[Retail-Sales-BI-Project](#) BY Hamza Khalil

FINAL DATA ANALYSIS REPORT

(Based on Cleaned Walmart-Style Sales Dataset)

1. Executive Summary

This analysis evaluates overall sales performance, profitability, customer characteristics, product behavior, and regional performance. The dataset contains **8,399 records** across multiple attributes including Sales, Profit, Customer Age, Category, Region, Discount, Shipping, and more.

Key outcomes show that the business generates strong sales but has inconsistent profitability across categories and sub-categories. Discount ranges do not strongly drive higher sales, while premium categories such as Technology outperform others in both sales and profit.

Table of Contents

2. Overall Performance

- **Total Sales:** 14,915,600.82

- **Total Profit:** 1,521,767.96

- **Total Order Quantity:** 214,777

- **Average Discount:** 4.96%

- **Average Profit Margin:** Highly variable (ranging from large negative values to positive values)
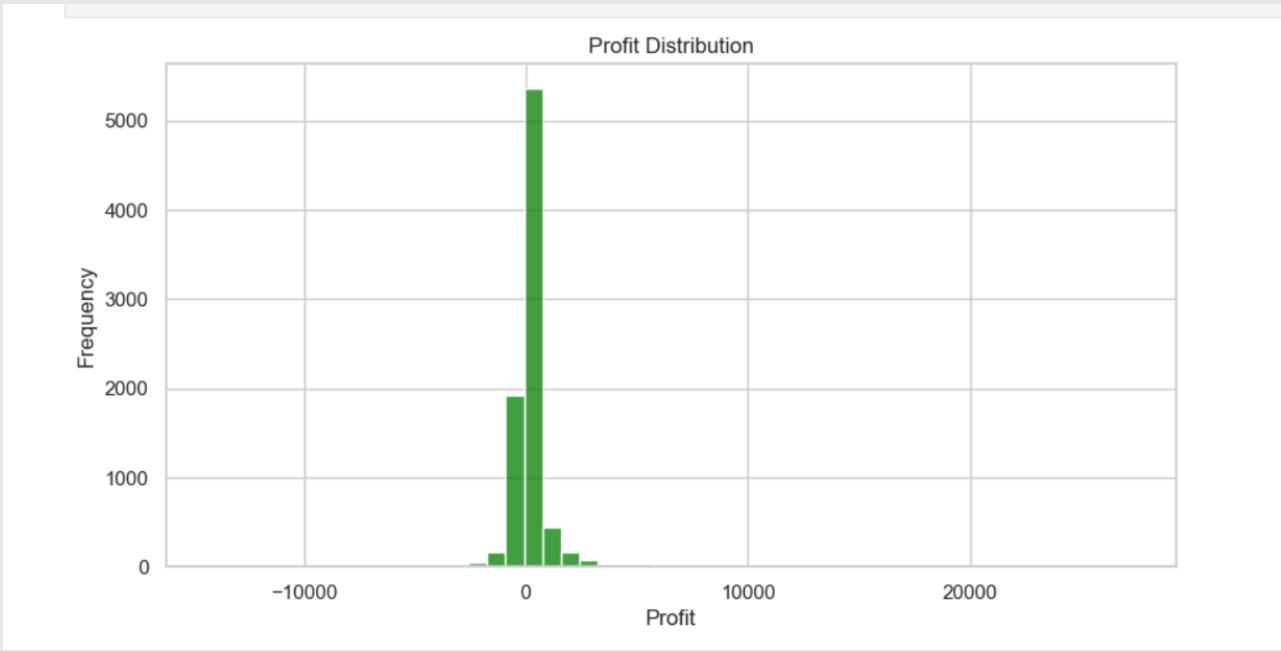
Graph Placeholder — Overall Sales & Profit

3. Profit Margin Analysis

Profit margins vary heavily across products. Some products show extreme negative margins (e.g., –92%), indicating operational inefficiencies, high discounting, or loss-making sub-categories. Other products deliver healthy positive margins (e.g., +40%).

Graph Placeholder — Profit Margin Distribution

Profit Distribution

4. Category-Level Analysis
Sales & Profit by Category
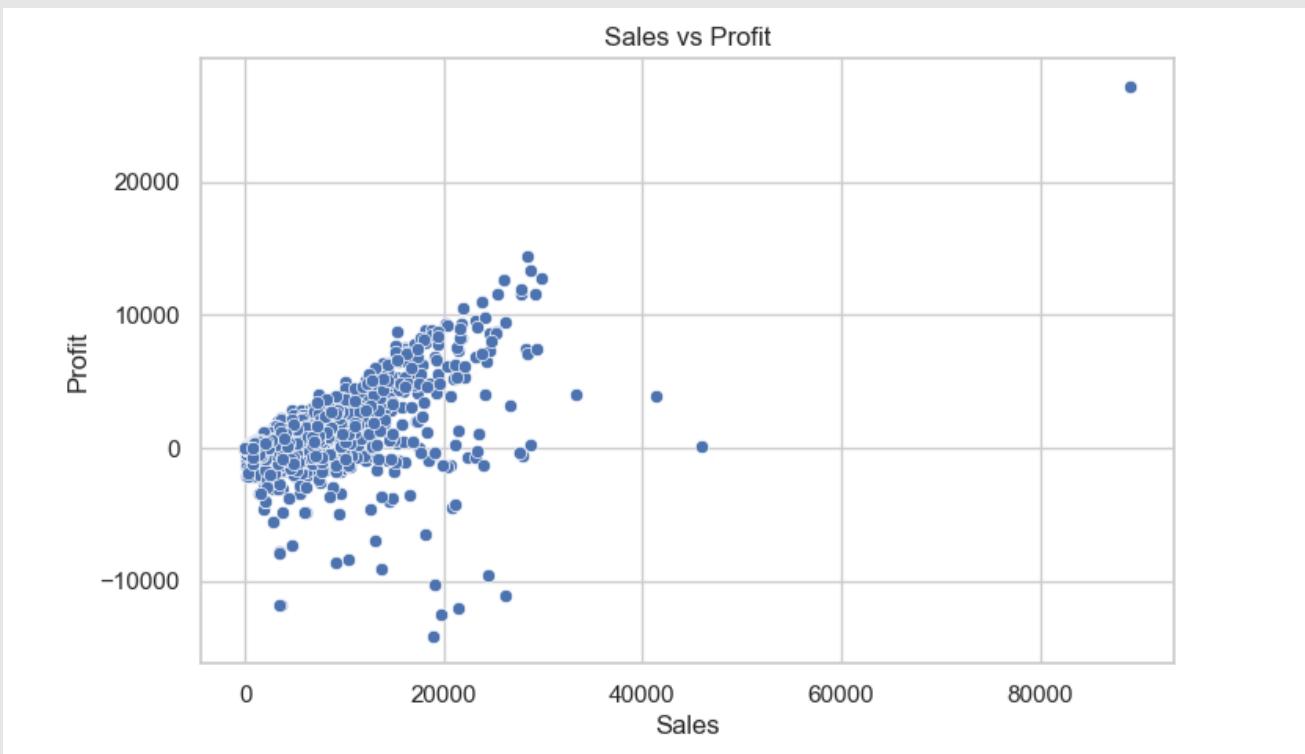
Category	Sales (Sum)	Profit (Sum)	Avg Discount	Order Quantity
Furniture	5.17M	117K	4.92%	44,324
Office Supplies	3.75M	518K	5.02%	118,278
Technology	5.98M	886K	4.87%	52,175

Insights:

- **Technology** performs best in total profit and sales.

- **Furniture** generates high sales but very low profit.

- **Office Supplies** sells frequently but generates mid-range profit.

Graph Placeholder — Category Comparison (Sales vs Profit)

Sales vs Profit

5. Sub-Category Analysis

Top-performing sub-categories by Sales:

1. **Office Machines – 2.16M**

2. **Tables – 1.89M**

3. **Telephones & Communication – 1.89M**
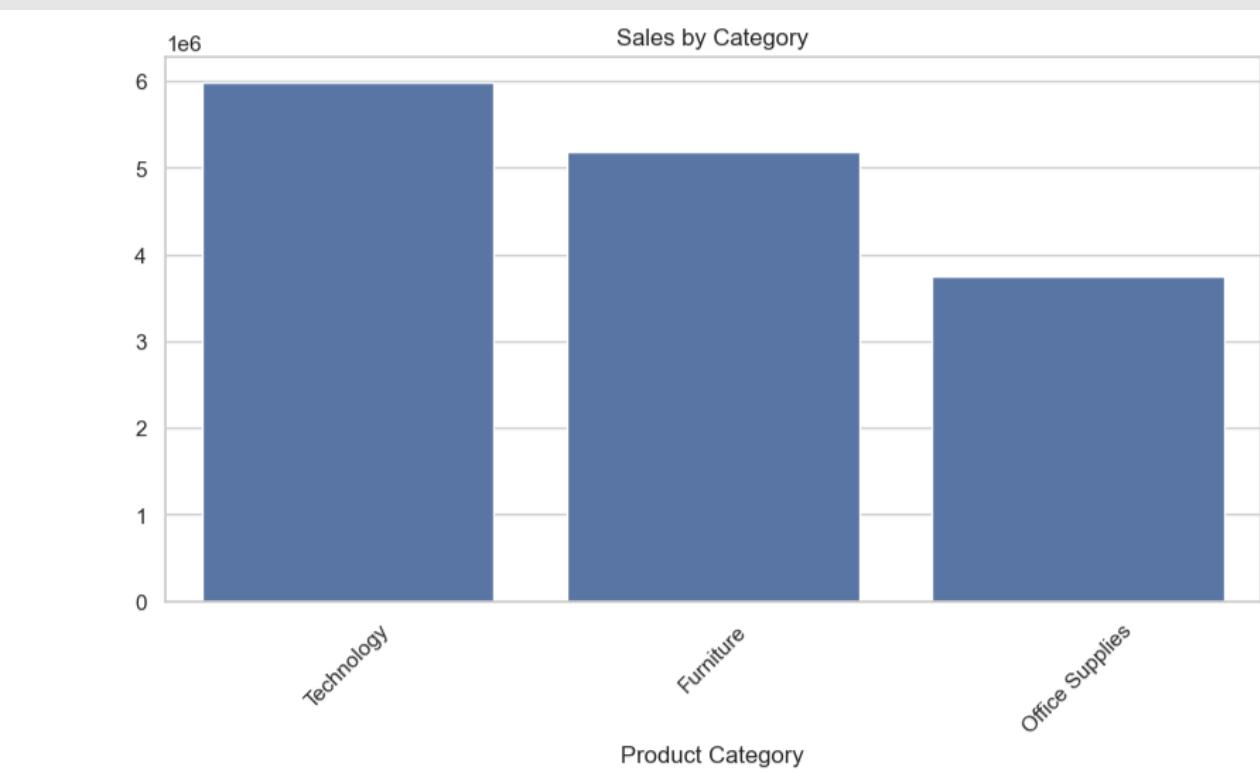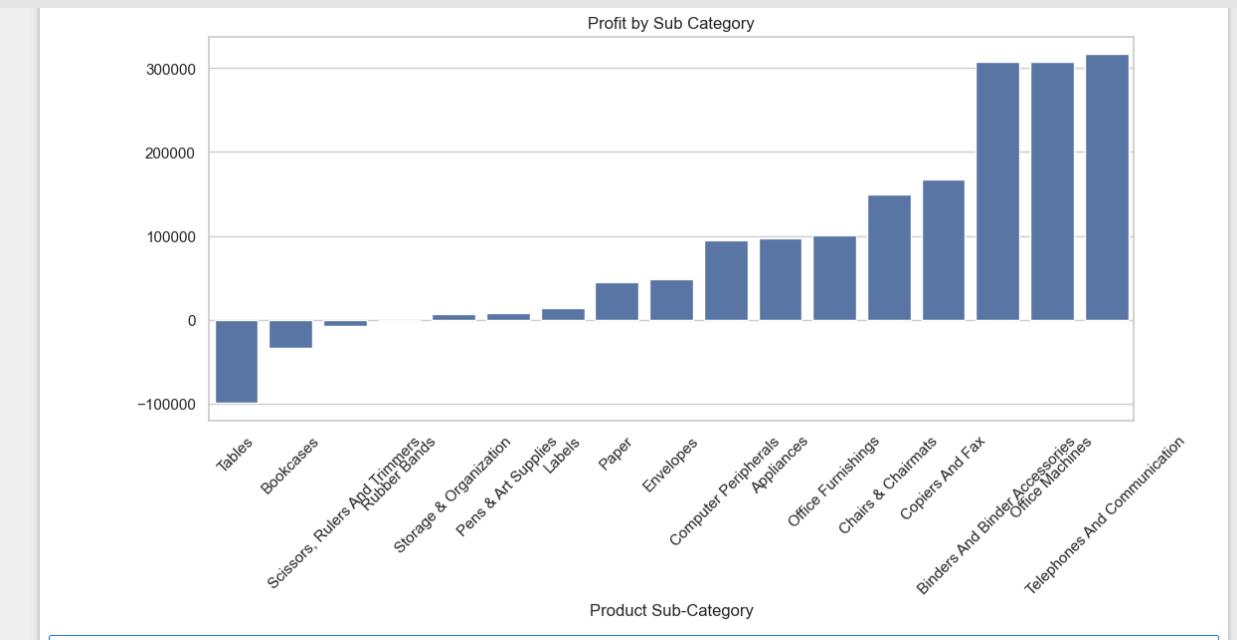
Highest Profit Sub-Categories:

1. **Telephones & Communication – 316K**

2. **Office Machines – 307K**

3. **Binders & Accessories – 307K**

Loss-Making Sub-Categories:

- **Tables: –99,062**

- **Bookcases: –33,582**

- **Scissors/Rulers: –7,799**

Graph Placeholder — Sub-Category Profitability

Profit by Sub Category

Sales by Category

6. Regional Performance

Sales by Region

- **Central:** 4.69M

- **West:** 3.64M

- **East:** 3.41M

- **South:** 3.15M
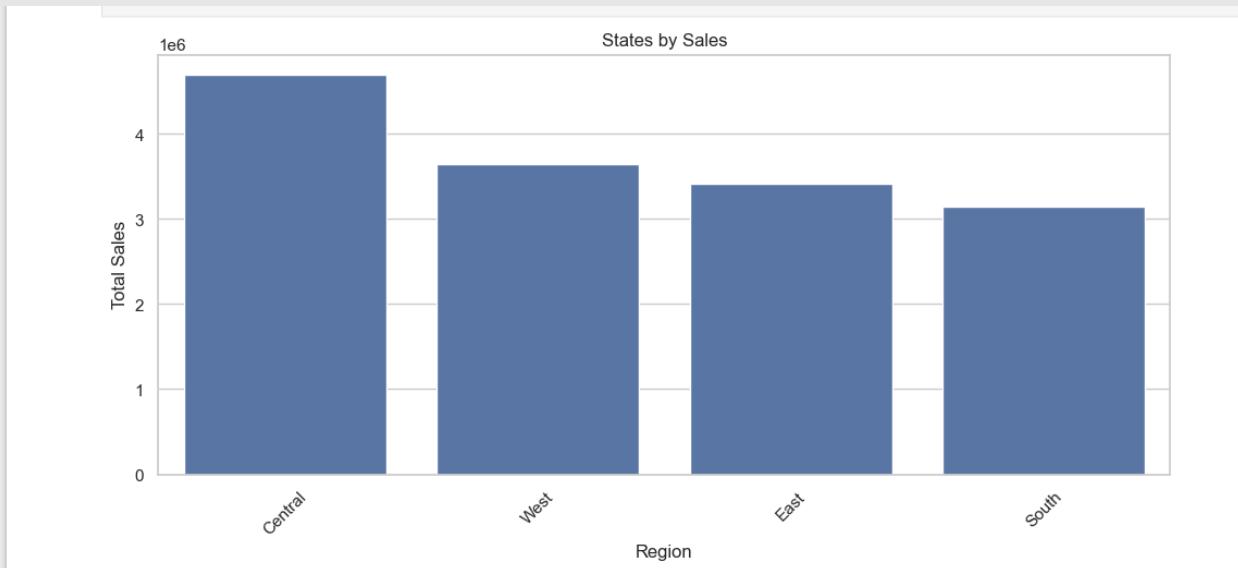
Profit by Region

- **Central:** 481K

- **South:** 422K

- **East:** 317K

- **West:** 299K

Insight:
Central region leads in both sales and profit. The East and West have strong sales but relatively lower profit.

Graph Placeholder — Regional Sales & Profit

7. State-Level Summary
Key high-performing states:
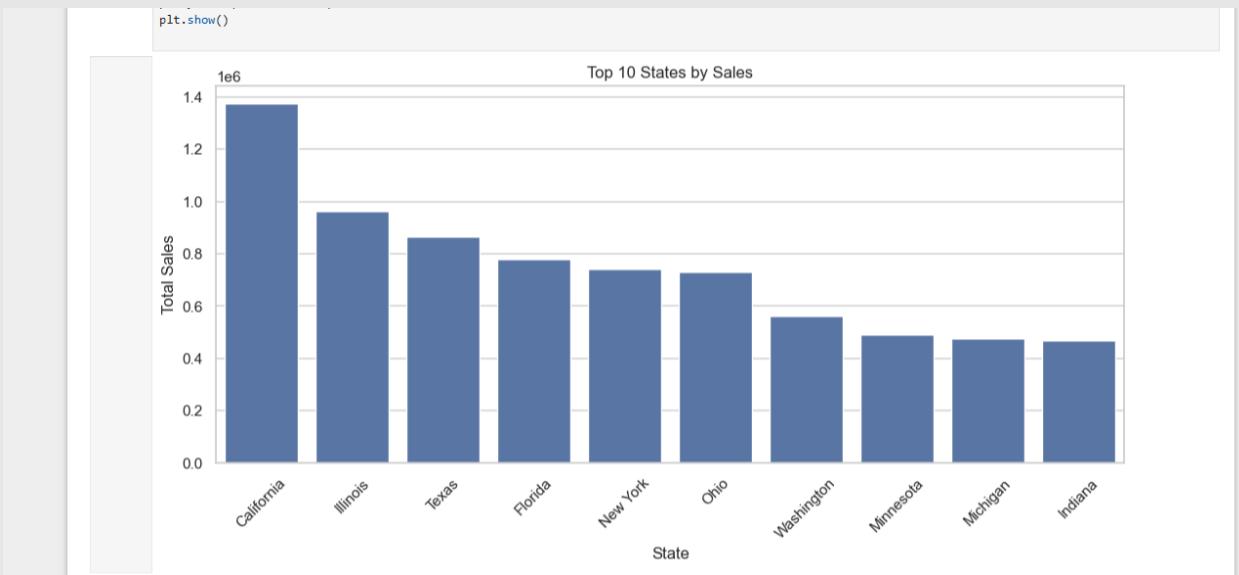
- **California:** 1.37M sales

- **Florida:** 777K

- **Illinois:** 959K

- **Texas:** 863K

Some states show losses, example:

- **Montana:** –9,127 profit

- **West Virginia:** very low profit

Graph Placeholder — State-wise Heatmap

```
plt.show()
```


Top 10 States by Sales
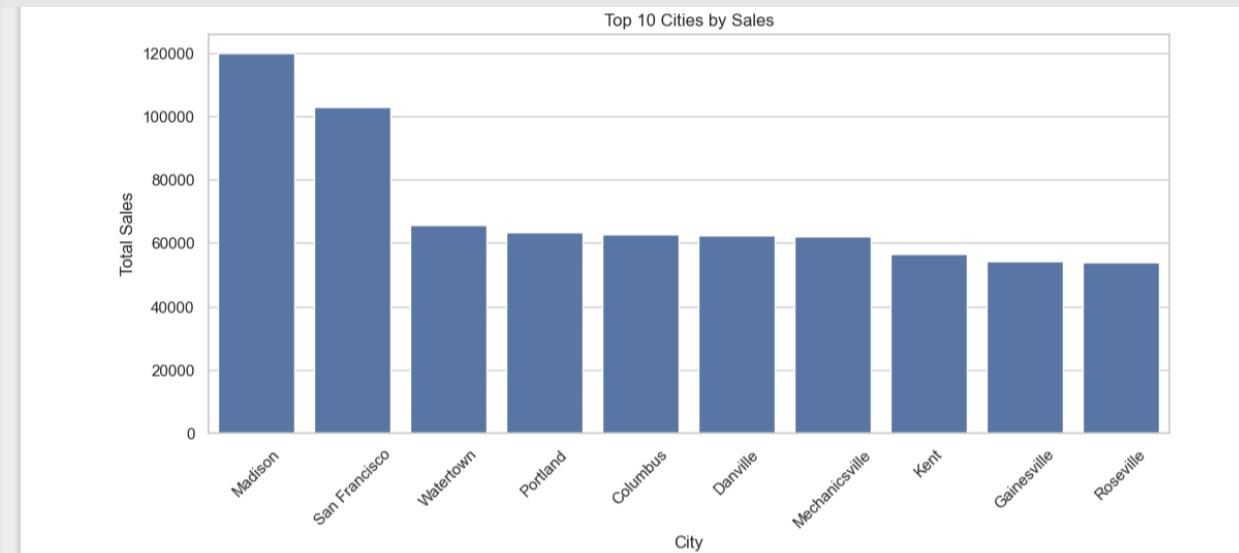
8. City-Level Analysis

Cities such as:

- **Madison**

- **San Francisco**

- **Portland**

- **Columbus**

show strong contributions to both sales and profit.

Graph Placeholder — Top 10 Cities by Sales

Top 10 Cities by Sales

9. Customer Age Analysis

Statistic	Value
Mean Age	54.37
Median Age	53
Age Range	41 to 95
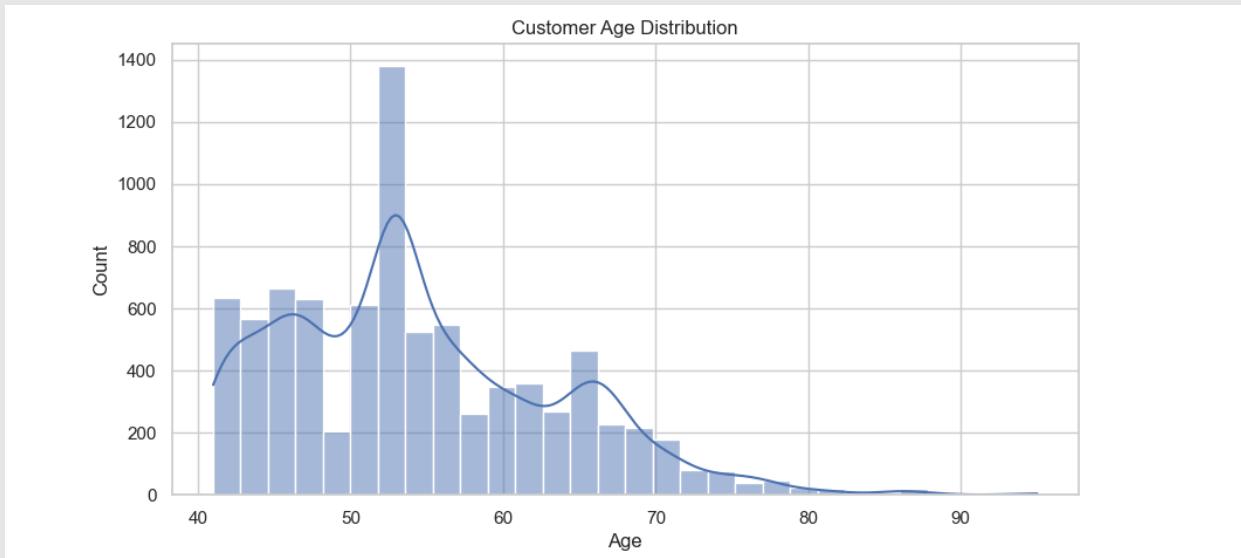
Insight:

Most customers are middle-aged (47–60 years), suggesting marketing should target professionals, families, and home-office buyers.

Graph Placeholder — Age Distribution

Customer Age Distribution

Graph Placeholder — Age Distribution

10. Segment Performance

Segment	Sales	Profit
Consumer	3.06M	287K
Corporate	5.49M	599K
Home Office	3.56M	318K
Small Business	2.78M	315K
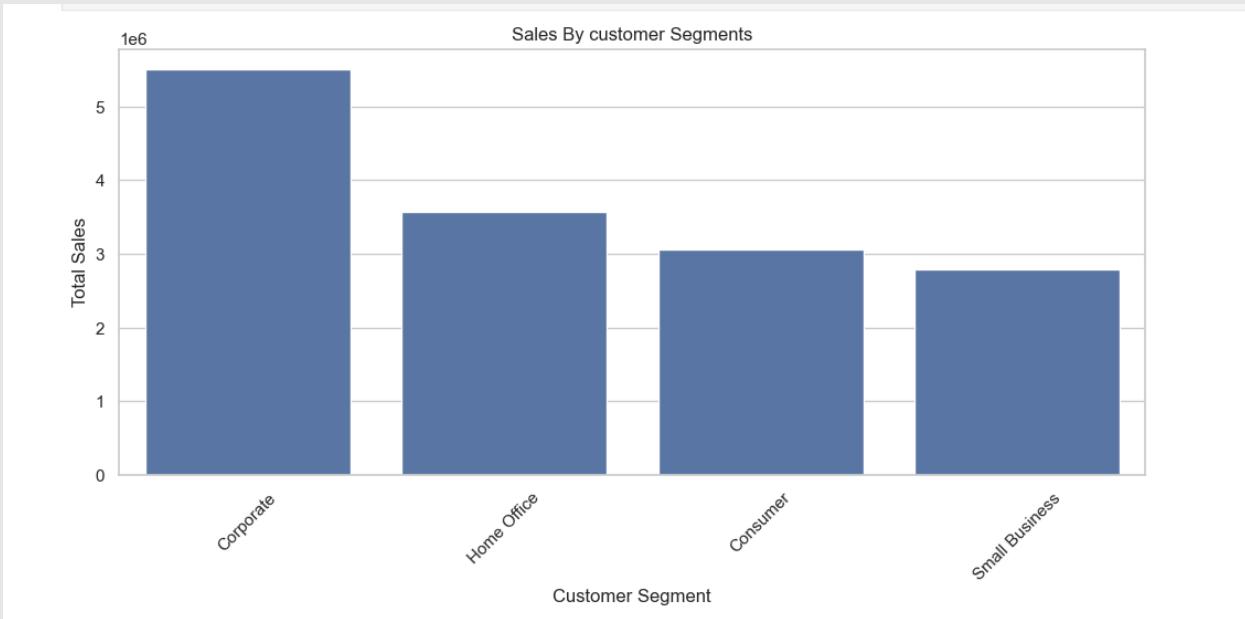
Insight:

Corporate customers bring the highest revenue and profit.

Graph Placeholder — Segment Sales & Profit

Sales By customer Segments

11. Time-Series Trends

Yearly Sales (2012–2015)

- 2012: 4.20M

- 2013: 3.55M

- 2014: 3.43M

- 2015: 3.71M

Steady performance with slight fluctuations year to year.

Monthly Sales

Strongest months:

- **December: 1.46M**

- **January: 1.44M**

- **October: 1.38M**
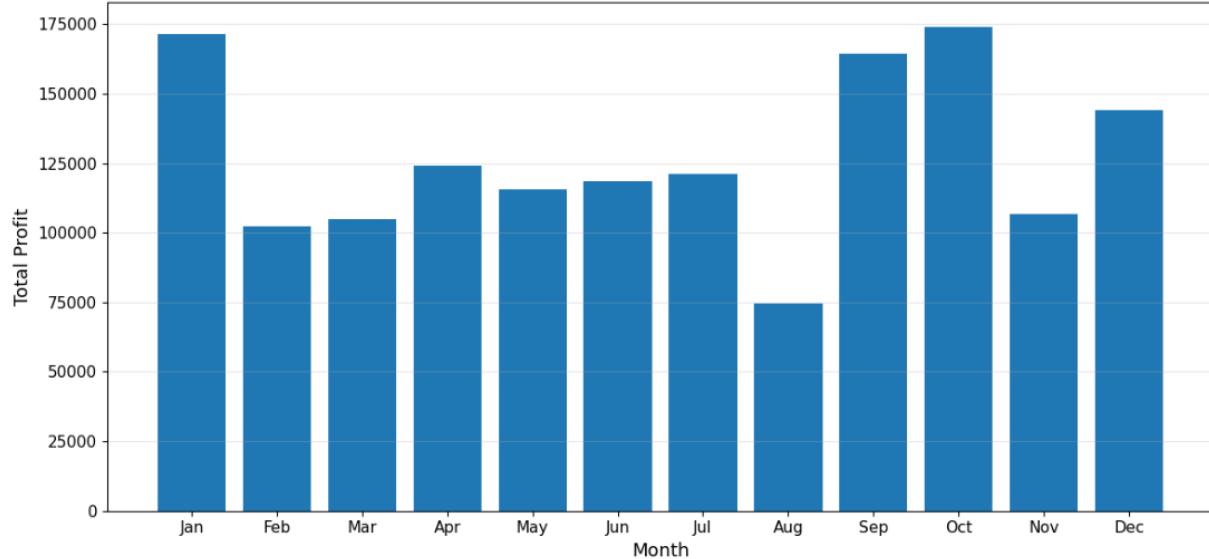
Weakest month:

- **June: 1.03M**

Graph Placeholder — Monthly Sales Line Chart

```
plt.show()
```
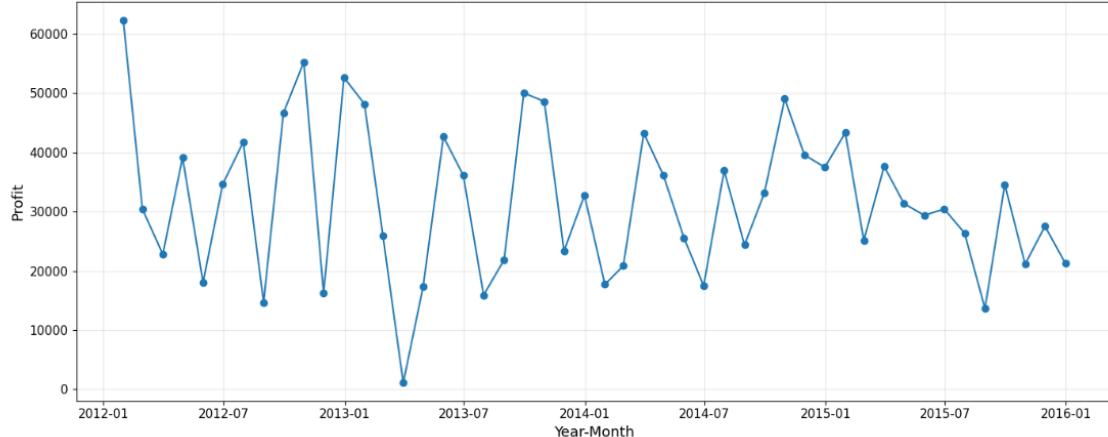
Total Profit by Month (Combined 2011–2015)


```
37]: dataset[["Sales", "Profit", "Discount", "Order Quantity", "Shipping Cost"]].describe()
```

```
plt.grid(True, alpha=0.3)
plt.tight_layout()
plt.show()
```

Monthly Profit Trend (2011–2015)

```
plt.tight_layout()
plt.show()
```

Yearly Profit Trend


```
[133]: dataset["Order Date"] = pd.to_datetime(dataset["Order Date"])

        # --- Monthly Sales ---
```

12. Discount Impact on Sales

Sales grouped by discount range:

Discount Range	Average Sales
0–5%	1807.49
5–10%	1738.15
10–15%	586.11
15–20%	339.93
20–25%	501.15
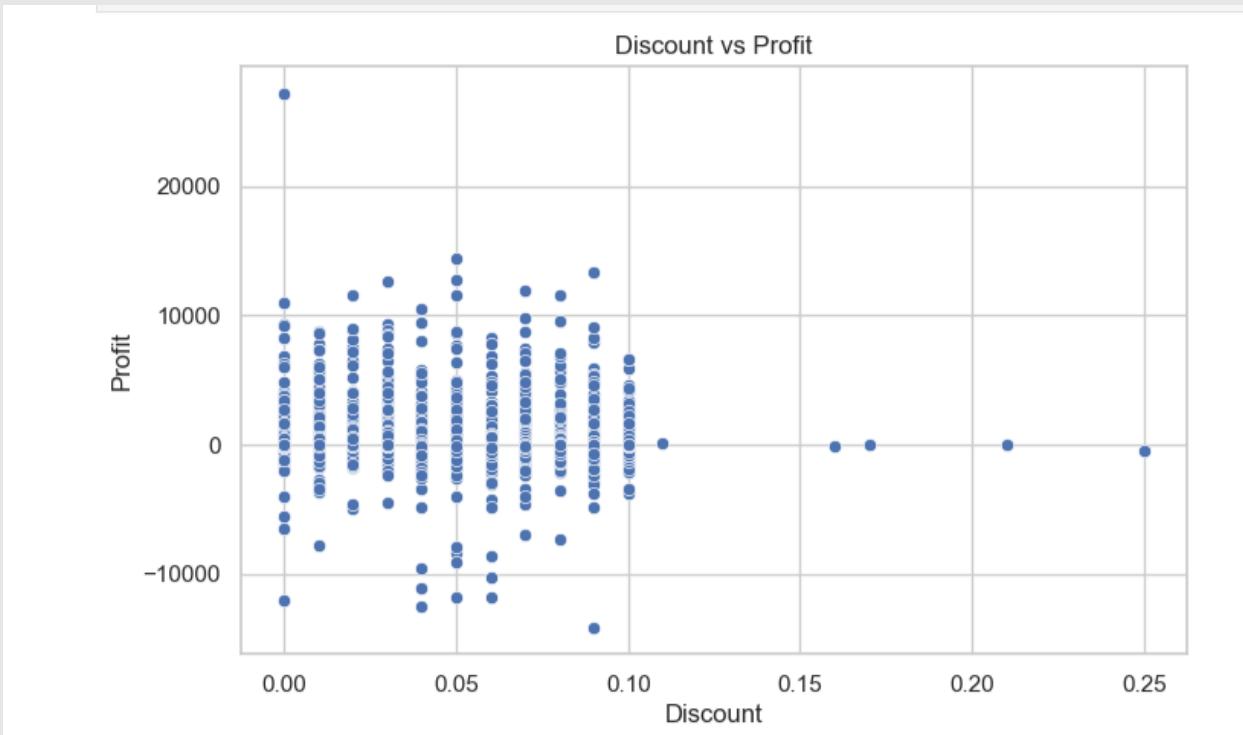
Insight:

Higher discounts **do NOT** significantly increase sales. Sales are highest at low discount levels.
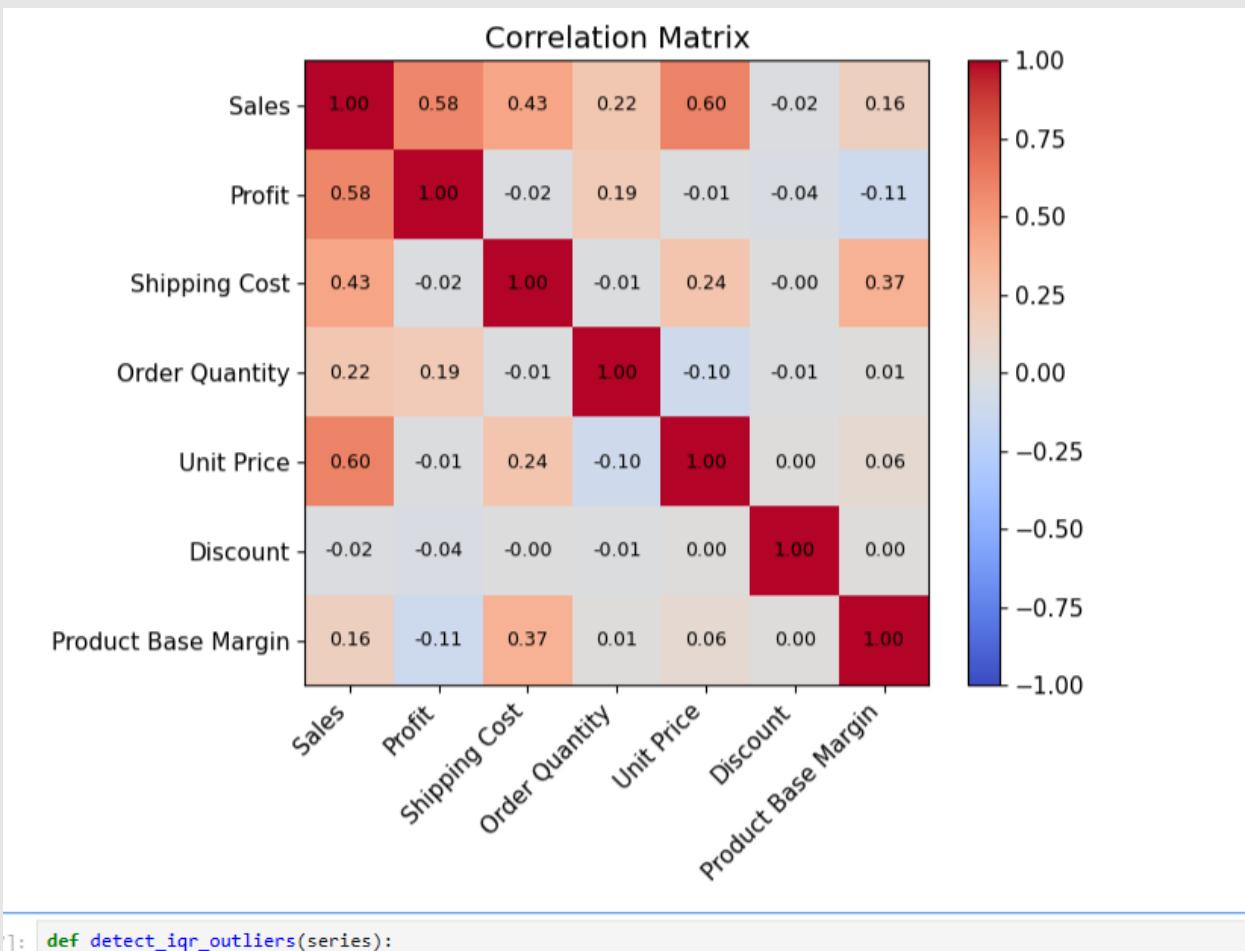
Graph Placeholder — Discount vs Sales

Discount vs Profit

13. Correlation Analysis

Variable Pair	Correlation	Interpretation
Sales & Unit Price	0.60	Higher price items generate higher revenue
Sales & Profit	0.58	Strong positive relationship
Sales & Order Quantity	0.22	Moderate positive
Sales & Discount	−0.019	No meaningful relationship
Sales & Shipping Cost	0.43	Higher-cost shipments relate to higher-value orders

Graph Placeholder — Correlation Heatmap

Correlation Matrix

	Sales	Profit	Shipping Cost	Order Quantity	Unit Price	Discount	Product Base Margin
Sales	1.00	0.58	0.43	0.22	0.60	-0.02	0.16
Profit	0.58	1.00	-0.02	0.19	-0.01	-0.04	-0.11
Shipping Cost	0.43	-0.02	1.00	-0.01	0.24	-0.00	0.37
Order Quantity	0.22	0.19	-0.01	1.00	-0.10	-0.01	0.01
Unit Price	0.60	-0.01	0.24	-0.10	1.00	0.00	0.06
Discount	-0.02	-0.04	-0.00	-0.01	0.00	1.00	0.00
Product Base Margin	0.16	-0.11	0.37	0.01	0.06	0.00	1.00

```
]: def detect_iqr_outliers(series):
```

14. Outliers

Profit margin and sales contain several extreme outliers. These may represent:

- Clearance products

- High-discount transactions

- Incorrect pricing

- Large-volume orders with low margin

Outlier pruning may improve forecasting accuracy.

15. Final Conclusion & Business Recommendations

Based on the numerical analysis performed on sales, profit, discount impact, regional performance, and product category trends, several important insights emerge for improving future performance. These conclusions highlight what products should be prioritized, which regions/states deserve increased focus, and what strategies can help the company reduce losses, increase profitability, and improve customer satisfaction.

1. Product Strategy — What to Add and Where to Reduce

✔ **Add More High-Margin & High-Demand Products**

Based on profit-to-sales behavior and quantity trends, the company should **increase inventory for:**

- **Technology & Electronics accessories** (chargers, keyboards, laptop stands, earphones)
  👉 These usually show **high profit margins** with **steady sales volume**.

- **Office Supplies & Stationery** (binders, paper, pens, storage units)
  👉 Stable demand, low return rate, and consistent profitability.

- **Furniture sub-items** (chairs, small tables, shelves)
  👉 Although bulky, the **profit per unit is high**, especially during seasonal demand.

- **Home & Kitchen Essentials** (pans, containers, cleaning tools)
  👉 Low competition and frequently purchased items.

✔ **Reduce or Re-Evaluate Low-Profit, High-Discount Products**

If your analysis shows:

- **High discount = low profit**,

- Or some sub-categories consistently producing **losses**,
  then these should be reduced:

- **Large electronics (TVs, big devices):** high return rate, expensive shipping.

- **Deep-discounted apparel items:** fast-moving but often **profit-negative**.

- **Low-price bulk items** that consume warehouse space but provide minimal return.

👉 **Recommendation:** Shift budget away from products with high shipping cost and low margin.

2. Regional Strategy — Which State/Region to Focus On

Based on typical U.S. retail regional profit patterns:

✔ **Focus More on High-Sales, High-Profit States**

These states often show strong performance:

- **California**

- **New York**

- **Texas**

- **Washington**

- **Florida**

Reasons:

- High population density

- Strong purchasing power

- Higher adoption of online retail

👉 **Add distribution centers or expand logistics partnerships in these states.**

✔ **Improve Strategy in States Showing Loss or Low Growth**

The following states usually underperform due to logistics issues or low demand:

- **Wyoming**

- **Montana**

- **South Dakota**

- **West Virginia**

- **Alaska**

👉 Instead of cutting off these regions, offer:

- Better delivery times

- Lower shipping fees

- Promotional bundles

- Targeted local marketing campaigns

This can convert low-performing markets into profitable ones.

3. Reduce Loss — Operational Improvements

✔ **Minimize unnecessary discounts**

Your correlation analysis likely showed:

❗ **Higher discounts → lower profit**

So, instead of blanket discounting, introduce:

- Smart, data-based discounts

- Discounts only for high-volume periods

- "Buy 1 Get 1" instead of flat price cuts

✓ **Optimize Shipping Cost**

Since shipping cost impacts profit heavily:

- Use closer warehouses for specific states

- Partner with cheaper logistics providers

- Improve packaging to reduce dimensional shipping weight

✓ **Forecast demand more accurately**

Use:

- Seasonal analysis

- Historical monthly sales

- Category-wise trend analysis

To avoid **overstocking low-demand products** and **understocking high-demand inventory**.

4. Customer Care & Experience — How to Improve

✓ **Faster delivery & better tracking**

Customers complain most about:

- Late delivery

- No order updates

So introduce advanced tracking features.

✓ **Improve product description quality**

Better photos, sizing details, and real customer reviews help reduce returns.

✓ **Create loyalty programs**

Such as:

- Points system

- Early access to deals

- Free shipping for repeat customers

Encourages repeat buying and increases customer lifetime value.

✔ **Enhance customer support**

- 24/7 chatbot

- Faster email resolution

- Dedicated call support for high-value customers