Mestrado em Data Science

Séries Temporais

Modelos de Séries Temporais

Conteúdo

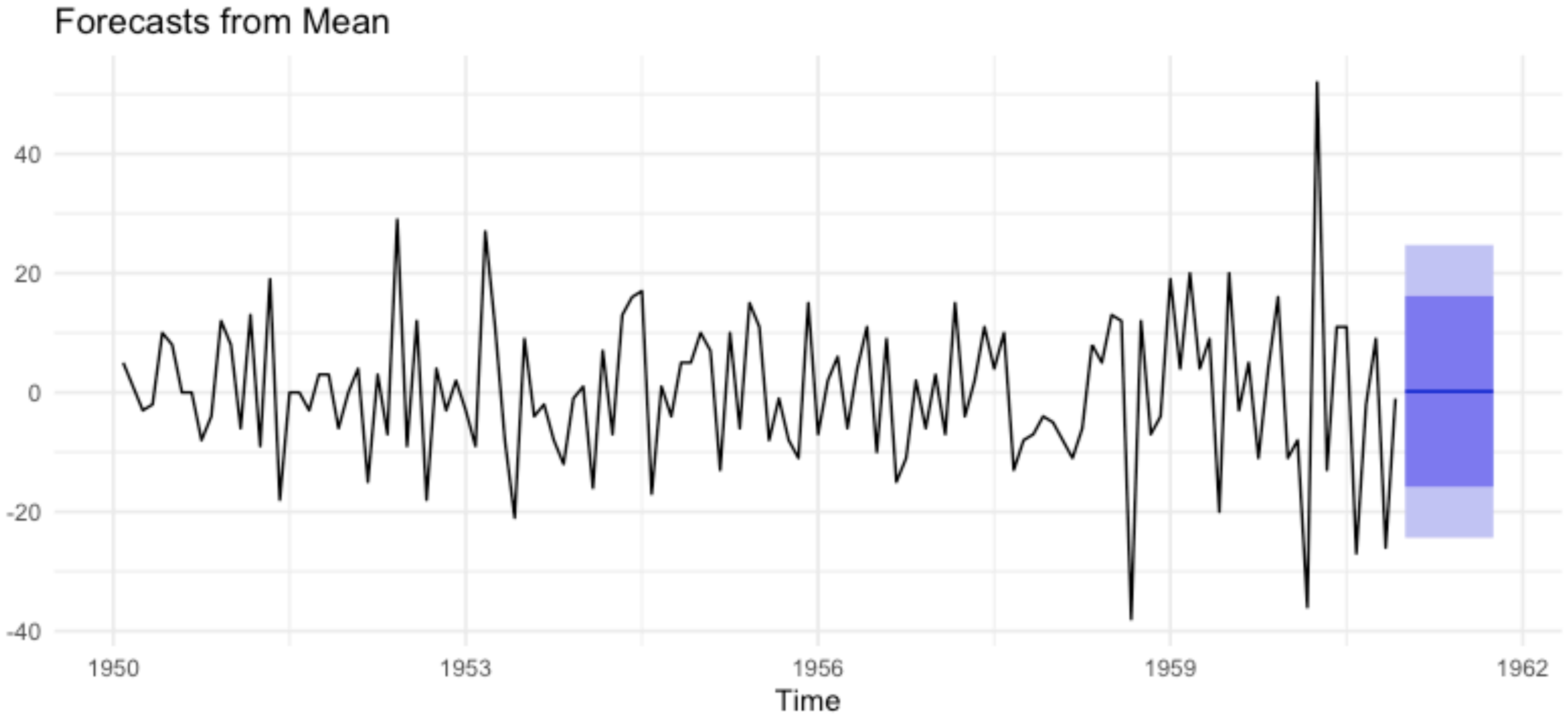
3.1

Modelos Simples

Método da Média

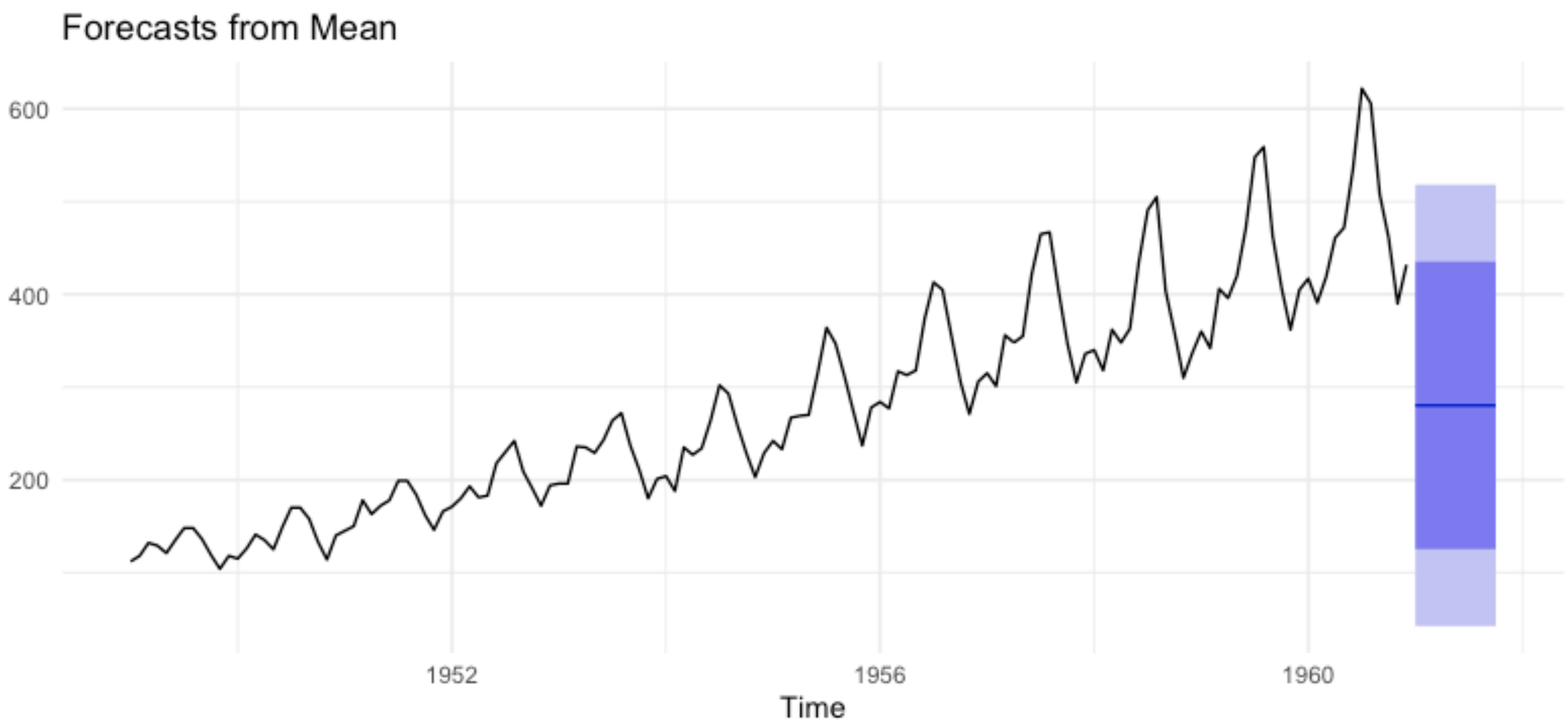
As previsões são iguais à média das observações históricas

$$\hat{y}_{t+h} = \bar{y} = (y_1 + y_2 + \cdots + y_n)/n$$



Exemplo: Importância das Transformações

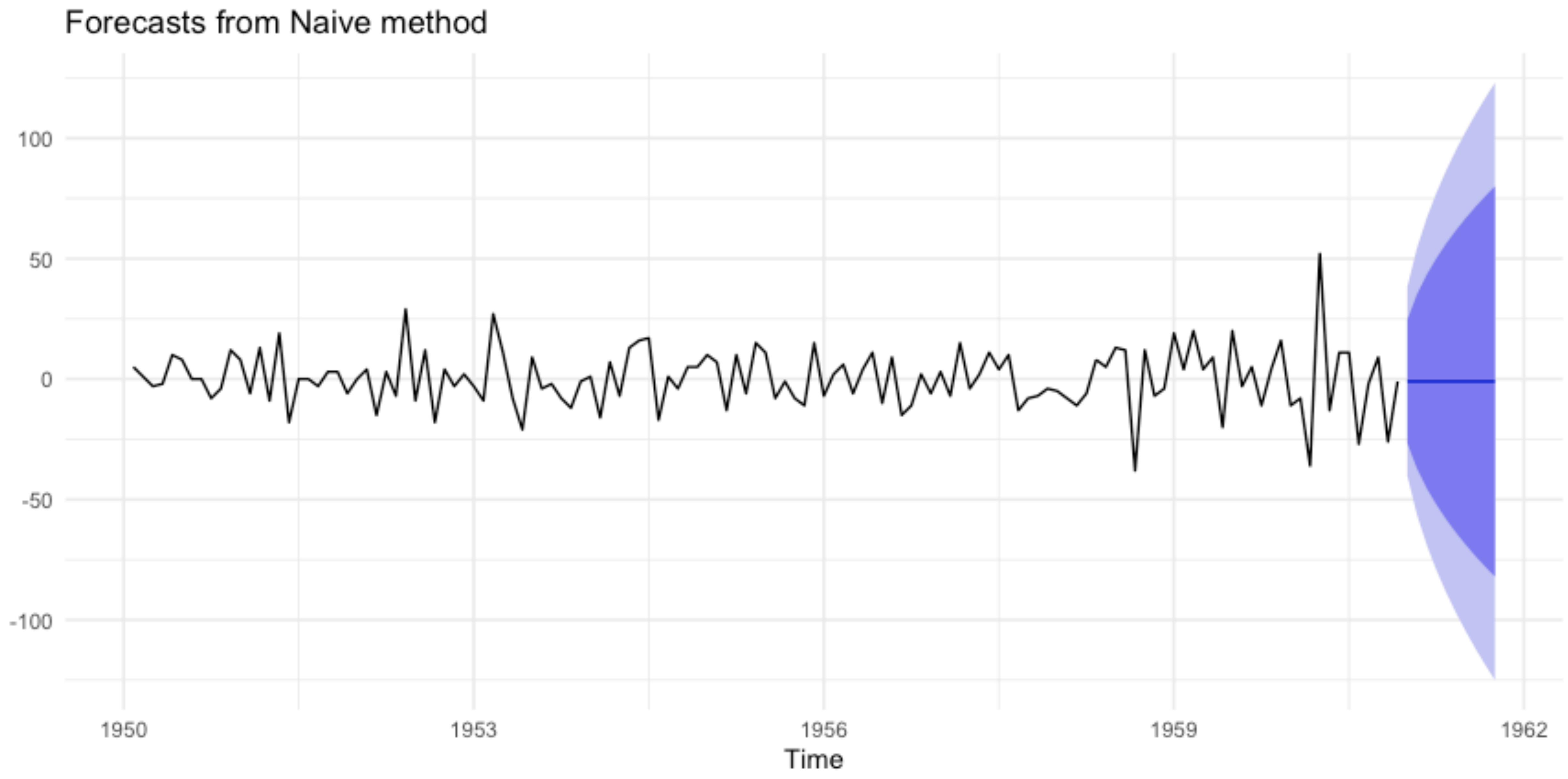
Aplicando o Método da Média na Série Temporal Original



Método Naive

As previsões são iguais à última observação conhecida.
Também conhecido como método persistente.

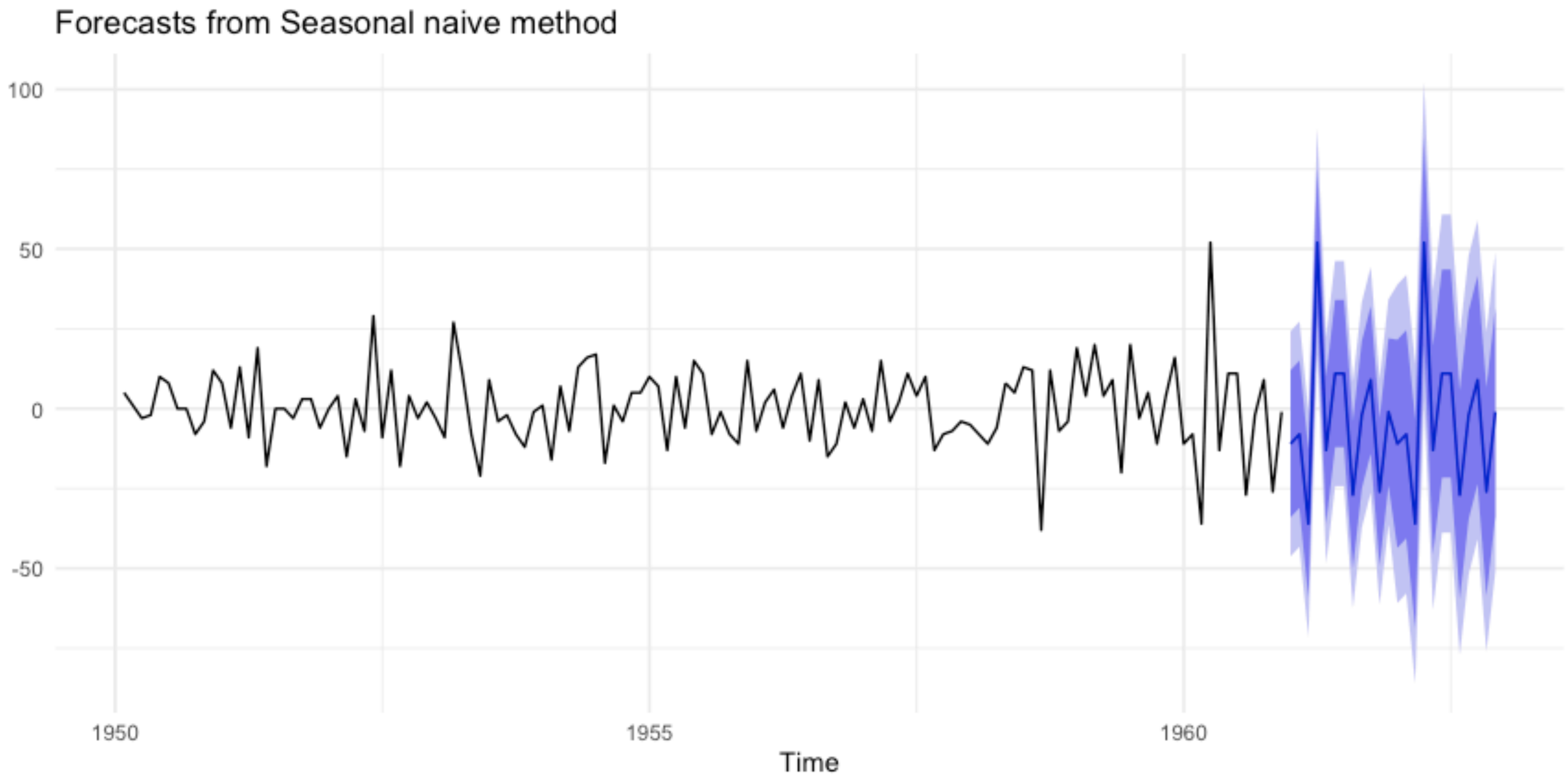
$$\hat{y}_{t+h} = y_t$$



Método Sazonal Naive

As previsões são iguais à última observação conhecida do mesmo período sazonal (indicado como **m**).

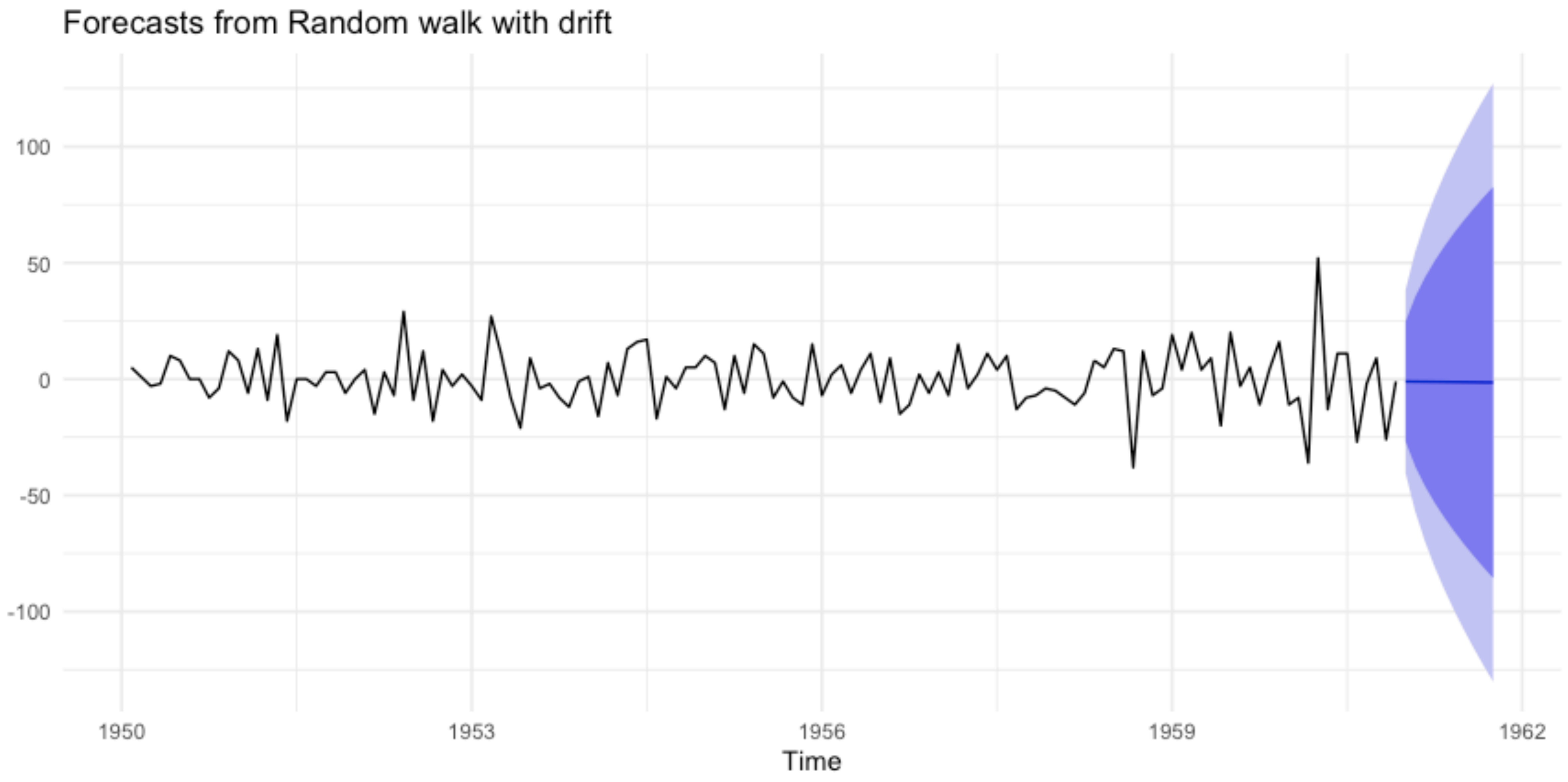
$$\hat{y}_{t+h} = y_{t+h-m}$$



Método Drift

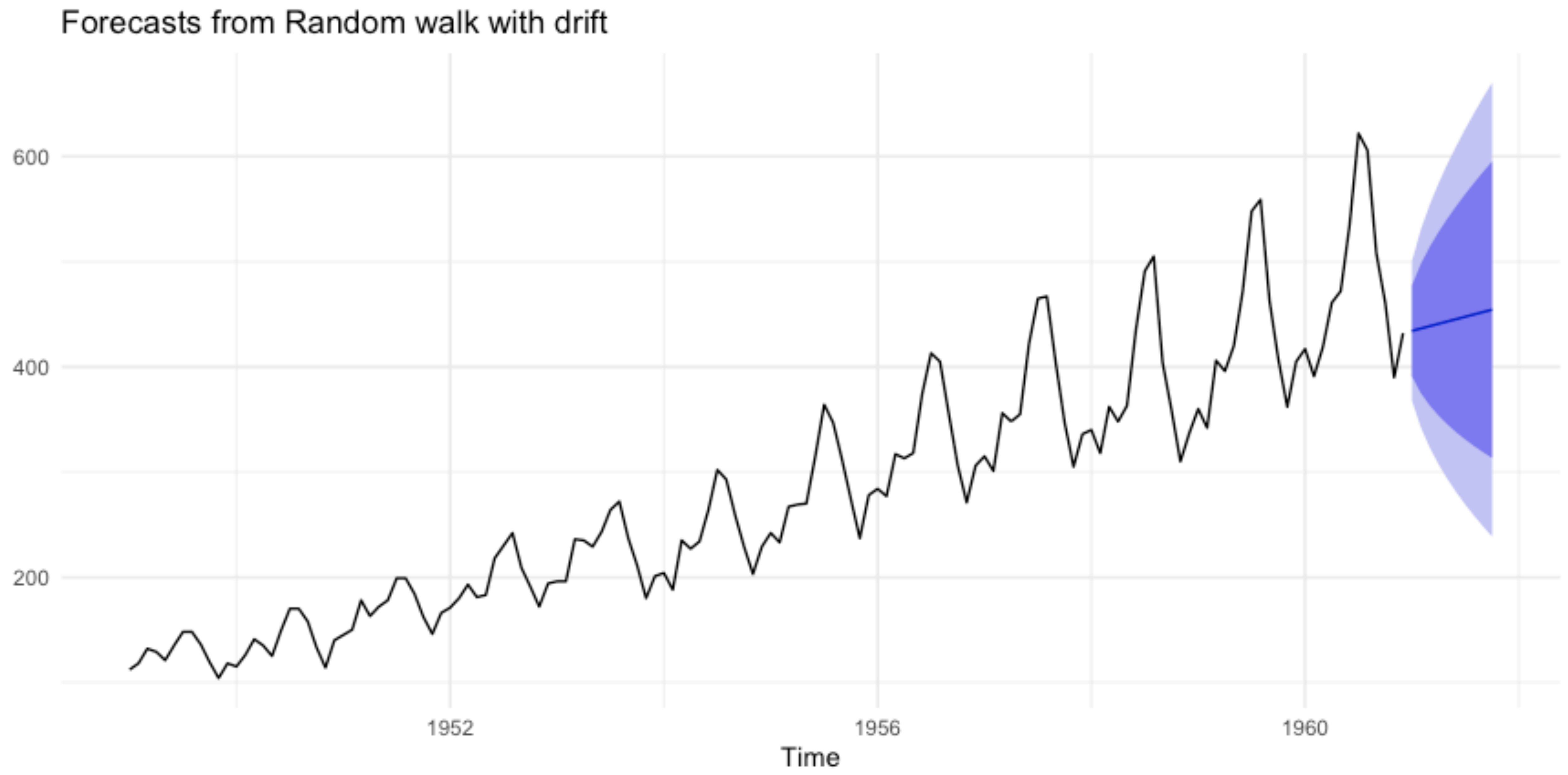
As previsões são iguais à última observação conhecida mais a variação média.

$$\hat{y}_{t+h} = y_t + h\left(\frac{y_n - y_1}{n-1}\right)$$

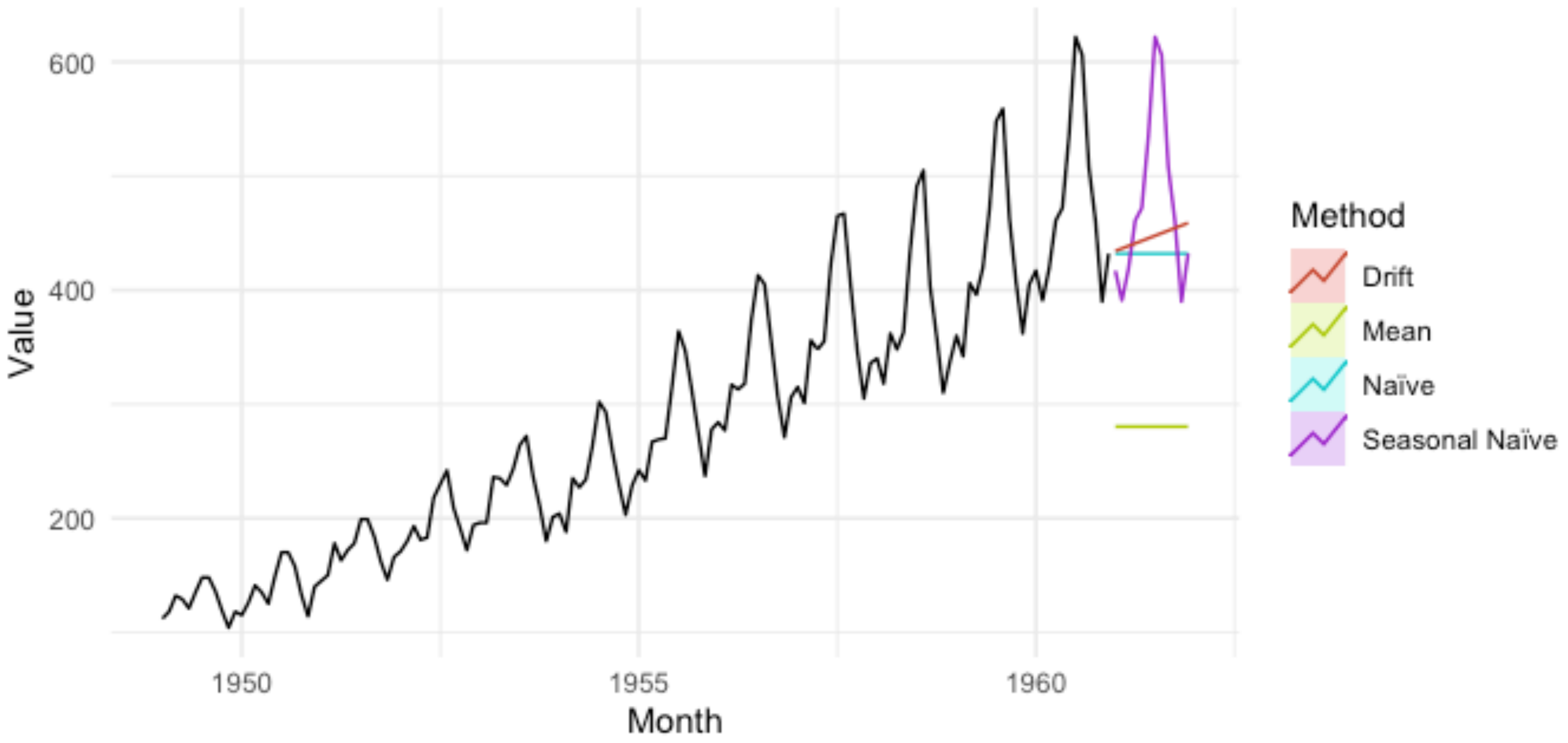


Método Drift

As previsões são iguais à última observação conhecida mais a variação média.



Comparando os Métodos



3.2
Exponential Smoothing

Simple Exponential Smoothing, Método Holt’s, ETS

Simple Exponential Smoothing (SES)

Método da Média

Média das observações conhecidas

Método Naive

Usando a última observação conhecida

SES

- *Trade-off* destes dois extremos
- Promove recência das observações

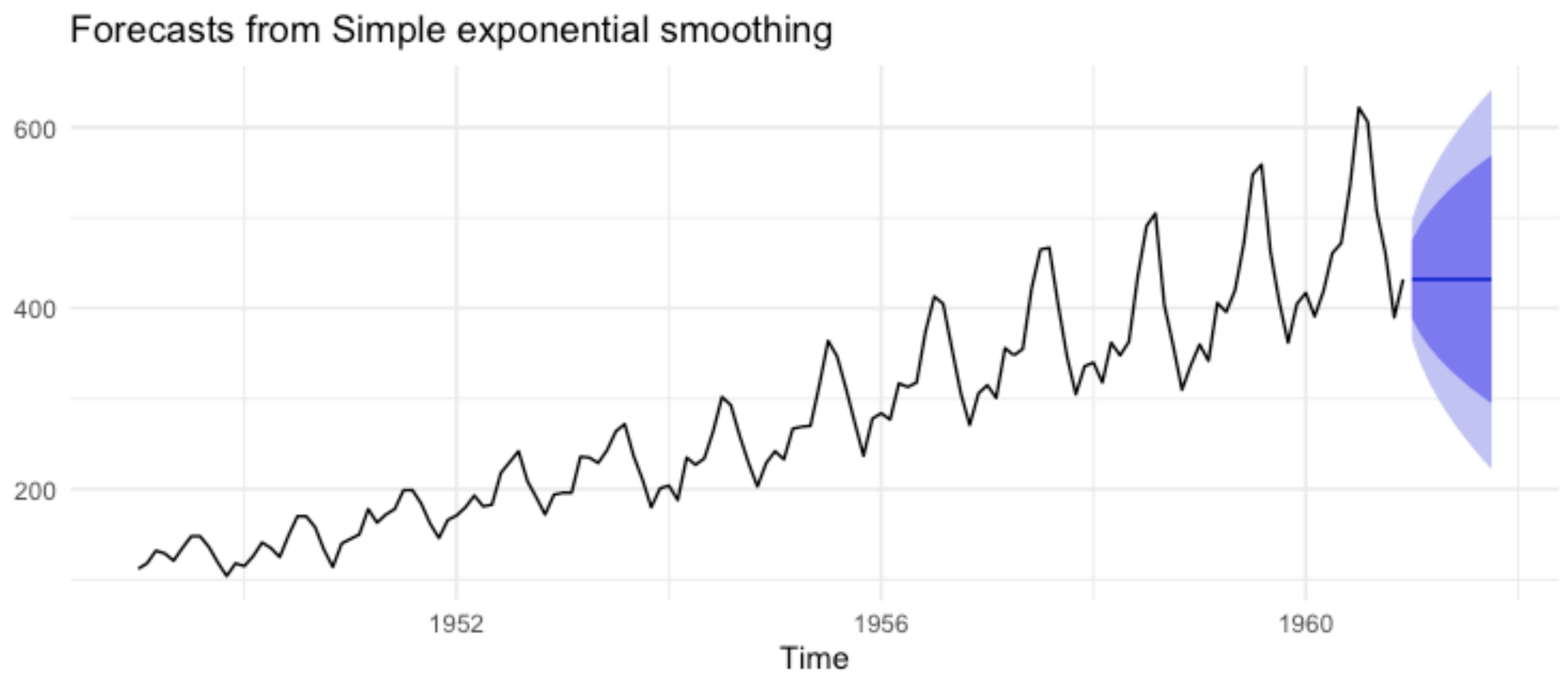
Simple Exponential Smoothing (SES)

Médias Pesadas

- Médias ponderadas de observações anteriores
  - Os pesos decaem exponencialmente à medida que as observações envelhecem
- ɑ assume um valor entre 0 e 1 que controla o decaimento
  - Podemos usar bibliotecas de software para estimar este parêmetro

$$\hat{y}_{t+1} = \alpha y_t + \alpha(1-\alpha) y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \dots$$

Simple Exponential Smoothing (SES)



Método Holt

Estendendo SES para modelar séries temporais com tendência

$$\hat{y}_{t+h} = l_t + hb_t$$

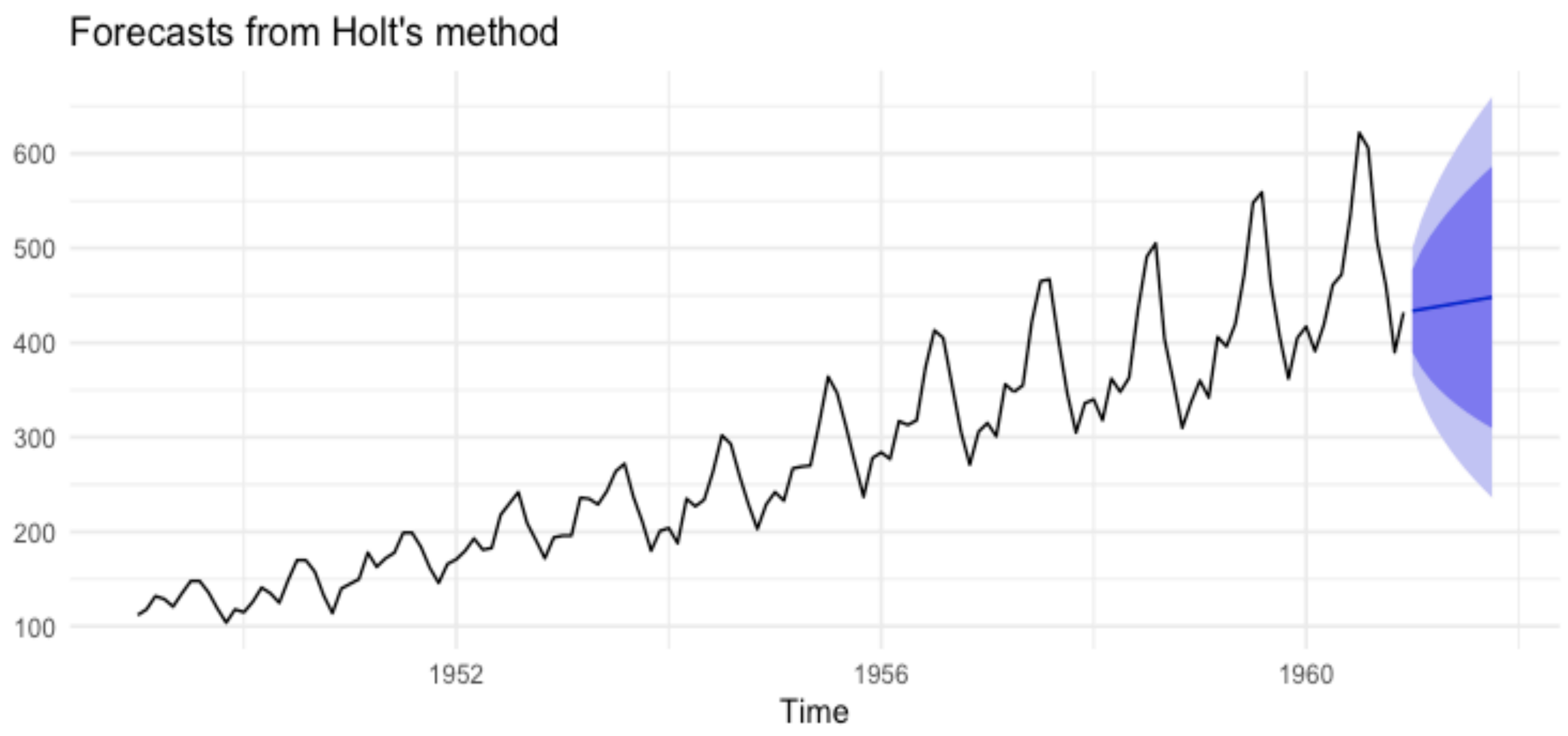
$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$$

$$b_t = \beta^*(l_t - l_{t-1}) + (1-\beta^*)b_{t-1}$$

- Contém 2 parâmetros de suavização: α e β. Ambos assumem valores entre 0 e 1
- Um modelo de duas partes que considera o nível (l_t)e a inclinação (b_t) da série
- Extensão possível: tendência amortecida (*damping*)
  - Diminuindo o efeito da tendência para horizontes mais longos

Método Holt

Estendendo SES para modelar séries temporais com tendência



Mais Métodos Exponential Smoothing

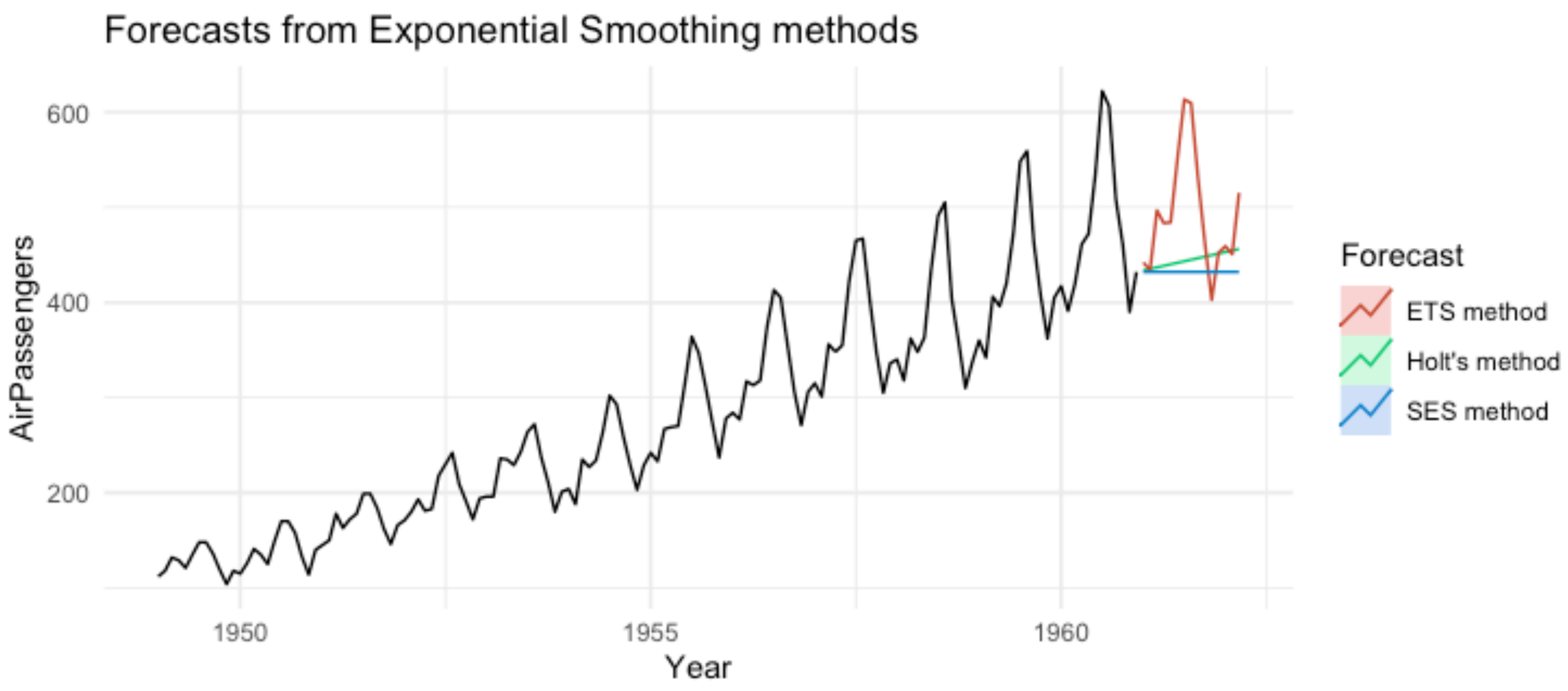
Holt-Winters

- Estende o método de Holt para capturar a sazonalidade
- Pode ser aditivo ou multiplicativo
  - Dependendo de como as variações sazonais afetam o nível da série
- Também pode ser amortecido

ETS (Error, Trend, Seasonality)

- Método *State-space*
- Considera 18 possíveis configurações
  - Seleção via Akaike's Information Criteria (AIC)

More Exponential Smoothing Methods



Métodos Exponential Smoothing

- Várias configurações possíveis
- Geralmente selecionado usando bibliotecas de software
- Eficaz na prática

Soluções Híbridas

- Pode ser combinado com métodos de aprendizagem máquina
- Parte da solução do vencedor da competição de previsão M4 (referência abaixo)
  - A suavização exponencial foi usada para normalizar séries temporais
  - Deep learning foi usado para as previsões

Smyl, Slawek. "A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting." *International Journal of Forecasting* 36.1 (2020): 75-85.

3.3 ARIMA

AR, MA, ARMA, ARIMA, ARIMAX

Método Auto-Regressivo (AR)

Regressão Múltipla

A variável objetivo é modelada de acordo com uma combinação linear de variáveis explicativas

AR(*p*)

- Regressão Múltipla para séries temporais
- A variável objetivo representa o próximo valor da série
- As variáveis explicativas são observações passadas
- ***p*** representa a ordem do modelo:
    - Número de observações passadas a usar

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$$

Método Auto-Regressive (AR)

Equação do Modelo

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$$

- *p*: número de observações passadas (lags)
- ϕ: parâmetros do modelo
- ϵ: termo de erro
- c: termo constante

Método Moving-Average (MA)

Moving Average

Não confundir com o processo de médias móveis.

MA(***q***)

- Parecido com AR(p), mas usa erros passados como variáveis explicativas
- ***q*** representa a ordem do modelo

MA(q) leva em conta os erros aleatórios que ocorreram nos últimos q pontos

$$y_t = c + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

Método Moving-Average (MA)

Equação MA(*q*)

$$y_t = c + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

- *q*: número de observações passadas (lags)
- Θ: parâmetros do modelo
- ϵ: termo de erro
- c: termo constante

Método Auto-Regressive Moving-Average (ARMA)

Combinação de AR(p) com MA(q) (Box-Jenkins)

ARMA(p, q) = AR(p) + MA(q)

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

Definido para séries temporais estacionárias

Auto-Regressive Integrated Moving-Average (ARIMA)

Adicionando um passo de integração

ARIMA(p, d, q)

- *d* é um parâmetro de integração de ordem **d**
- Aplicar *d* transformações de diferenciação (até obter estacionaridade)
- Depois, aplicar ARMA(p, q) à série transformada

Casos Especiais

ARIMA(0,0,0): Ruído branco

ARIMA(0,1,0) with c = 0: Passeio aleatório (Random Walk)

ARIMA(p,0,0): Auto-regressão

ARIMA(0,0,q): Moving-average

Auto-Regressive Integrated Moving-Average (ARIMA)

ARIMAX

- ARIMA com variáveis explicativas
- Útil quando existem variáveis adicionais relevantes para o problema

SARIMA e Auto ARIMA

SARIMA (Seasonal ARIMA)

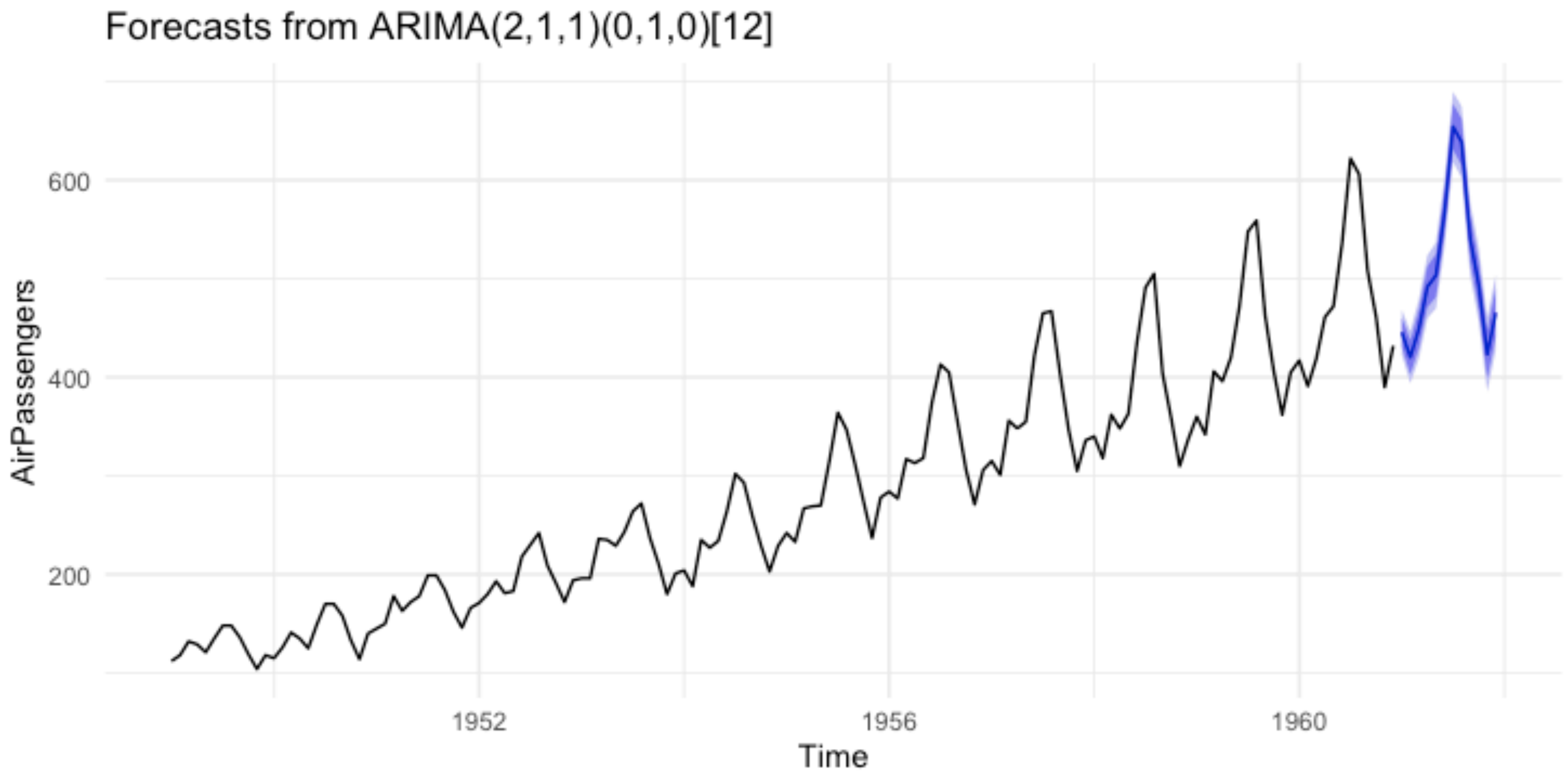
- ARIMA pode ter termos sazonais
- A estrutura sazonal é *idêntica à não sazonal* e para capturar padrões sazonais
- ARIMA$(p,d,q)(P,D,Q)_m$

Seleção da configuração do ARIMA

- Visualizar auto-correlações and determinar onde esta é significativa
- … Ou deixar que software faça isso por si
- Muitas implementações de Auto ARIMA, tanto em R como em Python

```python
from pmdarima.arima import auto_arima
model = auto_arima(series)
```

Auto ARIMA

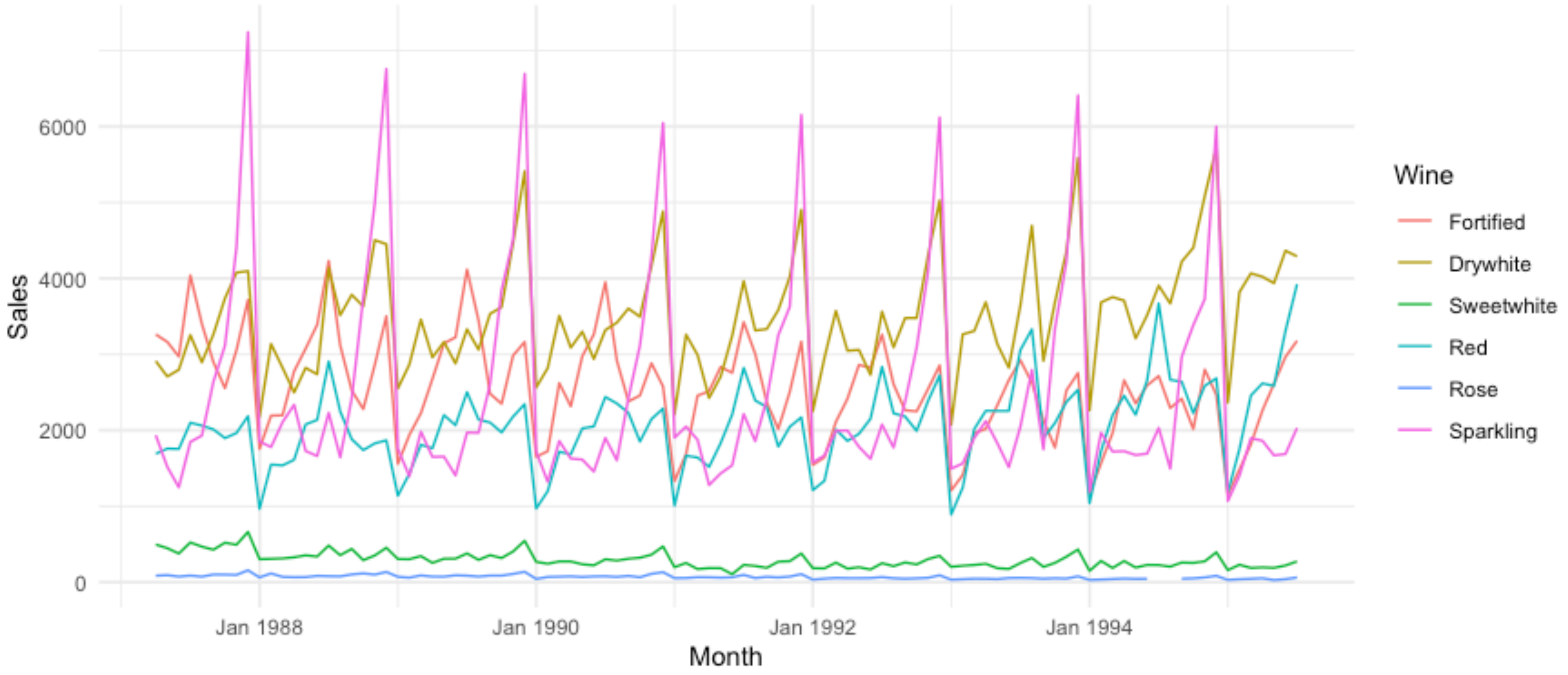


Forecasts from ARIMA(2,1,1)(0,1,0)[12]
600
400
200
AirPassengers
1952
1956
1960
Time

3.4
VAR, ARDL

Séries Temporais Multivariadas

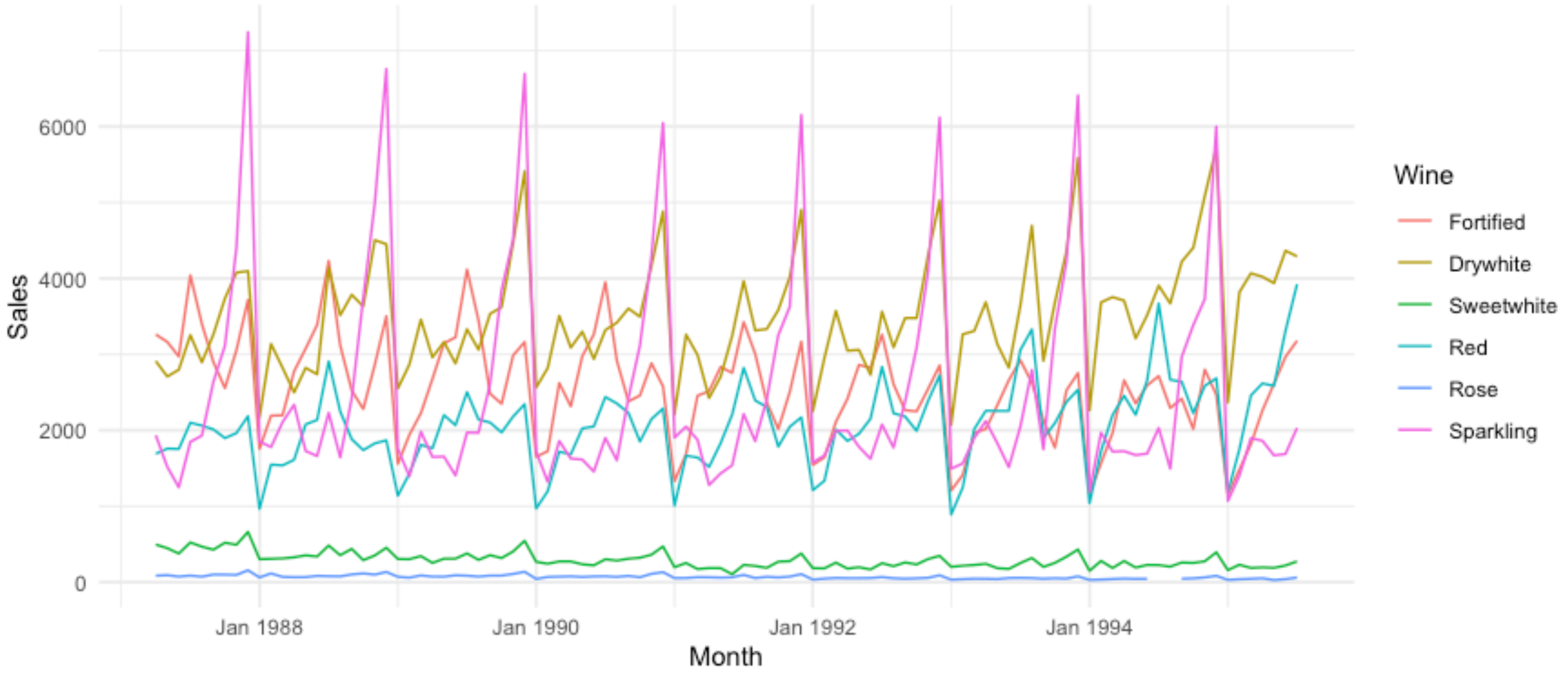
- Séries Temporais podem conter várias variáveis
  - Estas são chamadas de séries multivariadas
  - As variáveis adicionais podem ser de interesse para previsão ou podem simplemente ser usadas como variáveis explicativas



Vector Auto-Regression (VAR)

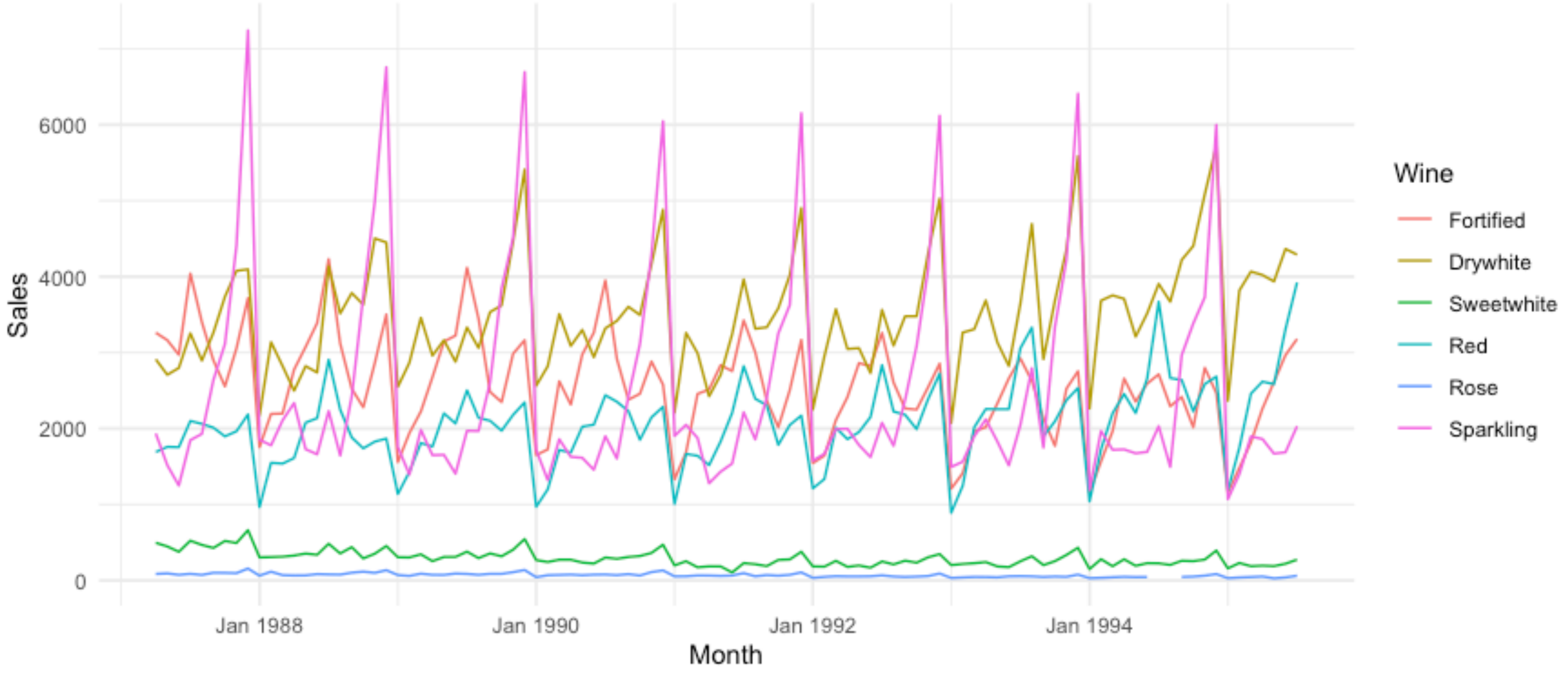
VAR é um método para séries temporais multivariadas

- "Auto-Regression" porque cada variável é modelada de acordo com as suas observações passadas e as observações passadas das restantes variáveis



Auto-Regressive Distributed Lag (ARDL)

- Parecido com VAR, mas apenas uma das variáveis é de interesse para previsão
  - As restantes podem ser usadas como variáveis explicativas



Modelos de Séries Temporais

Vitor Cerqueira
https://www.linkedin.com/in/vcerq/
cerqueira.vitormanuel@gmail.com

NUCLIO
DIGITAL SCHOOL