

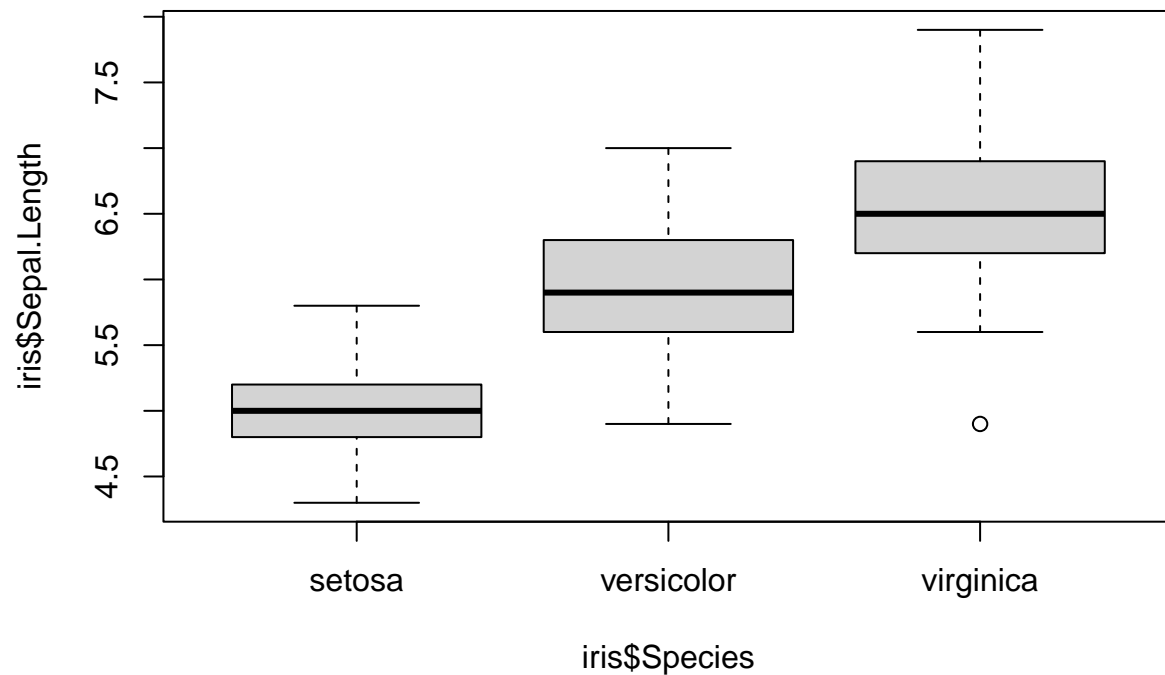
# LAB\_2\_CONTRASTE-DE-MEDIAS\_MAYA-TOVAR.R

RH

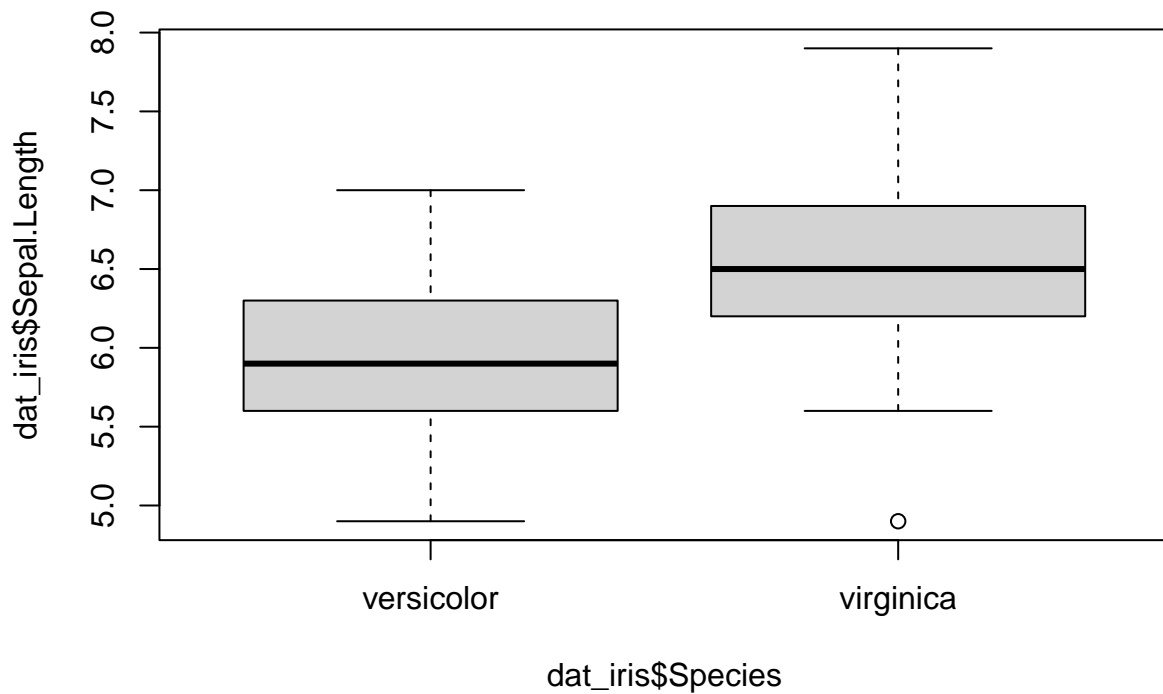
2025-09-04

```
#Maya Tovar  
#Semana5  
#04/09/2025
```

```
data("iris")  
boxplot(iris$Sepal.Length ~ iris$Species)
```



```
dat_iris <- subset(iris, Species %in% c("versicolor",  
                                         "virginica"))  
  
#Para eliminar un factor que se desecho de la base de datos original  
dat_iris$Species <- droplevels(dat_iris$Species)  
boxplot(dat_iris$Sepal.Length ~ dat_iris$Species)
```



*#H0 = la media de sepal length de virginica es mayor que la media de versicolor*

```
t.test(dat_iris$Sepal.Length ~ dat_iris$Species, alternative = "greater",
       var.equal = T)
```

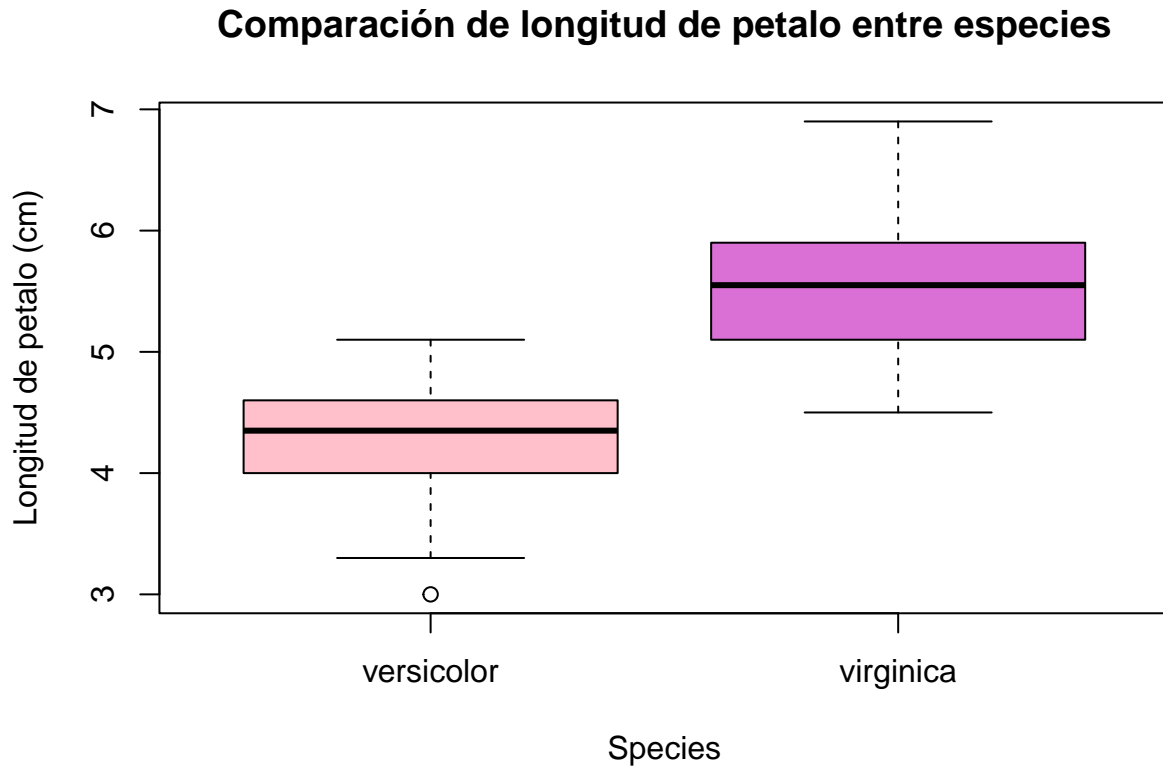
```
##
## Two Sample t-test
##
## data: dat_iris$Sepal.Length by dat_iris$Species
## t = -5.6292, df = 98, p-value = 1
## alternative hypothesis: true difference in means between group versicolor and group virginica is greater than 0
## 95 percent confidence interval:
## -0.8443339      Inf
## sample estimates:
## mean in group versicolor mean in group virginica
##          5.936          6.588
```

```
tapply(dat_iris$Sepal.Length, dat_iris$Species, mean)
```

```
## versicolor virginica
##      5.936      6.588
```

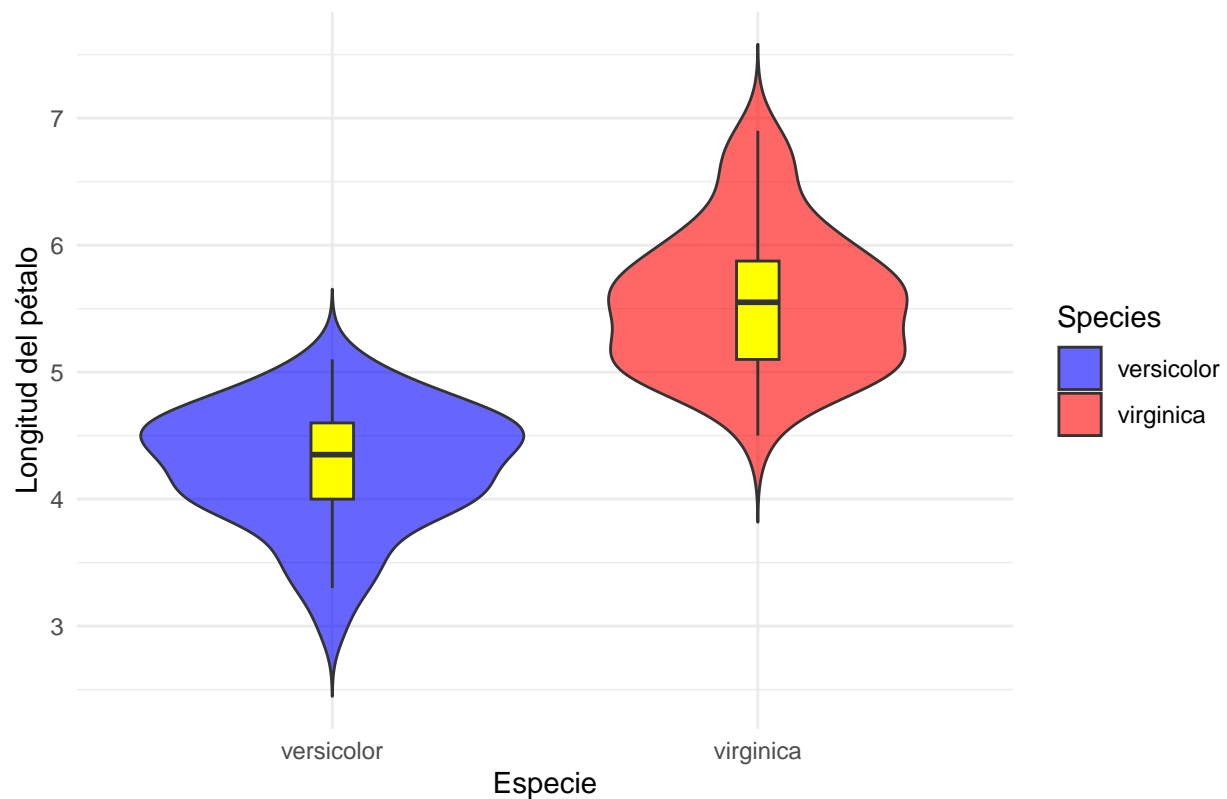
```
boxplot(Petal.Length ~ Species, data = dat_iris,
        col = c("pink", "orchid"),
```

```
main = "Comparación de longitud de petalo entre especies",
ylab = "Longitud de petalo (cm)")
```



```
# Gráfico de violín con boxplot superpuesto
library(ggplot2)
ggplot(dat_iris, aes(x = Species, y = Petal.Length, fill = Species)) +
  geom_violin(trim = FALSE, alpha = 0.6) +
  geom_boxplot(width = 0.1, fill = "yellow", outlier.shape = NA) +
  labs(title = "Distribución de longitud de pétalo por especie",
       x = "Especie",
       y = "Longitud del pétalo") +
  theme_minimal() +
  scale_fill_manual(values = c("blue", "red"))
```

## Distribución de longitud de pétalo por especie



```
#Prueba de igualdad de varianzas
```

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

```
##  
## 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
```

```
#Prueba t (Welch si varianzas distintas)
```

```
t_test_result <- t.test(Petal.Length ~ Species, data = dat_iris,  
                        alternative = "two.sided", var.equal = FALSE)  
print(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
```

```
cohens_efecto <- function(x, y) {
  n1 <- length(x); n2 <- length(y)
  s1 <- sd(x); s2 <- sd(y)
  sp <- sqrt(((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2))
  (mean(x) - mean(y)) / sp
}
data("iris")
df_versicolor <- subset(iris, Species == "versicolor")
df_virginica <- subset(iris, Species == "virginica")
d_cal <- cohens_efecto(df_versicolor$Petal.Length,
  df_virginica$Petal.Length)
d_cal
```

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

```
# Planteamiento del problema e hipótesis
# El objetivo de este análisis fue comparar las medias de características
# morfológicas entre dos especies del género Iris: versicolor y virginica, con
# énfasis en determinar si virginica presenta una mayor longitud de sépalo que
# versicolor.

# H0: La media de la longitud del sépalo en virginica no es mayor que en
#versicolor
# H1: La media de la longitud del sépalo en virginica es mayor que en versicolor
#
# Resultados numéricos y gráficos
# En cuanto a la longitud del sépalo, la media fue de 5.936 cm en versicolor y
# 6.588 cm en virginica. Se aplicó una prueba t unilateral para evaluar si
# virginica presenta sépalos más largos que versicolor.
# El resultado fue t = -5.6292 con un valor p = 1, lo que indica que no se
# rechaza la hipótesis nula, es decir, no hay evidencia suficiente para afirmar
# que virginica tiene sépalos significativamente más largos que versicolor.

# Respecto a la longitud del pétalo, las medias fueron de 4.260 cm en versicolor
# y 5.552 cm en virginica.
# El test de Welch arrojó un valor t = -12.604 con p < 2.2e-16, lo que evidencia
# una diferencia estadísticamente significativa entre ambas especies. La prueba
# de igualdad de varianzas (F = 0.7249, p = 0.2637) no mostró evidencia
# suficiente para asumir varianzas distintas.
# Además, el tamaño del efecto Cohen.d = -2.52 indica una diferencia muy grande
# en la longitud de los pétalos entre especies.

# Los gráficos de caja y violín reforzaron estos hallazgos, mostrando que
# virginica tiende a presentar valores más altos tanto en sépalos como en
# pétalos, con distribuciones desplazadas hacia longitudes mayores.
```

*# Interpretación estadística y biológica*

*# Desde el análisis estadístico, no se encontró una diferencia clara en la longitud del sépalo entre las especies.*

*# En cambio, la longitud del pétalo sí mostró una diferencia muy marcada, con resultados significativos y un efecto grande.*

*# Biológicamente, la longitud del pétalo es importante para distinguir entre especies de Iris, ya que ayuda a atraer polinizadores y se usa en su clasificación. En cambio, la longitud del sépalo no parece ser tan útil para diferenciar entre versicolor y virginica.*

*# En conclusión, los análisis realizados permiten confirmar que, aunque ambas especies presentan diferencias morfológicas en sépalos y pétalos, es principalmente en la longitud de los pétalos donde se observan diferencias significativas y biológicamente relevantes para la identificación y separación de especies dentro del género Iris.*