PRP AIML ASSIGNMNET-1

Mini-Project: E-Commerce Sales Data Analysis & Visualization

Objective: Analyze an e-commerce sales dataset using Pandas and NumPy, and create meaningful visualizations using Matplotlib, Seaborn, or ggplot.

Step 1: Data Preparation (Pandas) - Load CSV dataset with columns like: Order_ID, Customer_ID, Product, Category, Quantity, Price, Order_Date. - Explore dataset using: .head(), .tail(), .info(), .describe(), .value_counts(), .unique(), .nunique(). - Handle missing values and duplicates: .fillna(), .dropna(), .drop_duplicates().

Step 2: Data Analysis (NumPy) - Convert relevant columns to NumPy arrays (1-D for individual columns, 2-D for numeric features). - Calculate: - Total sales per product and per category. - Average order value per customer. - Monthly sales totals. - Percentage contribution of each category to total revenue.

Step 3: Data Visualization - Create meaningful visualizations: 1. Line Graph: Monthly sales trend over the year. 2. Bar Graph: Total sales per product category. 3. Pie Chart: Contribution of top 5 products to total sales. 4. Histogram: Distribution of order values. 5. Scatter Plot: Quantity vs Revenue per order.

Step 4: Insights Reporting - Summarize insights using Markdown cells or comments in Jupyter Notebook: - Best-selling products and categories. - Peak sales months. - Customer buying patterns. - Save graphs as images for reporting.

Submission Format: - A single Word or PDF document containing: - Screenshots - Short explanations – Diagrams. With your name and enrollment number as the file name and submit the hard copy.