Walmart Sales Project

EDS 6345

Dr. Lucy Nwosu

Sydney Ani

Caleb C Hairston

Ruthwik Reddy

Lujia Wu

August 11, 2025

Contents

1. Project Background and Objectives

The primary goal of this project was to use Tableau to visualize and analyze Walmart store sales data from 2010 to 2012, exploring the key factors influencing sales fluctuations, focusing specifically on the impact of holidays on sales, and analyzing quarterly, monthly, and semester sales patterns. Furthermore, a forecasting model was used to assess the impact of external factors such as the CPI, unemployment rate, and fuel prices on Store 1's sales, providing management with data-driven decision-making.

2. Project Task Assignment

Dashboard Creation: Sydney Ani, Caleb C Hairston, Ruthwik Reddy Karna, Lujia Wu

- Home Dashboard: Sydney Ani, Caleb C Hairston
- Total Sales KPI: Caleb C Hairston
- Sales Insights: Ruthwik Reddy
- Quarter Growth: Ruthwik Reddy
- Holiday Insights: Lujia Wu
- Monthly Insights: Sydney Ani
- Semester Insights: Sydney Ani
- Store 1 Prediction: Lujia Wu

3. Dataset Description

The dataset covers weekly sales data for 45 Walmart stores from 2010 to 2012, along with CPI, unemployment rate, fuel price, temperature, and holiday information. Key fields include:

- Date
- Store
- Weekly_Sales
- Holiday_Flag / Holiday_Type
- CPI
- Fuel Price
- Temperature
- Unemployment

3.1 Calculated Fields:

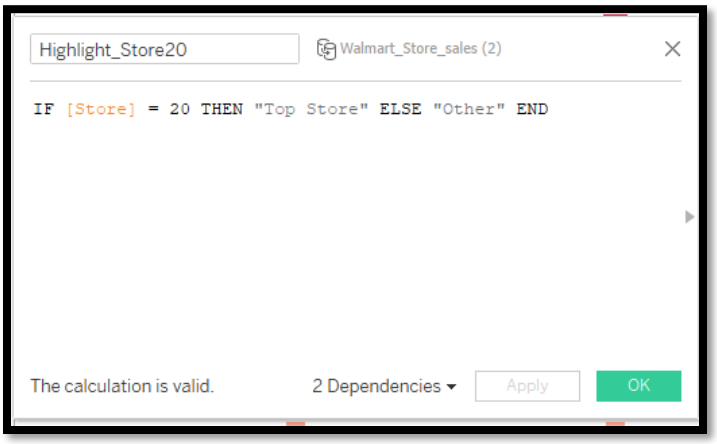
Holiday_Type:

- For this calculated field, we wanted to ensure that the Holiday_Type was specifically determining the correct seasons of the year. Hence the IF MONTH = 2 then Super Bowl, ELSEIF MONTH = 9 then Labor Day, etc. Every other Month that did not have a specific holiday in the Holiday_Type data was marked as “Other Holiday”.

Semester:

- For this calculated field, to divide the semesters successfully into 2 halves, we made it to where it will locate where the date is equal to 6 (June) and deemed that as “Semester 1” and ELSE it will be deemed “Semester 2”

Highlight_Store20:



- For this calculated field, we decided to create a function where if the store = 20, then it was the “Top Store” and ELSE every other store will be deemed as “Other”

Top_Store_Display:

- For this calculated field, we wanted to get the KPI for the Top Store that we have made previously in the “Highlight_Store20” table, so we created a function that determined if the the current store's total equals the maximum total across all stores. This is only TRUE for the store with the highest sales (Store 20). If it's the top store, creates a formatted string showing both the Store name/number and the Total sales amount.

Quarter:

- For this calculated field we wanted to ensure that the Date that was being spit out was the Quarter and nothing else.

Month:

- For this calculated field we wanted to ensure that the Date that was being spit out was the Month and nothing else.

4. Analytical Methods and Visualization Tools

Tableau was used as the primary analytical tool with the following approaches:

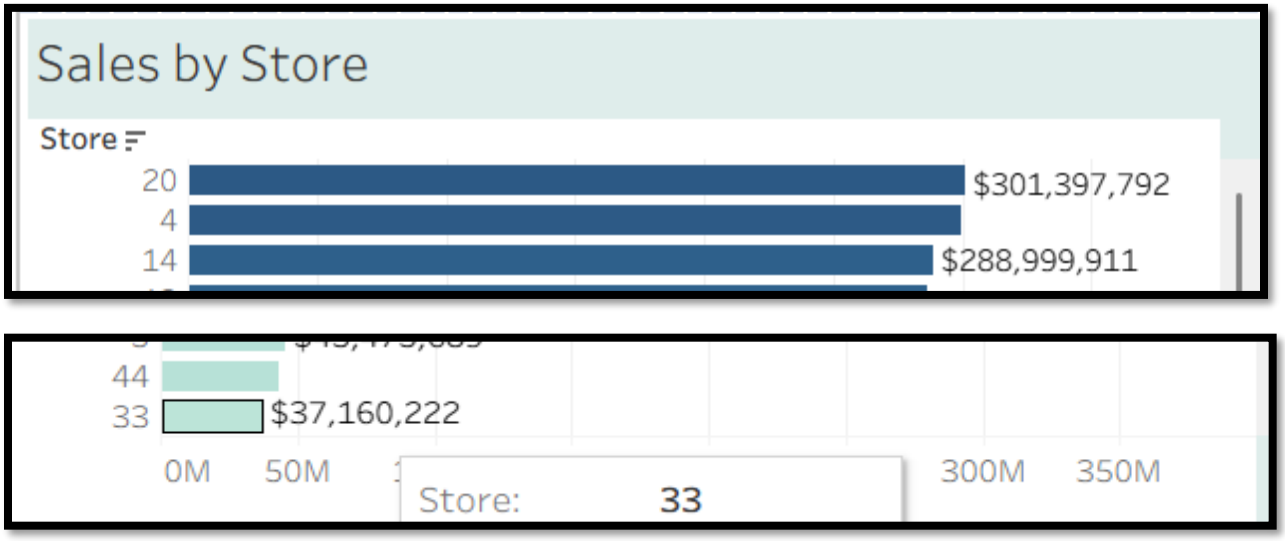
- Sales Comparison: Bar charts to compare total sales across stores.

- Sales Variability Analysis: Calculated standard deviation (STDEV) of Weekly Sales for each store and visualized results.
- Quarterly Growth Analysis: Focused on Q3 2012 store sales growth.
- Holiday Impact Analysis: Compared average sales during holidays vs. non-holidays.
- Monthly & Semester Sales Patterns: Line charts and heat maps to identify temporal trends.
- Predictive Modeling: Built a linear regression model for Store 1 to assess the effects of CPI, unemployment rate, and fuel price.

5. Key Findings

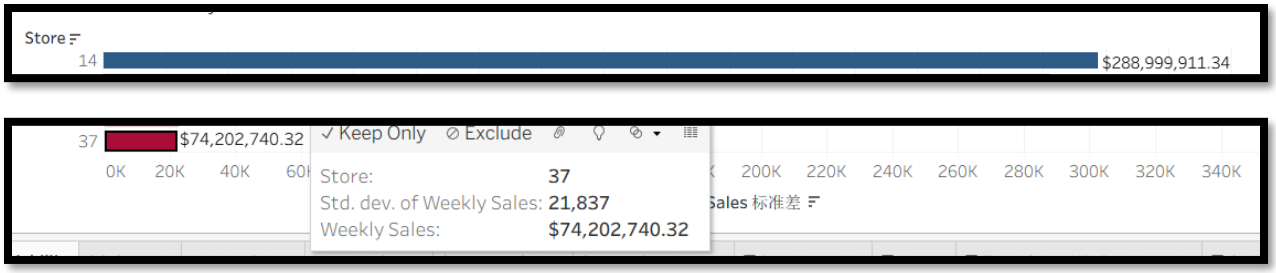
5.1 Sales by Store:

- Store 20 recorded the highest sales with a total of $301,397,792.
- Store 4 and Store 14 followed closely with $288,999,911 and $275,382,441 respectively.
- Store 33 had the lowest sales, totaling $37,160,222.



5.2 Sales Variability:

- Store 14 ($288,999,911) exhibited the highest sales variability (STDEV), indicating significant fluctuations over time.
- Store 37 ($74,202,740) showed the lowest variability, with relatively stable sales.



5.3 Q3 2012 Growth

- Stores 4, 13, and 20 showed outstanding performance in Q3 2012, with significantly higher sales than other stores.
- These stores may have benefited from effective promotions or favorable regional economic conditions.

5.4 Holiday vs. non-Holiday Sales

The comparison of weekly sales between different holiday types reveals notable variations in consumer spending behavior. Thanksgiving emerges as the most impactful holiday, with average weekly sales reaching approximately 1.47 million, significantly higher than any other holiday or non-holiday period. This suggests that Thanksgiving-related promotions and consumer demand drive a substantial sales surge.

In contrast, Christmas records the lowest average weekly sales among the holiday types, at about 0.96 million, which is even lower than non-holiday weeks. This could be attributed to factors such as early-season shopping in November or increased competition from other retailers, leading to reduced sales in the Christmas week itself.

Labor Day and non-holiday weeks both show similar performance, with average weekly sales of around 1.04 million, indicating that Labor Day promotions have only a modest effect compared to the baseline sales. The Super Bowl week averages 1.08 million, slightly outperforming both Labor Day and non-holiday weeks, likely due to increased spending on food, beverages, and related products for game-day events.

Overall, the data highlights the critical importance of tailoring promotional strategies for specific holidays. Thanksgiving offers the highest potential for revenue generation, while Christmas may require revised marketing approaches to boost sales during the holiday week.

5.5. Monthly & Semester Sales

- Monthly sales showed a clear peak in November and December, reflecting holiday promotion effects.
- Semester 1 sales were generally higher than Semester 2, possibly linked to annual consumption cycles.

5.6 Store 1 Prediction

The linear regression model estimated the relationship between year and weekly sales for Store 1 as:

- Weekly Sales = $-1.44377\times10^{8} \times \text{Year} + 2.92587\times10^{11}$.
- The negative coefficient suggests a decreasing trend in sales over the years. However, with $R^2 = 0.4041$ and a P-value of 0.5615, the model demonstrates limited explanatory power and fails to reach statistical significance, indicating that year alone is not a strong predictor of weekly sales. Furthermore, CPI, unemployment, and fuel price showed no significant predictive effect, likely due to the short-term dominance of store specific or seasonal factors such as promotions, holidays, and local events. Incorporating additional variables such as weather conditions, competitor activity, and promotional campaign data may improve forecasting accuracy.

6. Conclusions and Recommendations

In conclusion, this analysis has provided a comprehensive evaluation of the dataset through descriptive statistics, visualizations, and trend analysis, offering valuable insights into key patterns and relationships within the data. By systematically examining the variables and their interactions, we identified significant trends, anomalies, and correlations that can guide data-driven decision-making. The results reveal not only the central tendencies and variability but also potential areas for improvement, such as addressing missing values, refining data collection procedures, and monitoring for substantial fluctuations.

Based on these findings, several actionable recommendations are proposed. High-performing stores, such as Store 20, should maintain their current strategies while assessing their feasibility for replication across other locations. Stores experiencing high sales volatility should investigate the underlying causes of fluctuations and optimize both inventory management and marketing approaches. Holiday promotions, particularly during the Thanksgiving period, should be prioritized to maximize revenue. For Store 1, incorporating additional business and macroeconomic indicators into the predictive model is recommended to enhance forecasting accuracy.

Overall, this study underscores the importance of integrating statistical analysis with visual interpretation to derive meaningful conclusions, thereby establishing a strong foundation for further research and practical applications.

Link to Tableau Public Dashboard:

https://public.tableau.com/shared/GJWHNXD2H?:display_count=n&:origin=viz_share_link