

Séries Temporais

Fatec 2025

Definição

- dados coletados em intervalos regulares de tempo
 - Década, ano, mês, dia, hora, minuto, segundo
 - Dependência de ordem

Uma série temporal é um conjunto sequencial de pontos de dados, medido tipicamente em tempos sucessivos

- Usado para:
 - Compreender um fenômeno
 - Prever eventos

Diferença entre Análise e Previsão

Análise de Séries Temporais (Descritiva):

- • Busca entender o comportamento dos dados ao longo do tempo
- • Identifica **tendências** (crescimento ou queda ao longo do tempo), **sazonalidades** (padrões que se repetem, como estações do ano) e **ruído**
 - • **Ruído**: variações aleatórias ou imprevisíveis nos dados que não seguem um padrão claro. Exemplo: flutuações inesperadas nas vendas por causa de um evento não planejado

• Previsão (Forecasting):

- • Objetivo é estimar **valores futuros** com base nos dados passados
- • O futuro não pode ser observado diretamente, então usamos os dados históricos para **extrapolar**
 - • **Extrapolação**: processo de usar tendências observadas nos dados atuais para prever valores fora do intervalo conhecido (ou seja, no futuro)

Componentes

- Total de pessoas atendidas em um Hospital por dia

Medida

Fato

Unidade de
Tempo

Dependência da ordem

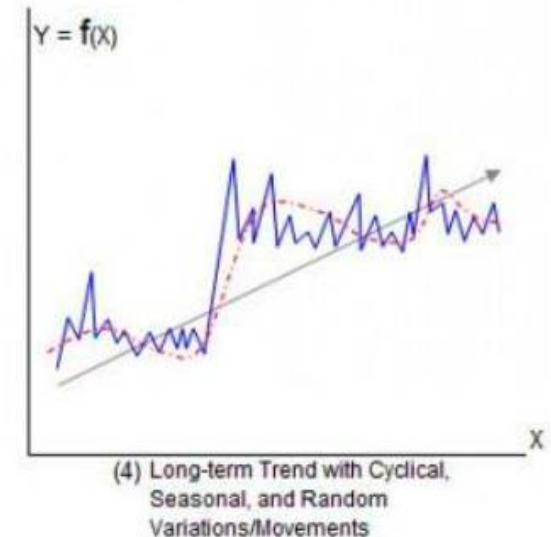
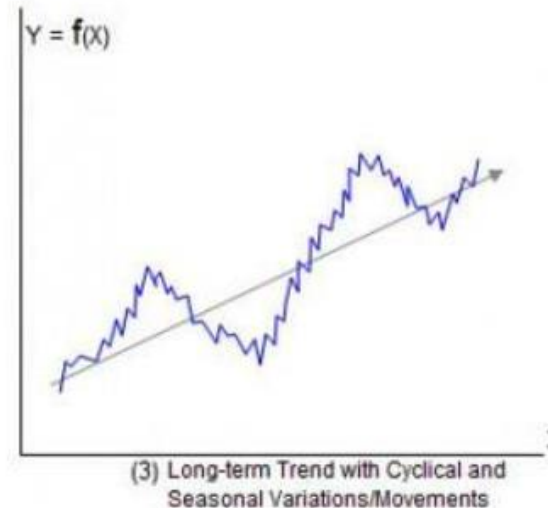
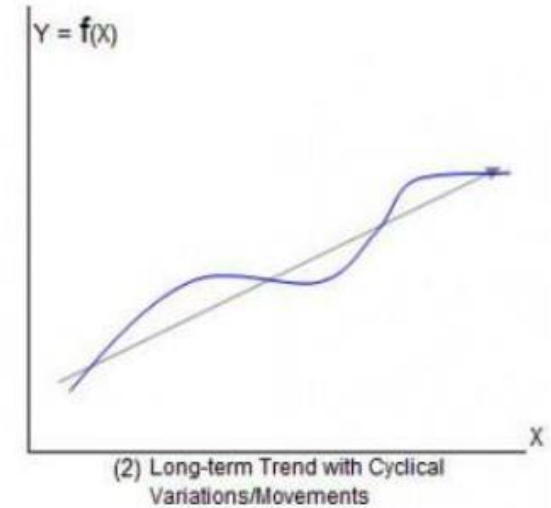
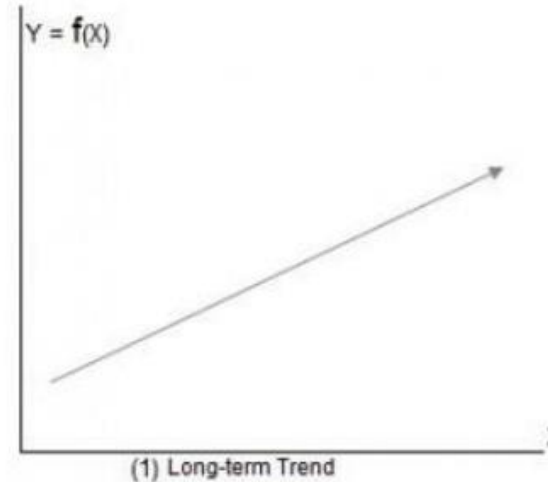
- Total de pessoas atendidas em um Hospital por dia



01/03	02/03	03/03
323	298	404

Componentes de séries temporais

- Tendência
 - Movimento de longo prazo: crescimento, queda ou estabilidade
 - Ex.: crescimento populacional
- Ciclo
 - Flutuações irregulares e de longo
 - Causa: fatores macroeconômicos ou sistêmicos
 - Frequência: irregular e imprevisível
 - Previsibilidade: baixa
 - Ex.: aumento de vendas em Dezembro
- Sazonalidade
 - Padrão regular e previsível que se repete em intervalos
 - Causas: calendário ou clima
 - Frequência: constante e conhecida
 - Previsibilidade: alta
 - Ex.: recessões e expansões econômicas
- Irregularidade ou ruído
 - Variações aleatórias sem padrão



Componentes

- Em **séries curtas**, muitas vezes **não se distingue ciclo de tendência** e o ciclo pode ser ignorado
- O **ruído** é o que sobra após remover os demais componentes
- Nem toda série tem todos os componentes
 - por exemplo, séries de dados climáticos podem ter apenas **tendência** e **sazonalidade**

Forma de representação

- **Modelo aditivo:**

$$Y(t) = T(t) + S(t) + C(t) + R(t)$$

- **Modelo multiplicativo:**

$$Y(t) = T(t) \times S(t) \times C(t) \times R(t)$$

Use o modelo aditivo quando as flutuações forem constantes em magnitude

Use o multiplicativo quando variarem proporcionalmente ao nível da série

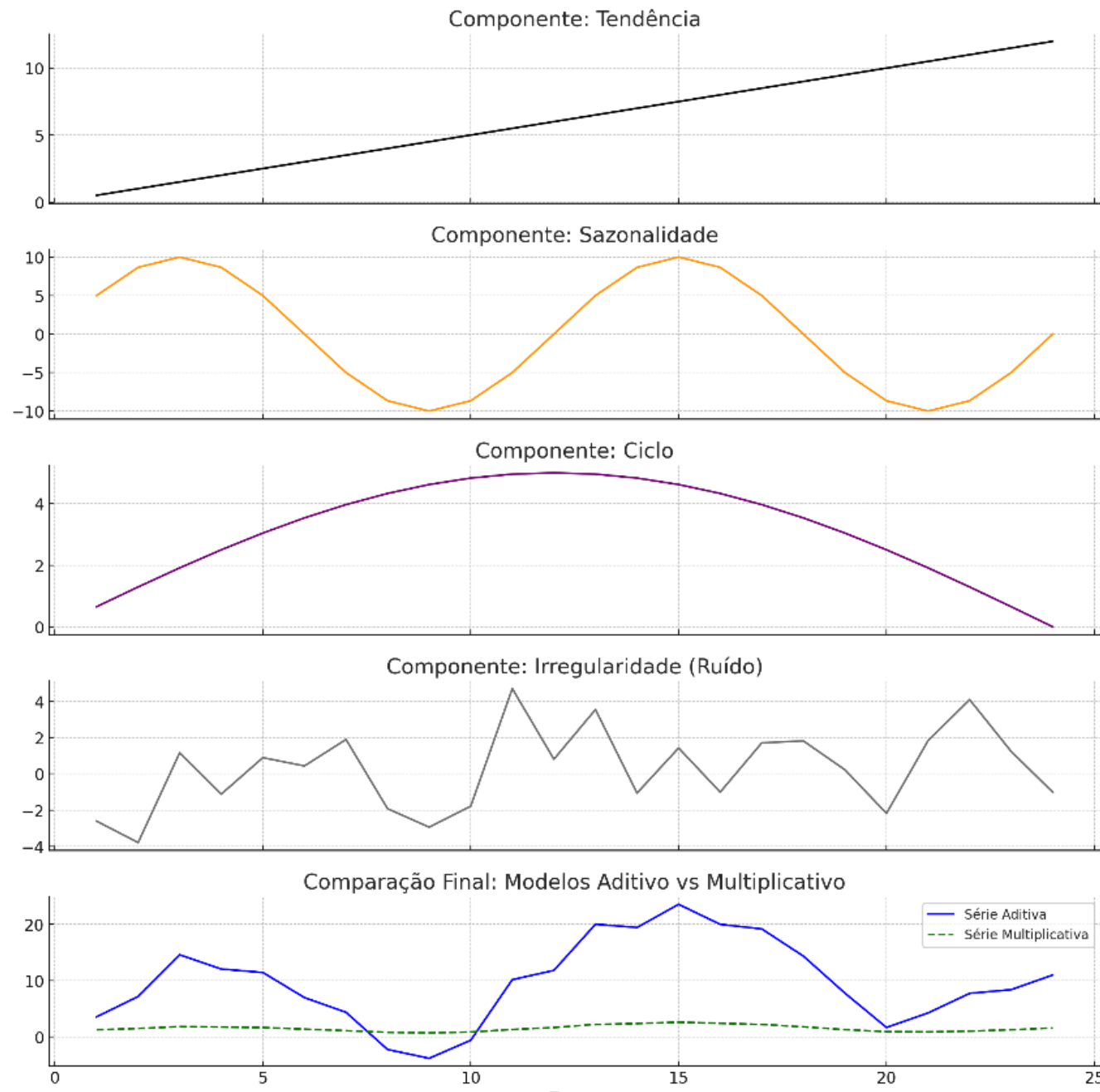
Comparando os modelos

- **Modelo Aditivo**

- Assume que os componentes da série temporal se somam
- É apropriado quando a amplitude das variações sazonais é constante ao longo do tempo

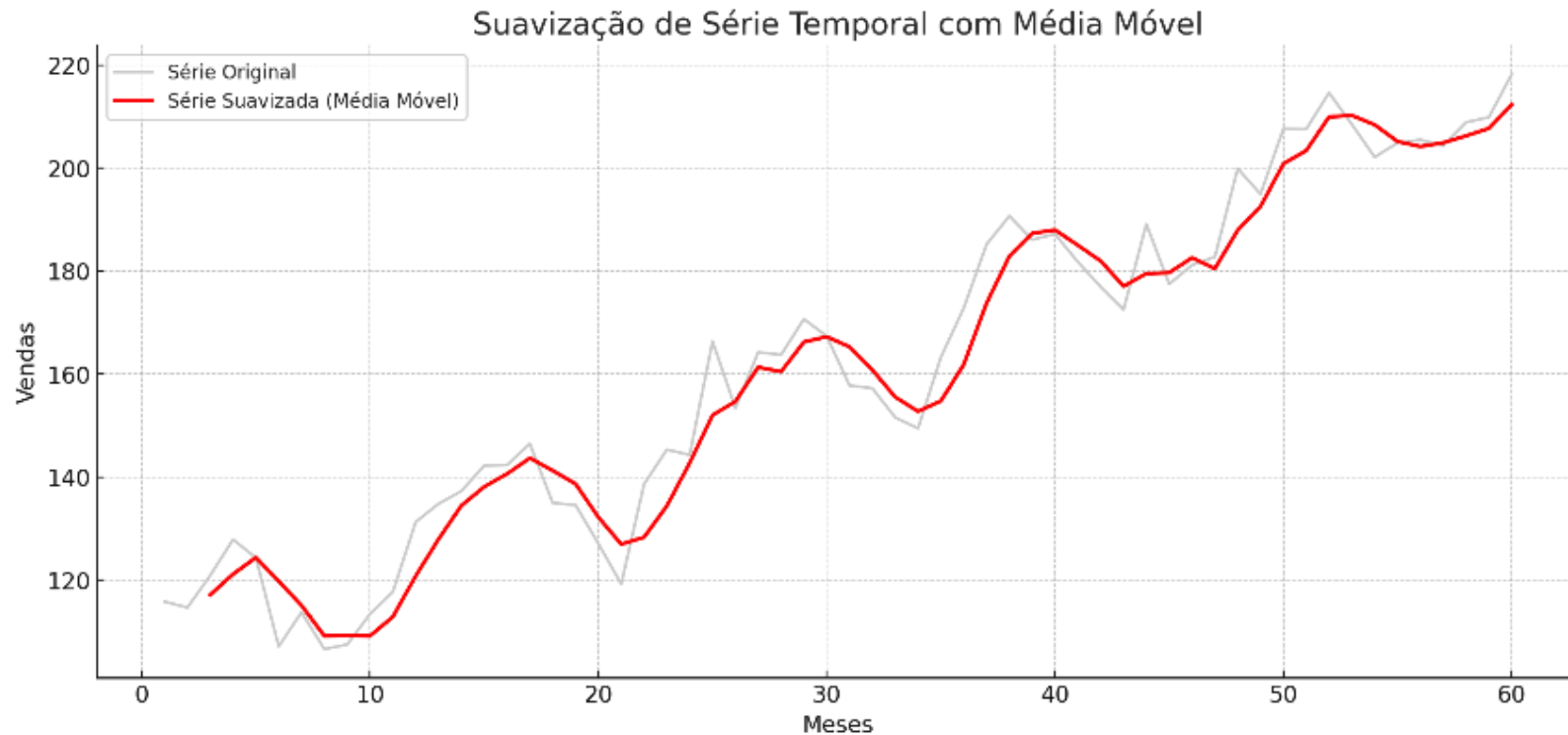
- **Modelo Multiplicativo**

- Assume que os componentes se multiplicam
 - É adequado quando a amplitude das variações sazonais aumenta ou diminui proporcionalmente ao nível da série
- Se as flutuações sazonais mantêm uma amplitude constante, o modelo aditivo é mais apropriado
 - Se as flutuações sazonais aumentam em amplitude conforme o nível da série aumenta, o modelo multiplicativo pode ser mais adequado



Média Móvel

- Técnica de suavização a fim de reduzir ruídos permitindo ver melhor as tendências e sazonalidades de uma série temporal



Média Móvel

Uma loja vende picolés e registra as vendas dos últimos 7 dias.

Dia	Vendas (unidades)
Segunda	30
Terça	28
Quarta	35
Quinta	33
Sexta	50
Sábado	60
Domingo	55

A cada dia, tiramos a **média dos 3 últimos dias** (incluindo o dia atual):

Dia	Vendas	Média Móvel (3 dias)
Segunda	30	—
Terça	28	—
Quarta	35	$(30+28+35)/3 = 31,0$
Quinta	33	$(28+35+33)/3 = 32,0$
Sexta	50	$(35+33+50)/3 = 39,3$
Sábado	60	$(33+50+60)/3 = 47,7$
Domingo	55	$(50+60+55)/3 = 55,0$

- Suaviza picos e vales (sábado e terça)
- Identifica tendências (aumento de vendas durante a semana)
- Pode ser usada para previsões curtas: aumentar o estoque aos finais de semana)

Média Móvel em R

```
# Instalar pacote, se necessário  
install.packages("zoo")
```

```
library(zoo)
```

```
# Dados de vendas por dia  
dias <- c("Seg", "Ter", "Qua", "Qui", "Sex", "Sáb", "Dom")  
vendas <- c(30, 28, 35, 33, 50, 60, 55)
```

```
# Calcular média móvel de 3 dias  
media_movel_3 <- rollmean(vendas, k = 3, align = "right", fill = NA)
```

```
# Gerar gráfico  
plot(vendas, type = "o", col = "blue", xaxt = "n",  
      xlab = "Dia da Semana", ylab = "Vendas (unidades)",  
      main = "Vendas Diárias vs Média Móvel (3 dias)",  
      ylim = c(min(vendas), max(vendas)))  
axis(1, at = 1:7, labels = dias)  
lines(media_movel_3, type = "o", col = "orange")  
# Adicionar legenda  
legend("topleft", legend = c("Vendas Diárias", "Média Móvel 3 dias"),  
      col = c("blue", "orange"), lty = 1, pch = 1)
```

Prática power Bi

- Carregar os dados
 - Criar uma coluna para o dia, mês e ano
 - Em DAX: day, month e year
- Pré-processamento:
 - Qual a menor data?
 - Qual a maior data?
 - Qual o menor valor de vendas? E o maior? E a média?

Prática Power BI

- Crie um gráfico de linhas por dia, mês e ano
- Crie a média móvel:

```
MediaMovel_3Dias :=  
AVERAGEX(  
    -- Cria uma tabela de datas de 3 dias, terminando na data atual do contexto  
    DATESINPERIOD(  
        dataset[Data],          -- Coluna de datas  
        MAX(dataset[Data]),     -- Data atual do contexto  
        -2,                     -- Pega os 2 dias anteriores + o dia atual = 3 dias  
        DAY  
    ),  
    CALCULATE(SUM(dataset[Total_vendas])) -- Soma das vendas para cada data da janela  
)
```

- Crie uma linha de tendência
- Faça a previsão para 2024

Tipos de séries temporais



Estacionárias x Não estacionárias

```
# Instalar, se necessário: install.packages("ggplot2")
library(ggplot2)
# fixa a semente de números aleatórios para os resultados sejam sempre os mesmos
set.seed(123)

# Série estacionária
tempo <- 1:100
estacionaria <- rnorm(100, mean = 10, sd = 2) # gera aleatórios com distrib normal
media_est <- mean(estacionaria)
dp_est <- sd(estacionaria) #desvio padrão, medida de dispersão

# Série não estacionária (com tendência)
nao_estacionaria <- cumsum(rnorm(100, mean = 0.5, sd = 2)) #soma acumulada
media_nao_est <- mean(nao_estacionaria)
dp_nao_est <- sd(nao_estacionaria)

# Data frame combinado
df <- data.frame(
  Tempo = rep(tempo, 2),
  Valor = c(estacionaria, nao_estacionaria),
  Tipo = rep(c("Estacionária", "Não Estacionária"), each = 100)
)

# Médias e desvios para adicionar faixas no gráfico
faixas <- data.frame(
  Tipo = c("Estacionária", "Não Estacionária"),
  media = c(media_est, media_nao_est),
  upper = c(media_est + dp_est, media_nao_est + dp_nao_est),
  lower = c(media_est - dp_est, media_nao_est - dp_nao_est)
)
```


Estacionárias x Não estacionárias

```
# Plot com ggplot2 e faixas de variância
ggplot(df, aes(x = Tempo, y = Valor)) +
  geom_line(color = "steelblue") +
  facet_wrap(~Tipo, scales = "free_y") +
  geom_hline(data = faixas, aes(yintercept = media), linetype = "dashed", color = "darkred") +
  geom_hline(data = faixas, aes(yintercept = upper), linetype = "dotted", color = "orange") +
  geom_hline(data = faixas, aes(yintercept = lower), linetype = "dotted", color = "orange") +
  labs(title = "Séries Temporais: Estacionária vs Não Estacionária",
        subtitle = "Linhas: média (vermelha tracejada) e  $\pm 1$  desvio padrão (laranja pontilhada)",
        y = "Valor", x = "Tempo") +
  theme_minimal()
```

Linear x Não Linear

```
# Definir uma semente para reprodutibilidade
set.seed(123)

# Período de tempo
tempo <- 1:100

# Série linear:  $y = a \cdot t + \text{erro}$ 
serie_linear <- 0.5 * tempo + rnorm(100, mean = 0, sd = 2)

# Série não linear:  $y = a \cdot t^2 + \text{erro}$ 
serie_nao_linear <- 0.05 * tempo^2 + rnorm(100, mean = 0, sd = 20)

# Criar os gráficos lado a lado
par(mfrow = c(1, 2)) # Dois gráficos na mesma janela

# Gráfico da série linear
plot(tempo, serie_linear, type = "l", col = "blue", lwd = 2,
     main = "Série Temporal Linear", xlab = "Tempo", ylab = "Valor")
abline(lm(serie_linear ~ tempo), col = "red", lty = 2) # Regressão linear

# Gráfico da série não linear
plot(tempo, serie_nao_linear, type = "l", col = "darkgreen", lwd = 2,
     main = "Série Temporal Não Linear", xlab = "Tempo", ylab = "Valor")
lines(lowess(tempo, serie_nao_linear), col = "orange", lty = 2) # Suavização local

par(mfrow = c(1,1)) # Reset layout
```

- **Série Linear:** A tendência segue uma linha reta, mesmo com ruído
- **Série Não Linear:** A tendência é curva, não pode ser ajustada adequadamente por uma reta
- **lm() e lowess():** usados para mostrar a tendência linear e suavizada, respectivamente

Série Sazonal e não sazonal

```
# Carregar a base AirPassengers (já vem com o R)
data("AirPassengers")

# Criar série não sazonal simulada com tendência e ruído
set.seed(123)
tempo <- 1:length(AirPassengers)
serie_nao_sazonal <- 2 * tempo + rnorm(length(tempo), mean = 0, sd = 20)
ts_nao_sazonal <- ts(serie_nao_sazonal, start = c(1949, 1), frequency = 12)

# Gráficos lado a lado: Sazonal x Não Sazonal
par(mfrow = c(1, 2))

# Série Sazonal: AirPassengers
plot(AirPassengers, main = "Série Sazonal: AirPassengers",
     col = "darkgreen", lwd = 2, ylab = "Nº de Passageiros", xlab = "Ano")

# Série Não Sazonal
plot(ts_nao_sazonal, main = "Série Não Sazonal (Simulada)",
     col = "blue", lwd = 2, ylab = "Valor Simulado", xlab = "Ano")

# Reset layout
par(mfrow = c(1, 1))
```

Série Sazonal e não sazonal

```
# Decomposição da série AirPassengers (aditiva)
# Como AirPassengers tem crescimento exponencial, aplicamos log para aproximar da aditividade
log_air <- log(AirPassengers)
decomp <- decompose(log_air, type = "additive")
```

```
# Gráfico da decomposição
plot(decomp, col = "darkred", xlab = "Ano")
```

- **Painel 1 (Observado):** Série original, já com crescimento e variação sazonal.
- **Painel 2 (Tendência):** Mostra o crescimento do número de passageiros ao longo do tempo.
- **Painel 3 (Sazonalidade):** Repetição **cíclica mensal**, destacando a **alta temporada**.
- **Painel 4 (Resíduo):** Variações que **não são explicadas** pela tendência nem pela sazonalidade.

Univariada x Multivariada

```
# Dataset real disponível no R
data("airquality")

# Preprocessamento: remover NAs e criar um vetor de datas
dados <- na.omit(airquality) # remove linhas com NA
tempo <- seq.Date(from = as.Date("1973-05-01"), by = "day", length.out = nrow(dados))

# Criar série univariada (apenas Ozone)
ozone_ts <- ts(dados$Ozone, start = c(1973, 5), frequency = 30) # aprox. diário

# Criar série multivariada com 4 variáveis
multi_ts <- ts(cbind(Ozone = dados$Ozone,
                     Solar = dados$Solar.R,
                     Wind = dados$Wind,
                     Temp = dados$Temp),
               start = c(1973, 5), frequency = 30)

# Gráficos lado a lado
par(mfrow = c(1, 2))

# Série Univariada
```

Univariada x Multivariada

Série Univariada

```
plot(ozone_ts, col = "blue", lwd = 2,  
     main = "Série Temporal Univariada",  
     ylab = "Ozone (ppb)", xlab = "Dias desde Maio/1973")  
legend("topright", legend = "Ozone", col = "blue", lty = 1, bty = "n")  
text(1973.3, max(ozone_ts), "1 variável: Ozone", col = "blue")
```

Série Multivariada

```
plot(multi_ts, main = "Série Temporal Multivariada",  
     col = c("blue", "orange", "darkgreen", "red"))  
legend("topright", legend = c("Ozone", "Solar.R", "Wind", "Temp"),  
     col = c("blue", "orange", "darkgreen", "red"), lty = 1, bty = "n")  
  
par(mfrow = c(1, 1)) # Reset layout
```