

1. Código em R

```
library(readxl)
library(ggplot2)

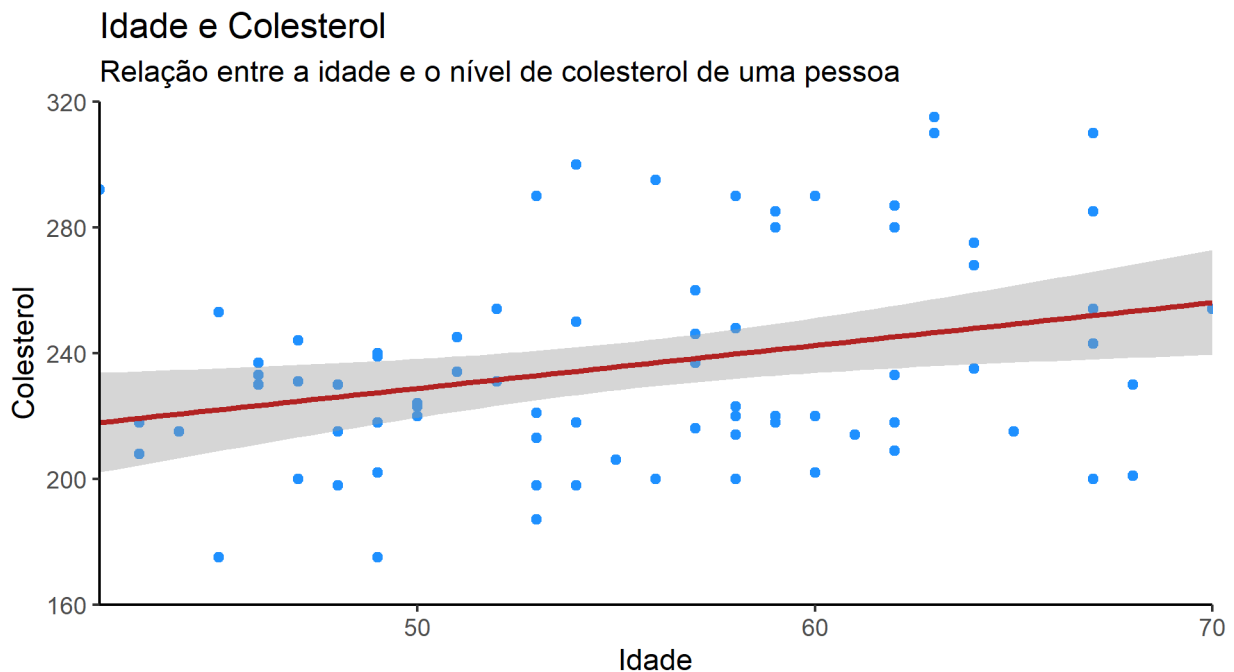
# Lê o ficheiro excel
Utentes <- read_excel("Utentes.xlsx")

# Calcula a correlação entre os dados
cor(Utentes$Idade, Utentes$Colesterol)

# Desenho do gráfico com os dados
plot <- ggplot(Utentes, aes(Idade, Colesterol))+
  geom_point(color = "dodgerblue")+
  geom_smooth(method = "lm", color = "firebrick")+
  labs(title = "Idade e Colesterol", x = "Idade", y = "Colesterol",
       subtitle = "Relação entre a idade e o nível de colesterol de uma pessoa") +
  scale_y_continuous(expand = c(0,0), limits = c(160, 320)) +
  scale_x_continuous(expand = c(0,0), limits = c(42, 70)) + theme_classic()

# Guarda o plot como imagem
ggsave("Plot4.png", plot, width = 1920, height = 1080, units = "px")
```

2. Gráfico de dispersão



3. Comentários

Através da observação do gráfico de dispersão e da sua regressão linear, é possível observar uma fraca correlação positiva ($= 0.3$) entre as variáveis. Isto quer dizer que estão vagamente relacionadas, não se podendo dizer ao certo que um aumento da idade causa um aumento de colesterol.