

# Customer Sales Data Analysis

---

## 1. Introduction

This report presents a comprehensive analysis of customer sales data, focusing on cleaning and exploring the dataset to extract meaningful insights. The dataset contains both numerical and categorical variables and includes missing values, duplicate entries, and outliers that needed to be addressed.

## 2. Dataset Overview

The dataset consists of 201 records and 7 features:

- Customer\_ID: Unique identifier for each customer
- Age: Numerical, customer's age
- Gender: Categorical, customer's gender (Male/Female/Other)
- Product\_Category: Categorical, type of product purchased
- Purchase\_Amount: Numerical, total purchase amount in dollars
- Discount\_Applied: Numerical, discount percentage applied
- Payment\_Method: Categorical, payment method used

## 3. Data Cleaning

- Handling Missing Values: Replaced missing values with median for numerical features and mode for categorical ones.
- Removing Duplicates: Identified and removed duplicate records.
- Treating Outliers: Used the Interquartile Range (IQR) method to cap extreme purchase amounts.
- Standardizing Categorical Values: Ensured consistent capitalization and formatting.

## 4. Exploratory Data Analysis (EDA)

### Univariate Analysis

- Age: Concentrated around 30–55 years
- Purchase Amount: Right-skewed, with most purchases around \$500
- Gender: Balanced distribution across Male and Female
- Product Category: Even spread across categories like Electronics, Clothing, and Furniture
- Payment Method: Credit Card and Debit Card most popular

### **Bivariate Analysis**

- Weak correlations among numerical features
- Higher purchase amounts observed for Electronics and Furniture categories
- Payment method impacts discount variability

### **Multivariate Analysis**

- No strong multivariate patterns among numerical variables
- Electronics and Furniture with Credit Card and PayPal show higher average purchase amounts

## **5. Key Insights and Observations**

- Majority of customers are aged 30–55 and prefer Credit or Debit Card payments.
- Electronics and Furniture drive higher purchase amounts.
- Discounts applied vary significantly based on the payment method used.

## **6. Conclusion**

This analysis provided a clear understanding of customer behavior and spending patterns. The cleaned and transformed data is now ready for further modeling or decision-making processes.