In [1]:
```
# Assignment-1 Group A
#Data Wrangling, I
#Perform the following operations using Python on any open source dataset (e.g.,
#1. Import all the required Python Libraries.
#2. Locate an open source data from the web (e.g. https://www.kaggle.com). Provi
#description of the data and its source (i.e., URL of the web site).
#3. Load the Dataset into pandas data frame.
#4. Data Preprocessing: check for missing values in the data using pandas insult
#function to get some initial statistics. Provide variable descriptions. Types o
#etc. Check the dimensions of the data frame.
#5. Data Formatting and Data Normalization: Summarize the types of variables by
#the data types (i.e., character, numeric, integer, factor, and logical) of the
#data set. If variables are not in the correct data type, apply proper type conv
#6. Turn categorical variables into quantitative variables in Python.
#In addition to the codes and outputs, explain every operation that you do in th
#explain everything that you do to import/read/scrape the data set.
```

In [7]:
```
import pandas as pd
import numpy as np
```

In [8]:
```
# Section-1 Data Discovering- What is your Data best Approch to Analyze it

# Section-2 Data Structuring- Diffrent Shapes and Size

# Section-3 Data Cleaning ->#  Understanding the data,
                #Checking the missing values,
                #get information without missing values etc.

dataset=pd.read_csv("C:\Users\swamini\iris.csv")

# Read Dataset File
```

In [9]:
```
dataset
```

Out[9]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
120	121	6.9	3.2	5.7	2.3	Iris-virginica
121	122	5.6	2.8	4.9	2.0	Iris-virginica

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
122	123	7.7	2.8	6.7	2.0	Iris-virginica
123	124	6.3	2.7	4.9	1.8	Iris-virginica
124	125	6.7	3.3	5.7	2.1	Iris-virginica
125	126	7.2	3.2	6.0	1.8	Iris-virginica
126	127	6.2	2.8	4.8	1.8	Iris-virginica
127	128	6.1	3.0	4.9	1.8	Iris-virginica
128	129	6.4	2.8	5.6	2.1	Iris-virginica
129	130	7.2	3.0	5.8	1.6	Iris-virginica
130	131	7.4	2.8	6.1	1.9	Iris-virginica
131	132	7.9	3.8	6.4	2.0	Iris-virginica
132	133	6.4	2.8	5.6	2.2	Iris-virginica
133	134	6.3	2.8	5.1	1.5	Iris-virginica
134	135	6.1	2.6	5.6	1.4	Iris-virginica
135	136	7.7	3.0	6.1	2.3	Iris-virginica
136	137	6.3	3.4	5.6	2.4	Iris-virginica
137	138	6.4	3.1	5.5	1.8	Iris-virginica
138	139	6.0	3.0	4.8	1.8	Iris-virginica
139	140	6.9	3.1	5.4	2.1	Iris-virginica
140	141	6.7	3.1	5.6	2.4	Iris-virginica
141	142	6.9	3.1	5.1	2.3	Iris-virginica
142	143	5.8	2.7	5.1	1.9	Iris-virginica
143	144	6.8	3.2	5.9	2.3	Iris-virginica
144	145	6.7	3.3	5.7	2.5	Iris-virginica
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [5]:
```
# read the first five rows

dataset.head()
```

Out[5]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

In [28]:
```
# read the last five rows

dataset.tail()
```

Out[28]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

In [12]:
```
# Read the Index

dataset.index
```

Out[12]:
```
RangeIndex(start=0, stop=150, step=1)
```

In [30]:
```
# Read Columns name

dataset.columns
```

Out[30]:
```
Index([u'Id', u'SepalLengthCm', u'SepalWidthCm', u'PetalLengthCm',
       u'PetalWidthCm', u'Species'],
      dtype='object')
```

In [74]:
```
# check out the dimension of the dataset

dataset.shape
```

Out[74]:
```
(150, 6)
```

In [75]:
```
# look at the data types for each column

dataset.dtypes
```

Out[75]:
```
Id                int64
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
Species           object
dtype: object
```

In [76]:
```
# return an array of column names

dataset.columns.values
```

Out[76]:
```
array(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm',
       'PetalWidthCm', 'Species'], dtype=object)
```

In [77]:
```
 # return a list of column names

dataset.columns.values.tolist()
```

Out[77]:
```
['Id',
 'SepalLengthCm',
 'SepalWidthCm',
 'PetalLengthCm',
 'PetalWidthCm',
 'Species']
```

In [13]:
```
# Read Data Column wise

dataset['Id']
```

Out[13]:
```
0        1
1        2
2        3
3        4
4        5
5        6
6        7
7        8
8        9
9       10
10      11
11      12
12      13
13      14
14      15
15      16
16      17
17      18
18      19
```

In [14]:
```
# Read Data Column wise

dataset['SepalLengthCm']
```

Out[14]:
```
0      5.1
1      4.9
2      4.7
3      4.6
4      5.0
5      5.4
6      4.6
7      5.0
8      4.4
9      4.9
10     5.4
11     4.8
12     4.8
13     4.3
14     5.8
15     5.7
16     5.4
17     5.1
18     5.7
19     5.1
20     5.4
21     5.1
22     4.6
23     5.1
24     4.8
25     5.0
26     5.0
27     5.2
28     5.2
29     4.7
      ...
120    6.9
121    5.6
122    7.7
123    6.3
124    6.7
125    7.2
126    6.2
127    6.1
128    6.4
129    7.2
130    7.4
131    7.9
132    6.4
133    6.3
134    6.1
135    7.7
136    6.3
137    6.4
138    6.0
139    6.9
140    6.7
```

```
141    6.9
142    5.8
143    6.8
144    6.7
145    6.7
146    6.3
147    6.5
148    6.2
149    5.9
Name: SepalLengthCm, Length: 150, dtype: float64
```

In [15]:
```
# Read Data Column wise

dataset['SepalWidthCm']
```

Out[15]:
```
0      3.5
1      3.0
2      3.2
3      3.1
4      3.6
5      3.9
6      3.4
7      3.4
8      2.9
9      3.1
10     3.7
11     3.4
12     3.0
13     3.0
14     4.0
15     4.4
16     3.9
17     3.5
18     3.8
```

In [16]:
```
#NumPy arrays have one dtype for the entire array,
#while pandas DataFrames have one dtype per column.

#it does not include the index or column labels in the output.

dataset.to_numpy()
```

Out[16]:
```
array([[1L, 5.1, 3.5, 1.4, 0.2, 'Iris-setosa'],
       [2L, 4.9, 3.0, 1.4, 0.2, 'Iris-setosa'],
       [3L, 4.7, 3.2, 1.3, 0.2, 'Iris-setosa'],
       [4L, 4.6, 3.1, 1.5, 0.2, 'Iris-setosa'],
       [5L, 5.0, 3.6, 1.4, 0.2, 'Iris-setosa'],
       [6L, 5.4, 3.9, 1.7, 0.4, 'Iris-setosa'],
       [7L, 4.6, 3.4, 1.4, 0.3, 'Iris-setosa'],
       [8L, 5.0, 3.4, 1.5, 0.2, 'Iris-setosa'],
       [9L, 4.4, 2.9, 1.4, 0.2, 'Iris-setosa'],
       [10L, 4.9, 3.1, 1.5, 0.1, 'Iris-setosa'],
       [11L, 5.4, 3.7, 1.5, 0.2, 'Iris-setosa'],
       [12L, 4.8, 3.4, 1.6, 0.2, 'Iris-setosa'],
       [13L, 4.8, 3.0, 1.4, 0.1, 'Iris-setosa'],
       [14L, 4.3, 3.0, 1.1, 0.1, 'Iris-setosa'],
       [15L, 5.8, 4.0, 1.2, 0.2, 'Iris-setosa'],
       [16L, 5.7, 4.4, 1.5, 0.4, 'Iris-setosa'],
       [17L, 5.4, 3.9, 1.3, 0.4, 'Iris-setosa'],
       [18L, 5.1, 3.5, 1.4, 0.3, 'Iris-setosa'],
       [19L, 5.7, 3.8, 1.7, 0.3, 'Iris-setosa'],
       [20L, 5.1, 3.8, 1.5, 0.3, 'Iris-setosa'],
       [21L, 5.4, 3.4, 1.7, 0.2, 'Iris-setosa'],
       [22L, 5.1, 3.7, 1.5, 0.4, 'Iris-setosa'],
       [23L, 4.6, 3.6, 1.0, 0.2, 'Iris-setosa'],
       [24L, 5.1, 3.3, 1.7, 0.5, 'Iris-setosa'],
       [25L, 4.8, 3.4, 1.9, 0.2, 'Iris-setosa'],
       [26L, 5.0, 3.0, 1.6, 0.2, 'Iris-setosa'],
       [27L, 5.0, 3.4, 1.6, 0.4, 'Iris-setosa'],
       [28L, 5.2, 3.5, 1.5, 0.2, 'Iris-setosa'],
       [29L, 5.2, 3.4, 1.4, 0.2, 'Iris-setosa'],
       [30L, 4.7, 3.2, 1.6, 0.2, 'Iris-setosa'],
       [31L, 4.8, 3.1, 1.6, 0.2, 'Iris-setosa'],
       [32L, 5.4, 3.4, 1.5, 0.4, 'Iris-setosa'],
       [33L, 5.2, 4.1, 1.5, 0.1, 'Iris-setosa'],
       [34L, 5.5, 4.2, 1.4, 0.2, 'Iris-setosa'],
       [35L, 4.9, 3.1, 1.5, 0.1, 'Iris-setosa'],
       [36L, 5.0, 3.2, 1.2, 0.2, 'Iris-setosa'],
       [37L, 5.5, 3.5, 1.3, 0.2, 'Iris-setosa'],
       [38L, 4.9, 3.1, 1.5, 0.1, 'Iris-setosa'],
       [39L, 4.4, 3.0, 1.3, 0.2, 'Iris-setosa'],
       [40L, 5.1, 3.4, 1.5, 0.2, 'Iris-setosa'],
       [41L, 5.0, 3.5, 1.3, 0.3, 'Iris-setosa'],
       [42L, 4.5, 2.3, 1.3, 0.3, 'Iris-setosa'],
       [43L, 4.4, 3.2, 1.3, 0.2, 'Iris-setosa'],
       [44L, 5.0, 3.5, 1.6, 0.6, 'Iris-setosa'],
       [45L, 5.1, 3.8, 1.9, 0.4, 'Iris-setosa'],
       [46L, 4.8, 3.0, 1.4, 0.3, 'Iris-setosa'],
       [47L, 5.1, 3.8, 1.6, 0.2, 'Iris-setosa'],
       [48L, 4.6, 3.2, 1.4, 0.2, 'Iris-setosa'],
```

```
[49L, 5.3, 3.7, 1.5, 0.2, 'Iris-setosa'],
[50L, 5.0, 3.3, 1.4, 0.2, 'Iris-setosa'],
[51L, 7.0, 3.2, 4.7, 1.4, 'Iris-versicolor'],
[52L, 6.4, 3.2, 4.5, 1.5, 'Iris-versicolor'],
[53L, 6.9, 3.1, 4.9, 1.5, 'Iris-versicolor'],
[54L, 5.5, 2.3, 4.0, 1.3, 'Iris-versicolor'],
[55L, 6.5, 2.8, 4.6, 1.5, 'Iris-versicolor'],
[56L, 5.7, 2.8, 4.5, 1.3, 'Iris-versicolor'],
[57L, 6.3, 3.3, 4.7, 1.6, 'Iris-versicolor'],
[58L, 4.9, 2.4, 3.3, 1.0, 'Iris-versicolor'],
[59L, 6.6, 2.9, 4.6, 1.3, 'Iris-versicolor'],
[60L, 5.2, 2.7, 3.9, 1.4, 'Iris-versicolor'],
[61L, 5.0, 2.0, 3.5, 1.0, 'Iris-versicolor'],
[62L, 5.9, 3.0, 4.2, 1.5, 'Iris-versicolor'],
[63L, 6.0, 2.2, 4.0, 1.0, 'Iris-versicolor'],
[64L, 6.1, 2.9, 4.7, 1.4, 'Iris-versicolor'],
[65L, 5.6, 2.9, 3.6, 1.3, 'Iris-versicolor'],
[66L, 6.7, 3.1, 4.4, 1.4, 'Iris-versicolor'],
[67L, 5.6, 3.0, 4.5, 1.5, 'Iris-versicolor'],
[68L, 5.8, 2.7, 4.1, 1.0, 'Iris-versicolor'],
[69L, 6.2, 2.2, 4.5, 1.5, 'Iris-versicolor'],
[70L, 5.6, 2.5, 3.9, 1.1, 'Iris-versicolor'],
[71L, 5.9, 3.2, 4.8, 1.8, 'Iris-versicolor'],
[72L, 6.1, 2.8, 4.0, 1.3, 'Iris-versicolor'],
[73L, 6.3, 2.5, 4.9, 1.5, 'Iris-versicolor'],
[74L, 6.1, 2.8, 4.7, 1.2, 'Iris-versicolor'],
[75L, 6.4, 2.9, 4.3, 1.3, 'Iris-versicolor'],
[76L, 6.6, 3.0, 4.4, 1.4, 'Iris-versicolor'],
[77L, 6.8, 2.8, 4.8, 1.4, 'Iris-versicolor'],
[78L, 6.7, 3.0, 5.0, 1.7, 'Iris-versicolor'],
[79L, 6.0, 2.9, 4.5, 1.5, 'Iris-versicolor'],
[80L, 5.7, 2.6, 3.5, 1.0, 'Iris-versicolor'],
[81L, 5.5, 2.4, 3.8, 1.1, 'Iris-versicolor'],
[82L, 5.5, 2.4, 3.7, 1.0, 'Iris-versicolor'],
[83L, 5.8, 2.7, 3.9, 1.2, 'Iris-versicolor'],
[84L, 6.0, 2.7, 5.1, 1.6, 'Iris-versicolor'],
[85L, 5.4, 3.0, 4.5, 1.5, 'Iris-versicolor'],
[86L, 6.0, 3.4, 4.5, 1.6, 'Iris-versicolor'],
[87L, 6.7, 3.1, 4.7, 1.5, 'Iris-versicolor'],
[88L, 6.3, 2.3, 4.4, 1.3, 'Iris-versicolor'],
[89L, 5.6, 3.0, 4.1, 1.3, 'Iris-versicolor'],
[90L, 5.5, 2.5, 4.0, 1.3, 'Iris-versicolor'],
[91L, 5.5, 2.6, 4.4, 1.2, 'Iris-versicolor'],
[92L, 6.1, 3.0, 4.6, 1.4, 'Iris-versicolor'],
[93L, 5.8, 2.6, 4.0, 1.2, 'Iris-versicolor'],
[94L, 5.0, 2.3, 3.3, 1.0, 'Iris-versicolor'],
[95L, 5.6, 2.7, 4.2, 1.3, 'Iris-versicolor'],
[96L, 5.7, 3.0, 4.2, 1.2, 'Iris-versicolor'],
[97L, 5.7, 2.9, 4.2, 1.3, 'Iris-versicolor'],
[98L, 6.2, 2.9, 4.3, 1.3, 'Iris-versicolor'],
[99L, 5.1, 2.5, 3.0, 1.1, 'Iris-versicolor'],
[100L, 5.7, 2.8, 4.1, 1.3, 'Iris-versicolor'],
[101L, 6.3, 3.3, 6.0, 2.5, 'Iris-virginica'],
[102L, 5.8, 2.7, 5.1, 1.9, 'Iris-virginica'],
[103L, 7.1, 3.0, 5.9, 2.1, 'Iris-virginica'],
[104L, 6.3, 2.9, 5.6, 1.8, 'Iris-virginica'],
[105L, 6.5, 3.0, 5.8, 2.2, 'Iris-virginica'],
```

```
[106L, 7.6, 3.0, 6.6, 2.1, 'Iris-virginica'],
[107L, 4.9, 2.5, 4.5, 1.7, 'Iris-virginica'],
[108L, 7.3, 2.9, 6.3, 1.8, 'Iris-virginica'],
[109L, 6.7, 2.5, 5.8, 1.8, 'Iris-virginica'],
[110L, 7.2, 3.6, 6.1, 2.5, 'Iris-virginica'],
[111L, 6.5, 3.2, 5.1, 2.0, 'Iris-virginica'],
[112L, 6.4, 2.7, 5.3, 1.9, 'Iris-virginica'],
[113L, 6.8, 3.0, 5.5, 2.1, 'Iris-virginica'],
[114L, 5.7, 2.5, 5.0, 2.0, 'Iris-virginica'],
[115L, 5.8, 2.8, 5.1, 2.4, 'Iris-virginica'],
[116L, 6.4, 3.2, 5.3, 2.3, 'Iris-virginica'],
[117L, 6.5, 3.0, 5.5, 1.8, 'Iris-virginica'],
[118L, 7.7, 3.8, 6.7, 2.2, 'Iris-virginica'],
[119L, 7.7, 2.6, 6.9, 2.3, 'Iris-virginica'],
[120L, 6.0, 2.2, 5.0, 1.5, 'Iris-virginica'],
[121L, 6.9, 3.2, 5.7, 2.3, 'Iris-virginica'],
[122L, 5.6, 2.8, 4.9, 2.0, 'Iris-virginica'],
[123L, 7.7, 2.8, 6.7, 2.0, 'Iris-virginica'],
[124L, 6.3, 2.7, 4.9, 1.8, 'Iris-virginica'],
[125L, 6.7, 3.3, 5.7, 2.1, 'Iris-virginica'],
[126L, 7.2, 3.2, 6.0, 1.8, 'Iris-virginica'],
[127L, 6.2, 2.8, 4.8, 1.8, 'Iris-virginica'],
[128L, 6.1, 3.0, 4.9, 1.8, 'Iris-virginica'],
[129L, 6.4, 2.8, 5.6, 2.1, 'Iris-virginica'],
[130L, 7.2, 3.0, 5.8, 1.6, 'Iris-virginica'],
[131L, 7.4, 2.8, 6.1, 1.9, 'Iris-virginica'],
[132L, 7.9, 3.8, 6.4, 2.0, 'Iris-virginica'],
[133L, 6.4, 2.8, 5.6, 2.2, 'Iris-virginica'],
[134L, 6.3, 2.8, 5.1, 1.5, 'Iris-virginica'],
[135L, 6.1, 2.6, 5.6, 1.4, 'Iris-virginica'],
[136L, 7.7, 3.0, 6.1, 2.3, 'Iris-virginica'],
[137L, 6.3, 3.4, 5.6, 2.4, 'Iris-virginica'],
[138L, 6.4, 3.1, 5.5, 1.8, 'Iris-virginica'],
[139L, 6.0, 3.0, 4.8, 1.8, 'Iris-virginica'],
[140L, 6.9, 3.1, 5.4, 2.1, 'Iris-virginica'],
[141L, 6.7, 3.1, 5.6, 2.4, 'Iris-virginica'],
[142L, 6.9, 3.1, 5.1, 2.3, 'Iris-virginica'],
[143L, 5.8, 2.7, 5.1, 1.9, 'Iris-virginica'],
[144L, 6.8, 3.2, 5.9, 2.3, 'Iris-virginica'],
[145L, 6.7, 3.3, 5.7, 2.5, 'Iris-virginica'],
[146L, 6.7, 3.0, 5.2, 2.3, 'Iris-virginica'],
[147L, 6.3, 2.5, 5.0, 1.9, 'Iris-virginica'],
[148L, 6.5, 3.0, 5.2, 2.0, 'Iris-virginica'],
[149L, 6.2, 3.4, 5.4, 2.3, 'Iris-virginica'],
[150L, 5.9, 3.0, 5.1, 1.8, 'Iris-virginica']], dtype=object)
```

In [17]:
```
#Transposing your data:

dataset.T
```

Out[17]:

	0	1	2	3	4	5	6	7	8	9	...
Id	1	2	3	4	5	6	7	8	9	10	...
SepalLengthCm	5.1	4.9	4.7	4.6	5	5.4	4.6	5	4.4	4.9	...
SepalWidthCm	3.5	3	3.2	3.1	3.6	3.9	3.4	3.4	2.9	3.1	...
PetalLengthCm	1.4	1.4	1.3	1.5	1.4	1.7	1.4	1.5	1.4	1.5	...
PetalWidthCm	0.2	0.2	0.2	0.2	0.2	0.4	0.3	0.2	0.2	0.1	...
Species	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	Iris-setosa	...

6 rows × 150 columns

In [18]:
```
# sorting an acess

dataset.sort_index(axis=1, ascending=False)
```

Out[18]:

	Species	SepalWidthCm	SepalLengthCm	PetalWidthCm	PetalLengthCm	Id
0	Iris-setosa	3.5	5.1	0.2	1.4	1
1	Iris-setosa	3.0	4.9	0.2	1.4	2
2	Iris-setosa	3.2	4.7	0.2	1.3	3
3	Iris-setosa	3.1	4.6	0.2	1.5	4
4	Iris-setosa	3.6	5.0	0.2	1.4	5
5	Iris-setosa	3.9	5.4	0.4	1.7	6
6	Iris-setosa	3.4	4.6	0.3	1.4	7
7	Iris-setosa	3.4	5.0	0.2	1.5	8
8	Iris-setosa	2.9	4.4	0.2	1.4	9
9	Iris-setosa	3.1	4.9	0.1	1.5	10
10	Iris-setosa	3.7	5.4	0.2	1.5	11
11	Iris-setosa	3.4	4.8	0.2	1.6	12
12	Iris-setosa	3.0	4.8	0.1	1.4	13
13	Iris-setosa	3.0	4.3	0.1	1.1	14
14	Iris-setosa	4.0	5.8	0.2	1.2	15
15	Iris-setosa	4.4	5.7	0.4	1.5	16
16	Iris-setosa	3.9	5.4	0.4	1.3	17
17	Iris-setosa	3.5	5.1	0.3	1.4	18
18	Iris-setosa	3.8	5.7	0.3	1.7	19
19	Iris-setosa	3.8	5.1	0.3	1.5	20
20	Iris-setosa	3.4	5.4	0.2	1.7	21
21	Iris-setosa	3.7	5.1	0.4	1.5	22
22	Iris-setosa	3.6	4.6	0.2	1.0	23
23	Iris-setosa	3.3	5.1	0.5	1.7	24
24	Iris-setosa	3.4	4.8	0.2	1.9	25
25	Iris-setosa	3.0	5.0	0.2	1.6	26
26	Iris-setosa	3.4	5.0	0.4	1.6	27
27	Iris-setosa	3.5	5.2	0.2	1.5	28
28	Iris-setosa	3.4	5.2	0.2	1.4	29
29	Iris-setosa	3.2	4.7	0.2	1.6	30
...	...	...	...	...	...	...
120	Iris-virginica	3.2	6.9	2.3	5.7	121
121	Iris-virginica	2.8	5.6	2.0	4.9	122

	Species	SepalWidthCm	SepalLengthCm	PetalWidthCm	PetalLengthCm	Id
122	Iris-virginica	2.8	7.7	2.0	6.7	123
123	Iris-virginica	2.7	6.3	1.8	4.9	124
124	Iris-virginica	3.3	6.7	2.1	5.7	125
125	Iris-virginica	3.2	7.2	1.8	6.0	126
126	Iris-virginica	2.8	6.2	1.8	4.8	127
127	Iris-virginica	3.0	6.1	1.8	4.9	128
128	Iris-virginica	2.8	6.4	2.1	5.6	129
129	Iris-virginica	3.0	7.2	1.6	5.8	130
130	Iris-virginica	2.8	7.4	1.9	6.1	131
131	Iris-virginica	3.8	7.9	2.0	6.4	132
132	Iris-virginica	2.8	6.4	2.2	5.6	133
133	Iris-virginica	2.8	6.3	1.5	5.1	134
134	Iris-virginica	2.6	6.1	1.4	5.6	135
135	Iris-virginica	3.0	7.7	2.3	6.1	136
136	Iris-virginica	3.4	6.3	2.4	5.6	137
137	Iris-virginica	3.1	6.4	1.8	5.5	138
138	Iris-virginica	3.0	6.0	1.8	4.8	139
139	Iris-virginica	3.1	6.9	2.1	5.4	140
140	Iris-virginica	3.1	6.7	2.4	5.6	141
141	Iris-virginica	3.1	6.9	2.3	5.1	142
142	Iris-virginica	2.7	5.8	1.9	5.1	143
143	Iris-virginica	3.2	6.8	2.3	5.9	144
144	Iris-virginica	3.3	6.7	2.5	5.7	145
145	Iris-virginica	3.0	6.7	2.3	5.2	146
146	Iris-virginica	2.5	6.3	1.9	5.0	147
147	Iris-virginica	3.0	6.5	2.0	5.2	148
148	Iris-virginica	3.4	6.2	2.3	5.4	149
149	Iris-virginica	3.0	5.9	1.8	5.1	150

150 rows × 6 columns

In [35]:
```
# Sorting by values:

dataset.sort_values(by="SepalLengthCm")
```

Out[35]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
13	14	4.3	3.0	1.1	0.1	Iris-setosa
42	43	4.4	3.2	1.3	0.2	Iris-setosa
38	39	4.4	3.0	1.3	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
41	42	4.5	2.3	1.3	0.3	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
47	48	4.6	3.2	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
45	46	4.8	3.0	1.4	0.3	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
30	31	4.8	3.1	1.6	0.2	Iris-setosa
57	58	4.9	2.4	3.3	1.0	Iris-versicolor
106	107	4.9	2.5	4.5	1.7	Iris-virginica
34	35	4.9	3.1	1.5	0.1	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
37	38	4.9	3.1	1.5	0.1	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
40	41	5.0	3.5	1.3	0.3	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
49	50	5.0	3.3	1.4	0.2	Iris-setosa
35	36	5.0	3.2	1.2	0.2	Iris-setosa
43	44	5.0	3.5	1.6	0.6	Iris-setosa
60	61	5.0	2.0	3.5	1.0	Iris-versicolor
7	8	5.0	3.4	1.5	0.2	Iris-setosa
93	94	5.0	2.3	3.3	1.0	Iris-versicolor
...	...	...	...	...	...	...
58	59	6.6	2.9	4.6	1.3	Iris-versicolor
75	76	6.6	3.0	4.4	1.4	Iris-versicolor

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
86	87	6.7	3.1	4.7	1.5	Iris-versicolor
140	141	6.7	3.1	5.6	2.4	Iris-virginica
108	109	6.7	2.5	5.8	1.8	Iris-virginica
77	78	6.7	3.0	5.0	1.7	Iris-versicolor
65	66	6.7	3.1	4.4	1.4	Iris-versicolor
144	145	6.7	3.3	5.7	2.5	Iris-virginica
145	146	6.7	3.0	5.2	2.3	Iris-virginica
124	125	6.7	3.3	5.7	2.1	Iris-virginica
143	144	6.8	3.2	5.9	2.3	Iris-virginica
76	77	6.8	2.8	4.8	1.4	Iris-versicolor
112	113	6.8	3.0	5.5	2.1	Iris-virginica
139	140	6.9	3.1	5.4	2.1	Iris-virginica
141	142	6.9	3.1	5.1	2.3	Iris-virginica
52	53	6.9	3.1	4.9	1.5	Iris-versicolor
120	121	6.9	3.2	5.7	2.3	Iris-virginica
50	51	7.0	3.2	4.7	1.4	Iris-versicolor
102	103	7.1	3.0	5.9	2.1	Iris-virginica
129	130	7.2	3.0	5.8	1.6	Iris-virginica
125	126	7.2	3.2	6.0	1.8	Iris-virginica
109	110	7.2	3.6	6.1	2.5	Iris-virginica
107	108	7.3	2.9	6.3	1.8	Iris-virginica
130	131	7.4	2.8	6.1	1.9	Iris-virginica
105	106	7.6	3.0	6.6	2.1	Iris-virginica
122	123	7.7	2.8	6.7	2.0	Iris-virginica
118	119	7.7	2.6	6.9	2.3	Iris-virginica
117	118	7.7	3.8	6.7	2.2	Iris-virginica
135	136	7.7	3.0	6.1	2.3	Iris-virginica
131	132	7.9	3.8	6.4	2.0	Iris-virginica

150 rows × 6 columns

In [80]:
```
# only want to know if there are any missing values

dataset.isnull().values.any()
```

Out[80]:
```
False
```

In [19]:
```
dataset.isnull()
```

Out[19]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
5	False	False	False	False	False	False
6	False	False	False	False	False	False
7	False	False	False	False	False	False
8	False	False	False	False	False	False
9	False	False	False	False	False	False
10	False	False	False	False	False	False
11	False	False	False	False	False	False
12	False	False	False	False	False	False
13	False	False	False	False	False	False
14	False	False	False	False	False	False
15	False	False	False	False	False	False
16	False	False	False	False	False	False
17	False	False	False	False	False	False
18	False	False	False	False	False	False
19	False	False	False	False	False	False
20	False	False	False	False	False	False
21	False	False	False	False	False	False
22	False	False	False	False	False	False
23	False	False	False	False	False	False
24	False	False	False	False	False	False
25	False	False	False	False	False	False
26	False	False	False	False	False	False
27	False	False	False	False	False	False
28	False	False	False	False	False	False
29	False	False	False	False	False	False
...	...	...	...	...	...	...
120	False	False	False	False	False	False
121	False	False	False	False	False	False
122	False	False	False	False	False	False

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
123	False	False	False	False	False	False
124	False	False	False	False	False	False
125	False	False	False	False	False	False
126	False	False	False	False	False	False
127	False	False	False	False	False	False
128	False	False	False	False	False	False
129	False	False	False	False	False	False
130	False	False	False	False	False	False
131	False	False	False	False	False	False
132	False	False	False	False	False	False
133	False	False	False	False	False	False
134	False	False	False	False	False	False
135	False	False	False	False	False	False
136	False	False	False	False	False	False
137	False	False	False	False	False	False
138	False	False	False	False	False	False
139	False	False	False	False	False	False
140	False	False	False	False	False	False
141	False	False	False	False	False	False
142	False	False	False	False	False	False
143	False	False	False	False	False	False
144	False	False	False	False	False	False
145	False	False	False	False	False	False
146	False	False	False	False	False	False
147	False	False	False	False	False	False
148	False	False	False	False	False	False
149	False	False	False	False	False	False

150 rows × 6 columns

In [81]:
```
#  knowling number of non-missing values for each variable

dataset.notnull().sum()
```

Out[81]:
```
Id               150
SepalLengthCm    150
SepalWidthCm     150
PetalLengthCm    150
PetalWidthCm     150
Species          150
dtype: int64
```

In [82]:
```
# knowing how many missing values in the data

dataset.isnull().sum().sum()
```

Out[82]:
```
0
```

In [12]:
```
dataset.iloc[5]

# Accesing Row using i location
```

Out[12]:
```
Id                           6
SepalLengthCm              5.4
SepalWidthCm               3.9
PetalLengthCm              1.7
PetalWidthCm               0.4
Species          Iris-setosa
Name: 5, dtype: object
```

In [37]:
```
#Selecting via [], which slices the rows.

dataset[0:3]
```

Out[37]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa

In [43]:
```
# Selection by label

dataset.loc[:, ["SepalLengthCm", "SepalWidthCm"]]
```

Out[43]:

	SepalLengthCm	SepalWidthCm
0	5.1	3.5
1	4.9	3.0
2	4.7	3.2
3	4.6	3.1
4	5.0	3.6
5	5.4	3.9
6	4.6	3.4
7	5.0	3.4
8	4.4	2.9
9	4.9	3.1
10	5.4	3.7
11	4.8	3.4
12	4.8	3.0
13	4.3	3.0
14	5.8	4.0
15	5.7	4.4
16	5.4	3.9
17	5.1	3.5
18	5.7	3.8
19	5.1	3.8
20	5.4	3.4
21	5.1	3.7
22	4.6	3.6
23	5.1	3.3
24	4.8	3.4
25	5.0	3.0
26	5.0	3.4
27	5.2	3.5
28	5.2	3.4
29	4.7	3.2
...	...	...
120	6.9	3.2
121	5.6	2.8

	SepalLengthCm	SepalWidthCm
122	7.7	2.8
123	6.3	2.7
124	6.7	3.3
125	7.2	3.2
126	6.2	2.8
127	6.1	3.0
128	6.4	2.8
129	7.2	3.0
130	7.4	2.8
131	7.9	3.8
132	6.4	2.8
133	6.3	2.8
134	6.1	2.6
135	7.7	3.0
136	6.3	3.4
137	6.4	3.1
138	6.0	3.0
139	6.9	3.1
140	6.7	3.1
141	6.9	3.1
142	5.8	2.7
143	6.8	3.2
144	6.7	3.3
145	6.7	3.0
146	6.3	2.5
147	6.5	3.0
148	6.2	3.4
149	5.9	3.0

150 rows × 2 columns

In [83]:
```
# a subset of the first 100 rows of the original data

dataset.iloc[:100, :]
```

Out[83]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
70	71	5.9	3.2	4.8	1.8	Iris-versicolor
71	72	6.1	2.8	4.0	1.3	Iris-versicolor

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
72	73	6.3	2.5	4.9	1.5	Iris-versicolor
73	74	6.1	2.8	4.7	1.2	Iris-versicolor
74	75	6.4	2.9	4.3	1.3	Iris-versicolor
75	76	6.6	3.0	4.4	1.4	Iris-versicolor
76	77	6.8	2.8	4.8	1.4	Iris-versicolor
77	78	6.7	3.0	5.0	1.7	Iris-versicolor
78	79	6.0	2.9	4.5	1.5	Iris-versicolor
79	80	5.7	2.6	3.5	1.0	Iris-versicolor
80	81	5.5	2.4	3.8	1.1	Iris-versicolor
81	82	5.5	2.4	3.7	1.0	Iris-versicolor
82	83	5.8	2.7	3.9	1.2	Iris-versicolor
83	84	6.0	2.7	5.1	1.6	Iris-versicolor
84	85	5.4	3.0	4.5	1.5	Iris-versicolor
85	86	6.0	3.4	4.5	1.6	Iris-versicolor
86	87	6.7	3.1	4.7	1.5	Iris-versicolor
87	88	6.3	2.3	4.4	1.3	Iris-versicolor
88	89	5.6	3.0	4.1	1.3	Iris-versicolor
89	90	5.5	2.5	4.0	1.3	Iris-versicolor
90	91	5.5	2.6	4.4	1.2	Iris-versicolor
91	92	6.1	3.0	4.6	1.4	Iris-versicolor
92	93	5.8	2.6	4.0	1.2	Iris-versicolor
93	94	5.0	2.3	3.3	1.0	Iris-versicolor
94	95	5.6	2.7	4.2	1.3	Iris-versicolor
95	96	5.7	3.0	4.2	1.2	Iris-versicolor
96	97	5.7	2.9	4.2	1.3	Iris-versicolor
97	98	6.2	2.9	4.3	1.3	Iris-versicolor
98	99	5.1	2.5	3.0	1.1	Iris-versicolor
99	100	5.7	2.8	4.1	1.3	Iris-versicolor

100 rows × 6 columns

In [45]:
```
#Acessing data by position

dataset.iloc[3]
```

Out[45]:
```
Id                        4
SepalLengthCm           4.6
SepalWidthCm            3.1
PetalLengthCm           1.5
PetalWidthCm            0.2
Species         Iris-setosa
Name: 3, dtype: object
```

In [84]:
```
# a subset of the first 3 columns of the original data

dataset.iloc[:, :3]
```

Out[84]:

	Id	SepalLengthCm	SepalWidthCm
0	1	5.1	3.5
1	2	4.9	3.0
2	3	4.7	3.2
3	4	4.6	3.1
4	5	5.0	3.6
5	6	5.4	3.9
6	7	4.6	3.4
7	8	5.0	3.4
8	9	4.4	2.9
9	10	4.9	3.1
10	11	5.4	3.7
11	12	4.8	3.4
12	13	4.8	3.0
13	14	4.3	3.0
14	15	5.8	4.0
15	16	5.7	4.4
16	17	5.4	3.9
17	18	5.1	3.5
18	19	5.7	3.8
19	20	5.1	3.8
20	21	5.4	3.4
21	22	5.1	3.7
22	23	4.6	3.6
23	24	5.1	3.3
24	25	4.8	3.4
25	26	5.0	3.0
26	27	5.0	3.4
27	28	5.2	3.5
28	29	5.2	3.4
29	30	4.7	3.2
...	...	...	...
120	121	6.9	3.2
121	122	5.6	2.8

	Id	SepalLengthCm	SepalWidthCm
122	123	7.7	2.8
123	124	6.3	2.7
124	125	6.7	3.3
125	126	7.2	3.2
126	127	6.2	2.8
127	128	6.1	3.0
128	129	6.4	2.8
129	130	7.2	3.0
130	131	7.4	2.8
131	132	7.9	3.8
132	133	6.4	2.8
133	134	6.3	2.8
134	135	6.1	2.6
135	136	7.7	3.0
136	137	6.3	3.4
137	138	6.4	3.1
138	139	6.0	3.0
139	140	6.9	3.1
140	141	6.7	3.1
141	142	6.9	3.1
142	143	5.8	2.7
143	144	6.8	3.2
144	145	6.7	3.3
145	146	6.7	3.0
146	147	6.3	2.5
147	148	6.5	3.0
148	149	6.2	3.4
149	150	5.9	3.0

150 rows × 3 columns

In [85]:
```
# a subset of the first 100 rows and the first 15 columns

dataset.iloc[:100, :15]
```

Out[85]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
70	71	5.9	3.2	4.8	1.8	Iris-versicolor
71	72	6.1	2.8	4.0	1.3	Iris-versicolor

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
72	73	6.3	2.5	4.9	1.5	Iris-versicolor
73	74	6.1	2.8	4.7	1.2	Iris-versicolor
74	75	6.4	2.9	4.3	1.3	Iris-versicolor
75	76	6.6	3.0	4.4	1.4	Iris-versicolor
76	77	6.8	2.8	4.8	1.4	Iris-versicolor
77	78	6.7	3.0	5.0	1.7	Iris-versicolor
78	79	6.0	2.9	4.5	1.5	Iris-versicolor
79	80	5.7	2.6	3.5	1.0	Iris-versicolor
80	81	5.5	2.4	3.8	1.1	Iris-versicolor
81	82	5.5	2.4	3.7	1.0	Iris-versicolor
82	83	5.8	2.7	3.9	1.2	Iris-versicolor
83	84	6.0	2.7	5.1	1.6	Iris-versicolor
84	85	5.4	3.0	4.5	1.5	Iris-versicolor
85	86	6.0	3.4	4.5	1.6	Iris-versicolor
86	87	6.7	3.1	4.7	1.5	Iris-versicolor
87	88	6.3	2.3	4.4	1.3	Iris-versicolor
88	89	5.6	3.0	4.1	1.3	Iris-versicolor
89	90	5.5	2.5	4.0	1.3	Iris-versicolor
90	91	5.5	2.6	4.4	1.2	Iris-versicolor
91	92	6.1	3.0	4.6	1.4	Iris-versicolor
92	93	5.8	2.6	4.0	1.2	Iris-versicolor
93	94	5.0	2.3	3.3	1.0	Iris-versicolor
94	95	5.6	2.7	4.2	1.3	Iris-versicolor
95	96	5.7	3.0	4.2	1.2	Iris-versicolor
96	97	5.7	2.9	4.2	1.3	Iris-versicolor
97	98	6.2	2.9	4.3	1.3	Iris-versicolor
98	99	5.1	2.5	3.0	1.1	Iris-versicolor
99	100	5.7	2.8	4.1	1.3	Iris-versicolor

100 rows × 6 columns

In [46]:
```
# Slicing of data

dataset.iloc[3:5, 0:2]
```

Out[46]:

	Id	SepalLengthCm
3	4	4.6
4	5	5.0

In [47]:
```
# By lists of integer position locations, similar to the NumPy/Python style:

dataset.iloc[[1, 2, 4], [0, 2]]
```

Out[47]:

	Id	SepalWidthCm
1	2	3.0
2	3	3.2
4	5	3.6

In [49]:
```
# For slicing rows explicitly:

dataset.iloc[1:3, :]
```

Out[49]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa

In [50]:
```
# For slicing column explicitly:

dataset.iloc[:, 1:3]
```

Out[50]:

	SepalLengthCm	SepalWidthCm
0	5.1	3.5
1	4.9	3.0
2	4.7	3.2
3	4.6	3.1
4	5.0	3.6
5	5.4	3.9
6	4.6	3.4
7	5.0	3.4
8	4.4	2.9
9	4.9	3.1
10	5.4	3.7
11	4.8	3.4
12	4.8	3.0
13	4.3	3.0
14	5.8	4.0
15	5.7	4.4
16	5.4	3.9
17	5.1	3.5
18	5.7	3.8
19	5.1	3.8
20	5.4	3.4
21	5.1	3.7
22	4.6	3.6
23	5.1	3.3
24	4.8	3.4
25	5.0	3.0
26	5.0	3.4
27	5.2	3.5
28	5.2	3.4
29	4.7	3.2
...	...	...
120	6.9	3.2
121	5.6	2.8

	SepalLengthCm	SepalWidthCm
122	7.7	2.8
123	6.3	2.7
124	6.7	3.3
125	7.2	3.2
126	6.2	2.8
127	6.1	3.0
128	6.4	2.8
129	7.2	3.0
130	7.4	2.8
131	7.9	3.8
132	6.4	2.8
133	6.3	2.8
134	6.1	2.6
135	7.7	3.0
136	6.3	3.4
137	6.4	3.1
138	6.0	3.0
139	6.9	3.1
140	6.7	3.1
141	6.9	3.1
142	5.8	2.7
143	6.8	3.2
144	6.7	3.3
145	6.7	3.0
146	6.3	2.5
147	6.5	3.0
148	6.2	3.4
149	5.9	3.0

150 rows × 2 columns

In [51]:
```
#For getting a value explicitly:

dataset.iloc[1, 1]
```

Out[51]: 4.9

In [14]:
```
dataset['SepalLengthCm'].iloc[5]

#Acessing Coloumn and Rows by position
```

Out[14]:
```
5.4
```

In [17]:
```
# Data Normalization-- Range Selection

dataset[['SepalLengthCm','PetalLengthCm']]
```

Out[17]:

	SepalLengthCm	PetalLengthCm
0	5.1	1.4
1	4.9	1.4
2	4.7	1.3
3	4.6	1.5
4	5.0	1.4
5	5.4	1.7
6	4.6	1.4
7	5.0	1.5
8	4.4	1.4
9	4.9	1.5
10	5.4	1.5
11	4.8	1.6
12	4.8	1.4
13	4.3	1.1
14	5.8	1.2
15	5.7	1.5
16	5.4	1.3
17	5.1	1.4
18	5.7	1.7
19	5.1	1.5
20	5.4	1.7
21	5.1	1.5
22	4.6	1.0
23	5.1	1.7
24	4.8	1.9
25	5.0	1.6
26	5.0	1.6
27	5.2	1.5
28	5.2	1.4
29	4.7	1.6
...	...	...
120	6.9	5.7

	SepalLengthCm	PetalLengthCm
121	5.6	4.9
122	7.7	6.7
123	6.3	4.9
124	6.7	5.7
125	7.2	6.0
126	6.2	4.8
127	6.1	4.9
128	6.4	5.6
129	7.2	5.8
130	7.4	6.1
131	7.9	6.4
132	6.4	5.6
133	6.3	5.1
134	6.1	5.6
135	7.7	6.1
136	6.3	5.6
137	6.4	5.5
138	6.0	4.8
139	6.9	5.4
140	6.7	5.6
141	6.9	5.1
142	5.8	5.1
143	6.8	5.9
144	6.7	5.7
145	6.7	5.2
146	6.3	5.0
147	6.5	5.2
148	6.2	5.4
149	5.9	5.1

150 rows × 2 columns

In [21]:
```
# Get Coloumn Name

cols_2_4=dataset.columns[2:4]
```

In [22]:
```
# Then get data from Columns

dataset[cols_2_4]
```

Out[22]:

	SepalWidthCm	PetalLengthCm
0	3.5	1.4
1	3.0	1.4
2	3.2	1.3
3	3.1	1.5
4	3.6	1.4
5	3.9	1.7
6	3.4	1.4
7	3.4	1.5
8	2.9	1.4
9	3.1	1.5
10	3.7	1.5
11	3.4	1.6
12	3.0	1.4
13	3.0	1.1
14	4.0	1.2
15	4.4	1.5
16	3.9	1.3
17	3.5	1.4
18	3.8	1.7
19	3.8	1.5
20	3.4	1.7
21	3.7	1.5
22	3.6	1.0
23	3.3	1.7
24	3.4	1.9
25	3.0	1.6
26	3.4	1.6
27	3.5	1.5
28	3.4	1.4
29	3.2	1.6
...	...	...
120	3.2	5.7
121	2.8	4.9

	SepalWidthCm	PetalLengthCm
122	2.8	6.7
123	2.7	4.9
124	3.3	5.7
125	3.2	6.0
126	2.8	4.8
127	3.0	4.9
128	2.8	5.6
129	3.0	5.8
130	2.8	6.1
131	3.8	6.4
132	2.8	5.6
133	2.8	5.1
134	2.6	5.6
135	3.0	6.1
136	3.4	5.6
137	3.1	5.5
138	3.0	4.8
139	3.1	5.4
140	3.1	5.6
141	3.1	5.1
142	2.7	5.1
143	3.2	5.9
144	3.3	5.7
145	3.0	5.2
146	2.5	5.0
147	3.0	5.2
148	3.4	5.4
149	3.0	5.1

150 rows × 2 columns

In [26]:
```
# same operation in 3 steps

cols_2_4=dataset.columns[2:4]

#then we get coloumn

cols_2_4=dataset[cols_2_4]

#Now select Row from Dataframe

cols_2_4.iloc[5:10]
```

Out[26]:

	SepalWidthCm	PetalLengthCm
5	3.9	1.7
6	3.4	1.4
7	3.4	1.5
8	2.9	1.4
9	3.1	1.5

In [20]:
```
# in one Expression all these 3 line

dataset[dataset.columns[2:4]].iloc[5:10]
```

Out[20]:

	SepalWidthCm	PetalLengthCm
5	3.9	1.7
6	3.4	1.4
7	3.4	1.5
8	2.9	1.4
9	3.1	1.5

In [53]:
```
# Data Preprocessing
# Preprocessing involves the following aspects:
# missing values
# data standardization
# data normalization
# data binning
# Check missing values in Data using pandas various function
# Pandas describe() is used to view some basic statistical details like percentil
# Whenthis method is applied to a series of string, it returns a different output

dataset.describe(percentiles=None, include=None, exclude=None)
```

Out[53]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

In [21]:
```
# print result

dataset
```

Out[21]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
120	121	6.9	3.2	5.7	2.3	Iris-virginica
121	122	5.6	2.8	4.9	2.0	Iris-virginica

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
122	123	7.7	2.8	6.7	2.0	Iris-virginica
123	124	6.3	2.7	4.9	1.8	Iris-virginica
124	125	6.7	3.3	5.7	2.1	Iris-virginica
125	126	7.2	3.2	6.0	1.8	Iris-virginica
126	127	6.2	2.8	4.8	1.8	Iris-virginica
127	128	6.1	3.0	4.9	1.8	Iris-virginica
128	129	6.4	2.8	5.6	2.1	Iris-virginica
129	130	7.2	3.0	5.8	1.6	Iris-virginica
130	131	7.4	2.8	6.1	1.9	Iris-virginica
131	132	7.9	3.8	6.4	2.0	Iris-virginica
132	133	6.4	2.8	5.6	2.2	Iris-virginica
133	134	6.3	2.8	5.1	1.5	Iris-virginica
134	135	6.1	2.6	5.6	1.4	Iris-virginica
135	136	7.7	3.0	6.1	2.3	Iris-virginica
136	137	6.3	3.4	5.6	2.4	Iris-virginica
137	138	6.4	3.1	5.5	1.8	Iris-virginica
138	139	6.0	3.0	4.8	1.8	Iris-virginica
139	140	6.9	3.1	5.4	2.1	Iris-virginica
140	141	6.7	3.1	5.6	2.4	Iris-virginica
141	142	6.9	3.1	5.1	2.3	Iris-virginica
142	143	5.8	2.7	5.1	1.9	Iris-virginica
143	144	6.8	3.2	5.9	2.3	Iris-virginica
144	145	6.7	3.3	5.7	2.5	Iris-virginica
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [60]:
```
# filling missing value using fillna()
# Filling null values with a single value

dataset.fillna(0)
```

Out[60]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa
20	21	5.4	3.4	1.7	0.2	Iris-setosa
21	22	5.1	3.7	1.5	0.4	Iris-setosa
22	23	4.6	3.6	1.0	0.2	Iris-setosa
23	24	5.1	3.3	1.7	0.5	Iris-setosa
24	25	4.8	3.4	1.9	0.2	Iris-setosa
25	26	5.0	3.0	1.6	0.2	Iris-setosa
26	27	5.0	3.4	1.6	0.4	Iris-setosa
27	28	5.2	3.5	1.5	0.2	Iris-setosa
28	29	5.2	3.4	1.4	0.2	Iris-setosa
29	30	4.7	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
120	121	6.9	3.2	5.7	2.3	Iris-virginica

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
121	122	5.6	2.8	4.9	2.0	Iris-virginica
122	123	7.7	2.8	6.7	2.0	Iris-virginica
123	124	6.3	2.7	4.9	1.8	Iris-virginica
124	125	6.7	3.3	5.7	2.1	Iris-virginica
125	126	7.2	3.2	6.0	1.8	Iris-virginica
126	127	6.2	2.8	4.8	1.8	Iris-virginica
127	128	6.1	3.0	4.9	1.8	Iris-virginica
128	129	6.4	2.8	5.6	2.1	Iris-virginica
129	130	7.2	3.0	5.8	1.6	Iris-virginica
130	131	7.4	2.8	6.1	1.9	Iris-virginica
131	132	7.9	3.8	6.4	2.0	Iris-virginica
132	133	6.4	2.8	5.6	2.2	Iris-virginica
133	134	6.3	2.8	5.1	1.5	Iris-virginica
134	135	6.1	2.6	5.6	1.4	Iris-virginica
135	136	7.7	3.0	6.1	2.3	Iris-virginica
136	137	6.3	3.4	5.6	2.4	Iris-virginica
137	138	6.4	3.1	5.5	1.8	Iris-virginica
138	139	6.0	3.0	4.8	1.8	Iris-virginica
139	140	6.9	3.1	5.4	2.1	Iris-virginica
140	141	6.7	3.1	5.6	2.4	Iris-virginica
141	142	6.9	3.1	5.1	2.3	Iris-virginica
142	143	5.8	2.7	5.1	1.9	Iris-virginica
143	144	6.8	3.2	5.9	2.3	Iris-virginica
144	145	6.7	3.3	5.7	2.5	Iris-virginica
145	146	6.7	3.0	5.2	2.3	Iris-virginica
146	147	6.3	2.5	5.0	1.9	Iris-virginica
147	148	6.5	3.0	5.2	2.0	Iris-virginica
148	149	6.2	3.4	5.4	2.3	Iris-virginica
149	150	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [22]:
```
dataset.mean()
```

Out[22]:
```
Id               75.500000
SepalLengthCm     5.843333
SepalWidthCm      3.054000
PetalLengthCm     3.758667
PetalWidthCm      1.198667
dtype: float64
```

In [67]:
```
#calculate median

dataset.median()
```

Out[67]:
```
Id               75.50
SepalLengthCm     5.80
SepalWidthCm      3.00
PetalLengthCm     4.35
PetalWidthCm      1.30
dtype: float64
```

In [70]:
```
# write mode values across rows and cloumn

dataset.mode(axis=0)
```

Out[70]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.0	3.0	1.5	0.2	Iris-setosa
1	2	NaN	NaN	NaN	NaN	Iris-versicolor
2	3	NaN	NaN	NaN	NaN	Iris-virginica
3	4	NaN	NaN	NaN	NaN	NaN
4	5	NaN	NaN	NaN	NaN	NaN
5	6	NaN	NaN	NaN	NaN	NaN
6	7	NaN	NaN	NaN	NaN	NaN
7	8	NaN	NaN	NaN	NaN	NaN
8	9	NaN	NaN	NaN	NaN	NaN
9	10	NaN	NaN	NaN	NaN	NaN
10	11	NaN	NaN	NaN	NaN	NaN
11	12	NaN	NaN	NaN	NaN	NaN
12	13	NaN	NaN	NaN	NaN	NaN
13	14	NaN	NaN	NaN	NaN	NaN
14	15	NaN	NaN	NaN	NaN	NaN
15	16	NaN	NaN	NaN	NaN	NaN
16	17	NaN	NaN	NaN	NaN	NaN
17	18	NaN	NaN	NaN	NaN	NaN
18	19	NaN	NaN	NaN	NaN	NaN
19	20	NaN	NaN	NaN	NaN	NaN
20	21	NaN	NaN	NaN	NaN	NaN
21	22	NaN	NaN	NaN	NaN	NaN
22	23	NaN	NaN	NaN	NaN	NaN
23	24	NaN	NaN	NaN	NaN	NaN
24	25	NaN	NaN	NaN	NaN	NaN
25	26	NaN	NaN	NaN	NaN	NaN
26	27	NaN	NaN	NaN	NaN	NaN
27	28	NaN	NaN	NaN	NaN	NaN
28	29	NaN	NaN	NaN	NaN	NaN
29	30	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...
120	121	NaN	NaN	NaN	NaN	NaN
121	122	NaN	NaN	NaN	NaN	NaN

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
122	123	NaN	NaN	NaN	NaN	NaN
123	124	NaN	NaN	NaN	NaN	NaN
124	125	NaN	NaN	NaN	NaN	NaN
125	126	NaN	NaN	NaN	NaN	NaN
126	127	NaN	NaN	NaN	NaN	NaN
127	128	NaN	NaN	NaN	NaN	NaN
128	129	NaN	NaN	NaN	NaN	NaN
129	130	NaN	NaN	NaN	NaN	NaN
130	131	NaN	NaN	NaN	NaN	NaN
131	132	NaN	NaN	NaN	NaN	NaN
132	133	NaN	NaN	NaN	NaN	NaN
133	134	NaN	NaN	NaN	NaN	NaN
134	135	NaN	NaN	NaN	NaN	NaN
135	136	NaN	NaN	NaN	NaN	NaN
136	137	NaN	NaN	NaN	NaN	NaN
137	138	NaN	NaN	NaN	NaN	NaN
138	139	NaN	NaN	NaN	NaN	NaN
139	140	NaN	NaN	NaN	NaN	NaN
140	141	NaN	NaN	NaN	NaN	NaN
141	142	NaN	NaN	NaN	NaN	NaN
142	143	NaN	NaN	NaN	NaN	NaN
143	144	NaN	NaN	NaN	NaN	NaN
144	145	NaN	NaN	NaN	NaN	NaN
145	146	NaN	NaN	NaN	NaN	NaN
146	147	NaN	NaN	NaN	NaN	NaN
147	148	NaN	NaN	NaN	NaN	NaN
148	149	NaN	NaN	NaN	NaN	NaN
149	150	NaN	NaN	NaN	NaN	NaN

150 rows × 6 columns

In [23]:
```
# Data Formating- Bring the data into standered format

dataset["SepalLengthCm"]=5/dataset["SepalLengthCm"]
```

In [88]:
```
# Print dataset

dataset
```

Out[88]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	0.980392	3.5	1.4	0.2	Iris-setosa
1	2	1.020408	3.0	1.4	0.2	Iris-setosa
2	3	1.063830	3.2	1.3	0.2	Iris-setosa
3	4	1.086957	3.1	1.5	0.2	Iris-setosa
4	5	1.000000	3.6	1.4	0.2	Iris-setosa
5	6	0.925926	3.9	1.7	0.4	Iris-setosa
6	7	1.086957	3.4	1.4	0.3	Iris-setosa
7	8	1.000000	3.4	1.5	0.2	Iris-setosa
8	9	1.136364	2.9	1.4	0.2	Iris-setosa
9	10	1.020408	3.1	1.5	0.1	Iris-setosa
10	11	0.925926	3.7	1.5	0.2	Iris-setosa
11	12	1.041667	3.4	1.6	0.2	Iris-setosa
12	13	1.041667	3.0	1.4	0.1	Iris-setosa
13	14	1.162791	3.0	1.1	0.1	Iris-setosa
14	15	0.862069	4.0	1.2	0.2	Iris-setosa
15	16	0.877193	4.4	1.5	0.4	Iris-setosa
16	17	0.925926	3.9	1.3	0.4	Iris-setosa
17	18	0.980392	3.5	1.4	0.3	Iris-setosa
18	19	0.877193	3.8	1.7	0.3	Iris-setosa
19	20	0.980392	3.8	1.5	0.3	Iris-setosa
20	21	0.925926	3.4	1.7	0.2	Iris-setosa
21	22	0.980392	3.7	1.5	0.4	Iris-setosa
22	23	1.086957	3.6	1.0	0.2	Iris-setosa
23	24	0.980392	3.3	1.7	0.5	Iris-setosa
24	25	1.041667	3.4	1.9	0.2	Iris-setosa
25	26	1.000000	3.0	1.6	0.2	Iris-setosa
26	27	1.000000	3.4	1.6	0.4	Iris-setosa
27	28	0.961538	3.5	1.5	0.2	Iris-setosa
28	29	0.961538	3.4	1.4	0.2	Iris-setosa
29	30	1.063830	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
120	121	0.724638	3.2	5.7	2.3	Iris-virginica

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
121	122	0.892857	2.8	4.9	2.0	Iris-virginica
122	123	0.649351	2.8	6.7	2.0	Iris-virginica
123	124	0.793651	2.7	4.9	1.8	Iris-virginica
124	125	0.746269	3.3	5.7	2.1	Iris-virginica
125	126	0.694444	3.2	6.0	1.8	Iris-virginica
126	127	0.806452	2.8	4.8	1.8	Iris-virginica
127	128	0.819672	3.0	4.9	1.8	Iris-virginica
128	129	0.781250	2.8	5.6	2.1	Iris-virginica
129	130	0.694444	3.0	5.8	1.6	Iris-virginica
130	131	0.675676	2.8	6.1	1.9	Iris-virginica
131	132	0.632911	3.8	6.4	2.0	Iris-virginica
132	133	0.781250	2.8	5.6	2.2	Iris-virginica
133	134	0.793651	2.8	5.1	1.5	Iris-virginica
134	135	0.819672	2.6	5.6	1.4	Iris-virginica
135	136	0.649351	3.0	6.1	2.3	Iris-virginica
136	137	0.793651	3.4	5.6	2.4	Iris-virginica
137	138	0.781250	3.1	5.5	1.8	Iris-virginica
138	139	0.833333	3.0	4.8	1.8	Iris-virginica
139	140	0.724638	3.1	5.4	2.1	Iris-virginica
140	141	0.746269	3.1	5.6	2.4	Iris-virginica
141	142	0.724638	3.1	5.1	2.3	Iris-virginica
142	143	0.862069	2.7	5.1	1.9	Iris-virginica
143	144	0.735294	3.2	5.9	2.3	Iris-virginica
144	145	0.746269	3.3	5.7	2.5	Iris-virginica
145	146	0.746269	3.0	5.2	2.3	Iris-virginica
146	147	0.793651	2.5	5.0	1.9	Iris-virginica
147	148	0.769231	3.0	5.2	2.0	Iris-virginica
148	149	0.806452	3.4	5.4	2.3	Iris-virginica
149	150	0.847458	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [89]:
```
# Rename the Column Name

dataset.rename(columns={"SepalLengthCm":"NewSepalLengthCm"},inplace=True)
```

In [90]:
```
# Print Dataset

dataset
```

Out[90]:

	Id	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	0.980392	3.5	1.4	0.2	Iris-setosa
1	2	1.020408	3.0	1.4	0.2	Iris-setosa
2	3	1.063830	3.2	1.3	0.2	Iris-setosa
3	4	1.086957	3.1	1.5	0.2	Iris-setosa
4	5	1.000000	3.6	1.4	0.2	Iris-setosa
5	6	0.925926	3.9	1.7	0.4	Iris-setosa
6	7	1.086957	3.4	1.4	0.3	Iris-setosa
7	8	1.000000	3.4	1.5	0.2	Iris-setosa
8	9	1.136364	2.9	1.4	0.2	Iris-setosa
9	10	1.020408	3.1	1.5	0.1	Iris-setosa
10	11	0.925926	3.7	1.5	0.2	Iris-setosa
11	12	1.041667	3.4	1.6	0.2	Iris-setosa
12	13	1.041667	3.0	1.4	0.1	Iris-setosa
13	14	1.162791	3.0	1.1	0.1	Iris-setosa
14	15	0.862069	4.0	1.2	0.2	Iris-setosa
15	16	0.877193	4.4	1.5	0.4	Iris-setosa
16	17	0.925926	3.9	1.3	0.4	Iris-setosa
17	18	0.980392	3.5	1.4	0.3	Iris-setosa
18	19	0.877193	3.8	1.7	0.3	Iris-setosa
19	20	0.980392	3.8	1.5	0.3	Iris-setosa
20	21	0.925926	3.4	1.7	0.2	Iris-setosa
21	22	0.980392	3.7	1.5	0.4	Iris-setosa
22	23	1.086957	3.6	1.0	0.2	Iris-setosa
23	24	0.980392	3.3	1.7	0.5	Iris-setosa
24	25	1.041667	3.4	1.9	0.2	Iris-setosa
25	26	1.000000	3.0	1.6	0.2	Iris-setosa
26	27	1.000000	3.4	1.6	0.4	Iris-setosa
27	28	0.961538	3.5	1.5	0.2	Iris-setosa
28	29	0.961538	3.4	1.4	0.2	Iris-setosa
29	30	1.063830	3.2	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
120	121	0.724638	3.2	5.7	2.3	Iris-virginica

	Id	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
121	122	0.892857	2.8	4.9	2.0	Iris-virginica
122	123	0.649351	2.8	6.7	2.0	Iris-virginica
123	124	0.793651	2.7	4.9	1.8	Iris-virginica
124	125	0.746269	3.3	5.7	2.1	Iris-virginica
125	126	0.694444	3.2	6.0	1.8	Iris-virginica
126	127	0.806452	2.8	4.8	1.8	Iris-virginica
127	128	0.819672	3.0	4.9	1.8	Iris-virginica
128	129	0.781250	2.8	5.6	2.1	Iris-virginica
129	130	0.694444	3.0	5.8	1.6	Iris-virginica
130	131	0.675676	2.8	6.1	1.9	Iris-virginica
131	132	0.632911	3.8	6.4	2.0	Iris-virginica
132	133	0.781250	2.8	5.6	2.2	Iris-virginica
133	134	0.793651	2.8	5.1	1.5	Iris-virginica
134	135	0.819672	2.6	5.6	1.4	Iris-virginica
135	136	0.649351	3.0	6.1	2.3	Iris-virginica
136	137	0.793651	3.4	5.6	2.4	Iris-virginica
137	138	0.781250	3.1	5.5	1.8	Iris-virginica
138	139	0.833333	3.0	4.8	1.8	Iris-virginica
139	140	0.724638	3.1	5.4	2.1	Iris-virginica
140	141	0.746269	3.1	5.6	2.4	Iris-virginica
141	142	0.724638	3.1	5.1	2.3	Iris-virginica
142	143	0.862069	2.7	5.1	1.9	Iris-virginica
143	144	0.735294	3.2	5.9	2.3	Iris-virginica
144	145	0.746269	3.3	5.7	2.5	Iris-virginica
145	146	0.746269	3.0	5.2	2.3	Iris-virginica
146	147	0.793651	2.5	5.0	1.9	Iris-virginica
147	148	0.769231	3.0	5.2	2.0	Iris-virginica
148	149	0.806452	3.4	5.4	2.3	Iris-virginica
149	150	0.847458	3.0	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [92]:

```
# Find the datatype - dataframe.dtypes()

# Convert the datatype - dataframe.astypes()

dataset["SepalWidthCm"]=dataset["SepalWidthCm"].astype("int")
```

In [93]:
```
#Print Result

dataset
```

Out[93]:

	Id	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	0.980392	3	1.4	0.2	Iris-setosa
1	2	1.020408	3	1.4	0.2	Iris-setosa
2	3	1.063830	3	1.3	0.2	Iris-setosa
3	4	1.086957	3	1.5	0.2	Iris-setosa
4	5	1.000000	3	1.4	0.2	Iris-setosa
5	6	0.925926	3	1.7	0.4	Iris-setosa
6	7	1.086957	3	1.4	0.3	Iris-setosa
7	8	1.000000	3	1.5	0.2	Iris-setosa
8	9	1.136364	2	1.4	0.2	Iris-setosa
9	10	1.020408	3	1.5	0.1	Iris-setosa
10	11	0.925926	3	1.5	0.2	Iris-setosa
11	12	1.041667	3	1.6	0.2	Iris-setosa
12	13	1.041667	3	1.4	0.1	Iris-setosa
13	14	1.162791	3	1.1	0.1	Iris-setosa
14	15	0.862069	4	1.2	0.2	Iris-setosa
15	16	0.877193	4	1.5	0.4	Iris-setosa
16	17	0.925926	3	1.3	0.4	Iris-setosa
17	18	0.980392	3	1.4	0.3	Iris-setosa
18	19	0.877193	3	1.7	0.3	Iris-setosa
19	20	0.980392	3	1.5	0.3	Iris-setosa
20	21	0.925926	3	1.7	0.2	Iris-setosa
21	22	0.980392	3	1.5	0.4	Iris-setosa
22	23	1.086957	3	1.0	0.2	Iris-setosa
23	24	0.980392	3	1.7	0.5	Iris-setosa
24	25	1.041667	3	1.9	0.2	Iris-setosa
25	26	1.000000	3	1.6	0.2	Iris-setosa
26	27	1.000000	3	1.6	0.4	Iris-setosa
27	28	0.961538	3	1.5	0.2	Iris-setosa
28	29	0.961538	3	1.4	0.2	Iris-setosa
29	30	1.063830	3	1.6	0.2	Iris-setosa
...	...	...	...	...	...	...
120	121	0.724638	3	5.7	2.3	Iris-virginica
121	122	0.892857	2	4.9	2.0	Iris-virginica

	Id	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
122	123	0.649351	2	6.7	2.0	Iris-virginica
123	124	0.793651	2	4.9	1.8	Iris-virginica
124	125	0.746269	3	5.7	2.1	Iris-virginica
125	126	0.694444	3	6.0	1.8	Iris-virginica
126	127	0.806452	2	4.8	1.8	Iris-virginica
127	128	0.819672	3	4.9	1.8	Iris-virginica
128	129	0.781250	2	5.6	2.1	Iris-virginica
129	130	0.694444	3	5.8	1.6	Iris-virginica
130	131	0.675676	2	6.1	1.9	Iris-virginica
131	132	0.632911	3	6.4	2.0	Iris-virginica
132	133	0.781250	2	5.6	2.2	Iris-virginica
133	134	0.793651	2	5.1	1.5	Iris-virginica
134	135	0.819672	2	5.6	1.4	Iris-virginica
135	136	0.649351	3	6.1	2.3	Iris-virginica
136	137	0.793651	3	5.6	2.4	Iris-virginica
137	138	0.781250	3	5.5	1.8	Iris-virginica
138	139	0.833333	3	4.8	1.8	Iris-virginica
139	140	0.724638	3	5.4	2.1	Iris-virginica
140	141	0.746269	3	5.6	2.4	Iris-virginica
141	142	0.724638	3	5.1	2.3	Iris-virginica
142	143	0.862069	2	5.1	1.9	Iris-virginica
143	144	0.735294	3	5.9	2.3	Iris-virginica
144	145	0.746269	3	5.7	2.5	Iris-virginica
145	146	0.746269	3	5.2	2.3	Iris-virginica
146	147	0.793651	2	5.0	1.9	Iris-virginica
147	148	0.769231	3	5.2	2.0	Iris-virginica
148	149	0.806452	3	5.4	2.3	Iris-virginica
149	150	0.847458	3	5.1	1.8	Iris-virginica

150 rows × 6 columns

In [96]:
```
# Turning categorical variables into quantitative variables in Python

# One hot Encoding

# Drop Id columns beacause it is not necessary since index will be populated auto

dataset.drop(columns='Id', inplace=True)
```

In [97]:
```
# Print result

dataset
```

Out[97]:

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	0.980392	3	1.4	0.2	Iris-setosa
1	1.020408	3	1.4	0.2	Iris-setosa
2	1.063830	3	1.3	0.2	Iris-setosa
3	1.086957	3	1.5	0.2	Iris-setosa
4	1.000000	3	1.4	0.2	Iris-setosa
5	0.925926	3	1.7	0.4	Iris-setosa
6	1.086957	3	1.4	0.3	Iris-setosa
7	1.000000	3	1.5	0.2	Iris-setosa
8	1.136364	2	1.4	0.2	Iris-setosa
9	1.020408	3	1.5	0.1	Iris-setosa
10	0.925926	3	1.5	0.2	Iris-setosa
11	1.041667	3	1.6	0.2	Iris-setosa
12	1.041667	3	1.4	0.1	Iris-setosa
13	1.162791	3	1.1	0.1	Iris-setosa
14	0.862069	4	1.2	0.2	Iris-setosa
15	0.877193	4	1.5	0.4	Iris-setosa
16	0.925926	3	1.3	0.4	Iris-setosa
17	0.980392	3	1.4	0.3	Iris-setosa
18	0.877193	3	1.7	0.3	Iris-setosa
19	0.980392	3	1.5	0.3	Iris-setosa
20	0.925926	3	1.7	0.2	Iris-setosa
21	0.980392	3	1.5	0.4	Iris-setosa
22	1.086957	3	1.0	0.2	Iris-setosa
23	0.980392	3	1.7	0.5	Iris-setosa
24	1.041667	3	1.9	0.2	Iris-setosa
25	1.000000	3	1.6	0.2	Iris-setosa
26	1.000000	3	1.6	0.4	Iris-setosa
27	0.961538	3	1.5	0.2	Iris-setosa
28	0.961538	3	1.4	0.2	Iris-setosa
29	1.063830	3	1.6	0.2	Iris-setosa
...	...	...	...	...	...
120	0.724638	3	5.7	2.3	Iris-virginica

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
121	0.892857	2	4.9	2.0	Iris-virginica
122	0.649351	2	6.7	2.0	Iris-virginica
123	0.793651	2	4.9	1.8	Iris-virginica
124	0.746269	3	5.7	2.1	Iris-virginica
125	0.694444	3	6.0	1.8	Iris-virginica
126	0.806452	2	4.8	1.8	Iris-virginica
127	0.819672	3	4.9	1.8	Iris-virginica
128	0.781250	2	5.6	2.1	Iris-virginica
129	0.694444	3	5.8	1.6	Iris-virginica
130	0.675676	2	6.1	1.9	Iris-virginica
131	0.632911	3	6.4	2.0	Iris-virginica
132	0.781250	2	5.6	2.2	Iris-virginica
133	0.793651	2	5.1	1.5	Iris-virginica
134	0.819672	2	5.6	1.4	Iris-virginica
135	0.649351	3	6.1	2.3	Iris-virginica
136	0.793651	3	5.6	2.4	Iris-virginica
137	0.781250	3	5.5	1.8	Iris-virginica
138	0.833333	3	4.8	1.8	Iris-virginica
139	0.724638	3	5.4	2.1	Iris-virginica
140	0.746269	3	5.6	2.4	Iris-virginica
141	0.724638	3	5.1	2.3	Iris-virginica
142	0.862069	2	5.1	1.9	Iris-virginica
143	0.735294	3	5.9	2.3	Iris-virginica
144	0.746269	3	5.7	2.5	Iris-virginica
145	0.746269	3	5.2	2.3	Iris-virginica
146	0.793651	2	5.0	1.9	Iris-virginica
147	0.769231	3	5.2	2.0	Iris-virginica
148	0.806452	3	5.4	2.3	Iris-virginica
149	0.847458	3	5.1	1.8	Iris-virginica

150 rows × 5 columns

In [99]:
```
# Print First 5 Row

dataset.head()
```

Out[99]:

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	0.980392	3	1.4	0.2	Iris-setosa
1	1.020408	3	1.4	0.2	Iris-setosa
2	1.063830	3	1.3	0.2	Iris-setosa
3	1.086957	3	1.5	0.2	Iris-setosa
4	1.000000	3	1.4	0.2	Iris-setosa

In [100]:
```
#Use get_dummies method from panda class to create dummy variable and store those

dummies = pd.get_dummies(dataset.Species)

#Display last 5 dummies

dummies.tail()
```

Out[100]:

	Iris-setosa	Iris-versicolor	Iris-virginica
145	0	0	1
146	0	0	1
147	0	0	1
148	0	0	1
149	0	0	1

In [110]:
```
#Concate the newly created dummy variables with the loaded data and store in merg
merged_data = pd.concat([dataset,dummies], axis=1)
```

In [111]:
```
merged_data
```

Out[111]:

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	Iris-setosa	Iris versicolc
0	0.980392	3	1.4	0.2	Iris-setosa	1	
1	1.020408	3	1.4	0.2	Iris-setosa	1	
2	1.063830	3	1.3	0.2	Iris-setosa	1	
3	1.086957	3	1.5	0.2	Iris-setosa	1	
4	1.000000	3	1.4	0.2	Iris-setosa	1	
5	0.925926	3	1.7	0.4	Iris-setosa	1	
6	1.086957	3	1.4	0.3	Iris-setosa	1	
7	1.000000	3	1.5	0.2	Iris-setosa	1	
8	1.136364	2	1.4	0.2	Iris-setosa	1	
9	1.020408	3	1.5	0.1	Iris-setosa	1	
10	0.925926	3	1.5	0.2	Iris-setosa	1	
11	1.041667	3	1.6	0.2	Iris-setosa	1	
12	1.041667	3	1.4	0.1	Iris-setosa	1	
13	1.162791	3	1.1	0.1	Iris-setosa	1	
14	0.862069	4	1.2	0.2	Iris-setosa	1	
15	0.877193	4	1.5	0.4	Iris-setosa	1	
16	0.925926	3	1.3	0.4	Iris-setosa	1	
17	0.980392	3	1.4	0.3	Iris-setosa	1	
18	0.877193	3	1.7	0.3	Iris-setosa	1	
19	0.980392	3	1.5	0.3	Iris-setosa	1	
20	0.925926	3	1.7	0.2	Iris-setosa	1	
21	0.980392	3	1.5	0.4	Iris-setosa	1	

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	Iris-setosa	Iris versicolo
22	1.086957	3	1.0	0.2	Iris-setosa	1	
23	0.980392	3	1.7	0.5	Iris-setosa	1	
24	1.041667	3	1.9	0.2	Iris-setosa	1	
25	1.000000	3	1.6	0.2	Iris-setosa	1	
26	1.000000	3	1.6	0.4	Iris-setosa	1	
27	0.961538	3	1.5	0.2	Iris-setosa	1	
28	0.961538	3	1.4	0.2	Iris-setosa	1	
29	1.063830	3	1.6	0.2	Iris-setosa	1	
...	...	...	...	...	...	...	.
120	0.724638	3	5.7	2.3	Iris-virginica	0	
121	0.892857	2	4.9	2.0	Iris-virginica	0	
122	0.649351	2	6.7	2.0	Iris-virginica	0	
123	0.793651	2	4.9	1.8	Iris-virginica	0	
124	0.746269	3	5.7	2.1	Iris-virginica	0	
125	0.694444	3	6.0	1.8	Iris-virginica	0	
126	0.806452	2	4.8	1.8	Iris-virginica	0	
127	0.819672	3	4.9	1.8	Iris-virginica	0	
128	0.781250	2	5.6	2.1	Iris-virginica	0	
129	0.694444	3	5.8	1.6	Iris-virginica	0	
130	0.675676	2	6.1	1.9	Iris-virginica	0	
131	0.632911	3	6.4	2.0	Iris-virginica	0	
132	0.781250	2	5.6	2.2	Iris-virginica	0	
133	0.793651	2	5.1	1.5	Iris-virginica	0	

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	Iris-setosa	Iris versicolc
134	0.819672	2	5.6	1.4	Iris-virginica	0	
135	0.649351	3	6.1	2.3	Iris-virginica	0	
136	0.793651	3	5.6	2.4	Iris-virginica	0	
137	0.781250	3	5.5	1.8	Iris-virginica	0	
138	0.833333	3	4.8	1.8	Iris-virginica	0	
139	0.724638	3	5.4	2.1	Iris-virginica	0	
140	0.746269	3	5.6	2.4	Iris-virginica	0	
141	0.724638	3	5.1	2.3	Iris-virginica	0	
142	0.862069	2	5.1	1.9	Iris-virginica	0	
143	0.735294	3	5.9	2.3	Iris-virginica	0	
144	0.746269	3	5.7	2.5	Iris-virginica	0	
145	0.746269	3	5.2	2.3	Iris-virginica	0	
146	0.793651	2	5.0	1.9	Iris-virginica	0	
147	0.769231	3	5.2	2.0	Iris-virginica	0	
148	0.806452	3	5.4	2.3	Iris-virginica	0	
149	0.847458	3	5.1	1.8	Iris-virginica	0	

150 rows × 8 columns

In [112]:
```
#Display merged_data

merged_data.head()
```

Out[112]:

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	Iris-setosa	Iris-versicolor
0	0.980392	3	1.4	0.2	Iris-setosa	1	0
1	1.020408	3	1.4	0.2	Iris-setosa	1	0
2	1.063830	3	1.3	0.2	Iris-setosa	1	0
3	1.086957	3	1.5	0.2	Iris-setosa	1	0
4	1.000000	3	1.4	0.2	Iris-setosa	1	0

In [113]:
```
#Drop the Species column since we've done one hot encoding, hence Species column

final_data = merged_data.drop(columns='Species')
```

In [114]:
```
# Print the Result

final_data
```

Out[114]:

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Iris-setosa	Iris-versicolor	Iri virginic
0	0.980392	3	1.4	0.2	1	0	
1	1.020408	3	1.4	0.2	1	0	
2	1.063830	3	1.3	0.2	1	0	
3	1.086957	3	1.5	0.2	1	0	
4	1.000000	3	1.4	0.2	1	0	
5	0.925926	3	1.7	0.4	1	0	
6	1.086957	3	1.4	0.3	1	0	
7	1.000000	3	1.5	0.2	1	0	
8	1.136364	2	1.4	0.2	1	0	
9	1.020408	3	1.5	0.1	1	0	
10	0.925926	3	1.5	0.2	1	0	
11	1.041667	3	1.6	0.2	1	0	
12	1.041667	3	1.4	0.1	1	0	
13	1.162791	3	1.1	0.1	1	0	
14	0.862069	4	1.2	0.2	1	0	
15	0.877193	4	1.5	0.4	1	0	
16	0.925926	3	1.3	0.4	1	0	
17	0.980392	3	1.4	0.3	1	0	
18	0.877193	3	1.7	0.3	1	0	
19	0.980392	3	1.5	0.3	1	0	
20	0.925926	3	1.7	0.2	1	0	
21	0.980392	3	1.5	0.4	1	0	
22	1.086957	3	1.0	0.2	1	0	
23	0.980392	3	1.7	0.5	1	0	
24	1.041667	3	1.9	0.2	1	0	
25	1.000000	3	1.6	0.2	1	0	
26	1.000000	3	1.6	0.4	1	0	
27	0.961538	3	1.5	0.2	1	0	
28	0.961538	3	1.4	0.2	1	0	
29	1.063830	3	1.6	0.2	1	0	
...	...	...	...	...	...	...	
120	0.724638	3	5.7	2.3	0	0	

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Iris-setosa	Iris-versicolor	Iri virginic
121	0.892857	2	4.9	2.0	0	0	
122	0.649351	2	6.7	2.0	0	0	
123	0.793651	2	4.9	1.8	0	0	
124	0.746269	3	5.7	2.1	0	0	
125	0.694444	3	6.0	1.8	0	0	
126	0.806452	2	4.8	1.8	0	0	
127	0.819672	3	4.9	1.8	0	0	
128	0.781250	2	5.6	2.1	0	0	
129	0.694444	3	5.8	1.6	0	0	
130	0.675676	2	6.1	1.9	0	0	
131	0.632911	3	6.4	2.0	0	0	
132	0.781250	2	5.6	2.2	0	0	
133	0.793651	2	5.1	1.5	0	0	
134	0.819672	2	5.6	1.4	0	0	
135	0.649351	3	6.1	2.3	0	0	
136	0.793651	3	5.6	2.4	0	0	
137	0.781250	3	5.5	1.8	0	0	
138	0.833333	3	4.8	1.8	0	0	
139	0.724638	3	5.4	2.1	0	0	
140	0.746269	3	5.6	2.4	0	0	
141	0.724638	3	5.1	2.3	0	0	
142	0.862069	2	5.1	1.9	0	0	
143	0.735294	3	5.9	2.3	0	0	
144	0.746269	3	5.7	2.5	0	0	
145	0.746269	3	5.2	2.3	0	0	
146	0.793651	2	5.0	1.9	0	0	
147	0.769231	3	5.2	2.0	0	0	
148	0.806452	3	5.4	2.3	0	0	
149	0.847458	3	5.1	1.8	0	0	

150 rows × 7 columns

In [115]:
```
#Display first 5 rows of the final_data

final_data.head()
```

Out[115]:

	NewSepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Iris-setosa	Iris-versicolor	Iris-virginica
0	0.980392	3	1.4	0.2	1	0	0
1	1.020408	3	1.4	0.2	1	0	0
2	1.063830	3	1.3	0.2	1	0	0
3	1.086957	3	1.5	0.2	1	0	0
4	1.000000	3	1.4	0.2	1	0	0

In [ ]:
```

```