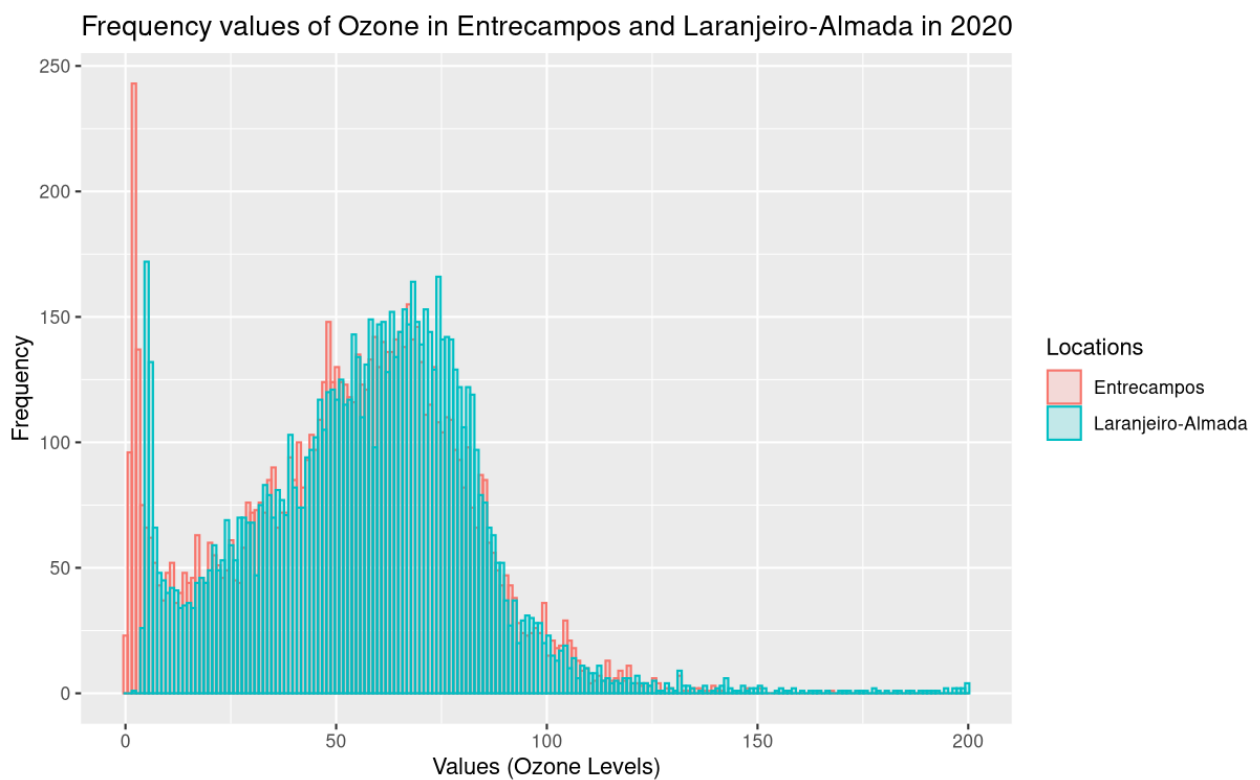


```
1 library(readxl)
2 library(ggplot2)
3
4 Values <- read_excel ("data/QualidadeARO3.xlsx")
5
6 Almada <- Values[, "Laranjeiro-Almada"]
7 Entrecampos <- Values[, "Entrecampos"]
8 colnames(Almada) <- c("Valores")
9 colnames(Entrecampos) <- c("Valores")
10
11 Entrecampos$Valores <- as.numeric(Entrecampos$Valores)
12
13 m <- rbind(Almada, Entrecampos)
14 df <- data.frame("Locations" = rep(c("Laranjeiro-Almada", "Entrecampos"), each=8784), "Valores" = m)
15
16 ggplot(df, aes(x=Valores, color=Locations, fill=Locations)) + geom_histogram(bins=200, alpha=0.2, position="identity")+
17   labs(title="Frequency values of Ozone in Entrecampos and Laranjeiro-Almada in 2020", x="Values (Ozone Levels)", y = "Frequency")
18
```



Pela observação do gráfico podemos reparar que em Entrecampos temos uma grande quantidade de dados com valor perto da casa das unidades. A partir da 1ª dezena vemos a frequência dos valores a aumentar até ao valor 75 onde atinge o pico. A partir do pico a frequência diminui rapidamente.

Para Laranjeiro-Almada vemos uma situação semelhante onde a única diferença encontra-se numa maior frequência de valores de ozono mais altos.