Data Science Assignment: eCommerce Transactions Dataset

Assignment Tasks:

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

1. Perform EDA on the provided dataset.

EDA codings:

```
import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt


# Load data

customers_path = '/mnt/data/Customers.csv'

products_path = '/mnt/data/Products.csv'

transactions_path = '/mnt/data/Transactions.csv'


customers = pd.read_csv(customers_path)

products = pd.read_csv(products_path)

transactions = pd.read_csv(transactions_path)


# Merge datasets

data = transactions.merge(customers, on='CustomerID').merge(products,
    on='ProductID')


# Perform analysis and visualization

# Example: Top-selling products
```

```
top_products =
    data.groupby('ProductName')['TotalValue'].sum().sort_values(ascending=Fals
    e).head(10)

top_products.plot(kind='bar', color='skyblue')

plt.title('Top-Selling Products')

plt.xlabel('Product Name')

plt.ylabel('Total Revenue')

plt.xticks(rotation=45, ha='right')

plt.tight_layout()

plt.show()

# Low-selling products by revenue

low_products =
    data.groupby('ProductName')['TotalValue'].sum().sort_values(ascending=True
    ).head(10)

plt.figure(figsize=(10, 6))

low_products.plot(kind='bar', color='lightcoral')

plt.title('Bottom 10 Selling Products by Revenue')

plt.ylabel('Total Revenue')

plt.xlabel('Product Name')

plt.xticks(rotation=45, ha='right')

plt.tight_layout()

plt.show()
```

2. Derive at least **5 business insights** from the EDA.
   ○ Write these insights in **short point-wise sentences** (maximum 100 words per insight).

1. Highest Revenue-Generating Products

The top 10 products generate a large share of the total revenue. These products are mainly in one or two categories, which shows customer preference and demand concentration.

2. Regional Variations in Revenue

Some regions, such as North America and Asia, are major revenue generators, indicating greater market penetration or spending power than regions like South America.

3. Monthly Revenue Trend

Revenue patterns show seasonal peaks in sales such as the holiday season, which can inform promotional activities and inventory replenishments.

4. Poor Performers

The bottom 10 represent an insignificant amount of revenue generated. These products show little interest from customers or weak marketing efforts. One may need to readjust pricing strategy and promotion or discontinue them if trends keep showing low performance.

5. Category Performance Analysis

Categories like electronics and books are dominating in terms of sales volume and revenue, indicating their popularity. Categories like home decor and clothing have growth potential if they are marketed strategically or bundled with other products.