DATASET – INTRODUCTION

----------------------------------------------

**Fisher’s Iris Dataset**



Iris Flowers

**Description:**

This famous (Fisher's or Anderson's) iris data set gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of 3 species of iris. The species are Iris setosa, versicolor, and virginica.

**Format**

iris is a data frame with 150 cases (rows) and 5 variables (columns) named Sepal.Length, Sepal.Width, Petal.Length, Petal.Width, and Species

**EXERCISES**

**Download Iris Dataset from kaggle and perform the following analysis**

1. Print the dataset *iris*
2. Print the structure of the dataset *iris*
3. Print the summary of all the variables of the dataset *iris*
4. How many of the variables (columns) are in the dataset *iris*
5. How many observations (rows) are in the dataset iris
6. Use *duplicated()* function to print the logical vector indicating the duplicate values present in the dataset *iris*
7. Extract duplicate elements from the dataset iris
8. Extract unique elements from the dataset *iris*
9. Print the indices of duplicate elements in the dataset iris
10. Print the indices of unique elements in the dataset iris
11. How many unique elements are in the dataset *iris*
12. How many duplicate elements are in the dataset iris
13. Print the sorted elements in the dataset iris(Ascending order)
14. find whether any missing values are in the dataset iris
15. Display how many missing values are present in each column.
16. replace all missing values with zero
17. Calculate Petal width mean ,median ,SD,Variance for the species setosa
18. Print from 10th row to 20th row of iris dataset.
19. Print Species and its corresponding Petal length and Width.
20. display records only with species "Iris-setosa"
21. Count number of times a particular species has occurred.
22. Identifying minimum and maximum Value of Sepal width.
23. Add new column to store sum of first four column values