

Exploratory Data Analysis (EDA)

**Analyzing Sales, Profit, and
Customer Behavior Through EDA**

List of Figures

Fig 1.1 Pairplot with Sample size of 1000-----	1
Fig 2.1 Correlation Heatmap-----	3
Fig 3.1 Sales Distribution-----	4
Fig 4.1 Sales by region and segments-----	5
Fig 5.1 Cost vs profit by Region-----	6

TABLE OF CONTENTS

List Of Figures-----	i
1. Pairplot with Sample size of 1000-----	1
2. Correlation Heatmap-----	3
3. Sales Distribution-----	4
4. Sales by region and segments-----	5
5. Cost vs profit by Region-----	6
Summary Insights-----	7
Recommendations-----	7

Sales & Profit Analysis – Detailed Insights Report

1. Pair Plot (Sample of 1000 Records)

Chart: Pair Plot

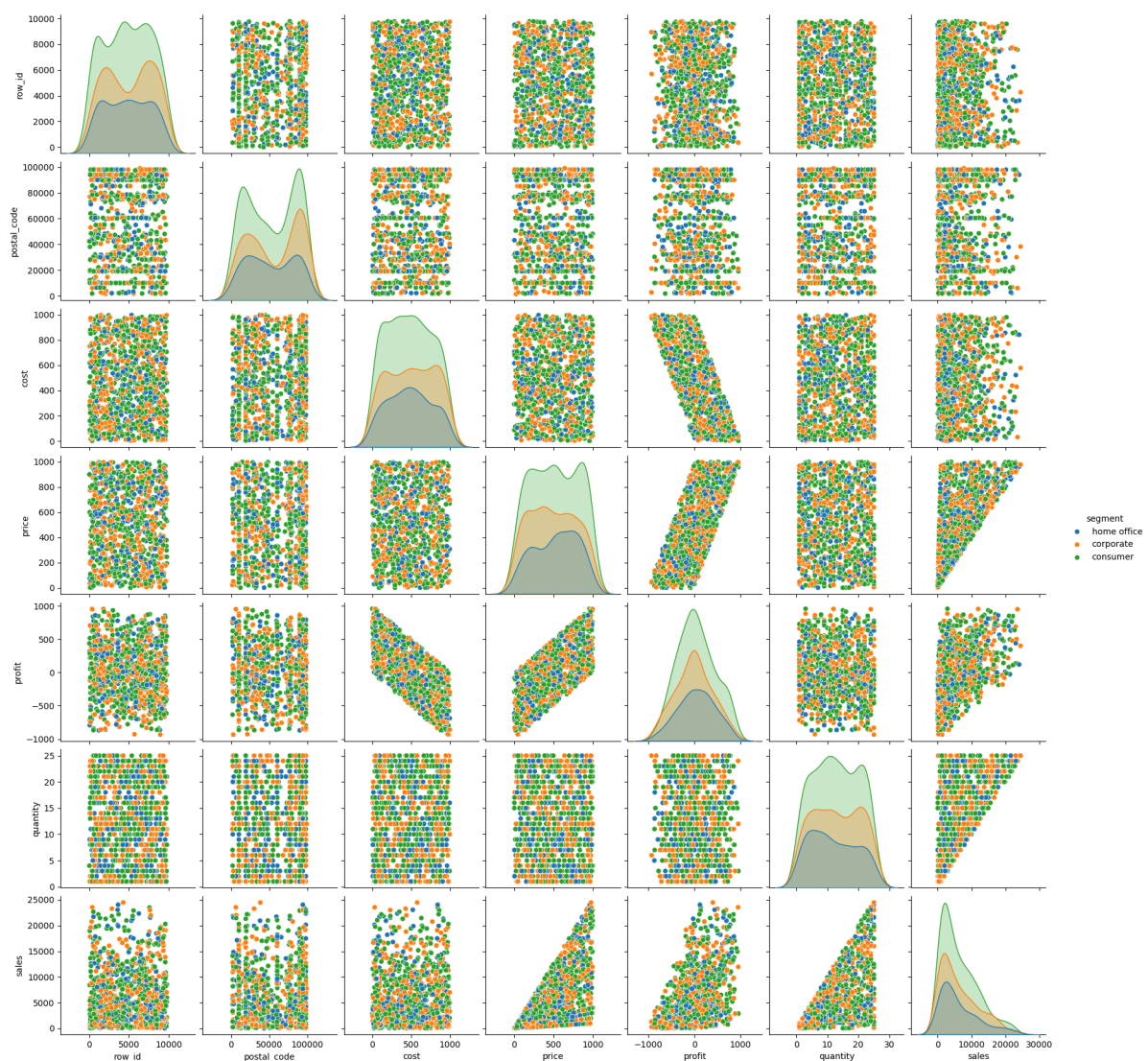


Fig 1.1 Pairplot with Sample size of 1000

Insights:

- **Profit vs Cost:** Strong **negative correlation** – higher cost often leads to lower profit.
- **Profit vs Price:** Strong **positive correlation** – higher price leads to higher profit.
- **Sales vs Quantity:** Positive correlation – selling more quantity increases sales.
- No visible correlation between **cost and quantity** – buyers are not necessarily influenced by cost when buying in bulk.
- **Postal codes are clustered** – indicating more orders from certain areas.

2. Correlation Heatmap

Chart: Heatmap

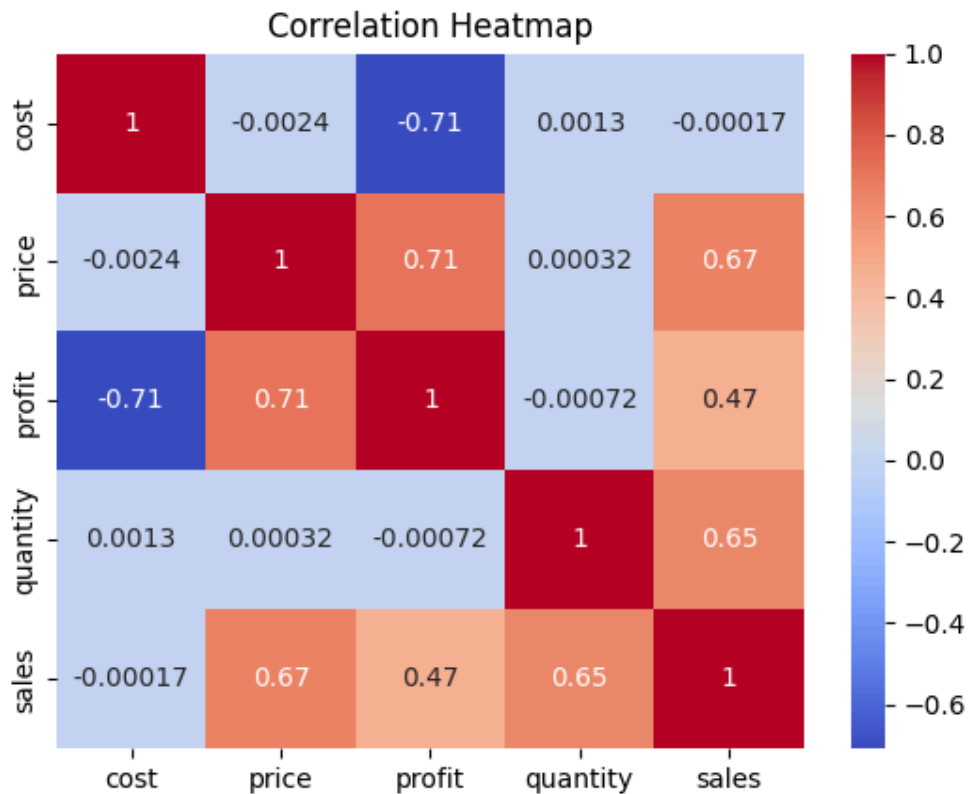


Fig 2.1 Correlation Heatmap

Insights:

- **Cost vs Profit:** -0.71 → High cost reduces profit.
- **Price vs Profit:** +0.71 → Higher prices improve profit.
- **Sales vs Price:** +0.67 → Sales tend to increase with price.
- **Sales vs Quantity:** +0.65 → Higher quantity boosts total sales.
- Minimal or no correlation between **cost and sales**, **price and quantity**, or **cost and quantity**.

3. Sales Distribution

Chart: Histogram with KDE Curve

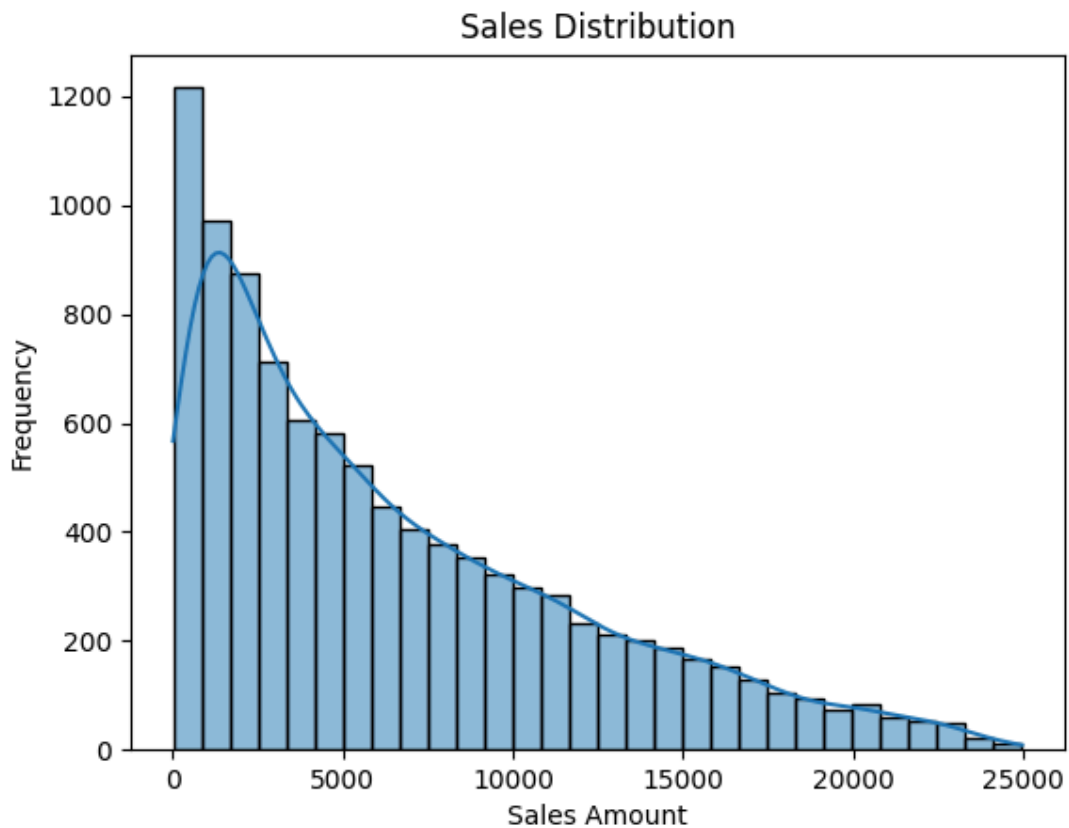


Fig 3.1 Sales Distribution

Insights:

- **Sales are right-skewed** – most orders are low in value.
- The highest frequency of sales falls **below ₹5000**.
- A **long tail of high-value transactions** indicates rare but significant big purchases.
- This suggests the business is **volume-driven**, relying on **frequent small sales**.

4. Sales by Region and Segment

Chart: Boxplot

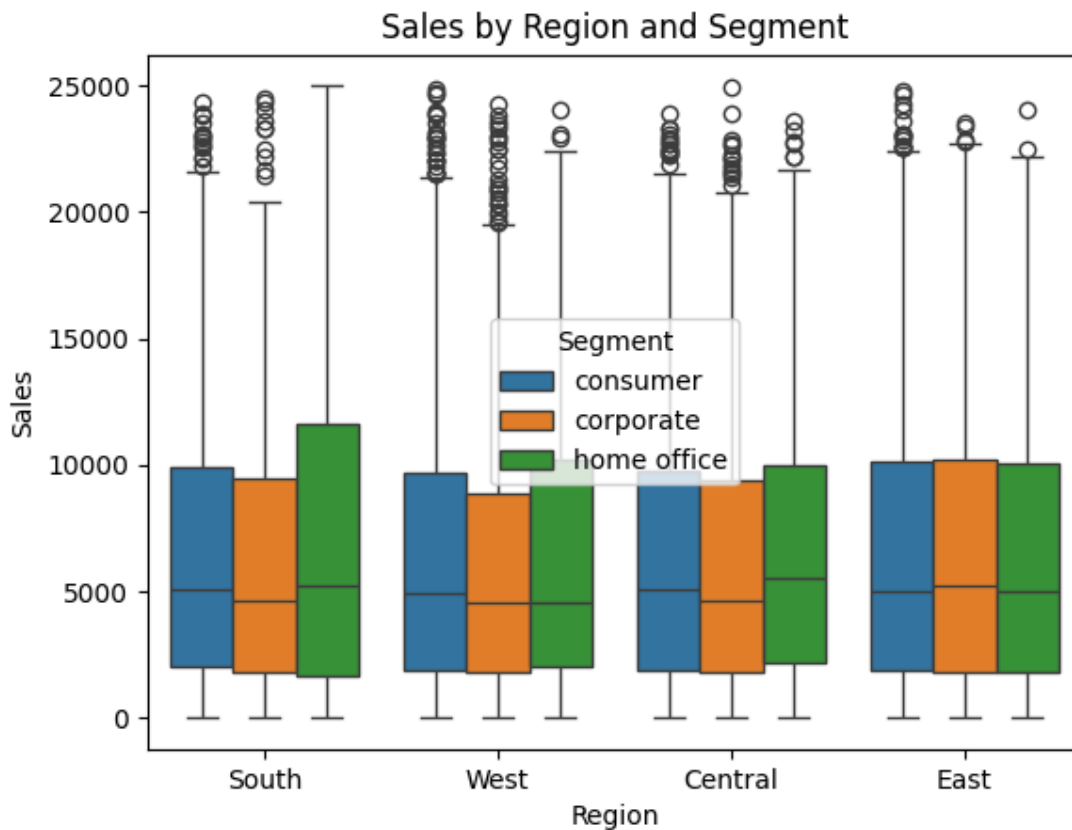


Fig 4.1 Sales by region and segments

Insights:

- Sales are spread similarly across all four regions: **South, West, Central, East**.
- All three segments — **Consumer, Corporate, Home Office** — show similar sales trends.
- **Most sales values are low to mid-range**, with a few high-value outliers.
- **The Home Office in the South** has slightly higher outliers than other segments.

5. Cost vs Profit by Region

Chart: Bubble Plot

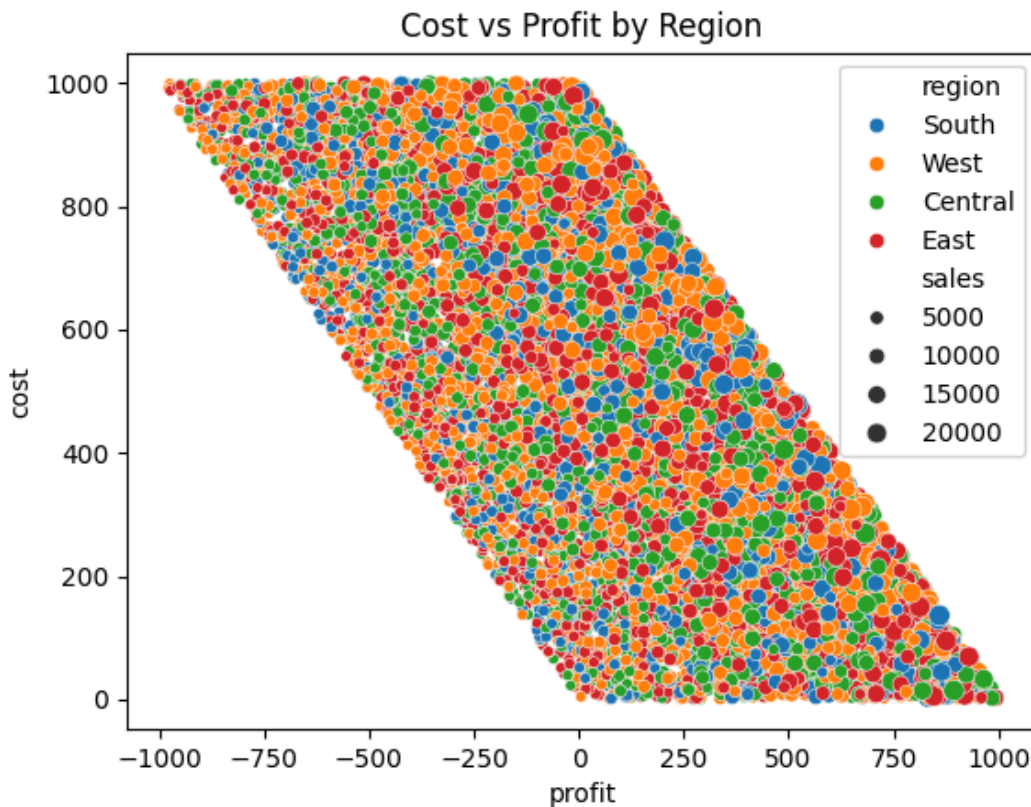


Fig 5.1 Cost vs profit by Region

Insights:

- A **clear inverse relationship** between cost and profit – lower cost, higher profit.
- **Bubble size = Sales** – large sales are mostly in low-cost, high-profit zones.
- All regions behave similarly, suggesting **customer behavior is consistent across locations**.
- Few data points show **high cost and high sales** – potentially premium products.

Summary Insights:

- **Most sales are low in value**, but occur frequently — the business relies on **volume, not price**.
- **Regions perform similarly**, but **South & West** show slightly more high-value transactions.
- **High profit is linked to higher pricing**, not higher cost.
- **Selling more items (quantity)** increases overall sales, regardless of cost.
- **Postal codes show concentration** – focus marketing efforts on these hotspots.
- **Large sales are achievable**, but rare — analyze and replicate high-value customer behavior.

Recommendations:

1. **Focus on cost-efficient products** that bring in consistent profit.
2. **Target regions with dense postal activity** for localized promotions.
3. **Promote bundles or quantity-based discounts** to increase volume sales.
4. **Monitor high-value outliers** – explore patterns for repeat opportunities.
5. **Keep pricing strategy smart** – adjust prices to balance volume and profit.