

```
pergunta3 > code.r
19 # Para criar uma pergunta em que analisa o código e o gráfico produzido
20 #pacotes
21 library(ggplot2)
22 #ler o ficheiro
23 dados <- read.csv("pergunta3/clima.csv", header = TRUE, sep = ",")
24 #converter a coluna Data para o formato datetime
25 dados$Data <- as.POSIXct(dados$Data, format = "%Y-%m-%d %H:%M:%S")
26 #filtrar só setembro de 2014
27 dados_set2014 <- dados[
28   format(dados$Data, "%Y") == "2014" & format(dados$Data, "%m") == "09",
29 ]
30 #criar coluna com dia e sem horas
31 dados_set2014$dia <- as.Date(dados_set2014$Data)
32 dados_set2014 <- dados_set2014[!is.na(dados_set2014$Poluição), ] #para tirar os com valores NA
33 #calcular mediana diária de Poluição
34 mediana_diaria <- aggregate(
35   Poluição ~ dia,
36   data = dados_set2014,
37   FUN = function(x) median(x, na.rm = TRUE)
38 )
39 #converter dia para POSIXct
40 mediana_diaria$dia <- as.POSIXct(mediana_diaria$dia)
41 #criar o gráfico
42 Sys.setlocale("LC_TIME", "pt_PT.UTF-8") #para meter em português
43 ggplot() +
44   geom_line(data = dados_set2014, aes(x = Data, y = Poluição), color = "blue", alpha = 0.5) +
45   geom_line(data = mediana_diaria, aes(x = dia, y = Poluição), color = "red", linewidth = 1) +
46   scale_x_datetime(date_labels = "%b %d") +
47   labs(
48     title = "Variação horária da Poluição em Setembro de 2014",
49     x = "Data e hora",
50     y = "Poluição",
51     caption = "Linha azul: variação horária | Linha vermelha: mediana diária"
52   ) +
53   theme_minimal()
54 ggsave("pergunta3/grafico.png", width = 8, height = 5)
```

Variação horária da Poluição em Setembro de 2014

