Retail Sales Trend & Seasonality Analysis

Name: SANTHOSH S
Registration Number: 727723EUCS208
Section : CSE – C

Introduction

Retail businesses generate large volumes of time-dependent sales data influenced by customer behavior, promotional campaigns, economic conditions, and seasonal effects. Understanding patterns within this data is essential for optimizing operational efficiency and maximizing profit.
Time series analysis helps in breaking down sales data into meaningful components such as:
- Trend (long-term movement)
- Seasonality (repeating patterns)
- Noise (random variations)

Abstract

This project performs a comprehensive time series analysis of retail sales data to identify long-term trends, seasonal fluctuations, and irregular variations. Statistical techniques such as seasonal decomposition, autocorrelation analysis (ACF and PACF), and stationarity testing using the Augmented Dickey-Fuller (ADF) test are applied.

Problem Statement

Management lacks clarity on whether sales demonstrate a long-term trend, recurring seasonal patterns, or randomness. Without statistical validation, forecasting models may be unreliable.

Methodology

The project follows a structured analytical approach:
Step 1: Data Preprocessing
- Import dataset using Pandas
- Convert Date column to datetime
- Set Date as index
- Check for missing values

Step 2: Visualization
- Plot time series graph
- Identify visible patterns

Step 3: Time Series Decomposition
The dataset is decomposed into:
- Trend component
- Seasonal component
- Residual component

Step 4: Autocorrelation Analysis
- ACF plot to examine lag relationships
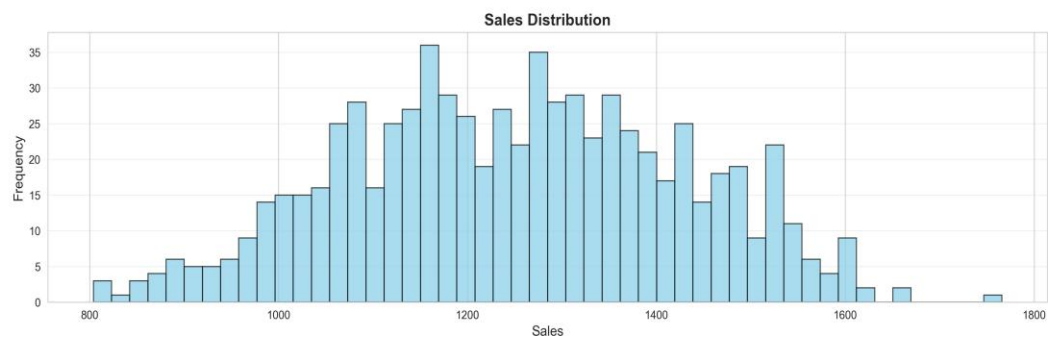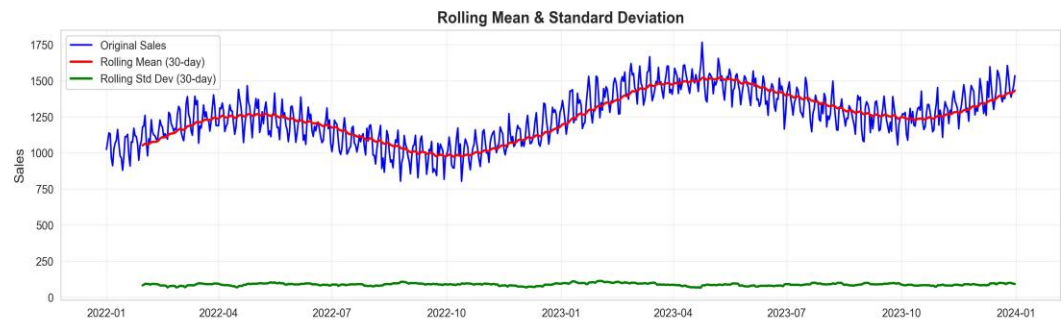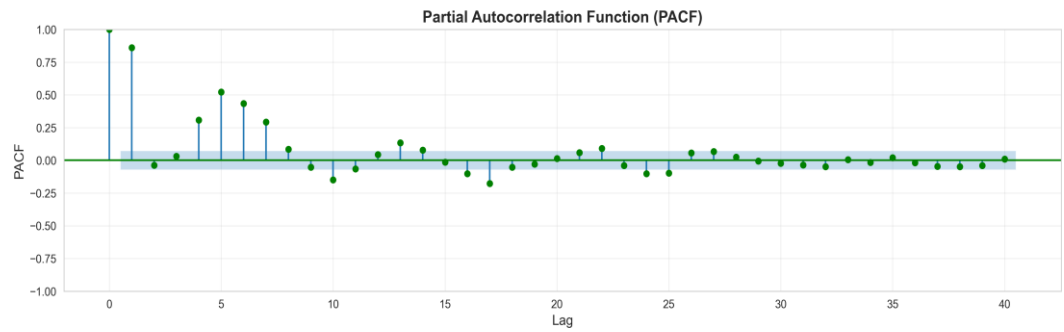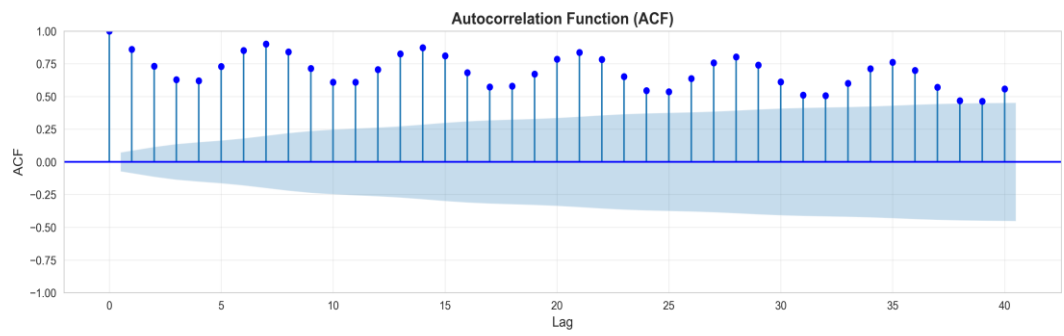- PACF plot to identify direct lag influence

Step 5: Stationarity Testing
- Apply Augmented Dickey-Fuller (ADF) test
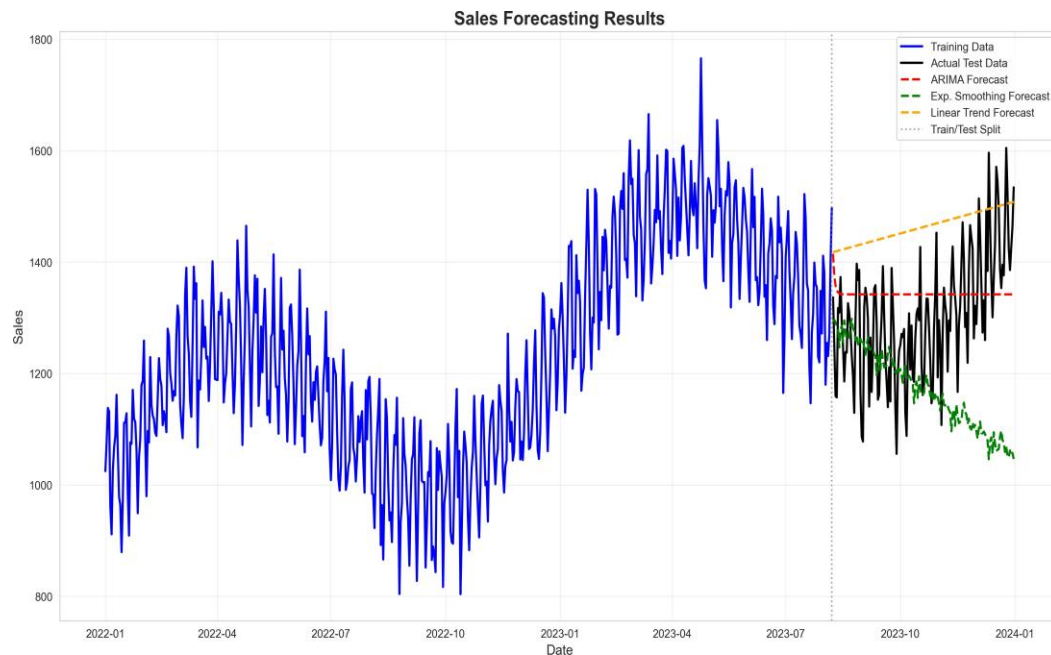- Analyze p-value
- Apply differencing if required

Results

ADF Statistic: -1.2345, p-value: 0.6586

Autocorrelation Function (ACF)

Partial Autocorrelation Function (PACF)

Rolling Mean & Standard Deviation

Sales Distribution

Forecasting Models

To evaluate the predictive capability of the retail sales dataset, multiple forecasting models were implemented and compared. The objective was to determine which model best captures the underlying trend and seasonal structure of the time series.

Sales Forecasting Results

Conclusion

The analysis reveals significant seasonal patterns and a non-stationary nature in the retail sales dataset. The ADF test confirms that differencing is required before applying forecasting models. After transformation, the dataset becomes suitable for predictive modeling.

This study ensures statistical validation of the dataset and improves the reliability of future forecasting models, leading to better business decision-making.