

# Séries Temporais

Disciplina: Modelos Estatísticos  
Professora: Jéssica Assunção

# Definição de Séries Temporais

# Definição

- Dados Coletados em Intervalo Regulares de Tempo
  - Intervalos:  
Milissegundos, hora, dia, trimestre, ano...
- Ordem < Dependência da ordem



# Exemplos

Evasão de alunos por mês



Internações por dia

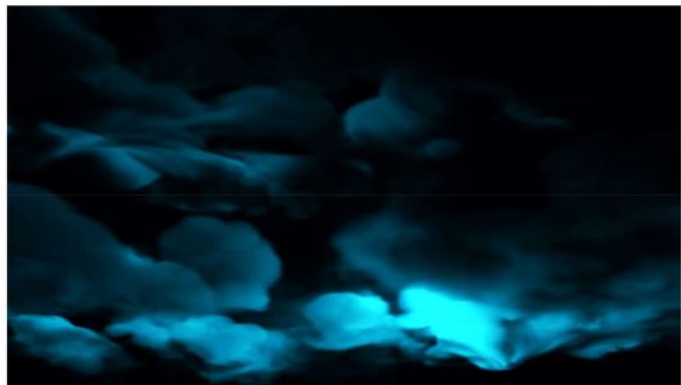


Mortes por problemas cardíacos por ano

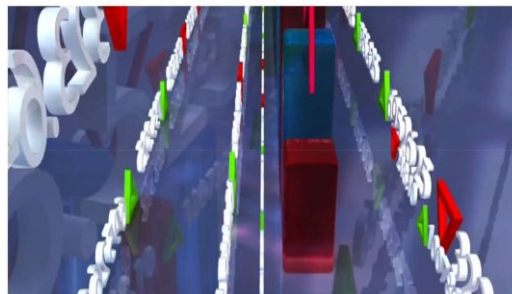


# Exemplos

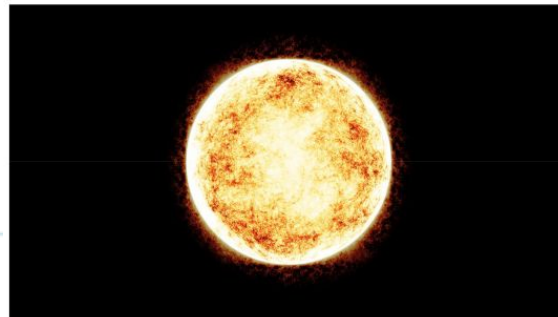
Índice pluviométrico mensal



Cotação da bolsa por minuto



Manchas solares por mês



# Componentes



# Ordem

■ Total de pessoas atendidas em um Hospital por dia

Medida

Fato

Unidade de  
Tempo



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

01/03	02/03	03/03
404		298



# Univariadas e Multivariadas

- Univariadas  
Apenas uma variável conectada ao tempo
- Multivariadas  
Duas ou mais variáveis conectadas ao tempo



# Por que estudarmos Séries Temporais

- Compreender fenômenos
- Prever eventos

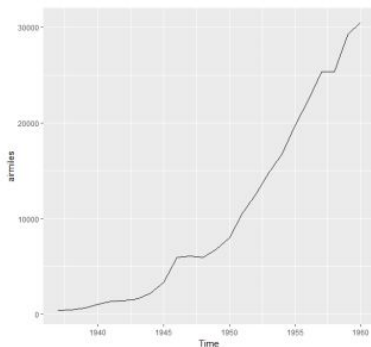


# Vamos ver na prática...

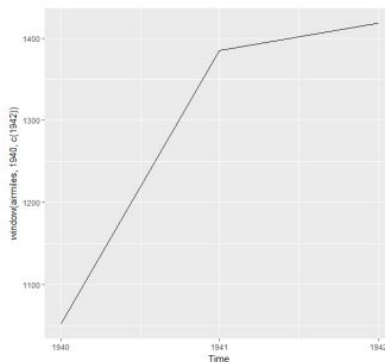
## airmiles

Milhas áreas de passageiros nos EUA de 1937 até 1960

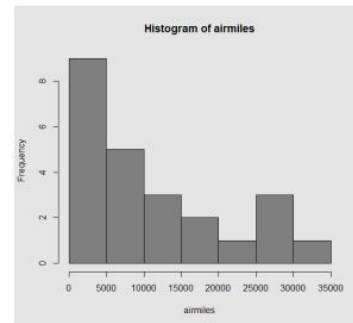
Gráfico



Janela 1940-1942



Histograma



Estatísticas

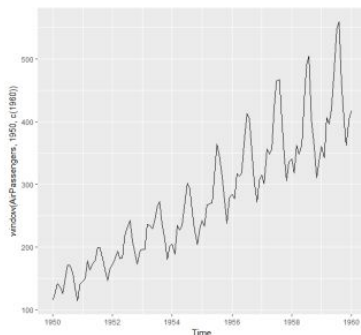
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
412	1580	6431	10528	17532	30514

# Vamos ver na prática...

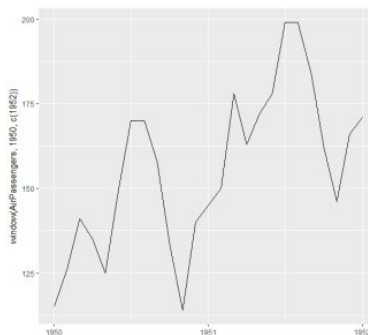
## AirPassengers

Passageiros de linha aérea mensais de 1949 até 1960

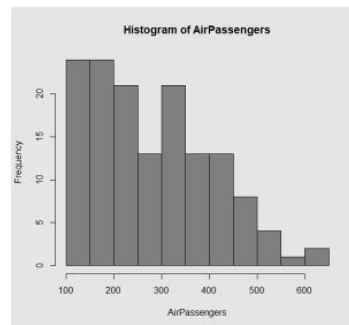
Gráfico



Janela 1950-1952



Histograma



Estatísticas

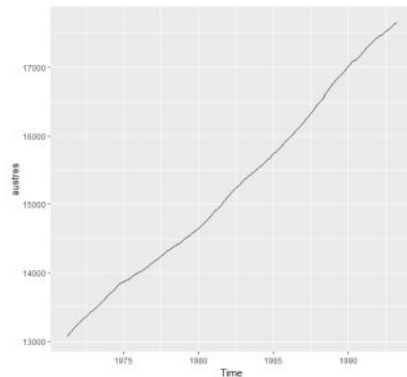
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
104.0	180.0	265.5	280.3	360.5	622.0

# Vamos ver na prática...

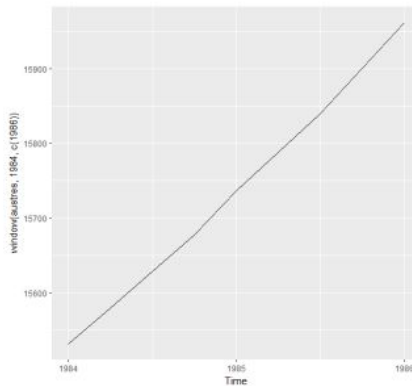
austres

Dados trimestrais de residentes australianos 1971- 1994

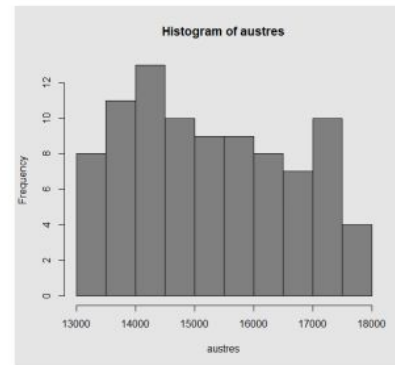
Gráfico



Janela 1984-1986



Histograma



Estatísticas

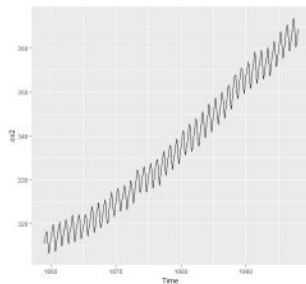
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
13067	14110	15184	15273	16399	17662

# Vamos ver na prática...

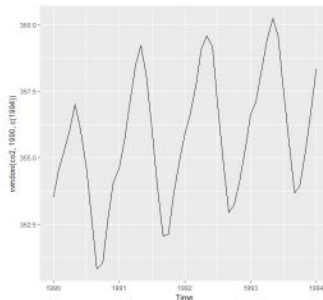
co2

Concentração de co2 em Mauna Loa de 1959 a 1997 (trimestral)

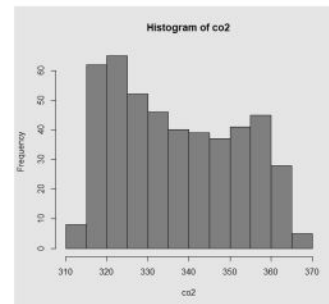
Gráfico



Janela 1990-1994



Histograma



Estatísticas

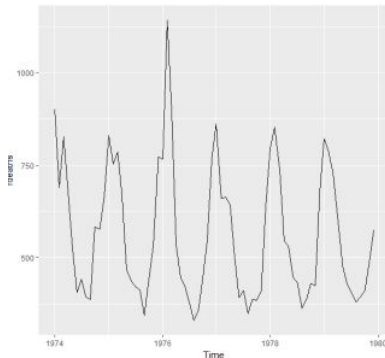
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
313.2	323.5	335.2	337.1	350.3	366.8

# Vamos ver na prática...

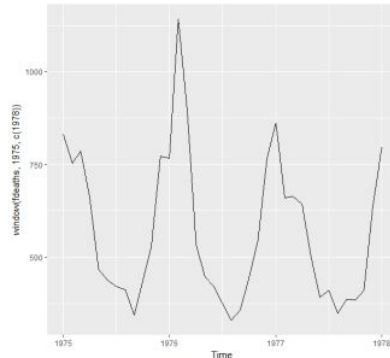
fdeaths

Mortes mensais de doenças pulmonares no Reino Unido de 1974-1979

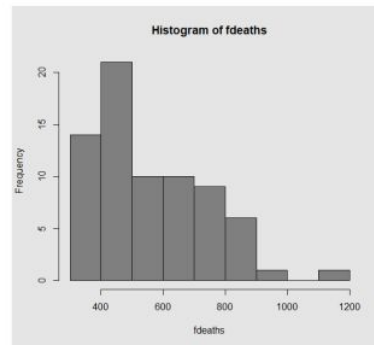
Gráfico



Janela 1975-1978



Histograma



Estatísticas

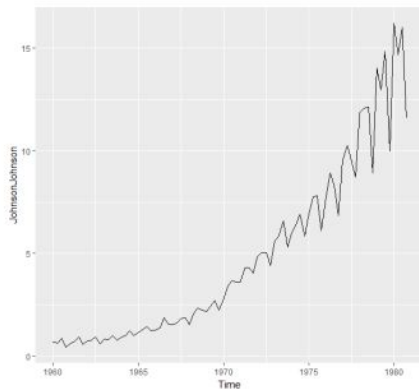
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
330.0	411.0	512.0	560.7	681.5	1141.0

# Vamos ver na prática...

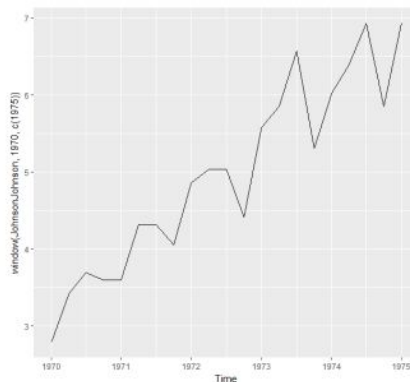
## JohnsonJohnson

Lucro trimestral por ação da Johnson & Johnson de 1960 a 1980

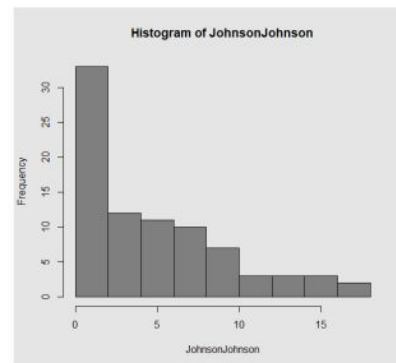
Gráfico



Janela 1970-1975



Histograma



Estatísticas

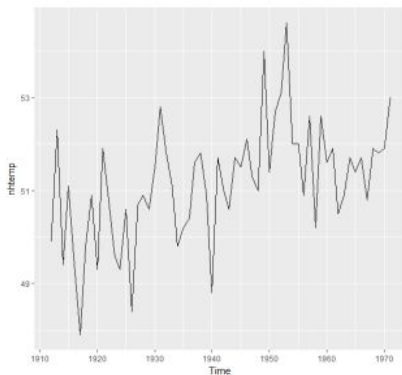
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.440	1.248	3.510	4.800	7.133	16.200

# Vamos ver na prática...

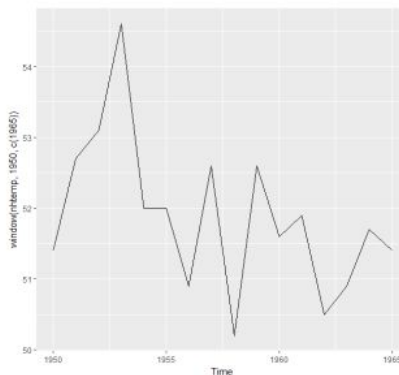
## nhtemp

Temperatura anual in Fahrenheit in New Heaven de 1912 até 1971

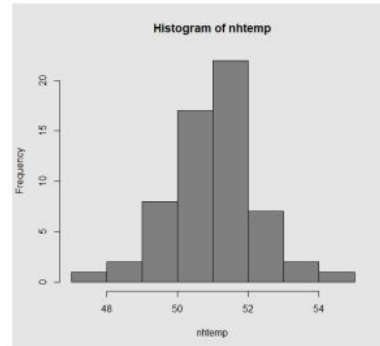
Gráfico



Janela 1950-1965



Histograma



Estatísticas

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
47.90	50.58	51.20	51.16	51.90	54.60

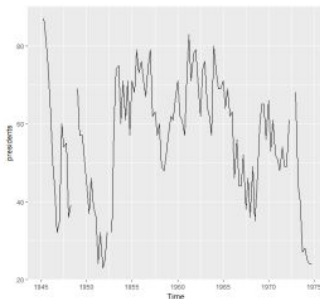


# Vamos ver na prática...

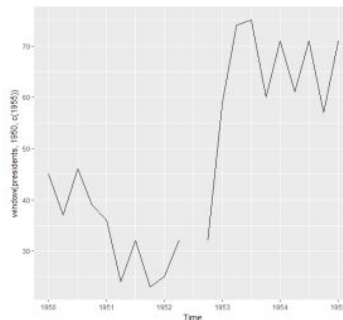
## presidents

Aprovação trimestral dos presidentes americanos de 1945 até 1974

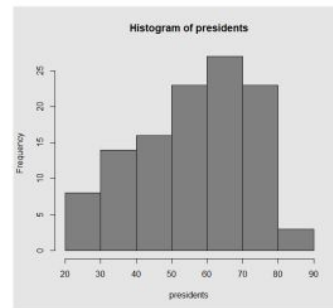
Gráfico



Janela 1950-1955



Histograma



Estatísticas

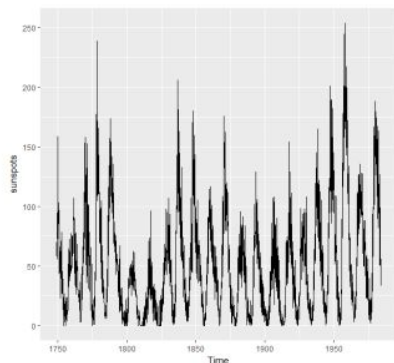
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
23.00	46.00	59.00	56.31	69.00	87.00	6

# Vamos ver na prática...

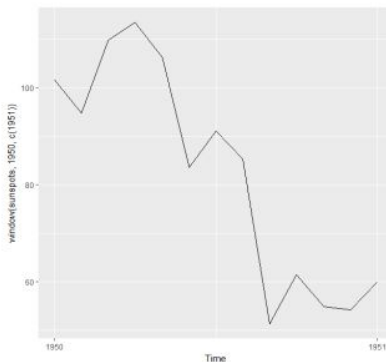
## sunspots

Número médio de manchas solares de 1749 a 1983

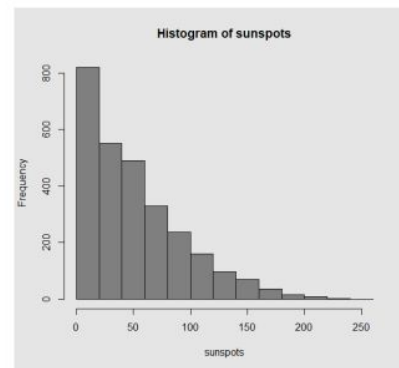
Gráfico



Janela 1950-1951



Histograma



Estatísticas

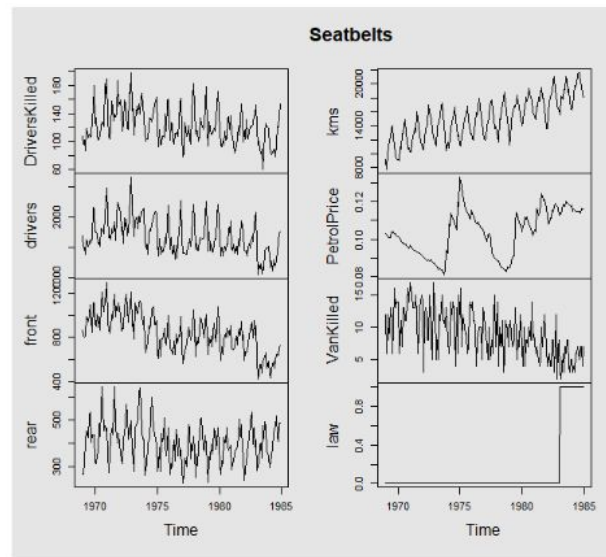
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	15.70	42.00	51.27	74.92	253.80

# Vamos ver na prática...

## Seatbelts

Mortes ou gravemente feridos em acidentes de carros de 1969 até 1984

- DriversKilled: motoristas mortos
- Front: passageiros do banco da frente mortos
- Rear: passageiros do banco traseiro mortos
- Kms: distância dirigida
- PetrolPrice: preço do combustível
- VanKilled: número de motoristas de vans
- Law: 0/1: vigência da lei de obrigatoriedade do uso do cinto



# Vamos ver na prática...

## EuStockMarkets

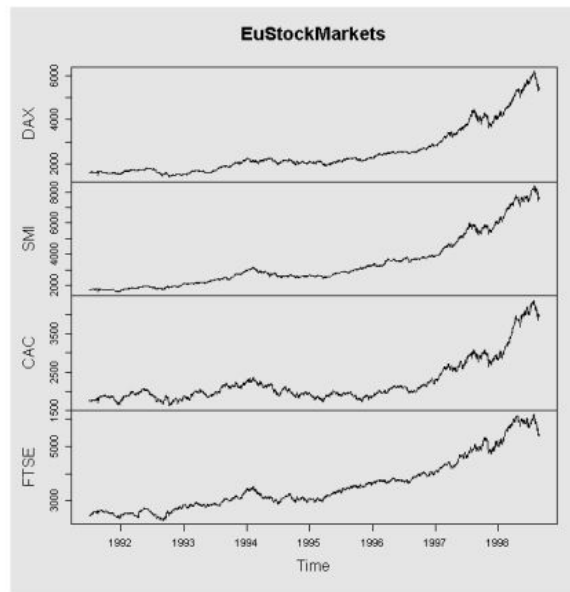
Preço de Fechamento das Bolsas Europeias de 1991 até 1998

Alemanha (DAX)

Suíça (SMI)

França (CAC)

Reino Unido (FTSE)

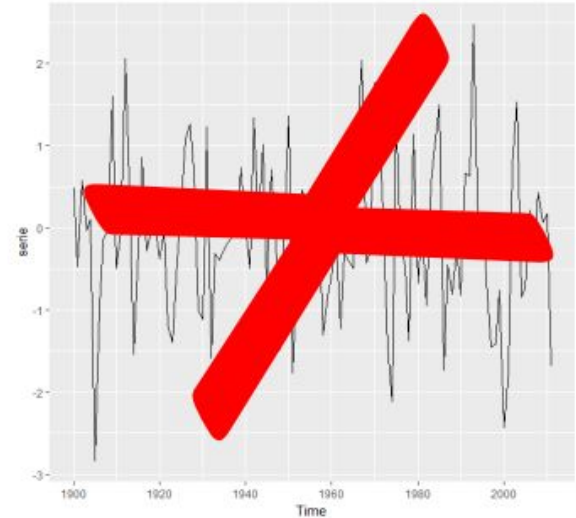
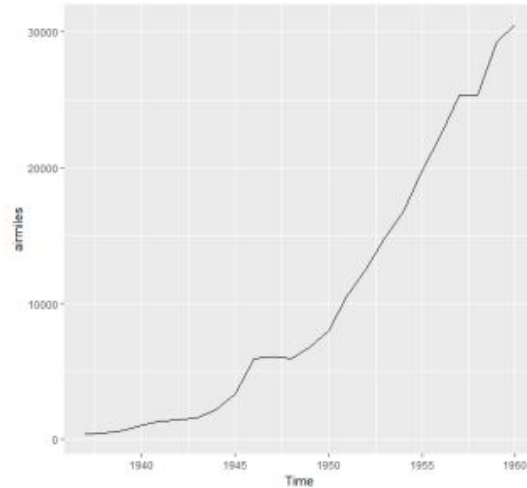
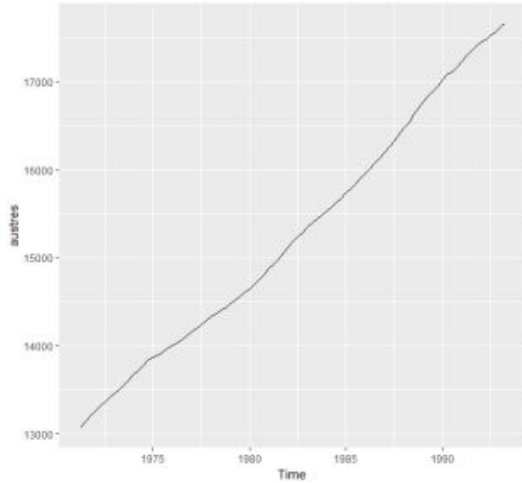


# Componentes e Padrões

- Tendência
- Sazonalidade
  - Ciclo
  - Erro

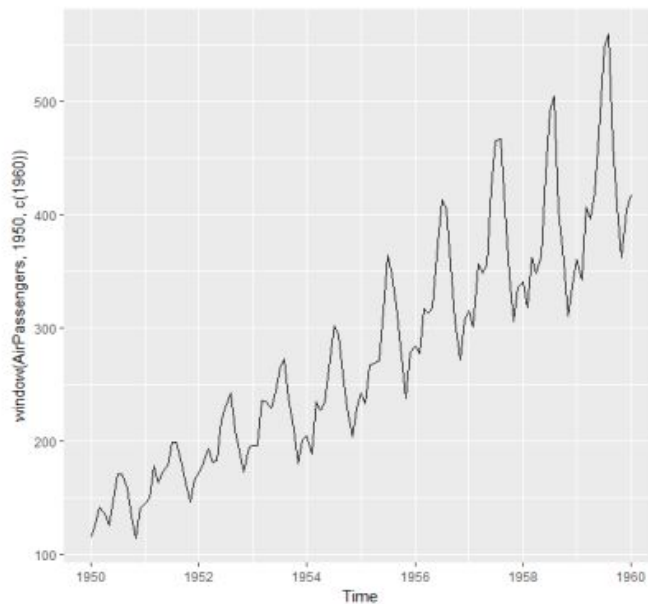
# Tendência

- Aumento ou redução a longo prazo



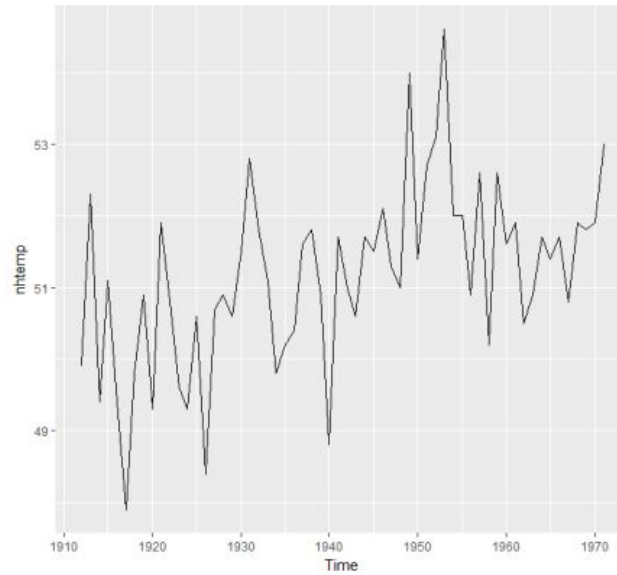
# Sazonalidade

- Padrões que ocorrem em intervalos fixos



# Ciclo

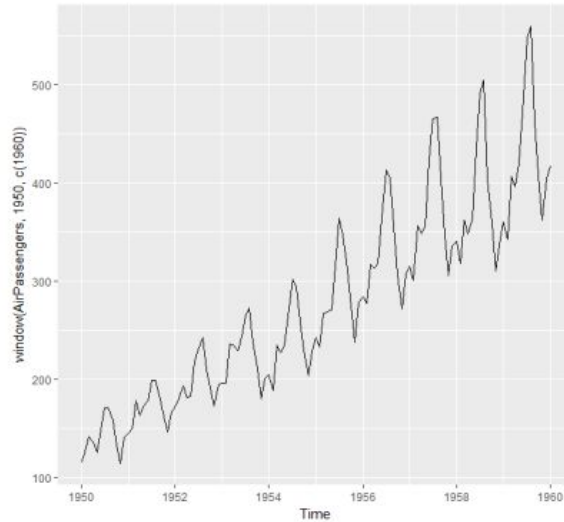
- Aumento ou redução de frequência sem intervalos fixos





# Erro

Tendência + Sazonalidade + Ciclo + ERRO =

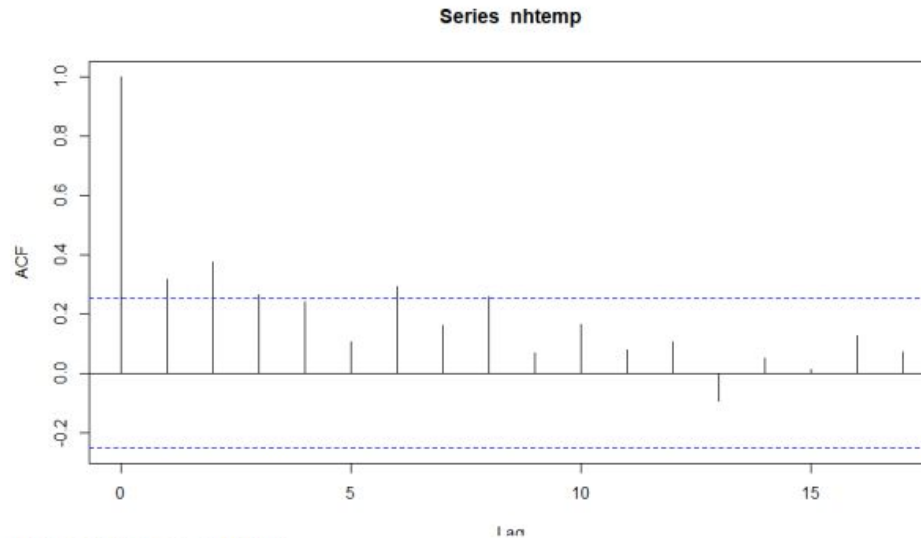


# Autocorrelação

- Lembra do correlação?
- A autocorrelação mede se existe uma relação matemática entre os intervalos da série temporal.
- Também deve estar entre -1 e 1.
- É medida em intervalos (lag):
  - De 1 intervalo: mede como valores de 1 período distantes estão relacionados.
  - De 2 intervalos: mede como valores de 2 períodos distantes estão relacionados.

# Autocorrelação - Diagramas

- ACF - Mostra a autocorrelação em uma série temporal

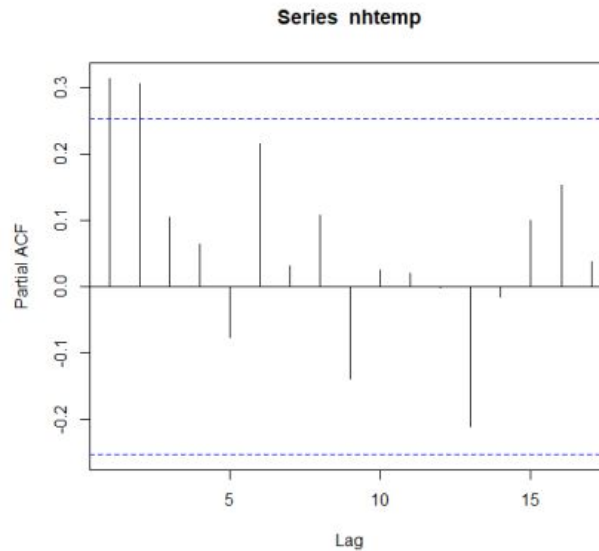


Autocorrelations of series 'nhtemp', by lag

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.000	0.315	0.375	0.264	0.241	0.106	0.292	0.161	0.259	0.068	0.165	0.077	0.107	-0.094	0.051	0.011	0.126	0.073

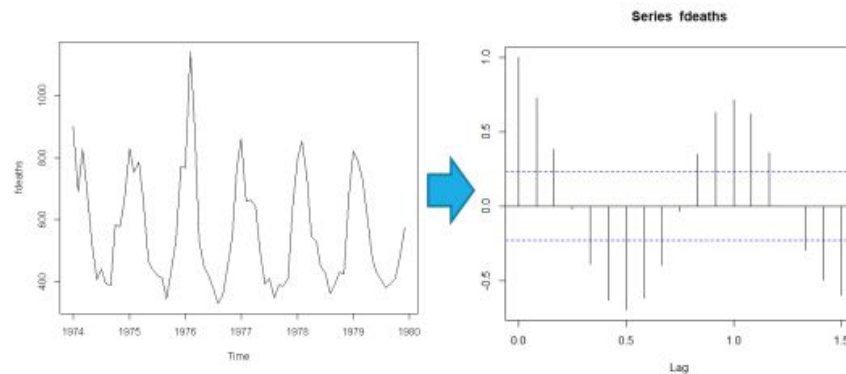
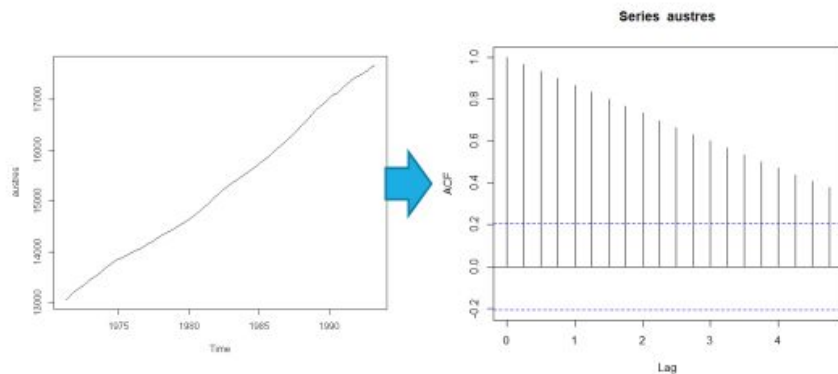
# Autocorrelação - Diagramas

- PACF - Mede a autocorrelação entre diferentes intervalos dos resíduos



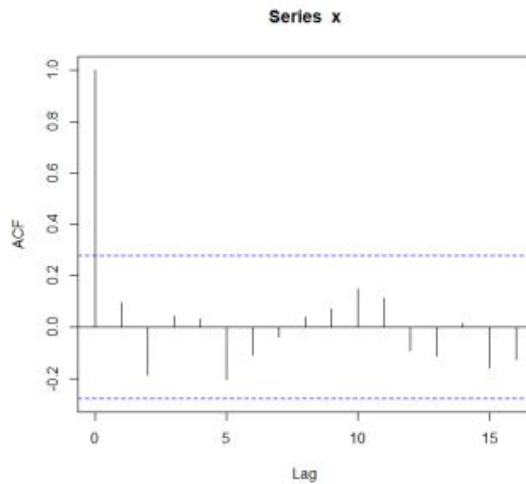
# Autocorrelação - com sazonalidade e tendência

- Com tendência: tende a ser grande e positiva
- Com sazonalidade: tende a ser maior nos intervalos com maiores valores

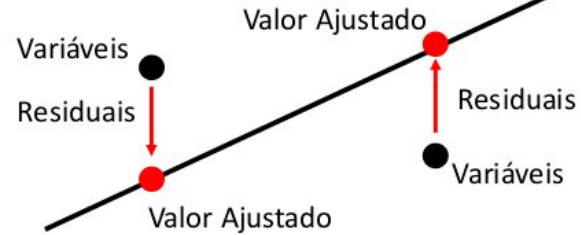
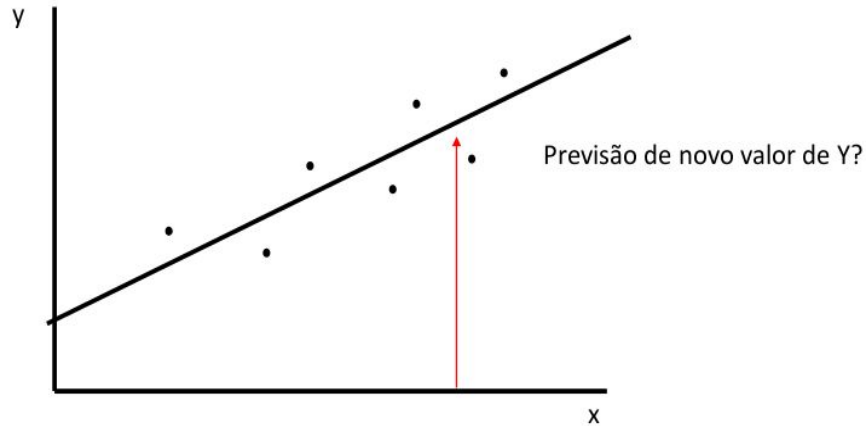


# Ruído Branco

- Série temporal que não apresenta autocorrelação
- Se mais de 5% dos intervalos estiverem fora da linha azul, não é ruído branco



# Regressão Linear - Erro



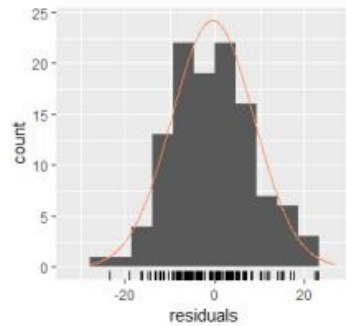
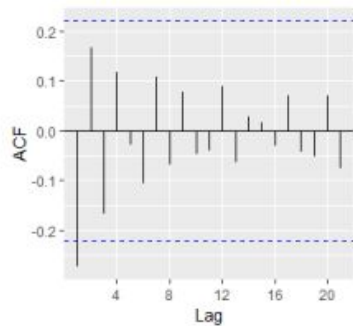
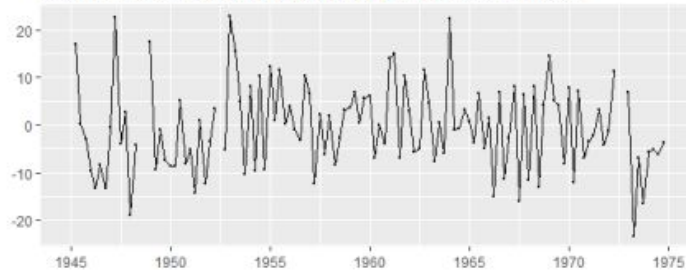
# Residuais: pressupostos para a previsão

- Não devem estar relacionados (autocorrelação)
  - Se existe autocorrelação, existe informação nos resíduos que deveria ser usada na previsão.
- Média deve ser zero (aproximadamente)
- Variância constante
- Distribuição Normal



# Analizando Residuais

Residuals from ARIMA(1,0,0)(1,0,0)[4] with non-zero mean



Ljung-Box test

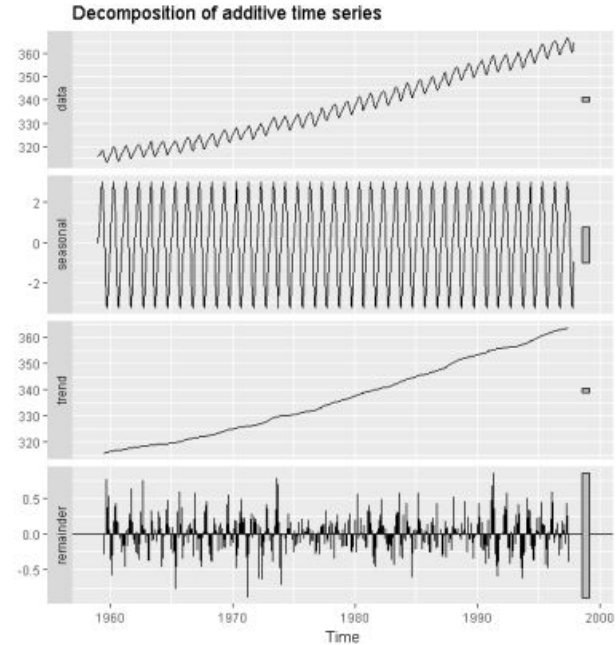
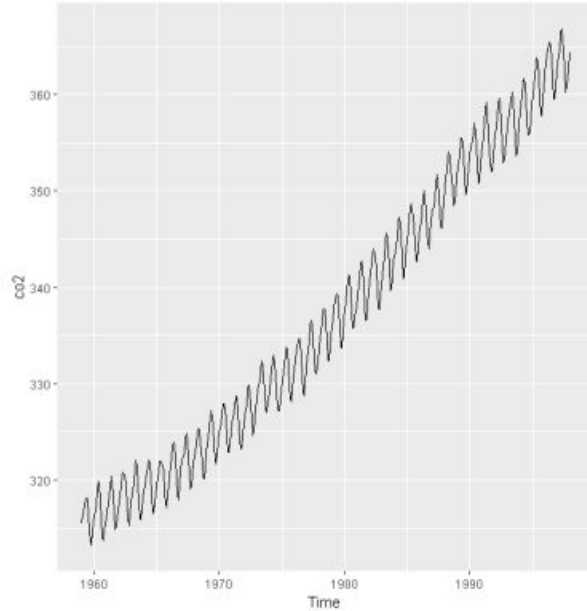
```
data: Residuals from ARIMA(1,0,1)(0,1,1)[12]  
Q* = 25.371, df = 21, p-value = 0.2315
```

```
Model df: 3. Total lags used: 24
```

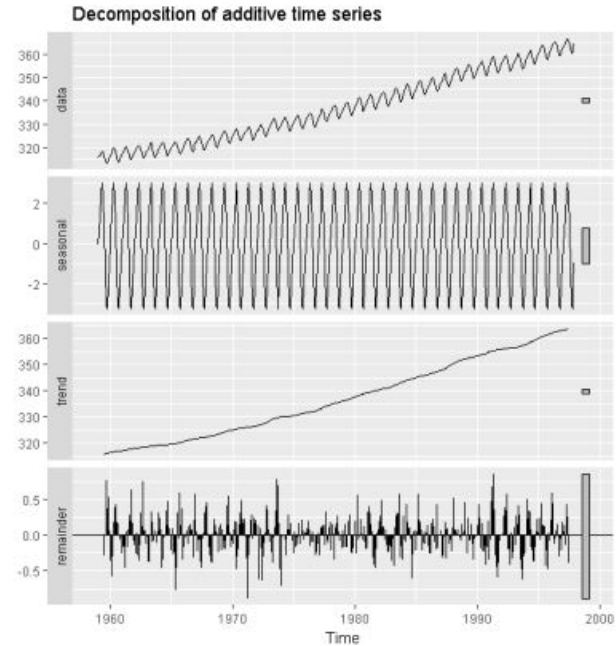
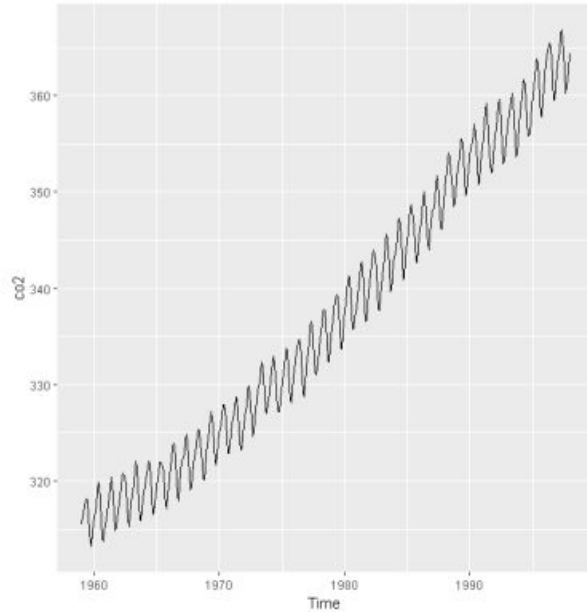
# Decomposição

- Processo de separar componentes de uma série temporal
- Por que?
  - Compreender a série
  - Fazer previsão
- Modelo Aditivo: melhor quando o elemento sazonal ou tendência não é proporcional com o nível da série.
- Modelo Multiplicativo: melhor quando o elemento sazonal ou tendência muda com o nível da série.
- Nível: valor médio da série

# Decomposição - Exemplo



# Decomposição - Exemplo



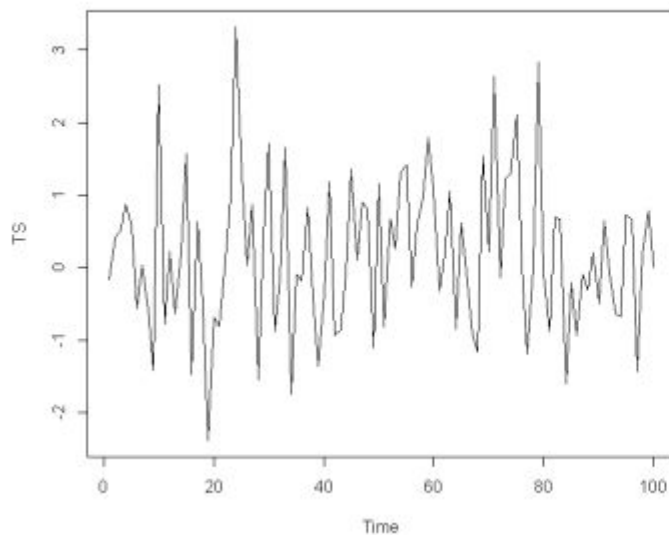
# Série Estacionária

- A média e a variância se mantêm constantes durante o tempo
- Em princípio, séries com tendência e sazonalidade não são estacionárias

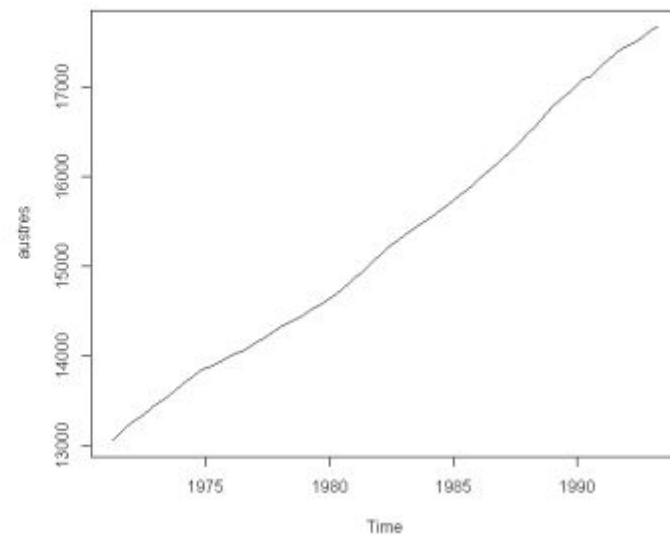


# Série Estacionária

Estacionária



Não estacionária



# Série Estacionária

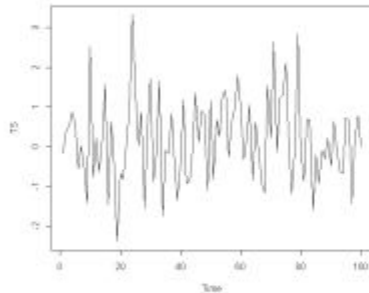
- Por que a estacionariedade é importante?
  - Existem técnicas analíticas de séries temporais que dependem da estacionariedade da série para funcionarem.
  - Se a série não é estacionária, pode-se aplicar algum tipo de transformação.



# Série Estacionária

- Como saber?
  - Visualmente
  - Testes Estatísticos (Dickey-Fuller)

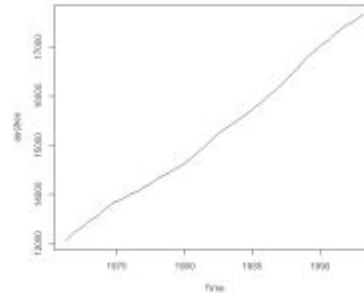
Estacionária



Augmented Dickey-Fuller Test

```
data: TS
Dickey-Fuller = -4.2082, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

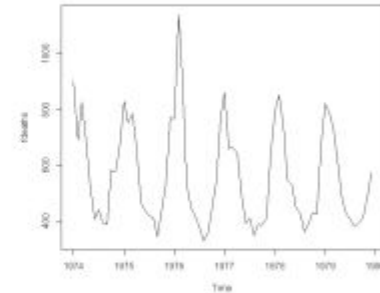
Não estacionária



Augmented Dickey-Fuller Test

```
data: austres
Dickey-Fuller = -2.5512, Lag order = 4, p-value = 0.3493
alternative hypothesis: stationary
```

Estacionária



Augmented Dickey-Fuller Test

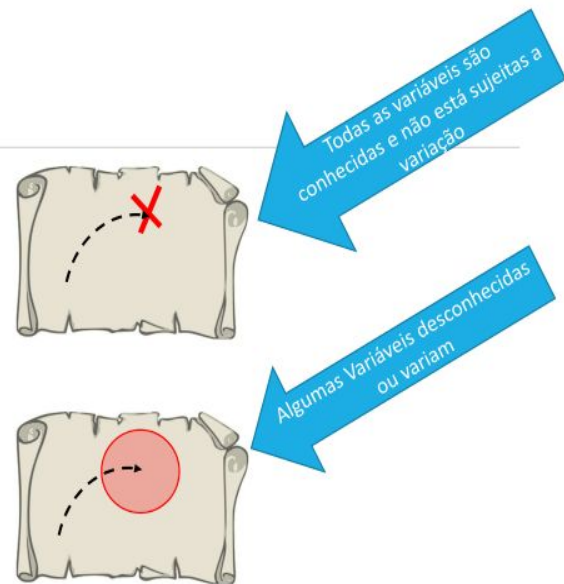
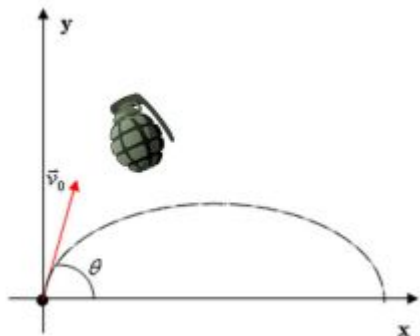
```
data: fdeaths
Dickey-Fuller = -6.7742, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```



# PREVISÕES

# Previsões

- Em estatística temos métodos eficientes para efetuar previsões de séries temporais
- Isso não significa que prever não esteja sujeito a erros
- Analogia da granada:
  - velocidade inicial, ângulo, resistência do ar, gravidade



# Previsões

Por que erros?



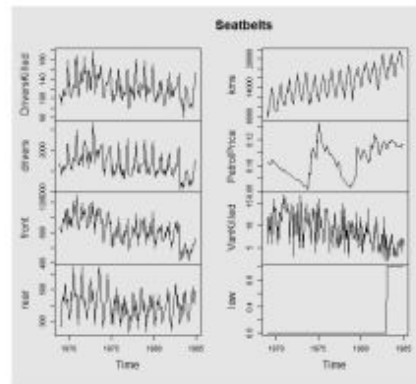
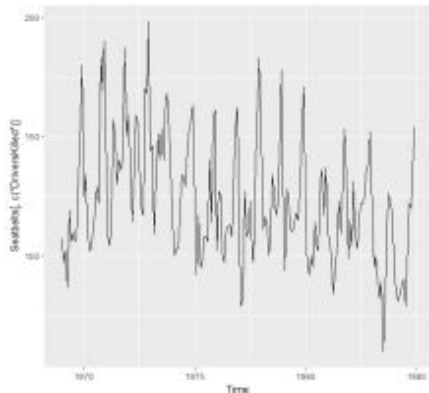
# Previsões

- Conceitualmente, toda previsão está errada
- O objetivo é minimizar o erro (torná-lo menor possível)
- Mas, o quão fácil é prever?



# Previsões

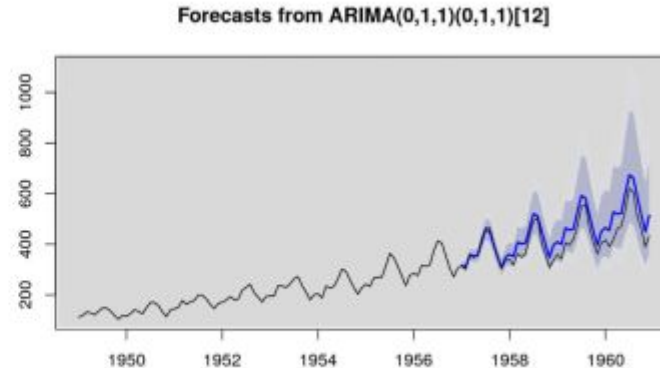
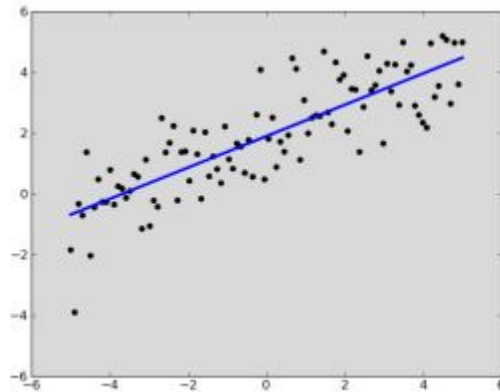
- Como prever usando séries temporais?
  - usando os próprio dados - séries temporais puras
  - usando outras variáveis - modelo explanatório
  - usando as duas técnicas - modelo misto



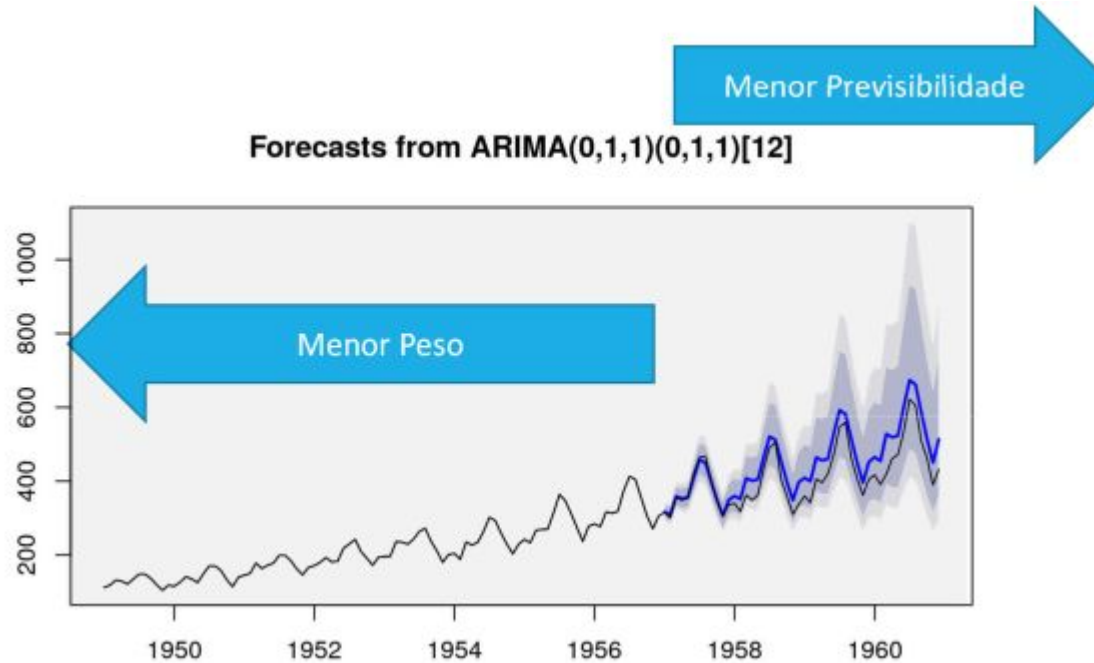
# Previsões

- Regressão vs Séries Temporais

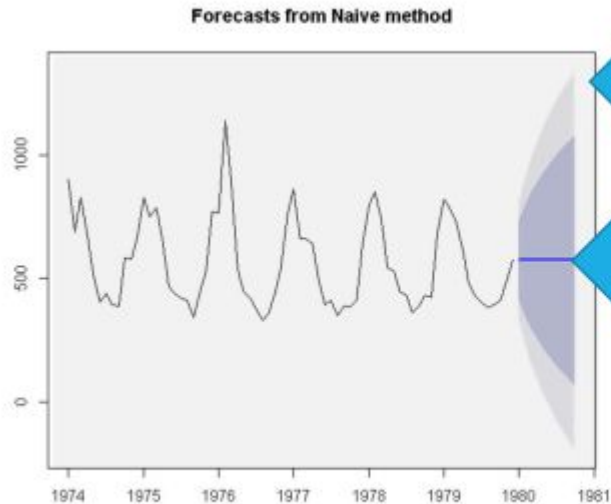
- Na regressão, em geral, a previsão pode significar extrapolação, mas também dados entre os intervalos já conhecidos
- Previsão em séries temporais univariadas significa extrapolação



# Previsões



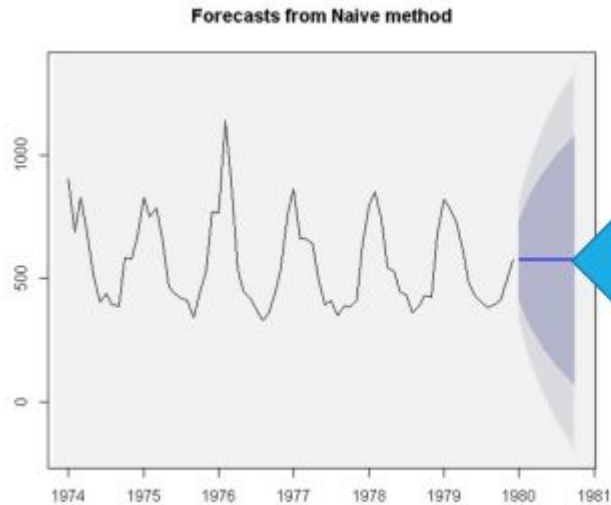
# Previsões



	Point Forecast	Lo 80	H1 80	Lo 95	H1 95
Jan 1980	574	412.71329	735.2867	327.33327	820.6667
Feb 1980	574	345.90615	802.0938	225.16056	922.8394
Mar 1980	574	294.64323	853.3568	146.76069	1001.2393
Apr 1980	574	251.42658	896.5734	80.66653	1067.3335
May 1980	574	213.35196	934.6480	22.43642	1125.5636
Jun 1980	574	178.92986	969.0701	-30.20763	1178.2076
Jul 1980	574	147.27548	1000.7245	-78.61883	1226.6188
Aug 1980	574	117.81230	1030.1877	-123.67888	1271.6789
Sep 1980	574	90.13988	1057.8601	-166.00020	1314.0002
Oct 1980	574	63.96665	1084.0334	-206.02870	1354.0287



# Previsões



Ocorreu fora do pontos de previsão

# Suavização Exponencial

# Suavização Exponencial

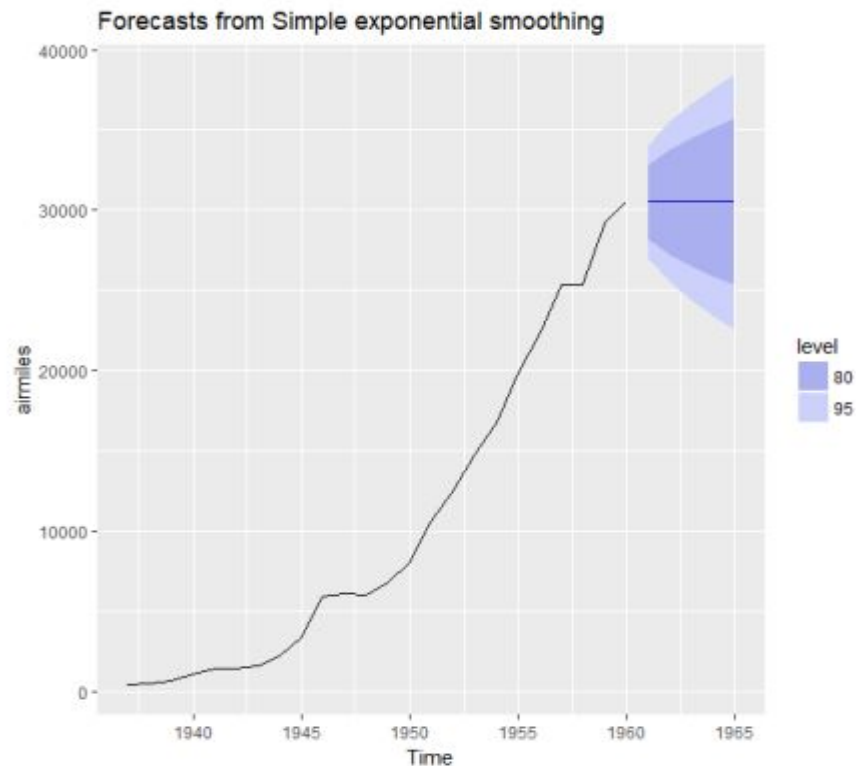
- Princípio básico
  - As observações passadas possuem pesos
  - Quanto mais recentes as observações, maiores seus pesos para as previsões
  - Utiliza médias que reduzem quanto mais distantes são as observações
  - O parâmetro  $\alpha$  determina o índice de redução: valor entre 0 e 1
    - Próximo de 0: observações antigas tem maior peso
    - Próximo de 1: observações recentes têm mais peso



# Suavização Exponencial

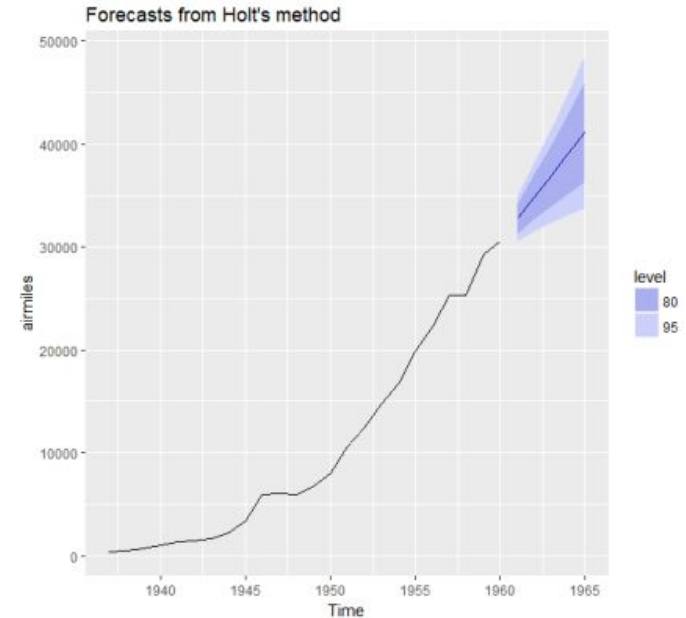
Intervalo	Peso
t-1	0,5
t-2	0,2
t-3	0,07
t-4	0,01
t-5	0,005
t-6	0,001
t-7	0,0004

# Suavização Exponencial Simples



# Tendência Linear de Hold

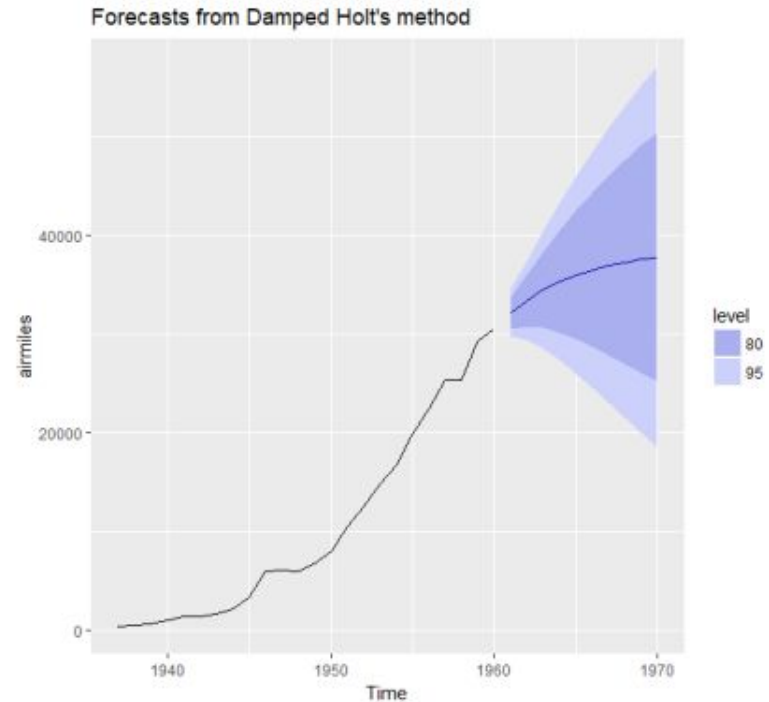
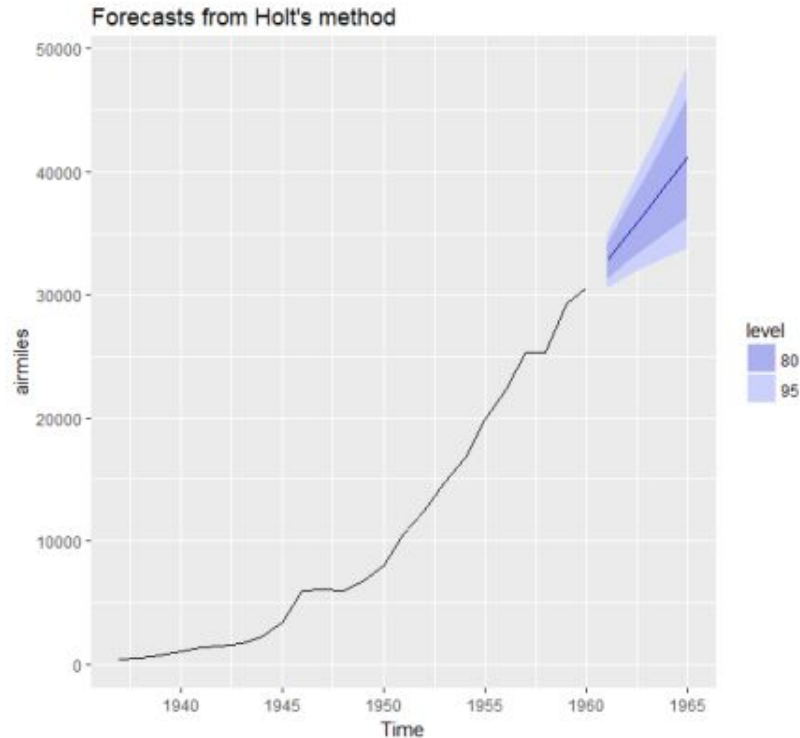
- Hold Linear Trend
- Suavização Exponencial para dados com tendência
- Gera uma tendência linear para o futuro
- Proposto por Hold(1957)



# Tendência Amortecida

- Damped
- Proposto por Gardner and McKenzie (1985)
- A tendência linear de Hold cresce ao infinito
- Na tendência amortecida, um novo parâmetro amortece a tendência conforme a previsão avança para o futuro
- Inclui um parâmetro de amortecimento:  $\Phi$
- $0 < \Phi < 1$

# Tendência Amortecida



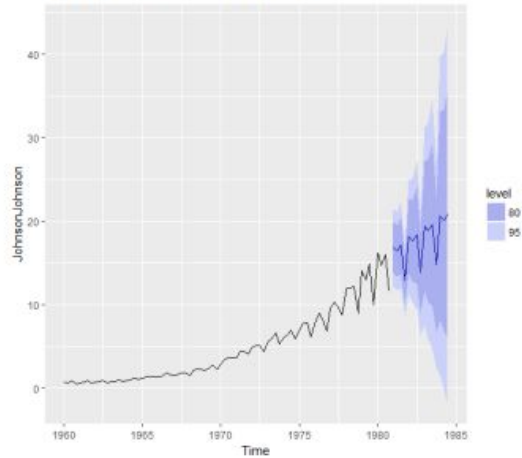


# Holt-Winters Sazonal

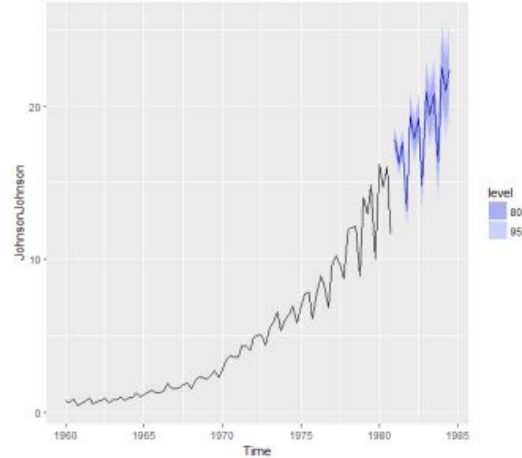
- Holt(1957) e Winters(1960)
- Inclui captura de sazonalidade
- Também ideal para tendência
- Aditivo: para variação sazonal constante
- Multiplicativo: variação sazonal varia na série

# Holt-Winters Sazonal

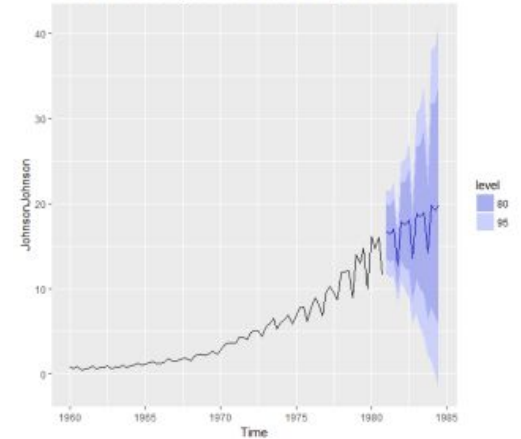
Forecasts from Holt-Winters' multiplicative method



Forecasts from Holt-Winters' additive method



Forecasts from Damped Holt-Winters' multiplicative method



# ETS (Error, trend, seasonal)

Erro	Tendência	Sazonalidade
A	N	N
M	A	A
Z	M	M
	Z	Z

A	Aditivo
M	Multiplicativo
N	Nenhum
Z	Automático

Tendência	
A <sub>a</sub>	Aditivo Amortecido
M <sub>a</sub>	Multiplicativo Amortecido

model="AAA"

# ARIMA

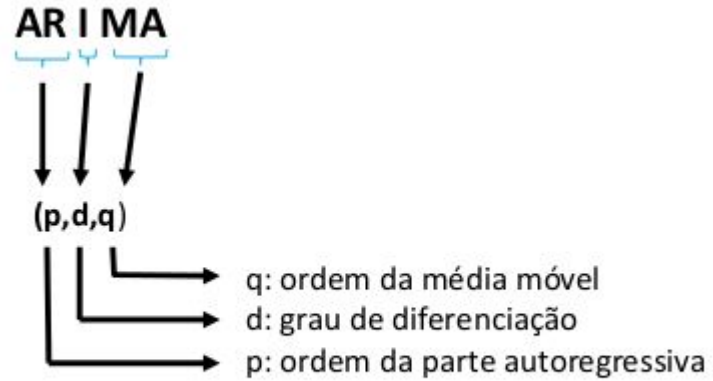
# ARIMA

- Robusto: Pode ser usado em praticamente qualquer tipo de ST
- Dados estáveis, com poucos outliers
- **Requer dados estacionários:** pode ser transformada usando diferenciação: remove tendências
- Subtrai a observação do período atual do período anterior
- A diferenciação pode ser feita 1x: diferenciação de primeira ordem
- Ou pode ser necessário uma segunda vez: diferenciação de segunda ordem (mais raro)

# ARIMA

- AR – Autoregressivo: avalia a relação entre os períodos (lags): autocorrelação. Extrai essa influência
- I – Integrated: Aplica a diferenciação, se necessária
- MA – Moving Average: avalia erros entre períodos e extrai estes erros (não tem relação com MA usados para suavização de st)

# ARIMA



# ARIMA

- $p = 1$ , significa que uma determinada observação pode ser explicada pela observação prévia + erro
- $p = 2$ , significa que uma determinada observação pode ser explicada por duas observações prévias + erro
- $d = 0$ , significa que não é aplicada diferenciação
- $d = 1$ , significa que será aplicada diferenciação de primeira ordem
- $d = 2$ , significa que será aplicada diferenciação de segunda ordem
- $q = 1$ , significa que uma determinada observação pode ser explicada pelo erro da observação prévia
- $q = 2$ , significa que uma determinada observação pode ser explicada pelo erro de duas observações prévias

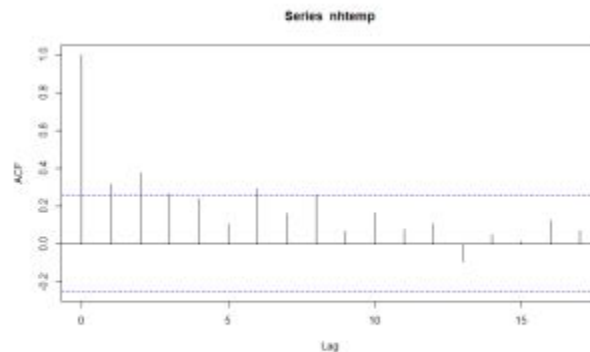
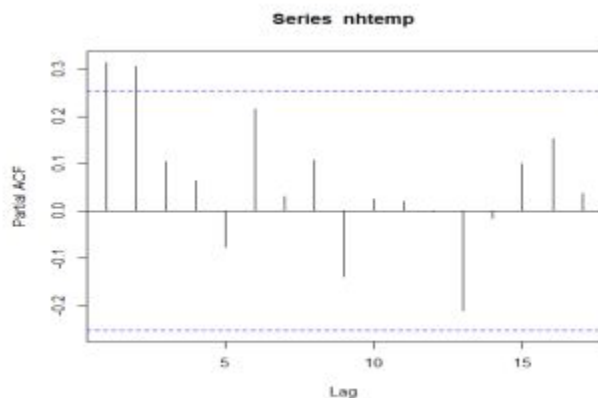


# ARIMA

- $AR(1)$  ou  $ARIMA(1,0,0)$  – Apenas elemento autoregressivo , de 1º ordem
- $AR(2)$  OU  $ARIMA(2,0,0)$  ) – Apenas elemento autoregressivo, de 2º ordem
- $MA(1)$  OU  $ARIMA(0,0,1)$  – Apenas média móvel
- $ARMA(1,1)$  ou  $Arima(1,0,1)$  – Autoregressão e média móvel de 1º ordem

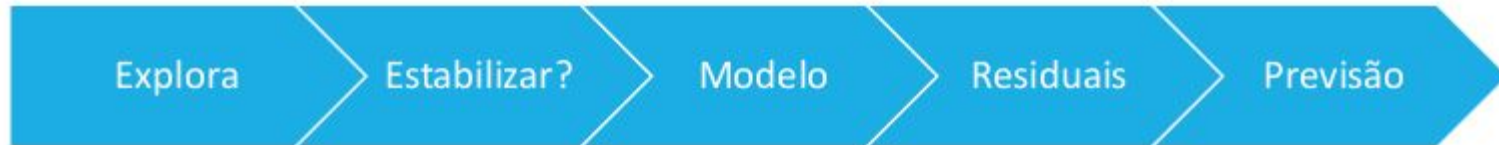
# ARIMA

- Como definir valores de  $p, d$  e  $q$ ?  
p: ordem da parte autoregressiva - PACF  
d: grau de diferenciação – Teste de Estacionariedade  
q: ordem da média móvel - ACF



# ARIMA

- Como saber qual o melhor modelo?
  - Akaike Information Criteria (AIC)
  - Bayesian Information Criteria (BIC)



# Referências Bibliográficas

- Introdução à Estatística - Mário F. Triola
- Noções de Probabilidade e Estatística - Marco Nascimento Magalhães
- Modelos de Regressão em R - Écio Souza Diniz
- Análise de Séries Temporais - Pedro A. Moretin / Clélia M. C. Toloí
- Análise e Previsões de Séries Temporais: Os modelos ARIMA - Reinaldo Castro Souza / Maria Emília Camargo