Correlaciones

Martha Alejandra Jiménez Sánchez

2022-06-27

#Utilidad

Esta metodología sirve para identificar la relacion entre dos variables de tipo **cuantitativo**. Los datos se distribuyen de forma normal. Area: Estadistica Parametrica Se utiliza la matriz penguins como ejemplo.

#librerias

```
library(readxl)
```

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

#Exploracion de variables

1.Dimensión

```
dim(penguins)
```

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

2. Tipos de variable

```
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. Identificar la especie adelie

```
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"
## [259] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [265] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [271] "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"    "Gentoo"
## [277] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [283] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [289] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [295] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [301] "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap" "Chinstrap"
## [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"
```

```r
colnames(penguins)
```

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

4.Seleccionar las observaciones de la 1 a la 152 y las variables cuantitativas

```
adelie<-penguins[1:152,4:7]
```
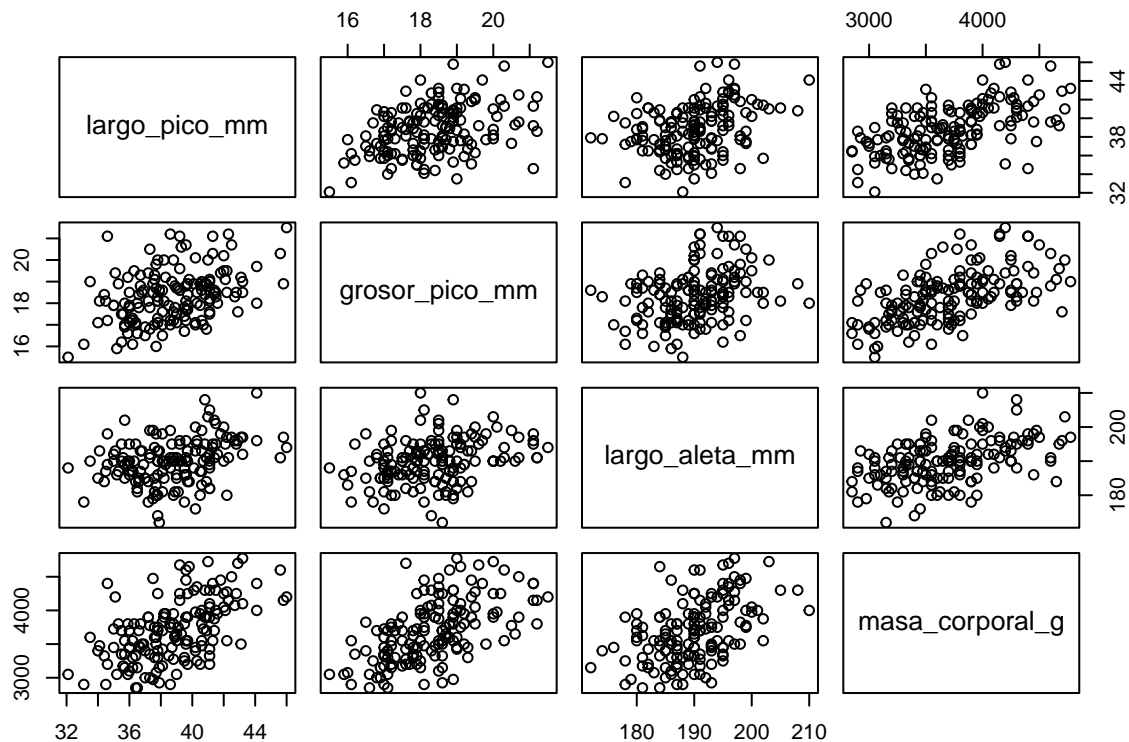
Verificación de la matriz

```
str(adelie)
```

```
## tibble [152 x 4] (S3: tbl_df/tbl/data.frame)
##  $ largo_pico_mm  : num [1:152] 39.1 39.5 40.3 37.8 36.7 39.3 38.9 39.2 34.1 42 ...
##  $ grosor_pico_mm : num [1:152] 18.7 17.4 18 18.1 19.3 20.6 17.8 19.6 18.1 20.2 ...
##  $ largo_aleta_mm : num [1:152] 181 186 195 190 193 190 181 195 193 190 ...
##  $ masa_corporal_g: num [1:152] 3750 3800 3250 3700 3450 ...
```

#Metodología 1.Calculo de la correlación

```
cor_adelie<-cor(adelie)
```

2. realización del plot de correlaciones

```
plot(adelie)
```


3. Organización de tabla de correlacion

Librerias

```
library(knitr)
```

Organización

```
kable(cor_adelie)
```

	largo_pico_mm	grosor_pico_mm	largo_aleta_mm	masa_corporal_g
largo_pico_mm	1.0000000	0.3917580	0.3256178	0.5486177
grosor_pico_mm	0.3917580	1.0000000	0.3075689	0.5760619

	largo_pico_mm	grosor_pico_mm	largo_aleta_mm	masa_corporal_g
largo_aleta_mm	0.3256178	0.3075689	1.0000000	0.4682015
masa_corporal_g	0.5486177	0.5760619	0.4682015	1.0000000

#Correlación Spearman

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

2. exploración de las variables

```
dim(marvel)
```

```
## [1] 39 11
```

3. identificar variables cuantitativas

```
str(marvel)
```

```
## tibble [39 x 11] (S3: tbl_df/tbl/data.frame)
##  $ ID                 : num [1:39] 1 2 3 4 5 6 7 8 9 10 ...
##  $ Original Title     : chr [1:39] "Iron Man" "The Incredible Hulk" "Iron Man 2" "Thor" ...
##  $ Company            : chr [1:39] "Marvel" "Marvel" "Marvel" "Marvel" ...
##  $ Rate               : num [1:39] 7.9 6.7 7 7 6.9 8 7.2 6.9 7.7 8 ...
##  $ Metascore          : num [1:39] 79 61 57 57 66 69 62 54 70 76 ...
##  $ Minutes            : chr [1:39] "126" "112 " "124 " "115" ...
##  $ Release            : num [1:39] 2008 2008 2010 2011 2011 ...
##  $ Budget             : chr [1:39] "140000000" "150000000" "200000000" "150000000 " ...
##  $ Opening Weekend USA: num [1:39] 9.86e+07 5.54e+07 1.28e+08 6.57e+07 6.51e+07 ...
##  $ Gross USA          : num [1:39] 3.19e+08 1.35e+08 3.12e+08 1.81e+08 1.77e+08 ...
##  $ Gross Worldwide    : num [1:39] 5.85e+08 2.63e+08 6.24e+08 4.49e+08 3.71e+08 ...
```

Nombre y posición de las variables

```
colnames(marvel)
```

```
##  [1] "ID"                 "Original Title"      "Company"
##  [4] "Rate"               "Metascore"           "Minutes"
##  [7] "Release"            "Budget"              "Opening Weekend USA"
## [10] "Gross USA"          "Gross Worldwide"
```

selección de las variables rate, minutos, budget y gross.worldwide
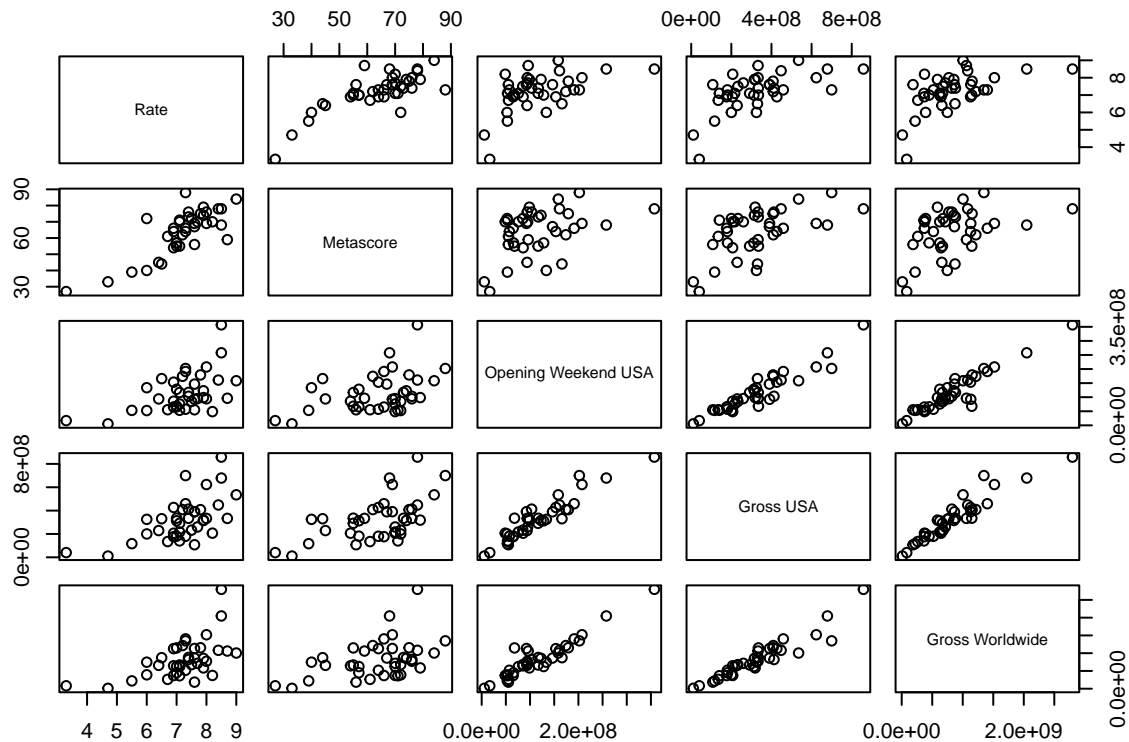
```
marvel<-marvel[,c(4,5,9,10,11)]
```

```
colnames(marvel)
```

```
## [1] "Rate"               "Metascore"           "Opening Weekend USA"
## [4] "Gross USA"          "Gross Worldwide"
```

Plot de exploración

```
plot(marvel)
```

correlación de Spearman

```
spearman<-cor(marvel, method = "spearman")
```

Visualizar el objeto

```
spearman
```

```
##                        Rate Metascore Opening Weekend USA Gross USA
## Rate              1.0000000 0.6938601          0.4711430 0.5830256
## Metascore         0.6938601 1.0000000          0.3733037 0.5201540
## Opening Weekend USA 0.4711430 0.3733037        1.0000000 0.8979757
## Gross USA         0.5830256 0.5201540          0.8979757 1.0000000
## Gross Worldwide   0.5289085 0.3926474          0.8779352 0.9536437
##                   Gross Worldwide
## Rate                    0.5289085
## Metascore               0.3926474
## Opening Weekend USA     0.8779352
## Gross USA               0.9536437
## Gross Worldwide         1.0000000
```

Librerias

```
library(knitr)
```

Se utiliza la funcion kable para tabla en formato markdown.

```
kable(spearman)
```

	Rate	Metascore	Opening Weekend USA	Gross USA	Gross Worldwide
Rate	1.0000000	0.6938601	0.4711430	0.5830256	0.5289085
Metascore	0.6938601	1.0000000	0.3733037	0.5201540	0.3926474

	Rate	Metascore	Opening Weekend USA	Gross USA	Gross Worldwide
Opening Weekend USA	0.4711430	0.3733037	1.0000000	0.8979757	0.8779352
Gross USA	0.5830256	0.5201540	0.8979757	1.0000000	0.9536437
Gross Worldwide	0.5289085	0.3926474	0.8779352	0.9536437	1.0000000