EDA and Insights from Transaction Data

This script performs comprehensive Exploratory Data Analysis (EDA) on customer, product, and transaction datasets. Key insights are derived using visualizations and metrics calculated through Python's pandas and matplotlib/seaborn libraries.

Key Features of the Code

1. Top-Selling Products Analysis:

- A bar chart showing the top 10 products based on total quantity sold, with an overlay for total revenue.
- Highlights the products that contribute the most to sales volume and revenue.
- Metrics derived:
 - ActiveWear Smartwatch sold the highest with 100 units, generating more than \$40,000 in revenue.

2. Regional Revenue and Transactions:

- A combined bar chart displaying revenue by region with transaction counts on a secondary axis.
- This highlights regions with high revenue vs. those with frequent transactions but lower revenue per transaction.
- Metrics derived:
 - South America generated the highest revenue, and had the highest transaction count.

3. Revenue by Region and Category:

• A grouped bar chart visualizing revenue by region and product category.

4. Monthly Trends in Transactions:

- Line charts illustrating trends in the number of transactions, revenue, and Average Order Value (AOV).
- Helps identify seasonal trends or irregularities in customer behavior.

5. Correlation Between Revenue and Customer Signups:

- Overlaid line plots of monthly revenue and customer signup trends.
- Calculated correlation coefficient between these variables.
- Metrics derived:

• Correlation coefficient = 0.14, indicating a weak relationship.

6. Customer-Level Insights:

- Bar charts showing the top 20 customers by revenue and transaction count.
- Highlights high-value customers and frequent buyers.

Usage

- Modules Used: pandas, matplotlib, seaborn
- Data Requirements: Customers.csv, Products.csv, Transactions.csv
- Instructions:
 - 1. Ensure the datasets are in the same directory as the script.
 - 2. Run the script to generate insights and visualizations.