

```

---
title: "Correlacion"
author: "Guillermo J.C"
output: html_document
date: "29-02-2024"
---

###Introduccion

```{r echo=TRUE}
library(readxl)
data <-as.data.frame(read_excel("C:/data.xlsx"))
data
```

### Dibujamos matriz de correlacion
#Para añadir coeficientes de correlacion
```{r}
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)
}

#Dibujamos matriz

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

```

```{r echo=TRUE}
cor.test(data$longitud, data$peso)
```

Este comando se utiliza para comprobar el azar, debemos fijarnos en el p-value, en el caso de que sea menor a 0,05 es que es fiable, hay menos posibilidades de que se cumpla al azar.

```{r warning=FALSE}
library(correlation)
resultados<-correlation(data)
resultados
```

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

Una correlacion positiva donde los puntos son la interrelacion entre el peso y la longitud de las piezas. La tendencia de la recta lo que marca es que a medida que aumenta una variabe aumenta la otra.

```{r }
library(corrplot)
corrplot(cor(data))
```

El mismo caso que el anterior, a medida que el punto se acerca a 1, es que cuanto mas aumenta una de las variables, la otra tambien.

```{r}

```

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

Y así se haría una gráfica más o menos mixta. Que muestra el coeficiente de correlación con su color.

VECTORES

```
```{r}
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(112,2,6,48,10,94,33,5,3,7)
```

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

```
```
##Calcula coeficiente de correlacion

```{r warning=FALSE}
library(correlation)
resultados<-correlation(dist_ncuent)
resultados
```
El resultado del coeficiente de la correlación es de -0,92, por tanto es negativo
```

```
##Calcula el nivel de significancia
```{r echo=TRUE}
cor.test(dist_ncuent$distancia, dist_ncuent$numero_cuentas)
```
```

El nivel de significancia es -0.9249824

```
##Intervalo de confianza al 95% en relación con el coeficiente de correlacion
```

##¿Qué intensidad y dirección presentan ambas variables?

```
```{r echo=TRUE}
library(ggpubr)
ggscatter(dist_ncuent, x = "distancia", y = "numero_cuentas",
 add = "reg.line", conf.int = TRUE,
 cor.coef = TRUE, cor.method= "pearson",
 xlab = "Distancia (Km)", ylab = "Número de piezas")
```
```

##¿Es significativa esta relación?

La probabilidad de que la correlación entre estas variables sea al azar es mínimo

##¿Resultaría apropiado afirmar la correlación (o no) entre variables con un tamaño muestral tan reducido (n=10)?

Lo apropiado sería obtener el mayor número de muestras posibles.