

LABORATORIO_4_ANGELICA_TORRES.R

acile

2023-02-27

```
#Angelica Torres García  
#22/02/2023  
#2173388
```

```
# LABORATORIO 4. RESUMIR DATOS GRAFICAMENTE
```

```
# IMPORTAR DATOS CSV -----
```

```
library(readr)
```

```
esp.url <- paste0("https://raw.githubusercontent.com/mgtagle/",  
                  "PrincipiosEstadistica2021/main/cuadro1.csv")
```

```
inventario <- read.csv(esp.url)
```

```
inventario
```

##	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
## 7	7	10	C	C	2	17.1	15.34
## 8	8	12	C	D	2	20.6	17.22
## 9	9	16	F	C	4	18.2	15.15
## 10	10	14	F	I	5	16.1	14.66
## 11	11	8	H	D	3	14.2	17.43
## 12	12	5	H	D	6	14.8	17.45
## 13	13	12	F	I	2	19.1	14.18
## 14	14	5	C	I	2	16.7	13.40
## 15	15	12	C	S	4	18.9	10.40
## 16	16	20	H	S	3	12.4	11.52
## 17	17	15	H	C	0	17.3	14.61
## 18	18	20	F	D	1	22.7	21.46
## 19	19	15	C	C	4	15.1	17.82
## 20	20	14	C	I	3	17.7	11.38
## 21	21	14	C	S	5	13.4	8.50
## 22	22	13	C	I	4	16.2	12.80
## 23	23	14	F	D	1	18.5	18.71
## 24	24	20	F	I	4	15.0	14.48
## 25	25	21	F	C	2	18.8	14.81
## 26	26	5	H	I	4	15.8	12.01
## 27	27	2	H	I	3	16.1	11.70
## 28	28	22	C	C	3	15.4	16.03
## 29	29	22	C	I	0	17.8	14.46
## 30	30	18	C	S	1	18.5	8.47
## 31	31	16	C	I	3	14.1	11.22
## 32	32	16	C	C	5	14.8	12.34
## 33	33	17	F	C	4	15.5	16.79
## 34	34	17	F	I	6	13.8	16.06
## 35	35	18	F	S	4	13.0	13.20
## 36	36	20	H	C	2	18.2	14.30
## 37	37	22	H	C	0	22.3	16.84
## 38	38	20	H	I	3	17.8	13.84
## 39	39	17	C	I	4	13.1	11.31
## 40	40	17	C	I	6	12.8	13.20
## 41	41	16	C	C	3	13.3	13.75
## 42	42	23	F	C	3	15.6	14.60
## 43	43	23	H	C	4	16.6	12.56
## 44	43	22	C	I	5	13.0	10.88
## 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

```
#funciones para inspeccionar sus propiedades y funciones y estructura básica:
```

```
#str(inventario): mostrar la estructura general de los datos
#dim(inventario): dimensiones (i.e. tamaño) del conjunto de datos
#head(inventario, n = 5): muestra las primeras n filas
#tail(inventario, n = 5): muestra las últimas n filas
#names(inventario): nombre de las columnas
#colnames(inventario): igual names(inventario)
#summary(inventario): resumen estadístico de las variables presentes en inventario

# mostrar la estructura general de los datos

str(inventario)
```

```
## 'data.frame': 50 obs. of 7 variables:
## $ Arbol : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Fecha : int 12 12 9 9 7 10 10 12 16 14 ...
## $ Especie : chr "F" "F" "C" "H" ...
## $ Posicion : chr "C" "D" "D" "S" ...
## $ Vecinos : int 4 3 5 4 6 3 2 2 4 5 ...
## $ Diametros: num 15.3 17.8 18.2 9.7 10.8 14.1 17.1 20.6 18.2 16.1 ...
## $ Altura : num 14.78 17.07 18.28 8.79 10.18 ...
```

```
# dimensiones (num filas y columnas)
```

```
dim(inventario)
```

```
## [1] 50 7
```

```
dim(inventario)
```

```
## [1] 50 7
```

```
#head(inventario, n = 5): muestra las primeras n filas
```

```
head(inventario, n=5)
```

```
## 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
```

```
#tail(inventario, n = 5): muestra las últimas n filas
```

```
tail(inventario, n=5)
```

```
##      Arbol Fecha Especie Posicion Vecinos Diametros Altura
## 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
```

```
# nombre de las primeras cinco columnas
```

```
names(inventario [1:5])
```

```
## [1] "Arbol"      "Fecha"      "Especie"    "Posicion"   "Vecinos"
```

```
names (inventario)
```

```
## [1] "Arbol"      "Fecha"      "Especie"    "Posicion"   "Vecinos"    "Diametros"
## [7] "Altura"
```

```
#colnames(inventario): igualnames(inventario)
```

```
colnames(inventario)
```

```
## [1] "Arbol"      "Fecha"      "Especie"    "Posicion"   "Vecinos"    "Diametros"
## [7] "Altura"
```

```
#summary(inventario): resumen estadístico de las variables presentes en inventario
```

```
summary(inventario)
```

```
##      Arbol      Fecha      Especie      Posicion
## Min.   : 1.00   Min.   : 2.00   Length:50   Length:50
## 1st Qu.:13.25   1st Qu.:12.00   Class :character   Class :character
## Median :25.50   Median :16.00   Mode  :character   Mode  :character
## Mean   :25.48   Mean   :15.94
## 3rd Qu.:37.75   3rd Qu.:20.75
## Max.   :50.00   Max.   :25.00
##      Vecinos      Diametros      Altura
## Min.   :0.00   Min.   : 7.70   Min.   : 8.47
## 1st Qu.:2.25   1st Qu.:13.88   1st Qu.:11.78
## Median :3.00   Median :15.70   Median :14.24
## Mean   :3.34   Mean   :15.79   Mean   :13.94
## 3rd Qu.:4.00   3rd Qu.:18.10   3rd Qu.:16.05
## Max.   :6.00   Max.   :22.70   Max.   :21.46
```

```
# Resumen estadístico básico de las columnas 3 a 5 columnas
```

```
summary(inventario[,3:5])
```

```
##      Especie      Posicion      Vecinos
## Length:50      Length:50      Min.    :0.00
## Class :character Class :character 1st Qu.:2.25
## Mode  :character Mode  :character Median :3.00
##                                     Mean  :3.34
##                                     3rd Qu.:4.00
##                                     Max.   :6.00
```

```
is.factor(inventario$Posicion)
```

```
## [1] FALSE
```

```
inventario$Posicion <- factor(inventario$Posicion)
is.factor(inventario$Posicion)
```

```
## [1] TRUE
```

```
summary(inventario[,3:5])
```

```
##      Especie      Posicion      Vecinos
## Length:50      C:14      Min.    :0.00
## Class :character D: 9      1st Qu.:2.25
## Mode  :character I:19      Median :3.00
##                                     S: 8      Mean   :3.34
##                                     3rd Qu.:4.00
##                                     Max.    :6.00
```

```
# TABLAS DE FRECUENCIA -----
```

#Una tabla de frecuencias muestra los recuentos de cada categoría. En R, tenemos la función table () para obtener este tipo de tablas.

```
freq_position <- table(inventario$Posicion)
freq_position
```

```
##
## C D I S
## 14 9 19 8
```

```
#frecuencias relativas.
```

```
prop_position <- freq_position / sum(freq_position)
prop_position
```

```
##
##      C      D      I      S
## 0.28 0.18 0.38 0.16
```

```
#Si desea expresar las proporciones como porcentajes, multiplique prop_position por 100:
```

```
perc_position = 100 * prop_position
perc_position
```

```
##
##  C  D  I  S
## 28 18 38 16
```

```
# GRAFICAS BARPLOT Y PIE -----
```

```
#Hay dos gráficos más comunes que se utilizan para visualizar frecuencias:
```

```
#Gráficas de barras (barplot)
```

```
#Gráficas de pastel (pie)
```

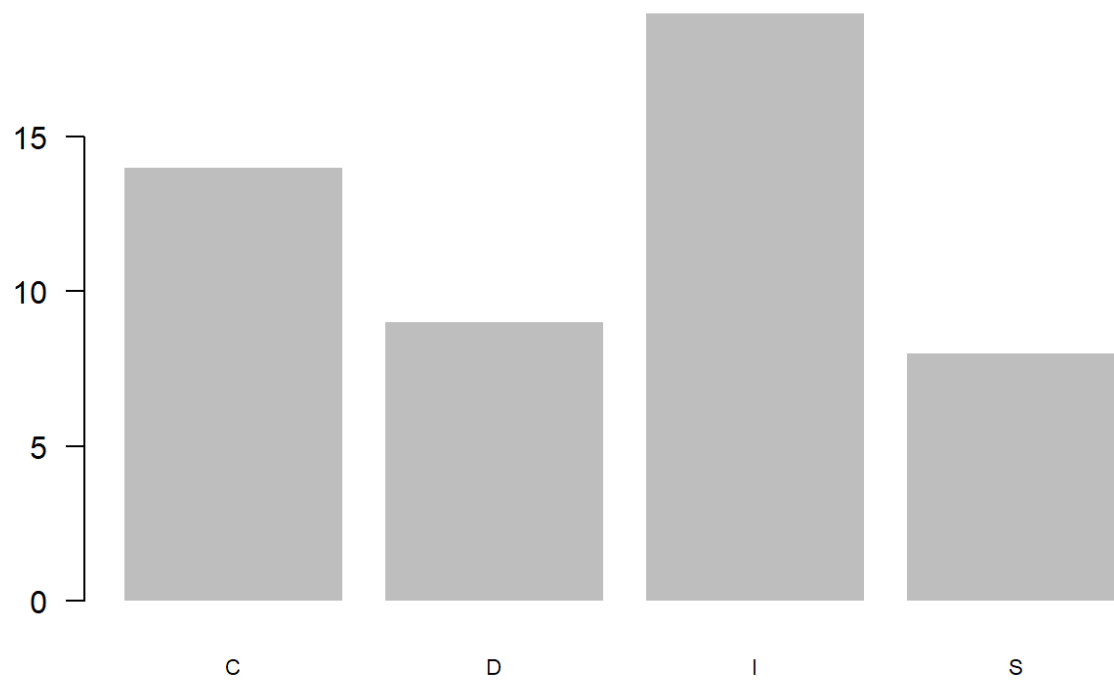
```
#El uso de barplot () incluye los argumentos las,border y cex.names:
```

```
#las = 1: muestra las frecuencias perpendiculares al eje-y.
```

```
#border = NA: elimina el borde negro alrededor de las barras.
```

```
#cex.names = 0.7: reduce los tamaños de las etiquetas de categoría (para que todas quepan en el gráfico).
```

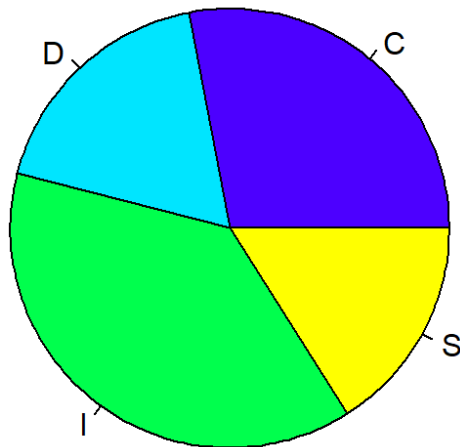
```
barplot(freq_position, las = 1, border = NA, cex.names = 0.7)
```



#Gráfico circular o pie.

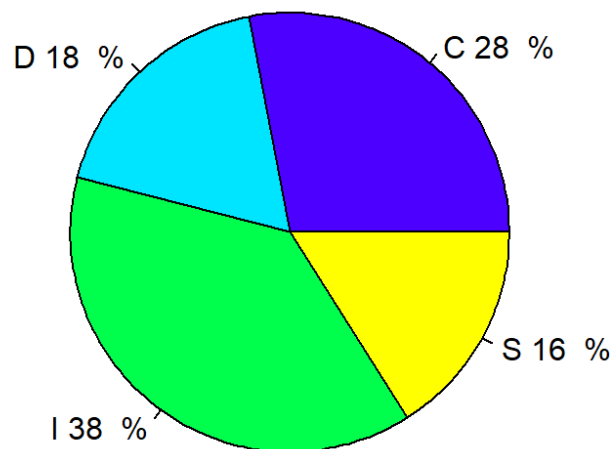
*# topo.colors es una paleta de colores pre establecidas en R y
el paréntesis indica el # de colores a usar*

```
pie(freq_position, col=topo.colors(4))
```



Si se desea mostrar las frecuencias, se puede hacer algo como esto:

```
pie(freq_position, col = topo.colors(4),  
    labels = paste(levels(inventario$Posicion), round(perc_position,)," %"))
```

```
# AUTOESTUDIO -----
```

```
#Completar una tabla de frecuencia y su representación gráfica (barplot y pie) para la variable Es
pecie del conjunto de datos inventario
```

```
#tablas de frecuencia
```

```
freq_position <- table(inventario$Especie)
freq_position
```

```
##
##  C  F  H
## 22 14 14
```

```
prop_position <- freq_position / sum(freq_position)
prop_position
```

```
##
##    C    F    H
## 0.44 0.28 0.28
```

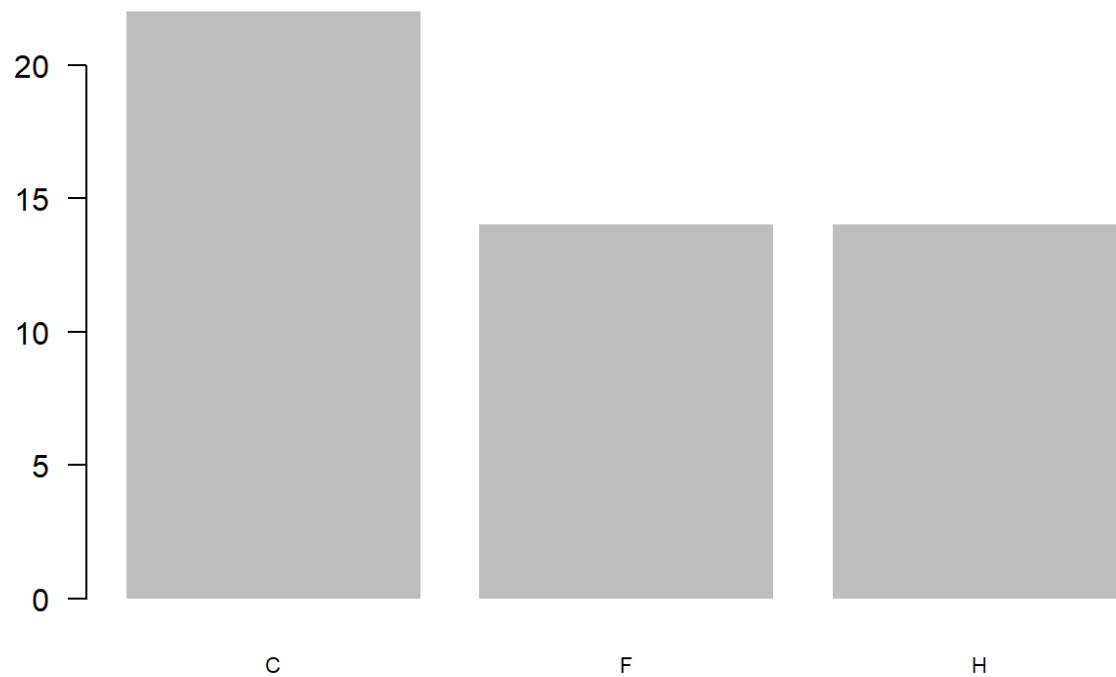
```
#proporciones como porcentajes, multiplique prop_position por 100:
```

```
perc_position = 100 * prop_position  
perc_position
```

```
##  
## C F H  
## 44 28 28
```

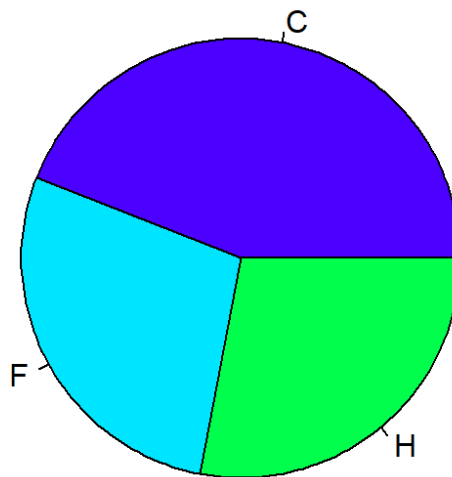
```
#Gráficas barplot y pie
```

```
barplot(freq_position, las = 1, border = NA, cex.names = 0.7)
```



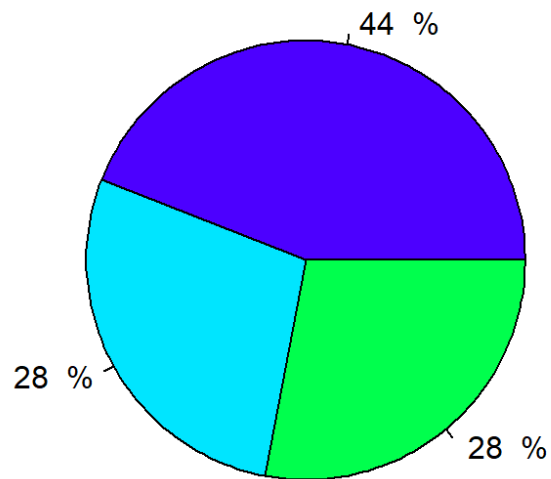
```
#Gráfico circular o pie.
```

```
pie(freq_position, col=topo.colors(4))
```



```
#mostrar frecuencias
```

```
pie(freq_position, col = topo.colors(4), labels = paste(levels(inventario$Especie), round(perc_pos  
ition, 2), " %"))
```



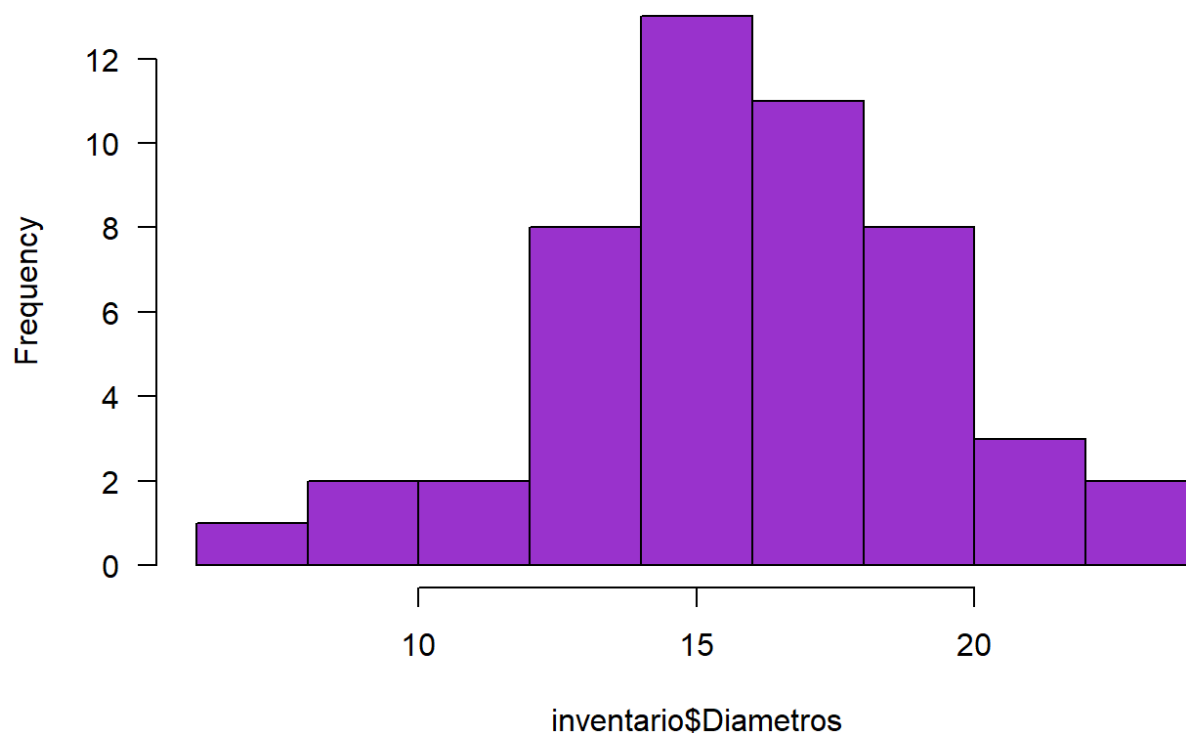
```
# REPRESENTACIÓN DE VARIABLES CUANTITATIVAS -----
```

```
#Histogramas
```

```
#Vamos a aplicar la función hist para la variable Diametros del conjunto inventario y guardar la salida en un objeto llamado diam_hist.
```

```
diam_hist <- hist(inventario$Diametros, las = 1, col = '#9932CC')
```

Histogram of inventario\$Diametros



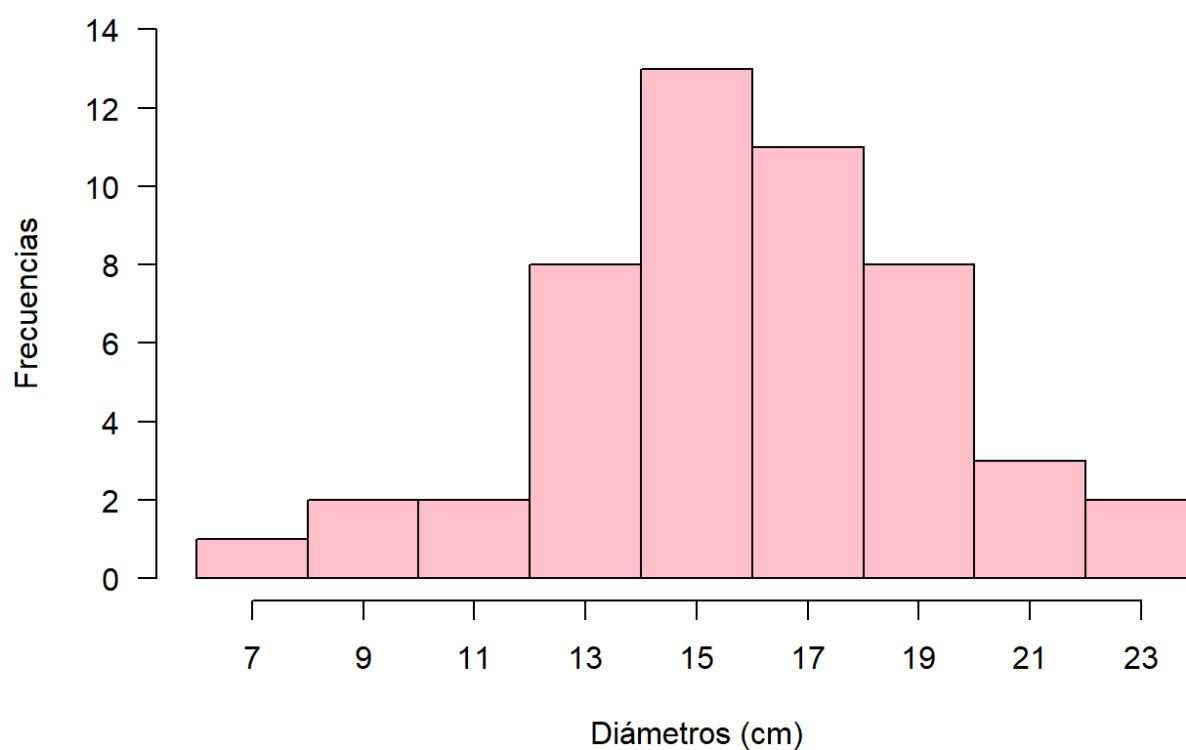
```
diam_hist
```

```
## $breaks
## [1]  6  8 10 12 14 16 18 20 22 24
##
## $counts
## [1]  1  2  2  8 13 11  8  3  2
##
## $density
## [1] 0.01 0.02 0.02 0.08 0.13 0.11 0.08 0.03 0.02
##
## $mids
## [1]  7  9 11 13 15 17 19 21 23
##
## $xname
## [1] "inventario$Diametros"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
```

```
diam_hist$breaks
```

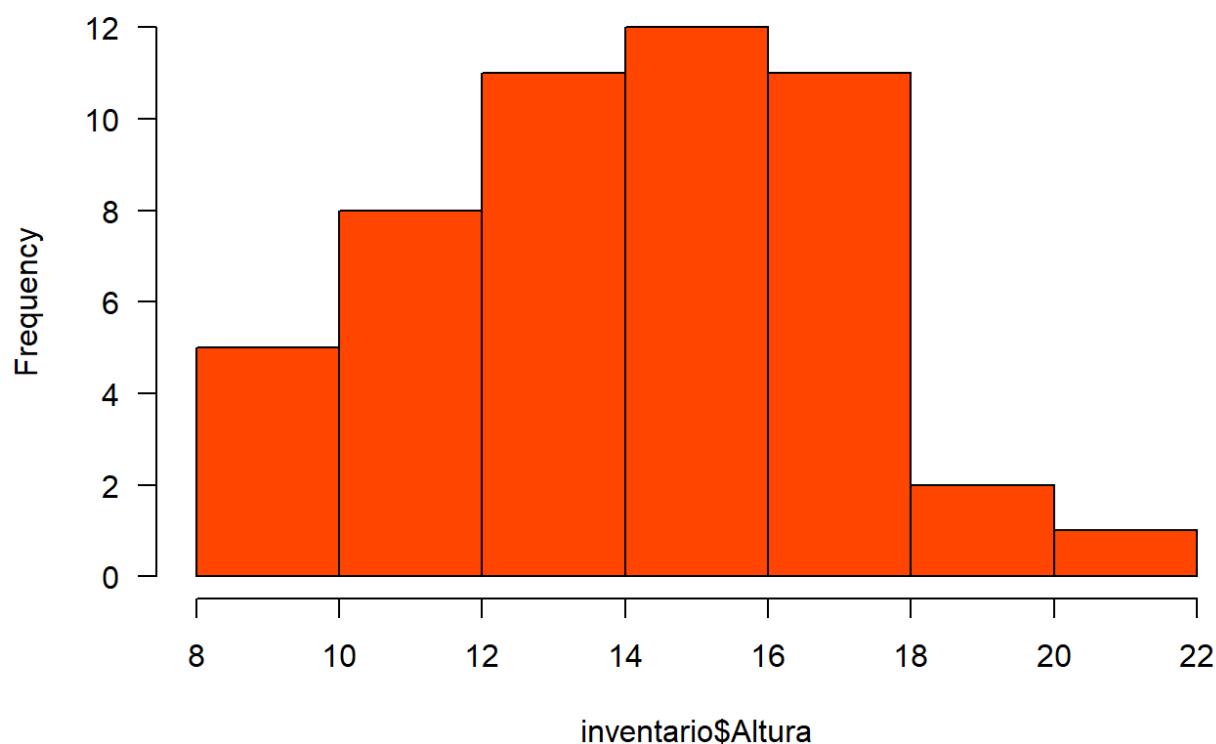
```
## [1] 6 8 10 12 14 16 18 20 22 24
```

```
h1 <- hist(inventario$Diametros, xaxt = "n",
           breaks = c(6, 8, 10, 12, 14, 16, 18, 20, 22, 24),
           col = "#FFC0CB", xlab="Diámetros (cm)",
           ylab= "Frecuencias",
           main = "",
           las = 1,
           ylim = c(0,14))
axis(1, h1$mids)
```



```
# AUTOESTUDIO -----
#Realizar el mismo procedimiento para la variable Altura
#HISTOGRAMA
alt_hist <- hist(inventario$Altura, las = 1, col = '#FF4500')
```

Histogram of inventario\$Altura



```
alt_hist
```

```
## $breaks
## [1]  8 10 12 14 16 18 20 22
##
## $counts
## [1]  5  8 11 12 11  2  1
##
## $density
## [1] 0.05 0.08 0.11 0.12 0.11 0.02 0.01
##
## $mids
## [1]  9 11 13 15 17 19 21
##
## $xname
## [1] "inventario$Altura"
##
## $equidist
## [1] TRUE
##
## attr(,"class")
## [1] "histogram"
```

```
alt_hist$breaks
```

```
## [1] 8 10 12 14 16 18 20 22
```

```
#personalizar grafica
```

```
h1 <- hist(inventario$Altura, xaxt = "n",  
           breaks = c(8, 10, 12, 14, 16, 18, 20, 22),  
           col = "#87cefa", xlab="Diámetros (cm)",  
           ylab= "Frecuencias",  
           main = "",  
           las = 1,  
           ylim = c(0,14))  
axis(1, h1$mids)
```

