

Walmart Time Series Analysis Project Report

Table of Contents

Problem Statement
Project Objective
Data Description
Data Pre-processing Steps and Inspiration
Choosing the Algorithm for the Project
Motivation and Reasons For Choosing the Algorithm
Assumptions
Model Evaluation and Techniques
Inferences from the Same
Future Possibilities of the Project
Conclusion
References

1. Problem Statement

The problem revolves around managing inventory across multiple outlets of a retail store, specifically Walmart. The challenge is to effectively match the demand for products with the available supply, taking into account factors such as seasonality, holidays, and economic indicators.

2. Project Objective

The objective of this project is to develop predictive models that can forecast sales for each store of Walmart and overall sales for the next 12 weeks. These forecasts will assist in inventory management, ensuring that the right products are available at the right stores to meet customer demand.

3. Data Description

The dataset contains information on weekly sales, holiday flags, temperature, fuel price, consumer price index (CPI), and unemployment for multiple Walmart stores across the country.

The dataset Walmart.csv contains 6435 rows and 8 columns, including the following features:

Store: Store number

Date: Week of Sales

Weekly_Sales: Sales for the given store in that week

Holiday_Flag: Indicates if it is a holiday week

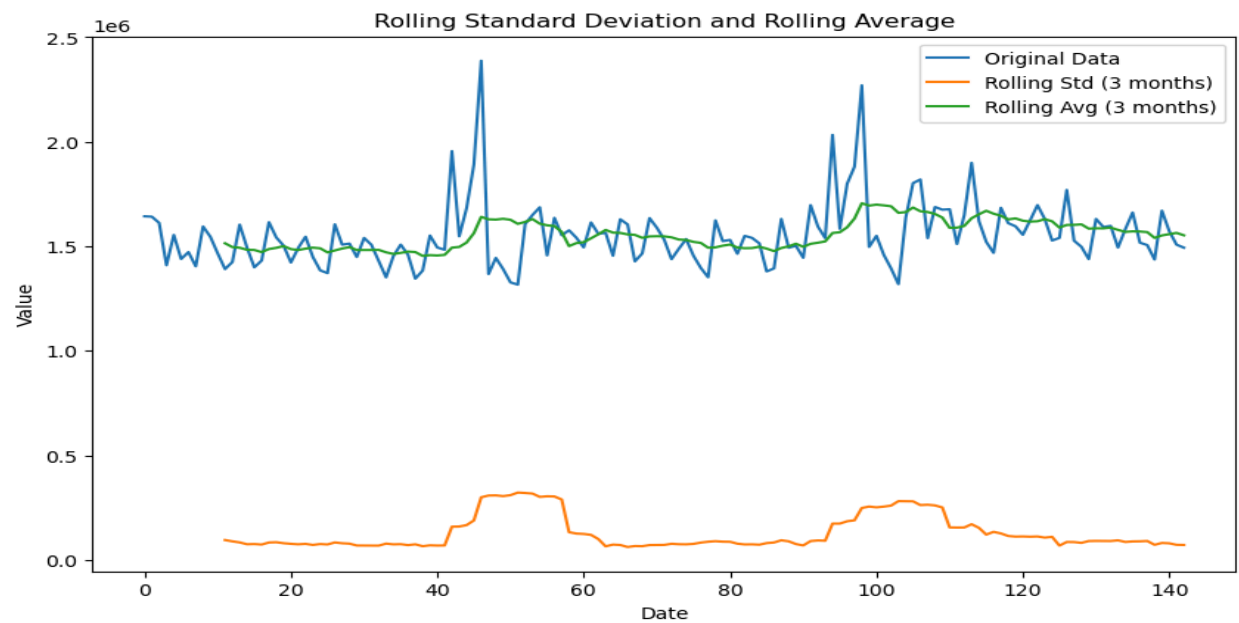
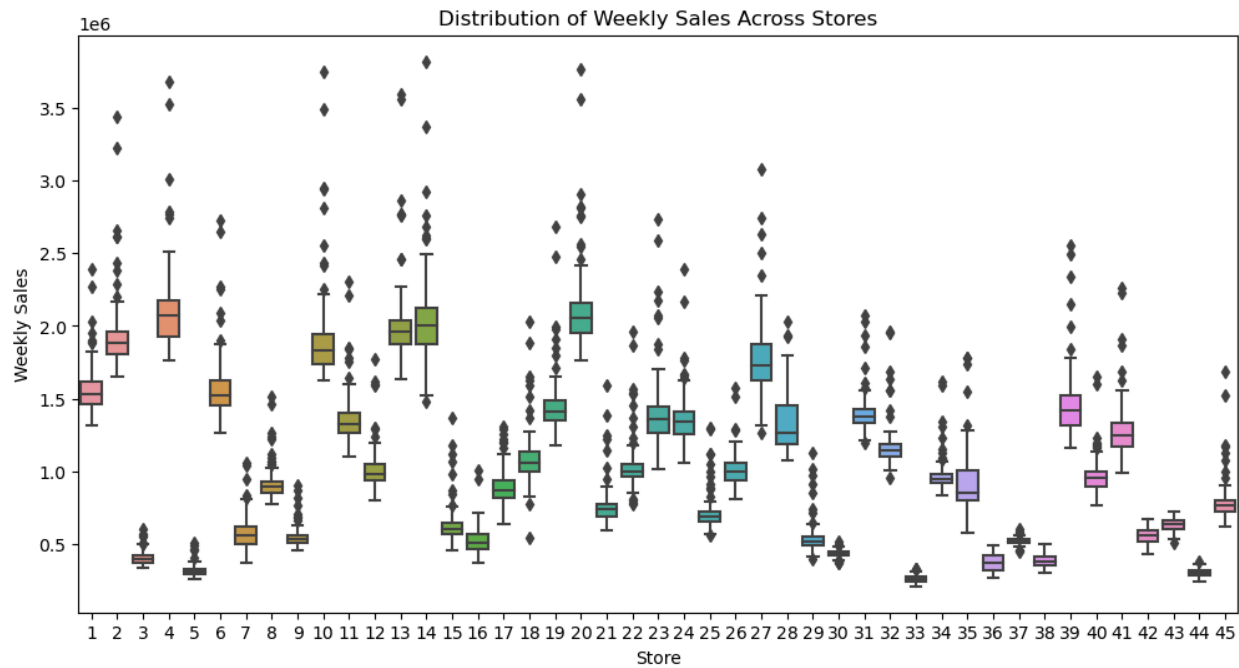
Temperature: Temperature on the day of the sale

Fuel_Price: Cost of fuel in the region

CPI: Consumer Price Index

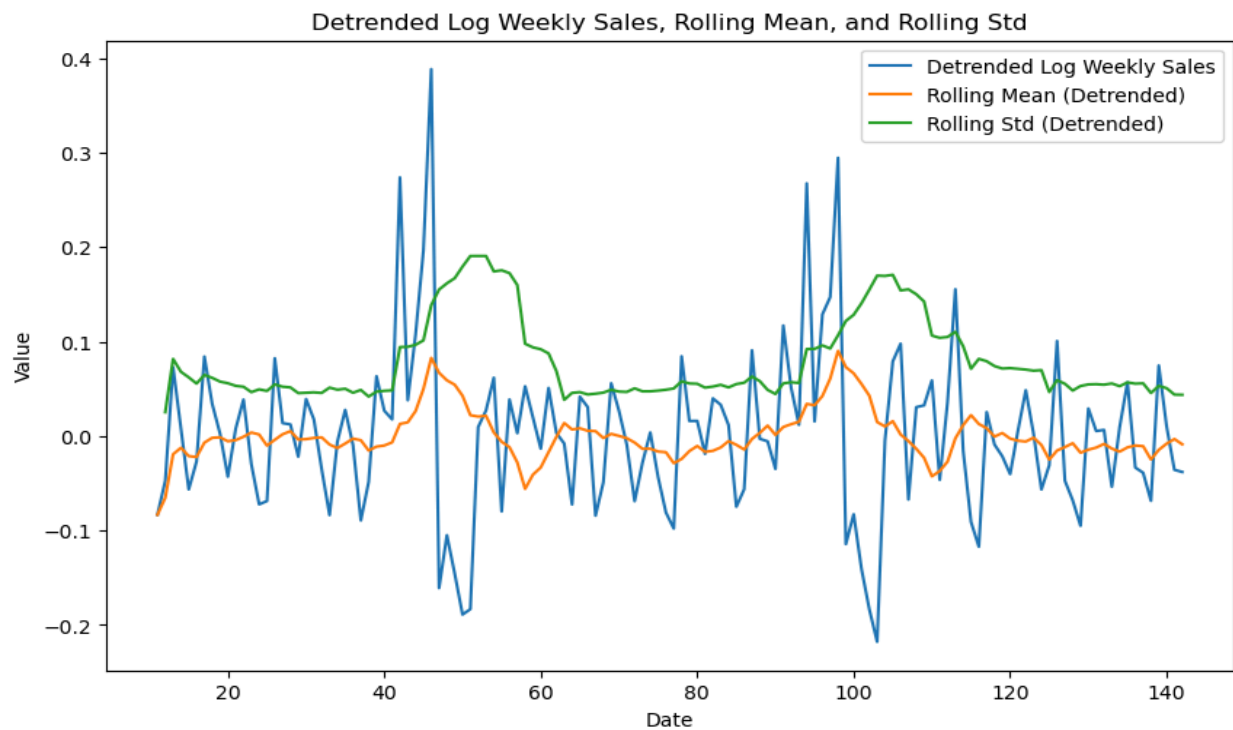
Unemployment: Unemployment Rate

4. Data Pre-processing Steps and Inspiration

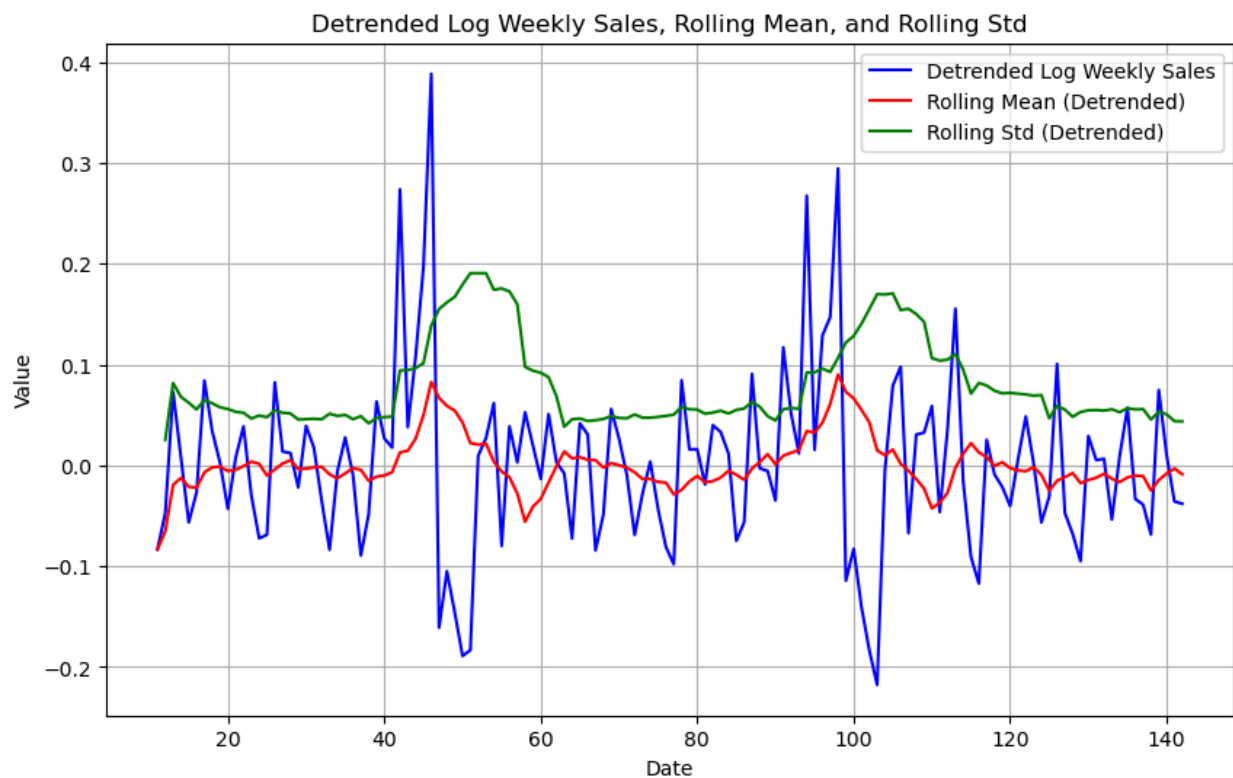


We can see data is not stationary as mean and std is fluctuating so we will take log of weekly sales and then will calculate again rolling mean of that log values and then will subtract that `log_rolling_mean` from `log_weekly_sales`

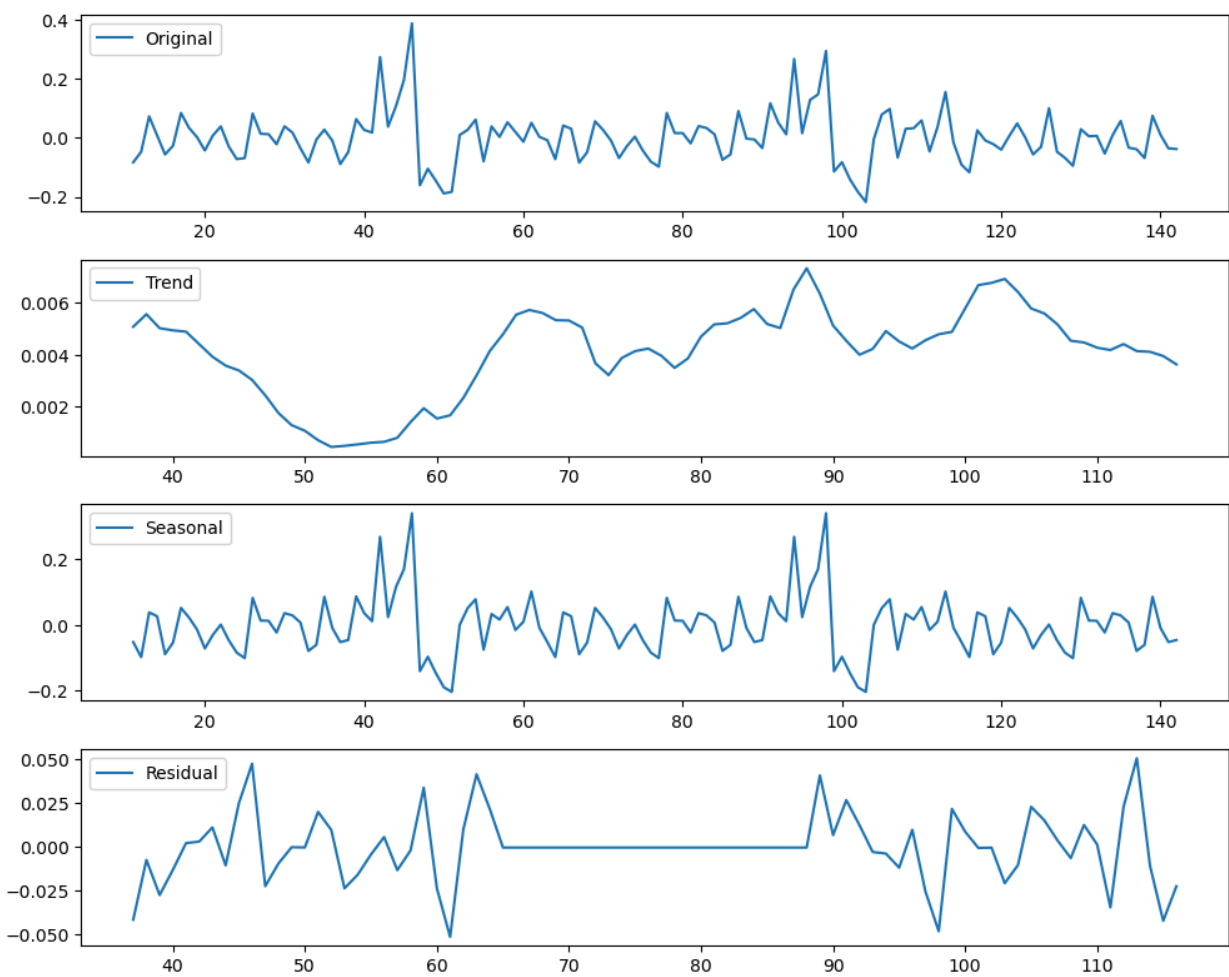
Detrended Log Weekly Sales, Rolling Mean, and Rolling Std

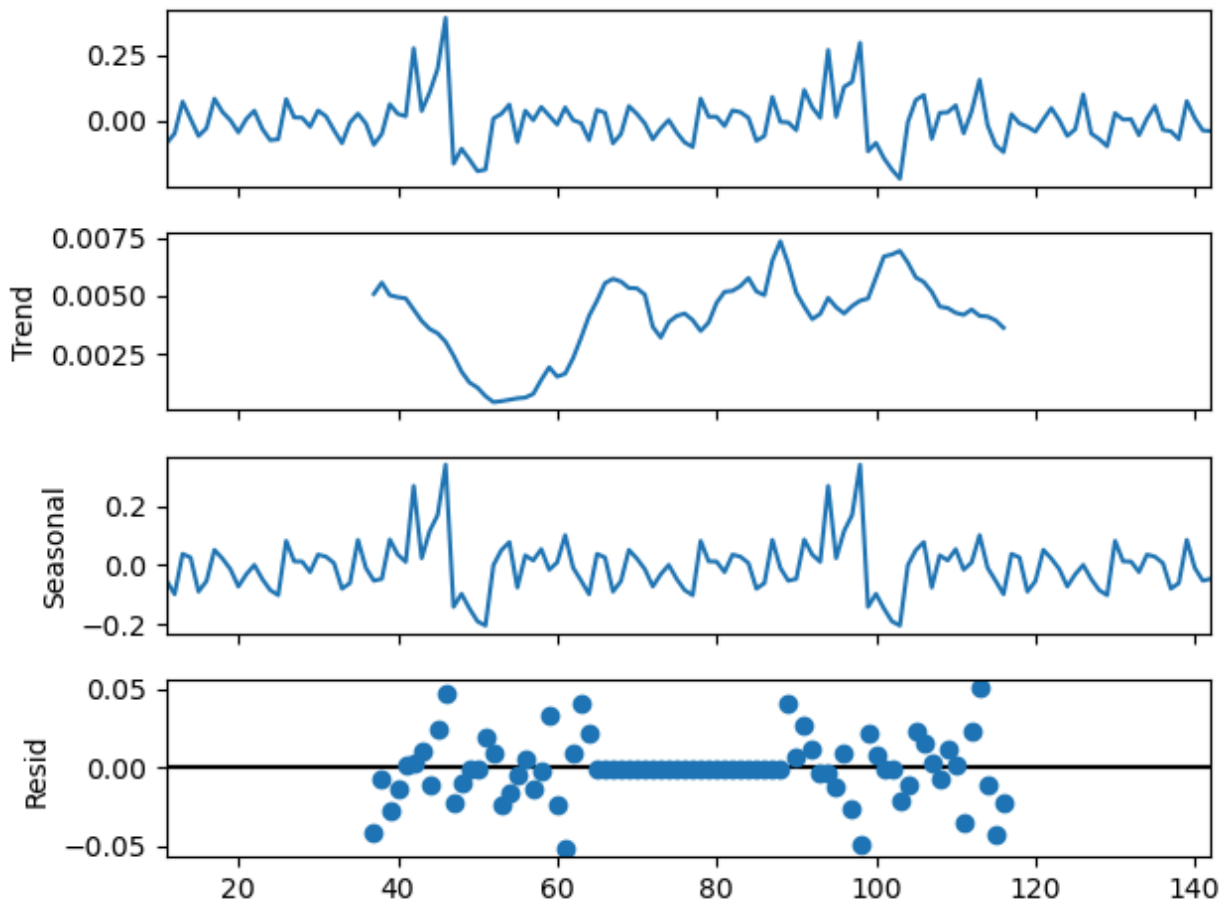


here we are checking seasonality, trend, randomness of our processed data



Decompose the detrended log weekly sales into trend, seasonal, and residual components



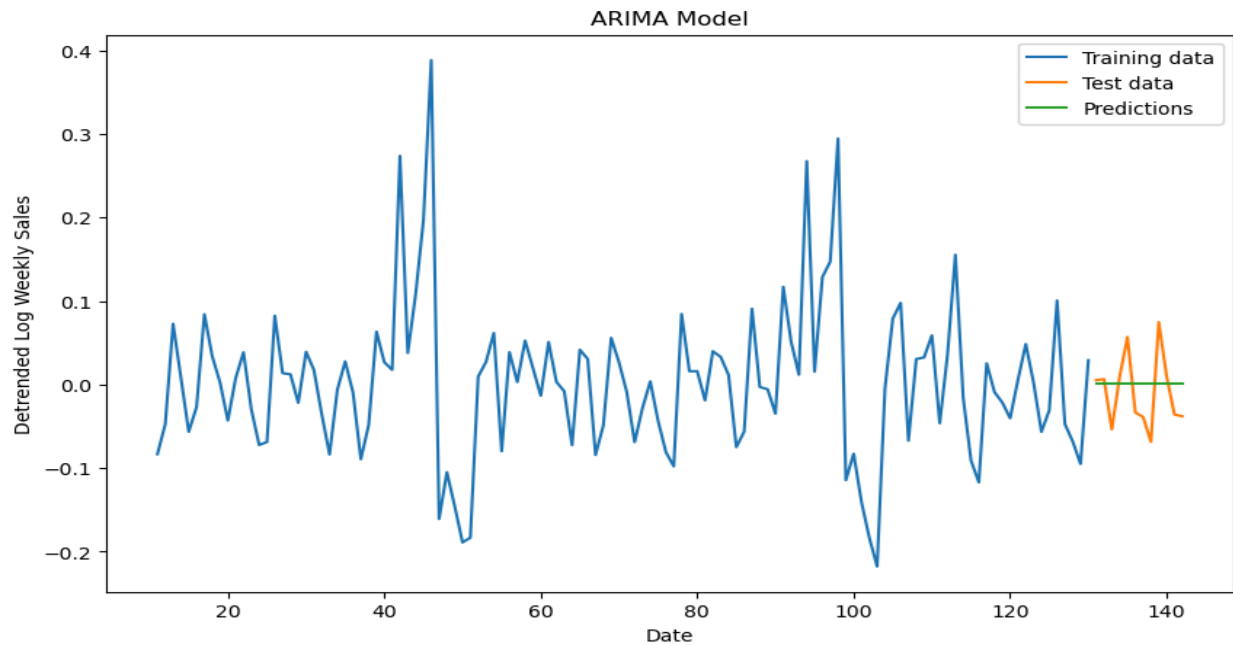


The data pre-processing steps include handling missing values, outlier detection, seasonal_decompose and stationarity. The inspiration behind these steps is to ensure the quality and relevance of the data for model training.

5. Choosing the Algorithm for the Project

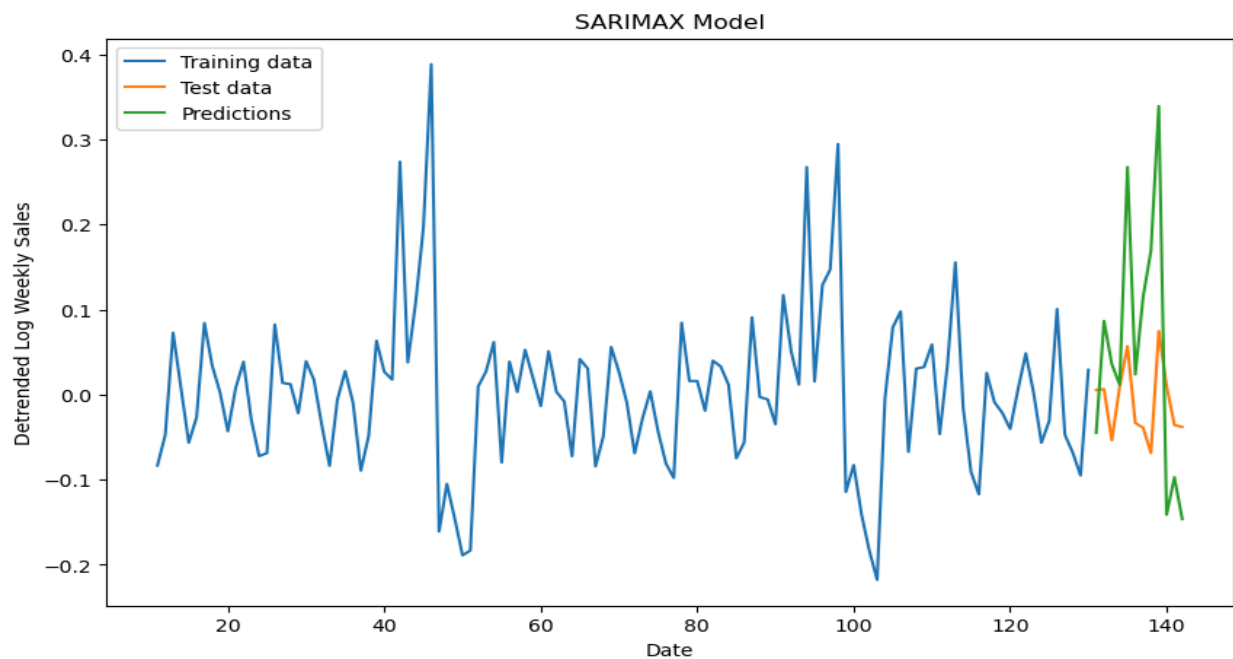
ARIMA Model

Generate predictions for the test set



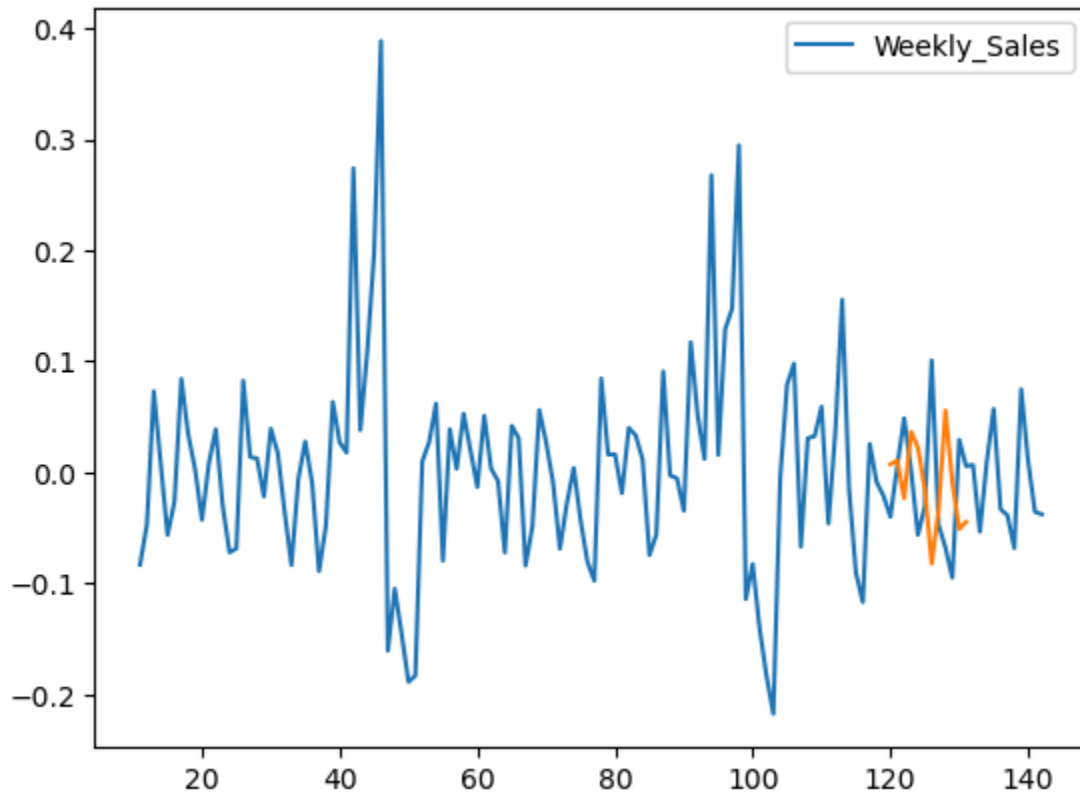
SARIMAX Model

Generate predictions for the test set



For this project, time series forecasting algorithms such as ARIMA, SARIMAX, and exponential smoothing are chosen due to their effectiveness in capturing temporal patterns and seasonality in sales data.

Generate forecasts for the next 12 weeks



6. Motivation and Reasons For Choosing the Algorithm

The motivation behind choosing time series forecasting algorithms is their ability to handle sequential data and capture patterns over time. These algorithms are well-suited for sales forecasting tasks where historical trends and seasonality play a crucial role.

7. Assumptions

The historical sales data is representative of future sales patterns.
External factors such as holidays, temperature, and economic indicators have a significant impact on sales.

8. Model Evaluation and Techniques

Mean Absolute Error (MAE): 0.1222703309873252
Mean Squared Error (MSE): 0.0211004576008912
Root Mean Squared Error (RMSE): 0.1452599655820254

Model evaluation involves assessing the performance of the ARIMA, SARIMAX, and exponential smoothing models using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Cross-validation techniques may also be employed to ensure the robustness of the models.

9. Inferences from the Same

Inferences from the model evaluation phase provide insights into the accuracy and reliability of the forecasts. These insights can help stakeholders make informed decisions regarding inventory management and resource allocation.

10. Future Possibilities of the Project

Future possibilities include enhancing the predictive models by incorporating additional features or exploring advanced machine learning techniques. Furthermore, integrating real-time data sources and improving model interpretability could further enhance the project's capabilities.

11. Conclusion

The project concludes with a summary of the findings, highlighting the effectiveness of the developed forecasting models in predicting sales for Walmart stores. Recommendations for future research and potential areas for improvement.

12. References

LMS, Google, Kaggle.

This comprehensive report outlines the Walmart project, from problem statement to conclusion, providing insights into the data, methodology, and findings. It serves as a valuable resource for stakeholders and researchers interested in retail inventory management and time series forecasting.