Solucion_Lab3.R

marco

2021-03-24

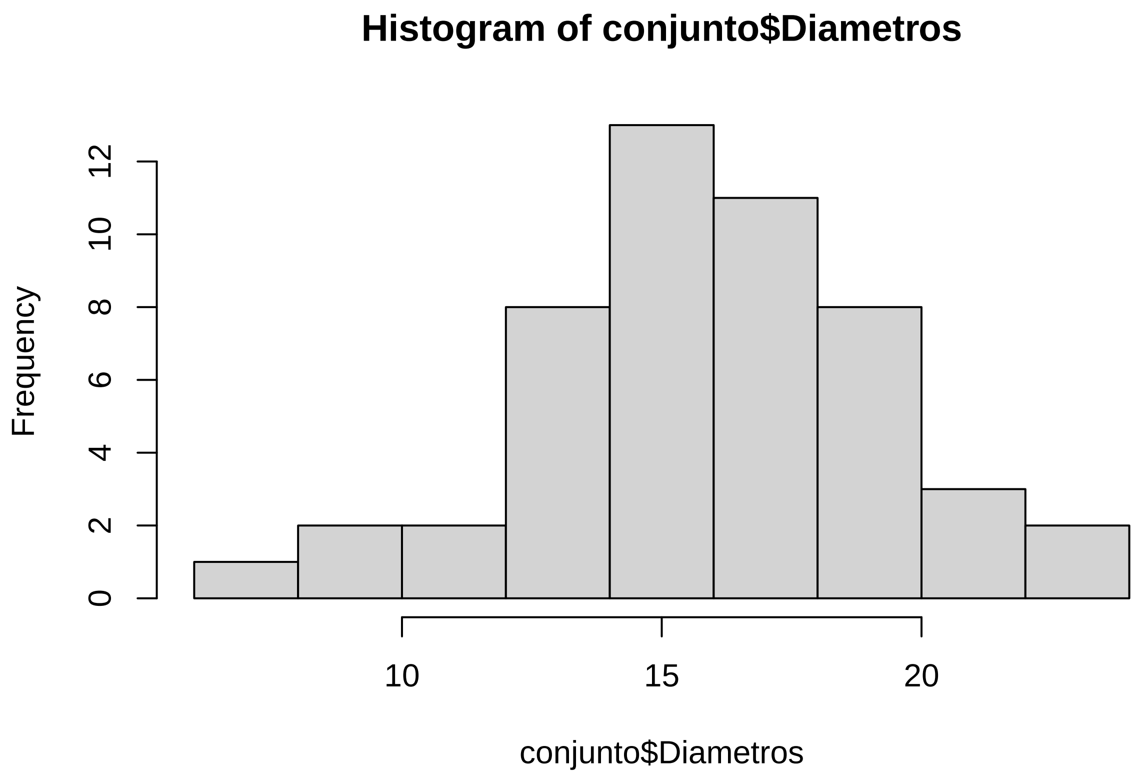
```r
# MAGT
# Laboratorio 3
#  03.03.2021


# Importar datos csv -----------------------------------------------------


conjunto <- read.csv("cuadro1.csv", header = TRUE)

head(conjunto)
```

```
##   Arbol Fecha Especie Posicion Vecinos Diametros Altura
## 1     1    12       F        C       4      15.3  14.78
## 2     2    12       F        D       3      17.8  17.07
## 3     3     9       C        D       5      18.2  18.28
## 4     4     9       H        S       4       9.7   8.79
## 5     5     7       H        I       6      10.8  10.18
## 6     6    10       C        I       3      14.1  14.90
```

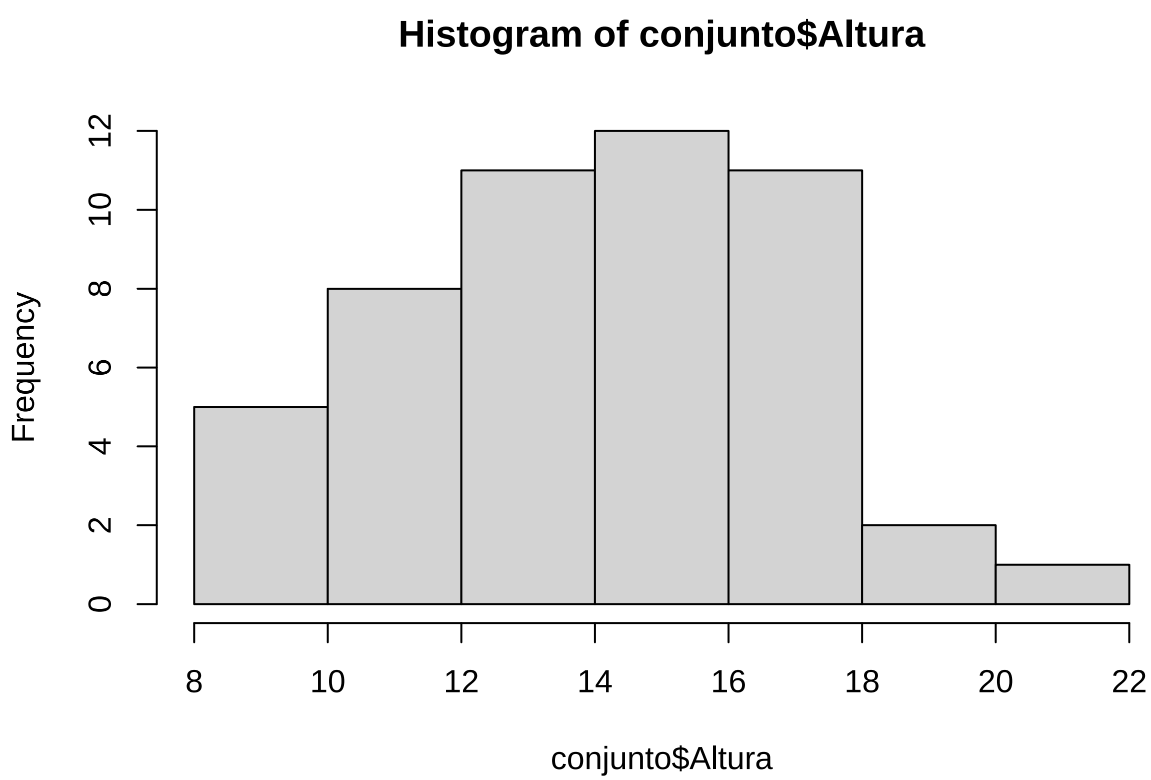
```r
tail(conjunto)
```

```
##    Arbol Fecha Especie Posicion Vecinos Diametros Altura
## 45    45    24       C        I       4      10.2  13.93
## 46    46    23       F        I       3      14.4  12.68
## 47    47    24       C        S       6       7.7  10.00
## 48    48    25       C        S       5       9.9   8.69
## 49    49    25       H        D       1      20.4  16.73
## 50    50    24       H        D       3      20.9  16.25
```

```r
hist(conjunto$Diametros)
```




```
hist(conjunto$Altura)
```




```
mean(conjunto$Diametros)
```

```
## [1] 15.794
```

```r
mean(conjunto$Vecinos)
```

```
## [1] 3.34
```

```r
range(conjunto$Vecinos)
```

```
## [1] 0 6
```

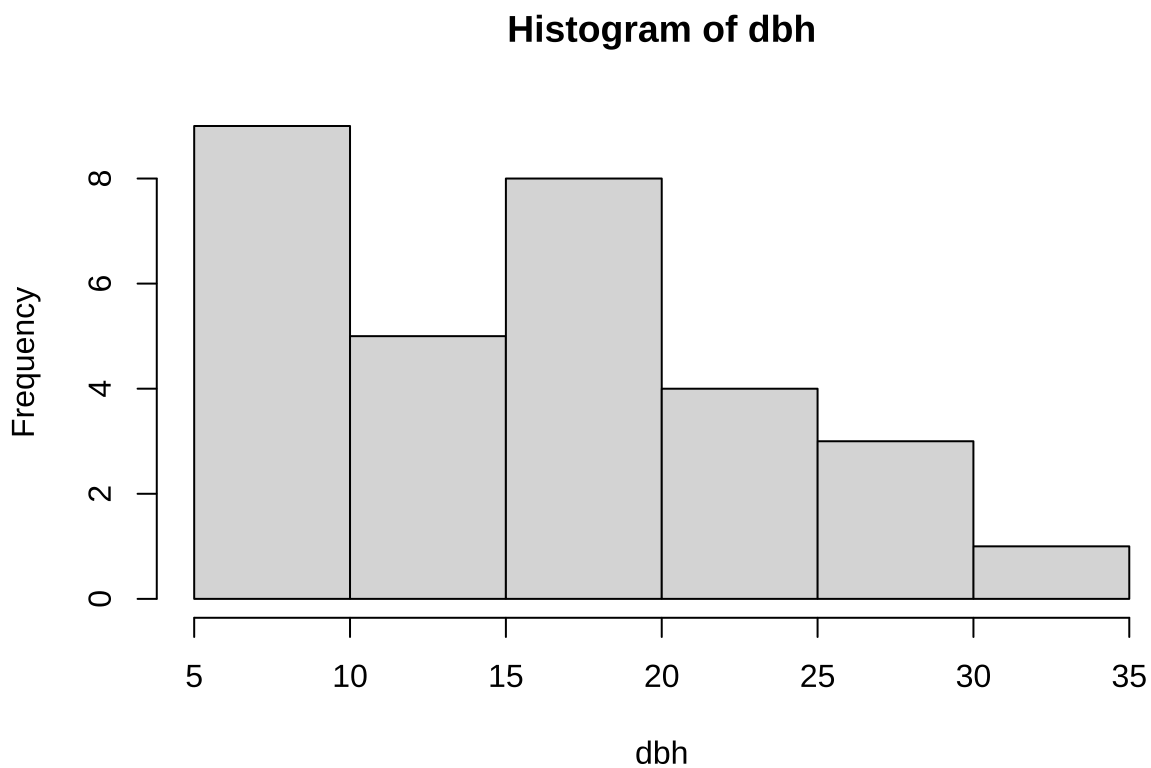
```r
#  Importar de la carpeta Datos el archivo cuadro2.csv

conjunto.2 <- read.csv("Datos/cuadro2.csv", header = TRUE)

# Directamente en consola ---------------------------------------------------

dbh <- c(16.5, 25.3, 22.1, 17.2, 16.1, 8.1, 34.3, 5.4, 5.7, 11.2, 24.1,
         14.5, 7.7, 15.6, 15.9, 10, 17.5, 20.5, 7.8, 27.3,
         9.7, 6.5, 23.4, 8.2, 28.5, 10.4, 11.5, 14.3, 17.2, 16.8)

sum(dbh)
```

```
## [1] 469.3
```

```r
prod(dbh)
```

```
## [1] 2.125828e+34
```

```r
hist(dbh)
```




```r
# Importar datos de una url -------------------------------------------------

prof_url <- "http://www.profepa.gob.mx/innovaportal/file/7635/1/accionesInspeccionfoanp.csv"
profepa <- read.csv(prof_url)
head(profepa)
```

```
##               Entidad Inspecciones Recorridos.de.vigilancia Operativos
## 1      Aguascalientes            0                        1          0
## 2     Baja California            0                        1          0
## 3 Baja California Sur            0                        0          0
## 4            Campeche            0                        0          0
## 5             Chiapas            0                        0          0
## 6           Chihuahua            3                        1          1
```

```
profepa
```

```
##                Entidad Inspecciones Recorridos.de.vigilancia Operativos
## 1       Aguascalientes            0                        1          0
## 2      Baja California            0                        1          0
## 3  Baja California Sur            0                        0          0
## 4             Campeche            0                        0          0
## 5              Chiapas            0                        0          0
## 6            Chihuahua            3                        1          1
## 7             Coahuila            1                        0          0
## 8               Colima            0                        0          0
## 9     Distrito Federal            0                        0          0
## 10             Durango            0                        0          0
## 11          Guanajuato            0                        0          0
## 12            Guerrero            0                        0          0
## 13             Hidalgo            0                        0          0
## 14             Jalisco            0                        0          0
## 15          M\xe9xico            2                        0          0
## 16       Michoac\xe1n            1                        3          1
## 17             Morelos            2                        0          1
## 18             Nayarit            0                        1          0
## 19       Nuevo Le\xf3n            0                        0          0
## 20              Oaxaca            0                        0          0
## 21              Puebla            0                        0          0
## 22       Quer\xe9taro            0                        0          0
## 23        Quintana Roo            0                        0          0
## 24  San Luis Potos\xed            0                        0          0
## 25             Sinaloa            0                        0          0
## 26              Sonora            0                        0          0
## 27             Tabasco            0                        0          0
## 28          Tamaulipas            0                        0          0
## 29            Tlaxcala            4                        2          0
## 30            Veracruz            0                        1          0
## 31          Yucat\xe1n            0                        0          0
## 32           Zacatecas            0                        1          0
## 33  Oficinas Centrales            6                       10          0
```

```
sum(profepa$Inspecciones)
```

```
## [1] 19
```

```
sum(profepa$Operativos)
```

```
## [1] 3
```

```
sum(profepa$Recorridos.de.vigilancia)
```

```
## [1] 21
```

```r
prof_url_2 <- paste0("http://www.profepa.gob.mx/innovaportal/",
                       "file/7635/1/accionesInspeccionfoanp.csv")
profepa2 <- read.csv(prof_url_2)
head(profepa2)
```

```
##                Entidad Inspecciones Recorridos.de.vigilancia Operativos
## 1      Aguascalientes            0                        1          0
## 2     Baja California            0                        1          0
## 3 Baja California Sur            0                        0          0
## 4            Campeche            0                        0          0
## 5             Chiapas            0                        0          0
## 6           Chihuahua            3                        1          1
```

```r
# Importar datos de url seguras -------------------------------------------

library(repmis)

conjunto.2 <- source_data("https://www.dropbox.com/s/hmsf07bbayxv6m3/cuadro1.csv?dl=1")
```

```
## Downloading data from: https://www.dropbox.com/s/hmsf07bbayxv6m3/cuadro1.csv?dl=1

## SHA-1 hash of the downloaded data file is:
## 2bdde4663f51aa4198b04a248715d0d93498e7ba
```

```r
sum(conjunto.2$Vecinos)
```

```
## [1] 167
```

```r
library(readr)

file <- paste0("https://raw.githubusercontent.com/mgtagle/",
                      "202_Analisis_Estadistico_2020/master/cuadro1.csv")
inventario <- read_csv(file)
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   Arbol = col_double(),
##   Fecha = col_double(),
##   Especie = col_character(),
##   Clase = col_character(),
##   Vecinos = col_double(),
##   Diametro = col_double(),
##   Altura = col_double()
## )
```

```r
sum(inventario$Vecinos)
```

```
## [1] 167
```

```r
# Parte II:  --------------------------------------------------------------

# Selección mediante restricciones

dbh
```

```
##  [1] 16.5 25.3 22.1 17.2 16.1  8.1 34.3  5.4  5.7 11.2 24.1 14.5  7.7 15.6 15.9
## [16] 10.0 17.5 20.5  7.8 27.3  9.7  6.5 23.4  8.2 28.5 10.4 11.5 14.3 17.2 16.8
```

```r
mean(dbh)
```

```
## [1] 15.64333
```

```r
dbh < 10
```

```
##  [1] FALSE FALSE FALSE FALSE FALSE  TRUE FALSE  TRUE  TRUE FALSE FALSE FALSE
## [13]  TRUE FALSE FALSE FALSE FALSE FALSE  TRUE FALSE  TRUE  TRUE FALSE  TRUE
## [25] FALSE FALSE FALSE FALSE FALSE FALSE
```

```r
sum(dbh < 10)
```

```
## [1] 8
```

```r
which(dbh < 10)
```

```
## [1]  6  8  9 13 19 21 22 24
```

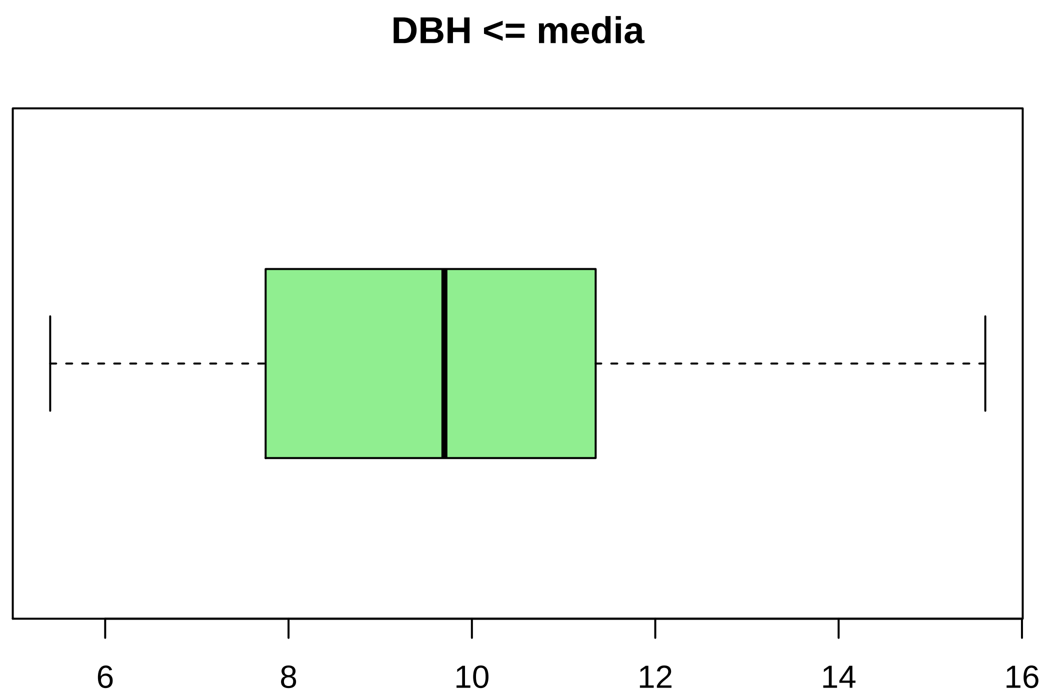
```r
dbh.url <- "https://raw.githubusercontent.com/mgtagle/PrincipiosEstadistica2021/main/DBH_1.csv"
parcelas <- read.csv(dbh.url)


tree.13 <- parcelas[!(parcelas$parcela == "2"),]

tree.23 <- parcelas[!(parcelas$parcela == "1"),]

tree.12 <- parcelas[!(parcelas$parcela == "3"),]


# Revisar las medias del dbh en cada combinación de parcelas

mean(tree.12$dbh); mean(tree.13$dbh); mean(tree.23$dbh)
```

```
## [1] 16.14
## [1] 15.42
## [1] 15.37
```

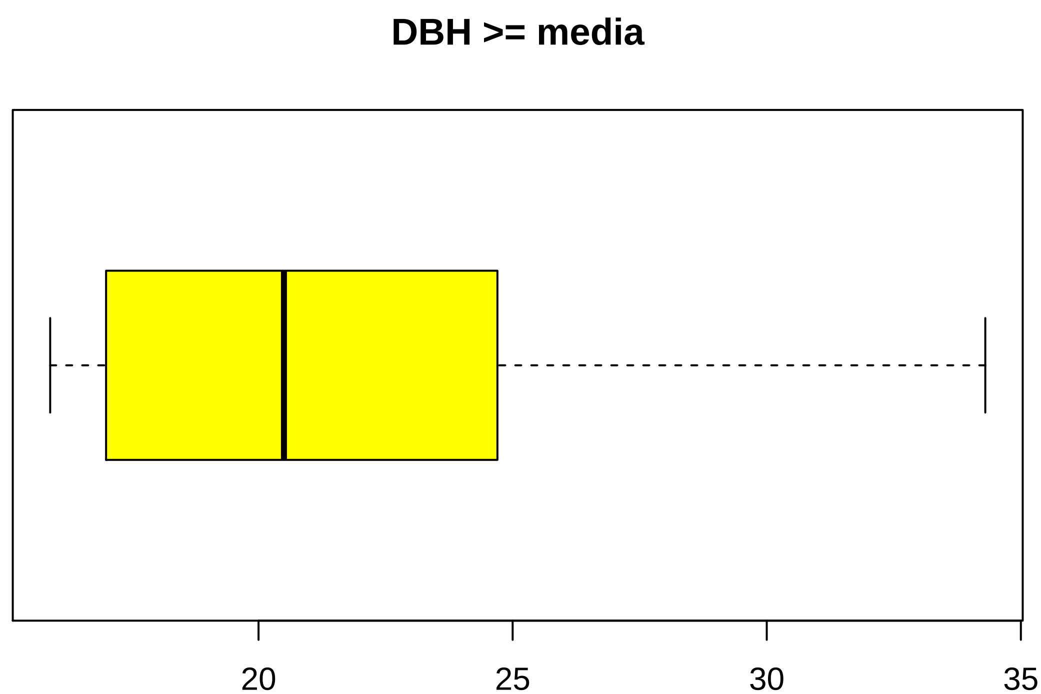
```r
# Selección de submuestras ------------------------------------------------

tree_mean <- subset(parcelas, dbh <= mean(parcelas$dbh))

tree.up <- subset(parcelas, dbh >= mean(parcelas$dbh))

mean(tree_mean$dbh); mean(tree.up$dbh)
```

```
## [1] 9.773333
## [1] 21.51333
```

```r
# Representación gráfica de los dos subconjuntos

boxplot(tree_mean$dbh, main = "DBH <= media", col = "lightgreen", horizontal = TRUE)
```




```
boxplot(tree.up$dbh, main = "DBH >= media", col = "yellow", horizontal = T)
```




```
quantile(tree_mean$dbh, 0.5)
```

```
## 50%
## 9.7
```

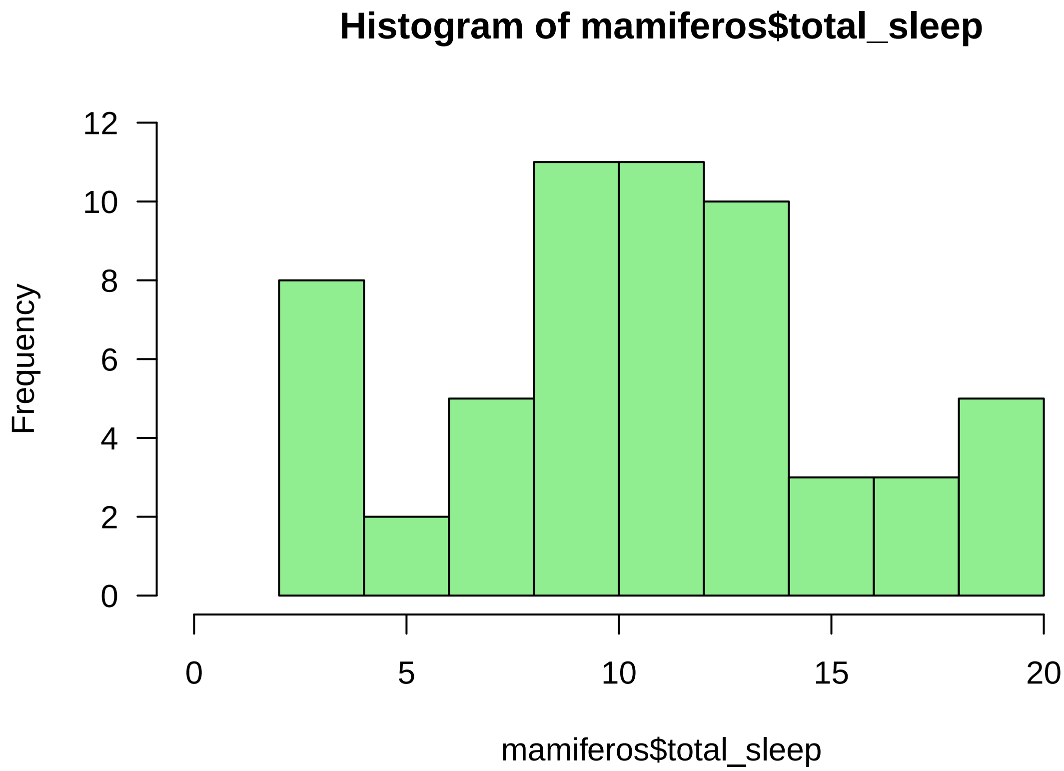
```
quantile(tree_mean$dbh, 0.75)
```

```
##   75%
## 11.35
```

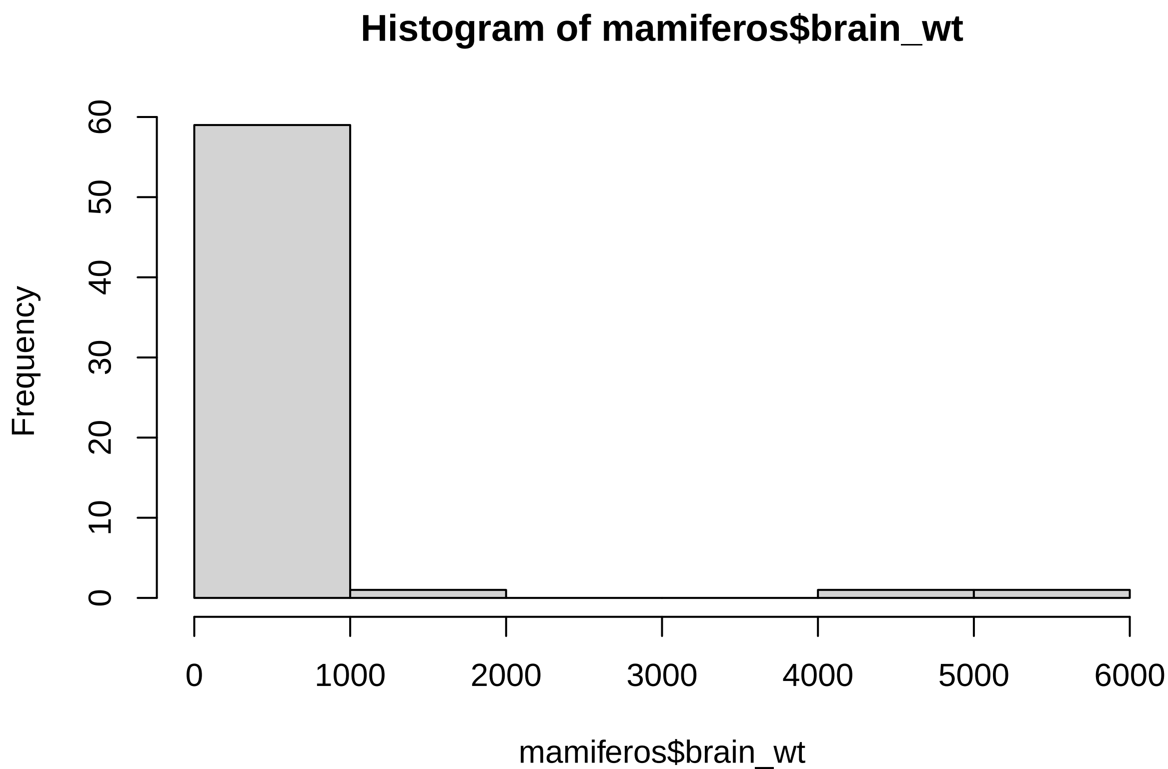
```r
# Parte 3: representación gráfica ----------------------------------------

mamiferos <- read.csv("https://www.openintro.org/data/csv/mammals.csv")

mean(mamiferos$brain_wt)
```

```
## [1] 283.1342
```

```r
hist(mamiferos$total_sleep, col="lightgreen",
     ylim = c(0, 12), xlim = c(0, 22),
     las = 1)
```




```r
hist(mamiferos$brain_wt)
```




```
fivenum(mamiferos$brain_wt)
```

```
## [1]    0.14    4.00   17.25  169.00 5712.00
```

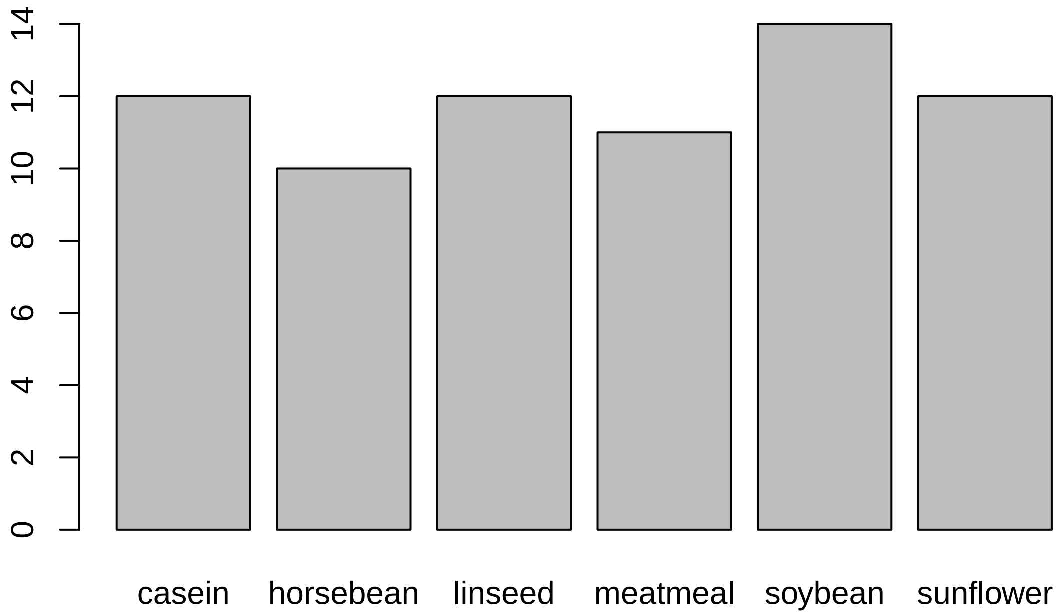
```
data("chickwts")
head(chickwts)
```

```
##   weight      feed
## 1    179 horsebean
## 2    160 horsebean
## 3    136 horsebean
## 4    227 horsebean
## 5    217 horsebean
## 6    168 horsebean
```

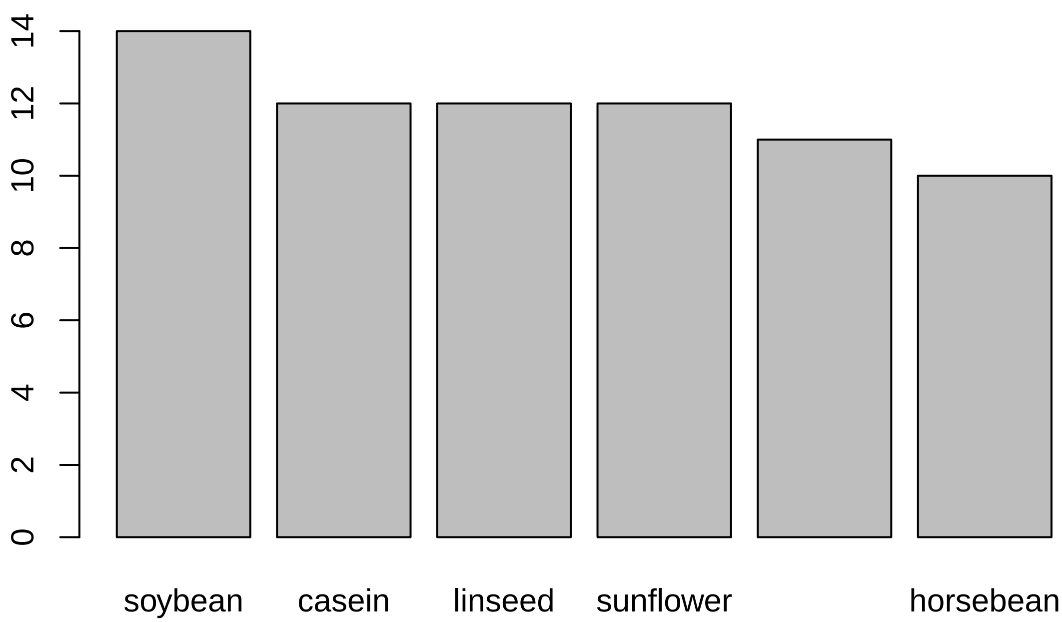
```
alimentación <- table(chickwts$feed)
alimentación
```

```
## 
##    casein horsebean   linseed  meatmeal   soybean sunflower 
##        12        10        12        11        14        12
```

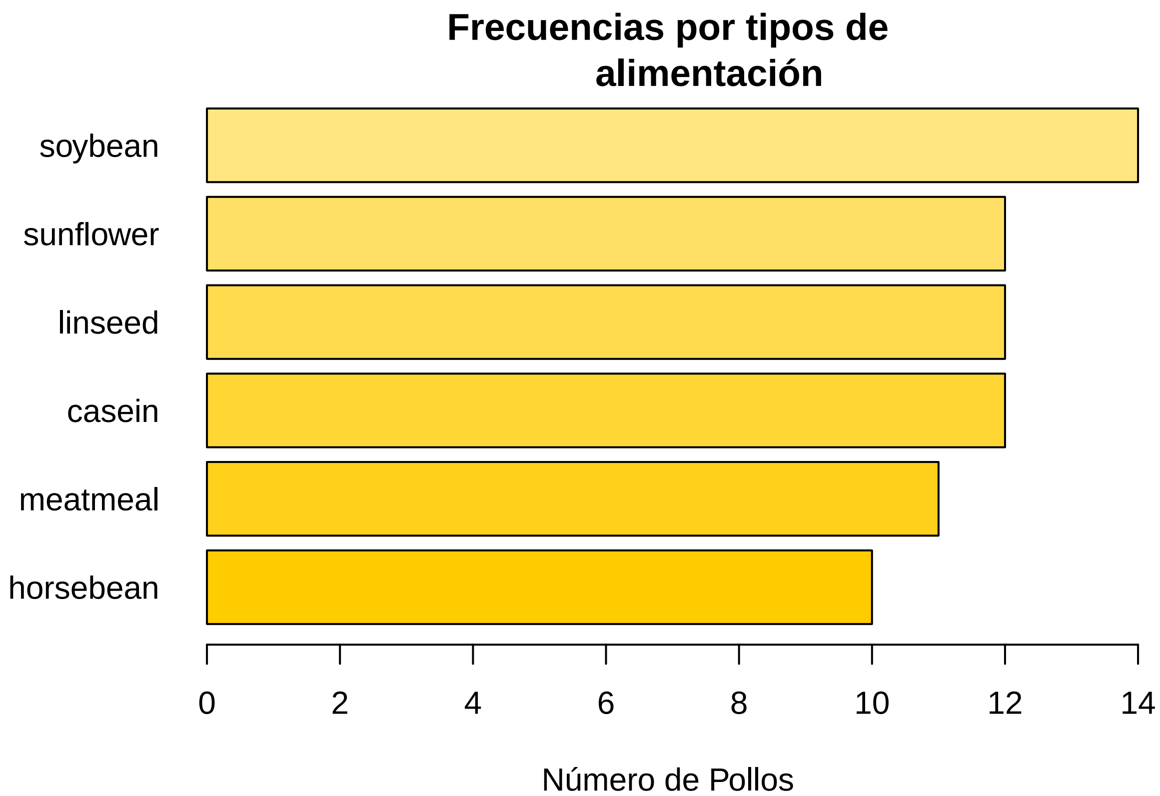
```
barplot(alimentación)
```




```
barplot(alimentación[order(alimentación, decreasing = TRUE)])
```




```
orig.par <- par() # Originales de las gráficas
par(oma=c(1,1,1,1))
par(mar=c(4,5,2,1))

barplot(alimentación[order(alimentación)],
        horiz = TRUE,
        las = 1,
        col = c("#ffcc00", "#ffd11a", "#ffd633", "#ffdb4d", "#ffe066", "#ffe680"),
        main = "Frecuencias por tipos de
        alimentación",
        xlab = "Número de Pollos")
```




```
orig.par
```

```
## $xlog
## [1] FALSE
##
## $ylog
## [1] FALSE
##
## $adj
## [1] 0.5
##
## $ann
## [1] TRUE
##
## $ask
## [1] FALSE
##
## $bg
## [1] "transparent"
##
## $bty
## [1] "o"
##
## $cex
## [1] 1
##
## $cex.axis
## [1] 1
##
```

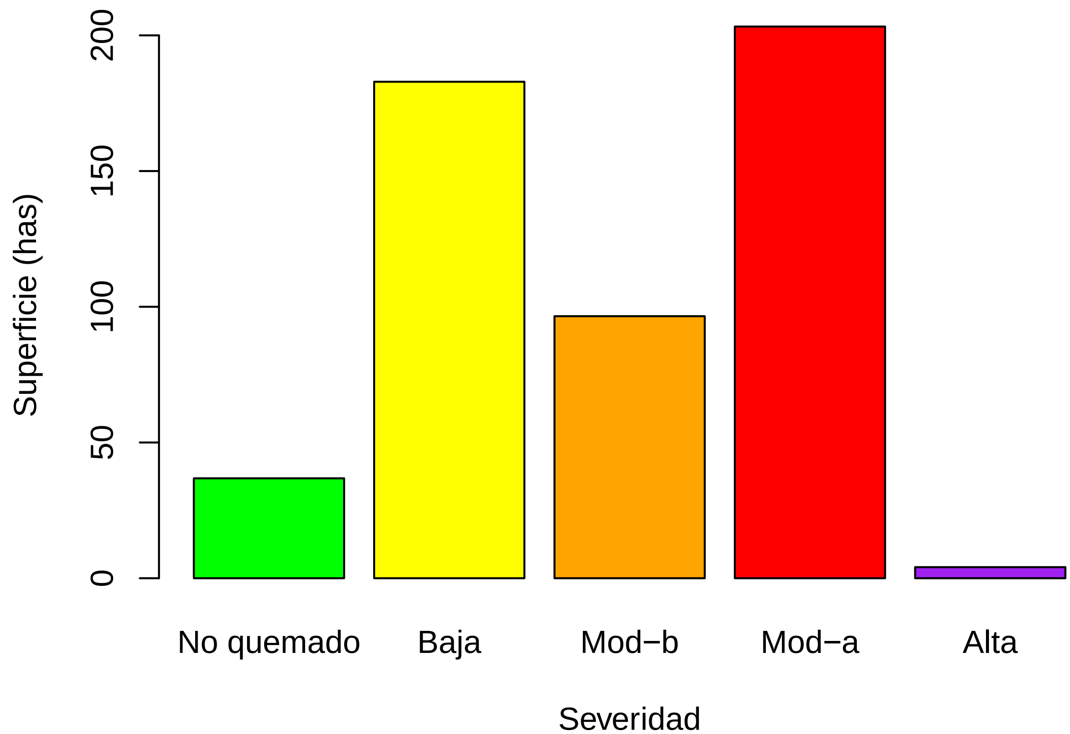
```
## $cex.lab
## [1] 1
##
## $cex.main
## [1] 1.2
##
## $cex.sub
## [1] 1
##
## $cin
## [1] 0.15 0.20
##
## $col
## [1] "black"
##
## $col.axis
## [1] "black"
##
## $col.lab
## [1] "black"
##
## $col.main
## [1] "black"
##
## $col.sub
## [1] "black"
##
## $cra
## [1] 10.8 14.4
##
## $crt
## [1] 0
##
## $csi
## [1] 0.2
##
## $cxy
## [1] 0.2155894 1.0631579
##
## $din
## [1] 6.5 4.5
##
## $err
## [1] 0
##
## $family
## [1] ""
##
## $fg
## [1] "black"
##
## $fig
## [1] 0 1 0 1
##
```

```
## $fin
## [1] 6.5 4.5
##
## $font
## [1] 1
##
## $font.axis
## [1] 1
##
## $font.lab
## [1] 1
##
## $font.main
## [1] 2
##
## $font.sub
## [1] 1
##
## $lab
## [1] 5 5 7
##
## $las
## [1] 0
##
## $lend
## [1] "round"
##
## $lheight
## [1] 1
##
## $ljoin
## [1] "round"
##
## $lmitre
## [1] 10
##
## $lty
## [1] "solid"
##
## $lwd
## [1] 1
##
## $mai
## [1] 1.02 0.82 0.82 0.42
##
## $mar
## [1] 5.1 4.1 4.1 2.1
##
## $mex
## [1] 1
##
## $mfcol
## [1] 1 1
##
```

```
## $mfg
## [1] 1 1 1 1
##
## $mfrow
## [1] 1 1
##
## $mgp
## [1] 3 1 0
##
## $mkh
## [1] 0.001
##
## $new
## [1] FALSE
##
## $oma
## [1] 0 0 0 0
##
## $omd
## [1] 0 1 0 1
##
## $omi
## [1] 0 0 0 0
##
## $page
## [1] TRUE
##
## $pch
## [1] 1
##
## $pin
## [1] 5.26 2.66
##
## $plt
## [1] 0.1261538 0.9353846 0.2266667 0.8177778
##
## $ps
## [1] 12
##
## $pty
## [1] "m"
##
## $smo
## [1] 1
##
## $srt
## [1] 0
##
## $tck
## [1] NA
##
## $tcl
## [1] -0.5
##
```

```
## $usr
## [1] -0.08  7.48 -0.14 14.00
##
## $xaxp
## [1] 0 7 7
##
## $xaxs
## [1] "r"
##
## $xaxt
## [1] "s"
##
## $xpd
## [1] FALSE
##
## $yaxp
## [1]  0 14  7
##
## $yaxs
## [1] "r"
##
## $yaxt
## [1] "s"
##
## $ylbias
## [1] 0.2
```

```r
superficie <- c(36.8, 182.9, 96.52, 203.24, 4.08)
nom <- c("No quemado", "Baja", "Mod-b", "Mod-a", "Alta")


# Instrucciones t.ly/98jG
data <- data.frame(
        name=nom,
        value=superficie
)

barplot(data$value, names= data$name,
        col=c("green", "yellow", "orange", "red", "purple"),
        ylab="Superficie (has)",
        xlab = "Severidad")
```




```
# 24 de marzo correlación ------------------------------------------------


data("anscombe")
anscombe
```

```
##    x1 x2 x3 x4    y1   y2    y3    y4
## 1  10 10 10  8  8.04 9.14  7.46  6.58
## 2   8  8  8  8  6.95 8.14  6.77  5.76
## 3  13 13 13  8  7.58 8.74 12.74  7.71
## 4   9  9  9  8  8.81 8.77  7.11  8.84
## 5  11 11 11  8  8.33 9.26  7.81  8.47
## 6  14 14 14  8  9.96 8.10  8.84  7.04
## 7   6  6  6  8  7.24 6.13  6.08  5.25
## 8   4  4  4 19  4.26 3.10  5.39 12.50
## 9  12 12 12  8 10.84 9.13  8.15  5.56
## 10  7  7  7  8  4.82 7.26  6.42  7.91
## 11  5  5  5  8  5.68 4.74  5.73  6.89
```

```
summary(anscombe[ , 1:4])
```

```
##        x1             x2             x3             x4
##  Min.   : 4.0   Min.   : 4.0   Min.   : 4.0   Min.   : 8
##  1st Qu.: 6.5   1st Qu.: 6.5   1st Qu.: 6.5   1st Qu.: 8
##  Median : 9.0   Median : 9.0   Median : 9.0   Median : 8
##  Mean   : 9.0   Mean   : 9.0   Mean   : 9.0   Mean   : 9
##  3rd Qu.:11.5   3rd Qu.:11.5   3rd Qu.:11.5   3rd Qu.: 8
##  Max.   :14.0   Max.   :14.0   Max.   :14.0   Max.   :19
```

```
sd(anscombe$x1)
```

```
## [1] 3.316625
```

```r
sd(anscombe$x2)
```

```
## [1] 3.316625
```

```r
summary(anscombe[ , 5:8])
```

```
##        y1               y2              y3              y4
##  Min.   : 4.260   Min.   :3.100   Min.   : 5.39   Min.   : 5.250
##  1st Qu.: 6.315   1st Qu.:6.695   1st Qu.: 6.25   1st Qu.: 6.170
##  Median : 7.580   Median :8.140   Median : 7.11   Median : 7.040
##  Mean   : 7.501   Mean   :7.501   Mean   : 7.50   Mean   : 7.501
##  3rd Qu.: 8.570   3rd Qu.:8.950   3rd Qu.: 7.98   3rd Qu.: 8.190
##  Max.   :10.840   Max.   :9.260   Max.   :12.74   Max.   :12.500
```

```r
sd(anscombe$y1)
```

```
## [1] 2.031568
```

```r
sd(anscombe$y3)
```

```
## [1] 2.030424
```

```r
cor.test(anscombe$x1, anscombe$y1)
```

```
##
##  Pearson's product-moment correlation
##
## data:  anscombe$x1 and anscombe$y1
## t = 4.2415, df = 9, p-value = 0.00217
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4243912 0.9506933
## sample estimates:
##       cor
## 0.8164205
```

```r
cor.test(anscombe$x2, anscombe$y2)
```

```
##
##  Pearson's product-moment correlation
##
## data:  anscombe$x2 and anscombe$y2
## t = 4.2386, df = 9, p-value = 0.002179
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4239389 0.9506402
## sample estimates:
##       cor
## 0.8162365
```

```r
cor.test(anscombe$x3, anscombe$y3)
```

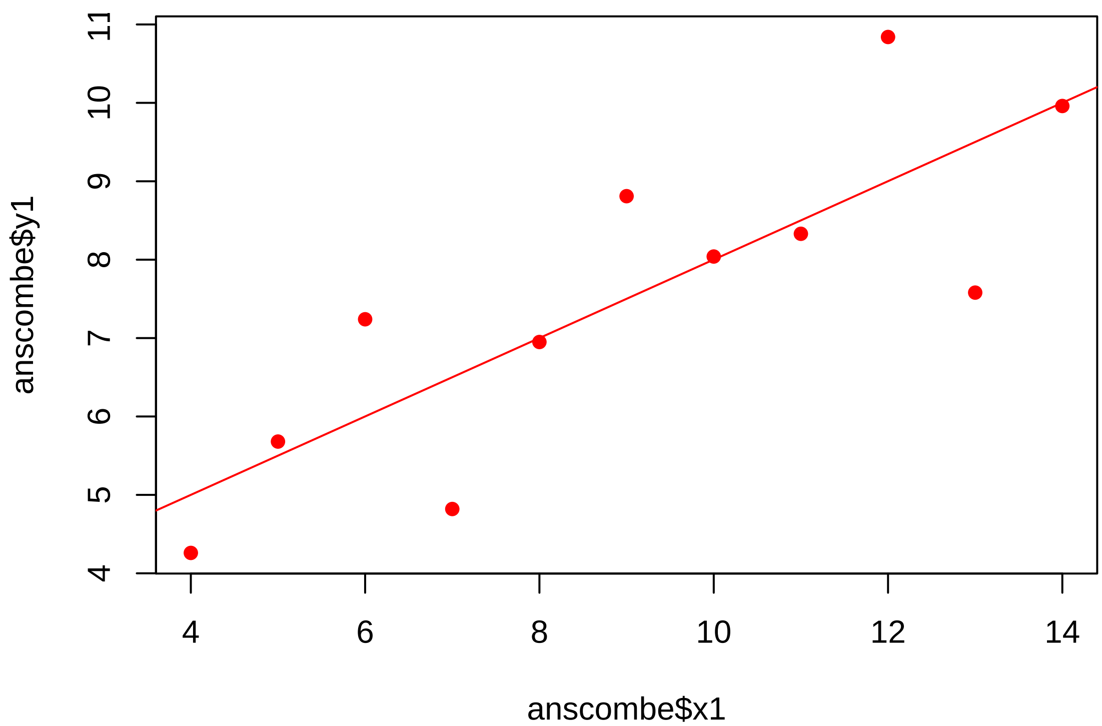
```
##
##  Pearson's product-moment correlation
##
## data:  anscombe$x3 and anscombe$y3
## t = 4.2394, df = 9, p-value = 0.002176
## alternative hypothesis: true correlation is not equal to 0
```

```
## 95 percent confidence interval:
##  0.4240623 0.9506547
## sample estimates:
##       cor
## 0.8162867
```

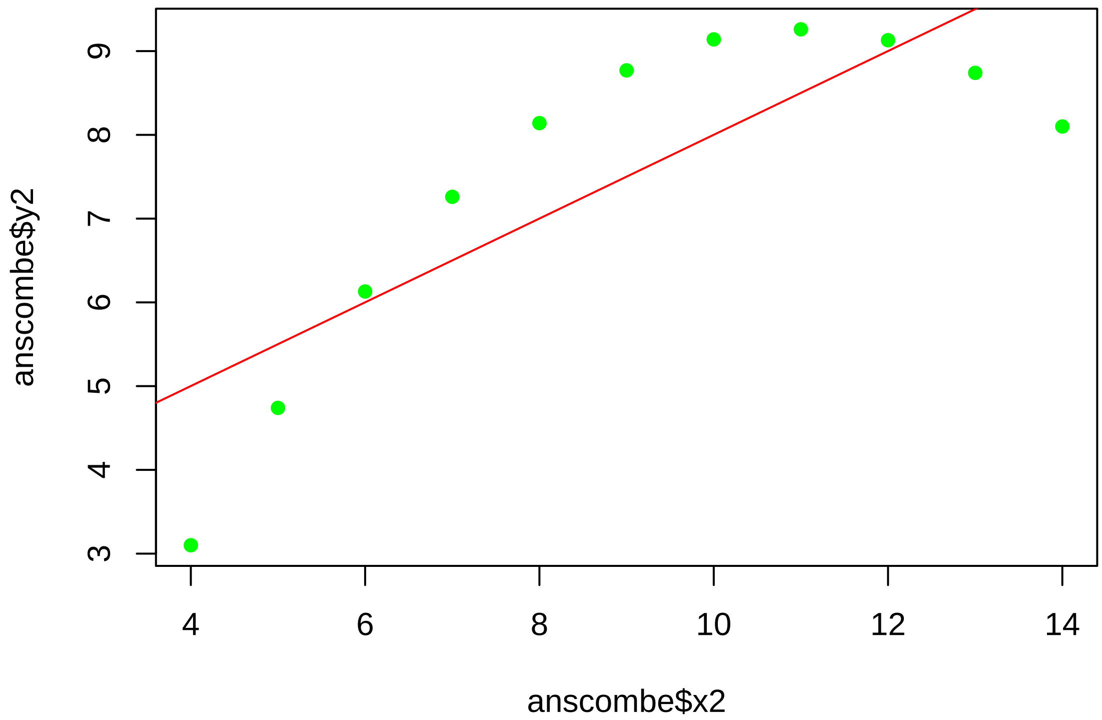
```r
cor.test(anscombe$x4, anscombe$y4)
```

```
##
##  Pearson's product-moment correlation
##
## data:  anscombe$x4 and anscombe$y4
## t = 4.243, df = 9, p-value = 0.002165
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4246394 0.9507224
## sample estimates:
##       cor
## 0.8165214
```

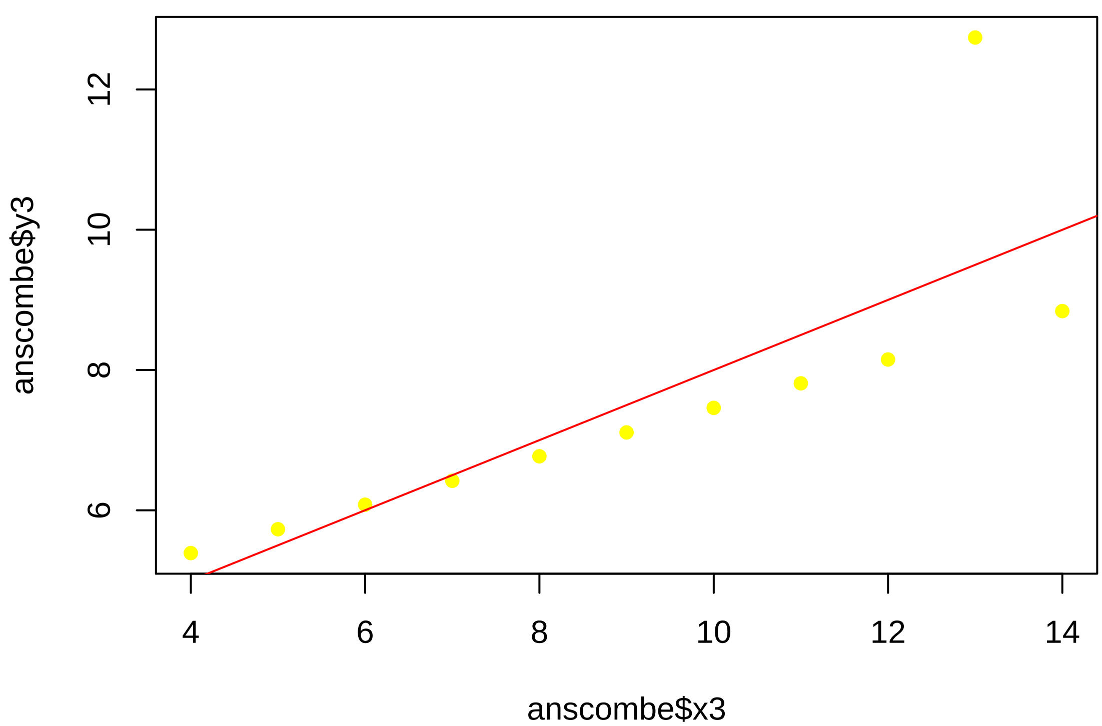
```r
plot(anscombe$x1, anscombe$y1, pch=16, col="red")
abline(lm(anscombe$y1 ~ anscombe$x1), col="red")
```



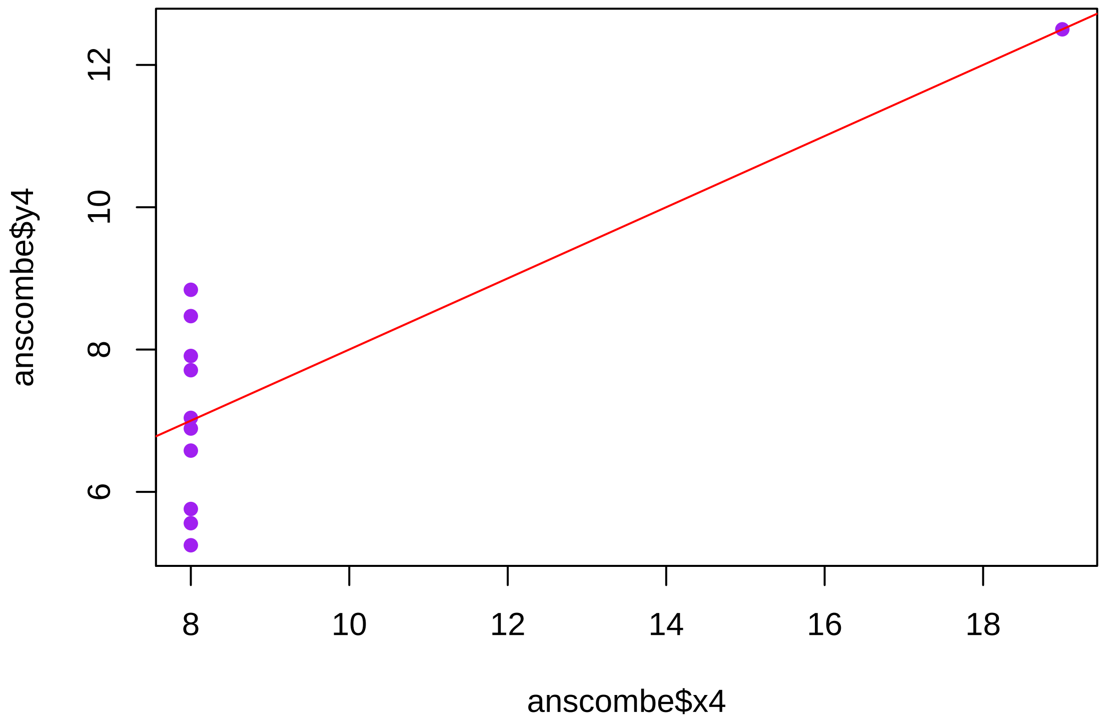

```r
plot(anscombe$x2, anscombe$y2, pch=16, col="green")
abline(lm(anscombe$y2~ anscombe$x2), col="red")
```




```
plot(anscombe$x3, anscombe$y3, pch=16, col="yellow")
abline(lm(anscombe$y3~ anscombe$x3), col="red")
```




```
plot(anscombe$x4, anscombe$y4, pch=16, col="purple")
abline(lm(anscombe$y4~ anscombe$x4), col="red")
```


anscombe$y4
12
10
8
6
8
10
12
14
16
18
anscombe$x4