# Project 2018 Programming and Scripting

# ADARSHA SACHAN 23/03/2108

Introduction Fisher’s Iris data set

Reference 1 <https://en.wikipedia.org/wiki/Iris_flower_data_set>

Reference 2 <https://www.techopedia.com/definition/32880/iris-flower-data-set>

Fisher's Iris data set introduced by the British statistician and biologist Ronald Fisher in his 1936.

This data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris virginica and Iris versicolor). Four features were measured from each sample

The length, width of the sepals and petals, in centimeters, Based on the combination of these four features, Fisher developed a linear discriminant model to distinguish the species from each other.

We have real data 3 species flower.

1 Irish Setosa

2 Irish Versicolor

3 Irish Virginica

In the given attribute data, we have petal length and petal width and Sepal Length and Sepal Width

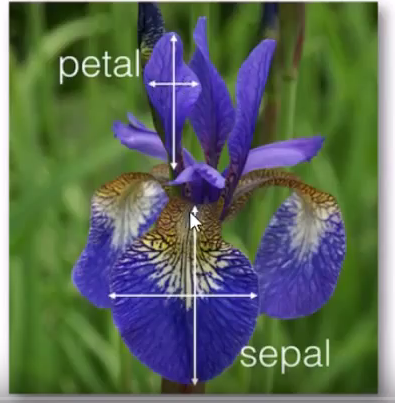
**Question** what are sepal, Petal Length and width

Reference 3 <https://www.youtube.com/watch?v=mE7MWa6N2Vc>

<https://www.ted.com/talks/linus_torvalds_the_mind_behind_linux/transcript?language=en>

Length

Width



**Research and Descriptive Analysis**

Based on my data analysis I am going to build mathematical way, while providing sepal, Petal length and width it’s will tell me the flower name.

I have written a code various way

**My first Analysis and code**

Right a code that will look for sepal, petal length and width in the csv file and return an output.

Project Name **Project.py**

While running Project.py it will ask user to type sepal, petal length and width and search the record in csv if it found then it will return else no record found will give the message.

But there is some limitation id the record did not found than???

I did my second Analysis

Project Name **myproject.py**

**While running this project it’s giving me flower name**

The logic I have written it will work only for this data set.

So my next opinion if I can represent this in Graph so I extend (**myproject.py**) project

In graph pattern

**I am writing logic in story in agile way**

**User Story 1**

As a user I want to identify what is the flower name

**I want** sepal, Petal Length and width so I can pass this info to my function.

**User story 2**

As I need a full Irish Data set.

**I want** to pass sepal, Petal Length and width in csv read search function.

Irish data set I can copy from wiki or <http://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data> copy from this link.

User story 3

As I need search logic functions.

**I want to** write a search function when user will pass input, it look in database or csv file and return the result.

If it dint match, Show the description that data not found

User story 4

As I need 4 input value petal, sepal length and width?

**I want to write** a function

While analysis the data I can see sepal length is less than 8 Cm and width is less than 5 Cm, Petal Length and width less than 7 and 3 in Cm for all the Data Set.

**So user can not provide any other input as per above petal, sepal length and width.**

To learn how this work I have gone various website

Reference Link

**Install pandas**

Reference <https://www.youtube.com/watch?v=khwJY8EeJtU>

In Cmd Type pip install pandas

<https://pandas.pydata.org/pandas-docs/stable/generated/pandas.read_table.html>

**Group series using mapper (function by a series of columns**.

<https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.groupby.html>

http://pandas.pydata.org/pandas-docs/version/0.16.2/generated/pandas.core.groupby.DataFrameGroupBy.plot.html

<https://www.youtube.com/watch?v=wfTABU8VeoY>

**For showing the graph need below library**

Import matplotlib.pyplot as Graph

<https://matplotlib.org/api/_as_gen/matplotlib.pyplot.plot.html>

<https://swcarpentry.github.io/python-novice-gapminder/09-plotting/>

http://pandas.pydata.org/pandas-docs/version/0.16.0/visualization.html

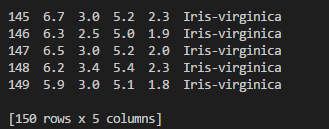
**Create graph using panda dataframe**

https://www.youtube.com/watch?v=F6kmIpWWEdU

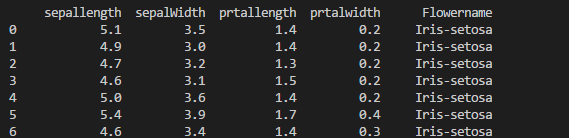
<https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.html>

**Create the graphs**

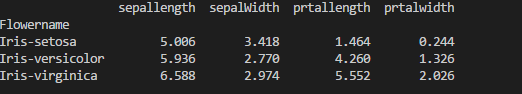
**First I am reading the csv file in Table form so while reading table I am getting below**



**But Column name was not there so I added Column after adding it looks like**

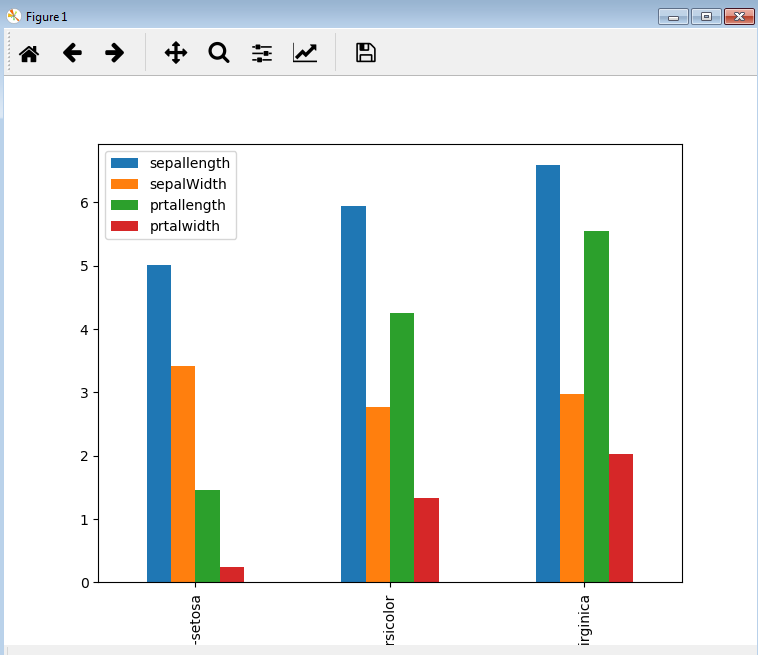


**After doing this I have created each flower mean value**



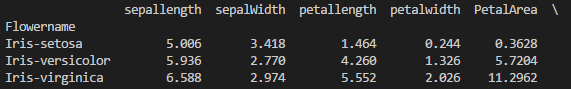
Once I got mean value than I can represent in graph

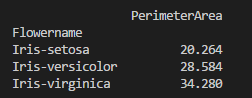
Currently I am representing in **bar Graph for each flower**



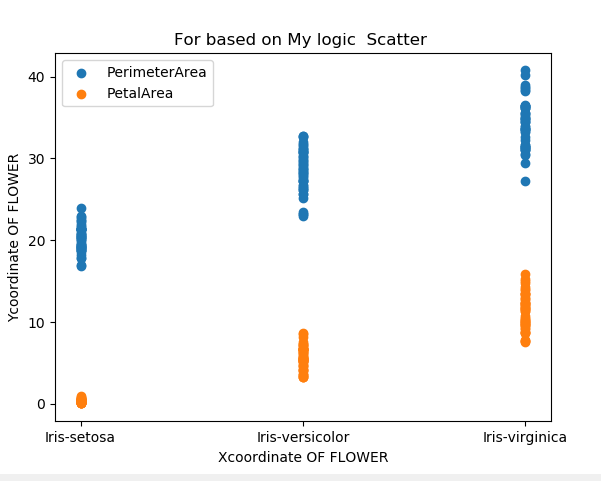
**Conclusion**

Now I am going to use Perimeter and Petal Area in the graph, so I have created two Colum.





After creating a mean and Graph I came to conclusion. Those are **Condition 1, 2, 3**



**My Second Analysis and code**

Project Name **myproject.py**

**Condition 1**

For Iris-SetosaI can see the petal area is less than 1.00

**If any time petal area is less than 2 then it’s Iris-Setos I can guess but if perimeter is less than 26 then 100 % it’s Iris-Setosa**

**Calculation algorithm** data analysis as per mine

**Petal area = Petal Length \* Petal Width**

**Perimeter = 2(Petal Length +Petal Width + Sepal Length+ Sepal Width)**

**What is perimeter???**

The **perimeter** of a shape is the distance around its edge. Finding the **perimeter** just means adding up the lengths of all its sides.

[**https://homeschool.rebeccareid.com/flower-garden-reviewing-area-perimeter/**](https://homeschool.rebeccareid.com/flower-garden-reviewing-area-perimeter/)

Below user story will be working base on my algorithm data set

User Story 5

As I need how to find **Iris-Setosa flower**

If any time **petal area** is less than 2 it’s Iris-Setosa

If perimeter is less than 25 then 100 % it’s Iris-Setosa

**Condition 1**

If **petal area** is less than 2 and perimeter is less than 25 then it is Iris-Setosa

User Story 6

As I need how to find **ris- virginica and** ris**- versicolor flower**

**Formula for Algorithm**

**Petal area = Petal Length \* Petal Width**

**Perimeter = 2(Petal Length +Petal Width + Sepal Length+ Sepal Width)**

**Condition 2**

If perimeter is less than 33 and petal area less than 9

**Versicolor flower**

**Condition 3**

If perimeter is less than 33 and petal area less than 9 and greater than 7

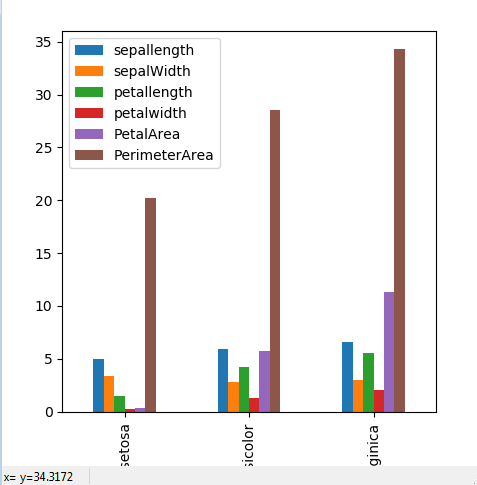
Than “80% Iris-Virginia and 20% Iris-versicolor”

**Else**

**Condition 4**

**Iris-virginica**

Bar Graph for all Colum

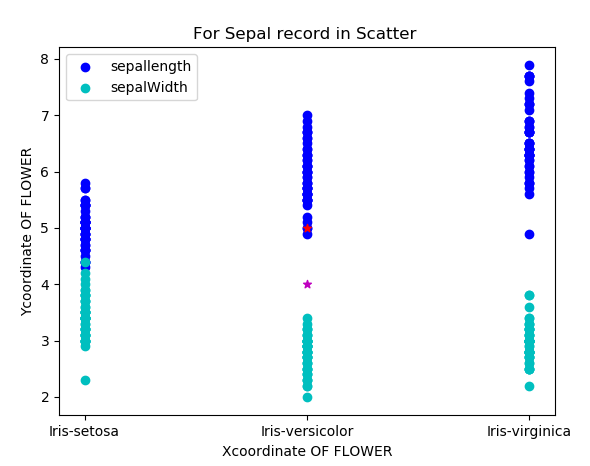


**Use Scatter graph and BAR**

For scatter graph I need X and Y axis where X is Flower Name

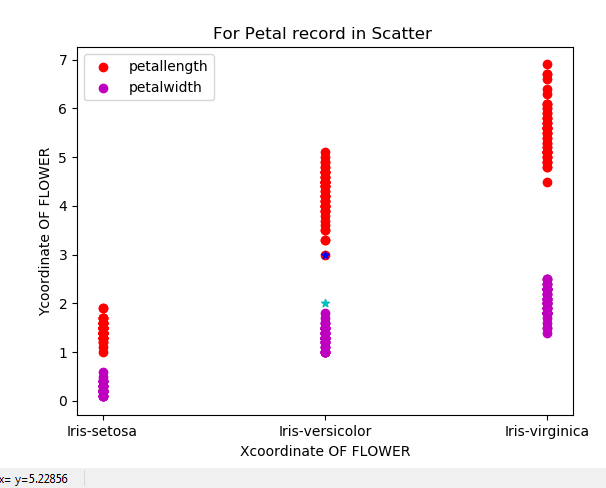
**First graph for Sepal length and Sepal Width**

Base on my logic I found flower and it represent in Star Symbol



**Second graph for Petal Length and Petal Width**

Base on my logic I found flower and it represent in Star Symbol



**Third graph for Sepal and Petal Both**

Base on my logic I found flower and it represent in O Symbol

