KI Service Zentrum by Hasso-Plattner-Institut

HPI

Gefördert durch:

Bundesministerium für Forschung, Technologie und Raumfahrt

Time Series Forecasting

1.3 Exploring Time Series Data

Mario Tormo Romero

Design IT.
Create Knowledge.

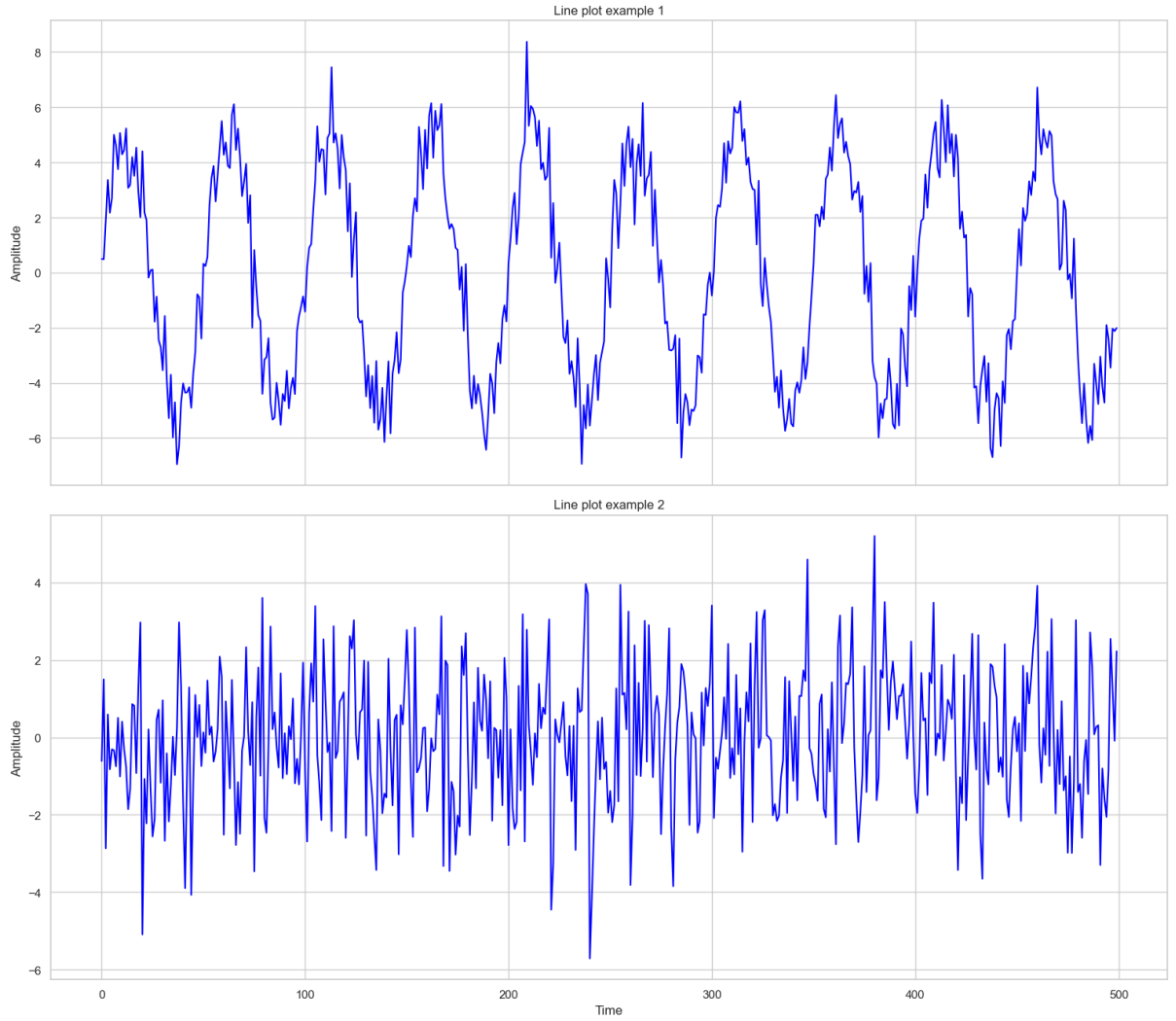
What we'll cover in this video

- Basic Visualization with Line Plots
- Identifying Trends
- Detecting Seasonality
- Recognizing Cycles
- Exploring Autocorrelation

Basic visualization

Line plots

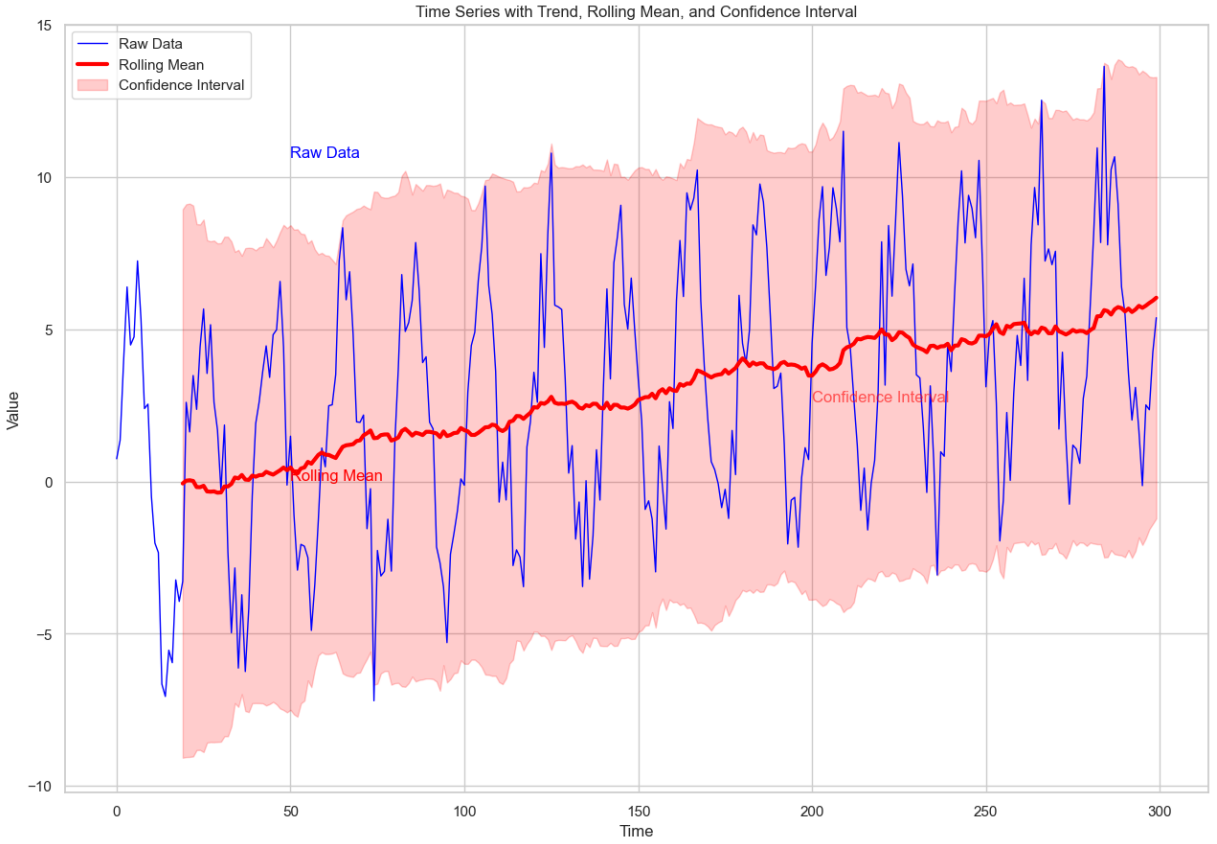
- The foundational tool for exploring time series data
- Connects data points in chronological order, revealing how values evolve over time
- Helps quickly identify:
  - Trends
  - Seasonality and cycles
  - Anomalies
- Easy to create and interpret, making it an essential first step before deeper analysis
- Provides context for applying more advanced visualization techniques



Identifying Trends

Rolling Mean & Confidence Intervals

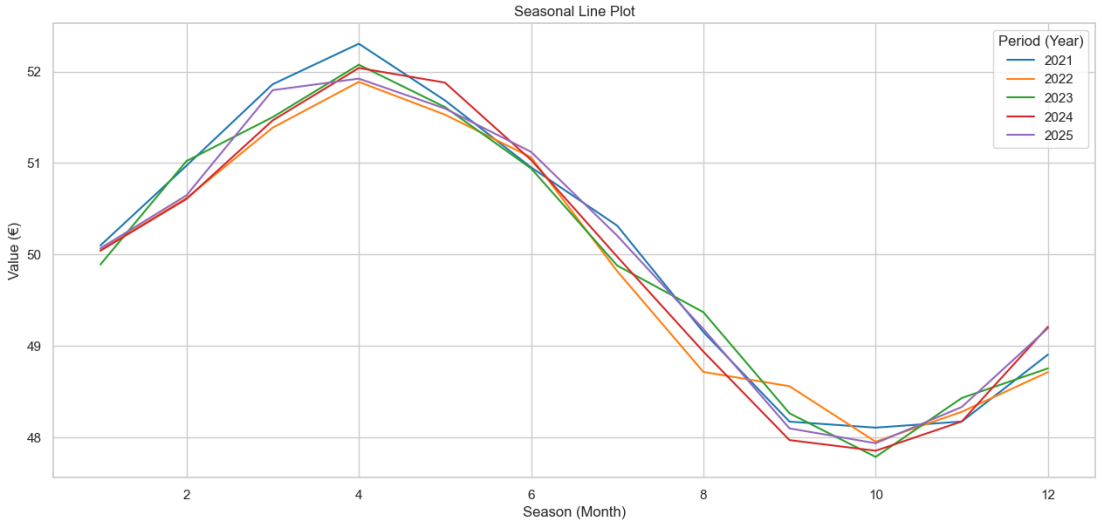
- Trends show the long-term direction in the data (upward, downward, or stable)
- Raw data can be noisy, making trends hard to spot
- A **Rolling Mean** (or **Moving Average**) smooths the series by averaging data points over a sliding window
- Helps highlight the underlying trend by reducing short-term fluctuations
- Adding **Confidence Intervals** shows the range where the true trend likely lies, indicating uncertainty

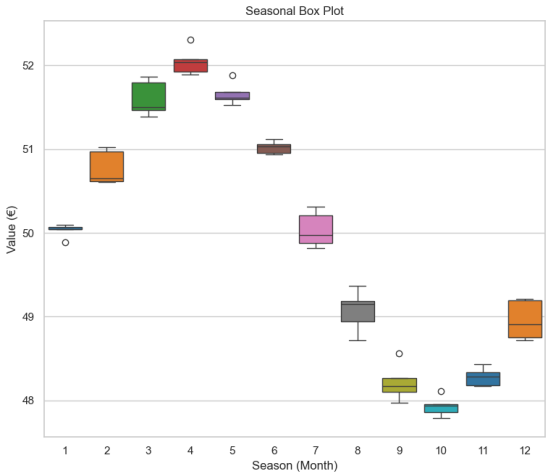


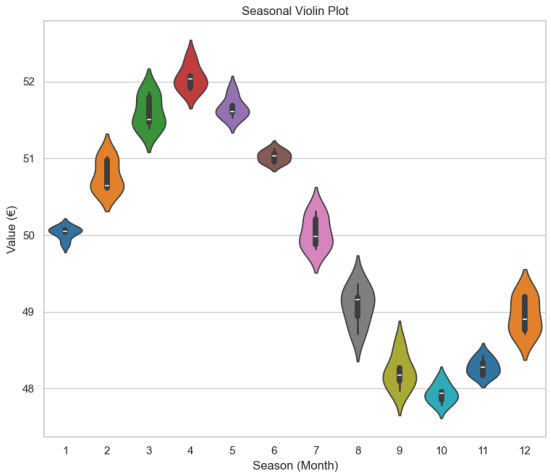
Visualizing seasonality

Seasonal plot lines, Box and Violin lines

- Seasonal Line Plots:
  - Plot data grouped by time period (e.g., all Januaries, all Mondays)
  - Helps detect repeating patterns across periods
- Box Plots:
  - Show the distribution of values within each season or period
  - Easily spot medians, variability, and outliers
- Violin Plots:
  - Similar to box plots but add a density estimate
  - Reveal the full shape of the data distribution per season



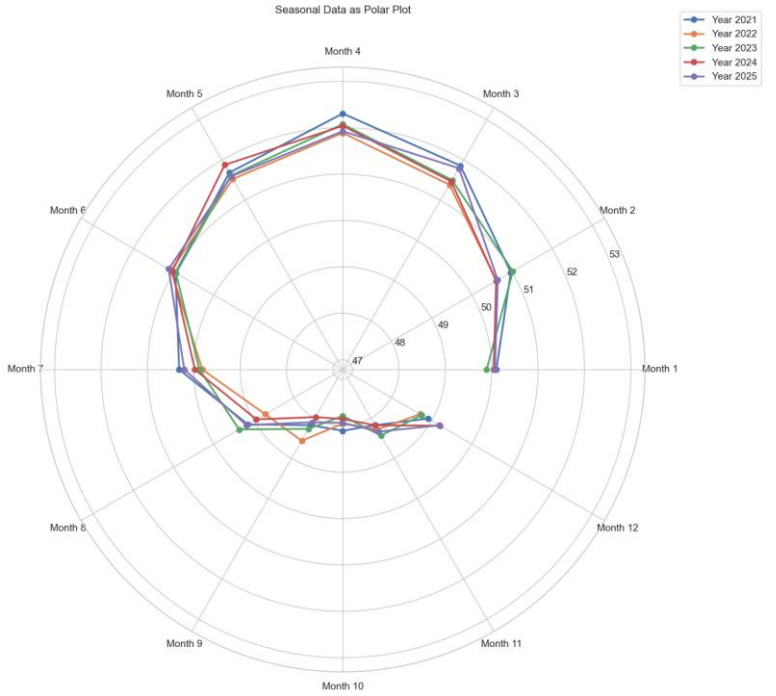


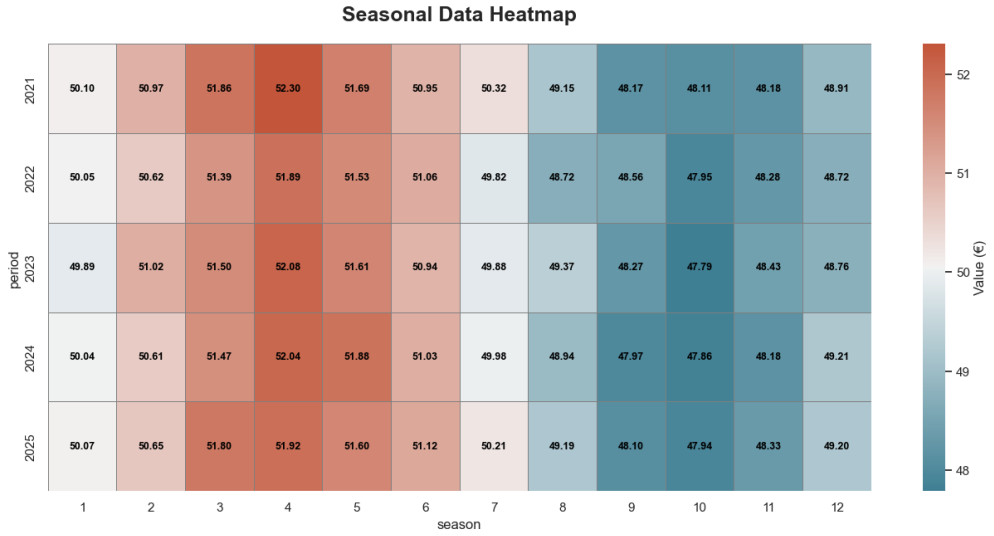


Visualizing seasonality

Polar plots and heatmaps

- Polar Plots:
  - Circular visualization of time-based data (e.g., months or hours)
  - Ideal for highlighting cyclical patterns like annual or daily seasonality
  - Makes repeating cycles visually intuitive
- Heatmaps:
  - Displays data values using color intensity over time dimensions (e.g., years vs months)
  - Excellent for spotting patterns, anomalies, and shifts in seasonality across years
  - Useful for long time series with multiple seasonal layers





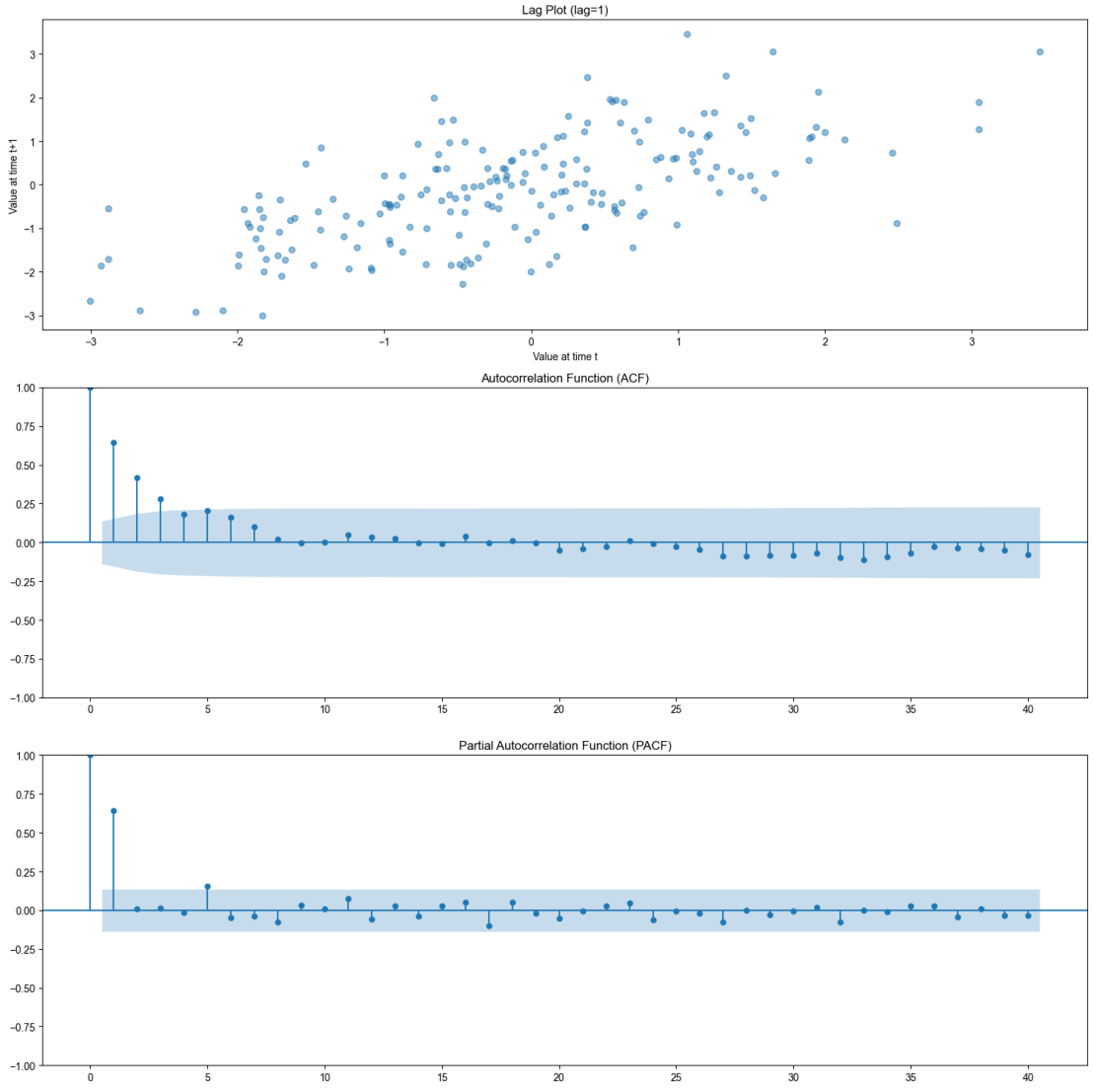
Recognizing cycles

Autocorrelation & Lag Plots

- Visual inspection and autocorrelation methods can help identify cycles

- Lags:
  - Lag 1 compares each point with the immediately preceding point, lag 2 compares points two steps apart, and so on
  - Seasonality often appears as spikes at multiples of a base lag (e.g., 12 for monthly data with yearly seasonality)

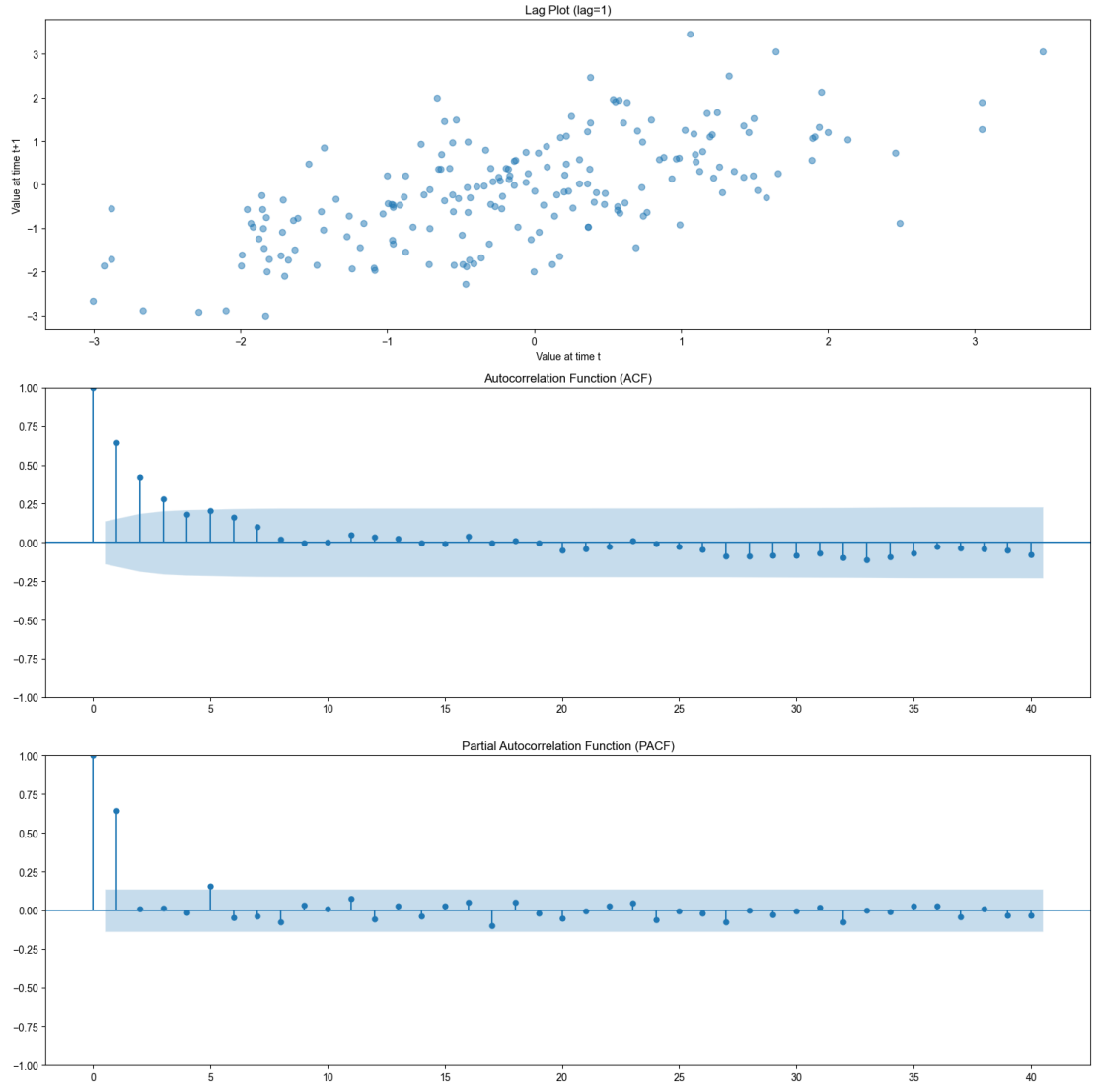
- Autocorrelation:
  - Measures how current values relate to past values at different lags
  - Helps identify repeating patterns and cycles in the data



Recognizing cycles

Autocorrelation & Lag Plots

- Autocorrelation Function (ACF):
  - Measures correlation between the series and its lagged values
  - Each lag represents the number of time steps between observations
  - Significant spikes at specific lags indicate correlation at those intervals
  - Useful to detect seasonality and persistence
- Partial Autocorrelation Function (PACF):
  - Measures direct correlation between the series and a given lag, controlling for correlations at shorter lags
  - Helps isolate the influence of each lag independently
  - Sharp spikes at particular lags show where the strongest direct relationships occur



What we've learnt

- Visualization is essential to understand trends, seasonality, and cycles
- Use:
  - Line plots and rolling means for trends
  - Seasonal plots, box plots, and violin plots for seasonality
  - Polar plots and heatmaps to visualize seasonal patterns across cycles
- Lag plots and ACF/PACF reveal correlations and repeating patterns
- These insights guide the choice of models and preprocessing steps