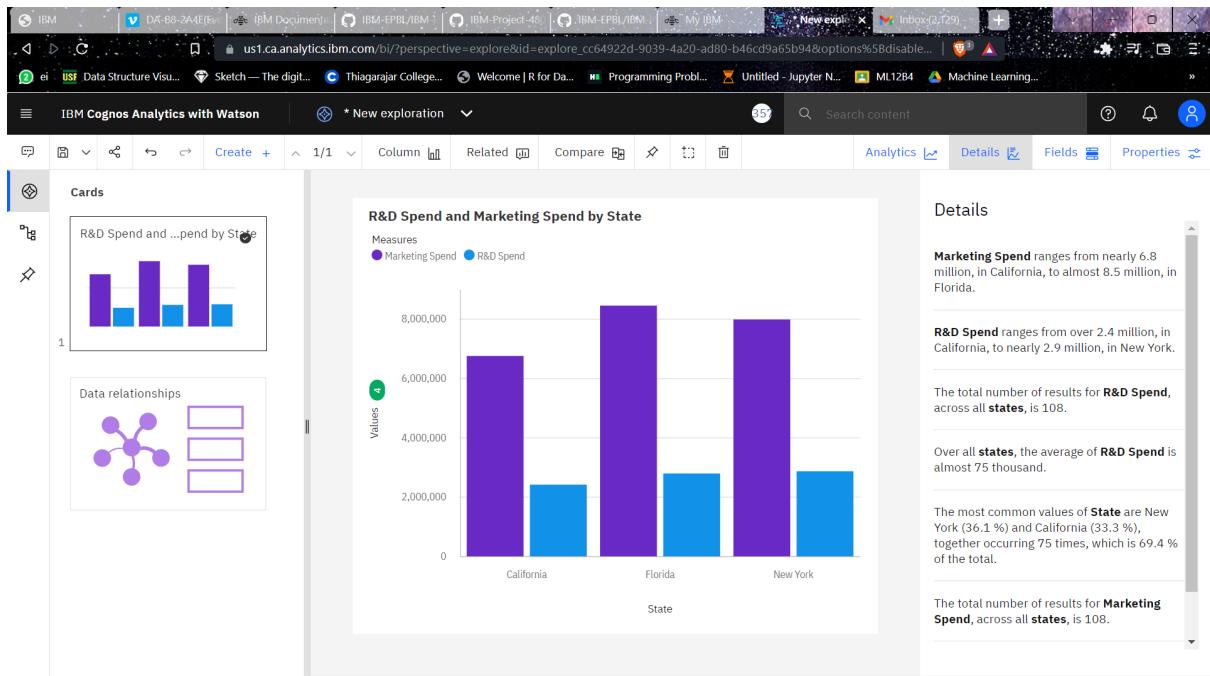
This screenshot shows the same interface as above, but with a different perspective. The sidebar now displays a 'Cards' section under 'R&D Spend', which includes a 'Data relationships' card showing a network of nodes and connections. The main 'Explore data relationships' panel still shows the same three nodes and their connections. The 'Select a visualization' panel remains the same, displaying the '8.1M' summary for 'R&D Spend'.

NAME: VINOODHINI D
REG NO: 917719IT118



Total amount R&D Spend is 108 and the value is 8.1 Million

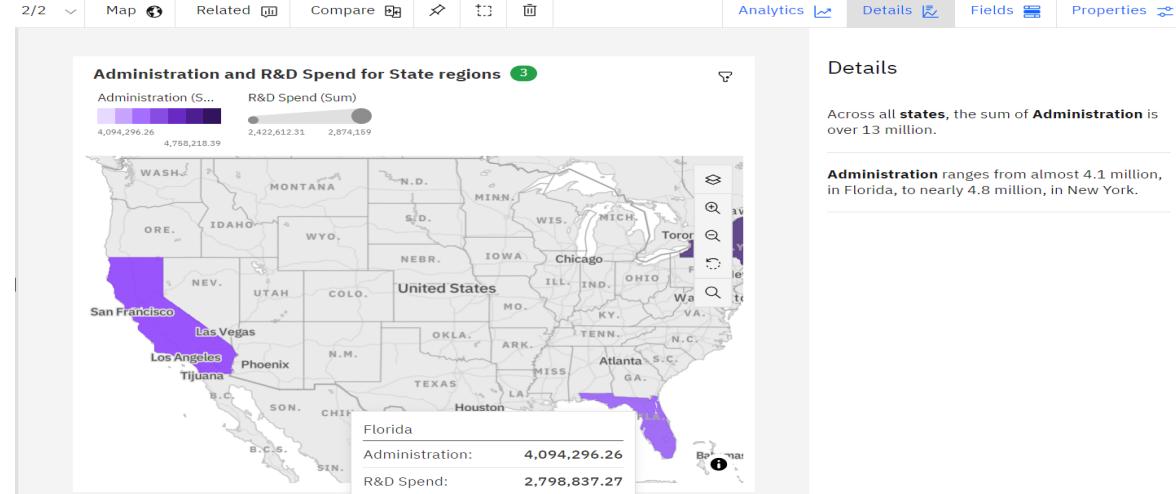
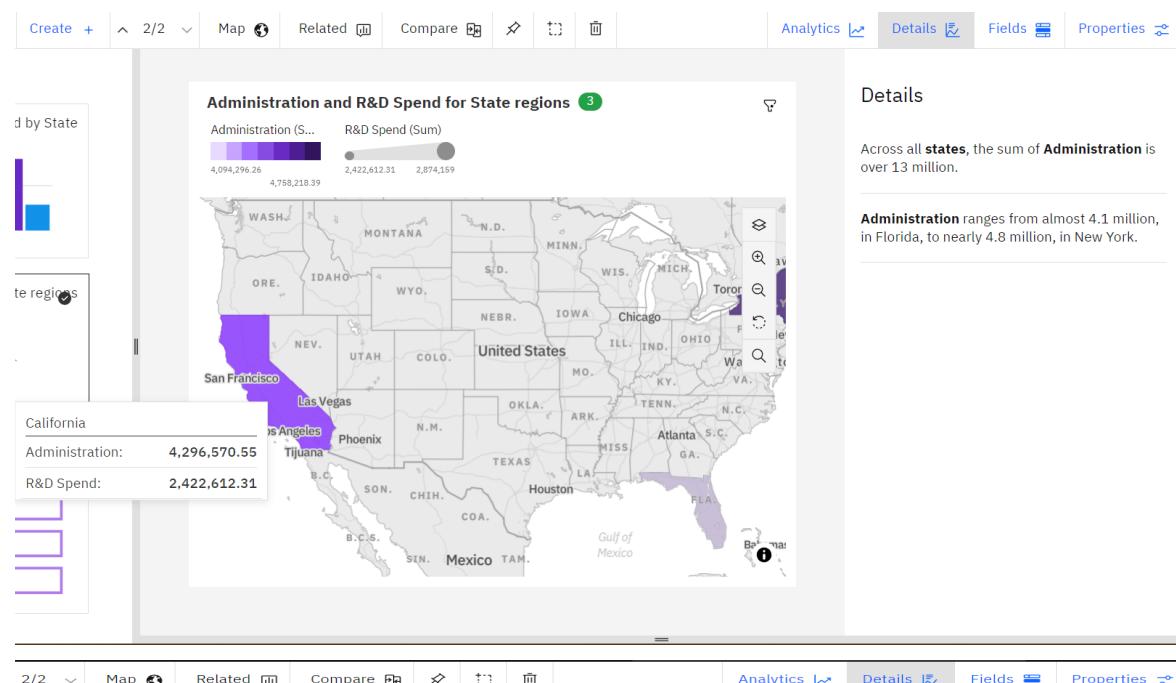
BAR CHART:



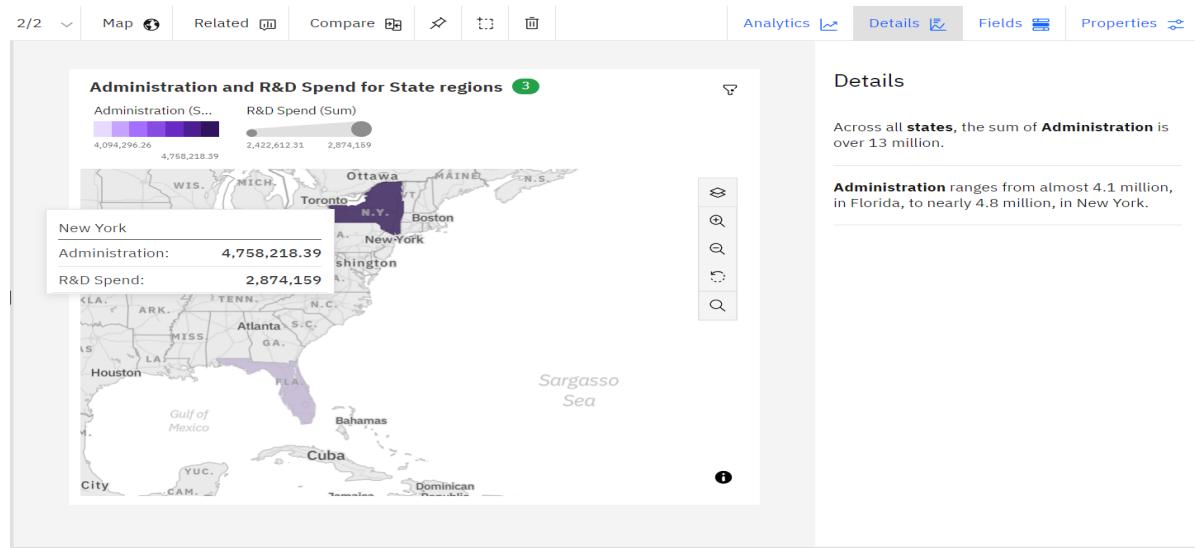
ADMINISTRATION AND R&D SPEND FOR STATE REGIONS:

Across all states, the sum of Administration is over 13 million.

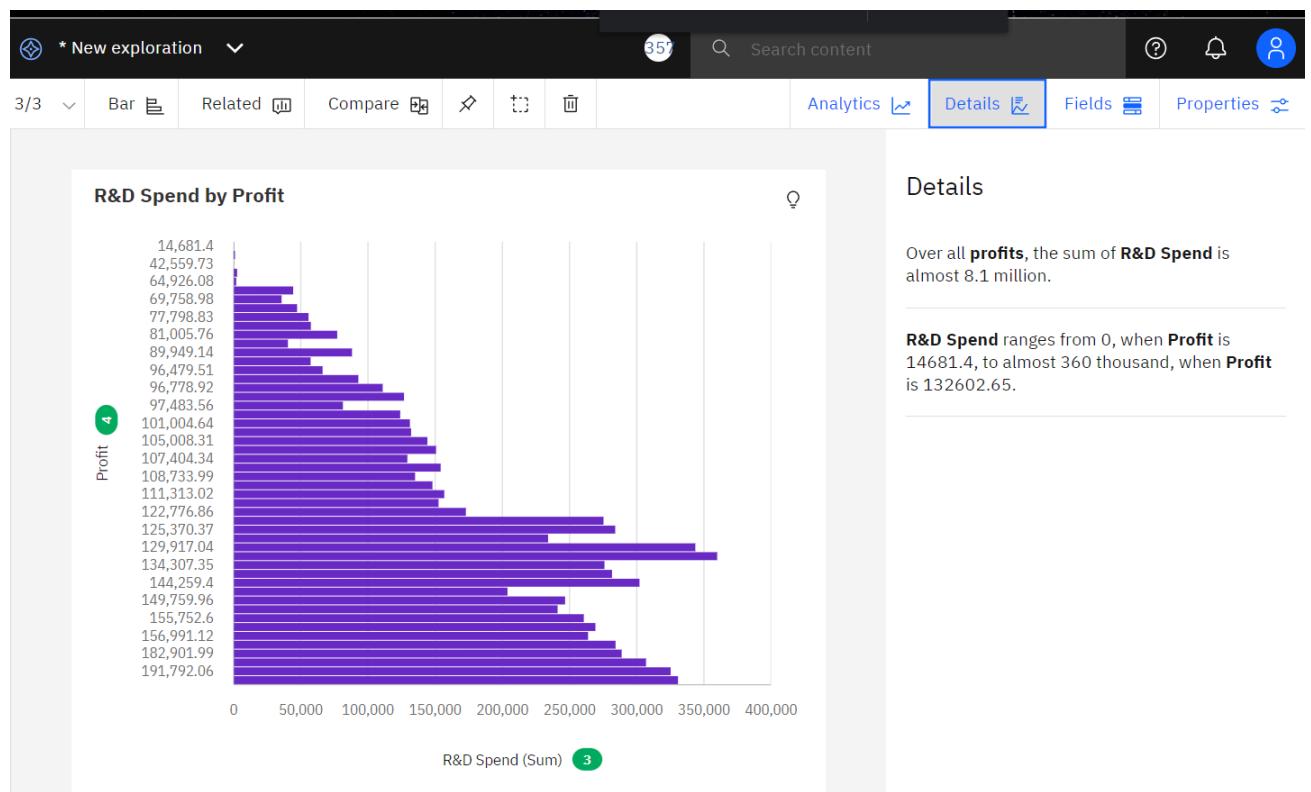
Administration ranges from almost 4.1 million, in Florida, to nearly 4.8 million, in New York.



NAME: VINOODHINI D
REG NO: 917719IT118



BAR GRAPH FOR PROFIT BASED ON R&D SPEND:



R&D SPENT BASED ON STATES:

NAME: VINOODHINI D
REG NO: 917719IT118



Details

Over all **states**, the sum of **R&D Spend** is nearly 8.1 million.

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.

R&D Spend ranges from over 2.4 million, in California, to nearly 2.9 million, in New York.

PROFIT FOR EACH STATES:



Details

Across all **states**, the sum of **Profit** is over 12 million.

Profit ranges from over 3.8 million, in California, to almost 4.4 million, in New York.

Correlated with R&D Spend and Marketing Spend (90%)

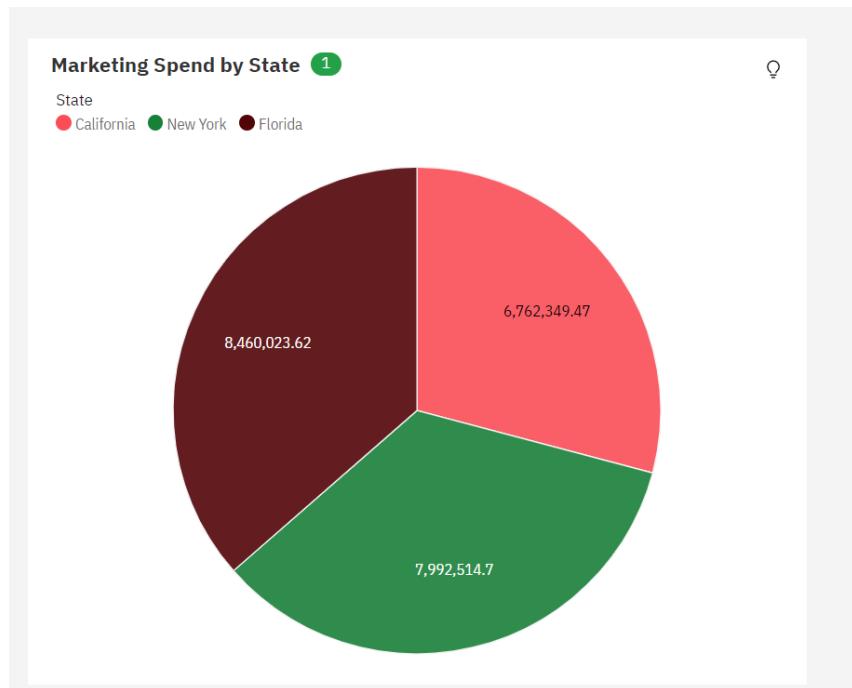
Correlated with R&D Spend and Administration (87%)

**NAME: VINOODHINI D
REG NO: 917719IT118**

Correlated with R&D Spend and State (86%)

Correlated with R&D Spend (84%)

PIE CHART DEPICTING THE MARKETING SPEND FOR EACH STATE:



Details

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 nearly 8.5 million, in Florida.

WORD CLOUD FOR ADMINISTRATION BASED ON THE STATES:



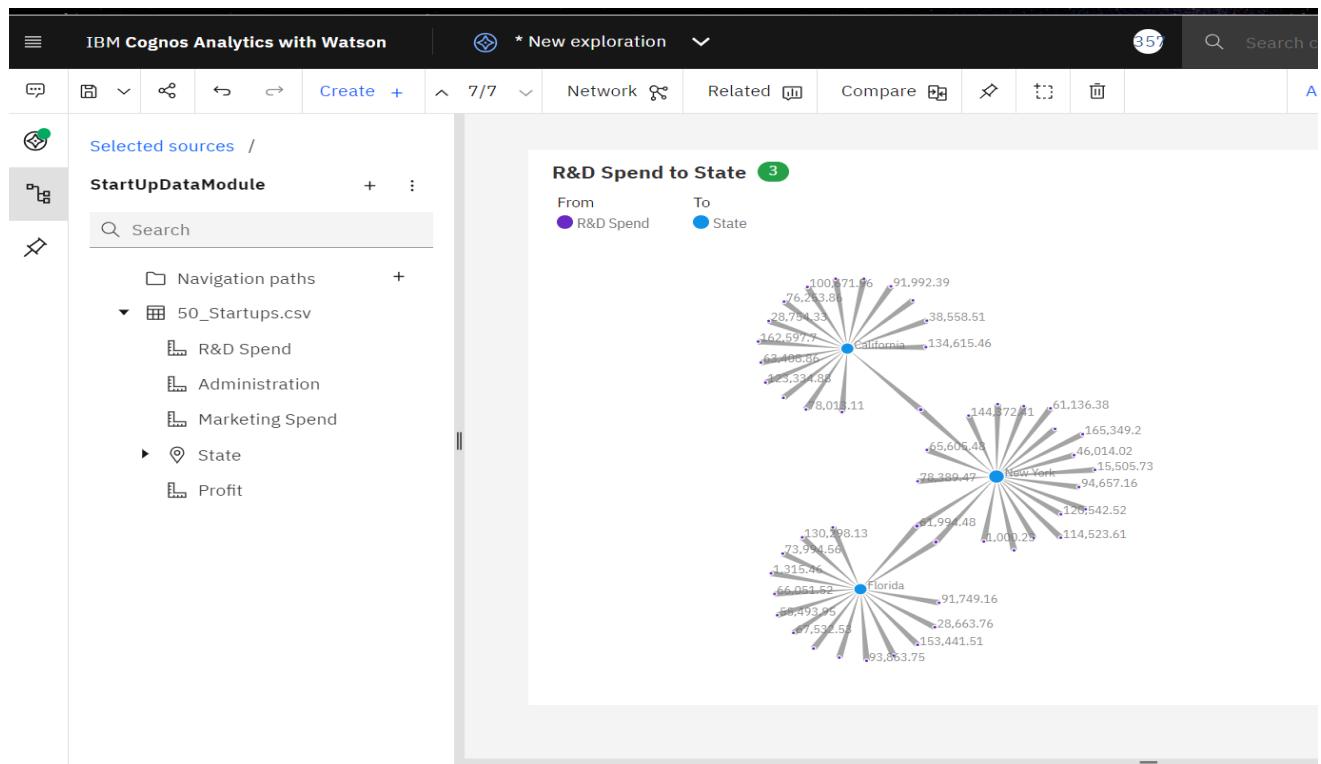
Details

The total number of results for **State**, across all **administrations**, is 108.

The most common values of **Administration** are 127320.38 (2.8 %), 122616.84 (2.8 %), 145077.58 (2.8 %), 114175.79 (2.8 %), and 121597.55 (2.8 %), together occurring 15 times, which is 13.9 % of the total.

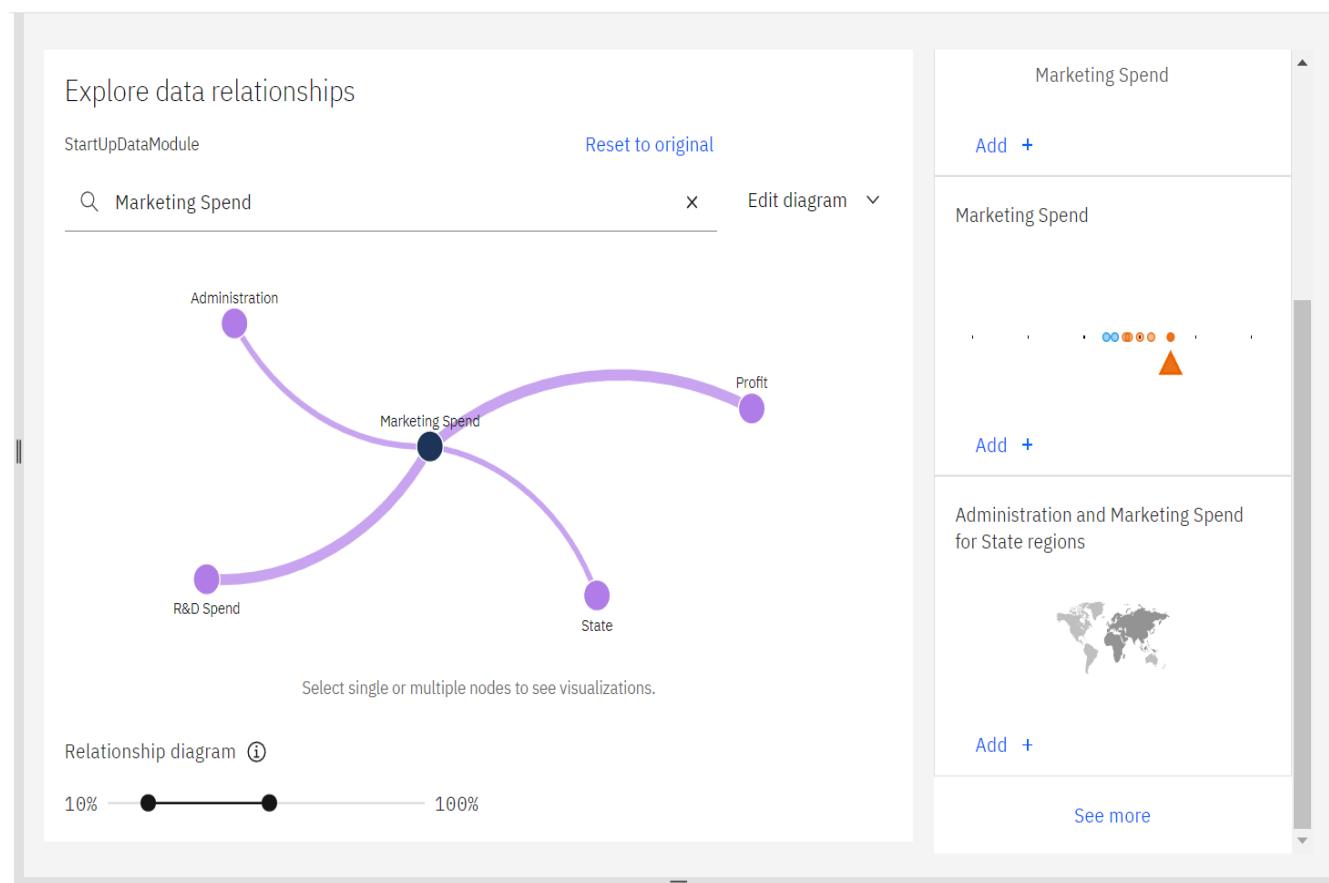
NAME: VINOODHINI D
REG NO: 917719IT118

NETWORK GRAPH BETWEEN R&D SPEND AND STATES:

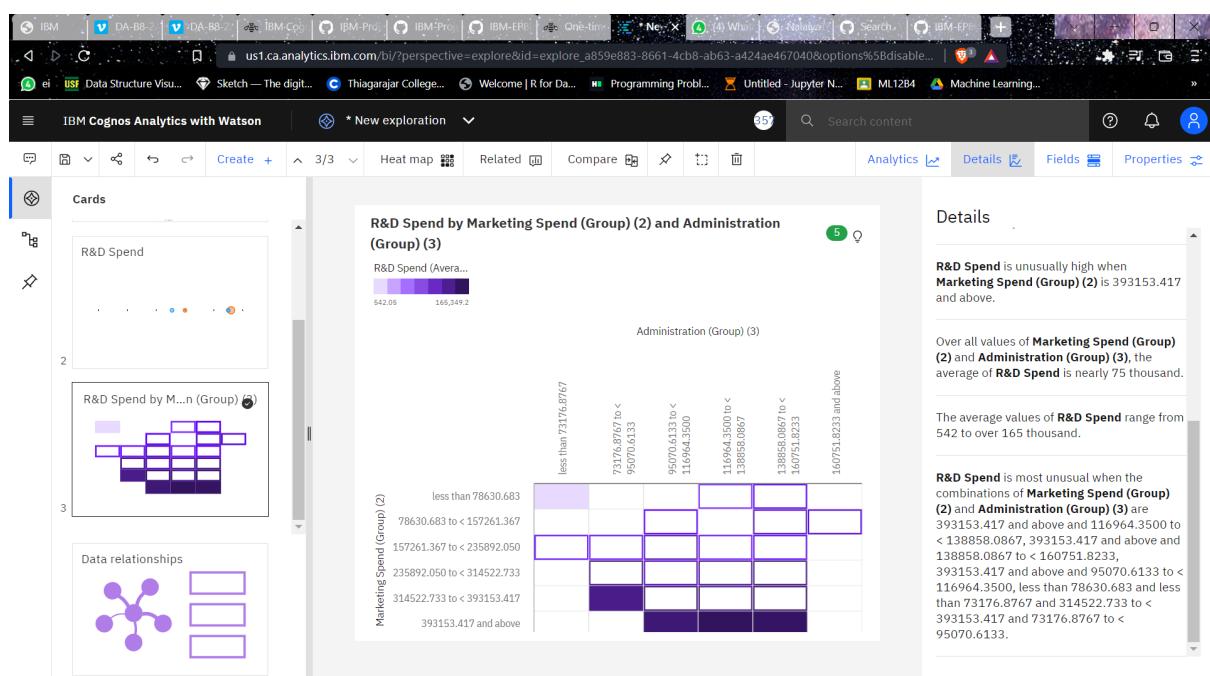


NAME: VINOODHINI D
REG NO: 917719IT118

RELATIONSHIP BETWEEN MARKETING SPEND AND OTHER COLUMNS:



R&D BY MARKETING SPEND AND ADMINISTRATION:



NAME: VINOODHINI D

REG NO: 917719IT118

R&D Spend is unusually high when Marketing Spend (Group) (2) is 393153.417 and above.

Over all values of Marketing Spend (Group) (2) and Administration (Group) (3), the average of R&D Spend is nearly 75 thousand.

The average values of R&D Spend range from 542 to over 165 thousand.

R&D Spend is most unusual when the combinations of Marketing Spend (Group) (2) and Administration (Group) (3) are 393153.417 and above and 116964.3500 to < 138858.0867, 393153.417 and above and 138858.0867 to < 160751.8233, 393153.417 and above and 95070.6133 to < 116964.3500, less than 78630.683 and less than 73176.8767 and 314522.733 to < 393153.417 and 73176.8767 to < 95070.6133.

EXPLORING PROFIT COLUMN:

NAME: VINOODHINI D

REG NO: 917719IT118

