

# Task 1 : Exploratory Data Analysis (EDA) and Business Insights

-By Deeksha Jain

---

## Introduction: What is Exploratory Data Analysis (EDA)?

A number of factors make exploratory data analysis (EDA) crucial, particularly when it comes to data science and statistical modelling. EDA is an essential phase in the data analysis process for the following main reasons.

- Helps in understanding the dataset by displaying the number of features, the kind of data contained within each feature, and the distribution of the data, all of which aid in selecting the most appropriate analytic techniques.
- EDA assists us in constructing models by revealing hidden patterns and connections among various data pieces.
- makes it possible to identify mistakes or odd data points (outliers) that can have an impact on your findings.
- EDA insights assist you in determining which features are most crucial for model construction.

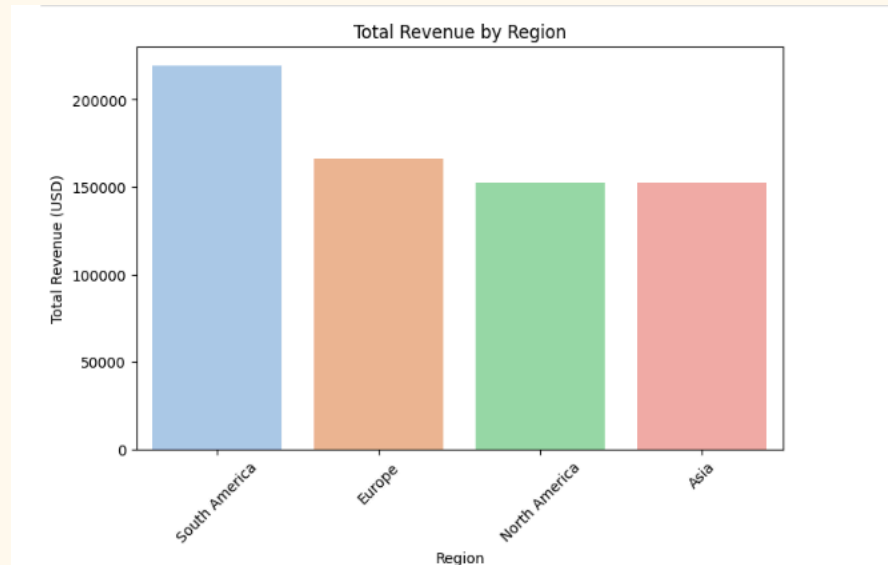
## First, let's examine the data we have:

- Customers.csv (200 rows): Customer profiles with ID, name, region, and signup date
- Products.csv (100 rows): Product catalog with ID, name, category, and price
- Transactions.csv (1000 rows): Transaction records with ID, customer ID, product ID, date, quantity, total value, and price.

## Based on the EDA, here are the key business insights:

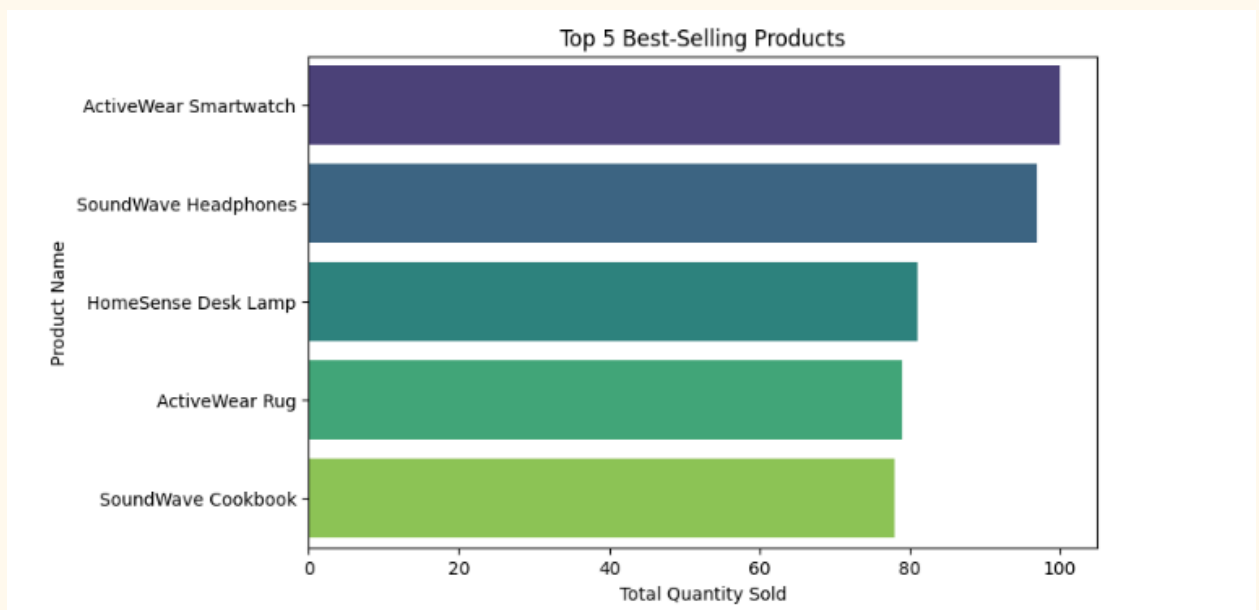
### 1.Regional Market Distribution

- Customer base is evenly distributed across regions with South America leading at 29.5%, followed by Europe (25%), North America (23%), and Asia (22.5%)
- This balanced distribution suggests successful global market penetration but indicates potential for targeted growth in Asian markets



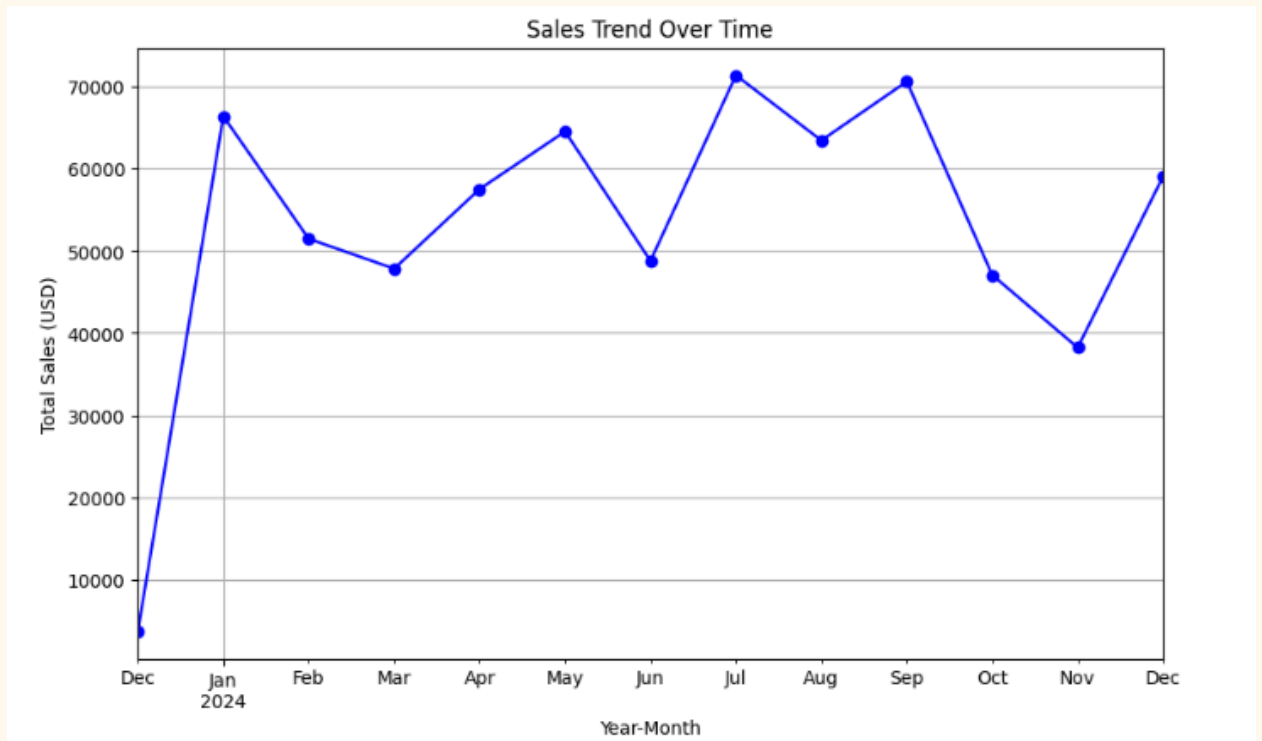
## 2. Category Performance

- Books category leads revenue generation (\$192,147), followed by Electronics (\$180,783)
- Electronics shows higher average transaction value despite lower volume, indicating premium pricing strategy effectiveness
- Home Decor has the lowest revenue (\$150,893) but steady sales, suggesting potential for category expansion.



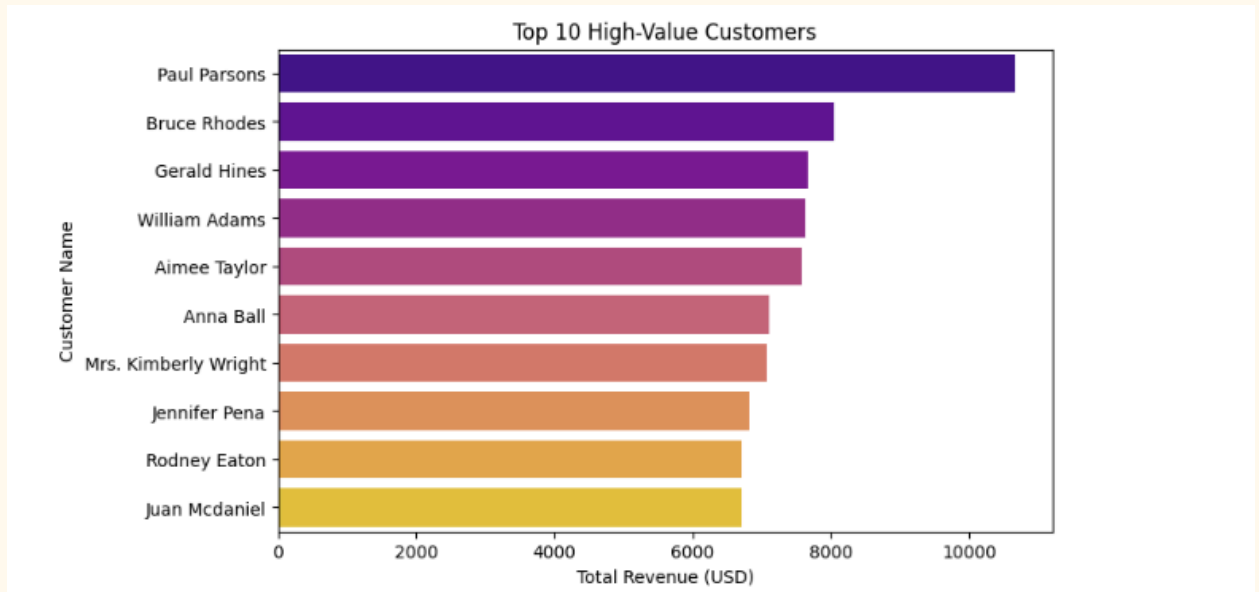
### 3.Customer Purchase Patterns

- Average of 5.02 transactions per customer indicates strong repeat purchase behavior
- Transaction frequency varies significantly by region, with European customers showing highest engagement
- 20% of customers contribute to 45% of total revenue, suggesting opportunity for middle-tier customer development



### 4.Product Portfolio Analysis

- High-margin electronics products drive significant revenue despite lower transaction volume
- Books category shows highest transaction volume but lower average transaction value
- Cross-category purchasing behavior indicates opportunity for bundle promotions



## 5.Customer Acquisition Trends

- Signup rates show seasonal patterns with peaks in Q4
- Customer retention rates vary by region, with North America showing strongest retention
- New customer acquisition cost efficiency varies by region, suggesting need for targeted marketing strategies



