DATA SCIENCE INSTITUTE

WOOLF

# ASSIGNMENT

## TIME SERIES ANALYSIS

Python

# REPORT

Pillar1-e

---


Prepared By

**William C. Phiri**

PG Dip in Data Science

**Student No. 4295158639**



+353-87-3502102

chenechoz@gmail.com

Athlone, IRELAND

**BACKGROUND:**

The data contains monthly sales (in million USD) of a leading processed food manufacturer in USA. Data for 3 business units, from February 2015 to January 2018, is provided in the same file.

Columns – Year, Month, BU1, BU2, BU3

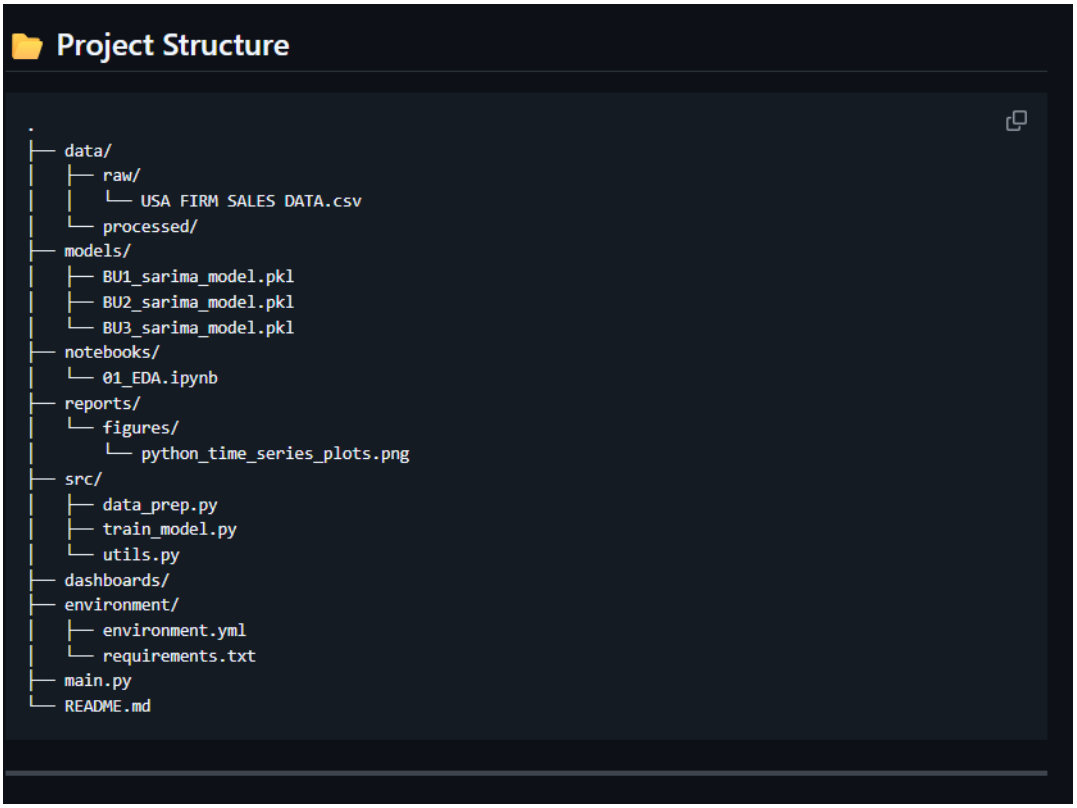
Number of observations – 36 for each business unit

**ASSIGNMENT OBJECTIVE:**

I. Import 'USA Firm Sales' data in R/PYTHON
II. Create time series objects of the data.
III. Check for stationarity for each of the three series.
IV. Obtain best model for each BU.
V. Predict sales for each BU for January 2018, February 2018, March 2018.

OUTPUT: Refer to GitHub repository at LINK HERE for the model and the associated notebooks.

**Project Repository Structure:**



# RESULTS:

**QUESTION 1:** USA Firm Sales' data.

**Data Snapshot**

| ... | Year | Month | BU1 | BU2 | BU3 |
|---|---|---|---|---|---|
| Date | | | | | |
| 2015-02-01 | 2015 | February | 125.1 | 115.5 | 113.8 |
| 2015-03-01 | 2015 | March | 123.6 | 115.7 | 113.8 |
| 2015-04-01 | 2015 | April | 123.1 | 116.5 | 114.0 |
| 2015-05-01 | 2015 | May | 123.1 | 117.7 | 114.1 |
| 2015-06-01 | 2015 | June | 123.4 | 118.3 | 114.7 |

**QUESTION 2:** Create time series objects of the data.

Data snapshot of time series objects

```
BU1 Sales Data:
 Date
2015-02-01    125.1
2015-03-01    123.6
2015-04-01    123.1
2015-05-01    123.1
2015-06-01    123.4
Name: BU1, dtype: float64
```

```
BU2 Sales Data:
 Date
2015-02-01    115.5
2015-03-01    115.7
2015-04-01    116.5
2015-05-01    117.7
2015-06-01    118.3
Name: BU2, dtype: float64
```

```
BU3 Sales Data:
 Date
2015-02-01    113.8
2015-03-01    113.8
2015-04-01    114.0
2015-05-01    114.1
2015-06-01    114.7
Name: BU3, dtype: float64
```

## Data Visualisation of the time Series decomposition.

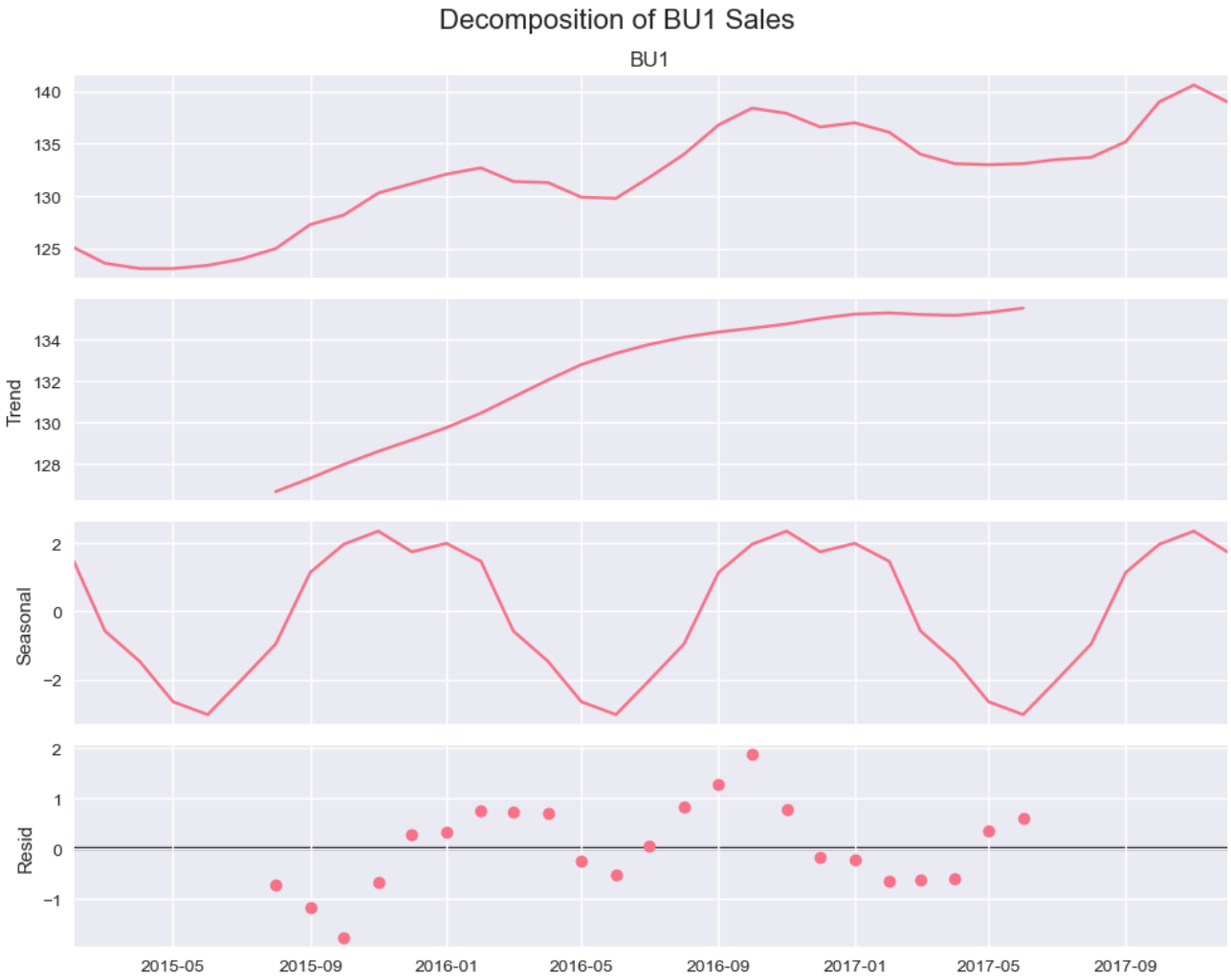
**BU1 sales show a clear upward trend with strong yearly seasonality.**

The decomposition model appears effective since the residuals show no obvious structure.

The trend component shows a gradual upward movement indicating long term growth.

As for the seasonal component it appears to be peaking once a year meaning that it is cyclic, likely caused by seasonal demand pattern.

Such a series is likely suitable for **seasonal ARIMA (SARIMA)**



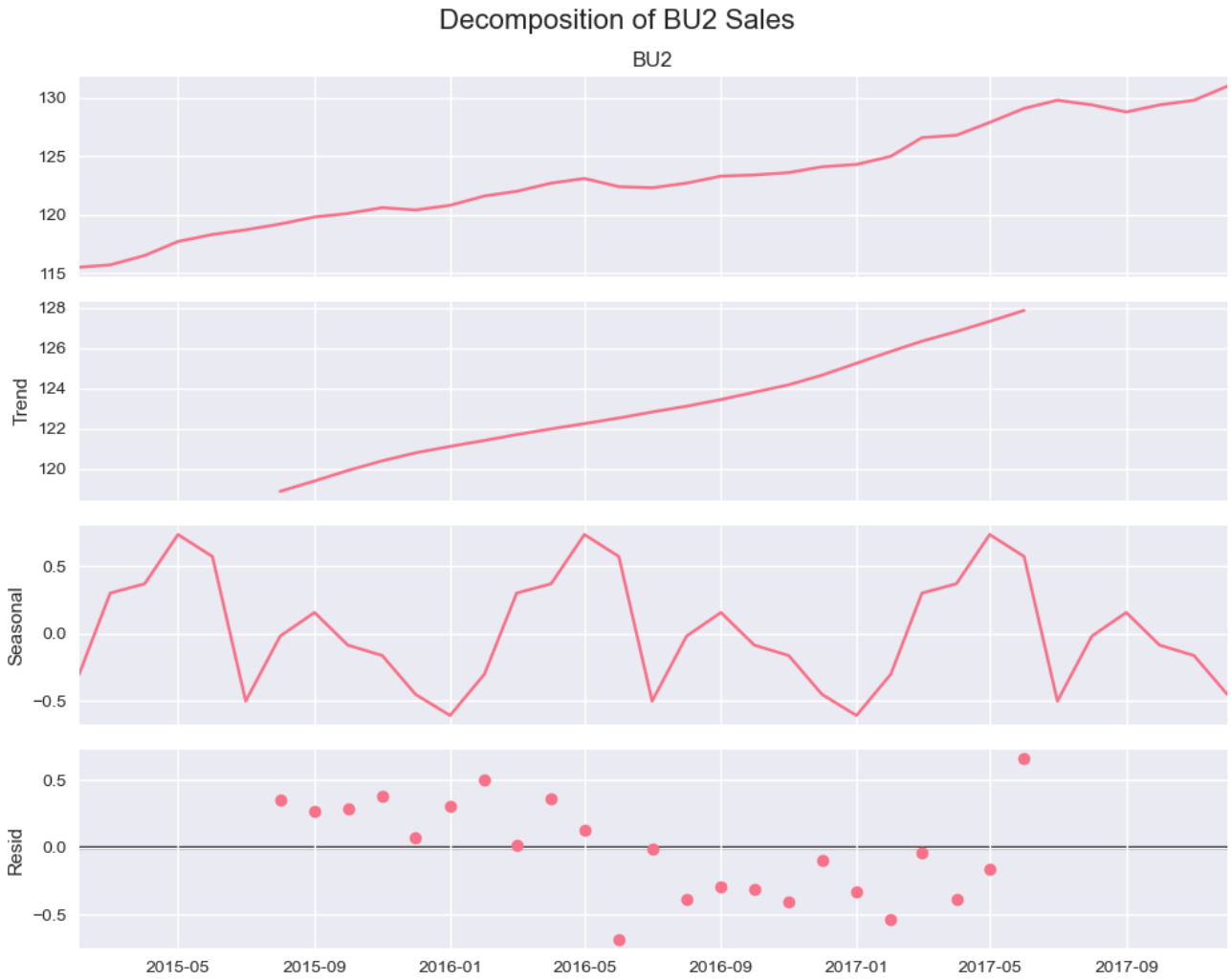
**BU2 sales shows a strong a clear upward trend with weak yearly seasonality.**

The decomposition model appears effective since the residuals show no obvious structure.

The trend component shows a gradual upward movement indicating long term growth.

Very mild seasonality maybe BU2 is less impacted by seasonal demand.

Such a series is likely suitable for **ARIMA** unless further diagnostics (like seasonal ACF spikes) reveal hidden seasonality.



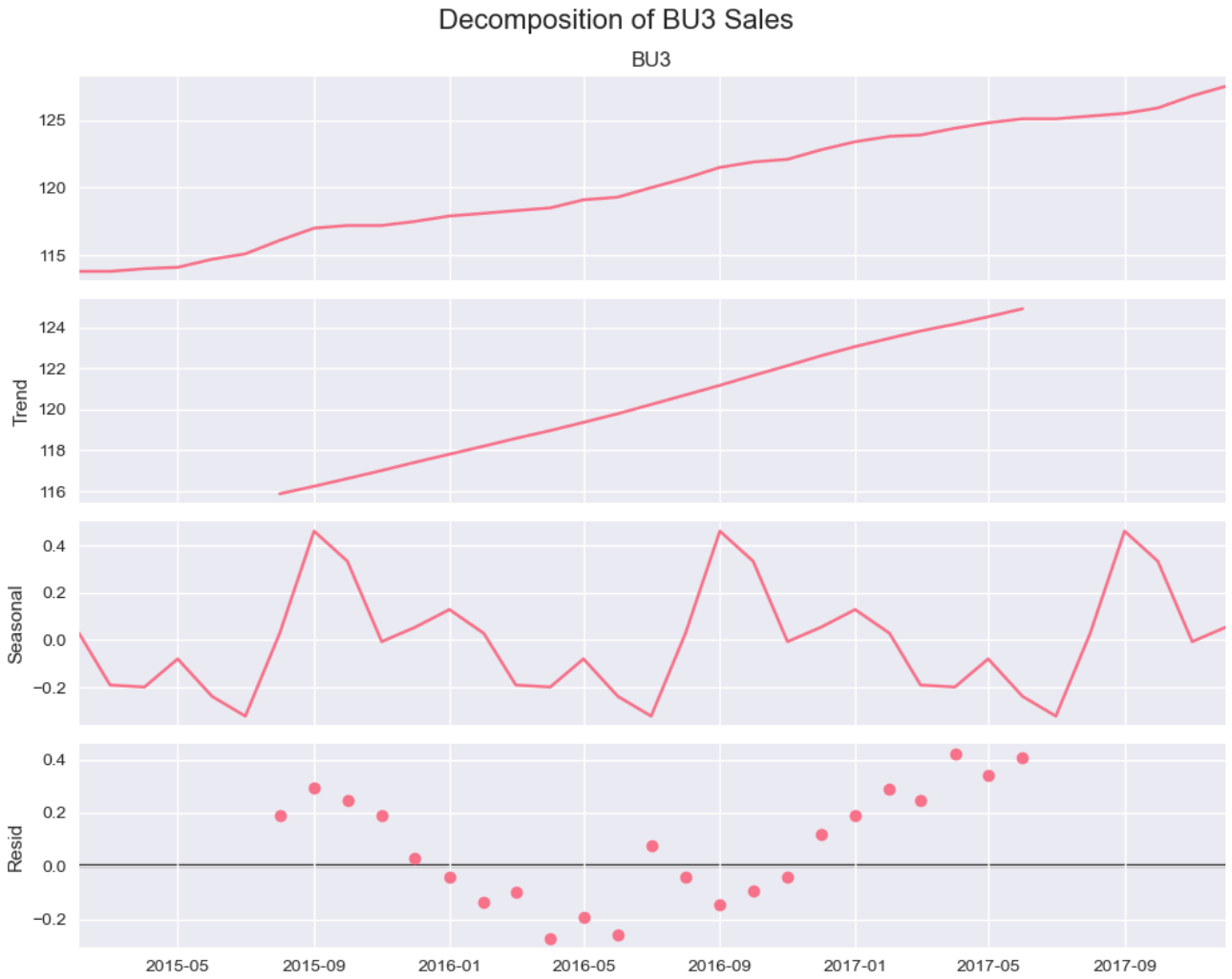
**BU3 sales shows a steady upward trend with low variability.**

The decomposition model appears effective since the residuals show no obvious structure.

The trend component shows a gradual upward movement indicating long term growth.

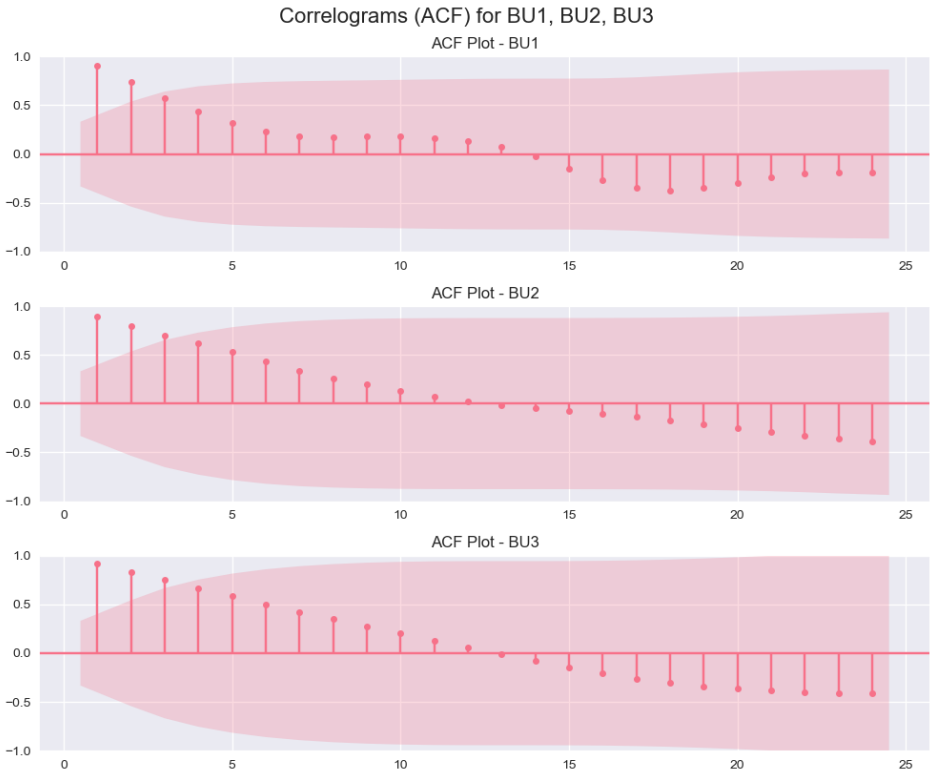
The seasonality is very weak, oscillating between **-0.2 and +0.4** suggesting mild and possibly negligible seasonal effects.

Such a series is likely suitable for **ARIMA** unless further diagnostics (like seasonal ACF spikes) reveal hidden seasonality.



## QUESTION 3: Check for stationarity for each of the three series.

Using Correlograms ACF plots to check for stationarity. From the plot it appears that all the business units have visible slow decay indicating that none of them are stationary.



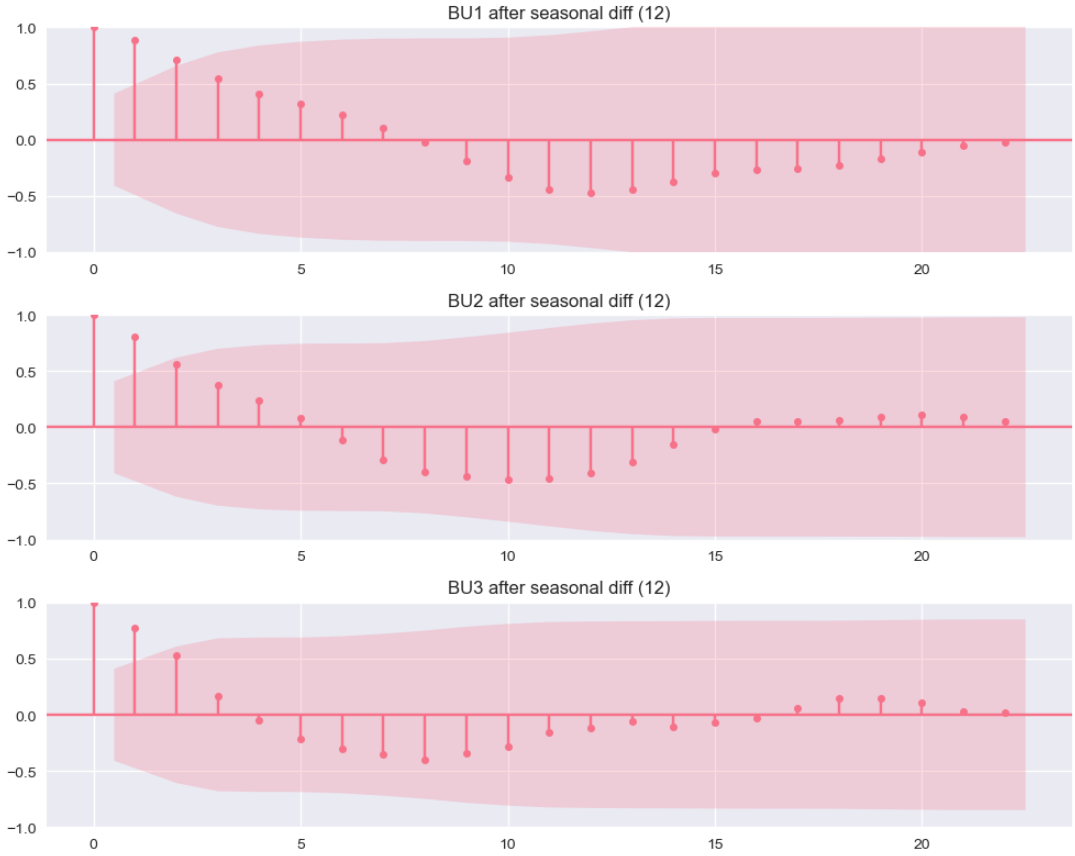

We Ran the augmented Dickey Fuller test and statistical test results revealed the same.

```
📉 Augmented Dickey-Fuller Test Results:

🔹 BU1:
   Test Statistic = -2.4760
   p-value        = 0.1214
   Critical Value (1%) = -3.7377
   Critical Value (5%) = -2.9922
   Critical Value (10%) = -2.6357
   → Non-stationary ❌
----------------------------------------
🔹 BU2:
   Test Statistic = -0.2875
   p-value        = 0.9273
   Critical Value (1%) = -3.6392
   Critical Value (5%) = -2.9512
   Critical Value (10%) = -2.6144
   → Non-stationary ❌
----------------------------------------
🔹 BU3:
   Test Statistic = -0.1352
   p-value        = 0.9458
   Critical Value (1%) = -3.6461
   Critical Value (5%) = -2.9541
   Critical Value (10%) = -2.6160
   → Non-stationary ❌
----------------------------------------
```

## Stationarity Confirmation:

No more slow decay indicating that stationarity has been achieved. We then went on to use auto-arima to determine the best model for the business units.




```
🔧 Training ARIMA/SARIMA model for BU1...
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(1,1,1)[12]             : AIC=inf, Time=0.41 sec
 ARIMA(0,1,0)(0,1,0)[12]             : AIC=78.215, Time=0.02 sec
 ARIMA(1,1,0)(1,1,0)[12]             : AIC=76.714, Time=0.03 sec
 ARIMA(0,1,1)(0,1,1)[12]             : AIC=74.500, Time=0.06 sec
 ARIMA(0,1,1)(0,1,0)[12]             : AIC=73.767, Time=0.03 sec
 ARIMA(0,1,1)(1,1,0)[12]             : AIC=74.500, Time=0.05 sec
 ARIMA(0,1,1)(1,1,1)[12]             : AIC=76.500, Time=0.06 sec
 ARIMA(1,1,1)(0,1,0)[12]             : AIC=inf, Time=0.07 sec
 ARIMA(0,1,2)(0,1,0)[12]             : AIC=inf, Time=0.05 sec
 ARIMA(1,1,0)(0,1,0)[12]             : AIC=76.954, Time=0.01 sec
 ARIMA(1,1,2)(0,1,0)[12]             : AIC=inf, Time=0.10 sec
 ARIMA(0,1,1)(0,1,0)[12] intercept   : AIC=75.364, Time=0.03 sec

Best model:  ARIMA(0,1,1)(0,1,0)[12]
Total fit time: 0.940 seconds
✅ Best ARIMA/SARIMA model for BU1: (0, 1, 1) x (0, 1, 0, 12)
```

```
🔧 Training ARIMA/SARIMA model for BU2...
Performing stepwise search to minimize aic
 ARIMA(2,0,2)(1,1,1)[12] intercept   : AIC=inf, Time=0.48 sec
 ARIMA(0,0,0)(0,1,0)[12] intercept   : AIC=84.447, Time=0.02 sec
 ARIMA(1,0,0)(1,1,0)[12] intercept   : AIC=56.347, Time=0.09 sec
 ARIMA(0,0,1)(0,1,1)[12] intercept   : AIC=inf, Time=0.15 sec
 ARIMA(0,0,0)(0,1,0)[12]             : AIC=142.871, Time=0.00 sec
 ARIMA(1,0,0)(0,1,0)[12] intercept   : AIC=57.858, Time=0.05 sec
 ARIMA(1,0,0)(2,1,0)[12] intercept   : AIC=58.351, Time=0.14 sec
 ARIMA(1,0,0)(1,1,1)[12] intercept   : AIC=inf, Time=0.28 sec
 ARIMA(1,0,0)(0,1,1)[12] intercept   : AIC=inf, Time=0.17 sec
 ARIMA(1,0,0)(2,1,1)[12] intercept   : AIC=inf, Time=0.43 sec
 ARIMA(0,0,0)(1,1,0)[12] intercept   : AIC=70.887, Time=0.08 sec
 ARIMA(2,0,0)(1,1,0)[12] intercept   : AIC=56.892, Time=0.09 sec
 ARIMA(1,0,1)(1,1,0)[12] intercept   : AIC=56.360, Time=0.12 sec
 ARIMA(0,0,1)(1,1,0)[12] intercept   : AIC=inf, Time=0.15 sec
 ARIMA(2,0,1)(1,1,0)[12] intercept   : AIC=inf, Time=0.28 sec
 ARIMA(1,0,0)(1,1,0)[12]             : AIC=inf, Time=0.07 sec

Best model:  ARIMA(1,0,0)(1,1,0)[12] intercept
Total fit time: 2.634 seconds
✅ Best ARIMA/SARIMA model for BU2: (1, 0, 0) x (1, 1, 0, 12)
```

```
🔧 Training ARIMA/SARIMA model for BU3...
Performing stepwise search to minimize aic
 ARIMA(2,0,2)(1,1,1)[12] intercept   : AIC=inf, Time=0.35 sec
 ARIMA(0,0,0)(0,1,0)[12] intercept   : AIC=41.888, Time=0.03 sec
 ARIMA(1,0,0)(1,1,0)[12] intercept   : AIC=23.058, Time=0.12 sec
 ARIMA(0,0,1)(0,1,1)[12] intercept   : AIC=30.627, Time=0.08 sec
 ARIMA(0,0,0)(0,1,0)[12]             : AIC=140.826, Time=0.01 sec
 ARIMA(1,0,0)(0,1,0)[12] intercept   : AIC=21.191, Time=0.06 sec
 ARIMA(1,0,0)(0,1,1)[12] intercept   : AIC=23.059, Time=0.10 sec
 ARIMA(1,0,0)(1,1,1)[12] intercept   : AIC=25.052, Time=0.25 sec
 ARIMA(2,0,0)(0,1,0)[12] intercept   : AIC=22.158, Time=0.08 sec
 ARIMA(1,0,1)(0,1,0)[12] intercept   : AIC=22.688, Time=0.07 sec
 ARIMA(0,0,1)(0,1,0)[12] intercept   : AIC=30.997, Time=0.04 sec
 ARIMA(2,0,1)(0,1,0)[12] intercept   : AIC=25.163, Time=0.16 sec
 ARIMA(1,0,0)(0,1,0)[12]             : AIC=inf, Time=0.07 sec

Best model:  ARIMA(1,0,0)(0,1,0)[12] intercept
Total fit time: 1.422 seconds
✅ Best ARIMA/SARIMA model for BU3: (1, 0, 0) x (0, 1, 0, 12)
```

**Overall Model Quality**

- **BU1**: Forecast reflects seasonal behavior reflecting a good use of SARIMA.
- **BU2 & BU3**: Forecasts continue steady upward trends which appropriate use of ARIMA.
- Forecasts are **visually reasonable** and **smoothly extend** each series without dramatic deviation.