

Repaso Faraway

Felipe Yebes

2023-10-01

Contents

0.1	Ejercicio 1	1
-----	-----------------------	---

0.1 Ejercicio 1

```
data(pima, package="faraway")
head(pima)
```

```
##   pregnant glucose diastolic triceps insulin   bmi diabetes age test
## 1         6     148         72      35         0 33.6    0.627  50    1
## 2         1      85         66      29         0 26.6    0.351  31    0
## 3         8     183         64       0         0 23.3    0.672  32    1
## 4         1      89         66      23        94 28.1    0.167  21    0
## 5         0     137         40      35       168 43.1    2.288  33    1
## 6         5     116         74       0         0 25.6    0.201  30    0
```

```
#Resumen#
summary(pima)
```

```
##      pregnant      glucose      diastolic      triceps
## Min.   : 0.000   Min.   : 0.0   Min.   : 0.00   Min.   : 0.00
## 1st Qu.: 1.000   1st Qu.: 99.0   1st Qu.: 62.00   1st Qu.: 0.00
## Median : 3.000   Median :117.0   Median : 72.00   Median :23.00
## Mean   : 3.845   Mean   :120.9   Mean   : 69.11   Mean   :20.54
## 3rd Qu.: 6.000   3rd Qu.:140.2   3rd Qu.: 80.00   3rd Qu.:32.00
## Max.   :17.000   Max.   :199.0   Max.   :122.00   Max.   :99.00
##      insulin      bmi      diabetes      age
## Min.   : 0.0   Min.   : 0.00   Min.   :0.0780   Min.   :21.00
## 1st Qu.: 0.0   1st Qu.:27.30   1st Qu.:0.2437   1st Qu.:24.00
## Median : 30.5   Median :32.00   Median :0.3725   Median :29.00
## Mean   : 79.8   Mean   :31.99   Mean   :0.4719   Mean   :33.24
## 3rd Qu.:127.2   3rd Qu.:36.60   3rd Qu.:0.6262   3rd Qu.:41.00
## Max.   :846.0   Max.   :67.10   Max.   :2.4200   Max.   :81.00
##      test
## Min.   :0.000
## 1st Qu.:0.000
## Median :0.000
```

```
## Mean :0.349
## 3rd Qu.:1.000
## Max. :1.000
```

Valores mínimos de 0 en la diástolica nos indican que algo anda mal. Ordenamos (clasificar, sort) los valores

```
sort(pima$diastolic)
```

```
## [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [19] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 24
## [37] 30 30 38 40 44 44 44 44 46 46 48 48 48 48 48 50 50 50
## [55] 50 50 50 50 50 50 50 50 50 50 52 52 52 52 52 52 52 52
## [73] 52 52 52 54 54 54 54 54 54 54 54 54 54 54 55 55 56 56
## [91] 56 56 56 56 56 56 56 56 56 56 58 58 58 58 58 58 58 58
## [109] 58 58 58 58 58 58 58 58 58 58 58 58 58 60 60 60 60 60
## [127] 60 60 60 60 60 60 60 60 60 60 60 60 60 60 60 60 60 60
## [145] 60 60 60 60 60 60 60 60 60 60 60 60 60 60 61 62 62 62
## [163] 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62 62
## [181] 62 62 62 62 62 62 62 62 62 62 62 62 62 64 64 64 64 64
## [199] 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64
## [217] 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64 64
## [235] 64 64 65 65 65 65 65 65 65 65 66 66 66 66 66 66 66 66
## [253] 66 66 66 66 66 66 66 66 66 66 66 66 66 66 66 66 66 66
## [271] 66 66 66 68 68 68 68 68 68 68 68 68 68 68 68 68 68 68
## [289] 68 68 68 68 68 68 68 68 68 68 68 68 68 68 68 68 68 68
## [307] 68 68 68 68 68 68 68 68 68 68 68 68 68 70 70 70 70 70
## [325] 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70
## [343] 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70
## [361] 70 70 70 70 70 70 70 70 70 70 70 70 70 70 70 72 72 72
## [379] 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72
## [397] 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72 72
## [415] 72 72 72 72 72 74 74 74 74 74 74 74 74 74 74 74 74 74
## [433] 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74
## [451] 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74 74
## [469] 74 74 74 75 75 75 75 75 75 75 75 75 76 76 76 76 76 76
## [487] 76 76 76 76 76 76 76 76 76 76 76 76 76 76 76 76 76 76
## [505] 76 76 76 76 76 76 76 76 76 76 76 76 76 76 78 78 78 78
## [523] 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78
## [541] 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78 78
## [559] 78 78 78 78 78 80 80 80 80 80 80 80 80 80 80 80 80 80
## [577] 80 80 80 80 80 80 80 80 80 80 80 80 80 80 80 80 80 80
## [595] 80 80 80 80 80 80 80 80 80 80 82 82 82 82 82 82 82 82
## [613] 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82
## [631] 82 82 82 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84
## [649] 84 84 84 84 84 84 84 84 84 85 85 85 85 85 85 86 86 86
## [667] 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 88
## [685] 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88
## [703] 88 88 88 88 88 88 90 90 90 90 90 90 90 90 90 90 90 90
## [721] 90 90 90 90 90 90 90 90 90 90 92 92 92 92 92 92 92 92
## [739] 94 94 94 94 94 94 95 96 96 96 96 98 98 98 100 100 100 102
## [757] 104 104 106 106 106 108 108 110 110 110 114 122
```

Establecemos los valores 0 como NA.

```
pima$glucose[pima$glucose==0]<- NA
pima$diastolic[pima$diastolic==0] <- NA
pima$triceps[pima$triceps==0]<- NA
pima$insulin[pima$insulin==0]<- NA
```

Pasar la variable test a factor

```
pima$test<-factor(pima$test)
summary(pima$test)
```

```
##    0    1
## 500 268
```

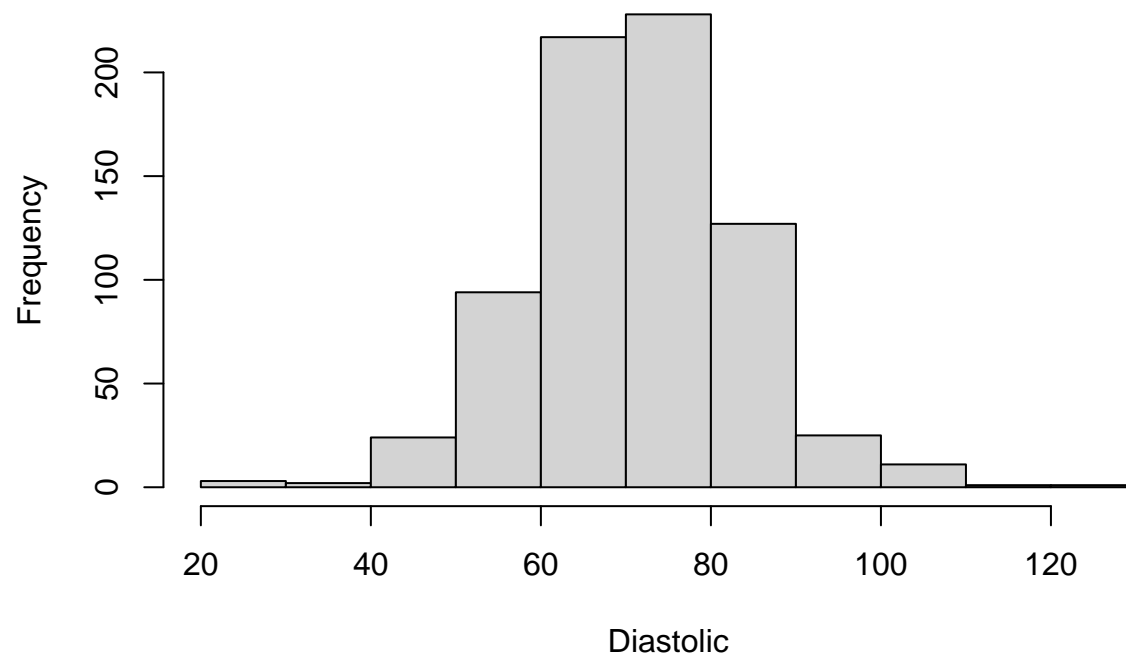
Ponemos datos descriptivos

```
levels(pima$test)<-c("negative","positive")
summary(pima)
```

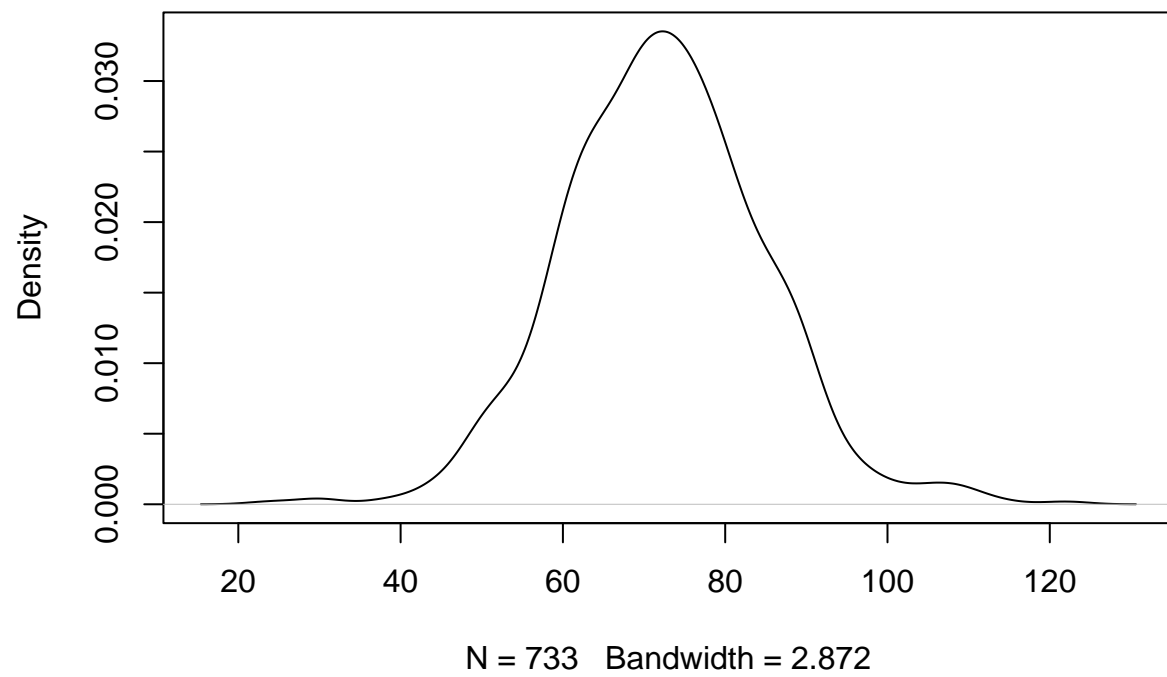
```
##      pregnant      glucose      diastolic      triceps
##  Min.   : 0.000   Min.   : 44.0   Min.   : 24.00   Min.   : 7.00
##  1st Qu.: 1.000   1st Qu.: 99.0   1st Qu.: 64.00   1st Qu.:22.00
##  Median : 3.000   Median :117.0   Median : 72.00   Median :29.00
##  Mean   : 3.845   Mean   :121.7   Mean   : 72.41   Mean   :29.15
##  3rd Qu.: 6.000   3rd Qu.:141.0   3rd Qu.: 80.00   3rd Qu.:36.00
##  Max.   :17.000   Max.   :199.0   Max.   :122.00   Max.   :99.00
##                      NA's    :5      NA's    :35      NA's    :227
##      insulin      bmi      diabetes      age
##  Min.   : 14.00   Min.   : 0.00   Min.   :0.0780   Min.   :21.00
##  1st Qu.: 76.25   1st Qu.:27.30   1st Qu.:0.2437   1st Qu.:24.00
##  Median :125.00   Median :32.00   Median :0.3725   Median :29.00
##  Mean   :155.55   Mean   :31.99   Mean   :0.4719   Mean   :33.24
##  3rd Qu.:190.00   3rd Qu.:36.60   3rd Qu.:0.6262   3rd Qu.:41.00
##  Max.   :846.00   Max.   :67.10   Max.   :2.4200   Max.   :81.00
##  NA's    :374
##      test
## negative:500
## positive:268
##
##
##
##
##
```

Graficamos

```
hist(pima$diastolic,xlab="Diastolic",main="")
```



```
plot(density(pima$diastolic,na.rm = T),main = "")
```



```
plot(sort(pima$diastolic),ylab="Sorted Diastolic")
```

