Walmart

Sales Analysis

Group 6

Team Introduction

Haley Henson
BAIM

Carly Sprenger
Marketing

Utkarsh Singh
BAIM

Kevin Peng
Marketing

Walmart

Contents

- Research Background
- Research Purpose
- Data Breakdown
- Data Reduction and Explanation
- Data Analysis
- Recommendation and Conclusion

Research Background

One of the leading retail stores in the US, Walmart, **would like to predict the sales and demand accurately.** There are certain events and holidays which impact sales on each day. There are sales data available for 45 Walmart stores. The business is facing a challenge due to unforeseen demands and runs out of stock some times, due to the inappropriate machine learning algorithm. An ideal ML algorithm will predict demand accurately and ingest factors like economic conditions including CPI, Unemployment Index, etc.

Walmart runs several promotional markdown events throughout the year. These markdowns precede prominent holidays, the four largest of all, which are the Super Bowl, Labour Day, Thanksgiving, and Christmas. The weeks including these holidays are weighted five times higher in the evaluation than non-holiday weeks. Part of the challenge presented by this competition is modeling the effects of markdowns on these holiday weeks in the absence of complete/ideal historical data. Historical sales data for 45 Walmart stores located in different regions are available.

Walmart

Research Purpose

- The objective behind our research is to test the hypothesis that our variables have an effect on **Walmart's Sales** and determine demand and stock during high sale seasons and holidays
- At the end of our analysis, we will arrive at a conclusion that will allow us to **recommend a marketing strategy** that will boost Walmarts sales by analyzing Weekly_Sales, Holiday_Flag, Temperature, Fuel_Price, CPI, Unemployment
- We will be able to provide insight on **changes of demand and sales** through our analysis of different seasonal and market changes

Walmart

Data Breakdown

The historical data covers sales from 2010-02-05 to 2012-11-01, Walmart_Store_sales and includes the following fields:

- Store - the store number
- Date - the week of sales
- Weekly_Sales - sales for the given store
- Holiday_Flag - whether the week is a special holiday week 1 – Holiday week 0 – Non-holiday week
- Temperature - Temperature on the day of sale
- Fuel_Price - Cost of fuel in the region
- CPI – Prevailing consumer price index
- Unemployment - Prevailing unemployment rate

Data Reduction and Explanation

- We have chosen to reduce our dataset, showing a selection of Walmart stores, from *45 down to 20 stores*

- The stores we include in the dataset currently are the first 20 stores on the dataset.

- We decided to reduce our dataset because of the large amount of stores to account for, and could lead to *overplotting*
  - *We will be focusing on a subsample of 20 Walmart Stores since these 20 stores are a representative sample of the dataset. Our reasoning for this is focusing only on 20 will enable us to avoid overplotting and this subset could provide insights into overall trend and pattern.*

Walmart

Initial Regression Analysis

(To see how Weekly_Sales are affected by independent variables)

```
Call:
lm(formula = log_transformed_weeklysales ~ Holiday_Flag + Temperature +

    Fuel_Price + CPI + Unemployment, data = selected_data)

Residuals:
     Min       1Q   Median       3Q      Max
-1.28411 -0.50259  0.08544  0.47888  1.50720

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   14.3579429  0.1282496 111.953  < 2e-16 ***
Holiday_Flag   0.1236365  0.0431974   2.862  0.00424 **
Temperature    0.0033077  0.0006168   5.362 8.88e-08 ***
Fuel_Price     0.0150287  0.0247504   0.607  0.54376
CPI           -0.0041653  0.0003053 -13.642  < 2e-16 ***
Unemployment  -0.0002028  0.0070958  -0.029  0.97720
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5807 on 2854 degrees of freedom
Multiple R-squared:  0.07812    Adjusted R-squared:  0.0765
```

- For the purpose of seeing percentage effect on weekly sales, it has been converted to log . Also, weekly sales had right skewed distribution to get better interpretation
- From the regression analysis what we discovered was that Holiday_Flag, Temperature, and CPI were statistically significant at 0.05 significant level
- However, Fuel_Price and Unemployment were not. Their p value were much higher
- In order to get a better model with significant predictors we performed regression analysis again.

Walmart

Final Regression Analysis

```
Call:
lm(formula = log_transformed_weeklysales ~ Holiday_Flag + Temperature +

    CPI + CPI_Un, data = Df)

Residuals:
    Min      1Q  Median      3Q     Max
-1.2734 -0.5113  0.1015  0.4978  1.4906

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.432e+01  6.054e-02 236.608  < 2e-16 ***
Holiday_Flag  1.190e-01  4.302e-02   2.766 0.005703 **
Temperature   3.318e-03  6.031e-04   5.501 4.11e-08 ***
CPI          -5.074e-03  3.578e-04 -14.183  < 2e-16 ***
CPI_Un        1.845e-04  4.785e-05   3.856 0.000118 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5791 on 2855 degrees of freedom
Multiple R-squared:  0.08277,   Adjusted R-squared:  0.08149
F-statistic: 64.41 on 4 and 2855 DF,  p-value: < 2.2e-16
```

- To get better model, we removed the insignificant variables Fuel_Price and Unemployment
- Created an interaction variable between CPI and Unemployment
- In this final model all the regressors are statistically significant at 5% level since their p value < alpha (0.05)
- Also, Adjusted R square increased to 0.08149

Walmart

# Temperature and Weekly Sales

Distribution of Total Weekly Sales by Temperature Category
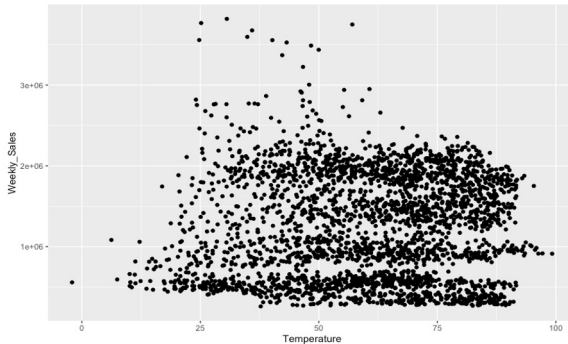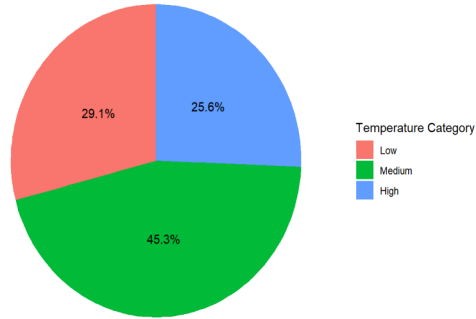
- Weekly sales are the highest when the temperature is in the medium range
- Most purchases are made between 50-75 degrees
- Although most holidays are correlated with hot or cold temperatures rather than mellow, knowing that holidays increase spending, we believe the holiday season temperatures are in the mid-range during the months of October and November leading up to big spending dates.

Walmart

Holiday_Flag and Weekly Sales

	Holiday_Flag	avg_weekly_sales
1	0	1041256
2	1	1122888

Effect of Holiday_Flag on Average Weekly Sales
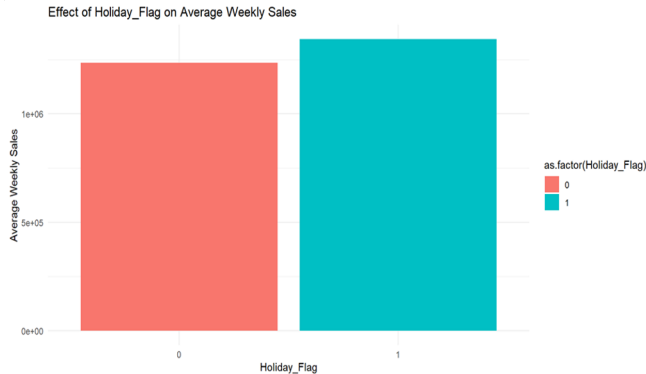
- Small increase in average weekly sales when there is a holiday week vs. None
- Customers are more likely to purchase more during those weeks when there is a holiday
- *Gifts for families and friends, groceries for holiday dinners and get togethers*

Walmart

CPI and Weekly Sales

Scatter Plot with Regression Line
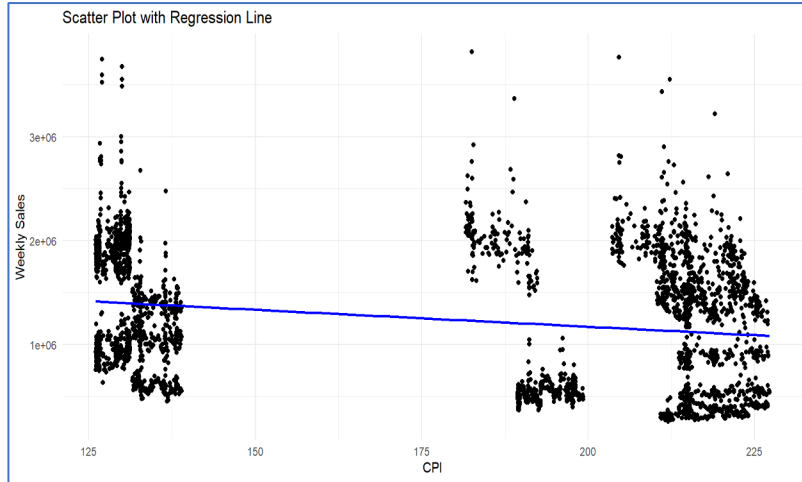
- There is a negative trend between CPI and Weekly Sales
- As we increase CPI by one-unit weekly sales decrease by 0.507%.
- There is a gap in this plot because we only focused 20 out of the 45 stores given to us in the data set.
- As we increase our CPI the weekly sales tends to decrease as observed from the trend in scatterplot.
- Since CPI refers to Consumer Price Index, if CPI increase it indicates inflation.
- This implies when prices of goods tend to increase in the store people might be more cautious and might tend to buy less.
- This is the potential reason when CPI increases, Weekly Sales tends to show declining trend.

Walmart

Weekly Sales and Macroeconomic trend

Scatterplot of Weekly Sales and CPI Unemployment
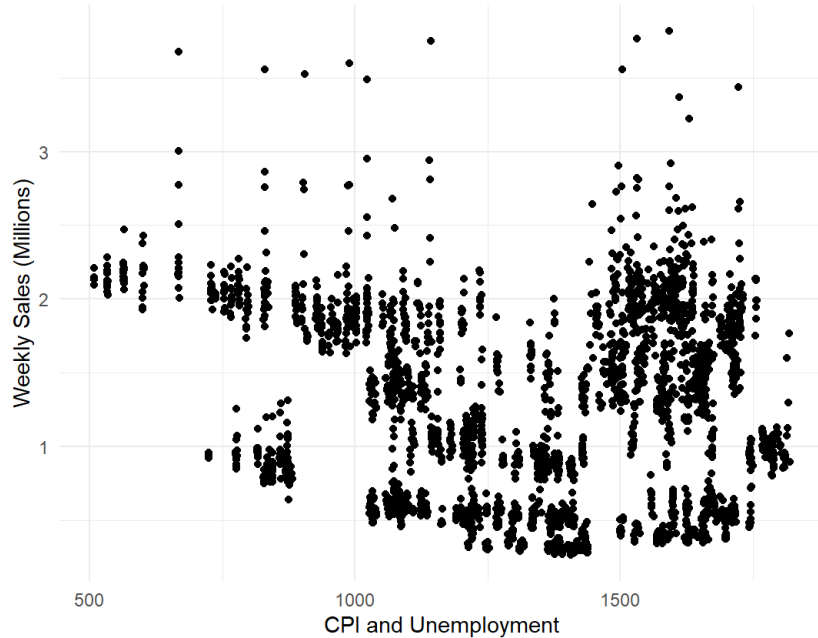
- Sales data points are highly concentrated towards the higher end of interaction variable which means that purchase volume and revenue tends to rise during periods of inflation and unemployment simultaneously.
- This suggests economic factors increases demand despite the consumer purchasing power.
- To find profit it's important to understand what kinds of products are driving sales in these periods aligning operations with consumer demand shift tied to macroeconomic trend.
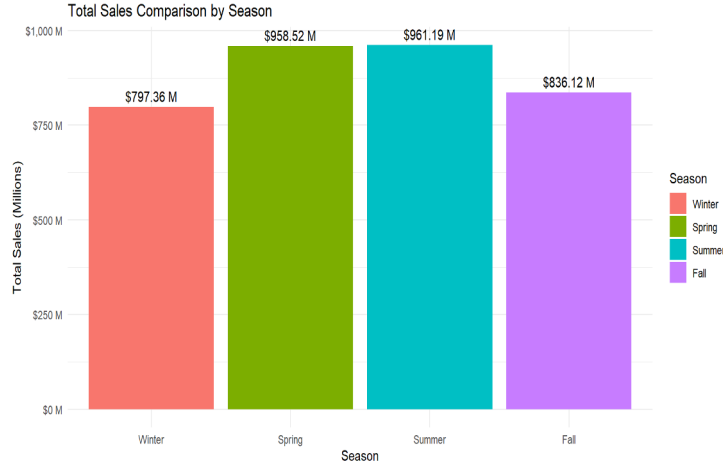
Store and Weekly Sales

Store vs Sales
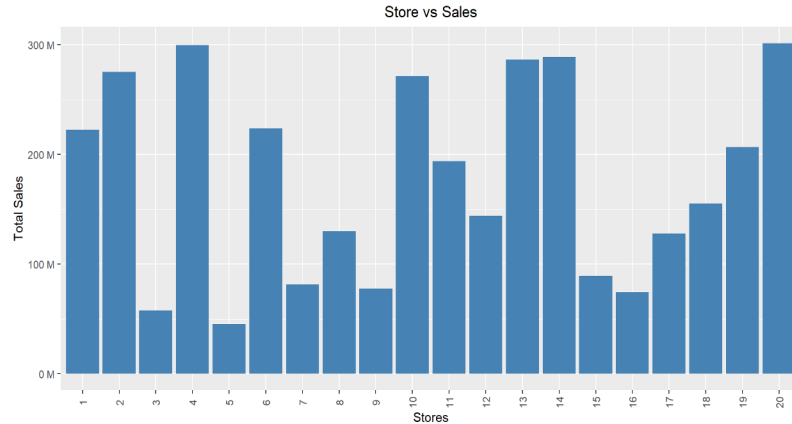
	Store	Weekly_Sales
20	20	301397792
4	4	299543953
14	14	288999911
13	13	286517704
2	2	275382441

Bar plot that shows how Weekly Sales (In Millions) varies with different Store Numbers (From 1-20). Shows top 5 stores with maximum total weekly sales. Since, these 5 stores show outstanding sales performance, we would recommend replicating merchandising strategies from these 5 stores to other Walmart Stores.

Walmart

Seasonality and Weekly Sales

Total Sales Comparison by Season

- The most profitable seasons are Summer and Spring
- Summer vacation shopping in the Spring, "Back to school" shopping in the Summer
- Aligns with temperature and weekly sales analysis; 50-75 degree weather is when Walmart is more profitable

Walmart

Overall Findings and Analysis

- We found that is surprising that the winter season is the least profitable despite there being multiple major holidays (Thanksgiving, Christmas)

- When unemployment rate is high, sales decrease. People are facing financial uncertainty and tend to cut on spending . People are more focused on essential goods.

- Seasons in general make the sales go up but spring and summer are observed to show highest weekly sales due pleasant weather , vacation season and the need to go out and shop

Walmart

Recommendations

- Since Winter is the least profitable season, scale back on marketing campaigns in December, January, and February, and capitalize on marketing campaigns in more profitable seasons ahead of holidays at that time

- Offer marketing campaigns slightly before profitable sales to optimize peak season

- Stay up to date on current news including when unemployment are spiking. These factors effect purchasing trends and vary randomly based on current events

- Since weekly sales depicted to go up during inflation and unemployment period, pricing of products should be optimized to balance customer affordability and making promotional campaign focused towards high selling products during these times of economic trend.

Walmart

Recommendations In Action

Knowing that the Spring and Summer are the most profitable seasons, we believe that Walmart should have a kickoff to summer or Spring-Cleaning Campaign in the Spring and a back-to-school campaign toward the end of Summer.

Kickoff to Summer: Campaign centered around floaties, drinks, sunscreen, barbeque materials, bathing suits etc.

Spring Cleaning: Discounts or coupons around cleaning supplies or organizational items as shoppers get ready for "Kickoff to Summer"

Back To School: Campaign around school supplies, backpacks, families, and sack lunches and getting ready to head back to school.

Walmart

Conclusion

We have found that all our variables effect the overall weekly sales in some way or another. We believe that if Walmart integrates our recommendations into their current or future marketing campaigns, they will capture more profits from consumers and be more successful when determining supply and demand throughout the year.

Walmart