Laboratorio-2.R

DELL LATITUDE 3510

2025-10-01

```
#Laboratorio-2
##carlos andrade
##2025-09-04
data("iris")
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

```
summary(iris)
```

```
##   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
##  1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
##  Median :5.800   Median :3.000   Median :4.350   Median :1.300
##  Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
##  3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
##  Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
##        Species
##  setosa    :50
##  versicolor:50
##  virginica :50
##
##
##
```

```
str(iris)
```

```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 ...
```

```r
data_sub <-subset(iris,Species %in% c("versicolor","virginica") )
table(data_sub$Species)
```

```
##
##     setosa versicolor  virginica
##          0         50         50
```

```r
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
```

```r
tapply(data_sub$Petal.Length,data_sub$Species,mean)
```

```
##     setosa versicolor  virginica
##         NA      4.260      5.552
```

```r
tapply(data_sub$Petal.Length, data_sub$Species, sd)
```

```
##     setosa versicolor  virginica
##         NA  0.4699110  0.5518947
```

```r
var.test(Petal.Length ~ Species,data = data_sub)
```

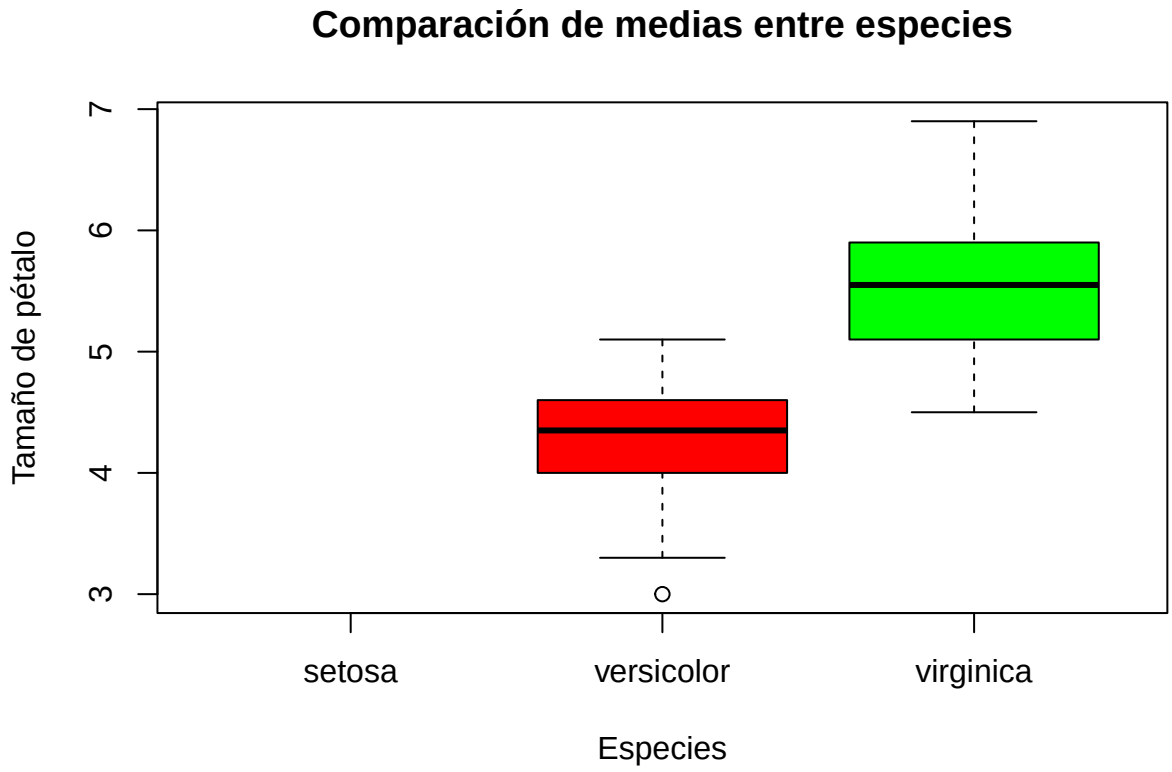
```
##
##  F test to compare two variances
##
## data:  Petal.Length by Species
## F = 0.72497, num df = 49, denom df = 49, p-value = 0.2637
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.411402 1.277530
## sample estimates:
## ratio of variances
##          0.7249678
```

```r
t.test(Petal.Length ~ Species, data = data_sub, var.equal = FALSE)
```

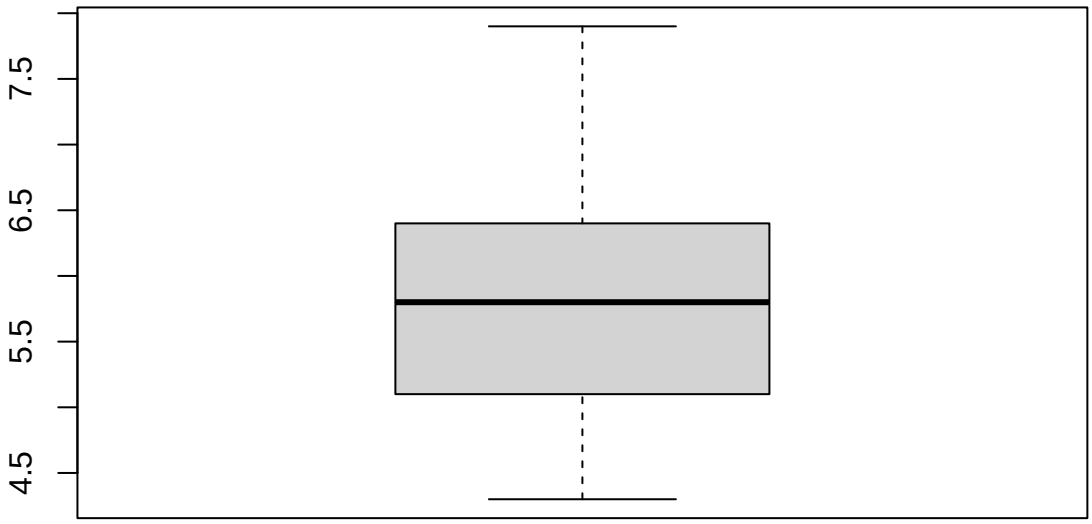
```
##
##  Welch Two Sample t-test
##
```

```
## 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
```

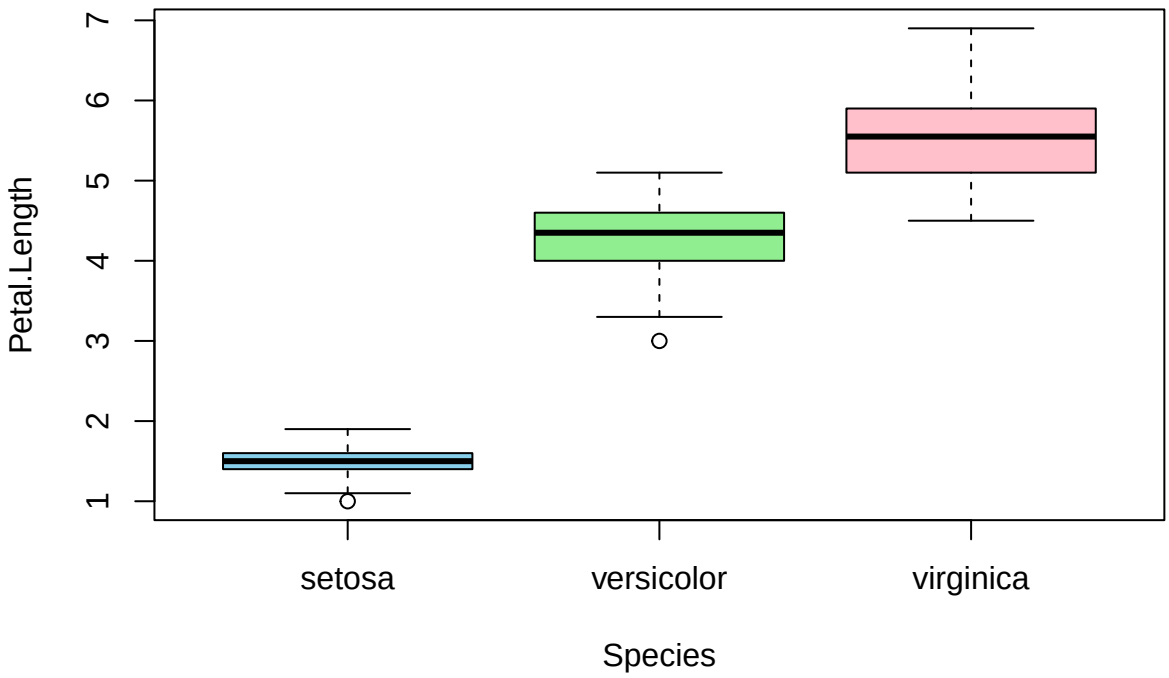
```
boxplot(Petal.Length ~ Species, data = data_sub,
col= c("green","red"),
main = "Comparación de medias entre especies",
xlab = "Especies",
ylab = "Tamaño de pétalo")
```



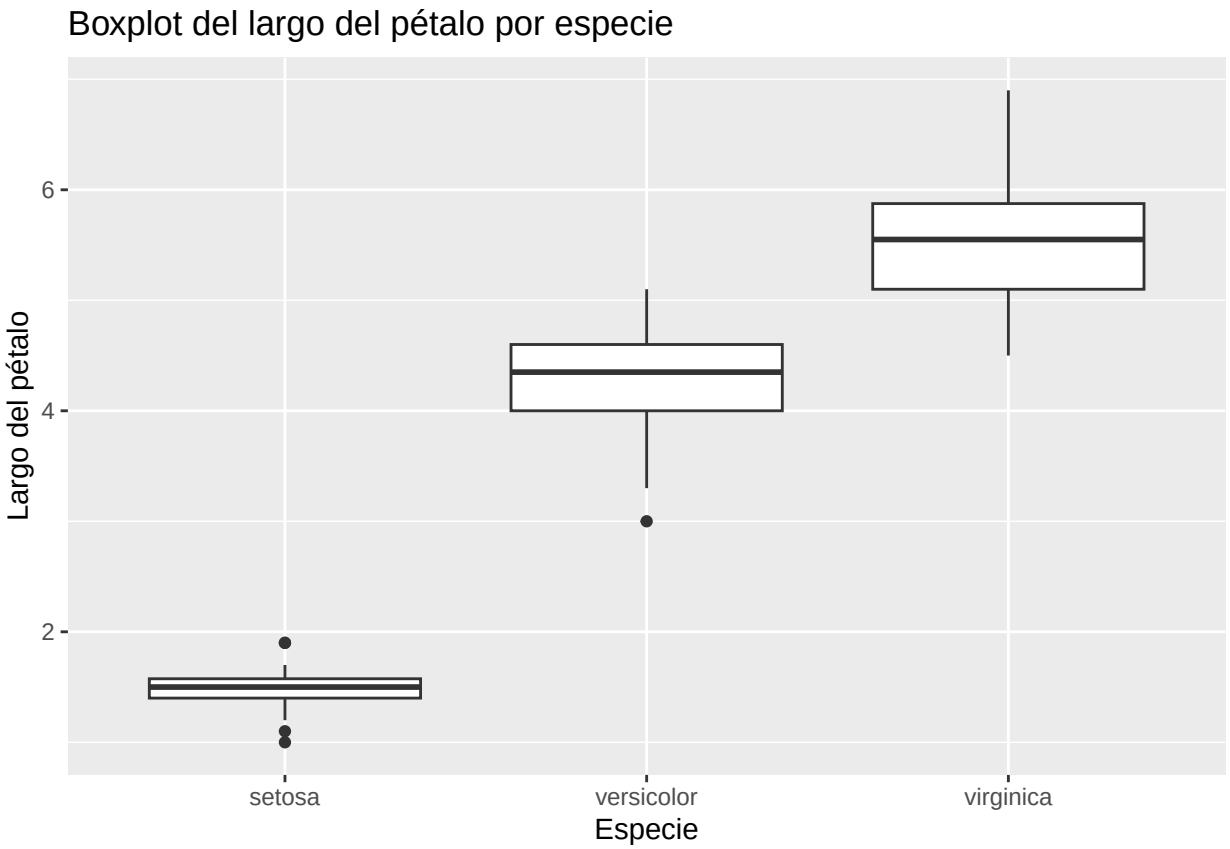

```
boxplot(iris$Sepal.Length)
```



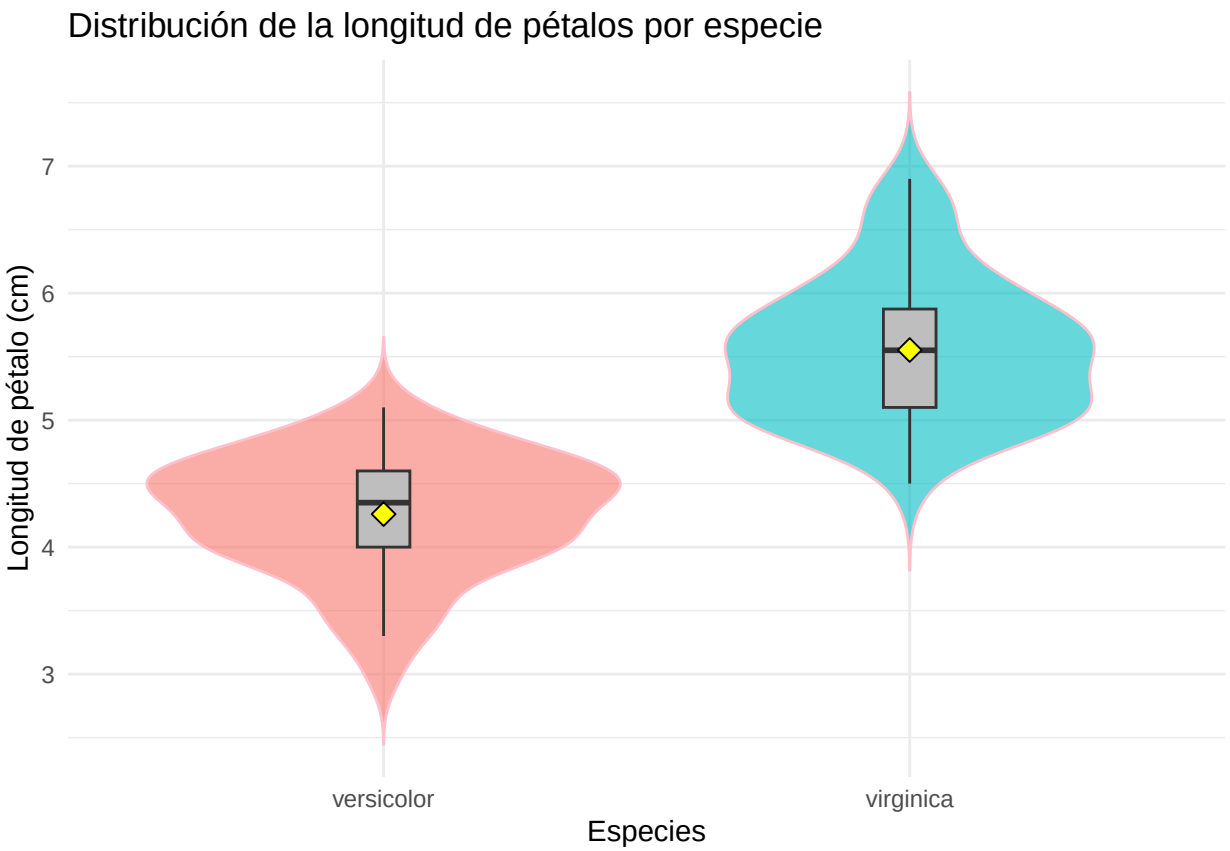

```
boxplot(Petal.Length ~ Species, data = iris, col= c("skyblue", "lightgreen","pink"))
```




```
library(ggplot2)
ggplot(iris, aes(x = Species, y = Petal.Length)) +
geom_boxplot()+ labs(title = "Boxplot del largo del pétalo por especie",
                     x = "Especie",
                     y = "Largo del pétalo")
```




```
ggplot(data_sub, aes(x = Species, y = Petal.Length, fill = Species)) +
geom_violin(trim = FALSE, color = "pink", alpha = 0.6) +
geom_boxplot(width = 0.1, fill = "gray", outlier.shape = NA) +
stat_summary(fun = mean, geom = "point", shape = 23, size = 3, fill = "yellow")+
theme_minimal() +
labs(title = "Distribución de la longitud de pétalos por especie",
       x = "Especies",
       y = "Longitud de pétalo (cm)") + theme(legend.position = "none")
```




```
library(effsize)
cohen.d(Petal.Length~Species,data = data_sub)
```

```
## Warning in cohen.d.default(d, f, subject = subject, ...): Factor with multiple
## levels, using only the two actually present in data

##
## Cohen's d
##
## d estimate: NaN (NA)
## 95 percent confidence interval:
## lower upper
##   NaN   NaN
```