

Time Series Data Analysis and Forecasting

2023 11 10

Sejin Chun



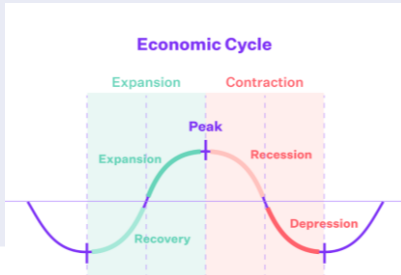
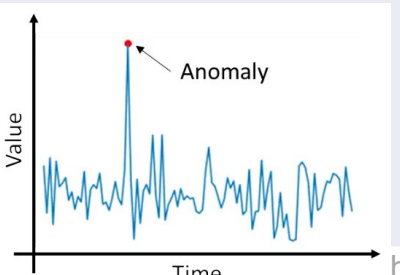
Contents

- What is Time Series Analysis?
- How to Analyze Time Series

What is Time Series Analysis?

- a specific way of analyzing a sequence of data points collected over time
- record data points at consistent intervals over a set period, rather than just recording the data points intermittently or randomly

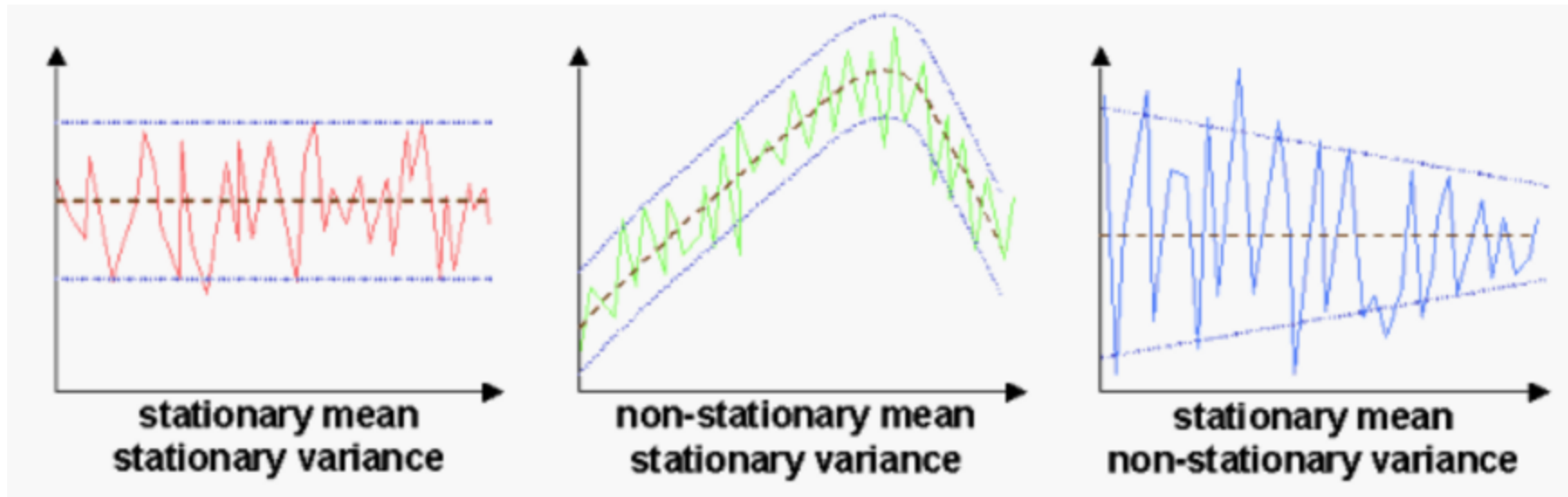
Components

	Trend	Seasonality	Cyclical	Irregularity
Time	Fixed Time Interval	Fixed Time Interval	Not Fixed Time Interval	Not Fixed Time Interval
Duration	Long and Short Term	Short Term	Long and Short Term	Regular/Irregular
Nature – I	Gradual	Swings between Up or Down	Repeating Up and Down	Errored or High Fluctuation
Nature – II	Upward/Down Trend	Pattern repeatable	No fixed period	Short and Not repeatable
Prediction Capability	Predictable	Predictable	Challenging	Challenging
Market Model				

Components

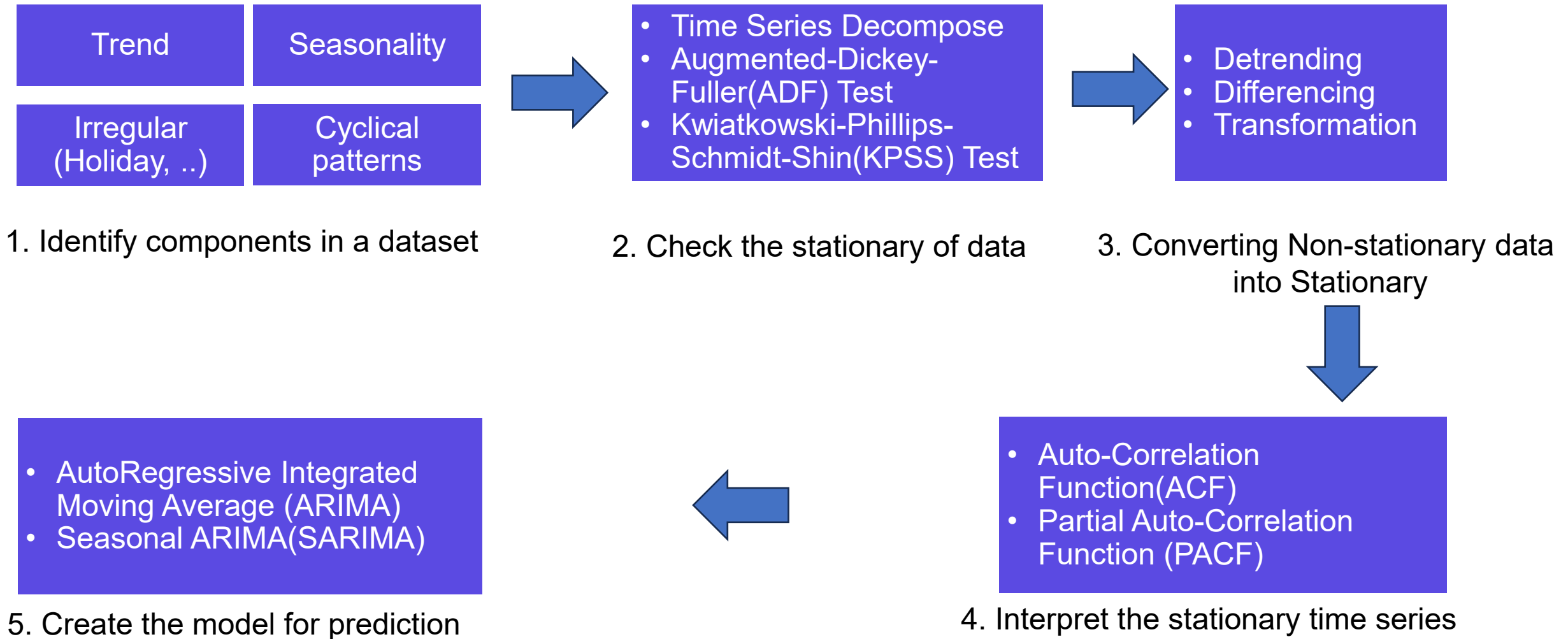
- **Trend:** 고정된 간격이 없고 주어진 데이터 세트 내의 모든 차이가 연속적인 타임라인. 추세는 부정적이거나 긍정적이거나 Null 추세임
- **Seasonality:** 연속적인 타임라인의 데이터 세트 내에서 정기적 또는 고정 간격이 이동합니다. 주로, 종형 곡선이거나 톱니 모양
- **Cyclical:** 고정된 간격이 없고 움직임과 패턴의 불확실성이 있는 경우
- **Irregularity:** 예상치 못한 상황/사건/시나리오가 단기간에 급증

Stationary



How to Analyze Time Series

Overview: Process Flows



Check the stationary of data

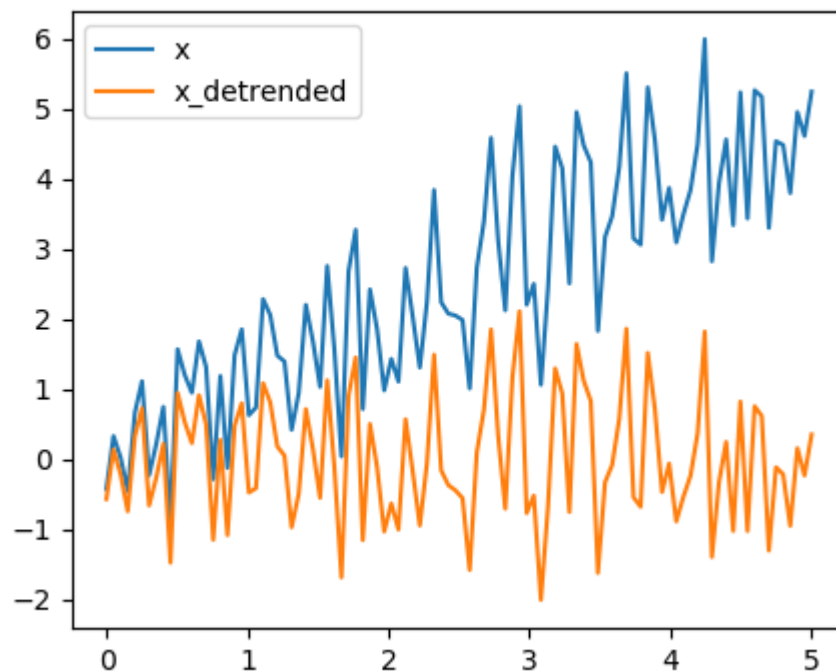
ADF Test

- Null Hypothesis (H_0): Series is non-stationary
- Alternate Hypothesis (H_A): Series is stationary
 - p-value > 0.05 Fail to reject (H_0)
 - p-value ≤ 0.05 Accept (H_1)

**Convert non-stationary
into stationary data**

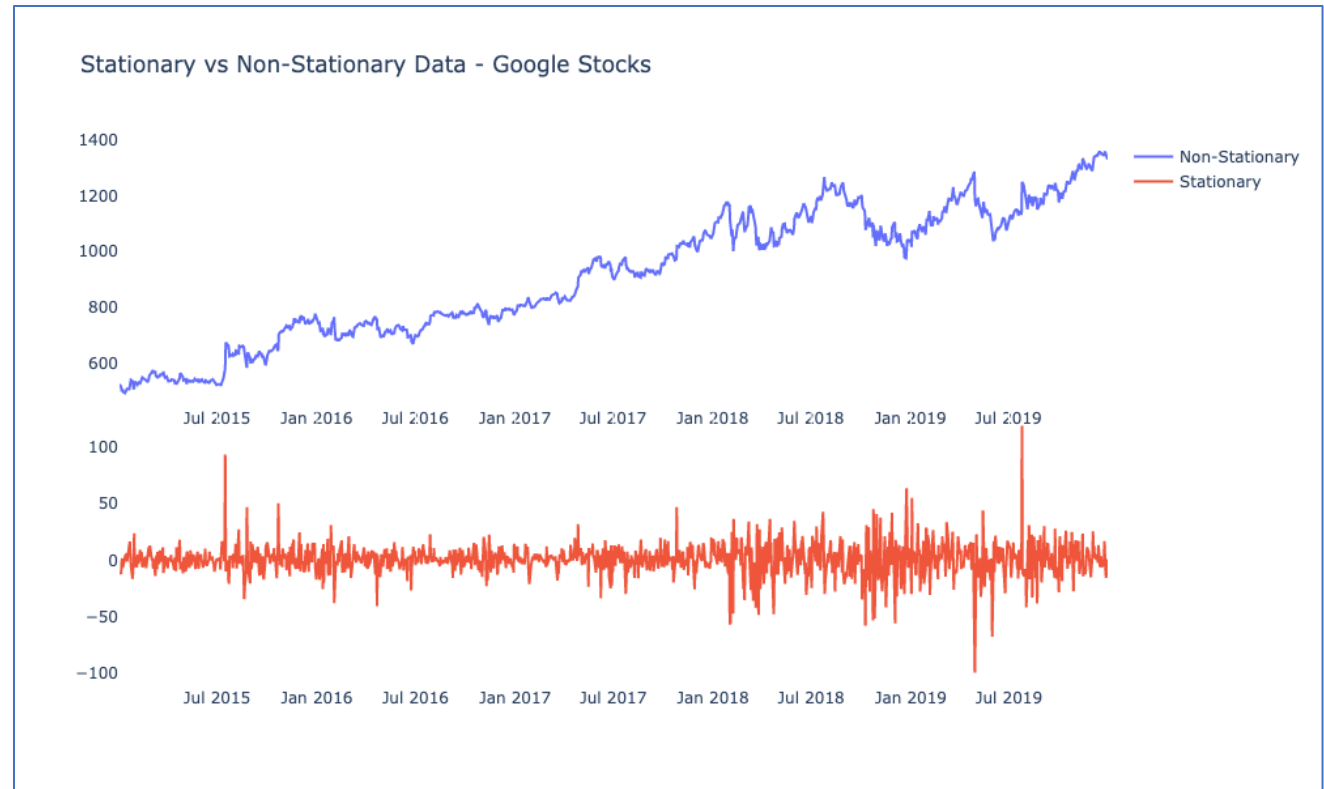
Detrending

- 데이터셋으로 trend effect를 제거



Differencing

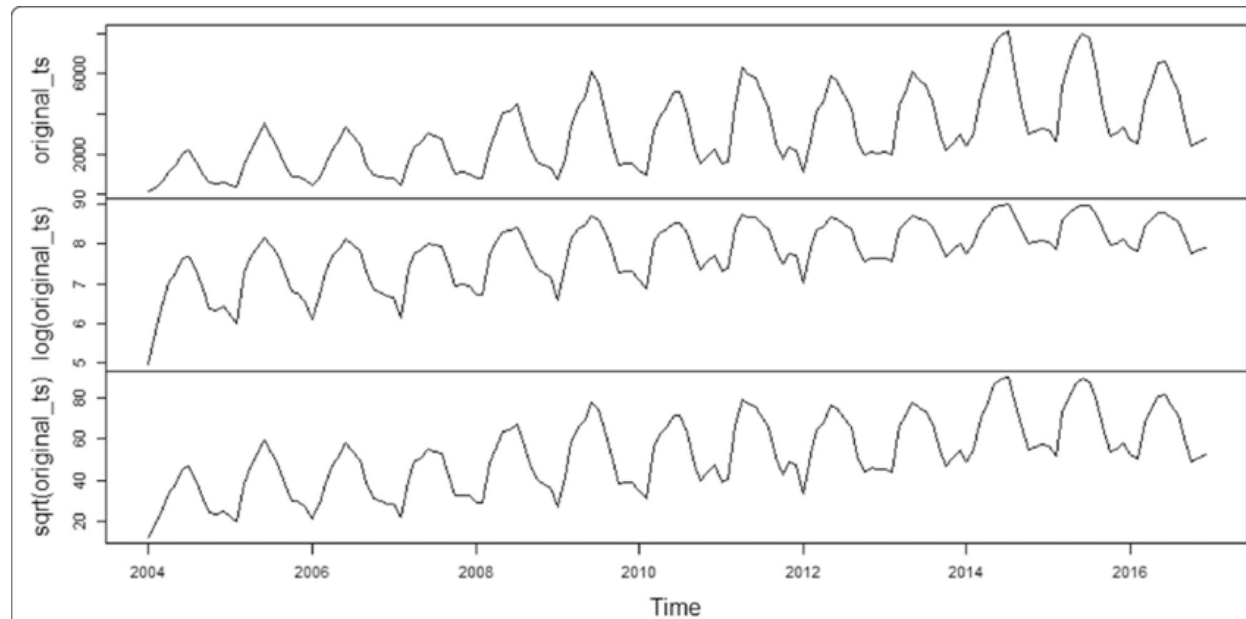
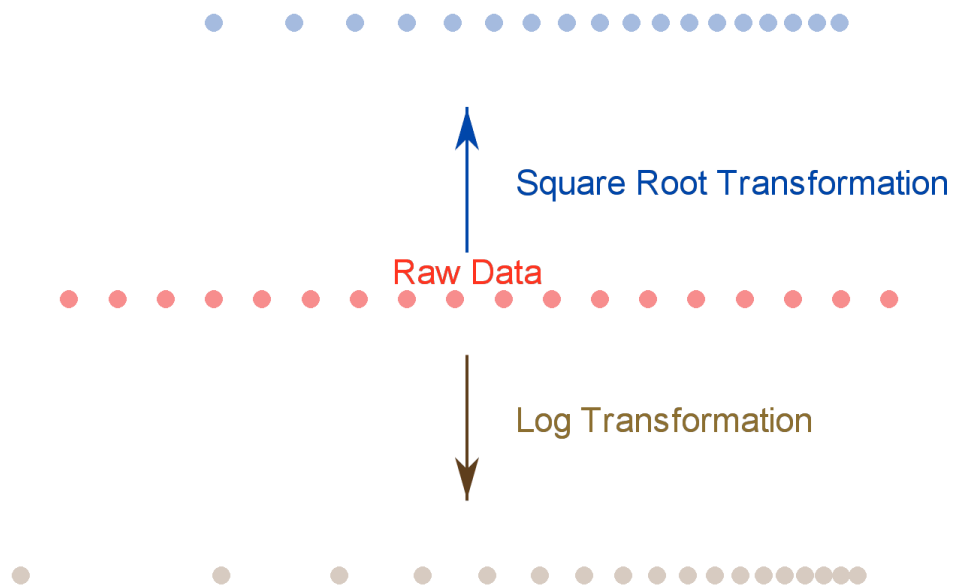
- Series을 새로운 Series 로 간단하게 변환
- 시간에 대한 계열 의존성 (series dependence on time)을 제거
- 시계열의 평균(mean)을 안정화 하여 trend 와 seasonality가 감소



$$\text{difference}(t) = \text{observation}(t) - \text{observation}(t-1)$$

Transform

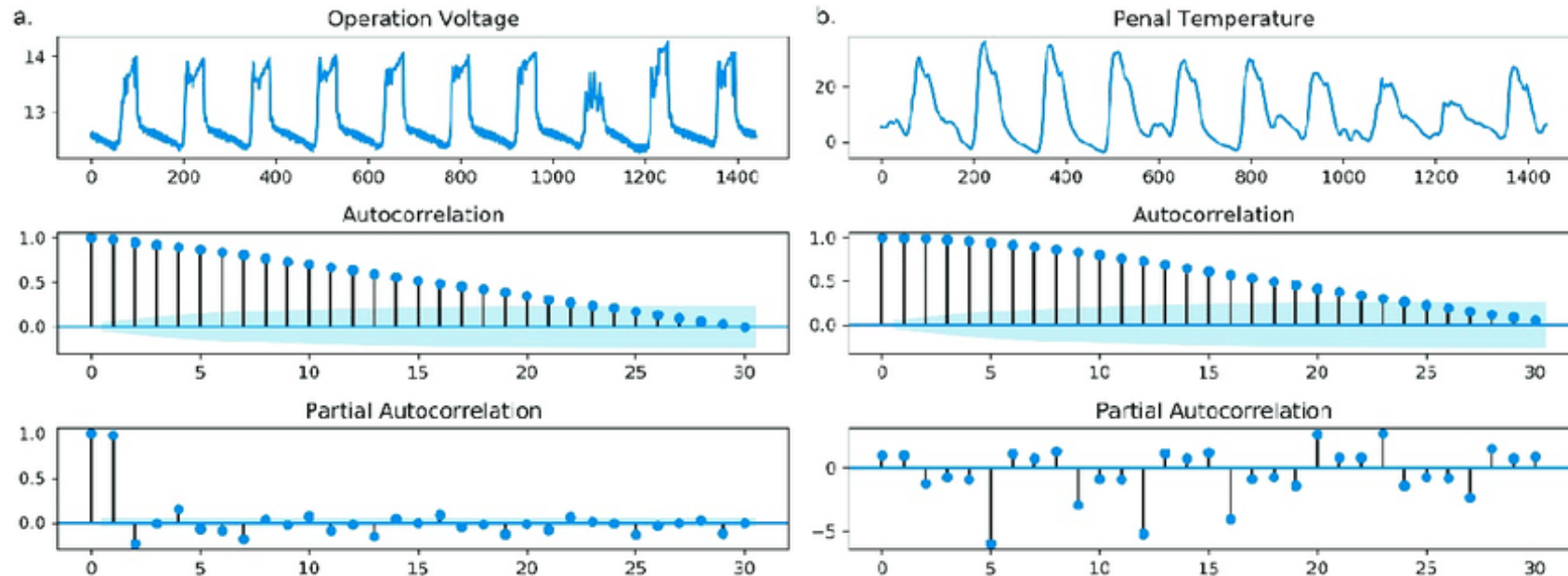
- Square Root and Log Transfer



Interpret the stationary time series

Auto Correlation Function(ACF)

- 시계열 내에서 값과 이전 값이 얼마나 유사한지를 측정
- 다양한 간격에서 주어진 시계열과 해당 시계열의 지연 버전 (lagged version) 간의 유사성 정도를 측정합니다.



Auto Correlation Function(ACF)

- 시계열 내에서 값과 이전 값이 얼마나 유사한지를 측정
- 다양한 간격에서 주어진 시계열과 해당 시계열의 지연 버전 (lagged version) 간의 유사성 정도를 측정합니다.

ACF	PACF	Perfect ML -Model
Plot declines gradually	Plot drops instantly	Auto Regressive model.
Plot drops instantly	Plot declines gradually	Moving Average model
Plot decline gradually	Plot Decline gradually	ARMA
Plot drop instantly	Plot drop instantly	You wouldn't perform any model