

# Report on Exploratory Data Analysis:

The main conclusions of the exploratory data analysis (EDA) carried out on the customer, product, and transaction datasets are compiled in this paper. Actionable insights were obtained from the investigation, which concentrated on product categories, sales patterns, consumer dispersion, and popular products.

## 1. Customer Distribution by Region

The analysis revealed variations in customer distribution across regions:

- **South America:** Dominated with the highest number of customers, exceeding **55**.
- **Asia:** Reported over **40** customers.
- **North America:** Slightly surpassed **45** customers.
- **Europe:** Fell between North America and South America, with customer numbers just under **50**.

This distribution highlights South America's dominance and offers a basis for region-focused marketing strategies.

## 2. Product Distribution by Category

Product preferences were analyzed based on category, revealing:

- **Books and Electronics:** Both categories were the most popular, each exceeding **25** in distribution.
- **Home Décor:** Ranked third, with values over **20** but lower than the top two categories.
- **Clothing:** Positioned below Books and Electronics but above Home Décor, with values between **20** and **25**.

These findings indicate strong customer interest in Books and Electronics, suggesting potential growth areas for targeted promotions.

## 3. Sales Trends Over Time

A time-series plot of **Total Value vs. Transaction Date** highlighted sales ranging from **0 to 2000** over time. This indicates seasonal or periodic trends in sales activity, which can be leveraged for inventory and campaign planning.

#### 4. Top-Selling Products

The analysis of top-selling products identified high-performing items based on Product IDs:

- **PO59**: Achieved the highest sales, exceeding **17.5**.
- **PO29**: Ranked second, with sales between **15 and 17.5**.
- **PO62, PO79, PO54, and PO61**: Recorded equal sales, all above **15**.
- **PO48, PO22, PO96, and PO49**: Formed the next tier, with sales of exactly **15**.

These insights provide valuable input for product stocking and promotional efforts.

#### 5. Most Popular Products (Merged Dataset)

After merging the datasets, the most popular products were identified as follows:

1. **ActiveWear Smartwatch**: Sold **100** units.
2. **SoundWave Headphones**: Sold **97** units.
3. **HomeSense Desk Lamp**: Sold **81** units.
4. **ActiveWear Rug**: Sold **79** units.
5. **SoundWave Cookbook**: Sold **78** units.

Additional popular items included **ActiveWear Jacket** (76), **BookWorld Biography** (71), and **TechPro T-Shirt** (66). This analysis highlights customer preferences, offering insights into inventory prioritization and promotional campaigns.

#### Conclusion:

Important patterns in consumer behaviour, product preferences, and sales performance were identified by the EDA. Notable sales trends, the great demand for books and electronics, and South America's dominance in client distribution are important lessons. Additionally, determining the best-selling and most well-liked products can help with marketing and inventory management strategy considerations.