

# Correlación Lineal

2024-02-26

```
nuevo_dir <- "C:/Ejercicios/Correlación" setwd(nuevo_dir)
```

```
#Correlación Lineal
```

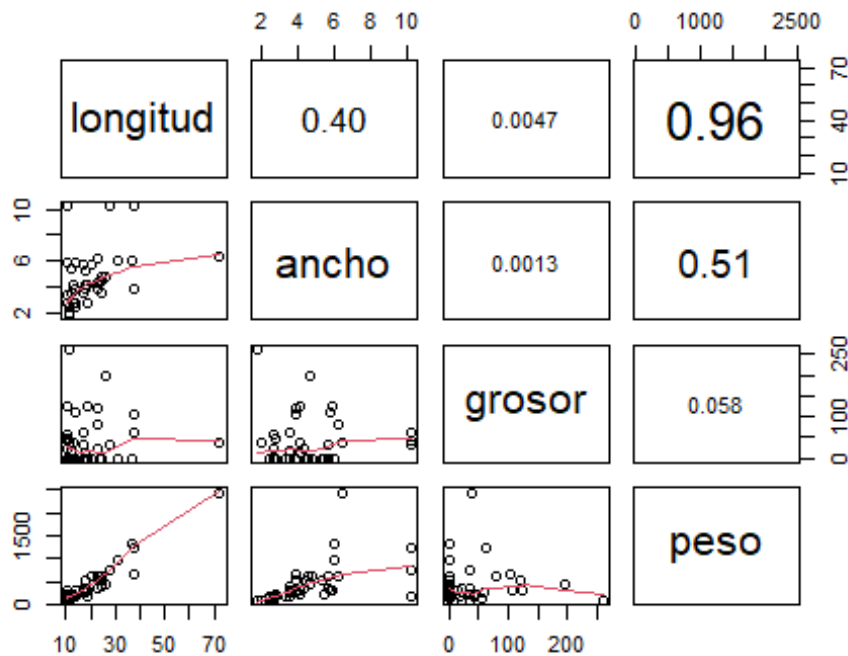
```
library(readxl)
data = as.data.frame(read_excel("C:/Ejercicios/Correlación/data.xlsx"))
data
```

##	longitud	ancho	grosor	peso
## 1	12.4	3.6	17.36	167.0
## 2	22.6	4.3	21.82	342.1
## 3	17.9	4.1	13.54	322.9
## 4	10.2	10.2	40.90	154.8
## 5	16.8	5.7	34.06	358.1
## 6	13.3	4.1	35.36	227.9
## 7	14.1	5.8	108.64	323.8
## 8	10.2	5.9	125.64	285.2
## 9	22.5	6.2	80.20	613.8
## 10	16.9	3.6	60.48	254.3
## 11	19.1	4.1	124.70	310.1
## 12	25.8	4.7	195.78	426.8
## 13	22.5	3.9	121.58	521.2
## 14	27.6	10.2	33.12	765.1
## 15	38.0	10.2	61.58	1217.2
## 16	72.4	6.4	38.48	2446.5
## 17	37.5	3.9	104.94	675.7
## 18	10.2	2.7	22.24	90.9
## 19	11.6	2.0	35.74	86.8
## 20	10.8	2.7	54.68	109.1
## 21	11.4	1.8	260.88	67.7
## 22	10.2	2.8	46.76	204.5
## 23	10.2	3.3	0.00	170.3
## 24	18.6	2.7	0.00	176.8
## 25	24.4	4.4	0.00	543.2
## 26	23.5	4.5	0.00	628.2
## 27	24.8	3.5	0.00	401.0
## 28	14.1	3.9	0.00	302.4
## 29	24.6	4.8	0.00	623.5
## 30	30.9	6.0	0.00	978.9
## 31	20.2	5.7	0.00	607.9
## 32	12.8	2.8	0.00	165.6
## 33	16.9	3.6	0.00	307.9
## 34	14.2	2.8	0.00	192.4
## 35	18.0	5.3	0.00	524.7

```
## 36      11.7    2.4    0.00   111.2
## 37      14.1    2.4    0.00   178.7
## 38      17.7    3.9    0.00   273.4
## 39      36.6    6.0    0.00  1304.4
## 40      12.3    5.4    0.00   233.8

panel.cor <- function(x, y, digits = 2, prefix = "", cex.cor, ...) {
  usr <- par("usr")
  on.exit(par("usr"))
  par(usr = c(0, 1, 0, 1))
  Cor <- abs(cor(x, y))
  txt <- paste0(prefix, format(c(Cor, 0.123456789), digits = digits)[1])
  if(missing(cex.cor)) {
    cex.cor = 0.4 / strwidth(txt)
  }
  text(0.5, 0.5, txt,
       cex = 1 + cex.cor * Cor)
}

pairs(data,
      upper.panel = panel.cor,
      lower.panel = panel.smooth)
```



```
cor.test(data$longitud, data$peso)

##
## Pearson's product-moment correlation
##
```

```
## data: data$longitud and data$peso
## t = 19.989, df = 38, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.9170685 0.9764377
## sample estimates:
##      cor
## 0.9555894

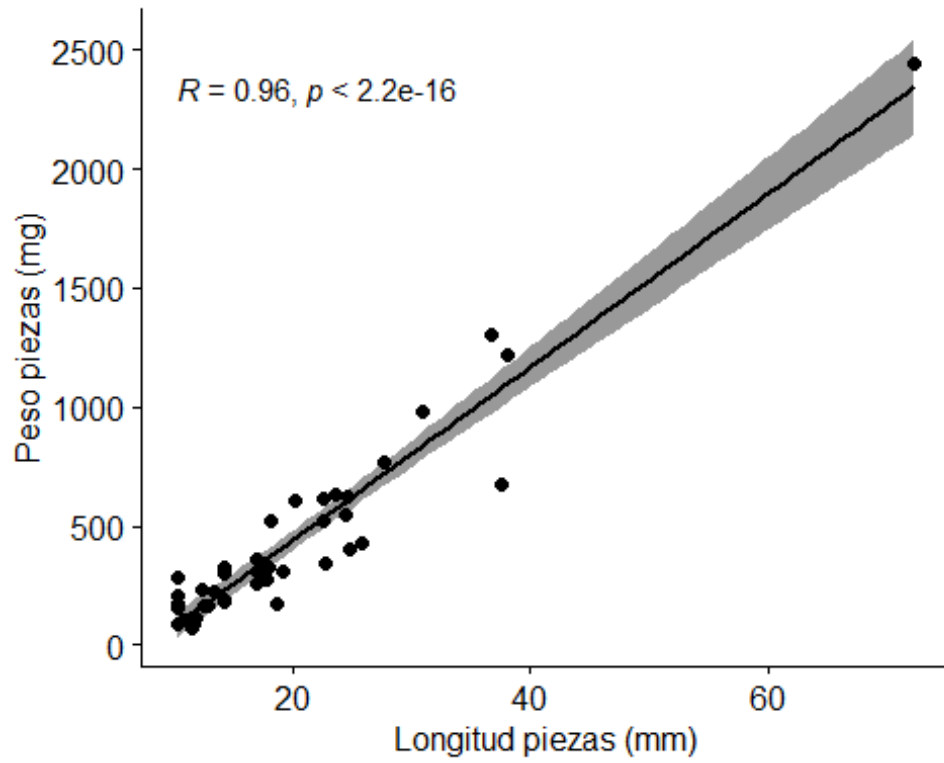
library(correlation)
resultados <- correlation(data)
resultados

## # Correlation Matrix (pearson-method)
##
## Parameter1 | Parameter2 |      r |      95% CI |      t(38) |
p
## -----
-----
## longitud   | ancho      | 0.40 | [ 0.10, 0.63] | 2.71 |
0.040*
## longitud   | grosor     | 4.68e-03 | [-0.31, 0.32] | 0.03 | >
.999
## longitud   | peso       | 0.96 | [ 0.92, 0.98] | 19.99 | <
.001***
## ancho      | grosor     | -1.29e-03 | [-0.31, 0.31] | -7.98e-03 | >
.999
## ancho      | peso       | 0.51 | [ 0.23, 0.71] | 3.64 |
0.004**
## grosor     | peso       | -0.06 | [-0.36, 0.26] | -0.36 | >
.999
##
## p-value adjustment method: Holm (1979)
## Observations: 40

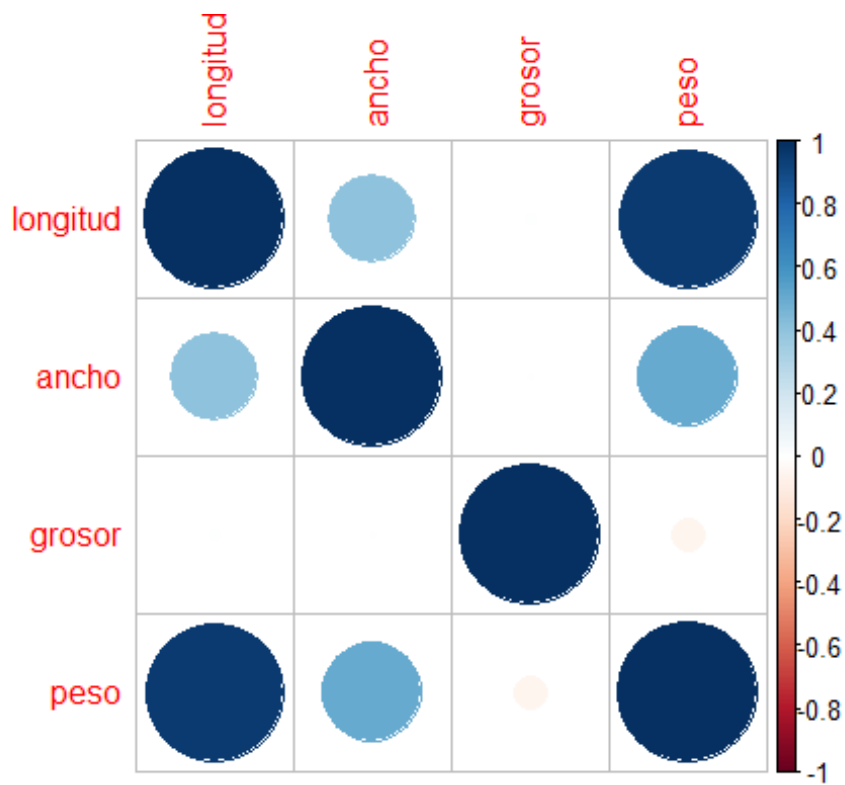
library(ggpubr)

## Loading required package: ggplot2

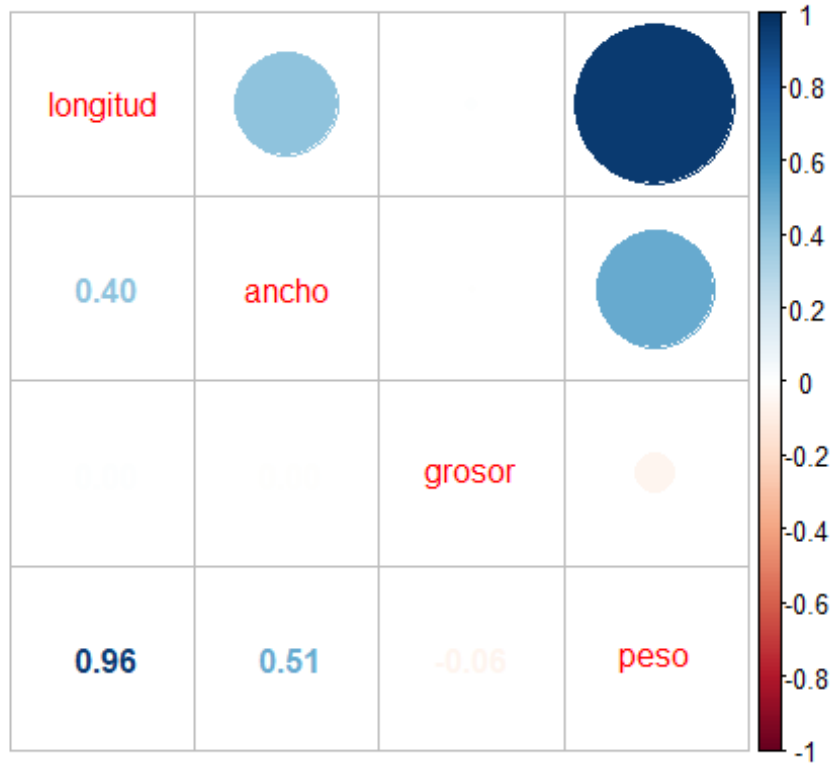
ggscatter(data, x = "longitud", y = "peso",
  add = "reg.line", conf.int = TRUE,
  cor.coef = TRUE, cor.method = "pearson",
  xlab = "Longitud piezas (mm)", ylab = "Peso piezas (mg)")
```



```
library(corrplot)
## corrplot 0.92 loaded
corrplot(cor(data))
```



```
corrplot.mixed(cor(data))
```



```
distancia <- c(1.1, 100.2, 90.3, 5.4, 57.5, 6.6, 34.7, 65.8, 57.9, 86.1)
n_piezas <- c(110, 2, 6, 98, 40, 94, 31, 5, 8, 10)
```

```
dist_ncuent <- data.frame(distancia, n_piezas)
knitr::kable(dist_ncuent)
```

distancia	n_piezas
1.1	110
100.2	2
90.3	6
5.4	98
57.5	40
6.6	94
34.7	31
65.8	5
57.9	8
86.1	10

#CALCULA EL COEFICIENTE DE CORRELACIÓN.

```
library(correlation)
resultado <- correlation(dist_ncuent)
resultado

## # Correlation Matrix (pearson-method)
##
## Parameter1 | Parameter2 |      r |          95% CI |   t(8) |          p
## -----
## distancia |   n_piezas | -0.92 | [-0.98, -0.71] | -6.88 | < .001***
##
## p-value adjustment method: Holm (1979)
## Observations: 10
```

#CALCULA EL NIVEL DE SIGNIFICANCIA

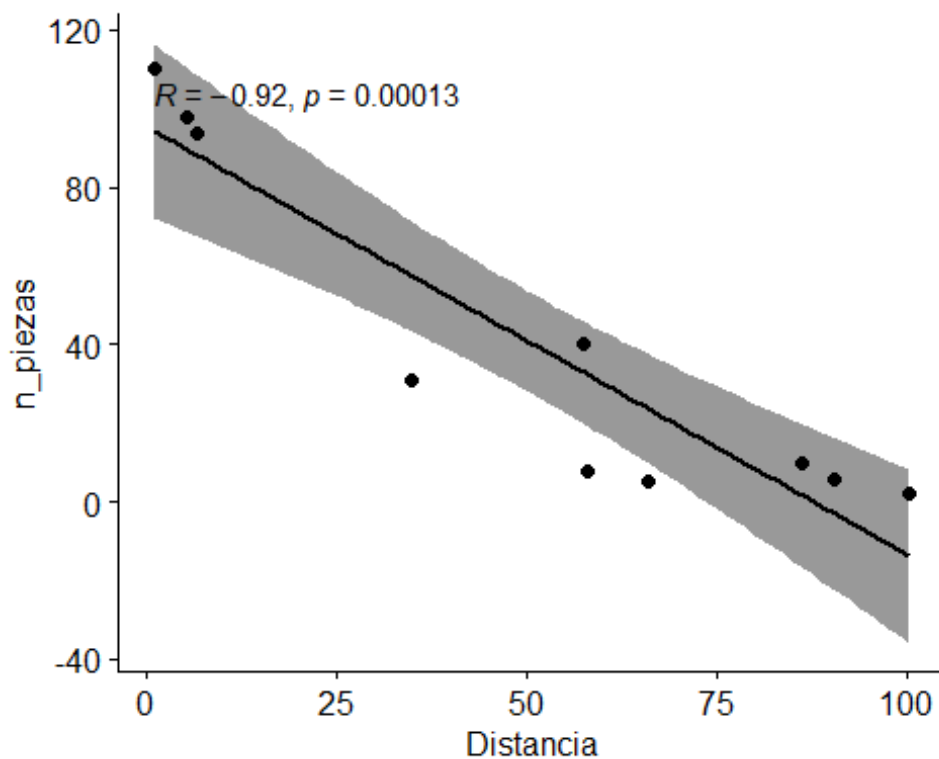
```
cor.test(dist_ncuent$distancia, dist_ncuent$n_piezas)

##
## Pearson's product-moment correlation
##
## data: dist_ncuent$distancia and dist_ncuent$n_piezas
## t = -6.8847, df = 8, p-value = 0.0001265
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.9824414 -0.7072588
## sample estimates:
## cor
## -0.9249824
```

#INTERVALO DE CONFIANZA AL 95% EN RELACIÓN CON EL COEFICIENTE DE CORRELACIÓN. Se observa que el P-VALUE proporciona un resultado de 0.0001265, concluyendo en un margen de error relativamente muy bajo e imposible.

#¿QUÉ INTENSIDAD Y DIRECCIÓN PRESENTAN AMBAS VARIABLES?

```
library(ggpubr)
ggscatter(dist_ncuent, x = "distancia", y = "n_piezas",
  add = "reg.line", conf.int = TRUE,
  cor.coef = TRUE, cor.method = "pearson",
  xlab = "Distancia", ylab = "n_piezas")
```



#¿ES

SIGNIFICATIVA ESTA RELACIÓN? Sí. Debido al P-VALUE tan reducido al porcentaje de confianza.

#¿RESULTARÍA APROPIADO AFIRMAR LA CORRELACIÓN (O NO) ENTRE VARIABLES CON UN TAMAÑO MUESTRAL TAN REDUCIDO (N=10)? No. Ya que un muestral tan reducido no puede generar un gran abanico de datos de confianza, siendo inapropiado.