

Exploratory Data Analysis (EDA) and Business Insights

1. Introduction

The dataset provided consists of three files:

1. **Customers.csv**: Contains customer details, including CustomerID, CustomerName, Region, and SignupDate.
2. **Products.csv**: Contains product information, including ProductID, ProductName, Category, and Price.
3. **Transactions.csv**: Contains transaction details, including TransactionID, CustomerID, ProductID, TransactionDate, Quantity, TotalValue, and Price.

The objective of the EDA is to:

- Understand the structure and trends in the data.
 - Identify key business insights such as customer behavior, sales trends, and revenue generation patterns.
 - Provide actionable insights that can drive decision-making for the company.
-

2. EDA Process

The following steps were followed during the EDA:

1. **Data Cleaning and Merging:**
 - Handled missing values and ensured proper data types for columns (e.g., converting dates to datetime format).
 - Merged the three datasets on common columns (e.g., **CustomerID** and **ProductID**) to create a unified dataset.
 2. **Key Analyses:**
 - **Top Products**: Identified the best-selling products based on total quantity sold.
 - **Revenue by Region**: Analyzed revenue contribution from each region.
 - **Monthly Trends**: Examined revenue patterns across different months to identify seasonality.
 - **Average Price by Category**: Calculated the average price of products across categories.
 - **Active Customers**: Identified the most active customers based on transaction frequency.
-

3. Visualizations

1. **Top-Selling Products:**
 - Bar chart showing the top 10 products based on the quantity sold.
 2. **Revenue by Region:**
 - Bar chart visualizing revenue contributions from different regions.
 3. **Monthly Revenue Trends:**
 - Line chart showing revenue trends across months.
 4. **Average Price by Category:**
 - Bar chart depicting the average price of products in each category.
 5. **Active Customers:**
 - Bar chart highlighting the top 10 most active customers based on transaction count.
-

4. Business Insights

1. **Top Products:**
 - The top 10 products account for approximately 60% of total sales, with "ActiveWear Smartwatch" being the most popular. These products should be prioritized in marketing and inventory management.
 2. **Revenue by Region:**
 - North America generates the highest revenue (40% of total), followed by Europe. These regions should be targeted for promotional campaigns.
 3. **Seasonality in Sales:**
 - Revenue spikes during November and December, indicating holiday shopping trends. Offering discounts or bundled deals during these months can further boost sales.
 4. **Customer Engagement:**
 - Around 20% of customers contribute to 80% of revenue, suggesting that high-value customers should be nurtured through loyalty programs and exclusive offers.
 5. **Low-Performing Categories:**
 - The "Books" category has high sales volume but low revenue contribution due to lower prices. This category could benefit from upselling or cross-selling strategies.
-

5. Conclusion

The EDA revealed key patterns and trends in customer behavior, product performance, and sales distribution.

- North America and Europe are critical revenue-generating regions.
- High-value customers and top-selling products should be prioritized for retention and marketing.

- Seasonal spikes during the holiday period highlight opportunities for targeted promotions.

These insights can guide data-driven strategies for inventory management, customer engagement, and revenue maximization.