**Where did you get the data? And how? What are features available with the data?**

**Answer:**

* I have used in built Iris dataset.
* The following syntax is used to import dataset:

**from sklearn.datasets import load\_iris**

* Features of Iris Data:

Iris data set Description:

Iris plants dataset

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

\*\*Data Set Characteristics:\*\*

:Number of Instances: 150 (50 in each of three classes)

:Number of Attributes: 4 numeric, predictive attributes and the class

:Attribute Information:

- sepal length in cm

- sepal width in cm

- petal length in cm

- petal width in cm

- class:

- Iris-Setosa

- Iris-Versicolour

- Iris-Virginica

:Summary Statistics:

============== ==== ==== ======= ===== ====================

Min Max Mean SD Class Correlation

============== ==== ==== ======= ===== ====================

sepal length: 4.3 7.9 5.84 0.83 0.7826

sepal width: 2.0 4.4 3.05 0.43 -0.4194

petal length: 1.0 6.9 3.76 1.76 0.9490 (high!)

petal width: 0.1 2.5 1.20 0.76 0.9565 (high!)

============== ==== ==== ======= ===== ====================

:Missing Attribute Values: None

:Class Distribution: 33.3% for each of 3 classes.

:Creator: R.A. Fisher

:Donor: Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)

:Date: July, 1988

**What is the question you want to answer through analysis?**

**Answer:**

Classification of Iris Flowers sepal width, sepal length, petal width and petal length.

**What preparation of data did you do, if any? And how?**

**Answer:**

The Iris Dataset is combined with features and data label.

So we have to split the data into train data and target data.

Using following code:

X = iris.data[:, :2] # we only take the first two features.

y = iris.target

**Citations:**

1.https://www.datacamp.com/community/tutorials