**Walmart Sales Analysis**

Walmart is an American founded retail store, dedicated to helping people save money and live better. It has thousands of stores in its home country, the United States of America, and also in various different countries in the world.

Walmart tends to approximately 255 million customers weekly both in their physical stores and on their e-commerce websites. They move more than 100 billion items per year (Walmart, 2025).

**Project Overview**

Contents

[Introduction 1](#_Toc190083666)

[Configuration 2](#_Toc190083667)

[Data Preprocessing 2](#_Toc190083668)

[Feature Engineering 2](#_Toc190083669)

[Data Analysis 2](#_Toc190083670)

[Results 2](#_Toc190083671)

# 

# Introduction

The purpose of this analysis was to study how different factors influence sales for various different Walmart Stores for a continuous period of time – from 2010 to 2012. From the analysis I was able to gain insights into the trend of sales and how these factors affect them. These factors include holiday flag – the presence or absence of a holiday, temperature, fuel price, consumer price index and unemployment rate.

The data set collected from 45 Walmart stores contain the following variables:

* Store: store number
* Date: Sales week start date
* Weekly\_Sales: Sales
* Holiday\_Flag: Mark on the presence or absence of holiday
* Temperature: Air temperature in the region
* Fuel\_Price: Fuel cost in the region
* CPI: Consumer Price Index
* Unemployment: Unemployment rate

# Configuration

Libraries: All the data manipulation and visualization were performed using Microsoft Excel 2016.

Dataset: The data set used included variables such as: store, date, Weekly\_Sales, Holiday\_Flag, Temperature, Fuel\_Price, CPI, and Unemployment.

# Data Preprocessing

The data was cleaned before starting the analysis by removing duplicate values, filling in missing values and converting data to other data types (data type conversions) and sorting. This was to ensure that the data was uniform before analyzing it and to help understand the structure of the data.

I then explored the data to find out how many rows and columns the data set contained.

# Feature Engineering

Date Features: Extracted additional information from date by extracting the date and month so as to provide insights into sales trends over yearly quarters for the period of the 3 years.

# Data Analysis

Different charts were used to visualize the data and draw meaning from them.

Bar charts and pie chart were used to visualize categorical variables of numerical data.

Correlation table was used to measure the strength and direction of relationships between variables.

Line charts were used to show trends between different variables over a continuous period of time.

# Results

Conclusion

Holiday distribution: The holidays take up only a small percentage (7%) of the total sales days.

Conclusion

The total sales for non-holiday days are higher compared to total sales of all holiday days. This suggests that sales performance tend to decline significantly on holidays. This may mean that a lot of people stay home and do not visit stores during holidays.

Top 10 performing stores.

Bottom 10 Performing Stores

Conclusion

Of the three years, 2011 was the best performing year in terms of total sales as a total of $2,455,437,098.73.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Correlation Heat map | | | |  |
|  | *Weekly\_Sales* | *Temperature* | *Fuel\_Price* | *CPI* | *Unemployment* |
| Weekly\_Sales | 1 |  |  |  |  |
| Temperature | -0.063810013 | 1 |  |  |  |
| Fuel\_Price | 0.009463786 | 0.144981806 | 1 |  |  |
| CPI | -0.072634162 | 0.176887676 | -0.17064 | 1 |  |
| Unemployment | -0.10617609 | 0.101157857 | -0.03468 | -0.30202 | 1 |

Conclusion

Weekly sales have a week negative correlation with temperature, CPI, and unemployment. This means that as these factors or variables increase, weekly sales tend to decrease slightly.

Weekly sales have a very week positive correlation with fuel price. This means that as fuel prices increase, weekly sales tend to increase slightly.