

# Estadística con R

Carlos Meneses

14/5/2020

## Tablas de Frecuencias

```
genero <- sample(c("H", "M"), size = 100, replace = TRUE)
respuesta <- sample(c("Si", "No"), size = 100, replace = TRUE)
ciudad <- sample(c("Quito", "Guayaquil", "Cuenca", "Ambato", "Ibarra"), size = 100, replace = TRUE)
f_absolutas <- table(genero, respuesta, ciudad)
f_absolutas
```

```
## , , ciudad = Ambato
##
##      respuesta
## genero No Si
##      H  2  4
##      M  6  6
##
## , , ciudad = Cuenca
##
##      respuesta
## genero No Si
##      H  4  2
##      M  5  1
##
## , , ciudad = Guayaquil
##
##      respuesta
## genero No Si
##      H  7  4
##      M  9  5
##
## , , ciudad = Ibarra
##
##      respuesta
## genero No Si
##      H  9  5
##      M  5  3
##
## , , ciudad = Quito
##
##      respuesta
## genero No Si
```

```
##      H  8  4
##      M  5  6
```

## Moda

```
moda <- function(x) {
  frec <- table(x)
  names(which(frec == max(frec)))
}

moda(genero)
```

```
## [1] "M"
```

```
moda(respuesta)
```

```
## [1] "No"
```

```
moda(ciudad)
```

```
## [1] "Guayaquil"
```

## Tabla de Frecuencias Absoluta

```
ftable(genero, respuesta, ciudad)
```

```
##              ciudad Ambato Cuenca Guayaquil Ibarra Quito
## genero respuesta
## H      No              2      4      7      9      8
##      Si              4      2      4      5      4
## M      No              6      5      9      5      5
##      Si              6      1      5      3      6
```

Para elegir las columnas en la función `ftable()` se especifica las columnas con el parámetro `col.vars`

```
ftable(genero, respuesta, ciudad, col.vars = c("genero", "respuesta"))
```

```
##      genero      H      M
##      respuesta No Si No Si
## ciudad
## Ambato      2  4  6  6
## Cuenca      4  2  5  1
## Guayaquil   7  4  9  5
## Ibarra      9  5  5  3
## Quito       8  4  5  6
```

## Frecuencias Relativas

```
prop.table(ftable(genero, respuesta, ciudad, col.vars = c("genero", "respuesta")))
```

```
##          genero      H      M
##          respuesta  No  Si  No  Si
## ciudad
## Ambato          0.02 0.04 0.06 0.06
## Cuenca          0.04 0.02 0.05 0.01
## Guayaquil       0.07 0.04 0.09 0.05
## Ibarra          0.09 0.05 0.05 0.03
## Quito           0.08 0.04 0.05 0.06
```

## Filtrar en las Tablas

### Por Ejemplo:

1. Cuantas mujeres de Quito repondieron Si

```
f_absolutas["M", "No", "Quito"]
```

```
## [1] 5
```

2. Cuantas personas respondieron No en Ibarra

```
f_absolutas[, "No", "Ibarra"]
```

```
## H M
## 9 5
```

## Tabla HairEyeColor

Datos del Color de Ojos, color de cabello, y el genero de personas

```
HairEyeColor
```

```
## , , Sex = Male
##
##      Eye
## Hair  Brown Blue Hazel Green
## Black   32   11   10    3
## Brown   53   50   25   15
## Red     10   10    7    7
## Blond    3   30    5    8
##
## , , Sex = Female
##
##      Eye
## Hair  Brown Blue Hazel Green
## Black   36    9    5    2
## Brown   66   34   29   14
## Red     16    7    7    7
## Blond    4   64    5    8
```

## Indexacion en la tabla HairEyeColor

Cuantos Hombre y mujeres tienen el color de cabello negro y os ojos Azules

```
HairEyeColor["Black", "Blue", ]
```

```
##   Male Female  
##    11      9
```

Calcular la tabla de frecuencias relativas agrupadas por el genero

```
prop.table(HairEyeColor, margin = 3)
```

```
## , , Sex = Male  
##  
##      Eye  
## Hair      Brown      Blue      Hazel      Green  
## Black 0.114695341 0.039426523 0.035842294 0.010752688  
## Brown 0.189964158 0.179211470 0.089605735 0.053763441  
## Red   0.035842294 0.035842294 0.025089606 0.025089606  
## Blond 0.010752688 0.107526882 0.017921147 0.028673835  
##  
## , , Sex = Female  
##  
##      Eye  
## Hair      Brown      Blue      Hazel      Green  
## Black 0.115015974 0.028753994 0.015974441 0.006389776  
## Brown 0.210862620 0.108626198 0.092651757 0.044728435  
## Red   0.051118211 0.022364217 0.022364217 0.022364217  
## Blond 0.012779553 0.204472843 0.015974441 0.025559105
```

Calcular la tabla de frecuencias relativas conformada por todos los colores de ojos y cabello filtrada por el género

```
colores <- aperm(HairEyeColor, perm = c("Sex", "Eye", "Hair"))  
prop.table(colores, margin = c(2, 3))
```

```
## , , Hair = Black  
##  
##      Eye  
## Sex      Brown      Blue      Hazel      Green  
## Male  0.4705882 0.5500000 0.6666667 0.6000000  
## Female 0.5294118 0.4500000 0.3333333 0.4000000  
##  
## , , Hair = Brown  
##  
##      Eye  
## Sex      Brown      Blue      Hazel      Green  
## Male  0.4453782 0.5952381 0.4629630 0.5172414  
## Female 0.5546218 0.4047619 0.5370370 0.4827586  
##
```

```
## , , Hair = Red
##
##      Eye
## Sex      Brown      Blue      Hazel      Green
## Male  0.3846154 0.5882353 0.5000000 0.5000000
## Female 0.6153846 0.4117647 0.5000000 0.5000000
##
## , , Hair = Blond
##
##      Eye
## Sex      Brown      Blue      Hazel      Green
## Male  0.4285714 0.3191489 0.5000000 0.5000000
## Female 0.5714286 0.6808511 0.5000000 0.5000000
```

## Representacion Grafica

### Diagrama de Barras

```
EnergyDrink <- read.table("../data/EnergyDrink", header = T)
```

Tabla de frecuencias por columnas del Dataframe

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

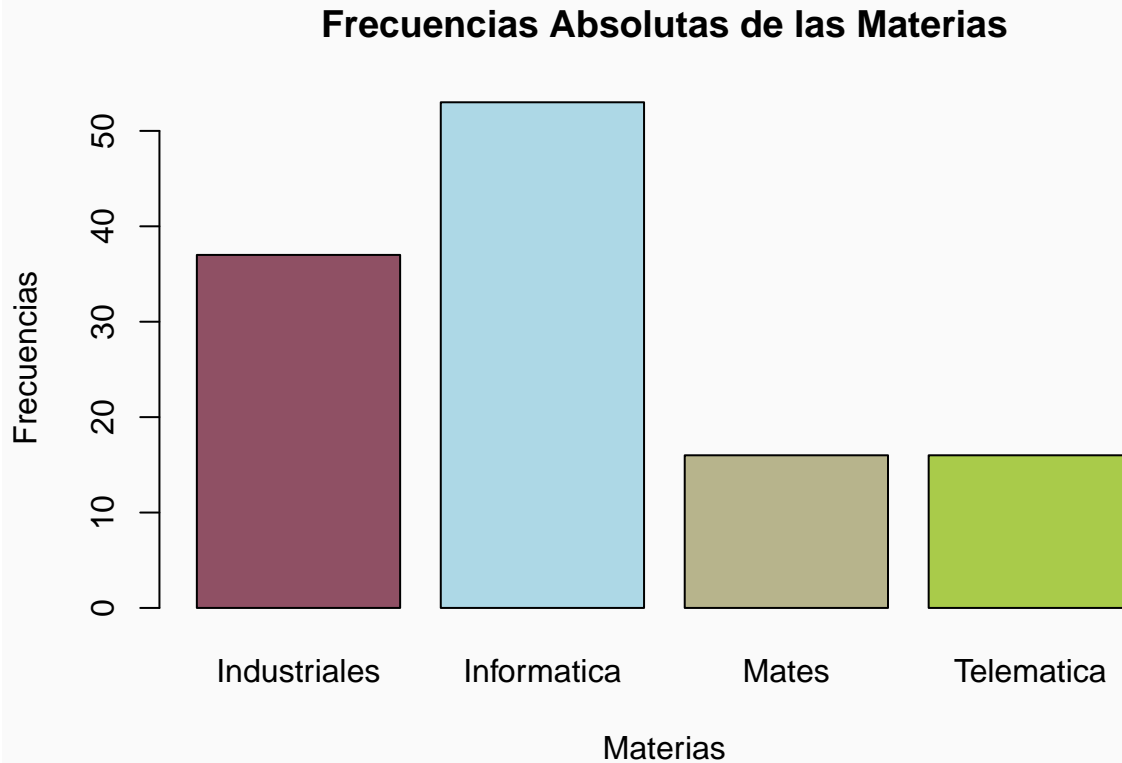
```
EnergyDrink %>% apply(MARGIN = 2, FUN = table)
```

```
## $estudio
##
## Industriales  Informatica      Mates      Telematica
##           37           53           16           16
##
## $bebe
##
## No Si
## 97 25
##
## $sexo
##
## Hombre  Mujer
##      83      39
```

Ahora para graficar las frecuencias de la variable `estudio` que son las materias, en un diagrama de barras, vamos a usar `barplot`

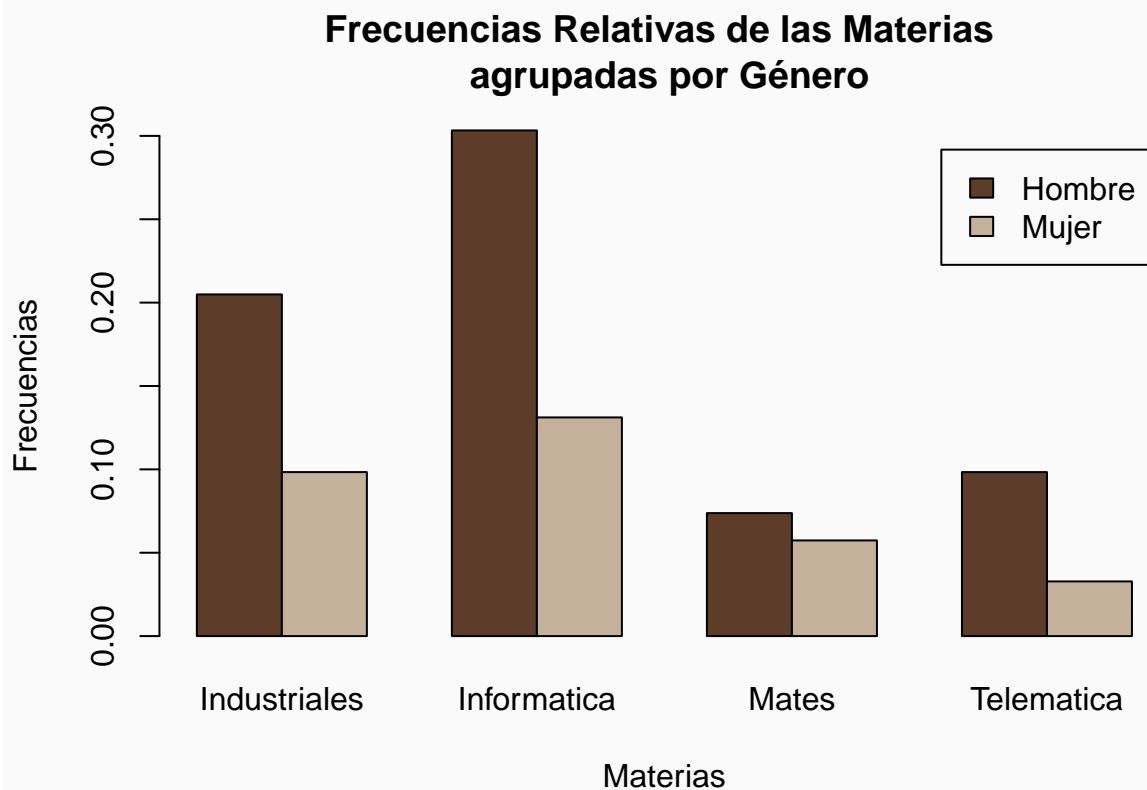
```
par(bg = "#fafafa")
EnergyDrink %>% select(estudio) %>%
  table %>%
```

```
barplot(col=c("#8F5064","lightblue", "#B7B48C", "#A9CB4A"),
        main ="Frecuencias Absolutas de las Materias",
        xlab = "Materias", ylab = "Frecuencias")
```



Ahora vamos a ver la tabla de frecuencias relativas de Las materias por el genero

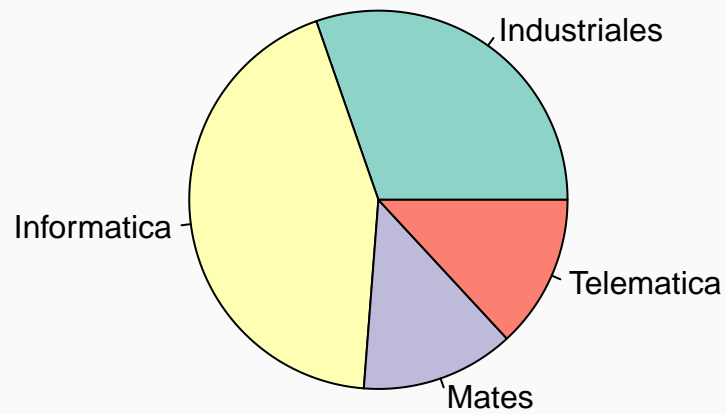
```
par(bg = "#fafafa")
EnergyDrink %>% select(sexo, estudio) %>%
  table %>%
  prop.table %>%
  barplot(col=c("#5E3C2A", "#C3B19B"),
          main ="Frecuencias Relativas de las Materias \n agrupadas por Género",
          xlab = "Materias", ylab = "Frecuencias",
          beside = T, legend.text = c("Hombre", "Mujer"))
```



#### Diagrama Circular(Pie Chart)

```
library(RColorBrewer)
frecuencias_materias <- table(EnergyDrink$estudio)
par(bg = "#fafafa")
pie(frecuencias_materias, main = "Frecuencias Absolutas de las Materias",
    col = brewer.pal(n = 4, name = "Set3"))
```

## Frecuencias Absolutas de las Materias



## Diagrama de Sectores(Mosaic Plot)

```
par(bg = "#fafafa")
EnergyDrink %>% select(estudio, sexo) %>% table %>%
  mosaicplot(main = "Frecuencias Absolutas de Materias por Genero",
             col = brewer.pal(n = 8, name = "Accent"))
```



## Frecuencias Absolutas de Materias por Genero

