**Walmart Project**

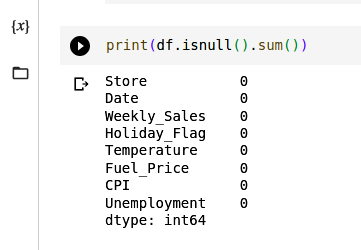
**Problem statement:** A retail store that has multiple outlets across the country is facing issues in managing the inventory - to match the demand with respect to supply.

**Project Objective:** The goal of this machine learning project is to forecast the sales for each store for the next 12 weeks. To achieve this, several Time series modelling techniques are explored, and the SARIMAX model yields the best Forecasting result.

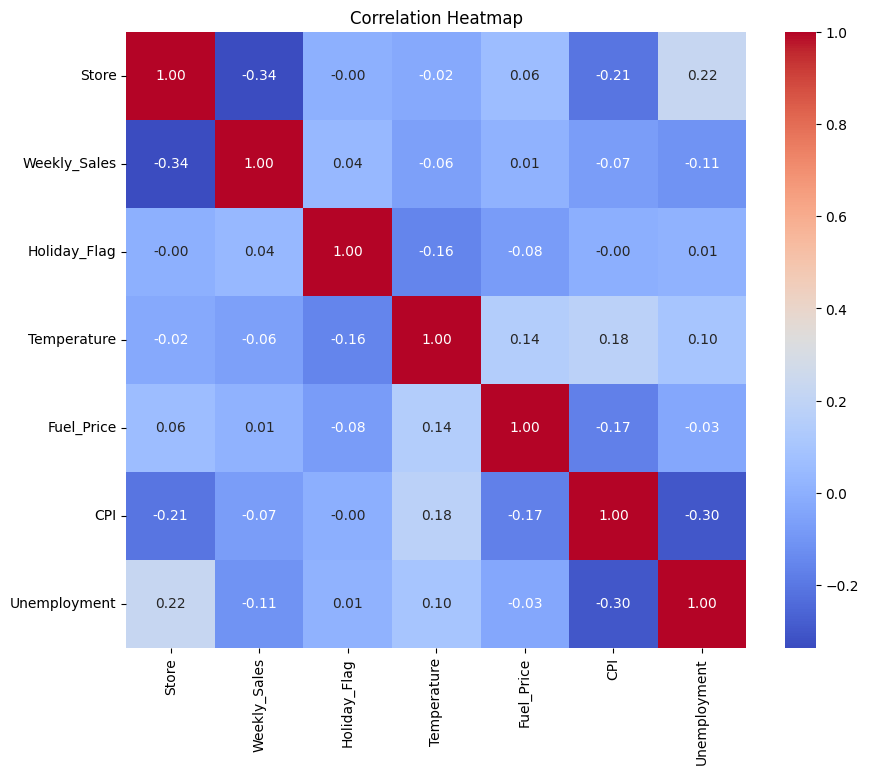
**Data Description:** The walmart.csv file contains Weekly sales information and other factors affecting sales for over 45 stores.

**Data Processing:**

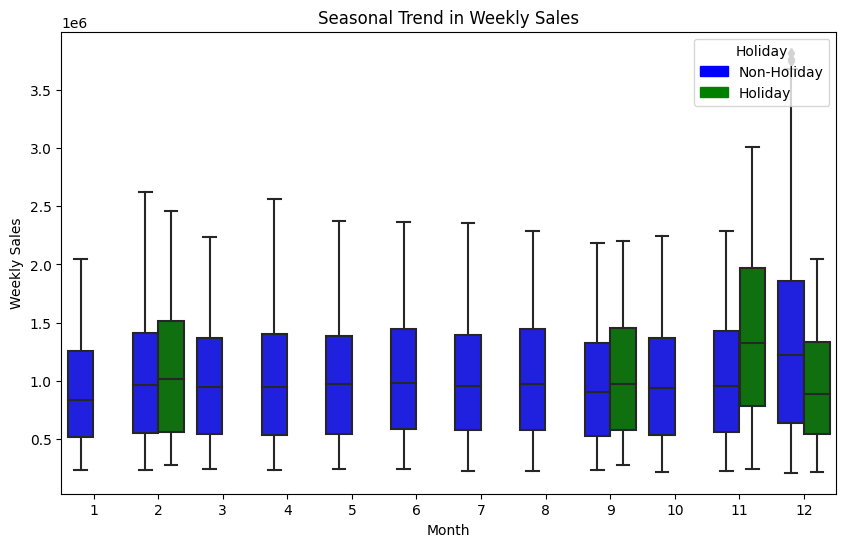
1.Checking for null values:



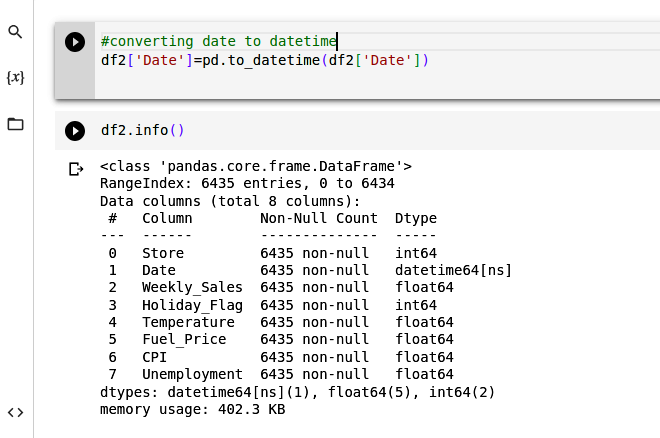
2.Correlation between factors:



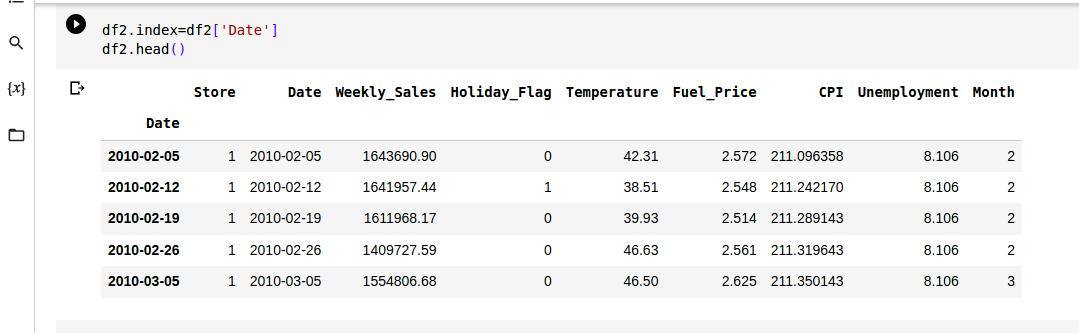
3.Seasonal Trend in Weekly sales:



4.Converting Date into Datetime for Time series model:

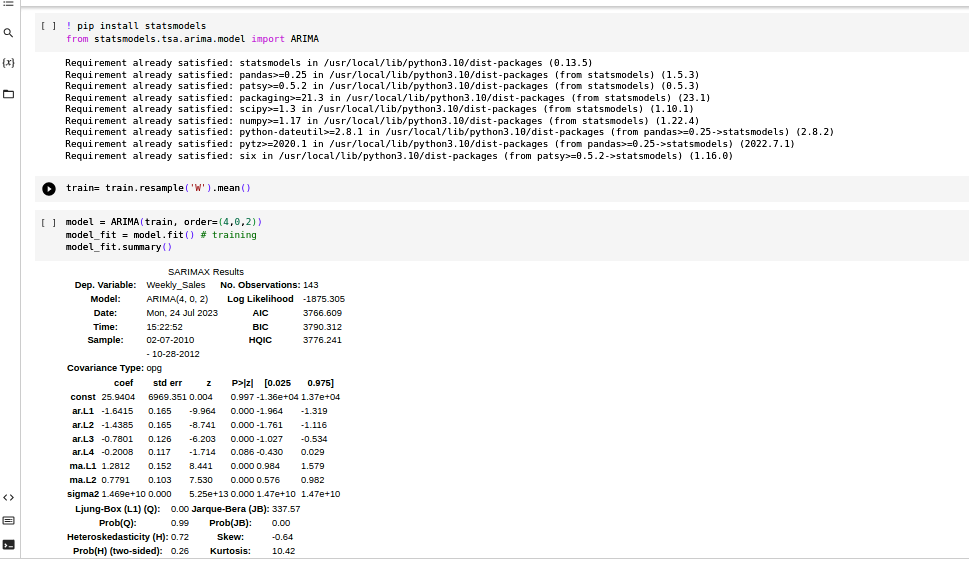


5.Making Date column as index column :

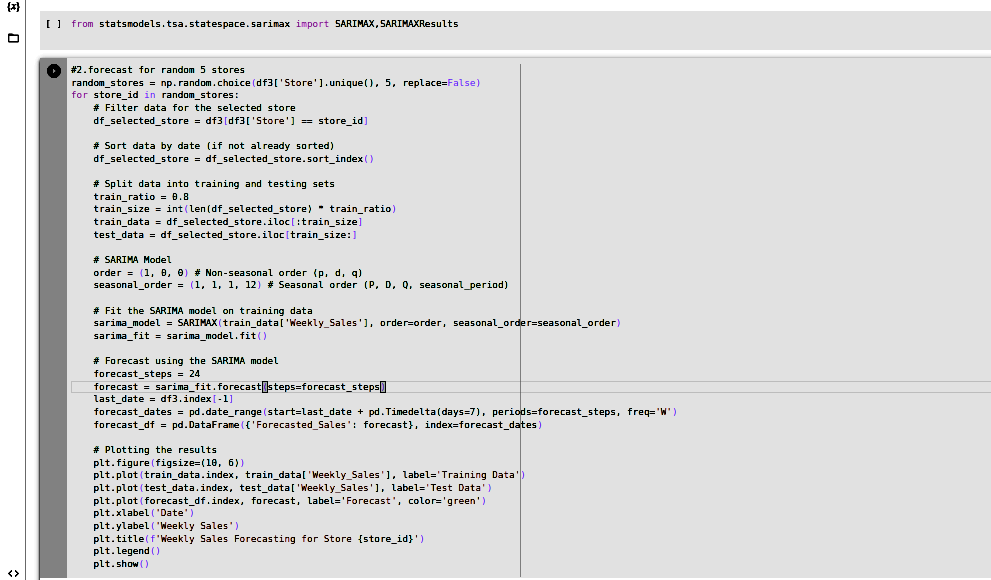


**Choosing Algorithm:**

1.ARIMA Model:



2.SARIMAX Model:

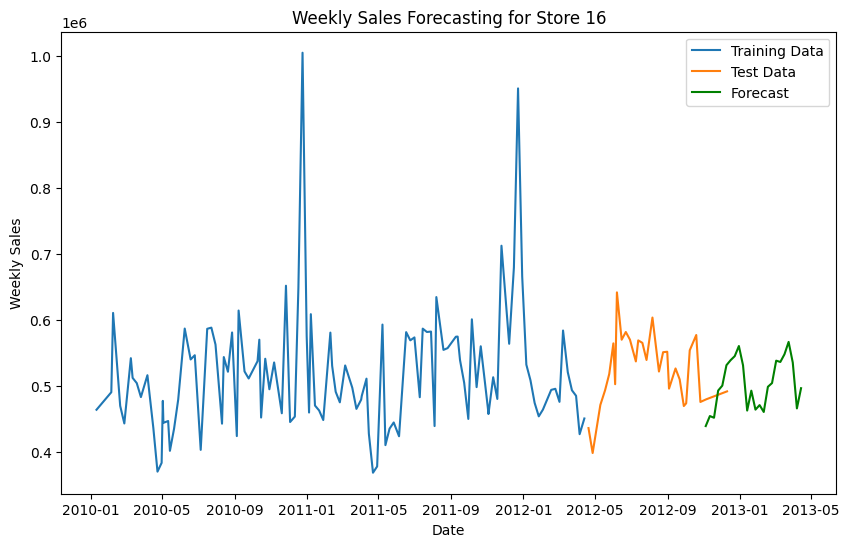


* ***After analyzing these algorithms, we can say SARIMAX is best fit model***

**Motivation and Reasons:** The reason for choosing SARIMAX as the best fit model is since it gives good forecasting for the given data.

**Inferences from same**: We had to perform a Time Series model to forecast the sales for each store for the next 12 weeks. So, we applied many algorithms like ARIMA, SARIMAX etc. We used certain libraries from statsmodels for our models and SARIMAX model gives us the best fit model.

**Model Evaluation:** We evaluate our model by plotting graph for our training data, test data and forecasting.



**Conclusion:**

1. The dataset doesn’t contain any null values.
2. Store number 36 is the store most affected by unemployment.
3. We can see seasonal trends and the effect of holidays on weekly sales of the stores.
4. The temperature shows a weak negative correlation with the weekly sales of the stores.
5. The consumer price index shows a weak negative correlation with the weekly sales of the stores.
6. Store 20 is the top-performing store according to the historical data.
7. Store 33 is the worst-performing store according to the historical data.

**References:** Statsmodel library are used, here is the link for further reference:

<https://www.statsmodels.org/stable/index.html>

For coding I took some help from previous lectures taught in session.