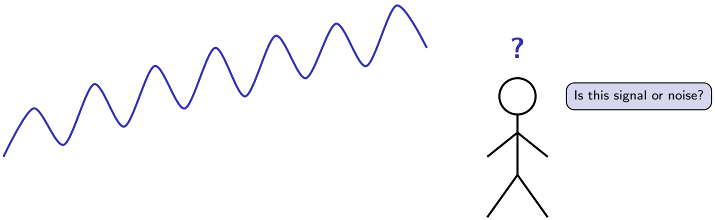
Time Series Analysis: A Visual Guide

A Formula-Free Introduction

Statistical Data Analysis

Lesson 5

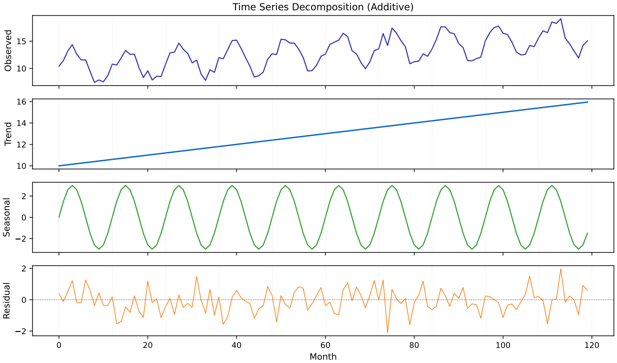
Is This Signal or Noise?



Data arrives as a wiggly line – can you spot the pattern?

Every time series hides a story – your job is to decode it.

What Hides Inside Every Time Series?



- Every series is a mix of **trend**, **seasonality**, and **noise**
- Decomposition separates these three layers
- Once separated, each part is easier to understand and model

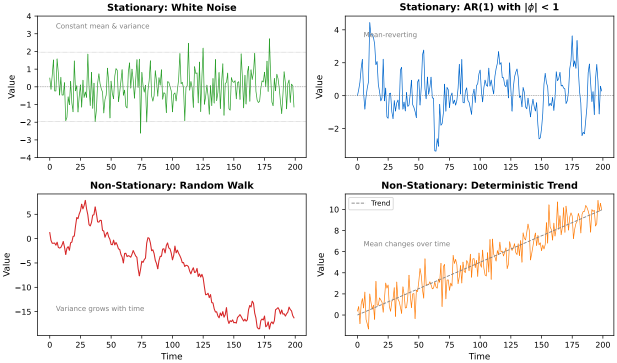
Decomposition is the first step in any time series analysis.

What Will You Learn Today?

1. **Decompose** a time series into trend, season, and noise
2. **Build ARIMA models** to forecast the future
3. **Know when ARIMA fails** and what to use instead

Three skills that cover 80% of real-world time series problems.

Why Does Stationarity Matter?



- A stationary series has **constant mean and variance** over time
- Most forecasting models **require stationarity** to work
- Differencing removes trends to make data stationary

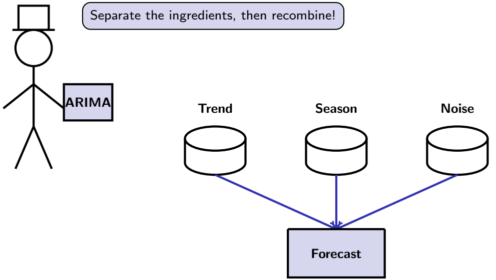
Stationarity is the foundation – without it, forecasts drift.

How Do You Make a Series Stationary?

- **Subtract the trend:** differencing means subtracting yesterday's value from today's
- **Remove seasonality:** seasonal differencing removes repeating calendar patterns
- **Check with the ADF test:** a small p-value confirms the series is now stationary

Differencing is the simplest and most common way to achieve stationarity.

Can You Follow the ARIMA Recipe?



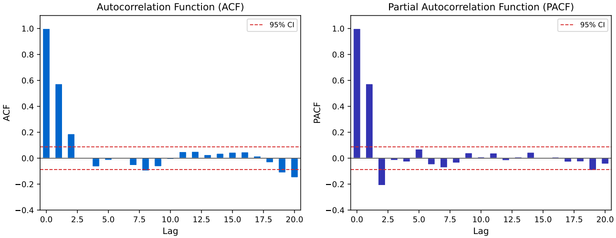
ARIMA separates the ingredients, models each one, and recombines them into a forecast.

Think of ARIMA as a recipe: decompose, model, recombine.

What Do p, d, and q Mean?

- $p =$ **how many past values** influence today (autoregressive order)
- $d =$ **how many times you difference** to achieve stationarity
- $q =$ **how many past forecast errors** matter (moving average order)

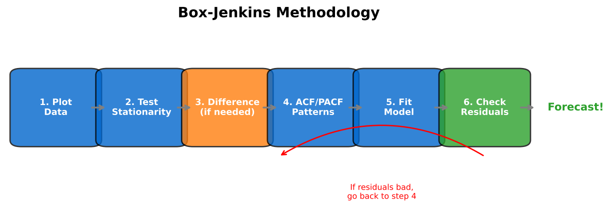
ARIMA(p,d,q) – three numbers that fully specify the model.



AR(2) process: PACF cuts off after lag 2, ACF decays gradually

- ACF shows overall correlation at each lag
- PACF shows direct correlation after removing intermediate effects
- The pattern of spikes and cutoffs tells you p and q

ACF and PACF are the fingerprints that identify your ARIMA model.



- **Identify** – check stationarity, read ACF/PACF
- **Estimate** – fit the ARIMA model to the data
- **Diagnose** – check that residuals look like white noise

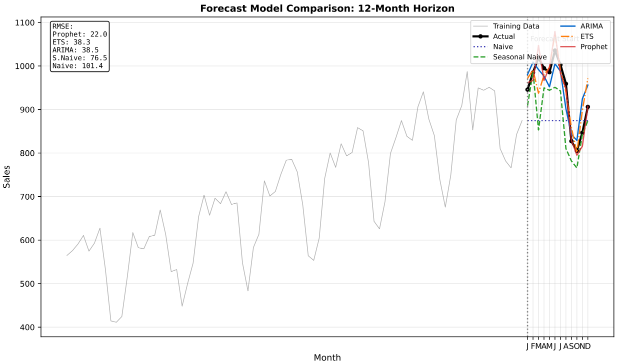
Box-Jenkins gives you a systematic three-step recipe for any ARIMA model.

How Do You Know If Your Model Is Good?

- **Residuals should look like random noise** – no patterns left behind
- **Compare models with AIC** – lower AIC means a better balance of fit and simplicity
- A good model **captures the signal** and leaves only noise behind

If the residuals still show patterns, your model is missing something.

Does ARIMA Actually Forecast Well?



- ARIMA captures trend and seasonality in the forecast
- The dashed line extends the pattern into the future
- Shaded bands show uncertainty – wider means less confident

ARIMA forecasts come with built-in uncertainty bands – always report them.

When Does ARIMA Shine?

- Works best for **short-to-medium range** forecasts
- Handles trend and seasonality via the **SARIMA extension**
- Struggles when the data has **changing variance** (volatility)

ARIMA is the workhorse – but every workhorse has its limits.

What Happens When Variance Is Not Constant?

- Financial data shows **calm periods and turbulent periods**
- ARIMA assumes constant variance – it **cannot adapt**
- This pattern is called **volatility clustering**

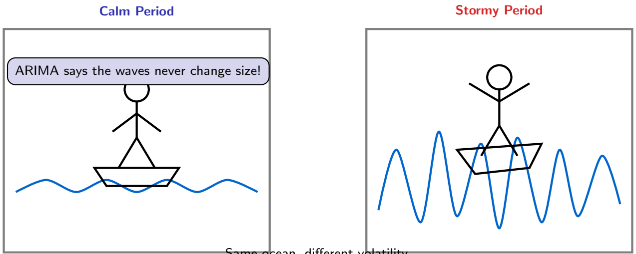
Constant variance is ARIMA's Achilles heel in financial data.

Why Do Calm and Stormy Periods Alternate?

- Large price moves tend to **follow large price moves**
- Small moves tend to **follow small moves**
- GARCH models capture this by letting **variance change over time**

GARCH extends ARIMA by modelling how volatility itself evolves.

Can ARIMA Handle a Storm?

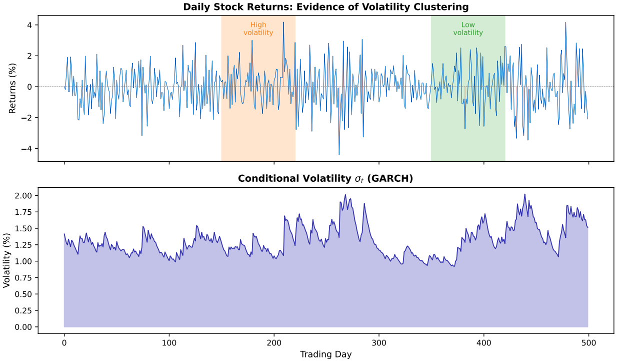


Same ocean, different volatility.

ARIMA assumes the ocean is always calm – GARCH knows storms come and go.

Volatility is not constant – models that ignore this miss the storms.

See Volatility Clustering in Action!



Stylized Facts: Large returns cluster together | Volatility is persistent | Returns are approximately uncorrelated

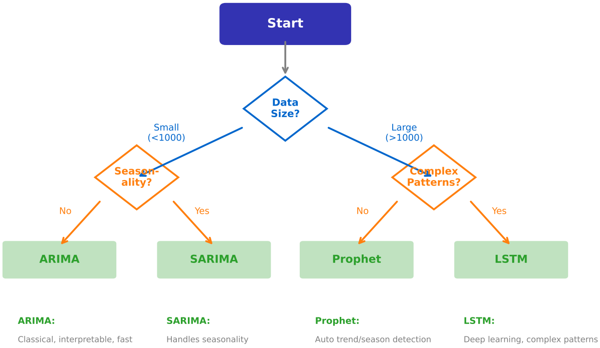
- Notice how **large swings cluster together**
- GARCH models this pattern by making today's variance depend on yesterday's

Volatility clustering is the signature pattern that calls for GARCH.

What About Really Complex Patterns?

- Some patterns are **too complex for any single equation**
- Machine learning (LSTM neural networks) can learn **nonlinear relationships**
- But ML needs **much more data** and is harder to interpret

Complexity is not free – more powerful models demand more data and more caution.



- Match the method to your **data characteristics**
- Start simple (ARIMA), add complexity **only if needed**

The best model is the simplest one that captures the key patterns.

Remember the Decision Hierarchy!

1. **Try ARIMA first** – it is interpretable and often sufficient
2. **Add GARCH** if you see volatility clustering in the residuals
3. **Try LSTM only** when you have thousands of observations and nonlinear patterns

---

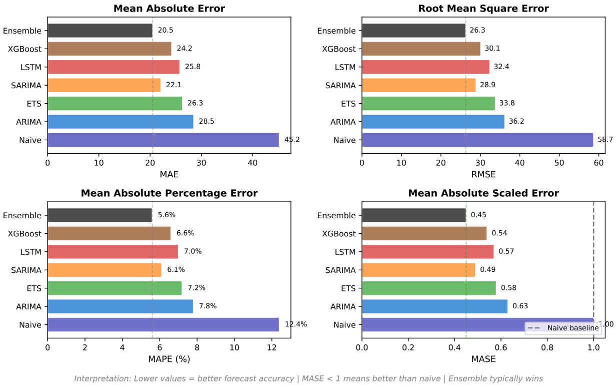
Simple models first, complex models last – always earn the complexity.

How Do You Know If a Forecast Is Good?

- A forecast is only useful if you can **measure its accuracy**
- Compare predicted values against actual outcomes on **held-out data**
- Never evaluate on the **same data you trained on**

Train-test splitting is not optional – it is the only honest evaluation.

Which Error Metric Should You Use?



- **MAE** = average absolute error (easy to interpret)
- **RMSE** = penalizes large errors more heavily
- **MAPE** = percentage error (good for comparing across scales)

Pick the metric that matches your business cost of being wrong.

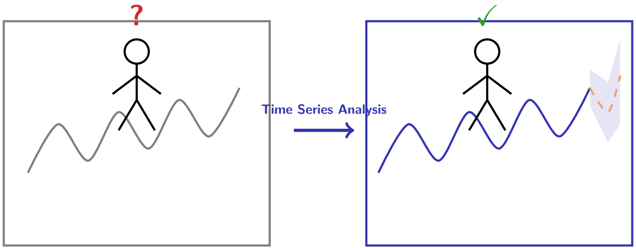
Remember the Three Big Takeaways!

1. **Decompose first** – every series has trend, season, and noise
2. **ARIMA is the workhorse** – follow Box-Jenkins to build and diagnose it
3. **Know its limits** – use GARCH for volatility, ML for truly complex patterns

---

Decompose, model, diagnose, and know when to upgrade – that is the workflow.

Can You Read the Wiggly Line Now?



From a confusing wiggly line to a confident forecast – that is time series analysis.

You started confused – now you have the tools to decode any time series.

Where Do You Go From Here?

- Try ARIMA on a real dataset: `auto.arima(data)` in R fits the best model automatically
- Experiment with seasonal decomposition: `decompose(data)` in R shows trend, season, and noise
- Read the full technical lecture for the mathematics and advanced models

---

The best way to learn time series is to apply it – start with your own data.