## **Pandas**

Pandas is an open source library which provides high-performance, easyto-use data structures and data analysis tools for the Python programming language. Pandas has a lot of functions that will help in reading and writing data and also for data manipulation. Thus we will be using pandas throughout the course.

```
#Import Pandas
In [2]:
        import pandas as pd
        #Loading data with read_csv() function. Here we are providing path to the csv file
        #If the file is in your system you can provide its path as well.
        iris = pd.read csv("https://archive.ics.uci.edu/ml/machine-learning-databases/iris/
        #Let's print and see its type
        print(type(iris))
```

<class 'pandas.core.frame.DataFrame'>

## **Pandas Dataframes**

DataFrame is an object for data manipulation. You can think of it as a 2D tabular structure, where every row is a dataset entry and columns represents features of data.

## Creating copy of DataFrame

```
df = iris
```

Above statement simply makes df refer to the data frame object that iris is referring to. So now both iris and df refer to the same dataframe object and any changes done via one will reflect in other.

So effectively this is not creating another dataframe object. And if we wish to create a copy then we will use copy() function for that

```
df = iris.copy()
```

```
#Ignoring header -> If you don't want first row to be treated as a header, you can
iris = pd.read csv("https://archive.ics.uci.edu/ml/machine-learning-databases/iris/
iris
```

Out[5]:

	0	1	2	3	4	
0	5.1	3.5	1.4	0.2	Iris-setosa	
1	4.9	3.0	1.4	0.2	Iris-setosa	
2	4.7	3.2	1.3	0.2	Iris-setosa	
3	4.6	3.1	1.5	0.2	Iris-setosa	
4	5.0	3.6	1.4	0.2	Iris-setosa	
5	5.4	3.9	1.7	0.4	Iris-setosa	
6	4.6	3.4	1.4	0.3	Iris-setosa	
7	5.0	3.4	1.5	0.2	Iris-setosa	
8	4.4	2.9	1.4	0.2	Iris-setosa	
9	4.9	3.1	1.5	0.1	Iris-setosa	
10	5.4	3.7	1.5	0.2	Iris-setosa	
11	4.8	3.4	1.6	0.2	Iris-setosa	
12	4.8	3.0	1.4	0.1	Iris-setosa	
13	4.3	3.0	1.1	0.1	Iris-setosa	
14	5.8	4.0	1.2	0.2	Iris-setosa	
15	5.7	4.4	1.5	0.4	Iris-setosa	
16	5.4	3.9	1.3	0.4	Iris-setosa	
17	5.1	3.5	1.4	0.3	Iris-setosa	
18	5.7	3.8	1.7	0.3	Iris-setosa	
19	5.1	3.8	1.5	0.3	Iris-setosa	
20	5.4	3.4	1.7	0.2	Iris-setosa	
21	5.1	3.7	1.5	0.4	Iris-setosa	
22	4.6	3.6	1.0	0.2	Iris-setosa	
23	5.1	3.3	1.7	0.5	Iris-setosa	
24	4.8	3.4	1.9	0.2	Iris-setosa	
25	5.0	3.0	1.6	0.2	Iris-setosa	
26	5.0	3.4	1.6	0.4	Iris-setosa	
27	5.2	3.5	1.5	0.2	Iris-setosa	
28	5.2	3.4	1.4	0.2	Iris-setosa	
29	4.7	3.2	1.6	0.2	Iris-setosa	
•••						
120	6.9	3.2	5.7	2.3	Iris-virginica	
121	5.6	2.8	4.9	2.0	Iris-virginica	
122	7.7	2.8	6.7	2.0	Iris-virginica	
123	6.3	2.7	4.9	1.8	Iris-virginica	
124	6.7	3.3	5.7	2.1	Iris-virginica	

	0	1	2	3	4
125	7.2	3.2	6.0	1.8	Iris-virginica
126	6.2	2.8	4.8	1.8	Iris-virginica
127	6.1	3.0	4.9	1.8	Iris-virginica
128	6.4	2.8	5.6	2.1	Iris-virginica
129	7.2	3.0	5.8	1.6	Iris-virginica
130	7.4	2.8	6.1	1.9	Iris-virginica
131	7.9	3.8	6.4	2.0	Iris-virginica
132	6.4	2.8	5.6	2.2	Iris-virginica
133	6.3	2.8	5.1	1.5	Iris-virginica
134	6.1	2.6	5.6	1.4	Iris-virginica
135	7.7	3.0	6.1	2.3	Iris-virginica
136	6.3	3.4	5.6	2.4	Iris-virginica
137	6.4	3.1	5.5	1.8	Iris-virginica
138	6.0	3.0	4.8	1.8	Iris-virginica
139	6.9	3.1	5.4	2.1	Iris-virginica
140	6.7	3.1	5.6	2.4	Iris-virginica
141	6.9	3.1	5.1	2.3	Iris-virginica
142	5.8	2.7	5.1	1.9	Iris-virginica
143	6.8	3.2	5.9	2.3	Iris-virginica
144	6.7	3.3	5.7	2.5	Iris-virginica
145	6.7	3.0	5.2	2.3	Iris-virginica
146	6.3	2.5	5.0	1.9	Iris-virginica
147	6.5	3.0	5.2	2.0	Iris-virginica
148	6.2	3.4	5.4	2.3	Iris-virginica
149	5.9	3.0	5.1	1.8	Iris-virginica

150 rows × 5 columns

```
In [6]: df = iris.copy()
        print(df)
        #Iris data has a total of 150 rows, where the first row is treated as a header
```

	0	1	2	3	4
0	5.1	3.5	1.4	0.2	· · · · · · · · · · · · · · · · · · ·
					Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa
5	5.4	3.9	1.7	0.4	Iris-setosa
6	4.6	3.4	1.4	0.3	Iris-setosa
7	5.0	3.4	1.5	0.2	Iris-setosa
8	4.4	2.9	1.4	0.2	Iris-setosa
9	4.9	3.1	1.5	0.1	Iris-setosa
10	5.4	3.7	1.5	0.2	Iris-setosa
11	4.8	3.4	1.6	0.2	Iris-setosa
12	4.8	3.0	1.4	0.1	Iris-setosa
13	4.3	3.0	1.1	0.1	Iris-setosa
14	5.8	4.0	1.2	0.2	Iris-setosa
15	5.7	4.4	1.5	0.4	Iris-setosa
16	5.4	3.9	1.3	0.4	Iris-setosa
17	5.1	3.5	1.4	0.3	Iris-setosa
18	5.7	3.8	1.7	0.3	Iris-setosa
19	5.1	3.8	1.5	0.3	Iris-setosa
20	5.4	3.4	1.7	0.2	Iris-setosa
21	5.1	3.7	1.5	0.4	Iris-setosa
22	4.6	3.6	1.0	0.2	Iris-setosa
23	5.1	3.3	1.7	0.5	Iris-setosa
24	4.8	3.4	1.9	0.2	Iris-setosa
25	5.0	3.0	1.6	0.2	Iris-setosa
26	5.0	3.4	1.6	0.4	Iris-setosa
27	5.2	3.5	1.5	0.2	Iris-setosa
28	5.2	3.4	1.4	0.2	Iris-setosa
29	4.7	3.2	1.6	0.2	Iris-setosa
		٥. ٢	0	~	1113 300030
				• • •	
					 Iris-virginica
					 Iris-virginica
 120	6.9	3.2	5.7	2.3	 Iris-virginica Iris-virginica
 120 121	6.9 5.6	3.2 2.8	5.7 4.9	2.3 2.0	 Iris-virginica Iris-virginica Iris-virginica
120 121 122 123	6.9 5.6 7.7 6.3	3.2 2.8 2.8 2.7	5.7 4.9 6.7 4.9	2.3 2.0 2.0 1.8	Iris-virginica Iris-virginica Iris-virginica Iris-virginica
120 121 122 123 124	6.9 5.6 7.7 6.3 6.7	3.2 2.8 2.8 2.7 3.3	5.7 4.9 6.7 4.9 5.7	2.3 2.0 2.0 1.8 2.1	Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica
120 121 122 123 124 125	6.9 5.6 7.7 6.3 6.7	3.2 2.8 2.8 2.7 3.3	5.7 4.9 6.7 4.9 5.7 6.0	2.3 2.0 2.0 1.8 2.1	Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica
120 121 122 123 124 125 126	6.9 5.6 7.7 6.3 6.7 7.2	3.2 2.8 2.8 2.7 3.3 3.2 2.8	5.7 4.9 6.7 4.9 5.7 6.0 4.8	2.3 2.0 2.0 1.8 2.1 1.8	Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica
120 121 122 123 124 125 126 127	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9	2.3 2.0 2.0 1.8 2.1 1.8 1.8	Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica
120 121 122 123 124 125 126 127 128	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6	2.3 2.0 2.0 1.8 2.1 1.8 1.8	Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica
120 121 122 123 124 125 126 127 128 129	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1	2.3 2.0 2.0 1.8 2.1 1.8 1.8 1.8 1.6 1.9	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4	2.3 2.0 2.0 1.8 2.1 1.8 1.8 1.8 2.1 1.6 1.9	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6	2.3 2.0 2.0 1.8 2.1 1.8 1.8 1.8 2.1 1.6 1.9 2.0 2.2	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.3	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.8	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.8 2.8 2.8 2.8 2.6 3.0	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 5.6 5.1 5.6 6.1	2.3 2.0 2.0 1.8 2.1 1.8 1.8 1.6 1.9 2.0 2.2 1.5 1.4 2.3	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.8 2.8 2.6 3.0	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1 5.6 6.1 5.6	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 2.8 2.8 2.8 2.6 3.0 3.4 3.1	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1 5.6 6.1 5.6	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 2.8 2.8 2.6 3.0 3.4 3.1	5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1 5.6 6.1 5.6 5.5 4.8	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.9	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.6 3.4 3.1 3.0 3.1	5.7 4.9 6.7 4.9 5.6 4.8 4.9 5.6 5.8 6.1 5.6 5.1 5.6 6.1 5.6 5.5 4.8 5.4	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.3 6.1 7.7 6.3 6.4 6.9 6.7	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 2.8 2.6 3.0 3.4 3.1 3.0 3.1	5.7 4.9 6.7 4.9 5.6 4.8 4.9 5.6 5.1 5.6 6.1 5.6 5.5 4.8 5.4	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.9	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.6 3.4 3.1 3.0 3.1	5.7 4.9 6.7 4.9 5.6 5.8 6.1 5.6 5.1 5.6 5.5 4.8 5.6 5.5 4.8 5.6 5.5	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.3 6.1 7.7 6.3 6.4 6.9 6.7	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 2.8 2.6 3.0 3.4 3.1 3.0 3.1	5.7 4.9 6.7 4.9 5.6 4.8 4.9 5.6 5.1 5.6 6.1 5.6 5.5 4.8 5.4	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141	6.9 5.6 7.7 6.3 6.7 7.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.9 6.7	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.0 2.8 3.0 3.1 3.1 3.1 3.1	5.7 4.9 6.7 4.9 5.6 5.8 6.1 5.6 5.1 5.6 5.5 4.8 5.6 5.5 4.8 5.6 5.5	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 2.1	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141	6.9 5.6 7.7 6.3 6.7 7.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.9 6.9 5.8	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.0 2.8 3.0 3.1 3.1 3.1 3.1 2.7	5.7 4.9 6.7 4.9 5.6 4.8 4.9 5.6 5.1 5.6 5.1 5.6 5.5 4.8 5.4 5.5 5.1 5.1	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.9 6.9 5.8 6.8	3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 2.8 2.8 2.6 3.0 3.1 3.1 3.1 3.1 3.1 3.1	5.7 4.9 6.7 4.9 5.6 4.8 4.9 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.7	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 1.8 2.1 2.4 2.3 2.4 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.0 3.1 3.1 3.1 3.1 3.1 3.1 3.1 3.2 3.3	5.7 4.9 6.7 4.9 5.6 4.8 4.9 5.6 5.1 5.6 6.1 5.6 5.5 4.8 5.6 5.1 5.6 5.1 5.6 5.7	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 1.8 2.1 2.3 2.4 1.8	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144	6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.7 6.7	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.0 3.1 3.1 3.1 3.1 3.1 3.1 3.2 3.3 3.0 3.1	5.7 4.9 6.7 4.9 5.6 5.8 6.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 2.1 2.3 2.4 1.8 2.1	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146	6.9 5.6 7.7 6.3 6.7 7.2 6.4 7.9 6.4 7.7 6.3 6.1 7.7 6.3 6.9 6.9 6.9 6.7 6.9 6.7 6.3	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.0 3.1 3.1 3.1 3.1 3.1 3.1 3.2 3.3 3.0 3.1 3.1 3.1 3.1 3.2	5.7 4.9 6.7 4.9 5.6 5.8 6.1 5.6 5.1 5.6 5.5 4.8 5.6 5.1 5.6 5.1 5.6 5.1 5.7 5.7 5.9 5.7 5.9 5.7 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0	2.3 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 2.1 2.4 2.3 2.4 2.3 2.4 2.3 2.1 2.1 2.1 2.1 2.1 2.1 2.1 2.1 2.1 2.1	Iris-virginica
120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147	5.6 7.7 6.3 6.7 7.2 6.1 6.4 7.2 7.4 7.9 6.3 6.1 7.7 6.3 6.9 6.9 6.7 6.8 6.7 6.3 6.7 6.3 6.7 6.9	2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.0 3.1 3.1 3.1 3.1 3.1 3.1 3.2 3.3 3.0 3.1	5.7 4.9 6.7 4.9 5.6 5.8 6.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.6 5.1 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7	2.3 2.0 2.0 1.8 2.1 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 2.1 2.3 2.4 1.8 2.1	Iris-virginica

[150 rows x 5 columns]

```
In [12]: # head() function is a good way to have a look at first few enteries of the data fire #You can write df.head(n) if you wish to have a look at first n enteries. #df.head()
```

```
      Out[12]:
      0
      1
      2
      3
      4

      0
      5.1
      3.5
      1.4
      0.2
      Iris-setosa

      1
      4.9
      3.0
      1.4
      0.2
      Iris-setosa

      2
      4.7
      3.2
      1.3
      0.2
      Iris-setosa

      3
      4.6
      3.1
      1.5
      0.2
      Iris-setosa

      4
      5.0
      3.6
      1.4
      0.2
      Iris-setosa
```

```
In [22]: #Column Headers
    df.columns # Will tell you what the current column headers are
    #Changing Column Headers
    df.columns = ['sl','sw','pl','pw','flower_type']
```

```
In [23]: #shape tells us about the shape of the structure.
print(df.shape)
#dtypes will tell you type of data each column is storing
print(df.dtypes)
```

```
(150, 5)
sl float64
sw float64
pl float64
pw float64
flower_type object
dtype: object
```

In [46]: #describe() function is a good way to get some statistics around the data. By defact
#Also the results exclude NaN enteries
print(df.describe())
#If you want to include all columns for result, then use df.describe(include='all')

```
sl
                                      pl
                          SW
count 150.000000 150.000000 150.000000 150.000000
mean
        5.843333
                    3.054000
                                3.758667
                                            1.198667
std
        0.828066
                    0.433594
                                1.764420
                                            0.763161
min
        4.300000
                    2.000000
                                1.000000
                                            0.100000
25%
        5.100000
                    2.800000
                                1.600000
                                            0.300000
50%
                    3.000000
                                4.350000
        5.800000
                                            1.300000
75%
        6.400000
                    3.300000
                                5.100000
                                            1.800000
        7.900000
                    4.400000
                                6.900000
                                            2.500000
max
```

150.000000

pw flower\_type

150

pΙ

Out[46]:

sl

count 150.000000 150.000000 150.000000

sw

```
NaN
                                                                       3
          unique
                       NaN
                                             NaN
                                                        NaN
                       NaN
                                  NaN
                                             NaN
                                                              Iris-virginica
             top
                                                        NaN
                       NaN
                                  NaN
                                             NaN
                                                        NaN
                                                                      50
            freq
           mean
                    5.843333
                               3.054000
                                          3.758667
                                                     1.198667
                                                                    NaN
                                          1.764420
                    0.828066
                               0.433594
                                                                    NaN
             std
                                                     0.763161
             min
                    4.300000
                               2.000000
                                          1.000000
                                                     0.100000
                                                                    NaN
            25%
                               2.800000
                                                                    NaN
                    5.100000
                                          1.600000
                                                     0.300000
            50%
                    5.800000
                               3.000000
                                          4.350000
                                                     1.300000
                                                                    NaN
            75%
                    6.400000
                               3.300000
                                          5.100000
                                                     1.800000
                                                                    NaN
                    7.900000
                               4.400000
                                          6.900000
                                                     2.500000
                                                                    NaN
            max
          #Accessing a particular column.
In [37]:
          #df.column_name or df['column_name']lets you access a particular column
          df.sl #OR
          df['sl']
          #You can also call describe() on a particular column
          df.sl.describe()
          #OR
          df['sl'].describe()
                   150.000000
          count
Out[37]:
          mean
                     5.843333
          std
                     0.828066
          min
                     4.300000
          25%
                     5.100000
          50%
                     5.800000
          75%
                     6.400000
          max
                     7.900000
          Name: sl, dtype: float64
In [42]:
          #Null enteries
          #df.isnull() will give you list of null enteries
          df.isnull()
          #Although df.isnull() in itself might not be very useful, we can use df.isnull().se
          df.isnull().sum()
                          0
          sl
Out[42]:
                          0
          SW
                          0
          p1
                          0
          рw
          flower_type
                          0
          dtype: int64
          #Selecting Rows and/or columns by position
In [52]:
          #iloc helps us in achieving this.
          #Selecting first 4 rows of data
          df.iloc[:4,:]
          #Selecting data for first 4 rows, and for first 2 columns
          df.iloc[:4,:2]
          #Selecting data for rows 2 to 5, and for columns 1 to 3
          df.iloc[2:6,1:4]
```

Out[52]:		sw	pl	pw
	2	3.2	1.3	0.2
	3	3.1	1.5	0.2
	4	3.6	1.4	0.2
	5	3.9	1.7	0.4