

Laboratorio_4.R

yesiv

2023-03-01

```
#Yesenia Villarreal Torres  
#1109559  
#Laboratorio 4  
#27 de febrero 2023
```

```
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 |

```
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 ...
```

```
dim(inventario)
```

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

```
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)
```

```
## 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
```

```
names(inventario)
```

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

```
colnames(inventario)
```

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

```
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
```

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

```
dim(inventario)
```

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

```
#nombre de las primeras 5 columnas
```

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

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

```
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
```

```
summary(inventario)
```

```
##   Arbol      Fecha      Especie      Posicion  Vecinos
## Min.   : 1.00  Min.   : 2.00  Length:50      C:14      Min.   :0.00
## 1st Qu.:13.25  1st Qu.:12.00  Class :character D: 9      1st Qu.:2.25
## Median :25.50  Median :16.00  Mode  :character I:19      Median :3.00
## Mean   :25.48  Mean   :15.94      S: 8      Mean   :3.34
## 3rd Qu.:37.75  3rd Qu.:20.75      3rd Qu.:4.00
## Max.   :50.00  Max.   :25.00      Max.   :6.00
## Diametros      Altura
## Min.   : 7.70  Min.   : 8.47
## 1st Qu.:13.88  1st Qu.:11.78
## Median :15.70  Median :14.24
## Mean   :15.79  Mean   :13.94
## 3rd Qu.:18.10  3rd Qu.:16.05
## Max.   :22.70  Max.   :21.46
```

```
# Tablas de frecuencia -----
```

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

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

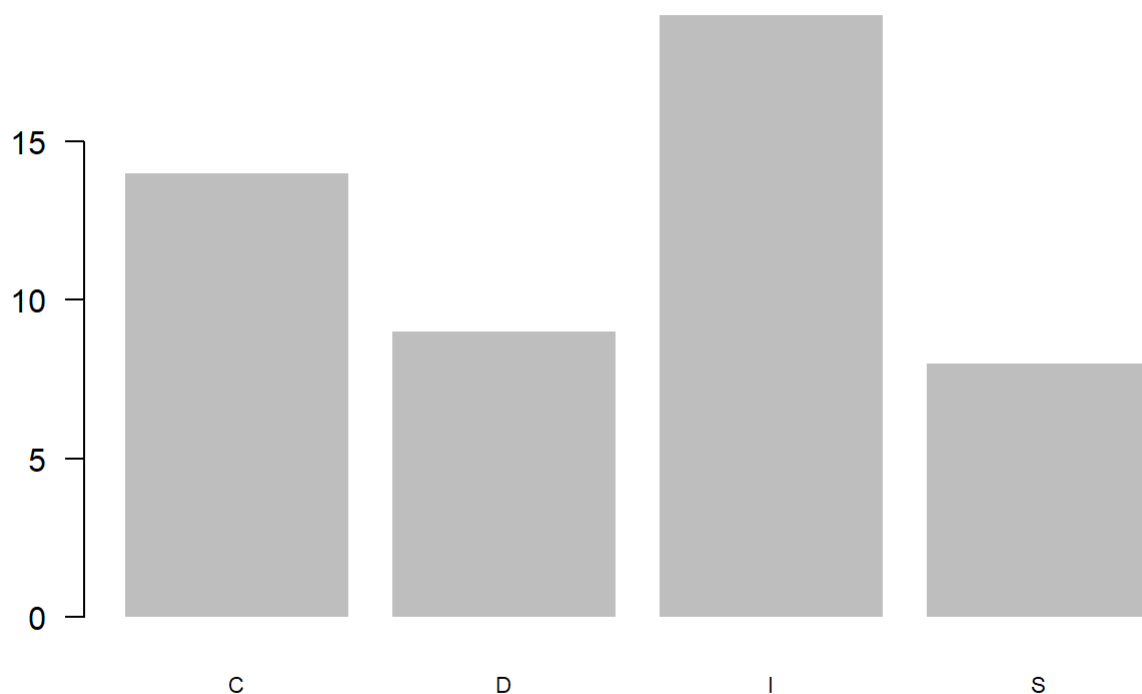
```
prop_position <- freq_position / sum(freq_position)
prop_position
```

```
##  
##      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
```

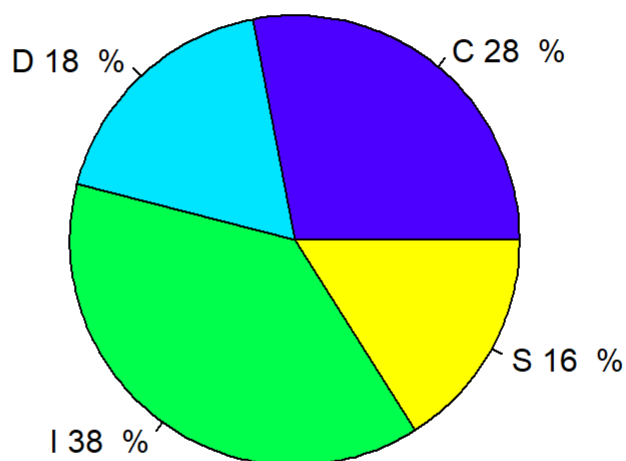
```
# Gráficas barplot y pie -----  
  
#Gráficas de barras (barplot)  
#Gráficas de pastel (pie)  
  
barplot(freq_position, las = 1, border = NA, cex.names = 0.7)
```



```
#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).
```

#Gráfico circular o pie. El otro tipo común de gráfico para ver frecuencias es un gráfico circular. R proporciona la función `pie()` para producir estos gráficos:

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



#Autoestudio

```
# topo.colors es una paleta de colores pre establecidas en R y
# el paréntesis indica el # de colores a usar
#Completar una tabla de frecuencia y su representación gráfica (barplot y pie) para la variable
Especie del conjunto de datos inventario
freq_Especie <- table(inventario$Especie)
freq_Especie
```

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

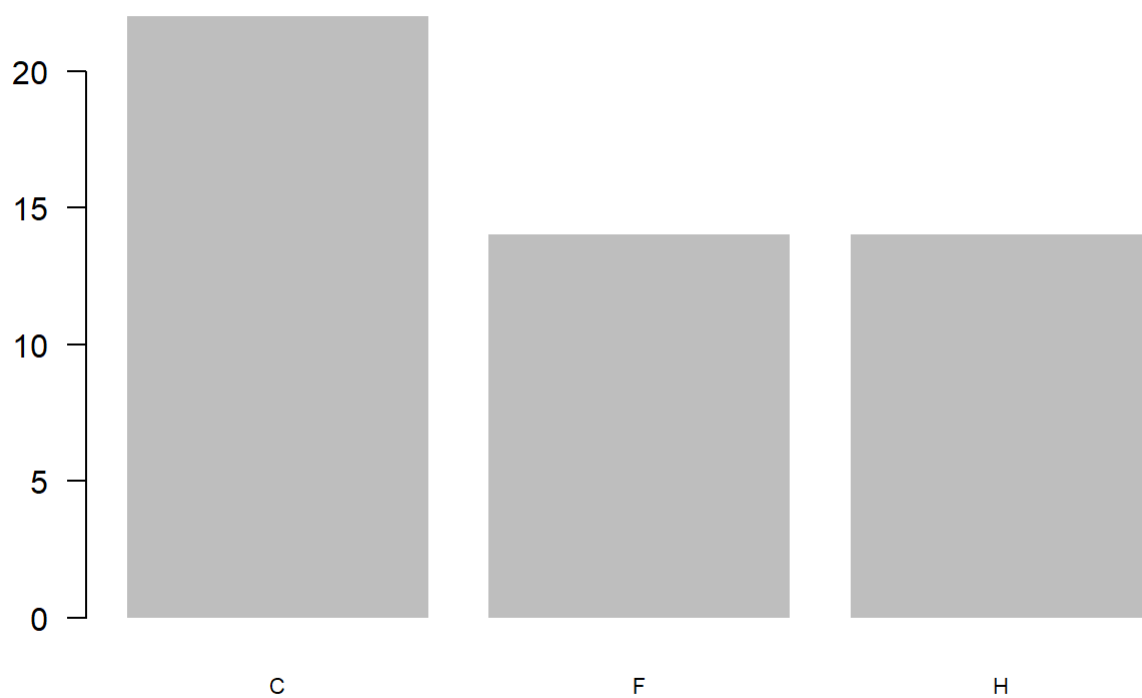
```
prop_Especie <- freq_Especie / sum(freq_Especie)  
prop_Especie
```

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

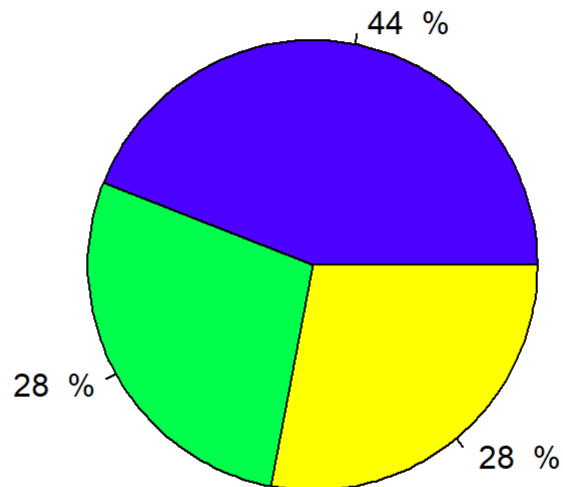
```
perc_Especie = 100 * prop_Especie  
perc_Especie
```

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

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




```
pie(freq_Especie, col = topo.colors(3),
    labels = paste(levels(inventario$Especie), round(perc_Especie, 2), "%"))
```



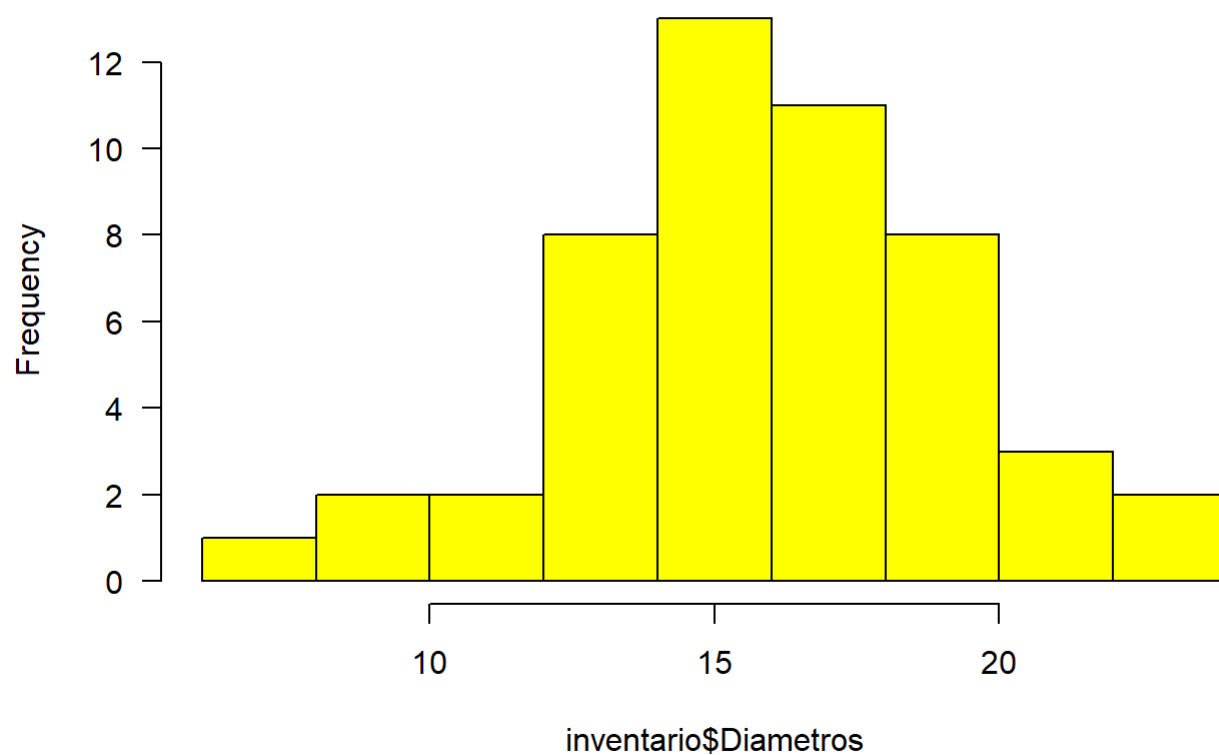
```
# Representación de variables cuantitativas -----

#histogramas
#boxplots o gráfica de cajas

#Histogramas
#Un histograma es un tipo de gráfico que muestra la distribución de datos numéricos

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

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"
```

```
#breaks: puntos de ruptura (corte) de los intervalos de clase  
#counts: número de observaciones en cada categoría  
#density: densidad  
#mids: punto central del intervalo  
#xname: nombre del objeto (variable) que se esta graficando  
#equidist: ¿Los categorías tienen el mismo ancho?  
# attr: Tipo de clase
```

```
diam_hist$breaks
```

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

```
diam_hist$mids
```

```
## [1] 7 9 11 13 15 17 19 21 23
```

```
diam_hist$counts
```

```
## [1] 1 2 2 8 13 11 8 3 2
```

```
diam_hist$density
```

```
## [1] 0.01 0.02 0.02 0.08 0.13 0.11 0.08 0.03 0.02
```

```
diam_hist$xname
```

```
## [1] "inventario$Diametros"
```

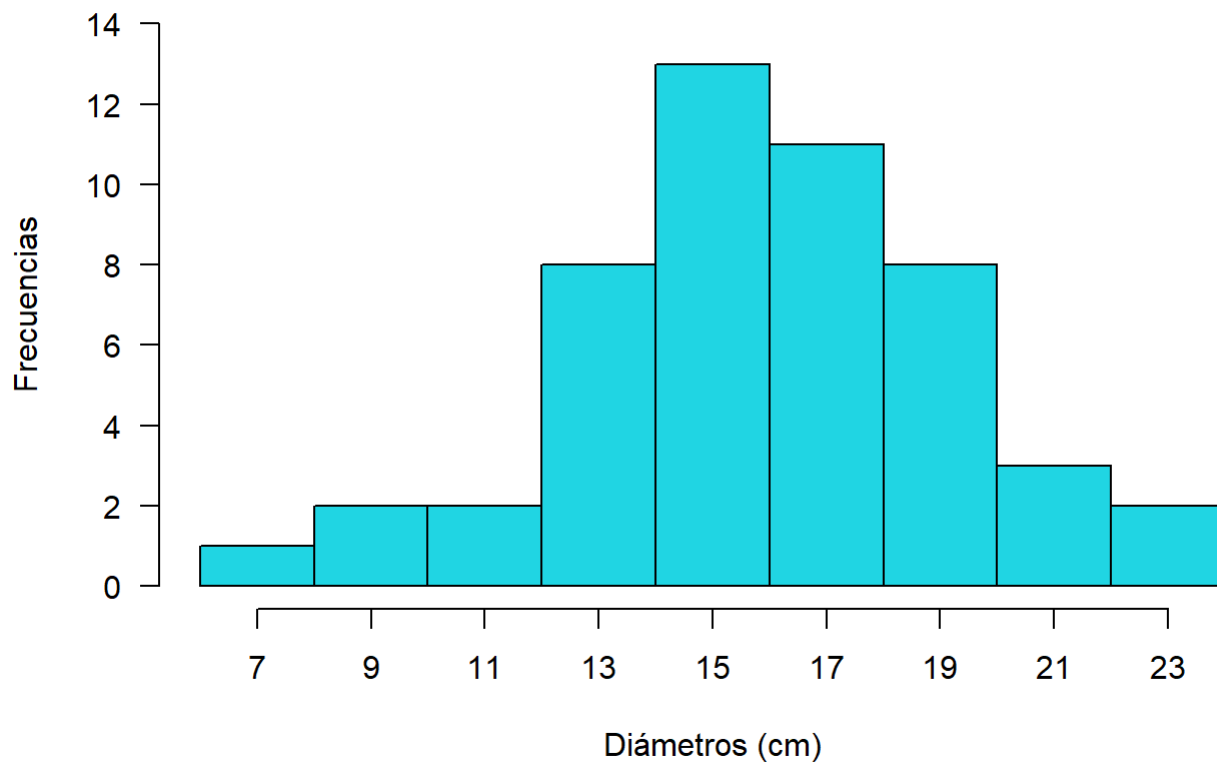
```
diam_hist$equidist
```

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

```
diam_hist$attr
```

```
## NULL
```

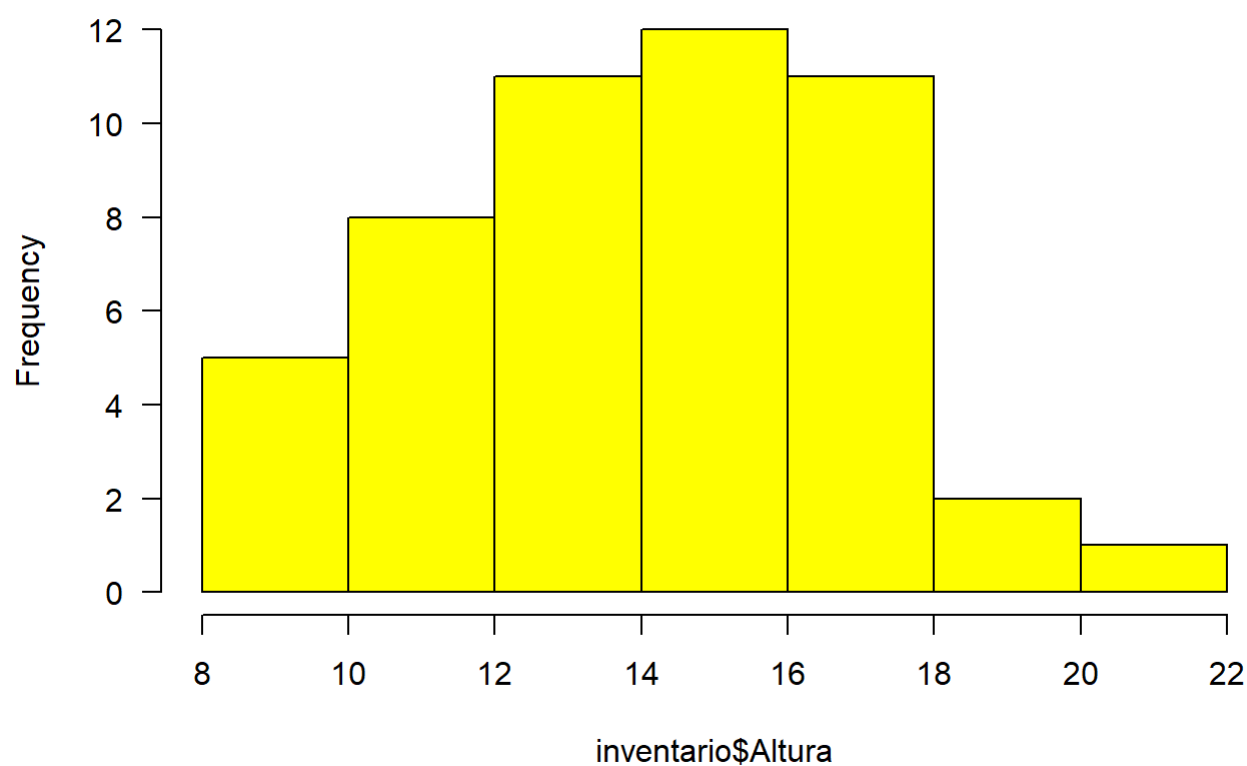
```
h1 <- hist(inventario$Diametros, xaxt = "n",
           breaks = c(6, 8, 10, 12, 14, 16, 18, 20, 22, 24),
           col = "#00cFDFDF", xlab="Diámetros (cm)",
           ylab= "Frecuencias",
           main = "",
           las = 1,
           ylim = c(0,14))
axis(1, h1$mids)
```



Autoestudio Realizar el mismo procedimiento para la variable Alt -----

```
altura_hist <- hist(inventario$Altura, las = 1, col = 'yellow')
```

Histogram of inventario\$Altura



altura_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"
```

```
altura_hist$breaks
```

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

```
altura_hist$mids
```

```
## [1] 9 11 13 15 17 19 21
```

```
altura_hist$counts
```

```
## [1] 5 8 11 12 11 2 1
```

```
altura_hist$density
```

```
## [1] 0.05 0.08 0.11 0.12 0.11 0.02 0.01
```

```
altura_hist$xname
```

```
## [1] "inventario$Altura"
```

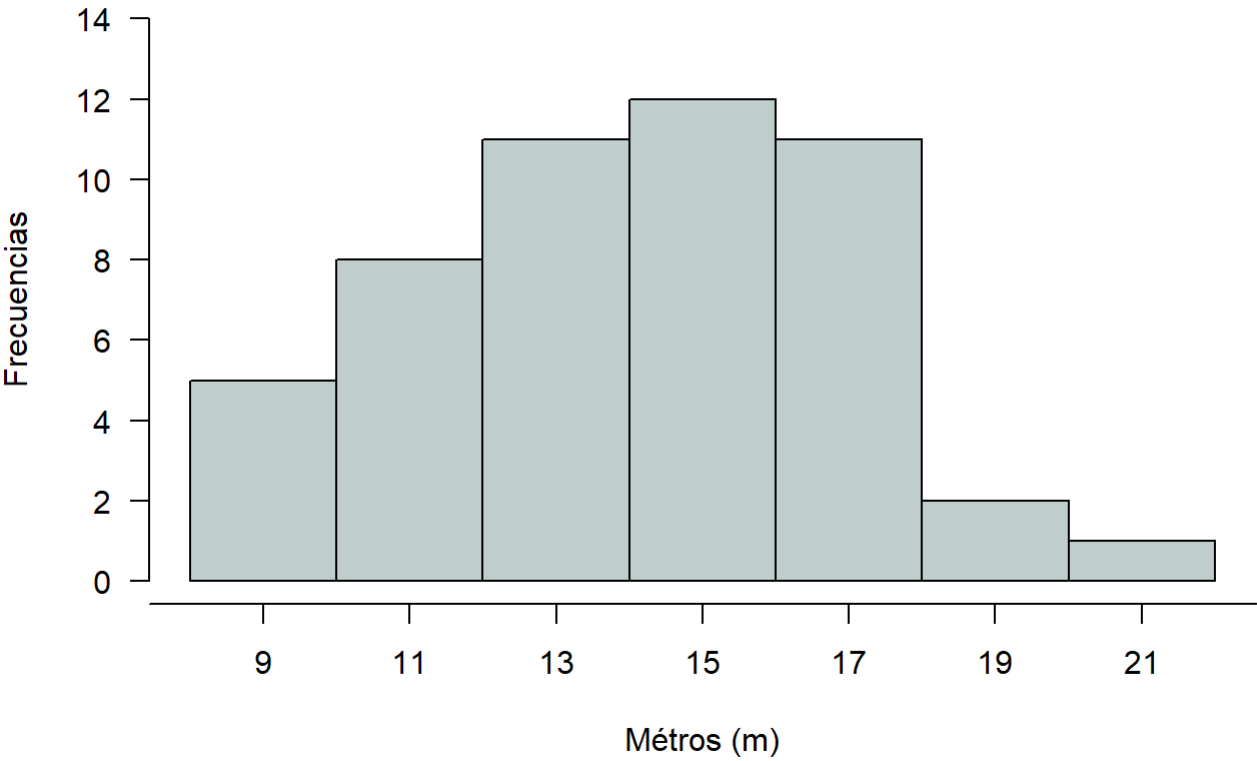
```
altura_hist$equidist
```

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

```
altura_hist$attr
```

```
## NULL
```

```
h2 <- hist(inventario$Altura, xaxt = "n",  
           breaks = c(8, 10, 12, 14, 16, 18, 20, 22),  
           col = "#12444444", xlab="Métros (m)",  
           ylab= "Frecuencias",  
           main = "",  
           las = 1,  
           ylim = c(0,14))  
axis(1, h1$mids)
```



#FIN 1