**Walmart Sales Data Analysis Report**

**Author:** Jhil Chatterjee  
**Project:** Sales Data Analysis – Commercial Store (Walmart)

**Introduction**

This project analyses Walmart's sales data to uncover insights into store performance, departmental trends, and the impact of holidays and economic factors on sales. Through data cleaning, merging, and visualization, this analysis supports data-driven business decisions.

**Objective**

* Analyse weekly sales trends over time
* Compare total sales across stores and departments
* Examine the effect of holidays on sales
* Investigate correlations with economic indicators (CPI, fuel prices, etc.)

**Dataset Overview**

The analysis is based on three CSV files:

* **train.csv** – Historical sales data
* **features.csv** – Contains Temperature, Fuel Price, CPI, and Unemployment
* **stores.csv** – Metadata about each store (Store number, Type, Size)

**Key Columns:**

* Store, Dept, Date, Weekly Sales, IsHoliday, Temperature, Fuel Price, CPI, Unemployment, Type, Size

**Data Preprocessing**

* Merged datasets on Store, Date, and IsHoliday
* Converted Date column to datetime format
* Filled missing values using **forward fill**
* Cleaned data was used for generating visual insights

**Exploratory Data Analysis (EDA)**

**1. Weekly Sales Trend**

* **Tool:** Line Plot
* **Color:** Dark Violet
* **Insight:** Identifies seasonal patterns; peaks around holidays

**2. Total Sales by Store**

* **Tool:** Bar Plot
* **Color:** Medium Sea Green
* **Insight:** Highlights top-performing and low-performing stores

**3. Total Sales by Department**

* **Tool:** Bar Plot
* **Color:** Salmon
* **Insight:** Some departments (like Grocery) consistently outperform others

**4. Holiday vs Non-Holiday Sales**

* **Tool:** Bar Plot with Hue
* **Colors:** Holiday – Medium Violet Red, Non-Holiday – Medium Turquoise
* **Insight:** Sales are higher during holiday periods

**5. Correlation Heatmap**

* **Tool:** Heatmap
* **Colormap:** Viridis (customizable)
* **Insight:** Weak negative correlation between sales and Unemployment/CPI

**Key Observations**

* Holiday periods show a significant rise in weekly sales
* Store #20 leads in total sales volume
* Economic factors like CPI and unemployment show low but noticeable correlations
* Store size and type affect total sales but need deeper analysis

**Conclusion**

Walmart's sales are influenced strongly by holiday events and store-level characteristics. Sales patterns exhibit seasonality, making holiday marketing and inventory planning crucial. While economic indicators are less influential, they still play a minor role.

**Future Scope**

* Build predictive models using time-series forecasting (e.g., ARIMA, Prophet)
* Create dashboards using Power BI or Tableau
* Deep-dive analysis of departments across store types
* Correlate weather and local events with sales spikes

**Tech Stack**

* Python 3.12
* Pandas, Matplotlib, Seaborn
* VS Code
* Git + GitHub

**Directory Structure**

sales\_data\_analysis/

├── data/ # Raw CSV files

├── outputs/

│ └── plots/ # Saved charts and visualizations

├── src/

│ ├── utils.py # Data functions and plotting code

│ └── main.py # Script to run entire workflow

├── report.md # This report file

**Appendix (Optional)**

* All plots are saved under outputs/plots/
* Plots can be regenerated anytime by running:
* python src/main.py