Superstore Sales Analysis

Superstore Sales Dataset

The **Superstore Sales dataset** contains **1,000 records** across **14 fields**, capturing detailed transaction data from a retail superstore. This dataset includes order details, customer information, product categories, sales amounts, discounts, and regional distribution.

Dataset Overview

- **Total Rows:** 1,000
- **Total Columns:** 14

Column Descriptions

Column Name	Description
Order ID	Unique identifier for each order
Order Date	The date when the order was placed
Ship Date	The date when the order was shipped
Customer Name	Name of the customer (anonymized)
Customer Segment	The customer category (e.g., Corporate, Home Office)
Category	The broad product category (e.g., Technology, Office Supplies)
Sub-Category	Specific sub-category of products (e.g., Binders, Printers)
Product Name	The name of the purchased product
Sales Amount	The revenue generated from the sale
Profit	The profit earned per transaction
Discount	The discount applied to the order
Quantity	The number of units sold in the order
Region	The geographical region of the sale (e.g., East, West)
State	The state where the order was placed

File Information

- File Name: Superstore_Sales_Dataset.xlsx
- Format: Excel (.xlsx)
- Size: Approximately 1000 rows

Data Analysis

1. Structure of the data

```
df.head()
```

	Order ID	Order Date	Ship Date	Customer Name	Customer Segment	Category	Sub-Category	Product Name	Sales Amount	Profit	Discount	Quantity	Region	State
0	ORD1000	2023-03-28	2023-03-30	Customer 86	Home Office	Office Supplies	Binders	Binders 96	611.31	39.71	0.16	8	South	Texas
1	ORD1001	2023-10-15	2023-10-17	Customer 449	Home Office	Technology	Phones	Phones 54	62.75	5.02	0.11	4	East	California
2	ORD1002	2023-01-24	2023-01-30	Customer 90	Corporate	Technology	Printers	Printers 47	454.63	52.64	0.18	3	West	Texas
3	ORD1003	2023-04-04	2023-04-09	Customer 383	Home Office	Technology	Printers	Printers 92	404.49	108.25	0.06	9	South	California
4	ORD1004	2023-06-11	2023-06-15	Customer 84	Small Business	Office Supplies	Paper	Paper 43	295.64	50.94	0.05	10	West	Florida

2. Shape of the Dataset

This Dataset Contain:

- **Total Rows:** 1,000
- **Total Columns:** 14

3. Identifying the Null Values

```
# Checking the null values in the dataset
df.isnull().sum()
```

```
Order ID            0
Order Date          0
Ship Date           0
Customer Name       0
Customer Segment    0
Category            0
Sub-Category        0
Product Name        0
Sales Amount        0
Profit              0
Discount            0
Quantity            0
Region              0
State               0
dtype: int64
```

4. Identifying the Data Types

```
#df.dtypes, Helps you to find the data types of each columns in the dataset
df.dtypes
```

```
Order ID                    object
Order Date          datetime64[ns]
Ship Date           datetime64[ns]
Customer Name               object
Customer Segment            object
Category                    object
Sub-Category                object
Product Name                object
Sales Amount               float64
Profit                     float64
Discount                   float64
Quantity                     int64
Region                      object
State                       object
dtype: object
```

5. Identifying the Unique Values

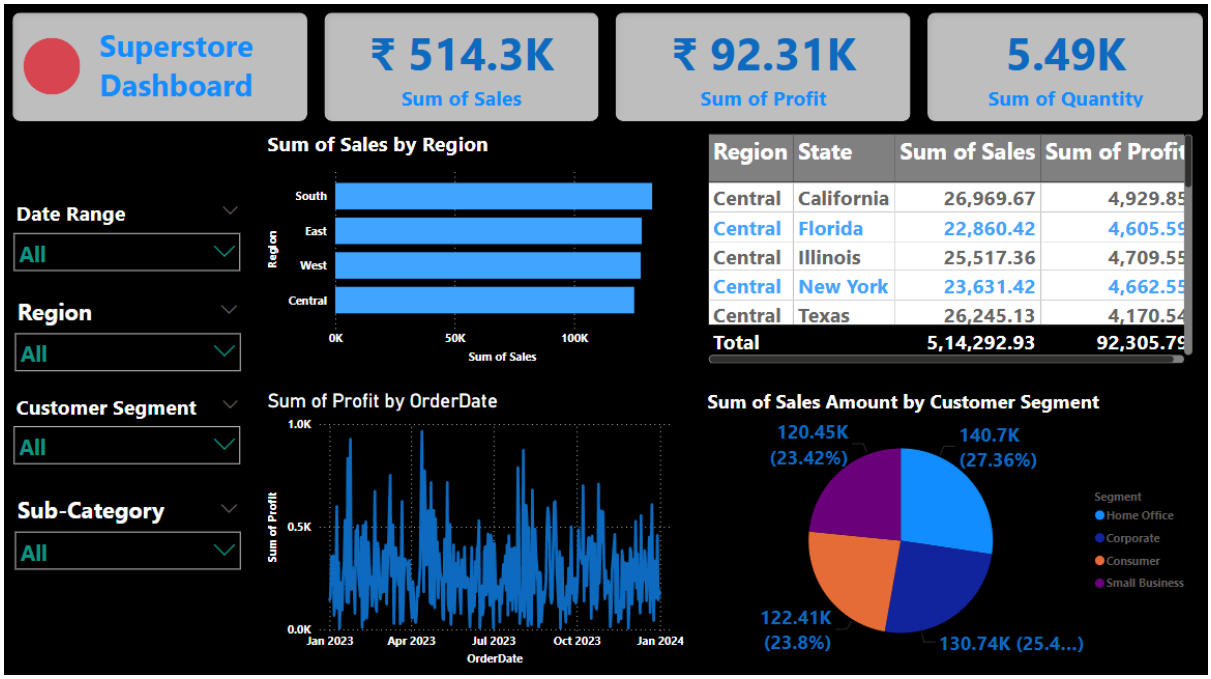
```
# df.nunique(), It shows the Number of unique values per column in the dataset
df.nunique()
```

```
Order ID            1000
Order Date           337
Ship Date            337
Customer Name        441
Customer Segment       4
Category               3
Sub-Category          12
Product Name         681
Sales Amount         994
Profit               983
Discount              21
Quantity              10
Region                 4
State                  5
dtype: int64
```

Superstore Sales Dashboard

This dashboard provides insights into **sales, profit, and customer segments.**

Superstore Dashboard

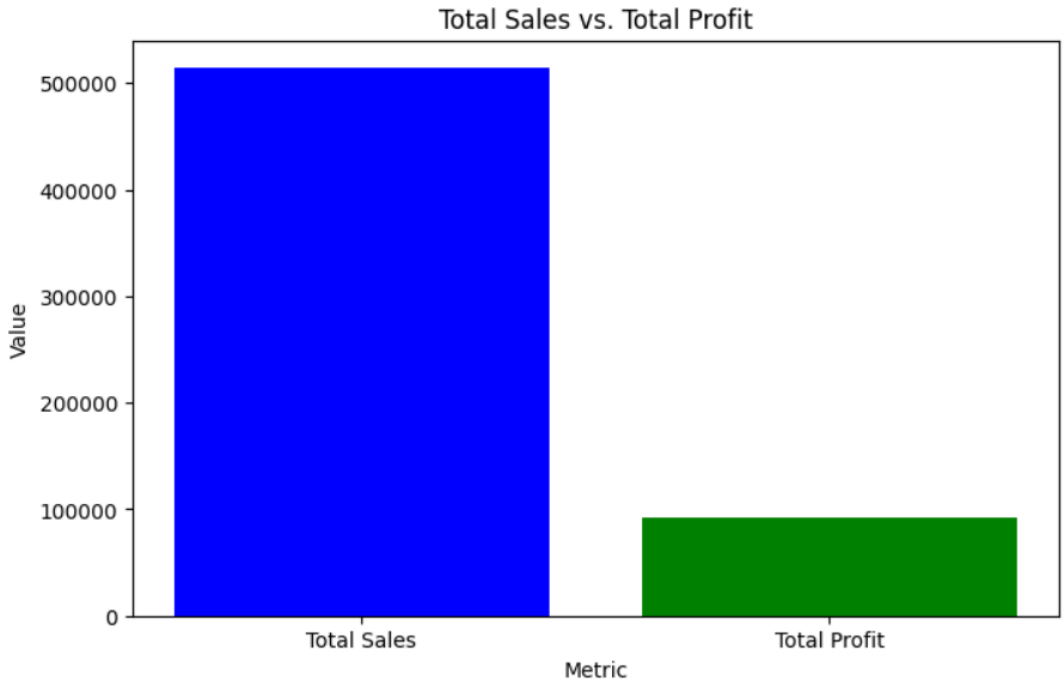


Analysis

Total Sales, Profit, and Quantity

Total Sales, Profit, and Quantity

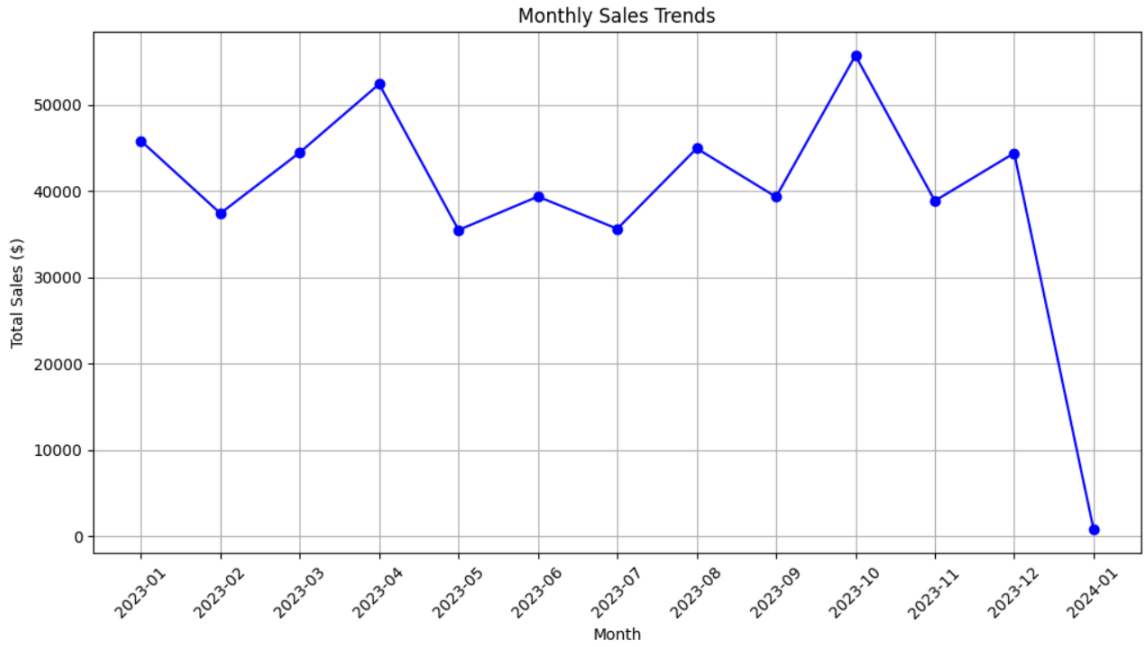
- The **Sum of Sales** in the Superstore Sales Dataset is **514292**
- The **Sum of Profit** in the Superstore Sales Dataset is **92305**
- The **Sum of Quantity** in the Superstore Sales Dataset is **5491**



Here, Total sales are 514292 and the total profit is 92305

Data Visualization

Line Chart: Monthly sales trends



This line chart represents **Monthly Sales Trends** over a period from **January 2023 to January 2024**.

Key Observations:

1.Fluctuations in Sales

- Sales are not consistent; they show **ups and downs** throughout the months.
- There are periods of high sales, followed by declines.

2. Peaks and lows

- **Highest sales** occur around **April 2023 and October 2023** (over $50,000)
- **Lowest sales** occur in **January 2024**, where sales drop significantly to almost zero.

3. End-of-Year Drop

- The steep decline in **December 2023 and January 2024** suggests a major reduction in sales.
- Possible reasons could be **seasonality, market conditions, or operational** issues affecting sales.

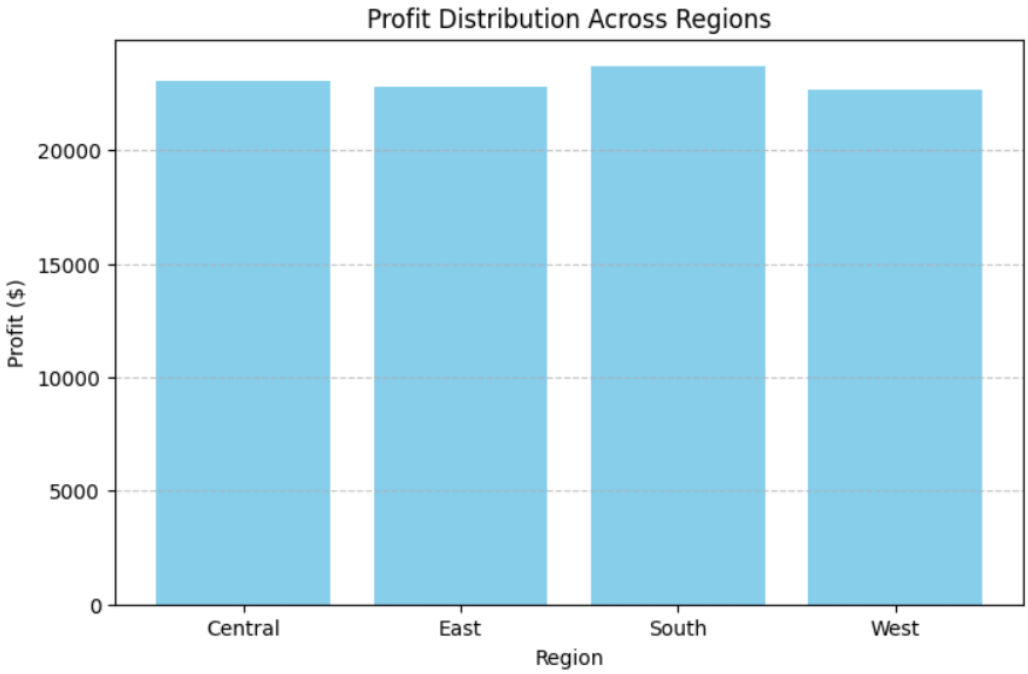
Possible Insights:

- **Sales Performance Needs Analysis** → Identify factors that drive high sales in peak months and replicate those strategies.
- Investigate and Decline → The sudden drop in January 2024 suggests an anomaly. Was it due to **holidays, reduced demand, stock issues, or company shutdowns**?
- **Forecasting and Strategy** → If this trend repeats yearly, the business should **plan for seasonal drops** and adjust marking or inventory strategies.

Profit distribution across regions

	Region	Profit
0	Central	23078.08
1	East	22834.36
2	South	23706.49
3	West	22686.86

Visualization



This bar chart represents the **Profit Distribution Across Regions.** Here's what it tells us:

Key Insights:

1.Almost Equal Profit Across Regions

- The profit values for **Central, East, South, and West regions** are quite similar, with only slight variations.
- No single region dominates in terms of profit.

2. Highest Profit in the South Region

- The **South** region has the highest profit, slightly higher than the other regions.

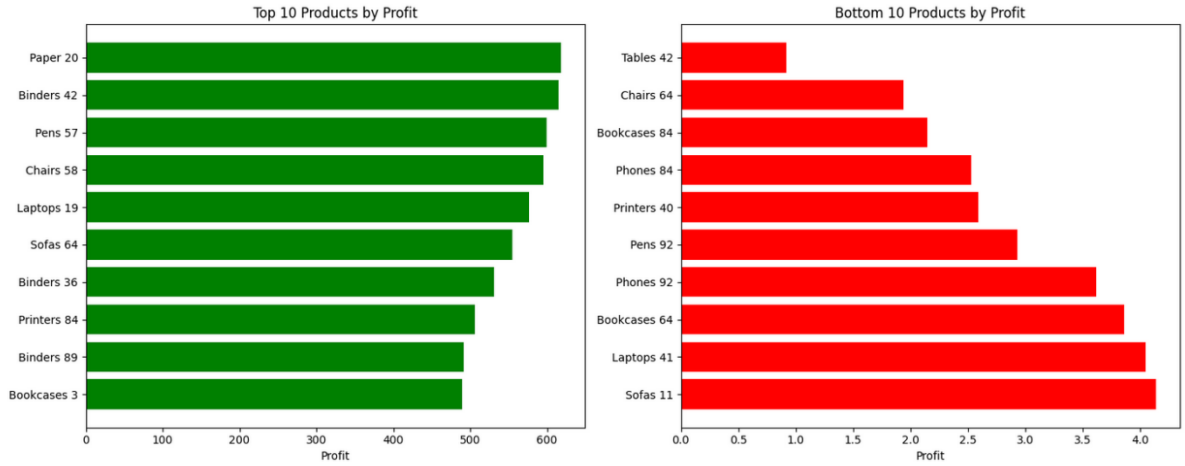
3. Lowest Profit in the West Region

- The **West** region has the lowest profit among the four, but the difference is not drastic.

Possible Business Insights:

- Since **profit is nearly equal across all regions**, the company has a balanced market presence.
- If the **West region has lower profitability**, further analysis could determine if this is due to **lower sales, higher costs, or market competition.**
- The **South region performing the best** suggests that marketing and operational strategies in that region might be more effective.

Bar Chart: Top & bottom products by profit



This chart presents two horizontal bar chart graphs side by side:

1.Left Chart (Top 10 Products by profit – Green Bars)

- It shows the top most profitable products.
- The products are listed on the y-axis, while the x-axis represents profit.
- The products at the top (e.g., Paper 20, Binders 42) have the highest profit.
- These products contribute significantly to overall profitability.

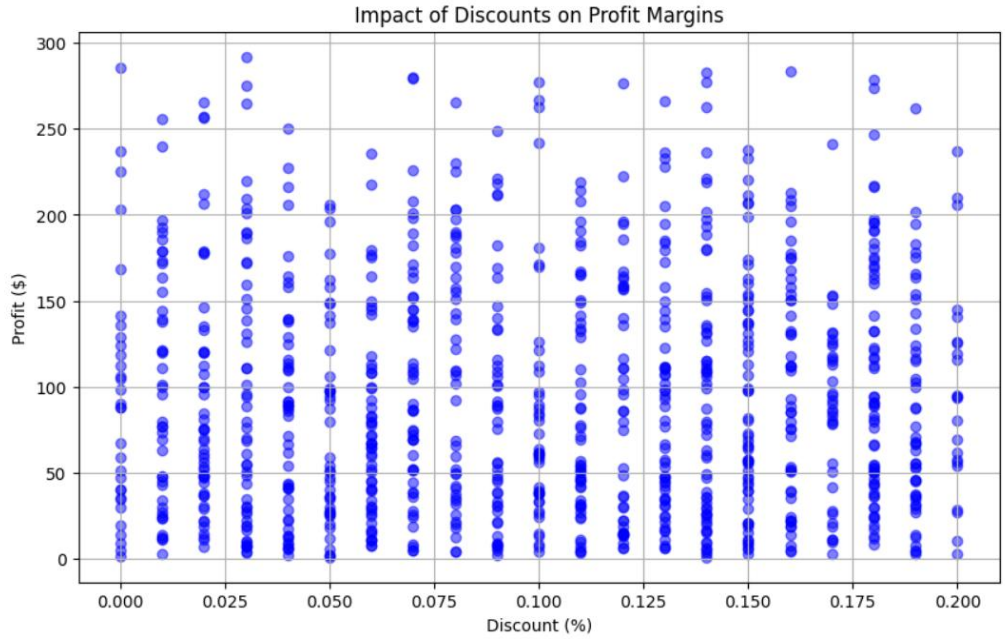
2.Right Chart (Bottom 10 products by Profit – Red Bars)

- It highlights the least profitable products.
- The x-axis represents profit, which is relatively lower compared to the left chart.
- The least profitable product is at the top (e.g., Tables 42), while Sofas 11 has the highest profit among the bottom 10.
- These products may be contributing less to the company's profitability or even incurring losses.

Key Takeaways:

- The green bar chart shows which products should be prioritized for sales and inventory management due to their high profitability.
- The red bar chart highlights products that might need review – either pricing adjustments, cost reduction, or discontinuation.'
- The Significant difference in profit values between the top and bottom products suggests opportunities for optimization in sales and marketing strategies.

Scatter Plot: Impact of discounts on profit margins



This scatter plot represents the **Impact of Discounts on Profit Margins**. Here's what it tells us:

1.X-Axis (Discount %):

- Represents the percentage of discount given on products.
- Ranges from 0% to around 20%.

2.Y-Axis (Profit $):

- Represents the profit in dollars.
- Ranges from $0 to above $300.

Key Observations:

1.Higher Discounts, Lower Profits:

- Many transactions with **low or no discount (left side)** have **higher profits.**
- As the discount **percentage increase (right side),** profit values become more scattered and generally lower.

2. Dense clustering at low Discounts:

- Most transactions are **concentrated near lower discount values,** suggesting that fewer products are sold at high discounts.

3. Some Outliers with High Profits at Higher Discounts:

- A few high-profits appear even at higher discount levels, indicating that some high-margin products remain profitable despite discounts.

Insights:

- **Higher discounts may reduce profitability,** as seen from the download spread of points.
- **Low or no discounts often correlate with higher profits**, suggesting that maintaining regular pricing might be beneficial.
- **Profitability varies even at the same discount level**, meaning other factors (e.g., product category, volume, or cost) influence profit.