**📊 Superstore Sales Data Analysis Report**

**Prepared by:** Nandkumar Bandivdekar  
**Education:** M.Sc. in Statistics, Shivaji University  
**Aspiring Role:** Data Analyst  
**Tools Used:** Python, Pandas, Excel, SQL (basic), Power BI (beginner)

**🎯 1. Objective**

As a student aspiring to start a career in data analytics, the aim of this project was to analyze sales data from a fictional Superstore to gain hands-on experience in data cleaning, analysis, and generating business insights.

**📁 2. Dataset Information**

* **Records:** 9,993
* **Features:** Order ID, Sales, Profit, Category, Region, Discount, Quantity, etc.
* **Time Period:** Multiple years (includes Order and Ship Dates)

**🧼 3. Data Cleaning**

* Converted Order Date and Ship Date to datetime format
* Removed 1 duplicate entry
* No missing values were found
* Dropped columns like Row ID, Postal Code, and Product ID for clarity

**📊 4. Key Findings**

| **Metric** | **Value** |
| --- | --- |
| Total Sales | $2.3 Million |
| Total Profit | $286,000 |
| Avg Discount | 15.6% |
| Avg Quantity/Order | 3.8 units |
| Highest Profit Line | $8,400 |
| Largest Loss | -$6,600 |

**🗂 5. Product & Sales Insights**

* **Top Categories by Sales:**
  + Technology > Furniture > Office Supplies
* **Most Profitable Sub-Categories:**
  + Copiers, Phones, and Accessories
* **Loss-Making Items:**
  + Tables and Bookcases had high sales but low or negative profit, likely due to heavy discounts

**🌎 6. Regional Performance**

* **Top Region:** West (most profitable and highest sales)
* **Low Profit Region:** South
* **Top Performing States:** California, New York, and Washington

**📉 7. Discount and Profit Relationship**

* High discounts (>30%) often resulted in losses
* Moderate discounts (10-20%) were more sustainable
* Suggests a need to manage discount strategy carefully

**💡 8. What I Learned**

* How to clean and analyze real-world datasets using Python
* The importance of visualizing and summarizing data before drawing conclusions
* How business decisions (like discounting) directly impact profits
* Gained confidence in using tools like Pandas, Excel, and basic SQL logic

**📝 9. Conclusion**

This project helped me apply theoretical knowledge from my M.Sc. in Statistics to a practical dataset. I developed a deeper understanding of how data drives decisions in business and how tools like Python and Power BI can help solve real-world problems.

I'm excited to grow further in data analytics and contribute to data-driven organizations like **IBM**.