

# Proyecto Final

Kristty Pamela Jiménez Montiel

## **Análisis Exploratorio del Conjunto de Datos de Plantas del Género Iris: Relación entre las Dimensiones de Sépalos y Pétalos en Diferentes Especies**

### **Objetivo del análisis**

El objetivo de este análisis es explorar las características morfológicas de las plantas del género *Iris* y cómo estas se relacionan con las tres especies presentes en el conjunto de datos: *setosa*, *versicolor* y *virginica*. A través de este análisis, se espera identificar patrones en las medidas de longitud y ancho de los sépalos y pétalos de las distintas especies, así como establecer diferencias significativas entre ellas en cuanto a sus características morfológicas.

### **Método**

El análisis se basa en el conjunto de datos `iris` del paquete `datasets` de R, que contiene 150 observaciones de tres especies de plantas *Iris* y las medidas de cuatro características morfológicas: longitud y ancho de los sépalos y los pétalos. Se utilizaron técnicas de análisis descriptivo y visualización de datos para investigar la distribución de las características por especie. Además, se aplicó un análisis de varianza (ANOVA) para evaluar si existen diferencias significativas entre las especies en cuanto a las medidas de las características. También se generará un gráfico de dispersión para observar la relación entre las variables.

### **Resultados**

El análisis descriptivo de las características morfológicas de las especies del género *Iris* reveló variaciones notables en las medidas de los sépalos y pétalos. La longitud del sépalo varió entre 4.3 cm y 7.9 cm, con un promedio de 5.84 cm, mientras que la anchura del sépalo osciló entre 2.0 cm y 4.4 cm, con una media de 3.06 cm. En cuanto a los pétalos, la longitud mostró una variabilidad entre 1.0 cm y 6.9 cm, con una media de 3.76 cm, y la anchura de los pétalos fluctuó entre 0.1 cm y 2.5 cm, con un valor promedio de 1.2 cm.

Los resultados del análisis de varianza (ANOVA) mostraron diferencias estadísticamente significativas entre las especies en cuanto a la longitud del sépalo y del pétalo. Para la longitud del sépalo, el valor de  $p$  fue inferior a  $2e-16$ , lo que indica que existen diferencias claras entre las especies en esta medida. Un hallazgo similar se observó para la longitud del pétalo, con un valor de  $p$  también menor a  $2e-16$ , sugiriendo que las especies presentan diferencias morfológicas significativas en esta característica.

El gráfico de dispersión entre la longitud del sépalo y la longitud del pétalo mostró una correlación positiva generalizada. Es decir, a medida que aumentaba la longitud del sépalo, también tendía a aumentar la longitud del pétalo. Las especies se agruparon en diferentes áreas del gráfico: *Iris setosa* (representada en rojo) mostró los valores más bajos tanto para la longitud del sépalo como para el pétalo. Por su parte, *Iris versicolor* (verde) y *Iris virginica* (azul) presentaron valores más altos, con *Iris virginica* destacándose por tener las dimensiones más grandes en ambas características. Aunque se observó una tendencia general, también se notó una cierta superposición entre las especies, lo que indica que la longitud del sépalo y del pétalo no son características completamente exclusivas de cada especie.

## Discusión

Los resultados obtenidos en este estudio evidencian patrones interesantes sobre la variabilidad morfológica entre las especies del género *Iris*, especialmente en las longitudes de sépalo y pétalo. En primer lugar, se observó una clara diferenciación entre las especies *Iris setosa*, *Iris versicolor* e *Iris virginica*, tanto en las medidas de los sépalos como en las de los pétalos. *Iris setosa* mostró las menores dimensiones en ambas características, lo que podría indicar una adaptación a un nicho ecológico distinto, posiblemente más árido o con condiciones de crecimiento más limitadas. Por otro lado, *Iris virginica* presentó las mayores longitudes de sépalo y pétalo, lo que podría asociarse con especies que se desarrollan en ambientes más húmedos o fértiles, donde el mayor tamaño de las flores podría facilitar la polinización por insectos o el éxito reproductivo.

En términos de la dispersión de los datos, la agrupación de las especies en el gráfico de dispersión sugiere que las longitudes del sépalo y del pétalo están correlacionadas de manera positiva, lo que indica que, en general, las especies con sépalos más largos tienden a presentar también pétalos más largos. Esta relación puede ser un indicio de un patrón evolutivo en el que un aumento en el tamaño floral puede ser ventajoso para la polinización o la competencia por recursos. Sin embargo, la superposición observada entre las especies, especialmente entre *Iris versicolor* y *Iris virginica*, resalta que las características morfológicas de estas especies no son completamente exclusivas, lo que podría estar relacionado con factores ambientales similares que influyen en su desarrollo y tamaño floral.

## Running Code

When

```
# Cargar datos
library(datasets)
data(iris)

# Análisis descriptivo
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
```

```
# Análisis de varianza (ANOVA) para comparar especies en cuanto a la longitud del sépalo
anova_sepal <- aov(Sepal.Length ~ Species, data = iris)
summary(anova_sepal)
```

```
              Df Sum Sq Mean Sq F value Pr(>F)
Species        2  63.21  31.606   119.3 <2e-16 ***
Residuals    147   38.96   0.265
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Análisis de varianza (ANOVA) para comparar especies en cuanto a la longitud del pétalo
anova_petal <- aov(Petal.Length ~ Species, data = iris)
summary(anova_petal)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Species	2	437.1	218.55	1180	<2e-16 ***
Residuals	147	27.2	0.19		

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
# Gráfico de dispersión
library(ggplot2)
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length, color = Species)) +
  geom_point() +
  labs(title = "Dispersión de Longitudes de Sépalo y Pétalo",
       x = "Longitud del Sépalo",
       y = "Longitud del Pétalo")
```

