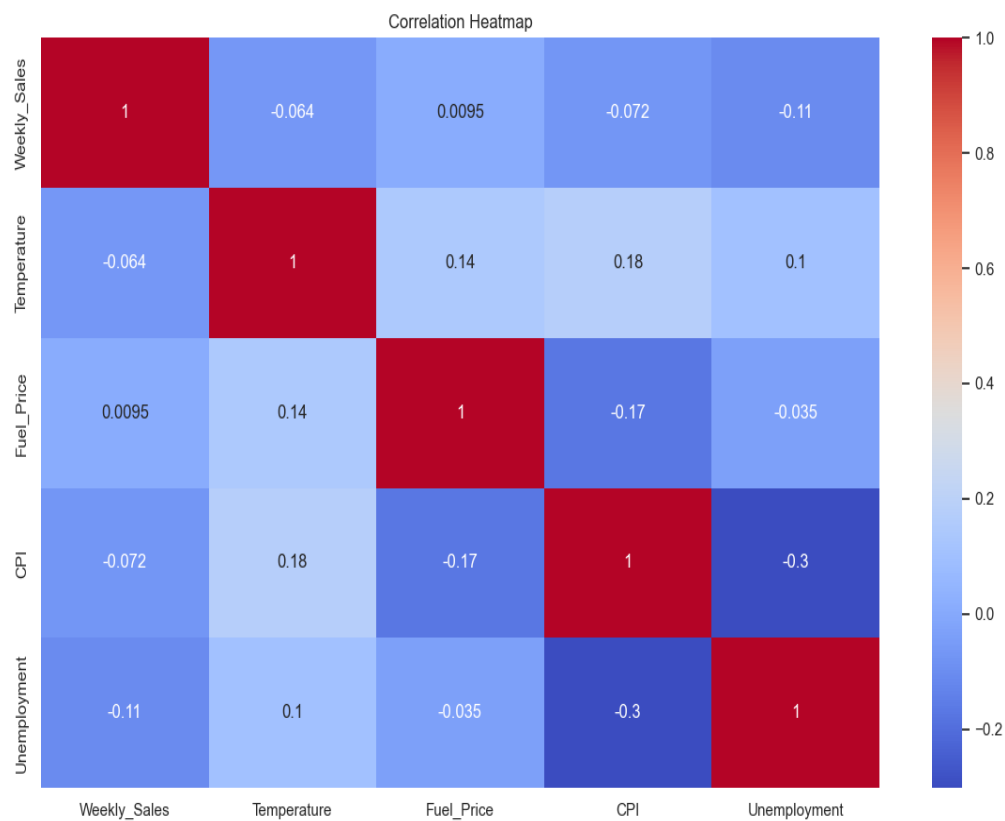


“WALMART SALES ANALYSIS “

BY: Payal Sharma

GRAPH DESCRIPTION

1. CORRELATION HEATMAP

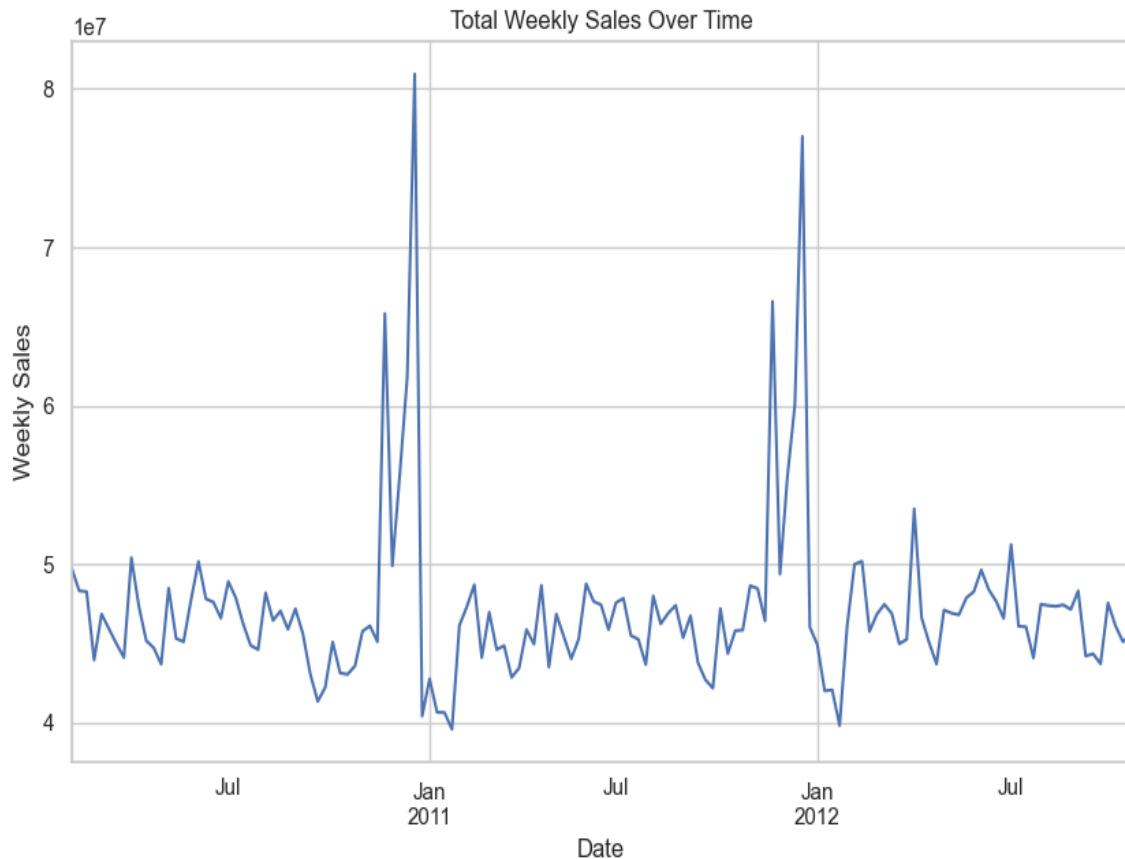


- **Explanation:** This heatmap shows the **correlation** between different variables in the dataset. Correlation is a statistical measure that indicates how closely two variables move together. The values range from -1 to +1. A value close to **+1** (dark red) indicates a strong positive correlation (when one variable increases, the other tends to increase). A value close to **-1** (dark blue) indicates a strong negative correlation (when one increases, the other decreases). A value close to **0** means there is no significant linear relationship.

- **Insights:** The most important insight from this map is the relationship between **Weekly_Sales** and other variables.
 - **Temperature** has a weak positive correlation with sales, meaning warmer weather might be associated with slightly higher sales.
 - **Fuel_Price** and **Unemployment** have a weak to moderate negative correlation with sales, suggesting that as fuel prices and unemployment rise, sales tend to fall. This makes intuitive sense as these factors reflect a weaker economy and reduced consumer spending.
 - **CPI (Consumer Price Index)** has a moderate negative correlation with sales, indicating that as the cost-of-living increases, sales might decrease.

- **Future Business Use:** The insights from this heatmap are crucial for building a strong predictive model. They confirm which external factors are most relevant to sales. In the future, this information can be used to:
 - **Develop Strategic Pricing:** Walmart could adjust prices or promotions in response to changes in fuel prices or unemployment rates.
 - **Refine the Predictive Model:** The features with the strongest correlations (like Unemployment and CPI) should be prioritized in more complex models to improve forecasting accuracy.
 - **Monitor Economic Indicators:** The company can keep a close eye on economic data to anticipate shifts in consumer behavior and adjust its business strategy accordingly.

2. TOTAL WEEKLY SALES OVER TIME



Insights

1. Overall Stability with Spikes

- Most weeks, sales stay between **\$40M–\$50M**.
- However, there are **significant spikes around late December and early January** in both 2011 and 2012, where sales peak above **\$80M**.

2. Seasonality

- The spikes align with **holiday shopping seasons (Christmas/New Year)**.
- This indicates **highly seasonal demand**, with sales doubling during these periods compared to the baseline.

3. Dips in Sales

- Noticeable dips just before and after the spikes, suggesting post-holiday slowdowns.

4. No Strong Upward/Downward Trend

- Apart from seasonal effects, the **baseline sales remain relatively flat** over time, showing **stable but not strongly growing sales**.

What It Tells for the Future

1. Inventory & Supply Chain Planning

- Ensure **extra stock, logistics, and staff** during December–January to capture peak demand.
- Avoid overstocking in **February–March** when sales drop back down.

2. Promotions & Marketing

- Invest more in **holiday campaigns** (Black Friday, Christmas, New Year) since these are proven high-revenue periods.
- Run **clearance or loyalty offers post-holiday** to smooth out the demand dip.

3. Forecasting

- Use **time-series forecasting models (ARIMA, Prophet, or LSTM)** that account for **seasonality**.
- Predict spikes more accurately for **better financial planning**.

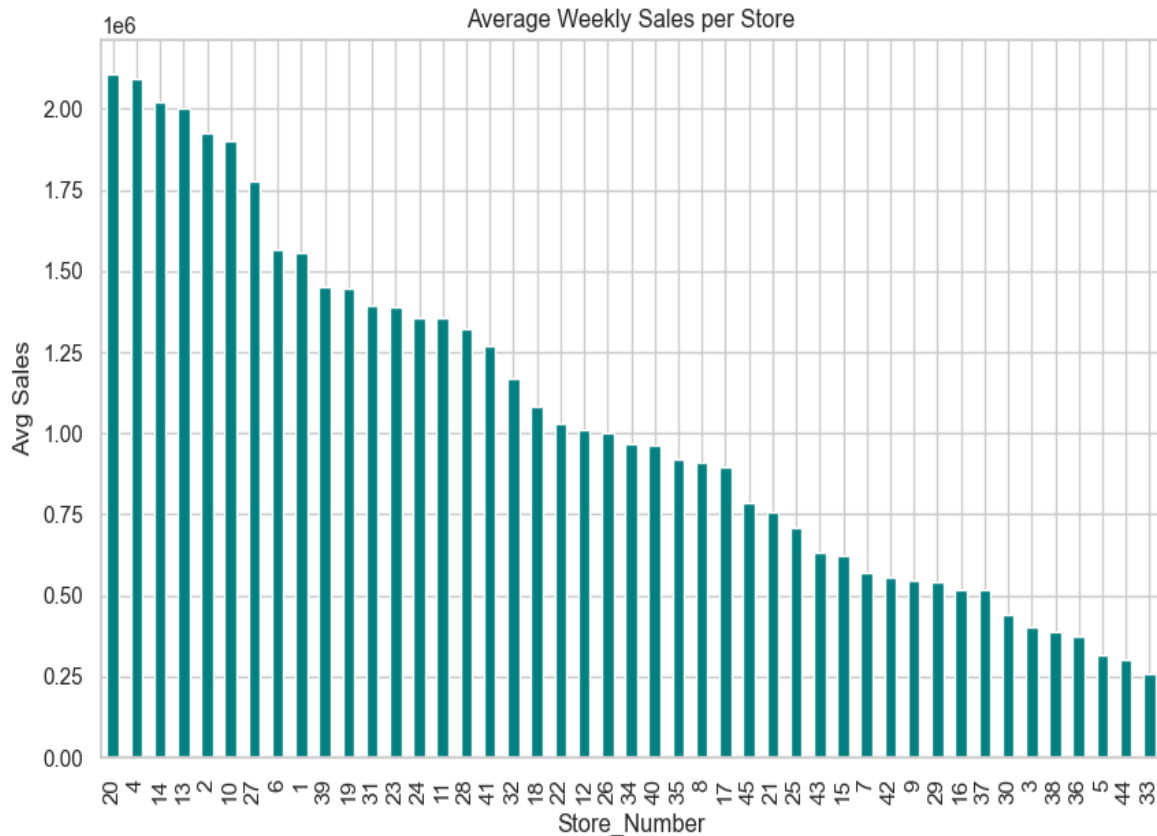
4. Diversification of Revenue

- Explore strategies to **boost sales in off-peak months** (April–September), e.g., summer promotions, product bundling, or targeting different customer needs.

5. Operational Efficiency

- Since sales are stable most of the year, the company can **optimize resources** in normal months and scale up during seasonal demand.

3. AVERAGE WEEKLY SALES PER STORE



Insights

1. High Variability Across Stores

- Some stores (like **Store 20, Store 4, Store 14**) achieve **over \$2M average weekly sales**,
- While others (like **Store 33, Store 5, Store 44**) average **only around \$250K–\$400K**.
- This indicates a **large performance gap** between top and bottom stores.

2. Pareto Effect (80/20 Rule)

- A small number of stores contribute disproportionately to overall sales.
- Top ~10 stores generate significantly more revenue compared to the bottom 20+ stores.

3. Potential Store Characteristics

- High-performing stores may be in **urban, high-traffic, or economically stronger regions**.
- Low-performing stores may face **location disadvantages, competition, or operational inefficiencies**.

4. Room for Optimization

- If low-performing stores can even partially match the performance of mid-level ones, overall revenue would increase substantially.

What It Tells for the Future

1. Best Practice Replication

- Study what **top-performing stores (20, 4, 14, etc.)** are doing right — location, promotions, product mix, staffing, or customer engagement.
- Apply these strategies to mid- and low-performing stores.

2. Targeted Interventions for Low-Performers

- Diagnose reasons why certain stores (33, 5, 44) underperform.
- Interventions may include:
 - Better inventory management
 - Regionalized promotions
 - Store layout improvements
 - Training for staff/customer service

3. Resource Allocation

- Invest **more marketing and logistics support** into top-performing stores, since they deliver higher returns.
- For underperformers, either improve efficiency or evaluate if resources should be reduced/reallocated.

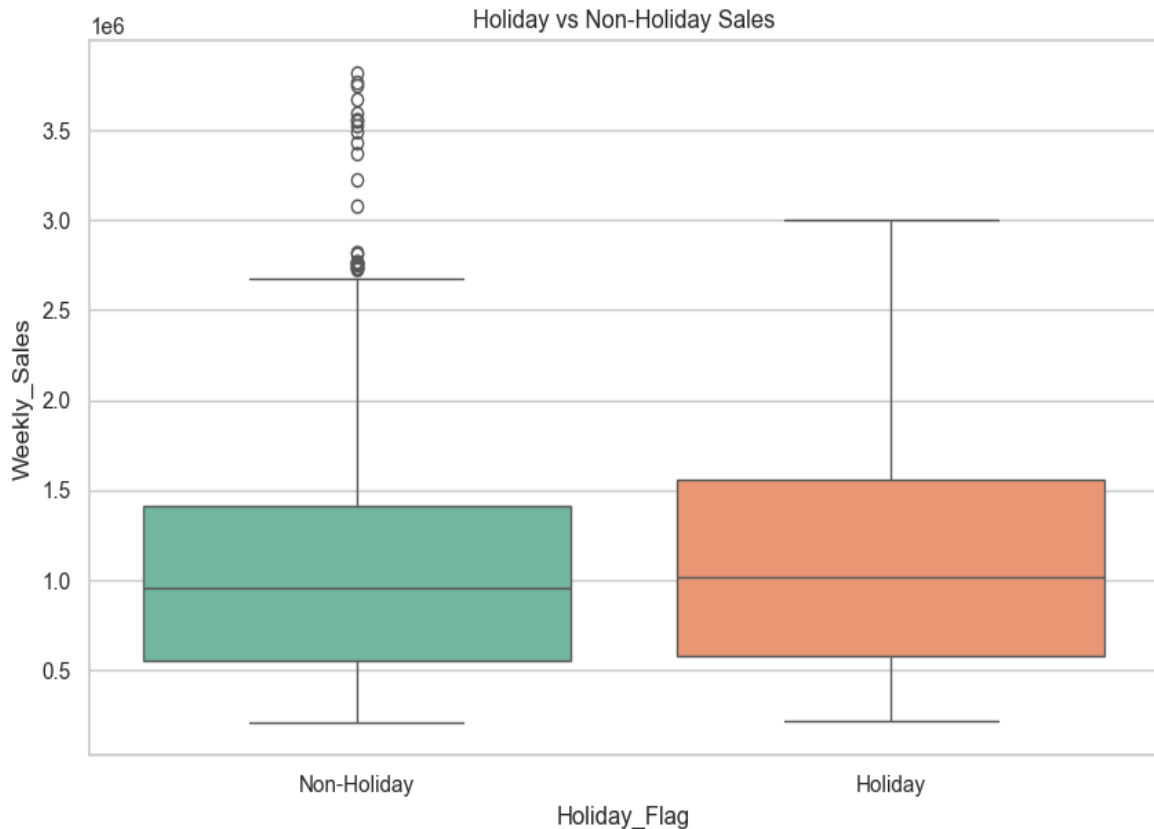
4. Customer Segmentation & Regional Insights

- Conduct **regional analysis** to see if demographics drive sales differences.
- Tailor product offerings (e.g., local preferences, seasonal items) by store.

5. Long-Term Strategy

- Consider **closing or rebranding persistently underperforming stores** if improvements don't yield results.
- Explore **new store openings in areas resembling top performers**.

4.HOLIDAY VS NON-HOLIDAY SALES



Insights

1. Median Sales Comparison

- **Holiday weeks** have a slightly **higher median sales** than non-holiday weeks (~\$1M vs ~\$0.9M).
- This confirms holidays do boost sales, but not drastically for the “average” week.

2. Spread (Variation in Sales)

- **Holiday weeks** show a **wider spread (greater variability)** in sales, meaning some holidays perform extremely well while others are closer to normal weeks.
- Non-holiday weeks have a **tighter spread**, showing more consistent performance.

3. Outliers in non-Holiday Weeks

- Interestingly, many **extreme outliers** (above \$3M) appear in the **non-holiday weeks**.
- These likely align with special events (e.g., Black Friday, promotions) that may not be tagged as “holiday” in the dataset.

4. Upper Bound of Sales

- Holiday weeks reach up to ~\$3M, while some non-holiday weeks reach even higher due to outliers.
- This suggests **holidays guarantee some boost, but exceptional spikes can also happen outside official holiday flags.**

What It Tells for the Future

1. Holiday Planning

- Holidays bring higher median sales, so stores should **prepare with inventory, staffing, and promotions** during holiday periods.

2. Event-Based Strategy

- Since large outliers occur in **non-holiday weeks**, the company should also capitalize on **promotional events (Black Friday, back-to-school, clearance sales)** that may not fall under “official holidays.”

3. Forecasting

- Models should account for **holiday flags** but also consider **other event-driven spikes**.
- A combined **seasonality + event calendar model** will give better accuracy.

4. Marketing Investments

- Strong marketing campaigns during **holidays and special events** can maximize the already higher demand.
- Focus particularly on **major shopping events** that outperform even holiday weeks.

5. ACTUAL VS PREDICTED SALES

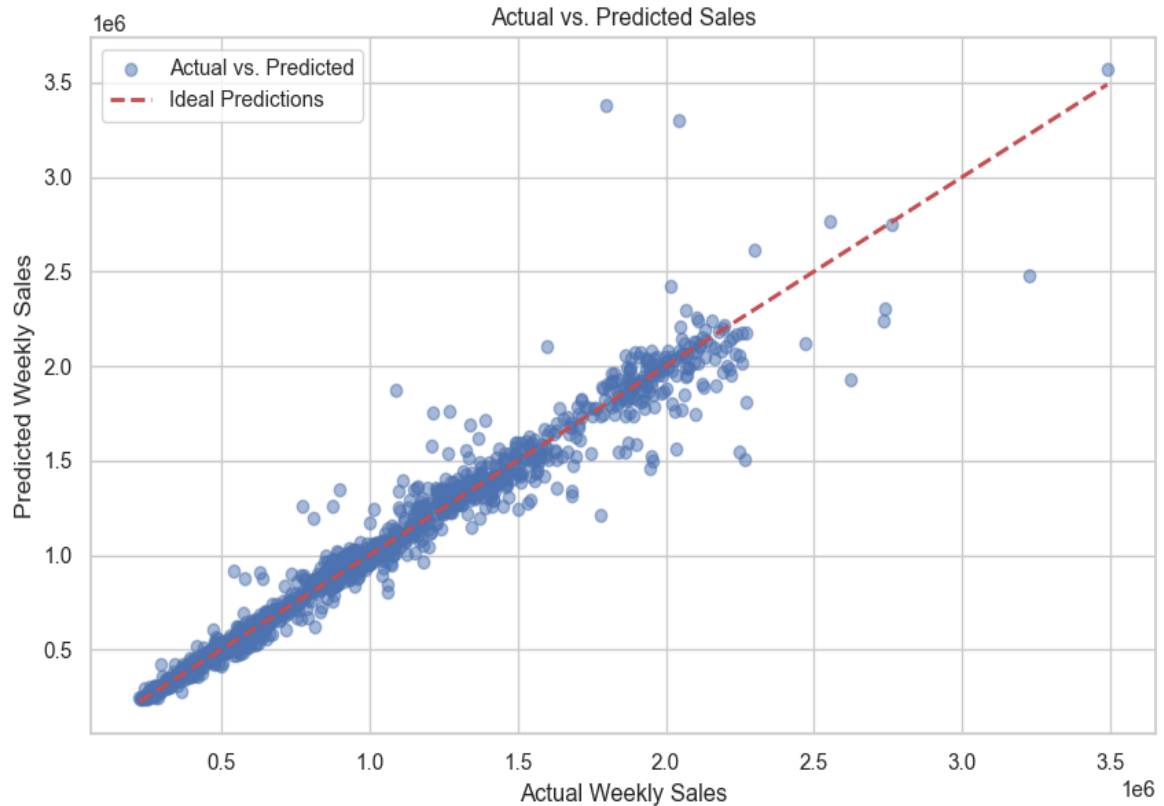
Insights

1. Strong Model Fit

- Most points lie close to the **red diagonal line (ideal predictions)**.
- This means the model predicts sales with **high accuracy** across most weeks.

2. Good Performance for Majority Cases

- For **low-to-medium sales ranges (up to ~\$2M)**, the predictions align very well with actuals.
- The model captures general sales behavior effectively.



3. Deviation at High Sales

- At **higher sales values (> \$2M)**, predictions start to **underestimate actuals**.
- The model struggles slightly to capture extreme spikes (e.g., holiday peaks, Black Friday).

4. Few Outliers

- Some scattered points show **larger prediction errors**. These are likely special event weeks not fully captured by the model.

What It Tells for the Future

1. Model Reliability

- The model can be trusted for **baseline and regular forecasting**.
- Useful for **inventory planning, staffing, and financial projections** in non-extreme weeks.

2. Improvement for High Sales Weeks

- Need to enhance the model's ability to capture **holiday/event-driven surges**.
 - Consider adding:
 - **Holiday/event flags**
 - **Promotional campaign data**
 - **External factors (weather, economic indicators)**
- 3. Business Use Case**
- Reliable forecasting allows **proactive decision-making**:
 - Avoids understocking/overstocking
 - Optimizes workforce allocation
 - Supports accurate revenue projections
- 4. Continuous Monitoring**
- Forecasting should be **iteratively retrained with new data** to adapt to changing consumer behavior.

CONCLUSION:

Walmart's sales data shows a **stable baseline throughout the year**, with **sharp spikes during holidays and special events**, highlighting strong seasonality. A few **top-performing stores drive a major share of revenue**, while many others underperform, suggesting room for targeted improvements. Holiday weeks deliver higher sales, but some of the **biggest spikes occur outside official holidays**, likely due to promotions or events like Black Friday. Forecasting models perform well overall but **underestimate peak sales**, meaning external factors like holidays, promotions, and regional demand should be incorporated.

What to Do Next:

- Prepare **inventory, staffing, and logistics** well ahead of holiday and event weeks.
- **Replicate best practices** from top stores and intervene in underperforming ones.
- Use **event-driven forecasting** to better capture peak sales.
- Explore **off-season promotions** to stabilize sales during slower months.