

# CLARK UNIVERSITY



SCHOOL OF  
PROFESSIONAL STUDIES

## GLOBAL INSIGHTS: VISUALIZING SALES AND PROFIT TRENDS IN E-COMMERCE

Group members:

NIKHIL KATTAGURI

SATHVIKA SHETTI

SPS Department, Clark University

MSDA3060-03-S25: DATA VISUALIZATION & STORY TELLING

Kayed Akkawi

4<sup>th</sup> May 2025

**List of Contents**

1. Introduction	3
2. Executive Summary	3
3. Overview of the dataset	4
4. Methodology	5
5. Intended Audience	6
6. Key Findings	6
7. Recommendations	14
8. Conclusion	15
9. References	16

## Introduction

This is an in-depth analysis of the Global Superstore dataset to determine patterns of sales, profit, and customer trends by product category and region. The dataset we have chosen consists of 51,290 rows and 24 columns. The dataset includes sales, profit, and customer data for a global superstore, covering multiple regions, product categories, and order details.

This report presents a detailed analysis of the Global Superstore dataset. The main goal of this analysis is to find ways to reduce financial losses and increase profits. By studying the data, we identified trends in sales, profit, customer behavior, and product performance. This helps businesses understand which areas are doing well and invest in areas that are already working, ultimately assisting businesses in optimizing their overall performance and delivering more results in the marketplace.

## Executive Summary

This report provides evidence-supported findings based on the dataset to help improve states, such as Illinois, North Carolina, and New York. Businesses can improve their overall performance by focusing on fixing these issues and supporting areas that are doing well.

This report also includes data-driven recommendations, supported by evidence from the dataset, to help improve financial outcomes. In addition to identifying losses, we also look at the best performing products, find the most profitable geographical areas, and search for loyal customers who are making meaningful efforts to create high revenue streams. By developing a clear and comprehensive picture of the determinants of positive as well as potential negative determinants, companies can act strategically with the vision of achieving long-term, sustainable business success.

## Overview of the dataset

The Global Superstore Dataset on Kaggle consists of 24 columns. Below is a detailed breakdown of the dataset:

Column name	Description	Data type
Order ID	Unique identifier for each order.	String
Order Date	The date the order was placed.	Date
Ship Date	The date of the order was shipped.	Date
Ship Mode	The mode of shipment (e.g., Standard Class, First Class).	String
Customer ID	Unique identifier for each customer.	String
Customer Name	Name of the customer.	String
Segment	Customer segment (e.g., Consumer, Corporate, Home Office).	String
City	City where the order was delivered.	String
State	State or region where the order was delivered.	String
Country	Country where the order was delivered.	String
Postal Code	Postal code of the delivery location.	Number
Market	Market region (e.g., APAC, EU, US).	String
Region	Geographic region (e.g., South, West, Central).	String
Product ID	Unique identifier for each product.	String

Category	Product category (e.g., Furniture, Office Supplies, Technology).	String
Sub-Category	Product sub-category (e.g., Chairs, Phones, Paper).	String
Product Name	Name of the product.	String
Sales	Total sales value for the order.	Number
Quantity	Quantity of products ordered.	Number
Discount	Discount applied to the order.	Number
Profit	Profit generated from the order.	Number
Shipping Cost	Cost of shipping the order.	Number
Order Priority	Priority level of the order (e.g., High, Medium, Low).	String

## Methodology

We have undertaken the following steps to successfully complete this project:

- **Data Cleaning:** We have checked the dataset for null or missing values. It is required because null values in the data can give wrong conclusions. We have ensured that all columns had correct data types and values.
- **Segmentation:** We segmented the data into different sets such as product category, customer segment, region, and year. By doing so, we were able to see trends which are not easily identifiable if we consider all the data at once.
- **Visualization:** We have used charts and graphs for visualization of data in a readable format. Visualization is easy to depict complex information in an efficient and concise format. We have created line charts, bar graphs, maps, and customer segments.

- **Insight Generation:** We examined points which had peak or bottom-of-the-dip performance, i.e., seasonality trends, high or low profitability, and loyal customers.
- **Recommendations:** We have drawn recommendations from the insights. These are actionable conclusions which can be adopted by a business to reduce losses and maximize profitability.

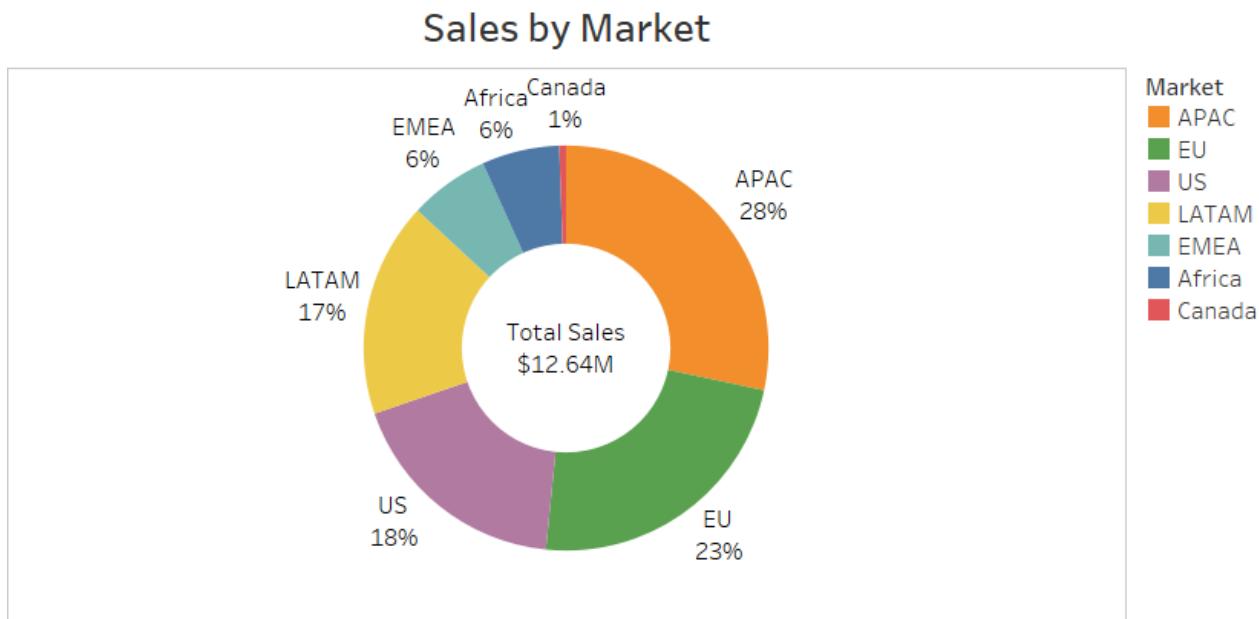
## Intended Audience

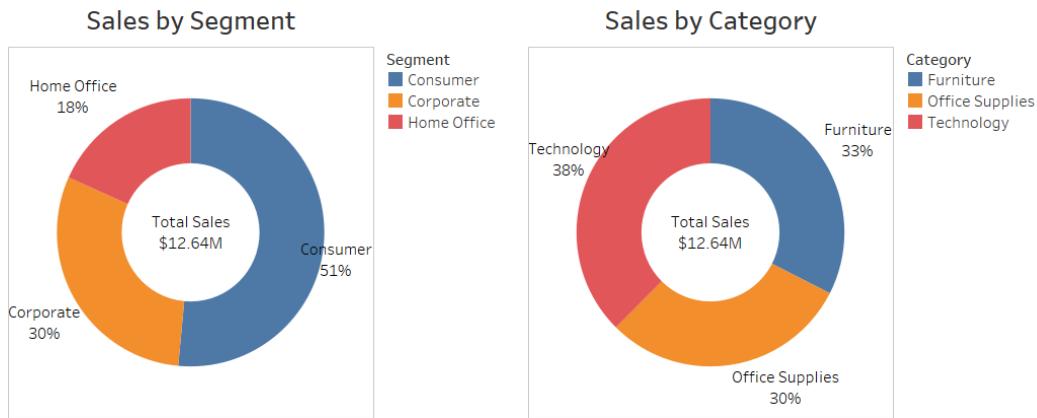
The intended audience is data analytics students, project graders who teach, and business stakeholders such as sales managers, marketing departments, and supply chain planners who can leverage the insights to drive operations and profitability.

## Key Findings

### SALES ANALYSIS

#### 1. Sales distribution based on different markets by Category and Segment.

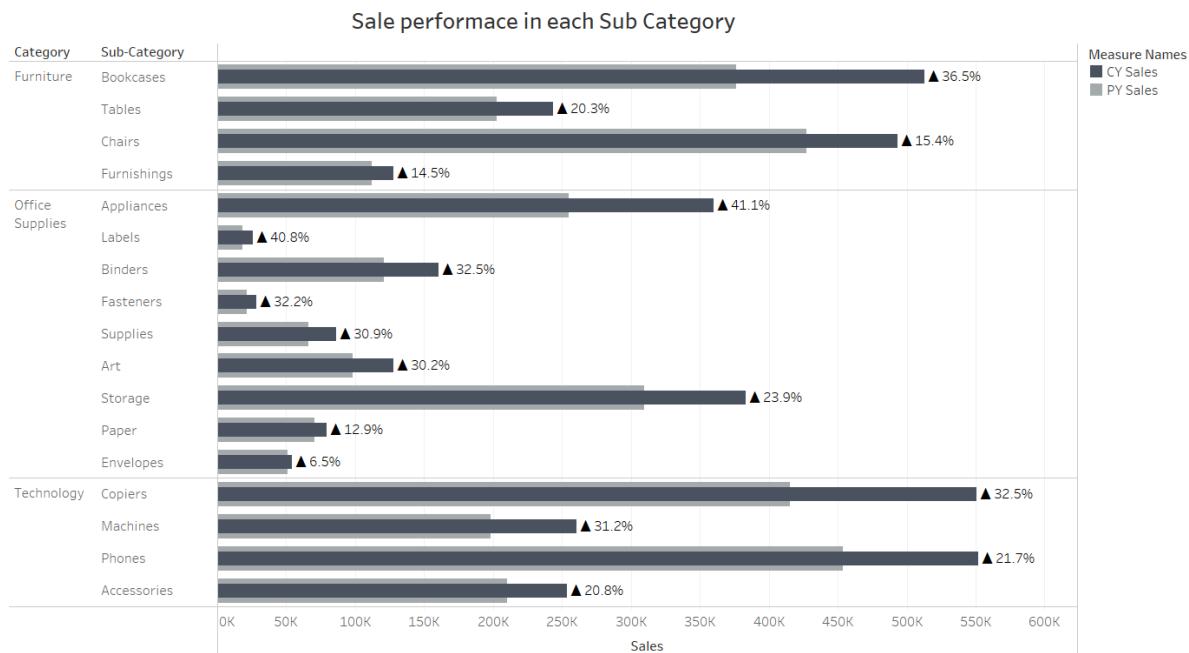




We found that the APAC region has the highest sales (28%). This means that a large part of the company's revenue comes from this region. The Technology category is the top-performing product group, making up 38% of sales. These products are in high demand and provide good profits. The Consumer segment accounts for more than half of total sales, showing that individual customers are very important to the business.

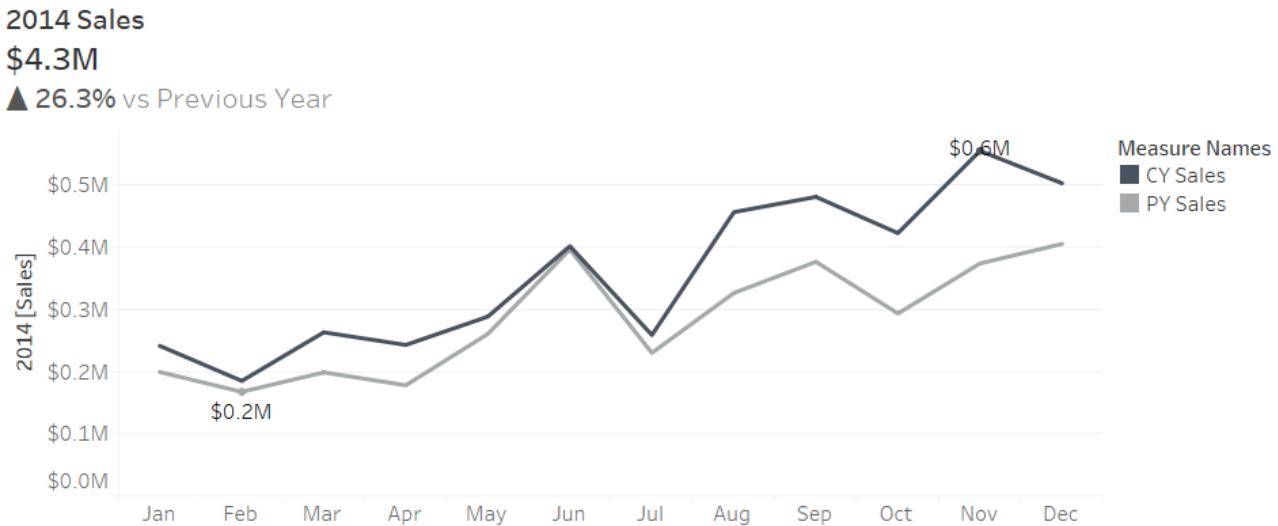
This suggests that the company should continue to invest in the APAC region and the Technology product category. These areas are already performing well and have room for more growth. Focusing on consumer-level sales may also lead to higher revenues.

## 2. Sales Distribution by Sub-Category.

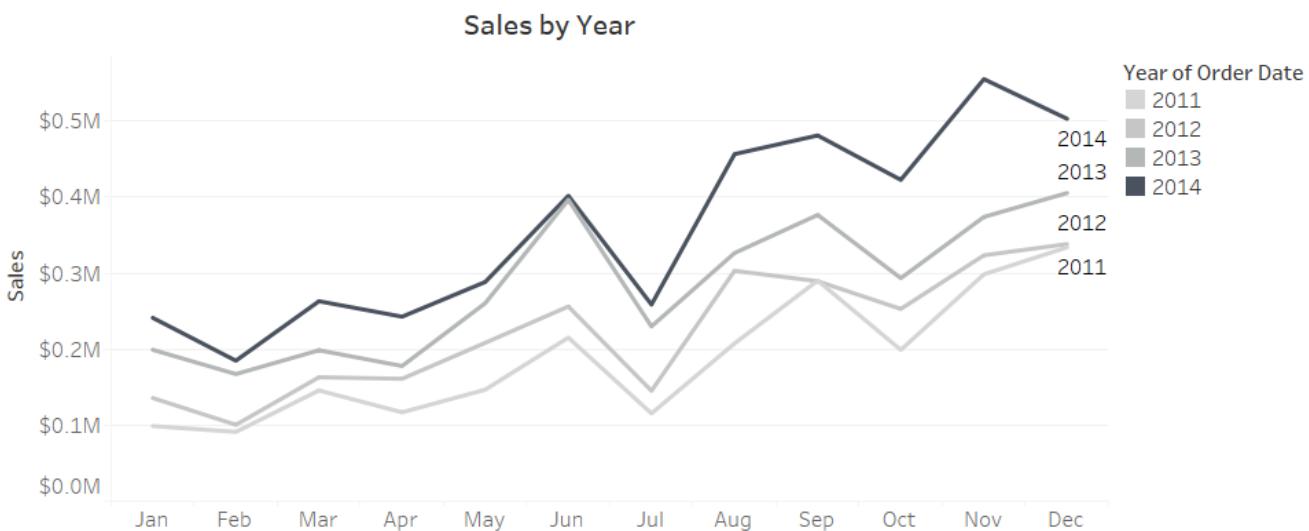


The chart compares sales performance across sub-categories in 2014, showing both current year (CY) and previous year (PY) sales. Most sub-categories experienced significant sales growth, with Appliances, Labels, and Bookcases showing the highest percentage increases.

### 3. Yearly Trends



Sales have been increasing each year from 2011 to 2014. In 2014, total sales were \$4.3 million, with a 26.3% increase from the previous year.



Year-over-year trends show consistent growth from 2011 to 2014, with each year outperforming the last, especially November had the highest sales, likely because of holiday shopping events like Black Friday. This seasonal trend shows that the company makes most of its money in the second half of the year.

Understanding these trends is important for planning. The company can prepare for high-demand months by increasing inventory and marketing. It can also try to boost sales in the first half of the year by offering special deals or launching new products.

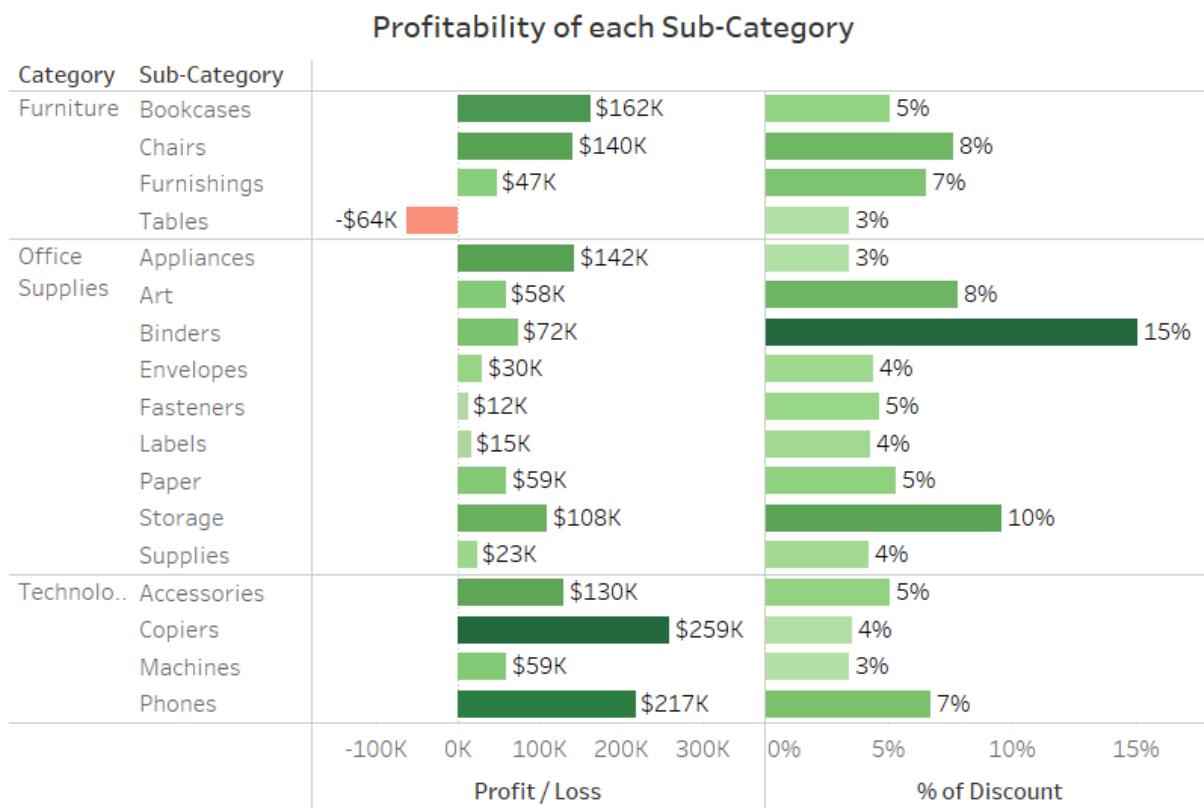
## PROFIT ANALYSIS

### 1. Profit by Category



This bump chart contrasts Technology, Furniture, and Office Supplies with profits. Technology tops the chart, followed by Office Supplies and Furniture which typically takes third place for a long time since 2011. Further in-depth analysis needs to be required in the furniture category, as it has remained the least profitable category for a long time.

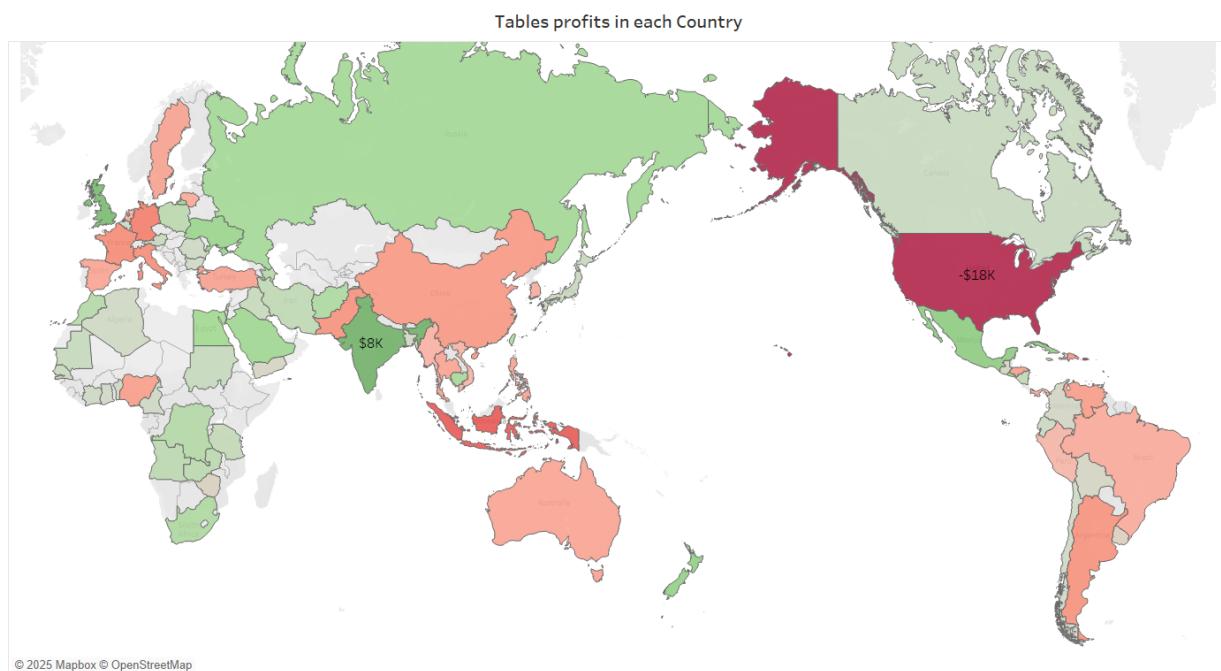
## 2. Profit by Subcategory



One of the most important findings is that the Tables sub-category is losing a lot of money. It has a total loss of around \$64,000. These losses are worse in certain places, especially in New York, Illinois, and North Carolina. This suggests that tables are not popular or not priced correctly in these regions. Selling them in these areas may cost more than the revenue they bring in.

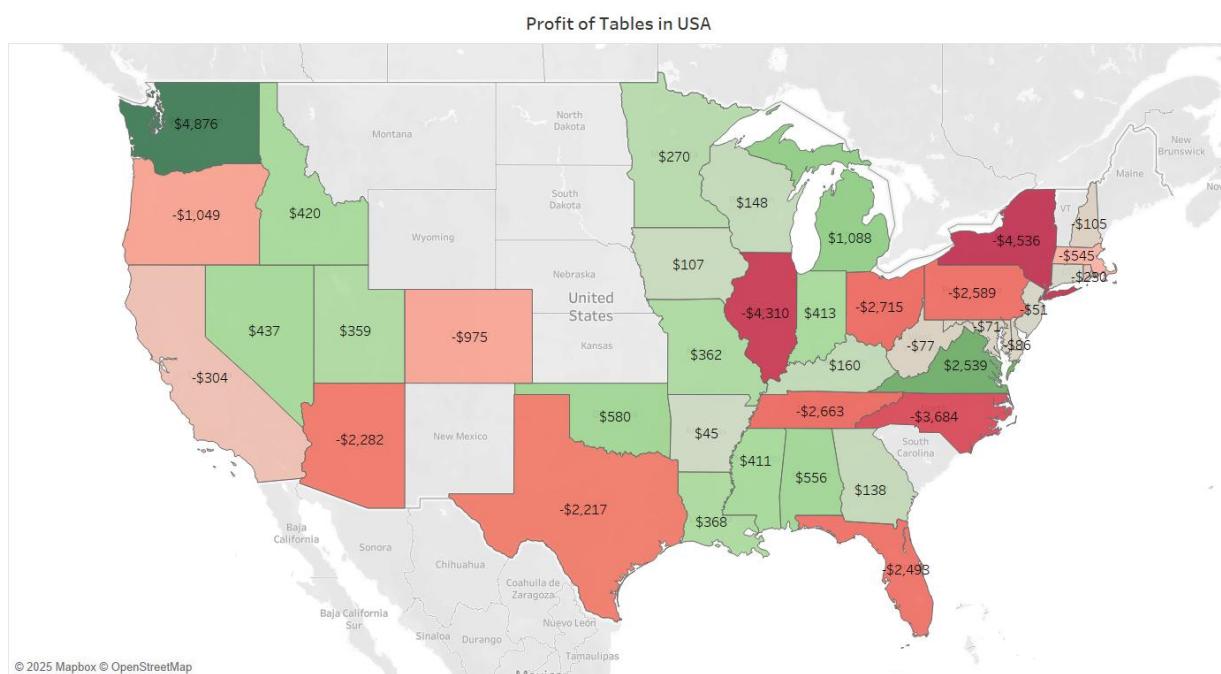
In contrast, products like Phones and Copiers are very profitable. These items have high sales and high profit margins. The company should consider shifting its focus from loss-making products to high-performing ones.

### 3. Profit Distribution of “Tables” Subcategory by Country



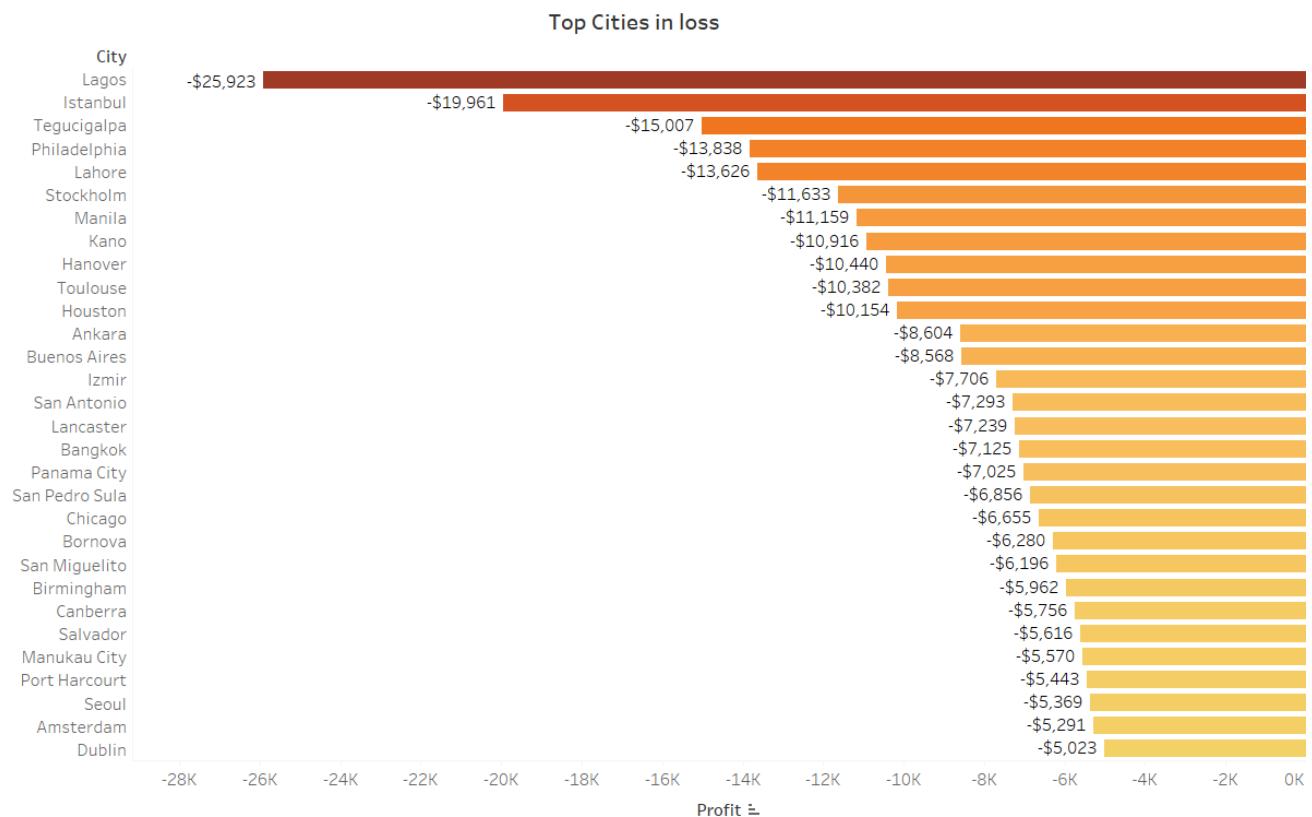
USA is the Country with the highest loss in the “Tables” subcategory of 18,000 dollars.

### 4. “Tables” Profits in USA by State



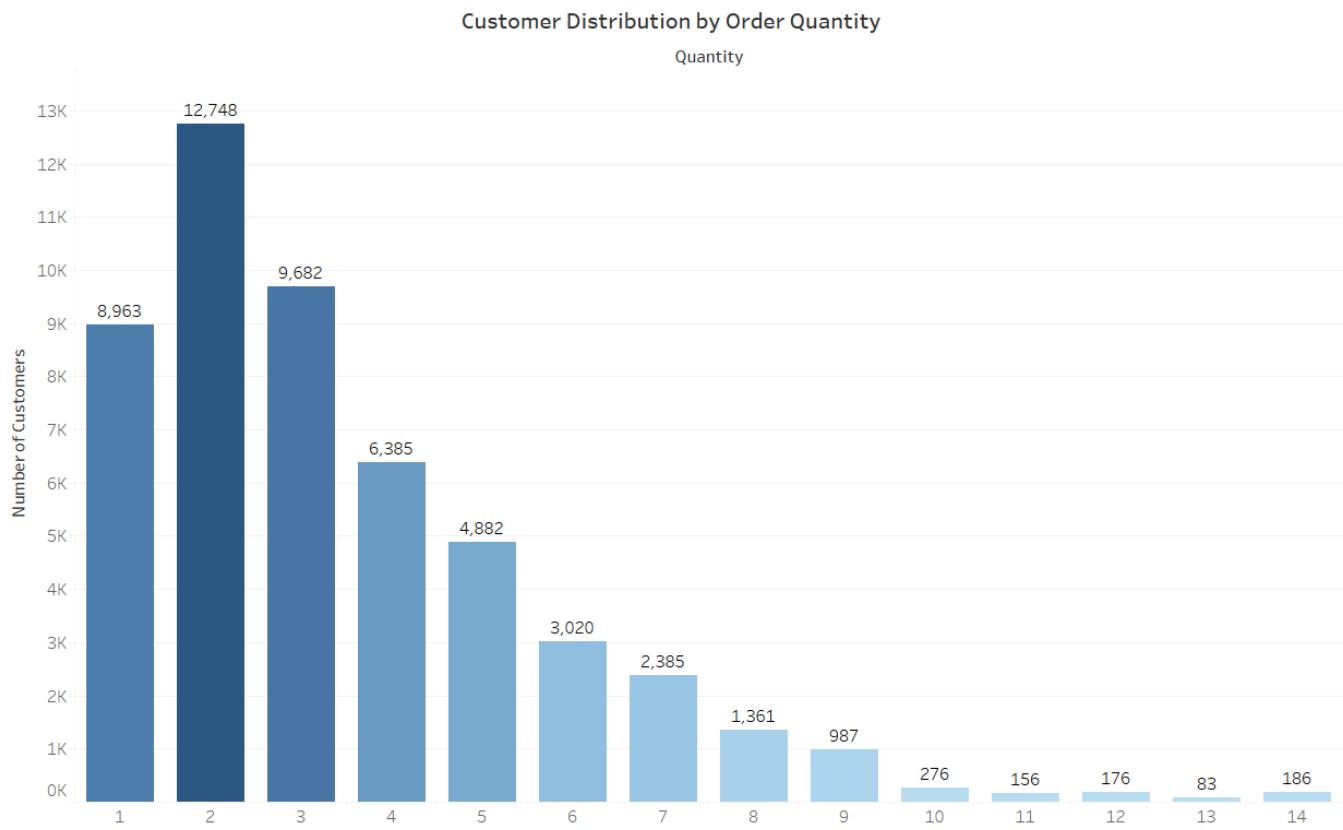
This map shows the profit of tables across different U.S. states. We can clearly see that many states, like Illinois, North Carolina, and New York, have large losses (highlighted in red) and profits in the negative thousands. On the other hand, a few states like California, Michigan, and Florida show positive gains, though these are fewer. Eliminating tables from the states having large losses like Illinois, North Carolina, and New York may reduce losses and put in profits.

## 5. Top Cities in Losses.



Lagos and Istanbul experienced the highest losses, with losses of \$25,923 and \$19,961 respectively. The chart highlights that several cities worldwide faced significant negative profits, with all listed cities incurring losses above \$5,000.

## CUSTOMER ANALYSIS



### Top Customers by Profit

INDEX..	Customer Name	Profit	Sales	No of Orders
#1	Tamara Chand	\$8,787	\$34,218	28
#2	Raymond Buch	\$8,524	\$29,198	25
#3	Sanjit Chand	\$8,106	\$25,603	28
#4	Bill Eplett	\$7,791	\$27,158	37
#5	Hunter Lopez	\$7,658	\$29,664	20
#6	Adrian Barton	\$6,913	\$22,967	33
#7	Susan Pistek	\$6,650	\$28,124	28
#8	Harry Marie	\$6,545	\$27,434	37
#9	Tom Ashbrook	\$6,275	\$35,668	25
#10	Sanjit Engle	\$5,864	\$29,533	36
#11	Bill Shonely	\$5,681	\$22,551	26
#12	Daniel Raglin	\$5,506	\$25,641	33
#13	Ellis Ballard	\$5,466	\$20,188	25
#14	Greg Tran	\$5,165	\$34,472	30
#15	Keith Dawkins	\$5,153	\$24,997	36
#16	Nathan Mautz	\$5,095	\$20,223	32
#17	Patrick Jones	\$4,889	\$13,640	19
#18	John Huston	\$4,830	\$23,587	31
#19	Carlos Daly	\$4,752	\$18,297	27
#20	Patrick O'Donn...	\$4,699	\$21,051	35

Most customers order 1 or 2 items. This is normal for retail businesses, where individual purchases are small. However, a few loyal customers, like Tamara Chand and Raymond Buch, make many purchases and bring in high profits. These customers are very valuable to the business.

Keeping these customers happy can lead to continued profits. The company can create special programs or offers for loyal customers to make sure they stay.



New York City and Los Angeles have the highest number of customers, with 915 and 747 respectively. This insight can help us decide where to invest in new stores or to introduce new products in these cities first to get a wide variety of customer feedback as these cities have a high number of customers.

## **Recommendations**

- **Stop or Reduce Sales of Tables**

Since Tables are losing money, especially in places like New York, the company should consider stopping sales in these areas. It may also want to stop selling Tables entirely if they continue to perform poorly.

- **Reduce Discounts on Low-Profit Products**

Some items like Binders are given big discounts but don't make much profit. Reducing discounts can help improve profits. Discounts should be given only when they lead to increased sales or customer loyalty.

- **Focus on Best-Selling Products**

Phones and Copiers are doing very well. The company should advertise and stock more of these products. This will lead to higher profits and better customer satisfaction.

- **Target Profitable Regions**

More effort should be put into regions like California and New York, which are bringing in the most money. Advertising and delivery services can be improved in these areas.

- **Keep Loyal Customers Happy**

Offer special deals to top customers to keep them coming back. These customers bring in a lot of profit and are more likely to refer to others.

- **Prepare for Seasonal Demand**

Since sales peak in November, the company should plan and have enough stock ready. Holiday campaigns should start early to attract more customers.

- **Re-evaluate Low-Performing Regions**

Areas like Illinois and North Carolina need to be studied more closely. If losses continue, the company should consider closing operations in those regions or changing its strategy.

- **Use Customer Segmentation**

By dividing customers into groups based on how often they buy or how much they spend, the company can create better marketing campaigns. This helps target the right people with the right offers.

## **Conclusion**

This report helps to show how a business can use data to make better decisions. By understanding which products and regions are making or losing money, the company can take action to fix problems. The most important finding is that “Tables” are causing large losses, especially in some regions of the United States like Illinois, North Carolina, and New York. By removing these loss-making items and focusing on profitable products and loyal customers, the business can increase its overall profits.

We also found that Phones and Copiers are performing very well and should be promoted. California and New York are high-performing regions that deserve more attention. At the same time, the company should be careful about where it gives discounts and how it manages low-performing products and regions.

In conclusion, our data analysis and visualization of the Global Superstore Dataset helps us see these patterns clearly. It also helps us explain them to others in a way that is easy to understand. This project shows the value of using data to understand these patterns can help businesses make informed decisions to drive growth.

## References

Harvard Business Review. (2018). How to make better data-driven decisions.

<https://hbr.org/2018/12/how-to-make-better-data-driven-decisions>

McKinsey & Company. (2022). Unlocking growth through analytics.

<https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights>

Tableau Software. (n.d.). Data visualization best practices.

<https://www.tableau.com/learn/articles/data-visualization-best-practices>