

# Homework 1

Daniel de Souza Sobrinho Macedo RA:813524

2025-06-20

## Inicialização das bibliotecas

```
#install.packages("gridExtra")
#install.packages("quantmod")
#install.packages("grid")
#install.packages("reshape2")
#install.packages("e1071")
#install.packages("BatchGetSymbols")
#install.packages("rugarch")
#install.packages("tidyverse")

library(BatchGetSymbols)
library(rugarch)
library(tidyverse)
library(ggthemes)
library(gridExtra)
library(quantmod)
library(grid)
library(reshape2)

rm(list=ls()) # Limpar variáveis da sessão
```

As 4 empresas selecionadas foram Vale, Embraer, Companhia Siderurgica Nacional e Ambev

```
empresas <- c('VALE3.SA', 'EMBR3.SA', 'CSNA3.SA', 'ABEV3.SA')

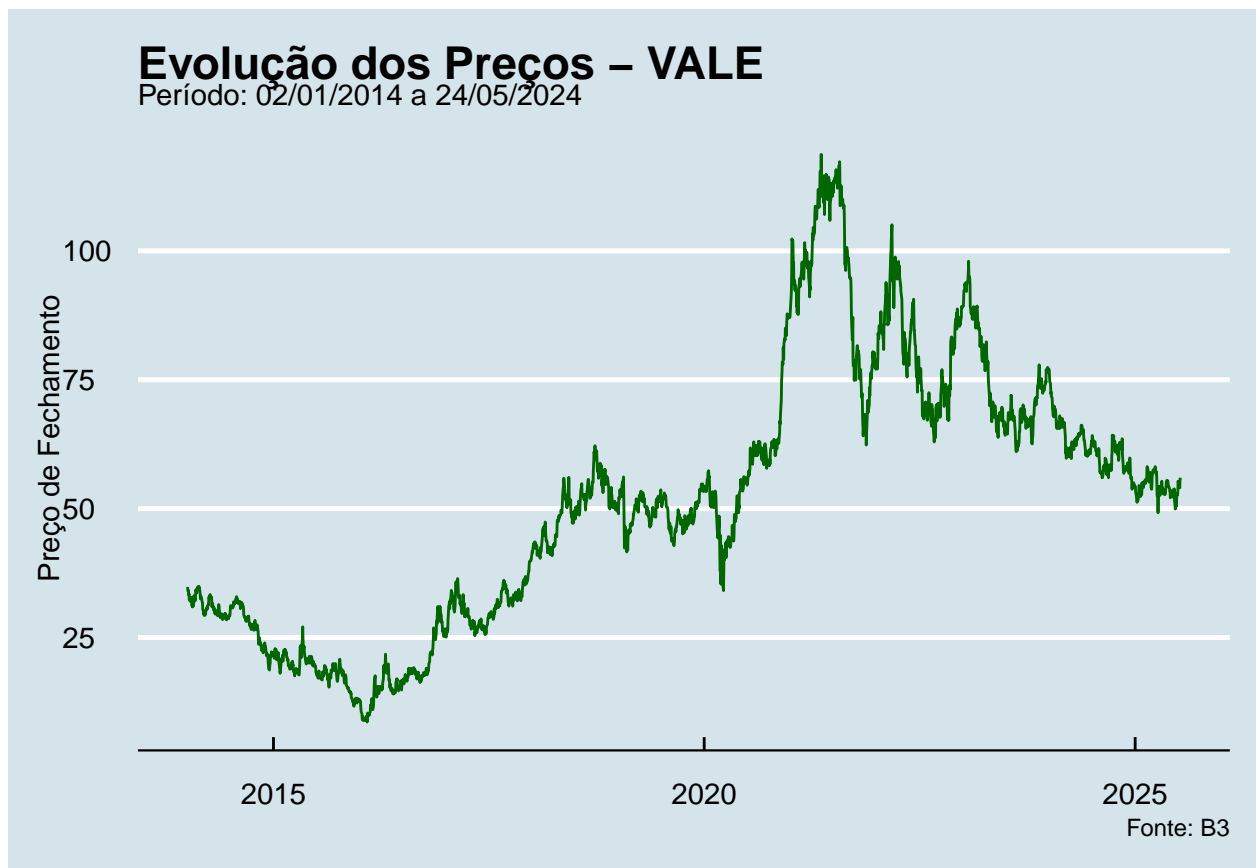
ativos <- BatchGetSymbols(tickers = empresas,
                          first.date = '2014-01-01',
                          last.date = Sys.time(),
                          type.return = "log",
                          freq.data = "daily")[[2]]
```

Agora vamos filtrar os dados:

```
dados_vale <- ativos %>% filter(ticker == 'VALE3.SA')
dados_embraer <- ativos %>% filter(ticker == 'EMBR3.SA')
dados_csn <- ativos %>% filter(ticker == 'CSNA3.SA')
dados_ambev <- ativos %>% filter(ticker == 'ABEV3.SA')
```

Vamos começar com os gráficos da Vale:

```
graf_vale_preco <- ggplot(dados_vale, aes(x = ref.date, y = price.close)) +  
  geom_line(color = '#006600') +  
  labs(  
    x = NULL, y = "Preço de Fechamento",  
    title = "Evolução dos Preços - VALE",  
    subtitle = "Período: 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_vale_preco
```



- Preços:

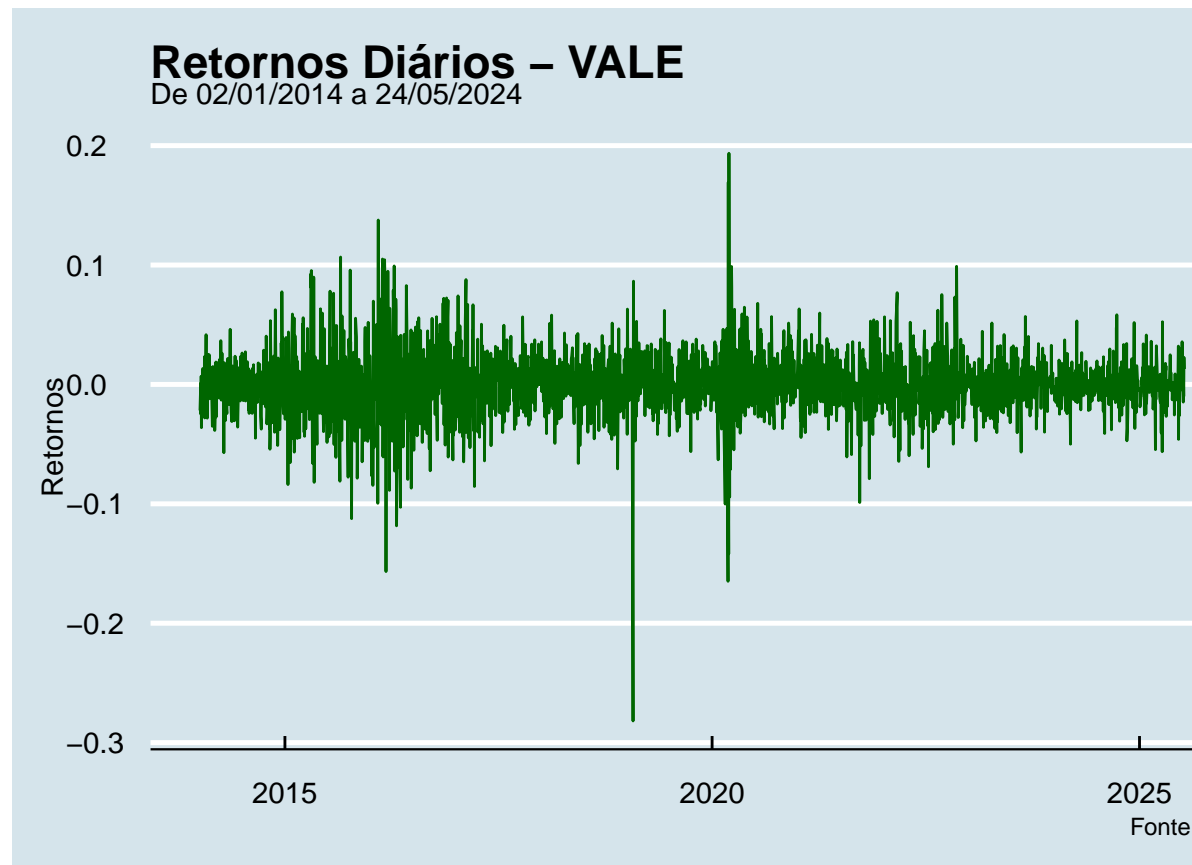
```
ret_vale <- dados_vale %>%  
  select(ref.date, ticker, ret.closing.prices)  
  
graf_vale_retorno <- ggplot(ret_vale, aes(x = ref.date, y = ret.closing.prices)) +  
  geom_line(color = '#006600') +  
  labs(  
    x = NULL, y = "Retorno",  
    title = "Retorno dos Preços - VALE",  
    subtitle = "Período: 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_vale_retorno
```

```

x = NULL, y = "Retornos",
title = "Retornos Diários - VALE",
subtitle = "De 02/01/2014 a 24/05/2024",
caption = "Fonte: B3"
) +
theme_economist()

graf_vale_retorno

```



- Retornos diários

```

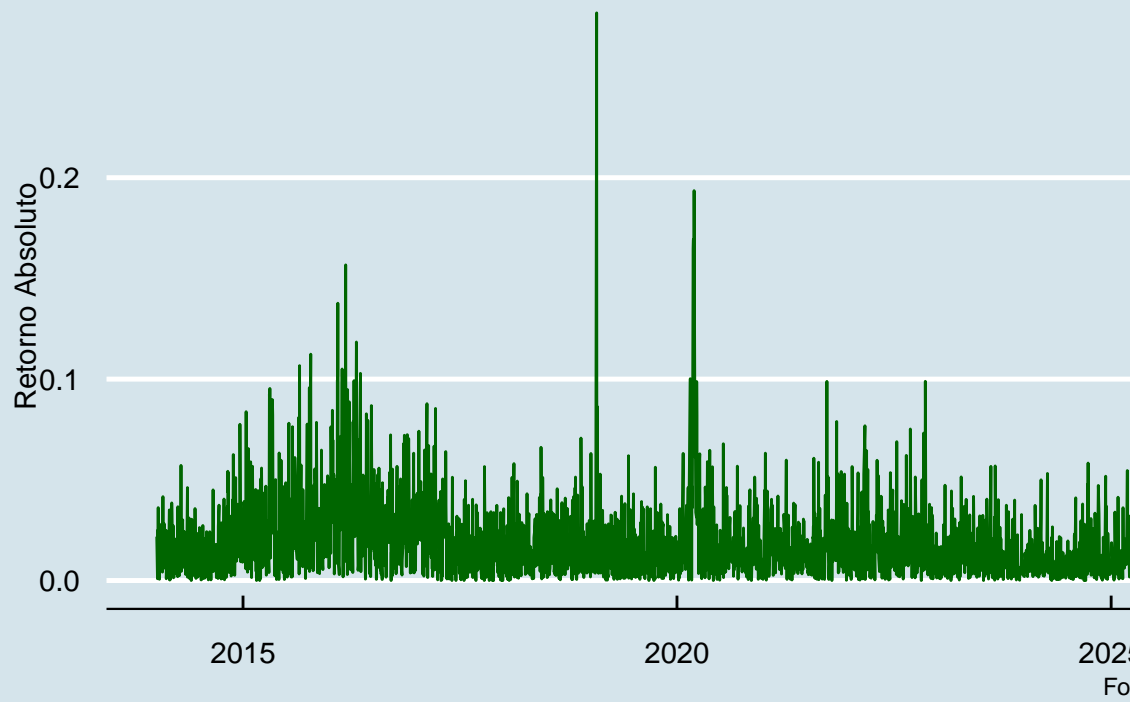
graf_vale_abs <- ggplot(ret_vale, aes(x = ref.date, y = abs(ret.closing.prices))) +
  geom_line(color = '#006600') +
  labs(
    x = NULL, y = "Retorno Absoluto",
    title = "Magnitude dos Retornos - VALE",
    subtitle = "Período: 02/01/2014 a 24/05/2024",
    caption = "Fonte: B3"
  ) +
  theme_economist()

graf_vale_abs

```

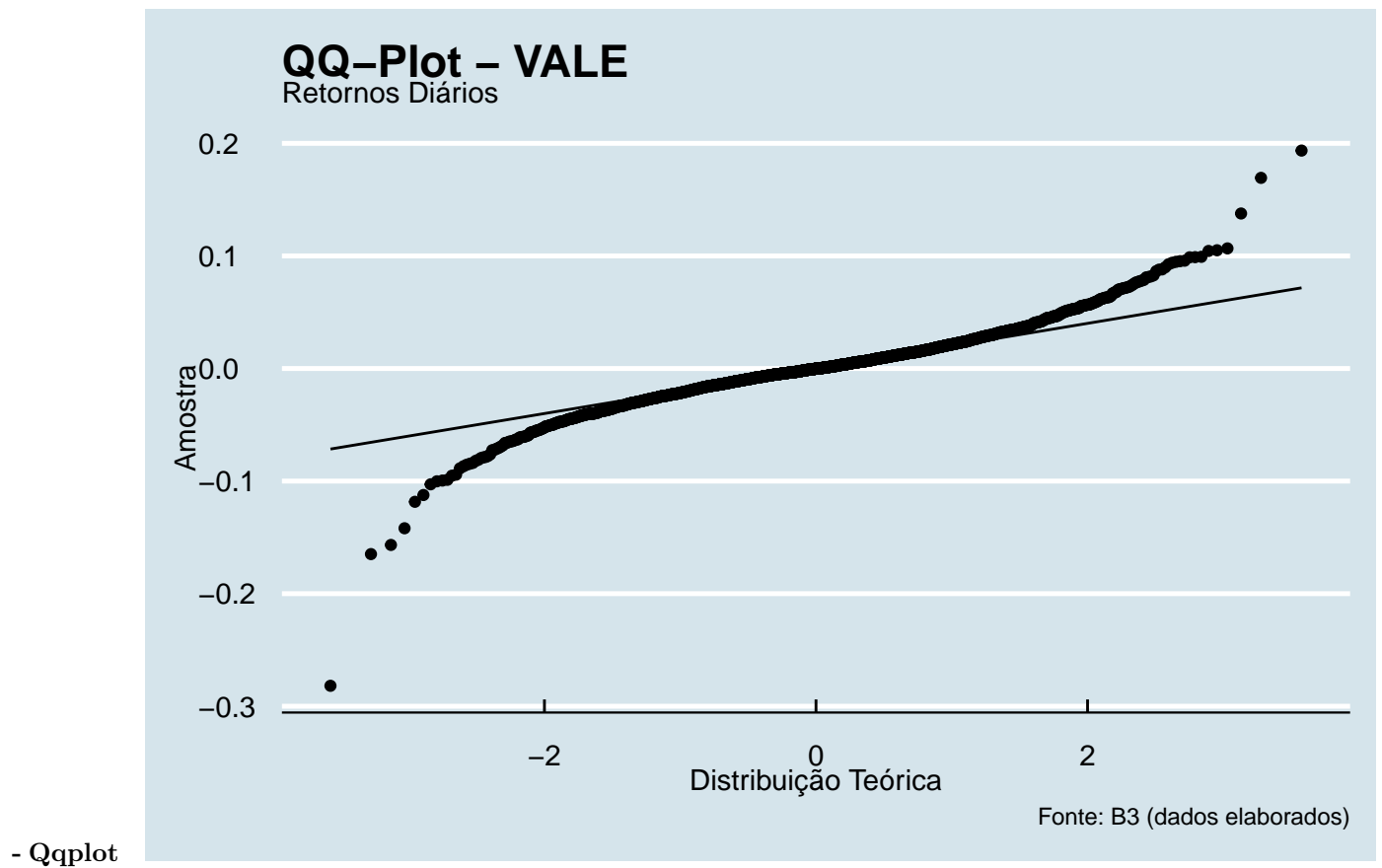
## Magnitude dos Retornos – VALE

Período: 02/01/2014 a 24/05/2024



- Retornos absolutos

```
qqplot_vale <- ggplot(ret_vale, aes(sample = ret.closing.prices)) +  
  stat_qq() +  
  stat_qq_line() +  
  labs(  
    x = "Distribuição Teórica", y = "Amostra",  
    title = "QQ-Plot - VALE",  
    subtitle = "Retornos Diários",  
    caption = "Fonte: B3 (dados elaborados)"  
  ) +  
  theme_economist()  
  
qqplot_vale
```

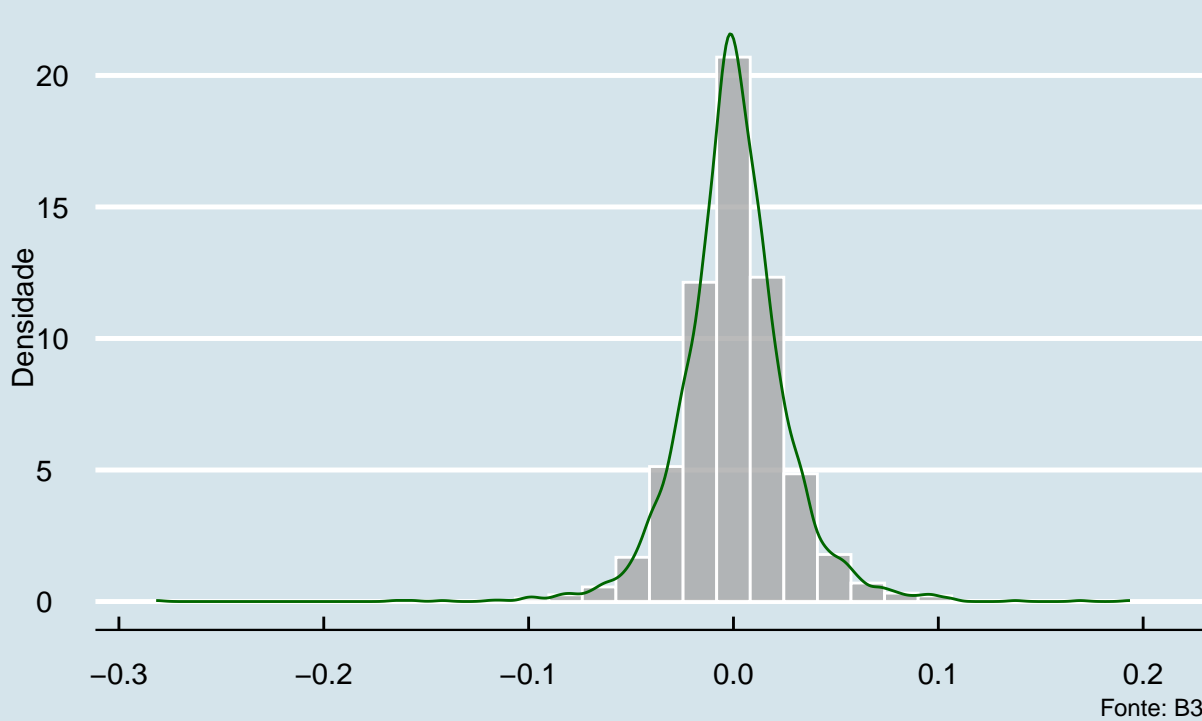


```
hist_vale <- ggplot(ret_vale) +
  geom_histogram(aes(x = ret.closing.prices, y = ..density..),
    fill = "darkgray", color = "white", alpha = 0.8) +
  geom_density(aes(x = ret.closing.prices), color = "#006600") +
  labs(
    x = NULL, y = "Densidade",
    title = "Distribuição dos Retornos - VALE",
    subtitle = "Histograma com Curva de Densidade",
    caption = "Fonte: B3"
  ) +
  theme_economist()

hist_vale
```

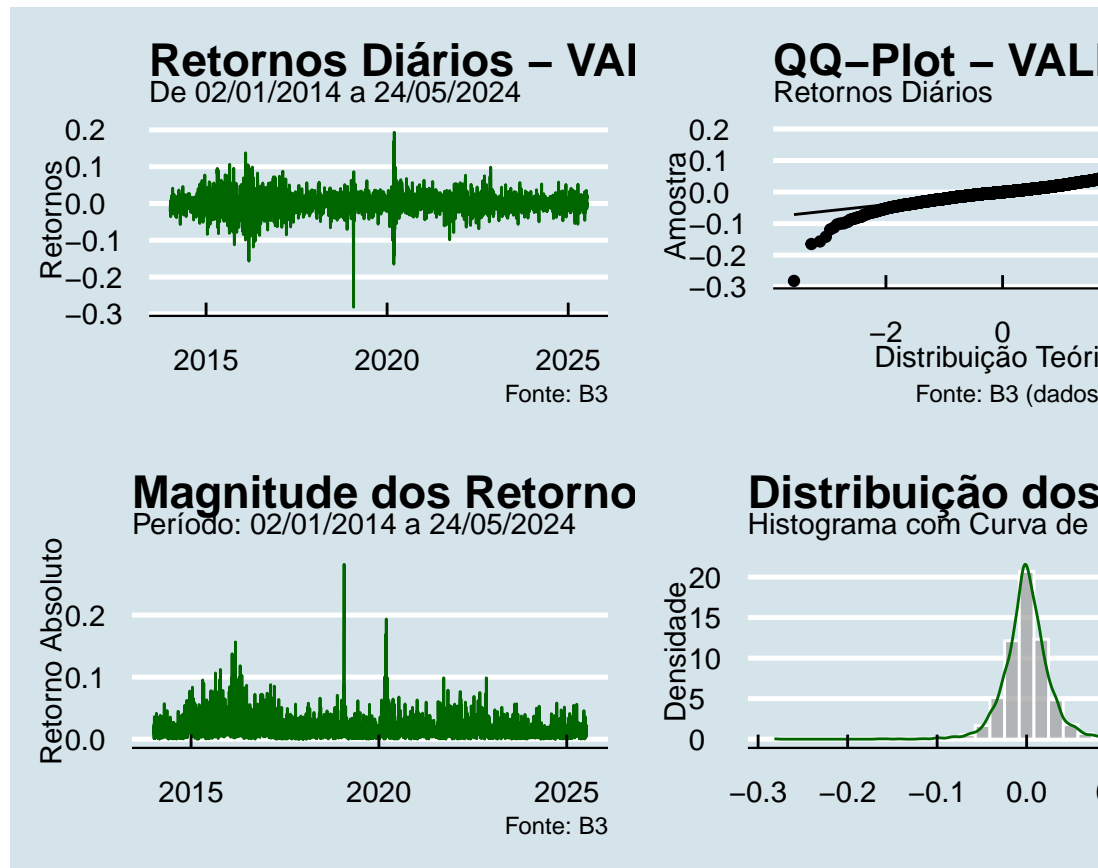
## Distribuição dos Retornos – VALE

Histograma com Curva de Densidade



- Histograma

```
cowplot::plot_grid(graf_vale_retorno, qqplot_vale, graf_vale_abs, hist_vale, nrow = 2)
```



- Todos os gráficos juntos

```
precos_vale <- dados_vale[[4]] # Coluna de preços

estatisticas_vale <- data.frame(
  Estatística = c("Média", "Mediana", "Desvio Padrão", "Curtose", "Assimetria",
                  "Valor Mínimo", "Valor Máximo", "1º Quartil", "3º Quartil"),
  Valor = c(
    mean(precos_vale, na.rm = TRUE),
    median(precos_vale, na.rm = TRUE),
    sd(precos_vale, na.rm = TRUE),
    e1071::kurtosis(precos_vale, na.rm = TRUE),
    e1071::skewness(precos_vale, na.rm = TRUE),
    min(precos_vale, na.rm = TRUE),
    max(precos_vale, na.rm = TRUE),
    quantile(precos_vale, 0.25, na.rm = TRUE),
    quantile(precos_vale, 0.75, na.rm = TRUE)
  )
)

tabela_vale <- tableGrob(estatisticas_vale)
grid.draw(tabela_vale)
```

	Estatística	Valor
1	Média	51.4814550
2	Mediana	51.9299984
3	Desvio Padrão	25.0342576
4	Curtose	-0.4770716
5	Assimetria	0.3943489
6	Valor Mínimo	8.6000004
7	Valor Máximo	118.7200012
8	1º Quartil	30.0100002
9	3º Quartil	67.1125011

- As principais estatísticas descritivas:

Os gráficos da Embraer:

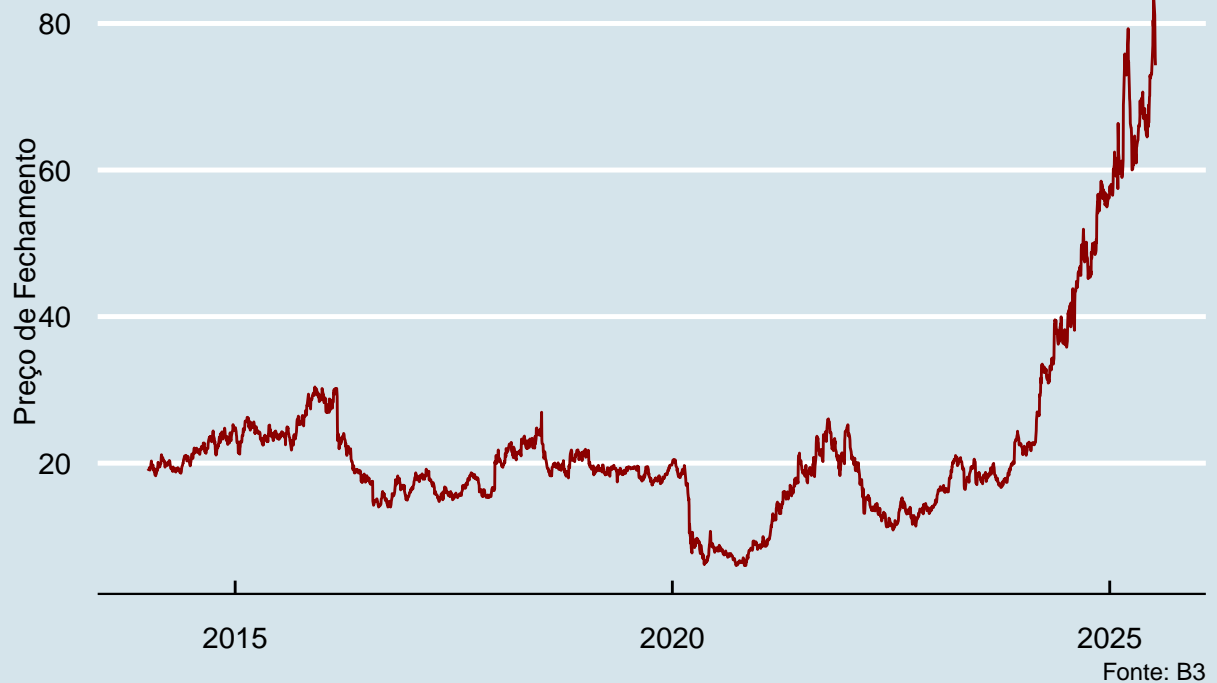
```
graf_embraer_preco <- ggplot(dados_embraer, aes(x = ref.date, y = price.close)) +
  geom_line(color = '#8b0000') +
  labs(
    x = NULL, y = "Preço de Fechamento",
    title = "Histórico de Preços - EMBRAER",
    subtitle = "Período: 02/01/2014 a 24/05/2024",
    caption = "Fonte: B3"
  ) +
  theme_economist()

graf_embraer_preco
```



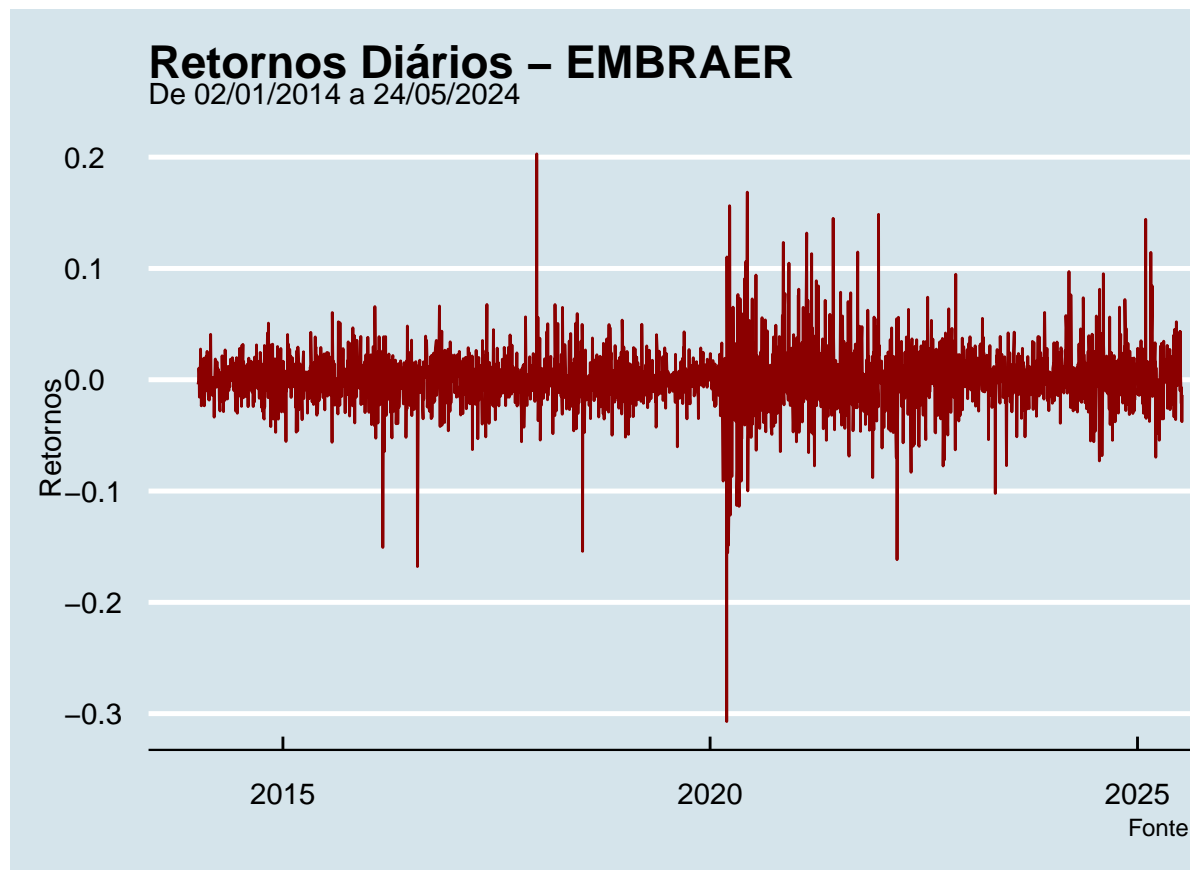
## Histórico de Preços – EMBRAER

Período: 02/01/2014 a 24/05/2024



- Preços:

```
ret_embraer <- dados_embraer %>%  
  select(ref.date, ticker, ret.closing.prices)  
  
graf_embraer_retorno <- ggplot(ret_embraer, aes(x = ref.date, y = ret.closing.prices)) +  
  geom_line(color = '#8b0000') +  
  labs(  
    x = NULL, y = "Retornos",  
    title = "Retornos Diários - EMBRAER",  
    subtitle = "De 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_embraer_retorno
```

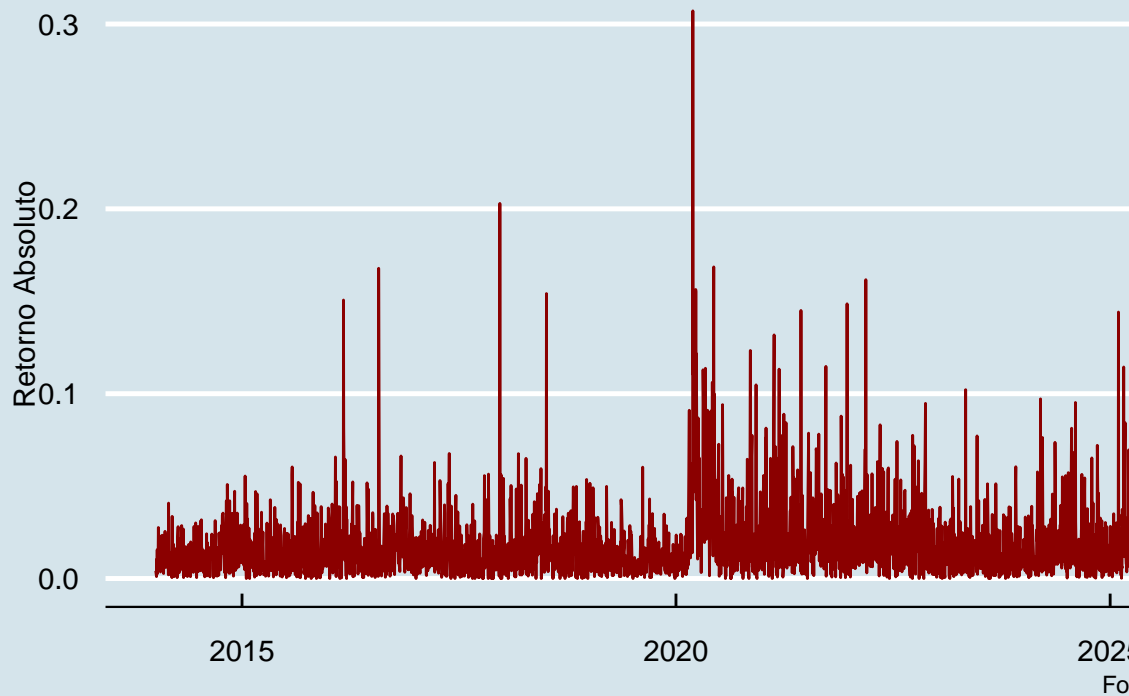


- Retornos diários

```
graf_embraer_abs <- ggplot(ret_embraer, aes(x = ref.date, y = abs(ret.closing.prices))) +  
  geom_line(color = '#8b0000') +  
  labs(  
    x = NULL, y = "Retorno Absoluto",  
    title = "Magnitude dos Retornos - EMBRAER",  
    subtitle = "Período: 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_embraer_abs
```

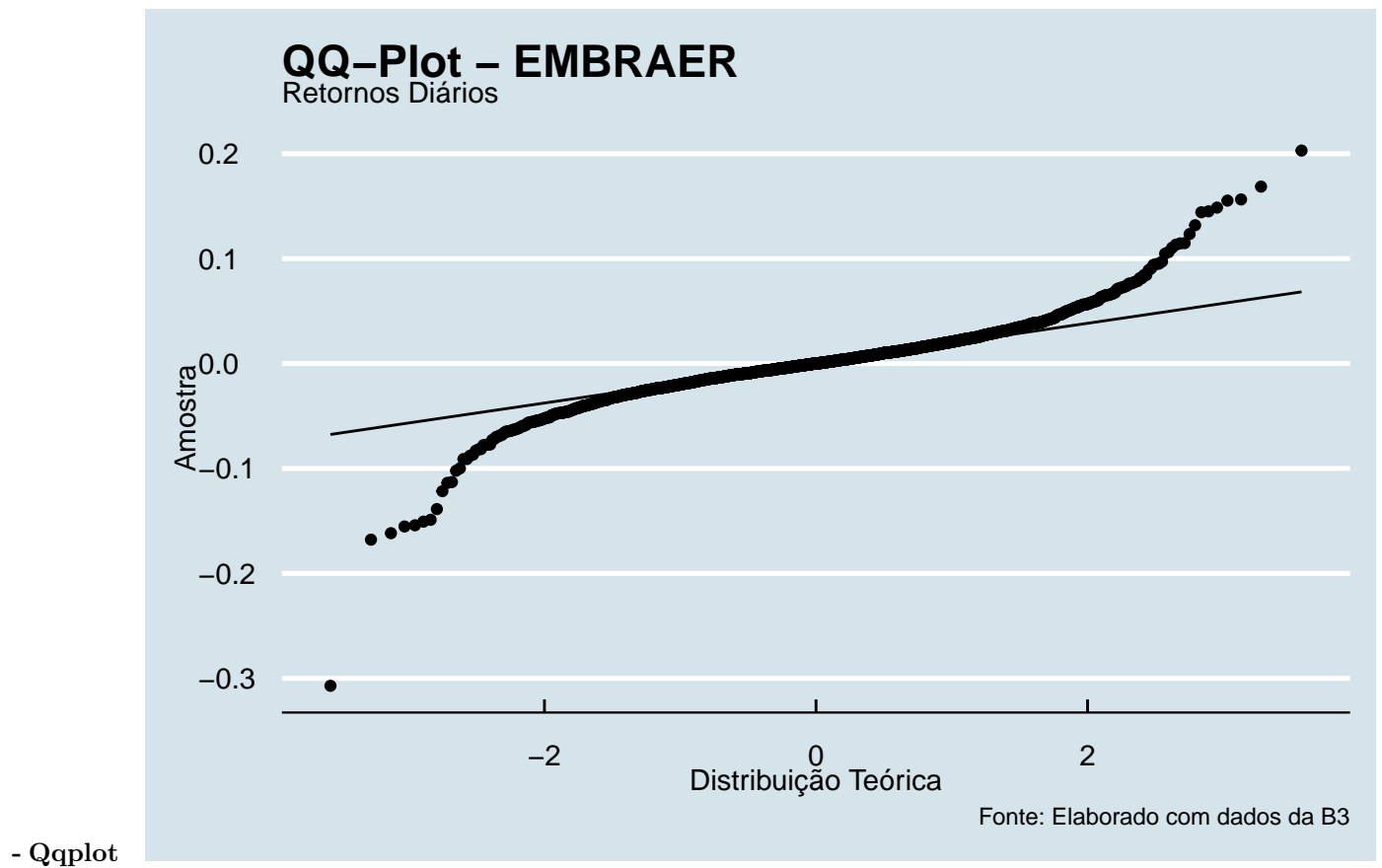
## Magnitude dos Retornos – EMBRAER

Período: 02/01/2014 a 24/05/2024



- Retornos absolutos

```
qqplot_embraer <- ggplot(ret_embraer, aes(sample = ret.closing.prices)) +  
  stat_qq() +  
  stat_qq_line() +  
  labs(  
    x = "Distribuição Teórica", y = "Amostra",  
    title = "QQ-Plot - EMBRAER",  
    subtitle = "Retornos Diários",  
    caption = "Fonte: Elaborado com dados da B3"  
  ) +  
  theme_economist()  
  
qqplot_embraer
```

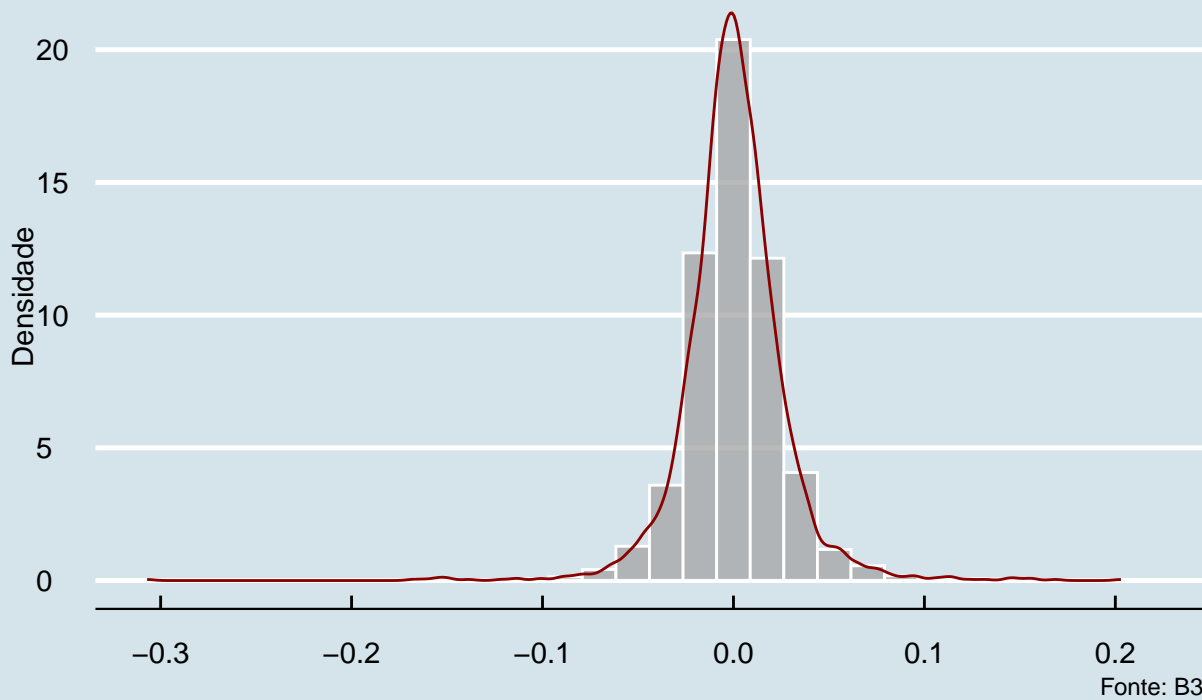


```
hist_embraer <- ggplot(ret_embraer) +
  geom_histogram(aes(x = ret.closing.prices, y = ..density..),
    fill = "darkgray", color = "white", alpha = 0.8) +
  geom_density(aes(x = ret.closing.prices), color = "#8b0000") +
  labs(
    x = NULL, y = "Densidade",
    title = "Distribuição dos Retornos - EMBRAER",
    subtitle = "Histograma com Densidade",
    caption = "Fonte: B3"
  ) +
  theme_economist()

hist_embraer
```

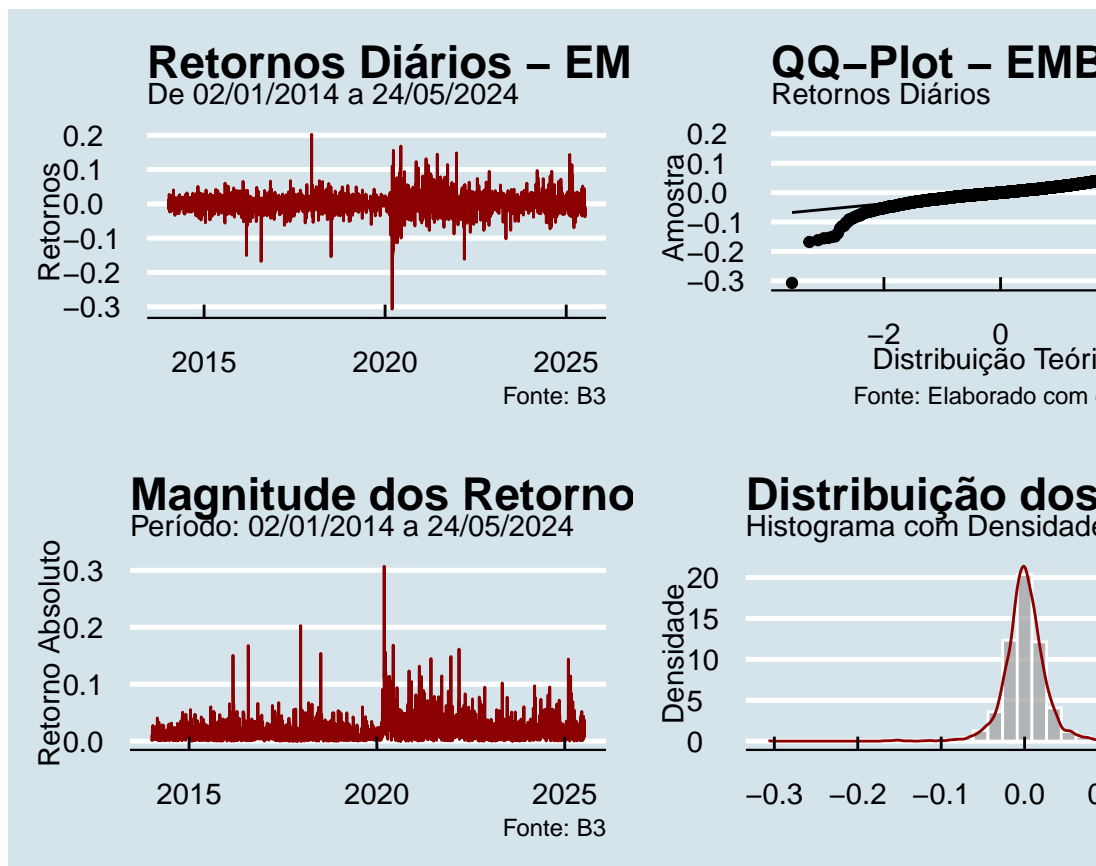
## Distribuição dos Retornos – EMBRAER

Histograma com Densidade



- Histograma

```
cowplot::plot_grid(graf_embraer_retorno, qqplot_embraer, graf_embraer_abs, hist_embraer, nrow = 2)
```



- Todos os gráficos juntos

```
precos_embraer <- dados_embraer[[4]]

estatisticas_embraer <- data.frame(
  Estatística = c("Média", "Mediana", "Desvio Padrão", "Curtose", "Assimetria",
                  "Valor Mínimo", "Valor Máximo", "1º Quartil", "3º Quartil"),
  Valor = c(
    mean(precos_embraer, na.rm = TRUE),
    median(precos_embraer, na.rm = TRUE),
    sd(precos_embraer, na.rm = TRUE),
    e1071::kurtosis(precos_embraer, na.rm = TRUE),
    e1071::skewness(precos_embraer, na.rm = TRUE),
    min(precos_embraer, na.rm = TRUE),
    max(precos_embraer, na.rm = TRUE),
    quantile(precos_embraer, 0.25, na.rm = TRUE),
    quantile(precos_embraer, 0.75, na.rm = TRUE)
  )
)

tabela_embraer <- tableGrob(estatisticas_embraer)
grid.draw(tabela_embraer)
```

	Estatística	Valor
1	Média	22.352756
2	Mediana	19.335000
3	Desvio Padrão	12.677857
4	Curtose	5.844707
5	Assimetria	2.344906
6	Valor Mínimo	6.030000
7	Valor Máximo	83.489998
8	1º Quartil	16.320000
9	3º Quartil	23.235000

- As principais estatísticas descritivas:

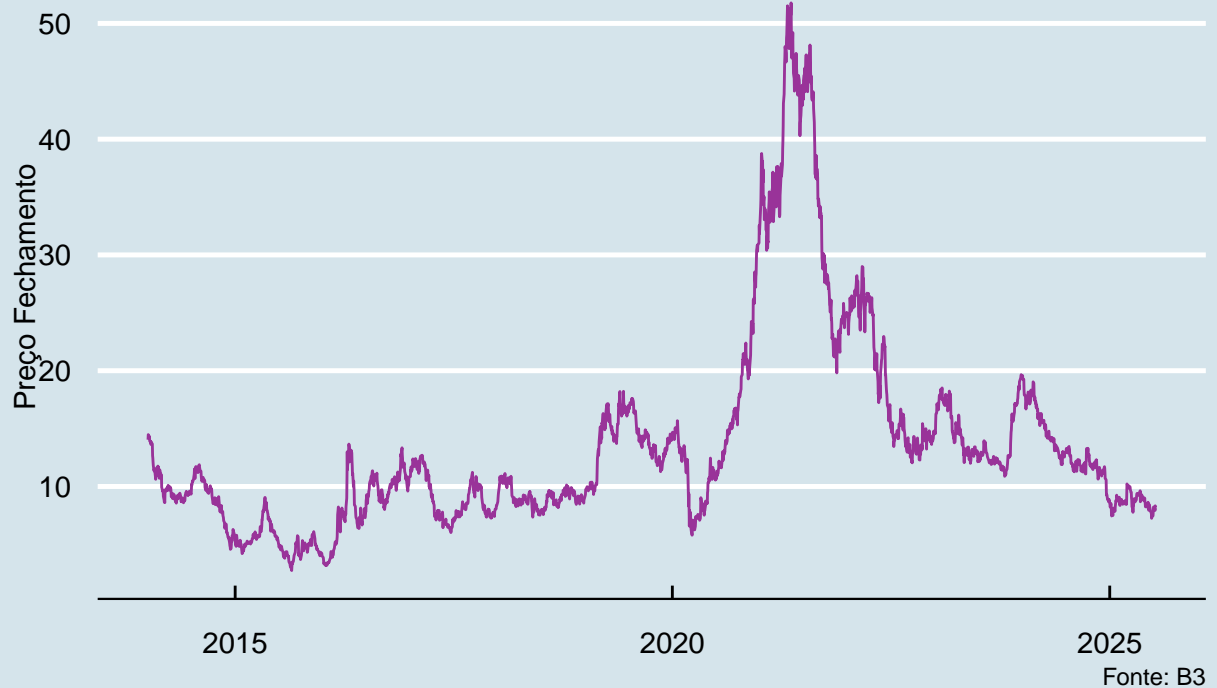
Os gráficos da Siderurgica:

```
graf_csn_preco <- ggplot(dados_csn, aes(x = ref.date, y = price.close)) +
  geom_line(color = '#993399') +
  labs(
    x = NULL, y = "Preço Fechamento",
    title = "Histórico Diário - Siderúrgica Nacional",
    subtitle = "De 02/01/2014 até 24/05/2024",
    caption = "Fonte: B3"
  ) +
  theme_economist()

graf_csn_preco
```

## Histórico Diário – Siderúrgica Nacional

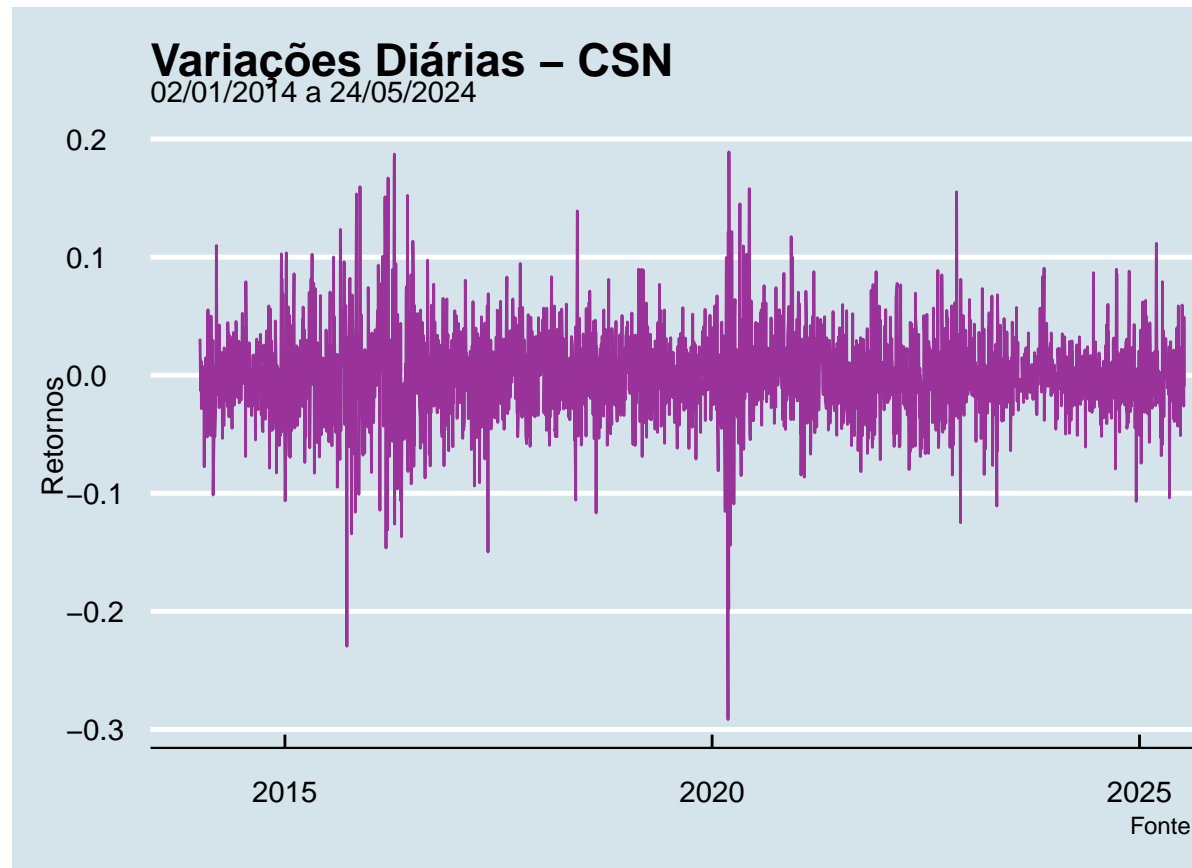
De 02/01/2014 até 24/05/2024



- Preços:

```
ret_csn <- dados_csn %>%  
  select(ref.date, ticker, ret.closing.prices)  
  
graf_csn_retorno <- ggplot(ret_csn, aes(x = ref.date, y = ret.closing.prices)) +  
  geom_line(color = '#993399') +  
  labs(  
    x = NULL, y = "Retornos",  
    title = "Variações Diárias - CSN",  
    subtitle = "02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_csn_retorno
```



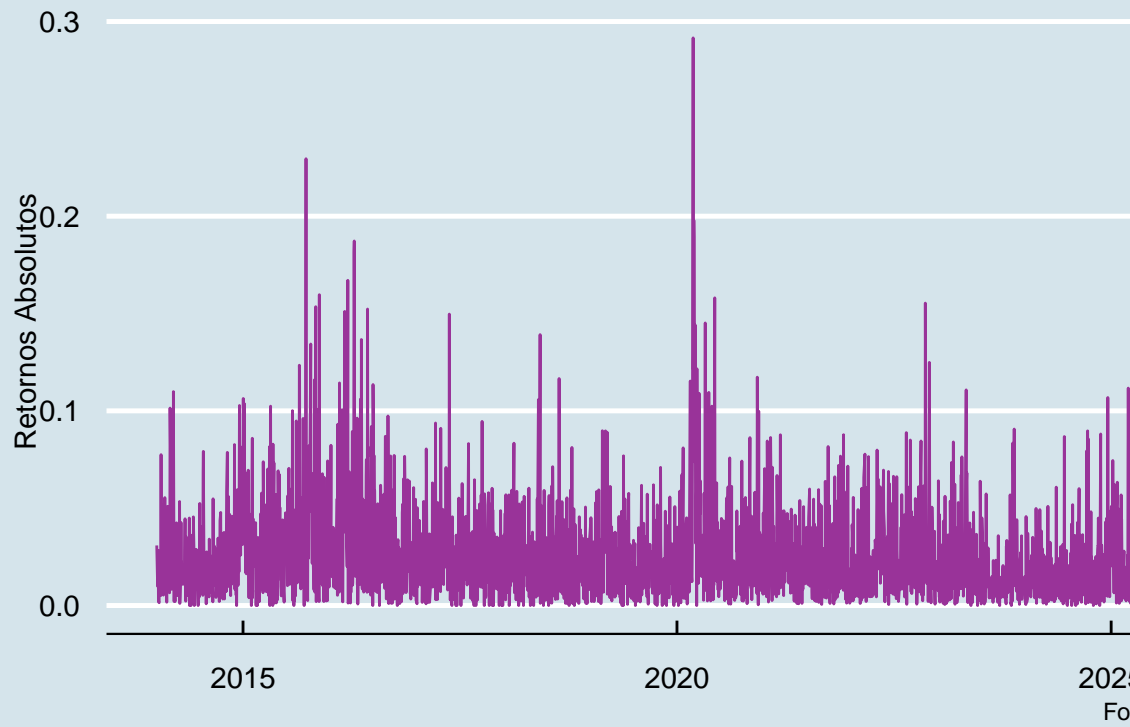


- Retornos diários

```
graf_csn_abs <- ggplot(ret_csn, aes(x = ref.date, y = abs(ret.closing.prices))) +  
  geom_line(color = '#993399') +  
  labs(  
    x = NULL, y = "Retornos Absolutos",  
    title = "Amplitude das Variações - CSN",  
    subtitle = "Período de 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_csn_abs
```

## Amplitude das Variações – CSN

Período de 02/01/2014 a 24/05/2024

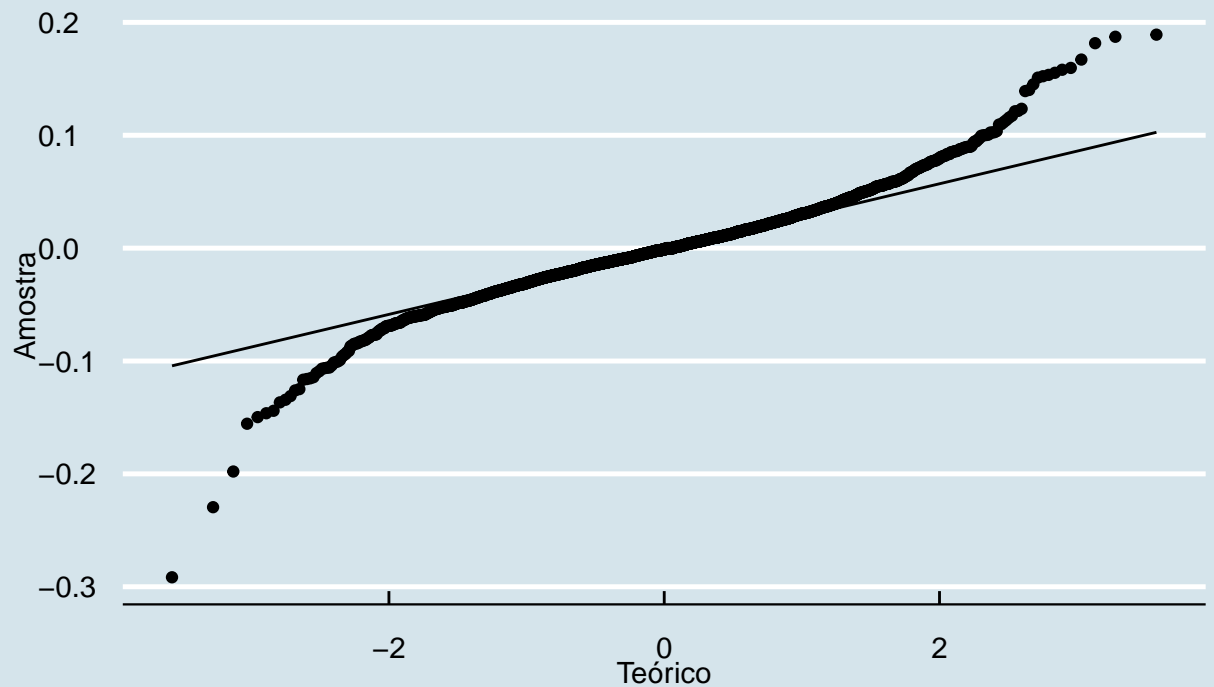


- Retornos absolutos

```
qqplot_csn <- ggplot(ret_csn, aes(sample = ret.closing.prices)) +  
  stat_qq() +  
  stat_qq_line() +  
  labs(  
    x = "Teórico", y = "Amostra",  
    title = "QQ-Plot - Retornos CSN",  
    subtitle = "Elaboração a partir de dados da B3",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
qqplot_csn
```

## QQ-Plot – Retornos CSN

Elaboração a partir de dados da B3



Fonte: B3

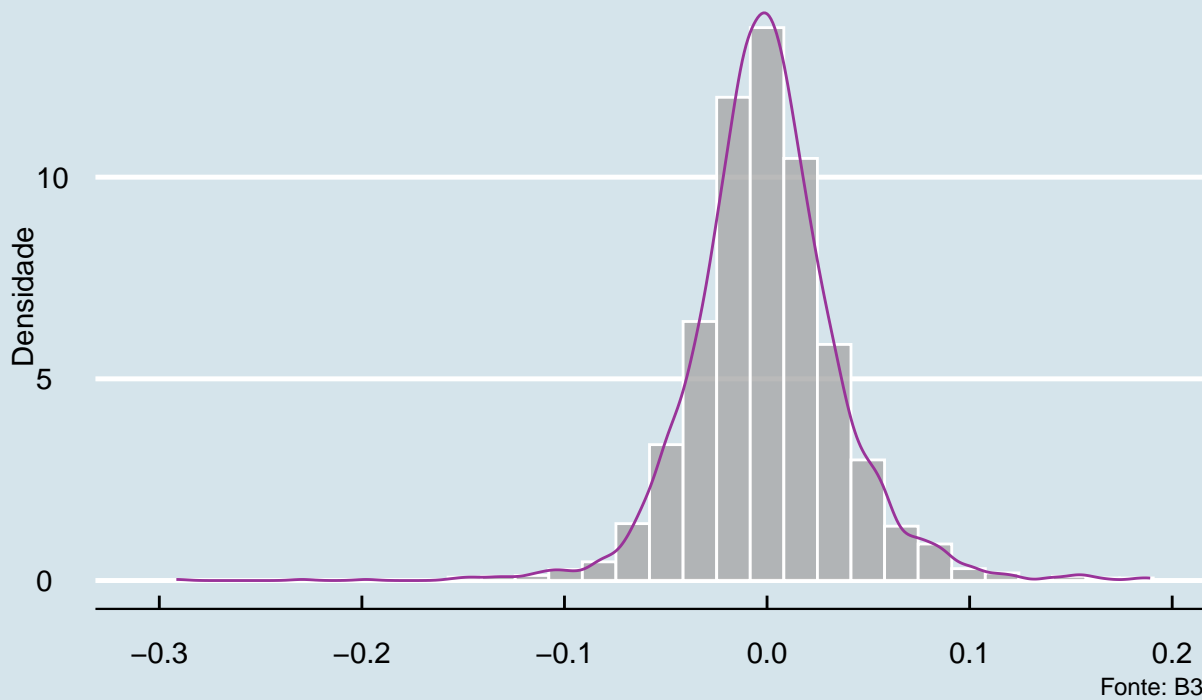
- Qqplot

```
hist_csn <- ggplot(ret_csn) +  
  geom_histogram(aes(x = ret.closing.prices, y = ..density..),  
                 fill = "darkgray", color = "white", alpha = 0.8) +  
  geom_density(aes(x = ret.closing.prices), color = "#993399") +  
  labs(  
    x = NULL, y = "Densidade",  
    title = "Distribuição dos Retornos - CSN",  
    subtitle = "Histograma com Curva de Densidade",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()
```

hist\_csn

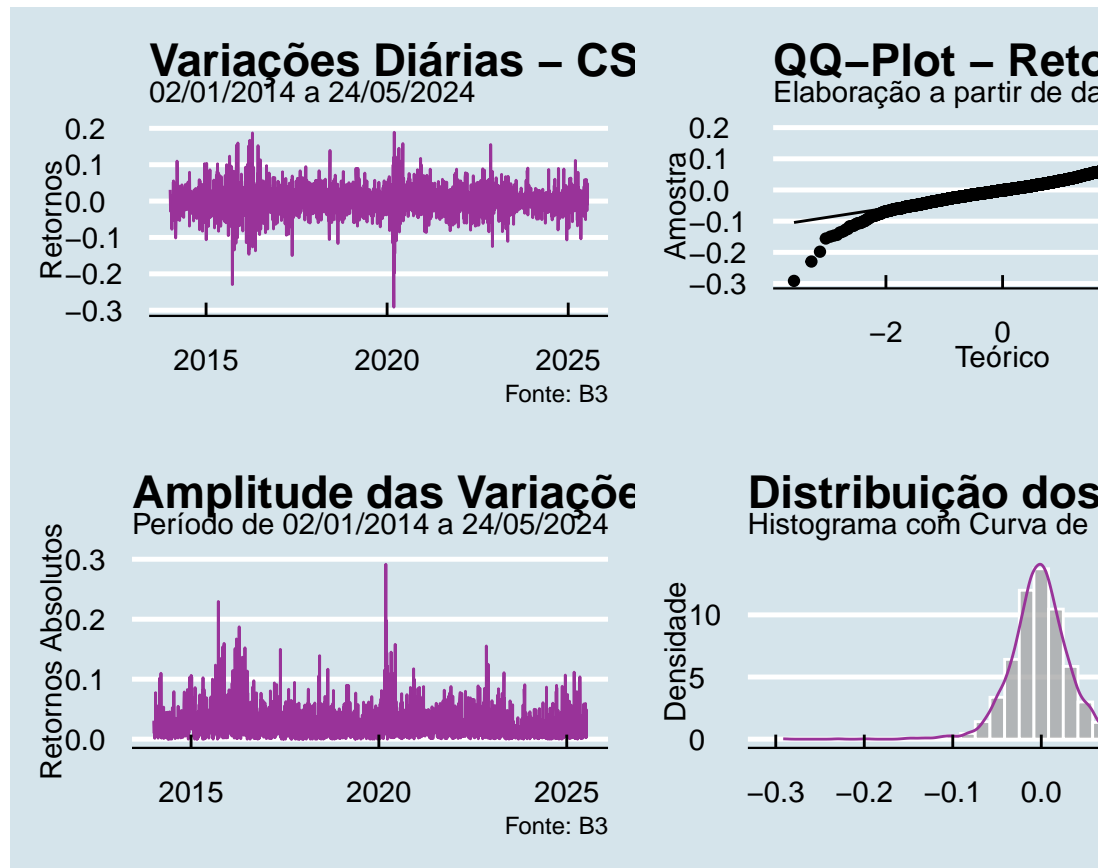
## Distribuição dos Retornos – CSN

Histograma com Curva de Densidade



- Histograma

```
cowplot::plot_grid(graf_csn_retorno, qqplot_csn, graf_csn_abs, hist_csn, nrow = 2)
```



- Todos os gráficos juntos

```
precos_csn <- dados_csn[[4]]

estatisticas_csn <- data.frame(
  Estatística = c("Média", "Mediana", "Desvio Padrão", "Curtose", "Assimetria", "Mínimo", "Máximo", "1º", "2º", "3º", "4º", "5º", "6º", "7º", "8º", "9º", "10º"),
  Valor = c(
    mean(precos_csn, na.rm = TRUE),
    median(precos_csn, na.rm = TRUE),
    sd(precos_csn, na.rm = TRUE),
    e1071::kurtosis(precos_csn, na.rm = TRUE),
    e1071::skewness(precos_csn, na.rm = TRUE),
    min(precos_csn, na.rm = TRUE),
    max(precos_csn, na.rm = TRUE),
    quantile(precos_csn, 0.25, na.rm = TRUE),
    quantile(precos_csn, 0.75, na.rm = TRUE)
  )
)

grid.draw(tableGrob(estatisticas_csn))
```

	Estatística	Valor
1	Média	13.609770
2	Mediana	11.385000
3	Desvio Padrão	8.687393
4	Curtose	4.839127
5	Assimetria	2.127917
6	Mínimo	2.740000
7	Máximo	51.770000
8	1º Quartil	8.580000
9	3º Quartil	14.945000

- As principais estatísticas descritivas:

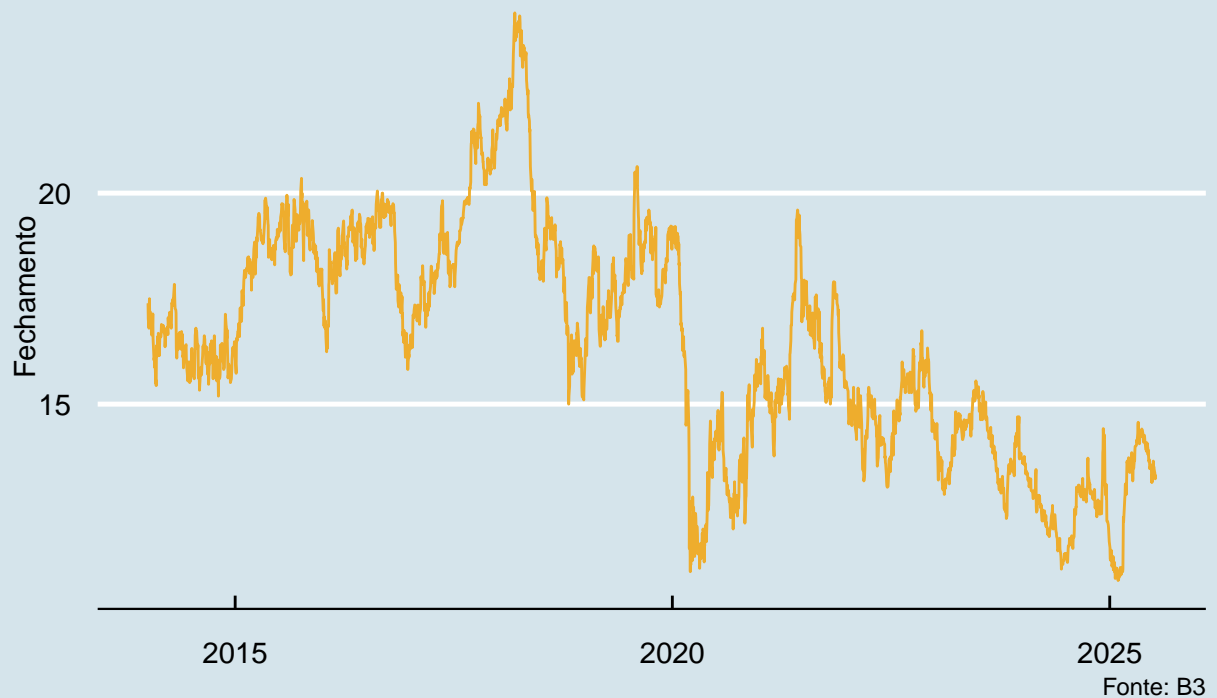
Os gráficos da Ambev:

```
graf_ambev_preco <- ggplot(dados_ambev, aes(x = ref.date, y = price.close)) +
  geom_line(color = '#eead2d') +
  labs(
    x = NULL, y = "Fechamento",
    title = "Histórico de Cotações - Ambev",
    subtitle = "De 02/01/2014 a 24/05/2024",
    caption = "Fonte: B3"
  ) +
  theme_economist()

graf_ambev_preco
```

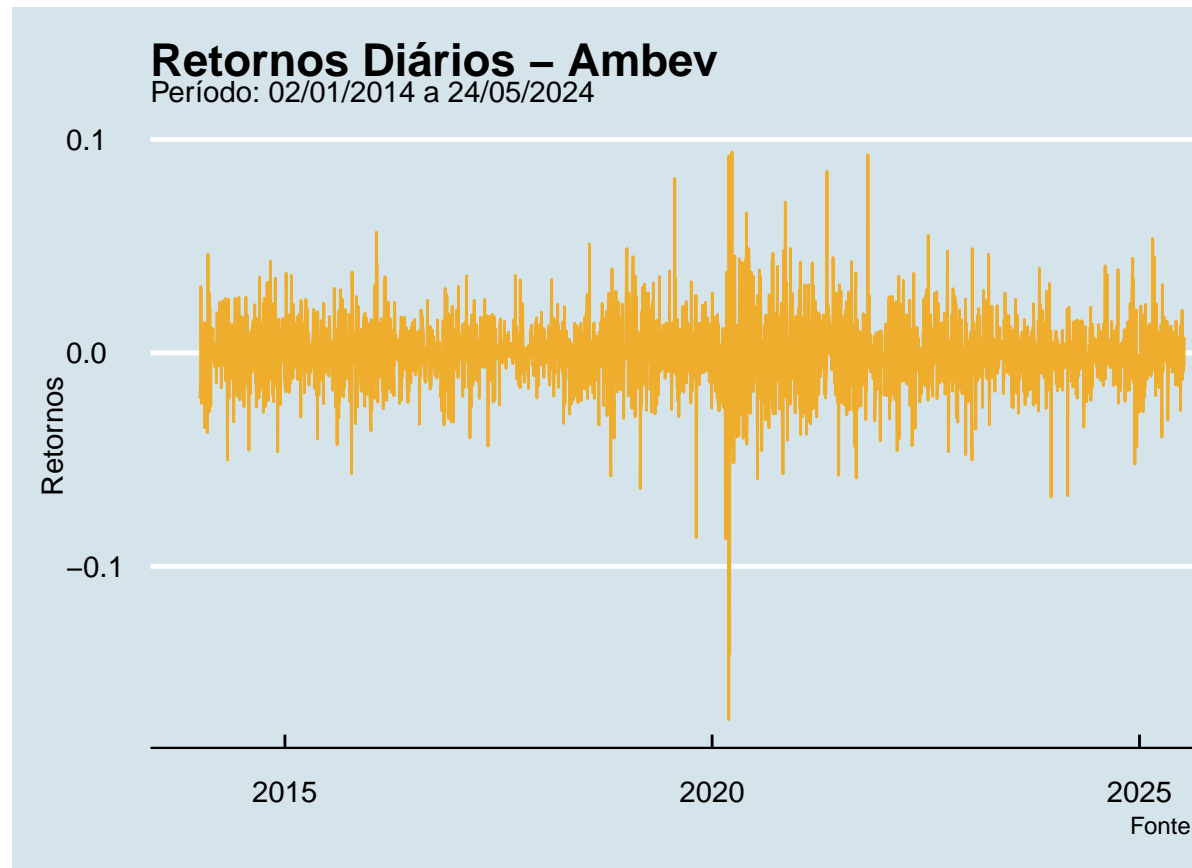
## Histórico de Cotações – Ambev

De 02/01/2014 a 24/05/2024



- Preços:

```
ret_ambev <- dados_ambev %>%  
  select(ref.date, ticker, ret.closing.prices)  
  
graf_ambev_retorno <- ggplot(ret_ambev, aes(x = ref.date, y = ret.closing.prices)) +  
  geom_line(color = '#eead2d') +  
  labs(  
    x = NULL, y = "Retornos",  
    title = "Retornos Diários - Ambev",  
    subtitle = "Período: 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_ambev_retorno
```



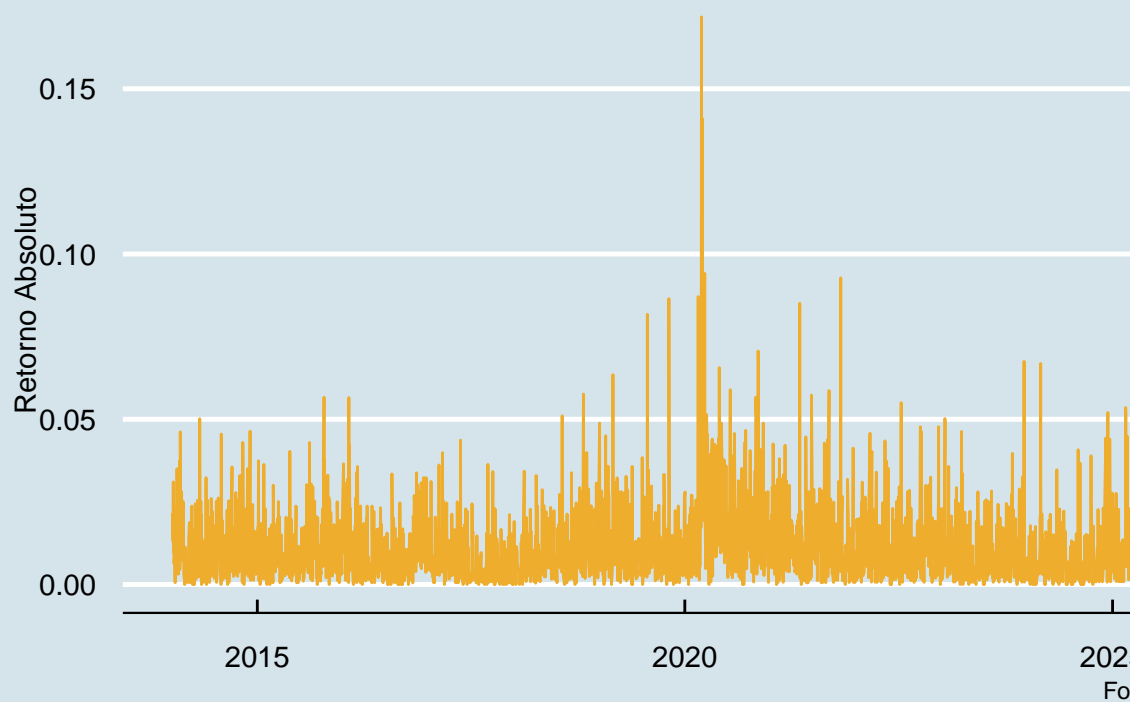
- Retornos diários

```
graf_ambev_abs <- ggplot(ret_ambev, aes(x = ref.date, y = abs(ret.closing.prices))) +  
  geom_line(color = '#eead2d') +  
  labs(  
    x = NULL, y = "Retorno Absoluto",  
    title = "Magnitude dos Retornos - Ambev",  
    subtitle = "Período de Análise: 02/01/2014 a 24/05/2024",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
graf_ambev_abs
```



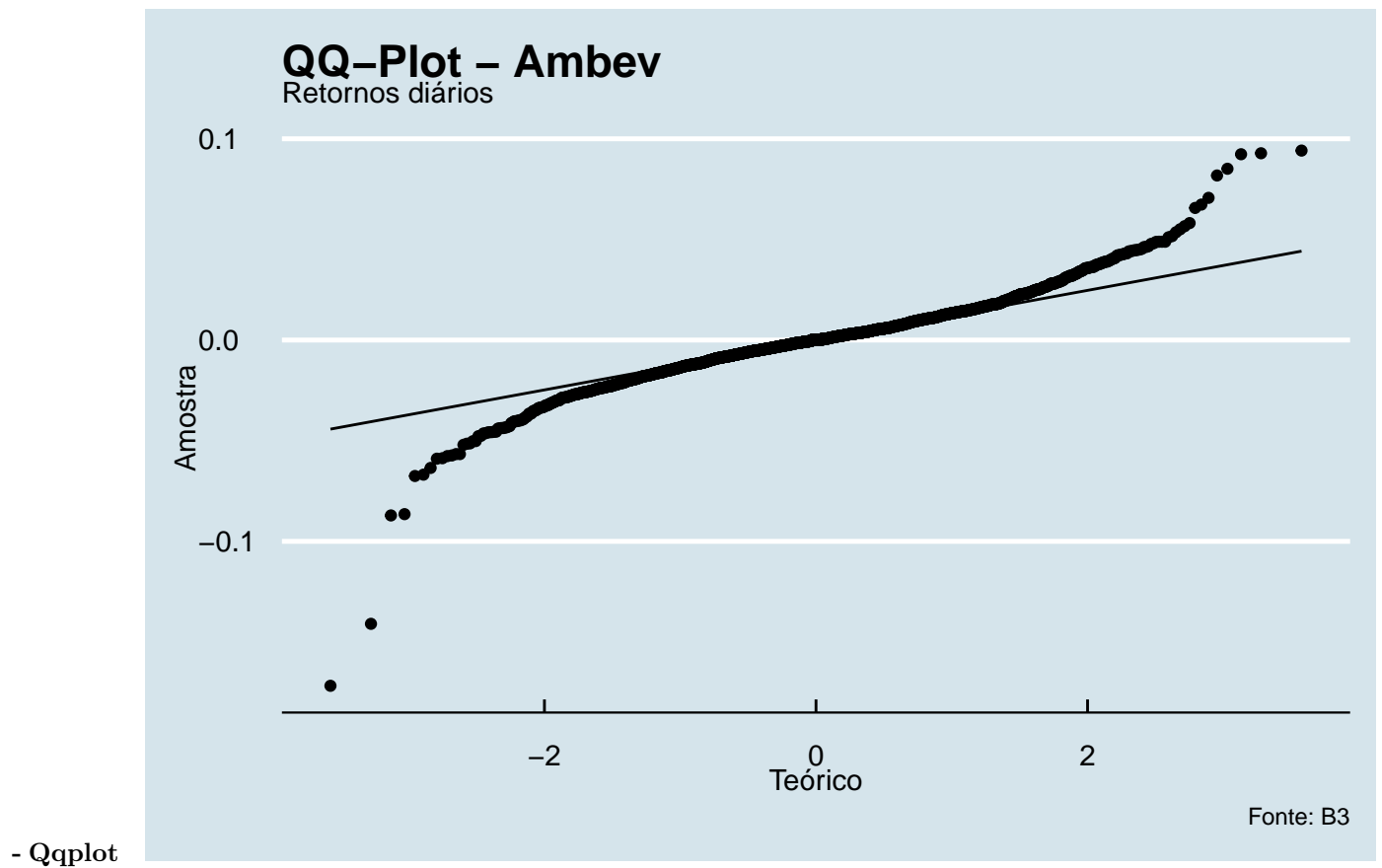
## Magnitude dos Retornos – Ambev

Período de Análise: 02/01/2014 a 24/05/2024



- Retornos absolutos

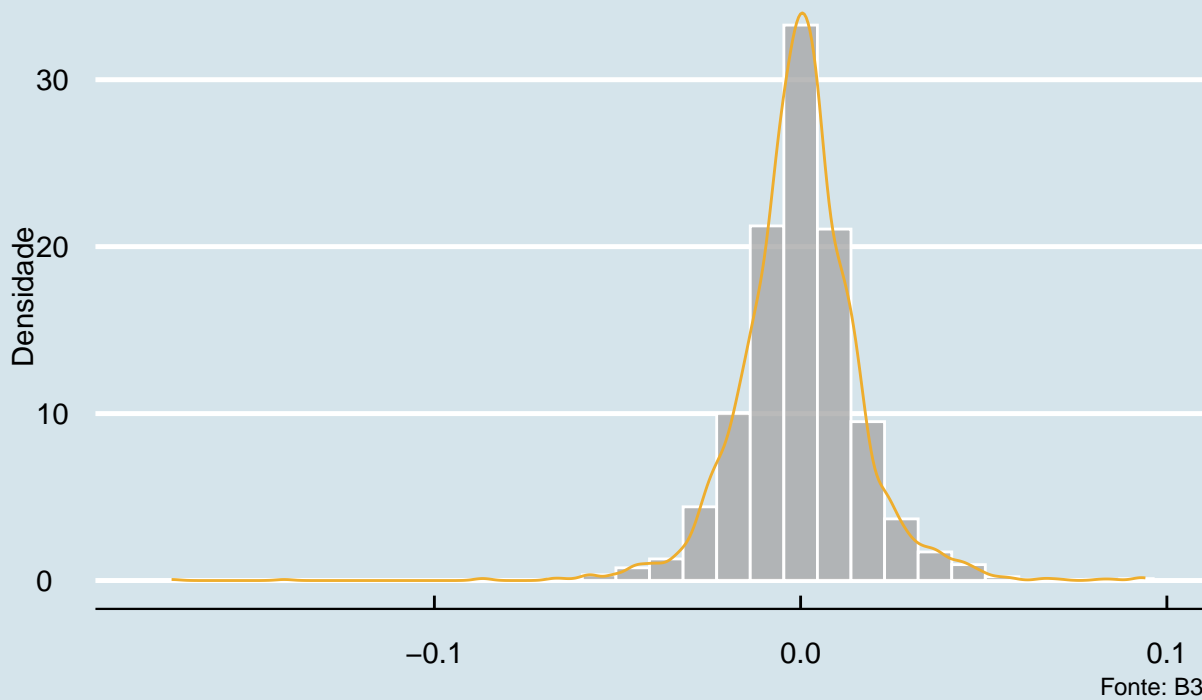
```
qqplot_ambev <- ggplot(ret_ambev, aes(sample = ret.closing.prices)) +  
  stat_qq() +  
  stat_qq_line() +  
  labs(  
    x = "Teórico", y = "Amostra",  
    title = "QQ-Plot - Ambev",  
    subtitle = "Retornos diários",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
qqplot_ambev
```



```
hist_ambev <- ggplot(ret_ambev) +  
  geom_histogram(aes(x = ret.closing.prices, y = ..density..),  
    fill = "darkgray", color = "white", alpha = 0.8) +  
  geom_density(aes(x = ret.closing.prices), color = "#eead2d") +  
  labs(  
    x = NULL, y = "Densidade",  
    title = "Distribuição de Retornos - Ambev",  
    subtitle = "Histograma com Densidade",  
    caption = "Fonte: B3"  
  ) +  
  theme_economist()  
  
hist_ambev
```

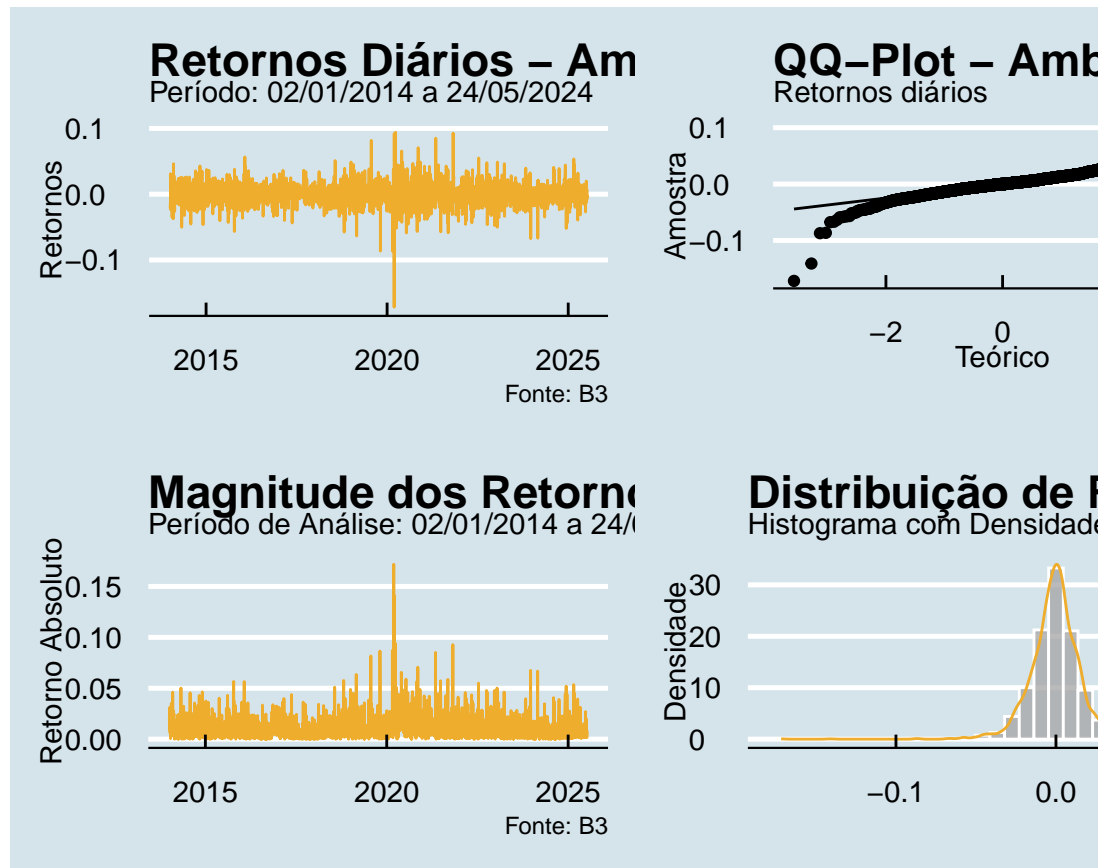
## Distribuição de Retornos – Ambev

Histograma com Densidade



- Histograma

```
cowplot::plot_grid(graf_ambev_retorno, qqplot_ambev, graf_ambev_abs, hist_ambev, nrow = 2)
```



- Todos os gráficos juntos

```
precos_ambev <- dados_ambev[[4]]

estatisticas_ambev <- data.frame(
  Estatística = c("Média", "Mediana", "Desvio Padrão", "Curtose", "Assimetria", "Mínimo", "Máximo", "1º", "2º", "3º", "4º", "5º", "6º", "7º", "8º", "9º", "10º"),
  Valor = c(
    mean(precos_ambev, na.rm = TRUE),
    median(precos_ambev, na.rm = TRUE),
    sd(precos_ambev, na.rm = TRUE),
    e1071::kurtosis(precos_ambev, na.rm = TRUE),
    e1071::skewness(precos_ambev, na.rm = TRUE),
    min(precos_ambev, na.rm = TRUE),
    max(precos_ambev, na.rm = TRUE),
    quantile(precos_ambev, 0.25, na.rm = TRUE),
    quantile(precos_ambev, 0.75, na.rm = TRUE)
  )
)

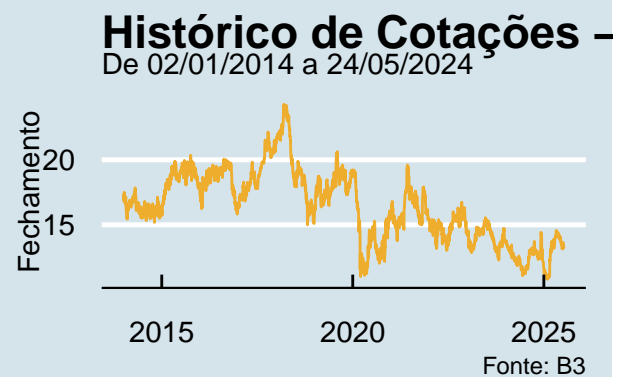
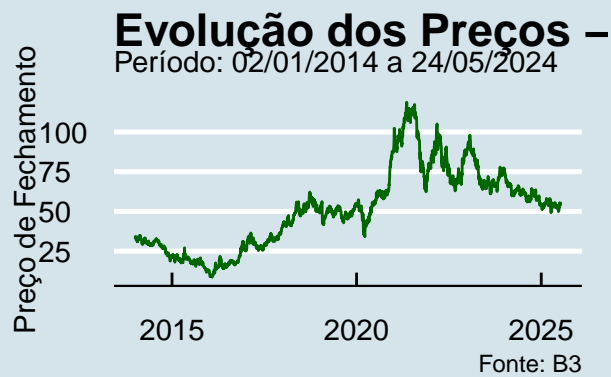
grid.draw(tableGrob(estatisticas_ambev))
```

	Estatística	Valor
1	Média	16.4092213
2	Mediana	16.3999996
3	Desvio Padrão	2.7370048
4	Curtose	-0.5286237
5	Assimetria	0.1393175
6	Mínimo	10.8199997
7	Máximo	24.2700005
8	1º Quartil	14.2399998
9	3º Quartil	18.5699997

- As principais estatísticas descritivas:

Agora a comparação entre cada empresa:

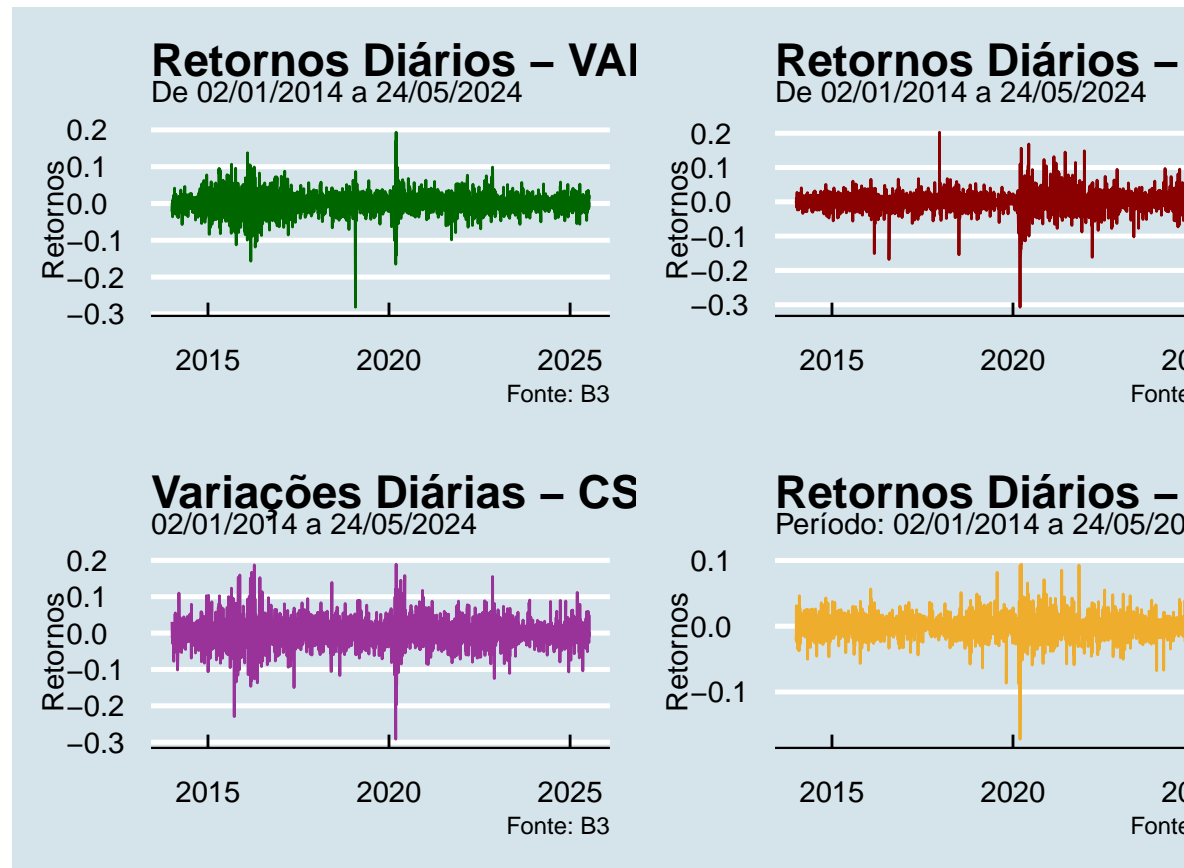
```
# Preços
cowplot::plot_grid(graf_vale_preco, graf_embraer_preco, graf_csn_preco, graf_ambev_preco, nrow = 2)
```



- Preços:

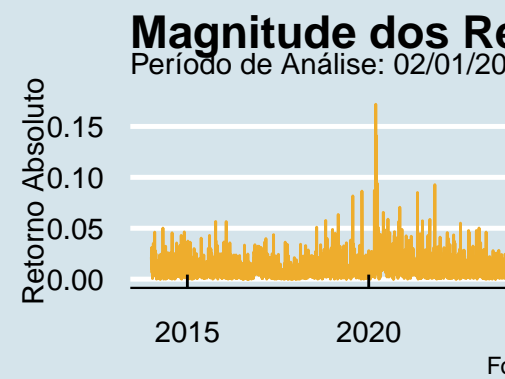
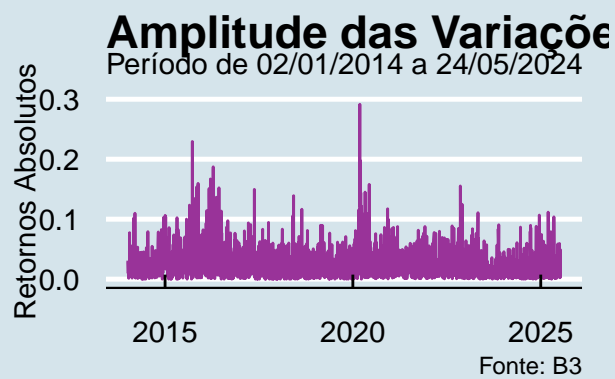
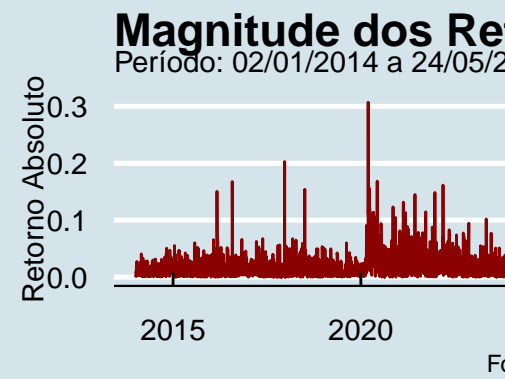
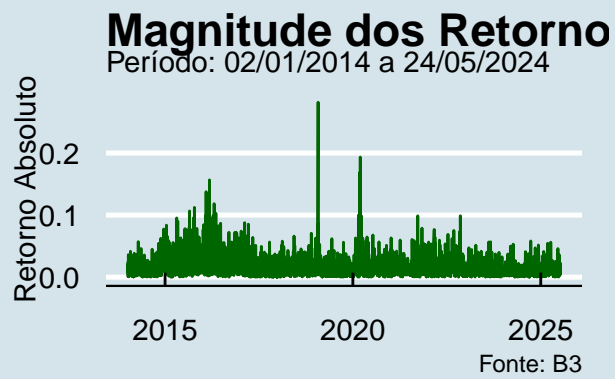
*# Retornos*

`cowplot::plot_grid(graf_vale_retorno, graf_embraer_retorno, graf_csn_retorno, graf_ambev_retorno, nrow = 2)`



- Retornos diários:

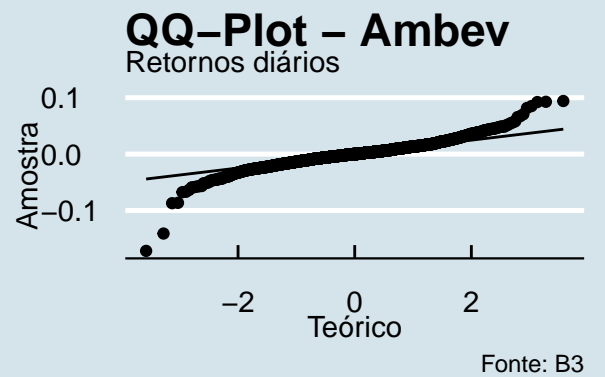
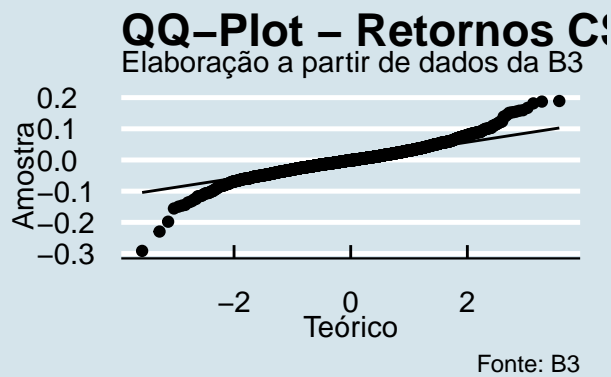
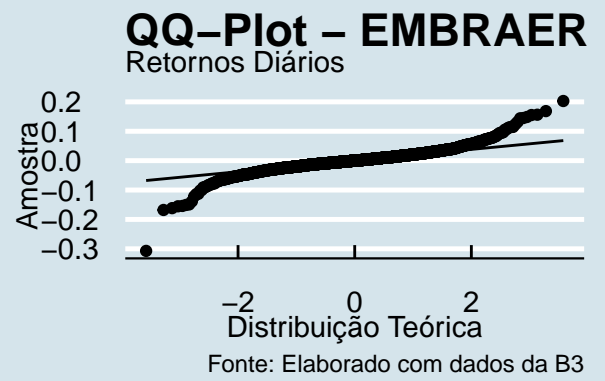
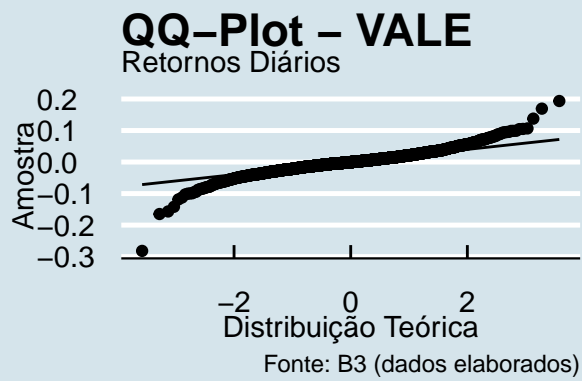
```
# Retornos absolutos
cowplot::plot_grid(graf_vale_abs, graf_embraer_abs, graf_csn_abs, graf_ambev_abs, nrow = 2)
```



- Retornos absolutos:

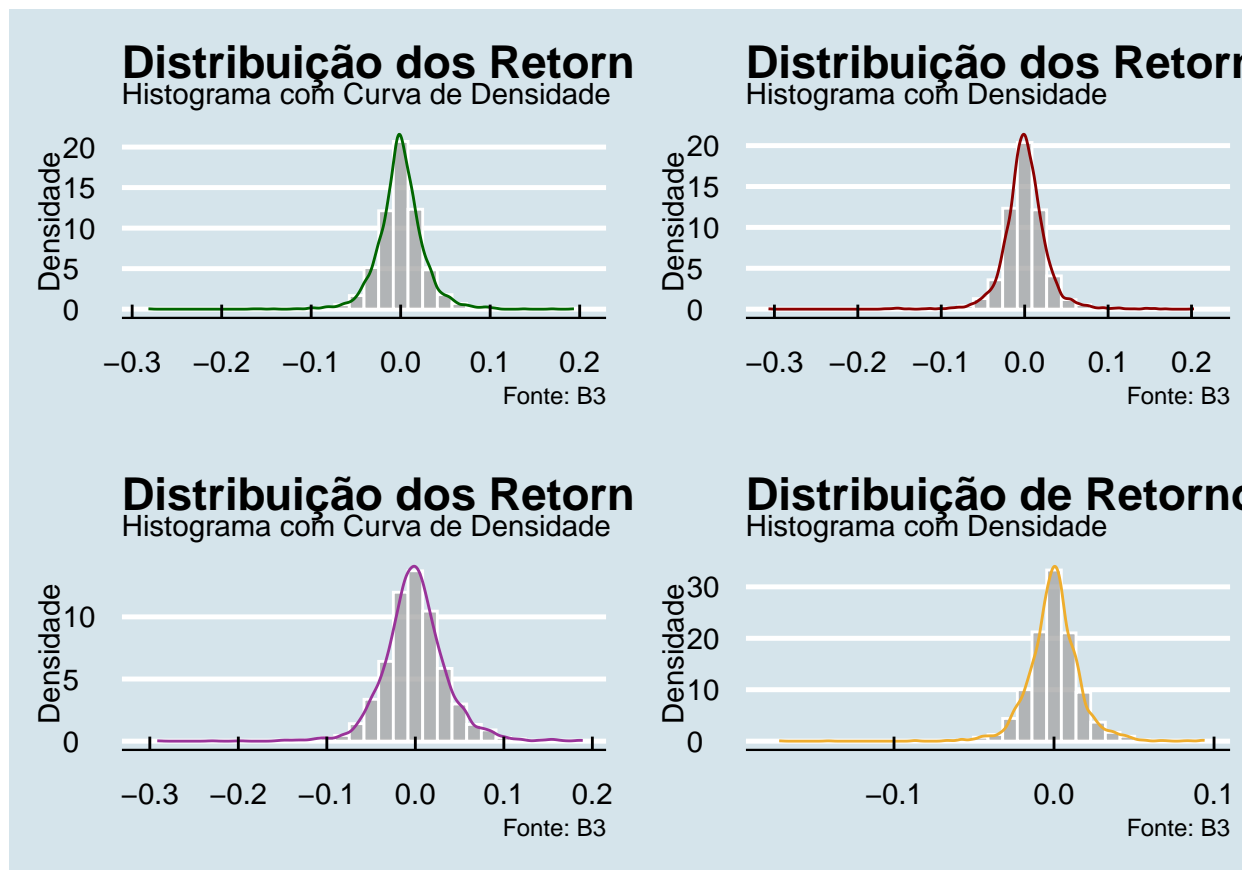
```
# QQ-Plots
cowplot::plot_grid(qqplot_vale, qqplot_embraer, qqplot_csn, qqplot_ambev, nrow = 2)
```





- Qqplot:

```
# Histogramas
cowplot::plot_grid(hist_vale, hist_embraer, hist_csn, hist_ambev, nrow = 2)
```



- Histograma:

Por fim podemos Ver a dispersão entre cada uma das empresas:

```
# Construção do dataframe de retornos
df_retornos <- data.frame(
  vale = ret_vale$ret.closing.prices,
  csu = ret_csu$ret.closing.prices,
  embraer = ret_embraer$ret.closing.prices,
  ambev = ret_ambev$ret.closing.prices
)

# Matriz de correlação
matriz_cor <- cor(df_retornos, use = "complete.obs")

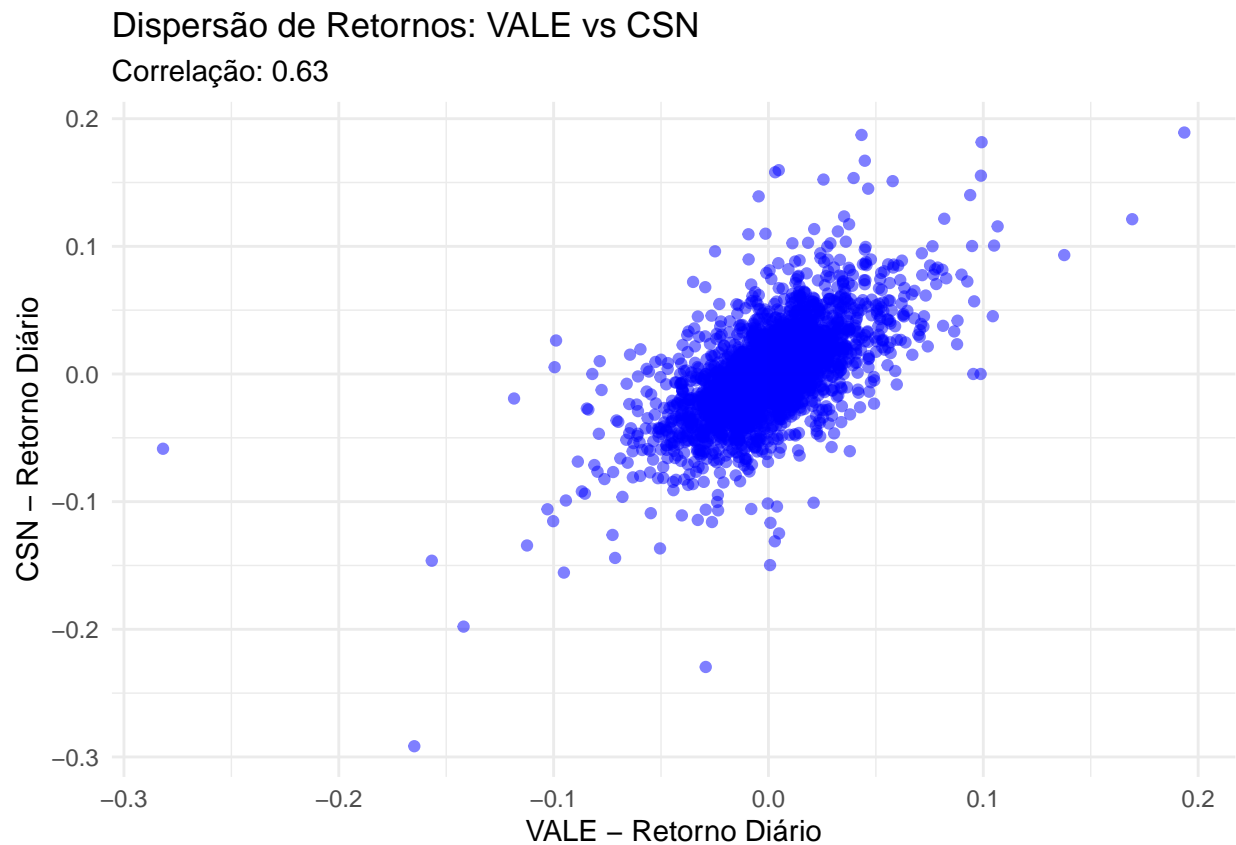
# Conversão para tabela
tabela_cor <- as.data.frame(matriz_cor)
tabela_cor <- cbind(Empresa = rownames(tabela_cor), tabela_cor)

# Exibir tabela gráfica
grid.draw(tableGrob(tabela_cor, rows = NULL))
```

Empresa	vale	csn	embraer	ambev
vale	1.0000000	0.6335523	0.2129896	0.2141344
csn	0.6335523	1.0000000	0.2373670	0.2778134
embraer	0.2129896	0.2373670	1.0000000	0.3017359
ambev	0.2141344	0.2778134	0.3017359	1.0000000

O gráfico de Dispersão entre a Vale e a Siderurgica por elas possuírem a maior coorelação:

```
ggplot(df_retornos, aes(x = vale, y = csn)) +
  geom_point(color = "blue", alpha = 0.5) +
  labs(
    title = "Dispersão de Retornos: VALE vs CSN",
    subtitle = paste("Correlação:", round(matriz_cor["vale", "csn"], 2)),
    x = "VALE - Retorno Diário",
    y = "CSN - Retorno Diário"
  ) +
  theme_minimal()
```



E o Heat Map entre todas as variaveis

```
library(reshape2)

# Derretendo a matriz
melt_cor <- melt(matriz_cor)

# Plot
ggplot(melt_cor, aes(Var1, Var2, fill = value)) +
  geom_tile(color = "white") +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
                      midpoint = 0, limit = c(-1, 1), space = "Lab",
                      name = "Correlação") +
  theme_minimal() +
  labs(
    title = "Mapa de Calor - Correlações entre Retornos",
    x = "Empresa",
    y = "Empresa"
  ) +
  theme(axis.text.x = element_text(angle = 0, vjust = 1, hjust = 1))
```

