Project Overview

This project analyzes customer transaction data for **Exploratory Data Analysis (EDA)**, a **Lookalike Model**, and **Customer Segmentation** using clustering techniques. It uses customer, product, and transaction data to uncover insights into customer behavior.

1. Dependencies

Required packages (pandas, numpy, matplotlib, seaborn, scikit-learn, requests) are installed if missing to ensure smooth execution.

2. Data Download and Loading

The project uses three datasets:

- **Customers.csv**: Customer details.
- **Products.csv**: Product info.
- **Transactions.csv**: Transaction records.

If not already present, these are downloaded and loaded into pandas Data Frames.

3. Exploratory Data Analysis (EDA)

- **Data Overview**: Display dataset samples.
- **Missing Values**: Check for missing data.
- **Customer Distribution**: Plot customer distribution by region.
- **Top Products**: Visualize the top 10 products by revenue.

4. Lookalike Model

- **Objective**: Find customers with similar purchasing patterns using the **Nearest Neighbours** algorithm.
- **Process**: Merge data, create a pivot table, and identify lookalike customers.
- **Output**: Save results in a CSV file showing similar customers.

5. Customer Segmentation (Clustering)

- **Objective**: Group customers by spending behaviour using **K Means**.
- **Process**: Aggregate spending, scale data, and find the optimal number of clusters using the **Davies-Bouldin index**.
- **Output**: Visualize customer clusters and segment customers based on their spending patterns.

6. Results

- **Lookalike Model**: Identify potential customers for targeted marketing.
- **Customer Segmentation**: Visualize segments to optimize marketing strategies.

Conclusion

This project provides insights into customer behaviour, helping businesses:

1. **Understand customer demographics** through EDA.
2. **Target lookalike customers** for marketing campaigns.
3. **Segment customers** to tailor marketing efforts.