Neekitasingh Rajput

2441339

Class: 3DM

Course: DSC261-3 - Data Visualization

Date: 07/07/2025

Exercise No: Lab Assignment-4
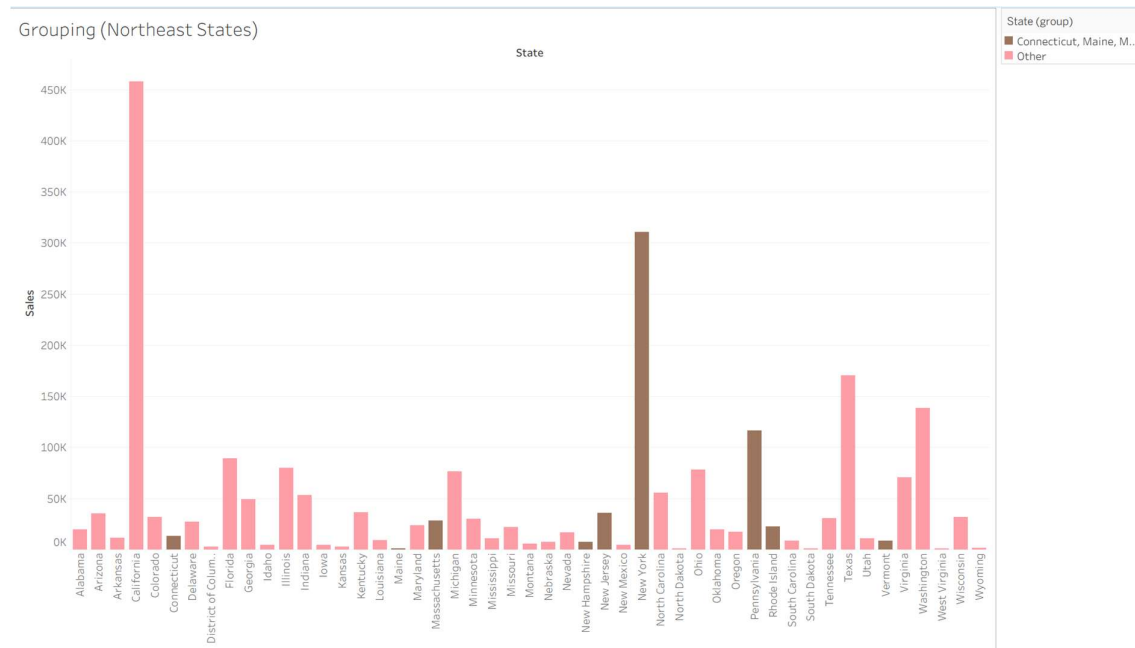
Superstore Dataset

About Dataset

The dataset pertains to a Superstore Giant aiming to optimize its business strategy amidst growing market demands and competition. It contains detailed transactional data to help analyse which products, regions, categories, and customer segments to target or avoid. The dataset also offers an opportunity to build a regression model to predict sales or profit, providing actionable business insights for improvement.

Attributes-

- **Row ID**: Unique identifier for each row. (Numerical - Integer)

- **Order ID**: Unique identifier for each customer's order. (Categorical - String)

- **Order Date**: Date when the product was ordered. (Date)

- **Ship Date**: Date when the product was shipped. (Date)

- **Ship Mode**: Shipping method specified by the customer (e.g., Standard, Express). (Categorical - String)

- **Customer ID**: Unique identifier for each customer. (Categorical - String)

- **Customer Name**: Name of the customer. (Categorical - String)

- **Segment**: Customer segment (e.g., Consumer, Corporate, Home Office). (Categorical - String)

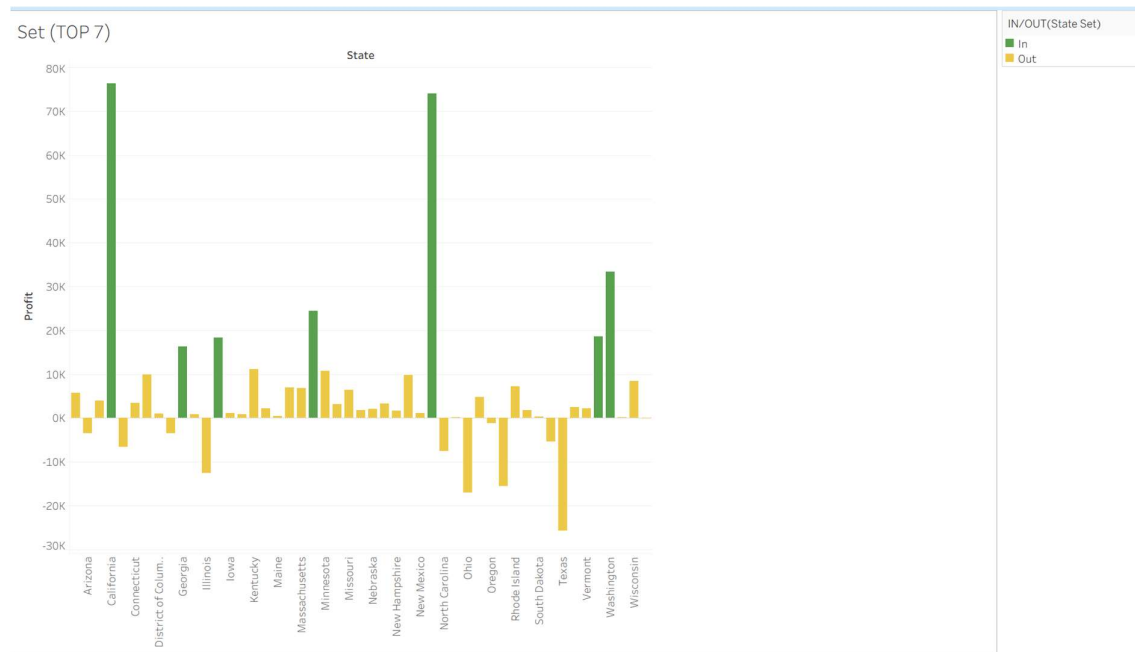- **Country**: Country of the customer's residence. (Categorical - String)

- **City**: City of the customer's residence. (Categorical - String)

- **State**: State of the customer's residence. (Categorical - String)

- **Postal Code**: Postal code of the customer's residence. (Numerical - Integer)

- **Region**: Region where the customer belongs (e.g., East, West). (Categorical - String)

- **Product ID**: Unique identifier for each product. (Categorical - String)

- **Category**: Category of the product ordered (e.g., Furniture, Office Supplies). (Categorical - String)

- **Sub-Category**: Sub-category of the product ordered (e.g., Chairs, Binders). (Categorical - String)

- **Product Name**: Name of the product. (Categorical - String)

- **Sales**: Sales amount of the product. (Numerical - Float)

- **Quantity**: Number of units of the product ordered. (Numerical - Integer)

- **Discount**: Discount percentage or amount provided. (Numerical - Float)

- **Profit**: Profit or loss incurred from the sale. (Numerical - Float)
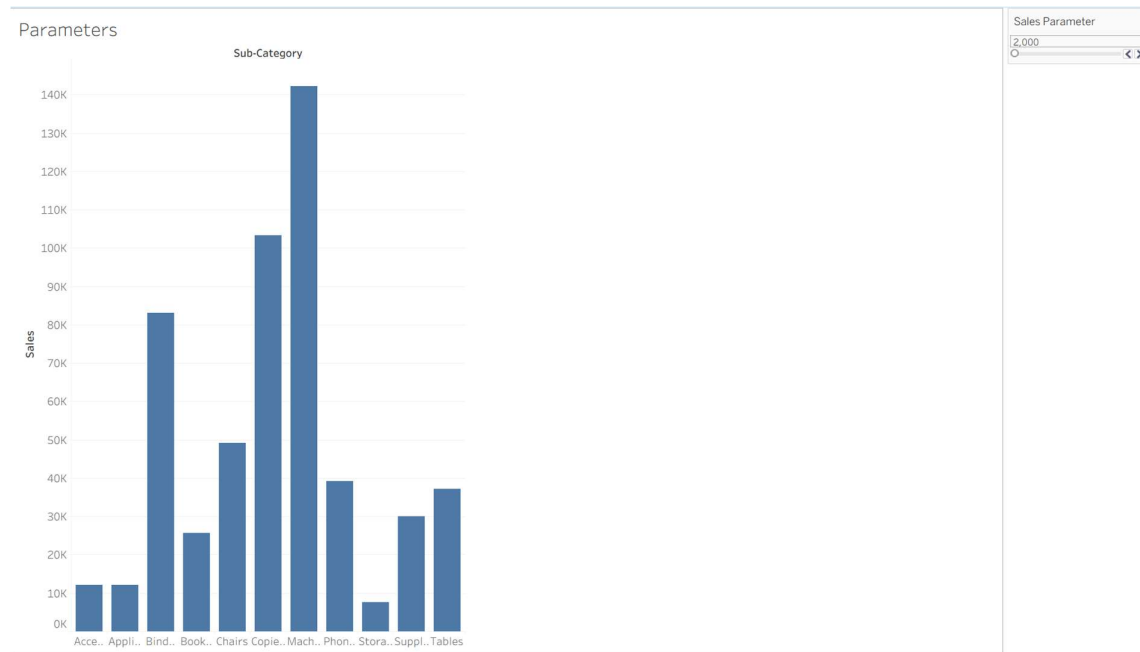
1.

Grouping (Northeast States)

This chart shows sales amounts for different states, grouped into Northeast states and others. California has the highest sales, while most other states have much lower sales. Here we have applied grouping to northeast states of the country and visualized it differently for better highlighting.

2.

Set (TOP 7)

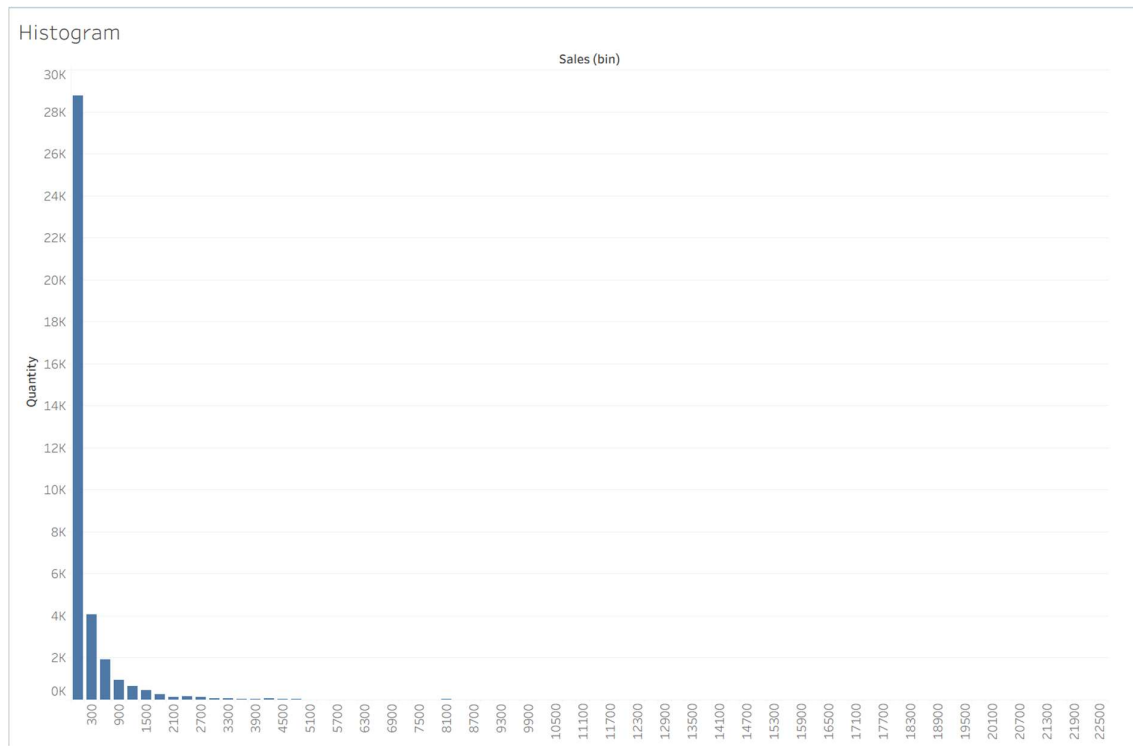IN/OUT(State Set)
- In
- Out

This chart shows the profits for the top 7 states, with the data split into in-state (green) and out-state (yellow) categories. Arizona stands out with the highest profit, reaching around 80K, while Massachusetts also does pretty well with a solid profit close to that. Other states like California, Illinois, and Kentucky have smaller profits, mostly between 10K and 20K, but some like Rhode Island and Oregon show losses, dipping down to around -20K or more. The chart gives a clear picture of which states are doing great and which ones are struggling a bit with their profits!
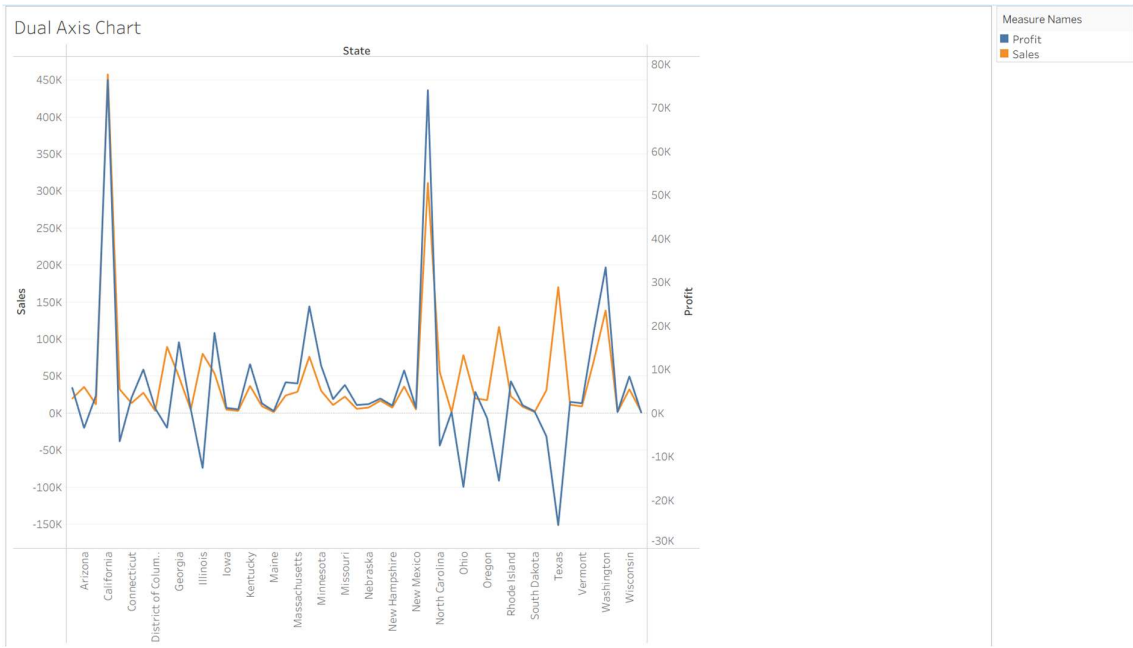
3.

This chart shows the sales for different sub-categories like Accessories, Appliances, and Binders, with a parameter set to 2,000. The sub-category "Tables" has the highest sales, reaching around 140K, while others like Accessories and Binders also do well with sales around 80K to 100K. The rest, like Phones and Storage, have much lower sales, dropping to around 10K to 40K, giving a clear idea of which items are selling the most!
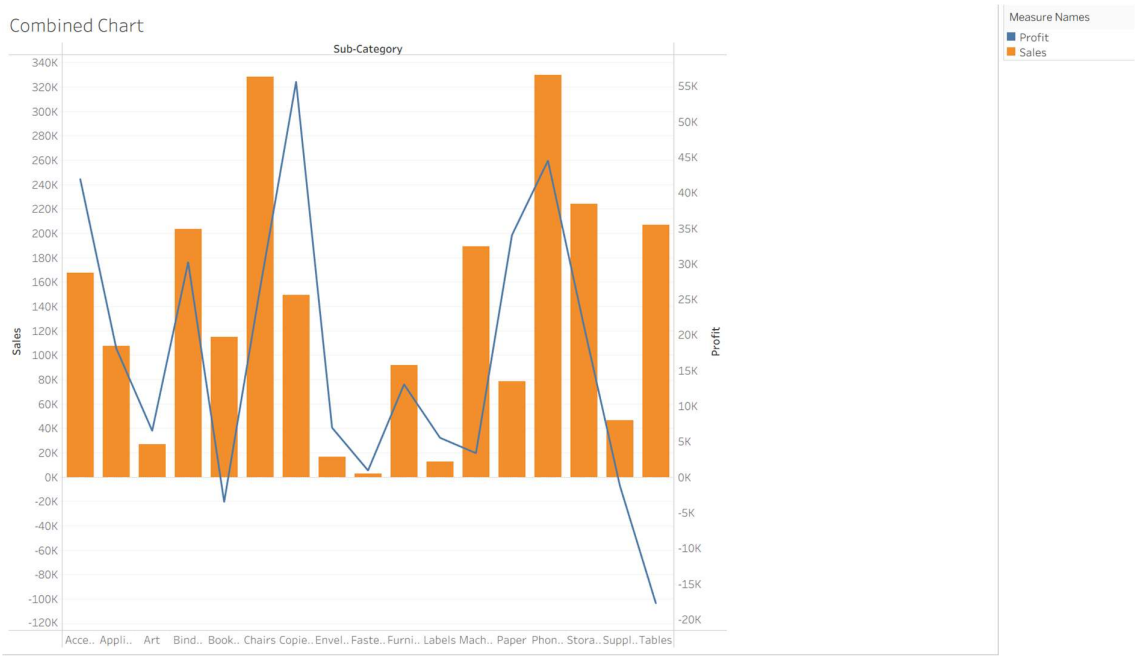
4.

Histogram

Sales (bin)

This chart is a histogram showing the quantity of sales grouped into bins of 600, starting from around 300 to 23,900. Most sales fall into the first bin around 300, with a huge spike of nearly 30K, while the rest of the bins have very few sales, dropping off quickly to almost zero. It's clear that the majority of sales are clustered at the lower end!

5.

This dual-axis chart compares sales (orange line) and profits (blue line) across different states. Arizona and Massachusetts have the highest sales, peaking around 400K and 350K, with profits also high at about 60K and 70K, while states like Vermont and Wisconsin show negative profits dipping below -10K despite moderate sales around 50K to 100K. It's a cool way to see how sales and profits don't always match up across the board!

6.

This combined chart shows sales (orange bars) and profits (blue line) for different sub-categories like Accessories, Appliances, and Tables. Tables and Binders lead with the highest sales, around 320K and 200K, and profits peak at about 50K for Tables, while some categories like Phones dip into negative profits around -100K despite decent sales. It's a fun way to see how sales and profits vary across the items!

----X----