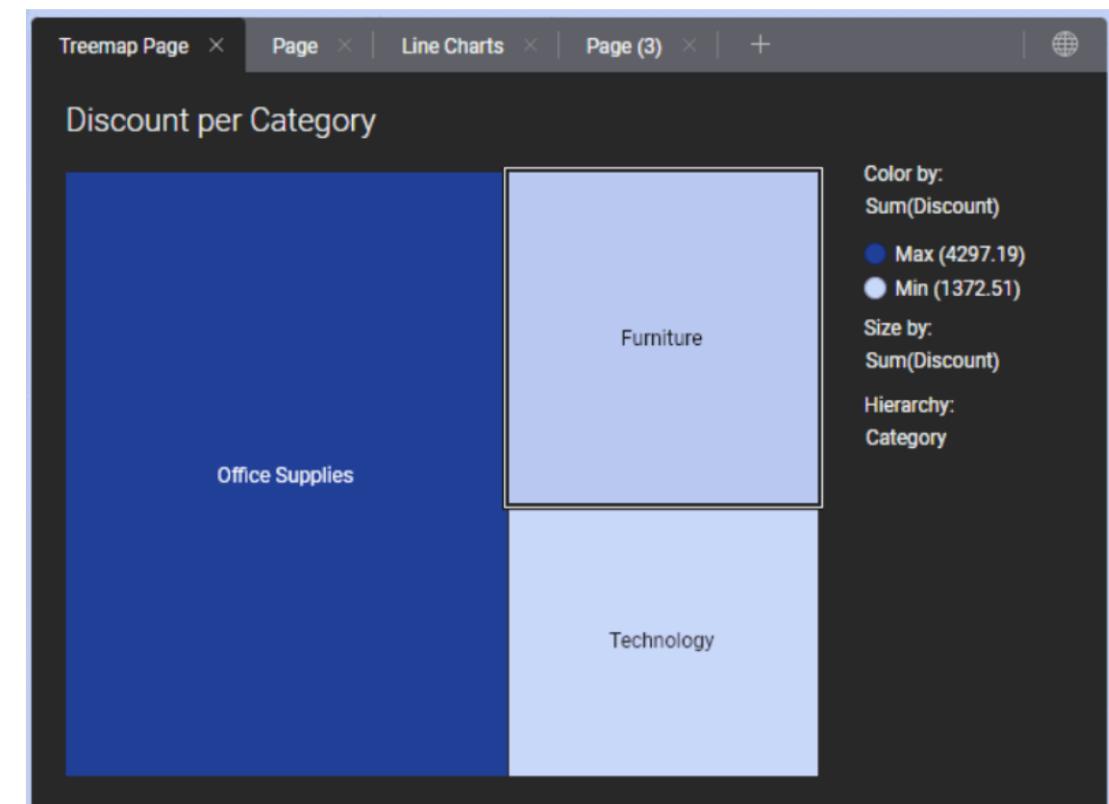
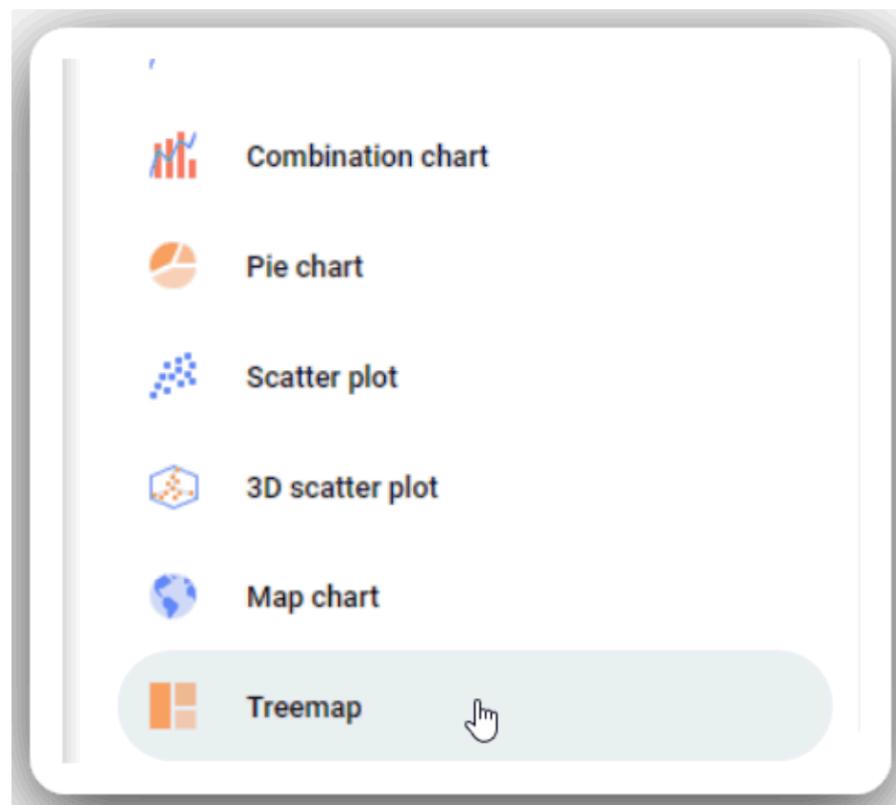


# How do I read a Tree Map?

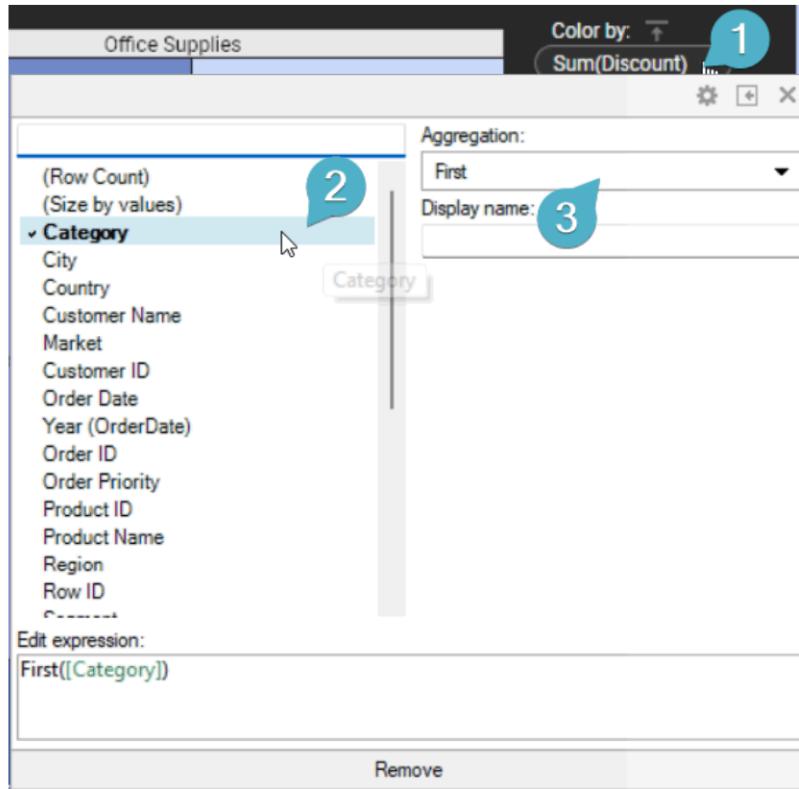
1. Sales Contribution by Category
2. Subcategory Performance
3. Visual Cues
4. Hierarchical Structure
5. Underperforming Areas



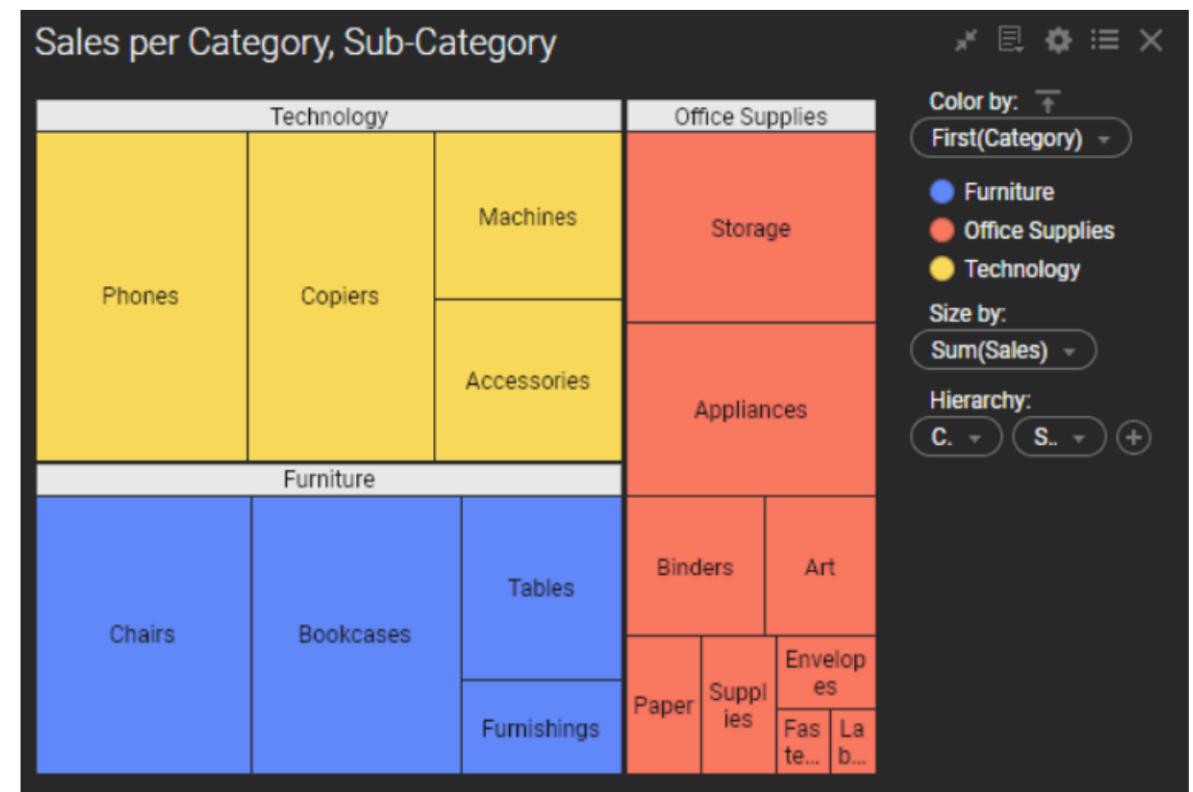
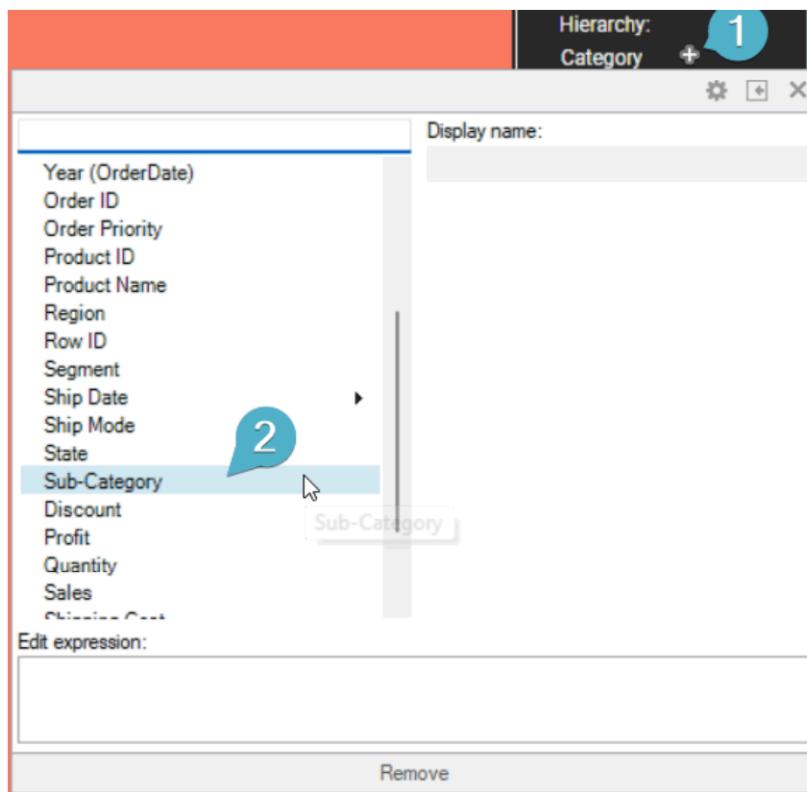
# Creating Treemaps



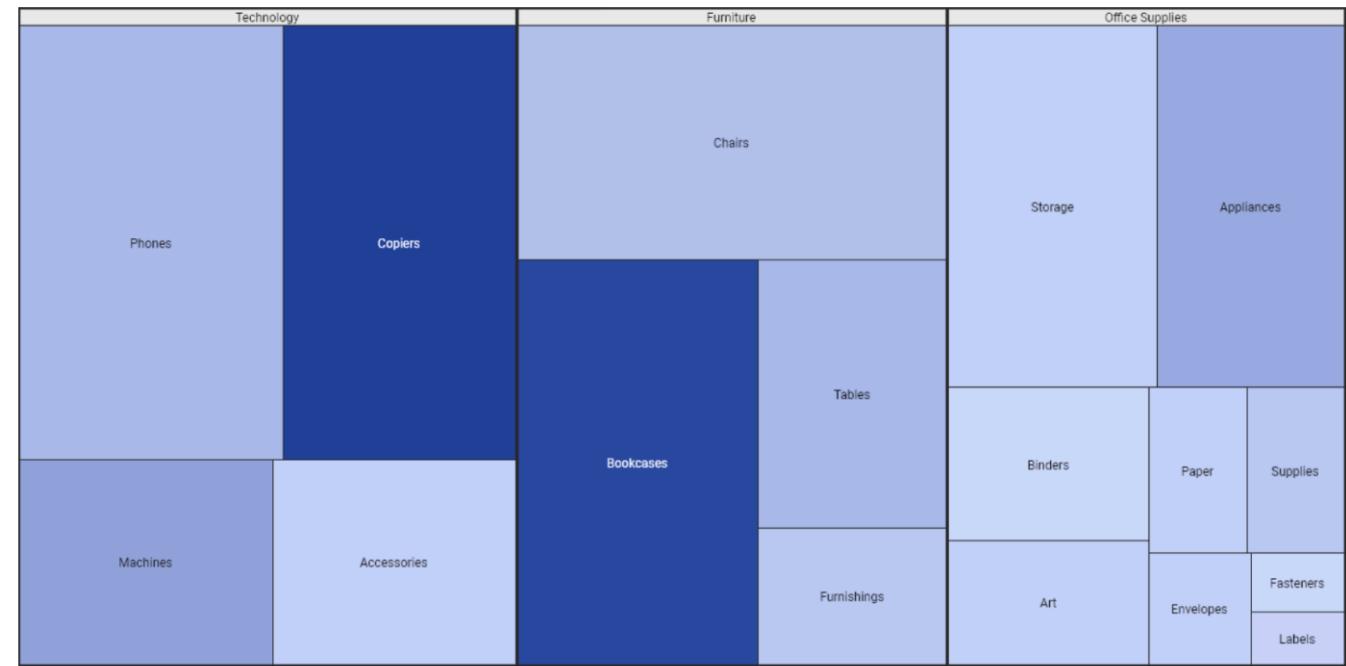
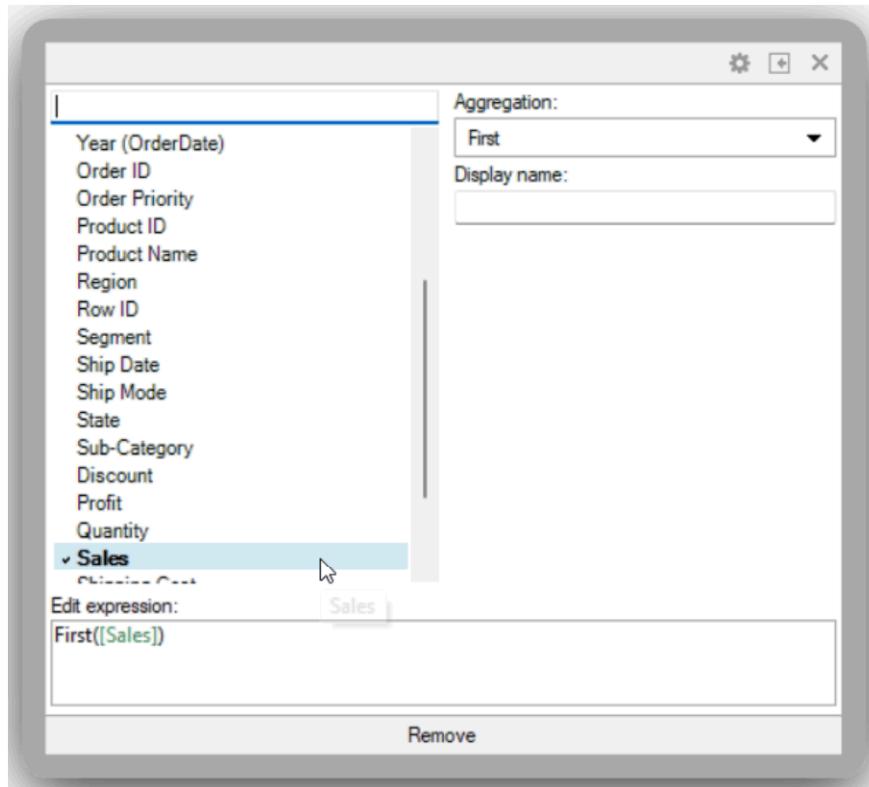
# Set Category Fields



# Add Sub-Category for Hierarchical View



# Visualize Sales with Color



# Scatter Plot

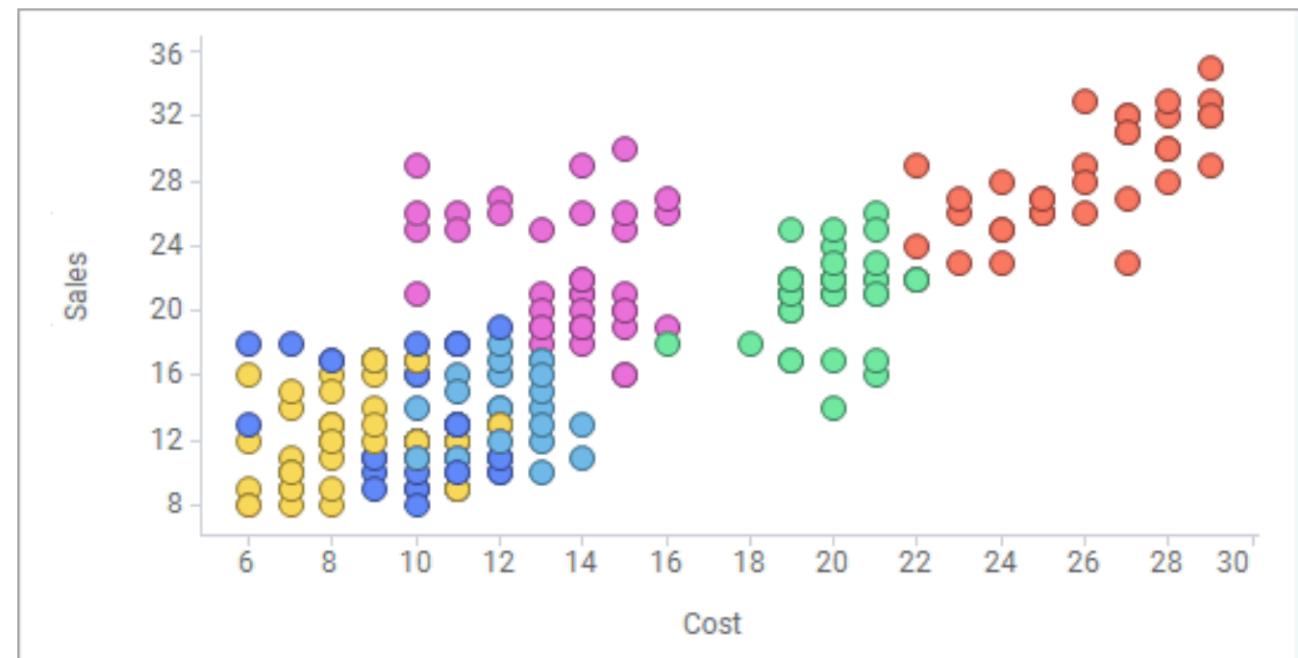
# Scatter Plot

## What Is a Scatter Plot?

- ✓ A **scatter plot** visualizes the relationship between two **numerical variables** by plotting data points on a two-dimensional graph.

## How Does It Work?

- ✓ X-axis: Represents the **independent variable**.
- ✓ Y-axis: Represents the **dependent variable**.
- ✓ **Data Points**: Each point's position reflects its values on both axes.
- ✓ **Trendline**: A line that indicates the general direction or pattern of the data, helping to identify correlations.



*What is a Scatter Plot? (2025). Tibco.com. [https://docs.tibco.com/pub/sfire-analyst/14.3.0/doc/html/en-US/TIB\\_sfire-analyst\\_UsersGuide/scat/scat\\_what\\_is\\_a\\_scatter\\_plot.htm](https://docs.tibco.com/pub/sfire-analyst/14.3.0/doc/html/en-US/TIB_sfire-analyst_UsersGuide/scat/scat_what_is_a_scatter_plot.htm)*

# Scatter Plot



## Adding More Dimensions

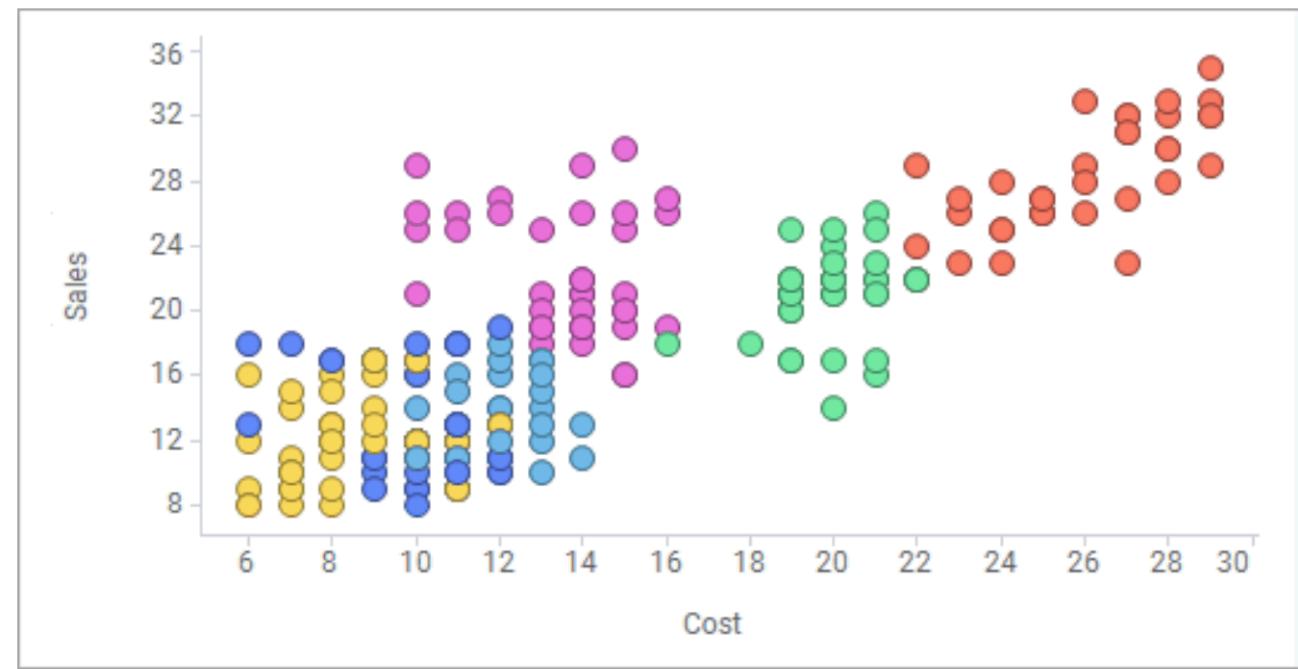
- A **third variable** can be encoded as:
  - Color (e.g., category, product)
  - Size (e.g., quantity, profit)
- Enables multi-dimensional analysis within a 2D space.

*What is a Scatter Plot?* (2025). Tibco.com. [https://docs.tibco.com/pub/sfire-analyst/14.3.0/doc/html/en-US/TIB\\_sf火Analyst\\_UsersGuide/scat/scat\\_what\\_is\\_a\\_scatter\\_plot.htm](https://docs.tibco.com/pub/sfire-analyst/14.3.0/doc/html/en-US/TIB_sf火Analyst_UsersGuide/scat/scat_what_is_a_scatter_plot.htm)

# Scatter Plot

## Correlation Between Variables

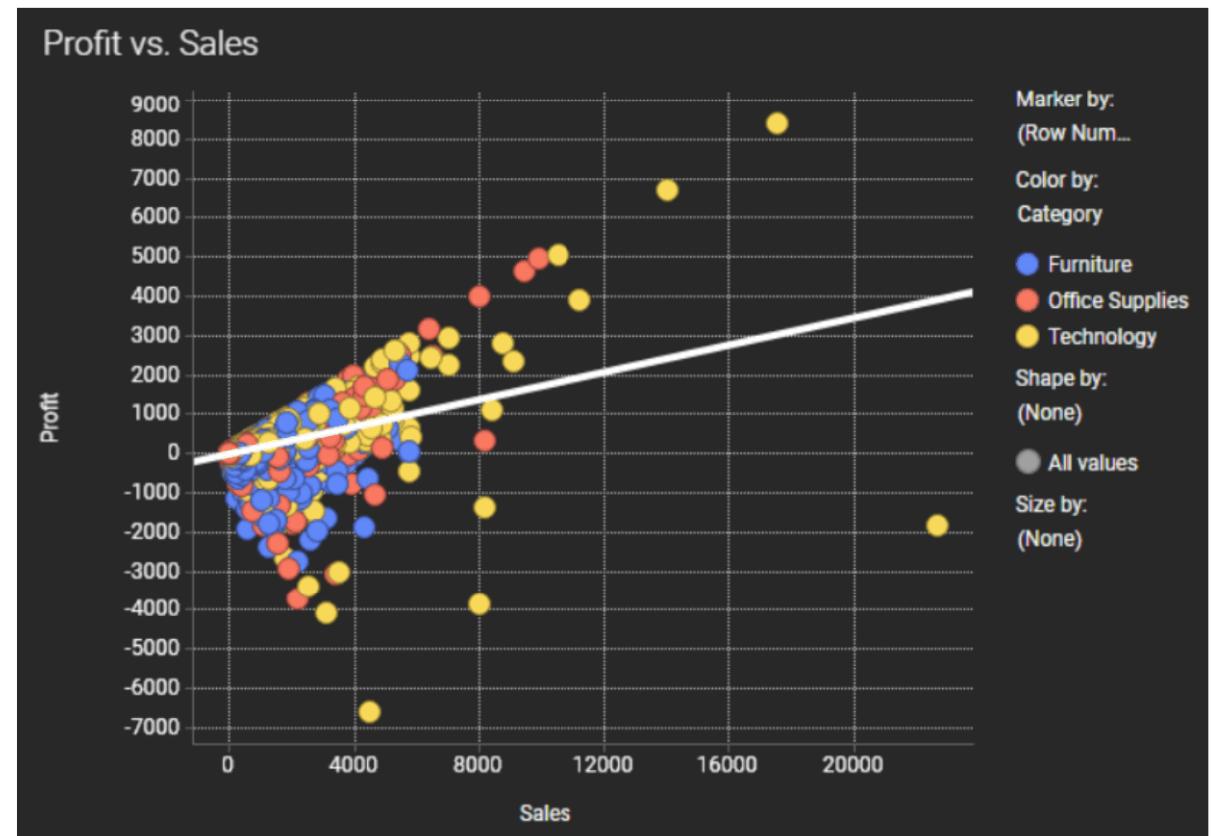
- The pattern of markers reflects the **relationship (correlation)** between variables:
  - Straight-line pattern** = High correlation
  - Scattered points** = Low or no correlation



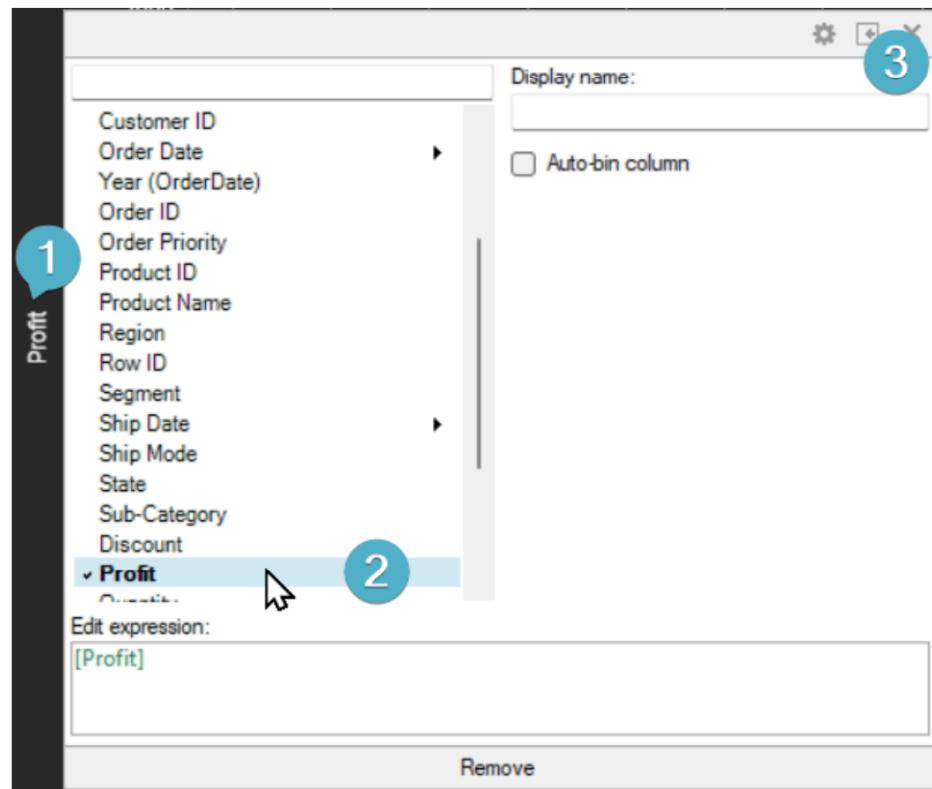
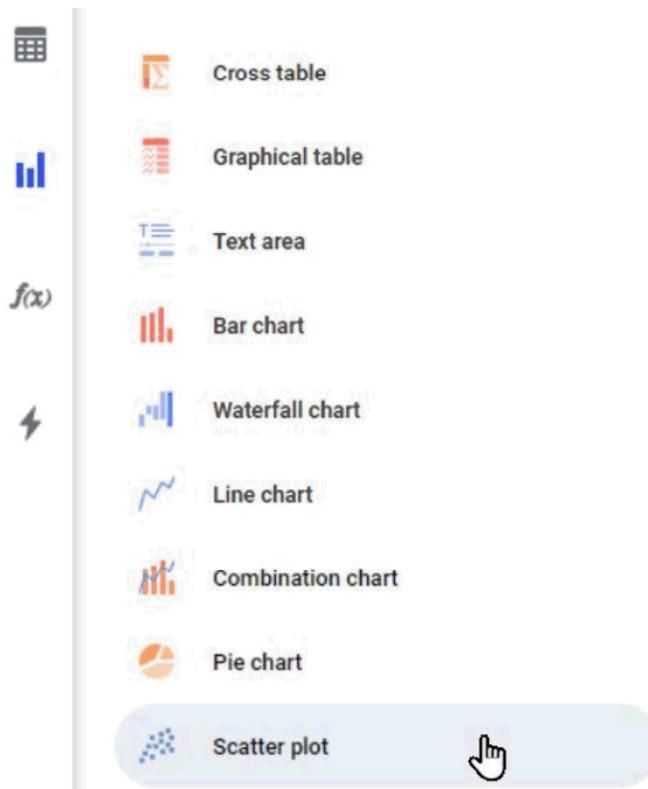
*What is a Scatter Plot?* (2025). Tibco.com. [https://docs.tibco.com/pub/sfire-analyst/14.3.0/doc/html/en-US/TIB\\_sf火Analyst\\_UsersGuide/scat/scat\\_what\\_is\\_a\\_scatter\\_plot.htm](https://docs.tibco.com/pub/sfire-analyst/14.3.0/doc/html/en-US/TIB_sf火Analyst_UsersGuide/scat/scat_what_is_a_scatter_plot.htm)

# Profit vs. Sales (Per Order or Product)

## 🔍 Key Insights

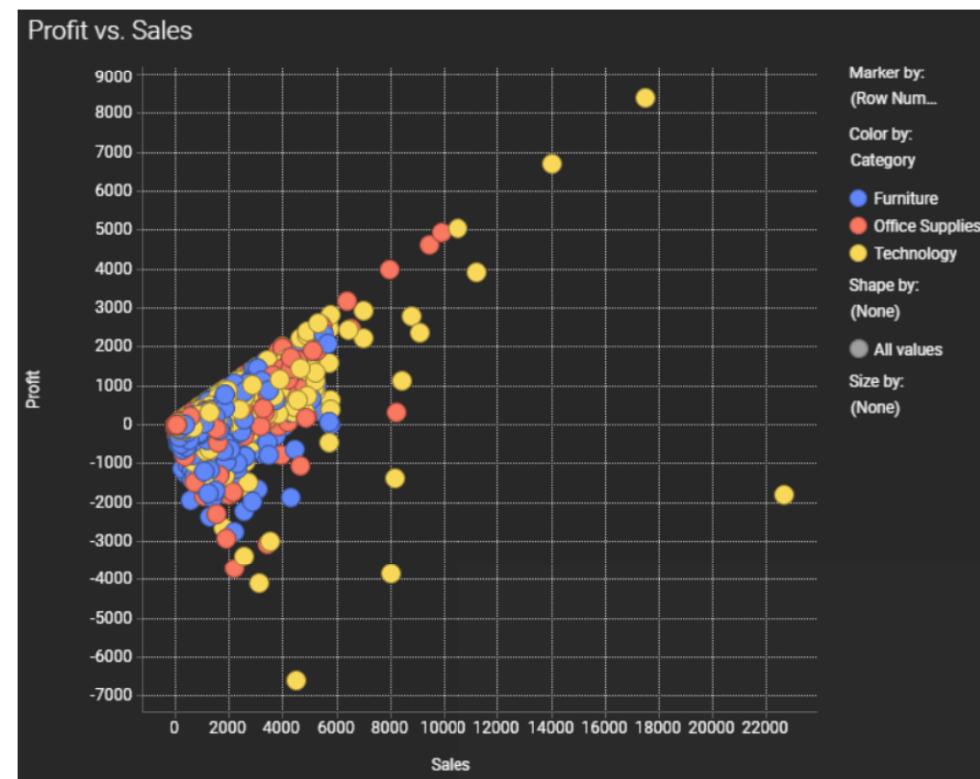


# Profit vs. Sales (Per Order or Product)

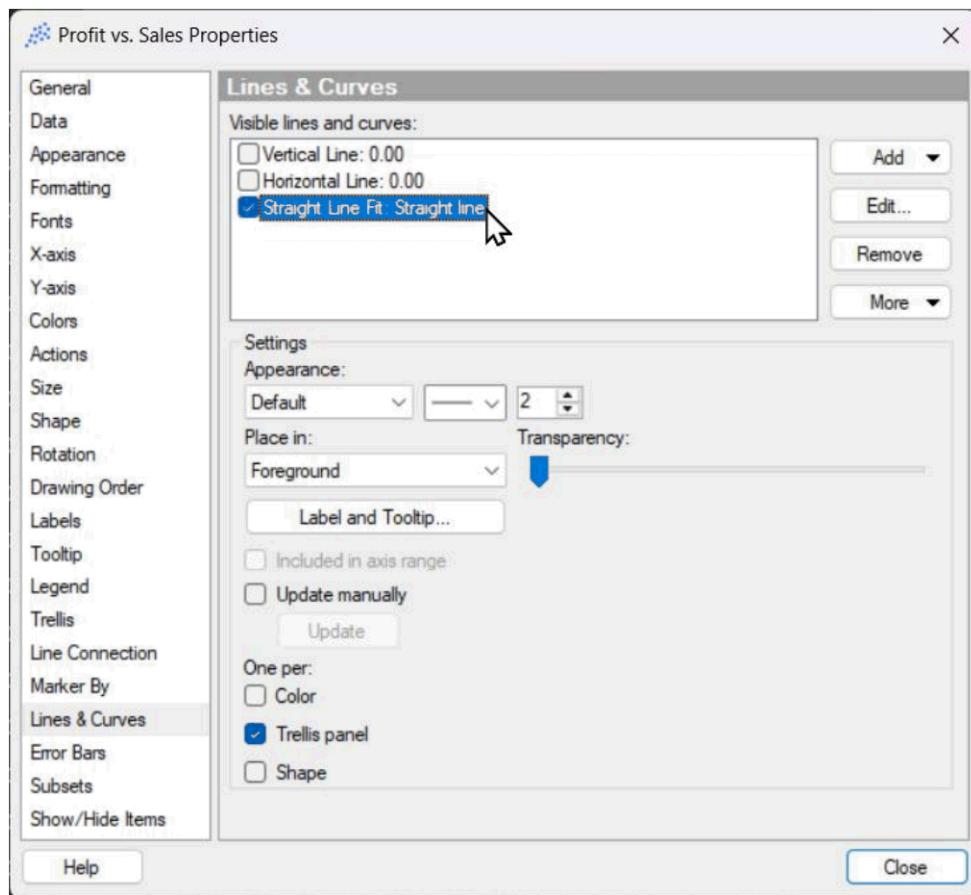


# Profit vs. Sales (Per Order or Product)

The screenshot shows the configuration interface for a scatter plot. The left pane lists various dimensions and measures: Year (OrderDate), Order ID, Order Priority, Product ID, Product Name, Region, Row ID, Segment, Ship Date, Ship Mode, State, Sub-Category, Discount, Profit, Quantity, Sales, and Profit. The 'Sales' measure is selected and highlighted with a blue circle labeled '1'. The middle pane shows the 'Edit expression:' field containing '[Sales]'. The right pane shows the 'Display name:' field and an unchecked checkbox for 'Auto-bin column'. A large blue circle labeled '3' is positioned at the top right of the configuration area.



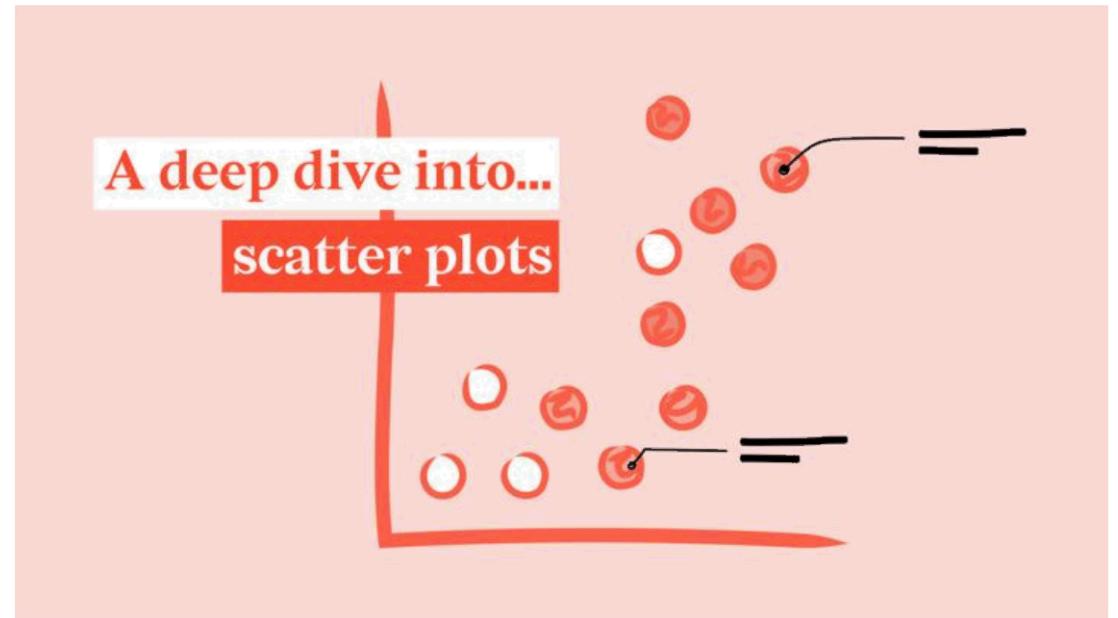
# Profit vs. Sales (Per Order or Product)



# Scatter Plot – Deep Dive

Correlation

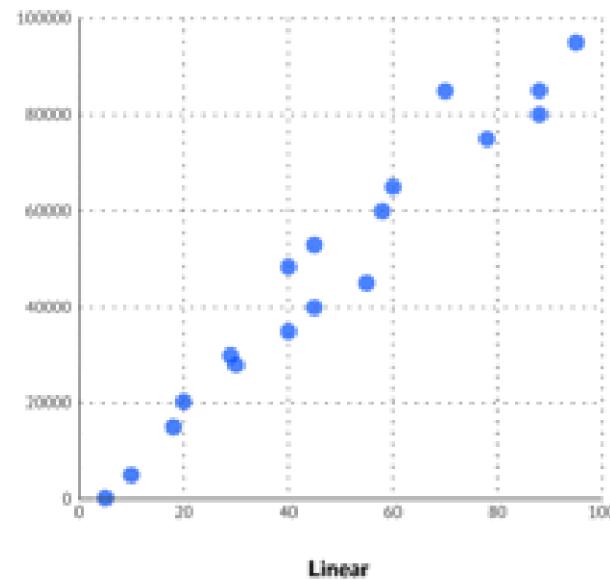
Cluster Detection



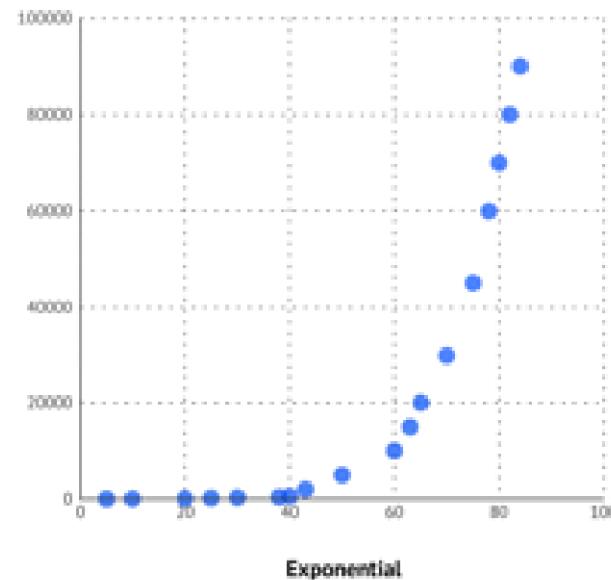
*A deep dive into... scatter plots | Blog | Datylon.* (2022). Datylon.com.  
<https://www.datylon.com/blog/scatter-plot-deep-dive>

# Correlation

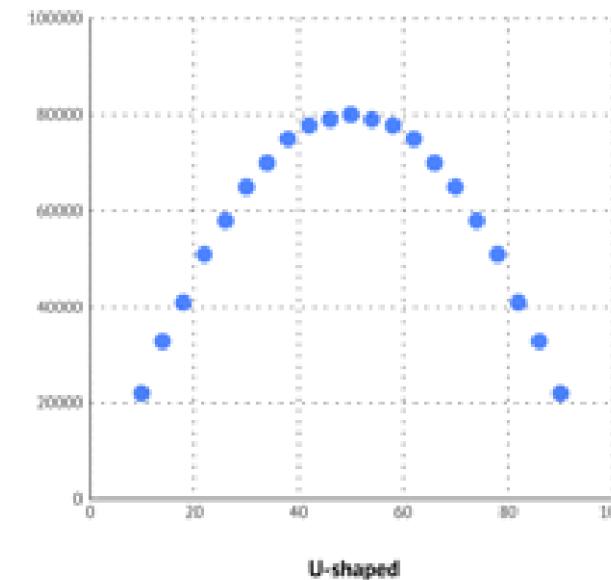
The forms of a scatter plot



Linear



Exponential

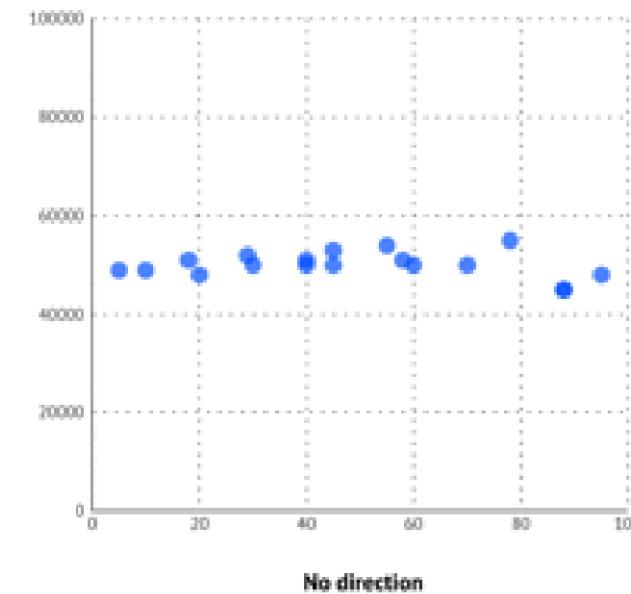
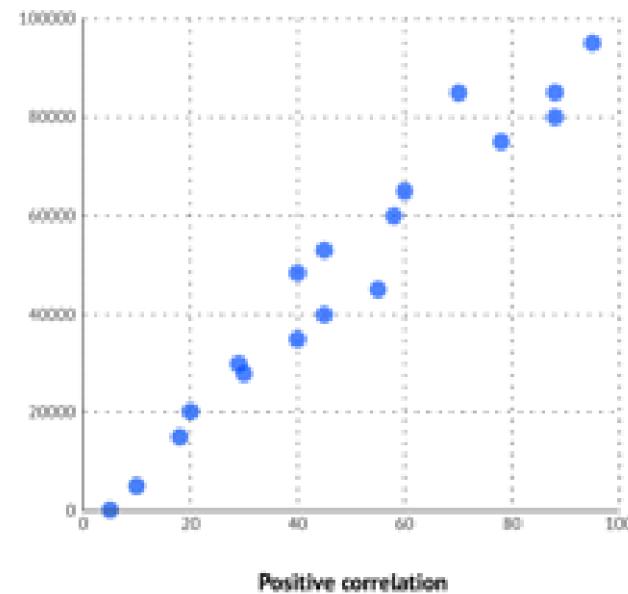
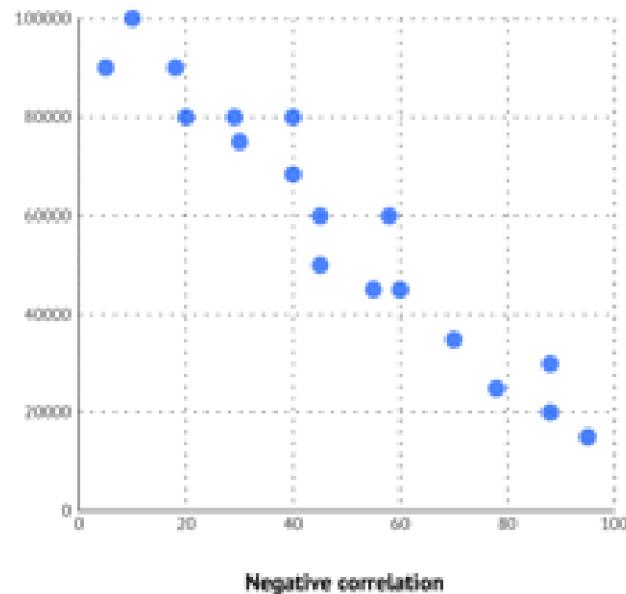


U-shaped

*A deep dive into... scatter plots | Blog | Datylon.* (2022). Datylon.com. <https://www.datylon.com/blog/scatter-plot-deep-dive>

# Correlation

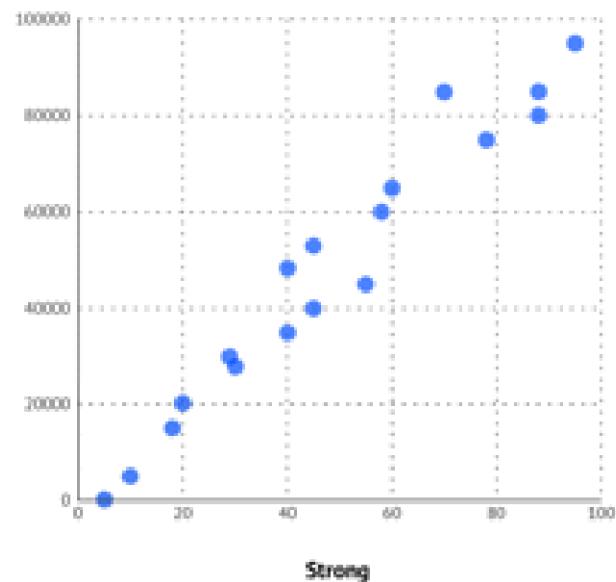
The directions of a scatter plot



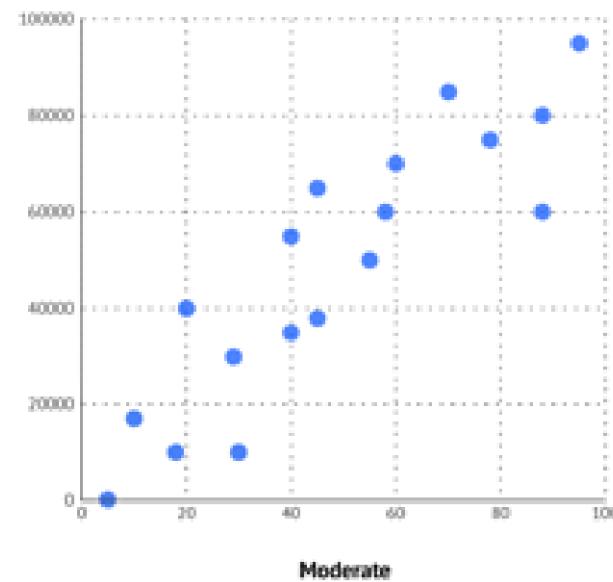
*A deep dive into... scatter plots | Blog | Datylon.* (2022). Datylon.com. <https://www.datylon.com/blog/scatter-plot-deep-dive>

# Correlation

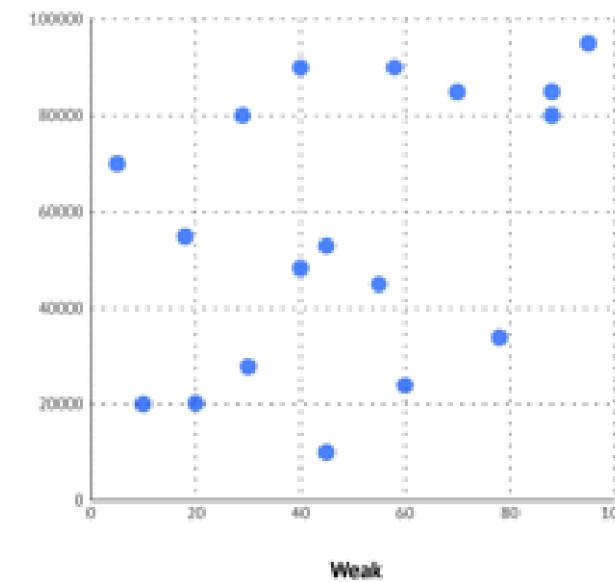
The strength of a scatter plot



Strong



Moderate

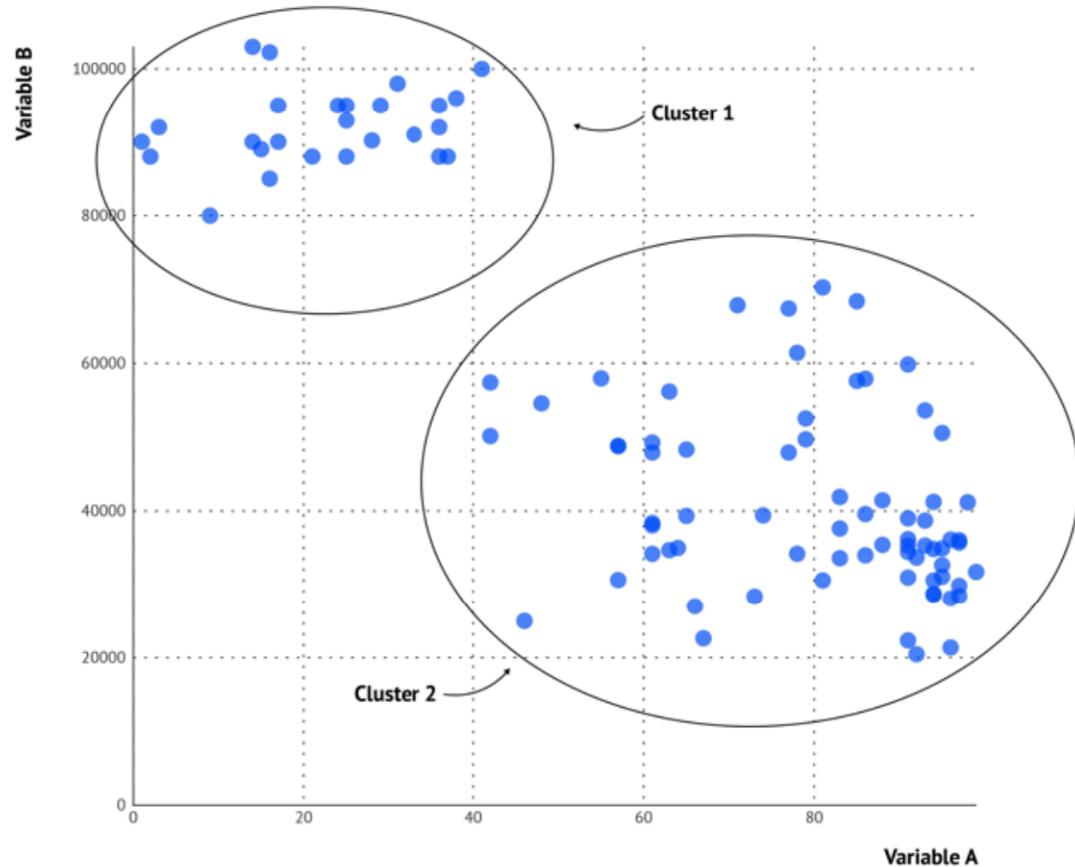


Weak

*A deep dive into... scatter plots | Blog | Datylon.* (2022). Datylon.com. <https://www.datylon.com/blog/scatter-plot-deep-dive>

# Cluster Detection

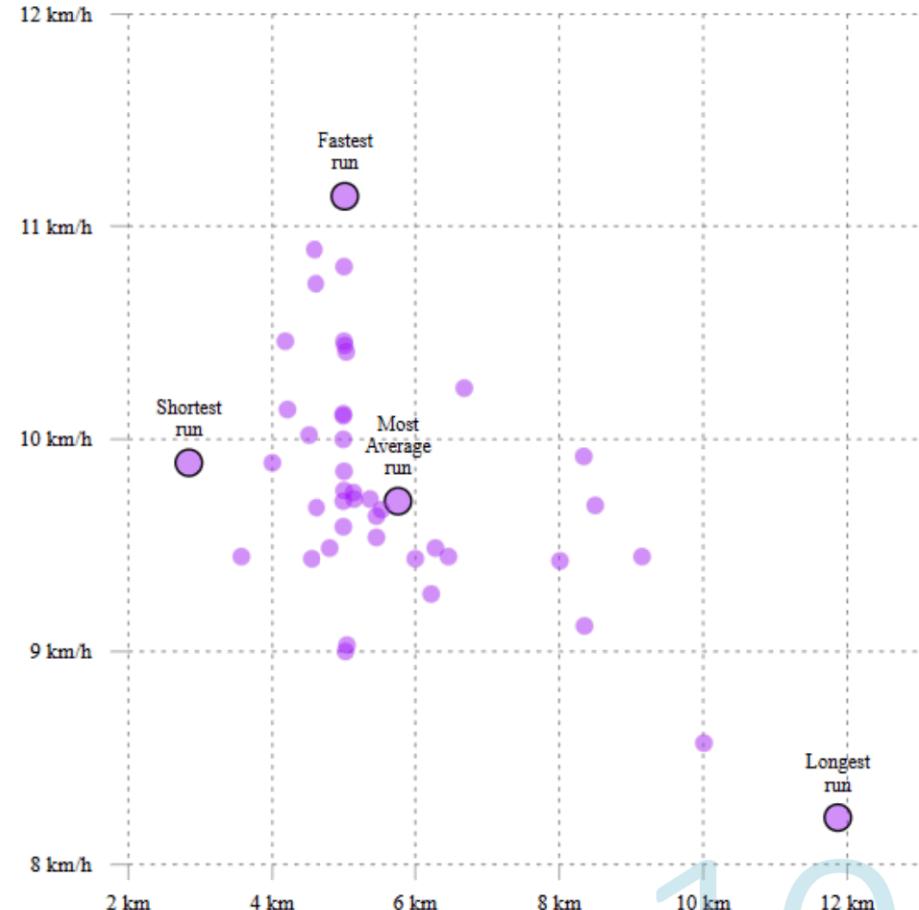
2 Clusters



# Types of Scatter Plots

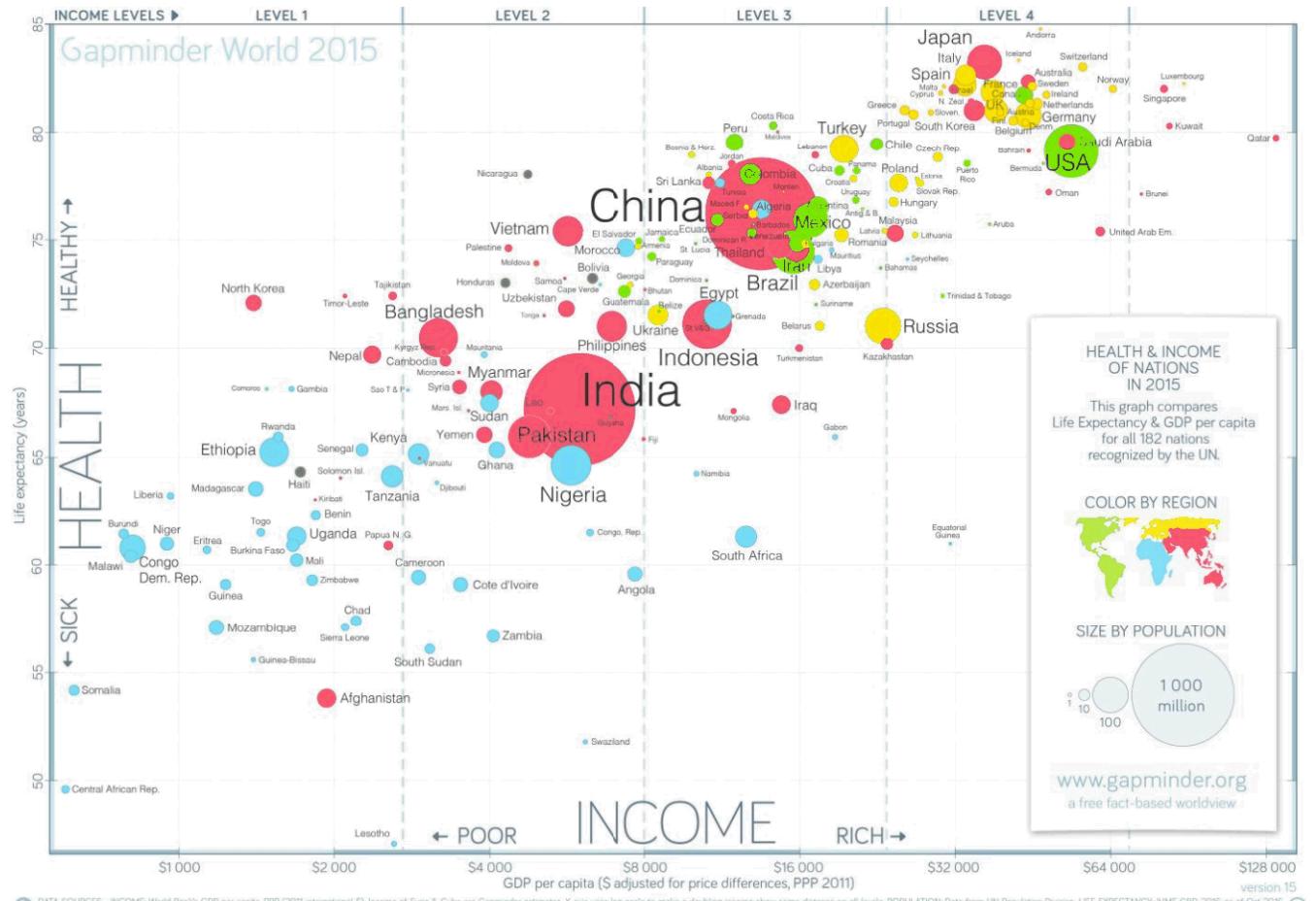
- ◆ 1. Basic Scatter Plot
- 2. Bubble Chart
- 3. Categorical Scatter Plot
- 4. Quadrant Chart

DISTANCE VS SPEED



# Types of Scatter Plots

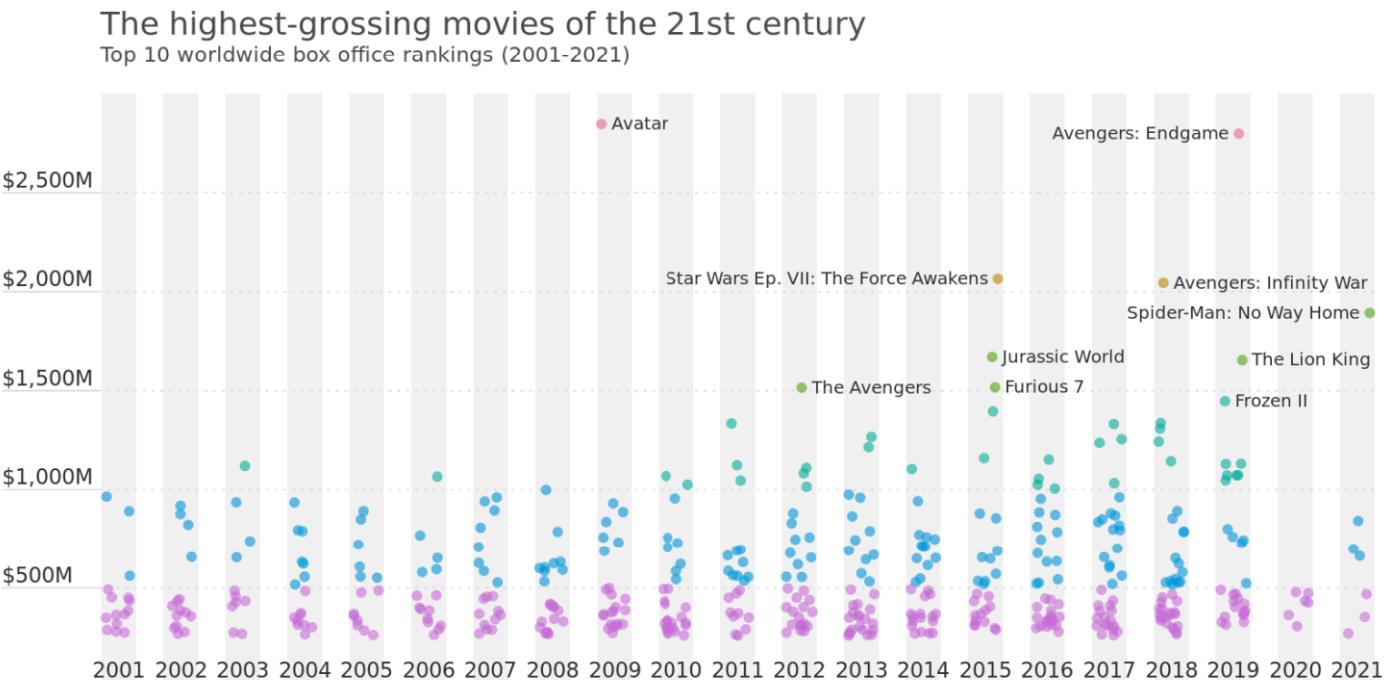
- ◆ 1. Basic Scatter Plot
- 2. Bubble Chart
- 3. Categorical Scatter Plot
- 4. Quadrant Chart



Updated Gapminder World Poster 2015! | Gapminder. (2015). Gapminder.org.  
<https://www.gapminder.org/downloads/updated-gapminder-world-poster-2015/>

# Types of Scatter Plots

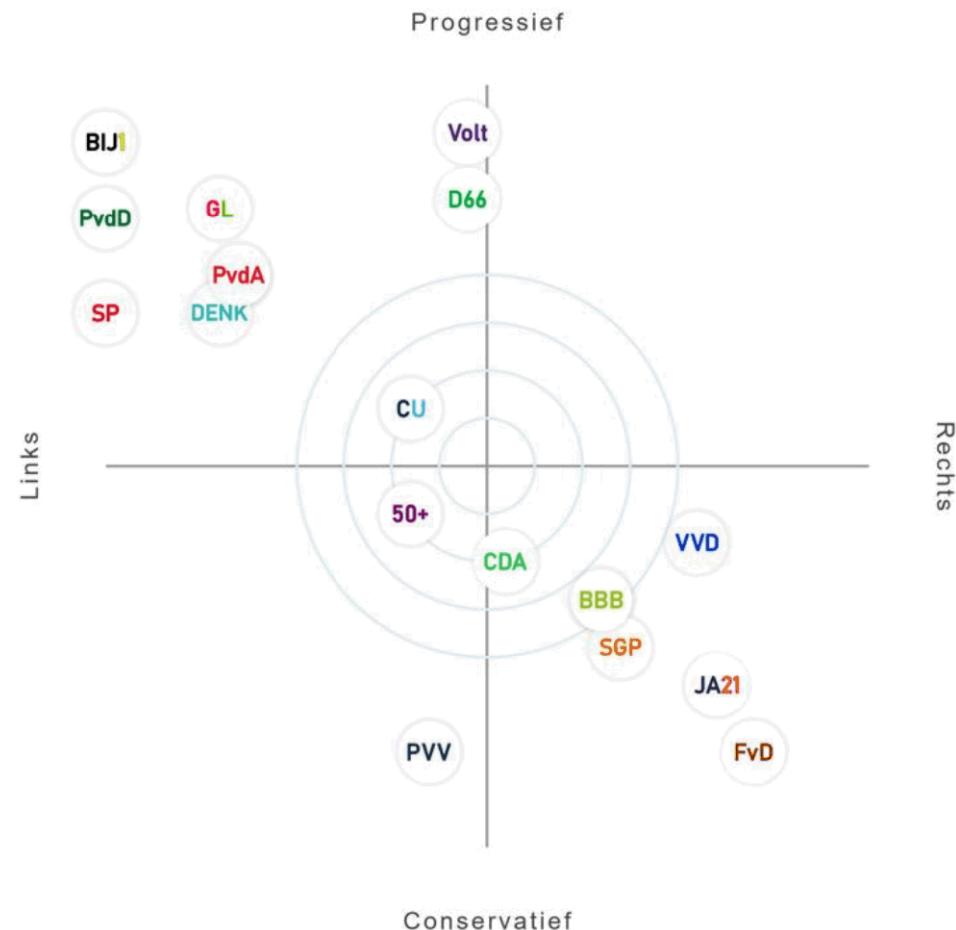
- ◆ 1. Basic Scatter Plot
- 2. Bubble Chart
- 3. Categorical Scatter Plot
- 4. Quadrant Chart



*Updated Gapminder World Poster 2015!* | Gapminder. (2015). Gapminder.org.  
<https://www.gapminder.org/downloads/updated-gapminder-world-poster-2015/>

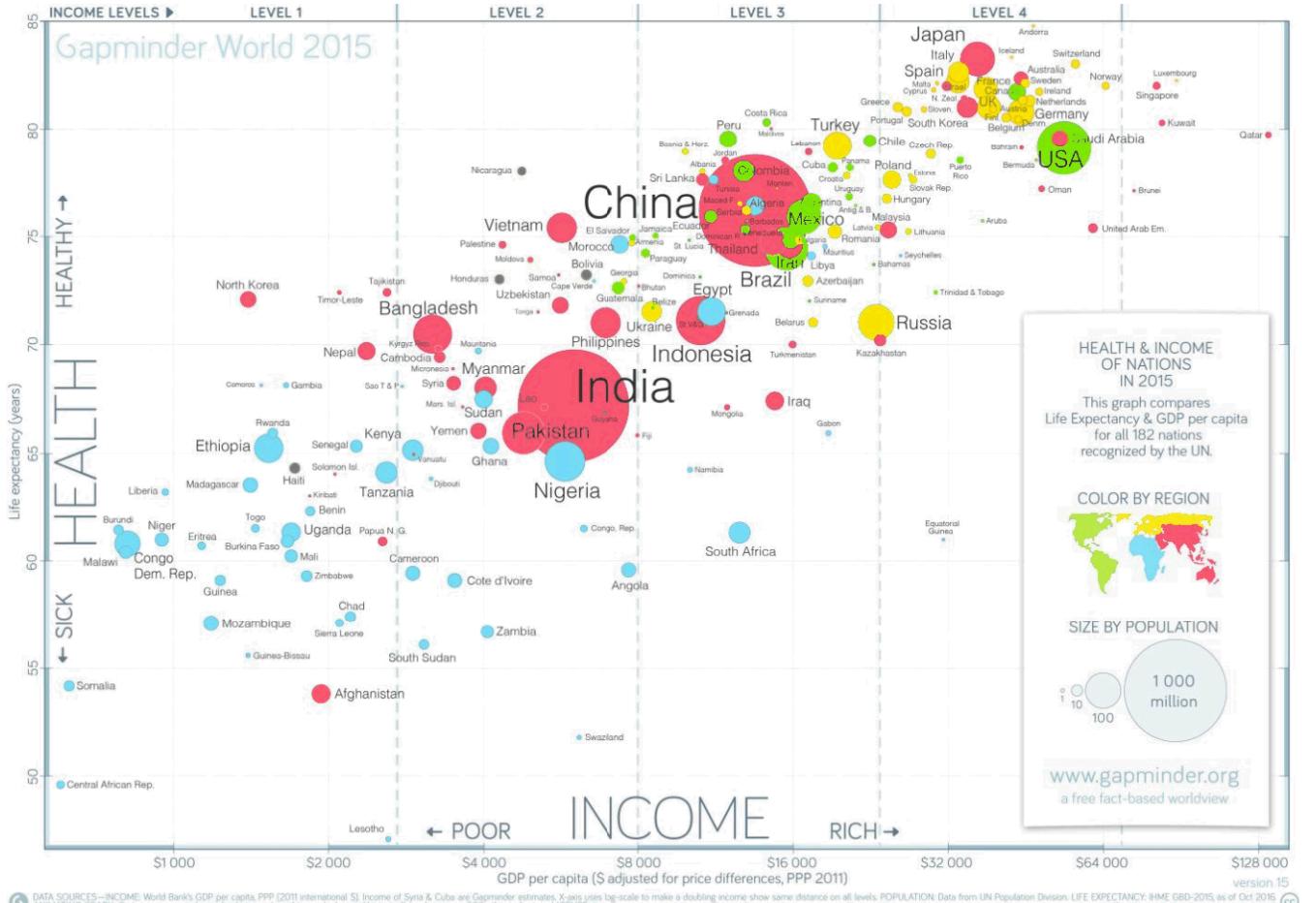
# Types of Scatter Plots

- ◆ 1. Basic Scatter Plot
- 2. Bubble Chart
- 3. Categorical Scatter Plot
- 4. Quadrant Chart



Updated Gapminder World Poster 2015! | Gapminder. (2015). Gapminder.org.  
<https://www.gapminder.org/downloads/updated-gapminder-world-poster-2015/>

# Bubble Chart



# Market vs Category sized by Market

## 🔍 What This Chart Shows

### 1. Top Performing Region-Category Combo

EMEA - Office Supplies shows the largest bubble

### 2. Technology Performs Consistently

Medium-to-large bubbles across all four markets

### 3. Furniture Is the Weakest Segment

Furniture bubbles are small across all markets

LATAM Is the Lowest Profit Market



# Market vs Category sized by Market

Insight	Recommended Action
🏆 EMEA Office Supplies	Scale further – strong profit center
🌐 Technology	Maintain/invest globally – consistent performance
📈 Furniture	Audit pricing, costs, or positioning across all markets
❗ LATAM	Assess market viability or adapt product strategy



# Project

# Day 2

14-May-25

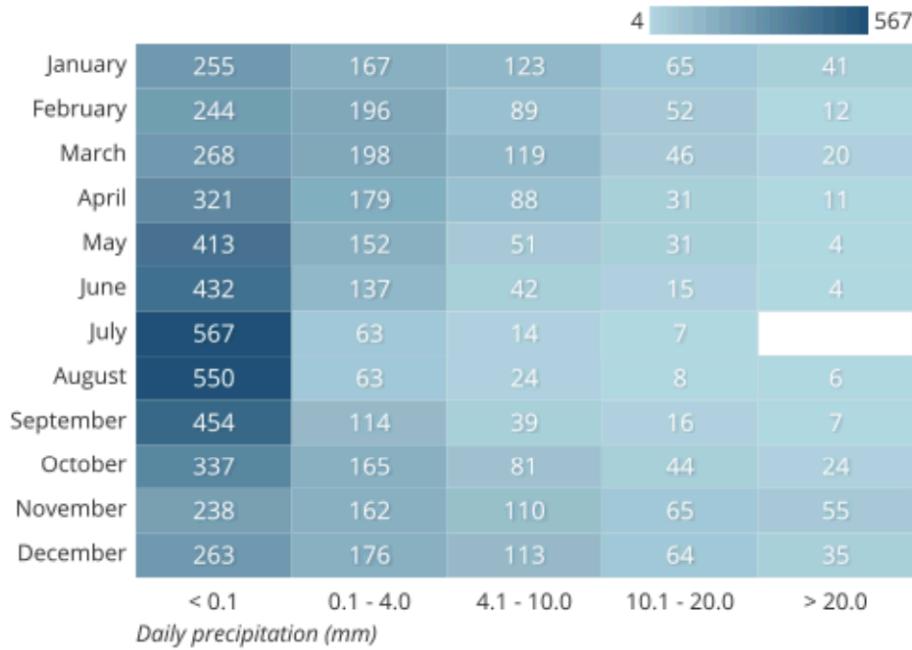
# Heatmaps

# Heat Map (heatmap)

## 🔍 What Is a Heatmap?

- ✓ A **heatmap** displays data values as colored cells on a grid.
- ✓ The grid is defined by **two axis variables** (e.g., time vs. category).
- ✓ Each cell's **color intensity** represents the value of a third (main) variable.

Seattle precipitation by month, 1998-2018



Atlassian. (2025). *A Complete Guide to Heatmaps* | Atlassian. Atlassian.  
<https://www.atlassian.com/data/charts/heatmap-complete-guide>

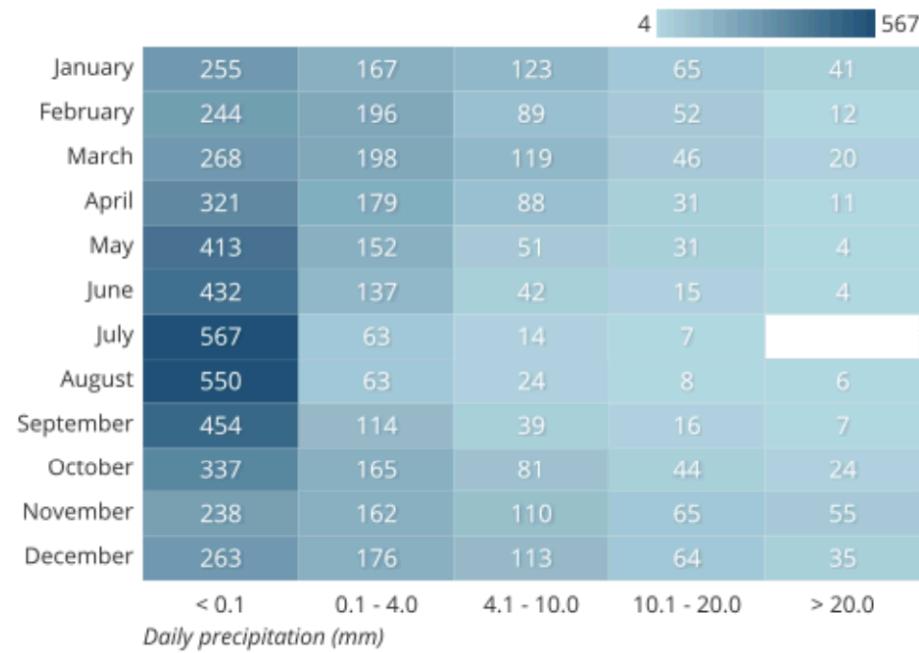
# Heatmap

## Example Insight

A heatmap showing **daily precipitation by month** in Seattle reveals:

- Most days had **no rain** (darkest left-most column).
- **Rain is most common from November to March.**
- **July and August** are typically **driest months**.

Seattle precipitation by month, 1998-2018



Atlassian. (2025). *A Complete Guide to Heatmaps* | Atlassian. Atlassian.  
<https://www.atlassian.com/data/charts/heatmap-complete-guide>

# Heatmap

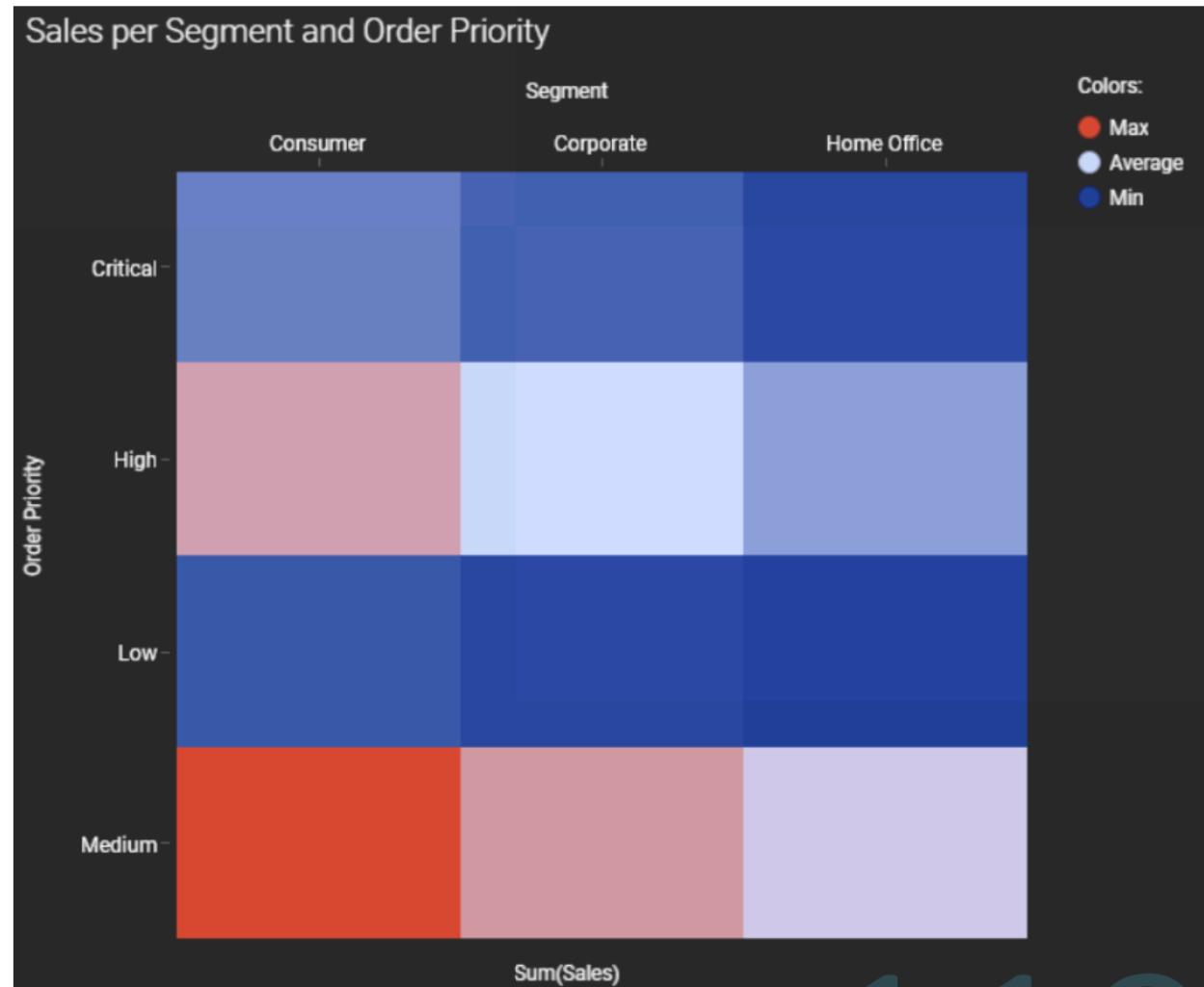
- 1. Highest Sales: Consumer – Medium Priority

*Strategic implication: Optimize inventory & delivery for medium-priority consumer orders.*

- 2. High Sales in Corporate – Medium

- 3. Lowest Sales Across Low and Critical Priorities

- 4. Home Office Segment is Generally Lower in Sales



# Heatmap

This screenshot shows the configuration interface for a heatmap. On the left, there is a list of dimensions and measures. The dimension 'Country' is selected and highlighted with a blue bar at the top. The measure 'Sum([Order Priority])' is selected and highlighted with a blue bar at the bottom. The 'Edit expression:' field contains the formula `[Order Priority]`. The 'Remove' button is located at the bottom right.

Display name: [ ]

Order Priority

- (Subsets)
- (Row Number)
- Category
- City
- Country
- Customer Name
- Market
- Customer ID
- Order Date
- Year (OrderDate)
- Order ID
- Order Priority
- Product ID
- Product Name
- Region
- Row ID

Edit expression:  
[Order Priority]

Remove

This screenshot shows the configuration interface for a heatmap. The dimension 'Sales' is selected and highlighted with a blue bar at the top. The measure 'Sum(Sales)' is selected and highlighted with a blue bar at the bottom. The 'Edit expression:' field contains the formula `Sum([Sales])`. The 'Remove' button is located at the bottom right.

Aggregation: Sum

Display name: [ ]

Region  
Row ID  
Segment  
Ship Date  
Ship Mode  
State  
Sub-Category  
Discount  
Profit  
Quantity

Sales

- Shipping Cost
- ProfitSign
- Percentage Category Sales
- Marked 10:16:47 AM
- Sub-Category Profit Margin

Edit expression:  
Sum([Sales])

Remove

Sum(Sales)

# Heatmap

The screenshot shows the Power BI Data view interface. At the top, there are three segments: Consumer, Corporate, and Home Office. A mouse cursor is hovering over the "Segment" button. Below the segments is a list of fields: City, Country, Customer Name, Market, Customer ID, Order Date, Year (OrderDate), Order ID, Order Priority, Product ID, Product Name, Region, Row ID, and Segment. The "Customer ID" field is currently selected. At the bottom, there is an "Edit expression:" section with the formula [Segment] and a "Remove" button.



# Graphical Table

# Graphical Table

## What Is a Graphical Table?

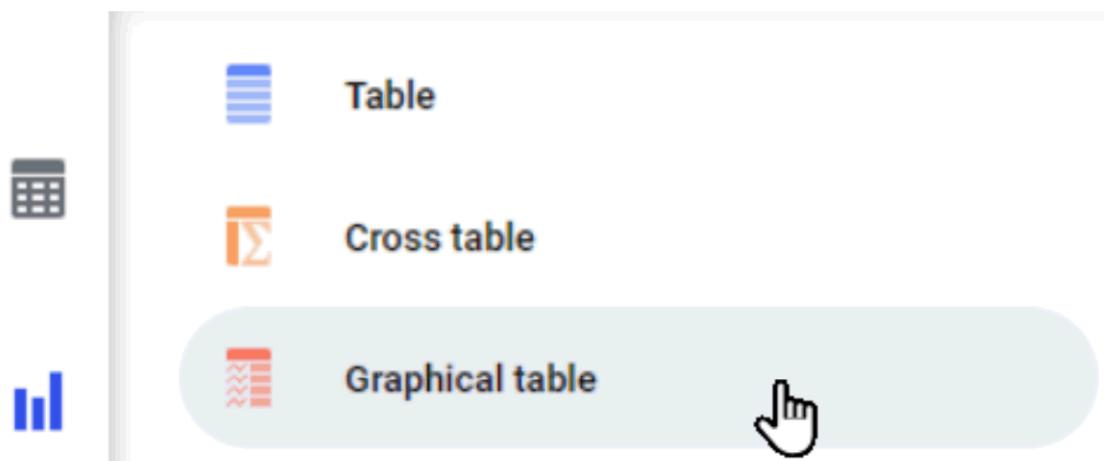
A **graphical table** is a **summary-style visualization** that combines **data** and **visuals** to present key insights at a glance.

Region	Trend	Change	Total	vs Target
Northeast	↗	↑	\$957,176	↓
South	↔	↑	\$1,201,056	↔
West	↙	↓	\$1,172,601	↓

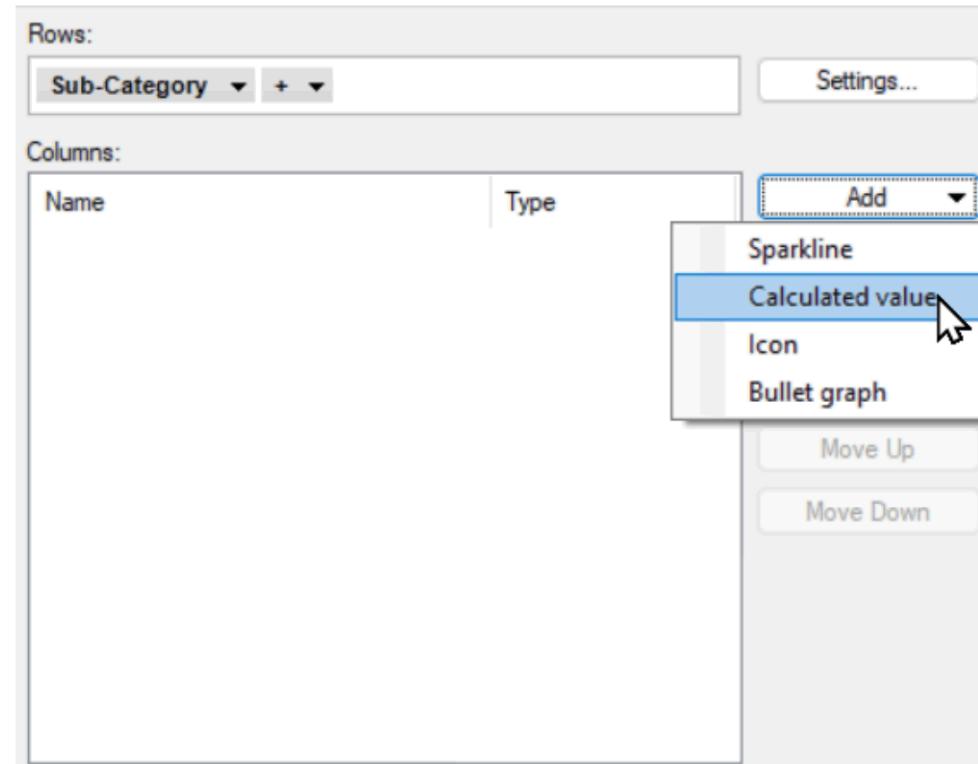
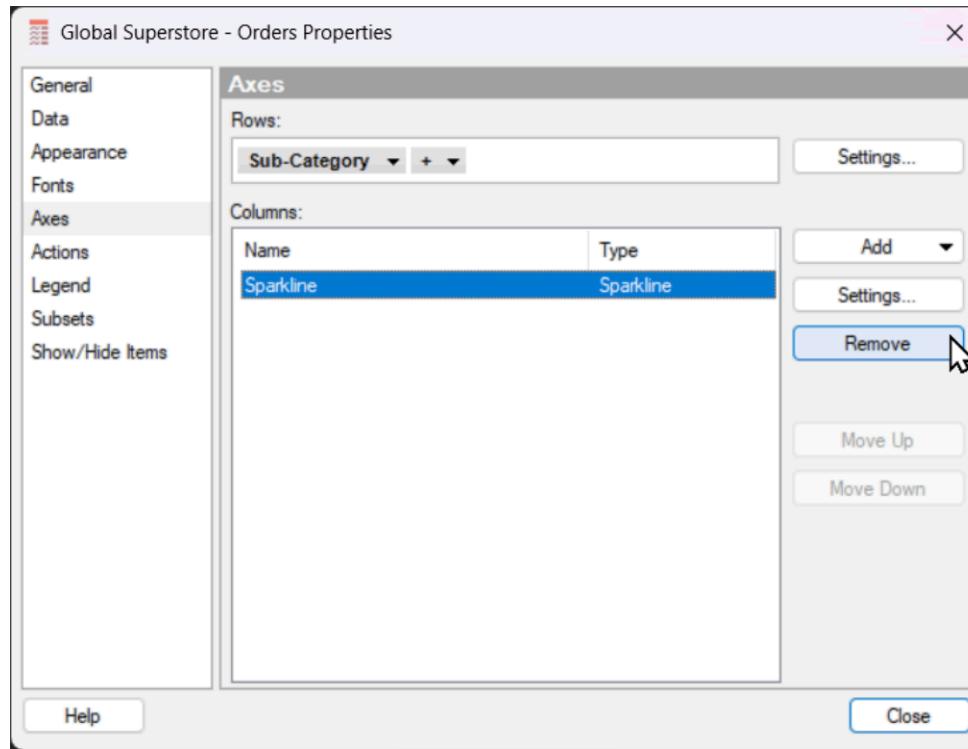
*What is a Graphical Table? (2025). Tibco.com.*

[https://docs.tibco.com/pub/spotfire/7.0.1/doc/html/gt/gt\\_what\\_is\\_a\\_graphical\\_table.htm](https://docs.tibco.com/pub/spotfire/7.0.1/doc/html/gt/gt_what_is_a_graphical_table.htm)

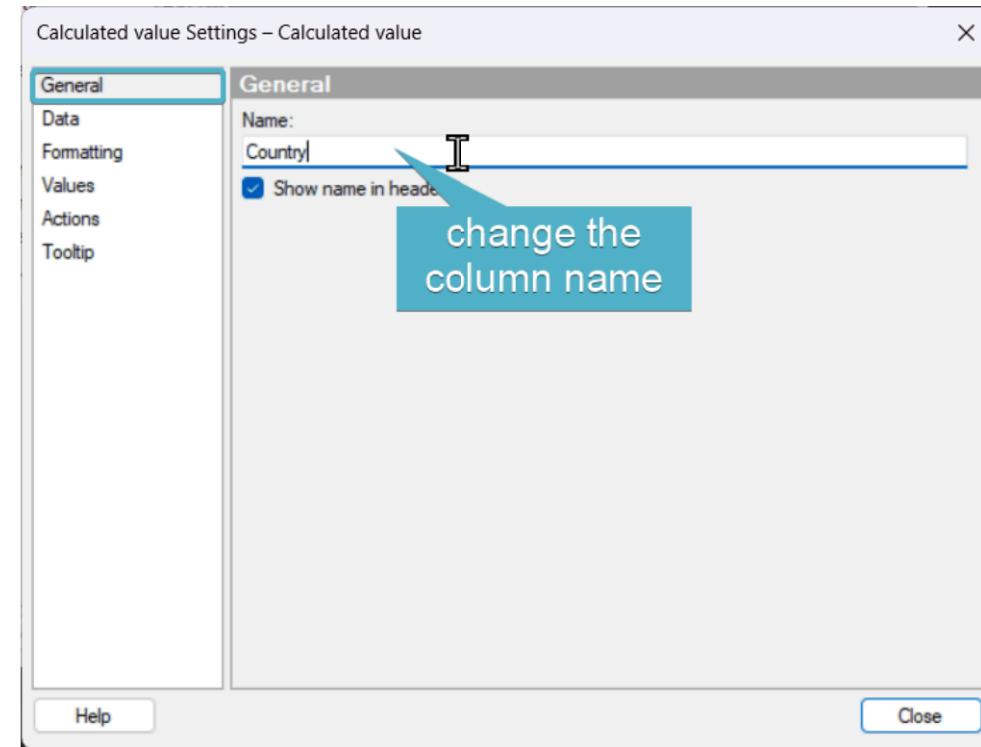
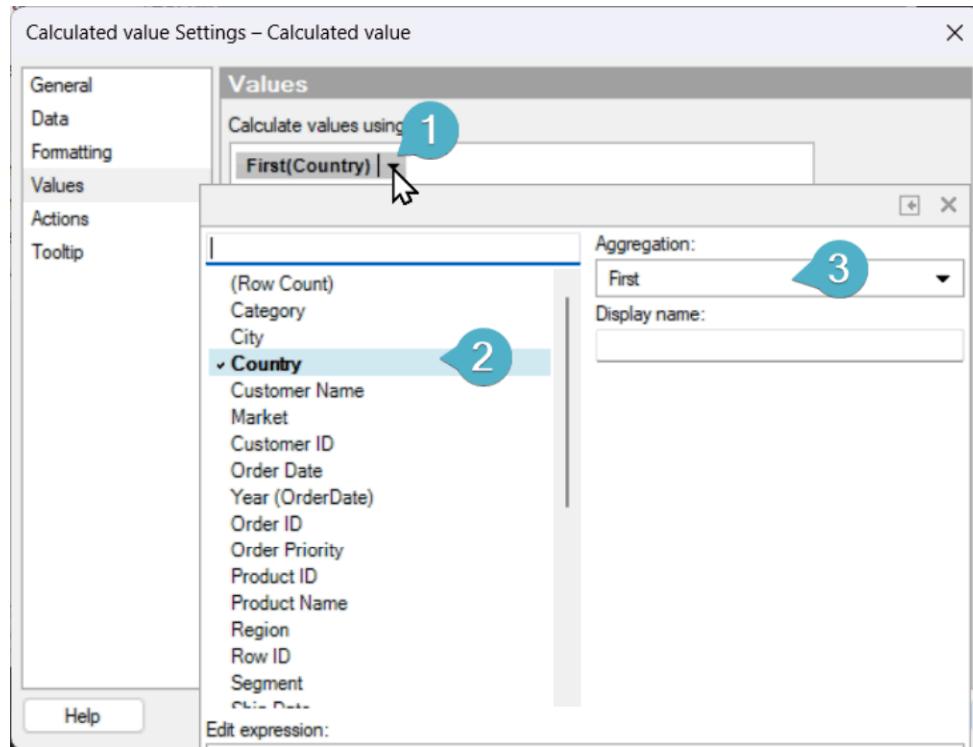
# Add Graphical Table



# Add Country as Calculated Column



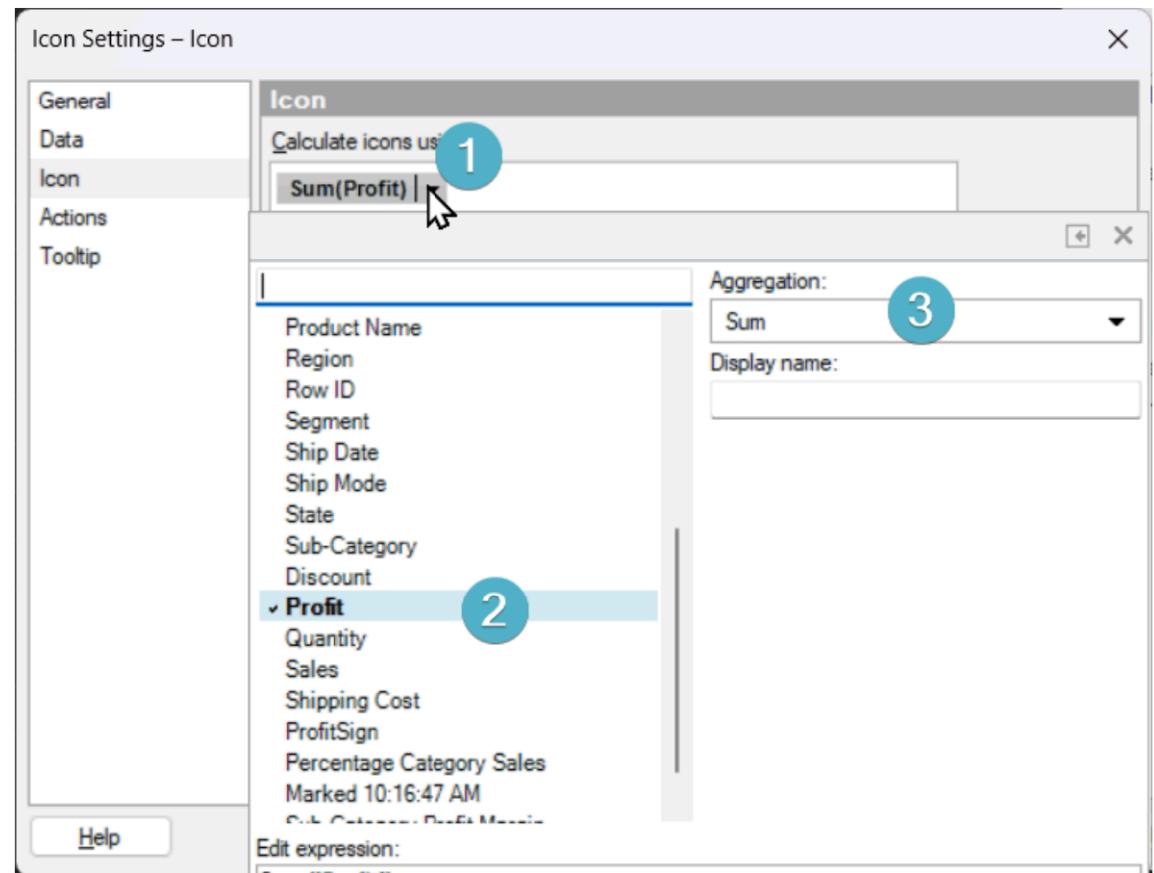
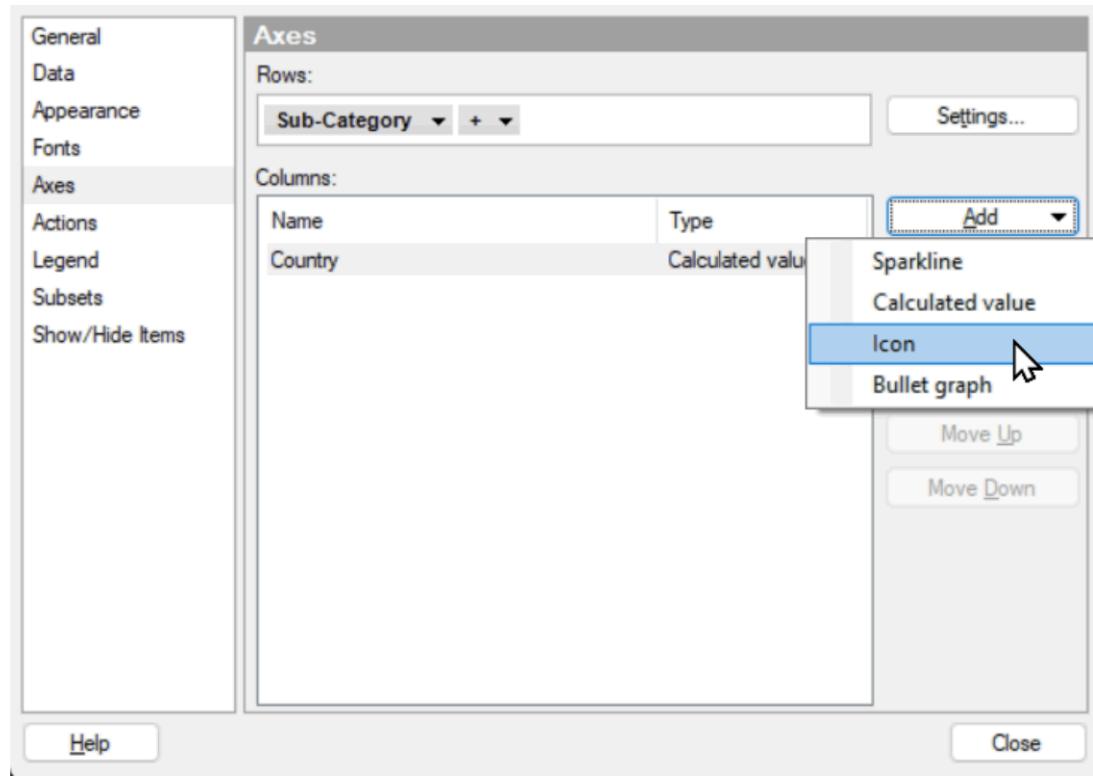
# Add Country as Calculated Column



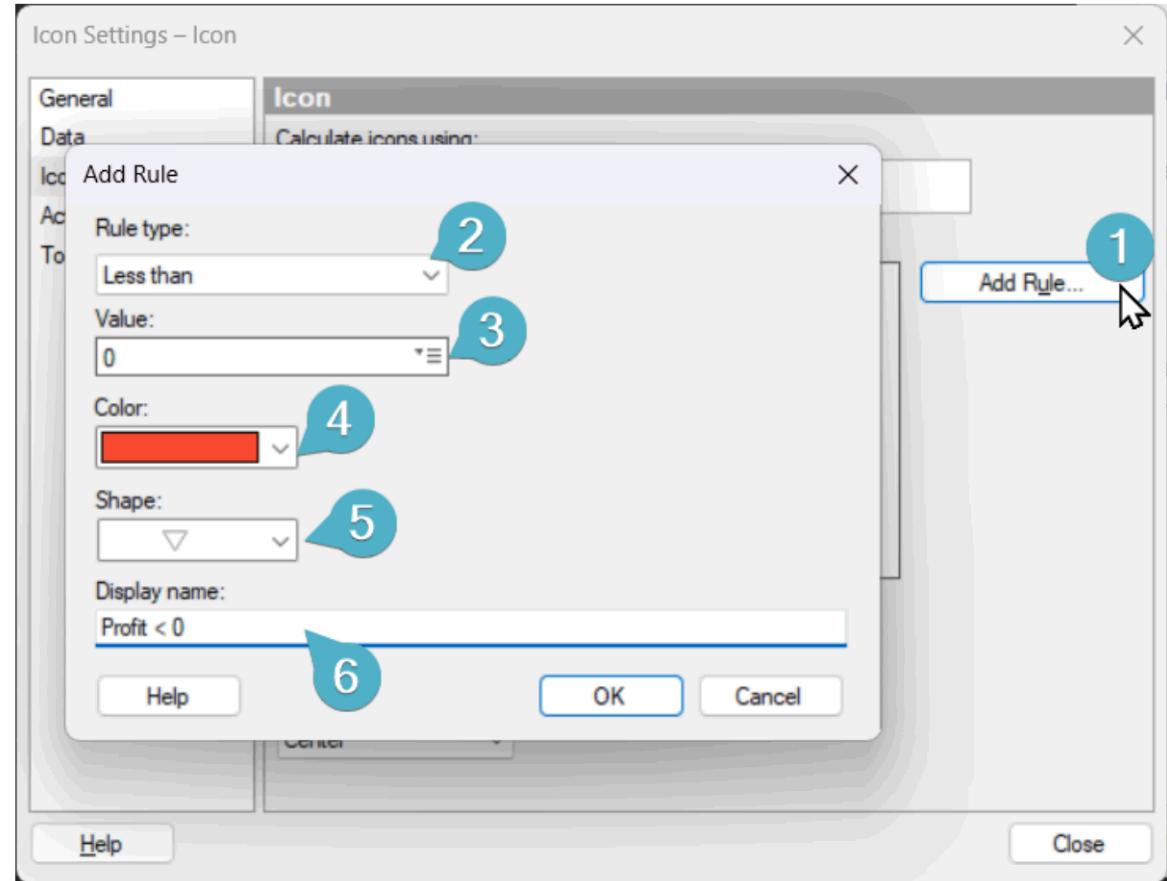
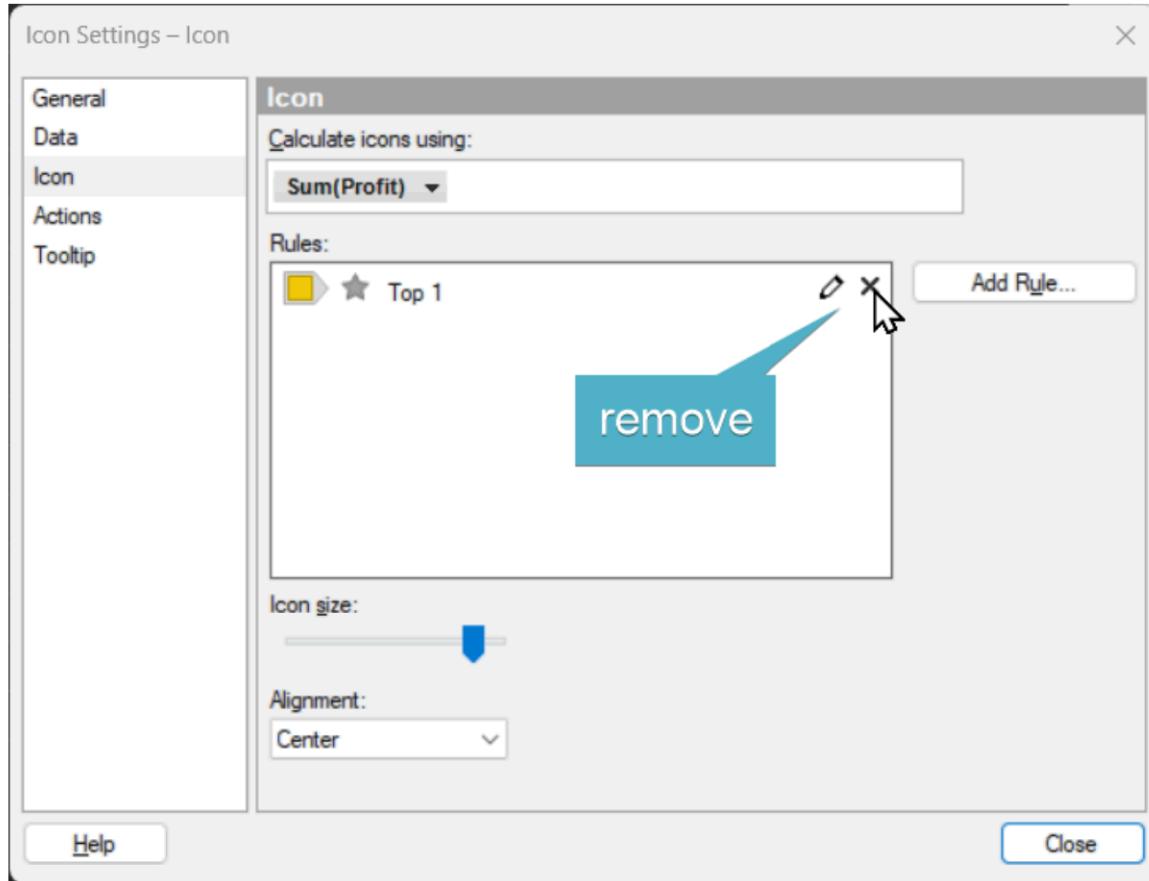
# Add Country as Calculated Column

Global Superstore - Orders	
Sub-Category	Country
Accessories	New Zealand
Appliances	Australia
Art	Iran
Binders	Sudan
Bookcases	United Kingdom
Chairs	New Zealand
Copiers	New Zealand
Envelopes	Iraq
Fasteners	Philippines
Furnishings	Australia
Labels	Australia
Machines	Canada
Paper	Australia
Phones	Guatemala
Storage	Hungary
Supplies	Australia
Tables	Philippines

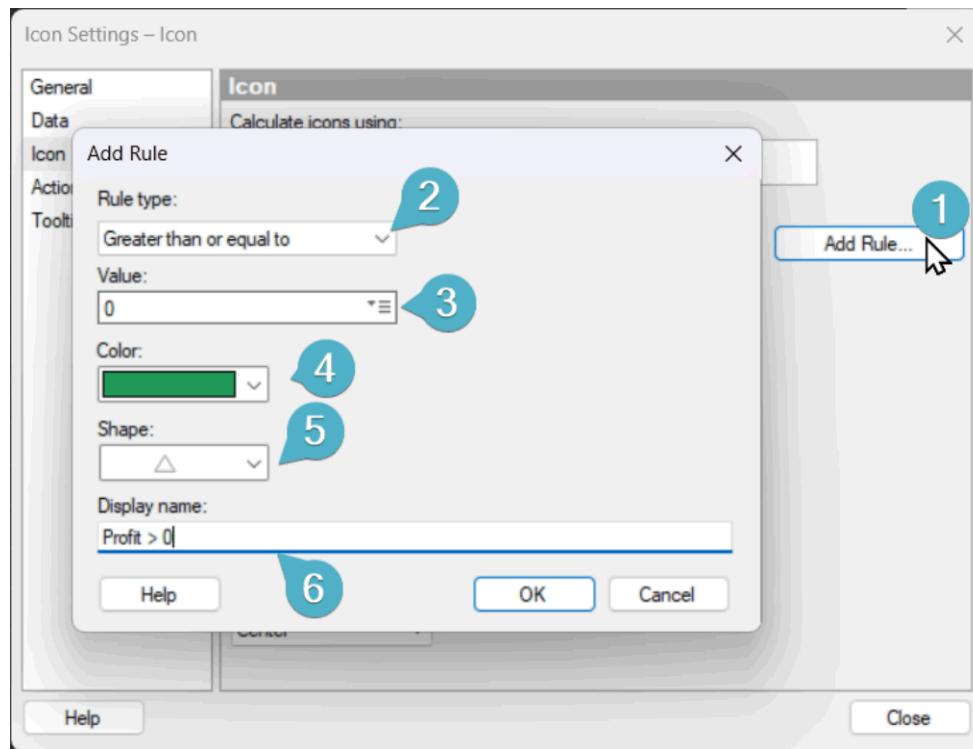
# Add Positive or Negative Profit as Icon



# Add Positive or Negative Profit as Icon

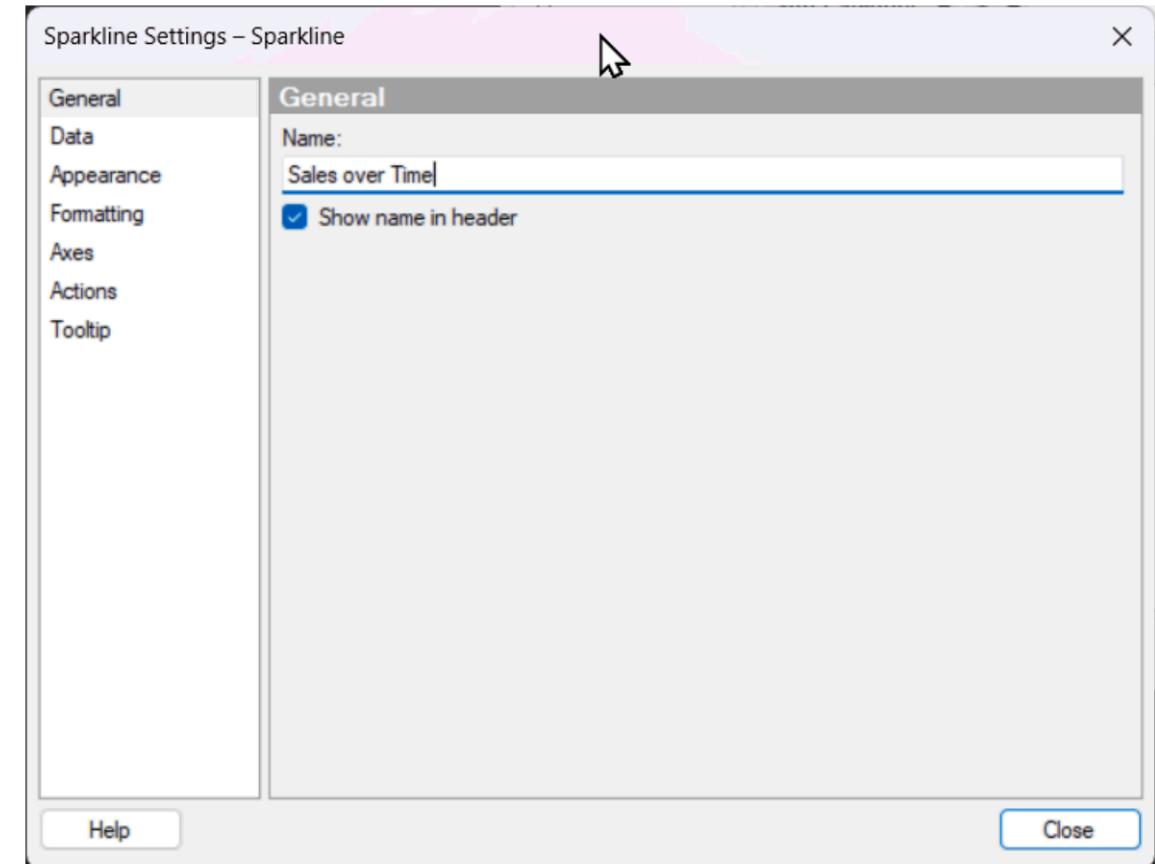
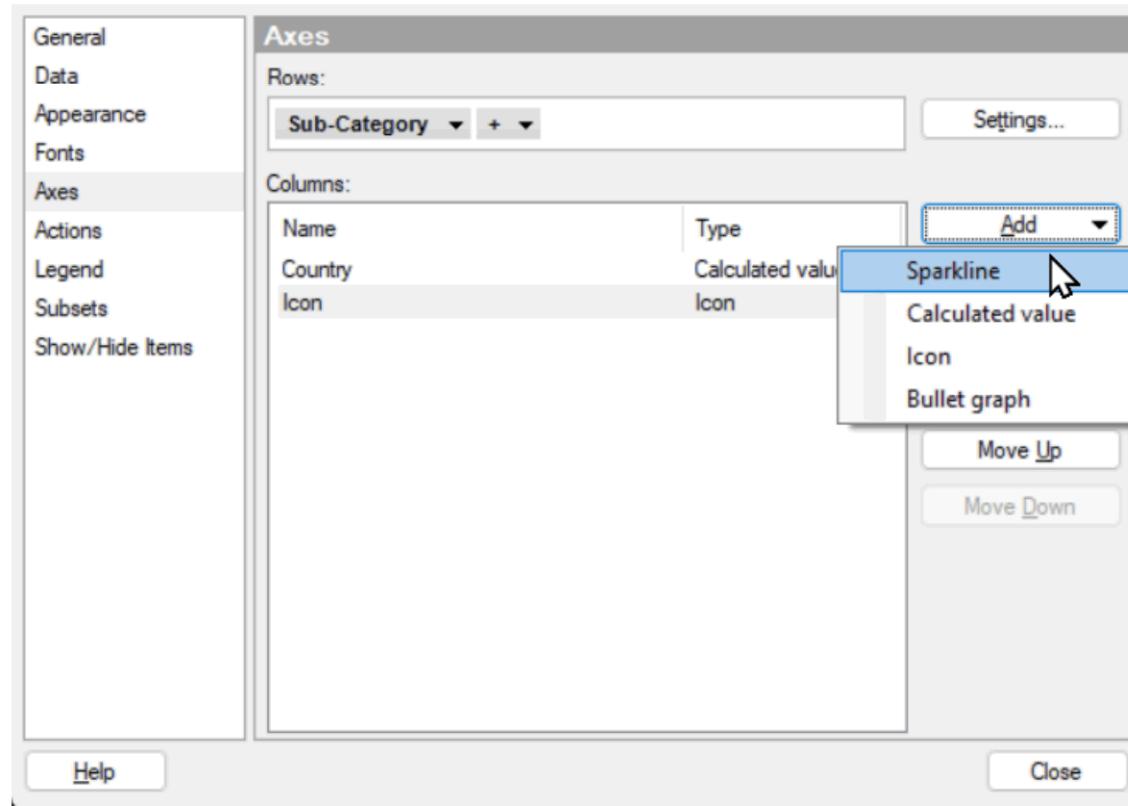


# Add Positive or Negative Profit as Icon

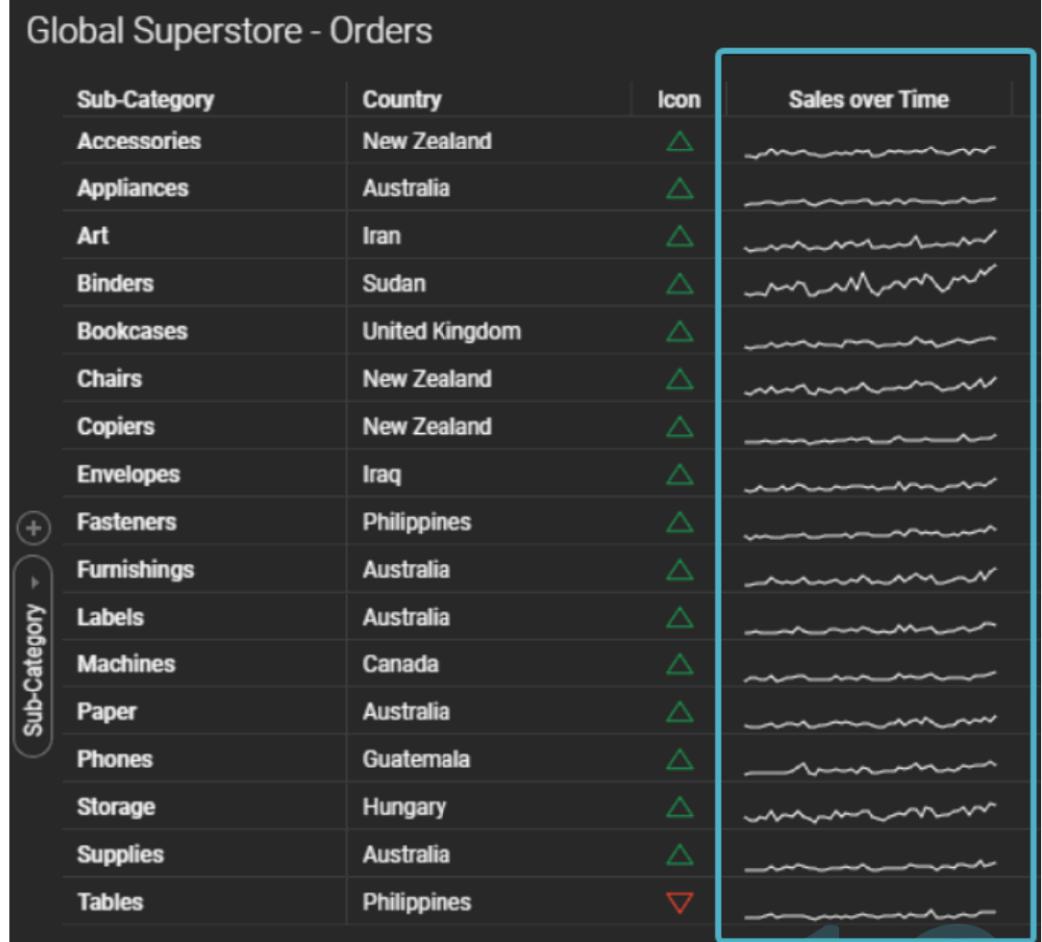
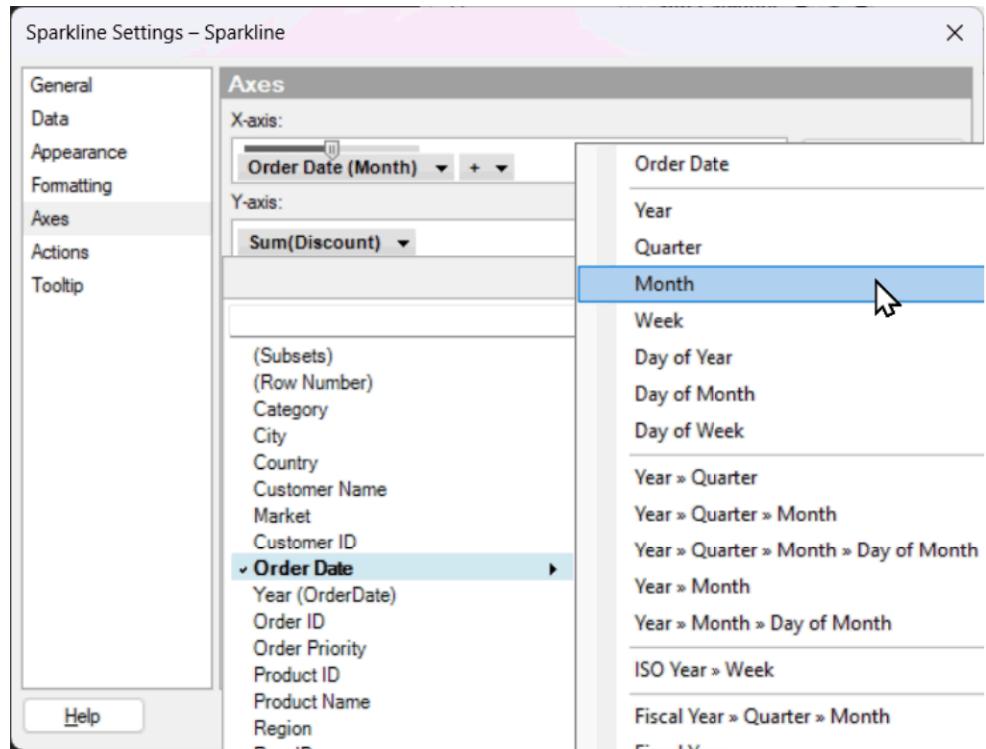


Global Superstore - Orders		
Sub-Category	Country	Icon
Accessories	New Zealand	▲
Appliances	Australia	▲
Art	Iran	▲
Binders	Sudan	▲
Bookcases	United Kingdom	▲
Chairs	New Zealand	▲
Copiers	New Zealand	▲
Envelopes	Iraq	▲
Fasteners	Philippines	▲
Furnishings	Australia	▲
Labels	Australia	▲
Machines	Canada	▲
Paper	Australia	▲
Phones	Guatemala	▲
Storage	Hungary	▲
Supplies	Australia	▲
Tables	Philippines	▼

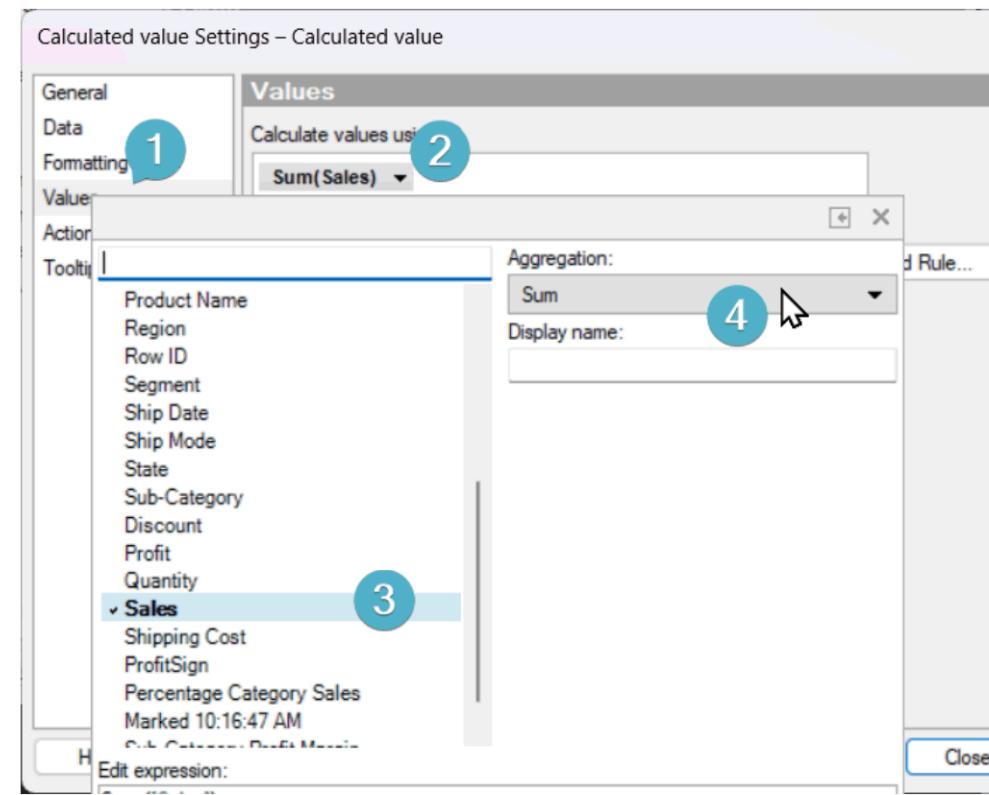
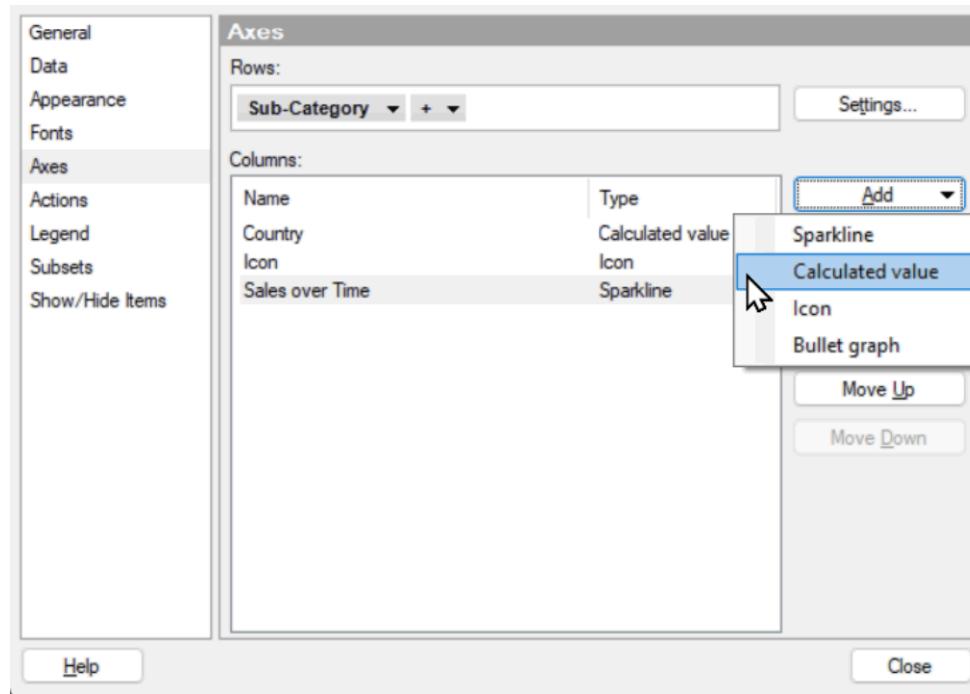
# Add Sales over Time as Sparkline



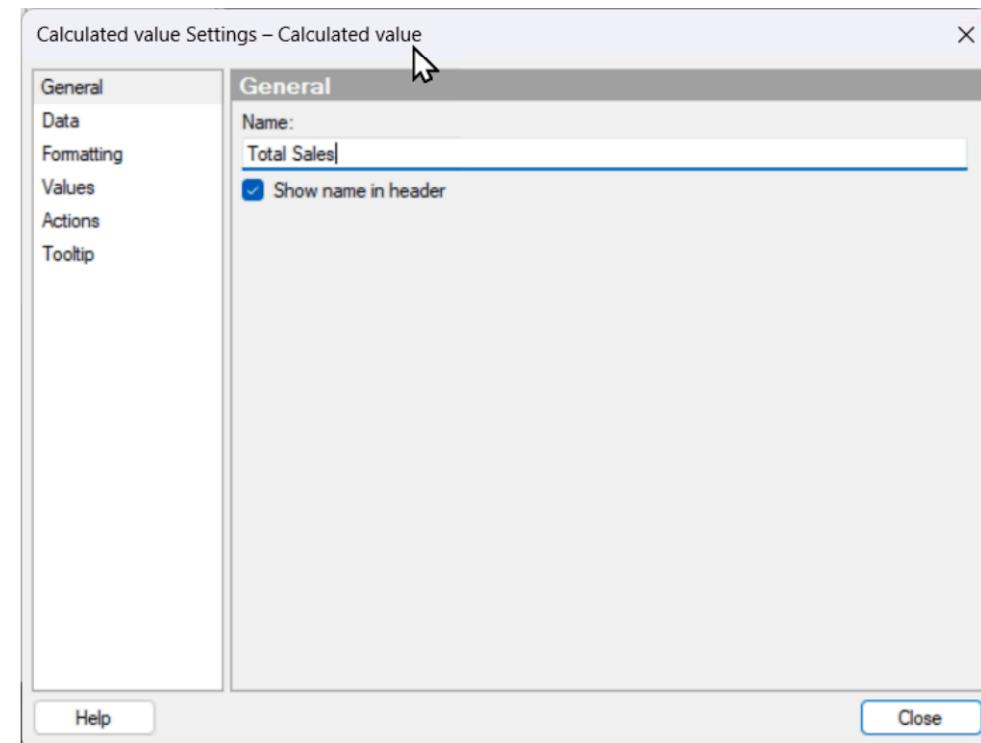
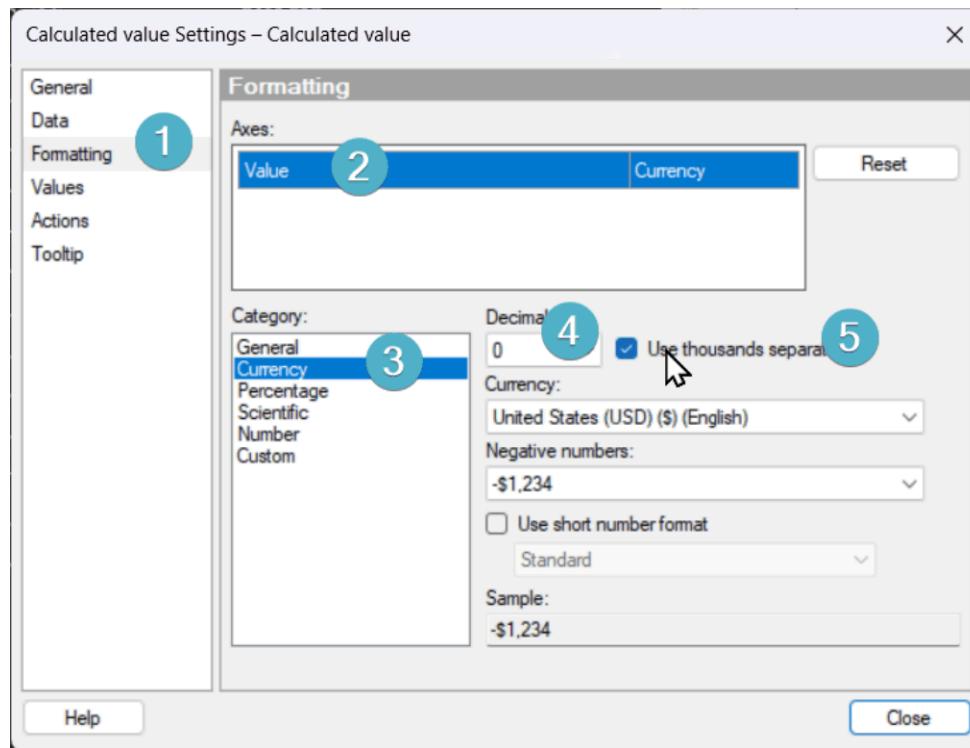
# Add Sales over Time as Sparkline



# Add Total Sales as Calculated Column

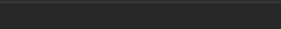
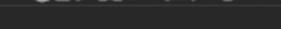
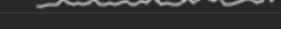


# Add Total Sales as Calculated Column

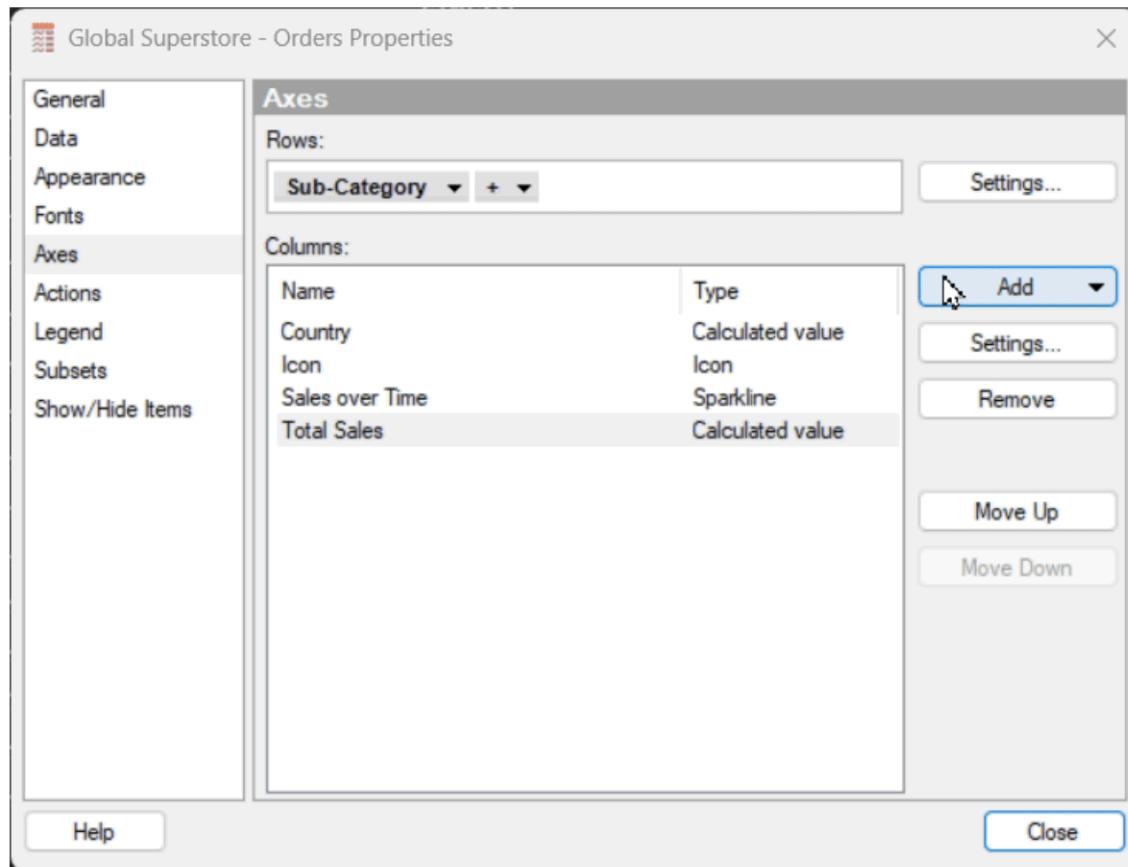


# Add Total Sales as Calculated Column

Notice Total Sales

Global Superstore - Orders				
Sub-Category	Country	Icon	Sales over Time	Total Sales
Accessories	New Zealand	▲		\$749,237
Appliances	Australia	▲		\$1,011,064
Art	Iran	▲		\$372,092
Binders	Sudan	▲		\$461,912
Bookcases	United Kingdom	▲		\$1,466,572
Chairs	New Zealand	▲		\$1,501,682
Copiers	New Zealand	▲		\$1,509,436
Envelopes	Iraq	▲		\$170,904
Fasteners	Philippines	▲		\$83,242
Furnishings	Australia	▲		\$385,578
Labels	Australia	▲		\$73,404
Machines	Canada	▲		\$779,060
Paper	Australia	▲		\$244,292
Phones	Guatemala	▲		\$1,706,824
Storage	Hungary	▲		\$1,127,086
Supplies	Australia	▲		\$243,074
Tables	Philippines	▼		\$757,042

# Add Total Profit as Calculated Column



Global Superstore - Orders				
Sub-Category	Country	Icon	Sales over Time	Total Sales
Accessories	New Zealand	▲		\$749,237
Appliances	Australia	▲		\$1,011,064
Art	Iran	▲		\$372,092
Binders	Sudan	△		\$461,912
Bookcases	United Kingdom	▲		\$1,466,572
Chairs	New Zealand	▲		\$1,501,682
Copiers	New Zealand	▲		\$1,509,436
Envelopes	Iraq	△		\$170,904
Fasteners	Philippines	▲		\$83,242
Furnishings	Australia	▲		\$385,578
Labels	Australia	▲		\$73,404
Machines	Canada	▲		\$779,060
Paper	Australia	▲		\$244,292
Phones	Guatemala	▲		\$1,706,824
Storage	Hungary	▲		\$1,127,086
Supplies	Australia	▲		\$243,074
Tables	Philippines	▼		\$757,042

# Add Sales vs Profit as Bullet Graph

The image shows two overlapping dialog boxes from a software application.

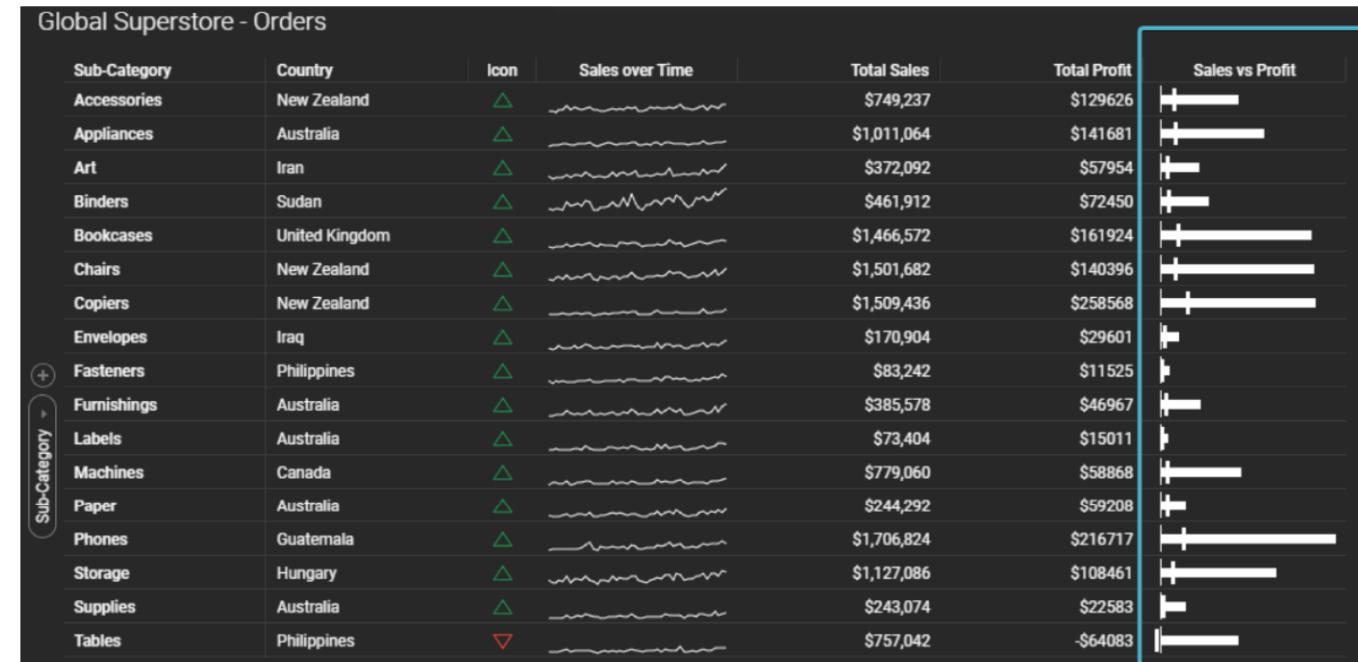
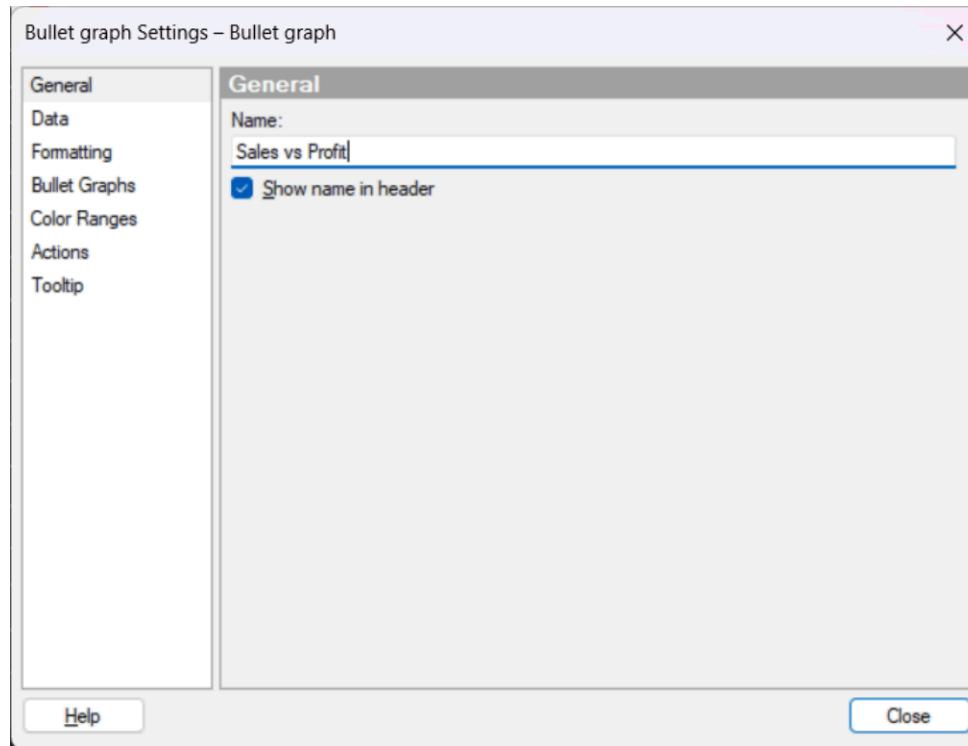
**Global Superstore - Orders Properties** (Left Dialog):

- Axes** tab selected.
- Rows:** Sub-Category
- Columns:** Name, Type:
  - Country: Calculated value
  - Icon: Icon
  - Sales over Time: Sparkline
  - Total Sales: Calculated value
  - Total Profit: Calculated value
- Add** dropdown menu:
  - Bullet graph (highlighted with a cursor)
  - Move Up
  - Move Down

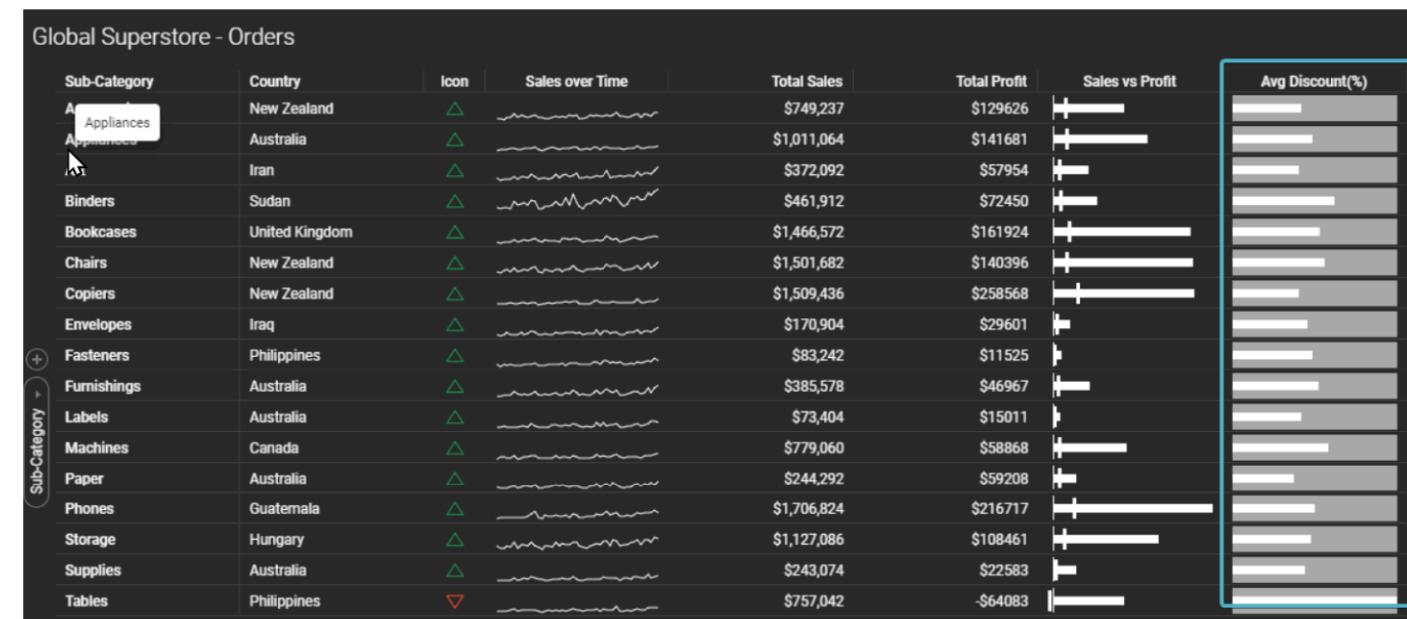
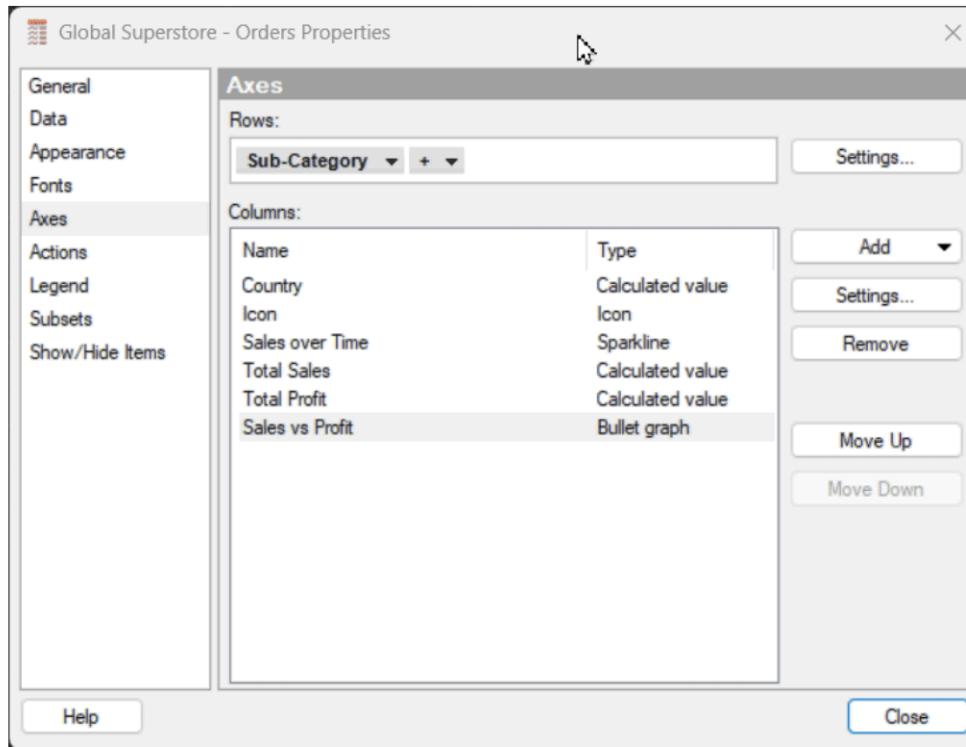
**Bullet graph Settings – Bullet graph** (Right Dialog):

- Bullet Graphs** tab selected (marked with 1).
- Calculate values using:** Sum(Sales) (marked with 2).
- Calculate comparative values using:** Sum(Profit) (marked with 3).
- Scaling:**
  - One scale for all bullet graphs in this column
  - Multiple scales
- Range**: Min: Automatic, Max: Automatic, Set to Current Range, Include origin (checked), Reverse scale (unchecked).

# Add Sales vs Profit as Bullet Graph



# Add Average Discount% As Bullet Graph



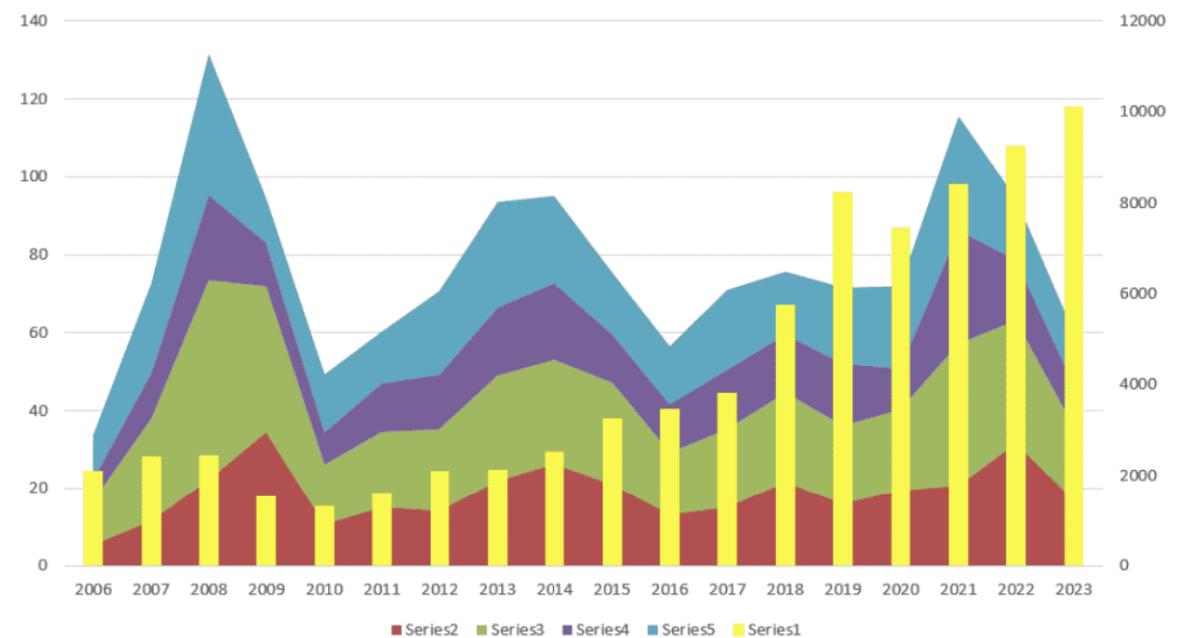
# Combination Charts

Combo Charts

# Combination Chart

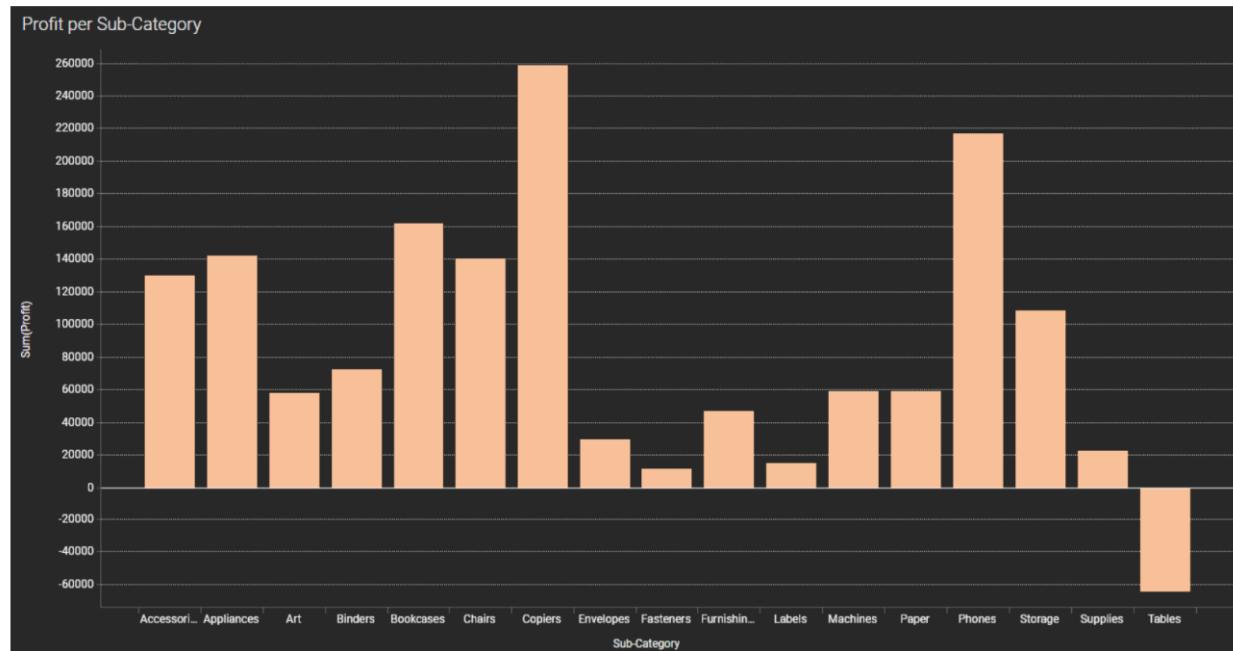
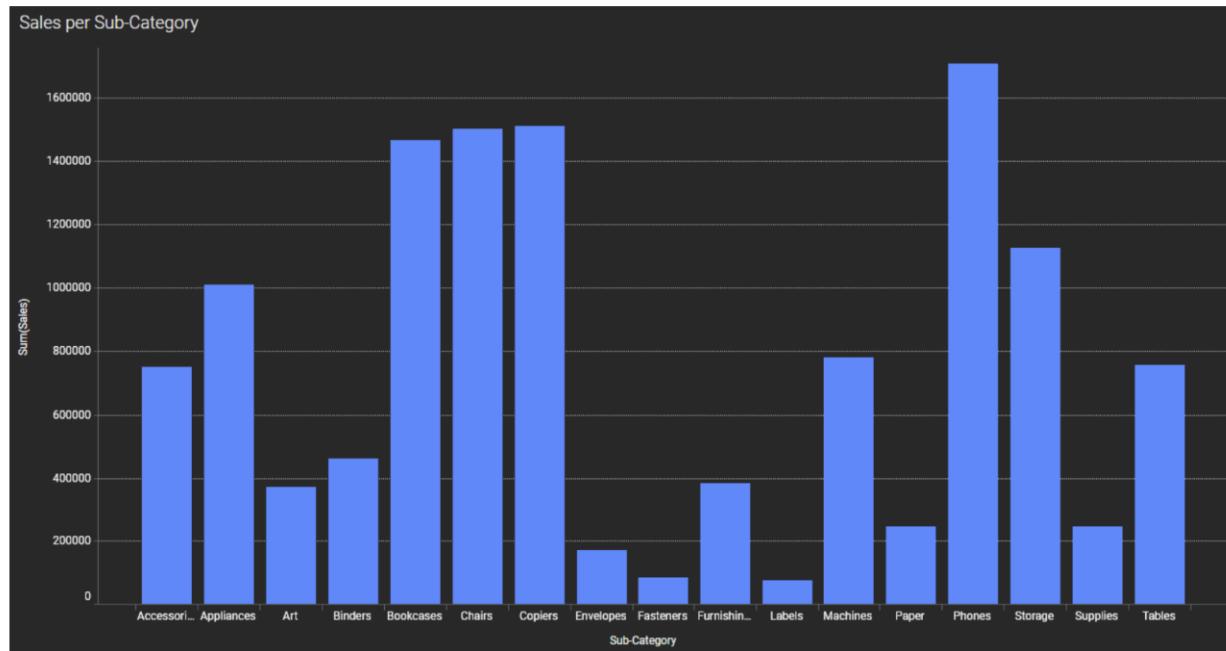
A **combo chart** (combination chart) merges two or more chart types (e.g., bar, line, area) into a **single chart**, typically:

- Sharing the **same X-axis** (e.g., time, categories)
- Using **dual Y-axes** for different metrics



Kosourova, E. (2024, June 12). *How to Create and Format a Combo Chart in Excel*. Datacamp.com; DataCamp. <https://www.datacamp.com/tutorial/mastering-combo-charts-in-excel>

# Sales vs Profit per Sub-Category



# Sales & Profit Per Sub-Category

## 🔍 Insights from the Chart

High Sales ≠ High Profit

*Tables & Bookcases: over-discounted? Costly to sell?*

Top Performers (High Sales & Profit)

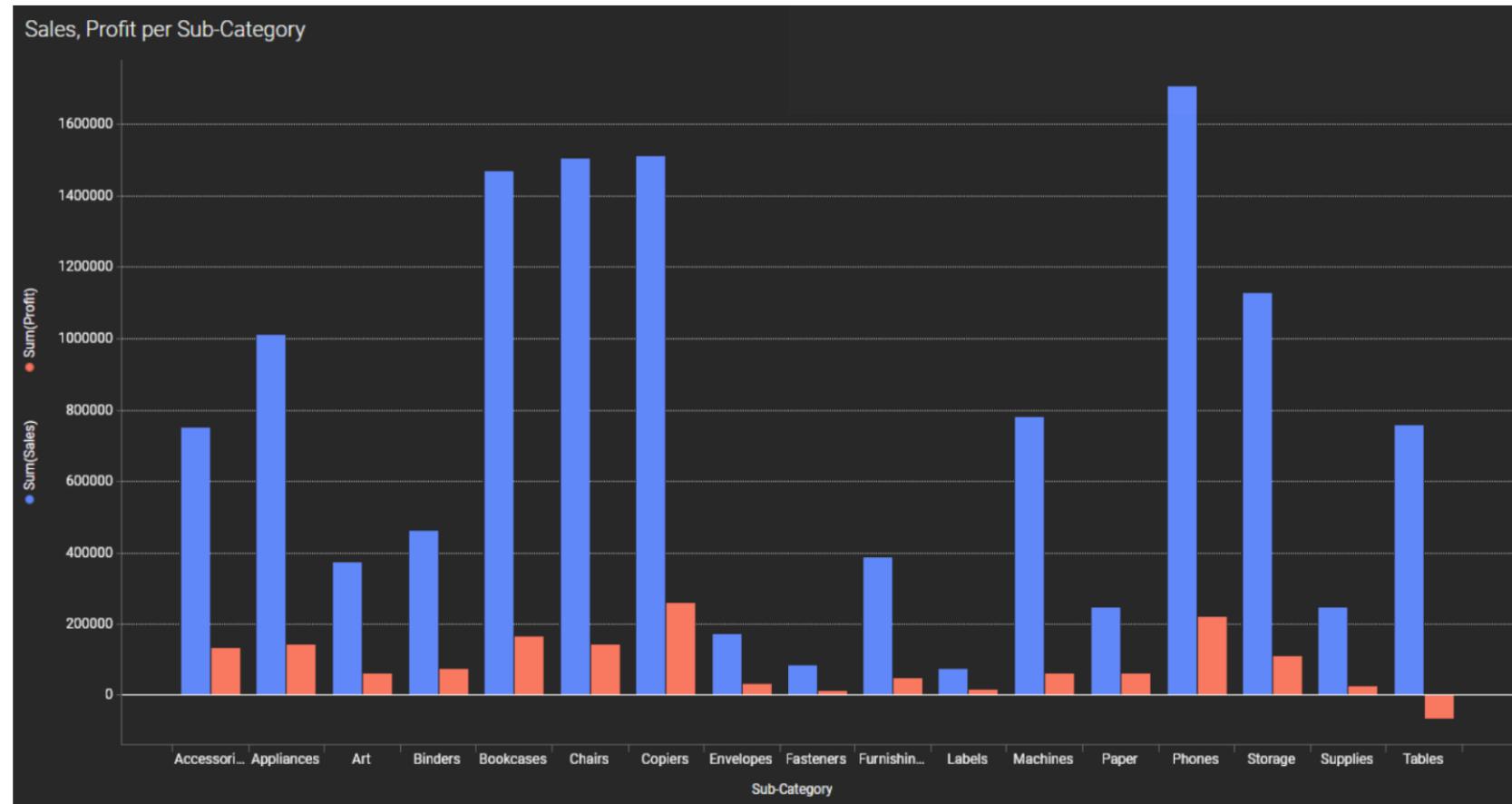
*Phones & Chairs*

Loss-Making Products

*Tables*

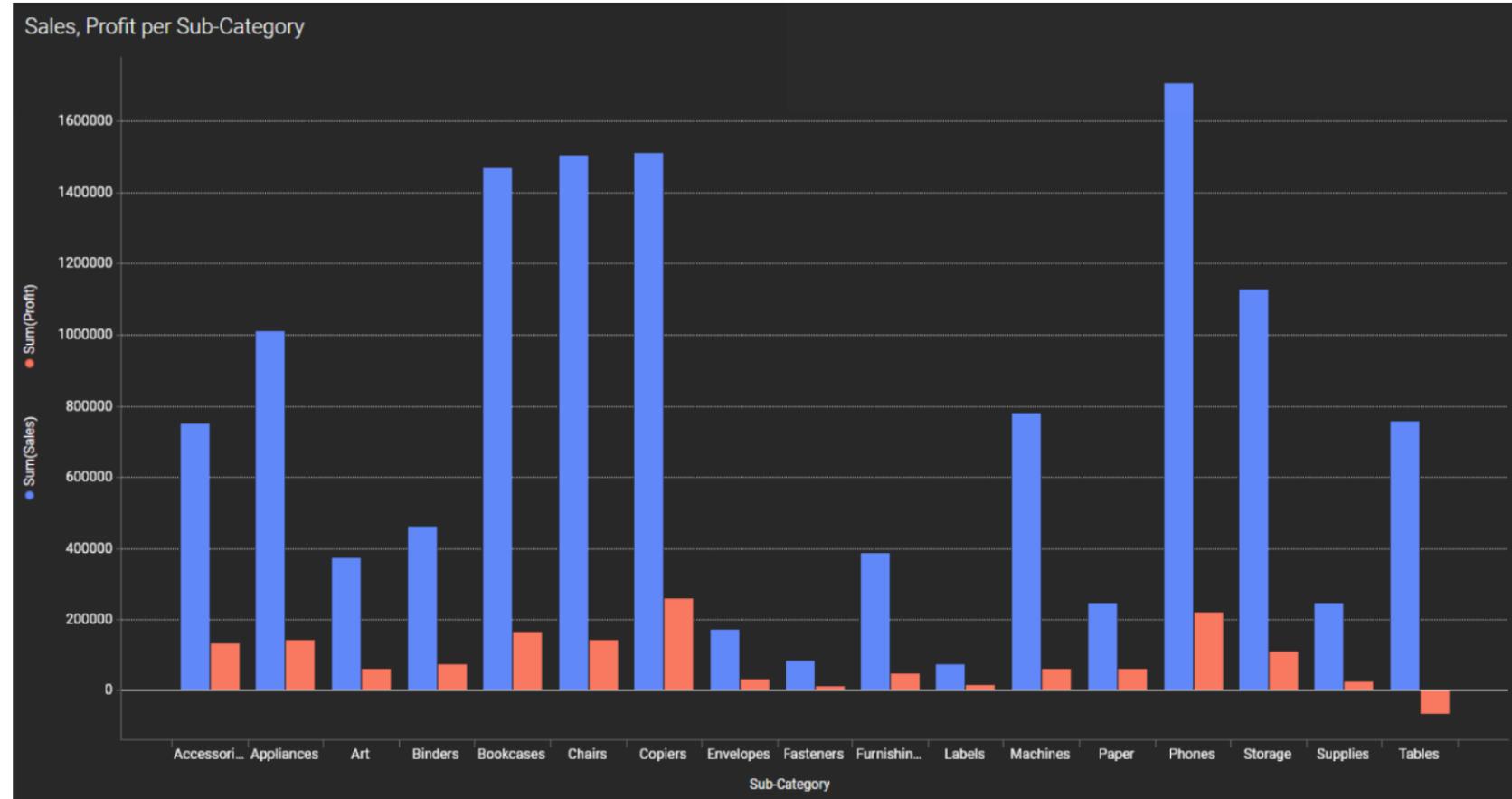
Low Sales, Low Profit Segments

*Labels, Fasteners & Envelopes: low-value or niche items.*



# Sales & Profit Per Sub-Category

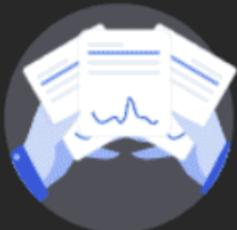
- 🧭 Business Recommendations
  - Investigate Loss-makers
  - Invest Bestsellers
  - Review Underperformance



Side by Side × Page × | Line Charts × | Page (2) × | Page (3) | < > ⌂ + | ⌂

# Your data is ready!

How do you want to continue?



Start from data  
Select data of interest and let us recommend visualizations



Start from visualizations  
Select a visualization type and continue from there



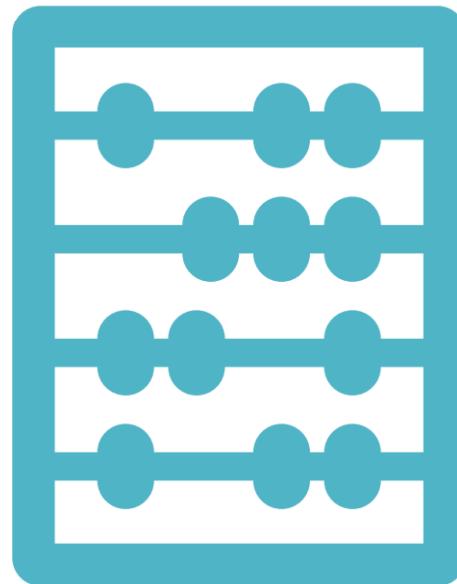
Explore by searching  
Search to explore your data and create visualizations

# Calculated Column



# Calculated Column

- Stored physically in the data model
- Calculated once **per row** during data load
- Suitable for **derived fields** like Profit = Sales - Cost



# Profit Margin over Sub-Categories

The screenshot shows the Spotfire interface with the title bar "Global Superstore - Orders.dxp - Spotfire". The left sidebar includes icons for Home, File, Edit, Data, Visualizations, View, Tools, Help, User, and a search bar. The main area displays the "Data tables" section with a table named "Global Superstore - Orders". A context menu is open over a calculated column, with numbered callouts pointing to specific items:

- 1: A blue circle on the far left of the interface.
- 2: "Add calculated column" button.
- 3: "Add binned column" button.
- 4: "Add hierarchy" button.

The table properties pane shows:

- Added calculated column: ProfitSign
- Added calculated column: Percentage
- Category Sales
- Added column for filtered items: Marked 10:16:47 AM

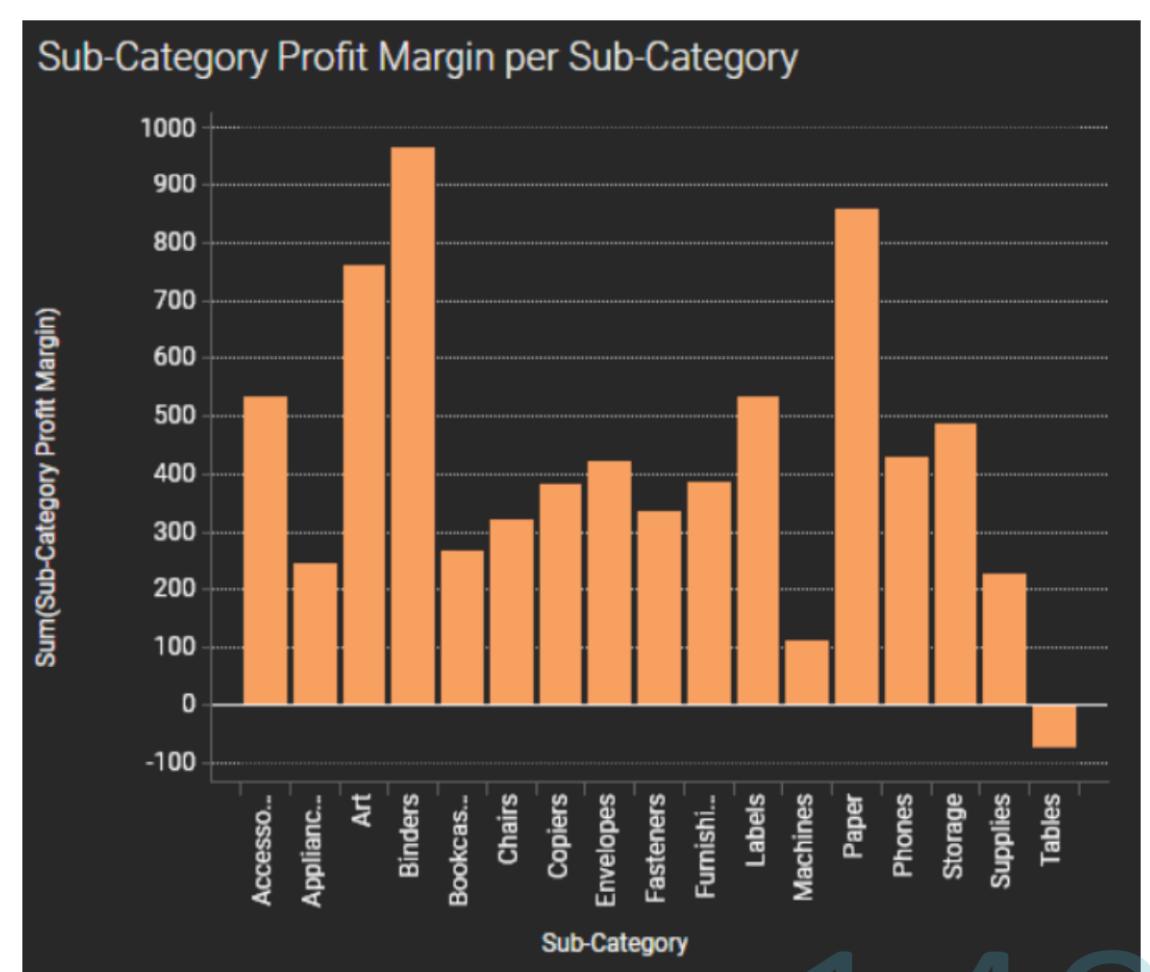
Information about the table:

- 51290 rows
- 27 columns

The screenshot shows the "Edit calculated column" dialog box with the following sections:

- Available columns:** A list of available columns with their names and data types.
- Available properties for column:** A list of properties for the selected column, including Name, Data Type, Property, and Value.
- Functions Category:** A dropdown menu showing "All functions".
- Function:** A list of mathematical operators: +, -, \*, /, %, &, =, and >=.
- Example:** Shows the expression  $3.5 + 2.5$ .
- Expression:** The current expression being built: `1 Sum({Profit}) OVER [Sub-Category] / Sum({Sales}) OVER [Sub-Category]`.
- Recent expressions:** Shows the recent expression: `Sum(Real([Profit])) / Sum(Real([Sales])) AS [Profit Margin]`.
- Resulting expression:** Displays "Not applicable."
- Column name:** Shows "Sub-Category Profit Margin".
- Sample result:** Shows "(Empty)".
- Type:** Shows "Real".
- Buttons:** OK and Cancel.

Percent...	Sub-Category Profit Margin
29.96	0.10
29.96	0.10
32.52	0.12
29.96	0.24
29.96	0.09
29.96	0.24
37.53	0.08
32.52	0.09
32.52	0.09
37.53	0.17
37.53	0.17
29.96	0.09
32.52	0.11
37.53	0.13
29.96	0.10
32.52	-0.08
36.26	0.14



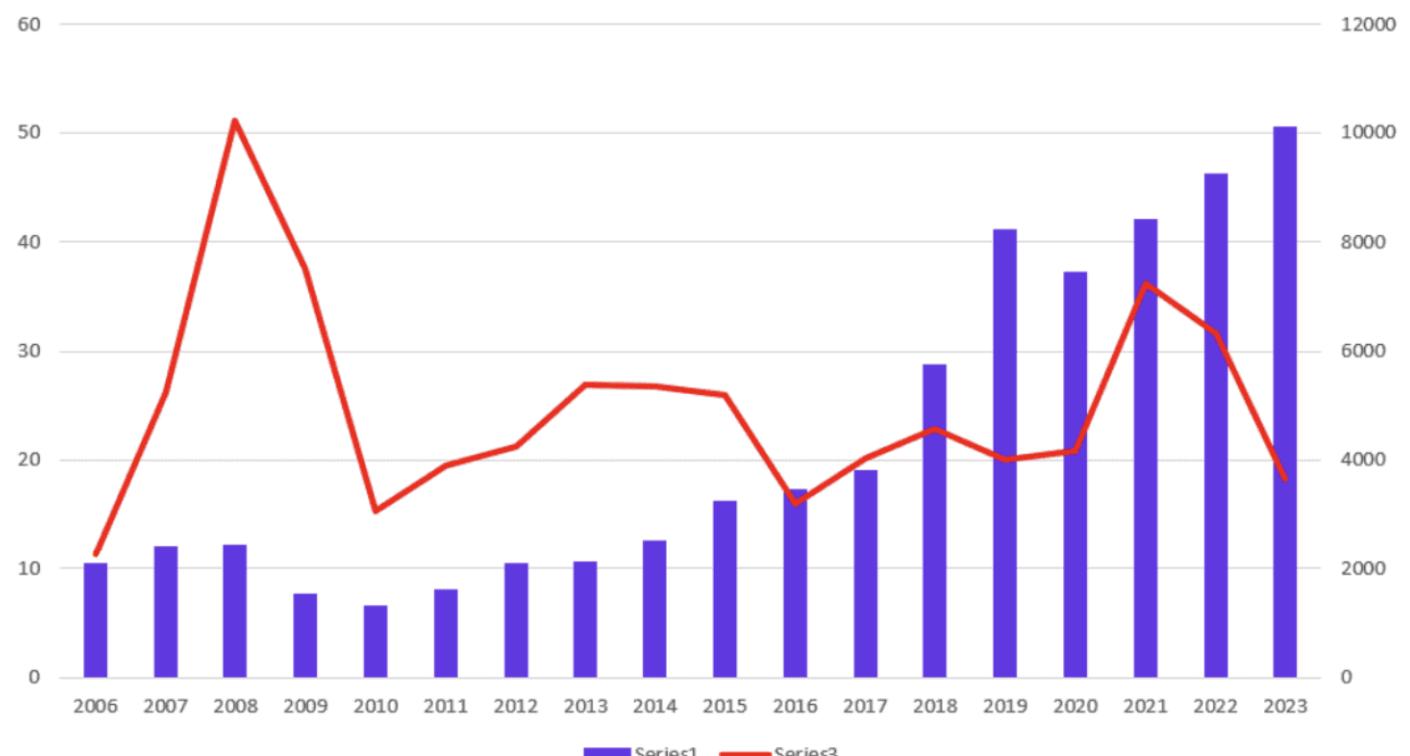
# Dual Axis

Using More than One Scale on an Axis

# Dual Axis

Dual axis allows you to plot two different scales (Y-axes) on the same X-axis in a single chart.

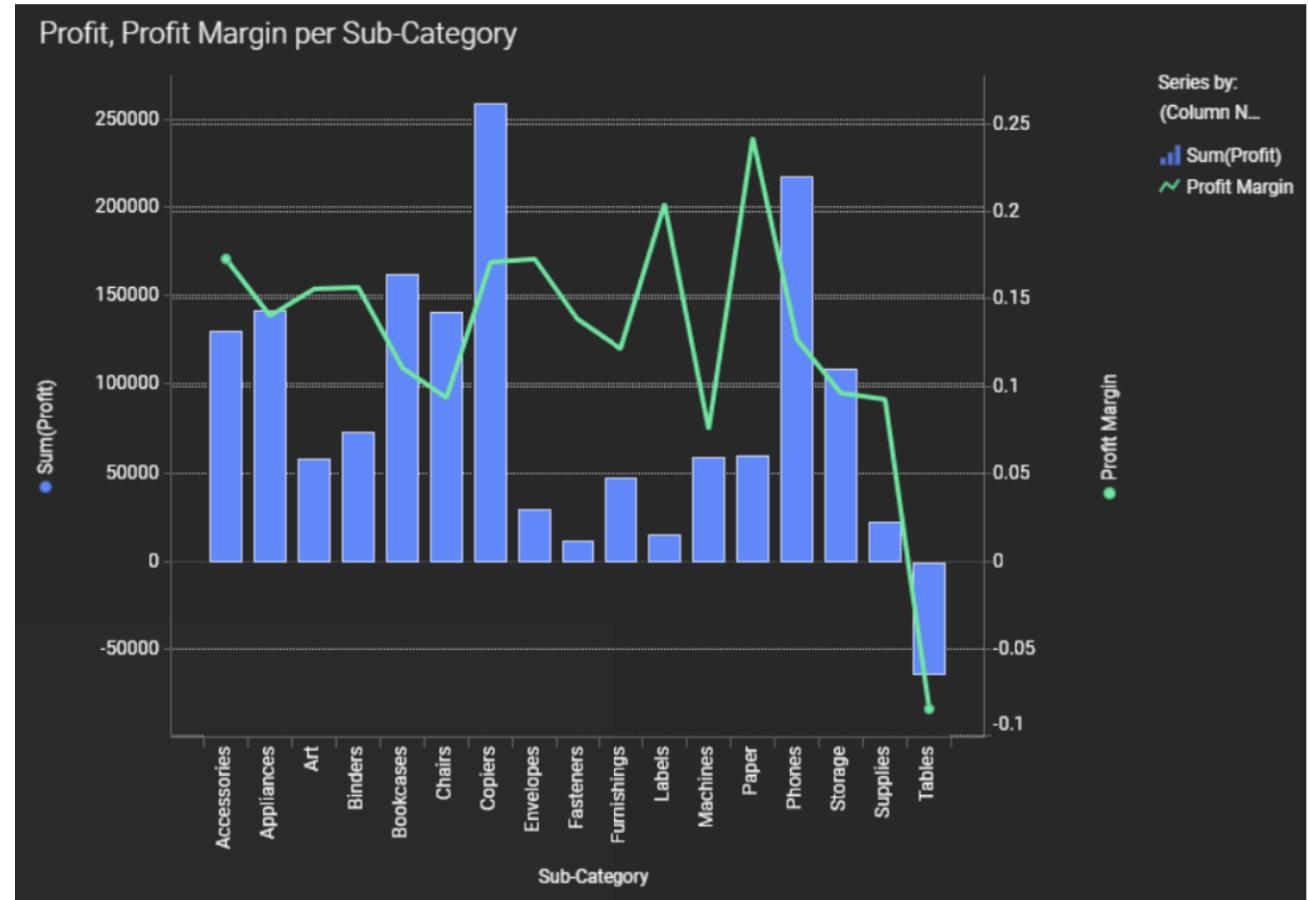
This is useful when comparing two different metrics with different units or magnitudes, but over the same category



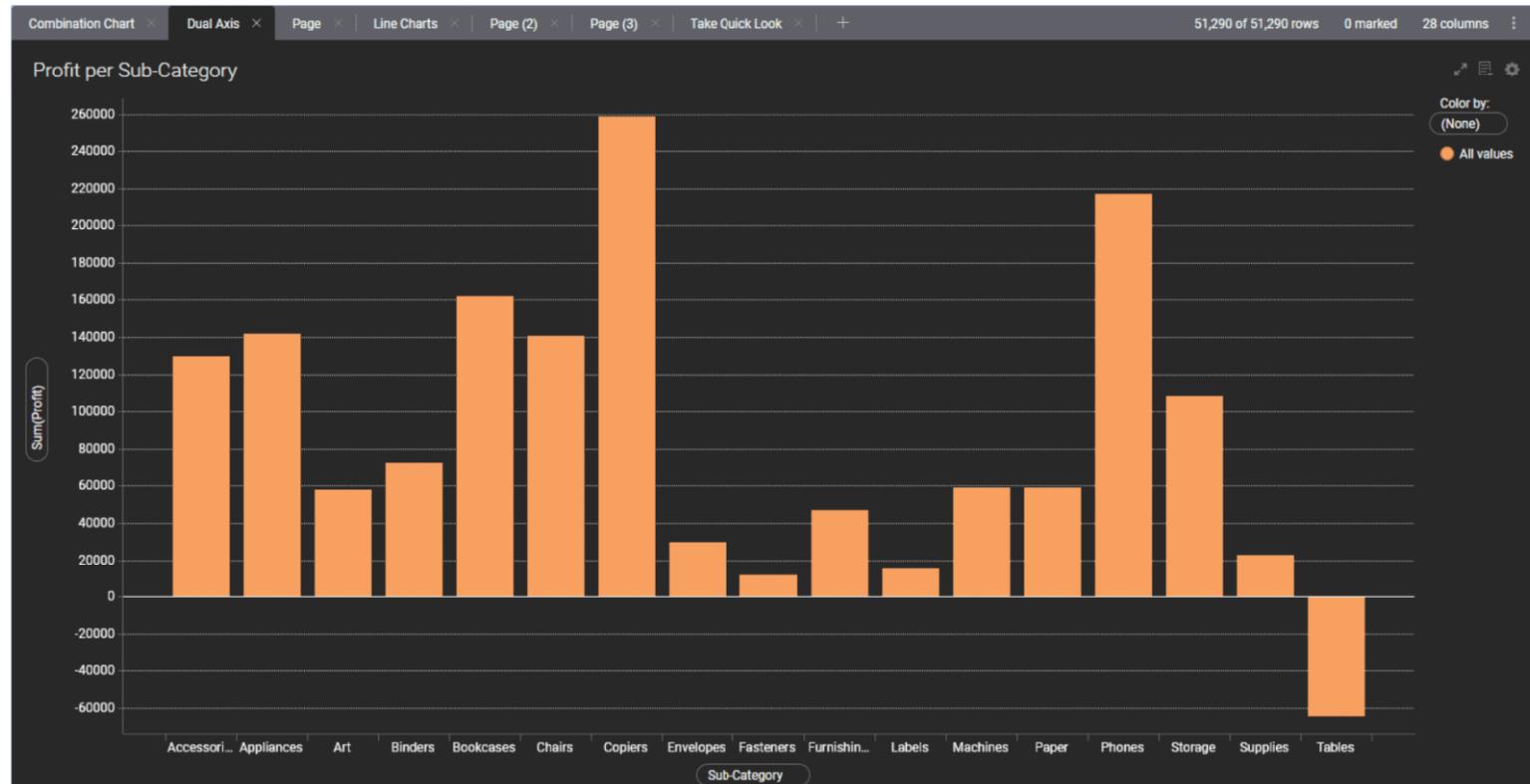
Kosourova, E. (2024, June 12). *How to Create and Format a Combo Chart in Excel*. Datacamp.com; DataCamp. <https://www.datacamp.com/tutorial/mastering-combo-charts-in-excel>

# Profit vs Profit Margin per Sub-Category

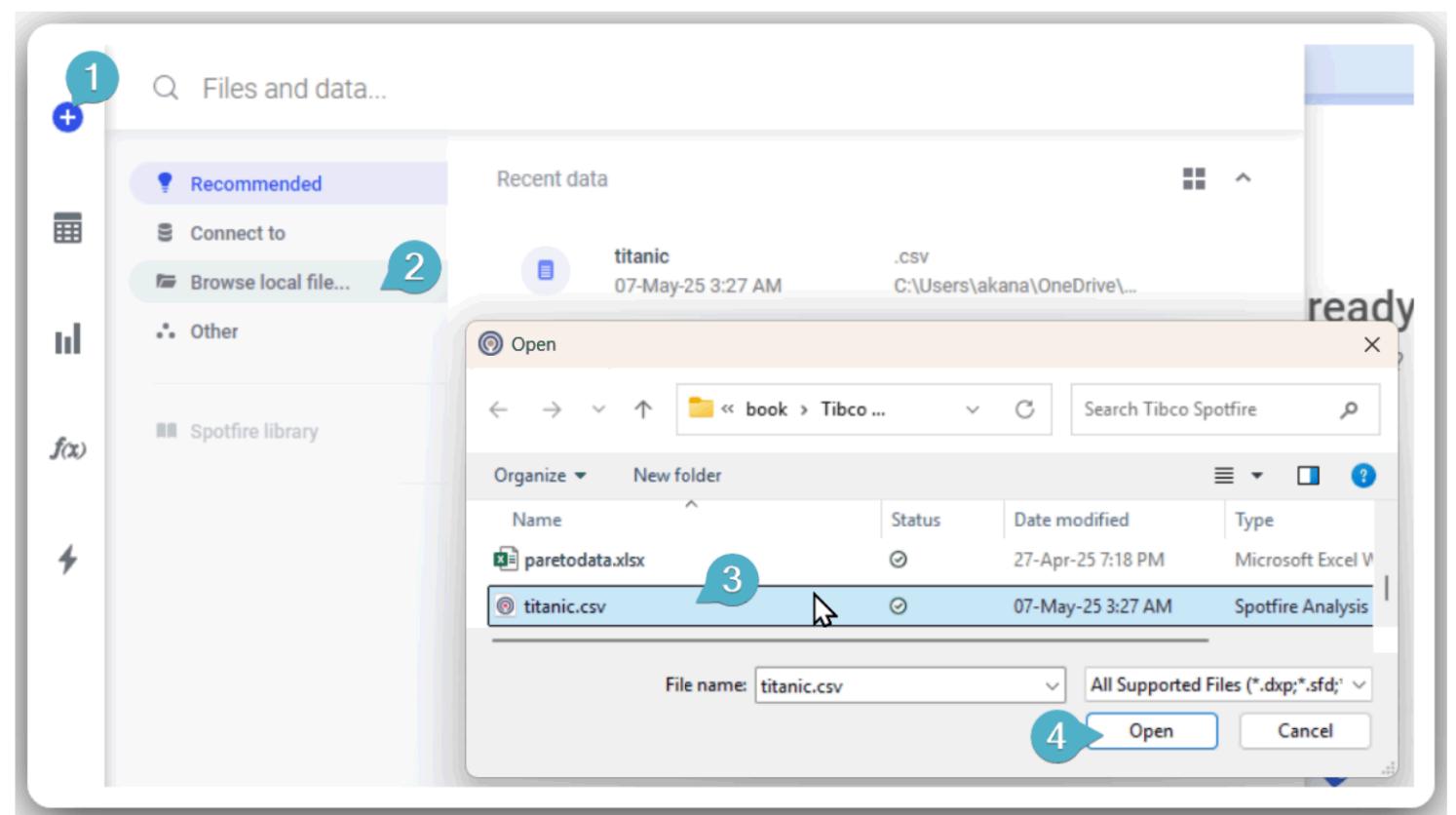
Using profit margin across sub-categories is essential for better decision-making because it tells you how efficiently each product category is turning sales into profit — not just how much it sells.



# Profit vs Profit Margin per Sub-Category



# Drilling into the data – details visualizations



# Take a quick look into Data Structure

titanic

Column	Sum	Avg	Min	Max	Median
PassengerId	397386.00	446.00	1	891	446.00
Survived	342.00	0.38	0	1	0.00
Pclass	2057.00	2.31	1	3	3.00
Age	21205.17	29.70	0.42	80.00	28.00
SibSp	466.00	0.52	0	8	0.00
Parch	340.00	0.38	0	6	0.00
Fare	28693.95	32.20	0.00	512.33	14.45

## Attribute Description

PassengerId:

Survived: Indicates whether the passenger survived (1) or not (0).

Pclass: Passenger's ticket class, used as a proxy for socio-economic status: 1st (upper), 2nd (middle), 3rd (lower).

Sex: Passenger's gender (male or female).

Age: Age of the passenger in years. Ages below 1 are fractional (e.g., 0.5). Estimated ages may be rounded to .5.

SibSp: Number of siblings or spouses the passenger had aboard the Titanic. Siblings include full/step-siblings; spouses include husbands and wives (excluding fiancés or mistresses).

Parch: Number of parents or children the passenger had aboard. Parents = mother/father; children = sons/daughters including stepchildren. Nannies are not counted.

Ticket: Passenger's ticket number (alphanumeric; may reflect booking group).

Fare: Fare paid for the ticket. May apply to one or more people traveling together.

Cabin: Cabin number assigned to the passenger. Some values may be missing.

Embarked: Port of embarkation where the passenger boarded: C = Cherbourg, Q = Queenstown, S = Southampton.

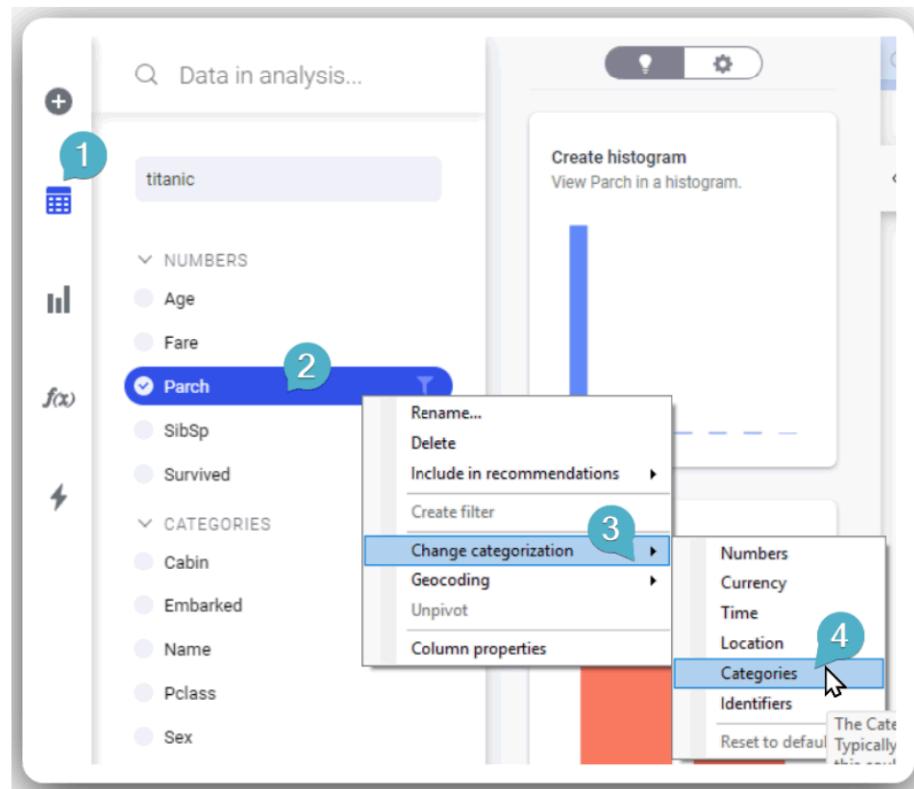
source: <https://www.kaggle.com/c/titanic/data>

# Take a quick look into Data Structure

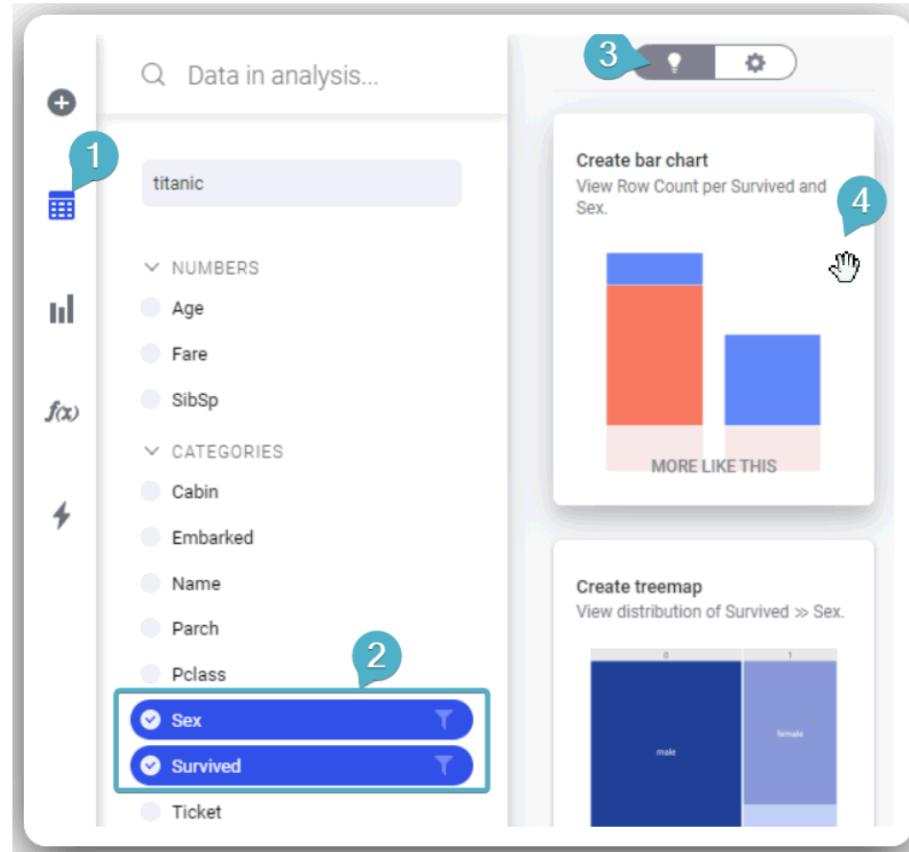
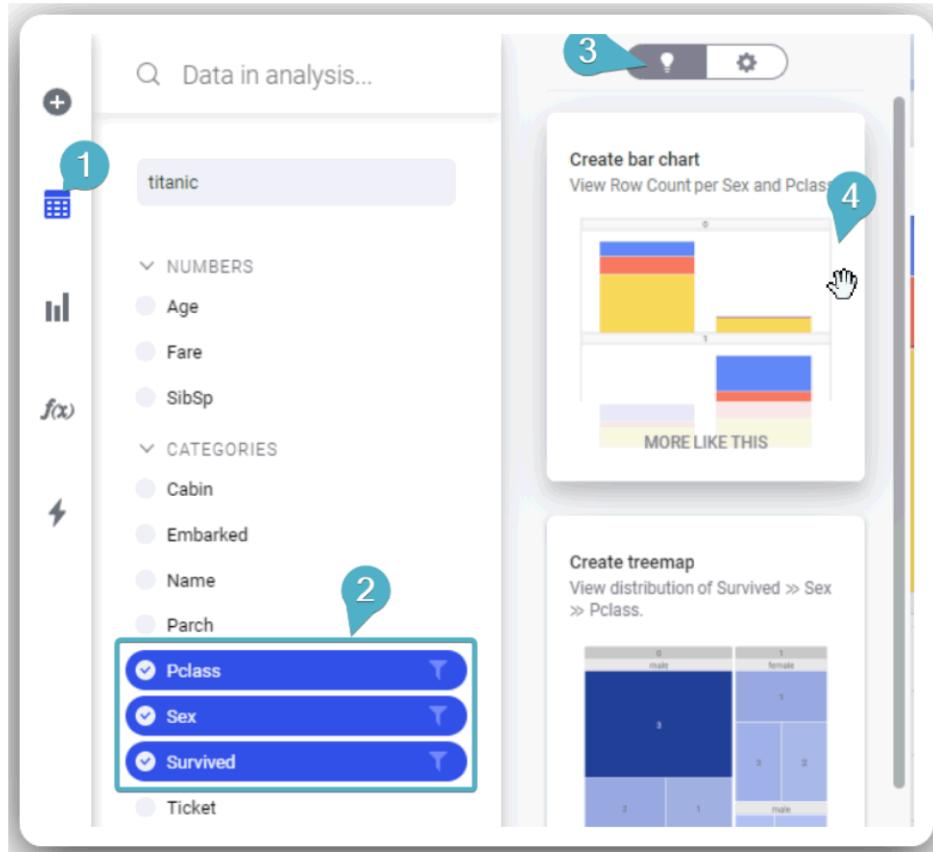
titanic						
PassengerId	Survived	Pclass	Name	Sex	Age	
1	0	3	Braund, Mr. O...	male	22.00	
2	1	1	Cumings, Mrs....	female	38.00	
3	1	3	Heikkinen, Mis...	female	26.00	
4	1	1	Futrelle, Mrs. ...	female	35.00	
5	0	3	Allen, Mr. Willi...	male	35.00	
6	0	3	Moran, Mr. Ja...	male		
7	0	1	McCarthy, Mr. ...	male	54.00	
8	0	3	Palsson, Mast...	male	2.00	
9	1	3	Johnson, Mrs....	female	27.00	
10	1	2	Nasser, Mrs. ...	female	14.00	
11	1	3	Sandstrom, Mi...	female	4.00	
12	1	1	Bonnell, Miss. ...	female	58.00	
13	0	3	Saundercock, ...	male	20.00	
14	0	3	Andersson, Mr...	male	39.00	
15	0	3	Vestrom, Miss...	female	14.00	
16	1	2	Hewlett, Mrs. (...	female	55.00	
17	0	3	Rice, Master. ...	male	2.00	
18	1	2	Williams, Mr. ...	male		
19	0	3	Vander Planke...	female	31.00	
...	...	...	...	...		

# Change Categorization

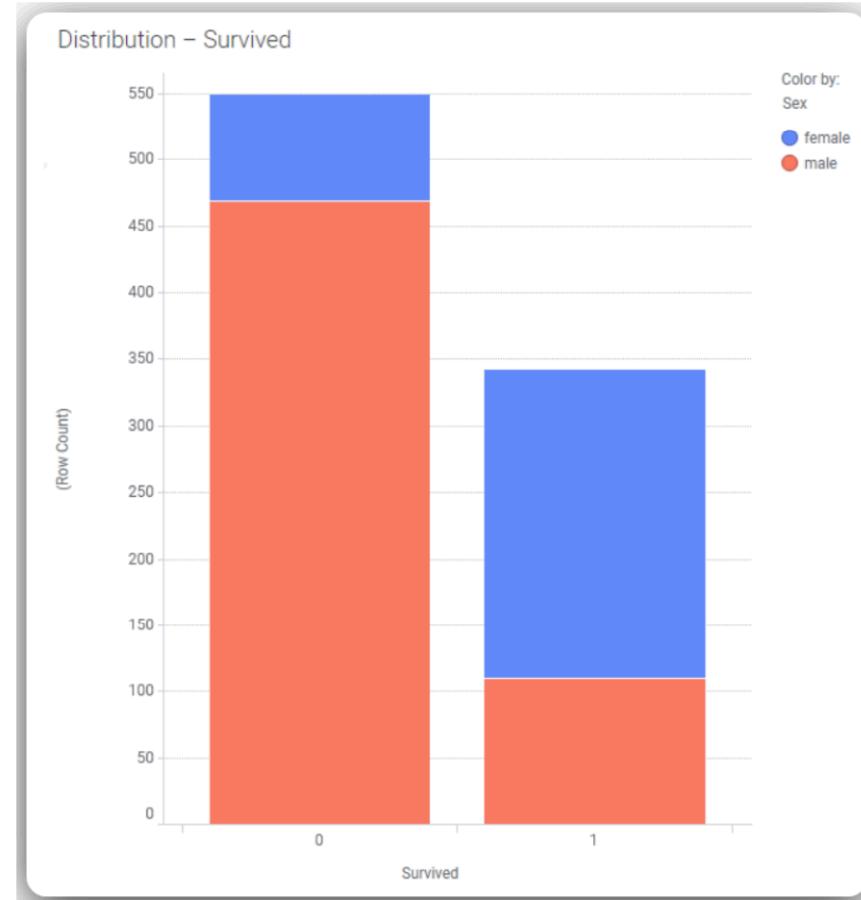
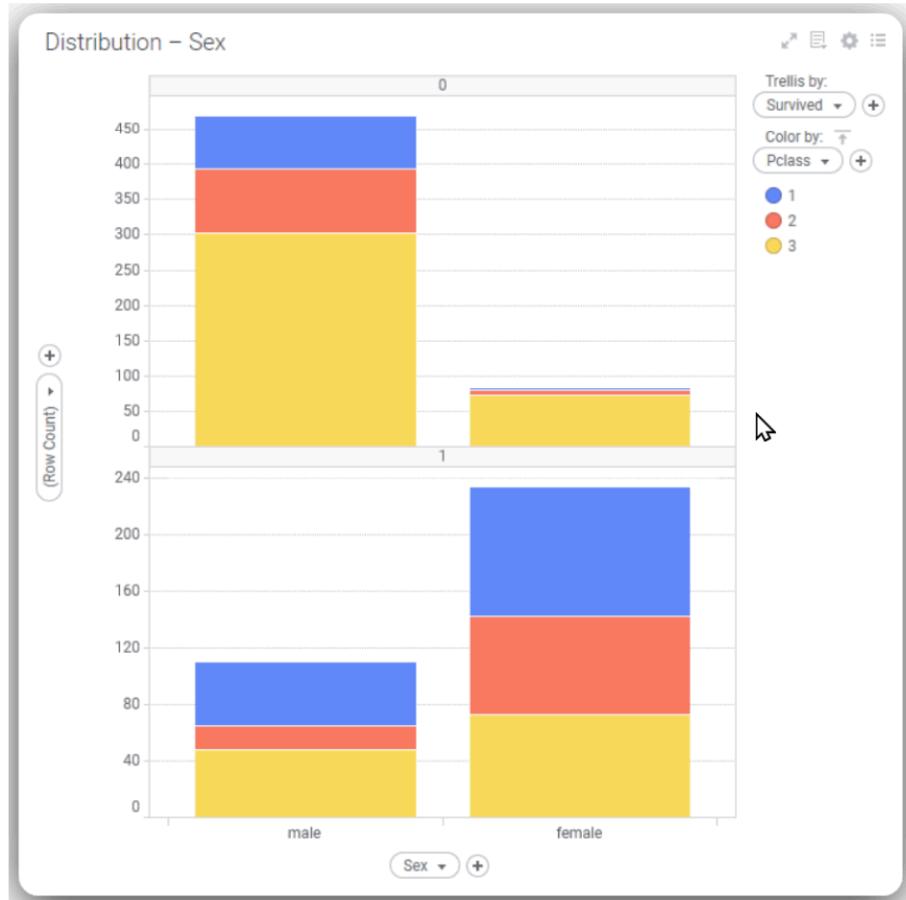
Do the same with Survived column



# Chart Suggestion

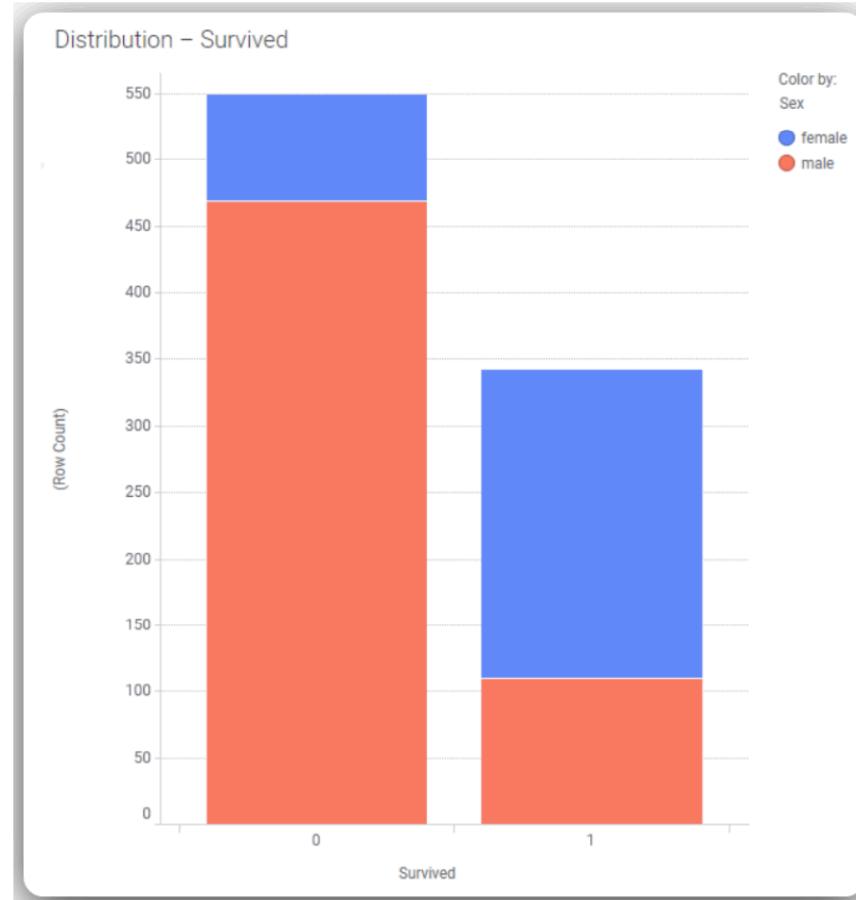


# Chart Suggestion



# Chart Suggestion

- ❖ Notice the disparity in gender
- ❖ More men on board than women.
- ❖ Higher male casualties: "women and children first" policy for lifeboat boarding



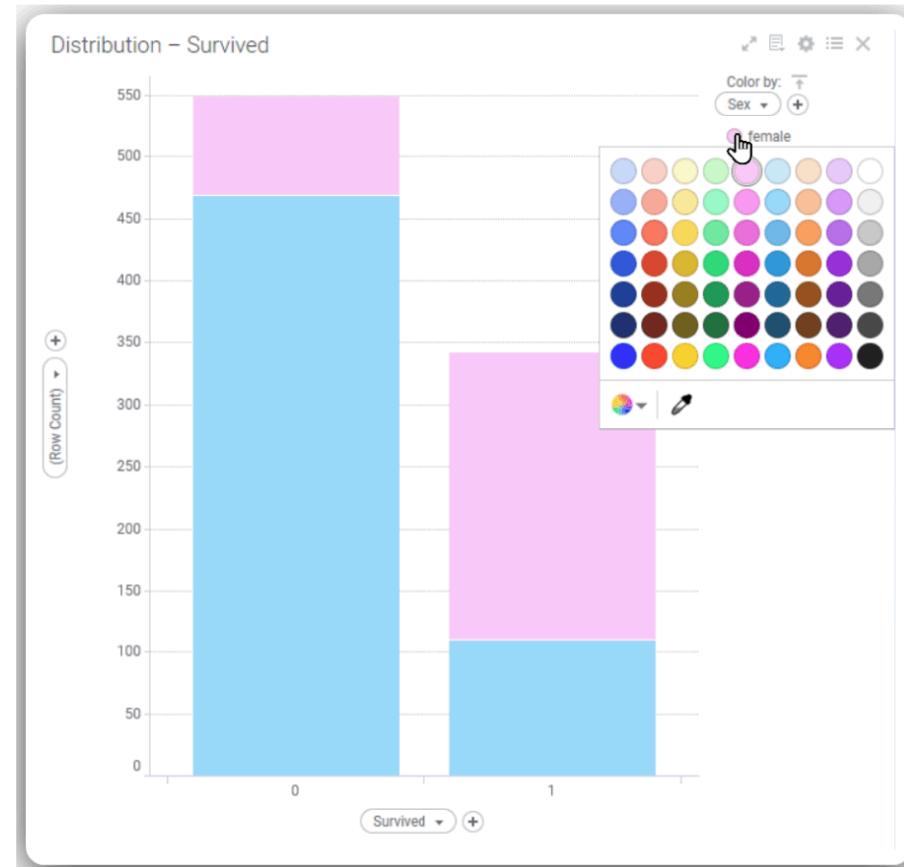
# Chart Suggestion

- ❖ Chart trellised by survival.
- ❖ First-class females had the highest survival chance.
- ❖ More third-class males survived than first- or second-class males (role of class and behaviors)

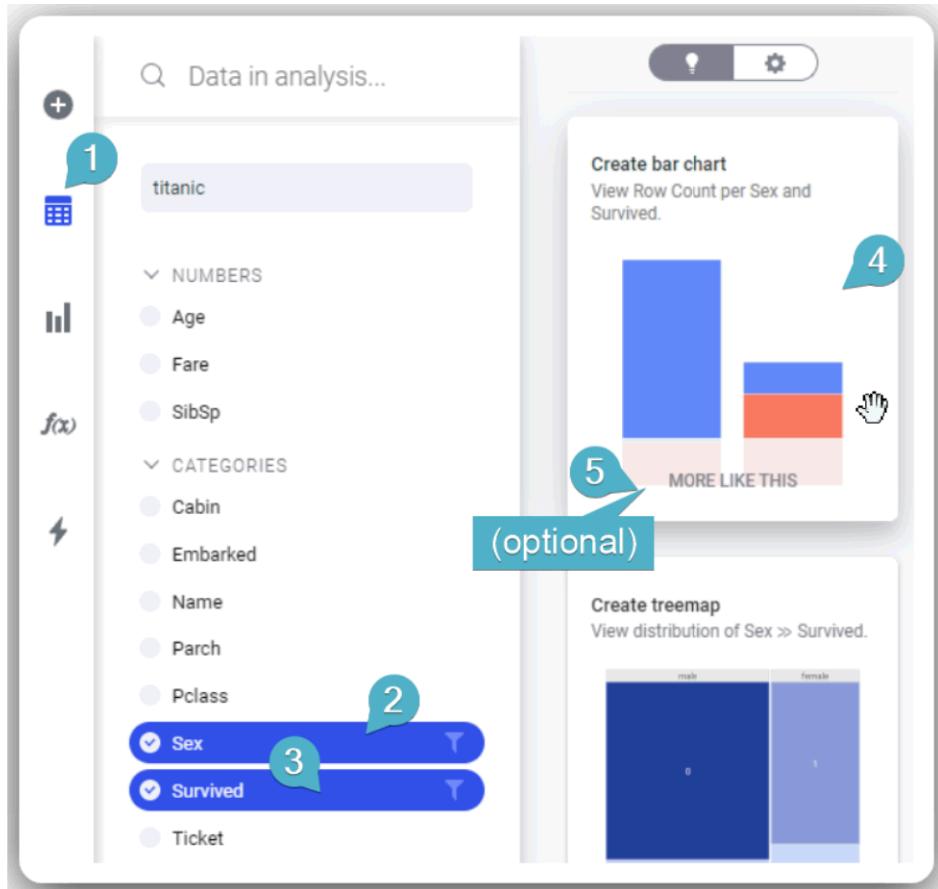


# Customizing Bar Chart Colors

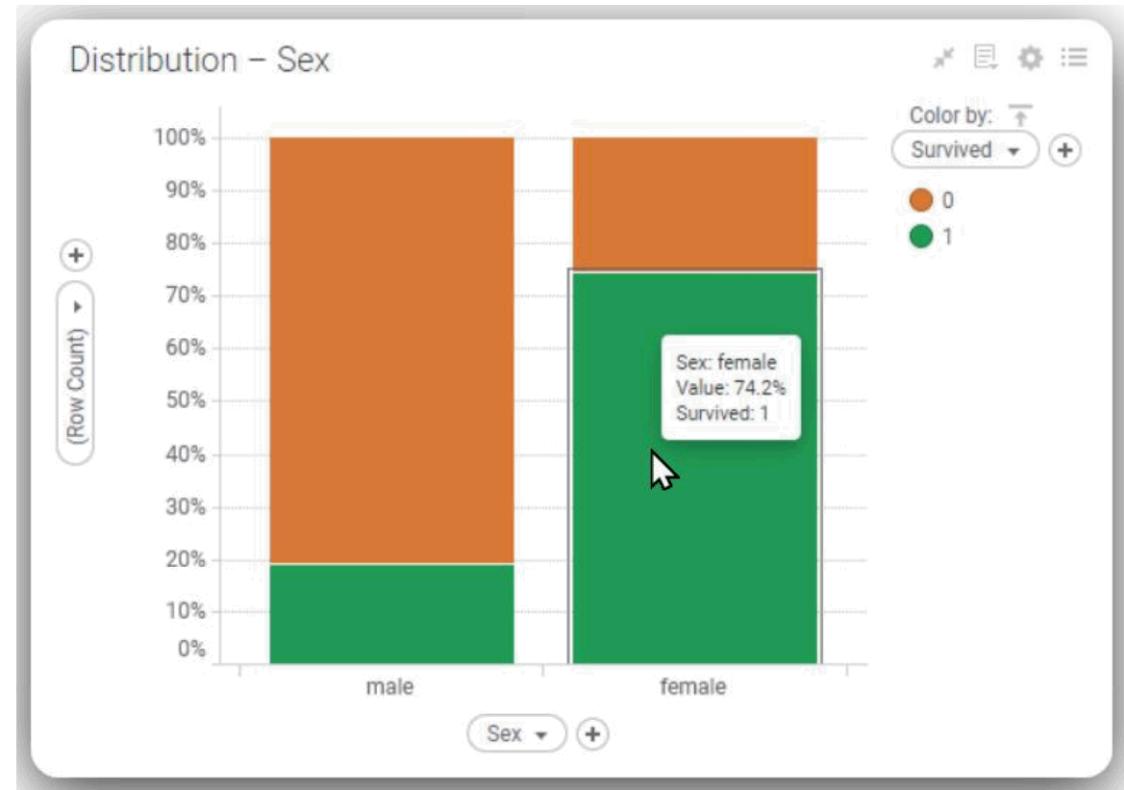
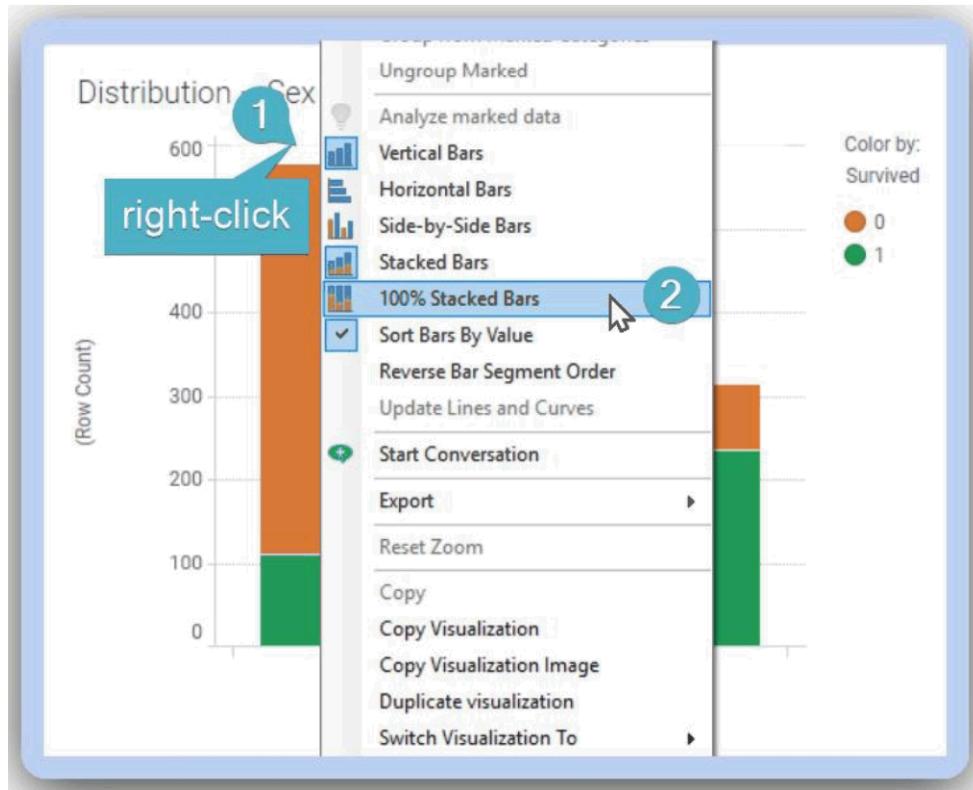
- ❑ Locate the legend
- ❑ Change female color
- ❑ Change male color



# Proportionality



# Proportionality



# Stacked vs 100% Stacked Bar Charts

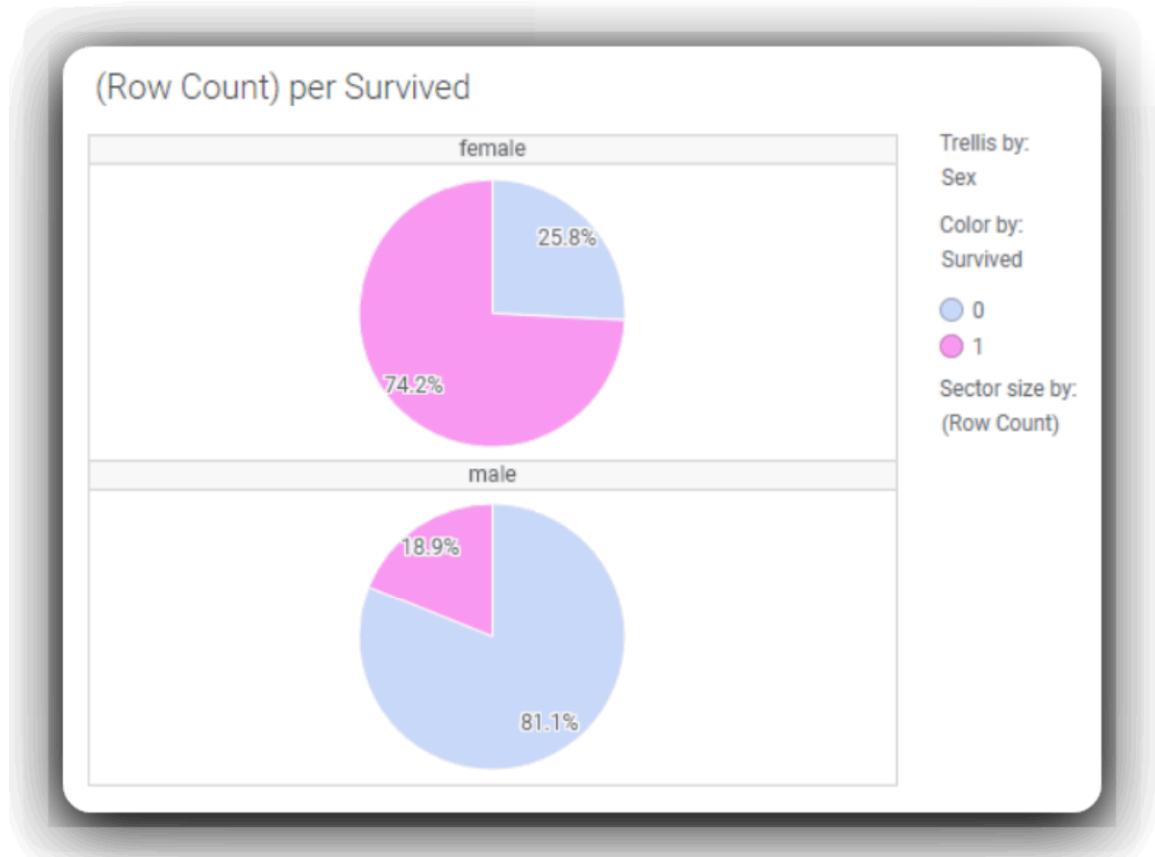
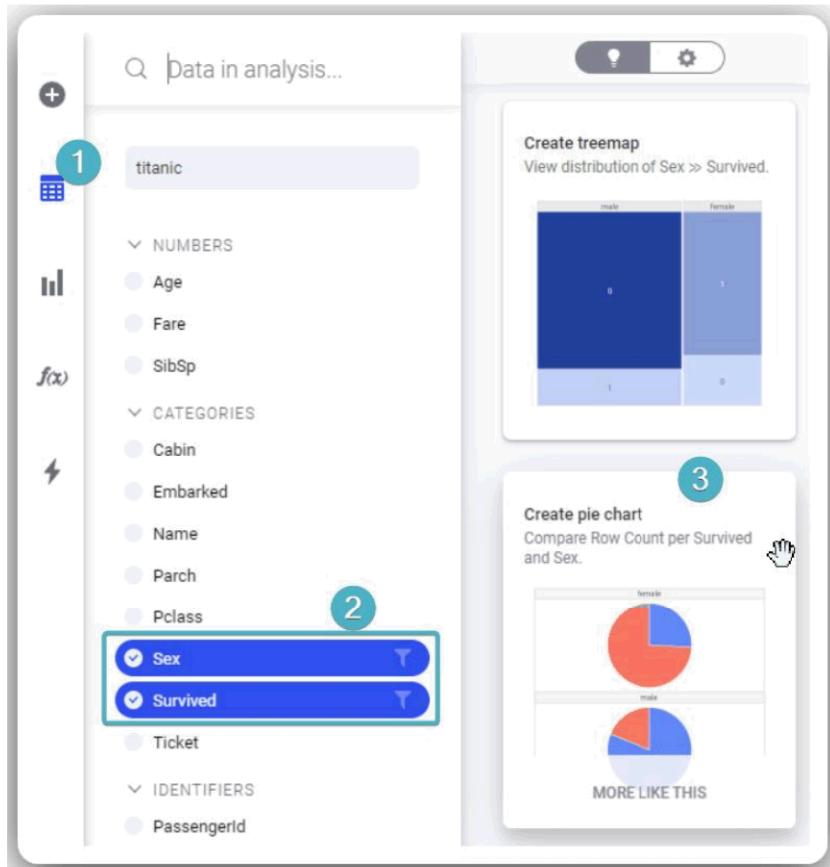
## STACKED BAR CHART

- ✓ Each bar shows the **actual values** of sub-categories stacked on top of each other.
- ✓ The **height of the bar represents the total sum** of values for each main category.
- ✓ Use when you want to show both:
  - ✓ Total size
  - ✓ Contribution of each part (absolute)

## 100% STACKED BAR CHART

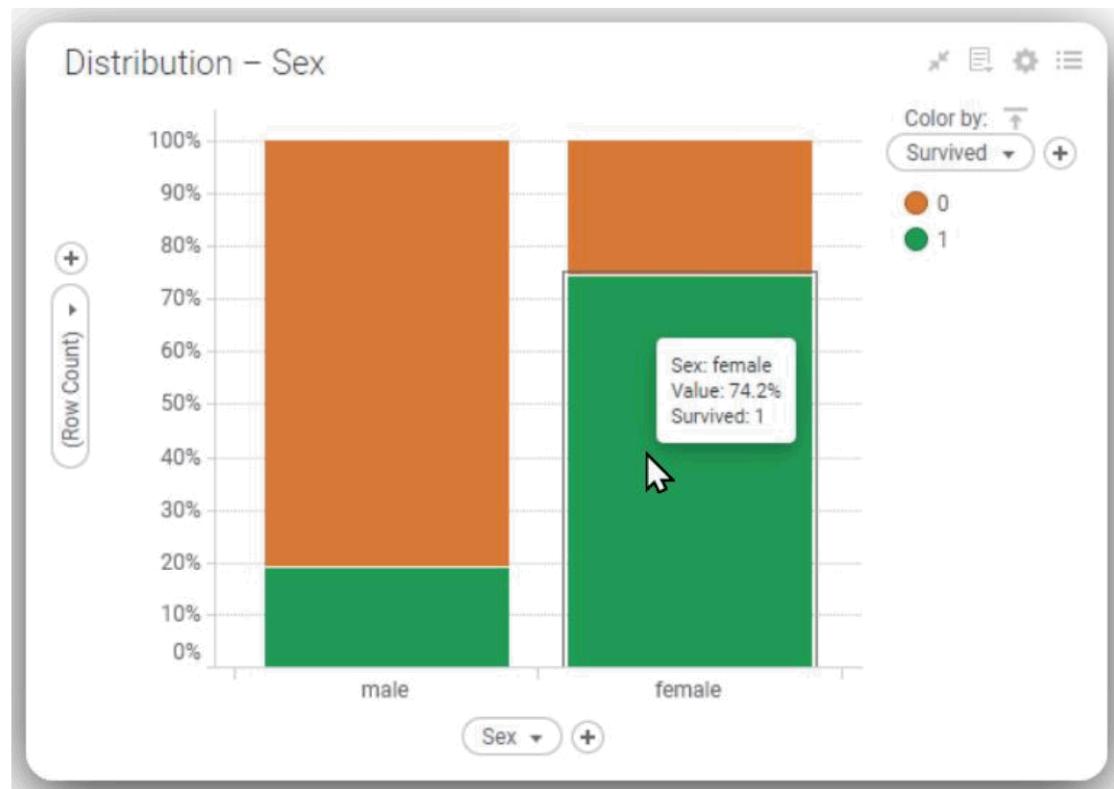
- ✓ Each bar shows **percentage composition** (not actual values).
- ✓ The bar height is always the **same (100%)**, and sub-categories are shown as a **percentage of the total**.
- ✓ Use when you're interested in **proportions**, not totals.

# Proportionality



# Proportionality

- 100% Bar Chart: Ideal for representing proportions, offering a clearer alternative to pie charts.
- Side-by-Side Bar Chart: Provides a clearer view of absolute numbers in each data category.

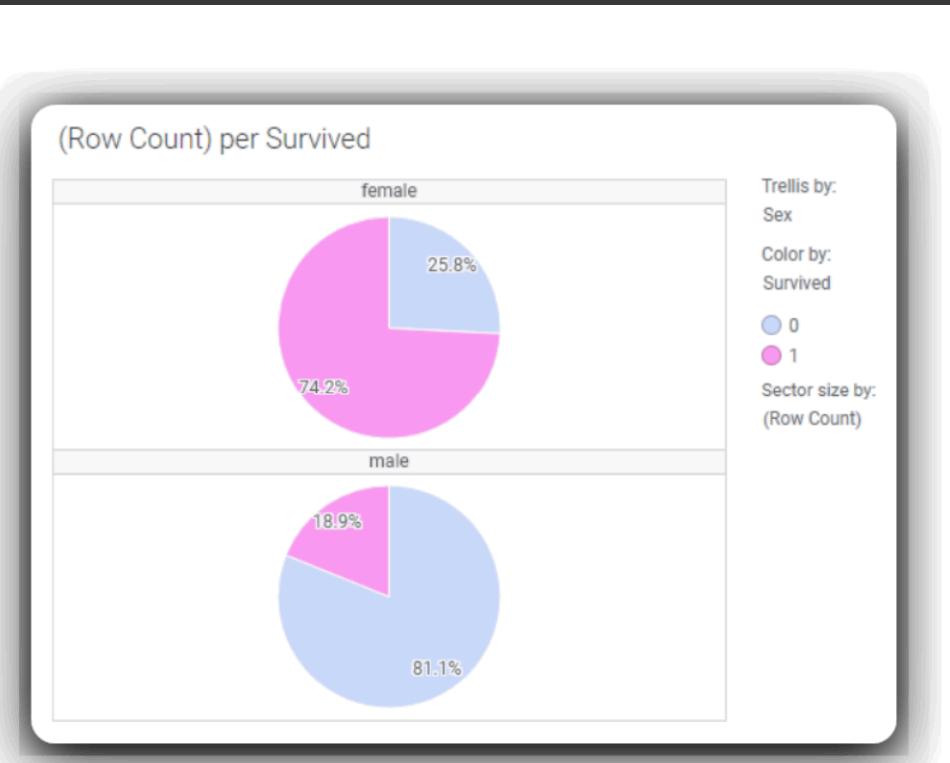
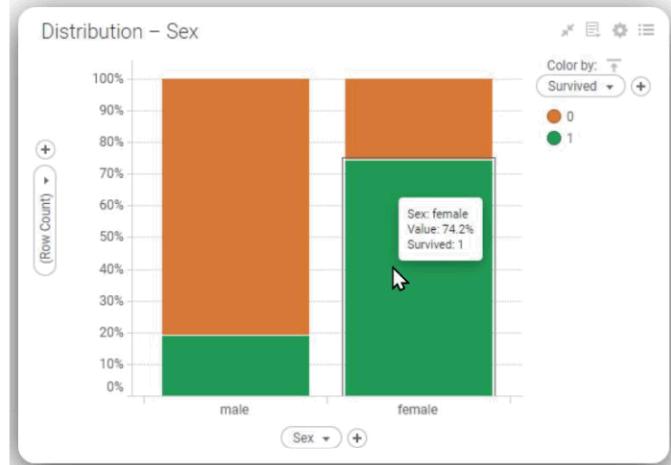


14-May-25

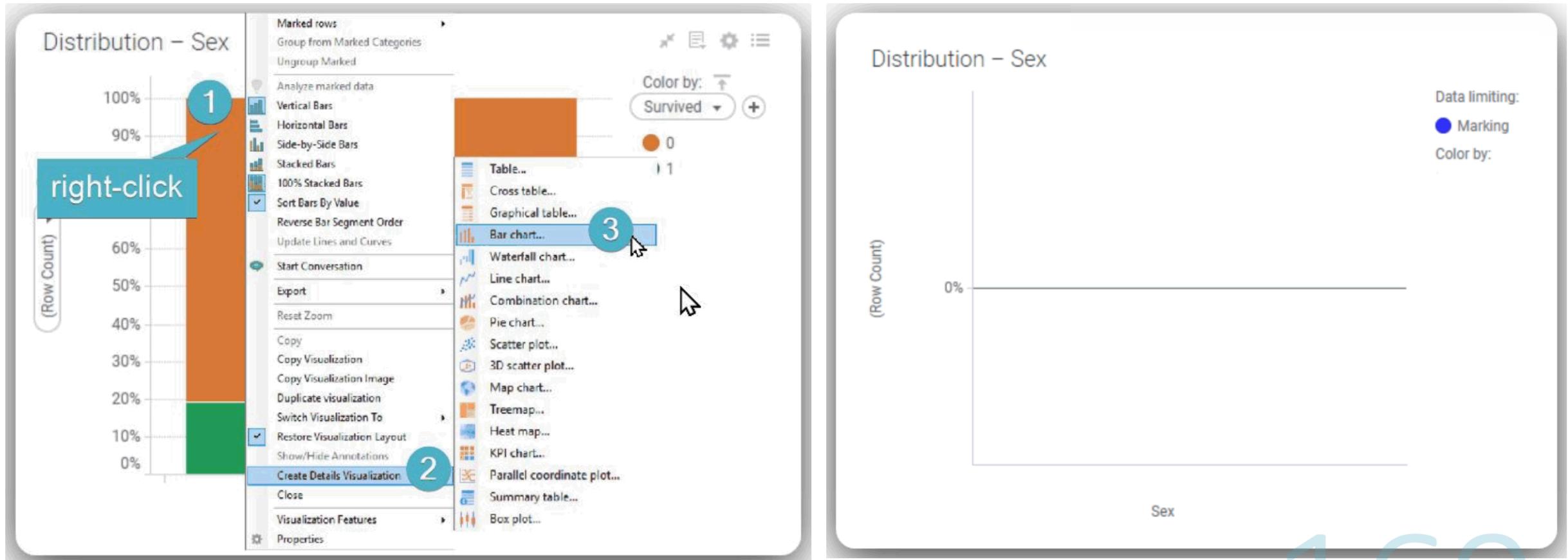
# Exploring "What" vs. "Why" in Data

What is happening?: The visualizations show proportions of male and female survivors.

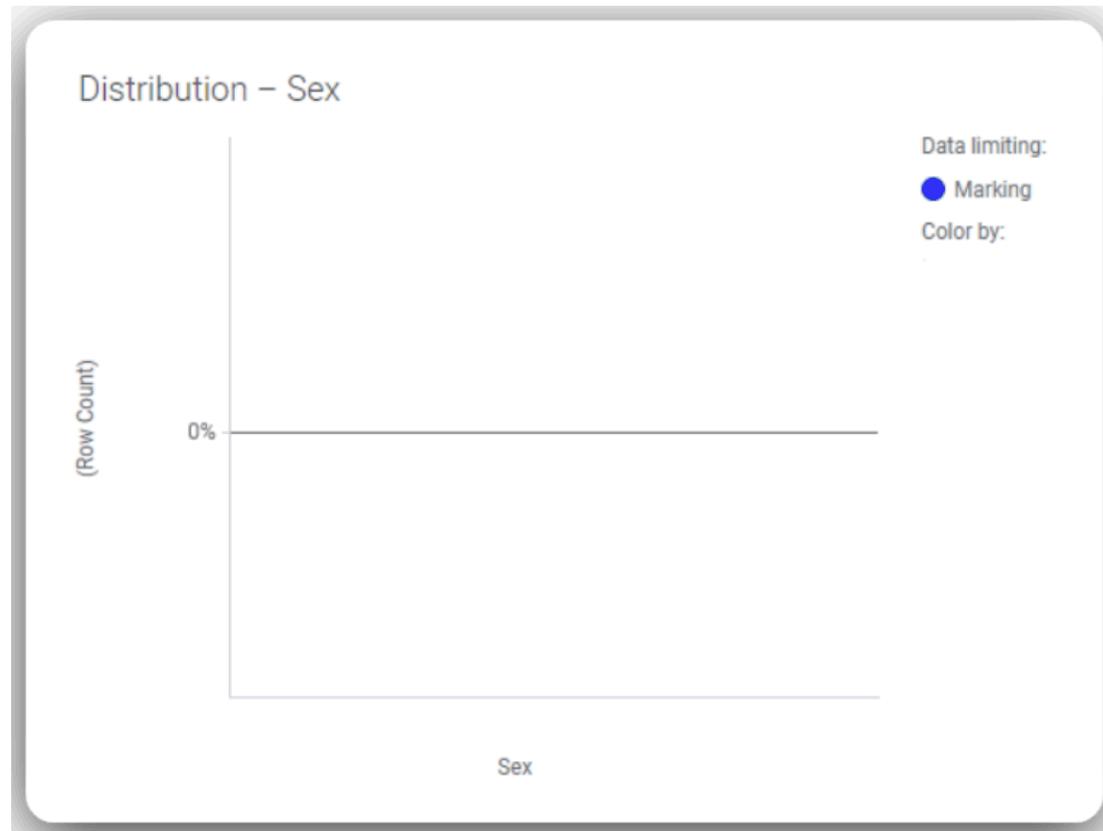
Why is it happening?: Drilling into the data allows us to explore deeper insights and understand the reasons behind the trends.



# Detailed Visualization

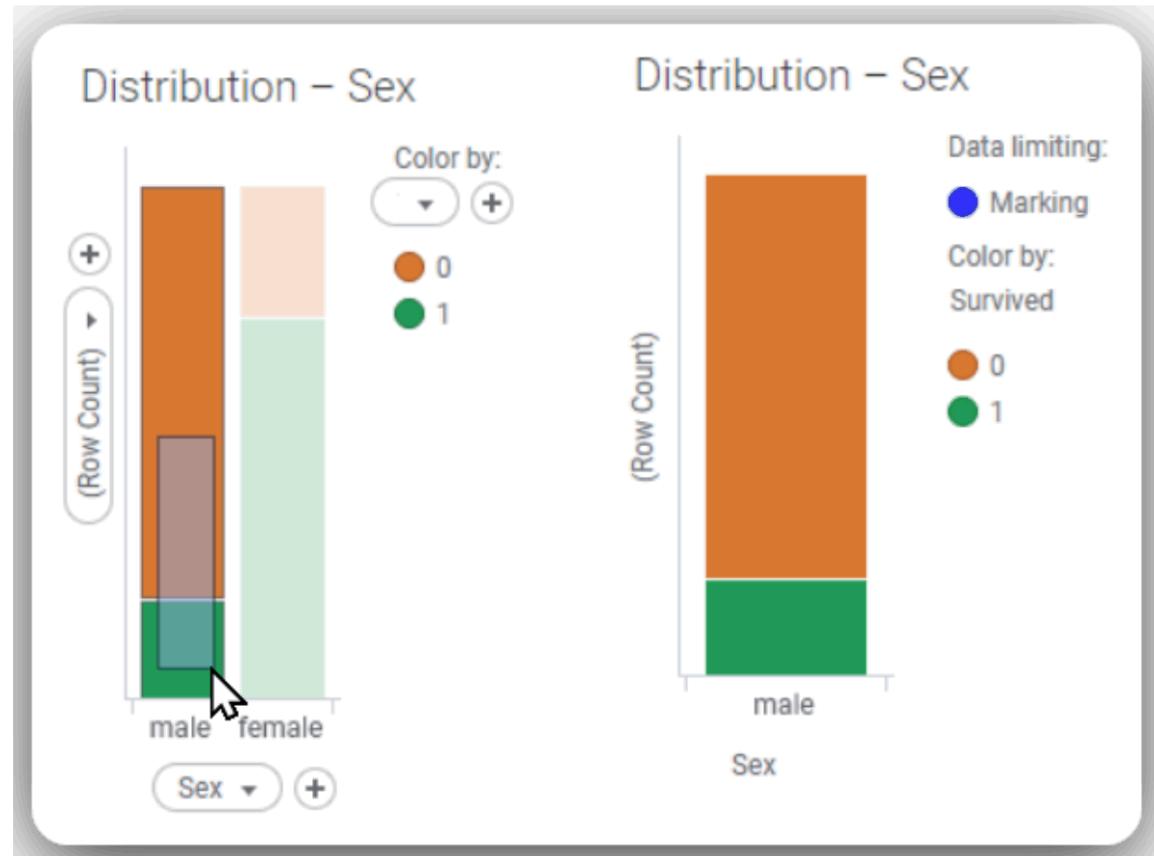


# Limiting Data & Marking

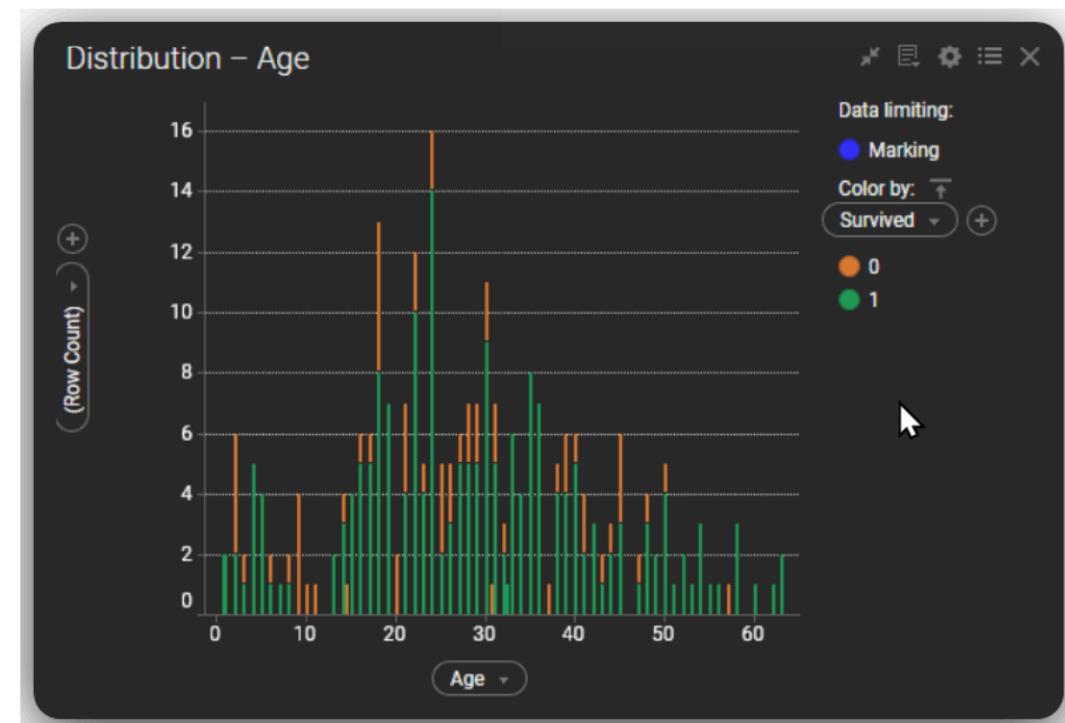
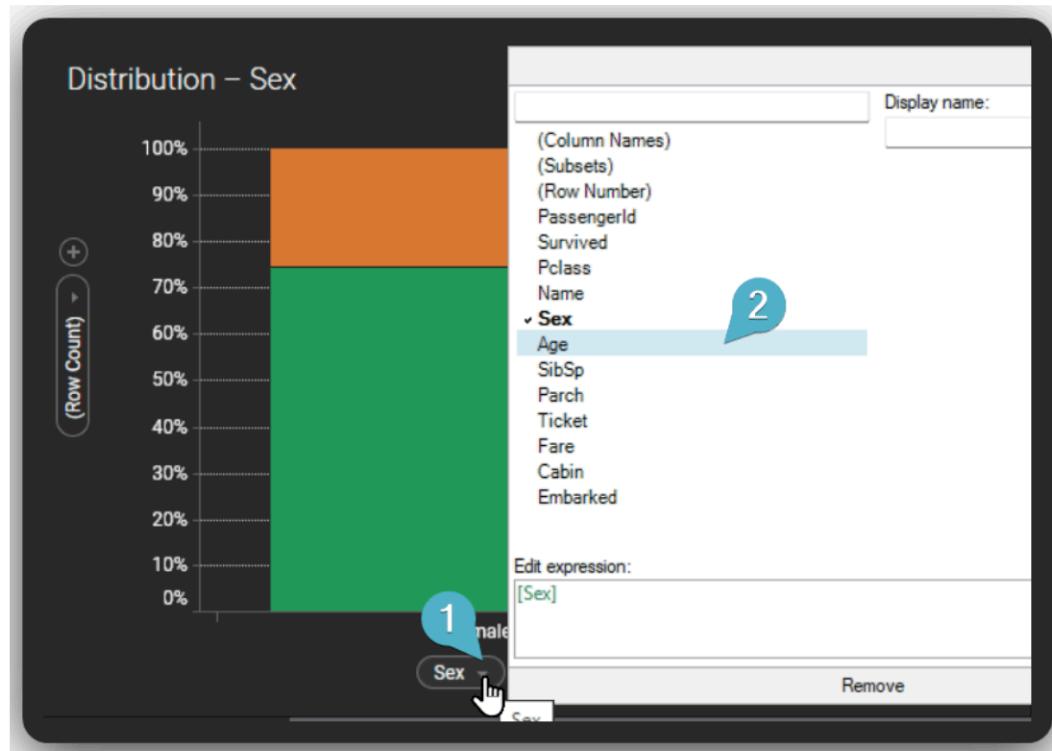


# Limiting Data & Marking

- ❑ Definition: By default, no data will be shown unless it's marked in another visualization.
- ❑ How it works: Marking data in one visualization will limit the data displayed in other visualizations.
- ❑ Purpose: Helps focus on specific subsets of data for deeper analysis.

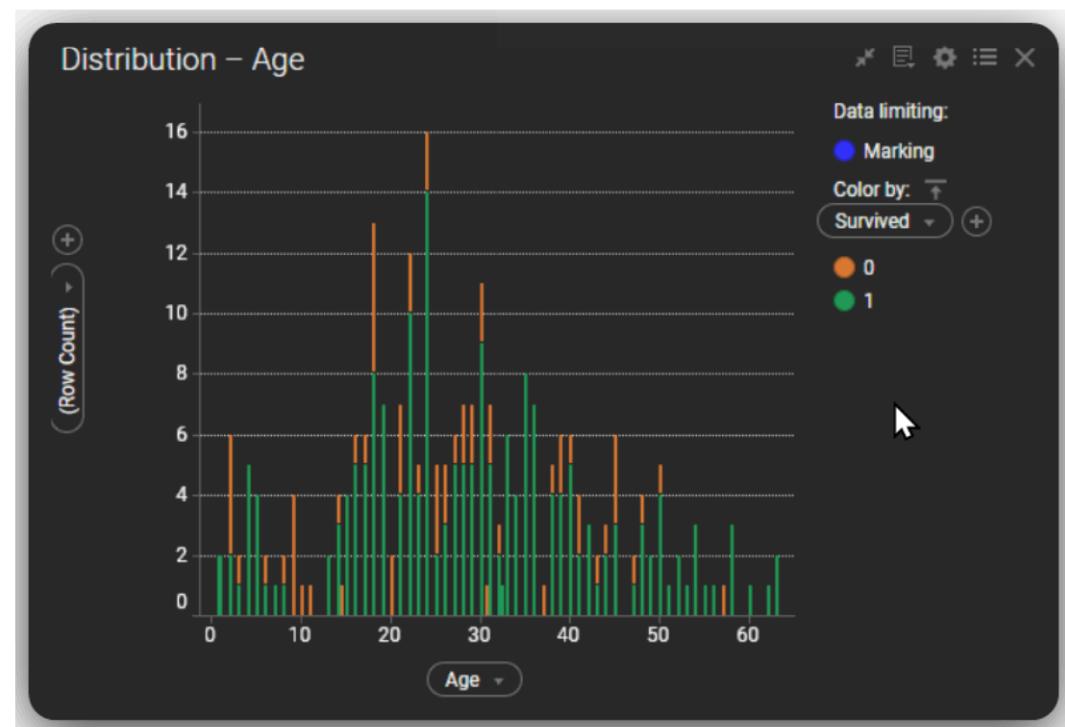


# Limiting Data & Marking

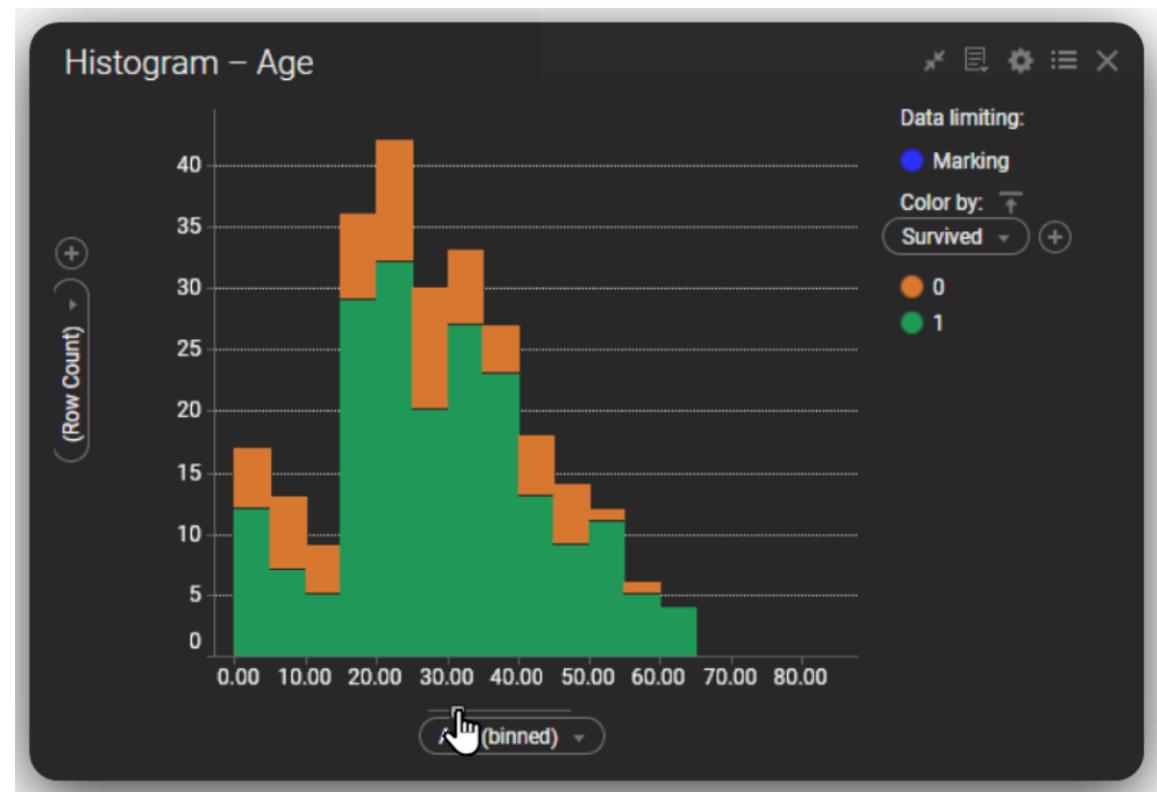
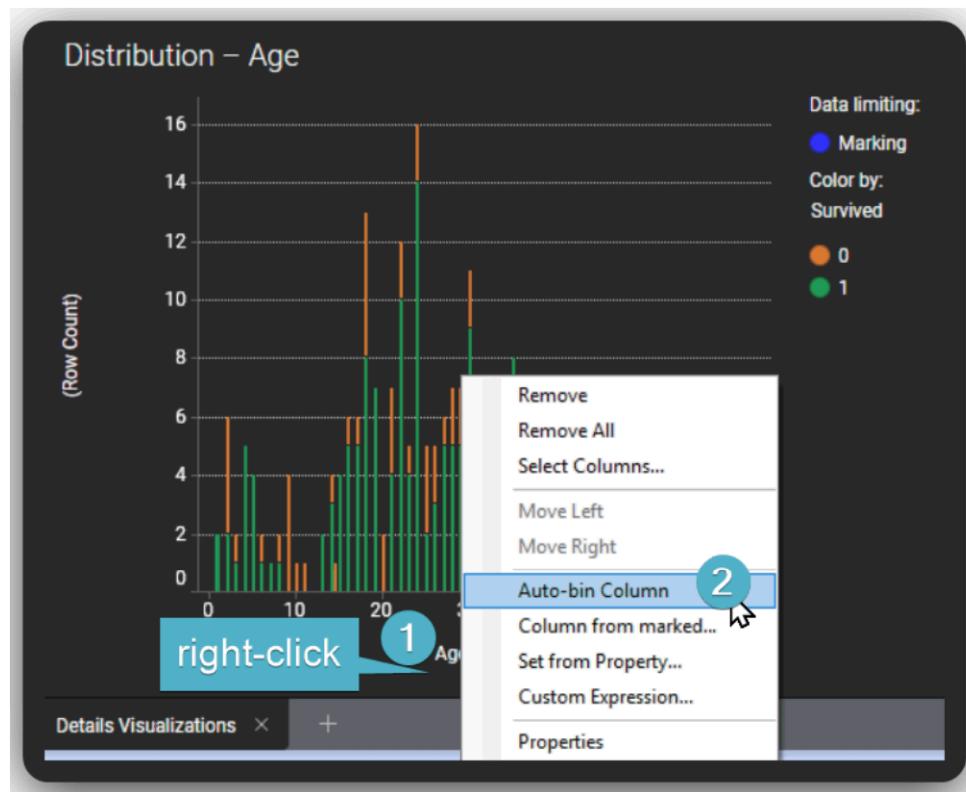


# Limiting Data & Marking

- ❖ The graph seems to show **a typical distribution**
- ❖ **Problem:** Tall bars interspersed with short bars in the histogram.
- ❖ **Cause:** Some ages recorded as fractions (e.g., 0.5 years for babies, 25.5 years for adults).
- ❖ **Impact:** Distorts the distribution and makes trend visualization unclear.
- ❖ **Solution:** Clean data by rounding fractional ages for better visualization.
- ❖ Note: marking is a **temporary highlight of selected data points**, which can influence the data displayed across multiple visualizations

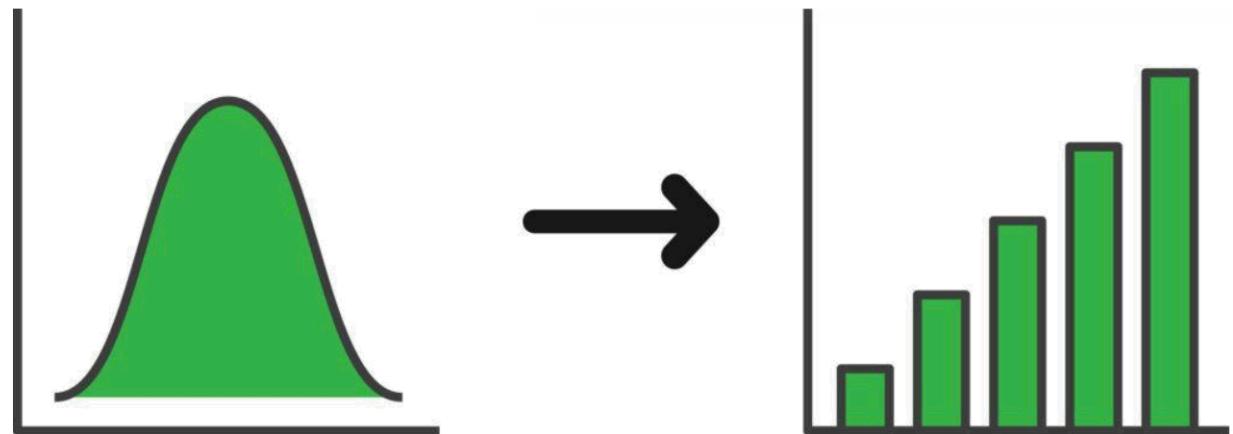


# Auto-Binning



# Binning (Discretization)

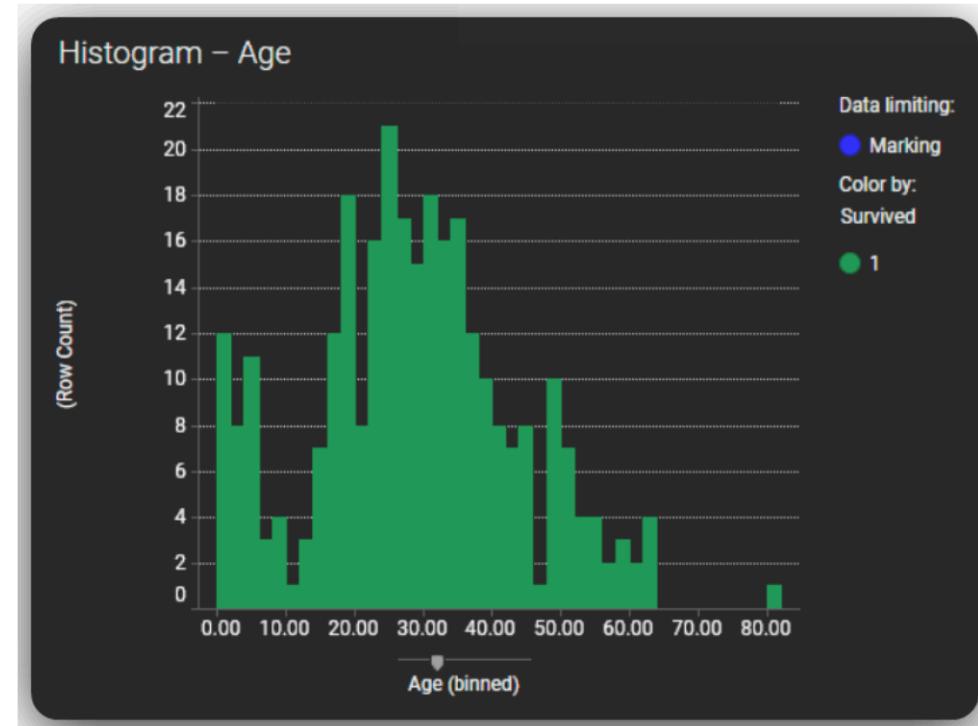
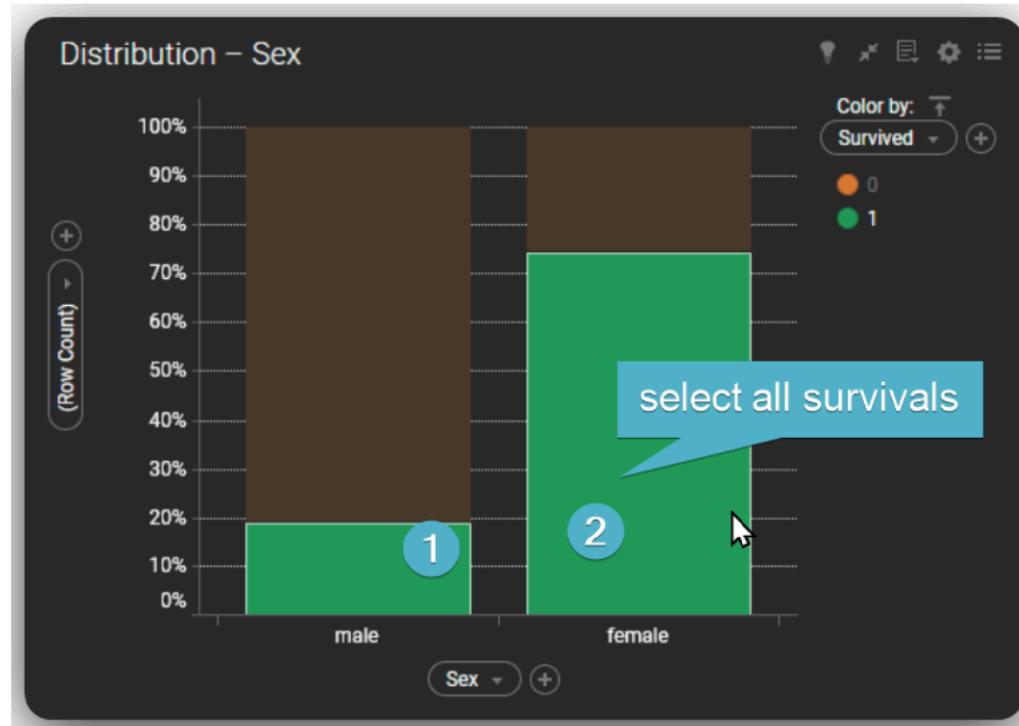
- ❖ Definition: Converting continuous data into discrete categories or bins.
- ❖ Purpose: Simplifies data analysis by grouping values into predefined intervals.



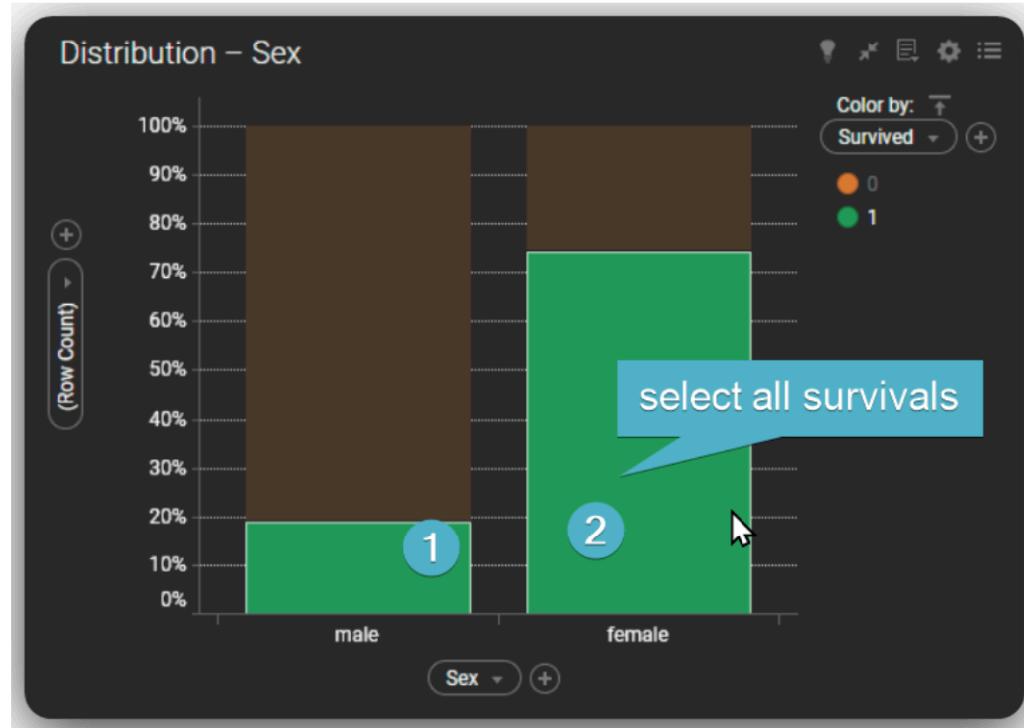
**Discretization Process**

Source: Fritz, & Fritz. (2023, September 21). *Hands-on with Feature Engineering Techniques: Variable Discretization - Fritz ai*. Fritz Ai; Fritz.ai. <https://fritz.ai/hands-on-with-feature-engineering-techniques-variable-discretization/>

# Limiting Data & Marking



# Limiting Data & Marking



Task: Try selecting the following groups from the dataset:

- All male passengers
- All female passengers
- All females that survived
- All males that survived

# Configuring Marking Properties.

How to display something while not selecting anything?

