

# Revisão e Notas sobre Séries Temporais

Pedro Rupf Pereira Viana

10 de fevereiro de 2026

# 1 Introdução e Objetivos

A análise de séries temporais representa um dos pilares mais consolidados e aplicados da estatística moderna, dedicando-se ao estudo de fenômenos que se manifestam e evoluem ao longo do tempo. Diferentemente das abordagens estatísticas tradicionais, que frequentemente tratam as observações como independentes e coletadas em um único instante ou período fixo, a análise temporal reconhece explicitamente a **ordem cronológica** como elemento estruturante central dos dados. Essa ordenação não é mero detalhe organizacional: ela introduz uma dependência intrínseca entre as observações sucessivas, de modo que o valor registrado em um dado momento tende a estar influenciado — direta ou indiretamente — pelos valores anteriores e pode, por sua vez, condicionar os valores futuros.

Formalmente, uma série temporal pode ser definida como um conjunto de observações sequenciais de uma ou mais variáveis, registradas em intervalos de tempo regulares ou (mais raramente) irregulares, e ordenadas de forma cronológica estrita. Essa sequência temporal permite capturar a dinâmica evolutiva de processos reais, revelando padrões que escapariam a análises estáticas. A importância dessa abordagem reside no fato de que a maioria dos fenômenos socioeconômicos, climáticos, biológicos, industriais e financeiros apresenta forte componente temporal: ignorá-la equivaleria a desconsiderar a própria essência do processo em estudo.

## 2 Desenvolvimento

### 2.1 Corte transversal, Painel e Séries Temporais

Para compreender adequadamente o escopo das séries temporais, é instrutivo contrastá-las com outras estruturas de dados comuns na estatística aplicada.

Os dados de **corte transversal** (ou *cross-sectional data*) consistem em observações coletadas sobre múltiplas unidades (indivíduos, empresas, regiões, países etc.) em um único instante ou período curto de tempo. Nessa estrutura, o tempo é fixo ou irrelevante, e o interesse recai nas diferenças entre as unidades no momento da coleta. Exemplos clássicos incluem uma pesquisa de renda domiciliar realizada em um único mês em diferentes municípios ou o levantamento das características financeiras de várias empresas em um balanço anual específico. Nesses casos, assume-se geralmente independência entre as observações (salvo estratificação ou agrupamento explícito), e as técnicas estatísticas predominantes são regressão linear, análise de variância ou métodos multivariados de redução de dimensionalidade.

Já os dados em **painel** (ou *panel data*, também chamados de dados longitudinais) combinam as duas dimensões anteriores: acompanham as mesmas unidades ao longo de múltiplos períodos de tempo. Essa estrutura permite controlar efeitos individuais fixos (heterogeneidade não observada) e capturar tanto variações intertemporais quanto diferenças entre unidades. Exemplos incluem o acompanhamento anual da produtividade de um conjunto fixo de fábricas ao longo de uma década ou o monitoramento longitudinal de indicadores de saúde em um painel de indivíduos. Os modelos para dados em painel (efeitos fixos, efeitos aleatórios, modelos dinâmicos etc.) exploram tanto a variação dentro das unidades quanto entre elas, oferecendo maior robustez inferencial em comparação com cortes transversais puros.

Por fim, as **séries temporais** puras concentram-se exclusivamente na dimensão tem-

poral, focando em uma (ou poucas) variáveis observadas repetidamente ao longo do tempo para a mesma unidade ou processo. Aqui, a ênfase recai na dependência serial (autocorrelação), na evolução dinâmica e nos padrões que emergem da sequência cronológica. A independência entre observações — pressuposto central em grande parte da estatística clássica — é sistematicamente violada, exigindo métodos específicos que incorporem explicitamente essa estrutura de dependência.

## 2.2 Tipos de Séries Temporais: Univariadas vs. Multivariadas

Uma distinção fundamental na análise de séries temporais refere-se ao número de variáveis envolvidas.

- **Séries temporais univariadas:** focam em uma única variável aleatória  $\{X_t\}$ . O objetivo central é modelar a estrutura de dependência interna do processo, explorando como  $X_t$  depende de seus próprios valores passados e de choques aleatórios passados. Modelos clássicos incluem os processos  $AR_{(p)}$ ,  $MA_{(q)}$  e  $ARMA_{(p,q)}$ , nos quais a dinâmica é governada exclusivamente pela história da própria série.

Exemplos:

- Índice mensal de preços ao consumidor (IPC);
- Preço diário de fechamento de uma ação;
- Consumo horário de energia elétrica;
- Temperatura média mensal em uma estação meteorológica.

- **Séries temporais multivariadas:** envolvem um vetor de variáveis aleatórias  $\{\mathbf{X}_t = (X_{1t}, X_{2t}, \dots, X_{kt})^\top\}$ . O interesse reside não apenas na dinâmica interna de cada componente, mas também nas inter-relações dinâmicas entre elas (efeitos cruzados, feedbacks, co-movimentos). Modelos representativos incluem vetores autotregressivos ( $VAR_{(p)}$ ) e modelos de correção de erros vetoriais (VECM) em presença de cointegração.

Exemplos:

- Sistema macroeconômico com PIB, inflação, taxa de juros e desemprego;
- Retornos de múltiplas ações de um mesmo setor;
- Indicadores climáticos multivariados (temperatura, precipitação, umidade).

## 3 Componentes clássicos de uma Série Temporal

A decomposição clássica de séries temporais constitui um dos pilares fundamentais da análise temporal, oferecendo uma estrutura conceitual intuitiva e amplamente utilizada para compreender a dinâmica subjacente aos dados observados ao longo do tempo. Essa abordagem, cujas raízes remontam ao início do século XX (com contribuições significativas de autores como Persons, 1919, e posteriormente sistematizada em métodos de suavização e decomposição), postula que qualquer série temporal observada pode ser decomposta em quatro componentes principais: **tendência**, **sazonalidade**, **ciclos** e **componente irregular** (também denominada resíduo, ruído ou variação aleatória). Esses componentes

interagem de forma sistemática para gerar o comportamento global da série, permitindo que o analista isole padrões previsíveis de flutuações imprevisíveis.

Formalmente, uma série temporal observada  $y_t$  (para  $t = 1, 2, \dots, T$ ) é modelada como uma combinação desses elementos, onde cada componente captura uma fonte distinta de variação:

- **Tendência ( $T_t$ ):** Representa o movimento de longo prazo da série, caracterizando a direção geral e persistente dos dados ao longo de períodos extensos (geralmente vários anos ou décadas). A tendência reflete forças estruturais profundas, como crescimento econômico secular, avanços tecnológicos, mudanças demográficas, urbanização progressiva ou alterações climáticas de longo alcance. Matematicamente,  $T_t$  é tipicamente uma função suave e monotonicamente crescente, decrescente ou aproximadamente constante no horizonte de análise. Exemplos clássicos incluem o aumento secular do PIB per capita em economias desenvolvidas, a ascensão histórica da expectativa de vida média global ou a tendência ascendente das concentrações atmosféricas de  $\text{CO}_2$  medidas na estação de Mauna Loa desde 1958.
- **Sazonalidade ( $S_t$ ):** Corresponde a flutuações regulares, previsíveis e repetitivas que ocorrem em intervalos fixos e constantes (diários, semanais, mensais, trimestrais ou anuais). Essas variações são impulsionadas por fatores calendarísticos, institucionais ou comportamentais recorrentes, como estações do ano, feriados, ciclos escolares, padrões de consumo sazonal ou horários de pico de demanda. A sazonalidade apresenta periodicidade fixa (denotada por  $m$ , o período sazonal:  $m = 12$  para dados mensais,  $m = 4$  para trimestrais,  $m = 7$  para diários semanais etc.) e, em geral, mantém amplitude e forma relativamente constantes ao longo dos anos. Exemplos incluem o pico anual de vendas de sorvetes no verão brasileiro, o aumento do consumo de energia elétrica no final da tarde (pico diário), maior incidência de gripes e resfriados nos meses de inverno ou o incremento das vendas varejistas em dezembro devido ao Natal.
- **Ciclos ( $C_t$ ):** Referem-se a oscilações de duração mais longa e menos previsível que a sazonalidade, geralmente associadas a expansões e contrações econômicas ou setoriais. Diferentemente da sazonalidade, os ciclos não possuem periodicidade fixa ou regular; sua duração varia tipicamente entre 2 e 10 anos (como os ciclos de negócios descritos por Burns & Mitchell, 1946), podendo estender-se a décadas em contextos específicos (ex.: ciclos de Kondratieff ou ondas longas tecnológicas). Esses movimentos são frequentemente ligados a fatores macroeconômicos (investimento, crédito, inovação) ou eventos globais. Exemplos incluem as recessões econômicas periódicas (como as de 2008-2009 ou 2020), os ciclos de boom e bust no mercado imobiliário ou as flutuações de longo prazo nos preços de commodities agrícolas.
- **Componente irregular ( $I_t$ ):** Engloba todas as variações residuais que não podem ser atribuídas às componentes anteriores. Trata-se de flutuações estocásticas, choques imprevisíveis, erros de medição, eventos extraordinários (guerras, pandemias, desastres naturais, crises financeiras repentinas) ou ruído aleatório genuíno. Essa componente é assumida como aleatória, com média zero e sem padrão sistemático identificável (idealmente, ruído branco ou próximo disso). Exemplos incluem o impacto imediato da pandemia de COVID-19 sobre o PIB trimestral de 2020, variações abruptas no preço do petróleo devido a conflitos geopolíticos ou erros de coleta de dados.

A decomposição clássica pode ser expressa em duas formas principais, dependendo da natureza da interação entre os componentes:

**1. Modelo aditivo:**

$$y_t = T_t + S_t + C_t + I_t$$

Nesse modelo, as componentes são somadas, implicando que a amplitude das variações sazonais e cíclicas permanece aproximadamente constante independentemente do nível da tendência. É adequado para séries em que as flutuações sazonais têm magnitude similar ao longo do tempo (ex.: temperaturas médias mensais, que variam em torno de  $\pm 10\text{--}15^\circ\text{C}$  independentemente da tendência de aquecimento global de longo prazo).

**2. Modelo multiplicativo:**

$$y_t = T_t \times S_t \times C_t \times I_t$$

Aqui, as componentes interagem de forma proporcional, de modo que a amplitude das variações sazonais e cíclicas cresce (ou diminui) à medida que o nível da tendência aumenta. Esse modelo é mais apropriado para séries econômicas ou comerciais em que as flutuações percentuais são relativamente constantes (ex.: vendas varejistas que crescem ao longo dos anos, mas mantêm picos sazonais proporcionais ao volume médio; ou preços de ações, onde a volatilidade tende a ser proporcional ao nível do preço). Frequentemente, aplica-se transformação logarítmica para converter o modelo multiplicativo em aditivo:

$$\log(y_t) = \log(T_t) + \log(S_t) + \log(C_t) + \log(I_t)$$

A escolha entre aditivo e multiplicativo depende principalmente da análise exploratória preliminar: se a amplitude das oscilações sazonais aumenta com o nível da série, o modelo multiplicativo é preferível; se a variação permanece estável, o aditivo é mais adequado. Em muitos softwares e pacotes de análise de séries temporais, a decomposição é realizada automaticamente com base em critérios de ajuste e diagnóstico, mas a compreensão conceitual desses modelos é essencial para interpretar corretamente os resultados e escolher a abordagem mais apropriada para cada contexto específico.

## 4 Análise Exploratória de Séries Temporais

A análise exploratória de dados (*Exploratory Data Analysis* – EDA) constitui a etapa inicial e indispensável em qualquer estudo de séries temporais. Diferentemente da análise estatística inferencial ou da modelagem preditiva, que buscam estimação de parâmetros ou previsão, a EDA visa compreender os padrões dominantes, identificar anomalias, detectar possíveis quebras estruturais e orientar as escolhas metodológicas subsequentes. Em séries temporais, essa etapa ganha relevância adicional devido à dependência serial inerente aos dados: ignorar ou subestimar padrões como tendência, sazonalidade ou heterocedasticidade pode comprometer toda a análise posterior.

A EDA de séries temporais combina ferramentas descritivas, visuais e sumárias, permitindo que o pesquisador passe de uma visão superficial dos números brutos para uma

compreensão estruturada do processo gerador subjacente. Seu objetivo primordial é responder a questões fundamentais: Há tendência clara de longo prazo? Existe sazonalidade pronunciada? Os padrões são estáveis ao longo do tempo? Há outliers ou quebras de regime? A série parece estacionária ou requer transformações? Essas respostas guiam a seleção de modelos (ex.: ARIMA vs. SARIMA vs. suavização exponencial), a escolha de transformações (logaritmo, diferenciação) e a detecção de problemas como multicolinearidade em séries multivariadas.

## 4.1 Principais Ferramentas e Técnicas de Análise Exploratória

1. **Gráfico da Série Bruta (Time Plot ou Line Plot):** O ponto de partida clássico é o gráfico da série observada  $y_t$  em função do tempo  $t$ . Esse plot revela de forma imediata a presença de tendência (crescente, decrescente ou estacionária), sazonalidade (oscilações regulares), ciclos (ondas de duração intermediária) e possíveis intervenções (quebras abruptas, choques ou outliers).

Recomenda-se plotar a série em escala linear e, quando apropriado, em escala logarítmica para avaliar se as variações são aditivas ou multiplicativas. Em séries de alta frequência (diária ou horária), zoom em subperíodos ajuda a identificar padrões intra-dia ou intra-semana. Exemplo: no consumo de energia elétrica horário, o gráfico bruto frequentemente mostra tendência ascendente de longo prazo (crescimento econômico), sazonalidade anual (picos no verão devido ao ar-condicionado) e sazonalidade diária (picos no final da tarde).

2. **Decomposição Clássica da Série:** A decomposição em tendência, sazonalidade e resíduo (ou irregular) permite isolar visualmente cada componente. Métodos clássicos incluem a decomposição aditiva ou multiplicativa via médias móveis (como no método de Cleveland ou o `decompose()` do *R base*). Versões mais robustas, como STL (Seasonal-Trend decomposition using LOESS), acomodam sazonalidade evolutiva e tendências não lineares, sendo preferíveis em séries longas ou com mudanças graduais nos padrões. A inspeção dos resíduos da decomposição é crucial: se os resíduos exibirem padrões sistemáticos (ex.: autocorrelação remanescente ou heteroscedasticidade), indica que a decomposição capturou apenas parte da estrutura, sugerindo necessidade de modelagem mais sofisticada.

### 3. Gráficos Sazonais e Sub-sazonais:

- **Gráficos sazonais:** Sobreposição das séries de cada ciclo sazonal (ex.: linhas para cada ano, alinhadas por mês) para avaliar a estabilidade da sazonalidade ao longo do tempo.
- **Gráficos Sub-sazonais:** agrupa os valores por mês, trimestre ou dia da semana, exibindo boxplots ou violinos. Isso destaca a magnitude e variabilidade da sazonalidade, identifica meses atípicos e revela se a amplitude sazonal muda com o tempo.

Exemplo: em vendas varejistas mensais, boxplots por mês frequentemente mostram picos em dezembro (Natal) e quedas em fevereiro, com variância maior nos meses de alta demanda.