

Laboratorio_sem_4.R

toryf

2023-02-27

```
# LABORATORIO 04: Resumir datos gráficamente
# Histogramas
# 27/03/2023
# Nombre: Luis Miguel Toribio Ferrer

# Importar datos -----

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

```
tail(inventario)
```

```
##   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
```

```
# funciones para revisar los conjuntos de datos -----

# 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 (i.e. tamaño) del conjunto de datos
```

```
dim(inventario)
```

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

```
# 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
```

```
# 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 columnas
```

```
names(inventario)
```

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

```
names(inventario[,4:7])
```

```
## [1] "Posicion" "Vecinos" "Diametros" "Altura"
```

```
# igual names(inventario)
```

```
colnames(inventario)
```

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

```
# 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
```

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

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

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

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

```
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 -----
```

```
# Frecuencia absoluta
```

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

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

```
# Frecuencia relativa
```

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

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

```
# Frecuencia en porcentaje
```

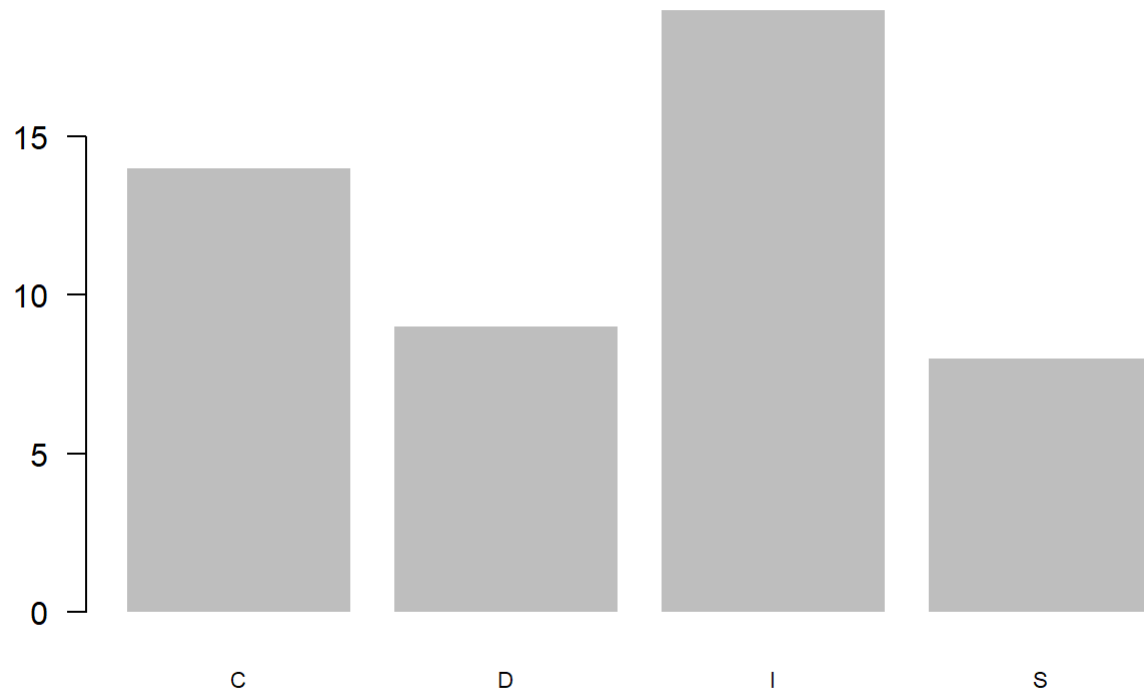
```
perc_position = 100 * prop_position
perc_position
```

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

```
# # Representación gráfica para variables cualitativas -----
```

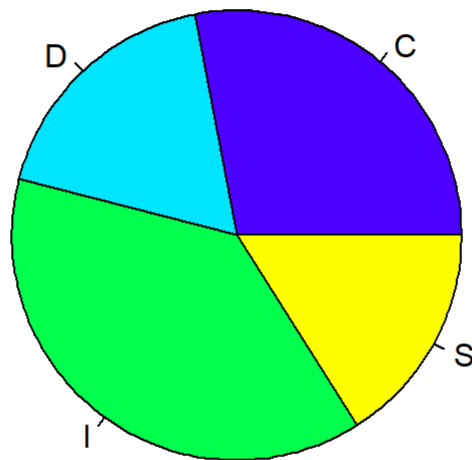
```
#Gráficas barplot
```

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

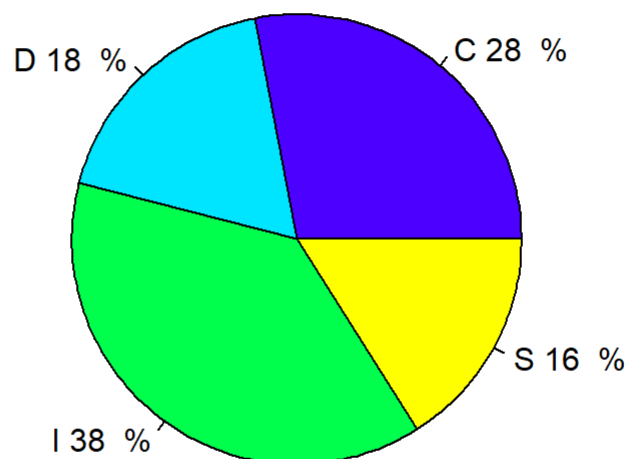


```
#grafico circula o de pie
```

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



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



Completar una tabla de frecuencia y su representación gráfica (barplot y pie)

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

```
prop_especie <- frep_especie / sum(frep_especie)
prop_especie
```

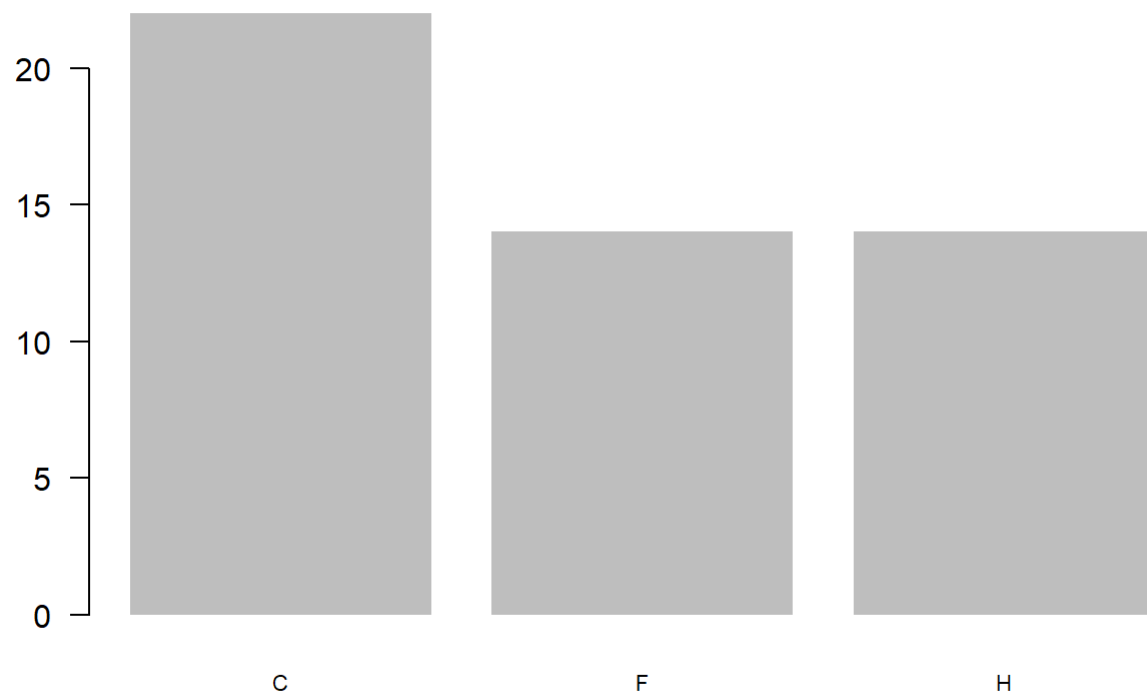
```
##
##      C      F      H
## 0.44 0.28 0.28
```

```
perc_especie = 100 * prop_especie
perc_especie
```

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

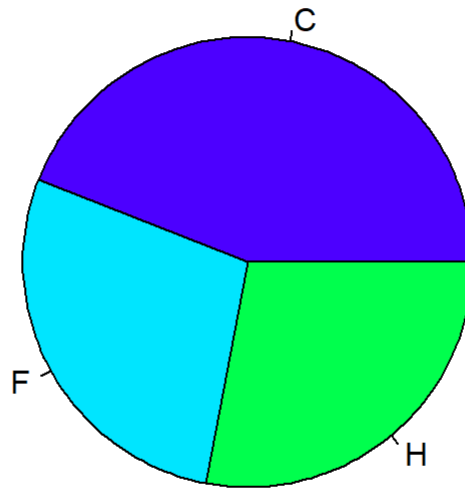
#Gráficas barplot

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

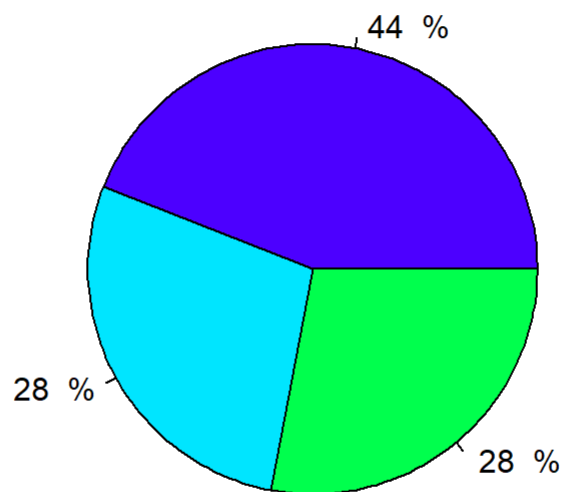


```
#grafico circular o de pie
```

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

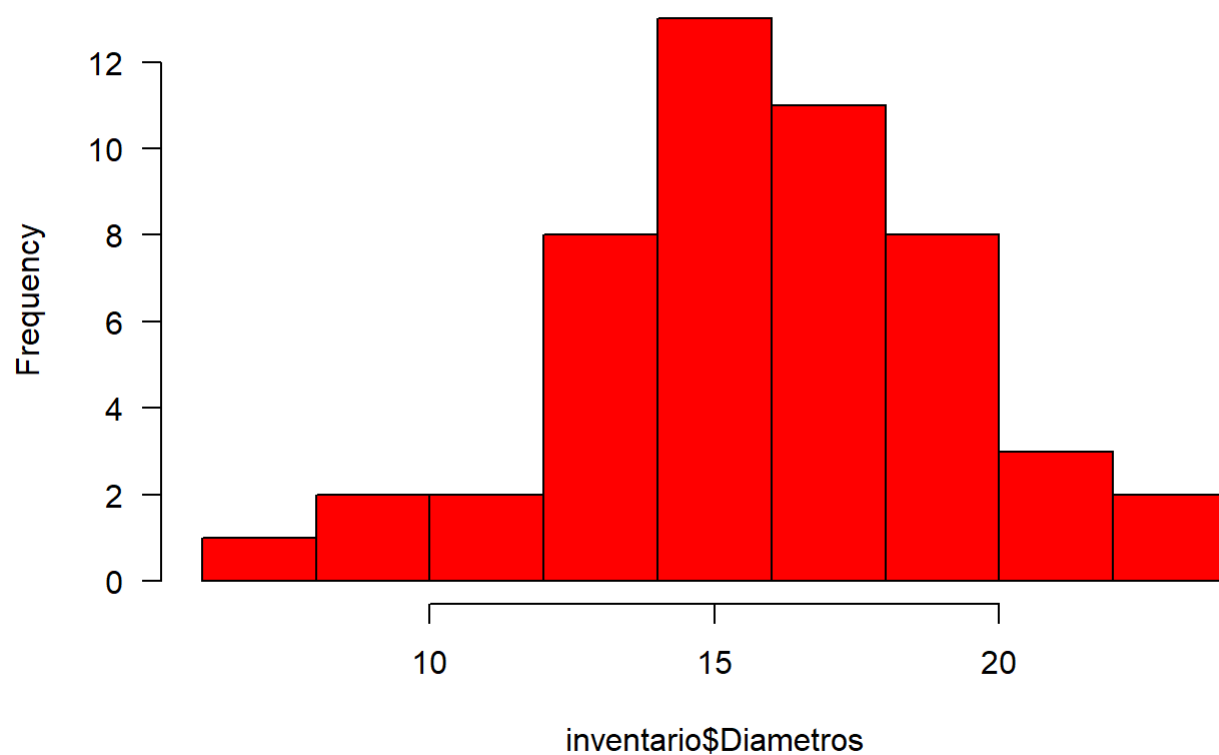



```
pie(frep_especie, col = topo.colors(4),  
    labels = paste(levels(inventario$Especie), round(perc_especie, 2), "%"))
```



```
# Representación gráfica para variables cuantitativas -----  
diam_hist <- hist(inventario$Diametros, las = 1, col = 'red')
```

Histogram of inventario\$Diametros



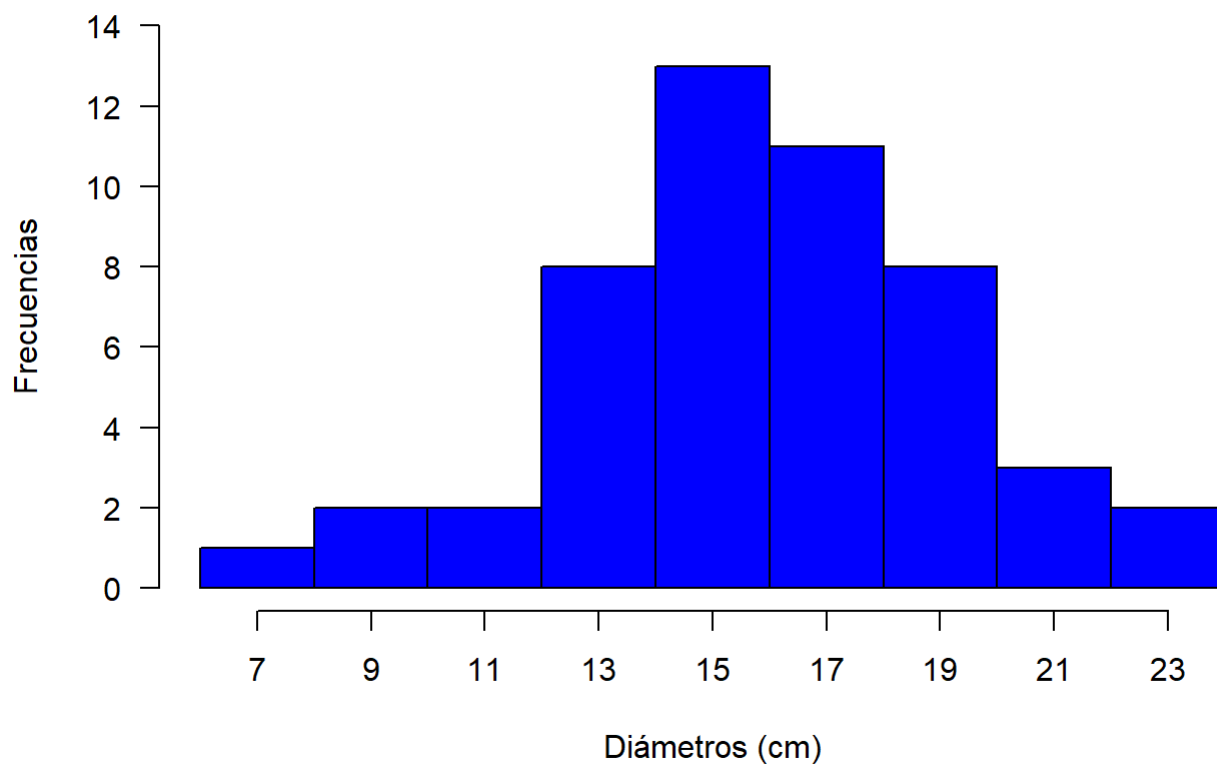
diam_hist

```
## $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
```

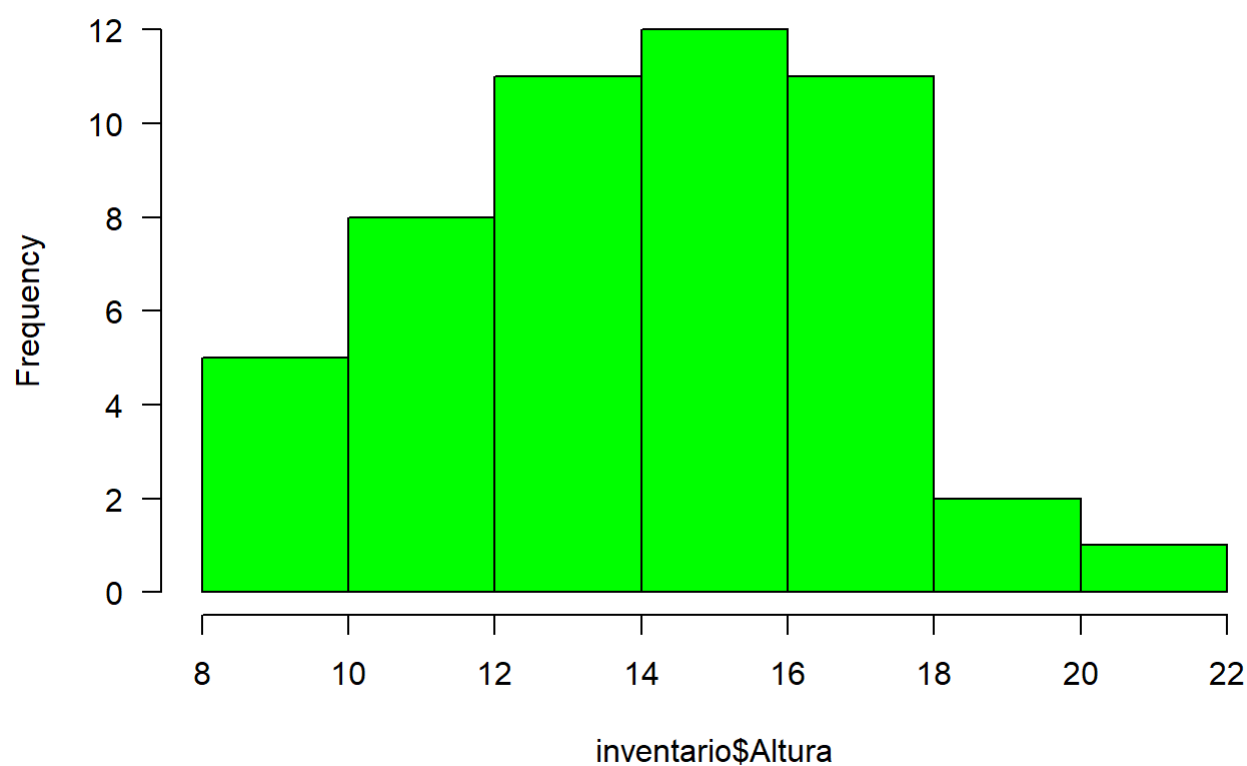
```
h1 <- hist(inventario$Diametros, xaxt = "n",  
           breaks = c(6, 8, 10, 12, 14, 16, 18, 20, 22, 24),  
           col = "blue", xlab="Diámetros (cm)",  
           ylab= "Frecuencias",  
           main = "",  
           las = 1,  
           ylim = c(0,14))  
  
axis(1, h1$mids)
```



#Realizar el mismo procedimiento para la variable Altura.

```
altura_hist <- hist(inventario$Altura, las = 1, col="green")
```

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
```

```
h2 <- hist(inventario$Altura, xaxt = "n",  
           breaks = c(8, 10, 12, 14, 16, 18, 20, 22),  
           col = "orange", xlab="Diámetros (cm)",  
           ylab= "Frecuencias",  
           main = "",  
           las = 1,  
           ylim = c(0,14))
```

```
axis(1, h2$mids)
```

