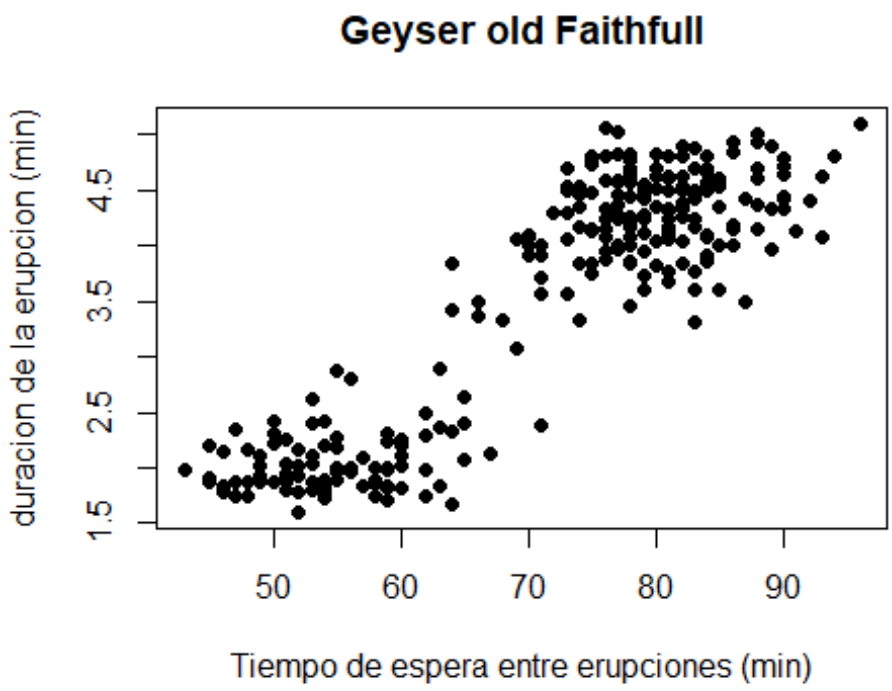
Clase_Semana_15_d1_Examen.R

Usuario

2022-05-11

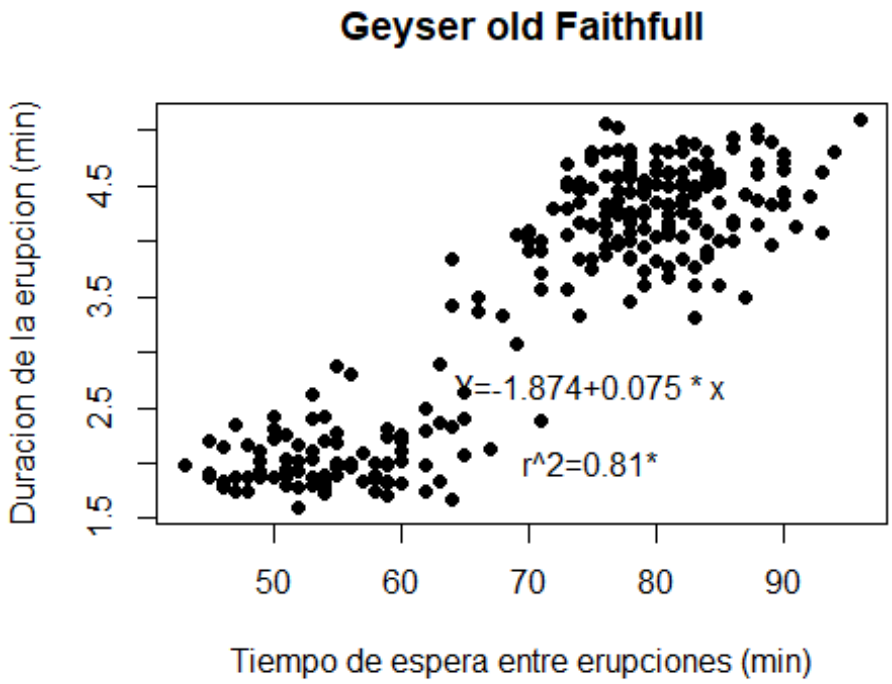
```
# J_Garcia
# 11/05/2022
# SEMANA 15_Examen: Datos de regresión lineal

geiser <- read.csv("erupciones.csv", header = T)
plot(geiser$waiting, geiser$eruptions, pch=19,
     xlab= "Tiempo de espera entre erupciones (min)",
     ylab = "duracion de la erupcion (min)" ,
     main= "Geyser old Faithfull")
```




```
mean(geiser$waiting)

## [1] 70.89706

mean(geiser$eruptions)

## [1] 3.487783

sd(geiser$waiting)
```

```
## [1] 13.59497

sd(geiser$eruptions)

## [1] 1.141371

var(geiser$waiting)

## [1] 184.8233

var(geiser$eruptions)

## [1] 1.302728

cor.test(geiser$waiting, geiser$eruptions)

##
##  Pearson's product-moment correlation
##
## data:  geiser$waiting and geiser$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

# lm para determinar alfa y beta

geiser.lm <- lm (geiser$waiting ~ geiser$eruptions)
summary(geiser.lm)

##
## Call:
## lm(formula = geiser$waiting ~ geiser$eruptions)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.0796  -4.4831   0.2122   3.9246  15.9719
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       33.4744     1.1549   28.98   <2e-16 ***
## geiser$eruptions  10.7296     0.3148   34.09   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.914 on 270 degrees of freedom
## Multiple R-squared:  0.8115, Adjusted R-squared:  0.8108
## F-statistic:  1162 on 1 and 270 DF,  p-value: < 2.2e-16
```

```r
plot(geiser$waiting, geiser$eruptions, pch=19,
     xlab= "Tiempo de espera entre erupciones (min)",
     ylab = "Duracion de la erupcion (min)" ,
     main= "Geyser old Faithfull")
abline(geiser.lm, col= "red")
text(75,2.7, "Y=-1.874+0.075 * x")
text(75,2, "r^2=0.81*")
```




```r
# Encontrar los siguientes valores aplicando la formula de regresión
valores <- c(80, 40, 45, 53, 61)

-1.874+0.075*valores
```

```
## [1] 4.126 1.126 1.501 2.101 2.701
```