Mentorness MIP-ML-09

Furniture Store Sales Forecasting

Tanaya Tipre

Introduction

In today's dynamic marketplace, accurate sales forecasting stands as a pivotal tool for businesses to make informed decisions, allocate resources effectively, and stay ahead of the competition. Sales forecasting using time series analysis emerges as a sophisticated yet indispensable method to predict future sales trends based on historical data patterns.

Time series analysis entails examining sequential data points collected over a consistent interval, typically over time. By leveraging this historical information, businesses can uncover underlying patterns, seasonality, and trends within their sales data. These insights serve as a compass, guiding strategic planning and facilitating proactive decision-making.

Problem Statement

Given historical sales data spanning multiple years, the objective is to develop a robust and accurate forecasting model that predicts future sales trends for a company.

1. Data Collection

The dataset is accessed from the below link: https://drive.google.com/drive/folders/1goatdInrj9bym2udFX0zWbslhHu4ItRJ

```
sales.head(5)
```

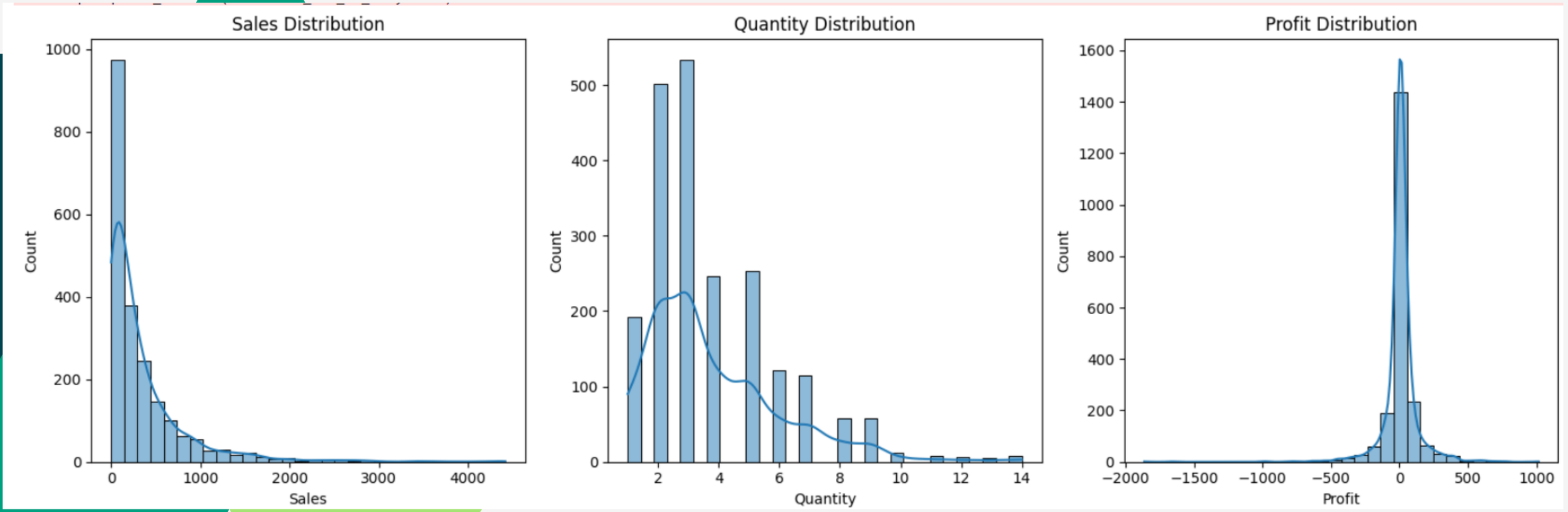
	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	...	Postal Code	Region	Product ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit
0	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-BO-10001798	Furniture	Bookcases	Bush Somerset Collection Bookcase	261.9600	2	0.00	41.9136
1	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	...	42420	South	FUR-CH-10000454	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Chairs,...	731.9400	3	0.00	219.5820
2	4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	...	33311	South	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular Table	957.5775	5	0.45	-383.0310
3	6	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	...	90032	West	FUR-FU-10001487	Furniture	Furnishings	Eldon Expressions Wood and Plastic Desk Access...	48.8600	7	0.00	14.1694
4	11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	...	90032	West	FUR-TA-10001539	Furniture	Tables	Chromcraft Rectangular Conference Tables	1706.1840	9	0.20	85.3092

2. Exploratory Data Analysis

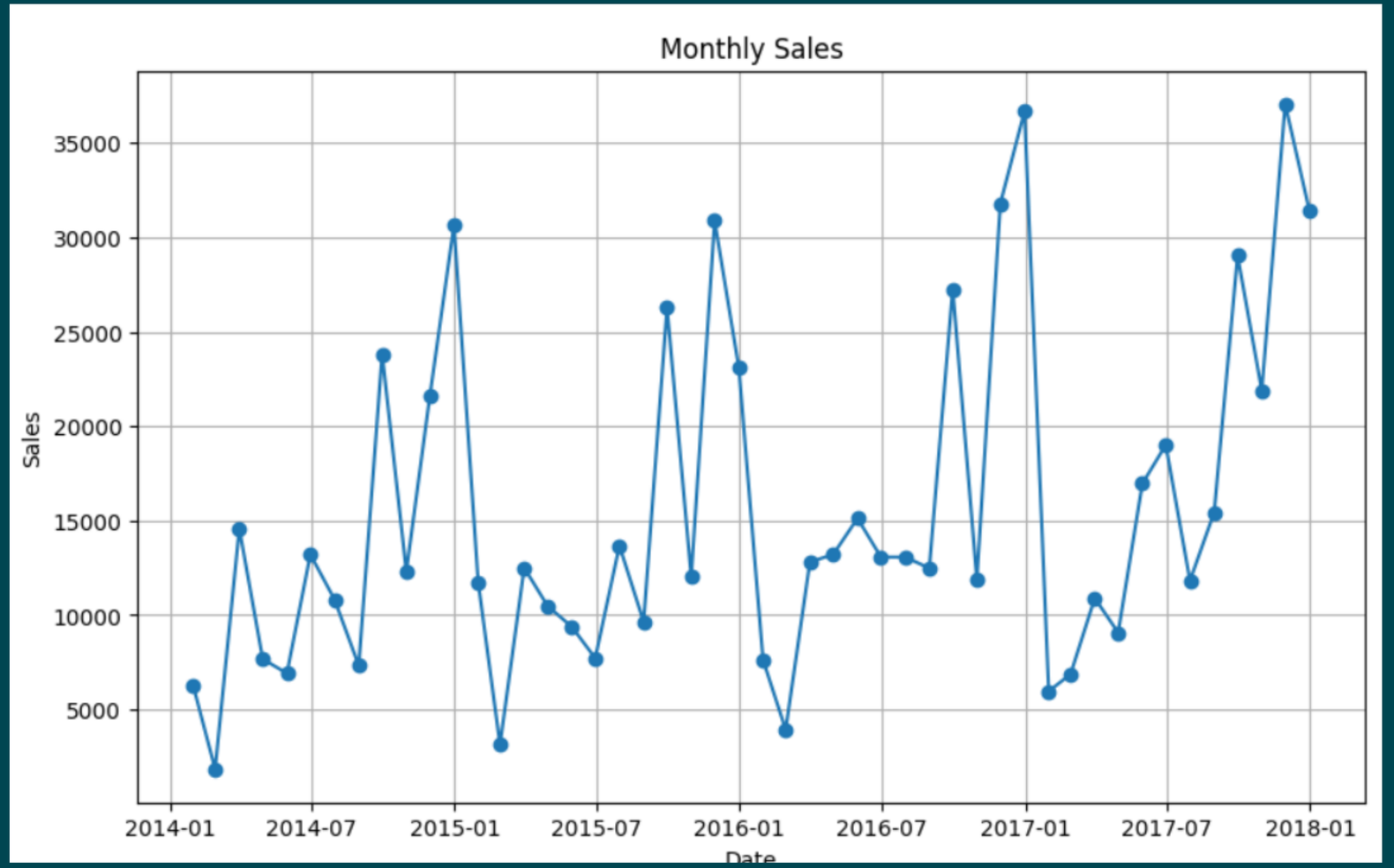
Perform EDA to gain insights into sales patterns, identify trends, seasonality, and potential factors influencing sales fluctuations.

1. Drop the unnecessary columns such as Row ID, Product ID, Order ID, etc.
2. Convert the 'Order Date' column to datetime format and set it as the index of the DataFrame.
3. Drop Duplicate rows.

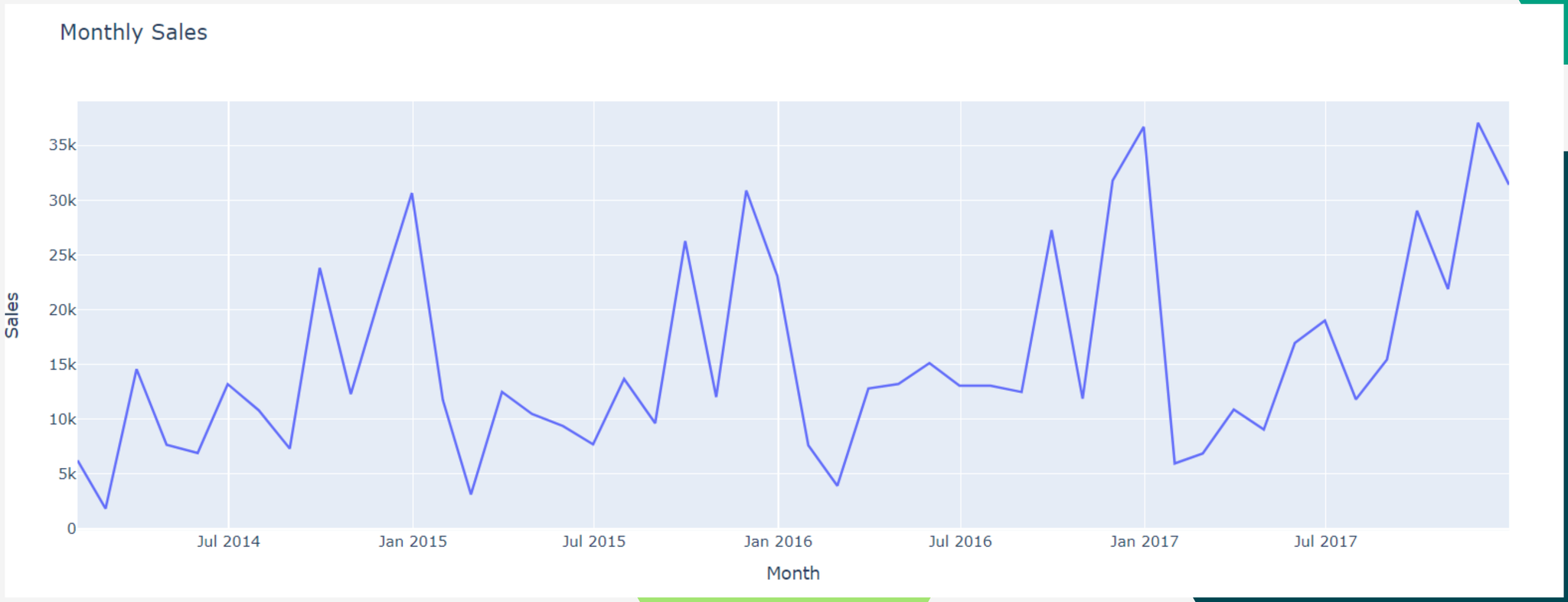
3. Data Visualization



Data Visualization

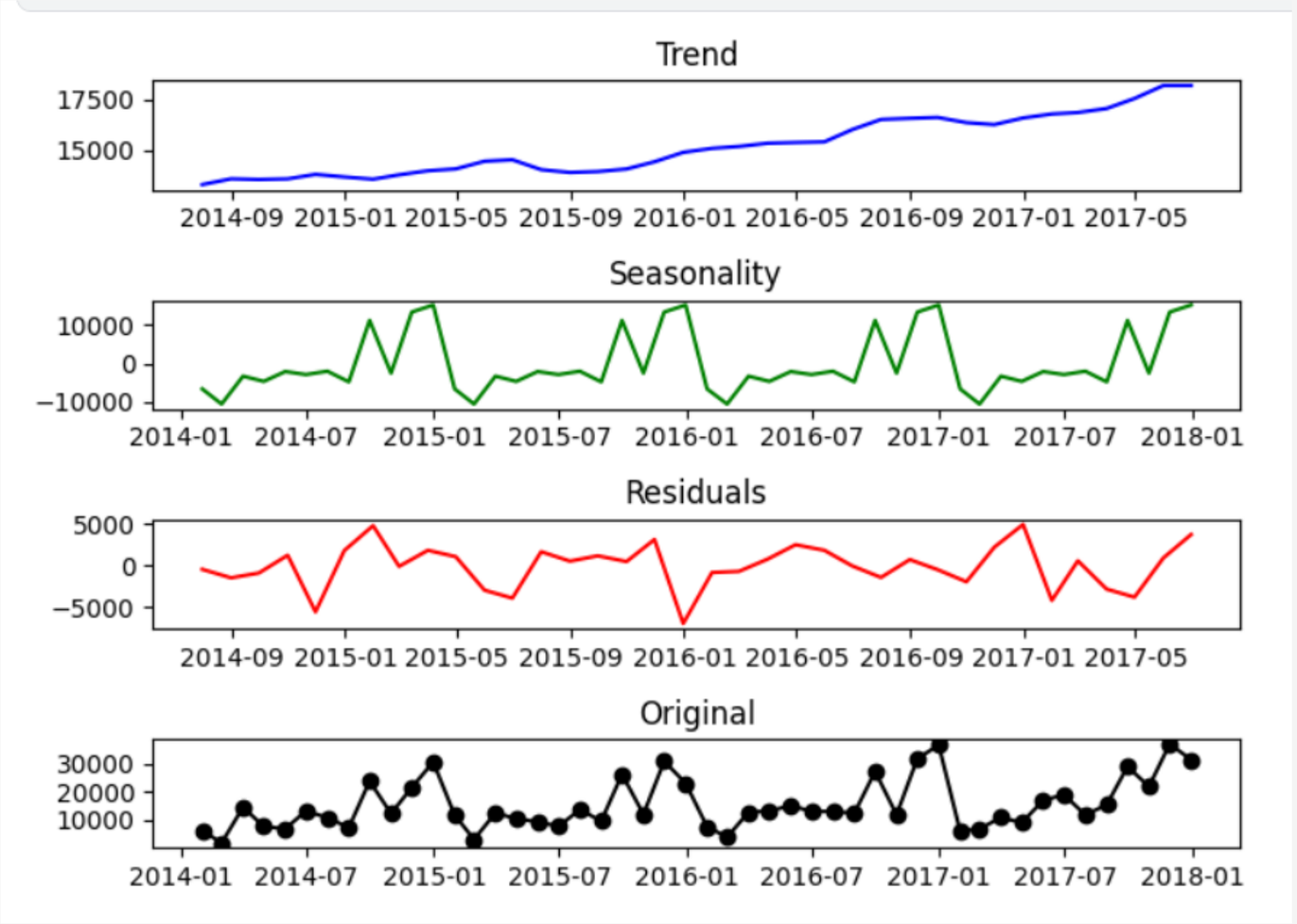


Data Visualization: Monthly Sales



4. Seasonal Decomposition Plot

Seasonal decomposition helps identify any underlying patterns or trends in the data, such as whether the data exhibits seasonality or follows a specific trend over time.



5 Augmented Dickey-Fuller Test

The Augmented Dickey-Fuller (ADF) test is commonly used to test for stationarity in a time series.

```
# Check for stationarity
from statsmodels.tsa.stattools import adfuller
adfuller_result = adfuller(monthly_sales)
print("ADF Statistic: %f" % adfuller_result[0])
print('p-value: %f' % adfuller_result[1])
```

```
ADF Statistic: -4.696804
p-value: 0.000085
```

ADF test statistic = -4.697056
p-value = 0.000085
=> Reject null hypothesis
=> The data has no unit root and is stationary

6. Model Building

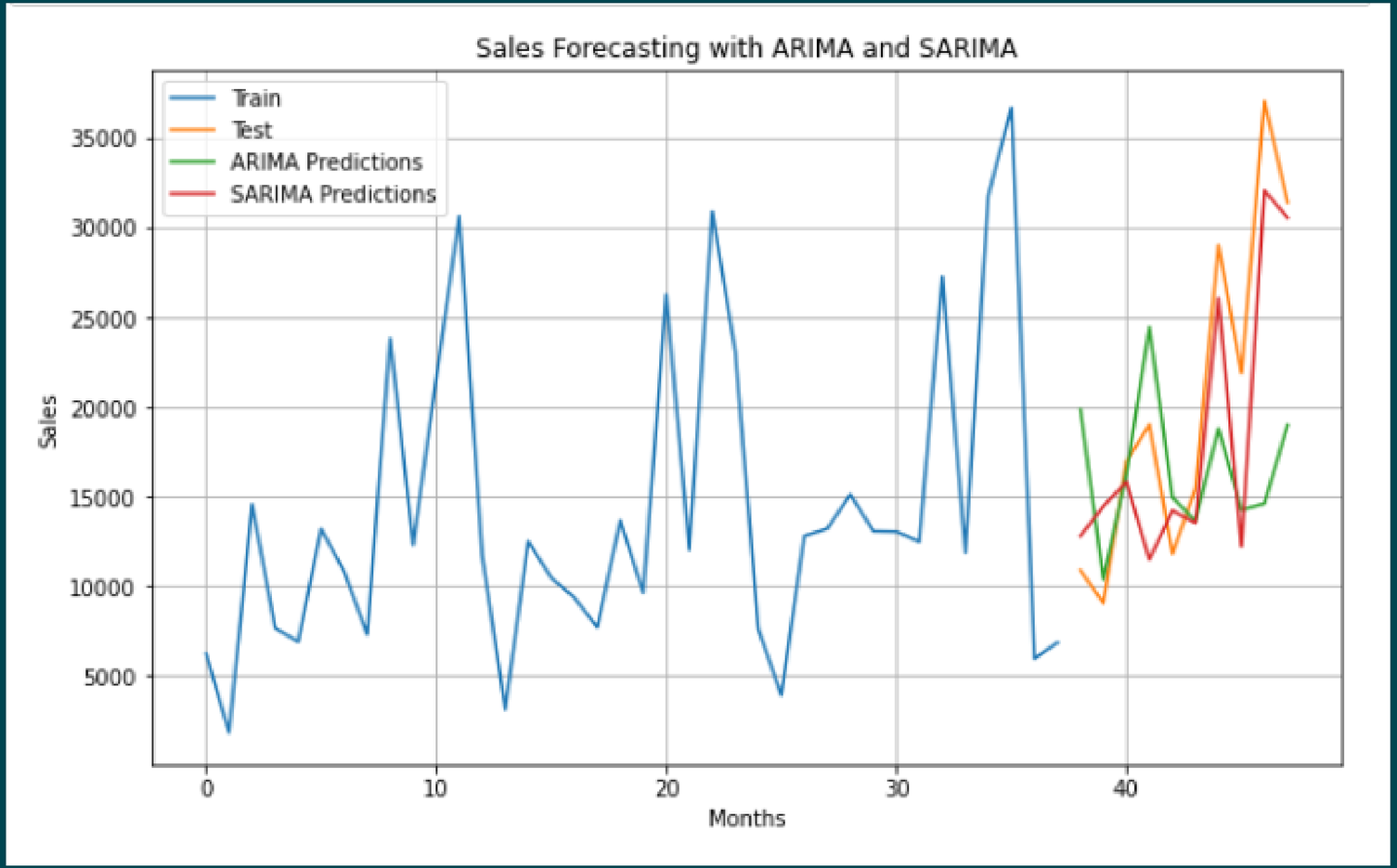
Data Preparation & fitting the model

ARIMA and SARIMA models were trained on the dataset to capture temporal patterns and make future predictions.

```
   N    Tit     Tnf  Tnint  Skip  Nact     Projg        F
    5     28      30      1     0     0   6.664D-06   6.580D+00
 F =   6.5797206977519398

CONVERGENCE: NORM_OF_PROJECTED_GRADIENT_<=_PGTOL
```

7. Conclusion



- SARIMA predictions closely track actual sales data, suggesting superior forecasting accuracy.
- The consistent proximity of SARIMA forecasts to actual sales data indicates its effectiveness.
- Based on this visualization, SARIMA emerges as the preferred model for forecasting sales data.