Task: Analyzing Customer Transactions and Sales Performance

Objective:

Analyze a customer transaction dataset to extract insights and visualize key trends using various chart types.

Dataset:

Use a dataset containing the following columns:

- Transaction_ID
- Date
- Customer_ID
- Product_Category
- Amount_Spent
- Units_Purchased
- Payment_Type
- Region

1. Data Preprocessing

- Load the dataset using Pandas.
- Convert the 'Date' column to datetime format.
- Check for and handle missing values.
- Remove duplicate transactions.
- Extract year, month, and weekday from the 'Date' column.

2. Data Analysis Using NumPy & Pandas

- Calculate total revenue and average transaction value.
- Identify the top 3 most popular product categories.
- Determine the most frequently used payment type.

- Compute monthly sales growth rate.

- Identify the top 5 highest spending customers.

- Determine the busiest sales day of the week.

3. Advanced Data Visualization

- Line Chart: Display monthly revenue trends for each product category.

- Pie Chart: Show the distribution of sales by payment type.

- Histogram: Illustrate the distribution of transaction amounts.

- Area Chart: Compare sales performance across different regions over time.

Bonus Challenges

- Calculate customer retention rate based on repeat purchases.

- Compare average spending per region.

- Analyze the correlation between the number of units purchased and the total amount spent.

- Create a heatmap using Seaborn to show sales trends by weekday and hour.

Customer Transactions Dataset Links

Available Datasets:

Customer Transactions Dataset:

https://www.kaggle.com/datasets/bkcoban/customer-transactions

Sample Sales Data:

https://www.kaggle.com/datasets/kyanyoga/sample-sales-data

Online Retail Dataset:

https://archive.ics.uci.edu/dataset/352/online%2Bretail