



## Table Of Contents:

1. Problem Statement
2. Project Objective
3. Data Description
4. Data Pre-processing steps and inspiration
5. Choosing the algorithm for the project
6. Assumptions
7. Inferences
8. Future possibilities of the project
9. Conclusion
10. References

## **PROBLEM STATEMENT:**

A retail store - "WALMART" that has multiple outlets across the country are facing issues in managing the inventory - to match the demand with respect to supply.

## **PROJECT OBJECTIVE:**

- The primary objective of this project is to develop a robust forecasting model for weekly sales data from 45 Walmart stores across the country, covering the period from 2010 to 2012.
- Analyzing historical sales data to identify trends and seasonal patterns, choosing the best model for forecasting and accurately predict weekly sales for the next 12 weeks.

## **DATA DESCRIPTION:**

- |                        |  |
|------------------------|--|
| 1. <b>Store</b>        | - Store number                           |
| 2. <b>Date</b>         | - Week of Sales                          |
| 3. <b>Weekly_Sales</b> | - Sales for the given store in that week |
| 4. <b>Holiday_Flag</b> | - If it is a holiday week                |
| 5. <b>Temperature</b>  | - Temperature on the day of the sale     |
| 6. <b>Fuel_Price</b>   | - Cost of the fuel in the region         |
| 7. <b>CPI</b>          | - Consumer Price Index                   |
| 8. <b>Unemployment</b> | - Unemployment Rate                      |

## **DATA PRE-PROCESSING STEPS AND INSPIRATION:**

- **Checking Data Types:** Verify the data types of all columns, ensuring the Date column is converted from an object to a datetime format to facilitate better visualization and time series analysis.
- **Handling Missing Values:** Identify and address any null values and duplicate entries in the dataset.
- **Removing Outliers:** Utilize the Interquartile Range (IQR) method to detect and eliminate outliers from the data.
- **Correlation Analysis:** Conduct a correlation analysis to explore the relationships among all columns in the dataset.
- **Visualization:** Use visualizations and statistical methods to examine the relationships between columns, revealing additional patterns and dependencies within the data

## CHOOSING THE ALGORITHM FOR THE PROJECT:

I have chosen **SARIMAX model** for forecasting Walmart's weekly sales.

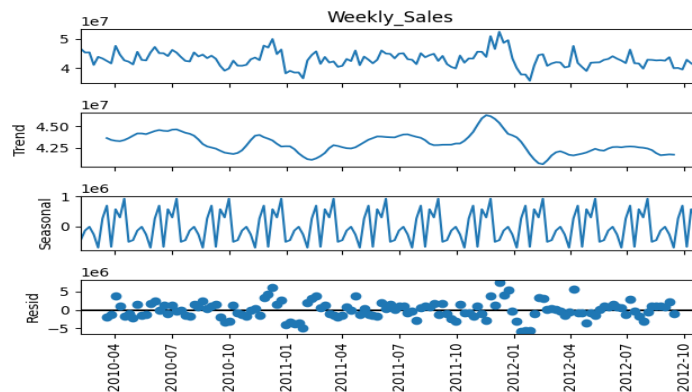
### *Reasons for Choosing SARIMAX:*

- SARIMAX is specifically designed to capture seasonality, making it well-suited for this type of data, since Walmart's sales data is weekly, it likely exhibits seasonal patterns.
- SARIMA allows for the integration of both autoregressive (AR) and moving average (MA) components, this flexibility helps model complex time series behaviors. It can handle different levels of noise and irregularities in the data, which is often present in retail sale

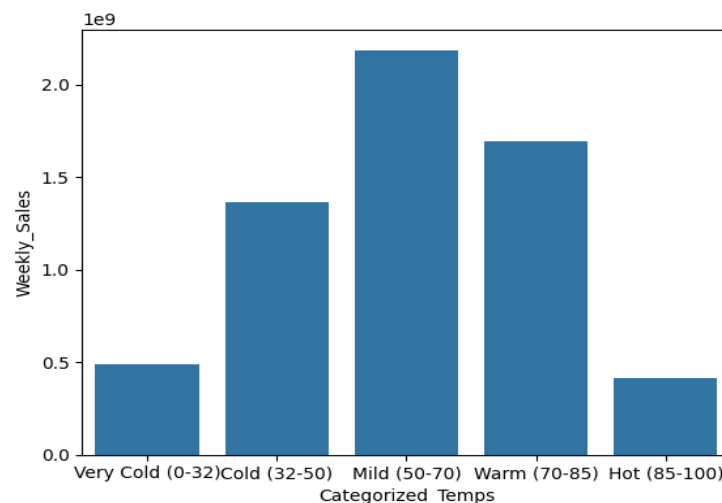
## ASSUMPTIONS:

No Assumptions were made.

## INFERENCES:



A notable **spike** in the trend is observed between **October 2011** and **January 2012**. This **increase** may be attributed to a significant **decline** in the Consumer Price Index (CPI) and inflation rate beginning in early October. Additionally, this period coincides with major sales events such as **Black Friday**, and the **Christmas holidays**, which likely contributed to the trend.



From the above graph, it reveals that

- sales performance is **notably stronger** during **mild to warm** temperature ranges.
- sales **decline** during **very cold and hot** temperatures,

Indicating that **consumer shopping behavior is influenced by weather conditions**

### *Consumer Price Index effect on Weekly sales:*

- The negative correlation of -0.81 between CPI and weekly sales suggests that higher prices are associated with lower sales.

### *Most affected store due to high unemployment rate*

	Store	Weekly_Sales
	6	14 2.761276e+08
	20	34 1.382498e+08
	8	18 1.375821e+08
	21	35 1.315207e+08
	3	10 1.309794e+08

### *Top Performing store*

	Store	Weekly_Sales
	19	20 2.800237e+08
	13	14 2.761276e+08
	1	2 2.687221e+08
	12	13 2.682025e+08
	9	10 2.556789e+08

## **FUTURE POSSIBILITIES OF THE PROJECT:**

We can implement machine learning models (e.g., Random Forest, XGBoost) to improve forecasting accuracy by incorporating additional features like customer demographics, local events, and economic indicators.

We can add on some more external datasets (e.g., weather patterns, economic data, competitor pricing) to refine forecasts and provide a more comprehensive view of sales drivers.

Analyze the effectiveness of promotions and marketing campaigns, helping to optimize future strategies and resource allocation.

We can use customer transaction data to segment sales forecasts by demographics or shopping behavior, providing tailored strategies for different customer groups.

We can develop an interactive dashboard that allow stakeholders to explore forecasts and insights visually, facilitating better decision-making.

## **CONCLUSION:**

In this project, the main objective was to analyze Walmart's sales performance

From my finding,

- Consumer shopping behavior is influenced by weather conditions, with sales peaking during mild to warm temperatures and declining in extreme heat or cold. Walmart could offer discounts on its online portal during extreme weather and promote pickup orders to enhance customer convenience.
- There is a negative correlation between CPI and weekly\_sales. This trend suggests that consumers are becoming increasingly price-sensitive, which may require Walmart to adapt its pricing strategies to maintain sales growth.

## **REFERENCES:**

1. NewDigits.org
2. YouTube
3. Investopedia
4. USA – CPI-Inflation-calculator
5. Intellipaat Live Class notes and code.

