Walmart Sales Data Analysis\*

Project Overview

This project focuses on \*analyzing Walmart sales data\* to uncover key trends, patterns, and store performance insights.

It involves \*data cleaning, exploratory data analysis (EDA), and visualizations\* to better understand pricing, promotions, and sales volumes.

Dataset Details

\*File Name:\* Walmart.csv

\*Number of Features:\* Varies based on dataset version

\*Main Columns:\*

| Column Name | Description |

| ---------------- | ---------------------------------- |

| store\_id | Unique store identifier |

| product\_name | Name of the product |

| unit\_price | Price per unit (numeric) |

| quantity\_sold | Quantity sold (numeric) |

| promotion\_type | Type of promotion applied (if any) |

Tools & Libraries Used

\* \*Python 3.x\*

\* \*pandas\* → For data manipulation and cleaning

\* \*numpy\* → For numerical calculations

\* \*matplotlib\* → For static visualizations

\* \*seaborn\* → For advanced plots and styling

Analysis Workflow

# Data Loading

\* Load Walmart.csv into a panda\_Data Frame.

\* Display the \*first\* and \*last\* 10 rows.

\* Inspect \*data types, \*\*shape, and \*\*column names\*.

# Data Cleaning

\* Identify and handle missing values in promotion\_type (replace with "None").

\* Ensure numeric columns (unit\_price, quantity\_sold) have correct data types.

#Exploratory Data Analysis (EDA)

\* \*Summary statistics\*: Mean, median, standard deviation for numeric columns.

\* \*Total quantity sold\* across all stores.

\* \*Average price per product\* using groupby.

# Visualization

\* \*Bar Chart\*: Average price per product (with labels).

\* \*Store Performance\*:

\* Calculate total sales by store\_id.

\* Determine median store sales.

\* Identify \*underperforming stores\* below median.

Key Insights

\* Products have varying price ranges and average unit prices.

\* Some stores consistently sell \*less than median\* sales volume.

\* Promotions may not be applied consistently across all products.

How to Run the Project

1. \*Clone\* the repository or download the .ipynb notebook.

2. \*Install dependencies\*:

bash

pip install pandas numpy matplotlib seaborn

3. \*Place\* Walmart.csv in the project folder.

4. Open the notebook in \*Jupyter Notebook\* or \*Google Colab\*.

5. Run each cell sequentially.

Future Enhancements

\* Add time-series analysis for sales trends over months/years.

\* Include profit margin calculations.

\* Explore correlation between promotions and sales lift.

\* Implement predictive modeling for future sales forecasting.