# Introdução aos modelos de previsão de séries temporais

VIII Encontro de Economia do Espírito Santo

Prof. Guilherme Pereira
Departamento de Economia - UFES

10 de novembro de 2021

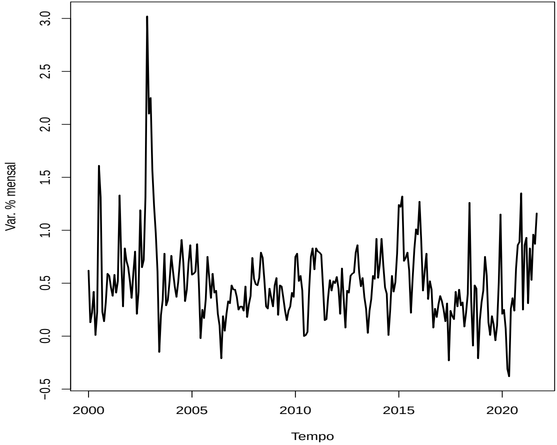
## Série temporal

- Variável aleatória observada ao longo do tempo.
    - dependência temporal!
- Escalas mais comuns em economia: diária, mensal, trimestral, anual etc.

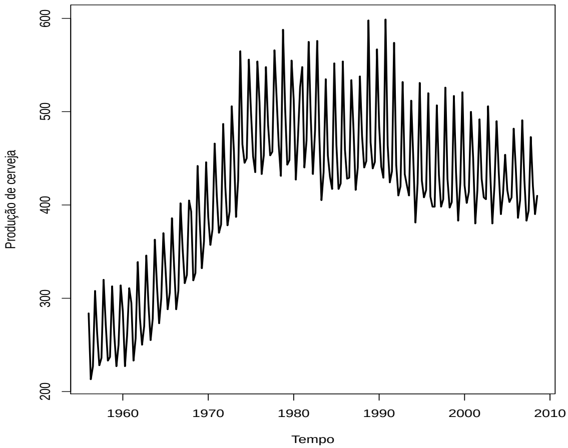
## Ex.

- Taxa de câmbio - R$/U$$ - comercial compra média (diária)
- Preços de commodities - carnes - (mensal)
- PIB - (trimestral)
- Renda - desigualdade - coeficiente de Gini - (anual)
- Vendas / Receitas/ Custos de determinada empresa.
- Retornos diários das cotações (PETROBRAS, IBOVESPA...)

Exemplo I - Índice nacional de preços ao consumidor-amplo (IPCA)
Frequência: mensal - 01.2000 até 09.2021.



## Exemplo II - Produção de cerveja na Austrália
## Frequência: trimestral 01.1956 até 03.2008



# Algumas características importantes

A análise de **séries temporais** busca identificar características importantes presentes nos dados tais como:

- tendência
- sazonalidade
- autocorrelação, dependência não-linear, etc.

Identificado tais padrões, estima-se o modelo mais adequado. Tal modelo pode ser utilizado para:

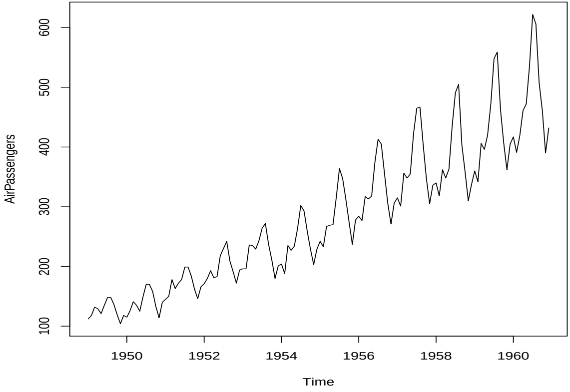
- previsão, simulação...

Identificando algumas características importantes...

- tendência
- estacionaridade
- sazonalidade
- autocorrelação, etc.

# Ex: Tendência

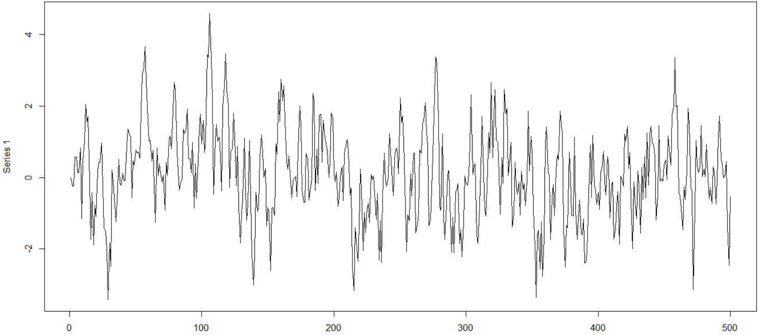
Número total de vendas de passagens aéreas entre 1949 e 1960.



Identificando algumas características importantes...

- tendência
- estacionaridade
- sazonalidade
- autocorrelação, etc.

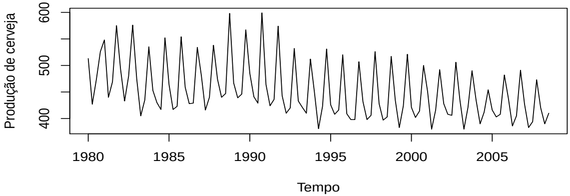
# Ex: Série Estacionária

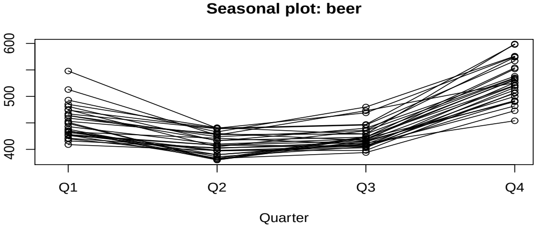


Identificando algumas características importantes...

- tendência
- estacionaridade
- sazonalidade
- autocorrelação, etc.

# Ex: Sazonalidade



Identificando algumas características importantes...

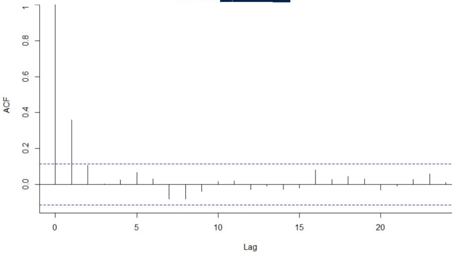
- tendência
- estacionaridade
- sazonalidade
- autocorrelação, etc.

Autocorrelação indica dependência linear de uma série temporal.

- dependência no tempo, isto é, o quanto $Y_t$ depende das observações defasadas $Y_{t-k}$.
- $\rho(k) = Corr(y_t, y_{t-k}) = \dfrac{Cov[Y_t, Y_{t-k}]}{\sqrt{Var(Y_T), Var(Y_{t-k})}}$.
- $-1 < \rho(k) < 1$

Função de autocorrelação:

No *R*: `> acf(serie)`



# Decomposição clássica de séries temporais

Um série temporal ($Y_t$) pode ser decomposta em componentes:

- $T_t$: tendência-ciclo
- $S_t$: sazonalidade
- $E_t$: componente aleatório

## Decomposição Aditiva

$$Y_t = T_t + S_t + E_t$$

## Decomposição Multiplicativa

$$Y_t = T_t \times S_t \times E_t$$

Na *decomposição clássica* o objetivo é estimar cada um dos componentes.

# Roteiro para decom. clássica

1. Estime $\hat{T}_t$ (médias móveis, regressão...).
2. Remova a tendência da série ($Y_t - \hat{T}_t$ ou $\frac{Y_t}{\hat{T}_t}$).
3. Estime os fatores sazonais a partir da *série sem tendência* (médias do período, regressão com dummies...)
4. Estime o componente irregular ('resíduo')
   - $\hat{E}_t = Y_t - \hat{T}_t - \hat{S}_t$
   - $\hat{E}_t = Y_t / (\hat{T}_t \times \hat{S}_t)$
5. A nova série $\hat{E}_t$ pode ser utilizada para análises, previsões e/ou estimação de modelos estacionários.

# Exemplo de estimação dos componentes

## $\hat{T}_t$ - Média Móvel(m)

$$\hat{T}_t = \frac{1}{m} \sum_{j=-k}^{k} y_{t+j}, \qquad m = 2k + 1.$$

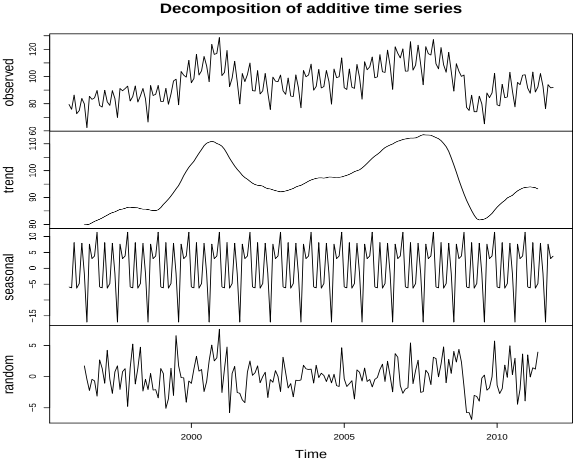
## $\hat{S}_t$ - média do período

Em uma série mensal, por exemplo, $\hat{S}_t$ pode ser a média(mediana) dos meses.

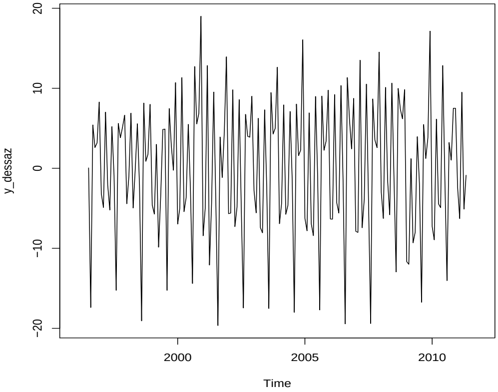
No R:

```
######################## Decomposição Clássica ######################

# Carregando os dados do pacote fpp. Série do índice de novos pedidos
# de fabricação de equipamentos eletrônicos.
data(elecequip)

decomp <- decompose(elecequip,type="additive")
plot(decomp)

# Podemos obter os componentes de forma isolada:
# Componente Sazonal
decomp$seasonal
# Componente de tendência
decomp$trend
# Compoenete "aleatório"
decomp$random

# Com base nos componentes e na série original, podemos obter, por exemplo,
# a série "sem tendência"
y_dessaz <- elecequip - decomp$trend
plot(y_dessaz)
```

```
data(elecequip)

decomp <- decompose(elecequip,type="additive")
```



```
y_dessaz <- elecequip - decomp$trend
plot(y_dessaz)
```



## Previsão de séries temporais

Projeção de valores futuros a partir do passado.

- Utilizando a série histórica $(y_1, y_2, ...y_T)$ realizamos previsões probabilísticas de $(\hat{y}_{T+1|T}, \hat{y}_{T+2|T}, \hat{y}_{T+h|T})$

## Horizonte de previsão - Previsão *h*-passos à frente

- Número de instantes de tempo (passos à frente) a serem previstos.
  - Ex: Previsão 1 passo à frente de uma série mensal de vendas significa que irei prever o volume de vendas para o próximo mês.

## Notação:

$\hat{y}_{T+1|T}$ é a previsão da variável $y$ para o instante $T+1$ tendo como origem o instante t $T$.

## Previsão de séries temporais

- previsão pontual
- previsão intervalar (intervalo de confiança)

Previsões podem ser de

- Curtíssimo prazo
- Curto prazo
- Médio prazo
- Longo prazo

A qualidade das previsões está sujeita

- principalmente ao grau de previsibilidade da série.
- ao horizonte de previsão.
    - incerteza aumenta conforme nos distanciamos do presente.

# Erro de previsão

- Uma vez estimado determinado modelo, podemos utilizar este para projetar os valores históricos da série em análise.
- Podemos comparar os valores reais vs. valores previstos/valores ajustados.
- Ao calcularmos os valores reais **menos** os valores previstos, obtemos os **erros de previsão** um passo à frente.
- O erro de previsão será útil para estabelecermos métricas de avaliação dos modelos.

# Erro de previsão

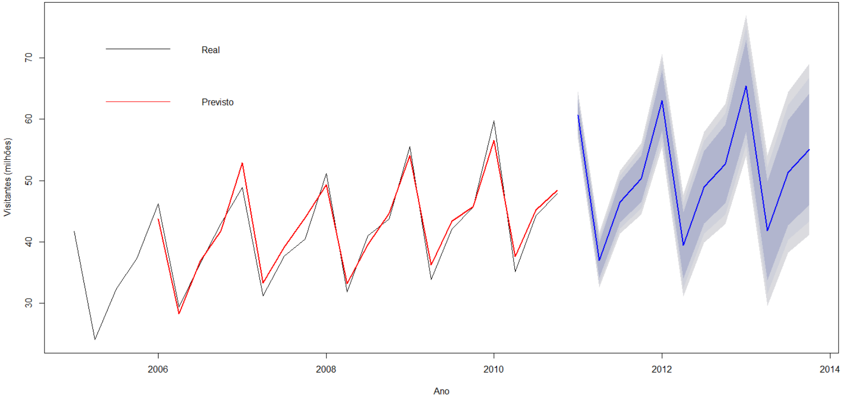
## Erro de previsão

$$y_{t,real} - \hat{y}_{t,previsto} = \hat{\varepsilon}_t$$

$$y_t - \hat{y}_t = \hat{\varepsilon}_t$$

É desejável que os erros sejam os menores possíveis !

Na figura abaixo podemos observar os valores reais, previstos (dentro da amostra), os valores previstos fora da amostra (futuro) e os intervalos de previsão.



# Métricas para avaliação

É importante termos em mente que as métricas de avaliação geralmente são calculadas com base no erro de previsão um passo à frente, i.e, $\hat{\varepsilon}_{t/t-1} = y_t - \hat{y}_{t/t-1}$.

Existem diversas métricas, contudo, neste minicurso iremos utilizar:

- MAPE
- RMSE

# Métricas para avaliação

## MAPE - mean absolute percentual erro

$$MAPE = \frac{\sum_{t=1}^{T} \frac{|\hat{\varepsilon}_{t/t-1}|}{y_t} \times 100}{T}$$

- Medida expressa o erro de previsão em termos percentuais.
- Pode ser utilizada para comparar modelos com escalas diferentes.
  - ex: $y_t$ e $log(y_t)$

## RMSE - root mean square error

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} \left(\hat{\varepsilon}_{t/t-1}\right)^2}{T}}$$

- Medida construída com base no erro quadrático.

1. Conceitos gerais
   - Decomposição clássica
   - Previsão e métricas para avaliação
   - Análise dos resíduos

2. Métodos de previsão
   - Naive
   - Amortecimento Exponencial Simples
   - Amortecimento Exponencial de Holt
   - Holt-Winters

A fase de diagnósticos de adequação é realizada sobre os resíduos estimados:

Resíduos estimados: $\hat{\varepsilon}_t = y_t - \hat{y}_t$ onde $\hat{y}_t$ denota o valor previsto para o modelo.

Em um primeiro momento, analisamos as seguintes características dos resíduos:

- autocorrelação;
- normalidade;

Existem diversos testes disponíveis para cada uma dessas características, contudo, iremos apresentar os principais.

## Normalidade - JB

Hipóteses:

- $H_0$ : Dados seguem uma distribuição normal.
- $H_1$ : Dados não seguem uma distribuição normal.

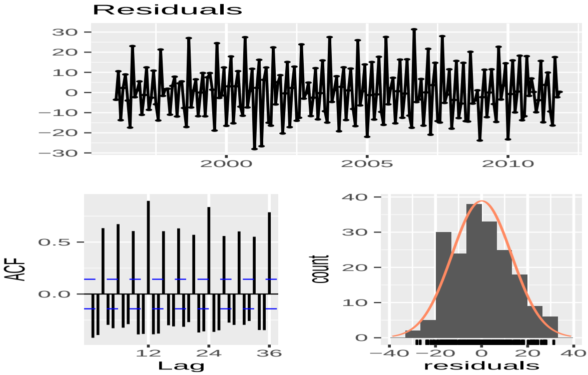
## Autocorrelação: Teste de Ljung-Box

Hipóteses:

- $H_0 : \rho_1 = \rho_2 = ... = \rho_k = 0$ .
- $H_1$ : pelo menos um dos $\rho$s é diferente de zero.

Exemplo: análise dos resíduos obtidos a partir de um modelo Naive.

```
############  Análise dos resíduos #############
data("elecequip") #lendo dados
modelo.naive <- naive(elecequip) # estimando o modelo Naive
residuo <- residuals(modelo.naive) # obtendo os resíduos do modelo

# Análise Gráfica
checkresiduals(residuo)
```



Saídas:

```
> # Teste de Ljung-Box
> Box.test(residuo,lag=10,  type="Lj")

        Box-Ljung test

data:  residuo
X-squared = 413.57, df = 10, p-value < 2.2e-16

> #Normalidade
> tseries::jarque.bera.test(na.remove(residuo))

        Jarque Bera Test

data:  na.remove(residuo)
X-squared = 5.0908, df = 2, p-value = 0.07844
```

Conclusão:

- Os resíduos são autocorrelacionados.
- Os resíduos seguem uma distribuição normal.

Lembre-se que o desejado é que os resíduos não sejam autocorrelacionados e que sigam aproximadamente uma dist. normal.

# Método ingênuo - naive

- Método bastante simples para previsão de séries temporais.
- A previsão para o instante $t$ ($\hat{y}_{t+1|t}$) é dada pelo último valor observado ($y_t$).
- Modelos mais complexos devem ser comparados com este método.

## Equação de previsão - método naive

- $\hat{y}_{t+1|t} = y_t,$
- $\hat{y}_{t+h|t} = y_t, \qquad h = 1, 2, 3...$

1. Conceitos gerais
   - Decomposição clássica
   - Previsão e métricas para avaliação
   - Análise dos resíduos

2. Métodos de previsão
   - Naive
   - Amortecimento Exponencial Simples
   - Amortecimento Exponencial de Holt
   - Holt-Winters

# Amortecimento Exponencial Simples (*Single Exponencial Smoothing*)

- Indicado para séries livres tendência e sazonalidade.
- $\hat{y}_{T+1|T}$ é a previsão para $y_{T+1}$ realizada em $T$.

O AES pondera todo o histórico da série para realizar a previsão:

$$\hat{y}_{T+1|T} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \ldots \tag{1}$$

onde $\alpha \in (0, 1)$ é a constante de amortecimento.

- Todas as informações são amortecidas.

É possível demostrar que a eq. 1 pode ser reescrita de forma sucinta:

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha)\hat{y}_{t|t-1}$$

- Quanto maior o valor de $\alpha$, maior será o peso dado a informação mais recente.
- Em séries mais "nervosas" geralmente utiliza-se um $\alpha$ maior.
- Em séries mais comportadas dá-se mais peso para informações passadas.
    - O modelo responde mais lentamente às mudanças no comportamento da série.

**Ideia geral do modelo**

$$\hat{y}_{t+1|t} = \alpha^{*}\text{presente} + (1 - \alpha)^{*}\text{passado}$$

# Amortecimento Exponencial Simples no R

```
# Amortecimento Exponencial
data("oil")
oildata <- window(oil, start=1996)

#Estimando o modelo e realizando as previsões
modelo <- ses(oildata, h=3)
modelo$model

# Gráfico
autoplot(modelo,ylab="Oil production in Saudi Arabia",xlab="Year") +
  autolayer(fitted(modelo), series="Fitted")

#Valores previstos
print(modelo)
```

# Amortecimento Exponencial Simples no R


```
> #Valores previstos
> print(modelo)
     Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
2011       469.2669 433.4186 505.1153 414.4416 524.0923
2012       469.2669 423.6608 514.8730 399.5184 539.0155
2013       469.2669 415.6505 522.8834 387.2676 551.2663
```

# Holt 2 parâmetros

- Similar ao AES porém possui o componente de tendência.
- Possui duas equações:
  - uma para o nível
  - uma para a tendência.

## Modelo

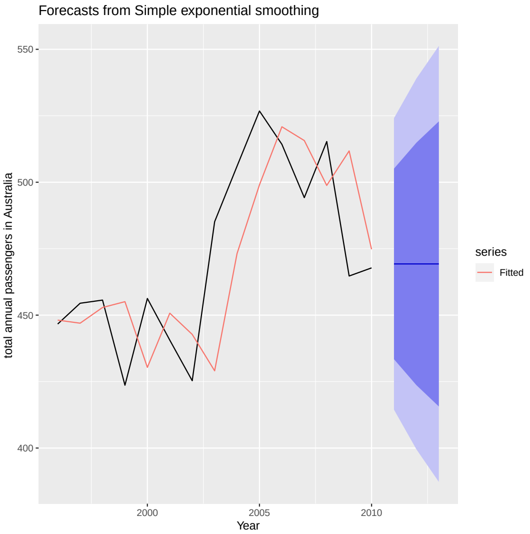
| | | | |
|---|---|---|---|
| Equação de previsão: | $\hat{y}_{t+h\|t}$ | $=$ | $\ell_t + hb_t$ |
| Equação do nível: | $\ell_t$ | $=$ | $\alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$ |
| Equação da tendência: | $b_t$ | $=$ | $\beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}$ |

onde $\ell_t$, $b_t$ denotam respectivamente o nível e a tendência prevista para o instante $t$, $0 < \alpha < 1$ é o parâmetro de suavização do nível e $0 < \beta < 1$ é o parâmetro de suavização da tendência.

# Holt 2 parâmetro no R

```
# Holt 2 Parâmetros
data("ausair")
air <- window(ausair, start=1990)
plot(air)

# Estimando o modelo & realizando as previsões.
modelo <- holt(air,h=5,level = c(80,90))
modelo$model

# Gráfico
autoplot(modelo,ylab="total annual passengers in Australia ",xlab="Year") +
  autolayer(fitted(modelo), series="Fitted")

#Valores previstos
print(modelo)
```

# Holt 2 parâmetro no R



1. Conceitos gerais
   - Decomposição clássica
   - Previsão e métricas para avaliação
   - Análise dos resíduos

2. Métodos de previsão
   - Naive
   - Amortecimento Exponencial Simples
   - Amortecimento Exponencial de Holt
   - Holt-Winters

# Método de Holt-Winters

Método adequado para se modelar séries que possuem os seguintes componentes:

- nível ($\ell_t$).
- tendência ($b_t$).
- sazonalidade ($s_t$)

A sazonalidade pode ser modelada da seguinte forma:

- aditiva
- multiplicativa

## Modelo Aditivo

Equação de previsão:

- $\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$

Equação do nível:

- $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$

Equação da tendência:

- $b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}$

Equação de suavização:

- $s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$

onde $\alpha, \beta, \gamma \in (0,1)$ denotam os parâmetros de suavização para nível, tendência e sazonalidade respectivamente. $k$ representa a parte inteira de $(h-1)/m$

## Modelo Multiplicativo

Equação de previsão:

- $\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$

Equação do nível:

- $\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1})$

Equação da tendência:

- $b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}$

Equação de suavização:

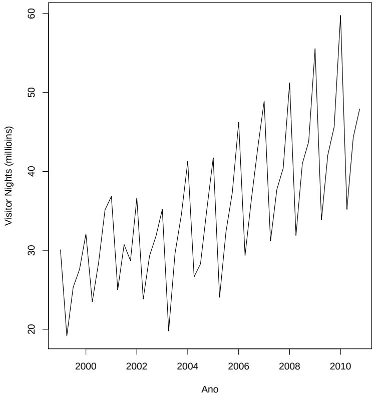
- $s_t = \gamma \frac{y_t}{\ell_{t-1} - b_{t-1}} + (1-\gamma)s_{t-m}$

# Holt-Winters no R

```
#Holt-Winters
data("austourists") # quarterly visitor nights in Australia spent by
#international tourists.
aust <- window(austourists,start=2005)

#Gráfico
plot(aust,ylab="Visitor Nights (millioins)",xlab="Year")
# Gráfico Sazonal
seasonplot(aust)

#Estimação e Previsão
modelo.ad <- hw(aust, h=12, seasonal = "additive",level = c(60,80, 95))
print(modelo.ad$model)
#modelo.mult <- hw(aust, h=12, seasonal = "multiplicative") - MULTIPLICATIVO


# Gráfico
autoplot(modelo.ad,ylab="Visitor Nights (millioins) ",xlab="Year") +
  autolayer(fitted(modelo.ad), series="Fitted")
```

# Holt-Winters no R

Gráfico da série:

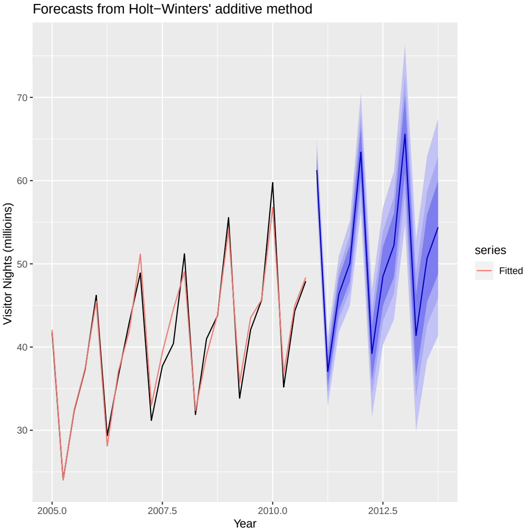


# Holt-Winters no R

Componentes e constantes de amortecimento.

```
> print(modelo.ad$model)
Holt-Winters' additive method

Call:
 hw(y = aust, h = 12, seasonal = "additive", level = c(60, 80,

 Call:
     95))

  Smoothing parameters:
    alpha = 0.3625
    beta  = 0.0754
    gamma = 0.6375

  Initial states:
    l = 31.6429
    b = 1.0028
    s = 2.0685 -2.0274 -9.4676 9.4264

  sigma:  1.9369

     AIC     AICc      BIC
116.2736 129.1308 126.8761
```

# Holt-Winters no R



# Referências

[1] Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on 08/11/2021.

[2] Hyndman , R.J., (2013). fpp: Data for "Forecasting: principles and practice". R package version 0.5. https://CRAN.R-project.org/package=fpp

[3] Ferreira, P. C., de Mattos, D. M., de Almeida, D. C. V., & de Oliveira, I. C. L. (2017). Análise de Séries Temporais em R: curso introdutório. Elsevier.