Marketing Analytics Team Project

Group 10
Aashna Rungta, Sneha Batchu, Sony Kumari, Udisha Madnani

Agenda

- Overview of the Company
- Situation and Objective of Analysis
- Insights from Exploratory Data Analysis
- Overview of Analysis
- What was learned from the Analysis?
- Conclusions and Recommendations

Overview of the Company

Walmart Inc (Walmart) is a retailer that operates grocery stores, supermarkets, hypermarkets, department and discount stores, and neighborhood markets. As of January 31, 2022, Walmart has 10,593 stores and clubs in 24 countries, operating under 48 different names. Walmart is the world's largest company by revenue, with US$548.743 billion, according to the Fortune Global 500 list in 2020. It is also the largest private employer in the world with 2.2 million employees.

Situation & Objective of Analysis

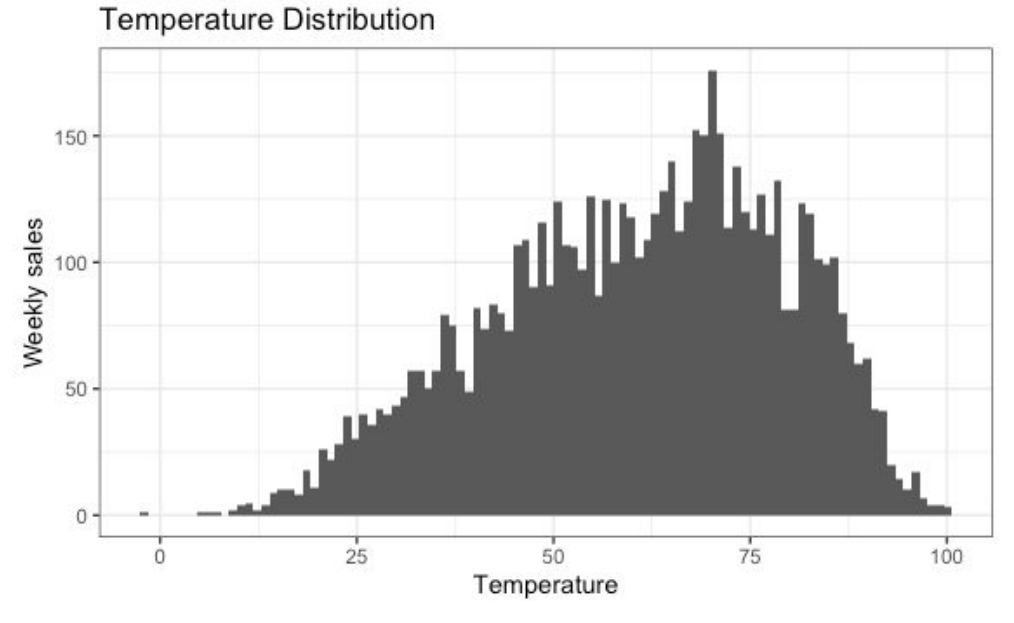
We want to study the effect of various factors on the weekly sales at Walmart, so that we can cater towards increasing revenue. Therefore, we want to explore the data with few features such as CPI, temperature, unemployment, fuel price, effect of holidays, etc, focusing on the weekly sale numbers.

Insights from Exploratory Data Analysis

The weekly sales are different in different stores, depending on the location.

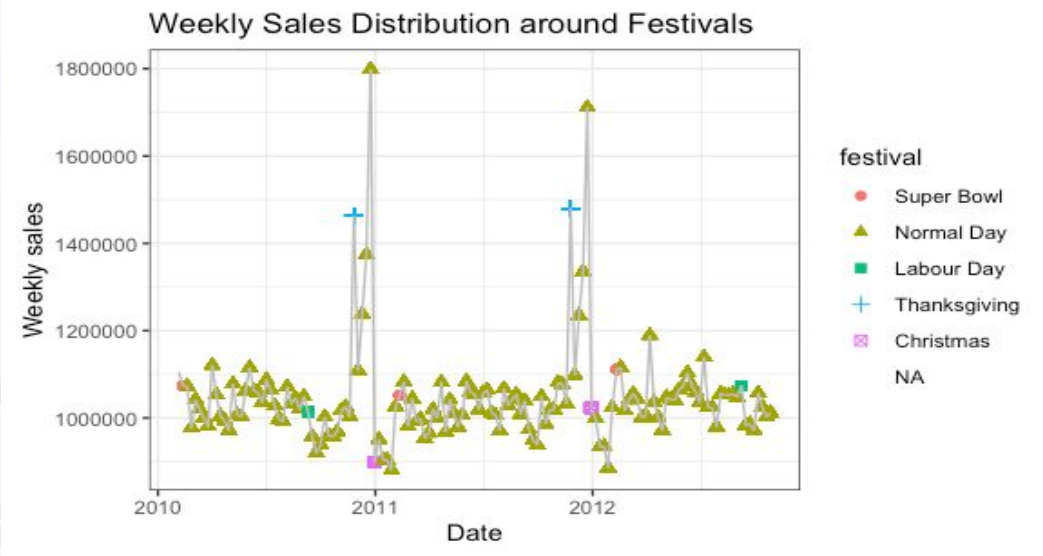
Insights from Exploratory Data Analysis

More sales were observed when the temperature was around 75 degrees.



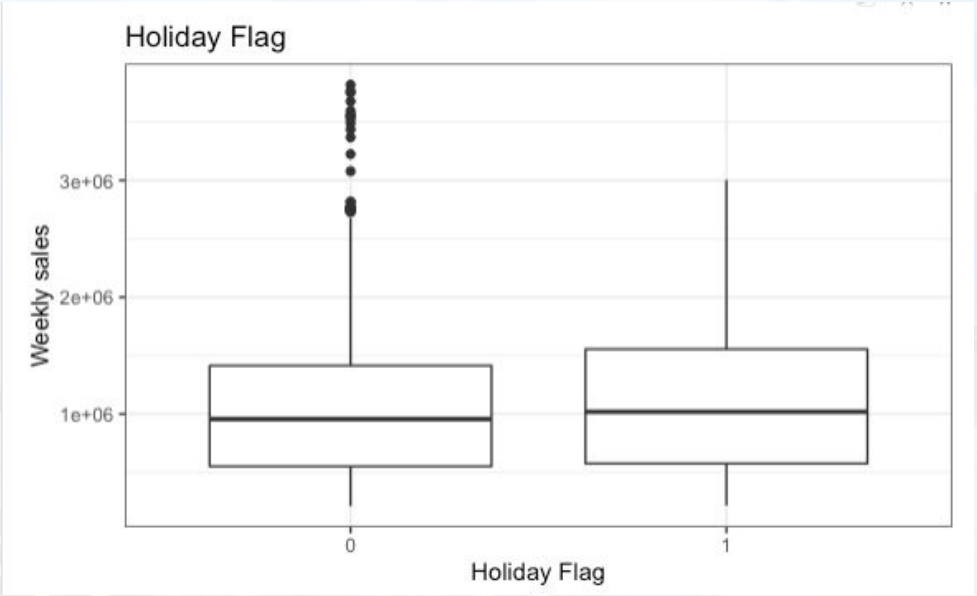
Insights from Exploratory Data Analysis

Spikes in sales were observed in weeks close to holidays. Our understanding is, this is because sales start a couple of weeks earlier and people want to shop before the holiday.



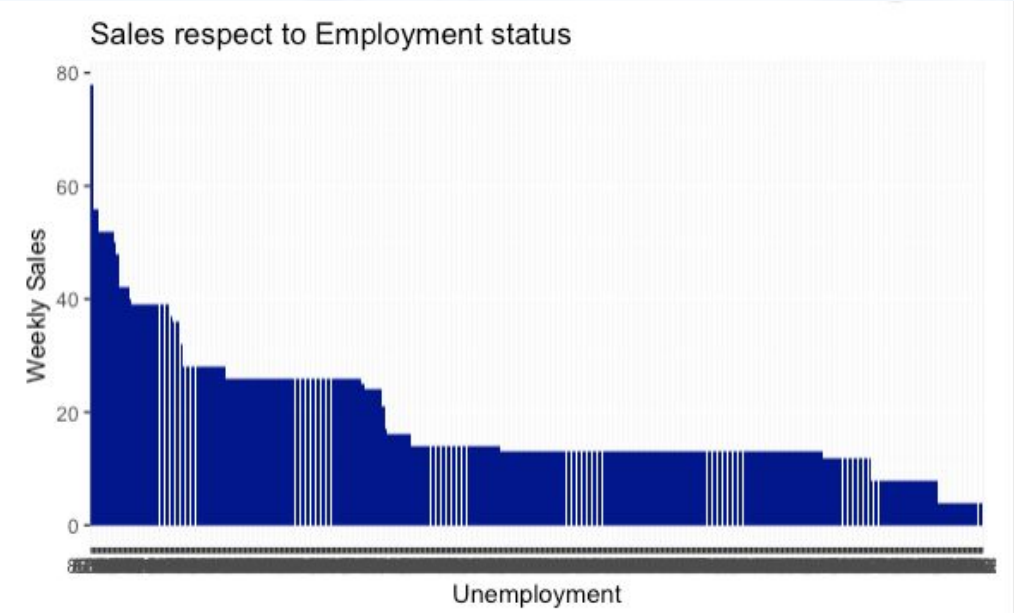
Insights from Exploratory Data Analysis

Extreme values were observed for non-holiday weeks. On average, there isn't a huge difference between the average sale for holidays as compared to non-holiday weeks.



Insights from Exploratory Data Analysis

The employed section of the society would prefer shopping from high end markets. Walmart is targeting mass over class.



Overview of Analysis

We have used Multi-linear Predictive Regression Analysis, in order to remove possible omitted variable bias, and use all the variables in the dataset and come up with a more accurate model. Using the multi - linear regression model was appropriate because we wanted to analyze the causal effect of different significant variables(Temperature, festival, CPI, Unemployment, and Fuel Prices) on average weekly sales of Walmart. Also, to magnify the significance of few variables, we have added interaction terms in the regression model. Linear regression model was found as a suitable technique to study impact of various numeric independent variables on a numeric dependent variable, which is our variable of interest.

What was learned from the analysis?

```
Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)         6648859.09  285789.84  23.265  < 2e-16 ***
Store                -19107.76     535.53 -35.680  < 2e-16 ***
Temperature             -63.23     378.54  -0.167   0.8673
festivalNormal Day    -44903.22   44981.75  -0.998   0.3182
festivalLabour Day    -19440.73   70148.87  -0.277   0.7817
festivalThanksgiving  389482.68   68709.82   5.669 1.50e-08 ***
festivalChristmas    -124767.01   68532.58  -1.821   0.0687 .
CPI                   -33508.77    1750.49 -19.142  < 2e-16 ***
Unemployment         -386993.07   18206.12 -21.256  < 2e-16 ***
Fuel_Price           -521081.11   65004.01  -8.016 1.29e-15 ***
CPI:Unemployment        2606.00     127.18  20.490  < 2e-16 ***
CPI:Fuel_Price          3412.62     377.87   9.031  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 503600 on 6378 degrees of freedom
  (45 observations deleted due to missingness)
Multiple R-squared:  0.2042,    Adjusted R-squared:  0.2028
F-statistic: 148.8 on 11 and 6378 DF,  p-value: < 2.2e-16
```

- RMSE was higher in training set than test set, showing that test set performed better
- RMSE and MAE were both quite high relative to the average weekly sales, indicating that the model is not the best predictor of weekly sales
- R2 was 0.2, which is also quite low - again showing the model may not be the most accurate.
- There is a huge upswing in sales during Thanksgiving, but not other holidays
- Unemployment has a negative correlation with weekly sales
- Economic factors don’t have a large effect on sales

Conclusions and Recommendations

Conclusion:

- During festivities, we always observe that there is huge rush at Walmart and our research from the data also backs the same hypothesis - Thanksgiving actually has extreme high weekly sales
- Our research shows that the sales reaches a peak just a week before Christmas
- Festivals like Labor day and Super bowl does not cause extremely high sales
- Unemployment rate negatively affects the weekly sales - the lower the unemployment rate, the higher will be the weekly sales
- Certain variables have little or no effect at all on the weekly sales, like - CPI and fuel price

Recommendation:

There has been a significant drop in the average weekly sales from 2011 to 2012 as per the data analysis, thus, as per our primary and secondary research, we come to the following recommendations for Walmart -

- Cater to supply chain disruption issues like, transportation delays, etc and kick in automation for activities like inventory management, etc
- Stores should restock shelves fast enough in order to avoid undue shrinkage and out-of-stocks
- Digital services should be integrated more with the physical stores

Thank You!