

## Caso practico 2

Laura Sudupe

24/09/2020

- a) Importar el fichero stressEchoMale.csv que corresponde a los datos de los hombres y guardarlos en un data frame llamado stressEchoMaleData.

```
stressEchoMaledata <-  
  read.csv('/Users/sudup/OneDrive/Documentos/UOC MATERIALES/SOFTWARE/LAB1_datos/SAD_LAB1_datos/stressEchoMale.csv')
```

- b) Importar el fichero stressEchoFem.dta que corresponde a los datos de las mujeres y guardarlos en un data frame llamado stressEchoFemaleData.

```
library(foreign)  
stressEchoFemaleData <-  
  read.dta('/Users/sudup/OneDrive/Documentos/UOC MATERIALES/SOFTWARE/LAB1_datos/SAD_LAB1_datos/stressEchoFemale.dta')  
head(stressEchoFemaleData)
```

##	idPatient	pkhr	dose	maxhr	age	gender	newMI	newPTCA	newCABG	death	hxofCig
## 1	4	118	30	118	57	female	0	0	0	0	non-smoker
## 2	7	98	40	98	81	female	0	0	0	0	non-smoker
## 3	8	144	40	144	90	female	0	0	0	0	non-smoker
## 4	9	115	40	113	81	female	0	0	0	0	non-smoker
## 5	11	171	40	171	61	female	0	0	0	0	non-smoker
## 6	14	110	40	110	29	female	0	0	0	0	non-smoker

- c) Especificar cómo cambiaríais el directorio donde guardar los ficheros de datos y de código que generaréis con R-Studio.

```
setwd('/Users/sudup/OneDrive/Documentos/UOC MATERIALES/SOFTWARE/')
```

- d) Crear un data frame llamado stressEchoDataSet que contenga los datos de stressEchoMaleData y stressEchoFemaleData (Pista: podéis utilizar la función rbind()).

```
stressEchoDataSet <- rbind(stressEchoMaledata, stressEchoFemaleData)  
head(stressEchoDataSet)
```

##	idPatient	pkhr	dose	maxhr	age	gender	newMI	newPTCA	newCABG	death	hxofCig
## 1	1	114	40	100	85	male	0	0	0	0	non-smoker
## 2	2	120	40	120	73	male	1	0	0	0	non-smoker
## 3	3	120	40	120	73	male	1	0	0	0	non-smoker
## 4	5	129	40	129	34	male	0	0	0	0	non-smoker
## 5	6	123	40	123	71	male	0	0	0	0	non-smoker
## 6	10	126	40	126	86	male	0	0	0	0	non-smoker

- e) Mostrar las características más importantes del data frame stressEchoDataSet, es decir, captura de los primeros y últimos registros de datos, nombre de las columnas, tipos de variables,...

```
head(stressEchoDataSet) #Para ver las primeras lineas
```

##	idPatient	pkhr	dose	maxhr	age	gender	newMI	newPTCA	newCABG	death	hxofCig
## 1	1	114	40	100	85	male	0	0	0	0	non-smoker

```
## 2      2 120 40 120 73 male 1 0 0 0 non-smoker
## 3      3 120 40 120 73 male 1 0 0 0 non-smoker
## 4      5 129 40 129 34 male 0 0 0 0 non-smoker
## 5      6 123 40 123 71 male 0 0 0 0 non-smoker
## 6     10 126 40 126 86 male 0 0 0 0 non-smoker
```

```
tail(stressEchoDataSet) #Para ver las ultimas lineas
```

```
##      idPatient pkhr dose maxhr age gender newMI newPTCA newCABG death  hxofCig
## 333      550  110  20  107  79 female  0      0      0      0 non-smoker
## 334      551  132  40  132  79 female  0      0      0      0 non-smoker
## 335      552   88  30  118  70 female  0      0      0      0 non-smoker
## 336      553  160  15  160  64 female  0      0      0      0 non-smoker
## 337      557  107  40  106  70 female  0      0      0      1 non-smoker
## 338      558  161  40  156  56 female  0      0      0      0 non-smoker
```

```
names(stressEchoDataSet) #Para ver el nombre de las columnas
```

```
## [1] "idPatient" "pkhr"      "dose"      "maxhr"      "age"        "gender"
## [7] "newMI"      "newPTCA"    "newCABG"    "death"      "hxofCig"
```

```
str(stressEchoDataSet) #Nos da la estructura del dataframe
```

```
## 'data.frame': 558 obs. of 11 variables:
## $ idPatient: int 1 2 3 5 6 10 12 13 21 26 ...
## $ pkhr : int 114 120 120 129 123 126 127 169 123 102 ...
## $ dose : int 40 40 40 40 40 40 30 40 40 25 ...
## $ maxhr : int 100 120 120 129 123 126 125 169 123 102 ...
## $ age : int 85 73 73 34 71 86 63 86 70 69 ...
## $ gender : chr "male" "male" "male" "male" ...
## $ newMI : int 0 1 1 0 0 0 0 0 0 0 ...
## $ newPTCA : int 0 0 0 0 0 0 0 0 0 0 ...
## $ newCABG : int 0 0 0 0 0 0 0 0 0 0 ...
## $ death : int 0 0 0 0 0 0 0 0 0 0 ...
## $ hxofCig : chr "non-smoker" "non-smoker" "non-smoker" "non-smoker" ...
```

```
dim(stressEchoDataSet) #Nos da la dimensión
```

```
## [1] 558 11
```

- f) Exportar el data frame stressEchoDataSet a un fichero en formato texto, es decir, mostrar cómo se realizaría en R desde RStudio y cómo se haría desde RCommander

```
#Primero en Rstudio
```

```
stressEchoDataSet_txt <-
```

```
  write.table(stressEchoDataSet, '/Users/sudup/OneDrive/Documentos/UOC MATERIALES/SOFTWARE/LAB1_datos/S
```

```
#En Rcommander ejecutar la siguiente instrucción
```

```
#library(Rcmdr)
```

```
#Una vez abierto R-Commander, clicamos en «No hay conjunto de datos activo» y
```

```
#seleccionamos stressEchoDataset. Así, una vez
```

```
#cargados los datos, seleccionamos Datos / Conjunto de datos activo / Exportar
```

```
#conjunto de datos activo.
```

- g) Realizar las siguientes consultas a partir del data frame stressEchoDataSet: • Calcular el número de pacientes que han realizado el estudio en stressEchoDataSet y cuántos parámetros se han analizado.

```
dim(stressEchoDataSet)
```

```
## [1] 558 11
```

- Mostrar en una tabla, la distribución por la variable age del data frame stressEchoDataSet y la ordenación de estas edades en modo ascendente (Pista: podéis consultar en la Ayuda de R-Studio la función ordered()).

```
#stressEchoDataSet_age <- ordered(stressEchoDataSet$age)
stressEchoDataSet_age <- factor(stressEchoDataSet$age,ordered=is.ordered(stressEchoDataSet$age))
stressEchoDataSet_age
```

```
## [1] 85 73 73 34 71 86 63 86 70 69 70 64 46 84 45 61 60 54 72 77 91 92 62 45 77
## [26] 57 73 54 67 69 67 79 77 69 63 82 53 59 81 65 81 79 59 74 88 67 59 70 88 50
## [51] 63 43 38 64 62 67 49 69 55 71 74 78 75 50 74 74 70 91 71 67 69 81 87 66 66
## [76] 71 63 78 49 79 88 57 65 83 66 86 65 82 67 50 72 74 70 75 54 58 72 30 82 73
## [101] 67 62 68 59 71 75 93 62 89 81 70 59 65 71 73 48 45 68 76 75 76 62 38 78 72
## [126] 59 71 63 75 81 74 66 78 65 68 81 64 67 68 71 76 65 58 54 75 67 71 74 82 69
## [151] 78 58 75 67 63 76 77 47 62 75 47 73 75 87 61 38 60 68 69 72 72 72 40 62 79
## [176] 68 51 56 59 65 40 50 74 63 81 65 78 77 53 61 68 59 51 54 73 69 68 73 83 67
## [201] 76 77 58 75 70 69 51 76 62 78 67 66 79 71 72 73 74 70 74 64 57 81 90 81 61
## [226] 29 71 63 90 47 65 81 83 91 59 84 74 26 72 42 67 47 46 69 52 74 53 60 70 70
## [251] 76 87 41 83 73 76 83 88 69 64 69 76 83 74 72 62 82 77 70 67 59 82 64 64 65
## [276] 72 71 57 54 70 70 47 77 47 52 69 72 65 85 88 63 56 64 58 83 28 52 67 78 73
## [301] 85 72 54 60 58 79 73 71 90 74 70 54 59 67 56 79 57 85 61 77 77 68 52 60 64
## [326] 77 66 59 54 69 64 57 73 61 40 73 48 59 69 69 71 86 73 85 57 77 81 75 88 58
## [351] 69 81 64 82 67 74 51 65 50 84 86 55 68 78 54 75 89 62 73 47 74 67 63 69 67
## [376] 74 55 63 70 83 65 68 88 61 76 80 48 39 68 69 76 66 67 59 70 71 73 67 68 84
## [401] 78 70 63 82 29 60 60 79 67 85 65 85 75 64 77 85 68 78 68 62 76 59 47 83 70
## [426] 55 63 67 83 68 71 45 50 67 75 53 49 65 69 68 59 34 74 55 71 64 33 79 69 75
## [451] 82 66 59 71 59 79 77 74 66 71 85 76 70 85 66 49 66 73 66 76 53 79 66 68 43
## [476] 69 70 52 34 60 48 64 53 73 65 62 77 67 49 85 69 66 64 73 44 72 53 60 53 50
## [501] 78 52 52 72 75 76 49 76 48 63 77 39 65 71 60 73 48 82 60 83 67 85 90 79 62
## [526] 83 54 72 72 62 59 67 62 77 56 53 64 69 71 64 73 72 76 60 75 78 59 70 59 72
## [551] 68 57 79 79 70 64 70 56
## 62 Levels: 26 28 29 30 33 34 38 39 40 41 42 43 44 45 46 47 48 49 50 51 ... 93
```

- Crear un vector que guarde los datos del paciente de menor edad e indicar la sintaxis para mostrar esta información por pantalla.

```
stressEchoDataSet_min_age <- c(min(stressEchoDataSet$age))
stressEchoDataSet_min_age
```

```
## [1] 26
```

- Crear un vector que incluya los id de los pacientes que tengan menos de 50 años, ¿cuántos pacientes hay?

```
stressEchoDataSet_menos50 <-
  stressEchoDataSet$idPatient[c(stressEchoDataSet$age < 50)]
stressEchoDataSet_menos50
```

```
## [1] 5 32 34 67 182 188 204 269 304 336 338 349 411 416 430 449 465 14 18
## [20] 30 35 37 38 51 100 102 116 164 167 208 234 237 259 291 323 330 339 350
## [39] 389 415 422 425 435 447 468 472 476 487
```

- Crear un vector que guarde las dosis otorgadas a cada paciente y especificar el valor mínimo y máximo de dosis asignadas. ¿A qué id de paciente o pacientes corresponde la menor dosis asignada? ¿Corresponde a hombres o a mujeres? Mostrar toda la información de estos pacientes.

```
dosis_paciente <- c(stressEchoDataSet$dose)
max_dosis <- max(dosis_paciente)
min_dosis <- min(dosis_paciente)
```

```
#Id paciente min dosis
id_paciente_min_dosis <-
  stressEchoDataSet$idPatient[stressEchoDataSet$dose== min_dosis]
id_paciente_min_dosis
```

```
## [1] 114 468
```

```
#Genero dosis min
genero_dosis_min <-
  stressEchoDataSet$gender[stressEchoDataSet$dose == min_dosis]
genero_dosis_min
```

```
## [1] "female" "female"
```

```
#Info correspondiente a los pacientes referidos
pacientes_dosis_min <-
  subset(stressEchoDataSet, stressEchoDataSet$dose == min_dosis)
pacientes_dosis_min
```

```
##      idPatient pkhr dose maxhr age gender newMI newPTCA newCABG death    hxofCig
## 741         114  105   10   105  58 female      0        0        0      0 non-smoker
## 287         468  108   10   108  49 female      0        0        0      0      heavy
```