

EDA Report for Task 1: Business Insights

Exploratory Data Analysis (EDA) Summary

1. Data Overview

- **Datasets Used:**
 - Customers.csv: Contains 200 customer records, including details like CustomerID, CustomerName, Region, and SignupDate.
 - Products.csv: Contains 100 product records with ProductID, ProductName, Category, and Price.
 - Transactions.csv: Contains 1000 transaction records with fields such as TransactionID, CustomerID, ProductID, TransactionDate, Quantity, and TotalValue.

2. Key Insights

- **Customer Region Distribution:**
 - The majority of customers are from South America (59 customers), followed by Europe (50), North America (46), and Asia (45).
 - **Visualization:** A bar chart was created to display the distribution of customers by region.
- **Monthly Sales Trends:**
 - Sales peaked in September 2024, reaching \$70,603.75.
 - Lowest sales occurred in December 2023, with \$3,769.52.
 - **Visualization:** A line chart was created to track monthly sales trends.
- **Top 10 Products by Sales:**
 - The most profitable product was **ActiveWear Smartwatch**, generating \$39,096.97 in total sales.
 - Other top products included **SoundWave Headphones** and **SoundWave Novel**.
 - **Visualization:** A horizontal bar chart was created to display sales of the top 10 products.
- **Average Order Value (AOV):**
 - The AOV across all transactions was **\$690.00**.

- **Category-Wise Sales:**

- The top-performing category was **Books**, contributing \$192,147.47 in total sales.
- Other major contributors included **Electronics** (\$180,783.50), **Clothing** (\$166,170.66), and **Home Decor** (\$150,893.93).
- **Visualization:** A bar chart was created to illustrate category-wise sales distribution.

3. Code and Outputs

- All steps, including data loading, merging, and analysis, were performed using Python libraries such as pandas, matplotlib, and seaborn.
- The outputs were visualized with appropriate charts to help understand trends and patterns.

OUTPUT





