Time Series Forecasting Project

UPI Transactions Volume & Value Forecasting

Guided By:
Monika Bhattacharjee
Assistant Professor, IIT Bombay

Problem Statement

❑ Objective:

Time Series Forecasting of UPI Transaction Volume and Value using ARIMA and SARIMA Models

❑ Emphasis:

Robust model diagnostics and validation were conducted to ensure reliable forecasting of UPI transaction volume and value, focusing on how well ARIMA and SARIMA models capture long-term trends, recurring monthly patterns, and the evolving dynamics of digital payment behavior in India

About the Dataset

The dataset, sourced from Reserve Bank of India (RBI), is having **~2,942 rows** , and 102 columns and we extracted following three columns for our further analysis.

Columns:

Column Name	Description
Date	The calendar date of observation, recorded in the format yyyy/mm/dd (e.g., 2020/06/01).
Vol (number of transactions)	Vol represents the number of UPI transactions per day (in lakhs)
Val (transaction value)	Val indicates the total value of those transactions (in crores).

Glimpse of the Data

	Date	Vol	Val
0	June 1, 2020	476.9671	10413.108975
1	June 2, 2020	476.78182	9951.298586
2	June 3, 2020	456.2593	9622.375213
3	June 4, 2020	463.04959	9639.502869
4	June 5, 2020	464.79398	9539.524729
5	June 6, 2020	458.64927	9119.199542

Preprocessing of the data

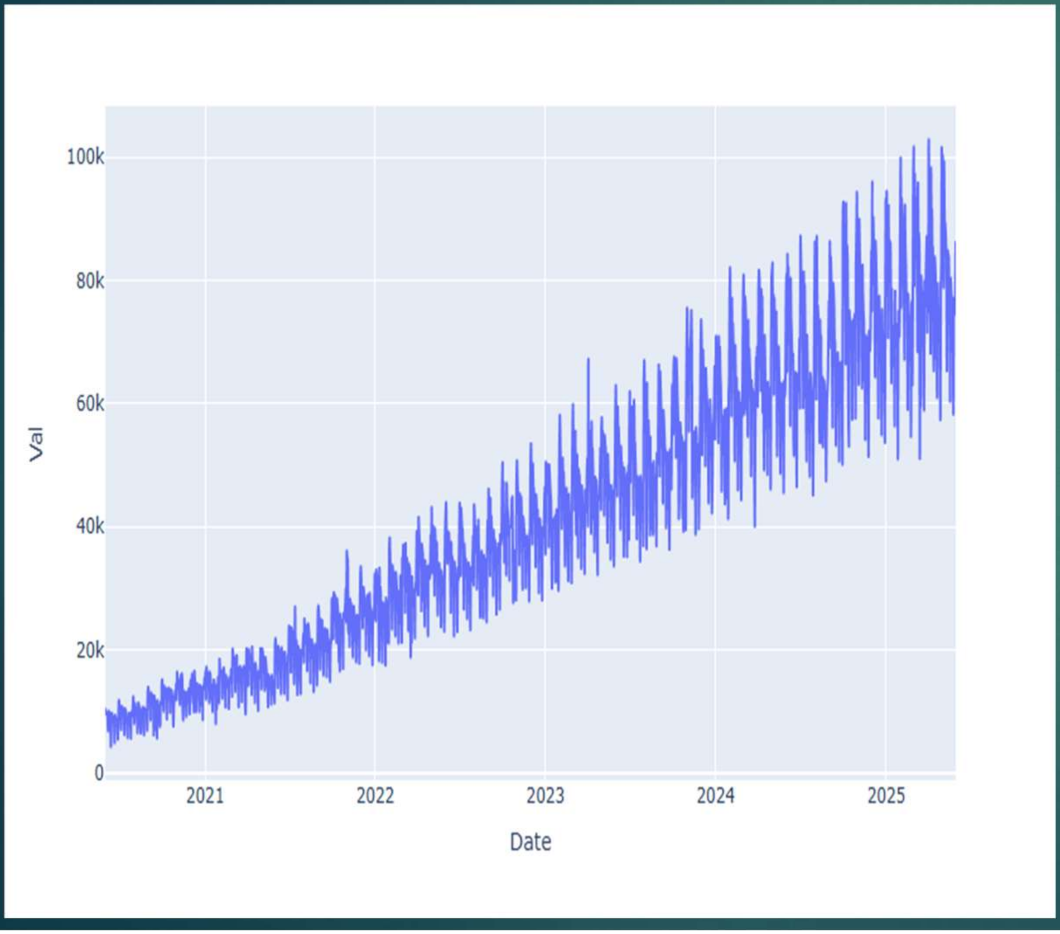
- There are 1112 missing/invalid values in the dataset.
- Converted Date strings to pandas datetime strings.
- Cast Vol and Val to float; no remaining nulls after conversion.
- Since missing values are invalid, hence we removed it and further resampled it in weekly data.

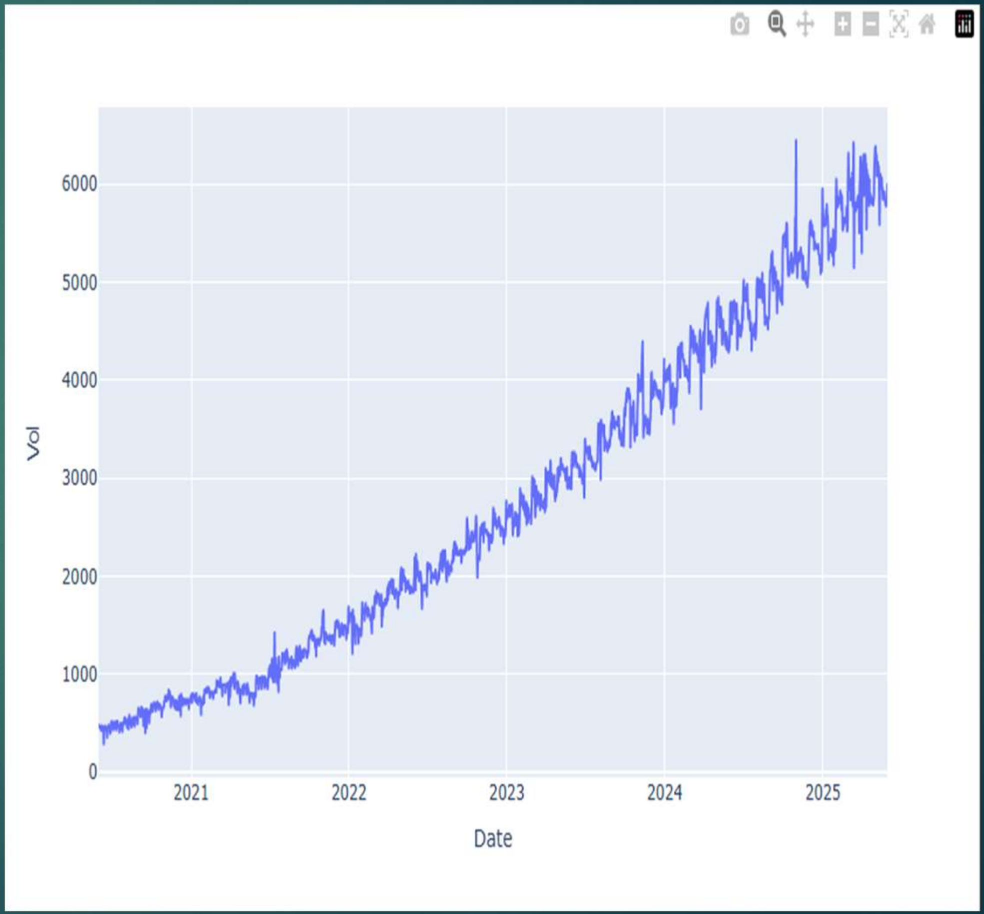
```
df.head()
```

Date	Vol	Val
2020-06-01	476.9671	10413.108975
2020-06-02	476.78182	9951.298586
2020-06-03	456.2593	9622.375213
2020-06-04	463.04959	9639.502869
2020-06-05	464.79398	9539.524729

Exploratory Data Analysis

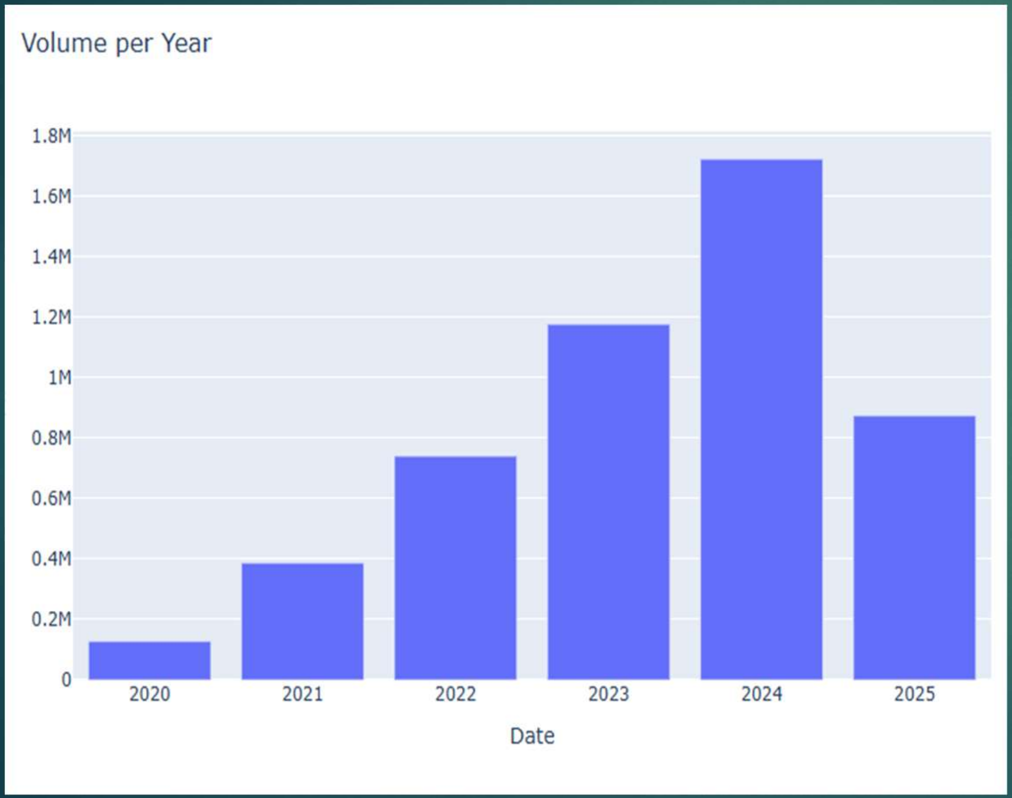
- Time Series Plot: Daily plots of UPI Volume and Value show a strong upward trends

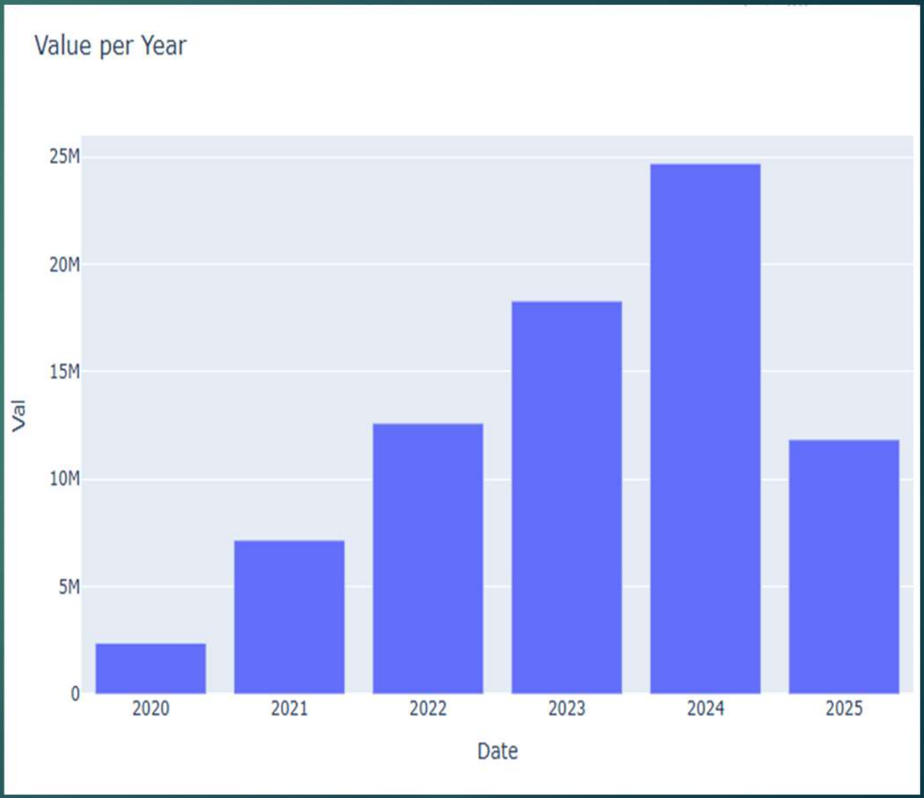




Exploratory Data Analysis

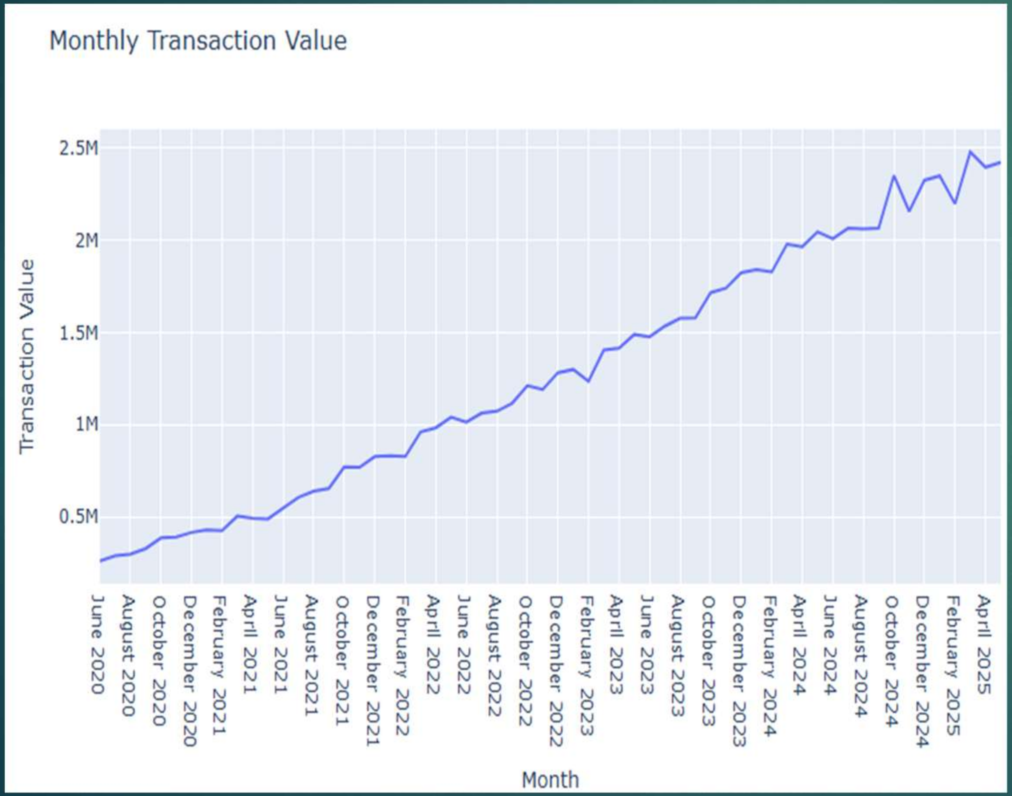
- **Yearly Summary: Bar charts of total Volume and Value by year highlight rapid growth (increasing bars each year).**

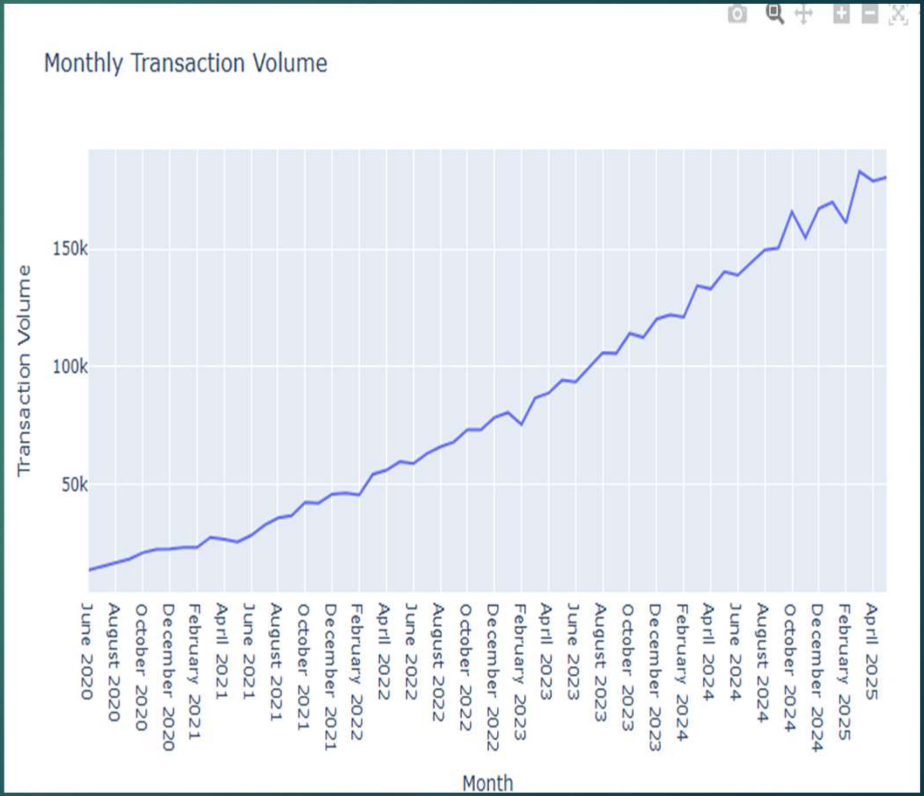




Exploratory Data Analysis

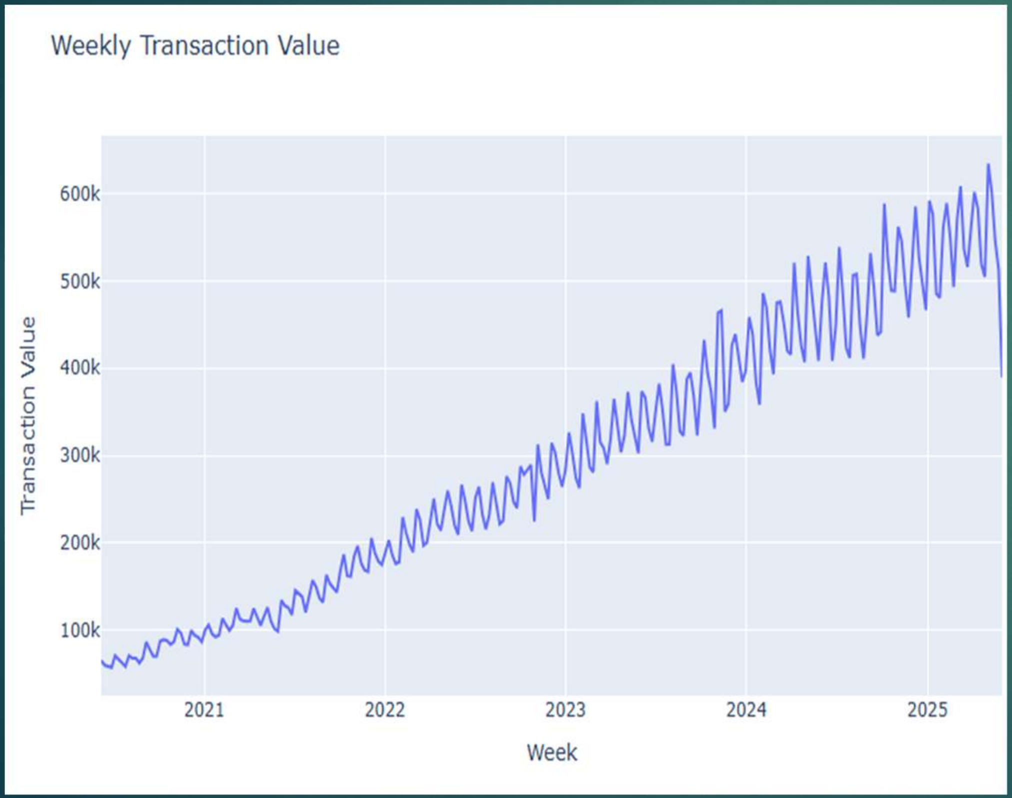
- **Monthly Trends: Line plots of monthly- summed Volume/Value plots show shorter-term fluctuations and confirm overall upward trajectory.**

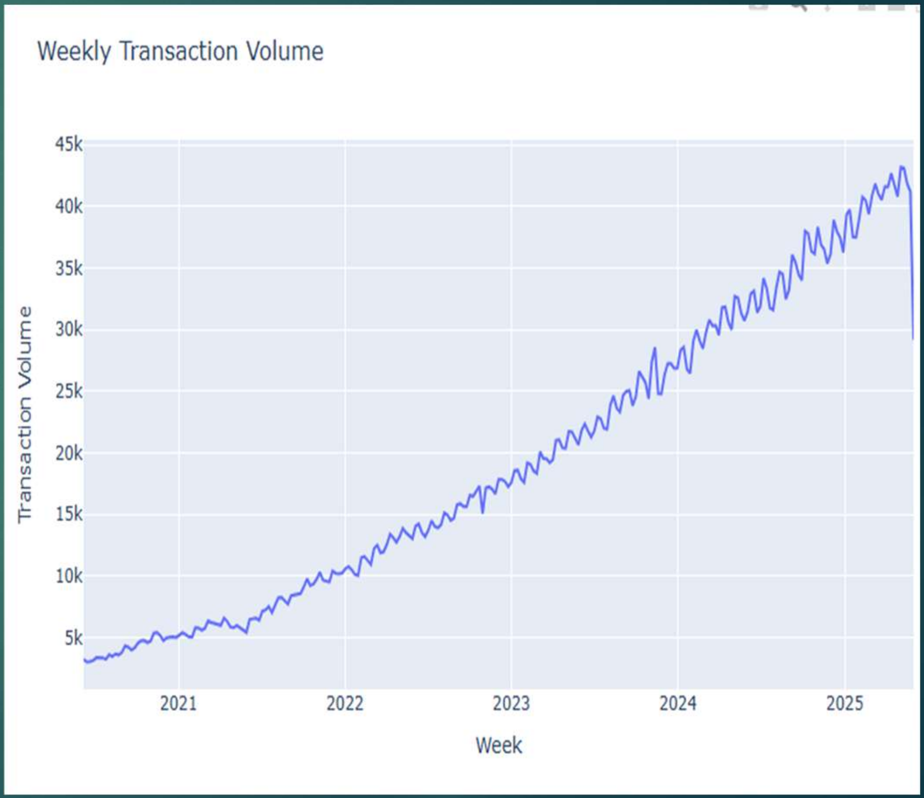




Exploratory Data Analysis

- **Weekly Patterns: Weekly-summed Volume/Value plots show shorter-term fluctuations and confirm overall upward trajectory.**

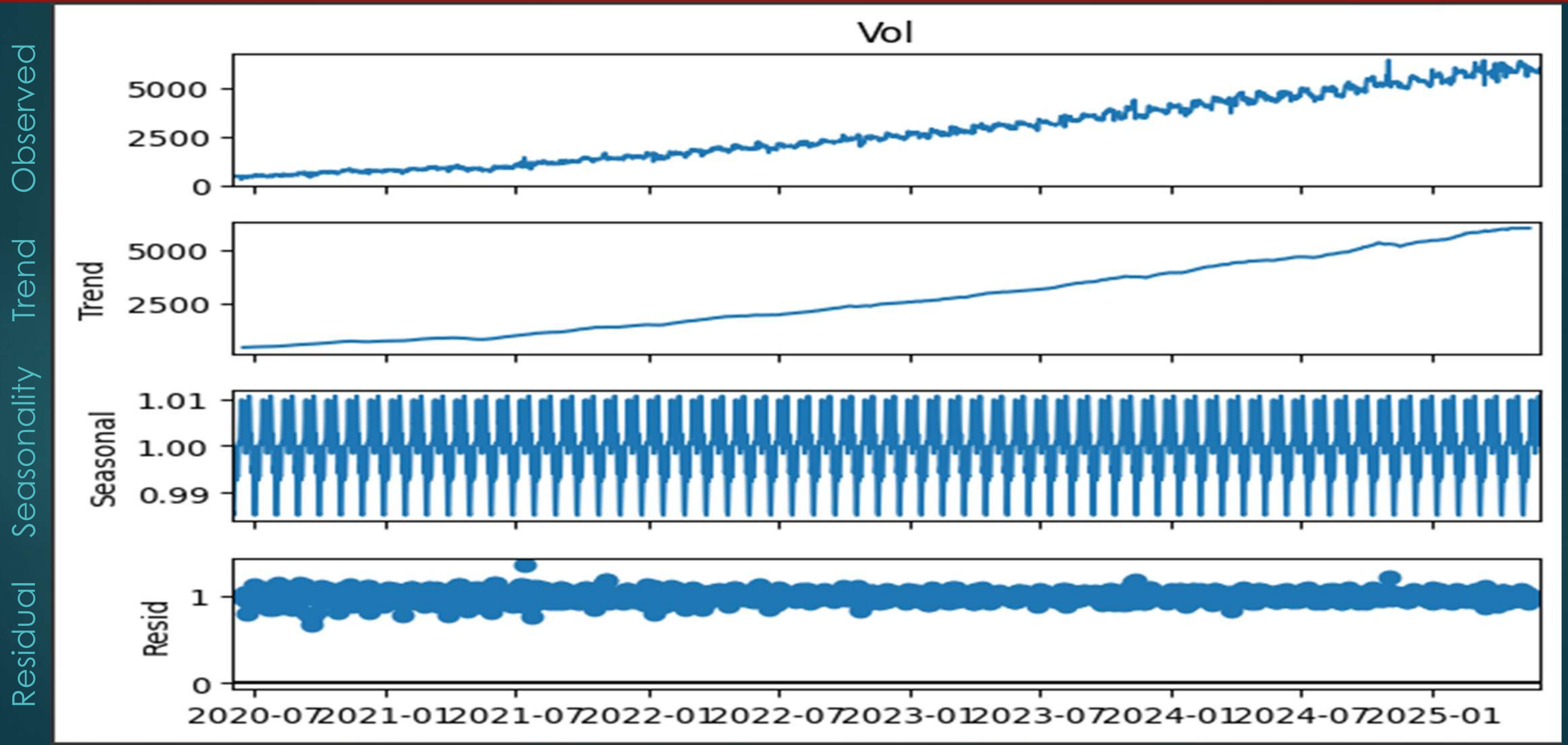




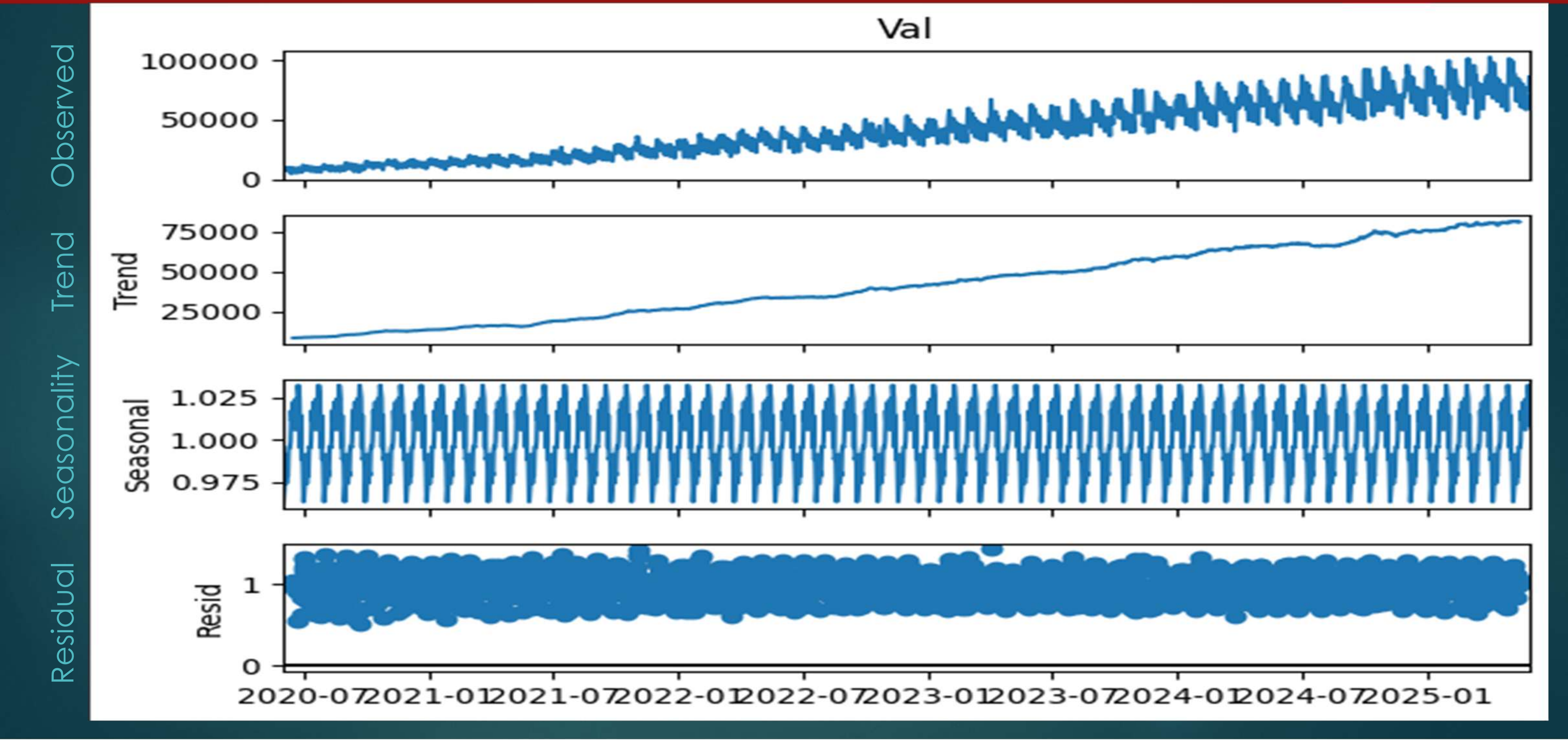
Seasonality Detection

- **FFT on Detrended Series:** Removed long-term trend then computed FFT to find periodicities.
- **Volume:** FFT periodogram of detrended *Vol* has dominant peak corresponding to ~1825 days (≈5 years), reflecting the strong multi-year trend.
- **Value:** FFT of detrended *Val* shows dominant period ~30.4 days, indicating a monthly cycle in transaction value.
- **Interpretation:** These frequencies guided choosing a seasonal period of 30 days for modeling.

Decomposition of the Series



Decomposition of the Series



ADF Test For Stationarity

Volume

- ❑ ADF Test Results
  - Test statistic: *1.8876*
  - p-value: *0.99*
- ❑ Since, *p-value > 0.05,* the time series is non-stationary

Value

- ❑ ADF Test Results
  - Test statistic: *-0.378*
  - p-value: *0.91*
- ❑ Since, *p-value > 0.05,* the time series is non-stationary

- Since it is non-stationarity we need to apply the differencing

Why Seasonal Differencing ?

- Removes repeating seasonal patterns (e.g., monthly cycles)
- Stabilizes the mean of a time series over time
- Helps make the series stationary — a key assumption for ARIMA/SARIMA models
- Enhances model performance by eliminating seasonal noise
- Enables better detection of true underlying trends and fluctuations

ADF Test For Stationarity After Applying Differencing

Volume

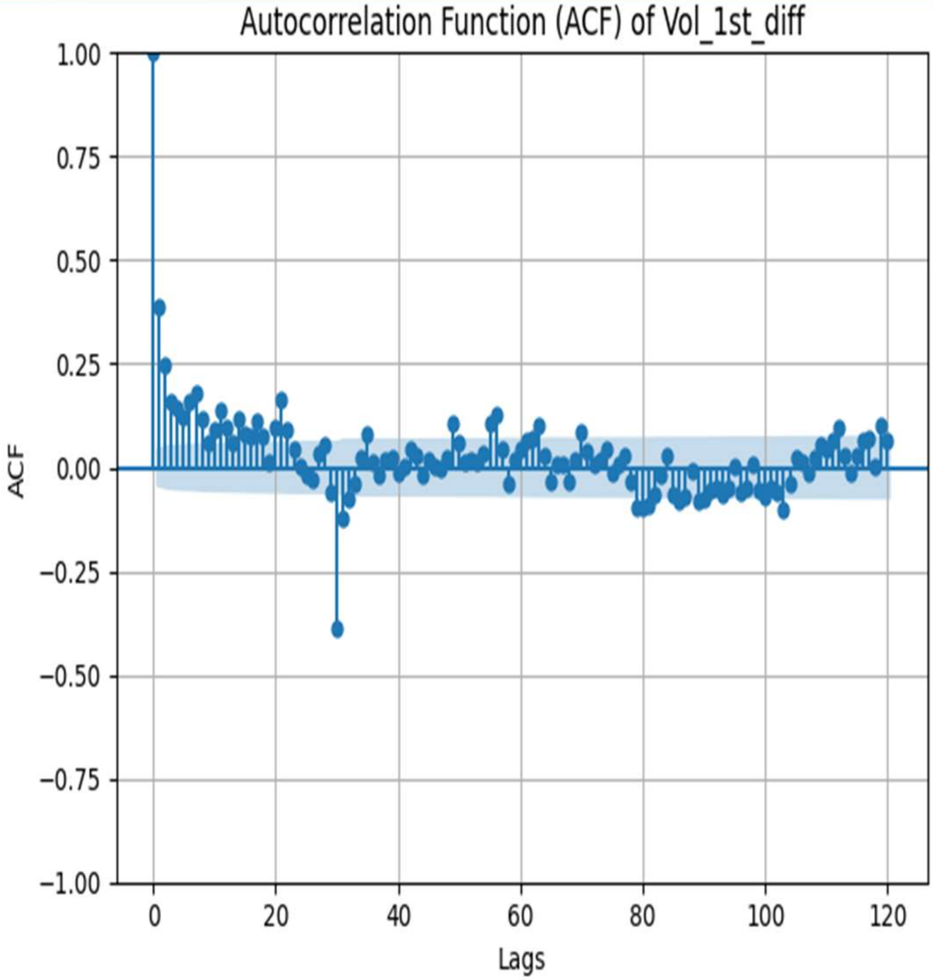
- ❑ ADF Test Results
  - Test statistic: *1.8876*
  - p-value: 8.468424127970887e-09
- ❑ Since, *p-value < 0.05,* the time series is stationary

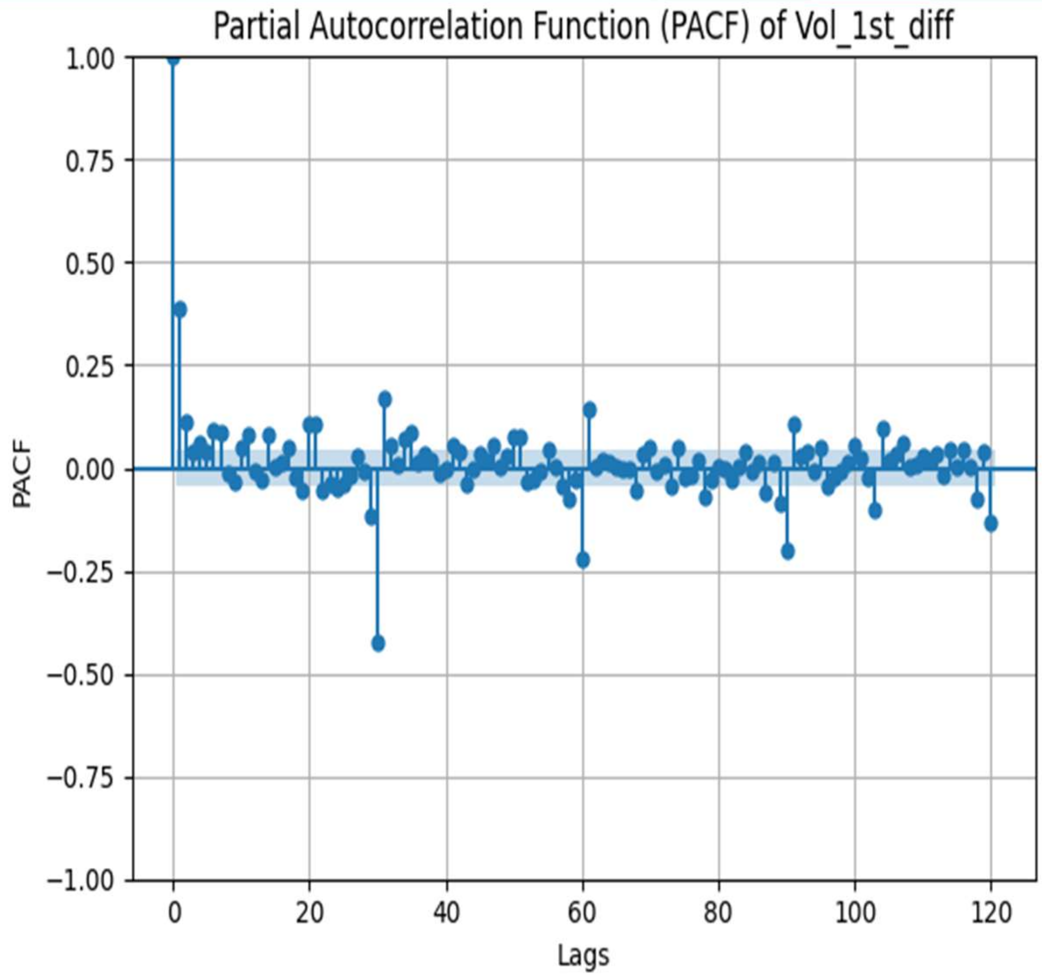
Value

- ❑ ADF Test Results
  - Test statistic: *-0.378*
  - p-value: 7.018026069290158e-11
- ❑ Since, *p-value < 0.05,* the time series is stationary

- Since it is now stationarity we proceed for Model Fitting Steps

ACF & PACF Plot





SARIMA Model Selection For Transaction Value & Volume

Volume

- Fitting various ARIMA model for different (p, d, q) values
- ARIMA (4, 1, 3) model has the least AIC value
- Thus, we concluded ARIMA (4, 1, 3) is the final ARIMA model for Transaction Volume

ARIMA(3,1,4)	-	AIC:	22887.8485
ARIMA(4,1,0)	-	AIC:	23113.4813
ARIMA(4,1,1)	-	AIC:	22924.1044
ARIMA(4,1,2)	-	AIC:	22927.1810
ARIMA(4,1,3)	-	AIC:	22853.6229
ARIMA(4,1,4)	-	AIC:	22888.6324

Value

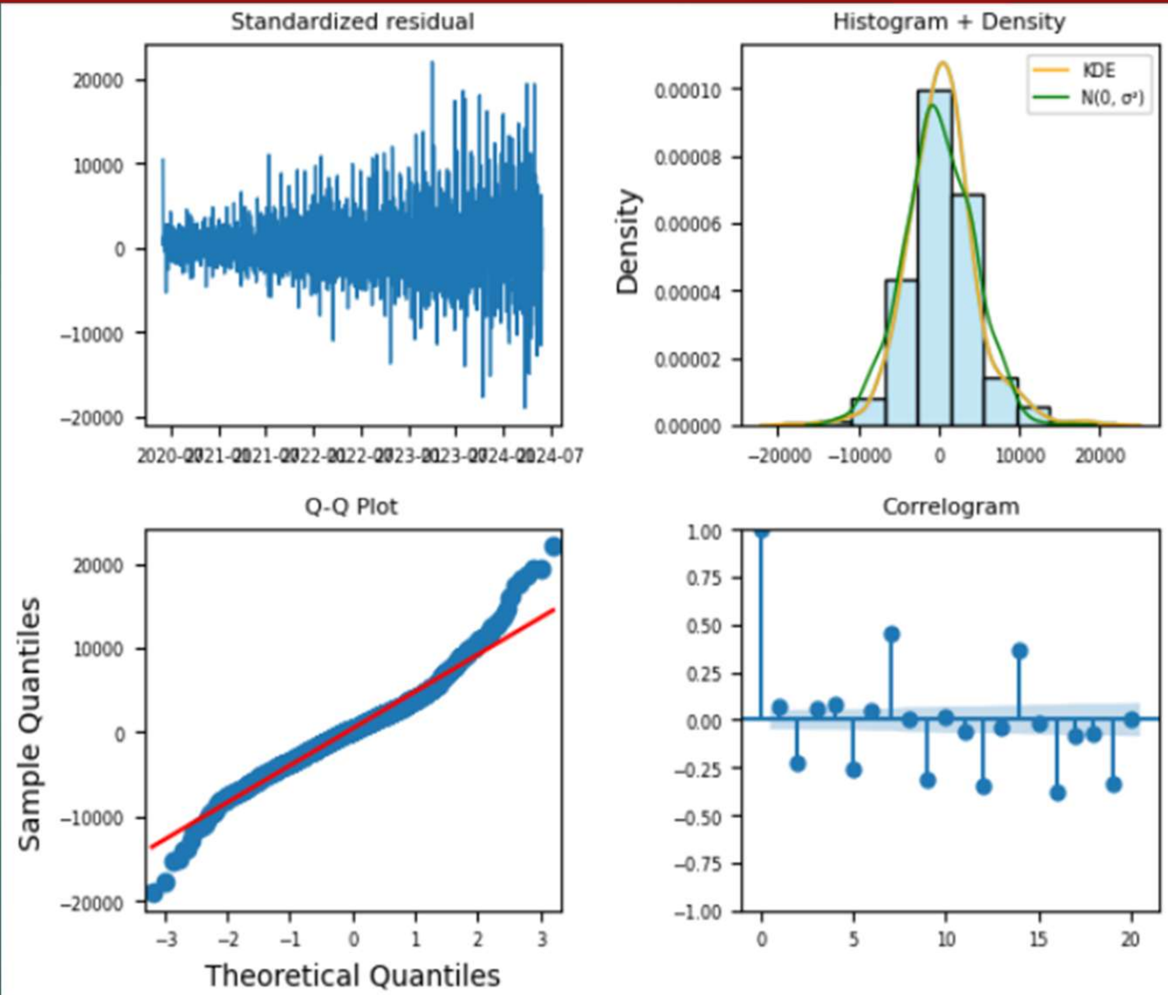
- Fitting various ARIMA model for different (p, d, q) values
- ARIMA (4, 1, 4) model has the least AIC value
- Thus, we concluded ARIMA (4, 1, 4) is the final ARIMA model for Transaction Value

ARIMA(4,1,0)	-	AIC:	37551.4461
ARIMA(4,1,1)	-	AIC:	37019.9099
ARIMA(4,1,2)	-	AIC:	36934.1305
ARIMA(4,1,3)	-	AIC:	36642.9607
ARIMA(4,1,4)	-	AIC:	36630.6818

Model Diagnostics ARIMA (4, 1, 3)

❑ **Remark**:

- According to the qualitative analysis, residuals approximately resemble white noise.
- In order to validate this, let's shift to quantitative analysis, namely Ljung-Box test

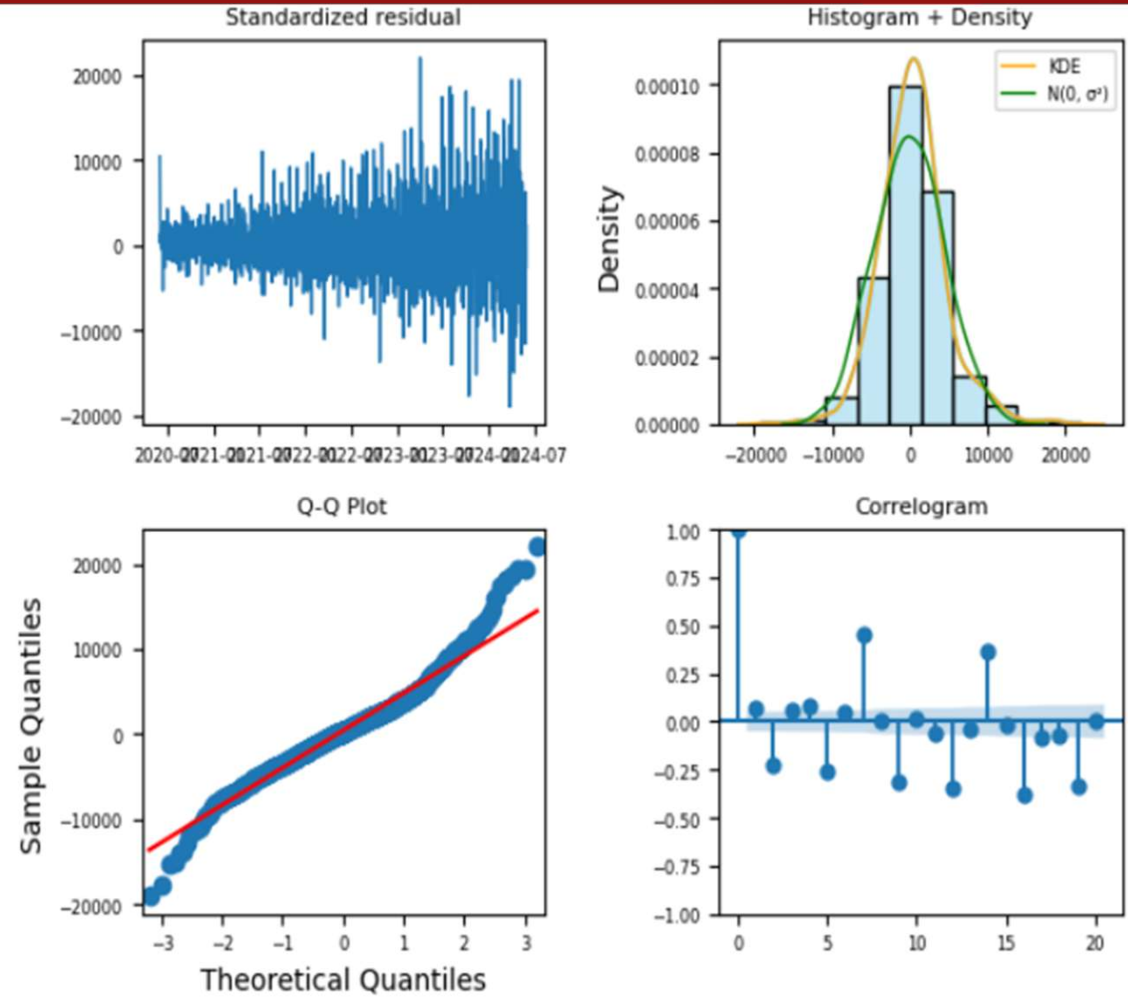


Model Diagnostics ARIMA (4, 1, 3)

❑ **Remark**:

- According to the qualitative analysis, residuals approximately resemble white noise.

- In order to validate this, let's shift to quantitative analysis, namely Ljung-Box test



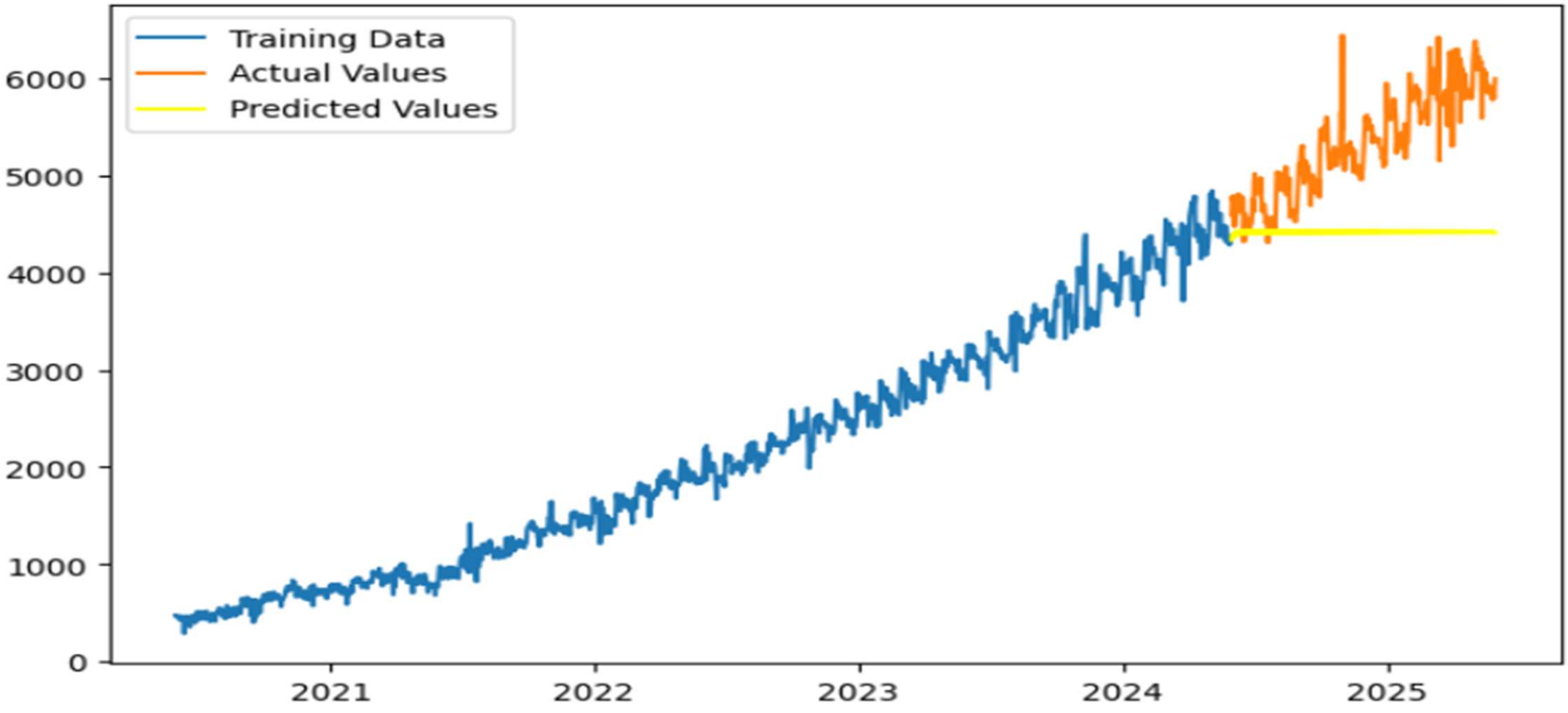
Ljung Box test

❑ **Remark**:

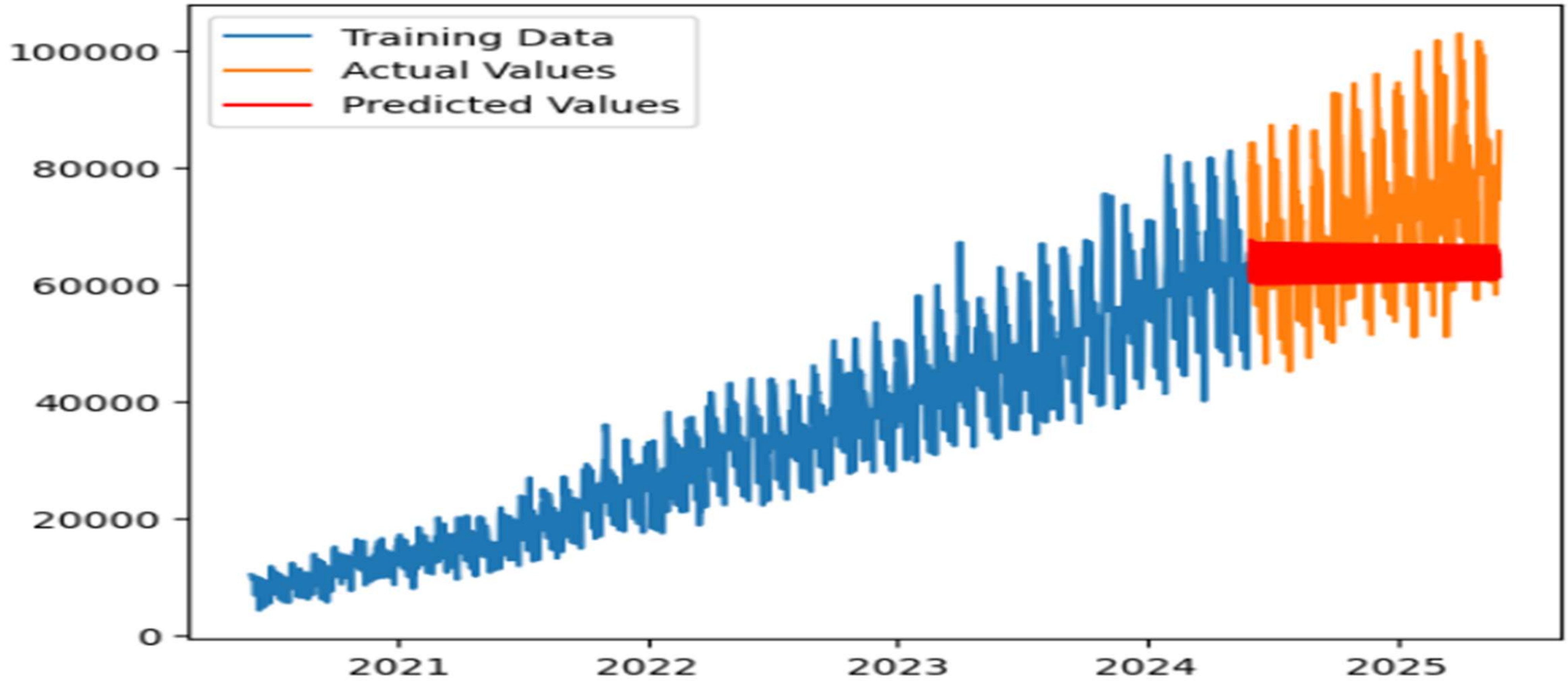
- We have checked Ljung Box test upto 5 lags
- p-values indicates that residuals are normal & not autocorrelated
- We can move towards forecasting using ARIMA (4,1,3)

	lb_stat_vol	lb_pvalue_vol
1	0.059034	0.808029
2	1.187947	0.552129
3	5.041165	0.168808
4	6.308064	0.177293
5	6.421829	0.267309

Forecasting ARIMA(4,1,3) for Volume



Forecasting ARIMA(4,1,4) for Value



A Team Project by:

- Lekhraj Meena (24N0049)
- Sachin Dimri (24N0083)
- Ratnesh Pati Tripathi (24N0065)

Thank you !!