Supermarket Sales Data Analysis

Project Report

Date: 2/14/2025
Author: Malak Ahmed Alaa

1-Introduction

This project involves the **analysis of supermarket sales data** to derive insights and improve business decisions. The dataset contains **transactional records** with product prices, payment methods, customer ratings, and total sales.

2-Data Cleaning Process

The following steps were taken to ensure **data quality**:

✓ **Data Inspection**

- Checked for **missing values, duplicates, and incorrect data types**.

- Identified **outliers and inconsistent values**.

✓ **Data Cleaning**

- **Handled missing values**:

    o Quantity → Replaced negative values with the **mean quantity**.

    o Rating → Replaced values **out of range (1-10)** with the **median rating**.

- **Converted incorrect data types** (Unit Price, Quantity, Total → Numeric).

- **Calculated new columns**:

    o Subtotal = Quantity × Unit Price

    o Total = Subtotal + Tax 5%

✓ **Outlier Treatment**

- **Applied IQR method** to remove extreme values in Total.

3-Data Analysis & Insights

The following business insights were extracted:

🛒 **Sales Insights**

- The **most frequently used payment method** is ... based on sales distribution.

- The **highest total sales** come from the ... category.

📊 **Correlation Analysis**

- **Total Sales** strongly correlate with **Quantity Sold**.

- **Boxplots & histograms** reveal trends in product pricing.

4-Technologies Used

Technology	Purpose
Python	Data Analysis
Pandas	Data Cleaning
Seaborn	Data Visualization
Matplotlib	Graphs & Plots

5-Conclusion

This project provided key insights into supermarket sales, enabling better business decision-making.