

```

1  ### =====
  ### Màster Universitari en Estadística i Investigació Operativa
  ### Computación en Estadística y en Optimización
  ### Test 2 con R (Grupo A): 5.11.2018
  ### Nombre:
  ### =====

  ## Ejercicio 1
  ## =====
10 setwd("W:/MESIO/Computacio/Examen2")
  load("CeoGrATestR2.RData")

  # 1a
  # --
  length(with(immulong, lympho[1500<=lympho & lympho<=2000]))
  # 181 pacientes
  # 1b
  # --
20 names1<-c("lympho", "cd4", "cd8", "nkiller")
  for ( i in names1){
    immulong[names1][i] [immulong[names1][i]==0]<-NA
  }
  summary(immulong)

  # 1c
  # --
  colSums(is.na(immulong))
  max(colSums(is.na(immulong)))
30 # El máximo es cd8 con 15 missings

  # Al intentar hacer todo el código en una línea me da un resultado incorrecto.
  # Alguna cosa no debo de estar controlando
  # which(colSums(is.na(immulong))==max(colSums(is.na(immulong))))

  # 1d
  # --
  sum(sapply(immulong, is.numeric))
  # 6 variables numéricas

40 # 1e
  # --
  lapply(immulong[sapply(immulong, is.numeric)], range, na.rm=TRUE)

  # 1f
  # --
  #install.packages("dplyr")
  library(dplyr)

50 immuno_1f<-immulong %>%
  filter(visit== 1|visit== 2) %>%
  select(-datvisit) %>%
  group_by(vol) %>%
  mutate(NASLympho=sum(is.na(lympho)))

  ## Ejercicio 2
  ## =====
  windows(20,10)
60 par(las = 1, font = 2, font.lab = 4, font.axis = 2, cex.main = 1.2,
  cex.lab = 1.25, cex.axis = 1.1, pch=16)
  with(immulong, plot(nkiller[visit==3]~cd8[visit==3], ylab="Natural killer cell count",
    xlab="CD8 cell count", main="Natural Killer Cells vs. CD8 Cells [3rd vi

  # Creo un data.frame auxiliar para identificar los 4 pacientes
  immuno4<-subset(immulong, visit==3 & (vol==26|vol==76|vol==91|vol==97))
  with(immuno4, text(cd8, nkiller-14, paste0("Vol.", vol)

```

```
70 ## Ejercicio.3
## =====
library(Hmisc)
states<-spss.get("States77WithoutWest.sav")
summary(states)
levels(states$State) [1:4]

fun_spss<-function(x){
80 dat<-spss.get(x)
require(stringr)
for (i in 1:ncol(dat)){
  if(!is.numeric(dat[,i])){
    dat[,i]<-droplevels(dat[,i])
    dat[,i]<-as.factor(str_trim(dat[,i]))
  }
}
return(dat)
90 }

prueba<-fun_spss("States77WithoutWest.sav")

table(prueba$Region)
levels(prueba$State) [1:4]
```

(-0.2) 1.85

(-0.05) dat[i] es un data frame y por ende
nunca un objeto numérico.

=> TODAS las variables se
convierten en factores (-0.4)

sustituir por dat[,i]