

# Codificación de Variables

Linda Estefany Bravo López

22/11/2021

\*\*\*\*\* RECODIFICACION DE VARIABLES \*\*\*\*\*

\*\*\*\*\*

## Importación de matriz

\*\*\*\*\*

### 1.-Lectura de la matriz

```
BD <-read.csv ("BD1.csv")
```

### 2.-Seleccionar las columnas 2 a la 9

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

### 3.- Se obtiene una copia de la matriz

```
BD2 <-BD1
```

\*\*\*\*\*

## Renombrar variables

\*\*\*\*\*

### 1.- Visualización del nombre de las variables

```
colnames (BD1)
```

```
## [1] "species"          "island"            "bill_length_mm"
## [4] "bill_depth_mm"    "flipper_length_mm" "body_mass_g"
## [7] "sex"              "year"
```

### 2.- Renombrar las variables en orden de las columnas.

```
names (BD2) = c ("Especies", "Isla", "Largo_pico_mm", "Ancho_pico_mm", "Largo_aleta_mm",
                 "Masa_corporal_g", "Sexo", "Año")
```

### 3.- Se vuelve a visualizar los nombres de las variables

```
colnames (BD2)
```

```
## [1] "Especies"          "Isla"              "Largo_pico_mm"    "Ancho_pico_mm"
## [5] "Largo_aleta_mm"    "Masa_corporal_g"   "Sexo"              "Año"
```

\*\*\*\*\*

## Tratamiento de las variables cualitativas

\*\*\*\*\*

### 1.-Factor

#a) Verificar el orden de la variable

```
str (BD2 $ Especies)
```

```
## chr [1:344] "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" "Adelie" ...
```



```
## [307] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [313] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [319] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [325] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [331] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [337] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [343] "Chinstrap" "Chinstrap"
```

1) Especie variable: originalmente está como caracter, pero se requiere convertir un factor de 3 niveles.

```
BD2$Especies<-factor (BD2 $ Especies,
                      levels = c ("Adelie", "Gentoo", "Chinstrap"))
```

Se verifica que la variable sea un factor.

```
str (BD2 $ Especies)
```

```
## Factor w/ 3 levels "Adelie","Gentoo",...: 1 1 1 1 1 1 1 1 1 1 ...
```

2) Isla variable

```
str (BD2 $ Isla)
```

```
## chr [1:344] "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen" ...
```

```
BD2 $ Isla
```

```
## [1] "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen"
## [7] "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen"
## [13] "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen"
## [19] "Torgersen" "Torgersen" "Biscoe" "Biscoe" "Biscoe" "Biscoe"
## [25] "Biscoe" "Biscoe" "Biscoe" "Biscoe" "Biscoe" "Biscoe"
## [31] "Dream" "Dream" "Dream" "Dream" "Dream" "Dream"
## [37] "Dream" "Dream" "Dream" "Dream" "Dream" "Dream"
## [43] "Dream" "Dream" "Dream" "Dream" "Dream" "Dream"
## [49] "Dream" "Dream" "Biscoe" "Biscoe" "Biscoe" "Biscoe"
## [55] "Biscoe" "Biscoe" "Biscoe" "Biscoe" "Biscoe" "Biscoe"
## [61] "Biscoe" "Biscoe" "Biscoe" "Biscoe" "Biscoe" "Biscoe"
## [67] "Biscoe" "Biscoe" "Torgersen" "Torgersen" "Torgersen" "Torgersen"
## [73] "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen"
## [79] "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen" "Torgersen"
## [85] "Dream" "Dream" "Dream" "Dream" "Dream" "Dream"
## [91] "Dream" "Dream" "Dream" "Dream" "Dream" "Dream"
```

```
## [97] "Dream"      "Dream"      "Dream"      "Dream"      "Biscoe"     "Biscoe"
## [103] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [109] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [115] "Biscoe"     "Biscoe"     "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"
## [121] "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"
## [127] "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"  "Torgersen"
## [133] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [139] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [145] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [151] "Dream"      "Dream"      "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [157] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [163] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [169] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [175] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [181] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [187] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [193] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [199] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [205] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [211] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [217] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [223] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [229] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [235] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [241] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [247] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [253] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [259] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [265] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [271] "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"     "Biscoe"
## [277] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [283] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [289] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [295] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [301] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [307] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [313] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [319] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [325] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [331] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [337] "Dream"      "Dream"      "Dream"      "Dream"      "Dream"      "Dream"
## [343] "Dream"      "Dream"
```

```
BD2 $ Isla <-factor (BD2 $ Isla,
                    levels = c ("Torgersen", "Sueño", "Biscoe"))
str (BD2 $ Isla)
```

```
## Factor w/ 3 levels "Torgersen","Sueño",...: 1 1 1 1 1 1 1 1 1 1 ...
```

### 3) Sexo variable

```
BD2$Sexo
```

```
## [1] "male" "female" "female" "female" "female" "male" "female" "male"
## [9] "female" "male" "female" "female" "female" "male" "male" "female"
## [17] "female" "male" "female" "male" "female" "male" "female" "male"
## [25] "male" "female" "male" "female" "female" "male" "female" "male"
## [33] "female" "male" "female" "male" "male" "female" "female" "male"
## [41] "female" "male" "female" "male" "female" "male" "male" "female"
## [49] "female" "male" "female" "male" "female" "male" "female" "male"
## [57] "female" "male" "female" "male" "female" "male" "female" "male"
## [65] "female" "male" "female" "male" "female" "male" "female" "male"
## [73] "female" "male" "female" "male" "female" "male" "female" "male"
## [81] "female" "male" "female" "male" "female" "male" "male" "female"
## [89] "male" "female" "female" "male" "female" "male" "female" "male"
## [97] "female" "male" "female" "male" "female" "male" "female" "male"
## [105] "female" "male" "female" "male" "female" "male" "female" "male"
## [113] "female" "male" "female" "male" "female" "male" "female" "male"
## [121] "female" "male" "female" "male" "female" "male" "female" "male"
## [129] "female" "male" "female" "male" "female" "male" "female" "male"
## [137] "female" "male" "female" "male" "female" "male" "female" "male"
## [145] "female" "male" "male" "female" "female" "male" "female" "male"
## [153] "female" "male" "female" "male" "male" "female" "female" "male"
## [161] "female" "male" "female" "male" "female" "male" "female" "male"
## [169] "female" "male" "female" "male" "male" "female" "female" "male"
## [177] "female" "male" "female" "male" "female" "male" "male" "female"
## [185] "female" "male" "female" "male" "female" "male" "female" "male"
## [193] "female" "male" "female" "male" "male" "female" "female" "male"
## [201] "female" "male" "female" "male" "female" "male" "female" "male"
## [209] "female" "male" "female" "male" "female" "male" "female" "male"
## [217] "female" "male" "female" "male" "female" "male" "female" "male"
## [225] "male" "female" "female" "male" "female" "male" "female" "male"
## [233] "female" "male" "female" "male" "female" "male" "female" "male"
## [241] "female" "male" "female" "male" "female" "male" "female" "male"
## [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"
```

```
BD2$Sexo<-factor(BD2$Sexo,
                  levels = c ("male", "female"),
                  labels = c ("Masculino", "Femenino"))
str (BD2$Sexo)
```

```
## Factor w/ 2 levels "Masculino","Femenino": 1 2 2 2 2 1 2 1 2 1 ...
```

#### 4) Año variable

```
BD2 $ 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
```

```
BD2 $ Año <-factor (BD2 $ Año, levels = c ("2007", "2008", "2009"))
str (BD2 $ Año)
```

```
## Factor w/ 3 levels "2007","2008",...: 1 1 1 1 1 1 1 1 1 1 ...
```

#### Para obtener los estadísticos descriptivos de las variables

```
summary(BD2)
```

```
##      Especies      Isla  Largo_pico_mm  Ancho_pico_mm
## Adelie   :152  Torgersen: 52    Min.   :32.10    Min.   :13.10
## Gentoo   :124   Sueño    : 0    1st Qu.:39.20    1st Qu.:15.60
## Chinstrap: 68   Biscoe   :168   Median :44.45    Median :17.30
##              NA's      :124   Mean    :43.92    Mean    :17.15
##              3rd Qu.:48.50    3rd Qu.:18.70
##              Max.    :59.60    Max.    :21.50
## Largo_aleta_mm Masa_corporal_g  Sexo      Año
```

```
## Min.      :172.0   Min.      :2700   Masculino:170   2007:110
## 1st Qu.:190.0   1st Qu.:3550   Femenino :174   2008:114
## Median :197.0   Median :4050           2009:120
## Mean    :200.9   Mean     :4202
## 3rd Qu.:213.2   3rd Qu.:4756
## Max.    :231.0   Max.     :6300
```

## Guardar matriz de datos final

```
BD3 <-as.data.frame(BD2)
```

## Activar la librería para realizar los gráficos

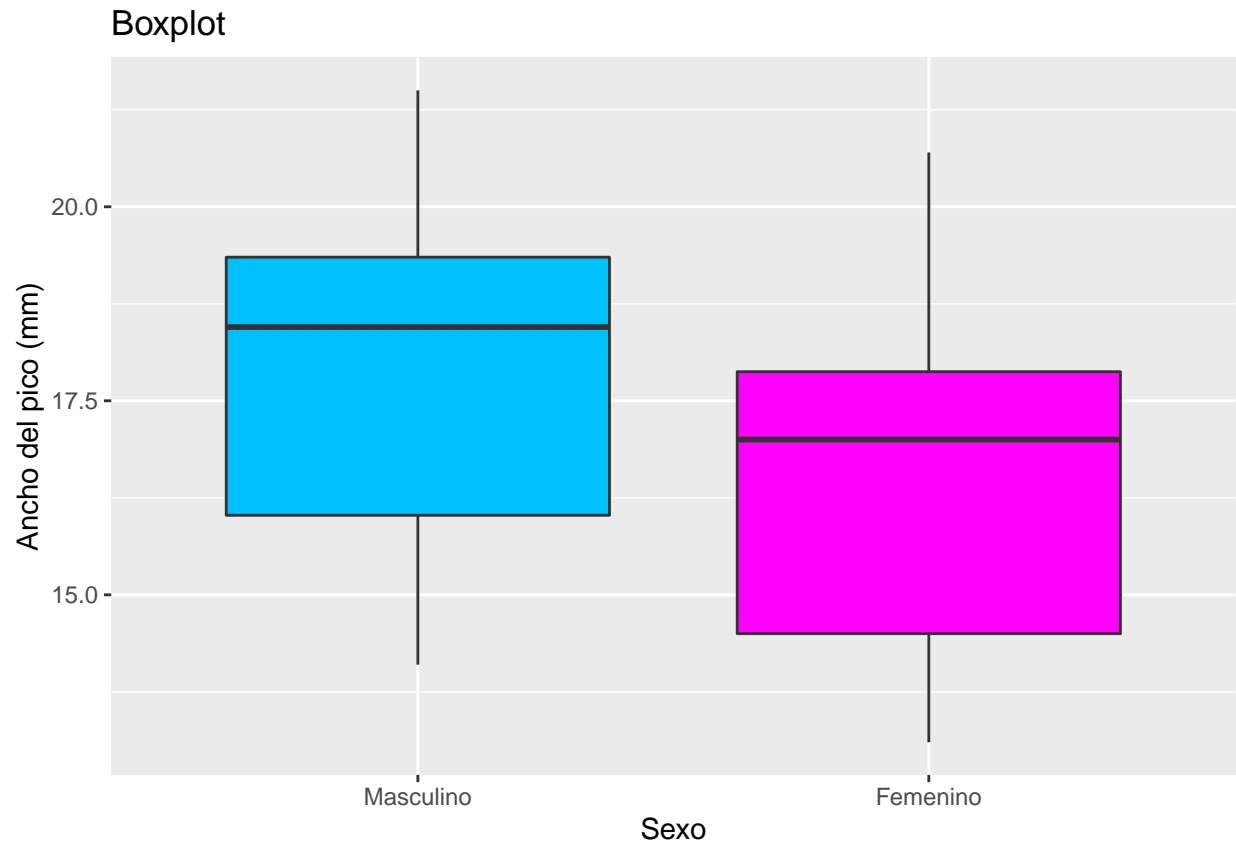
```
library (ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

## Gráfico boxplot de exploración

```
color = c ("deepskyblue1", "magenta")
BX <-ggplot (BD3, aes (x = Sexo, y = Ancho_pico_mm)) +
  geom_boxplot (fill = color) +
  ggtitle ("Boxplot") +
  xlab ("Sexo") +
  ylab ("Ancho del pico (mm)")
BX
```

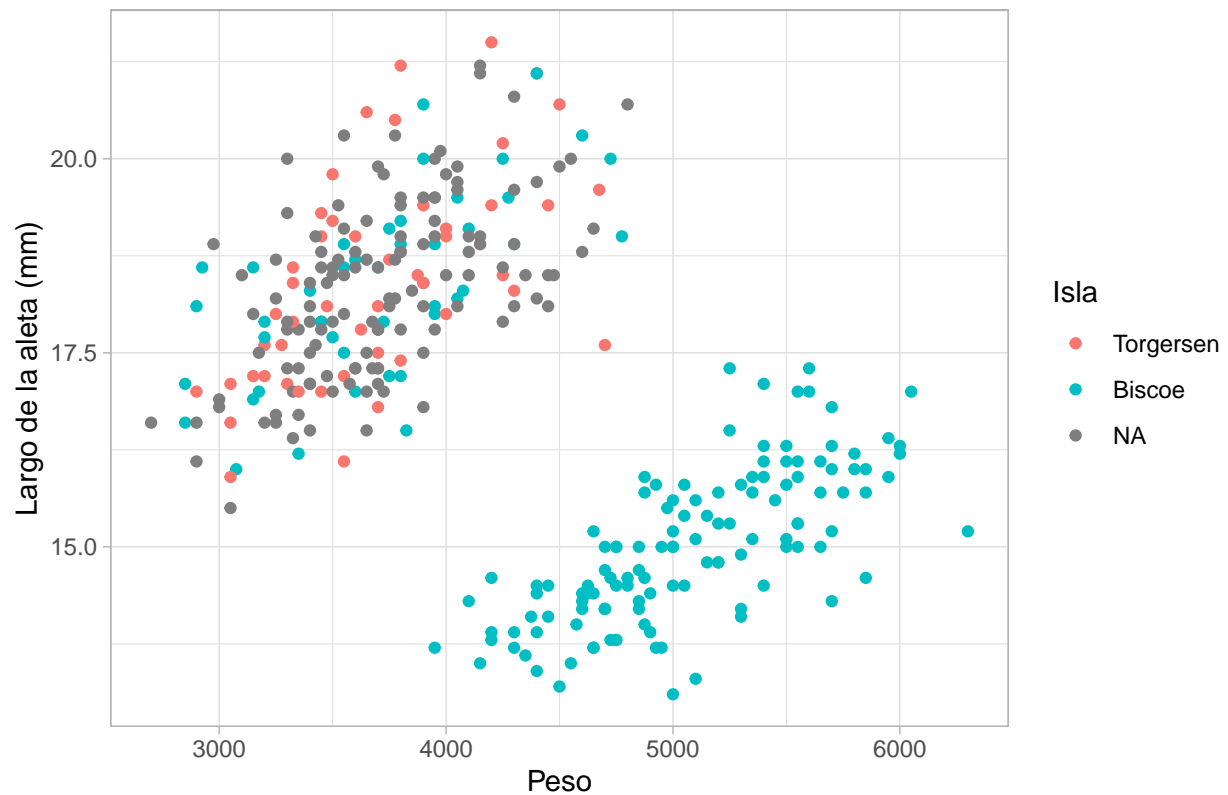




## Gráfico de dispersión

```
DD <-ggplot (BD3, aes (Masa_corporal_g, Ancho_pico_mm)) +  
  geom_point (aes (color = Isla)) +  
  labs (title = "Gráfico de dispersión") +  
  xlab ("Peso") +  
  ylab ("Largo de la aleta (mm)") +  
  theme_light ()  
DD
```

Gráfico de dispersión



**NOTA:** ggplot reconoce que masa corporal es el eje “ $x$ ” y el ancho del pico es el eje “ $y$ ”

Guardar la matriz de datos

```
write.csv ( BD3 , " BD3_penguins.csv " )
```