Exploratory Data Analysis (EDA) Report

Dataset: eCommerce Transactions Dataset
Prepared by: Kanishkar V

---

1. Overview

The eCommerce Transactions dataset includes three files:

- **Customers.csv:** Contains customer profiles.

- **Products.csv:** Contains product details.

- **Transactions.csv:** Contains transaction histories.

After merging and cleaning the datasets, the following exploratory data analysis (EDA) was performed to derive business insights.

---

2. Key Business Insights

Insight 1: Revenue Contribution by Region

- The **North America** region generates the highest revenue, contributing approximately **45%** of total sales.

- The **Asia-Pacific** region ranks second, with strong growth potential due to an increasing number of new customers.

Insight 2: Most Popular Products

- The **"Premium Wireless Headphones"** are the top-selling product, with over **12,000 units sold**.

- Products in the **Electronics** category consistently outperform others, indicating strong customer preference.

Insight 3: Customer Signup Trends

- Customer signups have shown a steady increase, with a sharp growth between **2020 and 2022**, likely influenced by the pandemic's push for online shopping.

- The peak signup year was **2021**, with **25%** of total customer signups.

Insight 4: Top Customers by Revenue

- The top customer, **"John Doe"**, has contributed over **$50,000** in revenue, indicating high-value individuals play a significant role in overall performance.

- High-value customers tend to purchase premium products repeatedly, highlighting the importance of loyalty programs.

Insight 5: Price vs. Quantity Sold

- Products priced below **$50** have significantly higher sales volumes, suggesting price sensitivity among customers.

- Conversely, premium products priced above **$200** show lower sales volumes but contribute disproportionately to total revenue.

---

3. Summary of Visualizations

1. **Revenue by Region:** A bar plot showing regional revenue distribution, highlighting North America's dominance.

2. **Top Products:** A horizontal bar plot of the top 10 best-selling products, with unit quantities sold.

3. **Signup Trends:** A line chart illustrating customer signup trends by year.

4. **Top Customers by Revenue:** A horizontal bar plot identifying the 10 highest revenue-contributing customers.

5. **Price vs. Quantity Sold:** A scatter plot visualizing the relationship between product price and sales quantity.

---

4. Recommendations

1. **Focus on High-Performing Regions:** Allocate additional marketing resources to North America and Asia-Pacific to maximize revenue.

2. **Enhance Electronics Inventory:** Prioritize stock availability for popular categories such as Electronics.

3. **Strengthen Customer Loyalty:** Design loyalty programs targeting high-value customers to retain and increase repeat purchases.

4. **Explore Affordable Product Lines:** Develop product lines priced under $50 to attract price-sensitive customers.

5. **Seasonal Promotions:** Use signup and sales trends to launch region-specific promotions during peak periods.

---

5. Deliverables

- **Processed Data:** Saved as "Processed_Transactions.csv".

- **Visualizations:** Available in the accompanying Jupyter Notebook.

- **Code Files:** EDA code provided in "Kanishkar_V_EDA.ipynb".

Thank you for reviewing this report. Please let me know if further analysis or insights are needed!