

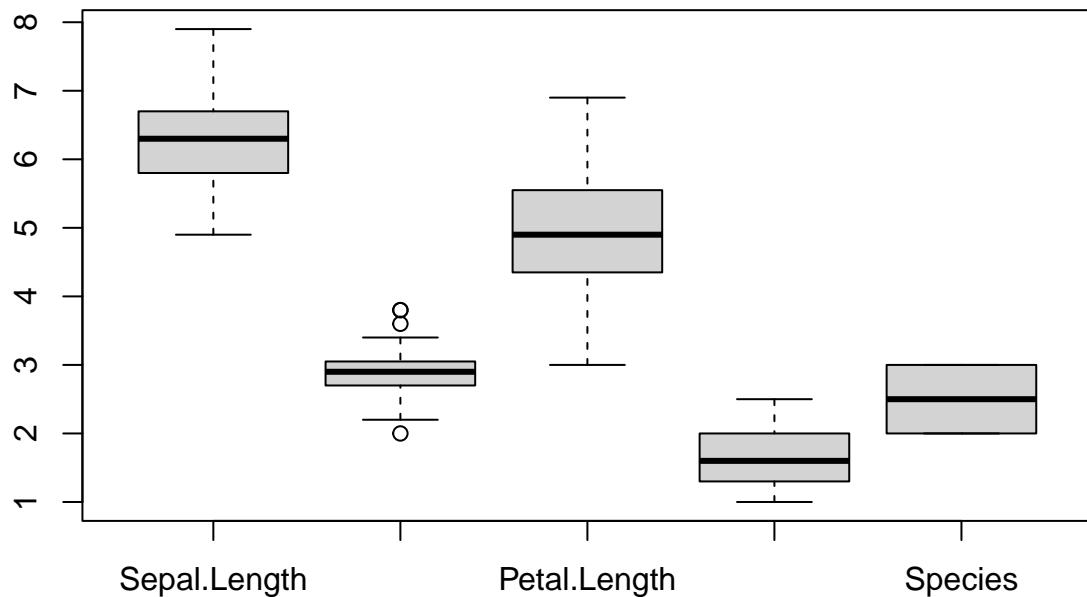
METES-tarea-2.R

Usuario

2025-09-03

```
#Se utiliza solamente dos especies de la base de datos "iris"
data_sub <- subset(iris,Species %in% c("versicolor","virginica"))

#Despues de separar las dos especies que necesito lo visualizo con un diagrama de caja
boxplot(data_sub)
```



```
#La función "head" se utiliza para visualizar las primeras filas de nuestra base de datos
head(data_sub,100) #En este caso le estoy pidiendo que muestre las primeras 100 filas (todas)
```

```
##      Sepal.Length Sepal.Width Petal.Length Petal.Width   Species
## 51          7.0       3.2        4.7       1.4 versicolor
## 52          6.4       3.2        4.5       1.5 versicolor
## 53          6.9       3.1        4.9       1.5 versicolor
```

## 54	5.5	2.3	4.0	1.3	versicolor
## 55	6.5	2.8	4.6	1.5	versicolor
## 56	5.7	2.8	4.5	1.3	versicolor
## 57	6.3	3.3	4.7	1.6	versicolor
## 58	4.9	2.4	3.3	1.0	versicolor
## 59	6.6	2.9	4.6	1.3	versicolor
## 60	5.2	2.7	3.9	1.4	versicolor
## 61	5.0	2.0	3.5	1.0	versicolor
## 62	5.9	3.0	4.2	1.5	versicolor
## 63	6.0	2.2	4.0	1.0	versicolor
## 64	6.1	2.9	4.7	1.4	versicolor
## 65	5.6	2.9	3.6	1.3	versicolor
## 66	6.7	3.1	4.4	1.4	versicolor
## 67	5.6	3.0	4.5	1.5	versicolor
## 68	5.8	2.7	4.1	1.0	versicolor
## 69	6.2	2.2	4.5	1.5	versicolor
## 70	5.6	2.5	3.9	1.1	versicolor
## 71	5.9	3.2	4.8	1.8	versicolor
## 72	6.1	2.8	4.0	1.3	versicolor
## 73	6.3	2.5	4.9	1.5	versicolor
## 74	6.1	2.8	4.7	1.2	versicolor
## 75	6.4	2.9	4.3	1.3	versicolor
## 76	6.6	3.0	4.4	1.4	versicolor
## 77	6.8	2.8	4.8	1.4	versicolor
## 78	6.7	3.0	5.0	1.7	versicolor
## 79	6.0	2.9	4.5	1.5	versicolor
## 80	5.7	2.6	3.5	1.0	versicolor
## 81	5.5	2.4	3.8	1.1	versicolor
## 82	5.5	2.4	3.7	1.0	versicolor
## 83	5.8	2.7	3.9	1.2	versicolor
## 84	6.0	2.7	5.1	1.6	versicolor
## 85	5.4	3.0	4.5	1.5	versicolor
## 86	6.0	3.4	4.5	1.6	versicolor
## 87	6.7	3.1	4.7	1.5	versicolor
## 88	6.3	2.3	4.4	1.3	versicolor
## 89	5.6	3.0	4.1	1.3	versicolor
## 90	5.5	2.5	4.0	1.3	versicolor
## 91	5.5	2.6	4.4	1.2	versicolor
## 92	6.1	3.0	4.6	1.4	versicolor
## 93	5.8	2.6	4.0	1.2	versicolor
## 94	5.0	2.3	3.3	1.0	versicolor
## 95	5.6	2.7	4.2	1.3	versicolor
## 96	5.7	3.0	4.2	1.2	versicolor
## 97	5.7	2.9	4.2	1.3	versicolor
## 98	6.2	2.9	4.3	1.3	versicolor
## 99	5.1	2.5	3.0	1.1	versicolor
## 100	5.7	2.8	4.1	1.3	versicolor
## 101	6.3	3.3	6.0	2.5	virginica
## 102	5.8	2.7	5.1	1.9	virginica
## 103	7.1	3.0	5.9	2.1	virginica
## 104	6.3	2.9	5.6	1.8	virginica
## 105	6.5	3.0	5.8	2.2	virginica
## 106	7.6	3.0	6.6	2.1	virginica
## 107	4.9	2.5	4.5	1.7	virginica

```

## 108      7.3      2.9      6.3      1.8  virginica
## 109      6.7      2.5      5.8      1.8  virginica
## 110      7.2      3.6      6.1      2.5  virginica
## 111      6.5      3.2      5.1      2.0  virginica
## 112      6.4      2.7      5.3      1.9  virginica
## 113      6.8      3.0      5.5      2.1  virginica
## 114      5.7      2.5      5.0      2.0  virginica
## 115      5.8      2.8      5.1      2.4  virginica
## 116      6.4      3.2      5.3      2.3  virginica
## 117      6.5      3.0      5.5      1.8  virginica
## 118      7.7      3.8      6.7      2.2  virginica
## 119      7.7      2.6      6.9      2.3  virginica
## 120      6.0      2.2      5.0      1.5  virginica
## 121      6.9      3.2      5.7      2.3  virginica
## 122      5.6      2.8      4.9      2.0  virginica
## 123      7.7      2.8      6.7      2.0  virginica
## 124      6.3      2.7      4.9      1.8  virginica
## 125      6.7      3.3      5.7      2.1  virginica
## 126      7.2      3.2      6.0      1.8  virginica
## 127      6.2      2.8      4.8      1.8  virginica
## 128      6.1      3.0      4.9      1.8  virginica
## 129      6.4      2.8      5.6      2.1  virginica
## 130      7.2      3.0      5.8      1.6  virginica
## 131      7.4      2.8      6.1      1.9  virginica
## 132      7.9      3.8      6.4      2.0  virginica
## 133      6.4      2.8      5.6      2.2  virginica
## 134      6.3      2.8      5.1      1.5  virginica
## 135      6.1      2.6      5.6      1.4  virginica
## 136      7.7      3.0      6.1      2.3  virginica
## 137      6.3      3.4      5.6      2.4  virginica
## 138      6.4      3.1      5.5      1.8  virginica
## 139      6.0      3.0      4.8      1.8  virginica
## 140      6.9      3.1      5.4      2.1  virginica
## 141      6.7      3.1      5.6      2.4  virginica
## 142      6.9      3.1      5.1      2.3  virginica
## 143      5.8      2.7      5.1      1.9  virginica
## 144      6.8      3.2      5.9      2.3  virginica
## 145      6.7      3.3      5.7      2.5  virginica
## 146      6.7      3.0      5.2      2.3  virginica
## 147      6.3      2.5      5.0      1.9  virginica
## 148      6.5      3.0      5.2      2.0  virginica
## 149      6.2      3.4      5.4      2.3  virginica
## 150      5.9      3.0      5.1      1.8  virginica

```

#Utilizo "summary" para tener las variables descriptivas de cada una de las columnas (minimo, maximo, media, etc)

```

##   Sepal.Length   Sepal.Width    Petal.Length   Petal.Width
## Min.   :4.900   Min.   :2.000   Min.   :3.000   Min.   :1.000
## 1st Qu.:5.800  1st Qu.:2.700  1st Qu.:4.375  1st Qu.:1.300
## Median :6.300  Median :2.900  Median :4.900  Median :1.600
## Mean   :6.262  Mean   :2.872  Mean   :4.906  Mean   :1.676
## 3rd Qu.:6.700  3rd Qu.:3.025  3rd Qu.:5.525  3rd Qu.:2.000
## Max.   :7.900  Max.   :3.800  Max.   :6.900  Max.   :2.500

```

```
##          Species
##  setosa      : 0
##  versicolor:50
##  virginica :50
##
##
```

#Utilizo tapply para tener un resumen de Petal.Length separado por especie

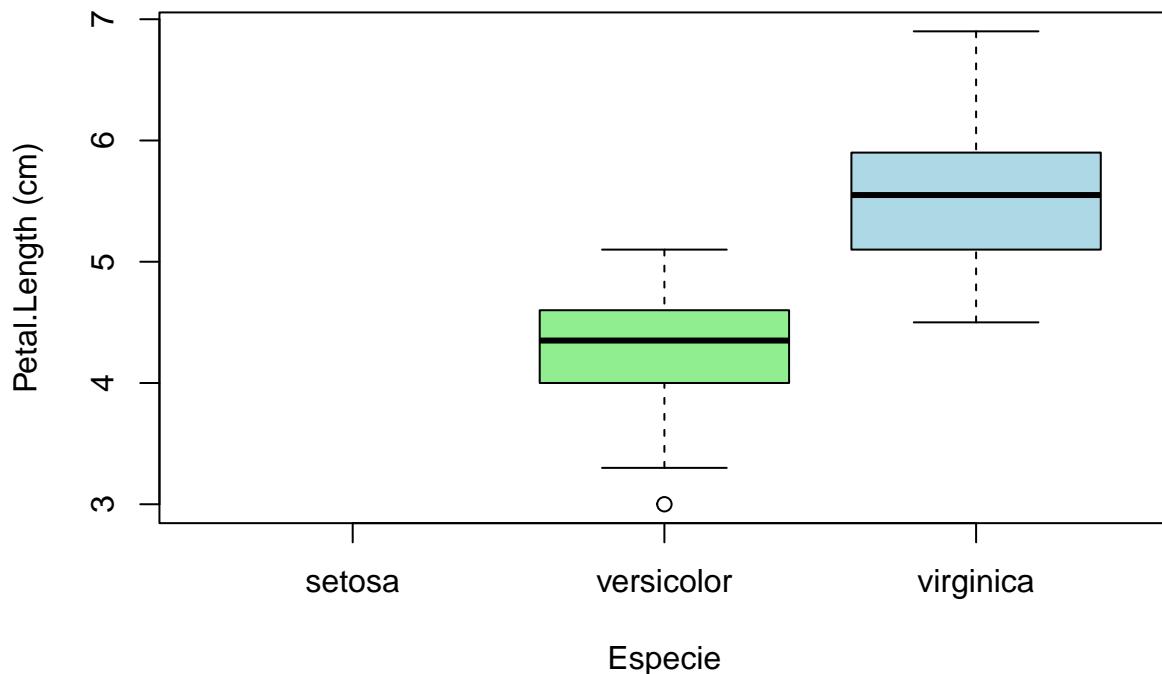
```
tapply(data_sub$Petal.Length, data_sub$Species, summary)
```

```
## $setosa
## NULL
##
## $versicolor
##   Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##   3.00    4.00   4.35    4.26    4.60    5.10
##
## $virginica
##   Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##   4.500   5.100   5.550    5.552   5.875   6.900
```

#Ahora utilizo el diagrama de caja para visualizar Petal.Length por especie

```
boxplot(Petal.Length ~ Species, data = data_sub,
        main = "Largo de Pétalo por Especie",
        xlab = "Especie", ylab = "Petal.Length (cm)",
        col = c("lightblue", "lightgreen"))
```

Largo de Pétalo por Especie



#¿Existe una diferencia significativa en la longitud promedio del pétalo entre versicolor y virginica?

#Hipótesis nula: No existe diferencia significativa entre la longitud media del pétalo de versicolor y virginica
#Hipótesis alternativa: existe diferencia significativa entre la longitud media del pétalo de versicolor y virginica

```
library(car) #Se descargó el paquete "car"
```

```
## Cargando paquete requerido: carData
```

```
#Prueba Levene para varianzas
leveneTest(Petal.Length ~ Species, data = data_sub)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##          Df F value Pr(>F)
## group    1  1.0674 0.3041
##         98
```

```
#Pueba F para varianzas
var.test(Petal.Length ~ Species, data = data_sub)
```

```
##
##  F test to compare two variances
##
## data: Petal.Length by Species
```

```

## F = 0.72497, num df = 49, denom df = 49, p-value = 0.2637
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.411402 1.277530
## sample estimates:
## ratio of variances
## 0.7249678

```

```

#Verificamos varianzas
var_test <- var.test(Petal.Length ~ Species, data = data_sub)
var_test$p.value # Si p > 0.05, varianzas iguales

```

```

## [1] 0.2637454

```

```

# Ejecutamos prueba t
t_test_result <- t.test(Petal.Length ~ Species,
                         data = data_sub,
                         alternative = "two.sided",
                         var.equal = FALSE)

```

```

t_test_result

```

```

##
## Welch Two Sample t-test
##
## data: Petal.Length by Species
## t = -12.604, df = 95.57, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group versicolor and group virginica is not
## 95 percent confidence interval:
## -1.49549 -1.08851
## sample estimates:
## mean in group versicolor mean in group virginica
## 4.260 5.552

```

```

# Calculamos Cohen's d manualmente
library(effectsizes)

cohens_d <- cohens_d(Petal.Length ~ Species,
                      data = data_sub,
                      pooled_sd = TRUE)

versicolor <- data_sub$Petal.Length[data_sub$Species == "versicolor"]
virginica <- data_sub$Petal.Length[data_sub$Species == "virginica"]

d <- (mean(virginica) - mean(versicolor)) /
  sqrt(((sd(versicolor)^2 + sd(virginica)^2)/2))

cohens_d

```

```

## Cohen's d | 95% CI
## -----

```

```
## -2.52      | [-3.04, -1.99]
##
## - Estimated using pooled SD.
```

```
d
```

```
## [1] 2.520756
```

```
boxplot(Petal.Length ~ Species,
        data = data_sub,
        main = "Longitud del Pétalo por Especie",
        xlab = "Especie",
        ylab = "Petal.Length (cm)",
        col = c("yellow", "peru"),
        border = "darkgray",
        notch = TRUE)
```

