



# Chapter 1: Introduction to Time Series

Seminar



## Today's Activities:

1. **Quick Review** – Key concepts recap
2. **Multiple Choice Quizzes** – Test your understanding
3. **True/False Questions** – Conceptual checks
4. **Calculation Exercises** – Hands-on practice
5. **Python Exercises** – Coding practice
6. **Discussion Questions** – Critical thinking

# Key Formulas to Remember

## Decomposition:

- Additive:  $X_t = T_t + S_t + \varepsilon_t$
- Multiplicative:  $X_t = T_t \times S_t \times \varepsilon_t$

## Exponential Smoothing:

- SES:  $\hat{X}_{t+1} = \alpha X_t + (1 - \alpha)\hat{X}_t$
- Holt: adds trend  $b_t$
- HW: adds seasonal  $S_t$

## Stationarity:

- $\mathbb{E}[X_t] = \mu$  (constant)
- $\text{Var}(X_t) = \sigma^2$  (constant)
- $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$

## Random Walk:

- $X_t = X_{t-1} + \varepsilon_t$
- $\text{Var}(X_t) = t\sigma^2$  (grows!)

# Key Concepts Summary

Concept	Key Point	When to Use
Additive decomp.	Constant seasonal amplitude	Stable variance
Multiplicative decomp.	Seasonal grows with level	Increasing variance
SES	Level only ( $\alpha$ )	No trend, no seasonality
Holt	Level + Trend ( $\alpha, \beta$ )	Trend, no seasonality
Holt-Winters	Level + Trend + Seasonal	Trend and seasonality
ADF Test	$H_0$ : unit root	Test for non-stationarity
KPSS Test	$H_0$ : stationary	Confirm stationarity
Differencing	Remove stochastic trend	Random walk, unit root
Regression	Remove deterministic trend	Linear/polynomial trend

## [QUIZ] Quiz 1: Time Series Basics

**Question:** Which of the following is NOT a characteristic of time series data?

- ☐ A. Observations are ordered in time
- ☐ B. Consecutive observations are typically correlated
- ☐ C. Observations are independent and identically distributed
- ☐ D. The data has a natural temporal ordering

*Think about it before moving to the next slide...*

## [QUIZ] Quiz 1: Answer

**Question:** Which is NOT a characteristic of time series data?

- ☐ A. Observations are ordered in time ✗
- ☐ B. Consecutive observations are typically correlated ✗
- ☒ C. **Observations are independent and identically distributed** ✓
- ☐ D. The data has a natural temporal ordering ✗

### Explanation

Time series observations are typically **dependent** (autocorrelated), not independent. The assumption of i.i.d. observations is fundamental to cross-sectional analysis but violated in time series. This temporal dependence is what makes time series analysis unique and requires specialized methods.

## [QUIZ] Quiz 2: Decomposition

**Question:** When should you use multiplicative decomposition instead of additive?

- ☐ A. When the seasonal pattern has constant amplitude
- ☐ B. When the variance of the series is stable over time
- ☐ C. When the seasonal fluctuations grow proportionally with the level
- ☐ D. When the time series has no trend component

*Think about it before moving to the next slide...*

## [QUIZ] Quiz 2: Answer

**Question:** When should you use multiplicative decomposition?

- A. When the seasonal pattern has constant amplitude ✗
- B. When the variance of the series is stable over time ✗
- C. **When the seasonal fluctuations grow proportionally with the level** ✓
- D. When the time series has no trend component ✗

### Explanation

In multiplicative decomposition  $X_t = T_t \times S_t \times \varepsilon_t$ , the seasonal component  $S_t$  is a ratio (e.g., 1.2 means 20% above average). This means the absolute seasonal effect **scales with the level**. Use when you see "fan-shaped" patterns where variance increases with the mean.



## [QUIZ] Quiz 3: Exponential Smoothing

**Question:** In Simple Exponential Smoothing with  $\alpha = 0.9$ , what happens?

- ☐ A. Forecasts are very smooth and stable
- ☐ B. Recent observations have very little weight
- ☐ C. Forecasts react quickly to recent changes
- ☐ D. The forecast is essentially a long-term average

*Think about it before moving to the next slide...*

## [QUIZ] Quiz 3: Answer

**Question:** In SES with  $\alpha = 0.9$ , what happens?

- ☐ A. Forecasts are very smooth and stable ✗
- ☐ B. Recent observations have very little weight ✗
- ☒ C. **Forecasts react quickly to recent changes ✓**
- ☐ D. The forecast is essentially a long-term average ✗

### Explanation

With  $\alpha = 0.9$ :  $\hat{X}_{t+1} = 0.9X_t + 0.1\hat{X}_t$

This means 90% weight on the most recent observation! High  $\alpha$  values make forecasts very responsive to new data. Low  $\alpha$  (e.g., 0.1) produces smoother, more stable forecasts that average over more history.

## [QUIZ] Quiz 4: Stationarity

**Question:** A random walk process  $X_t = X_{t-1} + \varepsilon_t$  is:

- ☐ A. Strictly stationary
- ☐ B. Weakly stationary
- ☐ C. Non-stationary because variance grows with time
- ☐ D. Stationary after adding a constant

*Think about it before moving to the next slide...*

## [QUIZ] Quiz 4: Answer

**Question:** A random walk is:

- A. Strictly stationary ✗
- B. Weakly stationary ✗
- C. **Non-stationary because variance grows with time ✓**
- D. Stationary after adding a constant ✗

### Explanation

For random walk:  $X_t = \sum_{i=1}^t \varepsilon_i$

- $\mathbb{E}[X_t] = 0$  (constant mean – OK)
- $\text{Var}(X_t) = t\sigma^2$  (variance depends on  $t$  – NOT OK!)

Since variance is not constant, the process violates the stationarity condition. Solution: **differencing** gives  $\Delta X_t = \varepsilon_t$  which IS stationary.

## [QUIZ] Quiz 5: Unit Root Tests

**Question:** You run ADF and KPSS tests. ADF fails to reject  $H_0$ , and KPSS rejects  $H_0$ . What do you conclude?

- ☐ A. The series is stationary
- ☐ B. The series has a unit root (non-stationary)
- ☐ C. The results are inconclusive
- ☐ D. You need to run more tests

*Think about it before moving to the next slide...*

## [QUIZ] Quiz 5: Answer

**Question:** ADF fails to reject, KPSS rejects. Conclusion?

- A. The series is stationary ✗
- B. **The series has a unit root (non-stationary) ✓**
- C. The results are inconclusive ✗
- D. You need to run more tests ✗

### Explanation

- ADF:  $H_0 = \text{unit root}$ . Fail to reject  $\Rightarrow$  evidence FOR unit root
- KPSS:  $H_0 = \text{stationary}$ . Reject  $\Rightarrow$  evidence AGAINST stationarity

Both tests agree: the series is **non-stationary**. You should difference the series before modeling with ARMA.

## [QUIZ] Quiz 6: Forecast Evaluation

**Question:** Which metric is most appropriate for comparing forecast accuracy across different time series with different scales?

- ☐ A. Mean Absolute Error (MAE)
- ☐ B. Root Mean Squared Error (RMSE)
- ☐ C. Mean Absolute Percentage Error (MAPE)
- ☐ D. Mean Squared Error (MSE)

*Think about it before moving to the next slide...*

## [QUIZ] Quiz 6: Answer

**Question:** Best metric for comparing across different scales?

- ☐ A. Mean Absolute Error (MAE) ✗
- ☐ B. Root Mean Squared Error (RMSE) ✗
- ☒ C. **Mean Absolute Percentage Error (MAPE)** ✓
- ☐ D. Mean Squared Error (MSE) ✗

### Explanation

$MAPE = \frac{100}{n} \sum \left| \frac{e_t}{X_t} \right|$  expresses errors as **percentages**.

- MAE, RMSE, MSE are **scale-dependent** (units of  $X_t$ )
- MAPE is **scale-independent** (always in %)
- Caveat: MAPE fails when  $X_t$  is near zero



## [QUIZ] Quiz 7: Trend Types

**Question:** A deterministic trend can be removed by:

- ☒ A. Differencing
- ☐ B. Regression on time
- ☐ C. Seasonal adjustment
- ☐ D. Moving average smoothing

## [QUIZ] Quiz 7: Trend Types

**Question:** A deterministic trend can be removed by:

- ☐ A. Differencing
- ☒ B. Regression on time
- ☐ C. Seasonal adjustment
- ☐ D. Moving average smoothing

**Answer:** B

Deterministic trends (e.g.,  $T_t = a + bt$ ) are removed by regression on time. Differencing is used for stochastic trends (unit roots). Knowing which type of trend you have is crucial for choosing the right method.

## [QUIZ] Quiz 8: ACF Interpretation

**Question:** If the ACF of a time series decays very slowly (remains significant for many lags), this suggests:

- ☐ A. The series is white noise
- ☐ B. The series is likely non-stationary
- ☐ C. The series has no autocorrelation
- ☐ D. The series is perfectly predictable

## [QUIZ] Quiz 8: ACF Interpretation

**Question:** If the ACF of a time series decays very slowly (remains significant for many lags), this suggests:

- ☐ A. The series is white noise
- ☒ B. The series is likely non-stationary
- ☐ C. The series has no autocorrelation
- ☐ D. The series is perfectly predictable

Answer: B

Slow ACF decay is a hallmark of non-stationarity. Stationary processes have ACF that decays to zero relatively quickly. A slowly decaying ACF suggests a unit root or long memory, requiring differencing.

## [QUIZ] Quiz 9: Holt's Method

**Question:** Holt's exponential smoothing differs from SES by adding:

- ☐ A. A seasonal component
- ☐ B. A trend component
- ☐ C. A cyclical component
- ☐ D. An irregular component

## [QUIZ] Quiz 9: Holt's Method

**Question:** Holt's exponential smoothing differs from SES by adding:

- ☐ A. A seasonal component
- ☒ B. A trend component
- ☐ C. A cyclical component
- ☐ D. An irregular component

Answer: B

Holt's method extends SES by adding a trend equation:  $b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$ . This allows forecasts to follow a trend rather than being flat. Holt-Winters further adds the seasonal component.

## [QUIZ] Quiz 10: White Noise

**Question:** Which property is NOT required for a process to be white noise?

- A.  $\mathbb{E}[\varepsilon_t] = 0$
- B.  $\text{Var}(\varepsilon_t) = \sigma^2$  (constant)
- C.  $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$  for  $t \neq s$
- D.  $\varepsilon_t \sim N(0, \sigma^2)$

## [QUIZ] Quiz 10: White Noise

**Question:** Which property is NOT required for a process to be white noise?

- A.  $\mathbb{E}[\varepsilon_t] = 0$
- B.  $\text{Var}(\varepsilon_t) = \sigma^2$  (constant)
- C.  $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$  for  $t \neq s$
- D.  $\varepsilon_t \sim N(0, \sigma^2)$

Answer: D

White noise requires zero mean, constant variance, and zero autocorrelation. Normality is NOT required—that would be Gaussian white noise. White noise can follow any distribution with these properties.



## [QUIZ] Quiz 11: Forecast Horizon

**Question:** As forecast horizon  $h$  increases, what typically happens to forecast intervals?

- ☐ A. They become narrower
- ☐ B. They stay the same width
- ☐ C. They become wider
- ☐ D. They disappear

## [QUIZ] Quiz 11: Forecast Horizon

**Question:** As forecast horizon  $h$  increases, what typically happens to forecast intervals?

- ☐ A. They become narrower
- ☐ B. They stay the same width
- ☒ C. They become wider
- ☐ D. They disappear

Answer: C

Forecast uncertainty increases with horizon. For most models, the forecast error variance grows with  $h$ , leading to wider confidence intervals. Only for stationary processes does uncertainty eventually stabilize at the unconditional variance.

## [QUIZ] Quiz 12: Seasonality Detection

**Question:** The ACF shows significant spikes at lags 12, 24, and 36 for monthly data. This suggests:

- ☐ A. No seasonality
- ☐ B. Annual seasonality
- ☐ C. Weekly seasonality
- ☐ D. Random noise

## [QUIZ] Quiz 12: Seasonality Detection

**Question:** The ACF shows significant spikes at lags 12, 24, and 36 for monthly data. This suggests:

- ☐ A. No seasonality
- ☒ B. Annual seasonality
- ☐ C. Weekly seasonality
- ☐ D. Random noise

Answer: B

Spikes at multiples of 12 in monthly data indicate annual (12-month) seasonality. The ACF captures correlation between observations separated by one year, two years, etc. This pattern is typical for economic and climate data.

## [QUIZ] Quiz 13: Cross-Validation in Time Series

**Question:** Why can't we use standard k-fold cross-validation for time series?

- ☐ A. Time series data is too small
- ☐ B. It would violate temporal ordering (future predicting past)
- ☐ C. Cross-validation is always invalid
- ☐ D. Time series don't need validation

## [QUIZ] Quiz 13: Cross-Validation in Time Series

**Question:** Why can't we use standard k-fold cross-validation for time series?

- ☐ A. Time series data is too small
- ☒ B. It would violate temporal ordering (future predicting past)
- ☐ C. Cross-validation is always invalid
- ☐ D. Time series don't need validation

**Answer:** B

Standard k-fold CV randomly shuffles data, which destroys temporal structure and causes data leakage. In time series, we use rolling-origin or expanding window CV to maintain the temporal order and never use future data to predict the past.

## [QUIZ] Quiz 14: MAPE Limitation

**Question:** MAPE (Mean Absolute Percentage Error) should NOT be used when:

- ☐ A. Comparing models on the same dataset
- ☐ B. The actual values can be zero or near zero
- ☐ C. Forecasting stock prices
- ☐ D. The data has a trend

## [QUIZ] Quiz 14: MAPE Limitation

**Question:** MAPE (Mean Absolute Percentage Error) should NOT be used when:

- ☐ A. Comparing models on the same dataset
- ☒ B. The actual values can be zero or near zero
- ☐ C. Forecasting stock prices
- ☐ D. The data has a trend

**Answer:** B

MAPE divides by actual values:  $|e_t/Y_t|$ . When  $Y_t$  is zero or near zero, MAPE becomes undefined or extremely large. Alternatives include SMAPE (symmetric MAPE) or scaled errors like MASE.



## [QUIZ] True or False? (Set 2)

Mark each statement as True (T) or False (F):

- ① A time series with constant mean is always stationary. \_\_\_\_\_
- ② The variance of a random walk increases linearly with time. \_\_\_\_\_
- ③ SES forecasts are always flat (constant for all horizons). \_\_\_\_\_
- ④ ADF and KPSS tests have the same null hypothesis. \_\_\_\_\_
- ⑤ Lower RMSE always means better forecasts. \_\_\_\_\_
- ⑥ Autocorrelation at lag 0 is always equal to 1. \_\_\_\_\_

## [QUIZ] True or False: Answers (Set 2)

- ❶ A time series with constant mean is always stationary. FALSE  
Also need constant variance and covariance depending only on lag.
- ❷ The variance of a random walk increases linearly with time. TRUE  
 $\text{Var}(X_t) = t\sigma^2$  for random walk starting at  $X_0$ .
- ❸ SES forecasts are always flat (constant for all horizons). TRUE  
SES has no trend component, so  $\hat{X}_{t+h} = L_t$  for all  $h$ .
- ❹ ADF and KPSS tests have the same null hypothesis. FALSE  
ADF:  $H_0 = \text{unit root}$ . KPSS:  $H_0 = \text{stationary}$ . Opposite nulls!
- ❺ Lower RMSE always means better forecasts. FALSE  
Depends on context. RMSE is scale-dependent; may overfit to outliers.
- ❻ Autocorrelation at lag 0 is always equal to 1. TRUE  
 $\rho(0) = \gamma(0)/\gamma(0) = 1$  by definition.

## [QUIZ] True or False?

Mark each statement as True (T) or False (F):

- ① The ACF of a stationary AR(1) process decays exponentially. \_\_\_\_\_
- ② White noise is always normally distributed. \_\_\_\_\_
- ③ Differencing can make a non-stationary series stationary. \_\_\_\_\_
- ④ The PACF of a MA(1) process cuts off after lag 1. \_\_\_\_\_
- ⑤ You should always use the test set for hyperparameter tuning. \_\_\_\_\_
- ⑥ Holt-Winters is appropriate for data with no seasonality. \_\_\_\_\_

*Answers on next slide...*

## [QUIZ] True or False: Answers

- ❶ The ACF of a stationary AR(1) decays exponentially. TRUE  
For AR(1):  $\rho(h) = \phi^h$ , which decays exponentially.
- ❷ White noise is always normally distributed. FALSE  
White noise only requires zero mean, constant variance, no autocorrelation. Gaussian white noise is a special case.
- ❸ Differencing can make a non-stationary series stationary. TRUE  
Differencing removes stochastic trends (unit roots).
- ❹ The PACF of a MA(1) cuts off after lag 1. FALSE  
It's the ACF that cuts off for MA. PACF decays for MA processes.
- ❺ You should always use the test set for hyperparameter tuning. FALSE  
Use validation set for tuning. Test set is for final evaluation only!
- ❻ Holt-Winters is appropriate for data with no seasonality. FALSE  
Use Holt's method (no seasonal component) or SES for non-seasonal data.

## Exercise 1: Simple Exponential Smoothing

**Problem:** Given the following data and  $\alpha = 0.3$ :

$t$	1	2	3	4	5
$X_t$	10	12	11	14	13

Starting with  $\hat{X}_1 = X_1 = 10$ , calculate:

- a) The forecasts  $\hat{X}_2, \hat{X}_3, \hat{X}_4, \hat{X}_5$
- b) The forecast for  $t = 6$ :  $\hat{X}_6$
- c) The forecast errors  $e_t = X_t - \hat{X}_t$  for  $t = 2, 3, 4, 5$
- d) The MAE and RMSE

**Formula:**  $\hat{X}_{t+1} = \alpha X_t + (1 - \alpha)\hat{X}_t$

## Exercise 1: Solution

Using  $\hat{X}_{t+1} = 0.3X_t + 0.7\hat{X}_t$ :

$t$	1	2	3	4	5	6
$X_t$	10	12	11	14	13	?
$\hat{X}_t$	10	10	10.6	10.72	11.70	<b>12.09</b>
$e_t$	–	2	0.4	3.28	1.30	–

Calculations:

- $\hat{X}_2 = 0.3(10) + 0.7(10) = 10$
- $\hat{X}_3 = 0.3(12) + 0.7(10) = 10.6$
- $\hat{X}_4 = 0.3(11) + 0.7(10.6) = 10.72$
- $\hat{X}_5 = 0.3(14) + 0.7(10.72) = 11.70$
- $\hat{X}_6 = 0.3(13) + 0.7(11.70) = \mathbf{12.09}$

$$\text{MAE} = \frac{|2| + |0.4| + |3.28| + |1.30|}{4} = 1.745 \quad \text{RMSE} = \sqrt{\frac{4 + 0.16 + 10.76 + 1.69}{4}} = 2.04$$

## Exercise 2: Autocovariance

**Problem:** For a stationary process with:

- $\mathbb{E}[X_t] = 5$
- $\gamma(0) = 4$  (variance)
- $\gamma(1) = 2$
- $\gamma(2) = 1$

Calculate:

- a) The autocorrelation function  $\rho(0), \rho(1), \rho(2)$
- b)  $\text{Cov}(X_t, X_{t-1})$
- c)  $\text{Corr}(X_5, X_7)$
- d) If  $X_t = 6$ , what is  $\mathbb{E}[X_{t+1} | X_t = 6]$  assuming AR(1)?

## Exercise 2: Solution

### a) Autocorrelations:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

- $\rho(0) = \gamma(0)/\gamma(0) = 1$
- $\rho(1) = \gamma(1)/\gamma(0) = 2/4 = 0.5$
- $\rho(2) = \gamma(2)/\gamma(0) = 1/4 = 0.25$

b)  $\text{Cov}(X_t, X_{t-1}) = \gamma(1) = 2$  (by stationarity, lag 1 covariance)

c)  $\text{Corr}(X_5, X_7) = \rho(|7 - 5|) = \rho(2) = 0.25$

d) For AR(1) with  $\phi = \rho(1) = 0.5$ :

$$\mathbb{E}[X_{t+1}|X_t] = \mu + \phi(X_t - \mu) = 5 + 0.5(6 - 5) = 5.5$$



## Exercise 3: Random Walk Properties

**Problem:** Consider a random walk  $X_t = X_{t-1} + \varepsilon_t$  where  $\varepsilon_t \sim WN(0, 4)$  and  $X_0 = 100$ .

Calculate:

- a)  $\mathbb{E}[X_{10}]$
- b)  $\text{Var}(X_{10})$
- c)  $\text{Cov}(X_5, X_{10})$
- d) The 95% confidence interval for  $X_{100}$
- e) After observing  $X_5 = 108$ , what is your best forecast for  $X_6$ ?

## Exercise 3: Solution

**Random walk:**  $X_t = X_0 + \sum_{i=1}^t \varepsilon_i$  with  $\sigma^2 = 4$

a)  $\mathbb{E}[X_{10}] = X_0 = 100$  (mean stays at starting value)

b)  $\text{Var}(X_{10}) = 10 \times \sigma^2 = 10 \times 4 = 40$

c)  $\text{Cov}(X_5, X_{10}) = \min(5, 10) \times \sigma^2 = 5 \times 4 = 20$

d) For  $X_{100}$ :

- $\mathbb{E}[X_{100}] = 100$ ,  $\text{Var}(X_{100}) = 400$ ,  $SD = 20$
- 95% CI:  $100 \pm 1.96 \times 20 = [60.8, 139.2]$

e) Best forecast:  $\hat{X}_6 = X_5 = 108$

(Random walk: best forecast is the last observed value)

## Python Exercise 1: Load and Plot

**Task:** Load S&P 500 data and create a basic time series plot.

### Starter Code

```
import yfinance as yf
import matplotlib.pyplot as plt

# Download S&P 500 data
sp500 = yf.download('^GSPC', start='2020-01-01', end='2025-01-01')

# TODO: Plot the closing prices
# TODO: Add title, labels, and grid
# TODO: Calculate and print basic statistics
```

### Questions:

- 1 What is the mean and standard deviation of returns?
- 2 Does the series appear stationary? Why or why not?

## Python Exercise 2: Decomposition

**Task:** Perform STL decomposition on airline passengers data.

### Starter Code

```
from statsmodels.tsa.seasonal import STL
import pandas as pd

# Load airline passengers
url = 'https://raw.githubusercontent.com/.../airline.csv'
airline = pd.read_csv(url, parse_dates=['Month'],
                      index_col='Month')

# TODO: Apply STL decomposition with period=12
# TODO: Plot all components
# TODO: What percentage of variance is explained by trend?
```

**Hint:** Use `STL(data, period=12).fit()`

## Python Exercise 3: Exponential Smoothing

**Task:** Compare SES, Holt, and Holt-Winters on real data.

### Starter Code

```
from statsmodels.tsa.holtwinters import (SimpleExpSmoothing,
    ExponentialSmoothing)

# Split data: 80% train, 20% test
train = airline[:'1958']
test = airline['1959':]

# TODO: Fit SES, Holt, and Holt-Winters
# TODO: Generate forecasts for test period
# TODO: Calculate RMSE for each method
# TODO: Which method performs best? Why?
```

## Python Exercise 4: Stationarity Testing

**Task:** Test for stationarity using ADF and KPSS tests.

### Starter Code

```
from statsmodels.tsa.stattools import adfuller, kpss

# Test S&P 500 prices
prices = sp500['Close']
returns = prices.pct_change().dropna()

# TODO: Run ADF test on prices and returns
# TODO: Run KPSS test on prices and returns
# TODO: Interpret the results

# ADF: adfuller(series)
# KPSS: kpss(series, regression='c')
```

### Questions:

- 1 Are prices stationary? Are returns stationary?
- 2 Do ADF and KPSS agree?

## Discussion Question 1

### Scenario

You are analyzing monthly sales data for a retail company. The data shows clear seasonality (high sales in December) and an upward trend. The seasonal peaks have been getting larger over time.

### Discuss:

- 1 Should you use additive or multiplicative decomposition? Why?
- 2 Which exponential smoothing method would you recommend?
- 3 How would you evaluate your forecast model?
- 4 What could go wrong if you used the wrong decomposition?

## Discussion Question 2

### Scenario

A colleague claims: "I ran the ADF test on my stock price data and got a p-value of 0.65, so my data is stationary and I can fit an ARMA model directly."

### Discuss:

- 1 What is wrong with this interpretation?
- 2 What do the ADF hypotheses actually test?
- 3 What should the colleague do before fitting an ARMA model?
- 4 How could the KPSS test help clarify the situation?



## Discussion Question 3

### Scenario

You're building a forecasting model and achieve excellent results: MAPE of 2% on your dataset. Your manager is impressed and wants to deploy the model immediately.

### Discuss:

- 1 What questions should you ask before deployment?
- 2 Did you use proper train/validation/test splits?
- 3 Could there be data leakage in your evaluation?
- 4 What additional diagnostics would you run?
- 5 How would you monitor the model in production?

## Discussion Question 4

### Scenario

You need to forecast daily electricity demand for the next week. The data shows: (1) strong daily patterns (peaks at 6pm), (2) weekly patterns (lower on weekends), and (3) annual patterns (higher in summer/winter).

### Discuss:

- 1 How would you handle multiple seasonal patterns?
- 2 Would Holt-Winters work here? Why or why not?
- 3 What's the advantage of Fourier terms in this case?
- 4 How would you set up your train/validation/test split?

## Key Takeaways from Today

- ① **Time series are dependent** – not i.i.d. like cross-sectional data
- ② **Choose decomposition wisely** – multiplicative when seasonal amplitude grows
- ③ **Understand smoothing parameters** – high  $\alpha$  = reactive, low  $\alpha$  = smooth
- ④ **Test for stationarity** – use both ADF and KPSS together
- ⑤ **Proper evaluation** – never tune on test set!
- ⑥ **Random walk is non-stationary** – variance grows with time

### Next Seminar

ARMA/ARIMA model identification, estimation, and forecasting

# Questions?

Good luck with the exercises!

Practice makes perfect.