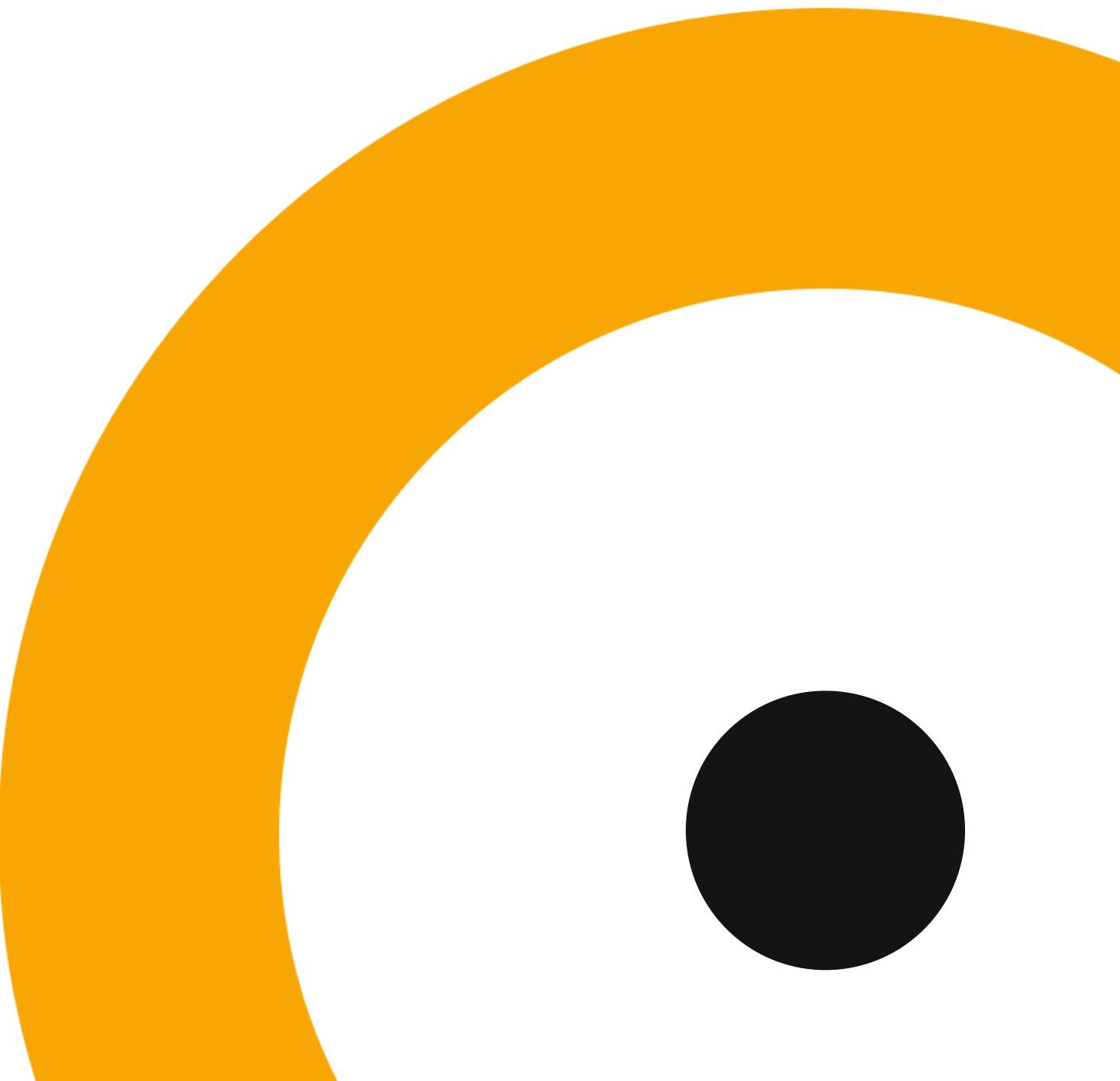


Mestrado em Data Science

Séries Temporais

Definições e Propriedades



Conteúdo

Definição de série temporal,
componentes básicas e propriedades
importantes

2.1

Definição de Série Temporal

2.2

Componentes Básicas

2.3

Propriedades Relevantes

2.1

Séries Temporais

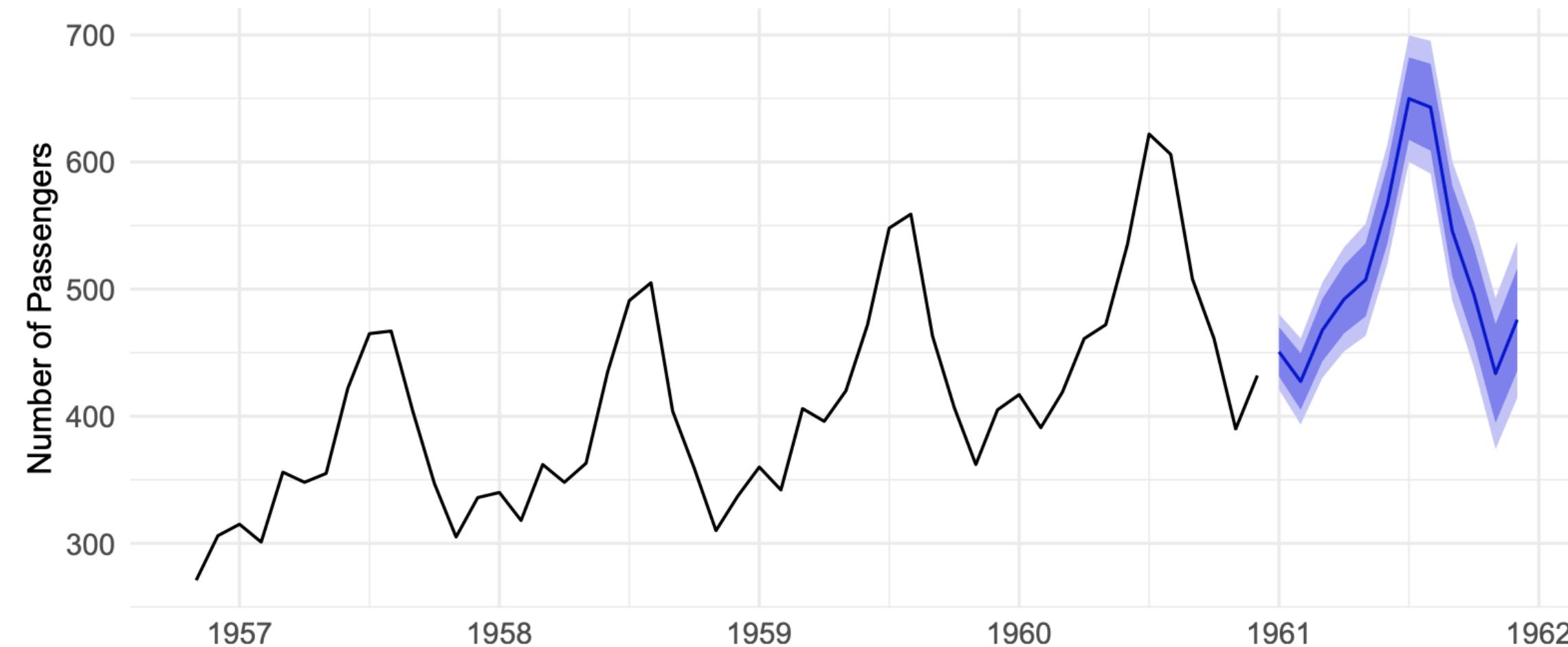
Definição e Objetivos



Série Temporal

Definição

Uma série temporal Y é uma sequência ordenada no tempo com valores $Y = y_1, y_2, \dots, y_n$, onde y_i é o valor no instante i .



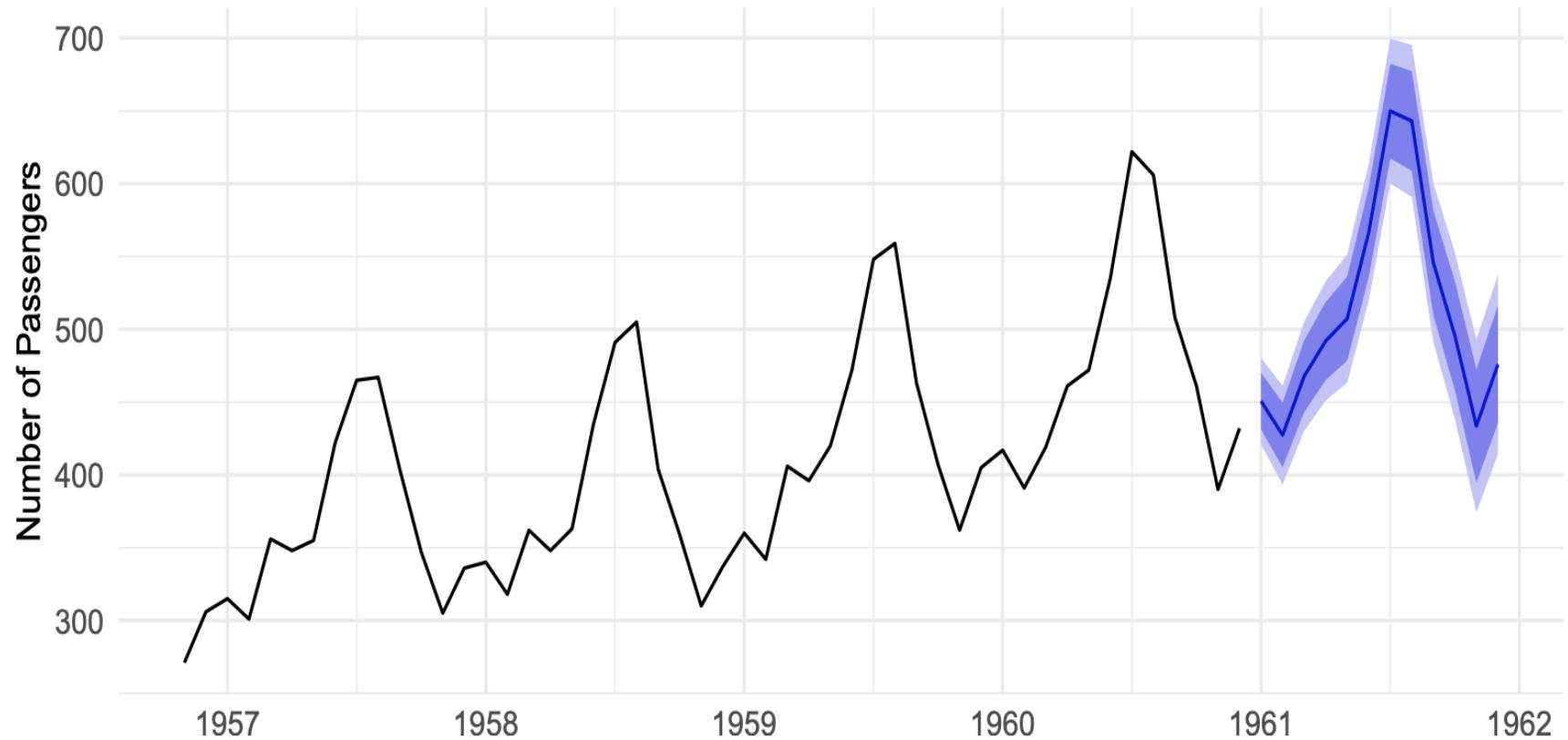
Premissas e Aspetos Essenciais

Premissas

1. Observações são colecionadas em períodos regulares, por exemplo diariamente;
2. Observações são numéricas.

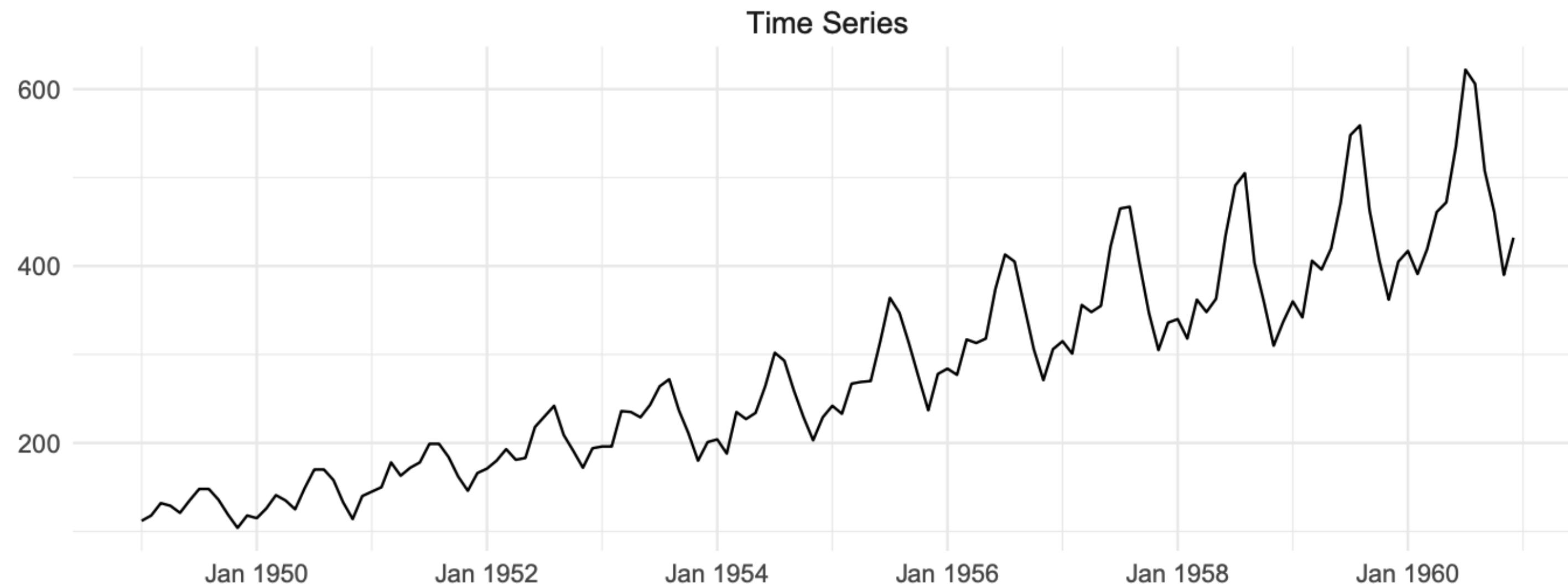
Aspetos Essenciais

1. Dependência Temporal
2. Susceptibilidade à Mudança



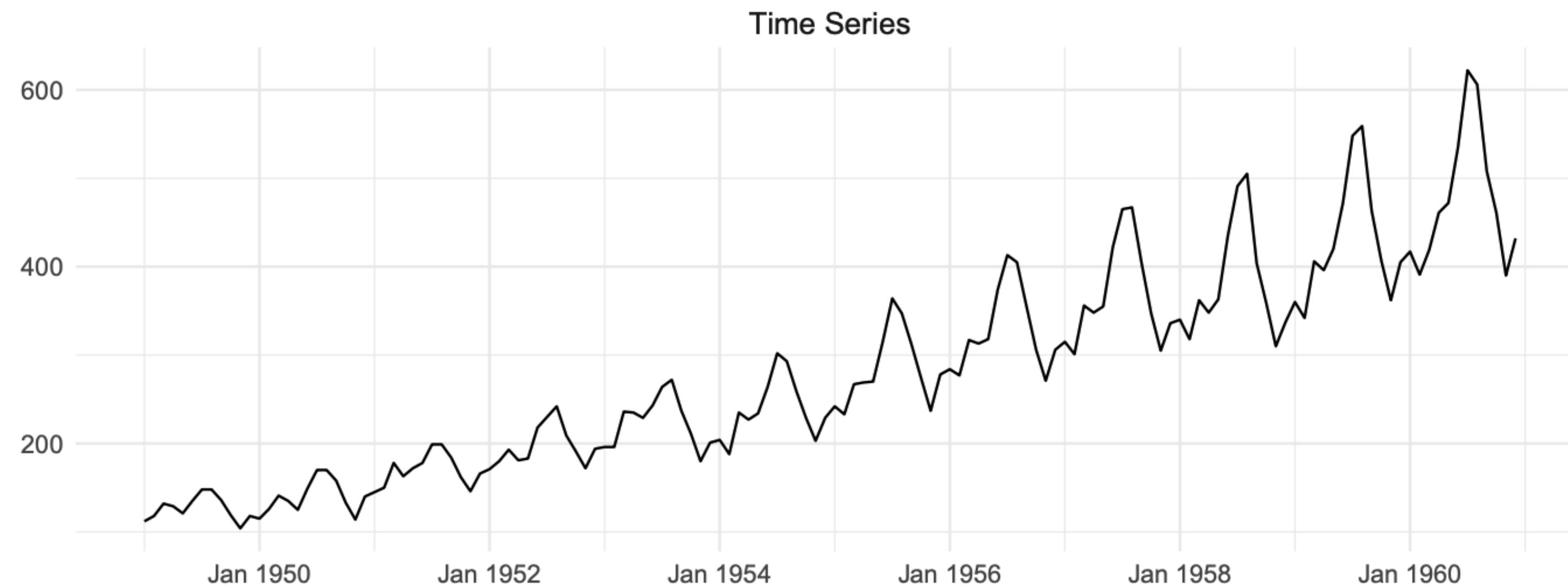
Objetivos da Análise de Séries

1. Descrever características ou padrões relevantes, por exemplo a **Tendência ou sazonalidade**



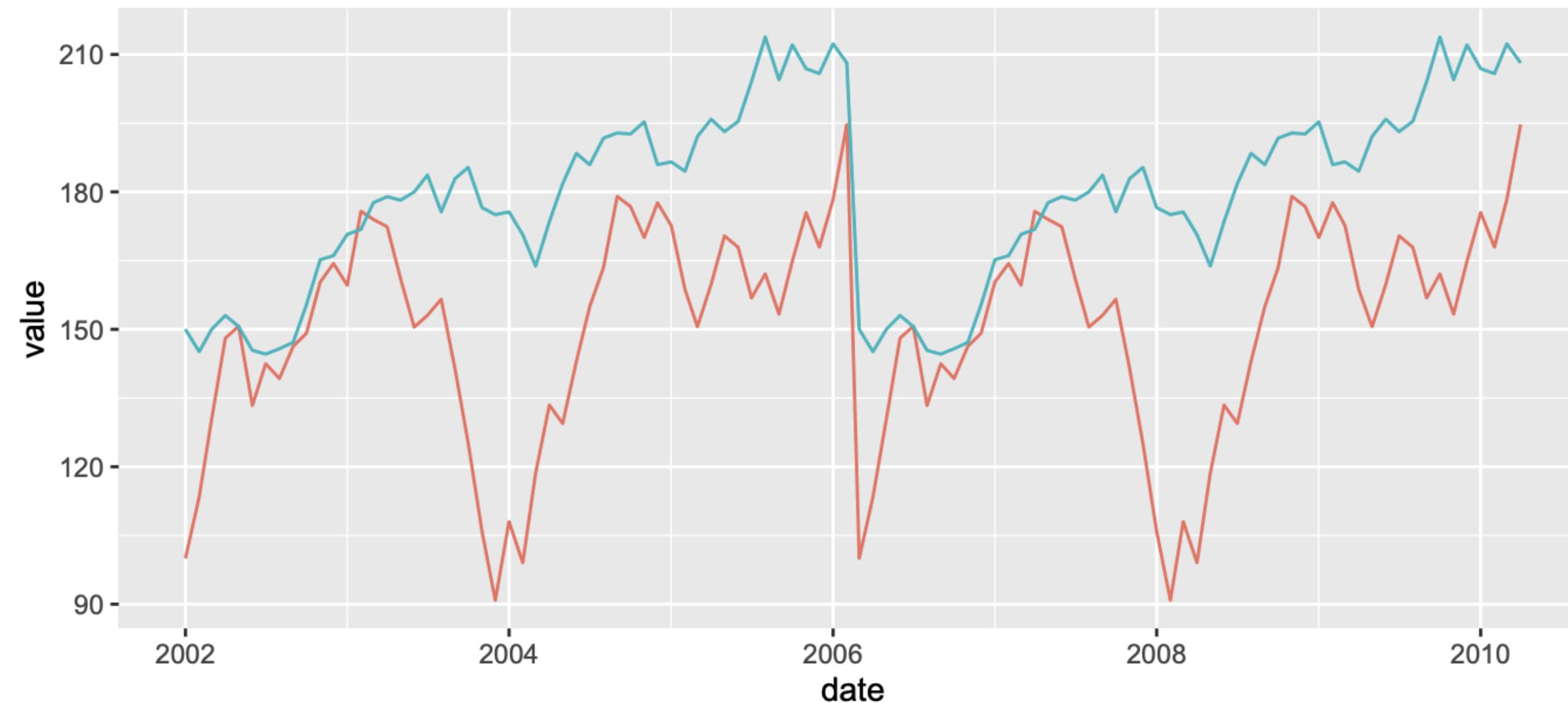
Objetivos da Análise de Séries Temporais

2. Explicar como o passado afeta o futuro da série



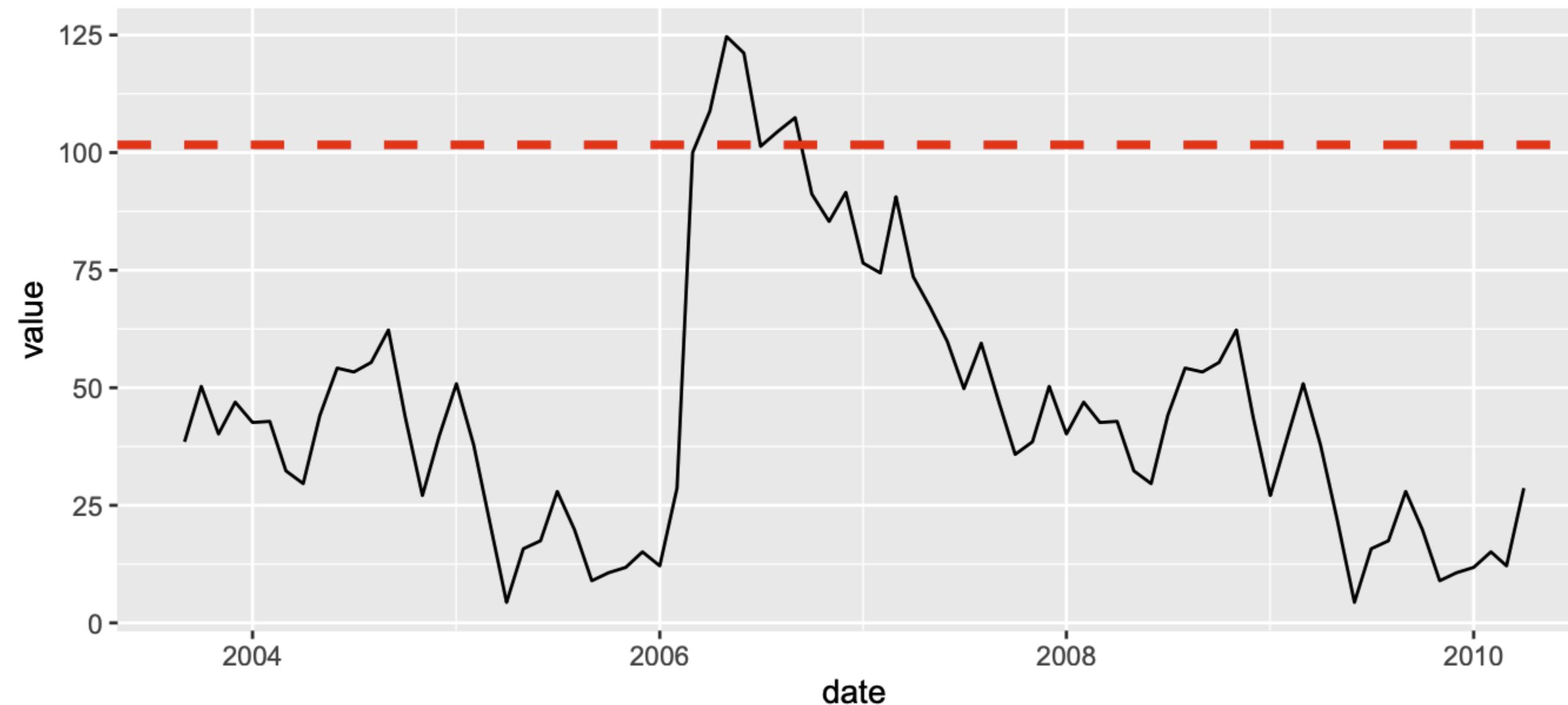
Objetivos da Análise de Séries Temporais

3. Explicar como duas séries se comportam ao longo do tempo



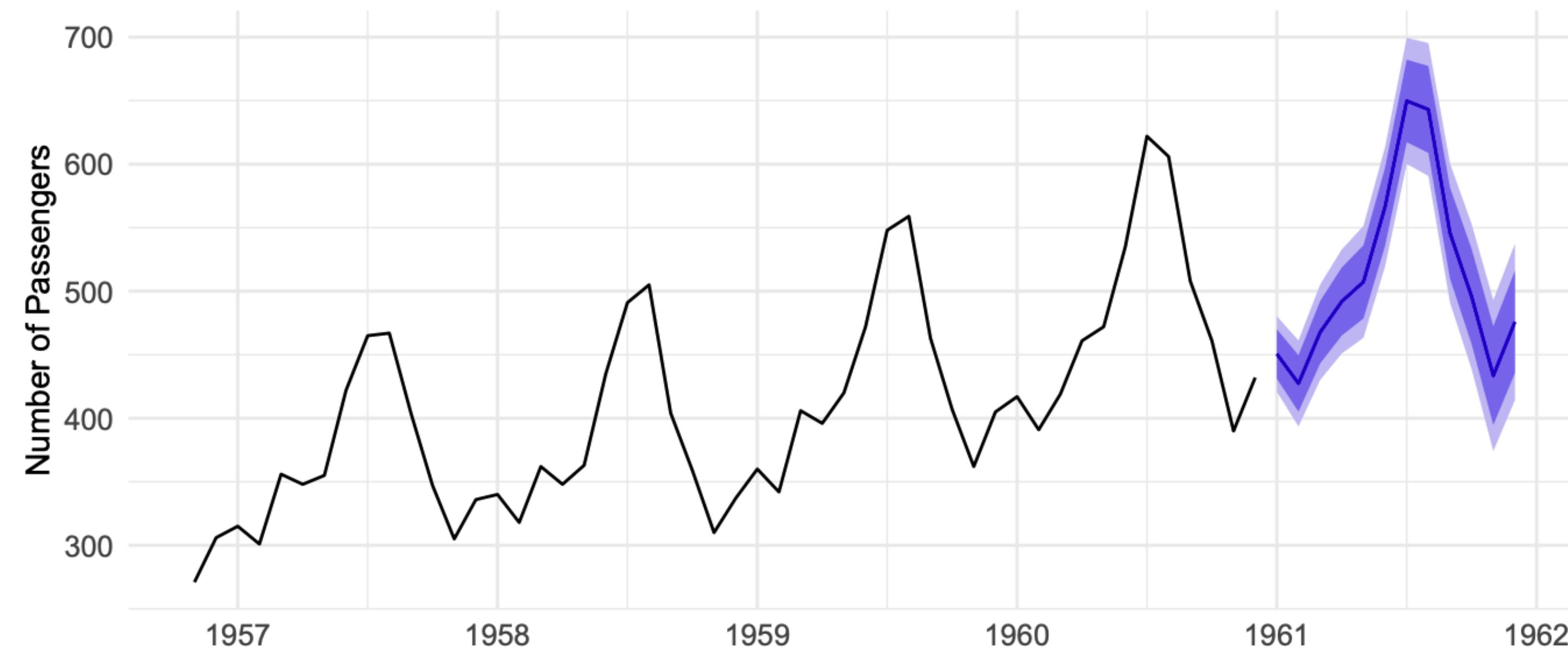
Objetivos da Análise de Séries

4. Componentes Temporais



Objetivos da Análise de Séries Temporais

5. Previsão e Validação



Objetivos da Análise de Séries

Temporais

1. Descrever características ou padrões relevantes, por exemplo a tendência ou sazonalidade
2. Explicar como o passado afeta o futuro da série
3. Explicar como duas séries interagem ao longo do tempo
4. Controlo de processos de séries temporais
5. Previsão de valores futuros

Objectives of Time Series Analysis

1. Descrever características ou padrões relevantes, por exemplo a tendência ou sazonalidade
2. Explicar como o passado afeta o futuro da série
3. Explicar como duas séries interagem ao longo do tempo
4. Controlo de processos de séries temporais
5. Previsão de valores futuros

Essencialmente, entender como a série evolui ao longo do tempo para melhorar os processos de decisão

2.2

Componentes Básicas

Tendência, Sazonalidade, e mais



Componentes Básicas

Tendência

Mudança do nível médio da série a longo-prazo

Sazonalidade e Ciclos

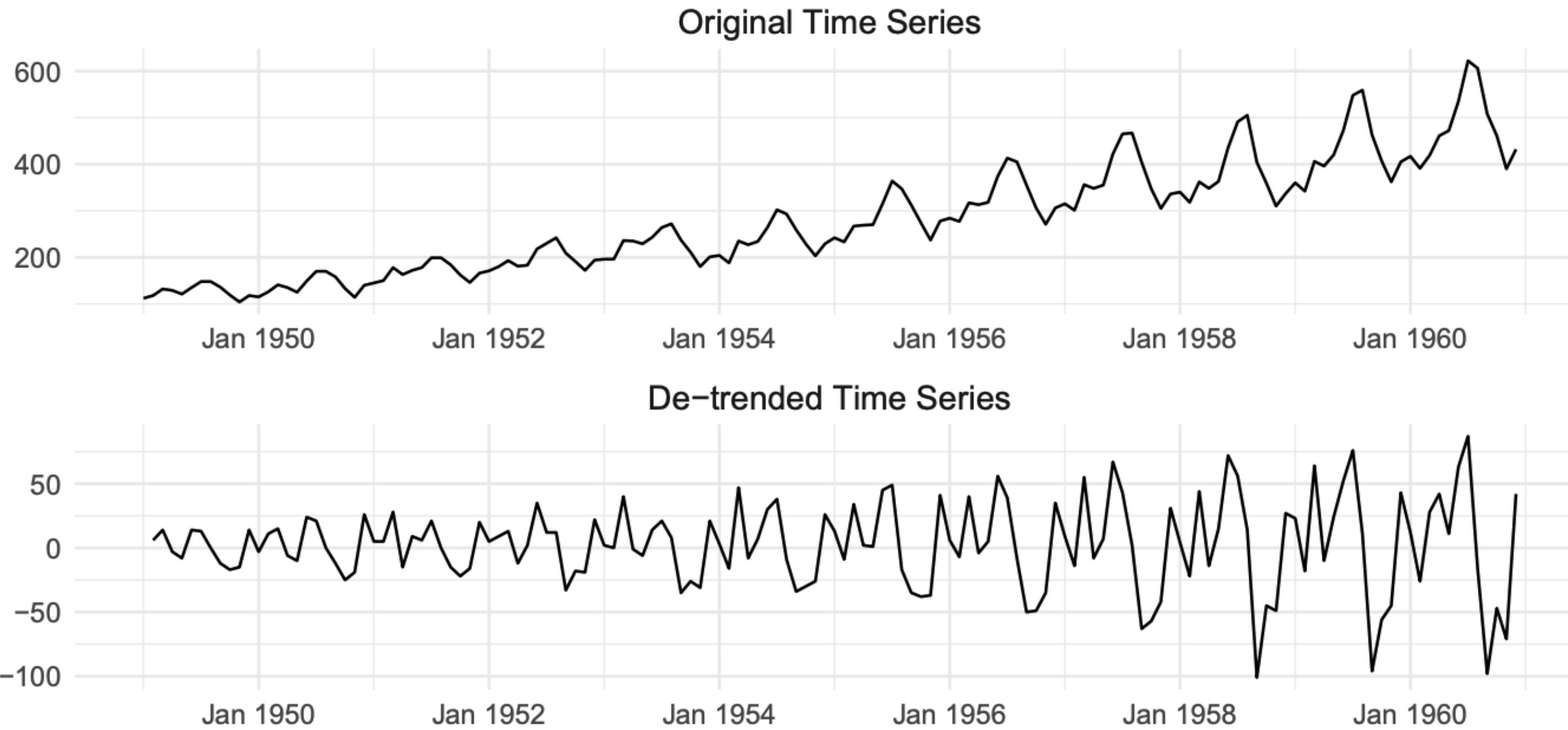
Oscilações regulares e previsíveis

Resíduos

Séries após remover tendência e sazonalidade

Tendência

Mudança do nível médio da série a longo-prazo



Tendência

Mudança do nível médio da série
a longo-prazo

```
import pandas as pd
from pmdarima.arima import ndiffs

data = pd.read_csv('air_passengers.csv')
```

Teste Estatístico

1. KPSS
2. Augmented Dickey-Fuller

```
series = pd.Series(data['V1'])
series.index = pd.to_datetime(data['Month'])

ndiffs(series, test='kpss')
# 1
```

Tendência

Mudança do nível médio da série a longo-prazo

```
import pandas as pd  
data = pd.read_csv('air_passengers.csv')
```

Diferenciação

Diferença entre valores consecutivos da série temporal:

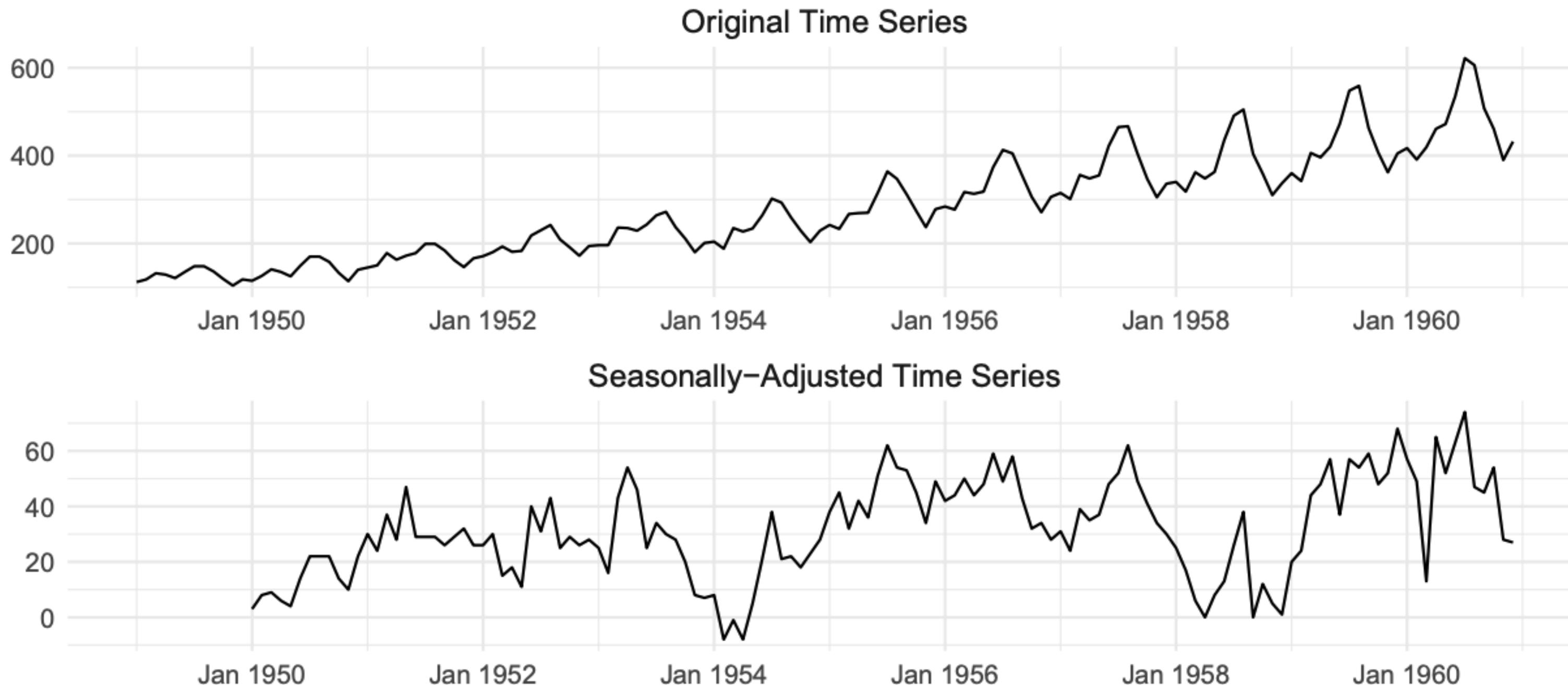
- Representa a mudança de um instante para o próximo
- Estabiliza a média

```
series = pd.Series(data['V1'])  
series.index = pd.to_datetime(data['Month'])  
series.diff(periods=1)
```

Sazonalidade

Mudanças regulares e previsíveis em períodos fixos (e.g. todos os meses)

- Quando as mudanças periódicas não têm um período fixo estas são chamadas de ciclos, por exemplo ciclos económicos



Sazonalidade

Mudanças regulares e previsíveis em períodos fixos (e.g. todos os meses)

Testes Estatísticos

- Teste OCSB;
- Teste Canova-Hansen

```
import pandas as pd
from pmdarima.arima import nsdiffs

data = pd.read_csv('air_passengers.csv')

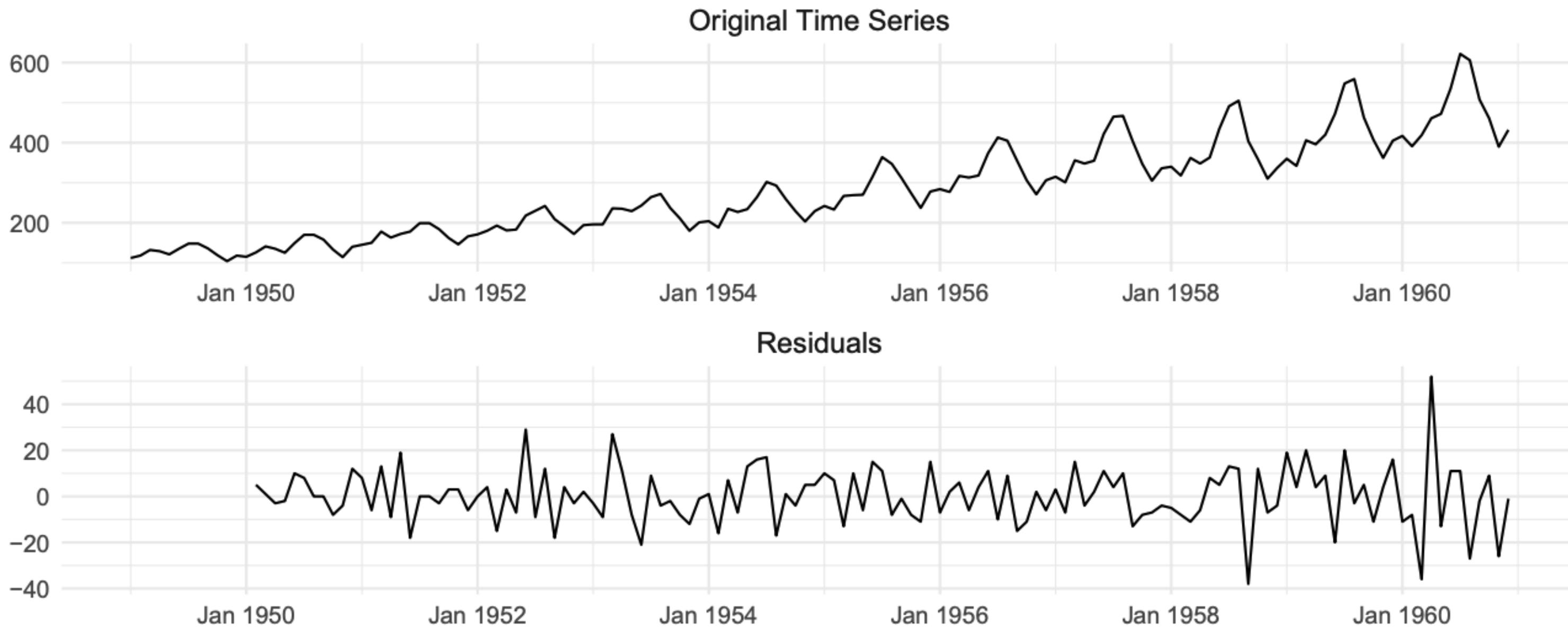
series = pd.Series(data['V1'])
series.index = pd.to_datetime(data['Month'])

nsdiffs(series, m=12, test='ocsb')
# 1

# apply seasonal differencing
series.diff(periods=12)
```

Resíduos

O que sobra da série temporal depois de remover a tendência e sazonalidade



Decomposição da Série Temporal

A cada instante i , uma série temporal pode ser decomposta da seguinte forma.

Decomposição Aditiva

$$y_i = \text{Tendência}_i + \text{Sazonalidade}_i + \text{Resíduos}_i$$

Decomposição Multiplicativa

$$y_i = \text{Tendência}_i \times \text{Sazonalidade}_i \times \text{Resíduos}_i$$

Time Series Decomposition

A cada instante i , uma série temporal pode ser decomposta nas três partes básicas: tendência, sazonalidade e resíduos.

Que decomposição devo usar?

Se as oscilações à volta da tendência e sazonalidade dependerem do nível da série: **Decomposição multiplicativa**

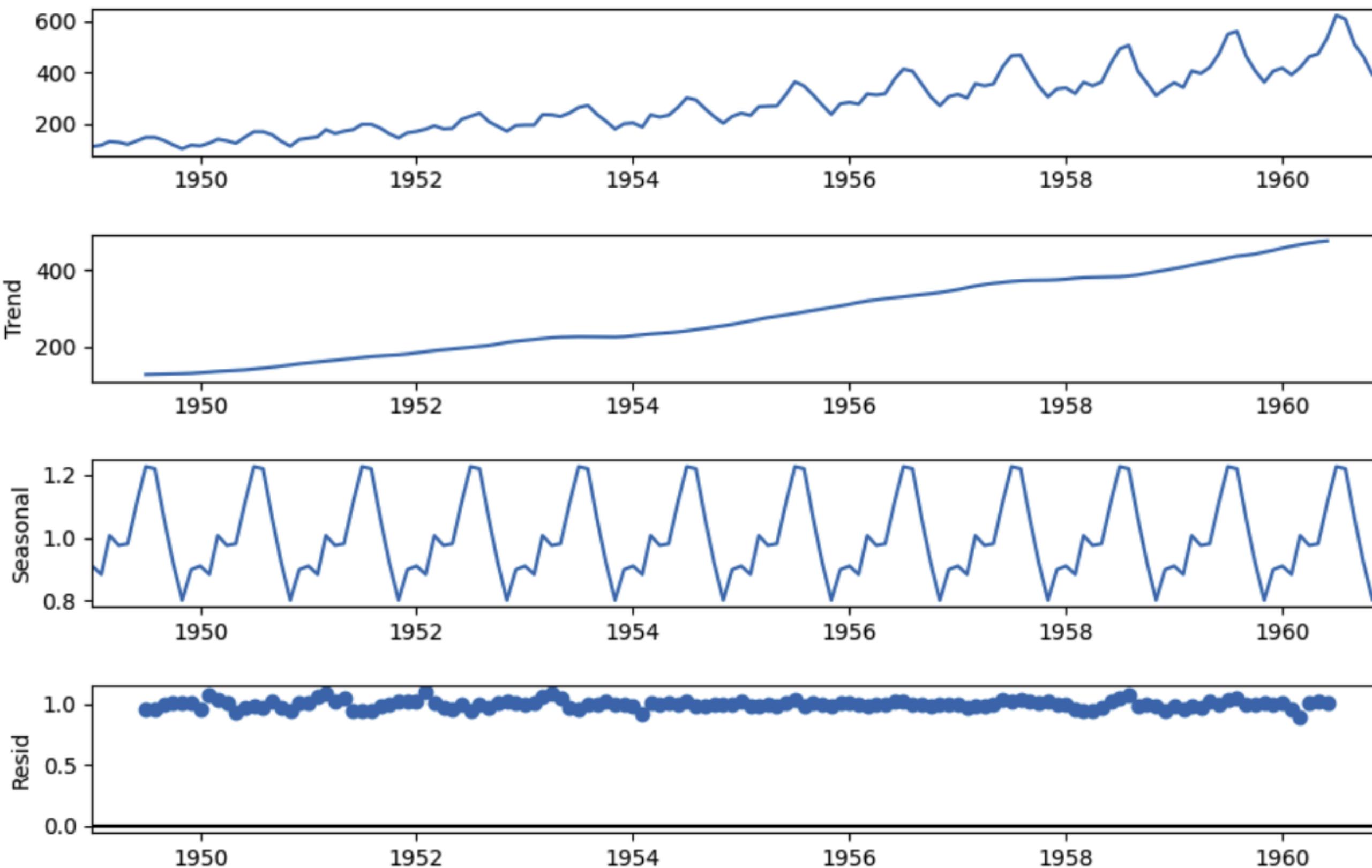
Log transformation

$$\log(y_i) = \log(\text{Tendência}_i) + \log(\text{Sazonalidade}_i) + \log(\text{Resíduos}_i)$$

Métodos de Decomposição

- Decomposição com Médias Móveis;
 - Decomposição X11;
 - Decomposição STL.

V1



2.3

Propriedades Chave



Propriedades Chave

- ✓ Estacionaridade
- ✓ Auto-correlação
- ✓ Heterocedasticidade
- ✓ Regularidade
- ✓ Frequência
- ✓ Outliers
- ✓ Regimes and Mudanças
- ✓ Reflexividade

Estacionaridade

Definição

Uma série temporal é estacionária se as suas propriedades não dependem do instante em que a série é observada.

- Não existem mudanças sistemáticas na média ou variância.

Estacionaridade Estrita

A distribuição conjunta de $\{y_1, \dots, y_i\}$ é idêntica à distribuição conjunta $\{y_{1+j}, \dots, y_{i+j}\}$, para todos $i, j \in \mathbb{N}$.

Estacionaridade Fraca

Média e variância constante e auto-covariância não depende do tempo.

Auto-correlação

Correlação

Grau de relação linear entre duas variáveis

Auto-correlação

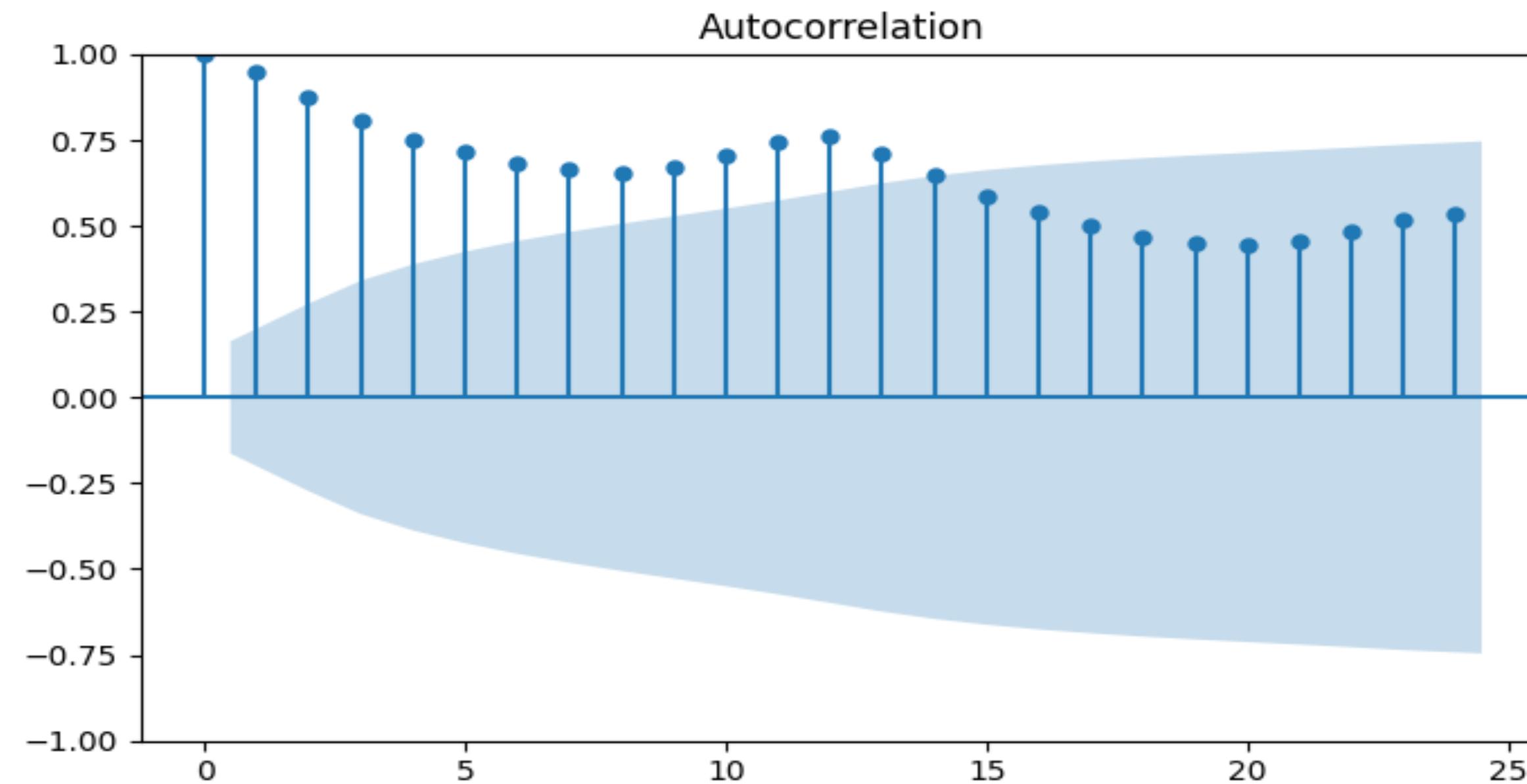
Grau de relação linear entre a série temporal e os seus valores passados

```
from statsmodels.graphics.tsaplots import plot_acf
```

```
plot_acf(series, lags=24)
```

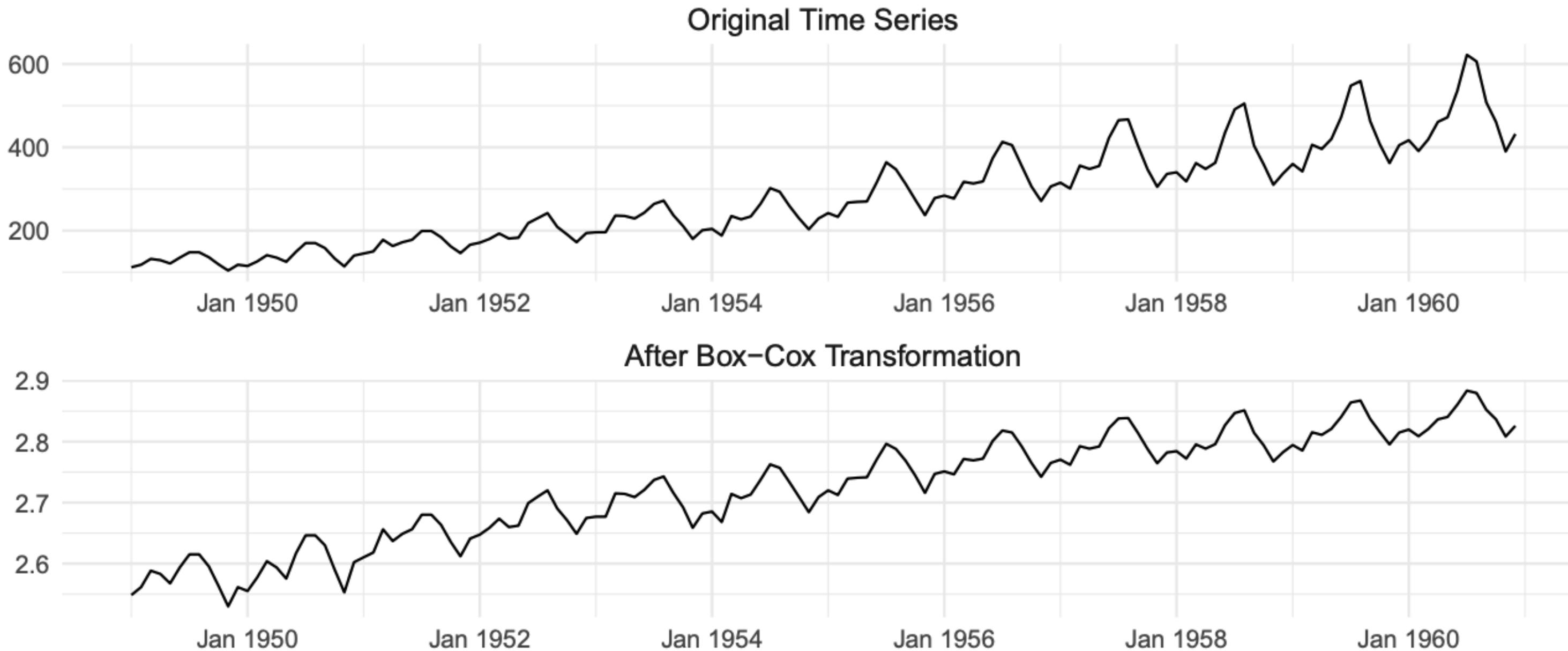
Auto-correlação

Grau de relação linear entre a série temporal e os seus valores passados



Heterocedasticidade

Uma série temporal é heterocedástica se a variância não for constante



Heterocedasticidade

Testes Estatísticos

- Teste Breusch-Pagan
- Teste White

Transformações de Potência

- Log
- Transformação Box-Cox

```
import pandas as pd
from heteroskedasticity import het_tests

data = pd.read_csv('air_passengers.csv')

series = pd.Series(data['V1'])
series.index = pd.to_datetime(data['Month'])

het_tests(series, test='white')
# < 0.001
```

Transformações Box-Cox

$$a_i = \begin{cases} \log(y_i) & \lambda = 0; \\ (y_i^\lambda - 1)/\lambda & \lambda \neq 0 \end{cases}$$

- $\lambda = 1$: Sem transformação;
- $\lambda = 0.5$: Raiz quadrada mais transformação linear;
- $\lambda = 0$: Logaritmo;
- $\lambda = -1$: Inversa mais 1

Heterocedasticidade

Testes Estatísticos

- Teste Breusch-Pagan
- Teste White

Transformações de Potência

- Log
- Transformação Box-Cox

```
import pandas as pd
from scipy.stats import boxcox
from scipy.special import inv_boxcox

data = pd.read_csv('air_passengers.csv')

series = pd.Series(data['V1'])
series.index = pd.to_datetime(data['Month'])

# aplicar a transformação box-cox
transformed_data, best_lambda = boxcox(series)

# reverter a transformação
inv_boxcox(transformed_data, best_lambda)
```

Frequência

Definição

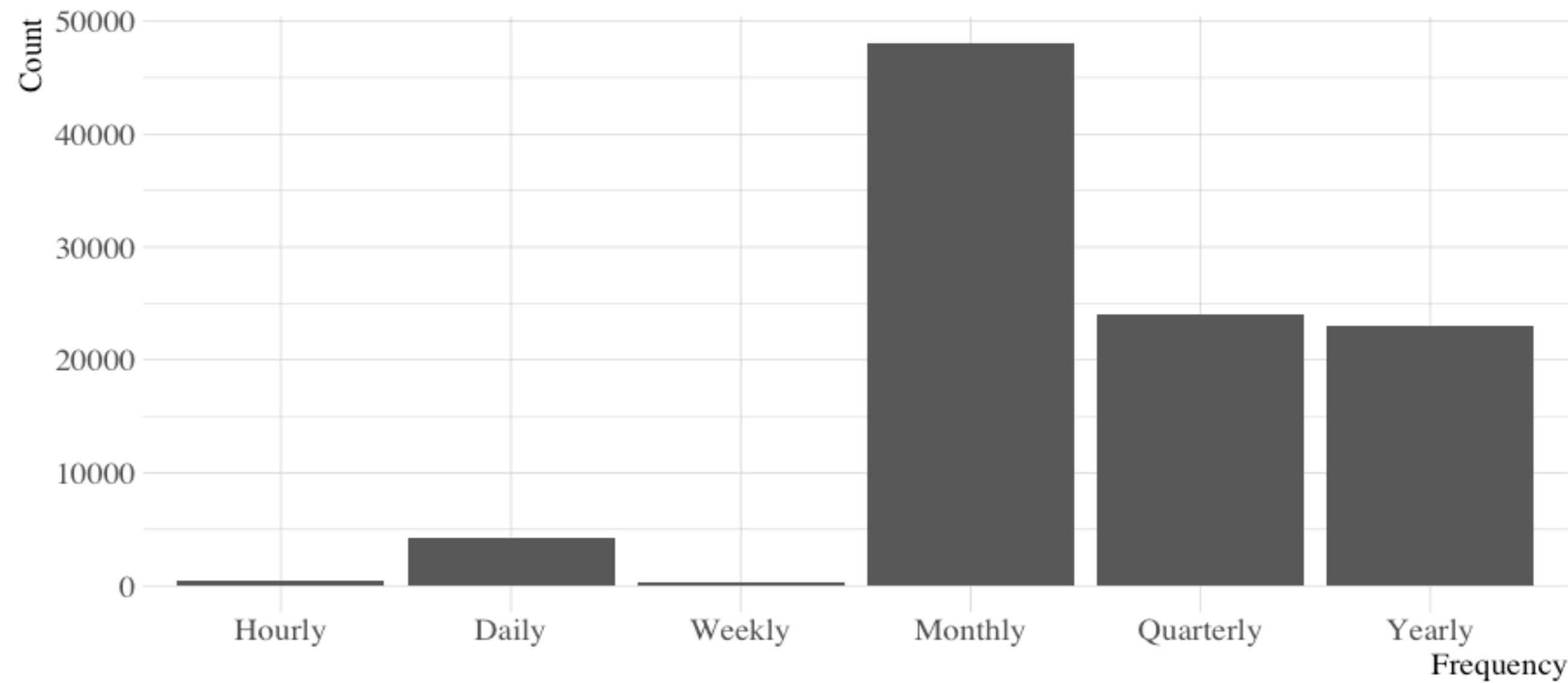
Regularidade com que a série é observada

- Por exemplo, diariamente ou semanalmente

Desafios

- Frequência elevada:
 - **Vários ciclos sazonais**
- Frequência baixa:
 - **Poucos dados**

Dist. Frequênci a do conjunto M4



<https://mofc.unic.ac.cy/m4/>

Regularidade

Definição

Séries temporais regulares são observadas com uma frequência regular

- por exemplo, todos os dias

No entanto, algumas séries temporais são naturalmente irregulares.

Por exemplo, as vendas de um produto.

2021-08-14 11:30:00

2021-08-14 15:18:00

2021-08-14 15:53:00

2021-08-14 18:55:00

2021-08-15 01:00:00

...

Agregação da Série Temporal

Conversão da série irregular numa série regular para análise

```
import pandas as pd
```

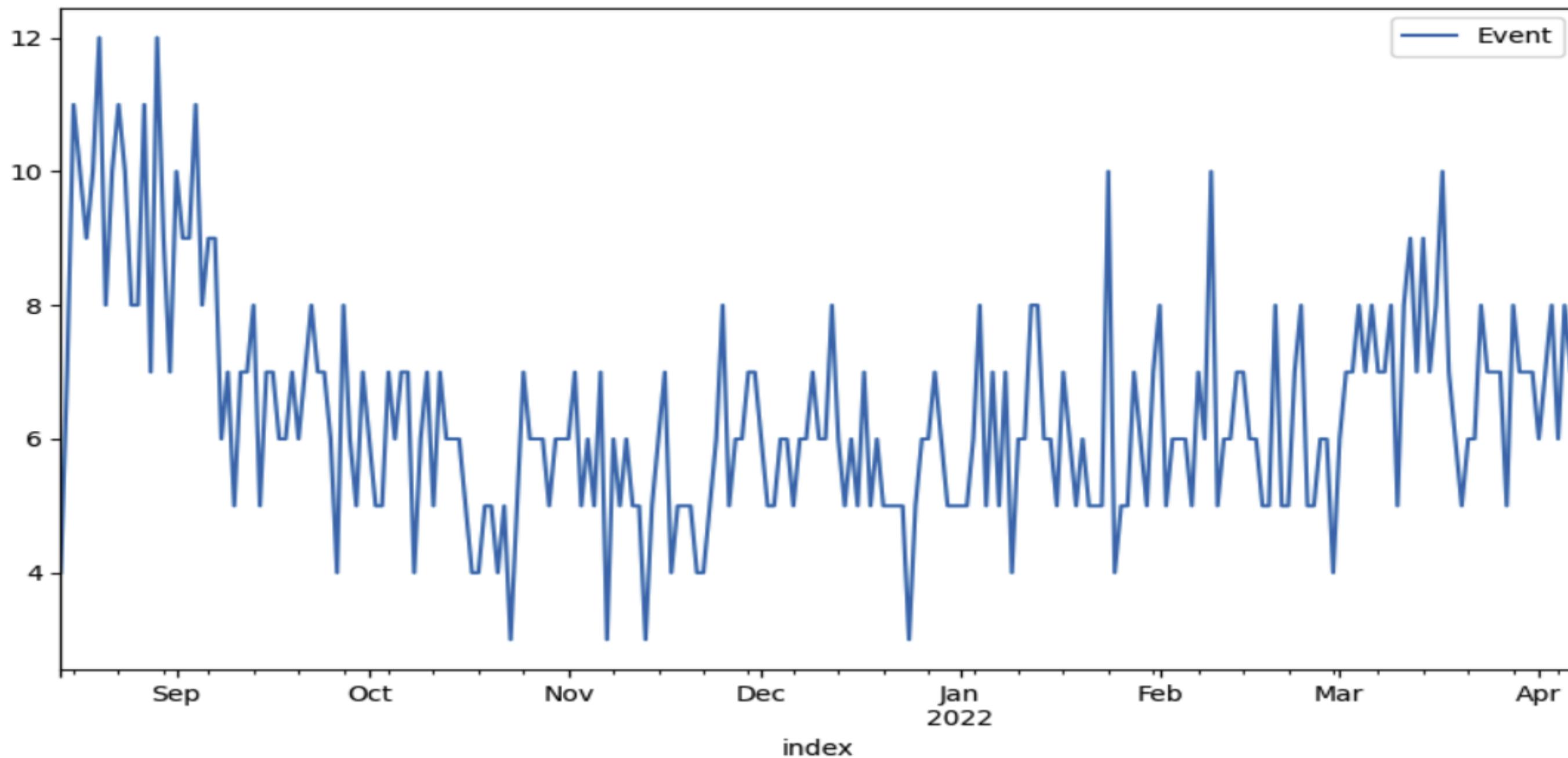
```
irregular = pd.read_csv('irregular_series.csv')
irregular['Evento'] = 'Venda'
irregular.set_index('Timestamp', inplace=True)
irregular.index = pd.DatetimeIndex(irregular.index)
```

```
regular = irregular.groupby(irregular.index.date).count()
```

```
# contar os zeros
index_range = pd.date_range('2021-08-14', '2022-04-07')
regular = regular.reindex(index_range, fill_value=0)
# visualizar
regular.reset_index().plot.line(x='index', y='Evento', rot=0)
```

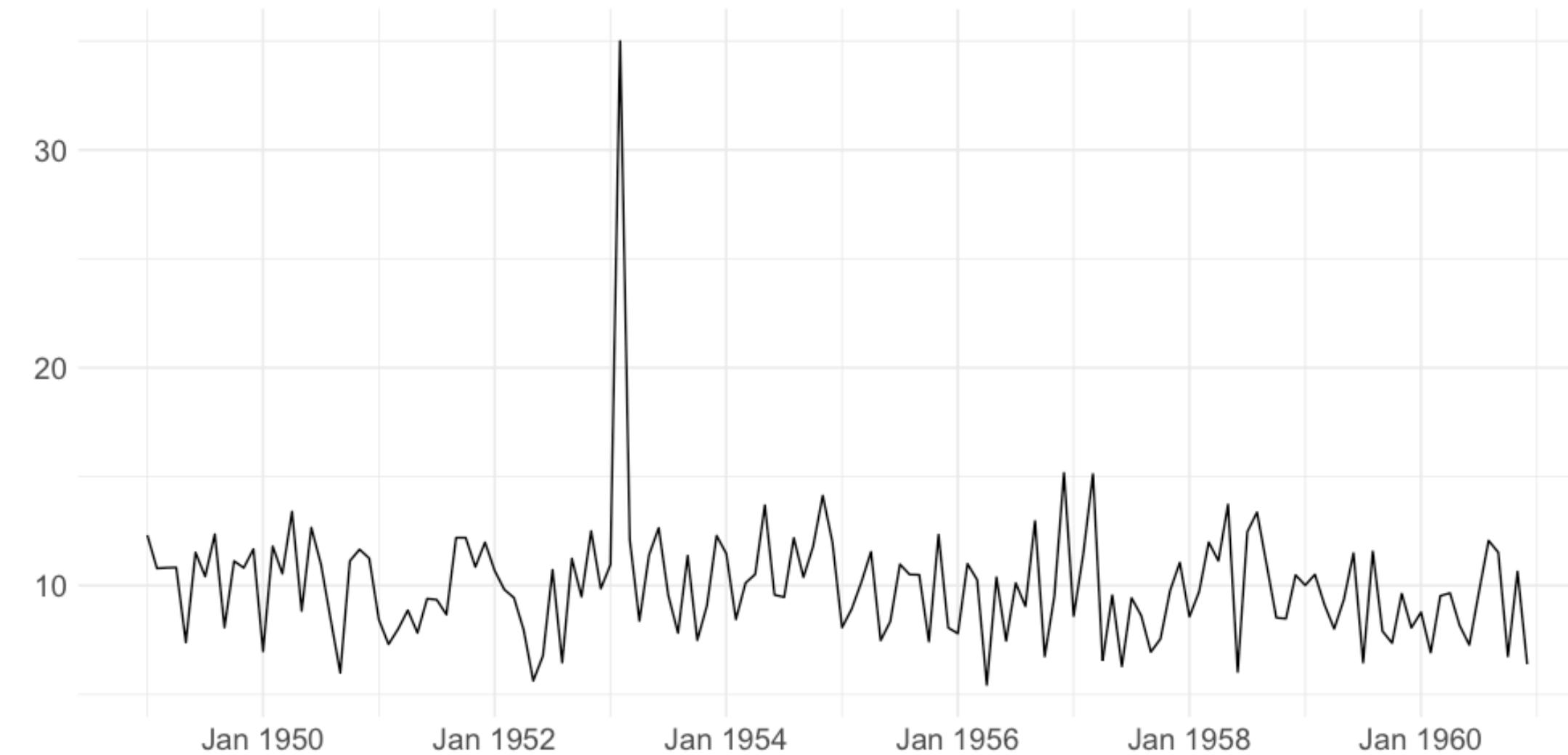
Agregação da Série Temporal

Conversão da série irregular numa série regular para análise



Outliers

Outliers ou anomalias são eventos raros que se desviam do comportamento normal da série temporal



Outliers

Diferentes tipos de outliers

- **Pontos de outlier:** ocorrem num instante de tempo
- **Subsequências de outliers:** abrangem várias observações

Como lidar com outliers

Depende da sua natureza

Dados não desejados

- Erros de leitura, ruído, sensores danificados, etc.

Eventos raros mas interessantes

- Por exemplo, quebras na bolsa financeira, atividade fraudulenta, etc.

Regimes and Mudança

Regimes

O processo da série temporal que origina os dados

Pontos de Mudança

A série temporal muda, e a sua distribuição muda consequentemente

Mudanças recorrentes ou permanentes

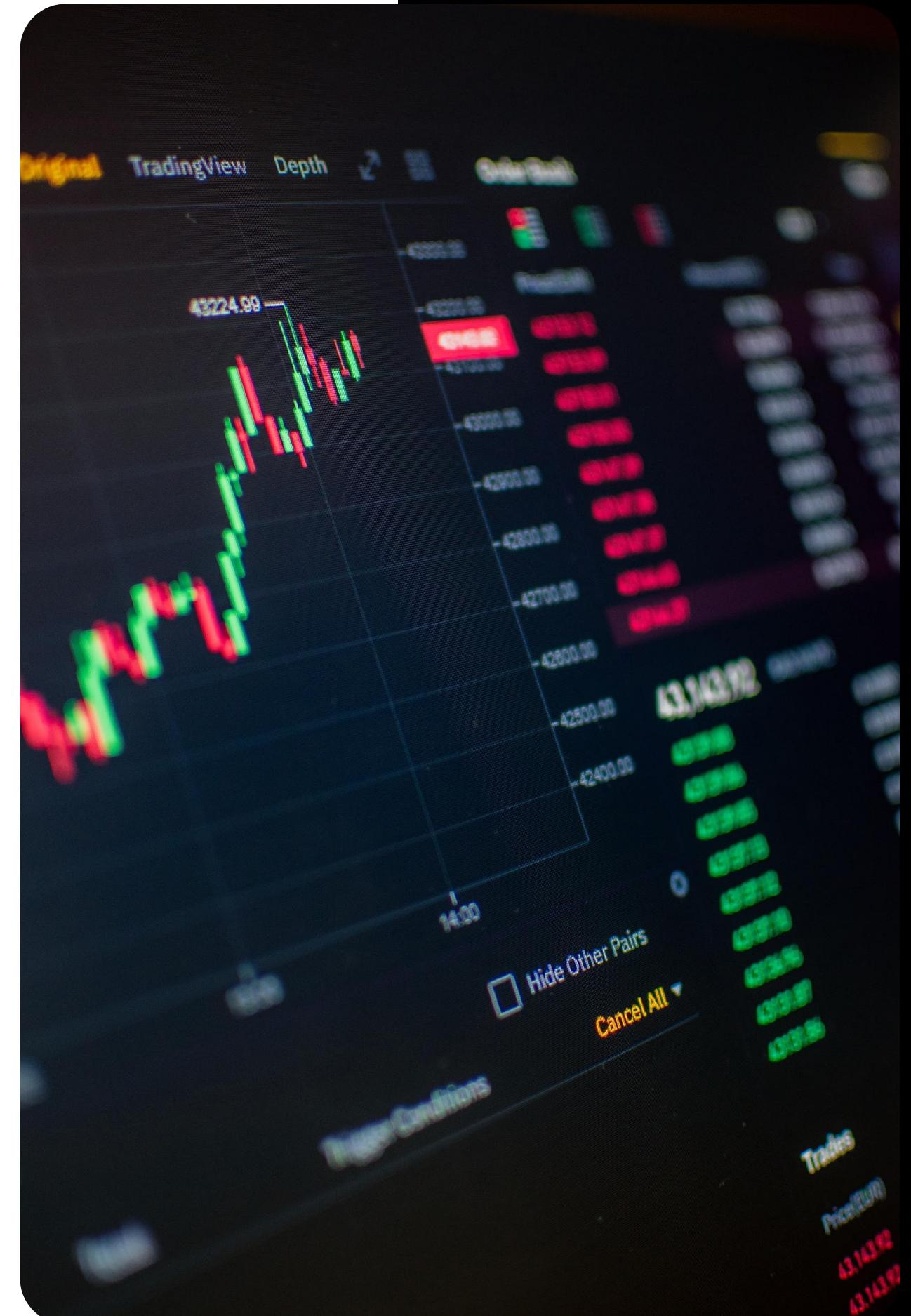
A distribuição pode mudar entre regimes (sazonalidade) ou mudar para algo nunca visto antes (quebras estruturais)

Reflexividade

Definição

Quando a previsão afeta o evento que está a ser previsto.

Mercados Financeiros: A previsão de um aumento no preço de uma ação atrai investidores, o que cria demanda e eleva o preço da ação.



Séries Temporais: Definições e Propriedades

Vitor Cerqueira

<https://www.linkedin.com/in/vcerq/>
cerqueira.vitormanuel@gmail.com