



UNIVERSIDAD
CATÓLICA DE
TEMUCO

DEPARTAMENTO DE
INGENIERÍA INFORMÁTICA
FACULTAD DE INGENIERÍA

Taller Formativo 10

Mario Castillo Sanhueza

Docente: Dr. Julio Rojas Mora

Departamento de Ingeniería Informática
Facultad de Ingeniería
Universidad Católica de Temuco

Agosto 19, 2022

1. Problemática.

- Utilizando la siguiente base de datos (BD) Compare las plaquetas (platelets) dependiendo de si la persona tiene anemia (anemia) y si tiene diabetes (diabetes), usando el gráfico que considere más adecuado.

2. Desarrollo

- Para obtener esta comparativa filtraremos las plaquetas basándose en las personas que padezcan anemia y las personas que padezcan diabetes, luego usaremos un gráfico de densidad para agrupar estos mismos datos según su conteo de plaquetas.

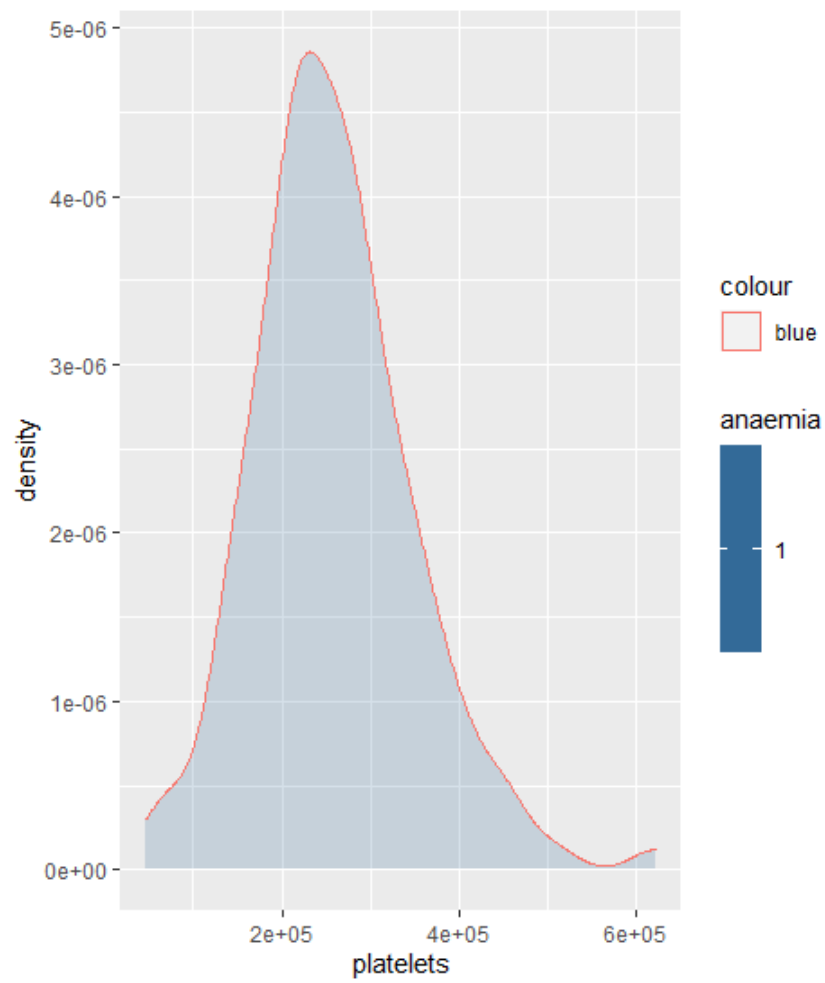


Figura 1: Pacientes con anaemia.

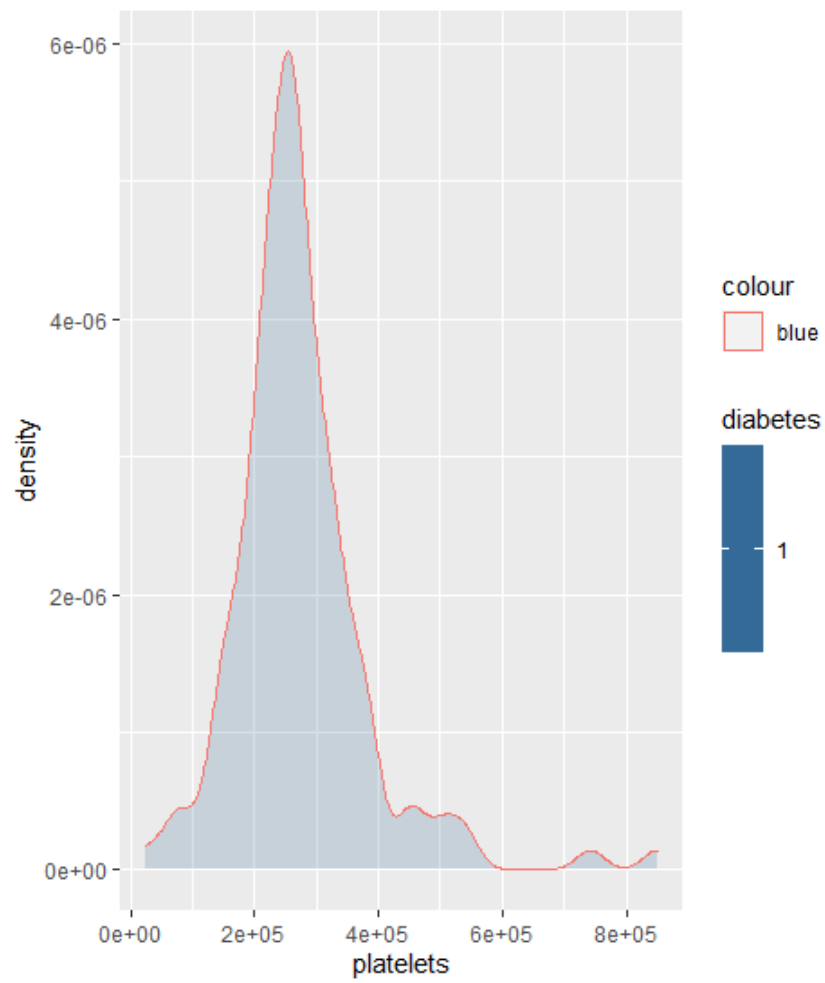


Figura 2: Pacientes con diabetes.

3. Código

```
library(ggplot2)

#We load the data
dataFH<-read.table("heart_failure_clinical_records_dataset.csv",
                  sep="," ,header=T)

#filer platelets of patients with anaemia
dataFHA<-subset(dataFH, dataFH$anaemia == 1,
                select=c(anaemia ,platelets))

#filer platelets of patients with diabetes
dataFHD<-subset(dataFH, dataFH$diabetes == 1,
                select=c(diabetes ,platelets))

#plot of density for anaemia patients.
ggplot(dataFHA, aes(x = platelets)) +
  geom_density(aes(fill=anaemia ,colour="blue"),alpha=0.2)

#plot of density for diabetes patients.
ggplot(dataFHD, aes(x = platelets)) +
  geom_density(aes(fill=diabetes ,colour="blue"),alpha=0.2)
```