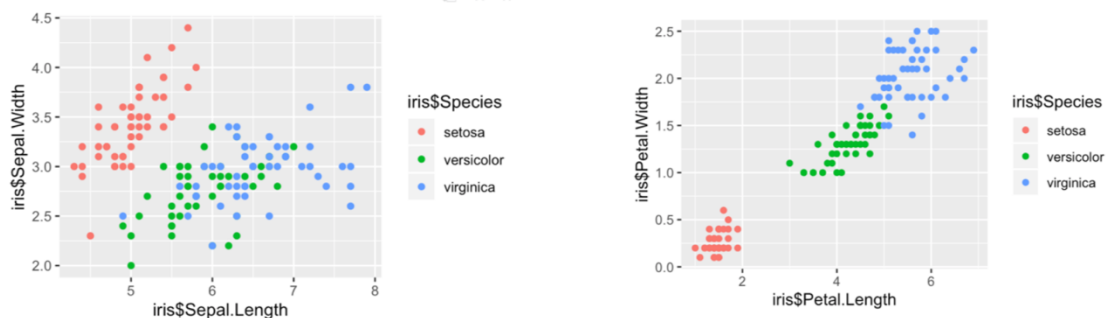


IRIS dataset:

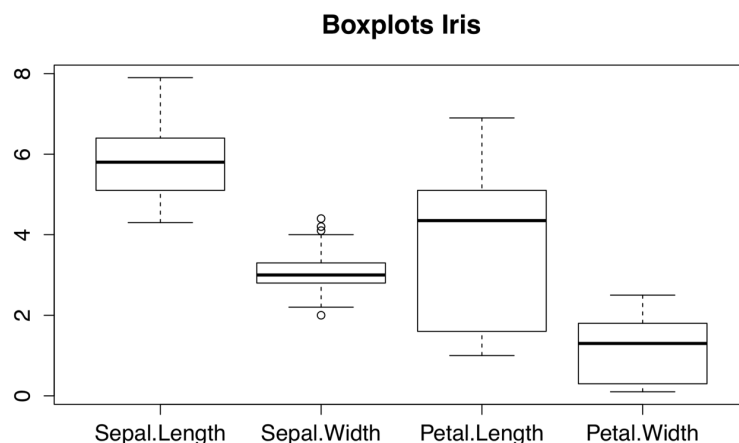
Víctor Marinas

Iris dataset, es un conjunto de datos que contiene 3 clases de 50 observaciones cada una. Fue utilizada para demostrar que estas medidas podrían utilizarse para diferenciar entre especies de plantas Iris.

En el Gráfico 1, podemos observar como la clasificación de las tres especies es visualmente óptima, ya que los datos de las tres especies apenas se entremezclan. Se puede apreciar que la clase Setosa, es linealmente independiente de las otras dos, pero no podemos decir lo mismo de Versicolor y virgínica. Mediante el rango de las variables, se determina si son iguales o distintas las observaciones. La longitud y ancho del pétalo, son mucho más pequeños en la clase Setosa, en cambio en las otras dos son parecidos, aunque son un poco más grandes en virgínica. En cuanto al tamaño del sépalo, la longitud es menor en setosa y mayor en virgínica, el ancho es menor en versicolor y mayor en setosa.



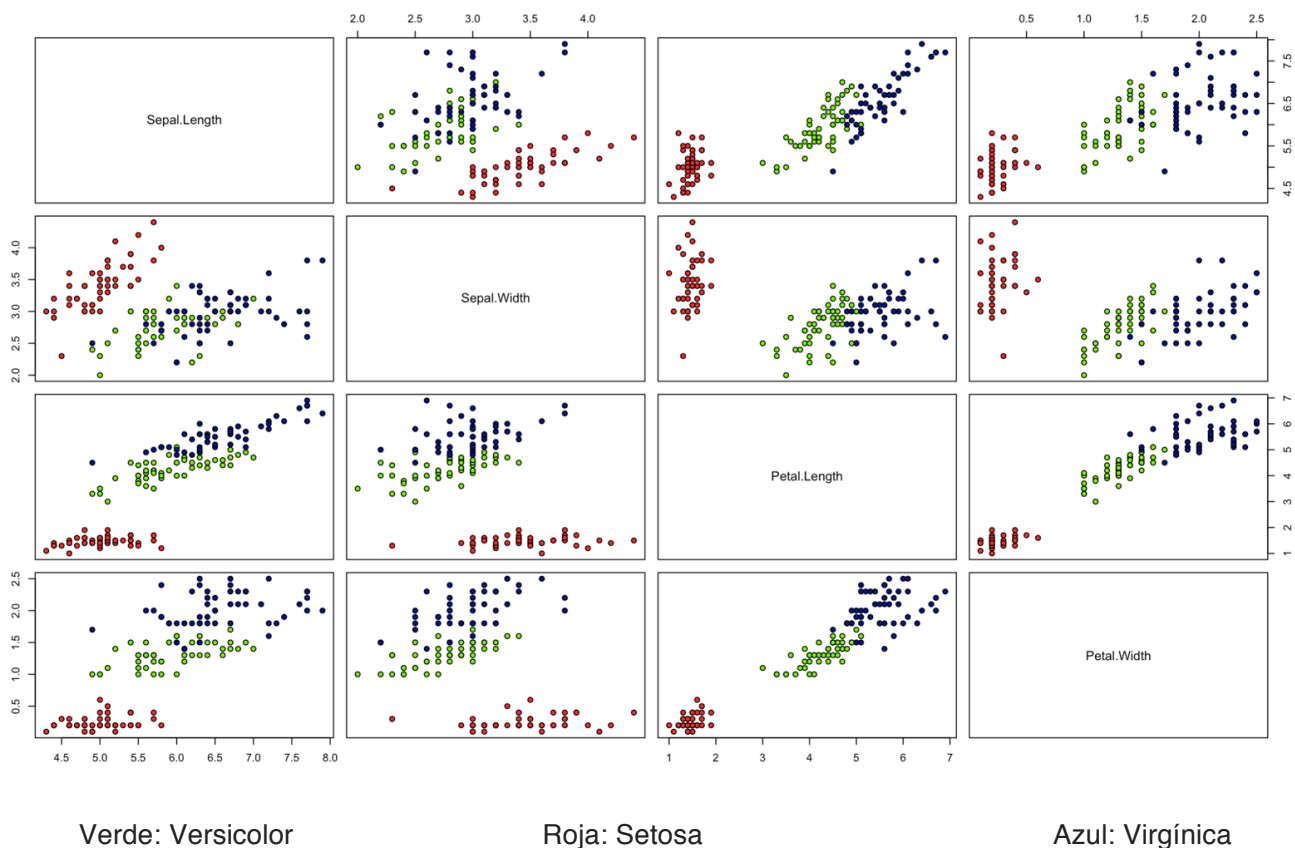
En el gráfico 2, observamos un Boxplots, de las variables del modelo. Podemos demostrar que hay variables con mayor rango, esta sería el caso de la longitud del pétalo y el menor sería el ancho del Sépalo. Excepto en la longitud del sépalo hay datos atípicos, ya que en los otros tres casos la mediana no se encuentra en el centro del cuadrado.



En el gráfico 3 observamos, la relación de las tres clases entre si, en una nube de puntos. El color de cada flor, muestra la especie a la que pertenece cada clase: (Verde: Versicolor, Azul: Virgínica, Rojo: Setosa). Según podemos ver el gráfico hay dos divisiones, por un lado está la clase Setosa, que como hemos mencionado anteriormente es linealmente independiente a las otras dos. Y por el otro lado están Virgínica y Versicolor, que mezclan sus puntos en algunas variables.

Por lo que gráficamente vemos correlación entre la longitud y el ancho del Pétalo, con la longitud del Sépalo. El ancho del sépalo apenas tiene correlación con el resto de variables, porque como podemos ver gráficamente, están muy separadas entre si las especies.

Gráfico 3:



Los atributos son: longitud del sépalo (Sepal.Length), ancho del sépalo (Sepal.Width), longitud del pétalo (Petal.Length), ancho del pétalo (Petal.Width), todos ellos medidos en centímetros.

Bibliografía:

<https://felipebravom.com/teaching/explora.pdf>

https://rpubs.com/StatGirl302/Manova_Boxplots_IRIS