

MRLS

Lucero Hernandez Mendez

2023-11-16

Modelo de regresion lineal simple

1.- Instalación de paqueteria

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

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

2.- Abrir libreria

```
library(readxl)
```

3.- Exportación de la matriz

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

4.- Exploración de la matriz

4.1 Dimensión de la matriz Con el codigo: dim(BD)

```
dim(penguins)
```

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

4.2 Nombre de las columnas Con el codigo str(BD)

```
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 ...
```

4.3 Tipo de variables Con el codigo: colnames(BD)

```
colnames(penguins)
```

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

4.4 En busca de datos perdidos Con el codigo: anyNA(BD)

```
anyNA(penguins)
```

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

Configuración de la matriz

1.- Visualización de las columnas: especie, isla, género y año.

1.1- 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"))
```

1.2 Visualización del tipo de variables Con el código str(BD)

```
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 ...
```

2.- Selección de variables 2.1 Seleccionar una variable de la matriz original

```
penguins$especie
```

```
##   [1] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##   [8] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [15] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [22] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [29] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [36] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [43] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [50] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [57] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [64] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [71] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [78] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [85] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [92] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
##  [99] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
## [106] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie  
## [113] Adelie    Adelie    Adelie    Adelie    Adelie    Adelie    Adelie
```

```
## [120] Adelie      Adelie      Adelie      Adelie      Adelie      Adelie      Adelie
## [127] Adelie      Adelie      Adelie      Adelie      Adelie      Adelie      Adelie
## [134] Adelie      Adelie      Adelie      Adelie      Adelie      Adelie      Adelie
## [141] Adelie      Adelie      Adelie      Adelie      Adelie      Adelie      Adelie
## [148] Adelie      Adelie      Adelie      Adelie      Adelie      Gentoo      Gentoo
## [155] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [162] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [169] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [176] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [183] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [190] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [197] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [204] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [211] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [218] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [225] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [232] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [239] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [246] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [253] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [260] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [267] Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo      Gentoo
## [274] Gentoo      Gentoo      Gentoo      Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [281] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [288] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [295] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [302] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [309] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [316] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [323] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [330] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [337] Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap  Chinstrap
## [344] Chinstrap
## Levels: Adelie Gentoo Chinstrap
```

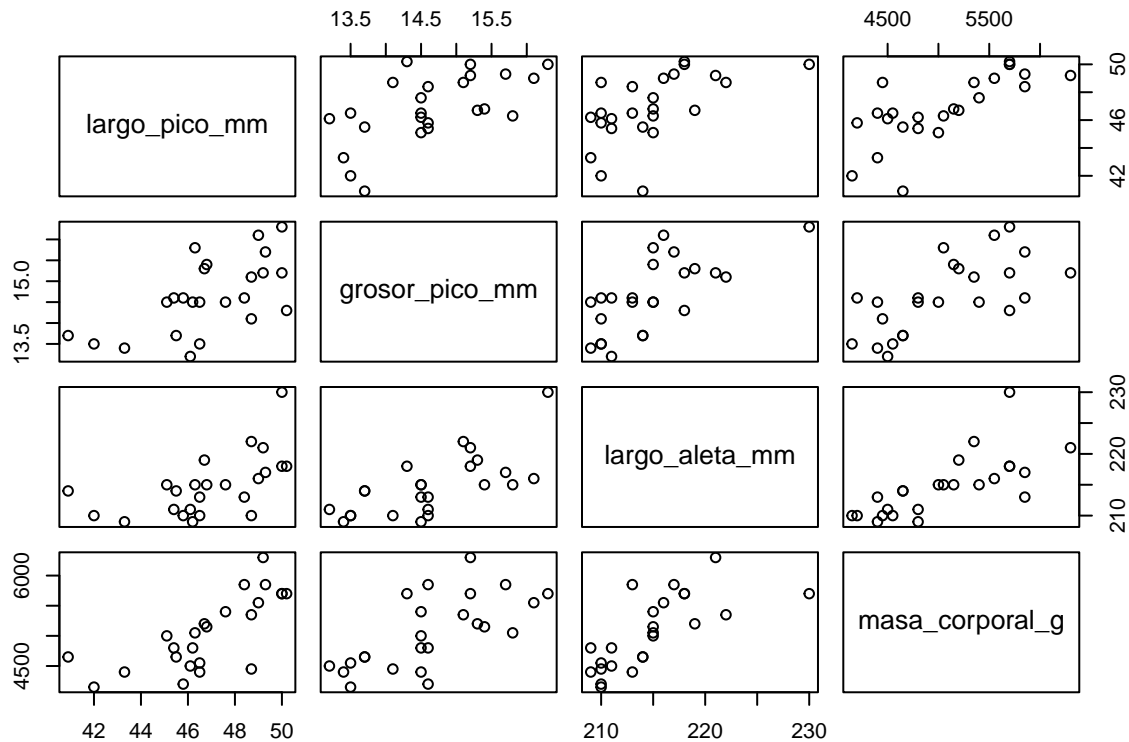
2.2 Seleccionar los datos con los que se van a trabajar y las variables numéricas que se ocuparán. El código es: `gentoo<-penguins[153:176, c(4,5,6,7)]`

```
gentoo<-penguins[153:176, c(4,5,6,7)]
```

Gráfico de dispersión

1.- Mediante la función `pairs` se creará un gráfico de dispersión para visualizar la nueva matriz.

```
pairs(gentoo)
```



Cálculo de la correlación de Pearson

1.- Mediante la función `cor` se calcula la correlación de Pearson. Con el código `cor(BD)`

```
cor(gentoo)
```

```
##                largo_pico_mm grosor_pico_mm largo_aleta_mm masa_corporal_g
## largo_pico_mm      1.0000000      0.6185638      0.5781154      0.7386365
## grosor_pico_mm      0.6185638      1.0000000      0.6931901      0.6735989
## largo_aleta_mm      0.5781154      0.6931901      1.0000000      0.7205205
## masa_corporal_g      0.7386365      0.6735989      0.7205205      1.0000000
```

Gráfico de dispersión con línea de regresión

1.- Se instala la paquetería `ggplot2`

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

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

2.- Se abre la librería

```
library(ggplot2)
```

3.- Se generará el gráfico de dispersión

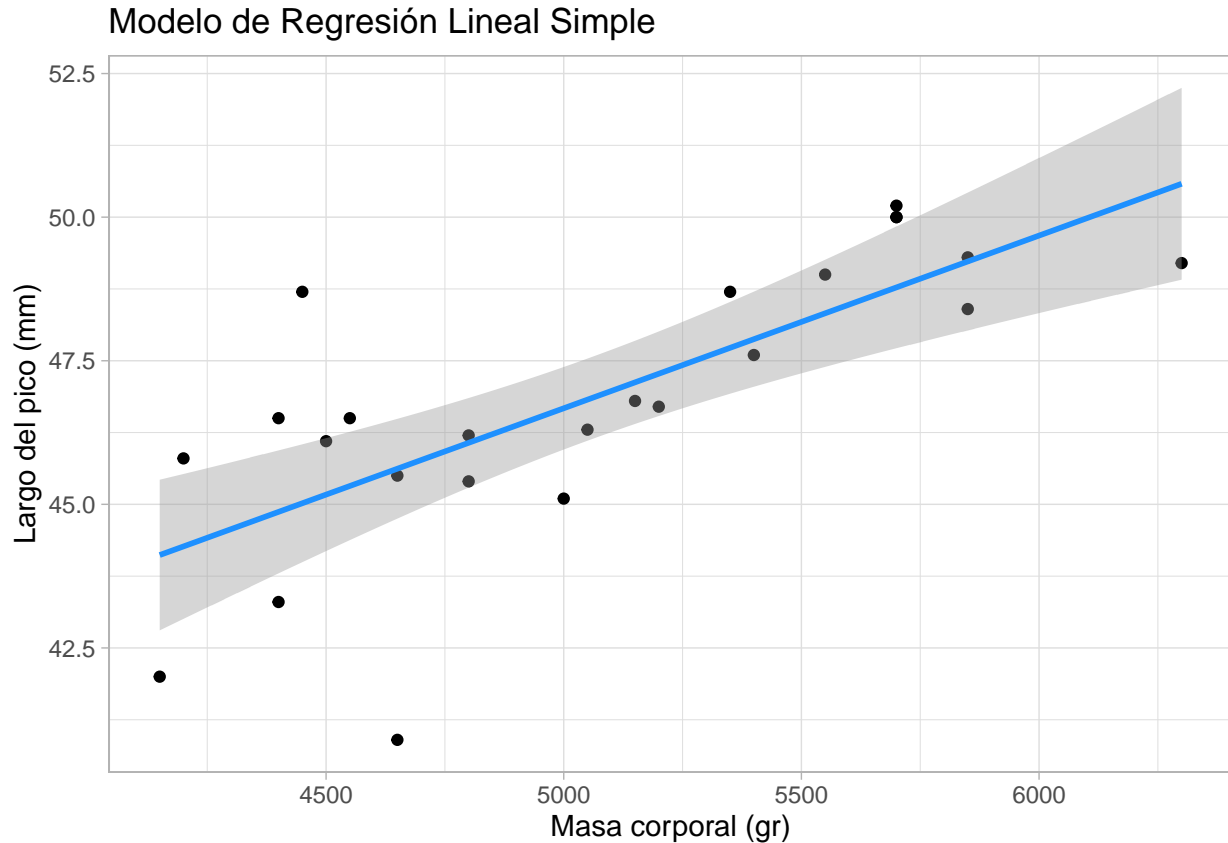
3.1 El objeto se llamará `MRL`, se ocupan las variables `masa corporal` y `largo del pico`.

```
MRL<-ggplot(gentoo, aes(x=masa_corporal_g, y=largo_pico_mm))+
  geom_point()+
  geom_smooth(method = "lm", formula=y~x, col="dodgerblue1")+
  ggtitle("Modelo de Regresión Lineal Simple")+
  xlab("Masa corporal (gr)")+
  ylab("Largo del pico (mm)")
```

```
ylab("Largo del pico (mm)") +
theme_light()
```

3.2 Visualizar el objeto

MRL



Cálculo y representación de la recta por mínimos cuadrados

1.- Con la función `lm` se encuentra el error y el intercepto.

```
regresion<-lm(gentoo$largo_pico_mm~gentoo$masa_corporal_g,
              data=gentoo)
```

2.- Se visualiza el objeto

```
summary(regresion)
```

```
##
## Call:
## lm(formula = gentoo$largo_pico_mm ~ gentoo$masa_corporal_g, data = gentoo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.7203 -0.7105 -0.0242  1.1910  3.6810
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   31.640147   2.977858  10.625 3.96e-10 ***
```

```
## gentoo$masa_corporal_g 0.003007 0.000585 5.139 3.76e-05 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.69 on 22 degrees of freedom
## Multiple R-squared: 0.5456, Adjusted R-squared: 0.5249
## F-statistic: 26.41 on 1 and 22 DF, p-value: 3.761e-05
```

Coefficiente de correlación de Pearson

1.- Con la función sqrt se saca la raíz cuadrada del dato Multiple R-squared para sacar el valor de la correlación de Pearson

```
r<-sqrt(0.5456)
```

2.- Visualización del objeto

```
r
```

```
## [1] 0.7386474
```