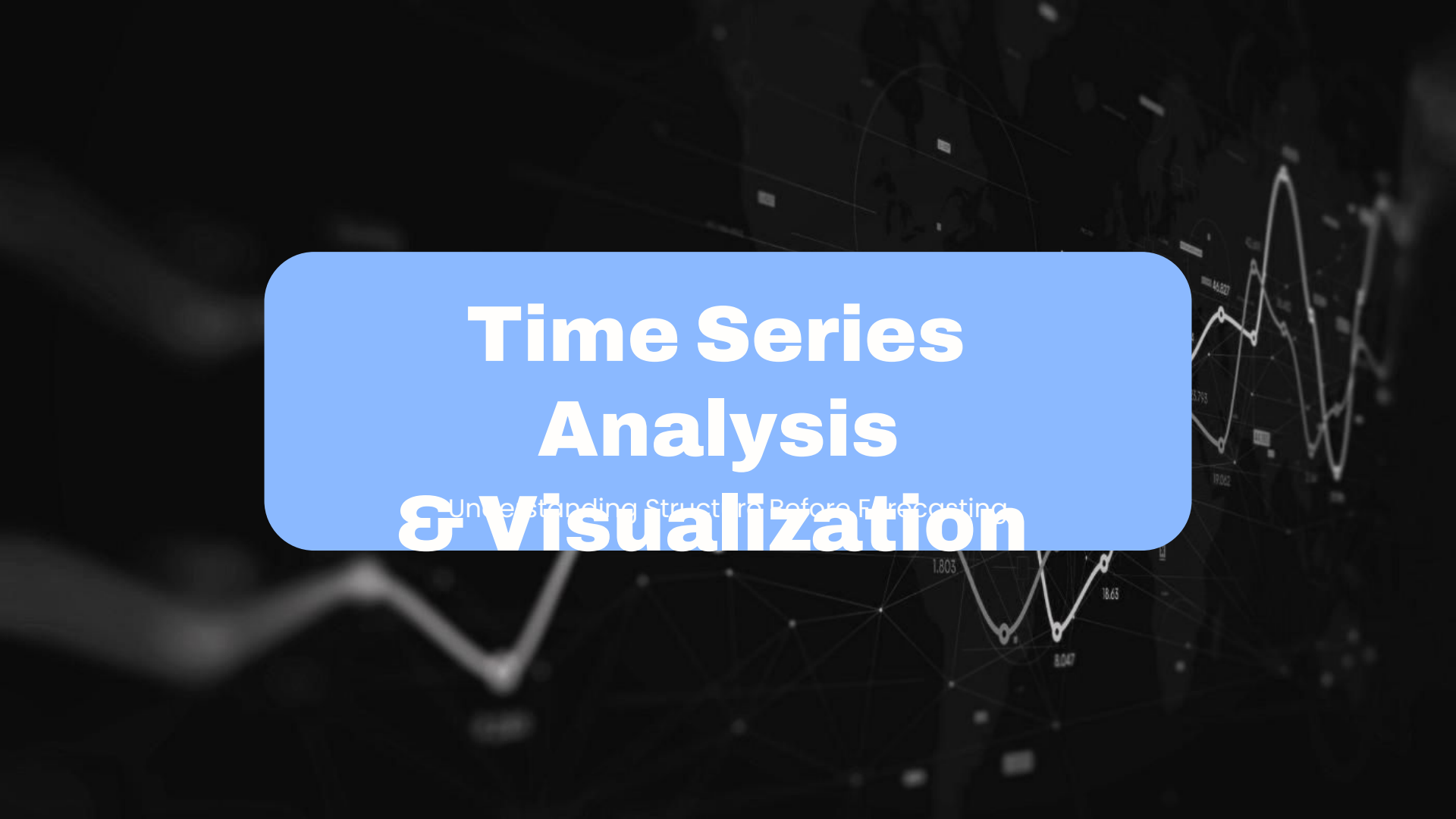


Time Series Analysis & Visualization

Understanding Structure Before Forecasting

ATTENDANC

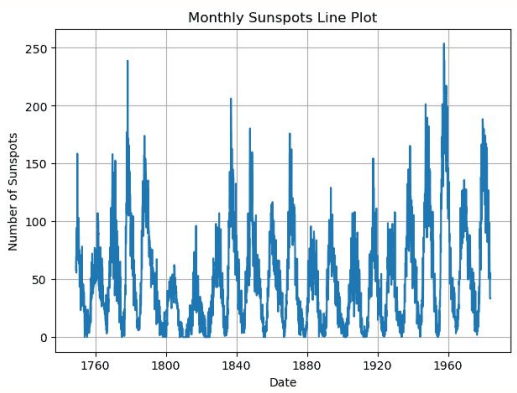
Meeting Name: “Time Series”

What Is Time Series Data?

- Data collected at consistent intervals to identify trends, cycles, and seasonal patterns
- Goal: detect seasonality & trends, forecast future values, and diagnose structure before modeling
- Key examples: Daily temperature, hourly website traffic, monthly sales, stock prices

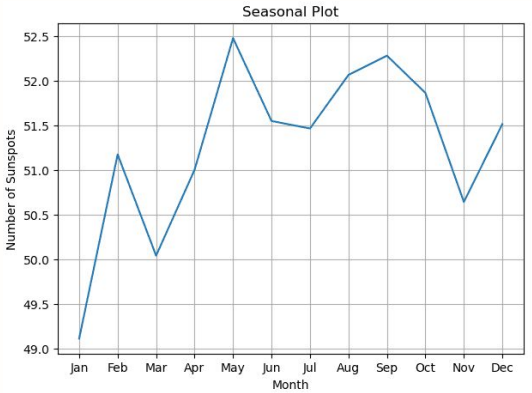
```
> head(data, 10)
      Month Passengers
1   1949-01        112
2   1949-02        118
3   1949-03        132
4   1949-04        129
5   1949-05        121
6   1949-06        135
7   1949-07        148
8   1949-08        148
9   1949-09        136
10  1949-10        118
```

Start With Visualization



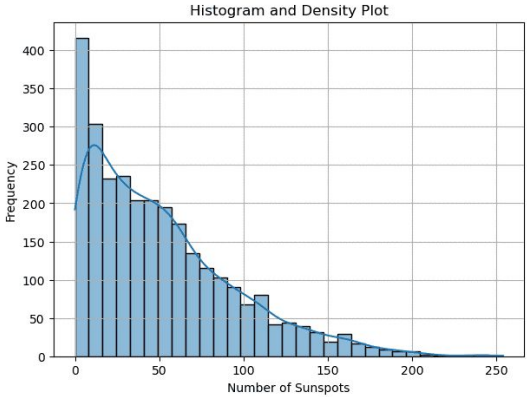
Line Plots

- Shows individual values over time
- Reveals trends and outliers



Seasonal Plots

- Aggregates same season (usually month or day) across years
- Confirms seasonality strength



Density Plots

- Distribution of values
- Shape
- Volatility changes
- Compares distributions across periods

Time Series Decomposition

Makes patterns obvious and helps with model selection

Series = Trend + Seasonality + Residual

Trend

Long-term movement

Ex. economic growth or climate change

Residual

Random variation or unexplained events

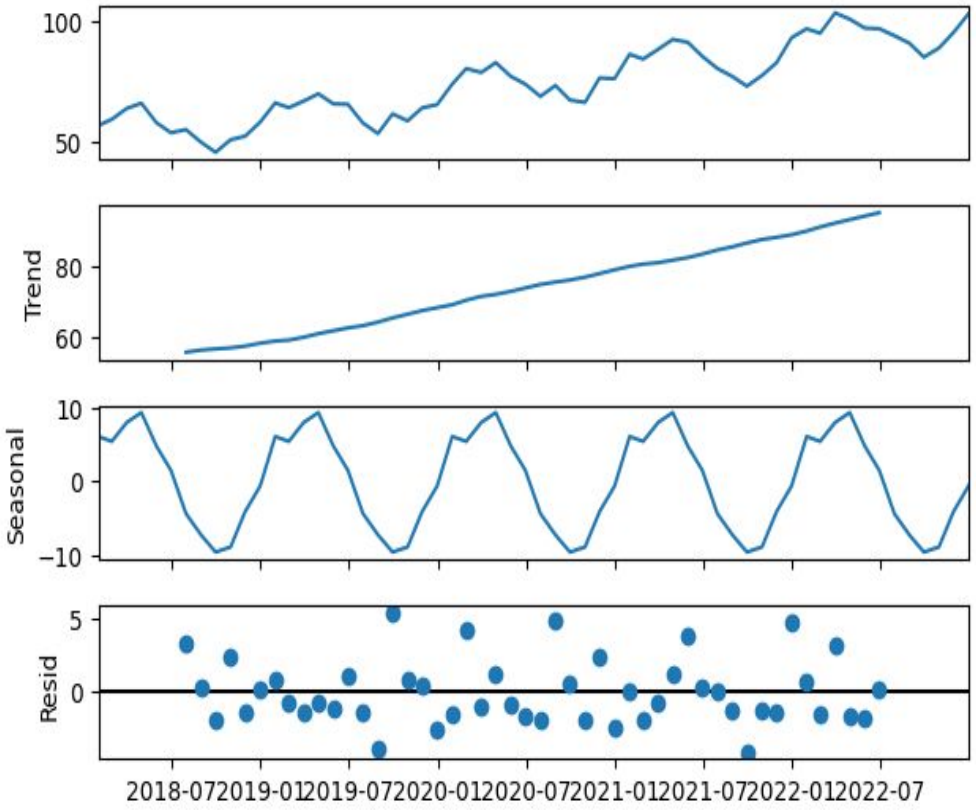
The “noise”

Seasonality

Repeating patterns

Ex. daily, weekly, or yearly cycles

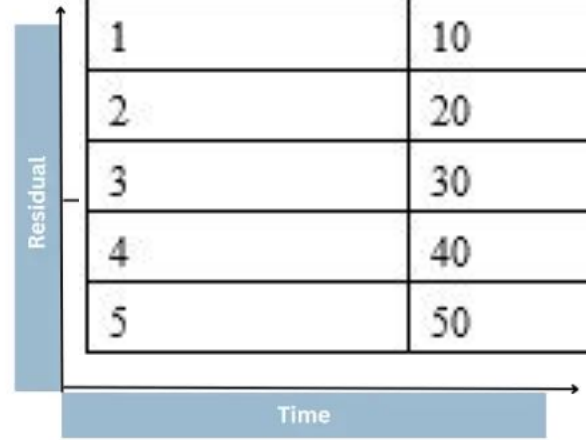
- Additive: seasonal effect is constant size
- Multiplicative: seasonal effect grows with level

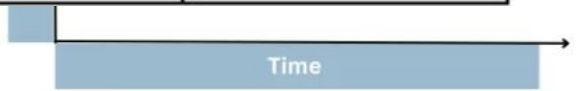


Autocorrelation

Measures how correlated a time series is with itself at different lags

Day	Observed Value	Lag-1	Lag-2
1	10		
2	20	10	
3	30	20	10
4	40	30	20
5	50	40	30

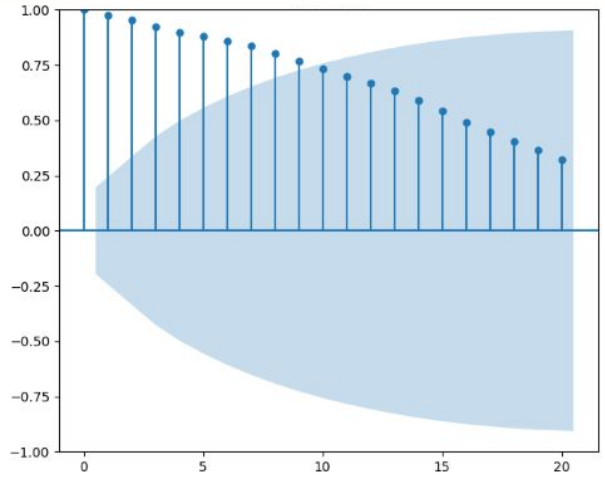




If data depends on its past, autocorrelation exists

If there is no autocorrelation, forecasting is very hard

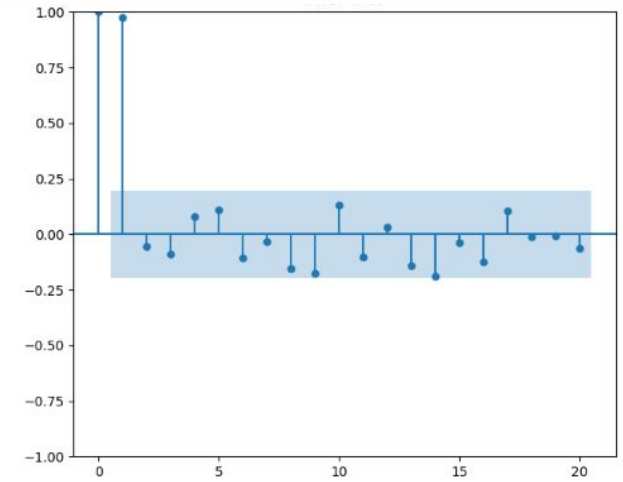
ACF Plot



X-axis: lag

Y-axis: correlation value

PACF Plot



- Bars outside of shaded region: statistically insignificant
- Shows <u>all</u> influence (indirect and direct)
- Captures full relationship between a value and all its past lags

- Helps determine how many past steps directly matter
- <u>Isolates direct</u> influence
- Removes influence of intermediate lags

Stationary

Statistical properties that remain constant over time

vs

Non-Stationary

Statistical properties that change over time, making it unpredictable and unsuitable for standard regression models

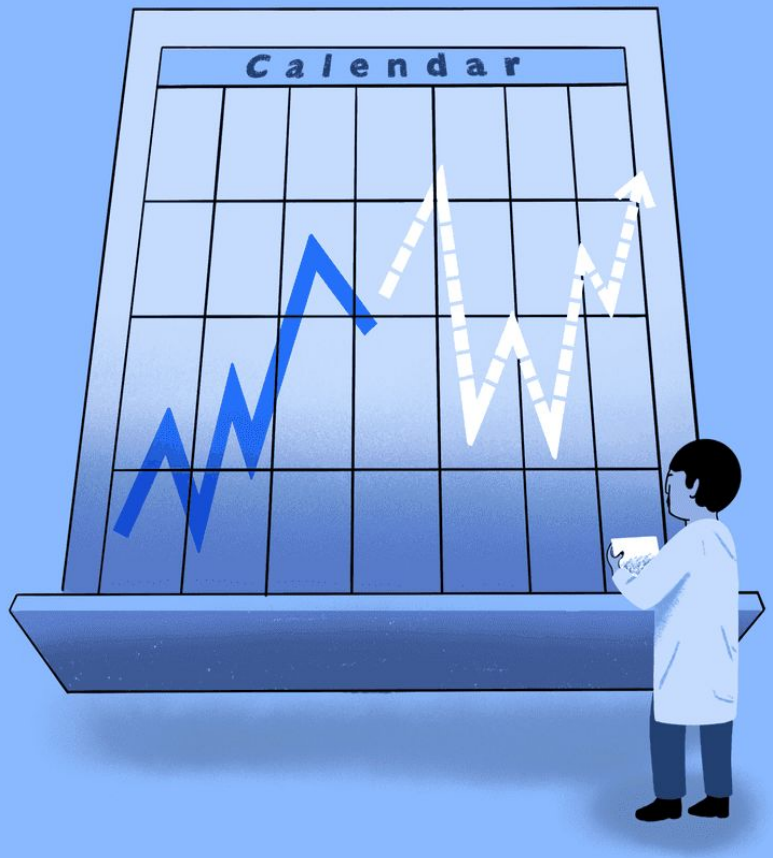
Most real-world time series are non-stationary

Time Series Analysis

- Time series data is ordered and dependent on past values
- Visualization is essential
- Decomposition clarifies structure
- ACF and PACF reveal dependence over time

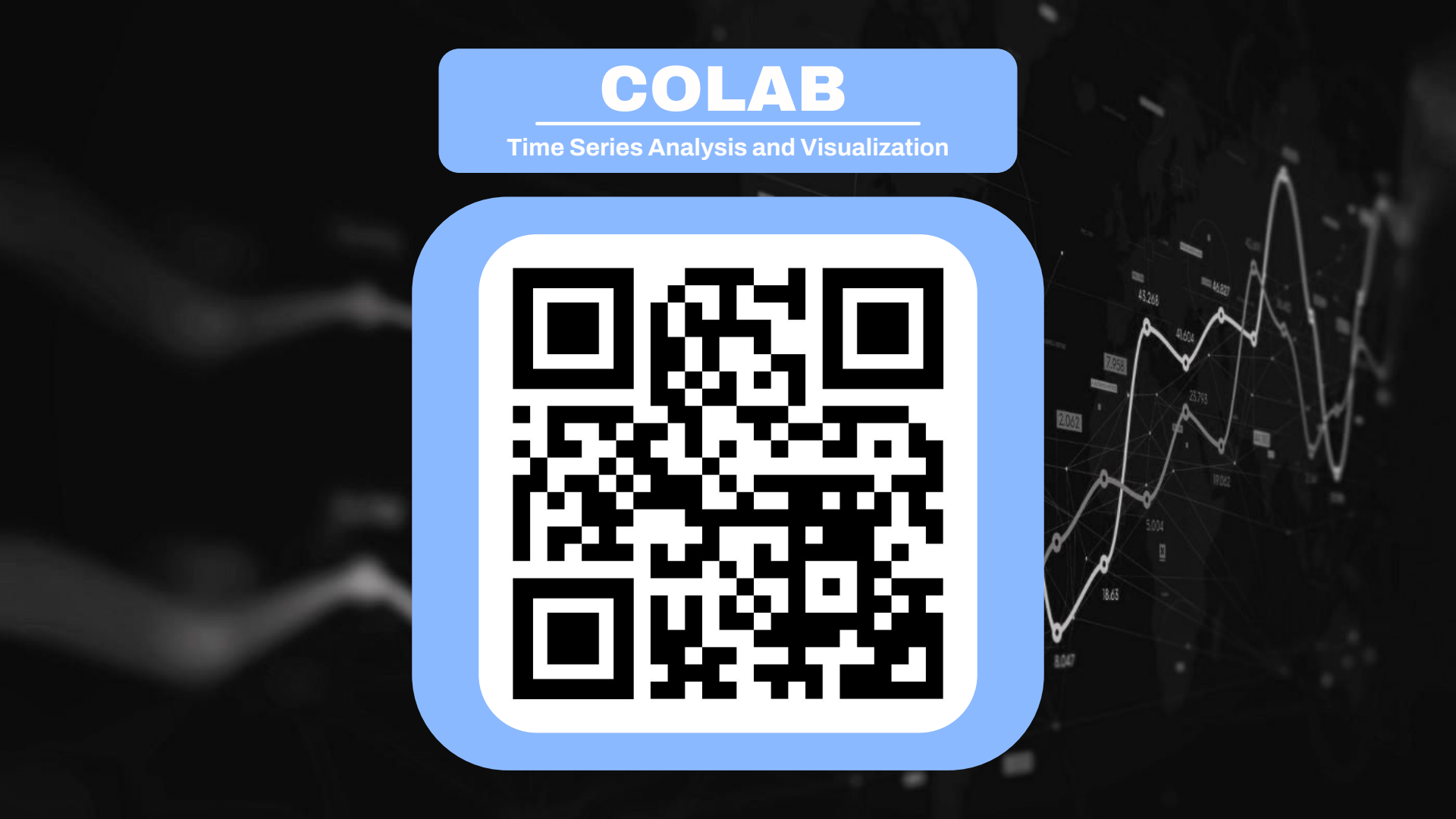
Next Week → ARIMA

Autocorrelation, PACF, and stationarity form the foundation of ARIMA modeling, which we'll implement in next week's lab



COLAB

Time Series Analysis and Visualization




Thank You
See you next week for ARIMA modeling