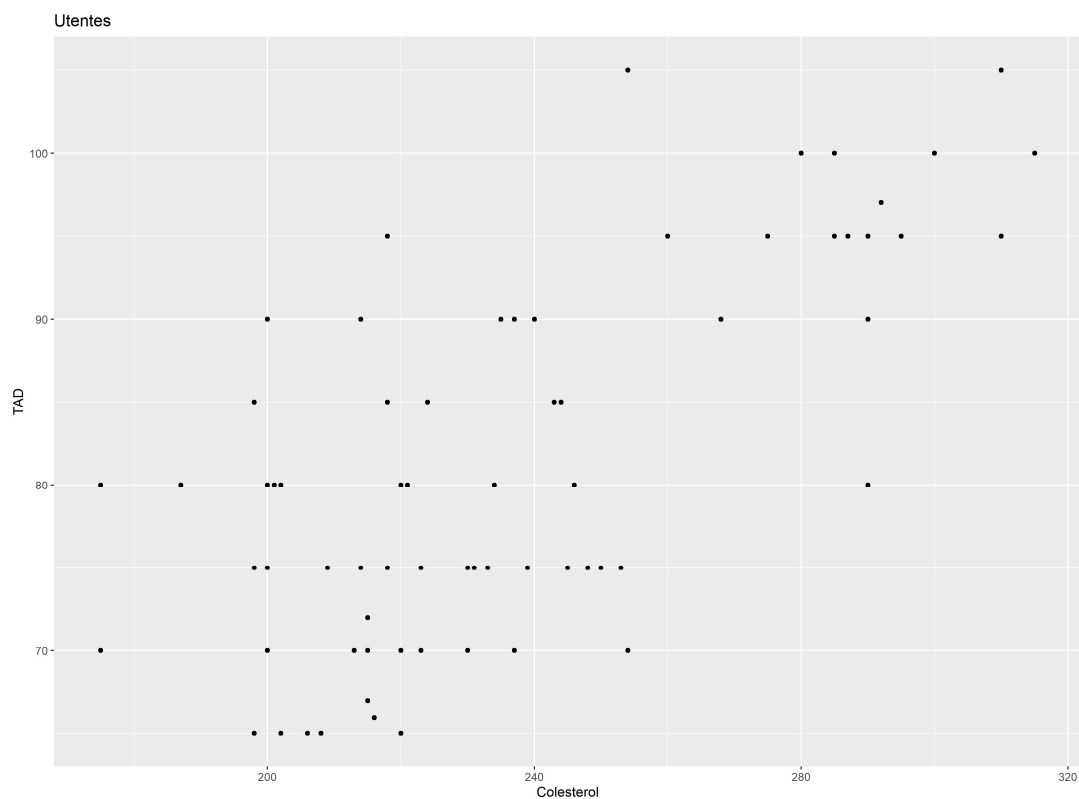


```
library(readxl)
library(tidyverse)
data <- read_excel("C:/Users/joao_/OneDrive/Documentos/PE/projeto/4/Utentes.xlsx")

keeps <- c("Colesterol")
Colesterol = data[keeps]

keeps<- c("TAD")
TAD = data[keeps]
data <- data.frame(Colesterol, TAD)

ggplot(data, aes(y = TAD, x=Colesterol)) +
  geom_point()+
  labs(title = "Utentes",
       x = "Colesterol",
       y = "TAD",
  )
ggsave("Ex4.png")
```



Analisando a gráfico de dispersão conseguimos identificar duas zonas de concentração principais, entre TAD = [70, 80], e Colesterol = [200, 260] e entre TAD = [90, 100], e Colesterol = [280, 300]. Apesar de não ser criada uma função fácil de analisar, conseguimos entender que existe uma relação de linearidade entre TAD e colesterol.