**Capstone Project :–** **Walmart Sales Prediction**

**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. Motivation and Reasons For Choosing the Algorithm

7. Assumptions

8. Model Evaluation and Techniques

9. Inferences from the Same

10. Future Possibilities of the Project

11. Conclusion

12. References

**Problem Statement**

A retail store that has multiple outlets across the country are facing issues in managing the inventory - to match the demand with respect to supply. You are a data scientist, who has to come up with useful insights using the data and make prediction models to forecast the sales for X number of months/years.

**Project Objective**

The objective of this project is to: -

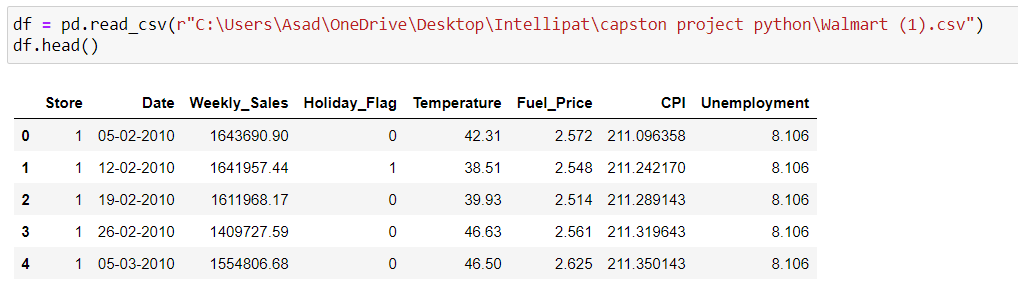
1. Get insights that can be used by each of the stores to improve in various areas.
2. And forecast the sales for each store for the next 12 weeks.

**Data Description**

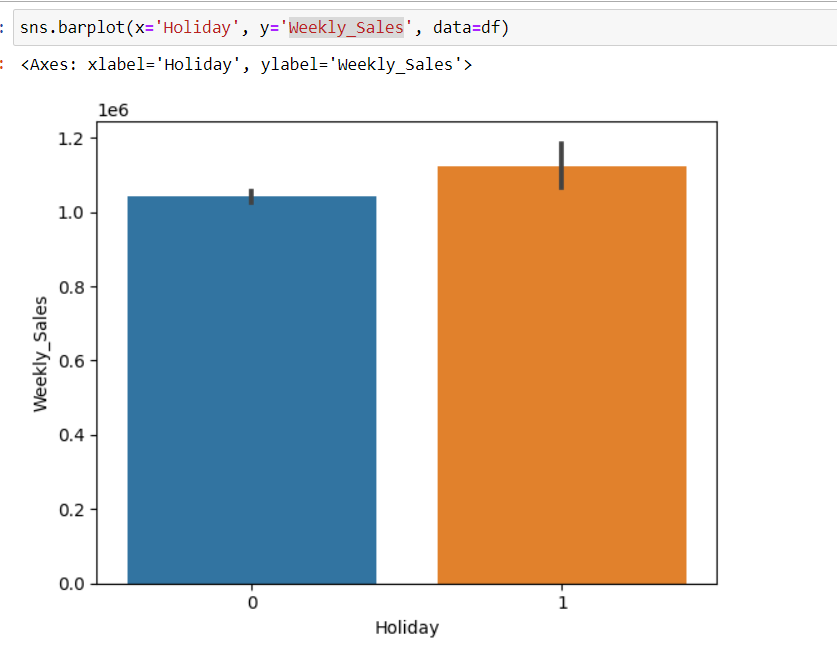
The dataset available is of Walmart sales data from which I had to predict the future sales of 12 weeks i.e., 4 months’ time

**Insights from the data: -**

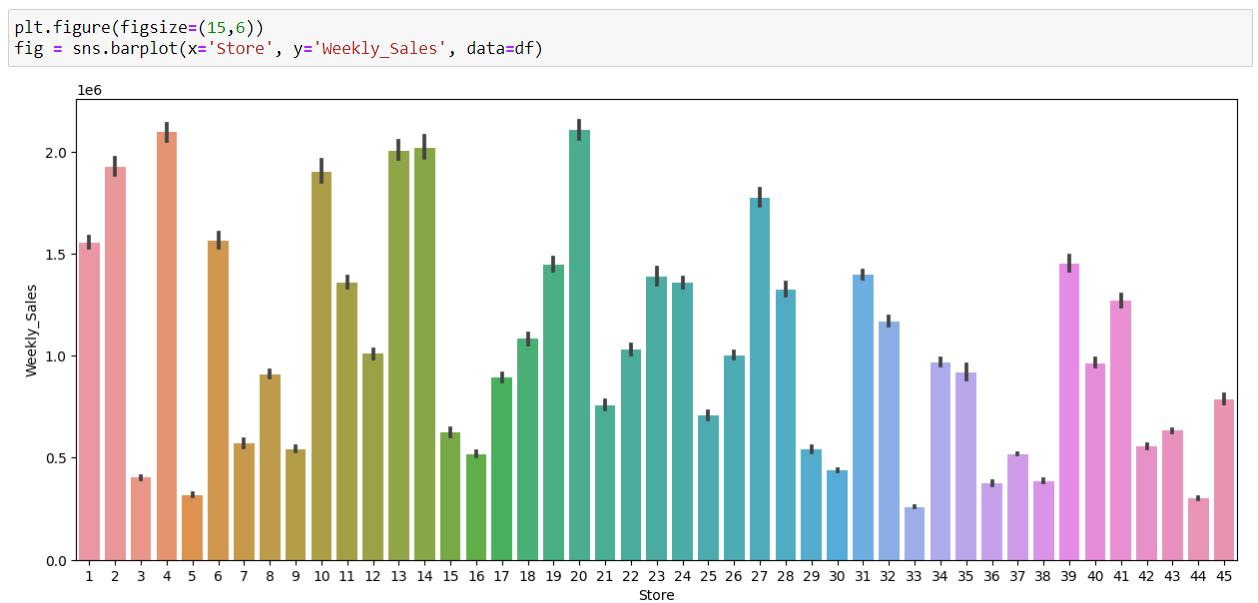
1. Data Contents.

****

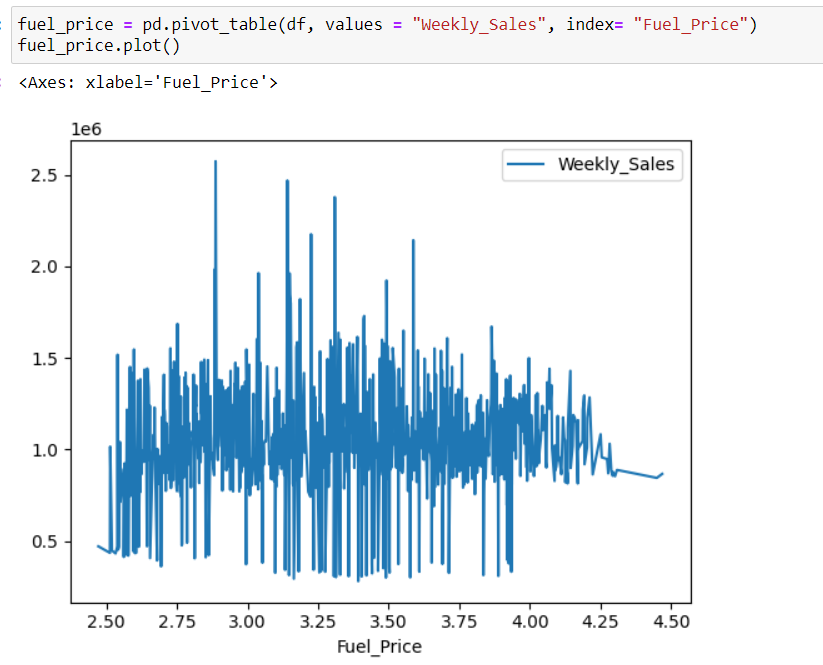
1. Sales on Holidays.

****

1. Weekly sales of each store.



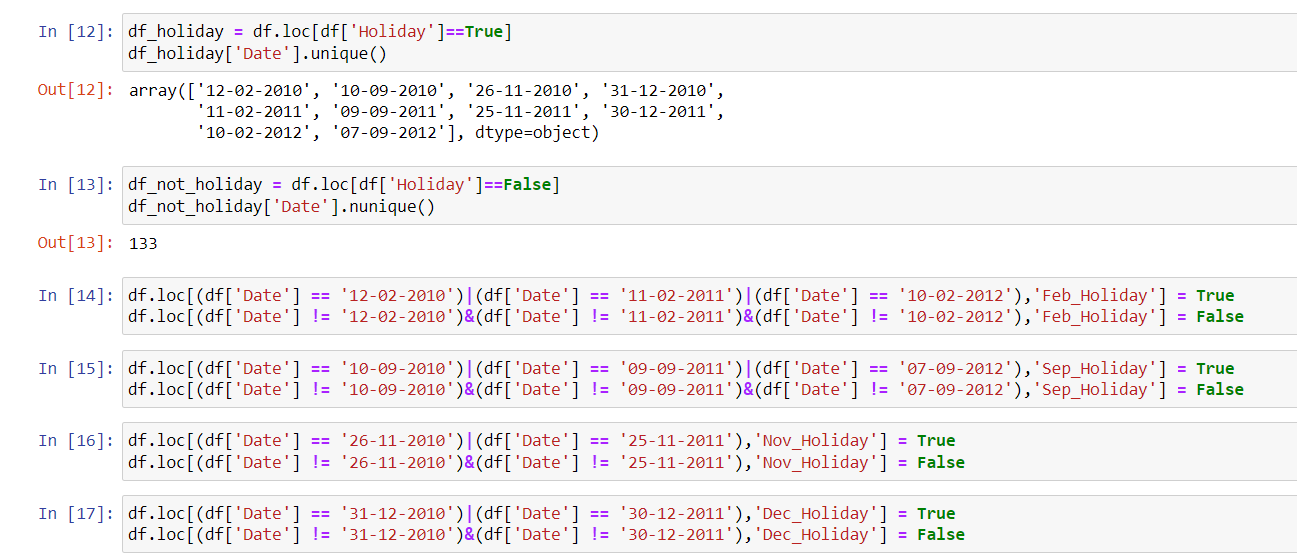
1. Sales with respect to Fuel Price.



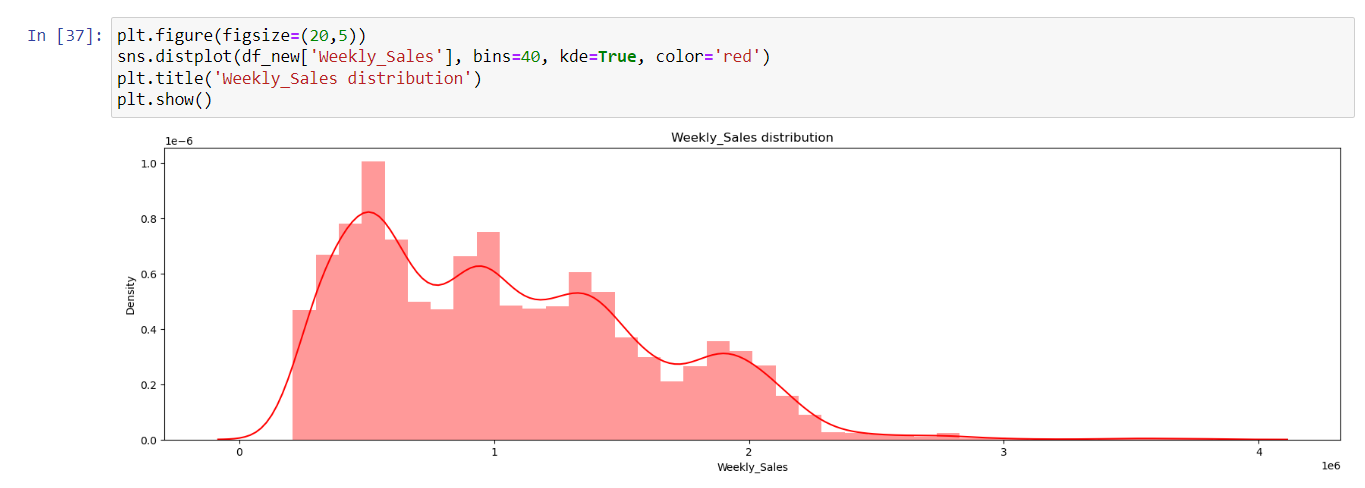
**Data Preprocessing**

The preprocessing of the data included the following steps:

1. First step: - to see the common dates of holidays and divide the accordingly for better understanding



1. Data Visualisation and Analysis: - try to understand the data and present in simple form



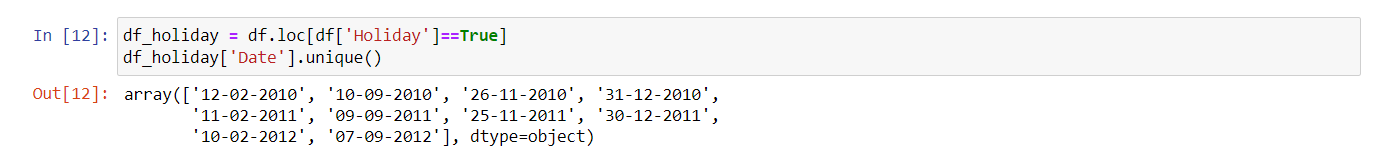
**Choosing the Prophet Algorithm For the Project**

I have chosen the Prophet algorithm for this project for the following reasons:

1. The model’s decisions are easy to interpret
2. It works best with time series that have strong seasonal effects and several seasons of historical data.
3. It is automatic and fast. It saves time for manual time series analysis and decomposition.
4. It produces reliable and accurate models.

**Assumptions**

In this project I assumed that the data has seasonal data, I came to this conclusion because of the following reason



Above we can see that the holidays are on February, September, November, December

**Model Evaluation and Technique**

The following techniques and steps were involved in the evaluation of the model

## Converting Time Series Data into Trend and Seasonality

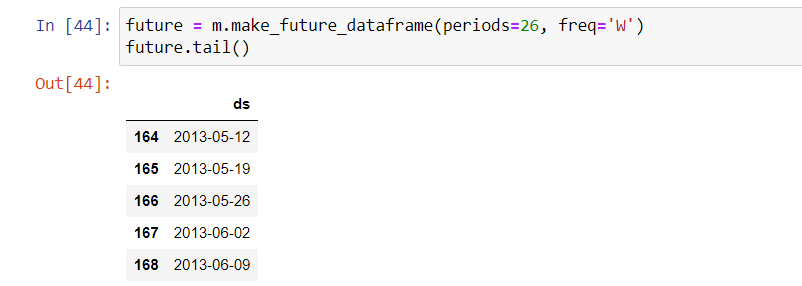
## Visualize the weekly sales by year

1. Fitting of the Prophet Model
2. Predict future sales
3. Visualize the forecasted sales

**Inferences from the Project**

In the Inference Phase, we forecast the future values at some time intervals.

Here we have a test CSV file containing date and hour columns. The demand column is missing since we have to predict the demand.



**Limitations**

1. It may not work well for all kinds of forecasting problems.
2. This is a time-series model so it doesn’t work well for other predictions
3. This model is only available in Python and R.
4. Limited plot styling
5. Doesn’t work with other popular machine learning libraries like TensorFlow, Pytorch, etc.

**Conclusion**

1. Prophet ease-of-use made it a very good baseline model when your timeseries easily breakdowns into simple temporal components. However, if your signal is noisy, fine-tuning the model’s performance can be a hassle.
2. The forecasts are pretty decent but, in some cases, certain parameters have to be tweaked compared to the default setting, which is easily done.

**References**

1. <https://www.kaggle.com/code/yehyachali/of-walmart-sales-forecasting#Model-2:-SARIMA>
2. <https://www.youtube.com/watch?v=2vF2xTUXJwM&t=283s>