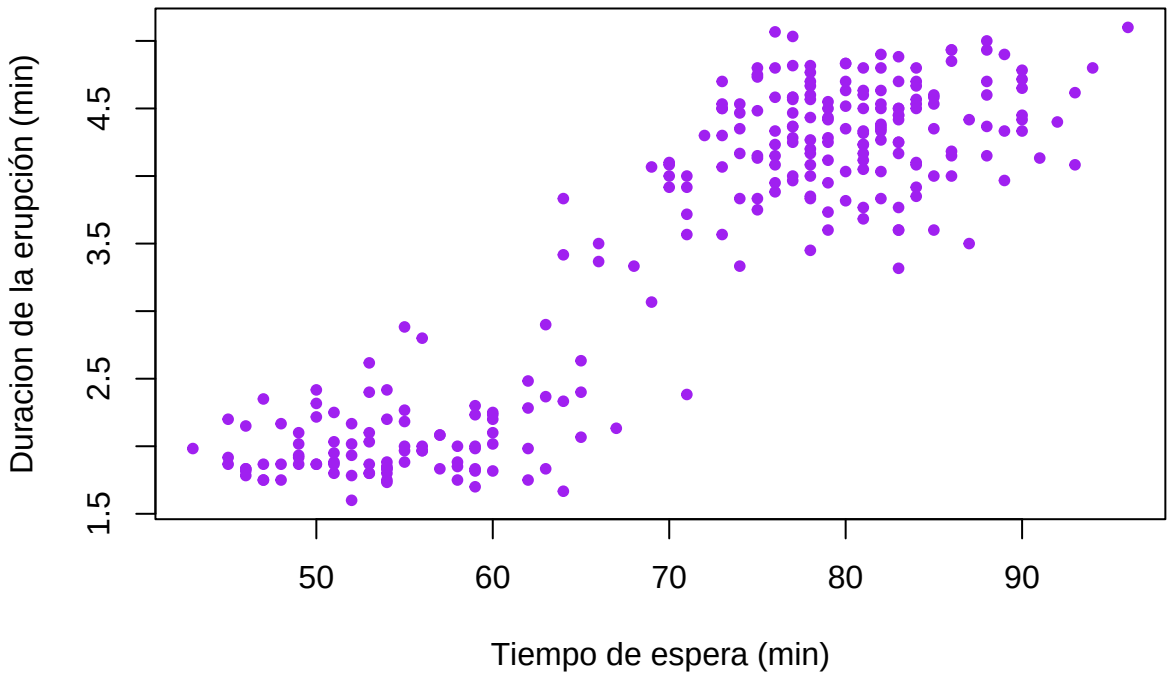
R6.R

Usuario

2025-11-27

```
#=============================================================================
#Correlación de Pearson
#Datos del geyser Old Faithful
#24/09/2025
#=============================================================================

data("faithful")

plot(faithful$waiting, faithful$eruptions,
     xlab= "Tiempo de espera (min)",
     ylab= "Duracion de la erupción (min)",
     col="purple",
     pch=20)
```



```
#############################################################################
#Correlacionar las dos variables ###############################################

shapiro.test(faithful$eruptions)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  faithful$eruptions
## W = 0.84592, p-value = 9.036e-16
```

```
shapiro.test(faithful$waiting)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  faithful$waiting
## W = 0.92215, p-value = 1.015e-10
```

```
#############################################################################
#Pearson solo se utiliza cuando tenemos datos normales#########################
cor.test(faithful$waiting, faithful$eruptions,
         method= "pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  faithful$waiting and faithful$eruptions
## t = 34.089, df = 270, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8756964 0.9210652
## sample estimates:
##       cor
## 0.9008112
```

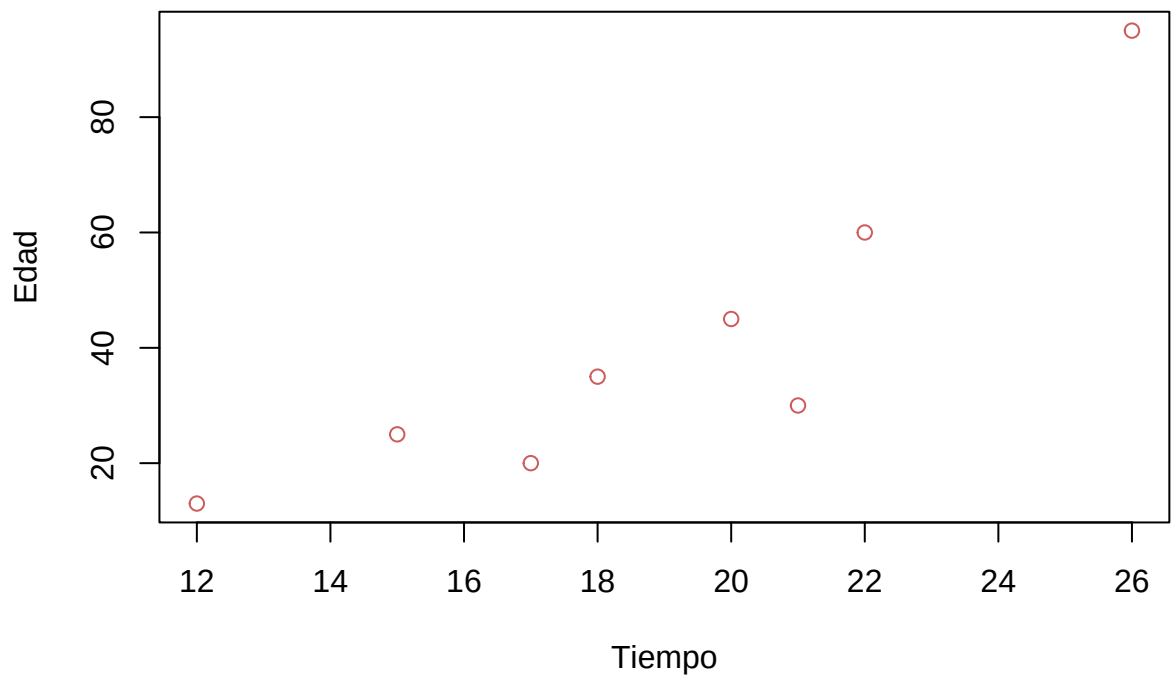
```
#############################################################################
#Spearman se utiliza como contraparte de datos no normales#####################
#(Ordena los datos de manera ascendente)

cor.test(faithful$waiting, faithful$eruptions,
         method= "spearman")
```

```
## Warning in cor.test.default(faithful$waiting, faithful$eruptions, method =
## "spearman"): Cannot compute exact p-value with ties
```

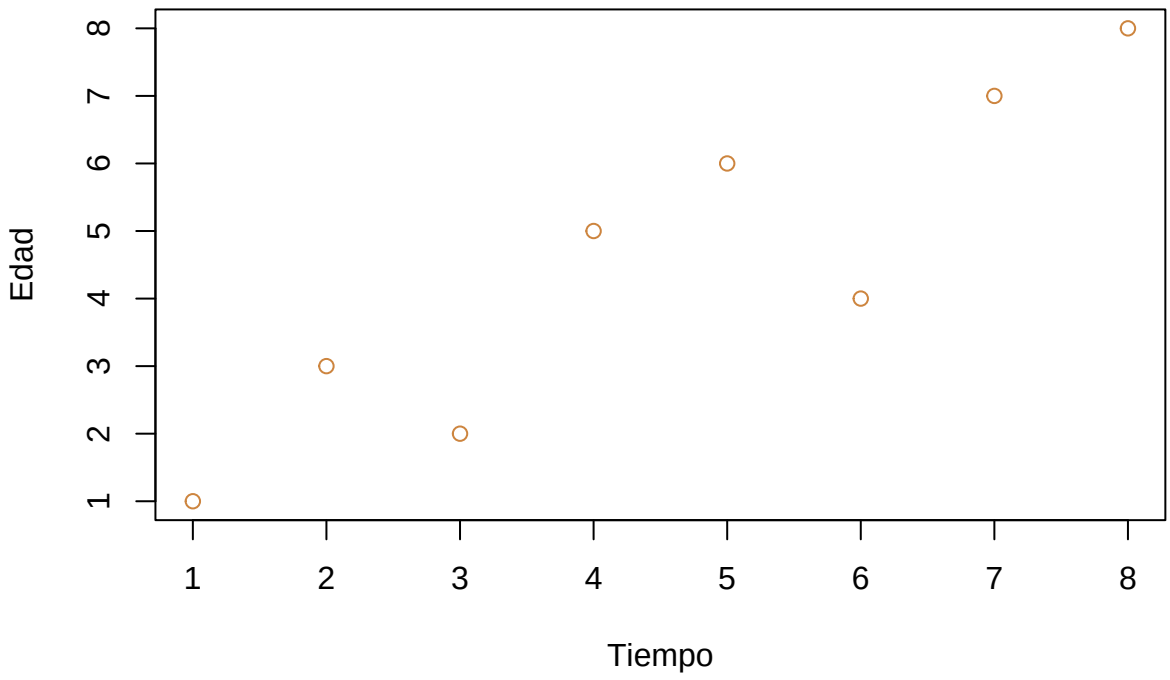
```
##
##  Spearman's rank correlation rho
##
## data:  faithful$waiting and faithful$eruptions
## S = 744659, p-value < 2.2e-16
```

```
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.7779721
```

```r
#=============================================================================
#=============================================================================
#25/09/2025

resp <- data.frame(
  Tiempo=c(12,15,17,18,20,21,22,26),
  Edad=c(13,25,20,35,45,30,60,95)
)

resp$Rango_Tiempo <- rank(resp$Tiempo, ties.method="first")
resp$Rango_Edad <- rank(resp$Edad, ties.method="first")

plot(resp$Tiempo, resp$Edad,
     xlab="Tiempo",
     ylab="Edad",
     col="indianred")
```




```r
plot(resp$Rango_Tiempo, resp$Rango_Edad,
     xlab="Tiempo",
     ylab="Edad",
     col="peru")
```




```
resp$dif <- resp$Rango_Tiempo - resp$Rango_Edad
resp$dif2 <- resp$dif^2
sum(resp$dif2)
```

```
## [1] 8
```

```
cor.test(resp$Rango_Tiempo, resp$Rango_Edad, method="spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  resp$Rango_Tiempo and resp$Rango_Edad
## S = 8, p-value = 0.004563
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9047619
```

```
cor.test(resp$Tiempo,resp$Edad,method="spearman")
```

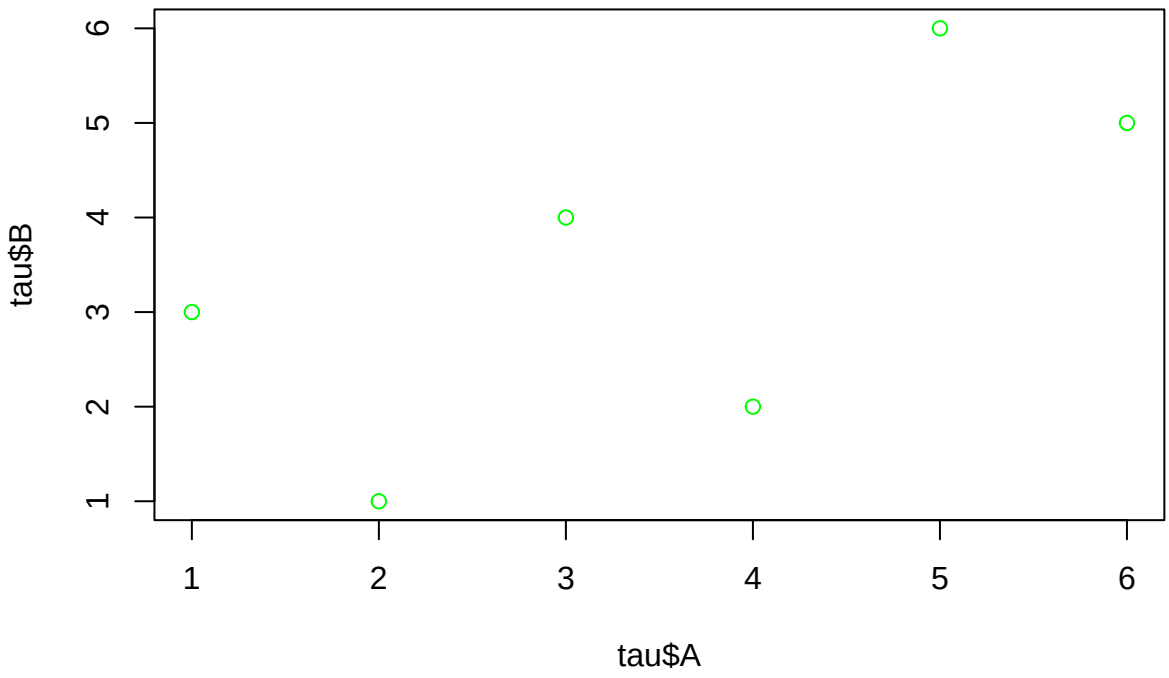
```
##
##  Spearman's rank correlation rho
##
## data:  resp$Tiempo and resp$Edad
```

```
## S = 8, p-value = 0.004563
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9047619
```

```
#=============================================================================
#=============================================================================
#=============================================================================
#Tau de Kendall
tau <- data.frame(
  A=c(1,2,3,4,5,6),
  B=c(3,1,4,2,6,5))
cor.test(tau$A, tau$B, method="kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  tau$A and tau$B
## T = 11, p-value = 0.2722
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.4666667
```

```
plot(tau$A,tau$B,
     col="green")
```




```
#=============================================================================
#=============================================================================
#=============================================================================
#Punto Biserial

set.seed(123)
n <- (20)
Horas_estudio <- sample(1:10,n,replace=T)
Resultado <- sapply(Horas_estudio, function(horas){ifelse(runif(1)<(horas/10),"Aprobado","Reprobado")})

estudio <- data.frame(
  Estudiante=1:n,
  Horas_estudio,
  Resultado
)

estudio$Resultado_bin <- ifelse(estudio$Resultado=="Aprobado",1,0)
head(estudio)
```

```
##   Estudiante Horas_estudio Resultado Resultado_bin
## 1          1             3 Reprobado             0
## 2          2             3 Reprobado             0
## 3          3            10  Aprobado             1
## 4          4             2 Reprobado             0
## 5          5             6  Aprobado             1
## 6          6             5  Aprobado             1
```

```r
cor.test(estudio$Horas_estudio, estudio$Resultado_bin,method="pearson")
```

```
##
##  Pearson's product-moment correlation
##
## data:  estudio$Horas_estudio and estudio$Resultado_bin
## t = 4.5536, df = 18, p-value = 0.0002461
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4275568 0.8869888
## sample estimates:
##       cor
## 0.7316478
```

```r
mean_aprobados <- mean(estudio$Horas_estudio[estudio$Resultado=="Aprobado"])
mean_aprobados
```

```
## [1] 7.733333
```

```r
mean_reprobados <- mean(estudio$Horas_estudio[estudio$Resultado=="Reprobado"])
mean_reprobados
```

```
## [1] 3
```