Task 1: Exploratory Data Analysis (EDA) and Business Insights

Objective

Exploratory Data Analysis (EDA) uncovers patterns, relationships, anomalies, and insights in data to guide informed business decisions.

Dataset Overview

1. Customers Dataset
- **Total Entries:** 200
- **Key Features:**
  - CustomerID: Unique identifier for each customer.
  - CustomerName: Name of the customer.
  - Region: Geographic location of the customer.
  - SignupDate: Date when the customer signed up.

2. Products Dataset
- **Total Entries:** 100
- **Key Features:**
  - ProductID: Unique identifier for each product.
  - ProductName: Name of the product.
  - Category: Product category (Books, Electronics, Clothing, etc.).
  - Price: Price of the product.

3. Transactions Dataset
- **Total Entries:** 1,000
- **Key Features:**
  - TransactionID: Unique identifier for each transaction.
  - CustomerID: Customer associated with the transaction.
  - ProductID: Product being purchased.
  - TransactionDate: Date and time of the transaction.
  - Quantity: Number of products purchased in a transaction.
  - TotalValue: Total cost of the transaction.

EDA Results & Visualizations

1. Data Cleaning & Preprocessing
- No missing values were found in any dataset.
- Data types were checked and converted where necessary (e.g., TransactionDate converted to DateTime).
- Merged datasets for a comprehensive analysis.

2. Key Findings

a) Customer Distribution by Region
- South America has the highest number of customers (59), followed by other regions.

- This distribution helps target marketing strategies based on location.

b) Product Price Distribution

- Prices range from **$16.08 to $497.76**, with an average price of **$267.55**.
- The majority of products fall within the mid-price range.

c) Transaction Value Distribution

- Most transactions have values between **$295 and $1011**.
- The highest recorded transaction was **$1991.04**.

d) Sales Trends Over Time

- Sales fluctuate across time, with notable spikes in certain periods.
- Seasonal trends could indicate high-demand periods.

e) Top-Selling Products

- The most sold product is "SoundWave Jeans": **46 units (1.81%)**
  "SoundWave Cookbook": **46 units (1.81%)**
- The top 10 best-selling products contribute to nearly **16.71%** of the total sales.
- Top 2 Most Popular Categories:
  - Books: **681 units**
  - Home Decor: **639 units**

Business Insights

1. **Regional Targeting:** Customers from **South America contribute the most** to sales. Focusing marketing campaigns here could further boost revenue.
2. **High-Value Transactions:** The average transaction value is **$690**. Offering discounts for bulk purchases may increase sales volume.
3. **Best-Selling Products: Electronics dominate sales**. Expanding inventory in this category could maximize profits.
4. **Seasonal Demand:** There are **sales peaks at specific times**. Planning marketing promotions around these periods can optimize revenue.
5. **Price Sensitivity:** Since most products fall in the **$150-$400 range**, pricing strategies should target this mid-range for maximum customer engagement.

Conclusion

The EDA findings highlight key business trends, such as customer demographics, product demand, and sales patterns. These insights can help the company refine its marketing strategies, improve inventory management, and boost overall sales performance.

```
Merged Data Preview:
  TransactionID CustomerID ProductID      TransactionDate  Quantity  \
0       T00001    C0199      P067  2024-08-25 12:38:23         1
1       T00112    C0146      P067  2024-05-27 22:23:54         1
2       T00166    C0127      P067  2024-04-25 07:38:55         1
3       T00272    C0087      P067  2024-03-26 22:55:37         2
4       T00363    C0070      P067  2024-03-21 15:10:10         3

   TotalValue  Price_x        Date      CustomerName         Region  \
0      300.68   300.68  2024-08-25   Andrea Jenkins         Europe
1      300.68   300.68  2024-05-27   Brittany Harvey           Asia
2      300.68   300.68  2024-04-25   Kathryn Stevens         Europe
3      601.36   300.68  2024-03-26   Travis Campbell  South America
4      902.04   300.68  2024-03-21    Timothy Perez         Europe

   SignupDate                      ProductName     Category  Price_y
0  2022-12-03  ComfortLiving Bluetooth Speaker  Electronics   300.68
1  2024-09-04  ComfortLiving Bluetooth Speaker  Electronics   300.68
2  2024-04-04  ComfortLiving Bluetooth Speaker  Electronics   300.68
3  2024-04-11  ComfortLiving Bluetooth Speaker  Electronics   300.68
4  2022-03-15  ComfortLiving Bluetooth Speaker  Electronics   300.68

Key Business Metrics:
Total Revenue: $689,995.56
Average Transaction Value: $690.00
Total Number of Transactions: 1,000
Total Number of Customers: 199
Total Number of Products: 100
```


Top 10 Best-Selling Products by Quantity Sold