

# Matheus3031

Matheus Freitas Martins

17/08/2021

## Relatório de Estatística Básica

Considere as medidas de temperatura de três cidades ao longo de 200 dias. Conforme apresentado abaixo:

```
set.seed(3031)#Acrescentar a matrícula como semente
#Dataframe com as medidas de temperatura das três cidades
df <- data.frame(x1=rnorm(200,30,10),
                  x2=rnorm(200,15,5),
                  x3=rnorm(200,0,15))
```

Observem que temos 3 colunas, cada coluna representa as temperaturas de uma cidade.

## Cálculo da média, mediana, variância, desvio padrão, coeficiente de variação e quartis para as 3 cidades.

### Cidade 1

```
mean(df$x1)#calculando a média
```

```
## [1] 30.01896
```

```
median(df$x1)#calculando a mediana
```

```
## [1] 30.25036
```

```
var(df$x1)#calculando a variancia
```

```
## [1] 84.18031
```

```
sd(df$x1)#calculando o desvio padrão
```

```
## [1] 9.174983
```

```
sd(df$x1)/mean(df$x1) #calculando o coeficiente de variação
```

```
## [1] 0.3056396
```

```
quantile(df$x1, probs = 0.25)#Quartis 1
```

```
##      25%  
## 23.5686
```

```
quantile(df$x1, probs = 0.50)#Quartis 2
```

```
##      50%  
## 30.25036
```

```
quantile(df$x1, probs = 0.75)#Quartis 3
```

```
##      75%  
## 36.19219
```

Observando a média e a mediana da primeira cidade, pode-se afirmar que nos 200 dias observados a temperatura média ficou aproximadamente em torno de 30°C. Como os valores da média e mediana foram próximos este condiz com essas observações.

Olhando para o desvio padrão (9,17) podemos afirmar que a menor temperatura foi de aproximadamente 21°C e a máxima foi de 39°C.

Por meio dos quartis, nota-se que nos primeiros dias apresentou-se uma baixa temperatura (23,5°C), o que possivelmente indica um clima de inverno. Por outro lado, na metade dos dias houve um significativo aumento da temperatura (30,25°C), o que aponta a chegada da temporada de primavera. Ademais, os últimos dias observados apresentam uma grande alta na temperatura (36,19°C), permitindo concluir que estava na temporada de verão.

## Cidade 2

```
mean(df$x2)#calculando a média
```

```
## [1] 14.9647
```

```
median(df$x2)#calculando a mediana
```

```
## [1] 14.71658
```

```
var(df$x2)#calculando a variancia
```

```
## [1] 24.58222
```

```
sd(df$x2)#calculando o desvio padrão
```

```
## [1] 4.958046
```

```
sd(df$x2)/mean(df$x2) #calculando o coeficiente de variação
```

```
## [1] 0.3313162
```

```
quantile(df$x2, probs = 0.25)#Quartis 1
```

```
##      25%  
## 11.72758
```

```
quantile(df$x2, probs = 0.50)#Quartis 2
```

```
##      50%  
## 14.71658
```

```
quantile(df$x2, probs = 0.75)#Quartis 3
```

```
##      75%  
## 18.68373
```

Observando a média e a mediana da segunda cidade, pode-se afirmar que nos 200 dias observados a temperatura média ficou aproximadamente em torno de 14°C. Como os valores da média e mediana foram próximos este condiz com essas observações.

Olhando para o desvio padrão (4,9) podemos afirmar que a menor temperatura foi de aproximadamente 10°C e a máxima foi de 19,8°C.

Por meio dos quartis, nota-se que nos primeiros dias apresentou-se uma baixa temperatura (11,7°C), o que possivelmente indica um clima de inverno. Por outro lado, na metade dos dias houve um significativo aumento da temperatura (14,7°C), o que aponta a chegada da temporada de primavera. Ademais, os últimos dias observados apresentam uma pequena alta na temperatura (18,68°), permitindo concluir a chegada gradativa da temporada de verão.

## Cidade 3

```
mean(df$x3)#calculando a média
```

```
## [1] -0.2786445
```

```
median(df$x3)#calculando a mediana
```

```
## [1] 1.458539
```

```
var(df$x3)#calculando a variancia
```

```
## [1] 269.3661
```

```
sd(df$x3)#calculando o desvio padrão
```

```
## [1] 16.41238
```

```
sd(df$x3)/mean(df$x3) #calculando o coeficiente de variação
```

```
## [1] -58.90077
```

```
quantile(df$x3, probs = 0.25)#Quartis 1
```

```
##      25%  
## -10.48415
```

```
quantile(df$x3, probs = 0.50)#Quartis 2
```

```
##      50%  
##  1.458539
```

```
quantile(df$x3, probs = 0.75)#Quartis 3
```

```
##      75%  
## 10.38523
```

Como a média é sensível a valores extremos, no caso da cidade 3, o valor que melhor descreve a temperatura média observada nos 200 dias é a mediana que é de 1,45°C

Olhando para o desvio padrão (16,4) podemos afirmar que a menor temperatura foi de aproximadamente -16,61°C e a máxima foi de 16,21°C.

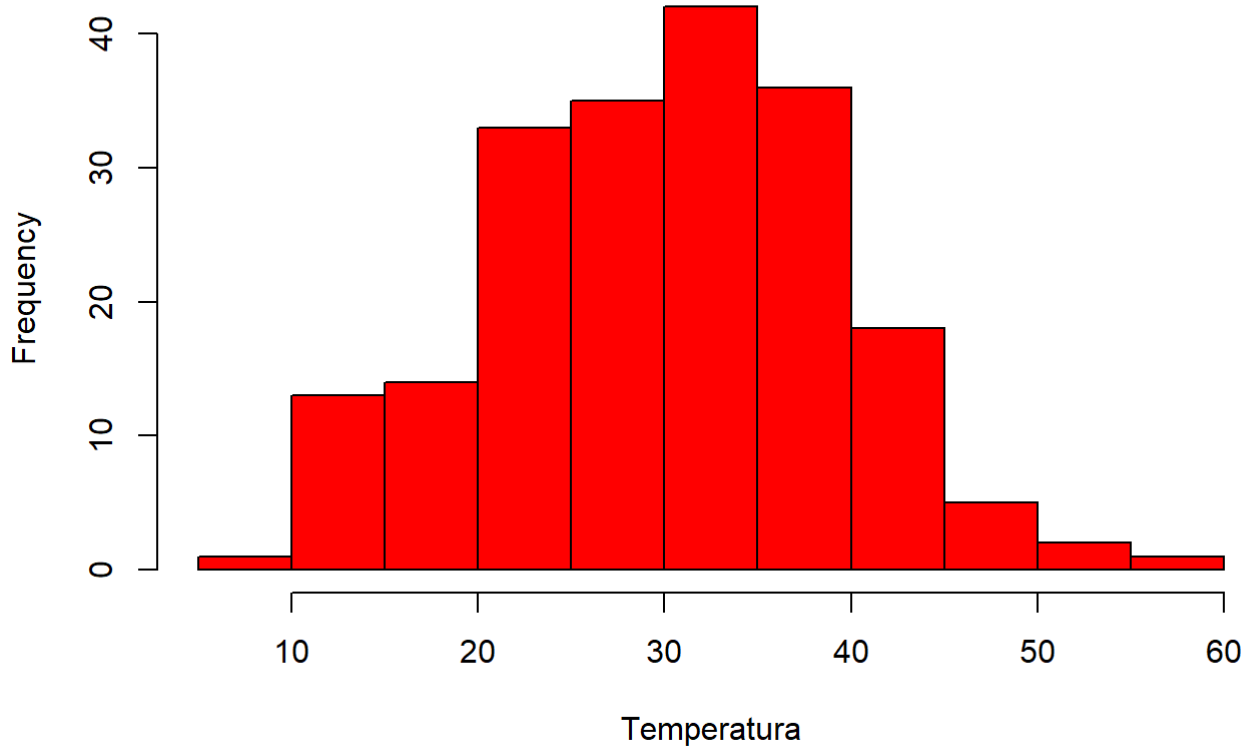
Por meio dos quartis, nota-se que nos primeiros dias apresentou-se uma temperatura negativa (-10,4°C), o que possivelmente indica um rigoroso clima de inverno. Por outro lado, na metade dos dias houve um significativo aumento da temperatura (1,45°C), o que aponta a chegada da temporada de primavera. Ademais, os últimos dias observados apresentam uma pequena alta na temperatura (10,68°), permitindo concluir a chegada gradativa da temporada de verão.

Por fim, com os dados obtidos até aqui conclui-se que houve uma maior variabilidade nas temperaturas da terceira cidade, sendo ela de 58%.

## Histograma da cidade 1

```
hist(df$x1, main = "Histograma da Cidade 1", xlab = "Temperatura", col="red")
```

## Histograma da Cidade 1

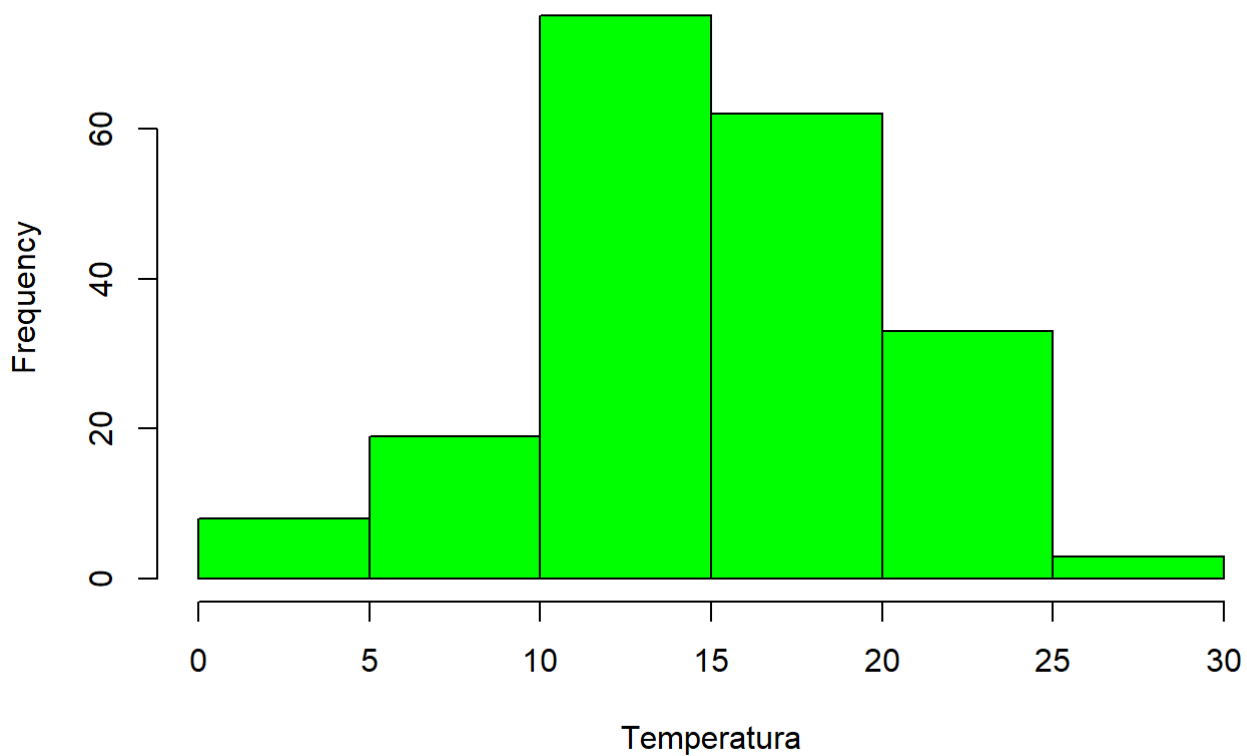


Conforme observado anteriormente com os cálculos da cidade 1, pode-se constatar realmente uma grande predominância de temperaturas que giram em torno de 20°C, 30°C e 40°C. Sendo esta considerada a cidade mais quente na maior parte dos dias.

## Histograma da cidade 2

```
hist(df$x2, main = "Histograma da Cidade 2", xlab = "Temperatura", col="green")
```

## Histograma da Cidade 2

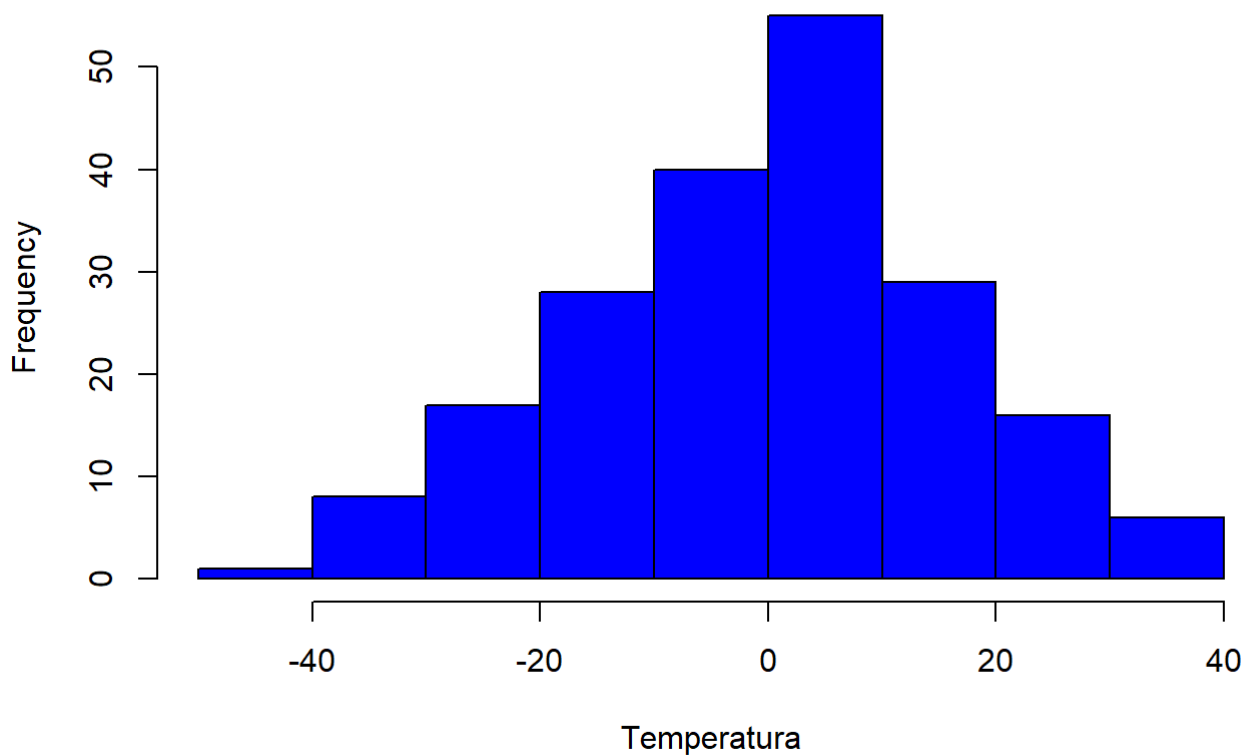


Conforme observado anteriormente com os cálculos da cidade 2, pode-se constatar realmente uma grande predominância de temperaturas que giram em torno de 10°C, 15°C e 20°C. Sendo esta a cidade com o clima mais moderado.

## Histograma da cidade 3

```
hist(df$x3, main = "Histograma da Cidade 3", xlab = "Temperatura", col="blue")
```

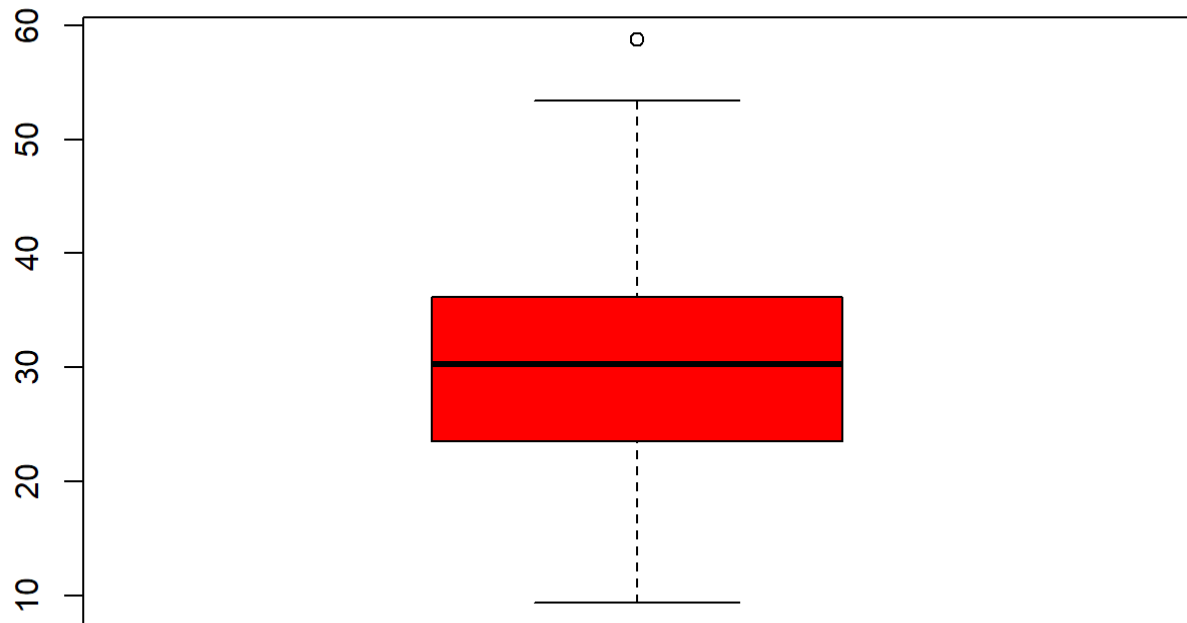
## Histograma da Cidade 3



Conforme observado anteriormente com os cálculos da cidade 3, é notável de que se trata de uma cidade relativamente mais fria do que as demais na maior parte dos dias, possuindo uma maior frequência que gira em torno de 0°C a 10°C, e nota-se também que possui uma maior variabilidade em relação as temperaturas.

## Boxplot da cidade 1

```
boxplot(df$x1, col="red")
```

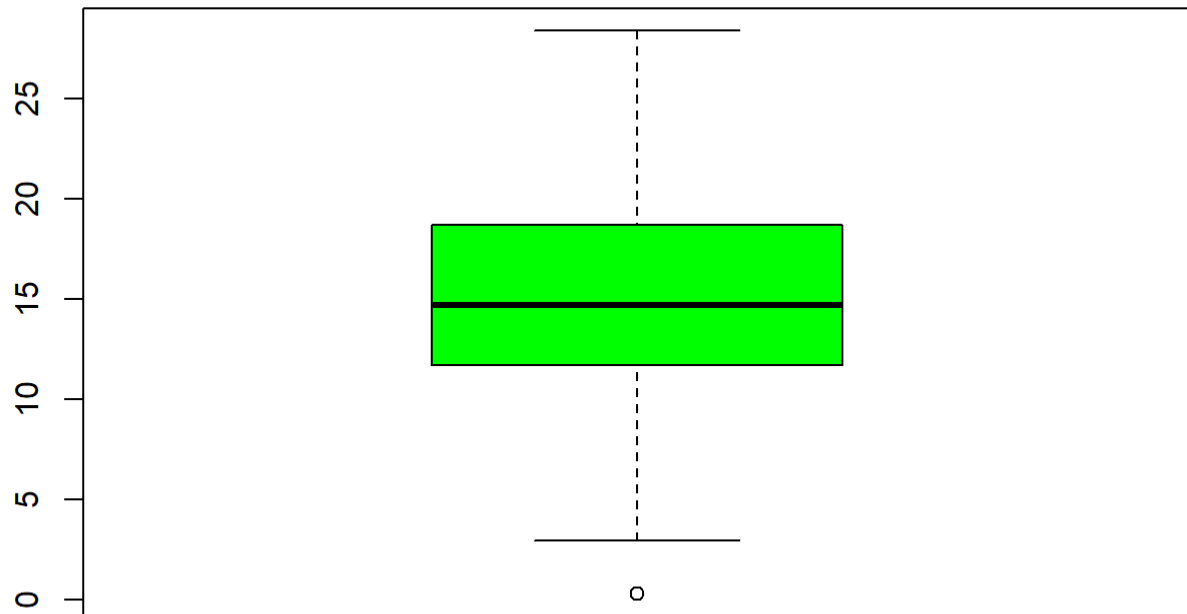


Observando o respectivo boxplot da cidade 1, é visto que existe um outlier de temperatura elevada. Além disso, vale dizer que os dados observados estão mais condensados aproximadamente entre 37° e 23°C

## Boxplot da cidade 2

```
boxplot(df$x2, col="green")
```

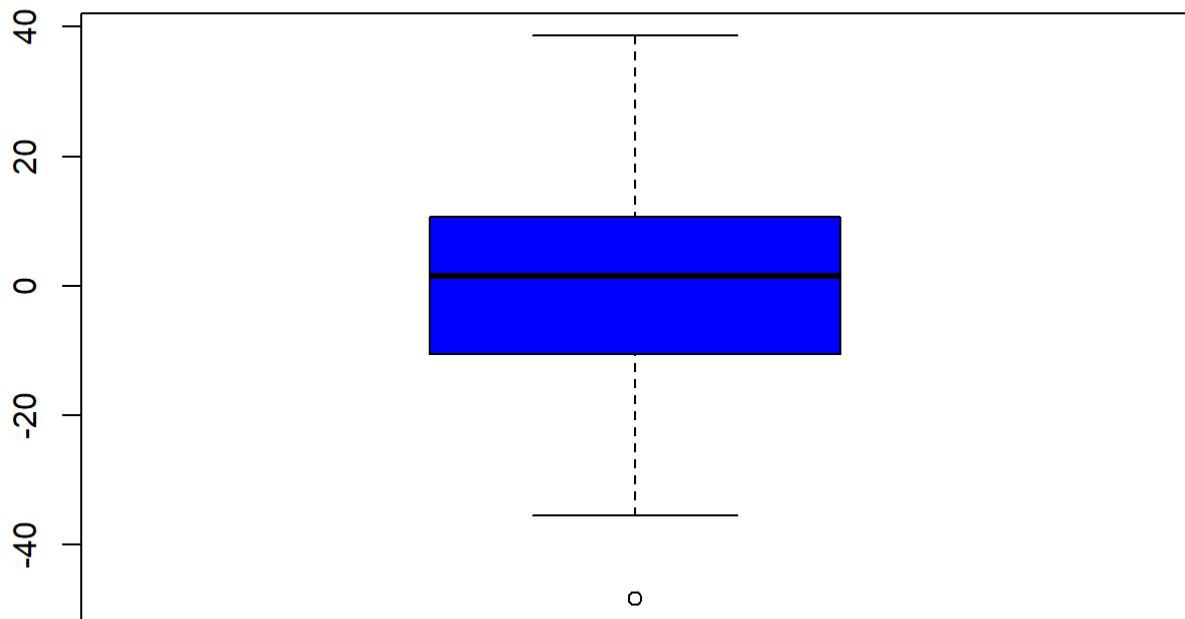




Observando o respectivo boxplot da cidade 2, é visto que existe um outlier de temperatura inferior. Além disso, vale dizer que os dados observados estão mais condensados aproximadamente entre 18° e 13°C

## Boxplot da cidade 3

```
boxplot(df$x3, col="blue")
```



Observando o respectivo boxplot da cidade 3, é visto que existe um outlier de temperatura inferior. Além disso, vale dizer que os dados observados estão mais condensados aproximadamente entre  $-10^{\circ}$  e  $15^{\circ}\text{C}$

## Assimetria dos dados

Se o coeficiente de assimetria for maior do que zero, indica que a maioria dos valores está concentrado no lado esquerdo do gráfico.

Se o coeficiente de assimetria for igual ou próximo de zero, indica que os dados são normalmente distribuídos.

Se o coeficiente de assimetria for menor do que zero, indica que a maioria dos valores está concentrado no lado direito do gráfico.

Cidade 1

```
library(moments) #utilizando a biblioteca "moments"
skewness(df$x1) #Assimetria dos dados da cidade 1
```

```
## [1] 0.06818247
```

Observando o valor resultante da função skewness e o histograma da cidade 1 é possível observar que a maioria dos valores está concentrada do lado esquerdo do gráfico.

Cidade 2

```
library(moments) #utilizando a biblioteca "moments"
skewness(df$x2) #Assimetria dos dados da cidade 2
```

```
## [1] -0.1402911
```

Observando o valor resultante da função skewness e o histograma da cidade 2 é possível observar que a maioria dos valores está concentrada ao lado direito.

Cidade 3

```
library(moments) #utilizando a biblioteca "moments"  
skewness(df$x3) #Assimetria dos dados da cidade 3
```

```
## [1] -0.1127597
```

Observando o valor resultante da função skewness e o histograma da cidade 3 é possível observar que a maioria dos valores está concentrada ao lado direito.