

Exercício 2

Nota: Para minimizar erros, convém garantir que os ficheiros usados como base de dados (HFC.xlsx) estão na mesma pasta que o ficheiro com o código.

```
#bibliotecas

library(tidyverse)
library(openxlsx)
library(dplyr)
library(tidyr)
library(reshape)

#importar dados

hfc <- read.xlsx("HFC.xlsx")
names(hfc)[1] <- 'dummy'

g <- hfc %>% separate(dummy, c("Idade","Anemia", "C.F.", "Diabetes", "F.E.",
                             "Hipertensão Arterial", "Plaquetas", "C.S", "Sódio",
                             "Sexo", "Fumador", "Tempo", "Morte"), sep=",")

grafico <- select(g, c('Hipertensão Arterial', 'Sódio'))

#long list to wide list

grafico.wide <- pivot_wider(grafico, names_from="Hipertensão Arterial",
                           values_from="Sódio")

#separar em vetores com hipertensão 1 e 0

sodio1=as.numeric(unlist(grafico.wide$`1`))
sodio0=as.numeric(unlist(grafico.wide$`0`))

#desenhar o grafico
#parametros esteticos standard do hist() pareceram aceitáveis

hist(sodio1, main="Com hipertensão arterial",
     xlab="Quantidade de sódio no sangue", xlim=c(110,150),
     ylab="Frequência dos valores",breaks=10)
mtext("Presença de sódio em pessoas com ou sem pressão arterial", side = 3, line=-1,
     outer = TRUE)#breaks 10 e 17 é para deixar os histogramas +- semelhantes
hist(sodio0, main="Sem hipertensão arterial",
     xlab="Quantidade de sódio no sangue",xlim=c(110,150),
     ylab="Frequência dos valores",breaks=17)
```

Comentários

Os valores médios de ambos os casos aproximam-se do intervalo de valores de sódio no sangue considerados normais (135-145), o que leva a crer que a hipertensão arterial não está diretamente associada aos níveis de sódio no sangue. Esta conclusão pode perfeitamente não ser válida: por um lado a amostra recolhida é muito pequena (299 dados recolhidos), por outro não existe um número igual de amostras nos dois casos, havendo mais amostras de pessoas sem hipertensão do que com hipertensão.