

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
```

```
In [3]: df = sns.load_dataset('iris')
```

```
In [4]: df
```

Out[4]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...
145	6.7	3.0	5.2	2.3	virginica
146	6.3	2.5	5.0	1.9	virginica
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

150 rows × 5 columns

```
In [5]: df.head()
```

Out[5]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

```
In [6]: df.tail()
```

```
Out[6]:
```

	sepal_length	sepal_width	petal_length	petal_width	species
145	6.7	3.0	5.2	2.3	virginica
146	6.3	2.5	5.0	1.9	virginica
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

```
In [7]: df.dtypes
```

```
Out[7]: sepal_length    float64
sepal_width    float64
petal_length    float64
petal_width    float64
species        object
dtype: object
```

```
In [9]: # filtering the data by the species
setosa = df[df['species'] == 'setosa']
versicolor = df[df['species'] == 'versicolor']
virginica = df[df['species'] == 'virginica']
```

In [10]: setosa

Out[10]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
5	5.4	3.9	1.7	0.4	setosa
6	4.6	3.4	1.4	0.3	setosa
7	5.0	3.4	1.5	0.2	setosa
8	4.4	2.9	1.4	0.2	setosa
9	4.9	3.1	1.5	0.1	setosa
10	5.4	3.7	1.5	0.2	setosa
11	4.8	3.4	1.6	0.2	setosa
12	4.8	3.0	1.4	0.1	setosa
13	4.3	3.0	1.1	0.1	setosa
14	5.8	4.0	1.2	0.2	setosa
15	5.7	4.4	1.5	0.4	setosa
16	5.4	3.9	1.3	0.4	setosa
17	5.1	3.5	1.4	0.3	setosa
18	5.7	3.8	1.7	0.3	setosa
19	5.1	3.8	1.5	0.3	setosa
20	5.4	3.4	1.7	0.2	setosa
21	5.1	3.7	1.5	0.4	setosa
22	4.6	3.6	1.0	0.2	setosa
23	5.1	3.3	1.7	0.5	setosa
24	4.8	3.4	1.9	0.2	setosa
25	5.0	3.0	1.6	0.2	setosa
26	5.0	3.4	1.6	0.4	setosa
27	5.2	3.5	1.5	0.2	setosa
28	5.2	3.4	1.4	0.2	setosa
29	4.7	3.2	1.6	0.2	setosa
30	4.8	3.1	1.6	0.2	setosa
31	5.4	3.4	1.5	0.4	setosa
32	5.2	4.1	1.5	0.1	setosa
33	5.5	4.2	1.4	0.2	setosa
34	4.9	3.1	1.5	0.2	setosa

	sepal_length	sepal_width	petal_length	petal_width	species
35	5.0	3.2	1.2	0.2	setosa
36	5.5	3.5	1.3	0.2	setosa
37	4.9	3.6	1.4	0.1	setosa
38	4.4	3.0	1.3	0.2	setosa
39	5.1	3.4	1.5	0.2	setosa
40	5.0	3.5	1.3	0.3	setosa
41	4.5	2.3	1.3	0.3	setosa
42	4.4	3.2	1.3	0.2	setosa
43	5.0	3.5	1.6	0.6	setosa
44	5.1	3.8	1.9	0.4	setosa
45	4.8	3.0	1.4	0.3	setosa
46	5.1	3.8	1.6	0.2	setosa
47	4.6	3.2	1.4	0.2	setosa
48	5.3	3.7	1.5	0.2	setosa
49	5.0	3.3	1.4	0.2	setosa

In [11]: `versicolor`

Out[11]:

	sepal_length	sepal_width	petal_length	petal_width	species
50	7.0	3.2	4.7	1.4	versicolor
51	6.4	3.2	4.5	1.5	versicolor
52	6.9	3.1	4.9	1.5	versicolor
53	5.5	2.3	4.0	1.3	versicolor
54	6.5	2.8	4.6	1.5	versicolor
55	5.7	2.8	4.5	1.3	versicolor
56	6.3	3.3	4.7	1.6	versicolor
57	4.9	2.4	3.3	1.0	versicolor
58	6.6	2.9	4.6	1.3	versicolor
59	5.2	2.7	3.9	1.4	versicolor
60	5.0	2.0	3.5	1.0	versicolor
61	5.9	3.0	4.2	1.5	versicolor
62	6.0	2.2	4.0	1.0	versicolor
63	6.1	2.9	4.7	1.4	versicolor
64	5.6	2.9	3.6	1.3	versicolor
65	6.7	3.1	4.4	1.4	versicolor
66	5.6	3.0	4.5	1.5	versicolor
67	5.8	2.7	4.1	1.0	versicolor
68	6.2	2.2	4.5	1.5	versicolor
69	5.6	2.5	3.9	1.1	versicolor
70	5.9	3.2	4.8	1.8	versicolor
71	6.1	2.8	4.0	1.3	versicolor
72	6.3	2.5	4.9	1.5	versicolor
73	6.1	2.8	4.7	1.2	versicolor
74	6.4	2.9	4.3	1.3	versicolor
75	6.6	3.0	4.4	1.4	versicolor
76	6.8	2.8	4.8	1.4	versicolor
77	6.7	3.0	5.0	1.7	versicolor
78	6.0	2.9	4.5	1.5	versicolor
79	5.7	2.6	3.5	1.0	versicolor
80	5.5	2.4	3.8	1.1	versicolor
81	5.5	2.4	3.7	1.0	versicolor
82	5.8	2.7	3.9	1.2	versicolor
83	6.0	2.7	5.1	1.6	versicolor
84	5.4	3.0	4.5	1.5	versicolor

	sepal_length	sepal_width	petal_length	petal_width	species
85	6.0	3.4	4.5	1.6	versicolor
86	6.7	3.1	4.7	1.5	versicolor
87	6.3	2.3	4.4	1.3	versicolor
88	5.6	3.0	4.1	1.3	versicolor
89	5.5	2.5	4.0	1.3	versicolor
90	5.5	2.6	4.4	1.2	versicolor
91	6.1	3.0	4.6	1.4	versicolor
92	5.8	2.6	4.0	1.2	versicolor
93	5.0	2.3	3.3	1.0	versicolor
94	5.6	2.7	4.2	1.3	versicolor
95	5.7	3.0	4.2	1.2	versicolor
96	5.7	2.9	4.2	1.3	versicolor
97	6.2	2.9	4.3	1.3	versicolor
98	5.1	2.5	3.0	1.1	versicolor
99	5.7	2.8	4.1	1.3	versicolor

In [13]: verginica

Out[13]:

	sepal_length	sepal_width	petal_length	petal_width	species
100	6.3	3.3	6.0	2.5	virginica
101	5.8	2.7	5.1	1.9	virginica
102	7.1	3.0	5.9	2.1	virginica
103	6.3	2.9	5.6	1.8	virginica
104	6.5	3.0	5.8	2.2	virginica
105	7.6	3.0	6.6	2.1	virginica
106	4.9	2.5	4.5	1.7	virginica
107	7.3	2.9	6.3	1.8	virginica
108	6.7	2.5	5.8	1.8	virginica
109	7.2	3.6	6.1	2.5	virginica
110	6.5	3.2	5.1	2.0	virginica
111	6.4	2.7	5.3	1.9	virginica
112	6.8	3.0	5.5	2.1	virginica
113	5.7	2.5	5.0	2.0	virginica
114	5.8	2.8	5.1	2.4	virginica
115	6.4	3.2	5.3	2.3	virginica
116	6.5	3.0	5.5	1.8	virginica
117	7.7	3.8	6.7	2.2	virginica
118	7.7	2.6	6.9	2.3	virginica
119	6.0	2.2	5.0	1.5	virginica
120	6.9	3.2	5.7	2.3	virginica
121	5.6	2.8	4.9	2.0	virginica
122	7.7	2.8	6.7	2.0	virginica
123	6.3	2.7	4.9	1.8	virginica
124	6.7	3.3	5.7	2.1	virginica
125	7.2	3.2	6.0	1.8	virginica
126	6.2	2.8	4.8	1.8	virginica
127	6.1	3.0	4.9	1.8	virginica
128	6.4	2.8	5.6	2.1	virginica
129	7.2	3.0	5.8	1.6	virginica
130	7.4	2.8	6.1	1.9	virginica
131	7.9	3.8	6.4	2.0	virginica
132	6.4	2.8	5.6	2.2	virginica
133	6.3	2.8	5.1	1.5	virginica
134	6.1	2.6	5.6	1.4	virginica

	sepal_length	sepal_width	petal_length	petal_width	species
135	7.7	3.0	6.1	2.3	virginica
136	6.3	3.4	5.6	2.4	virginica
137	6.4	3.1	5.5	1.8	virginica
138	6.0	3.0	4.8	1.8	virginica
139	6.9	3.1	5.4	2.1	virginica
140	6.7	3.1	5.6	2.4	virginica
141	6.9	3.1	5.1	2.3	virginica
142	5.8	2.7	5.1	1.9	virginica
143	6.8	3.2	5.9	2.3	virginica
144	6.7	3.3	5.7	2.5	virginica
145	6.7	3.0	5.2	2.3	virginica
146	6.3	2.5	5.0	1.9	virginica
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

In [14]: setosa.describe()

Out[14]:

	sepal_length	sepal_width	petal_length	petal_width
count	50.00000	50.000000	50.000000	50.000000
mean	5.00600	3.428000	1.462000	0.246000
std	0.35249	0.379064	0.173664	0.105386
min	4.30000	2.300000	1.000000	0.100000
25%	4.80000	3.200000	1.400000	0.200000
50%	5.00000	3.400000	1.500000	0.200000
75%	5.20000	3.675000	1.575000	0.300000
max	5.80000	4.400000	1.900000	0.600000

In [15]: versicolor.describe()

Out[15]:

	sepal_length	sepal_width	petal_length	petal_width
count	50.000000	50.000000	50.000000	50.000000
mean	5.936000	2.770000	4.260000	1.326000
std	0.516171	0.313798	0.469911	0.197753
min	4.900000	2.000000	3.000000	1.000000
25%	5.600000	2.525000	4.000000	1.200000
50%	5.900000	2.800000	4.350000	1.300000
75%	6.300000	3.000000	4.600000	1.500000
max	7.000000	3.400000	5.100000	1.800000

In [16]: virginica.describe()

Out[16]:

	sepal_length	sepal_width	petal_length	petal_width
count	50.000000	50.000000	50.000000	50.000000
mean	6.588000	2.974000	5.552000	2.026000
std	0.63588	0.322497	0.551895	0.27465
min	4.900000	2.200000	4.500000	1.400000
25%	6.225000	2.800000	5.100000	1.800000
50%	6.500000	3.000000	5.550000	2.000000
75%	6.900000	3.175000	5.875000	2.300000
max	7.900000	3.800000	6.900000	2.500000

```
In [28]: # group data by categorical variable
grouped_data = df.groupby('species')
# Calculate the summary statistics for each group
stats = grouped_data['sepal_length'].agg(['mean', 'median', 'min', 'max', 'std'])
stats
```

Out[28]:

	mean	median	min	max	std
species					
setosa	5.006	5.0	4.3	5.8	0.352490
versicolor	5.936	5.9	4.9	7.0	0.516171
virginica	6.588	6.5	4.9	7.9	0.635880

```
In [29]: stats = grouped_data['sepal_width'].agg(['mean', 'median', 'min', 'max', 'std'])
stats
```

Out[29]:

	mean	median	min	max	std
species					
setosa	3.428	3.4	2.3	4.4	0.379064
versicolor	2.770	2.8	2.0	3.4	0.313798
virginica	2.974	3.0	2.2	3.8	0.322497

```
In [30]: stats = grouped_data['petal_width'].agg(['mean', 'median', 'min', 'max', 'std'])
stats
```

Out[30]:

	mean	median	min	max	std
species					
setosa	0.246	0.2	0.1	0.6	0.105386
versicolor	1.326	1.3	1.0	1.8	0.197753
virginica	2.026	2.0	1.4	2.5	0.274650

```
In [31]: stats = grouped_data['petal_length'].agg(['mean', 'median', 'min', 'max', 'std'])
stats
```

Out[31]:

	mean	median	min	max	std
species					
setosa	1.462	1.50	1.0	1.9	0.173664
versicolor	4.260	4.35	3.0	5.1	0.469911
virginica	5.552	5.55	4.5	6.9	0.551895

In []: