

In [1]:

```
import pandas as pd
import numpy as np
import statistics as st
```

In [2]:

```
df = pd.read_csv('iris.csv')
```

In [3]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149 entries, 0 to 148
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   5.1              149 non-null    float64
1   3.5              149 non-null    float64
2   1.4              149 non-null    float64
3   0.2              149 non-null    float64
4   Iris-setosa      149 non-null    object
dtypes: float64(4), object(1)
memory usage: 5.9+ KB
```

In [4]:

```
df.describe()
```

Out[4]:

	5.1	3.5	1.4	0.2
count	149.000000	149.000000	149.000000	149.000000
mean	5.848322	3.051007	3.774497	1.205369
std	0.828594	0.433499	1.759651	0.761292
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.400000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

In [5]:

```
df.columns
```

Out[5]:

```
Index(['5.1', '3.5', '1.4', '0.2', 'Iris-setosa'], dtype='object')
```

In [6]:

```
df
```

Out[6]:

	5.1	3.5	1.4	0.2	Iris-setosa
0	4.9	3.0	1.4	0.2	Iris-setosa
1	4.7	3.2	1.3	0.2	Iris-setosa
2	4.6	3.1	1.5	0.2	Iris-setosa
3	5.0	3.6	1.4	0.2	Iris-setosa
4	5.4	3.9	1.7	0.4	Iris-setosa
...
144	6.7	3.0	5.2	2.3	Iris-virginica
145	6.3	2.5	5.0	1.9	Iris-virginica
146	6.5	3.0	5.2	2.0	Iris-virginica
147	6.2	3.4	5.4	2.3	Iris-virginica
148	5.9	3.0	5.1	1.8	Iris-virginica

149 rows × 5 columns

In [7]:

```
df=df.set_index('Unnamed: 0')
```

```
-----
-----
KeyError                                Traceback (most recent call
last)
/tmp/ipykernel_6879/2784185274.py in <module>
----> 1 df=df.set_index('Unnamed: 0')

~/.local/lib/python3.10/site-packages/pandas/util/_decorators.py in wr
apper(*args, **kwargs)
    329             stacklevel=find_stack_level(),
    330         )
--> 331         return func(*args, **kwargs)
    332
    333     # error: "Callable[[VarArg(Any), KwArg(Any)], Any]" ha
s no

~/.local/lib/python3.10/site-packages/pandas/core/frame.py in set_inde
x(self, keys, drop, append, inplace, verify_integrity)
    5999
    6000         if missing:
-> 6001             raise KeyError(f"None of {missing} are in the colu
mns")
    6002
    6003         if inplace:

KeyError: "None of ['Unnamed: 0'] are in the columns"
```

In [9]:

```
missing_count=df.isnull().sum()
```

In [10]:

```
missing_count
```

Out[10]:

```
5.1      0
3.5      0
1.4      0
0.2      0
Iris-setosa  0
dtype: int64
```

In [12]:

```
df[df['5.1'].isnull()]
```

Out[12]:

```
5.1  3.5  1.4  0.2  Iris-setosa
```

In [13]:

```
df['5.1']
```

Out[13]:

```
0      4.9
1      4.7
2      4.6
3      5.0
4      5.4
...
144    6.7
145    6.3
146    6.5
147    6.2
148    5.9
Name: 5.1, Length: 149, dtype: float64
```

In [14]:

```
df.isnull().sum()
```

Out[14]:

```
5.1      0
3.5      0
1.4      0
0.2      0
Iris-setosa  0
dtype: int64
```

In [15]:

```
df.sort_values(by='5.1',ascending=False)
```

Out[15]:

	5.1	3.5	1.4	0.2	Iris-setosa
130	7.9	3.8	6.4	2.0	Iris-virginica
134	7.7	3.0	6.1	2.3	Iris-virginica
121	7.7	2.8	6.7	2.0	Iris-virginica
116	7.7	3.8	6.7	2.2	Iris-virginica
117	7.7	2.6	6.9	2.3	Iris-virginica
...
40	4.5	2.3	1.3	0.3	Iris-setosa
7	4.4	2.9	1.4	0.2	Iris-setosa
41	4.4	3.2	1.3	0.2	Iris-setosa
37	4.4	3.0	1.3	0.2	Iris-setosa
12	4.3	3.0	1.1	0.1	Iris-setosa

149 rows × 5 columns

In [16]:

```
df.sort_values(by='0.2',ascending=True)
```

Out[16]:

	5.1	3.5	1.4	0.2	Iris-setosa
36	4.9	3.1	1.5	0.1	Iris-setosa
31	5.2	4.1	1.5	0.1	Iris-setosa
12	4.3	3.0	1.1	0.1	Iris-setosa
11	4.8	3.0	1.4	0.1	Iris-setosa
8	4.9	3.1	1.5	0.1	Iris-setosa
...
135	6.3	3.4	5.6	2.4	Iris-virginica
113	5.8	2.8	5.1	2.4	Iris-virginica
143	6.7	3.3	5.7	2.5	Iris-virginica
99	6.3	3.3	6.0	2.5	Iris-virginica
108	7.2	3.6	6.1	2.5	Iris-virginica

149 rows × 5 columns

In [17]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149 entries, 0 to 148
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   5.1              149 non-null   float64
1   3.5              149 non-null   float64
2   1.4              149 non-null   float64
3   0.2              149 non-null   float64
4   Iris-setosa      149 non-null   object
dtypes: float64(4), object(1)
memory usage: 5.9+ KB
```

In [18]:

```
df['5.1']=df['5.1'].astype(int)
```

In [19]:

```
df
```

Out[19]:

	5.1	3.5	1.4	0.2	Iris-setosa
0	4	3.0	1.4	0.2	Iris-setosa
1	4	3.2	1.3	0.2	Iris-setosa
2	4	3.1	1.5	0.2	Iris-setosa
3	5	3.6	1.4	0.2	Iris-setosa
4	5	3.9	1.7	0.4	Iris-setosa
...
144	6	3.0	5.2	2.3	Iris-virginica
145	6	2.5	5.0	1.9	Iris-virginica
146	6	3.0	5.2	2.0	Iris-virginica
147	6	3.4	5.4	2.3	Iris-virginica
148	5	3.0	5.1	1.8	Iris-virginica

149 rows × 5 columns

In []: