

Gráficos

Cielo Darlene Barrios Mixteco

2023-11-07

Se va a trabajar con la matriz penguins

Imprtación de la matriz penguins

Import data / from excel /browse / seleccionar archivo / aceptar

1.- Instalar paqueteria

```
install.packages("readxl")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'  
## (as 'lib' is unspecified)
```

2.-Abrir librería

```
library("readxl")
```

3.- Exportación de la matriz de datos

```
penguins<-read_excel("penguins.xlsx")
```

Exploración

1.- Dimensión de la matriz.

```
dim(penguins)
```

```
## [1] 344 9
```

2.- Tipo de variables

```
str(penguins)
```

```
## tibble [344 x 9] (S3: tbl_df/tbl/data.frame)  
## $ ID : chr [1:344] "i1" "i2" "i3" "i4" ...  
## $ especie : chr [1:344] "Adelie" "Adelie" "Adelie" "Adelie" ...  
## $ isla : chr [1:344] "Torgersen" "Torgersen" "Torgersen" "Torgersen" ...  
## $ largo_pico_mm : num [1:344] 39.1 39.5 40.3 37.8 36.7 39.3 38.9 39.2 34.1 42 ...  
## $ grosor_pico_mm : num [1:344] 18.7 17.4 18 18.1 19.3 20.6 17.8 19.6 18.1 20.2 ...  
## $ largo_aleta_mm : num [1:344] 181 186 195 190 193 190 181 195 193 190 ...  
## $ masa_corporal_g: num [1:344] 3750 3800 3250 3700 3450 ...  
## $ genero : chr [1:344] "male" "female" "female" "female" ...  
## $ año : num [1:344] 2007 2007 2007 2007 2007 ...
```

3.-En busca de datos perdidos

```
anyNA(penguins)
```

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

Configuración de la matriz

Visualización de la columna

```
penguins$especie
```

```
## [1] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [7] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [13] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [19] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [25] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [31] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [37] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [43] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [49] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [55] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [61] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [67] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [73] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [79] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [85] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [91] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [97] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [103] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [109] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [115] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [121] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [127] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [133] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [139] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [145] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie"
## [151] "Adelie" "Adelie" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [157] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [163] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [169] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [175] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [181] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [187] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [193] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [199] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [205] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [211] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [217] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [223] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [229] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [235] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [241] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [247] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
## [253] "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo" "Gentoo"
```



```
## [249] "male" "female" "female" "male" "female" "male" "female" "male"
## [257] "female" "male" "female" "male" "female" "male" "female" "male"
## [265] "female" "male" "female" "male" "female" "male" "female" "male"
## [273] "female" "male" "female" "male" "female" "male" "male" "female"
## [281] "male" "female" "female" "male" "female" "male" "female" "male"
## [289] "female" "male" "female" "male" "male" "female" "female" "male"
## [297] "female" "male" "female" "male" "female" "male" "female" "male"
## [305] "female" "male" "female" "male" "female" "male" "male" "female"
## [313] "female" "male" "female" "male" "male" "female" "male" "female"
## [321] "female" "male" "female" "male" "male" "female" "female" "male"
## [329] "female" "male" "female" "male" "female" "male" "male" "female"
## [337] "male" "female" "female" "male" "female" "male" "male" "female"
```

```
penguins$año
```

```
## [1] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [16] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [31] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [46] 2007 2007 2007 2007 2007 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [61] 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [76] 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [91] 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2009 2009 2009 2009 2009
## [106] 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
## [121] 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
## [136] 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
## [151] 2009 2009 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [166] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [181] 2007 2007 2007 2007 2007 2007 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [196] 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [211] 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [226] 2008 2008 2008 2008 2008 2008 2008 2009 2009 2009 2009 2009 2009 2009 2009
## [241] 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
## [256] 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
## [271] 2009 2009 2009 2009 2009 2009 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [286] 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007 2007
## [301] 2007 2007 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008 2008
## [316] 2008 2008 2008 2008 2008 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
## [331] 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009
```

1.- Convertir las variables categóricas a factores

```
penguins$especie<-factor(penguins$especie,
                        levels=c("Adelie", "Gentoo", "Chinstrap"))
```

```
penguins$isla<-factor(penguins$isla,
                     levels=c("Torgersen", "Biscoe", "Dream"))
```

```
penguins$genero<-factor(penguins$genero,
                       levels=c("male", "female"))
```

```
penguins$año<-factor(penguins$año,
                    levels=c("2007", "2008", "2009"))
```

1.1 Tipo de variables

```
str(penguins)
```

```
## tibble [344 x 9] (S3: tbl_df/tbl/data.frame)
## $ ID          : chr [1:344] "i1" "i2" "i3" "i4" ...
## $ especie     : Factor w/ 3 levels "Adelie","Gentoo",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ isla        : Factor w/ 3 levels "Torgersen","Biscoe",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ largo_pico_mm : num [1:344] 39.1 39.5 40.3 37.8 36.7 39.3 38.9 39.2 34.1 42 ...
## $ grosor_pico_mm : num [1:344] 18.7 17.4 18 18.1 19.3 20.6 17.8 19.6 18.1 20.2 ...
## $ largo_aleta_mm : num [1:344] 181 186 195 190 193 190 181 195 193 190 ...
## $ masa_corporal_g: num [1:344] 3750 3800 3250 3700 3450 ...
## $ genero      : Factor w/ 2 levels "male","female": 1 2 2 2 2 1 2 1 2 1 ...
## $ año         : Factor w/ 3 levels "2007","2008",...: 1 1 1 1 1 1 1 1 1 1 ...
```

1.2 Media y mediana

```
summary(penguins)
```

```
##      ID          especie      isla      largo_pico_mm
## Length:344      Adelie    :152  Torgersen: 52      Min.    :32.10
## Class :character  Gentoo    :124  Biscoe   :168  1st Qu.:39.20
## Mode  :character  Chinstrap: 68  Dream    :124  Median  :44.45
##                                     Mean    :43.92
##                                     3rd Qu.:48.50
##                                     Max.    :59.60
## grosor_pico_mm  largo_aleta_mm  masa_corporal_g  genero      año
## Min.    :13.10  Min.    :172.0  Min.    :2700    male :170    2007:110
## 1st Qu.:15.60  1st Qu.:190.0  1st Qu.:3550    female:174  2008:114
## Median :17.30  Median :197.0  Median :4050                                2009:120
## Mean    :17.15  Mean    :200.9  Mean    :4202
## 3rd Qu.:18.70  3rd Qu.:213.2  3rd Qu.:4756
## Max.    :21.50  Max.    :231.0  Max.    :6300
```

1.3 Visualización

```
penguins
```

```
## # A tibble: 344 x 9
##   ID especie isla      largo_pico_mm grosor_pico_mm largo_aleta_mm
##   <chr> <fct> <fct>          <dbl>          <dbl>          <dbl>
## 1 i1 Adelie Torgersen      39.1            18.7           181
## 2 i2 Adelie Torgersen      39.5            17.4           186
## 3 i3 Adelie Torgersen      40.3            18            195
## 4 i4 Adelie Torgersen      37.8            18.1           190
## 5 i5 Adelie Torgersen      36.7            19.3           193
## 6 i6 Adelie Torgersen      39.3            20.6           190
## 7 i7 Adelie Torgersen      38.9            17.8           181
## 8 i8 Adelie Torgersen      39.2            19.6           195
## 9 i9 Adelie Torgersen      34.1            18.1           193
## 10 i10 Adelie Torgersen      42             20.2           190
## # i 334 more rows
## # i 3 more variables: masa_corporal_g <dbl>, genero <fct>, año <fct>
```

2.- Creamos una nueva matriz de datos donde se seleccionan

Las columnas de la 2 a la 9

```
BD1<-penguins[,2:9]
```

```
colnames(BD1)
```

```
## [1] "especie"      "isla"          "largo_pico_mm"  "grosor_pico_mm"  
## [5] "largo_aleta_mm" "masa_corporal_g" "genero"         "año"
```

Librerías

1.- Instalar paquetería

```
install.packages("ggplot2")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'  
## (as 'lib' is unspecified)
```

2.- Abrir librería

```
library(ggplot2)
```

Boxplot

1.- Creación de un vector de color

```
color=c("lightpink2","lightblue1")
```

2.- Creacion del grafico

```
BX<-ggplot(BD1, aes(x=genero, y=largo_pico_mm))+  
  geom_boxplot(fill=color)+  
  ggtitle("Boxplot")+  
  xlab("Género")+  
  ylab("largo de la aleta (mm)")+  
  theme_bw()
```

3.- Visualización del boxplot

```
BX
```

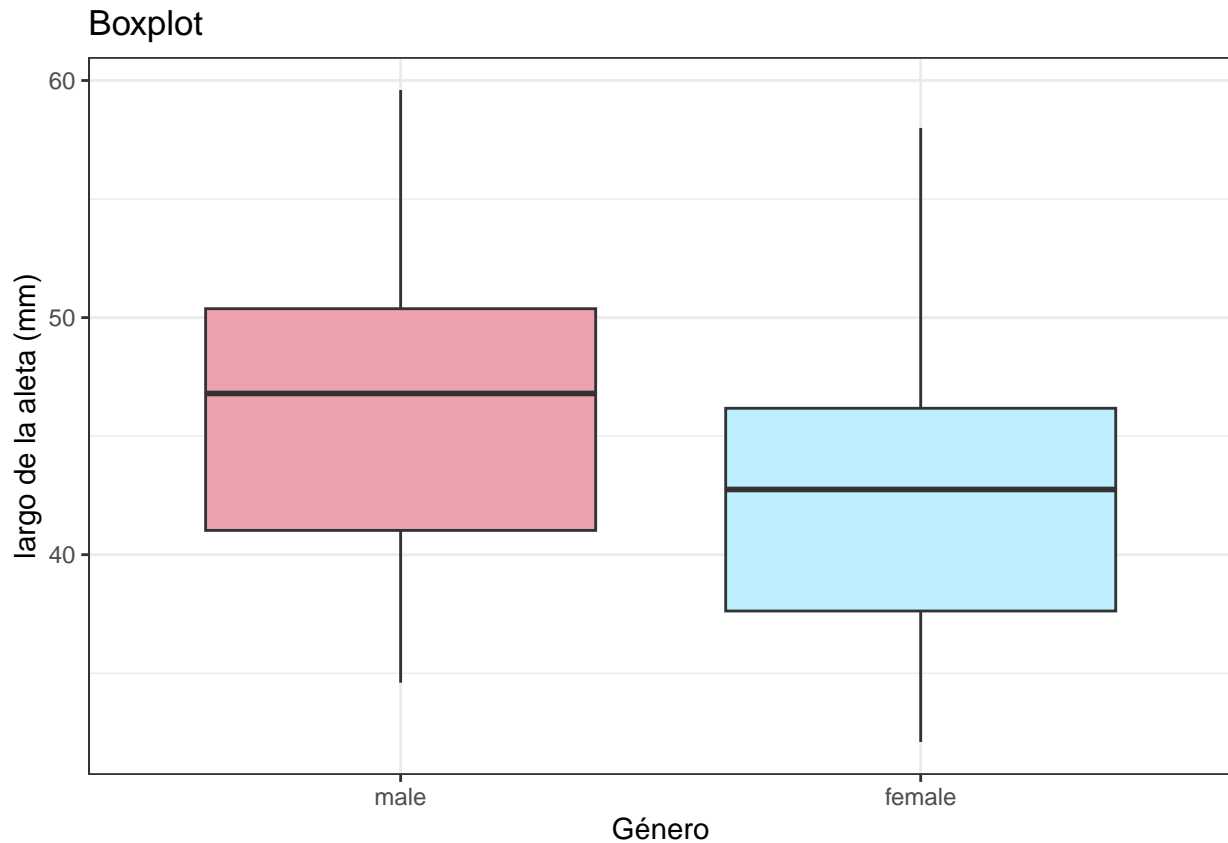


Gráfico de barras

1.- Creación de un vector de color

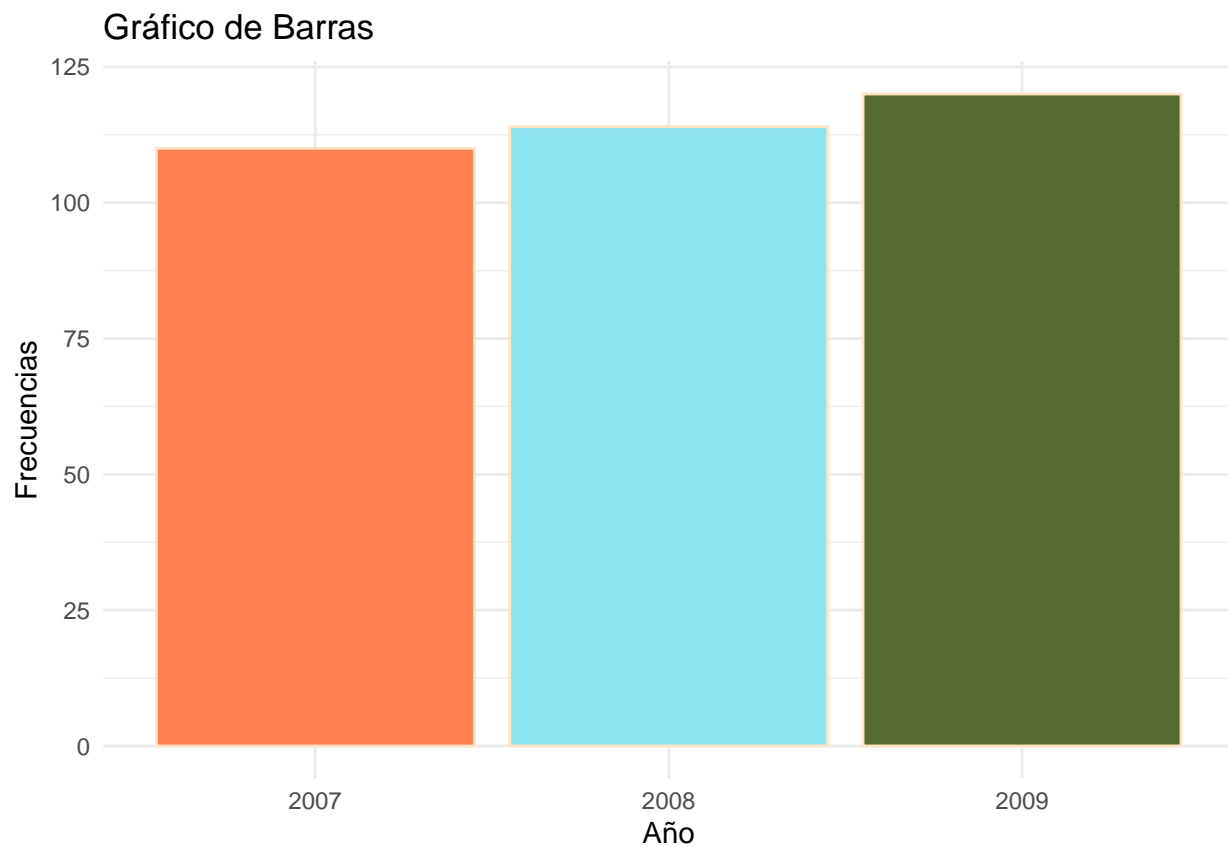
```
color=c("coral", "cadetblue2", "darkolivegreen")
```

2.- Creación del gráfico

```
GB1<-ggplot(BD1, aes(x=año))+
  geom_bar(colour= "bisque", fill=color)+
  ggtitle("Gráfico de Barras")+
  xlab("Año")+
  ylab("Frecuencias")+
  theme_minimal()
```

3.- Visualizacion del grafico

```
GB1
```

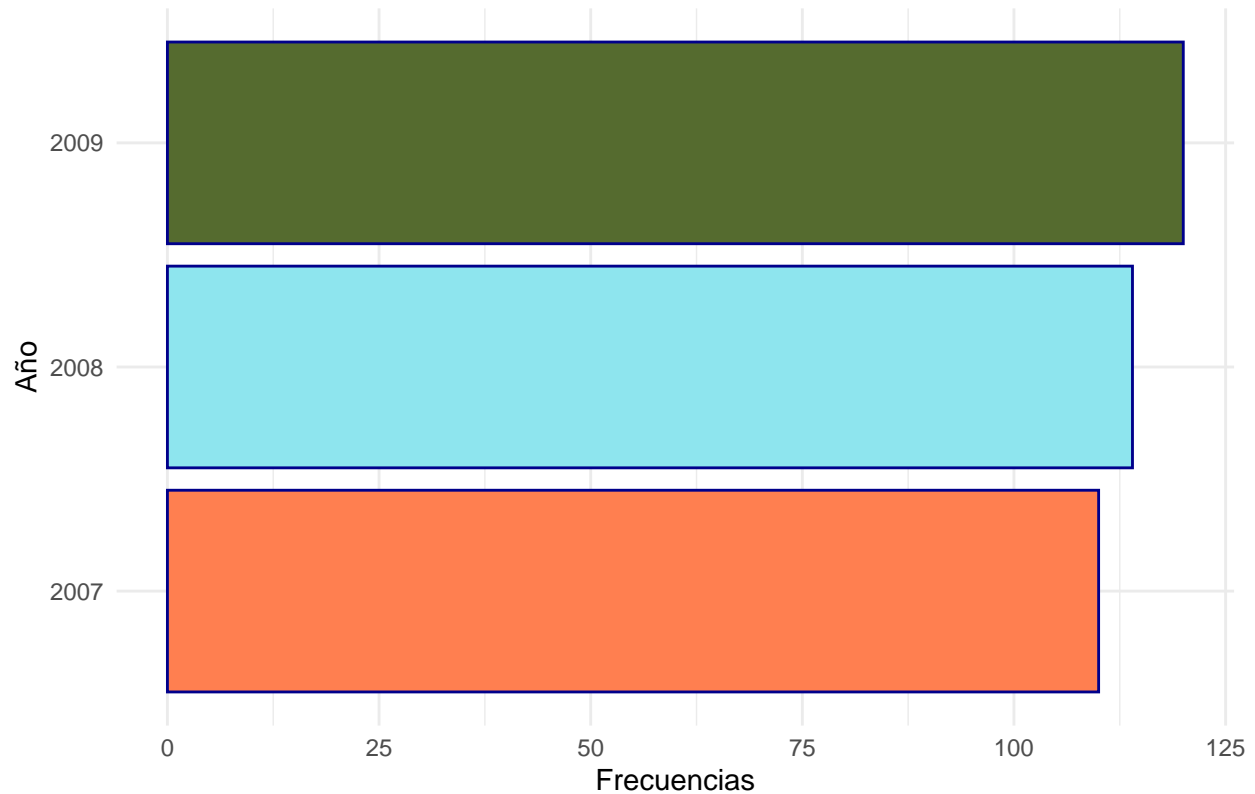
4.- Barras verticales

```
GB2<-ggplot(BD1, aes(x=año))+  
  geom_bar(colour= "darkblue", fill=color)+  
  ggtitle("Gráfico de Barras")+  
  xlab("Año")+  
  ylab("Frecuencias")+  
  coord_flip()+  
  theme_minimal()
```

5. Visualizacion del objeto

GB2

Gráfico de Barras



Histograma

1.- Construcción del gráfico

```
HG<-ggplot(BD1, aes(x=largo_aleta_mm))+  
  geom_histogram(col="chocolate1", fill="darkseagreen3")+  
  ggtitle("Histograma")+  
  xlab("Largo de la aleta (mm)")+  
  ylab("Frecuencias")+  
  theme_classic()
```

2.- Visualización del gráfico

HG

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

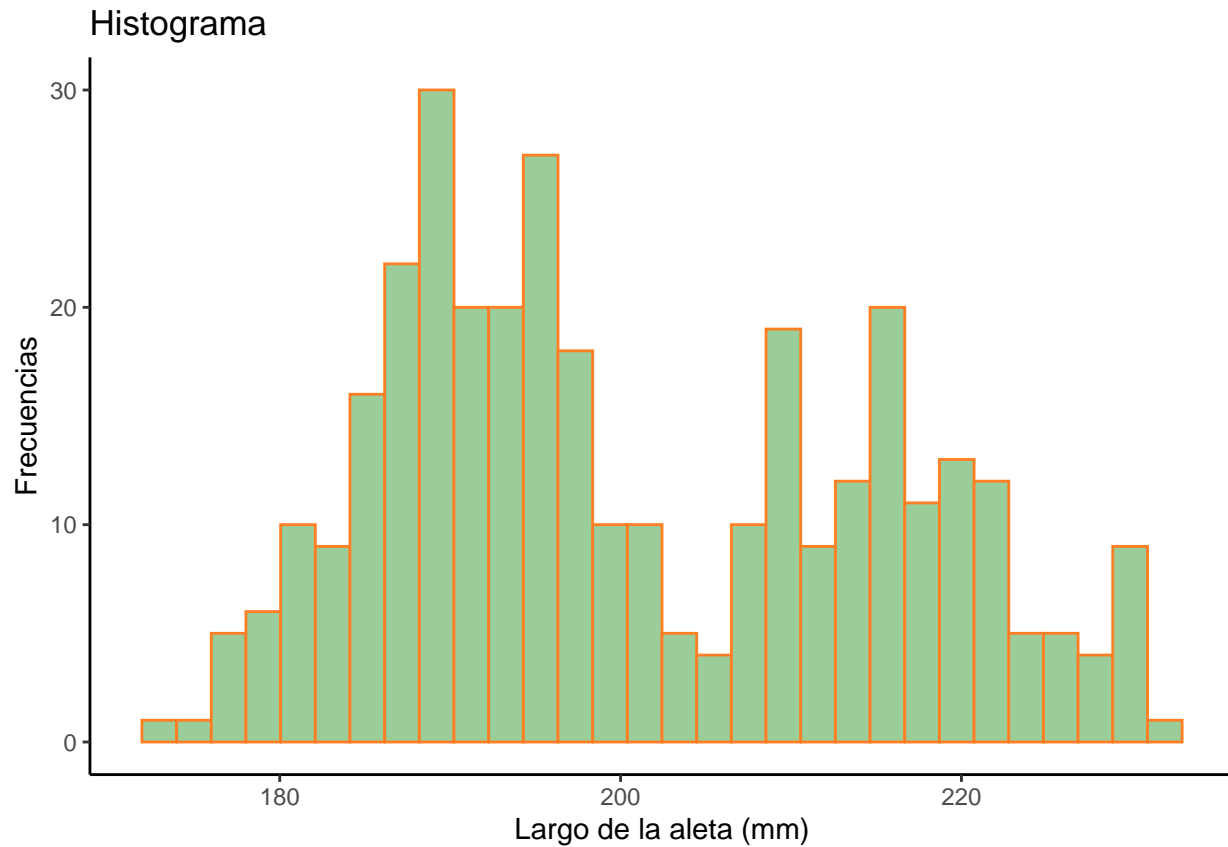


Gráfico de dispersión

1.- Construcción del gráfico

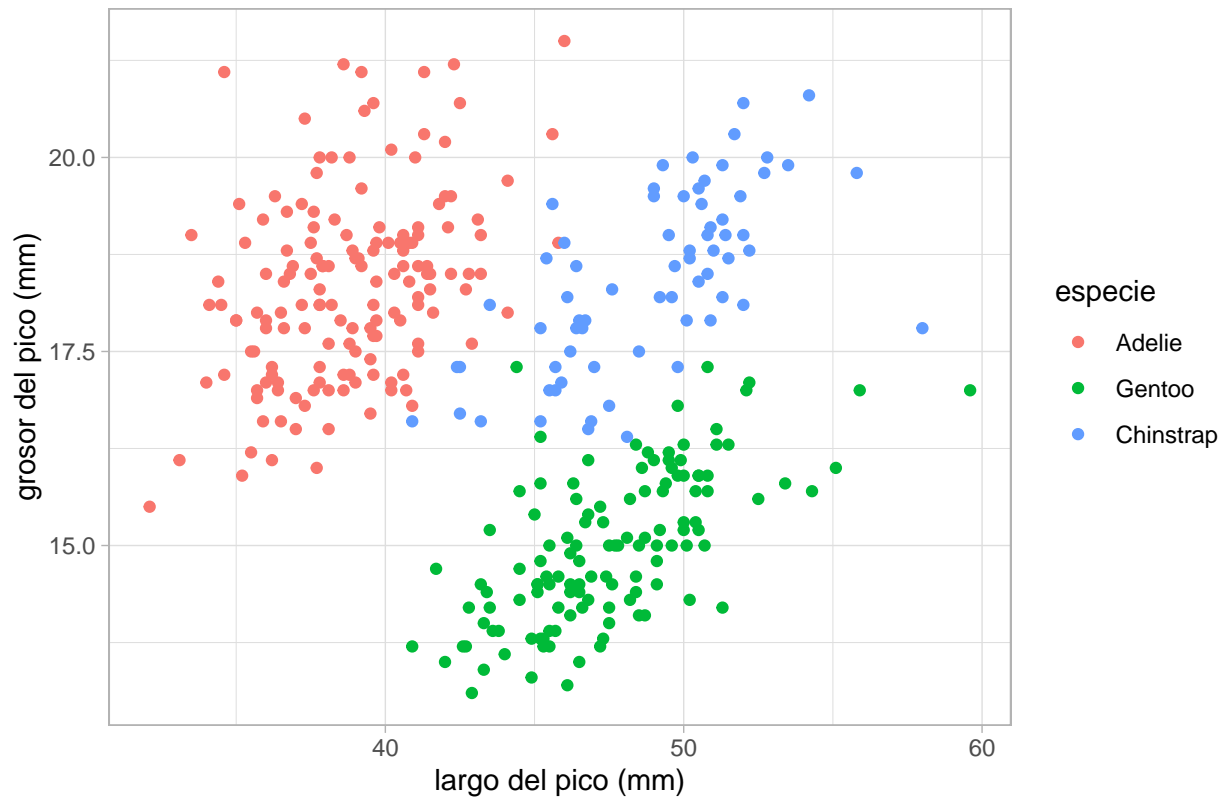
```
GD<-ggplot(BD1, aes(x=largo_pico_mm, y=grosor_pico_mm))+
  geom_point(aes(color=especie))+
  ggtitle("Gráfico de dispersión")+
  xlab("largo del pico (mm)") +
  ylab("grosor del pico (mm)") +
  theme_light()
```

Nota: es bd1 porque es nuestra matriz modificada, es decir, eliminamos la primera columna de penguins

2.- Visualización del objeto

```
GD
```

Gráfico de dispersión

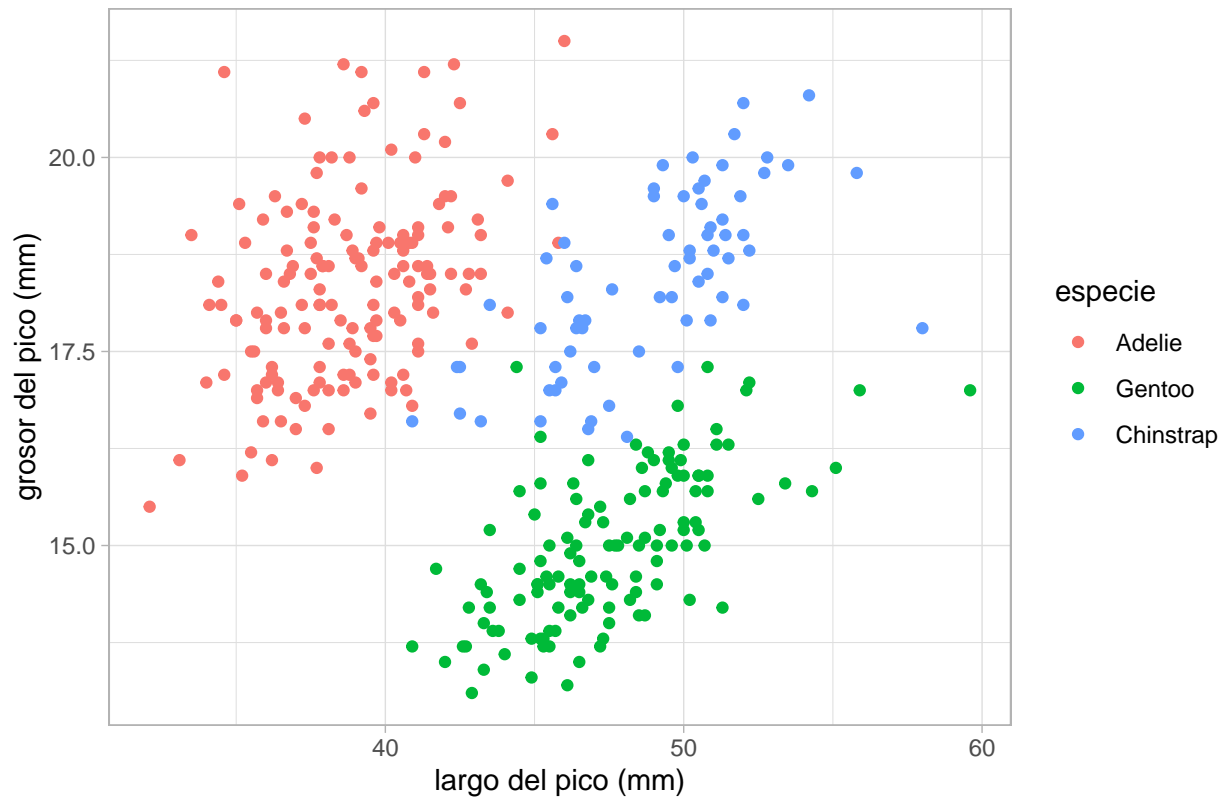


Otros 2 ejemplos

```
GD2<-ggplot(BD1, aes(x=largo_pico_mm, y=grosor_pico_mm))+  
  geom_point(aes(color=especie))+  
  ggtitle("Gráfico de dispersión")+  
  xlab("largo del pico (mm)")+  
  ylab("grosor del pico (mm)")+  
  theme_light()
```

GD2

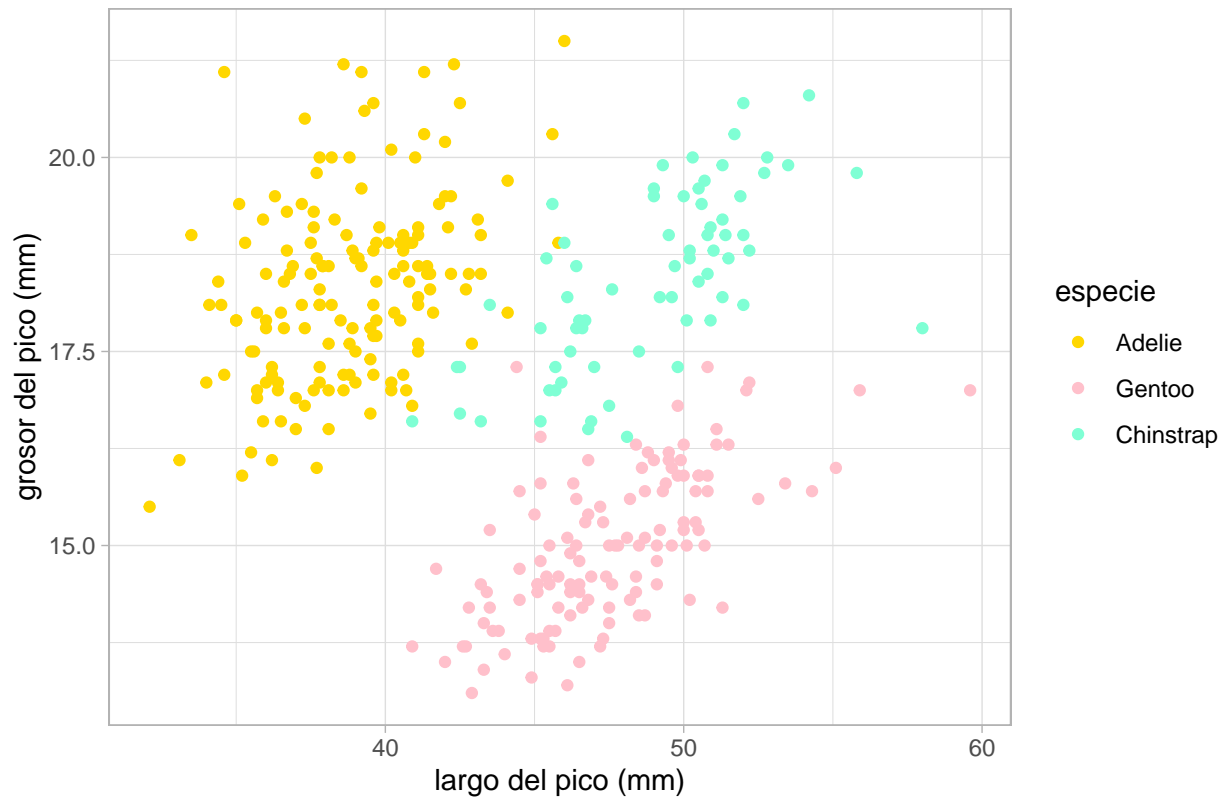
Gráfico de dispersión



```
GD3<-ggplot(BD1, aes(x=largo_pico_mm, y=grosor_pico_mm))+  
  geom_point(aes(color=especie))+  
  scale_color_manual(values=c("gold","pink","aquamarine"))+  
  ggtitle("Gráfico de dispersión")+  
  xlab("largo del pico (mm)")+  
  ylab("grosor del pico (mm)")+  
  theme_light()
```

GD3

Gráfico de dispersión



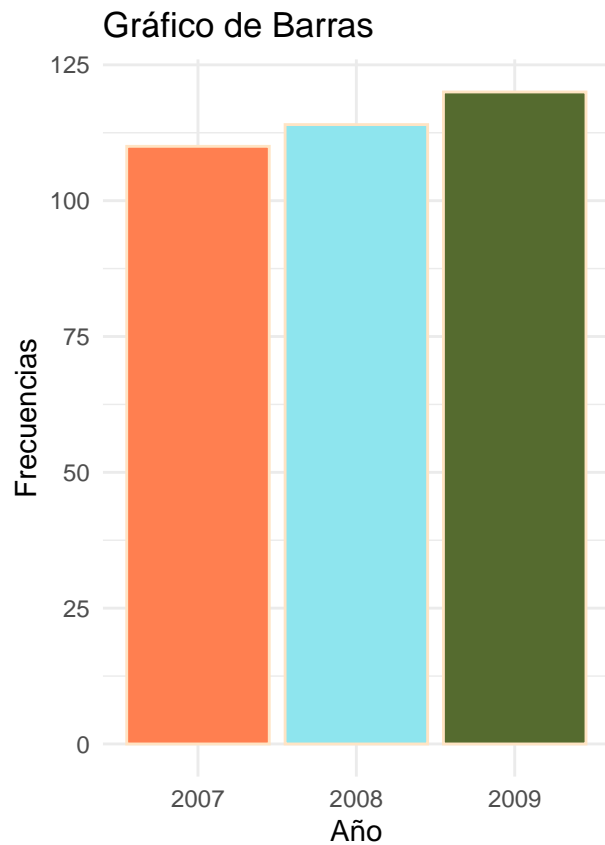
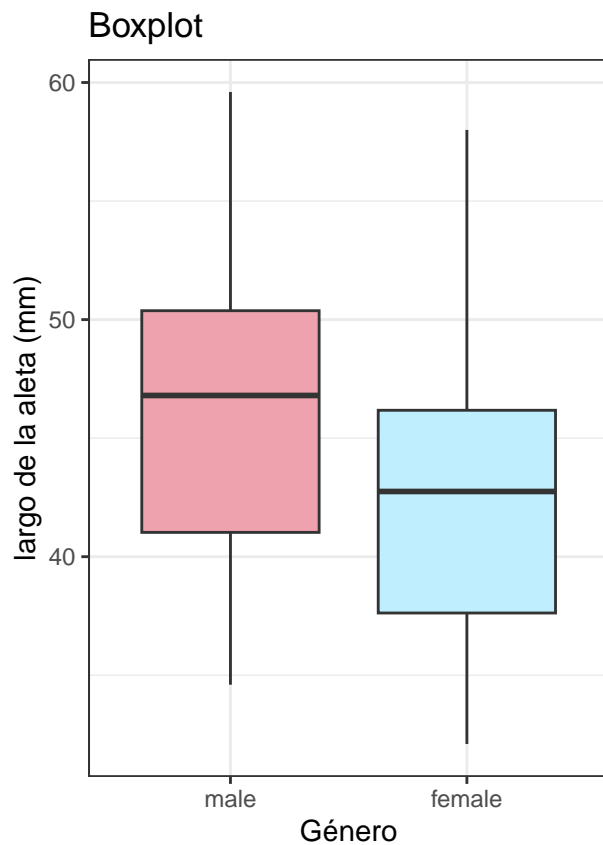
Organización de gráficos

1.- Abrir librería de paquete gridExtra

```
library(gridExtra)
```

2.- Organización 2 gráficos en una fila y dos columnas

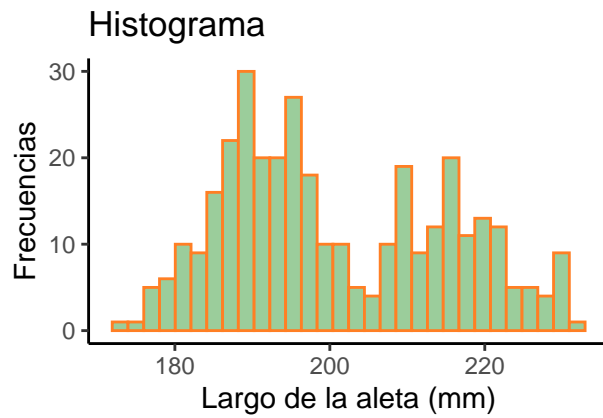
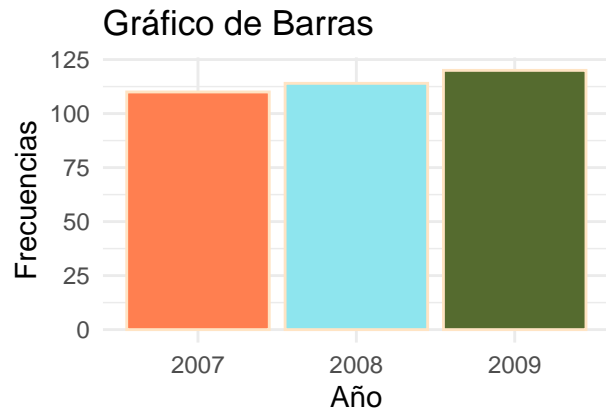
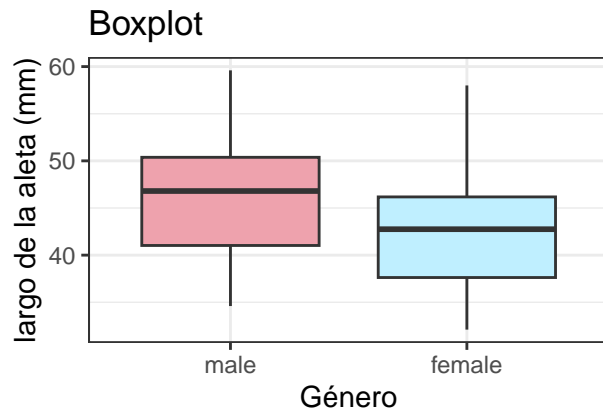
```
grid.arrange(BX,GB1, nrow=1, ncol=2)
```



3.-Organizacion 3 graficos en dos filas y dos columnas

```
grid.arrange(BX,GB1,HG, nrow=2, ncol=2)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



4.-Organizacion 4 graficos en dos filas y dos columnas

```
grid.arrange(BX,GB1,HG,GD3, nrow=2, ncol=2)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```