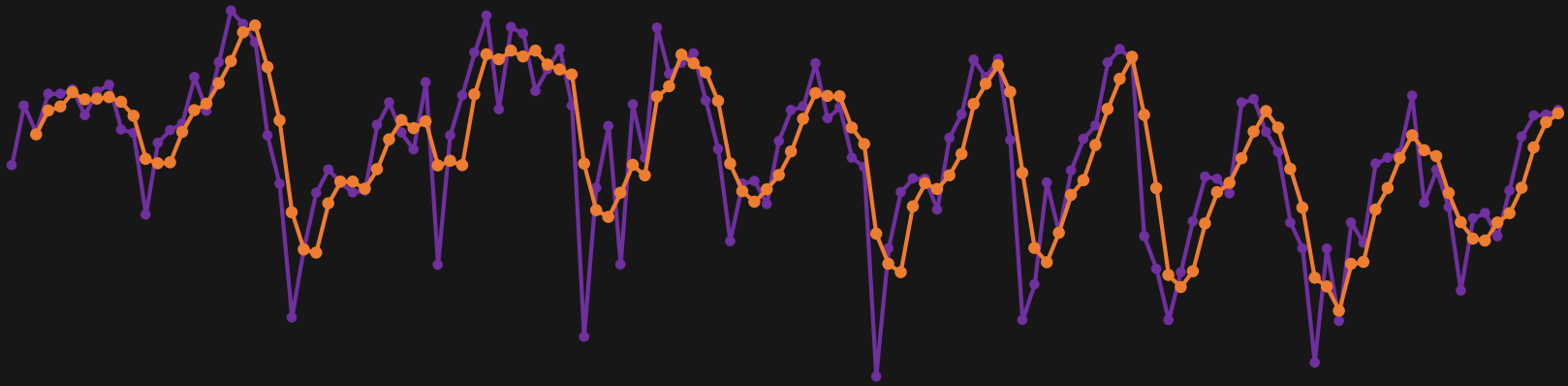


# Modelagem de Séries Temporais em R



Prof. Eng. Ícaro Agostino

Engenheiro de Produção - Ceuma

Mestrando em Engenharia de Produção - UFSM

## 1. Linguagem R + RStudio

- Utilizando linguagem R
- O ambiente do RStudio
- pacotes
- importando dados

## 2. O que são series temporais?

- Conceitos básicos
- Componentes
- Séries temporais no mundo real
- Realizar previsões?
- Abordagens e métodos
- O processo de previsão

## 3. Análise e Modelagem

- Modelos quantitativos
- Holt-Winters
- ARIMA
- Redes Neurais Artificiais
- Medidas de avaliação

## 4. Ajustando modelos com dados reais

# Linguagem R + RStudio

# Utilizando linguagem R

---

- Desenvolvida por **cientistas**
- Foco em análise de dados
- Linguagem livre
- Altamente documentada\*
- Estatística



# O ambiente do RStudio

---

- Ambiente de desenvolvimento integrado
- Gerenciador de arquivos
- Visualizador de variáveis
- Leitor de script (Editor) e Console
- Visualizador de gráficos
- Gratuito



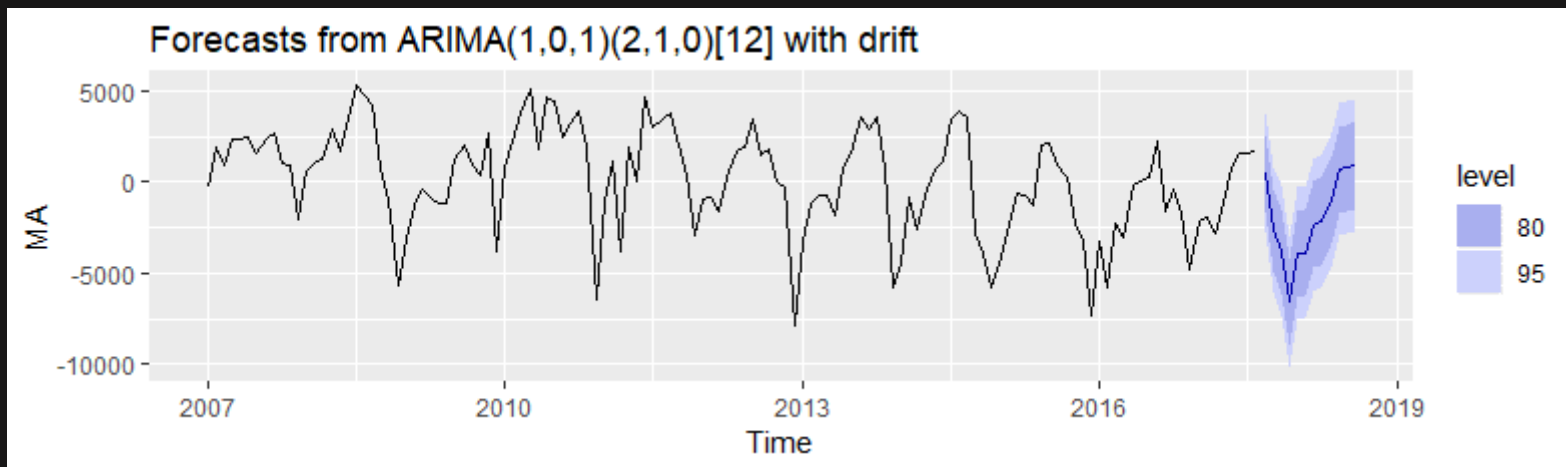


O que são series  
temporais?

# Conceitos básicos

## O que são series temporais?

Séries temporais são sequências de observações sobre uma variável em **diferentes instantes de tempo**, em que os dados são observados em momentos **discretos**, usualmente **equidistantes**.





# Componentes

---

As séries temporais são compostas por quatro elementos:

**Tendência:** verifica o sentido de deslocamento da série temporal ao longo do tempo

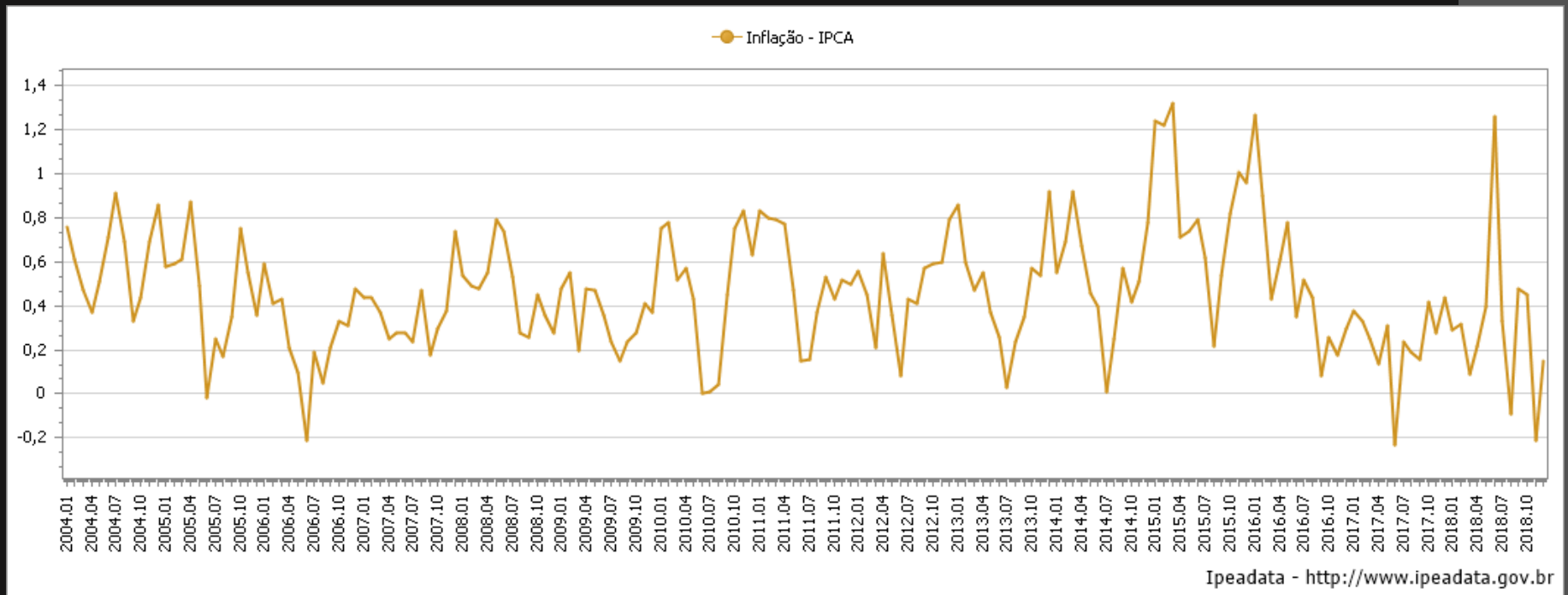
**Ciclo:** movimento ondulatório da série temporal, que ao longo de vários anos tende a ser periódico

**Sazonalidade:** movimento ondulatório de curta duração, normalmente inferior a um ano, que, em geral, está associado a mudanças climáticas

**Ruído aleatório ou erro:** compreende a variabilidade intrínseca à série temporal, não podendo ser modelado.

# Séries temporais reais

## Inflação Mensal - Índice Nacional de Preços ao Consumidor Amplo (IPCA)



Fonte: Ipea data

# Séries temporais reais

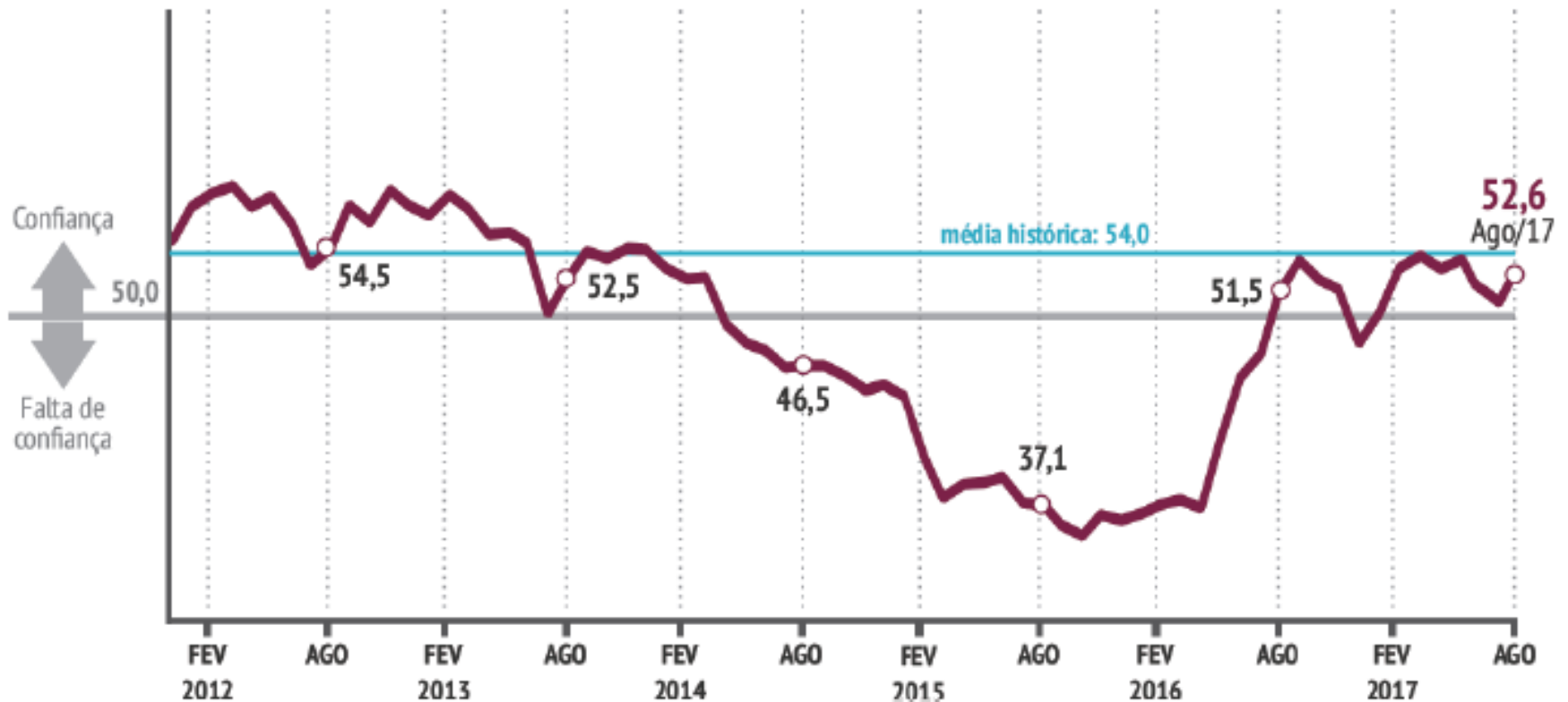
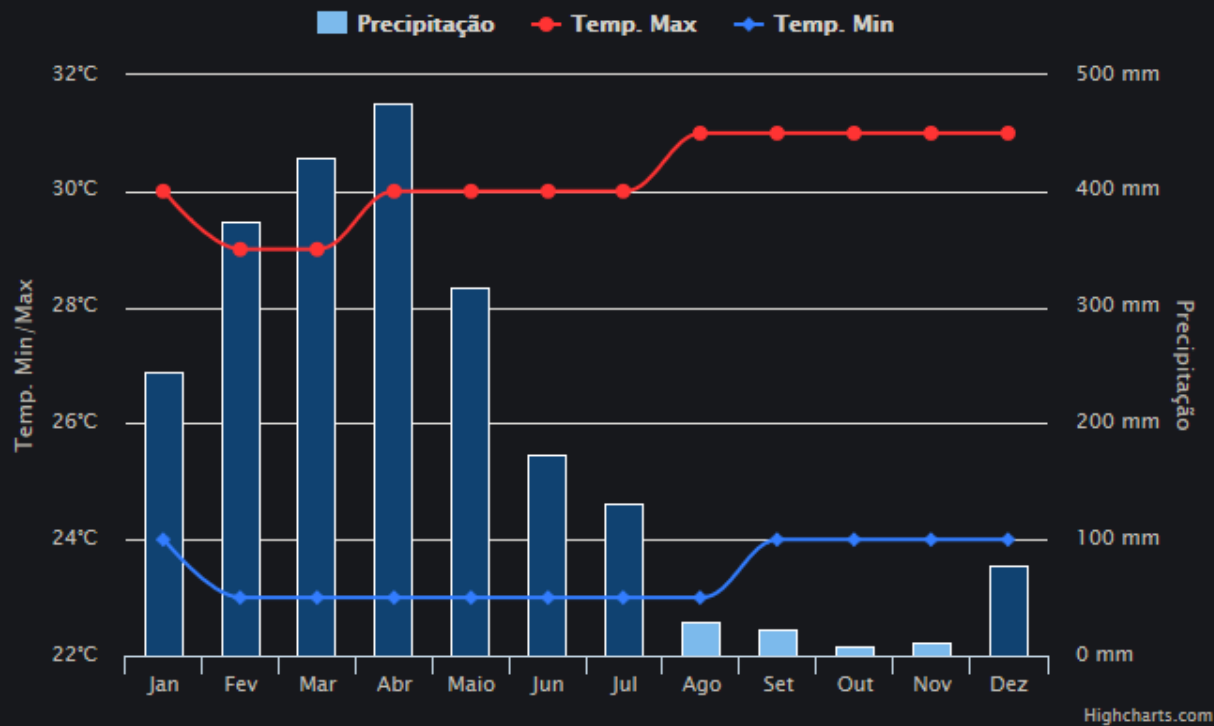


Figura 1 – Índice de Confiança do Empresário Industrial

Fonte: Adaptado de CNI (2017)

# Séries temporais reais

Temperaturas e índice pluviométrico médio mensal do Maranhão.



Fonte: Climatempo

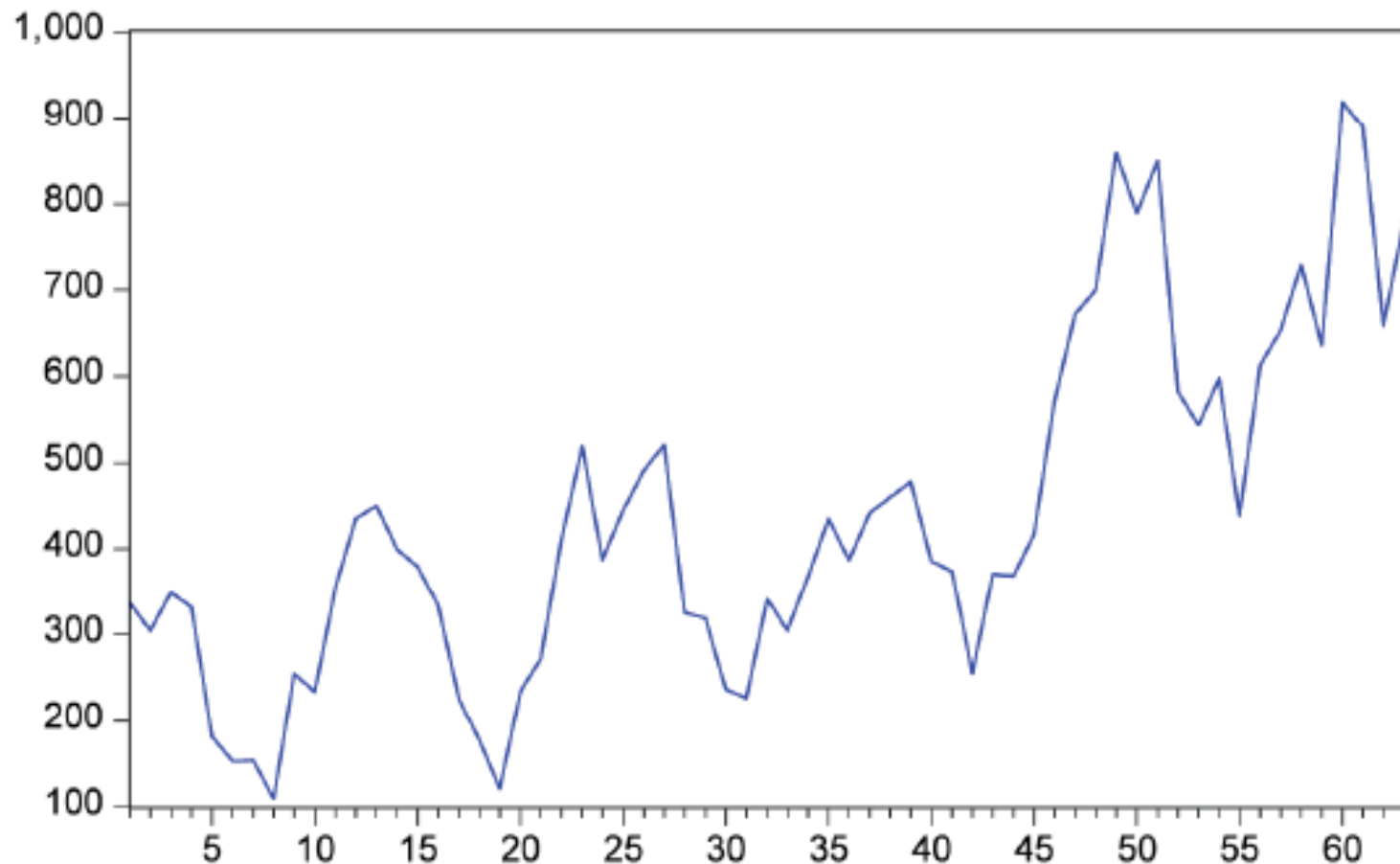
# Séries temporais reais

Veja o desmatamento da amazônia ao longo da história (em km<sup>2</sup>)



Fonte: Sistema Prodes/Inpe

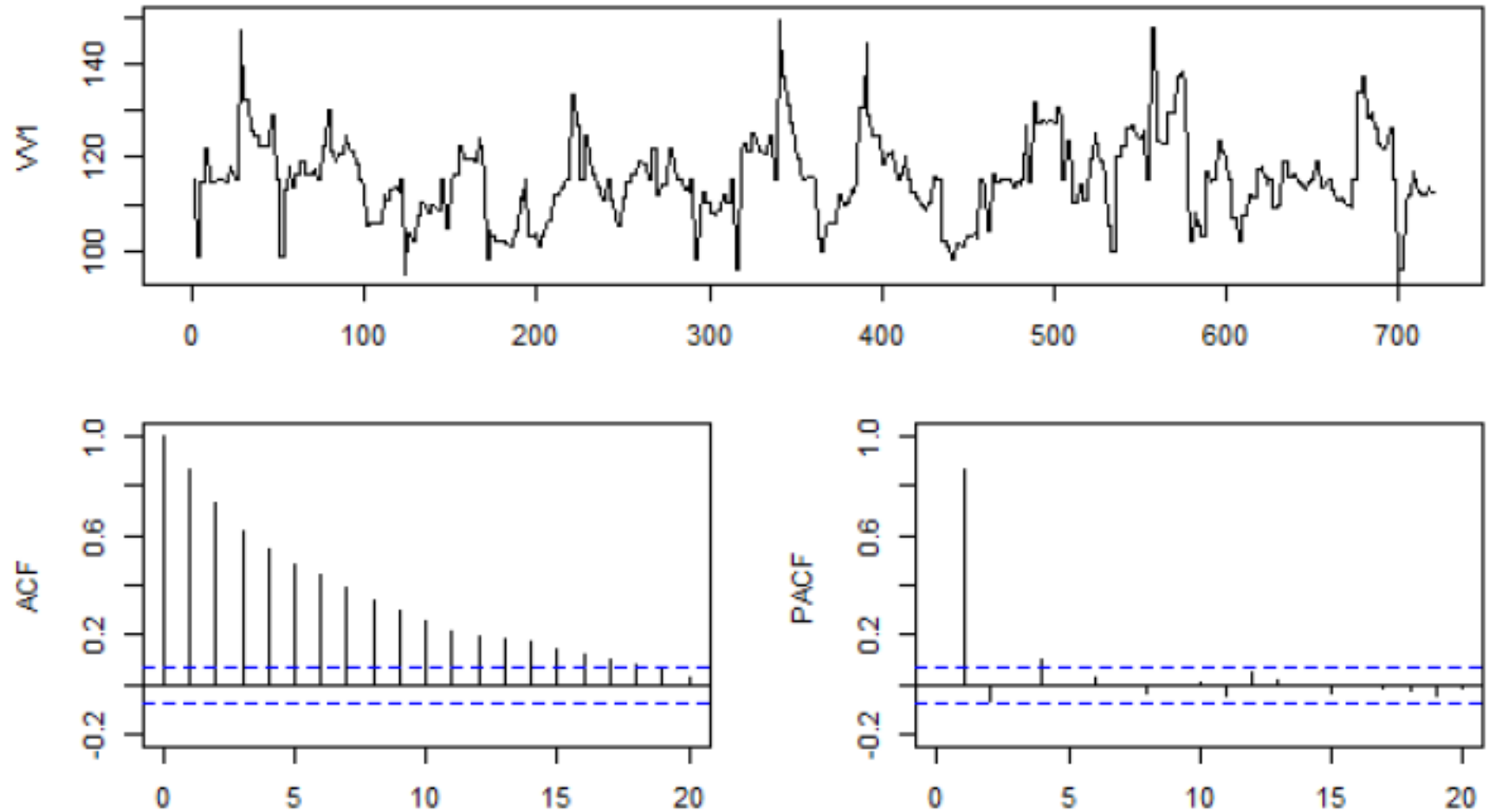
# Séries temporais reais



**Figura 2: Demanda por rações de peixe, no período de jan/2010 a mar/2015**

Fonte: Os autores.

# Séries temporels reals







# Realizar previsões?

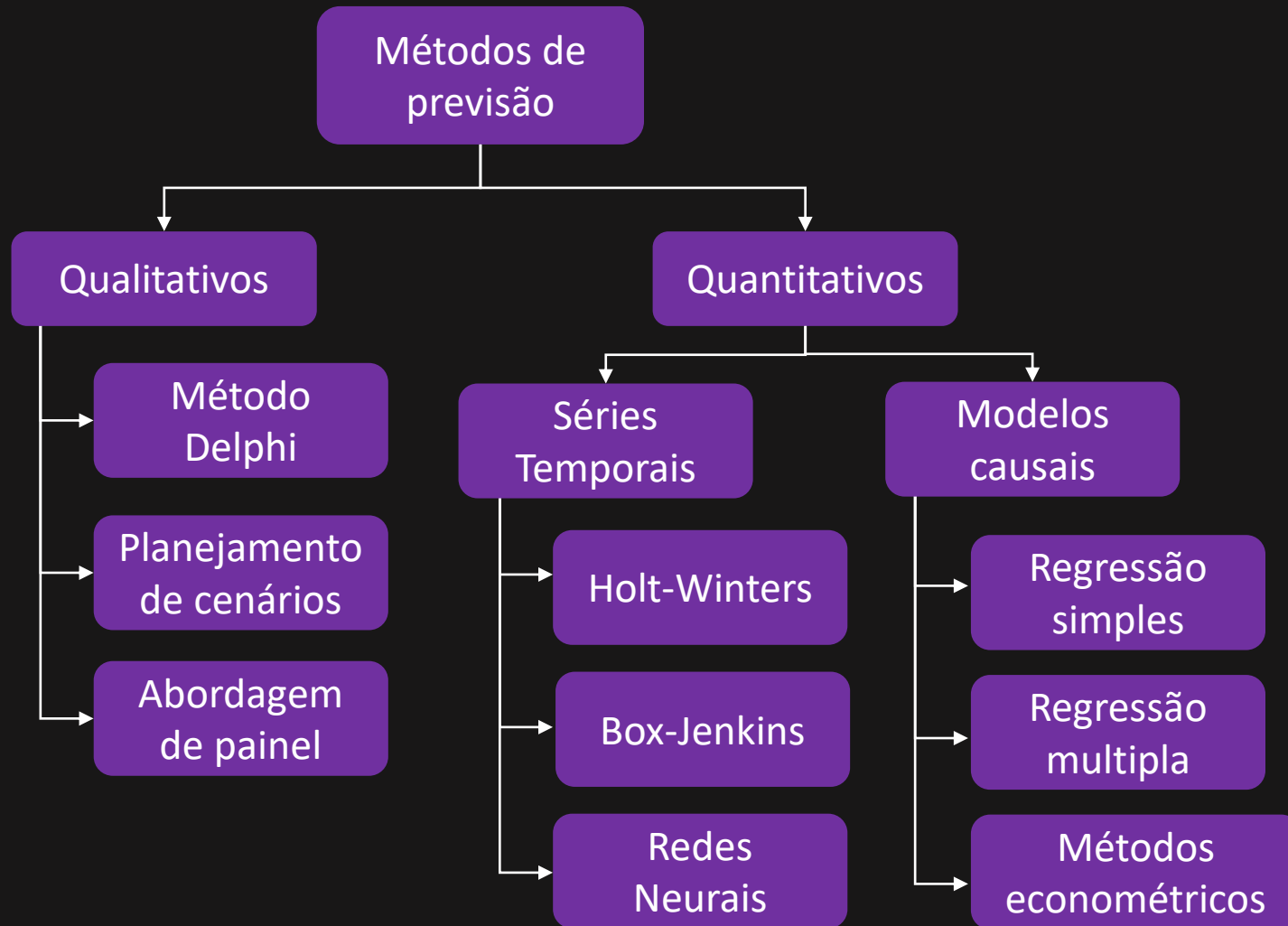
---

Realizar previsões é uma atividade **indispensável** no planejamento de atividades e nos processos de tomada decisão.

A previsão pode ser entendida como a busca de informações sobre o **comportamento futuro** de uma variável através de um **processo racional** envolvendo incerteza.

Previsão consiste em um **processo metodológico** que objetiva definir os dados futuros com base em **modelos**.

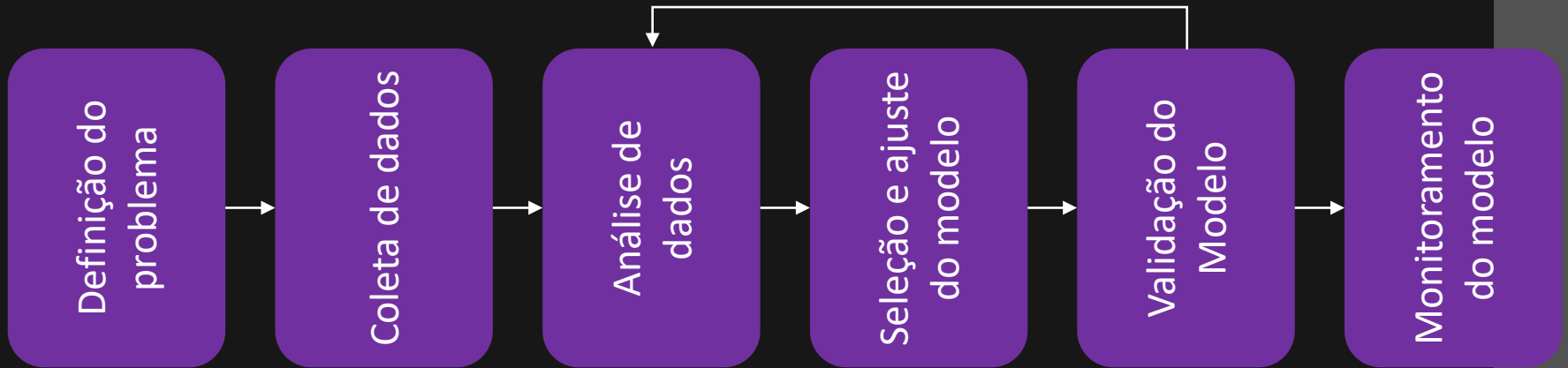
# Abordagens e métodos



# O processo de previsão

---

O processo de realização de previsões é composto de etapas estruturadas




# Análise e Modelagem

# Modelos quantitativos

---

As técnicas quantitativas utilizam **dados históricos** para a realização de previsões.

Modelos de séries temporais utilizam o comportamento dos dados ao longo do tempo para **extrair um padrão** que possa **explicar o comportamento** futuro da variável estudada.

- Holt-Winters
- ARIMA 
- Redes Neurais Artificiais

# ARIMA

---

- A metodologia de Box-Jenkins (1970) é uma abordagem difundida para a realização de previsão em séries temporais. A ênfase está na análise das propriedades da própria série temporal, capturando a correlação entre os valores da série ao longo do tempo. (Gujarati; Porter, 2011).
- Três filtros: o componente Autorregressivo (AR), o filtro de Integração (I) e o componente de Médias Móveis (MA).

# ARIMA

---

Em aplicações reais, raramente as séries temporais são estacionárias,, sendo necessário a aplicação de diferenças. ARIMA (p,d,q):

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Onde:  $w_t = \Delta^d Z_t$ ;  $\phi_p$  é o parâmetro autorregressivo de ordem “ $p$ ”;  $\theta_q$  é o parâmetro de médias móveis de ordem “ $q$ ”;  $\varepsilon_t$  representa o erro ou ruído aleatório  $\sim (0, \sigma^2)$ .

# Redes Neurais Artificiais

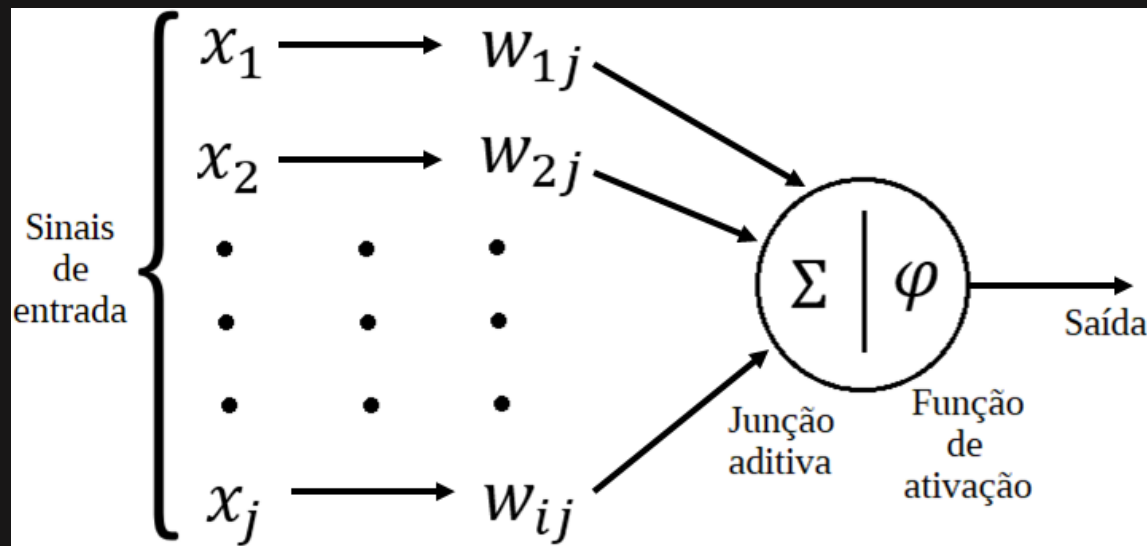
---

- Modelos de Redes Neurais possuem diversas aplicações, com capacidade de reconhecer e aprender padrões complexos e não-lineares (Haykin, 2001)
- Foi desenvolvida inspirada no processo de informação do cérebro humano, por meio de algoritmos de treinamentos, sendo também uma abordagem bastante difundida para realização de previsões em séries temporais (Montgomery; Jennings; Kulahci, 2007)



# Redes Neurais Artificiais

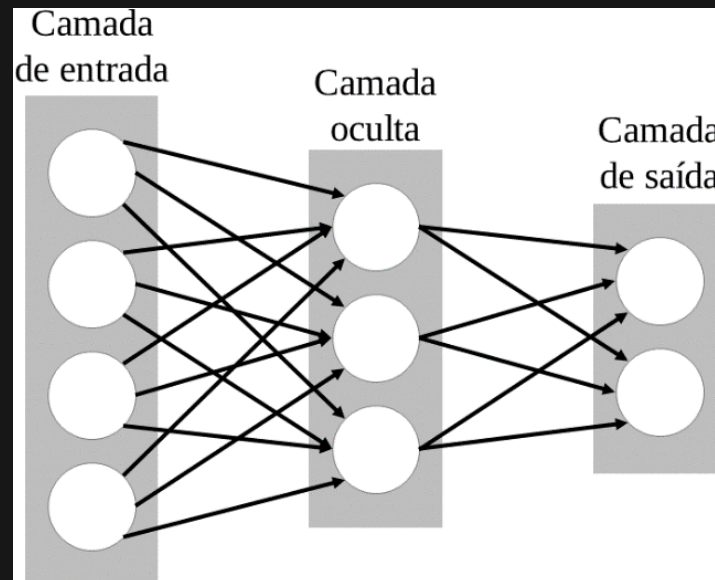
- A estrutura básica de um neurônio artificial compõe quatro componentes: conexões, junção aditiva, função de ativação e o viés ou bias (Martins; Werner, 2014):



# Redes Neurais Artificiais

---

- Os modelos de RNA são divididos em classes de acordo com o formato de funcionamento, as redes alimentadas por múltiplas camadas são consideradas as mais utilizadas, tendo superior capacidade preditiva em modelos temporais (Qi; Zhang, 2001).



# Redes Neurais Artificiais

---

- A realização de previsões com modelos univariados, utiliza os períodos defasados da própria série na camada de entrada, podendo ser considerado um modelo genérico autorregressivo não-linear (Zhang; Patuwo; Hu, 2001; Jacobs, 2014):

$$\hat{Z}_t = f(Z_{t-1}, Z_{t-2}, \dots, Z_{t-p})$$

- Onde:  $Z_t$  corresponde ao valor da série no período “t”; “p” o número de camadas de entrada na RNA; “f” a função linear ou não-linear.

# Medidas de avaliação

Sigla	Descrição	Equação
MAE	Erro Médio Absoluto	$\frac{\sum_{t=1}^n  \varepsilon_t }{n}$
MAPE	Erro Percentual Médio Absoluto	$\frac{\sum_{t=1}^n  \varepsilon_{t_t}/Z_t  * 100}{n}$
RMSE	Raiz do Erro Quadrático Médio	$\frac{\sqrt{\sum_{t=1}^n (\varepsilon_t)^2}}{n}$
U-Theil	Coeficiente U de Theil	$\frac{\sqrt{\sum_{t=1}^n (\varepsilon_t)^2}}{\sqrt{\sum_{t=1}^n (Z_t - Z_{t-1})^2}}$

# Critérios de seleção

---

De acordo com Box e Jenkins (1976), a com a inclusão de muitos parâmetros no modelo ajustado pode ser prejudicial, pois o mesmo deve ser o mais parcimonioso possível.

Para decidir sobre a inclusão ou não de parâmetros, os critérios Akaike Information Criteria – AIC (AKAIKE, 1973) e Bayesian Information Criteria – BIC (SCHWARZ, 1978) são utilizados.

Tais critérios são considerados penalizadores visto que levam em consideração o número de parâmetros dos modelos e a variância dos erros gerados, dessa forma o modelo que apresentar menores valores para os critérios e AIC e BIC terá o melhor ajuste (MORETTIN, 2008).

# Critérios de seleção

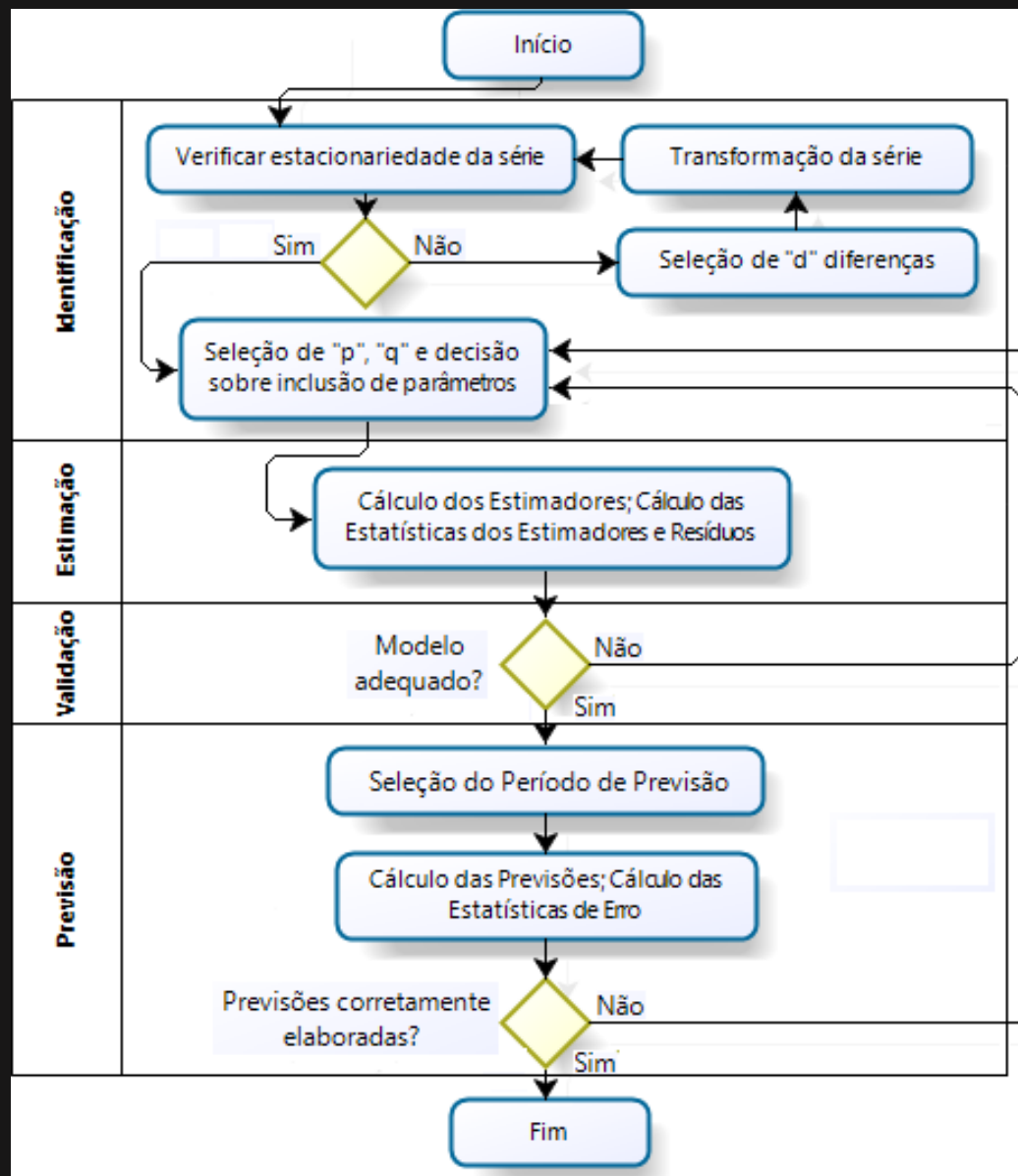
---

$$AIC(p, q) = \ln \sigma_{p.q}^2 + \frac{2(p + q)}{n}$$

$$BIC(p, q) = \ln \sigma_{p.q}^2 + (p + q) \frac{\ln(n)}{n}$$

Onde:  $p$  e  $q$  são os parâmetros conhecidos,  $n$  é o tamanho da amostra,  $\ln$  o logaritmo neperiano e  $\sigma^2$  a variância estimada dos erros.

# Box & Jenkins na prática



# Identificação

**Padrões teóricos de FAC e FACP para modelos ARMA**

	<b>FAC</b>	<b>FACP</b>
AR(1)	Barras decaem exponencialmente barras positivas se $\phi_1 > 0$ barras de sinais alternados se $\phi_1 < 0$	Apenas uma barra, em $k=1$ positiva se $\phi_1 > 0$ negativa se $\phi_1 < 0$
AR( $p$ )	Barras decaem exponencialmente ou em forma de senóide amortecida	Há $p$ barras, de $k=1$ a $k=p$
MA(1)	Apenas uma barra, em $k=1$ positiva se $\theta_1 < 0$ negativa se $\theta_1 > 0$	Barras decaem exponencialmente barras negativas se $\theta_1 > 0$ barras de sinais alternados se $\theta_1 < 0$
MA( $q$ )	Há $q$ barras, de $k=1$ a $k=q$	Barras decaem exponencialmente ou em forma de senóide amortecida
ARMA( $p, q$ )	Sequência infinita de barras, dominada por exponenciais ou senóides amortecidas, para $k > p - q$	Sequência infinita de barras, dominada por exponenciais ou senóides amortecidas, para $k > p - q$





# Links úteis

---

<https://github.com/icaroagostino>

<https://www.r-project.org/>

<https://www.rstudio.com/>

<https://otexts.com/fpp2/>

Ajustando modelos  
com dados reais

