

Exercício 4

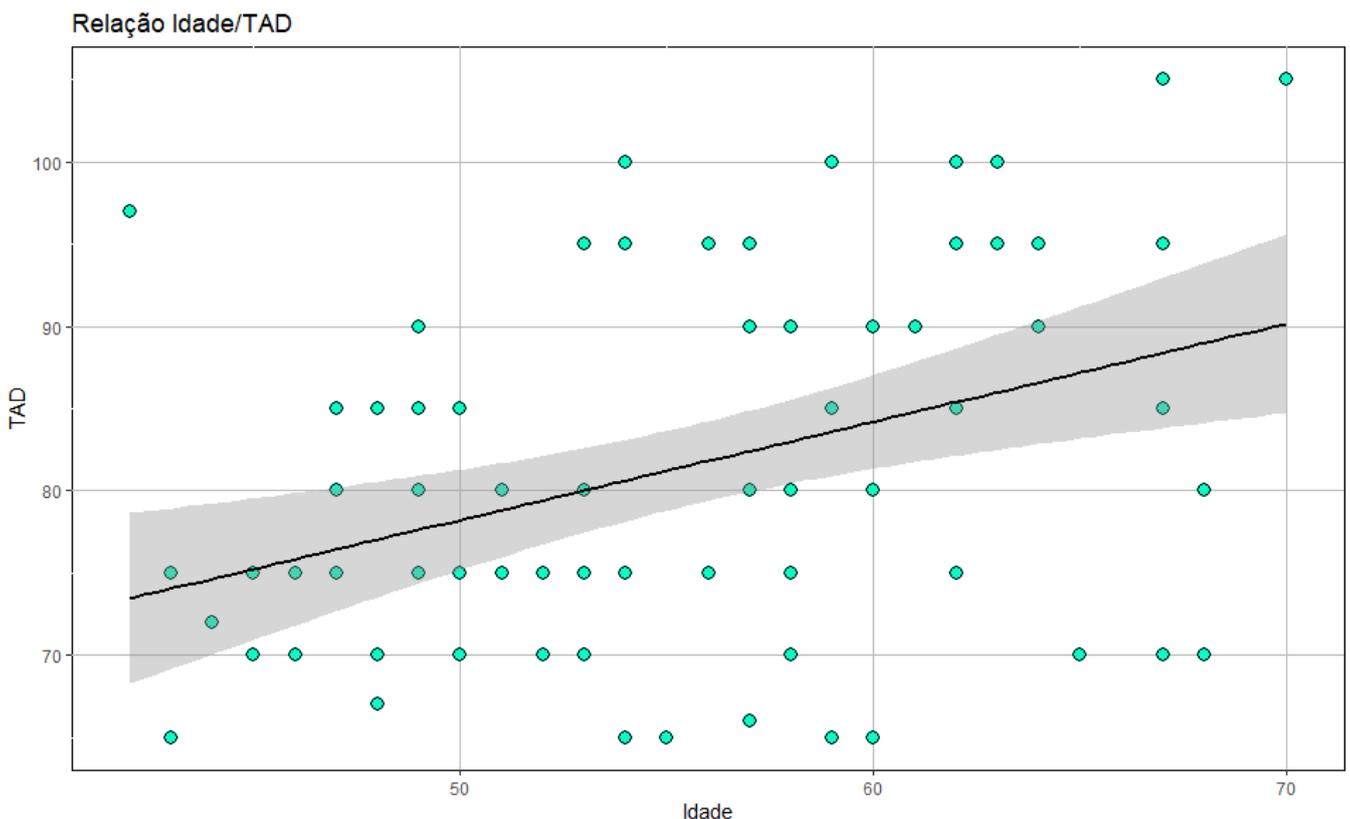
Henrique Pocinho IST199952

```
library(ggplot2) #Representação grafica
library(readxl) #Leitura de ficheiros
library(tidyverse) #Tratamento de dados
library(dplyr)

#Leitura do ficheiro
ficheiro <- read_excel(file.choose())
#tratamento de dados
graf <- ficheiro %>% select(Idade,TAD)
#A data frame está pronta para ser representada gráficamente

plot <- ggplot(graf, aes(x=Idade, y=TAD)) +
  geom_point(shape=21, colour="black",fill="#0dfcc4",size=3,stroke = 1) +
  geom_smooth(method=lm, color="black") +
  labs(x = "Idade", y = "TAD") +
  ggtitle("Relação Idade/TAD") +
  theme(legend.key.size = unit(1, 'cm'),
        legend.key.height = unit(1, 'cm'),
        legend.key.width = unit(1, 'cm'),
        legend.title = element_text(size=14),
        legend.text = element_text(size=10),
        panel.background = element_rect(fill = "white", colour = "black",
                                         size = 0, linetype = "solid"),
        panel.grid.major = element_line(size = 0.5, linetype = 'solid',
                                         colour = "grey"),
        panel.grid.minor = element_line(size = 0.25, linetype = 'solid',
                                         colour = "white"))
```

plot



O gráfico de dispersão indica que aparenta haver uma correlação entre a idade e TAD.

Essa correlação pode ser explicitada ao colocar uma reta de regressão que nos indica que com o aumento da idade há um aumento da TAD.

Obviamente que a idade não é o único fator que determina a TAD dado que existem dados de pessoas novas com TAD alta e pessoas mais velhas com uma TAD menor.

A cinzento está o intervalo de confiança da regressão, que é relativamente largo, no entanto indica sempre uma correlação positiva.