Time Series Forecasting

1.2 Time Series Analysis Concepts

Mario Tormo Romero

Design IT.
Create Knowledge.

www.hpi.de

KI Service Zentrum by Hasso-Plattner-Institut
Gefördert durch:
Bundesministerium für Forschung, Technologie und Raumfahrt

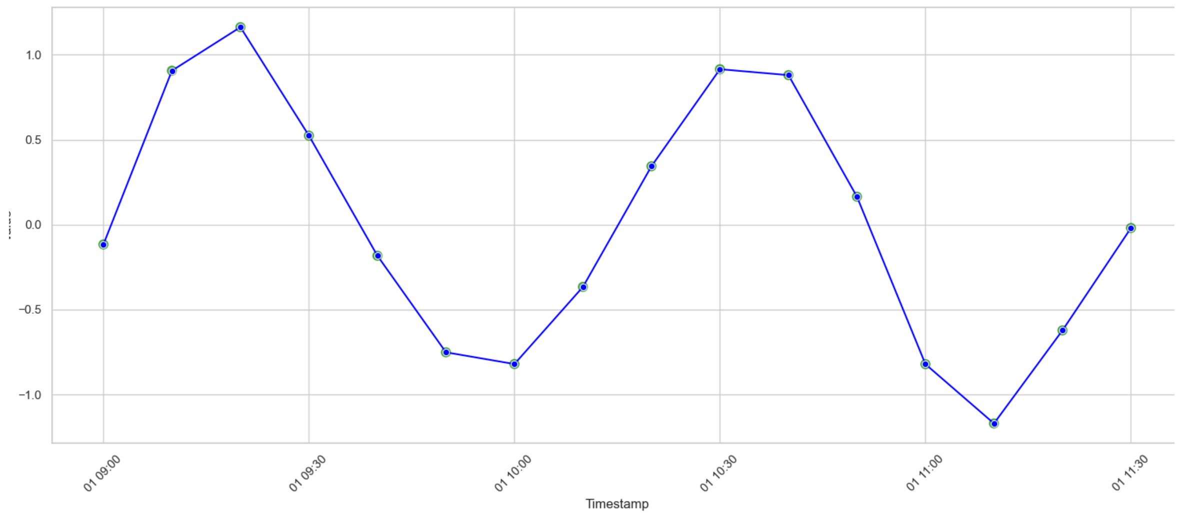
What we`ll cover in this video

- Types of Time Series Data
- Core Concepts:
  - Seasonality and Cyclical signals
  - Trends
  - Autoregressive signals
  - Stationarity and Unit roots
  - Endogenous vs. Exogenous Variables

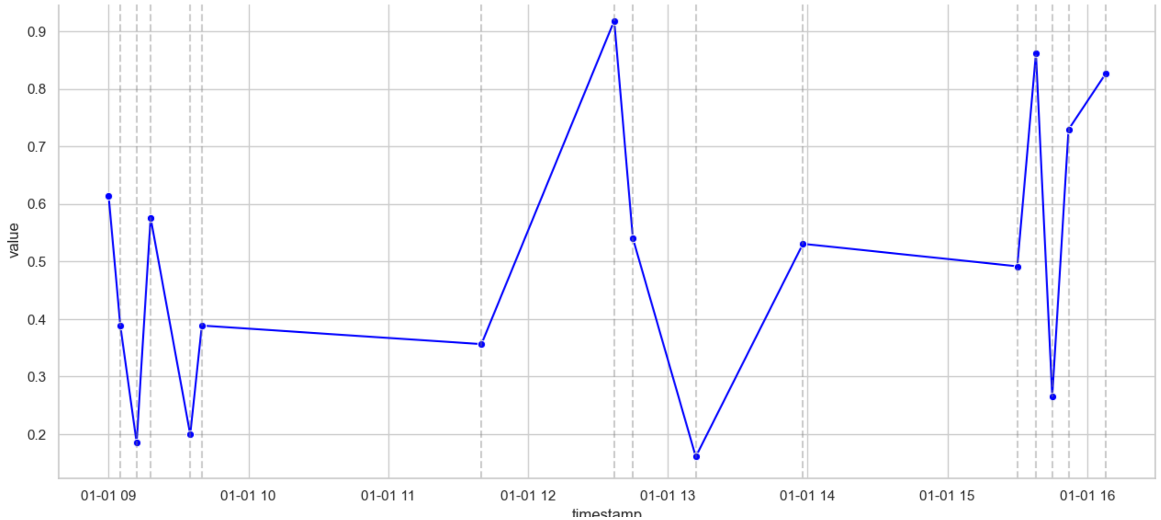
Types of Time Series Data

Regular vs. Irregular Time Series

- **Regular Time Series** (*Evenly spaced observations*)
    - Observations are collected at consistent intervals
    - Easier to model and align with standard forecasting techniques
    - Common in:
        - Weather monitoring
        - Energy consumption tracking
        - Financial time series (e.g., daily stock prices)
    - Uniform timestamps



- **Irregular Time Series** (*Unevenly spaced observations*)
    - Observations occur at uneven or event-driven intervals
    - May require resampling or special models
    - Common in:
        - Medical events (e.g., when symptoms occur)
        - System logs and alerts
        - User activity data
    - Variable time gaps



Different Types of Time Series

Univariate vs. Multivariate | Simple vs. Multiple

	Simple	**Multiple**
Univariate	One variable, one entity over time. *E.g.:* Sales of one product	One variable, many entities. *E.g.:* Sales of many products over time
Multivariate	Several **related** variables for one entity. *E.g.:* Heart rate & blood pressure of one patient	Several **related** variables for many entities. *E.g.:* Vitals for multiple patients

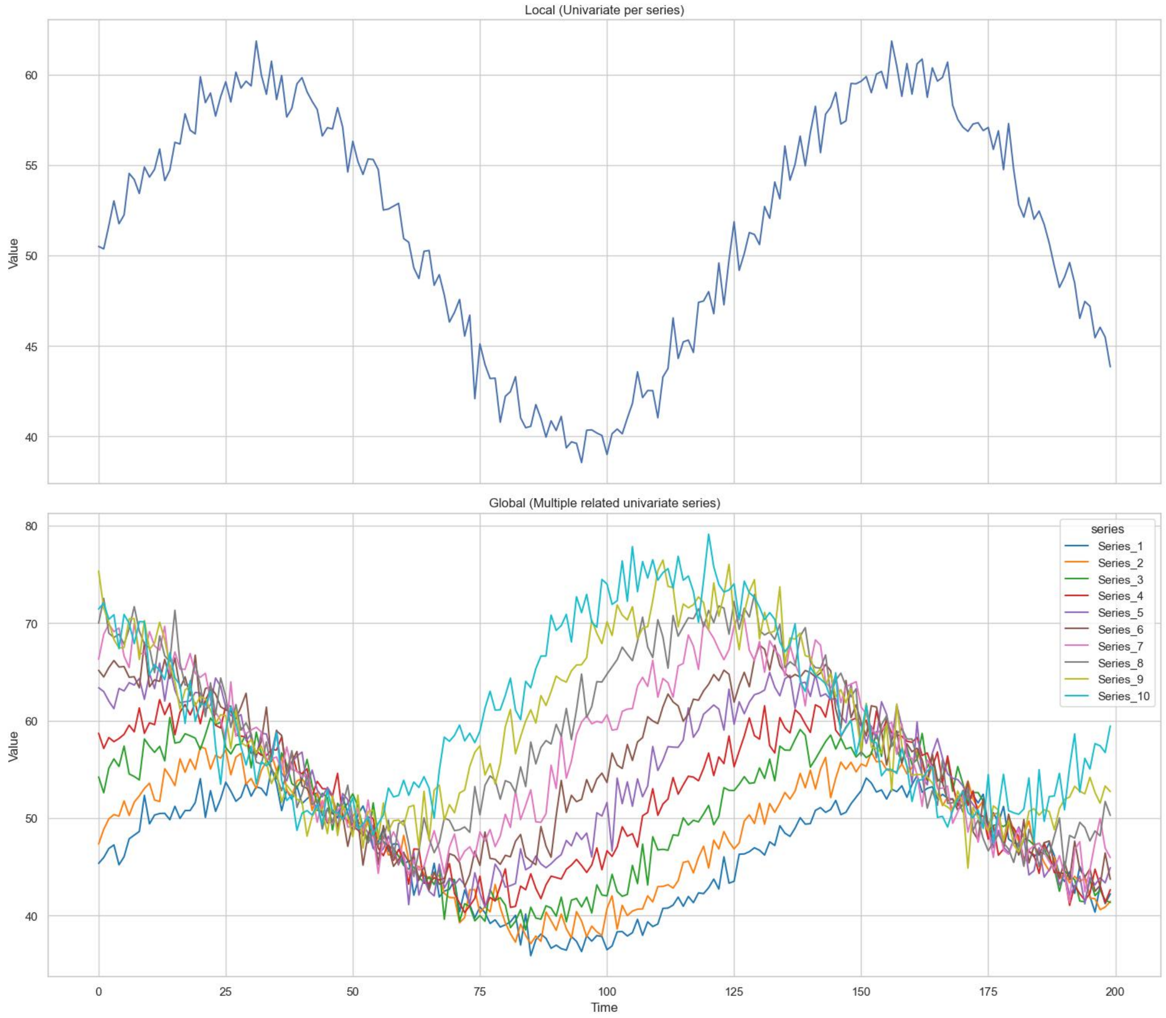
Important: In multivariate time series, the variables are usually **interdependent** — meaning they influence or respond to each other.
If variables are unrelated and analyzed separately, it's better treated as **multiple univariate series**, even if recorded together.

Types of Time Series Data

Local vs Global Time Series

- **Local Forecasting**
  - Focuses on individual time series (usually univariate)
  - Models trained separately for each series
  - Captures specific local patterns and behaviors
  - Examples: forecasting sales of a single product, weather for one city

- **Global Forecasting**
  - Uses one model across multiple related series (multiple univariate series)
  - Learns shared patterns and relationships across series
  - Helps improve accuracy by leveraging cross-series information
  - Examples: forecasting demand for multiple stores, energy use across regions

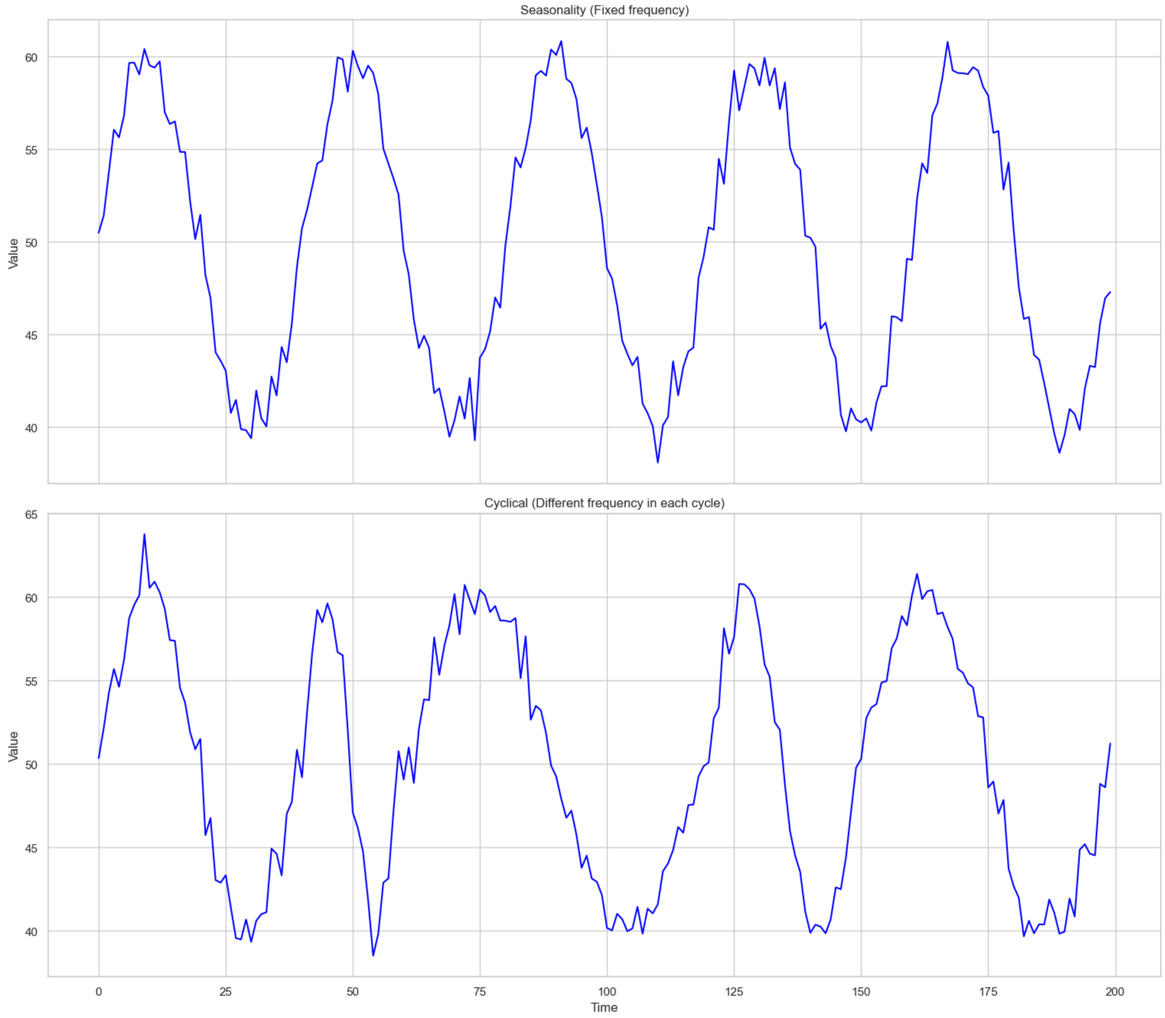


Core concepts

Seasonal and cyclical signals

- **Seasonality**
  - Regular, repeating patterns that occur at fixed, known intervals.
  - Examples:
    - Retail sales spikes every December.
    - Electricity usage peaking daily in the evenings.
  - Key Feature: Frequency is predictable, tied to calendar or time intervals.

- **Cyclical signals**
  - Fluctuations that repeat over longer, irregular periods, often driven by economic or external factors.
  - Examples:
    - Economic expansions and recessions.
    - Commodity price cycles.
  - Key Feature: No fixed period; cycles can vary in length and intensity.



Core concepts

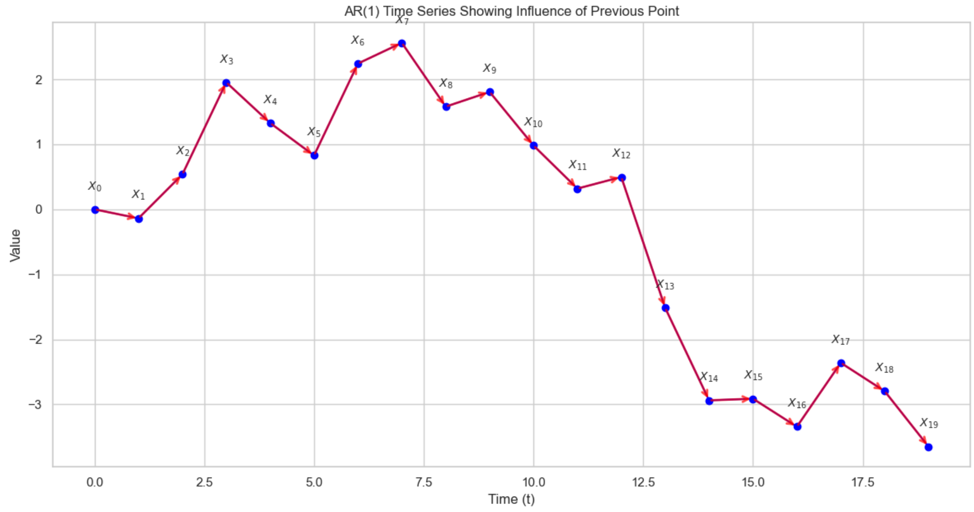
Autoregressive (AR) Signals

- The current value of a time series depends on its own previous values.

- Example:
  - AR(1) model — today's value depends on yesterday's value plus some noise.
  - Stock prices or temperature readings where past behavior influences the present.

- Key Feature: Captures persistence or momentum in data, making past observations predictive of future ones.

$$X_t = \phi X_{t-1} + \varepsilon_t$$



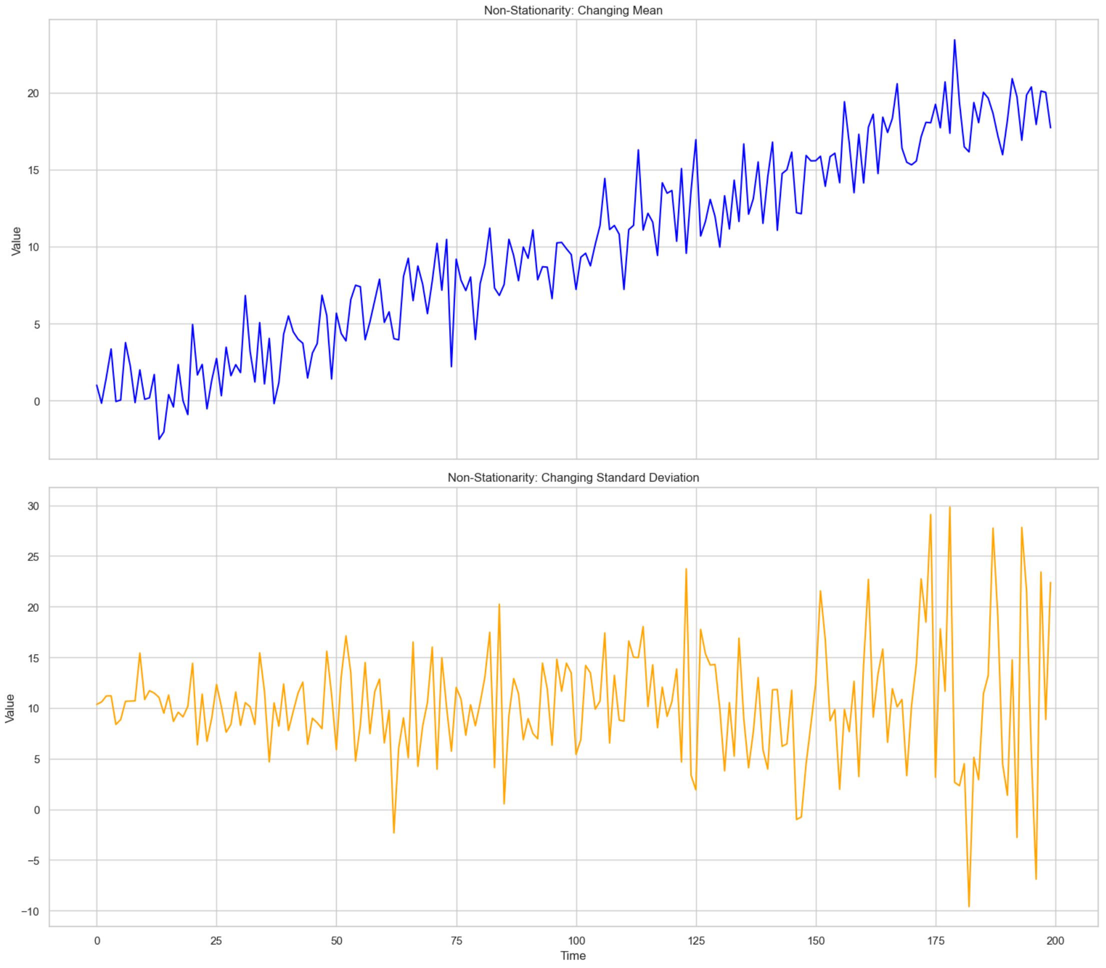
Core concepts

Stationarity

- A time series is **stationary** if its key statistical properties—mean, variance, and autocorrelation—stay constant over time.

- Stationarity means the data fluctuates around a constant average without long-term trends or changing volatility.

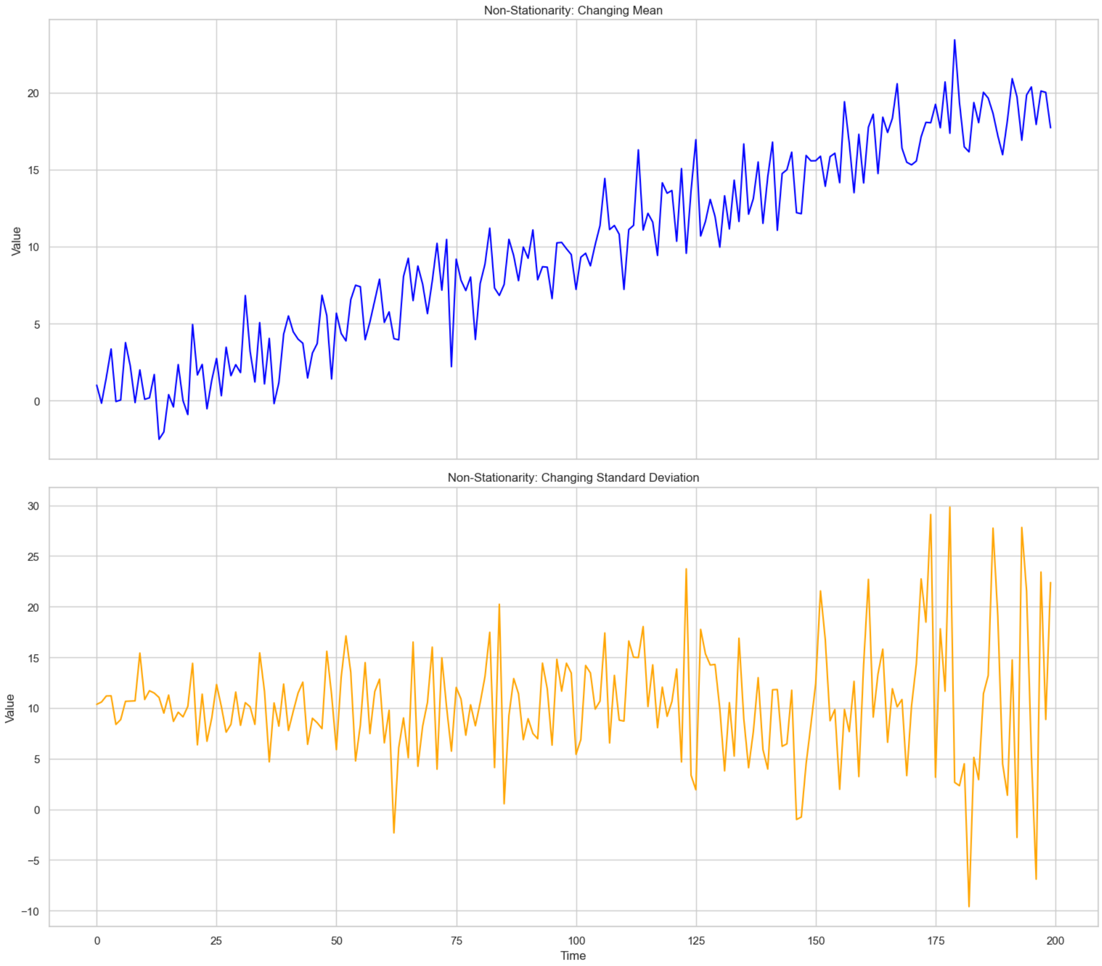
- Types of Stationarity:
  - **Hard (Strict) Stationarity**: The entire distribution of the series remains the same over time, meaning all statistical properties are constant. This is a strong condition and harder to satisfy.
  - **Soft (Weak) Stationarity**: Only the first two moments (mean and variance) and autocorrelation structure are constant over time. This is more common and sufficient for many models.



Core concepts

Stationarity

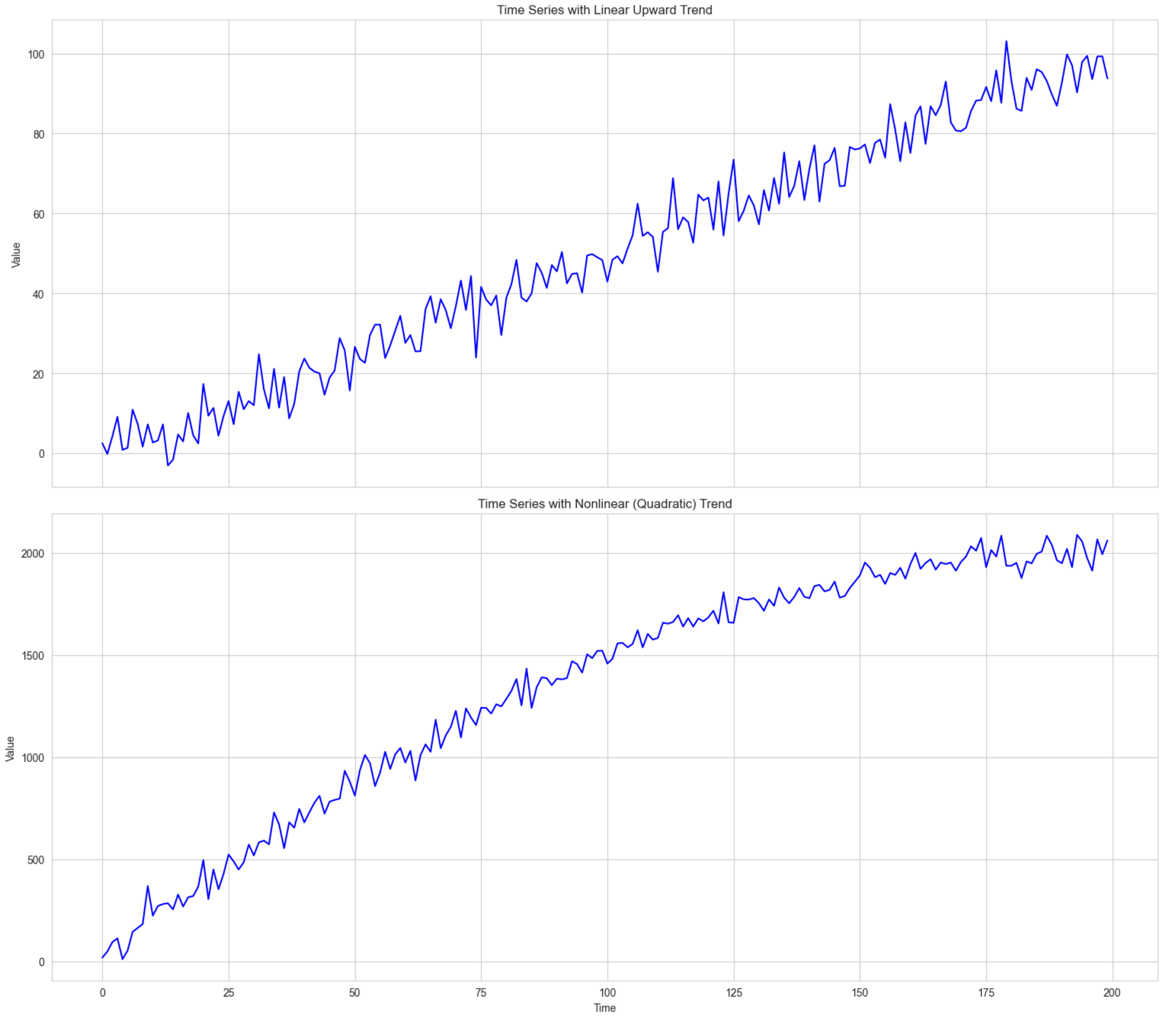
- Why is Stationarity Important?
  - Most time series models assume stationarity to make reliable forecasts.
  - Non-stationary data can lead to misleading or unstable predictions.
- Examples:
  - The daily returns of a stock (percentage changes in price) are often soft-stationary because their mean and variance remain stable.
  - Raw stock prices are usually non-stationary.



Core concepts

Trends

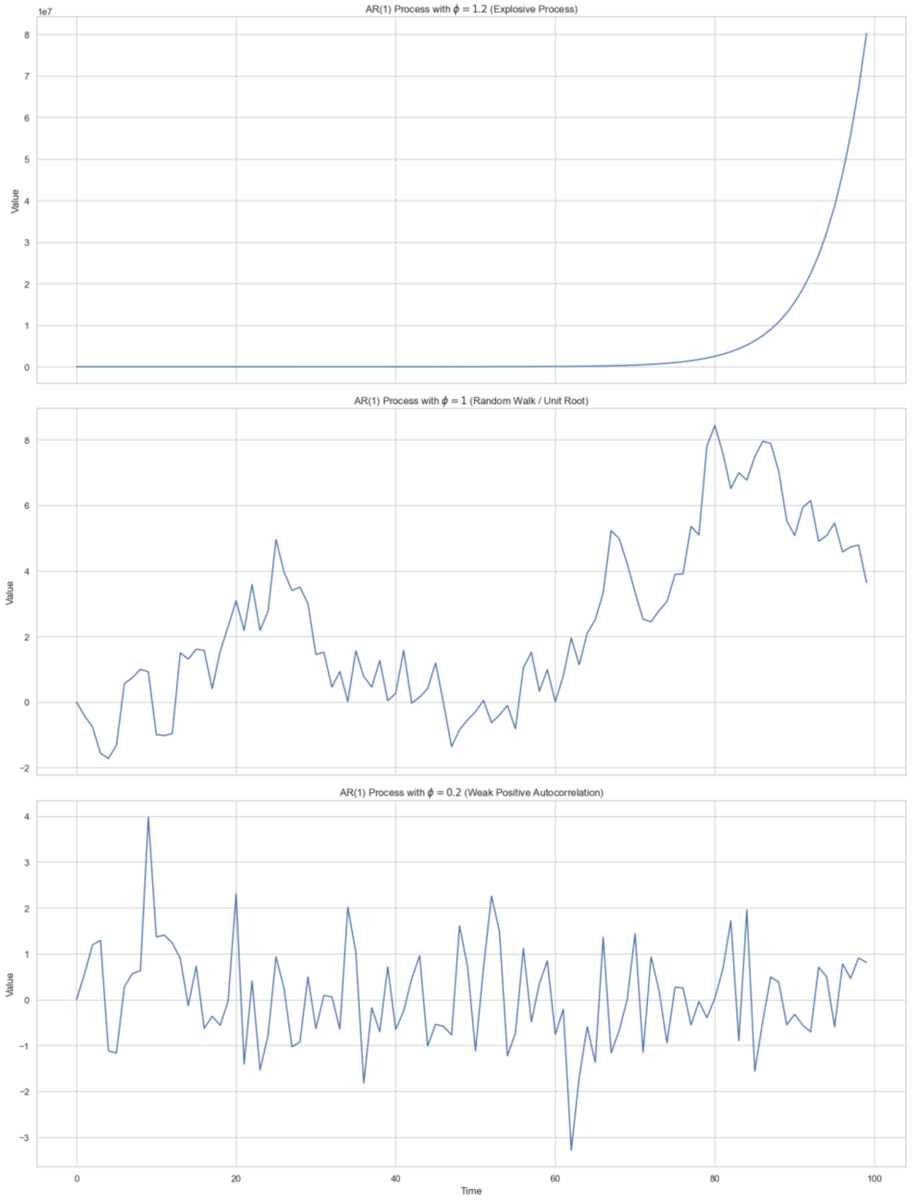
- What is a Trend?
  - A long-term increase or decrease in the data over time.
  - Represents systematic, consistent changes that persist across periods.
  - Can be linear or nonlinear.
- Why Trends Matter
  - Trends affect the stationarity of a time series.
  - Detecting and modeling trends is essential for accurate forecasting.
  - Removing trends (detrending) is often necessary for some models.
- Examples
  - Global average temperature increasing over decades (upward trend).
  - Product lifecycle sales that rise and then decline (nonlinear trend).



Core concepts

Unit roots

- A unit root is a property that makes a time series non-stationary.
- It implies the series depends strongly on past values, leading to a random walk behavior.
- Example AR(1) $X_t = \phi X_{t-1} + \varepsilon_t$
  - $|\phi|>1$: Explosive — the series diverges exponentially over time.
  - $|\phi|=1$: Unit Root — the series is a random walk, non-stationary.
  - $|\phi|<1$: Stationary — the series reverts to the mean over time.
- Example:
  - Stock prices: Often exhibit a unit root ($\phi \approx 1$), leading to unpredictable, non-reverting behavior.



Core concepts

Exogenous and endogenous variables

- Endogenous Variables
  - Variables that are determined within the system being modeled.
  - In time series, the past values of the target variable itself are considered endogenous inputs.
  - Example: Predicting future energy consumption using its own historical values.
- Exogenous Variables (Exogenous Regressors)
  - Variables that are external to the system but still impact the target variable.
  - Also called covariates or external regressors.
  - Example: Forecasting sales by incorporating advertising spend, holidays, or weather data.

- Endogenous-only models rely solely on the internal dynamics of the series.
- Including exogenous variables can significantly improve forecasting when external factors influence the target.

What we‵ve learnt

- Time series can be classified in multiple ways, including:
  - Regular vs. Irregular — depending on the timing of observations.
  - Univariate vs. Multivariate — based on the number of variables observed.
  - Simple vs. Multiple — regarding the complexity or number of series analyzed together.
  - Local vs. Global — whether modeling focuses on individual series or groups/collections.
- Understanding seasonality, cyclical signals, stationarity, and unit roots is essential for building effective models.
- Differentiating endogenous (internal) and exogenous (external) variables helps improve forecasting accuracy.