# Walmart

# Sales Analysis

Sample **Full Data Science** Project by Marat Karlin

# Introduction

## The objectives of this project were:

- Analyse historical data from different Walmart stores. Because it can help us understand data and improve the company's operating system.

- Find and understand key features influencing sales. These insights are a vital part of understanding data in general and will be used in the future as a tool to maximise sales.

- Create predictive model capable of forecasting weekly sales. Accurate sales forecasts are essential for making informed business decisions and improving the supply chain efficiency. By predicting weekly sales, Walmart can better manage inventory levels, reduce costs and understand how company potentially will be doing in the future.

# Exploratory Data Anailysis

**EDA is an important step in understanding and summarising data, and the following features were analysed in detail:**

- **Fuel Price.** What is the price of fuel in the following week.

- **Temperature.** What is average temperature in the following week.

- **Holiday Flag.** Is this a holiday week or not.

- **Unemployment.** What is unemployment rate in the following week.

- **Store.** Which store is taken into account.

These features were chosen because they are the most influential on our data (according to the importance score of the predictive model) and these features showed the most interesting and informative correlations with the weekly sales ( target feature ).

# Fuel Price


Fuel Price and Weekly Sales Plot
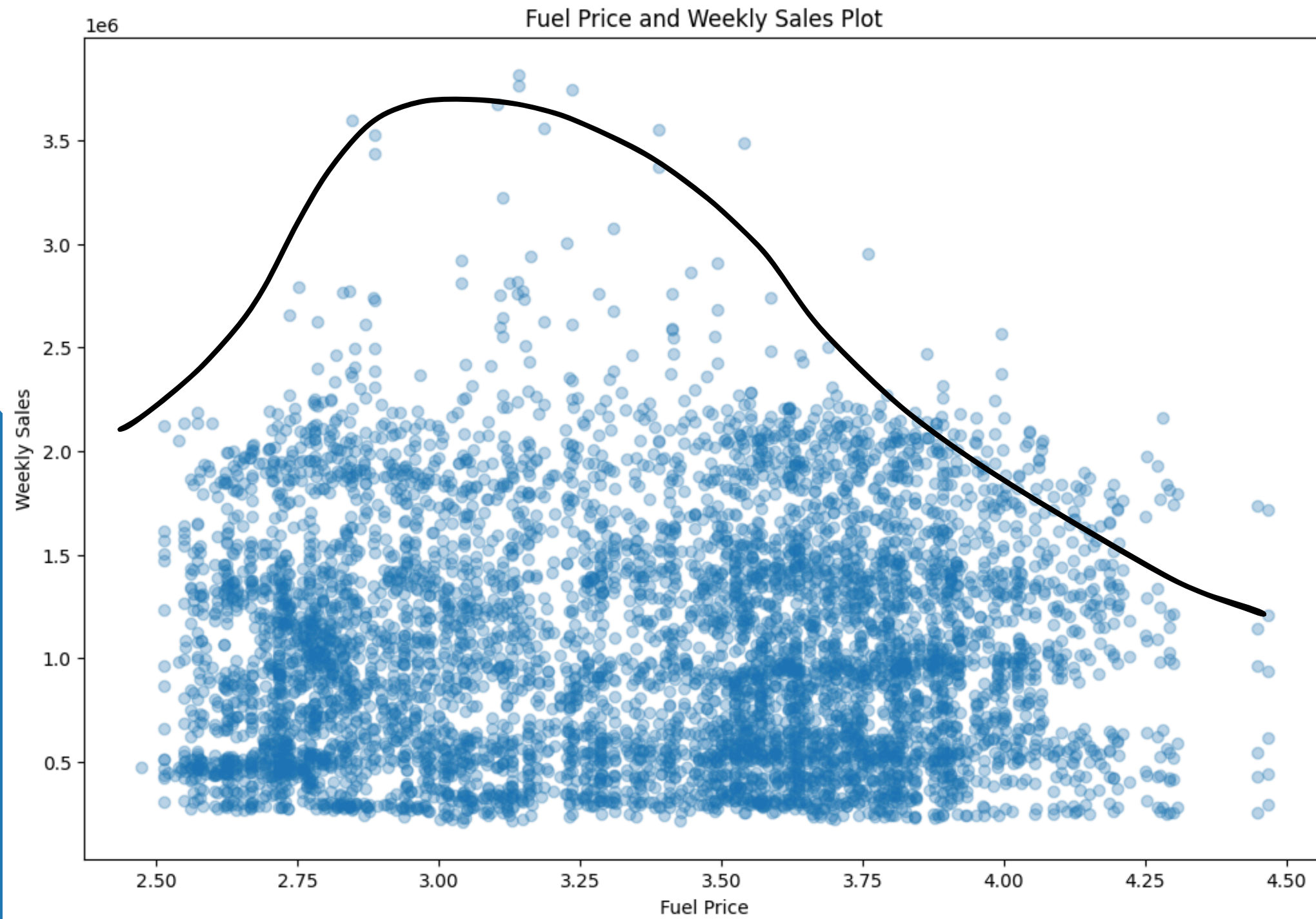
This scatter plot roughly shows us how fuel price correlates with sales. We clearly can see higher density of data points between $2.75 and $3.60 ( X-axis ). When fuel price fall into this range we may expect higher sales, as less will be spent on transportation, so Walmart should focus on more aggressive marketing campaign and as it has more sales and reduced costs on transportation during low fuel price time, the company can reinvest these savings into further improving the supply chain efficiency.

# Temperature

This bar plot confirms that people tend to spend more during winter season, as it is a time when people are buying, for example, winter clothes which are more expensive than in any other season.

During the winter season, Walmart can expect higher sales, which can be reinvested in promotions and selling more winter products during low temperatures.

# Holiday Flag



Average sales in holiday week or not
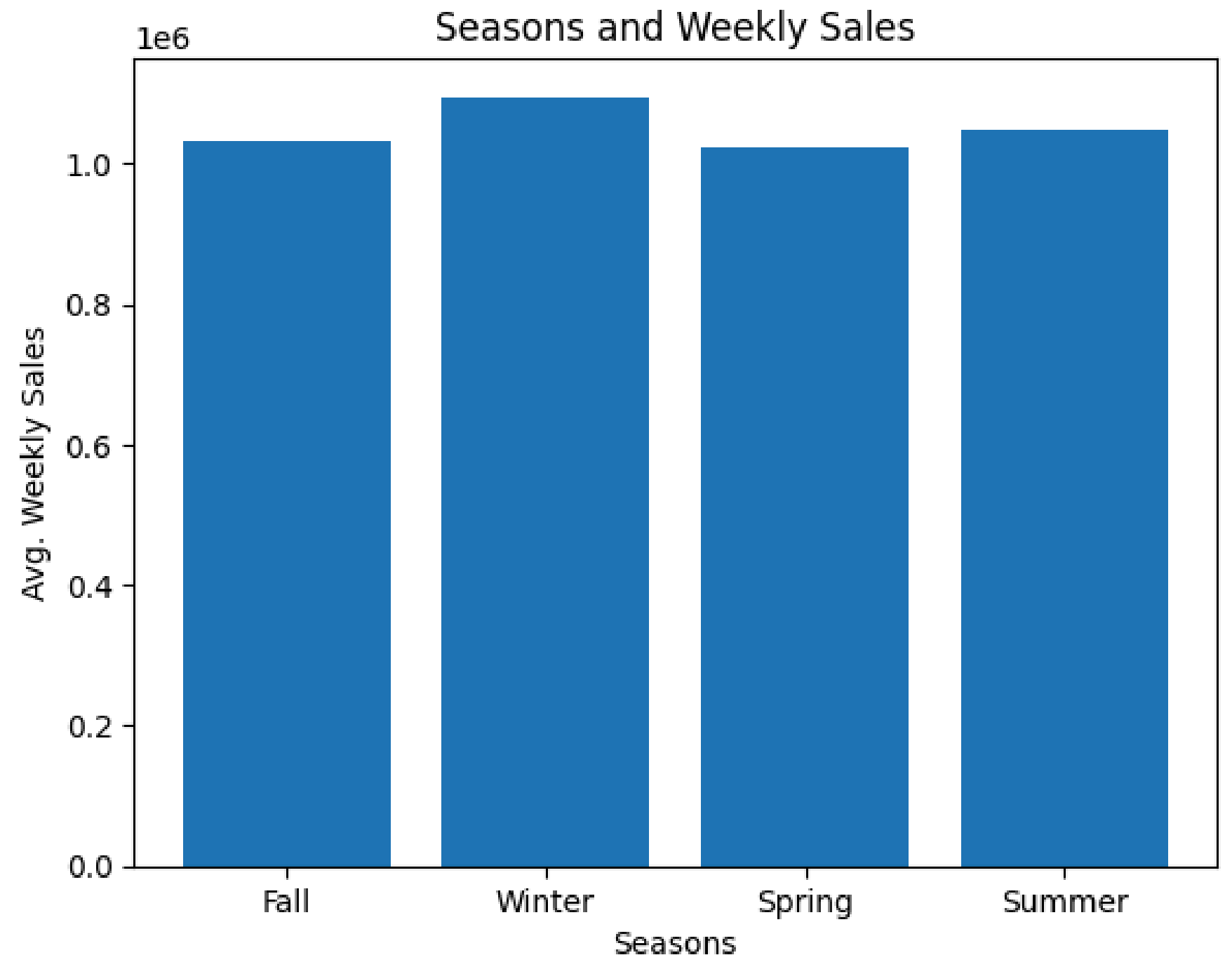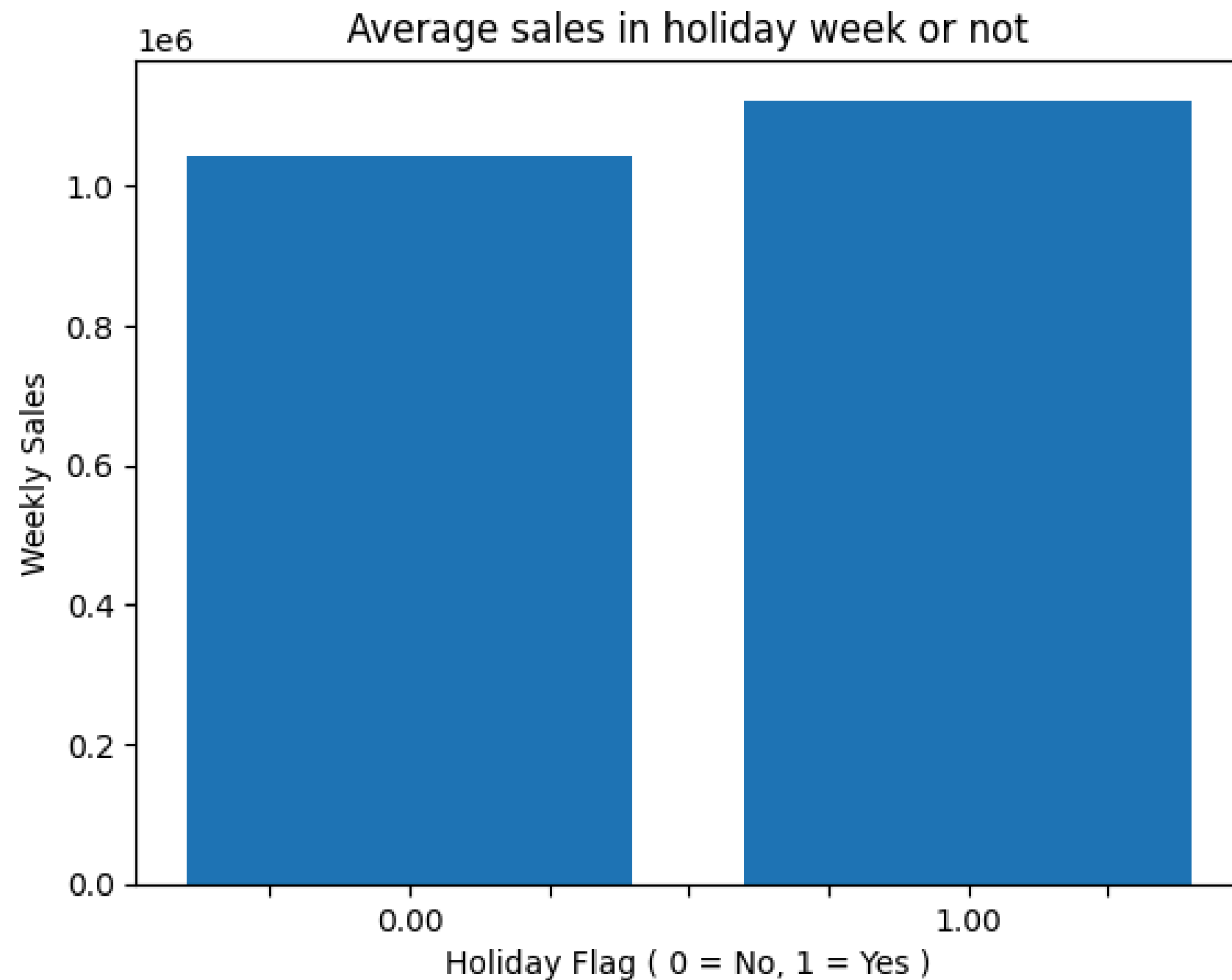
The finding on this bar chart is pretty self explanatory, people are spending more during holidays in order to meet their needs and wants ( food, presents and etc. ).

During the holidays, the company can offer more deals to customers, ensure that high-demand holiday items are well stocked and promoted, and outside of the holidays, Walmart can focus on other aspects that affect sales, as it's the time when shoppers aren't spending as much on average.

# Unemployment

As we might expect, when the unemployment rate is higher, fewer people will spend a lot in retail stores like Walmart, and this issue is more complicated to deal with, but Walmart successfully keeps its prices low, ensuring that sales are consistent in periods of high unemployment. In addition, I could suggest that Walmart can increase the number of temporary job opportunities during peak seasons such as the holiday season.


Unemployment and Weekly Sales Plot
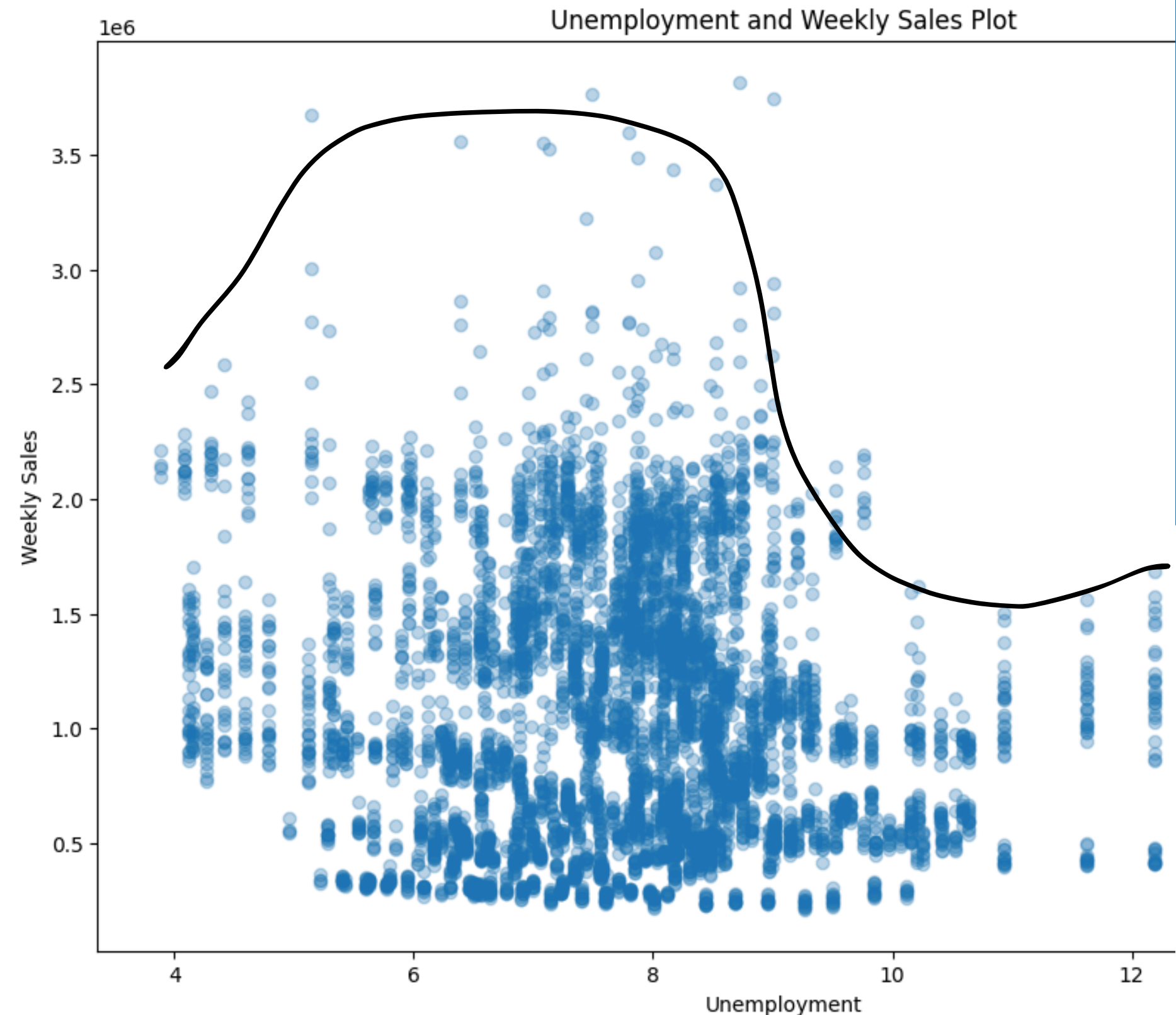
# Store

As we can see, some stores perform much better than others, which can be influenced by a number of factors such as location, demographics and the products sold in the store.

Walmart can compare different stores and determine what is actually making one store outperform another and address this issue to improve sales in specific stores and ultimately improve the company's overall sales.

## Top Performing Stores

| Store # | Total Weekly Sales |
|---------|--------------------|
| 20 | $301.397.800 |
| 4 | $299.544.000 |
| 14 | $288.999.900 |

## Worst Performing Stores

| Store # | Total Weekly Sales |
|---------|--------------------|
| 33 | $37.160.220 |
| 44 | $43.293.090 |
| 5 | $45.475.690 |

# Predictive Model

Several models were tested and the one that performed best was the **Random Forest regressor**, with an R-squared of 97%, a mean error of 111367.73 and an accuracy of approximately 96.43%.

This model is used to predict Weekly Sales based on these features: **Store, Holiday Flag, Temperature, Fuel Price, CPI, Unemployment , month and year.**

# Output Can Be Used

**For Inventory Management,** Walmart can use sales predictions to ensure that each store has the right amount of inventory, reducing overstock and stockouts.

**For Financial Planning and Budgeting**, the company may use sales forecasts to allocate budgets more effectively across different stores and departments.

In general, these predictions will help business to make more data-driven decisions, operate company more efficiently and reduce excessive costs.

# Example Strategy

- **Model Predicted high sales week for some specific shop.**

- **We look for potential features that might affect the prediction,** for example, we know there are holidays next week, the unemployment rate is low, the store we are considering is near the expensive neighbourhood.

- We inform other parts of the business and decide what can be done to maximise revenue, and at the same time we can take the model's predictions into account and use them to plan budgets and costs more efficiently.
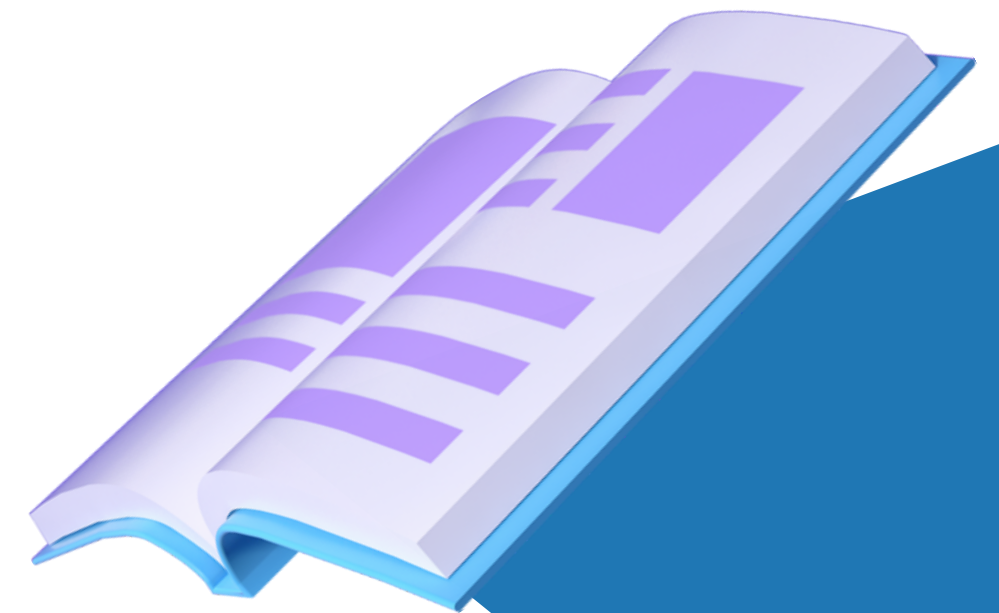
# Conclusions

**This Full Data Science project successfully has achieved its objectives:** analysis of the relationships between the features was made, and it gives a better understanding of what and how influences the weekly sales of the Walmart, a predictive machine learning model has been created that is capable of predicting weekly sales with a relatively high accuracy based on new unseen data.

In the future, the project may be continued and conducted again with new features, data, problems and ideas, as well as further improvement of the model and new relevant understanding of the data.

# Thank You !

by Marat Karlin, Data Scientist