**A walmart store front with a blue sign

Description automatically generatedWalmart Sales Analysis Report**

**Introduction:**

Greetings and welcome to my Personalised Walmart Store Sales Analysis Report. Here, we will explore retail sales data to gain useful insights and practical suggestions. My objective as the Walmart Store Sales Analyst is to give you thorough insights that help you make strategic decisions and advance the company in the competitive retail market.

**Business Goal:**

The primary goal of this analysis is to uncover patterns and trends in Walmart's sales data that can help improve performance and drive strategic decision-making. By understanding factors influencing sales, Walmart can optimize inventory management, marketing strategies, and resource allocation to enhance profitability.

**Recommendations Based on Analysis:**

1. **Optimize Inventory Management:** Use sales trends and forecasts to adjust inventory levels, ensuring sufficient stock availability while minimizing excess inventory costs.
2. **Targeted Marketing Campaigns:** Leverage insights from holiday sales analysis to develop targeted marketing campaigns during peak shopping seasons, maximizing revenue opportunities.
3. **Adaptation to External Factors:** Monitor and respond to external factors such as temperature fluctuations, fuel prices, and economic indicators (CPI, unemployment rate) to adjust pricing, promotions, and product offerings accordingly.

**Exploration (About Data):** The dataset comprises sales data from Walmart stores, including weekly sales, holiday flags, temperature, fuel prices, CPI, and unemployment rates. Exploratory data analysis was conducted to understand the distribution, trends, and relationships within the dataset, providing insights into sales performance drivers.

**Analysis:**

1. **Data Cleaning:** I downloaded the CSV file from Kaggle, and opened it using Microsoft sql server via Azure Data Studio, I performed data cleaning to ensure there was no missing values, inconsistencies, and data format issues to ensure the integrity and reliability of the analysis.
2. --Perform some data cleaning
3. --check for missing values
4. SELECT \*
5. FROM sales
6. WHERE Store IS NULL OR Date IS NULL OR Weekly\_Sales IS NULL OR Holiday\_Flag IS NULL OR Temperature IS NULL OR Fuel\_Price IS NULL OR CPI IS NULL OR Unemployment IS NULL;
7. -- there is no missing value
8. --checking for duplicate rows
9. SELECT \*
10. FROM sales
11. GROUP BY Store, Date, Weekly\_Sales, Holiday\_Flag, Temperature, Fuel\_Price, CPI, Unemployment
12. HAVING COUNT(\*) > 1;
13. -- there are no duplicated values

I also performed some summary statistics to provide insights into the distribution and variability of each numerical column in your Walmart sales data.

--summary statistics for numeric columns

SELECT

MIN(Weekly\_Sales) AS Min\_Weekly\_Sales,

MAX(Weekly\_Sales) AS Max\_Weekly\_Sales,

AVG(Weekly\_Sales) AS Avg\_Weekly\_Sales,

MIN(Temperature) AS Min\_Temperature,

MAX(Temperature) AS Max\_Temperature,

AVG(Temperature) AS Avg\_Temperature,

MIN(Fuel\_Price) AS Min\_Fuel\_Price,

MAX(Fuel\_Price) AS Max\_Fuel\_Price,

AVG(Fuel\_Price) AS Avg\_Fuel\_Price,

MIN(CPI) AS Min\_CPI,

MAX(CPI) AS Max\_CPI,

AVG(CPI) AS Avg\_CPI,

MIN(Unemployment) AS Min\_Unemployment,

MAX(Unemployment) AS Max\_Unemployment,

AVG(Unemployment) AS Avg\_Unemployment

FROM sales;

--The weekly sales range from approximately $209,986.25 to $3,818,686.45, with an average weekly sale of approximately $1,046,964.88.

-- The temperature ranges from -2.06°F to 100.14°F, with an average temperature of approximately 60.66°F

-- The fuel price ranges from approximately $2.472 to $4.468, with an average fuel price of approximately $3.36.

-- The CPI ranges from approximately 126.064 to 227.233, with an average CPI of approximately 171.578.

-- The unemployment rate ranges from approximately 3.879% to 14.313%, with an average unemployment rate of approximately 7.999%.

1. **Key Performance Indicators (KPIs):** Identified KPIs such as total sales, sales growth rate, holiday sales ratio, average temperature, and unemployment rate to measure performance and track progress towards business goals.
2. --total revenue
3. SELECT FORMAT(SUM(Weekly\_Sales), 'C', 'en-US') AS Total\_Revenue
4. FROM sales;

**$6,737,218,987.11**

--no of stores

SELECT COUNT(DISTINCT Store) AS Number\_of\_Stores

FROM sales;

**45**

--unemployment rate

SELECT AVG(Unemployment) AS Average\_Unemployment\_Rate

FROM sales;

**8.0**

--avg temperature

SELECT AVG(Temperature) AS Average\_Temperature

FROM sales;

**The average temperature is approximately 60.66 degrees.**

--yearly sales growth rate

WITH YearlySales AS (

SELECT

YEAR(Date) AS Sales\_Year,

SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY YEAR(Date)

)

SELECT

Sales\_Year,

Total\_Sales,

LAG(Total\_Sales) OVER (ORDER BY Sales\_Year) AS Previous\_Year\_Sales,

ROUND(((Total\_Sales - LAG(Total\_Sales) OVER (ORDER BY Sales\_Year)) / LAG(Total\_Sales) OVER (ORDER BY Sales\_Year)) \* 100, 2) AS Yearly\_Growth\_Rate

FROM YearlySales

ORDER BY Sales\_Year;

*2010 2288886120.41 NULL NULL*

*2011 2448200007.3499975 2288886120.41 6.96*

*2012 2000132859.3500023 2448200007.3499975 -18.3*

This table shows the total sales for each year (2010, 2011, and 2012) and the corresponding percentage change (growth rate) compared to the previous year.

- In 2011, the total sales increased by approximately 6.96% compared to 2010.

- However, in 2012, there was a decrease in total sales by approximately 18.3% compared to 2011.

This indicates a positive sales growth from 2010 to 2011 but a decline in sales in the following year, 2012.

--yearly holiday vs non holiday sales ratio

SELECT

YEAR(Date) AS Year,

SUM(CASE WHEN Holiday\_Flag = 1 THEN Weekly\_Sales ELSE 0 END) / SUM(Weekly\_Sales) AS Holiday\_Sales\_Ratio,

1 - (SUM(CASE WHEN Holiday\_Flag = 1 THEN Weekly\_Sales ELSE 0 END) / SUM(Weekly\_Sales)) AS Non\_Holiday\_Sales\_Ratio

FROM sales

GROUP BY YEAR(Date);

*2010 0.0874768717956726 0.9125231282043273*

*2011 0.08444387141546353 0.9155561285845365*

*2012 0.049166467502542846 0.9508335324974572*

These figures represent the holiday sales ratio and non-holiday sales ratio for each year:

- In 2010, approximately 8.75% of total sales occurred during holiday weeks, while around 91.25% of sales were during non-holiday weeks.

- In 2011, the holiday sales ratio decreased slightly to about 8.44%, while the non-holiday sales ratio increased to approximately 91.56%.

- In 2012, there was a significant decrease in the holiday sales ratio to around 4.92%, with the non-holiday sales ratio increasing to about 95.08%.

This indicates the proportion of sales generated during holiday weeks compared to non-holiday weeks for each year.

1. **Exploratory Data Analysis:** I explored various aspects of the data, including total sales distribution, sales trends over time, and the impact of external factors such as holidays and temperature.

**SALES TREND and STORE**

--TOTAL SALES BY STORE

SELECT Store, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY Store;

**Sales Distribution**: The weekly sales vary significantly across different stores, with some stores having notably higher sales than others.

--Top Performing Stores: Stores such as Store 27, Store 10, and Store 2 stand out as top performers, with exceptionally high weekly sales.

-**-Challenges**: Some stores, such as Store 39 and Store 14, seem to face challenges in generating substantial weekly sales compared to others.

**--Opportunities**: Stores with relatively lower sales could explore strategies to enhance their performance and increase their weekly sales.

--**Overall Trends**: While there are fluctuations in weekly sales across stores, overall trends indicate a mix of high-performing and moderate-performing stores.

-- monthly sales trend

SELECT

DATEPART(YEAR, Date) AS Year,

DATEPART(MONTH, Date) AS Month,

SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY

DATEPART(YEAR, Date),

DATEPART(MONTH, Date)

ORDER BY

Year,

Month;

--Sales exhibit seasonal variations, with December consistently emerging as the highest sales month across all three years, likely influenced by holiday shopping.

--The summer months, particularly June and July, also demonstrate notable sales peaks, possibly due to seasonal promotions and increased consumer activity during warmer weather.

--January appears to be a consistently low-sales month, which could be attributed to reduced consumer spending post-holiday season and harsh winter weather conditions.

--yearly sales trend

SELECT

DATEPART(YEAR, Date) AS Year,

SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY

DATEPART(YEAR, Date)

ORDER BY

Year;

*2010 2288886120.41*

*2011 2448200007.3499975*

*2012 2000132859.3500023*

--2010 vs. 2011: Sales increased from 2010 to 2011, with a growth of approximately $159 million. This growth could be attributed to various factors such as economic recovery, expanded product offerings, or improved marketing strategies.

--2011 vs. 2012: Sales experienced a slight decline in 2012 compared to 2011, with a decrease of approximately $448 million. Possible factors contributing to this decline could include economic downturn, changes in consumer behaviour, or increased competition.

-- month that generated highest sales

SELECT TOP 5

YEAR(Date) AS Year,

MONTH(Date) AS Month,

SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY

YEAR(Date),

MONTH(Date)

ORDER BY

Total\_Sales DESC;

*2010 12 288760532.71999997*

*2011 12 288078102.4800002*

*2012 6 240610329.28999984*

*2012 8 236850765.67999992*

*2010 7 232580125.9800001*

--Seasonal Trends: December appears to be the month with consistently high sales across multiple years, possibly due to holiday shopping and year-end promotions.

--Yearly Variations: While December consistently ranks high, other months such as June and August in 2012 also show significant sales, indicating potential seasonal variations or one-time events influencing sales performance.

--Implications for Business: Understanding the months with the highest sales can help businesses plan marketing campaigns, promotions, and inventory management strategies to capitalize on peak sales periods.

--TOTAL SALES DURING HOLIDAY AND NON HOLIDAY

SELECT Holiday\_Flag, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY Holiday\_Flag;

--Total Sales During Non-Holiday Periods: $6,231,919,435.55

--Total Sales During Holiday Periods: $505,299,551.56

--These figures highlight the significant difference in sales between holiday and non-holiday periods, indicating the potential impact of holidays on consumer spending behaviour.

-- Top selling dates (by total sales):

SELECT top 5 Date, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY Date

ORDER BY Total\_Sales DESC;

***Date Total Sales***

*2010-12-24 80931415.60000001*

*2011-12-23 76998241.31*

*2011-11-25 66593605.25999998*

*2010-11-26 65821003.23999999*

*2010-12-17 61820799.85000001*

These dates represent peak periods of sales activity, likely corresponding to significant holidays or promotional events.

--Total sales trend over time for each store

SELECT top 5 Store, Date, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY Store, Date

ORDER BY Total\_Sales desc, Date;

***Store Date Total Sales***

*14 2010-12-24 3818686.45*

*20 2010-12-24 3766687.43*

*10 2010-12-24 3749057.69*

*4 2011-12-23 3676388.98*

*13 2010-12-24 3595903.2*

--The top-performing stores experienced peak sales on specific dates, with December 24, 2010, being a significant sales day for multiple stores.

--Store 14, Store 20, and Store 10 recorded exceptionally high sales on December 24, 2010, indicating strong holiday shopping activity.

--Store 4 saw a surge in sales on December 23, 2011, suggesting effective sales strategies or promotions during the holiday season.

--Store 13 also experienced notable sales on December 24, 2010, highlighting its performance during peak shopping periods.

--BOTTOM 10

SELECT TOP 10 Store, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

GROUP BY Store

ORDER BY Total\_Sales ASC;

*33 37160221.960000016*

*44 43293087.83999999*

*5 45475688.9*

*36 53412214.969999984*

*38 55159626.420000024*

*3 57586735.07*

*30 62716885.12000001*

*37 74202740.32000001*

*16 74252425.40000002*

*29 77141554.30999997*

--The bottom 10 stores exhibit varying levels of total sales, with Store 33 having the lowest total sales among them.

--Store 29 has the highest total sales among the bottom 10 stores, indicating relatively better performance compared to others in this group.

--Factors contributing to lower total sales could include location, market demographics, competition, and store-specific strategies.

**HOLIDAY VS NON HOLIDAY**

--Average sales per store and holiday type:

SELECT top 5 Store, Holiday\_Flag, AVG(Weekly\_Sales) AS Avg\_Weekly\_Sales

FROM sales

GROUP BY Store, Holiday\_Flag

ORDER BY Avg\_Weekly\_Sales ASC;

***Store Holiday Flag Avg Weekly Sales***

*33 0 259656.21631578956*

*33 1 262594.51900000003*

*44 1 296035.60099999997*

*44 0 303253.6227819548*

*5 0 314892.27691729314*

--The average weekly sales per store vary across different holiday types, with slight fluctuations observed between holiday and non-holiday periods.

--Stores 33 and 44 exhibit lower average weekly sales compared to Store 5, indicating potential differences in market demographics, store location, or promotional strategies.

--While there is a slight increase in average weekly sales during holidays for Stores 33 and 44, the difference is not significant compared to Store 5.

--TOTAL SALES DURING HOLIDAY AND NON HOLIDAY by store

SELECT top 5

Store,

SUM(CASE WHEN Holiday\_Flag = 1 THEN Weekly\_Sales ELSE 0 END) AS Total\_Sales\_Holiday,

SUM(CASE WHEN Holiday\_Flag = 0 THEN Weekly\_Sales ELSE 0 END) AS Total\_Sales\_Non\_Holiday

FROM

sales

GROUP BY

Store;

***Store Holiday Non\_Holiday***

*23 14625422.94 184125194.90999994*

*29 6069578.889999999 71071975.42*

*9 5889508.21 71899 710.77999999*

*15 7064060.18 82069623.73999998*

*3 4378110.5 53208624.57*

**Observations:**

--Total sales vary significantly between holiday and non-holiday periods for each store.

--Some stores demonstrate higher sales during holidays, while others perform better during non-holidays.

--Understanding sales patterns during different periods can help businesses tailor their marketing and sales strategies accordingly.

**CPI**

--cpi for each quarter

SELECT

YEAR(Date) AS Year,

DATEPART(QUARTER, Date) AS Quarter,

AVG(CPI) AS Average\_CPI

FROM sales

GROUP BY

YEAR(Date),

DATEPART(QUARTER, Date)

ORDER BY

Year,

Quarter;

**Observations**:

--CPI shows a general increasing trend over the years, with slight fluctuations between quarters.

--Q4 of each year tends to have the highest CPI values, indicating potential seasonal or year-end economic factors.

-- Impact of CPI on sales:

SELECT top 5 CPI, AVG(Weekly\_Sales) AS Avg\_Weekly\_Sales

FROM sales

GROUP BY CPI

ORDER BY CPI DESC;

**Observations**:

--The analysis suggests a correlation between CPI values and average weekly sales, with higher CPI values generally corresponding to higher average weekly sales.

--Fluctuations in CPI values may influence consumer purchasing power and behaviour, impacting sales performance.

**Implications**:

--Understanding the relationship between CPI and sales can help businesses anticipate and respond to changes in economic conditions.

--Monitoring CPI trends enables businesses to make informed decisions regarding pricing strategies, inventory management, and promotional activities.

-- Total sales for each store during weeks with low CPI (< X150), sorted by store:

SELECT top 5 Store, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

WHERE CPI < 150

GROUP BY Store

ORDER BY Total\_Sales DESC;

*4 299543953.38*

*13 286517703.80000013*

*10 271617713.8899999*

*27 253855916.88000003*

*19 206634862.1*

--Despite challenging economic conditions characterized by low CPI values,

--the top-performing Walmart stores have demonstrated resilience and strong sales performance.

--By leveraging insights from this analysis, Walmart can refine its strategies, optimize resource allocation, and maintain its competitive edge in the retail market.

**TEMPERATURE**

-- Total sales distribution across different temperature ranges

SELECT Temperature\_Range, SUM(Weekly\_Sales) AS Total\_Sales

FROM (

SELECT CASE

WHEN Temperature < 50 THEN 'Cold'

WHEN Temperature >= 50 AND Temperature < 70 THEN 'Moderate'

ELSE 'Hot'

END AS Temperature\_Range,

Weekly\_Sales

FROM sales

) AS SalesByTemperatureRange

GROUP BY Temperature\_Range;

***Temp Total Sales***

*Hot 2288529494.9499984*

*Moderate 2415104472.4600053*

*Cold 2033585019.699996*

--The analysis reveals that moderate temperature ranges contribute the highest total sales, followed by hot and cold temperature ranges.

--This suggests a potential correlation between moderate temperatures and increased consumer activity and spending compared to extreme hot or cold temperatures.

--Retail strategies and promotions tailored to moderate temperature conditions may be more effective in driving sales during various seasons.

**FUEL PRICE**

-- Total sales for each store during weeks with high fuel prices (> 4), sorted by store:

SELECT TOP 5 Store, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

WHERE Fuel\_Price > 4

GROUP BY Store

ORDER BY Total\_Sales DESC;

*Fuel Price Total Sales*

*27 51045520.53000001*

*10 48761665.79000001*

*28 43596113.910000004*

*19 41654036.85000001*

*24 39596415.67*

--Impact of High Fuel Prices: Despite high fuel prices, several stores have managed to maintain respectable sales figures, indicating resilience and effective sales strategies.

--Store Performance: Stores 10 and 27 stand out with notably high sales figures during weeks with high fuel prices, suggesting successful management practices or strong customer loyalty.

--Sales Strategies: Further analysis is needed to understand the specific sales strategies or promotions implemented by these stores to mitigate the impact of high fuel prices on consumer spending.

**UNEMPLOYMENT**

-- Total sales for each store during weeks with high unemployment (> 10), sorted by store:

SELECT Store, SUM(Weekly\_Sales) AS Total\_Sales

FROM sales

WHERE Unemployment > 10

GROUP BY Store

ORDER BY Store;

***Store Total Sales***

*12 144287230.14999998*

*28 189263680.57999986*

*29 32767838.680000003*

*33 2299155.24*

*34 64964338.48*

*38 55159626.420000024*

*43 40802240.249999985*

The analysis of total sales during weeks with high unemployment (> 10) reveals varying performance across stores, with Store 28 standing out with approximately $189.26 million in sales, followed by Store 12 with $144.29 million. Meanwhile, Store 33 demonstrates the lowest sales performance at approximately $2.30 million, emphasizing the need for targeted strategies to navigate economic challenges.

**Thinking About Stakeholders:**

1. **Primary Stakeholder Interest:** The primary stakeholder is focused on maximizing sales revenue and profitability while maintaining customer satisfaction. They seek insights into sales performance drivers, market trends, and opportunities for growth.
2. **Information Provided:** I provide detailed analysis and visualizations on key metrics such as total sales, sales trends, and the impact of external factors on sales performance.
3. **Defining KPIs:** KPIs are defined to measure performance and track progress towards business objectives. These KPIs include total sales, sales growth rate, holiday sales ratio, average temperature, and unemployment rate.

**Conclusion and Key Takeaways:** Through my analysis of Walmart sales data, I identified actionable insights and recommendations to optimize performance and drive growth. By leveraging data-driven strategies and monitoring key metrics, Walmart can enhance sales effectiveness, improve resource allocation, and better serve its customers' needs.

RELATED IMAGES

