**Taller IRIS**

**Nombre:** Adrian Hoyos

**NRC:** 13816

**Análisis estadísticos:**

1. Visualizando los datos de la desviación estándar de cada una de las especies de las plantas, se puede observar que la especie versicolor es la que menos varia sus medidas de los sépalos referente a las otras dos especies, por otro lado, setosa referente a las otras dos especies es la que mas varia en las medidas de sus sépalos.
2. Se puede deducir según los resultados obtenidos, que le mayoría de las flores tienen una longitud del sépalo de 5 centímetros, al igual que el ancho de los sépalos donde la mayoría de las flores tienen 3 centímetros de longitud, por otra parte, la longitud de los pétalos en la mayoría de las flores corresponde a 1.4 centímetros y, la mayoría de las flores tiene un ancho de pétalo de 0.2 centímetros.
3. Según el histograma, se puede observar que 44 tipos de iris tienen una longitud de pétalo de 1.5 centímetros aproximadamente, por otro lado 1 especie de iris tiene un tamaño de 7 centímetros aproximadamente.
4. Interpretando los boxplots, se puede deducir que la especie de iris setosa tiene menos variabilidad con respecto a las otras dos especies, debido a que el rango de la caja es mucho menor, aunque se pueden ver dos datos atípicos o outlayers lo que nos indican valores nulos. La especie versicolor tiene una variabilidad de los datos menor a la especie viriginica debido a que el rango es mucho menor que esta, pero es mayor que la especie setosa.

Haciendo referencia al boxplot de viriginica, se puede también observar que la mediana es de 2 centímetros, lo cual indica que hay más especies viriginica con mayor a esta longitud, que especies virginica que estén por debajo a esta longitud.

Virginica es el boxplot mas elevado, dado que son las especies con mayor anchura de los pétalos.

También se puede interpretar que de la especie versicolor el valor mínimo es de 1 centímetro aproximadamente y el valor máximo es de 1.8 aproximadamente.

1. Para este punto, se crearon dos columnas que calculan el área del sépalo y del pétalo de cada una de las flores para facilitar el desarrollo, luego se crearon dos variables, una con el top 5 de los pétalos con mayor área y otra variable con los sépalos con mayor área. Después de haber creado el top 5 de cada área, se creó un dataframe el cual une el top 5 de los sépalos con mayor área y los pétalos con mayor área. Por último, dado que hay flores que cumplen con ambas características, es decir, tanto el pétalo como el sépalo entran dentro del top, se creo una ultima variable que quitara los datos repetidos, dejando así en total 8 flores dentro del dataframe de top 5.