

Pergunta 3

Código:

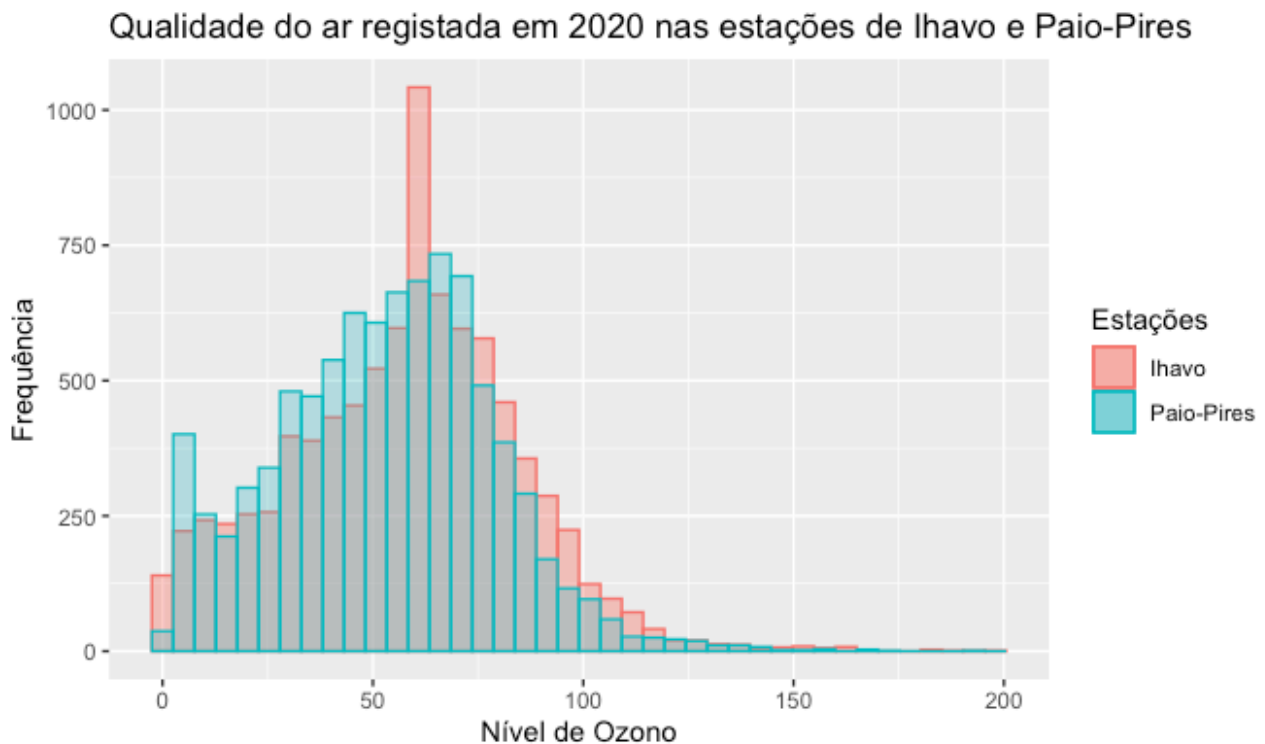
```
library(ggplot2)
library(readxl)
QualidadeARO3 <- read_excel("QualidadeARO3.xlsx");
lhavo = as.numeric(QualidadeARO3$lhavo);
Paio_Pires = as.numeric(QualidadeARO3$Paio-Pires)

data = data.frame(lhavo, Paio_Pires)

Estações = c(lhavo, Paio_Pires)

ggplot(data, aes(fill=Estações, color=Estações)) +
  geom_histogram(aes(x = lhavo, fill = "lhavo", color="lhavo"), bins = 40, alpha=0.4) +
  geom_histogram(aes(x = Paio_Pires, fill = "Paio-Pires", color="Paio-Pires"), bins = 40, alpha=0.3) +
  labs(title = "Qualidade do ar registada em 2020 nas estações de lhavo e Paio-Pires",
       x = "Nível de Ozono", y = "Frequência")
```

Gráfico obtido:



Comentário:

O histograma apresenta as frequências dos valores de Nível de Ozono nas estações de Paio-Pires e lhavo no ano 2020.

Analisando o histograma reparamos que Paio-Pires “domina” quando os Níveis de Ozono se compreendem entre os valores 5 e 60 utilizadores (com uma exceção não muito significativa), e entre os valores 65 e 75, enquanto que lhavo domina quando os Níveis de Ozono se compreendem entre os valores 75 e 120. lhavo tem no entanto muitos registos apresentados entre 60 e 65 (mais de 1000 registos, o que significa que mais de 10% dos registos ficam compreendidos entre estes valores, tendo em conta que existiram no total 8784 registos).

Em geral, a estação de lhavo aparenta ter maiores Níveis de Ozono que a estação de Paio-Pires.