Sales Data Analysis Project Documentation

# 1. Objective

The objective of this project is to perform an exploratory data analysis (EDA) on sales data. The analysis aims to uncover trends, customer preferences, and insights into different product categories, payment methods, and gender-based purchasing patterns.

# 2. Data Source

The dataset used in this project is a sales dataset containing 1000 rows and 17 columns. It includes details of sales transactions such as Invoice ID, Customer type, Gender, Product line, Unit price, Quantity, Sales totals, Date, Time, Payment methods, and customer ratings.

# 3. Libraries Used

The following Python libraries were used for data analysis and visualization:   
• Pandas: For data manipulation and analysis  
• NumPy: For numerical operations  
• Matplotlib: For generating plots and visualizations  
• Seaborn: For creating attractive statistical graphics

# 4. Steps Involved in the Analysis

## 4.1 Data Loading

The dataset was loaded into a Pandas DataFrame using the read\_csv() function, which allows for easy manipulation of data.

## 4.2 Data Cleaning

The 'Date' column was converted from string format to a datetime object to facilitate time-based analysis. No missing values were found in the dataset.

## 4.3 Exploratory Data Analysis (EDA)

### 4.3.1 Monthly Sales Trends

The dataset was grouped by month to examine total sales over time. A line plot was created to visualize monthly sales trends, showing fluctuations in sales across different months.

### 4.3.2 Sales by Product Line

Sales totals were grouped by product line to identify which categories generated the most revenue. A horizontal bar chart was used to visualize sales distribution by product category.

### 4.3.3 Gender-Based Purchasing Patterns

Sales data was grouped by gender to understand purchasing behavior differences between male and female customers. A pie chart was used to show the percentage of total sales contributed by each gender.

### 4.3.4 Preferred Payment Methods

The frequency of each payment method (Cash, Ewallet, Credit Card) was analyzed. A count plot was generated to visualize the most commonly used payment methods.

# 5. Conclusion

This project provided valuable insights into sales trends, customer demographics, and product performance. The analysis revealed key information that can be used to optimize sales strategies and better understand customer preferences.

Shivanshu Tiwari