

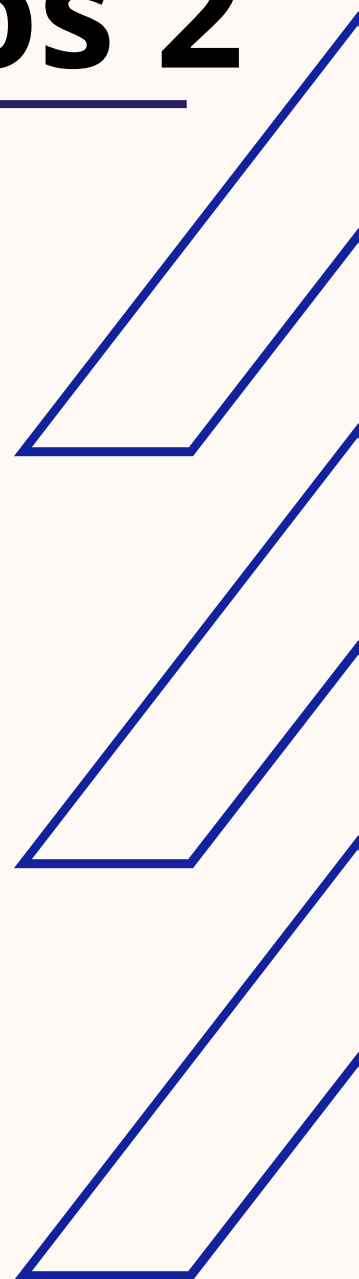
# **Trabalho pratico**

# **Analise de dados 2**

---

---

---



# Sumário

**Contextualização** Página - 03

**Autores** Página - 04

**Visualizando a serie temporal** Página - 05

**Características da serie** Página - 06

**Decompondo a serie** Página - 08

**Previsao** Página - 11

**Calculando erros** Página - 13

**Residuos** Página - 14

**Graficos relevantes** Página - 15

## Proposta:

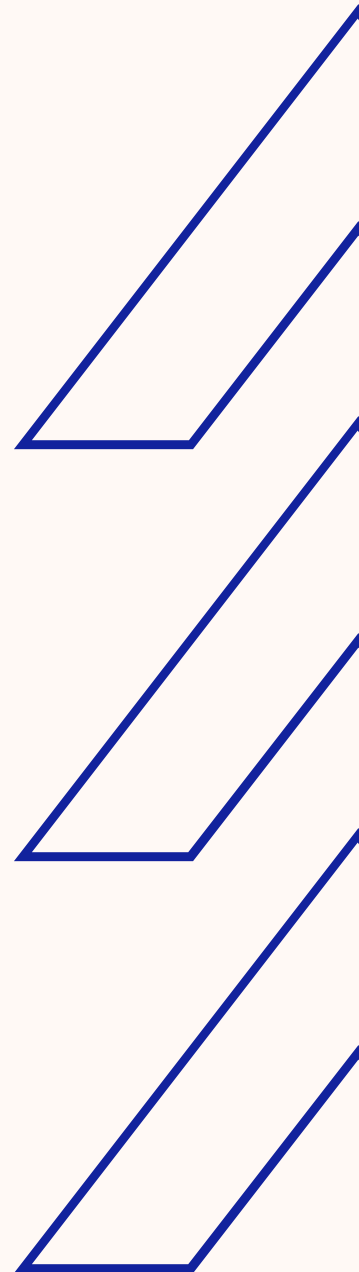
A essência desse trabalho é selecionar, analisar e interpretar um conjunto de dados utilizando conceitos teóricos e práticos lecionados na classe de análise de dados (2023 - 2024) do Instituto politécnico de Bragança (IPB).

## Metodologia:

A manipulação dos dados presentes neste trabalho como: gráficos, estatísticas e textos interpretativos. São da autoria de: Dennis de Sousa Farias e Pedro Eunísio Vieira de Souza. E seguem os parâmetros ensinados em sala aplicados ao software "R".

## Critério de escolha:

Escolhemos a amostra de dados que preenchesse os requisitos da proposta e que a afinidade fosse maior com nossos conhecimentos.



## Docente:



Maria Prudência Gonçalves Martins

## Alunos:



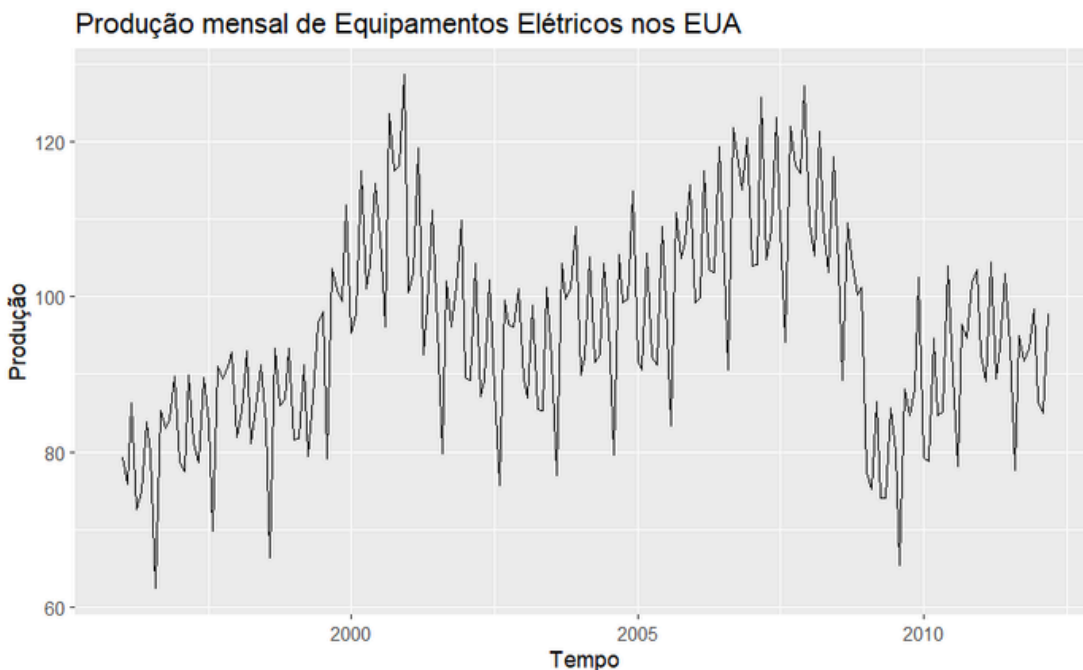
Dennis de Sousa Farias



Pedro Eunísio Vieira de Souza

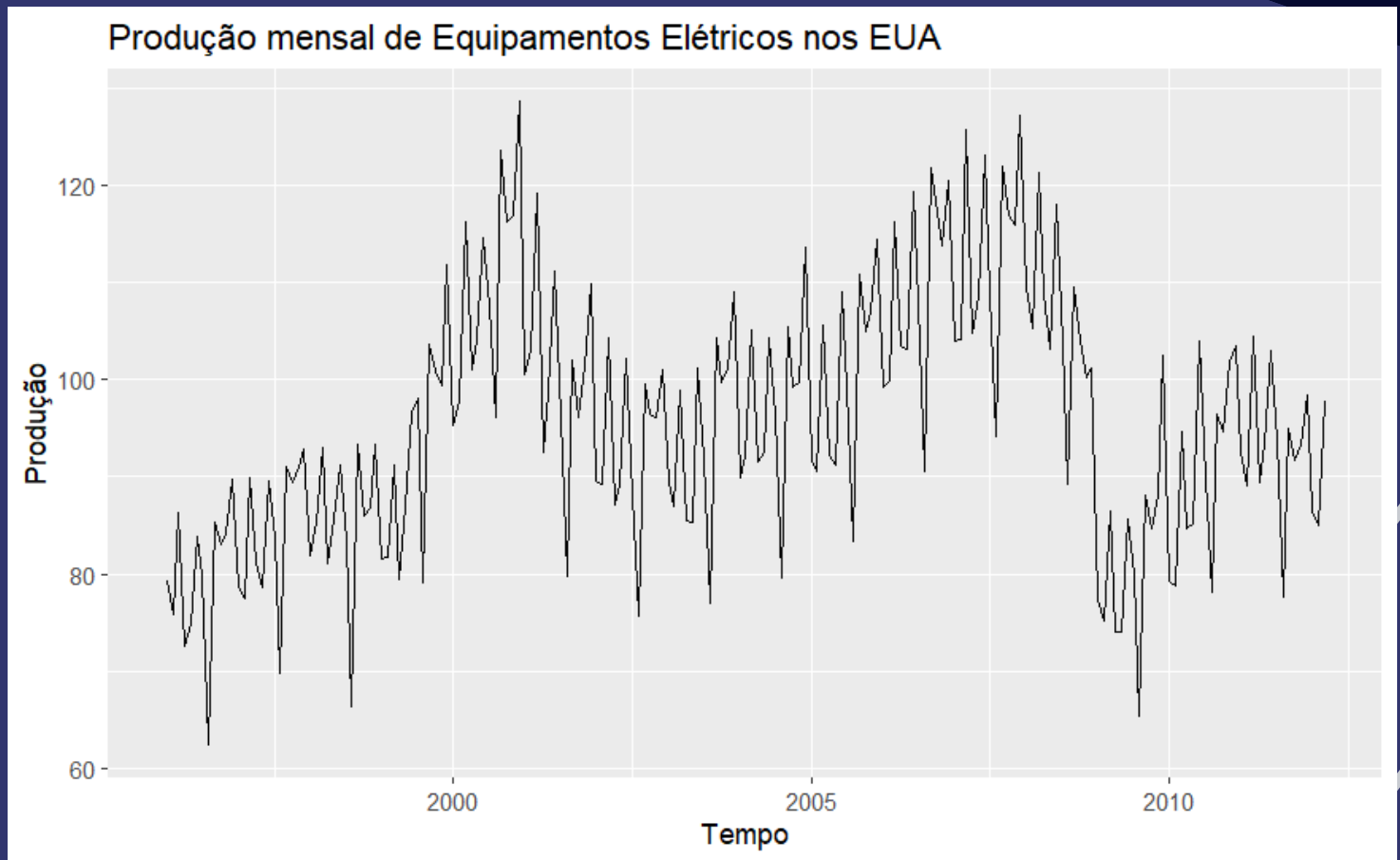
# Visualizando serie temporal

```
#Carregando a série temporal 'elecequip'.  
#Esta série temporal representa a produção mensal de  
equipamentos elétricos nos EUA de 1996 a 2012.  
  
data(elecequip)  
  
#Visualizando a série temporal  
autoplot(elecequip, main="Produção mensal de Equipamentos  
Elétricos nos EUA", ylab="Produção", xlab="Tempo")
```



**Para este trabalho escolhemos a serie temporal “elecequip” que apresenta a produção mensal de equipamentos eletricos nos Estados Unidos da America. Período (1996 -> 2012)**

## Características da serie :



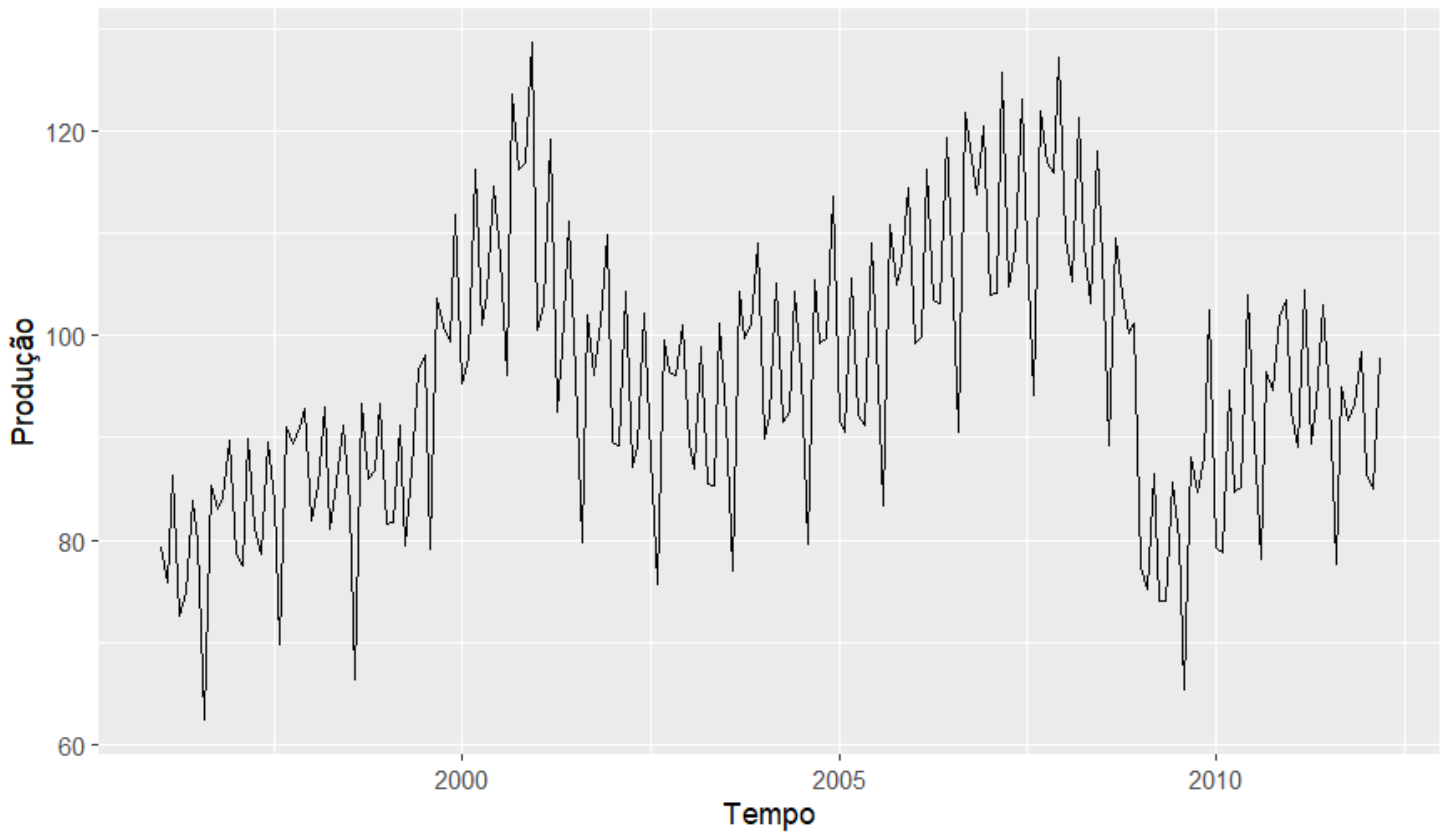
### Tendência:

-> Tendência Crescente: A série mostra um aumento consistente na produção de equipamentos elétricos ao longo do período, refletindo o crescimento industrial e econômico dos EUA durante esse tempo.

### Sazonalidade:

-> Sazonalidade Multiplicativa: Padrões sazonais evidentes indicam meses específicos com variações regulares na produção, possivelmente devido a fatores como demandas de mercado ou ciclos de manutenção.

Produção mensal de Equipamentos Elétricos nos EUA



#### Ciclo:

-> Durante os anos 2000 é possível visualizar um ciclo de aumento na produção de equipamentos elétricos, isso pode ser devido ao “boom” tecnológico dos anos 2000 e a demanda por novos equipamentos.

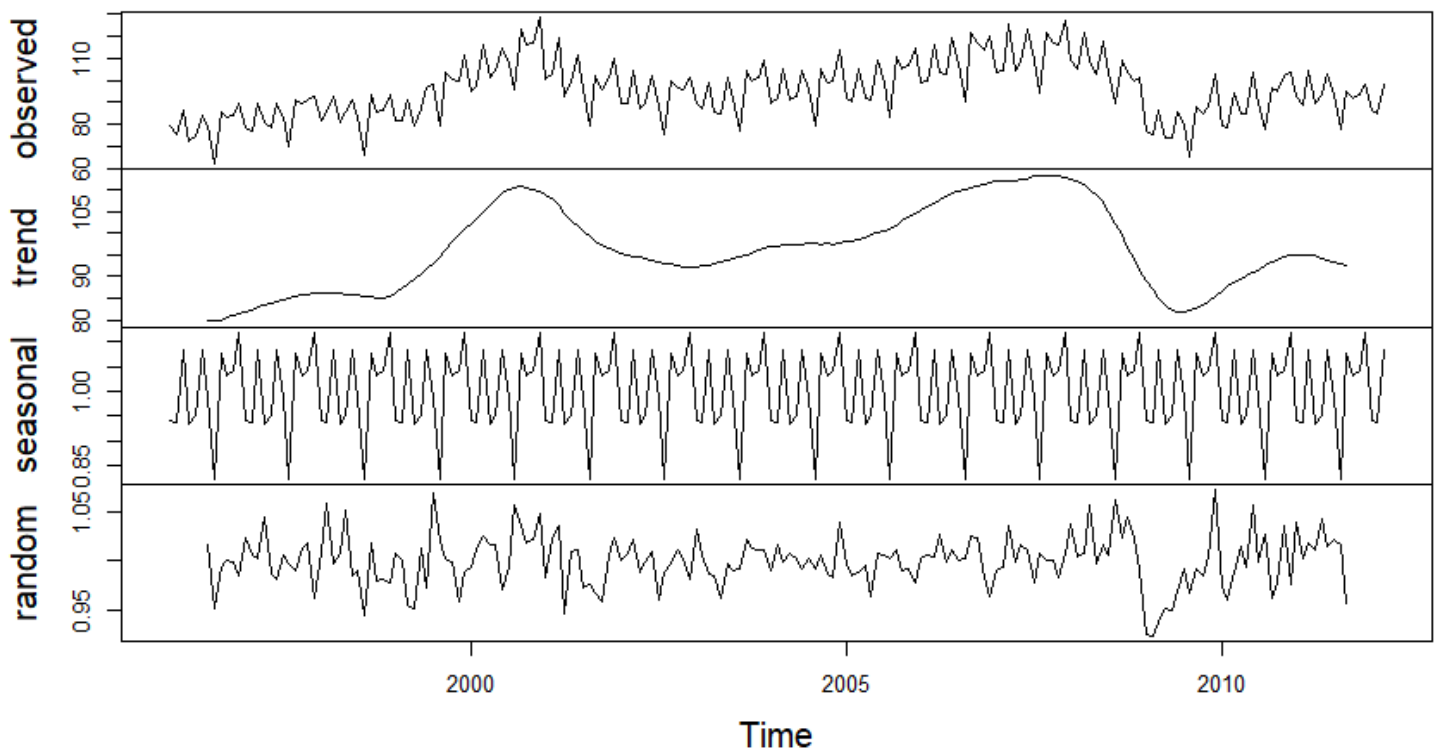
-> Por volta de 2008 é observado um ciclo de queda significativa na produção como resultado da crise financeira global, seguida por um período de recuperação econômica nos anos subsequentes.

# Decompondo a serie

```
#Decompondo a série:
```

```
serie_decomposta = decompose(elecequip, 'multiplicative')  
plot(serie_decomposta)
```

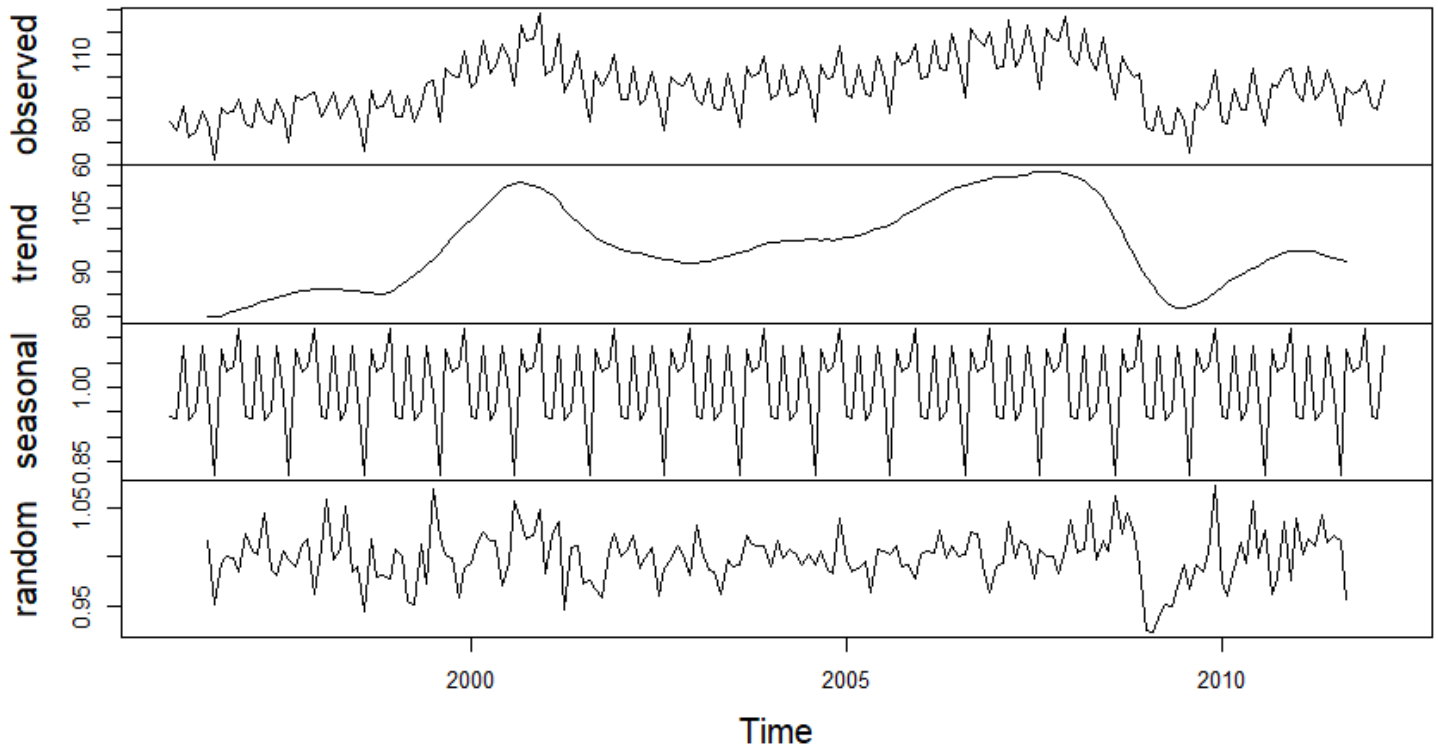
Decomposition of multiplicative time series



**Tendência:**  
Identificamos uma tendência geral crescente na produção de equipamentos elétricos ao longo do período de 1996 a 2012, refletindo um aumento constante na produção.



## Decomposition of multiplicative time series



### Sazonalidade:

-> Observamos padrões sazonais claros na série, com picos e vales que se repetem a cada ano, indicando flutuações regulares na produção relacionadas a fatores sazonais.

### Resíduos:

-> Após remover a tendência e a sazonalidade, os resíduos ainda mostram variações estacionárias, indicando a presença de flutuações não explicadas por esses componentes.

## Decompondo a serie

Interpretação:

-> A decomposição sugere uma forte influência de fatores sazonais na produção, juntamente com uma tendência geral de crescimento. Eventos anômalos ou influências externas podem ser observados nos resíduos.

Conclusão:

-> A análise da decomposição revela padrões consistentes na produção de equipamentos elétricos, destacando a importância de considerar tanto os fatores sazonais quanto os de longo prazo na compreensão da dinâmica da série temporal.



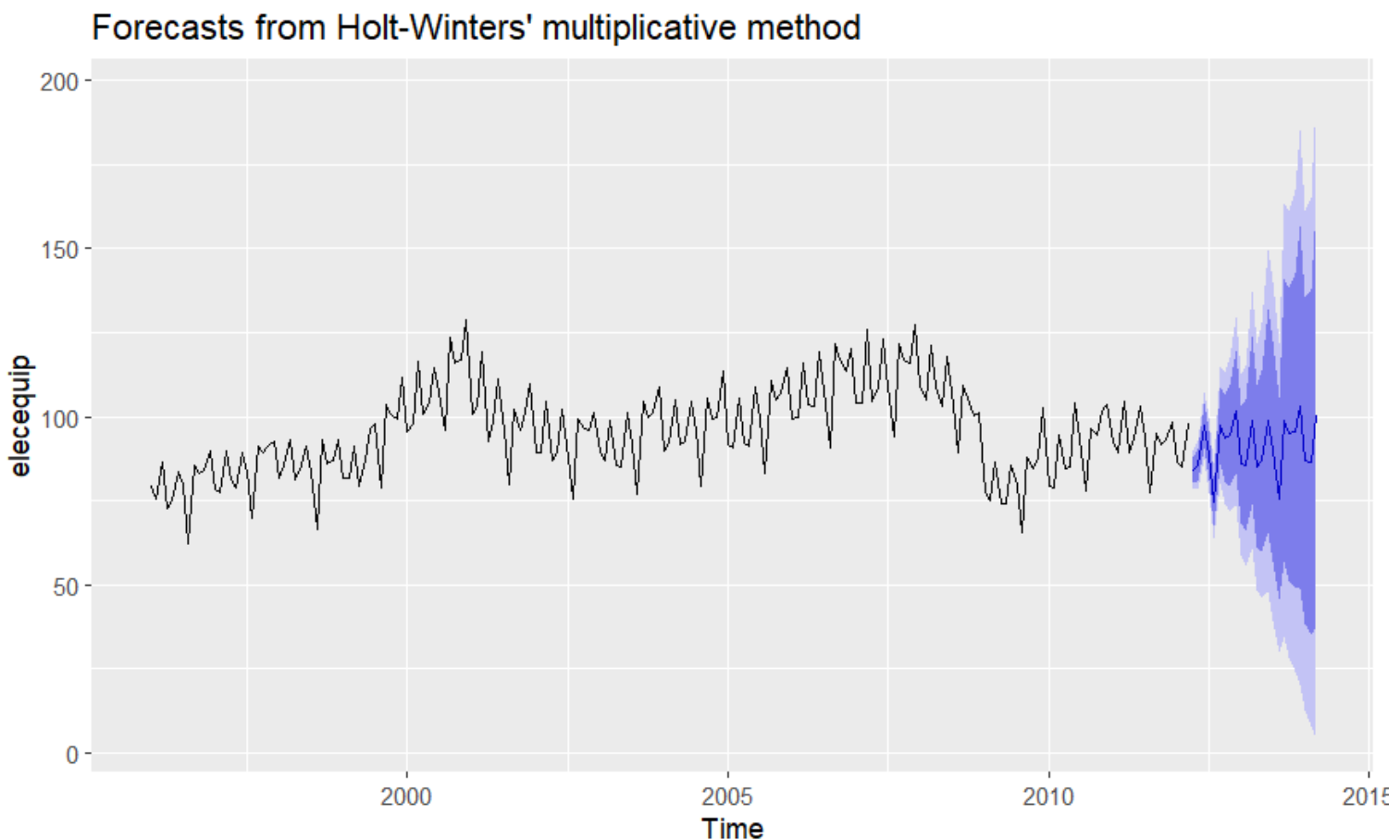
# Previsao:

Espaço de dois periodos

```
# Carregando o pacote 'forecast'
library(forecast)

# Modelo Holt-Winters com sazonalidade multiplicativa
modelo_hw <- hw(elecequip, seasonal="multiplicative")

# Fazendo previsões para os próximos 2 períodos sazonais completos
previsoes <- forecast(modelo_hw, h=2*frequency(elecequip))
previsoes
autoplot(previsoes)
```



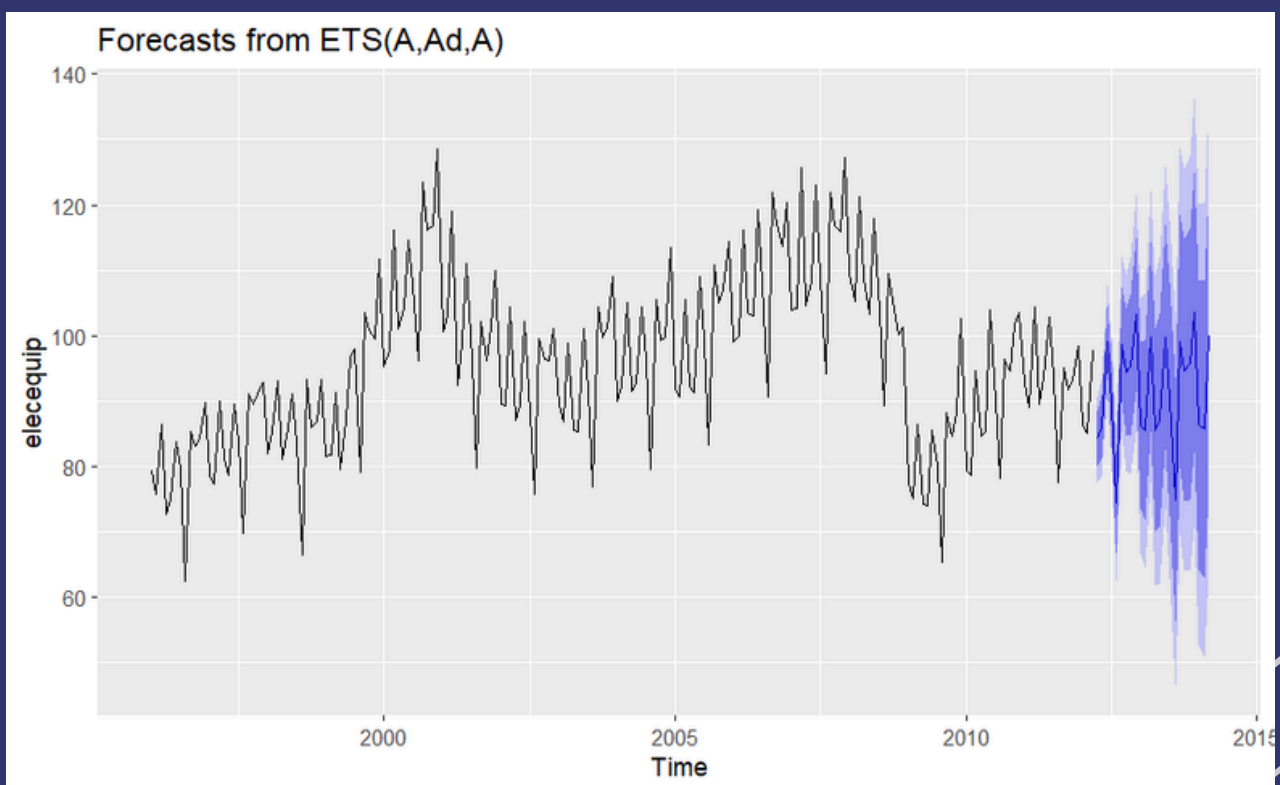
# Previsao

Justificativa do método de amortecimento de Holt-Winters:

-> O método de Holt-Winters leva em conta tanto a tendência como a sazonalidade presentes na série temporal, permitindo uma modelagem precisa das variações observadas. Ele também é capaz de se adaptar a mudanças nessas características, o que é importante para prever esta série, que tem grandes flutuações. Este método é amplamente utilizado e demonstrou ser eficaz na previsão de séries temporais, especialmente aquelas com sazonalidade multiplicativa, por esses motivos estamos utilizando ele.

## Otimizando parametros

```
# Ajustando o modelo Holt-Winters com otimização automática dos  
# parâmetros usando ETS  
modelo_ets <- ets(elecequip, model="AAA")  
  
# Fazendo previsões para os próximos 2 anos (2 períodos sazonais  
# completos)  
previsoes2 <- forecast(modelo_ets, h=2*frequency(elecequip))  
previsoes2  
autoplot(previsoes2)
```



# Calculando erros:

```
mse <- mean((elecequip - previsoes_1passo)^2)
mape <- mean(abs((elecequip - previsoes_1passo)/elecequip)) * 100
mae <- mean(abs(elecequip - previsoes_1passo))
print(paste("Erro Quadrático Médio (MSE):", mse))
print(paste("Erro Absoluto Percentual Médio (MAPE):", mape))
print(paste("Erro Absoluto Médio (MAE):", mae))
```

```
[1] "Erro Quadrático Médio (MSE): 10.1142820418206"
[1] "Erro Absoluto Percentual Médio (MAPE): 2.58225518109558"
[1] "Erro Absoluto Médio (MAE): 2.4340525457889"
```

## Erro Quadrático Médio (MSE):

- > O MSE mede a média dos quadrados das diferenças entre os valores observados e previstos. Um valor menor de MSE indica um melhor ajuste do modelo.
- > O MSE pode ser sensível a outliers, pois os erros são elevados ao quadrado. É útil para entender a magnitude geral dos erros, mas deve ser interpretado com cautela se houver valores atípicos na série.

## Erro Absoluto Percentual Médio (MAPE):

- > O MAPE mede a média dos erros percentuais absolutos entre os valores observados e previstos. É expresso em porcentagem, tornando-o fácil de interpretar em termos relativos.
- > O MAPE é útil para entender o erro em termos relativos, mas pode ser distorcido por valores observados muito pequenos, resultando em porcentagens muito altas.

## Erro Absoluto Médio (MAE):

- > O MAE mede a média das diferenças absolutas entre os valores observados e previstos. É uma medida simples e intuitiva de erro.
- > O MAE é menos sensível a outliers em comparação com o MSE, oferecendo uma visão clara do erro médio em termos absolutos.

## Conclusão:

- > A baixa magnitude dos erros (MSE, MAPE, MAE) sugere que o modelo é adequado para prever a série temporal, indicando sua eficácia em prever a produção de equipamentos elétricos.

# Resíduos:

```
# Resíduos
residuos = residuals(modelo_ets)

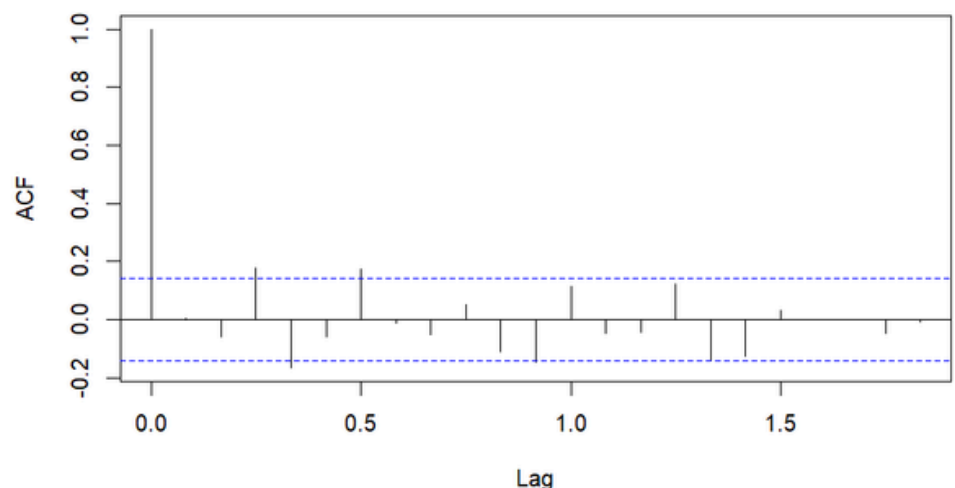
# Teste de Normalidade
shapiro_test = shapiro.test(residuos)
shapiro_test
# H0: Os dados seguem uma distribuição normal
# H1: Os dados não seguem uma distribuição normal
# p-valor = 0.10 > 0.05, não rejeitamos H0, os dados seguem uma
distribuição normal

# Funções de Autocorrelação (ACF e PACF)
acf(residuos, main="Função de Autocorrelação dos Resíduos")
pacf(residuos, main="Função de Autocorrelação Parcial dos
Resíduos")
# As funções mostram que os resíduos não têm autocorrelação
significativa, o que indica que o modelo capturou todas as
dependências temporais presentes na série.
```

Shapiro-wilk normality test

data: residuos  
W = 0.98807, p-value = 0.1014

Função de Autocorrelação dos Resíduos



# Graficos relevantes:

```
# Resíduos
```

```
plot(residuos, main="Resíduos do Modelo", ylab="Resíduos",  
xlab="Tempo")
```

```
# Histograma dos resíduos
```

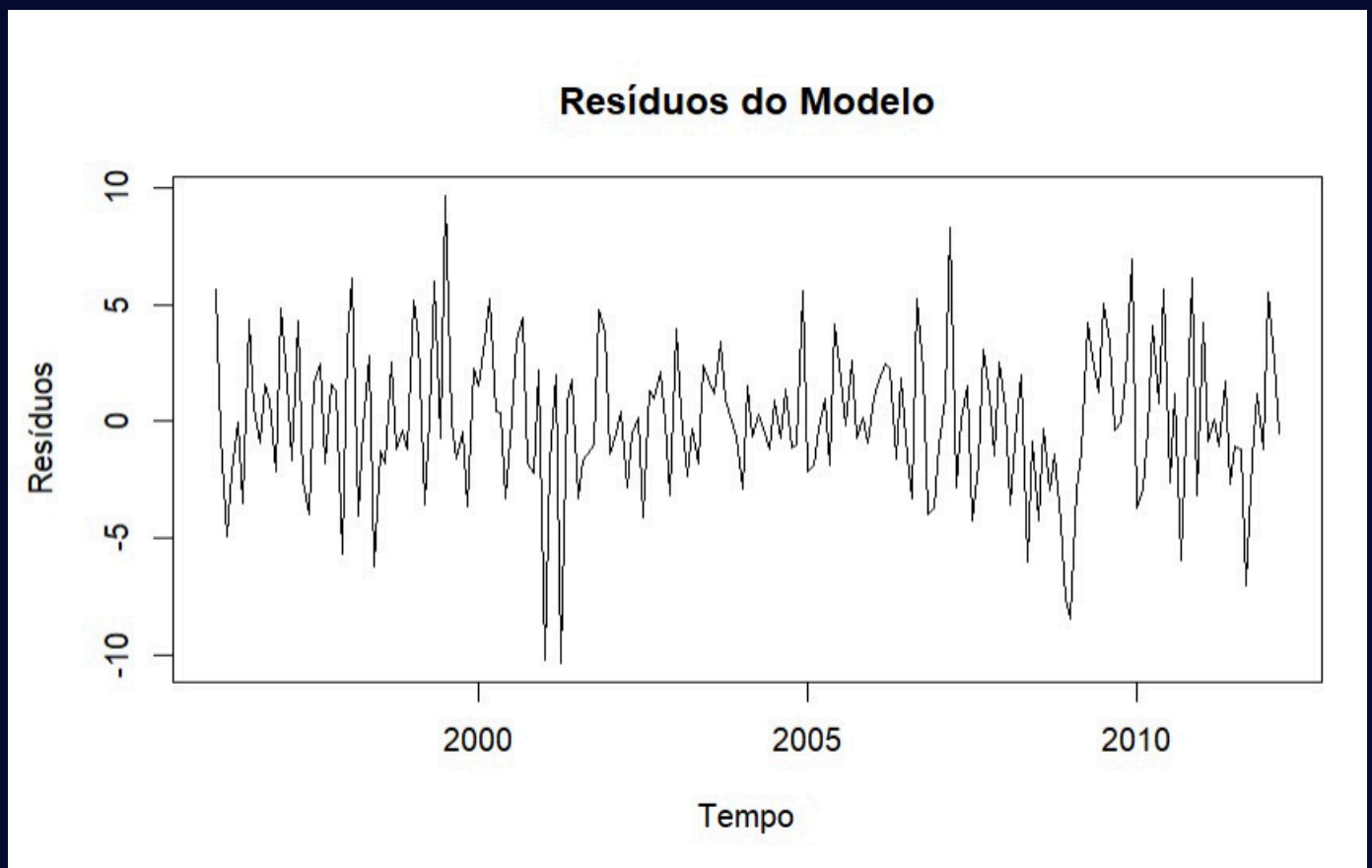
```
hist(residuos, breaks=20, main="Histograma dos Resíduos",  
xlab="Resíduos", col="blue")
```

```
# Q-Q plot dos resíduos
```

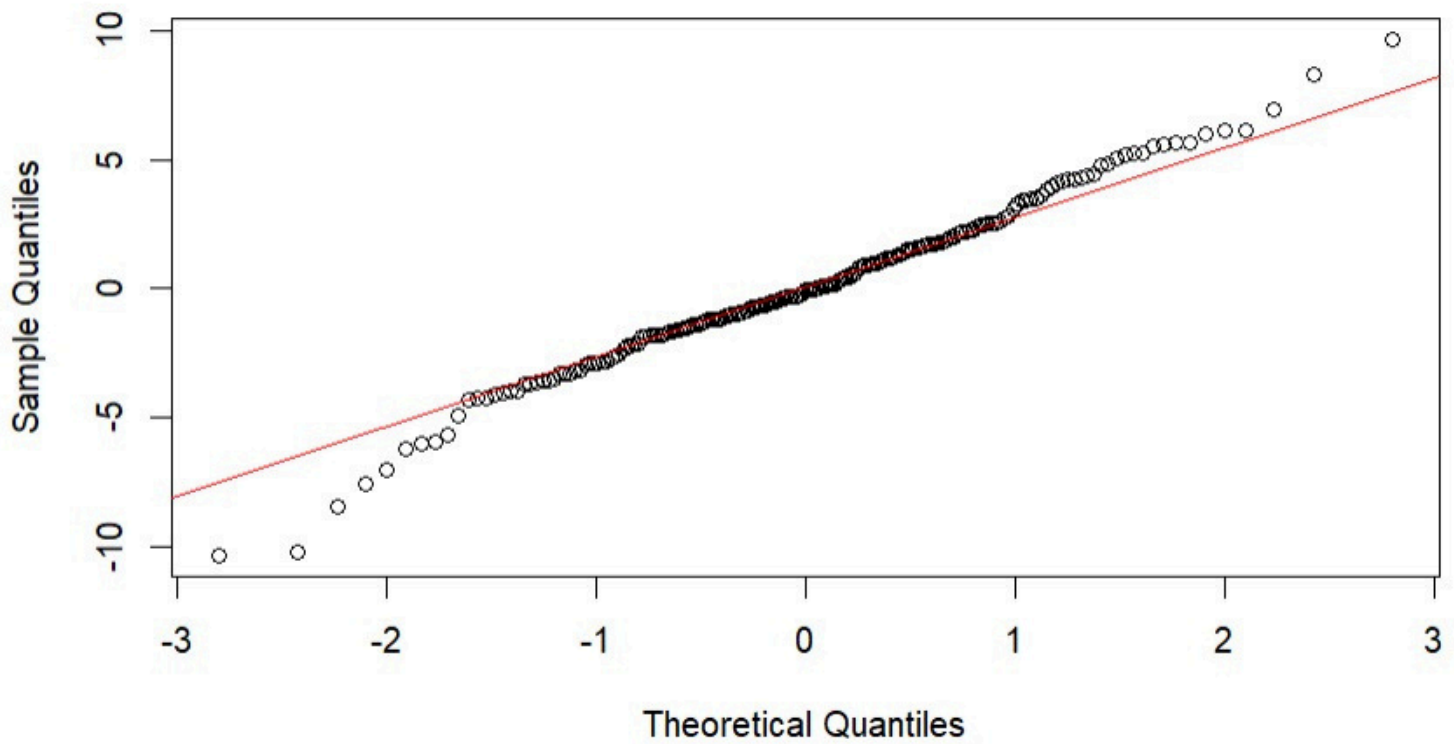
```
qqnorm(residuos)
```

```
qqline(residuos, col="red")
```

# A análise dos resíduos é crucial para validar o modelo ajustado. Como os resíduos se comportam como ruído branco (sem autocorrelação e normalmente distribuídos), o modelo é considerado adequado.



Normal Q-Q Plot



Histograma dos Resíduos

