EXPLORATORY DATA ANALYSIS REPORT

Business Understanding :

The business problem revolves around understanding **weekly sales performance** across multiple Walmart stores .
Walmart wants to identify:

- How revenue varies across stores
- Impact of holidays on sales
- Effect of economic indicators such as
  – Temperature
  – Fuel Price
  – CPI
  – Unemployment
- Seasonal and monthly sales behavior
- Long-term sales trends over the years

The goal is to help the retail business make data-driven decisions regarding **promotions, staffing, inventory, store operations, and pricing strategies**.

Data Set Overview :

The dataset contains weekly transactional data from multiple Walmart stores.

Columns & Meanings

Column	Description
Store	Store ID (1–45)
Date	Weekly sale date
Weekly_Sales	Total revenue generated in that week
Holiday_Flag	1 = holiday week, 0 = normal week
Temperature	Temperature in that region
Fuel_Price	Fuel price in that region
CPI	Consumer Price Index
Unemployment	Regional unemployment rate

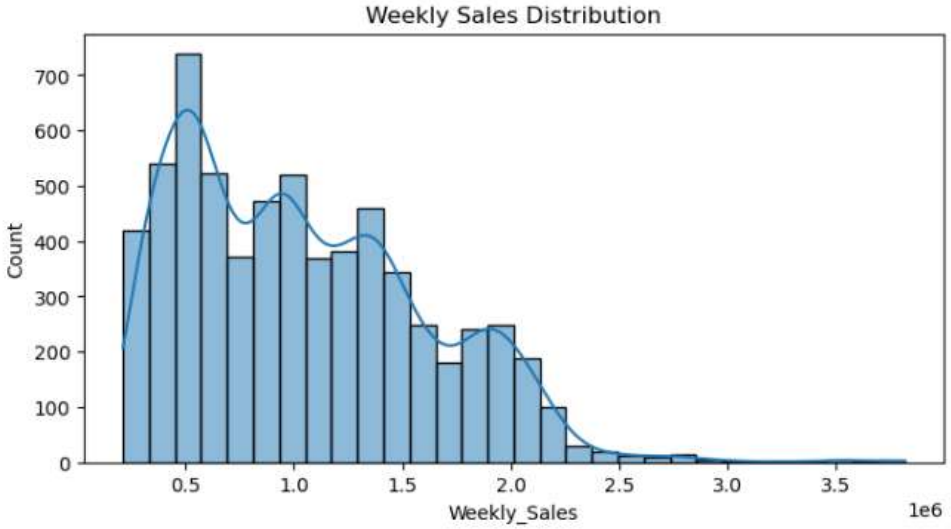
Created Columns

- Month – Month name
- Year – Year extracted from date
- Week – Week number (1–52)

Univariate Analysis :

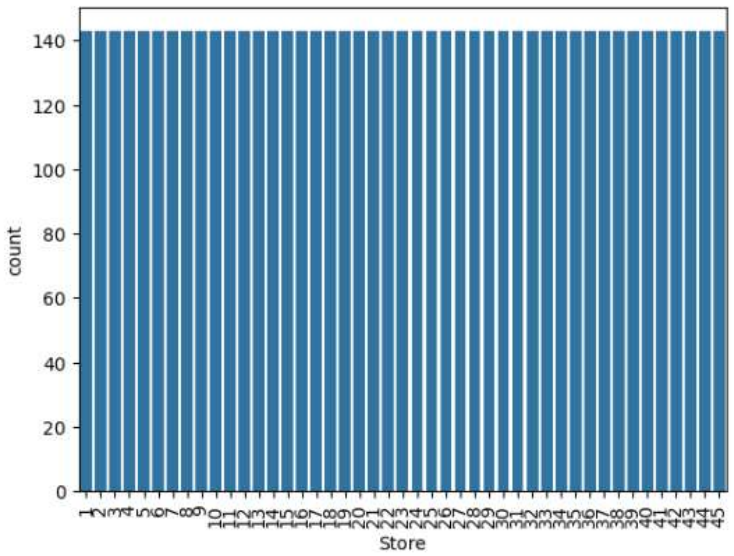
3.1 Weekly Sales Distribution

- Weekly sales are right-skewed, meaning most weeks have moderate sales but few weeks have extremely high revenue.
- Indicates seasonal or promotional spikes.



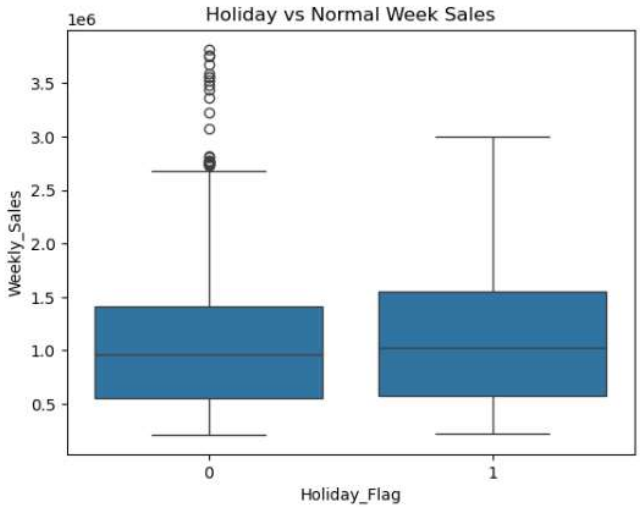
3.2 Store Count Distribution

- All stores have nearly equal participation.
- Helps ensure uniformity in comparison.



3.3 Holiday vs Non-Holiday Sales

- Holiday weeks show higher median sales, confirming holiday impact.



3.4 Summary Statistics

- Large range in weekly sales shows high variability.
- CPI, fuel price, and unemployment vary across regions, affecting performance.

	Store	Date	Weekly_Sales	Holiday_Flag	Temperature	Fuel_Price	CPI	Unemployment	Year	Week
count	6435.000000	6435	6.435000e+03	6435.000000	6435.000000	6435.000000	6435.000000	6435.000000	6435.000000	6435.0
mean	23.000000	2011-06-17 00:00:00	1.046965e+06	0.069930	60.663782	3.358607	171.578394	7.999151	2010.965035	25.818182
min	1.000000	2010-02-05 00:00:00	2.099862e+05	0.000000	-2.060000	2.472000	126.064000	3.879000	2010.000000	1.0
25%	12.000000	2010-10-08 00:00:00	5.533501e+05	0.000000	47.460000	2.933000	131.735000	6.891000	2010.000000	14.0
50%	23.000000	2011-06-17 00:00:00	9.607460e+05	0.000000	62.670000	3.445000	182.616521	7.874000	2011.000000	26.0
75%	34.000000	2012-02-24 00:00:00	1.420159e+06	0.000000	74.940000	3.735000	212.743293	8.622000	2012.000000	38.0
max	45.000000	2012-10-26 00:00:00	3.818686e+06	1.000000	100.140000	4.468000	227.232807	14.313000	2012.000000	52.0
std	12.988182	NaN	5.643666e+05	0.255049	18.444933	0.459020	39.356712	1.875885	0.797019	14.129201

SQL With Business Questions :

1. Total Sales for Each Store –

```
SELECT Store, SUM(Weekly_Sales) AS Total_Sales
FROM Walmart
GROUP BY Store
ORDER BY Total_Sales DESC;
```

[21]:

	Store	Total_Sales
0	20	3.013978e+08
1	4	2.995440e+08
2	14	2.889999e+08
3	13	2.865177e+08
4	2	2.753824e+08
5	10	2.716177e+08
6	27	2.538559e+08
7	6	2.237561e+08
8	1	2.224028e+08
9	39	2.074455e+08
10	19	2.066349e+08
11	31	1.996139e+08
12	23	1.987506e+08
13	24	1.940160e+08
14	11	1.939628e+08
15	28	1.892637e+08
16	41	1.813419e+08
17	32	1.668192e+08
18	18	1.551147e+08
19	22	1.470756e+08
20	12	1.442872e+08
21	26	1.434164e+08
22	34	1.382498e+08
23	40	1.378703e+08
24	35	1.315207e+08
25	8	1.299512e+08
26	17	1.277821e+08
27	45	1.123953e+08
28	21	1.081179e+08
29	25	1.010612e+08
30	43	9.056544e+07
31	15	8.913368e+07
32	7	8.159828e+07
33	42	7.956575e+07
34	9	7.778922e+07
35	29	7.714155e+07
36	16	7.425243e+07
37	37	7.420274e+07
38	30	6.271689e+07
39	3	5.758674e+07
40	38	5.515963e+07
41	36	5.341221e+07
42	5	4.547569e+07
43	44	4.329309e+07
44	33	3.716022e+07

2. Which month has highest average sales ?

```
SELECT MONTH (Date) AS MONTH, AVG(Weekly_Sales) AS Avg_Sales
FROM Walmart
GROUP BY MONTH
ORDER BY Avg_Sales DESC;
```

	Month	Avg_Sales
0	12	1.281864e+06
1	11	1.147266e+06
2	06	1.064325e+06
3	02	1.053200e+06
4	08	1.048017e+06
5	07	1.031748e+06
6	05	1.031714e+06
7	04	1.026762e+06
8	03	1.013309e+06
9	10	9.996321e+05
10	09	9.893353e+05
11	01	9.238846e+05

3. Does Holiday Week Increase Revenue ?

```
SELECT Holiday_Flag, AVG(Weekly_Sales) As Avg_Sales
FROM Walmart
GROUP BY Holiday_Flag;
```

	Holiday_Flag	Avg_Sales
0	0	1.041256e+06
1	1	1.122888e+06

4. Top 10 Highest Revenue Weeks ?

```
SELECT Date, Store, Weekly_Sales
From Walmart
Order By Weekly_Sales Desc;
Limit 10;
```

	Date	Store	Weekly_Sales
0	2010-12-24 00:00:00	14	3818686.45
1	2010-12-24 00:00:00	20	3766687.43
2	2010-12-24 00:00:00	10	3749057.69
3	2011-12-23 00:00:00	4	3676388.98
4	2010-12-24 00:00:00	13	3595903.20
5	2011-12-23 00:00:00	13	3556766.03
6	2011-12-23 00:00:00	20	3555371.03
7	2010-12-24 00:00:00	4	3526713.39
8	2011-12-23 00:00:00	10	3487986.89
9	2010-12-24 00:00:00	2	3436007.68

5. Fuel Price vs Sales

```
SELECT Fuel_Price, AVG(Weekly_Sales) AS Avg_Sales
FROM Walmart
GROUP BY Fuel_Price
ORDER BY Fuel_Price;
```

	Fuel_Price	Avg_Sales
0	2.472	4.702810e+05
1	2.513	4.344714e+05
2	2.514	1.015099e+06
3	2.520	4.475194e+05
4	2.533	4.312945e+05
...	...	...
887	4.294	8.758470e+05
888	4.301	8.531141e+05
889	4.308	8.875630e+05
890	4.449	8.440218e+05
891	4.468	8.659708e+05

892 rows × 2 columns

6. Yearly Sales Comparison

```
SELECT YEAR(Date) AS YEAR , SUM(Weekly_Sales) AS Total_Sales
FROM Walmart
GROUP BY YEAR;
```

	Year	Total_Sales
0	2010	2.288886e+09
1	2011	2.448200e+09
2	2012	2.000133e+09

7. CPI Impact on Weekly Sales

```
SELECT CPI, AVG(Weekly_Sales) AS Avg_Sales
FROM Walmart
GROUP BY CPI
ORDER BY CPI;
```

	CPI	Avg_Sales
0	126.064000	1.016970e+06
1	126.076645	1.021504e+06
2	126.085452	1.056102e+06
3	126.089290	9.676731e+05
4	126.101935	1.040635e+06
...	...	...
2140	227.018417	9.003098e+05
2141	227.036936	8.916714e+05
2142	227.169392	5.584648e+05
2143	227.214288	5.420095e+05
2144	227.232807	5.497315e+05

2145 rows × 2 columns

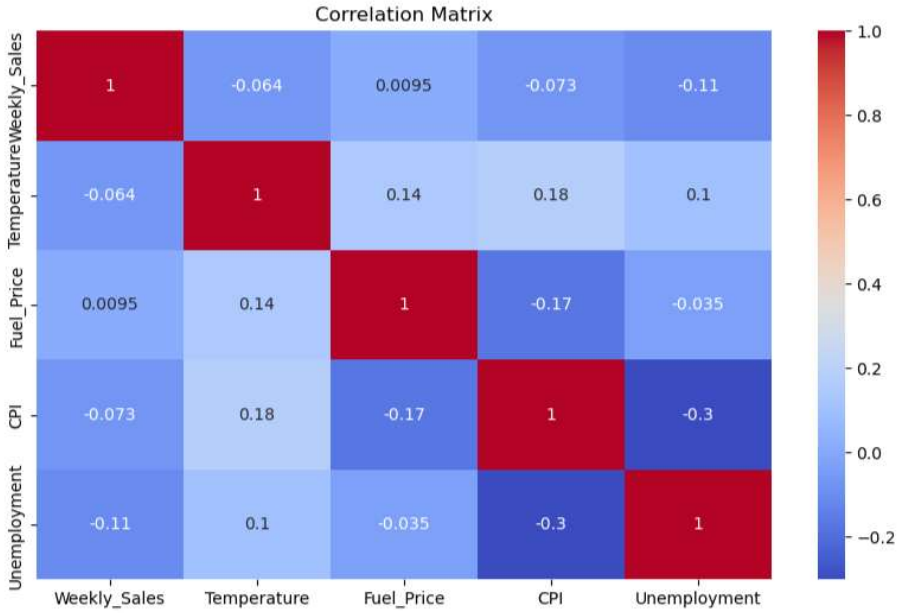
MULTIVARIATE ANALYSIS:

Correlation Matrix :

Key Insights :

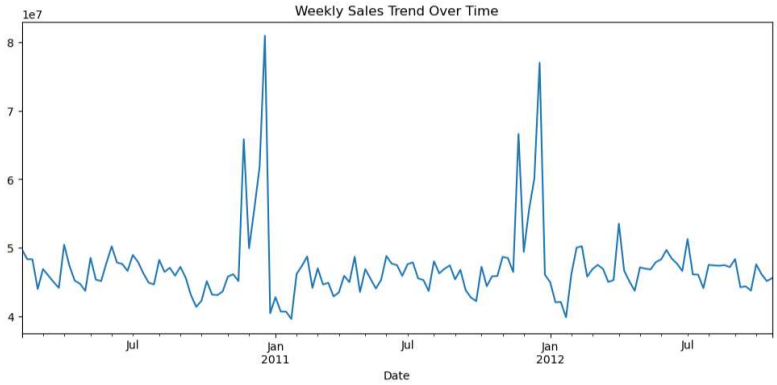
CPI & Unemployment are strongly related (economic conditions).

- Weak correlation between sales and fuel price indicates **price fluctuation does not heavily affect spending**.
- Holiday flag shows mild positive correlation with weekly sales.



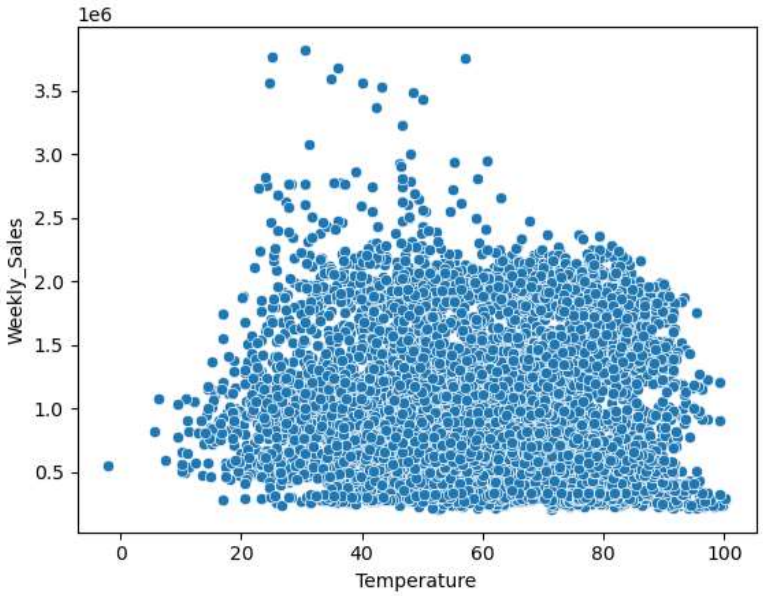
Trend Over Time :

- Clear seasonal spikes every year.
- Sales rise during festivals and drop in off-season months.



Temperature vs Weekly Sales :

Slight negative trend → Hotter weeks may see lower sales**.**



Fuel Price vs Weekly Sales :

Weak relationship → Customers continue to purchase despite fuel price variations.