

# **Data Science Assignment: ECommerce Transactions Dataset**

Report

On

**Exploratory Data Analysis (EDA)**

by

**Himanshu Raj**

**[himuraj2003@gmail.com](mailto:himuraj2003@gmail.com)**

**BE (4th Year)**

**Computer Science and Engineering**

**Chandigarh University, Mohali**

# Exploratory Data Analysis (EDA)

## Introduction

This report presents exploratory data analysis (EDA) performed on the e-commerce dataset, comprising three files: Customers.csv, Products.csv, and Transactions.csv. The objective is to identify patterns, derive actionable insights, and suggest strategies for business growth.

## Data Overview

- 1. **Customers.csv**  
Contains customer profiles: unique identifiers, names, regions, and signup dates.
- 2. **Products.csv**  
Includes product details: unique identifiers, names, categories, and prices.
- 3. **Transactions.csv**  
Records transactional data: unique transaction IDs, customer-product mappings, transaction dates, quantities, and total values.

The datasets were merged on **CustomerID** and **ProductID** to create a unified dataset for analysis.

## Exploratory Data Analysis

### 1. Monthly Revenue Trends

- **Observation:** Revenue peaks during December and January, likely due to holiday season shopping.
- **Graph:** A line plot depicting monthly revenue trends highlights seasonal spikes.

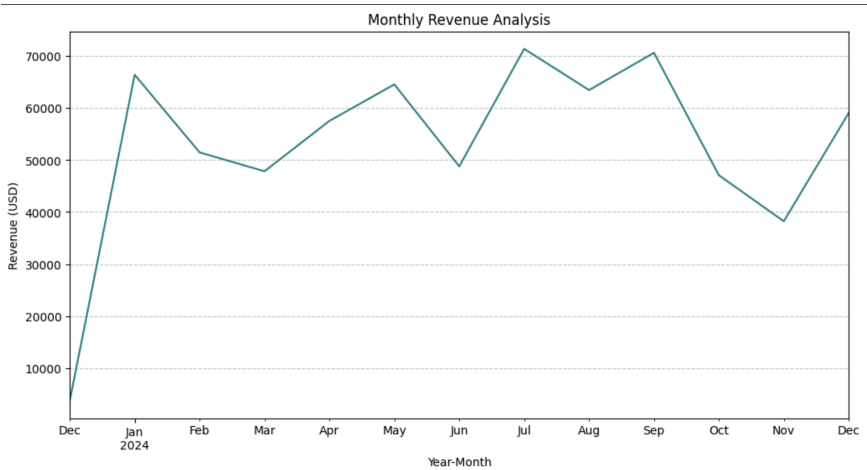


Fig: Monthly Revenue Trends

## 2. Customer Distribution by Region

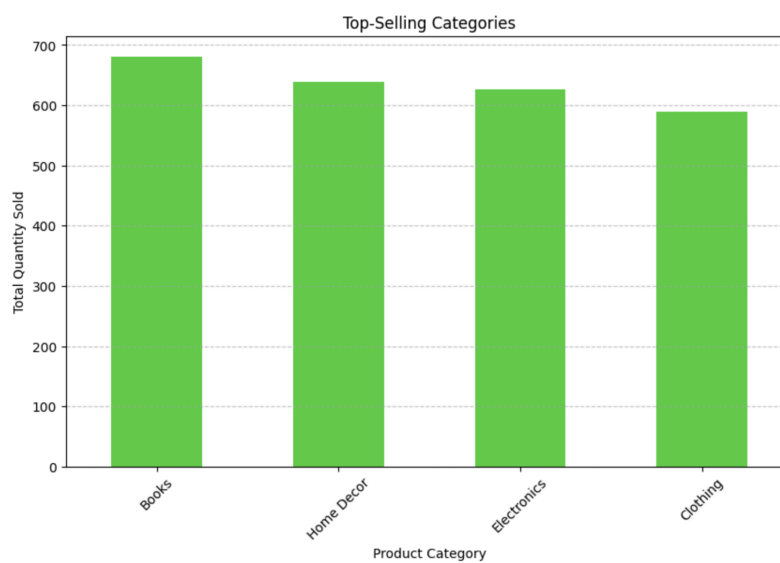
- **Observation:** South America has the largest number of customers, while North America generates higher revenue per customer.
- **Graph:** A bar chart visualizing customer distribution by region.



*Fig: Customer Distribution by Region*

## 3. Top-Selling Product Categories

- **Observation:** Electronics and Books lead in sales volume, while Home Decor underperforms, presenting a growth opportunity with better promotions.
- **Graph:** A bar chart showing the total quantity sold by category.



*Fig: Top-Selling Categories*

4. Price Distribution by Category

- **Observation:** Products priced between \$125 and \$375 dominate sales, reflecting a balance between affordability and profitability.
- **Graph:** A box plot visualizing price distribution across categories segmented by region.

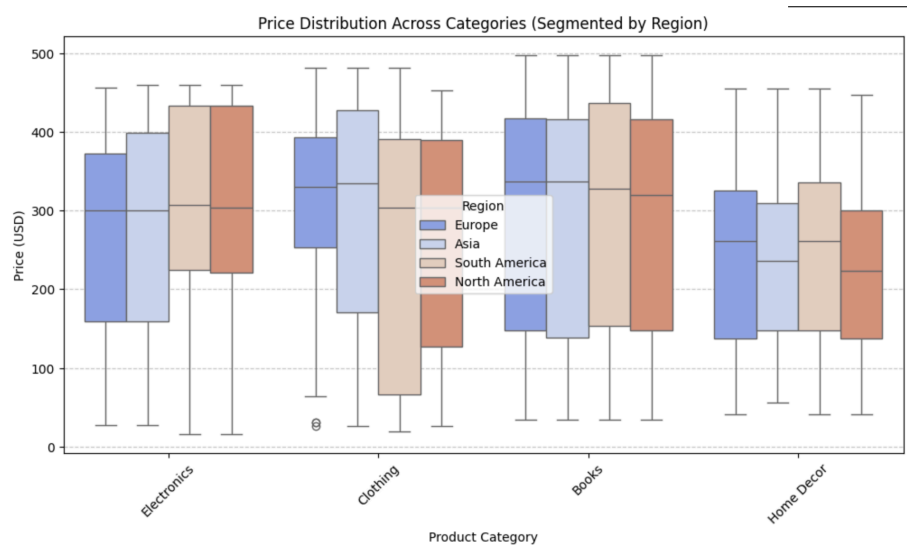


Fig: Price Distribution by Category

5. High-Value Customers

- **Observation:** The top 10 customers contribute significantly to overall revenue, underlining the importance of customer loyalty programs.
- **Graph:** A bar chart highlighting the revenue generated by the top customers.

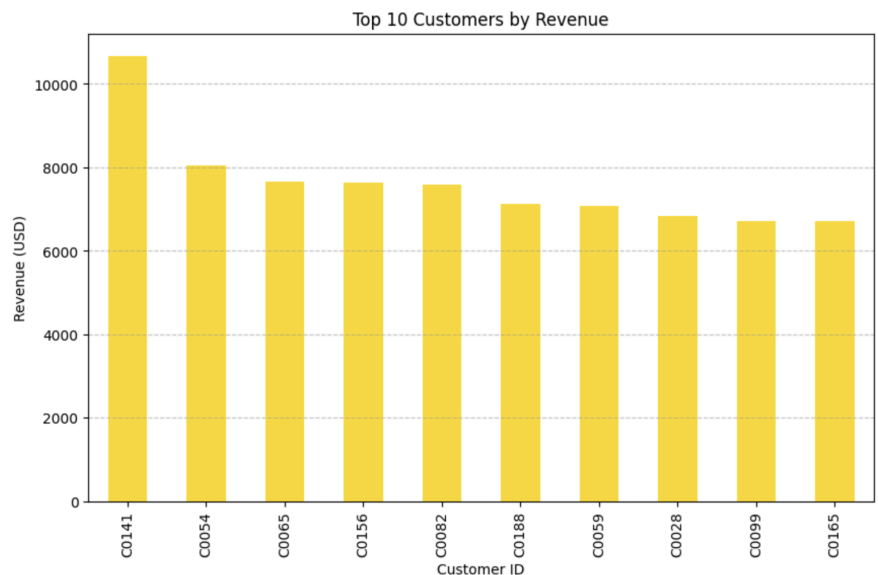


Fig: High-Value Customers Revenue

6. Regional Revenue Contribution

- **Observation:** North America generates the highest revenue despite a smaller customer base, demonstrating high spending power.
- **Graph:** A bar chart showing revenue by region.

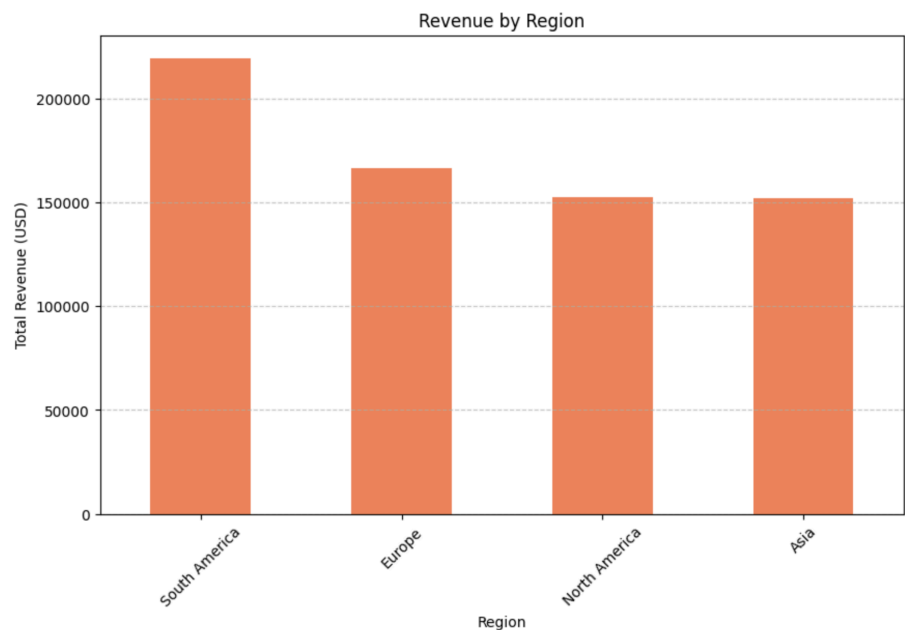


Fig: Regional Revenue

7. Seasonal Trends by Category

- **Observation:** Electronics and Books show consistent seasonal trends, while categories like Fashion exhibit sporadic sales patterns.
- **Graph:** A line plot displaying monthly revenue trends by product category.

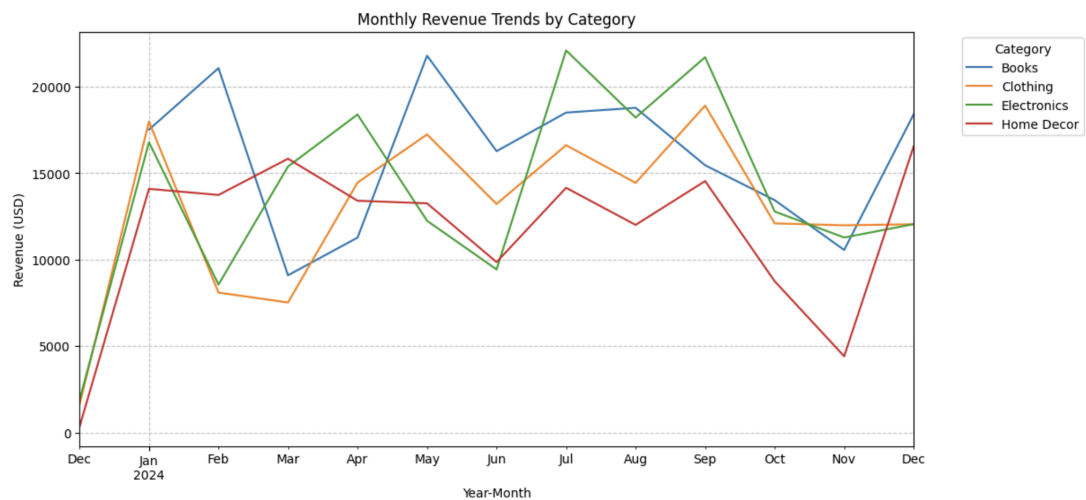


Fig: Seasonal Revenue Trends

## Business Insights

1. **Seasonal Revenue Peaks:**  
Monthly revenue spikes during December and January suggest opportunities for seasonal marketing campaigns and exclusive discounts.
2. **Dominant Product Categories:**  
Electronics and Books are top-performing categories. Niche categories like Home Decor could benefit from targeted promotions to boost sales.
3. **Regional Variations:**  
South America has the largest customer base, but North America drives higher revenue per customer. Customized strategies by region could enhance profitability.
4. **Price Sensitivity:**  
Products in the \$125–\$375 range perform best. Introducing products in this price range for underperforming categories might improve sales.
5. **Customer Loyalty:**  
The top 10 customers contribute disproportionately to revenue. Implementing loyalty programs and personalized offers could retain these high-value customers.

## Conclusion

The analysis highlights seasonal trends, category performance, and regional variations. Strategies such as seasonal campaigns, regional pricing, and customer loyalty programs are recommended to drive growth and revenue.