```
import numpy as np
import pandas as pd
```

```
from google.colab import files
uploaded = files.upload()
```

Choose Files StudentsPe…ceTest1.xlsx

- **StudentsPerformanceTest1.xlsx**(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 9301 bytes, last modified: 2/14/2023 - 100% done

```
Saving StudentsPerformanceTest1.xlsx to StudentsPerformanceTest1.xlsx
```

```
df = pd.read_excel('StudentsPerformanceTest1.xlsx')
```

```
df
```

	gender	math score	reading score	writing score	Placement Score	placement offer count	Region
0	female	72	72	74.0	78.0	1	Pune
1	female	69	90	88.0	NaN	2	na
2	female	90	95	93.0	74.0	2	Nashik
3	male	47	57	NaN	78.0	1	Na
4	male	na	78	75.0	81.0	3	Pune
5	female	71	Na	78.0	70.0	4	na
6	male	12	44	52.0	12.0	2	Nashik
7	male	NaN	65	67.0	49.0	1	Pune
8	male	5	77	89.0	55.0	0	NaN

```
df_stats = df.describe()
```

```
df_stats
```

	writing score	Placement Score	placement offer count
count	8.000000	8.000000	9.000000
mean	77.000000	62.125000	1.777778
std	13.416408	23.295846	1.201850
min	52.000000	12.000000	0.000000
25%	72.250000	53.500000	1.000000
50%	76.500000	72.000000	2.000000
75%	88.250000	78.000000	2.000000
max	93.000000	81.000000	4.000000

```
from google.colab import files
uploaded = files.upload()
```

Choose Files IRIS.csv

- **IRIS.csv**(text/csv) - 4617 bytes, last modified: 3/16/2023 - 100% done

```
Saving IRIS.csv to IRIS.csv
```

```
df2 = pd.read_csv('IRIS.csv')
```

```
df2.describe()
```

sepal_length sepal_width petal_length petal_width

```
df3 = df2[df2['species'] == 'Iris-setosa']
```

```
df3
```

	sepal_length	sepal_width	petal_length	petal_width	species
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
```

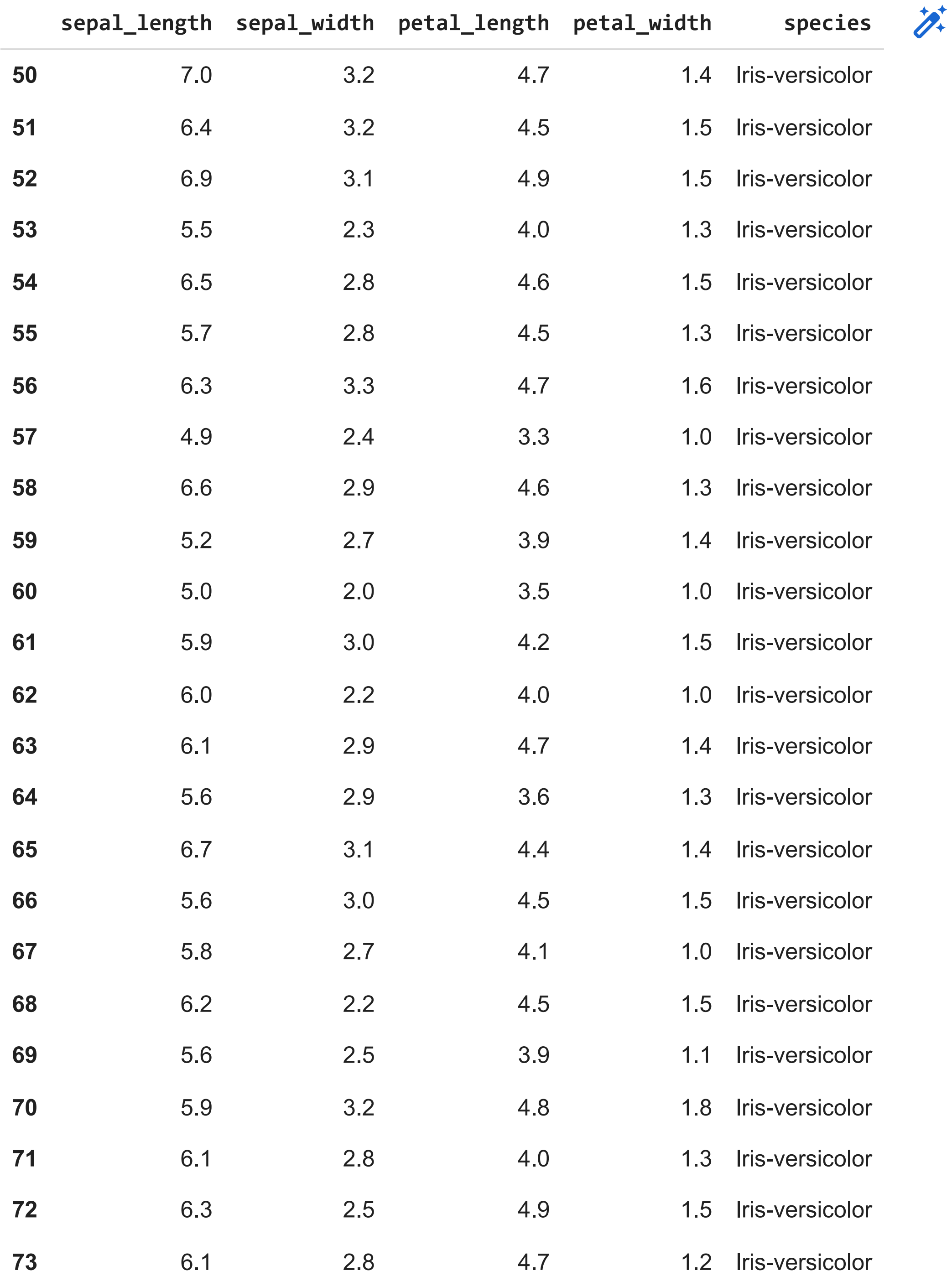
```
df3.describe()
```

	sepal_length	sepal_width	petal_length	petal_width
count	50.00000	50.000000	50.000000	50.00000
mean	5.00600	3.418000	1.464000	0.24400
std	0.35249	0.381024	0.173511	0.10721
min	4.30000	2.300000	1.000000	0.10000
25%	4.80000	3.125000	1.400000	0.20000
50%	5.00000	3.400000	1.500000	0.20000
75%	5.20000	3.675000	1.575000	0.30000
max	5.80000	4.400000	1.900000	0.60000

```
35    5.0    3.2    1.2    0.2  Iris-setosa
```

```
df4 = df2[df2['species'] == 'Iris-versicolor']
```

```
df4
```



	sepal_length	sepal_width	petal_length	petal_width	species
50	7.0	3.2	4.7	1.4	Iris-versicolor
51	6.4	3.2	4.5	1.5	Iris-versicolor
52	6.9	3.1	4.9	1.5	Iris-versicolor
53	5.5	2.3	4.0	1.3	Iris-versicolor
54	6.5	2.8	4.6	1.5	Iris-versicolor
55	5.7	2.8	4.5	1.3	Iris-versicolor
56	6.3	3.3	4.7	1.6	Iris-versicolor
57	4.9	2.4	3.3	1.0	Iris-versicolor
58	6.6	2.9	4.6	1.3	Iris-versicolor
59	5.2	2.7	3.9	1.4	Iris-versicolor
60	5.0	2.0	3.5	1.0	Iris-versicolor
61	5.9	3.0	4.2	1.5	Iris-versicolor
62	6.0	2.2	4.0	1.0	Iris-versicolor
63	6.1	2.9	4.7	1.4	Iris-versicolor
64	5.6	2.9	3.6	1.3	Iris-versicolor
65	6.7	3.1	4.4	1.4	Iris-versicolor
66	5.6	3.0	4.5	1.5	Iris-versicolor
67	5.8	2.7	4.1	1.0	Iris-versicolor
68	6.2	2.2	4.5	1.5	Iris-versicolor
69	5.6	2.5	3.9	1.1	Iris-versicolor
70	5.9	3.2	4.8	1.8	Iris-versicolor
71	6.1	2.8	4.0	1.3	Iris-versicolor
72	6.3	2.5	4.9	1.5	Iris-versicolor
73	6.1	2.8	4.7	1.2	Iris-versicolor

```
74    6.4    2.9    4.3    1.3  Iris-versicolor
```

```
df4.describe()
```

	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

```
85    6.0    3.4    4.5    1.6  Iris-versicolor
```

```
df5 = df2[df2['species'] == 'Iris-virginica']
```

```
df5
```

	sepal_length	sepal_width	petal_length	petal_width	species
100	6.3	3.3	6.0	2.5	Iris-virginica
101	5.8	2.7	5.1	1.9	Iris-virginica
102	7.1	3.0	5.9	2.1	Iris-virginica
103	6.3	2.9	5.6	1.8	Iris-virginica
104	6.5	3.0	5.8	2.2	Iris-virginica
105	7.6	3.0	6.6	2.1	Iris-virginica
106	4.9	2.5	4.5	1.7	Iris-virginica
107	7.3	2.9	6.3	1.8	Iris-virginica
108	6.7	2.5	5.8	1.8	Iris-virginica
109	7.2	3.6	6.1	2.5	Iris-virginica
110	6.5	3.2	5.1	2.0	Iris-virginica
111	6.4	2.7	5.3	1.9	Iris-virginica
112	6.8	3.0	5.5	2.1	Iris-virginica
113	5.7	2.5	5.0	2.0	Iris-virginica
114	5.8	2.8	5.1	2.4	Iris-virginica
115	6.4	3.2	5.3	2.3	Iris-virginica
116	6.5	3.0	5.5	1.8	Iris-virginica
117	7.7	3.8	6.7	2.2	Iris-virginica
118	7.7	2.6	6.9	2.3	Iris-virginica
119	6.0	2.2	5.0	1.5	Iris-virginica
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

```
df5.describe()
```

	sepal_length	sepal_width	petal_length	petal_width
count	50.00000	50.000000	50.000000	50.00000
mean	6.58800	2.974000	5.552000	2.02600
std	0.63588	0.322497	0.551895	0.27465
min	4.90000	2.200000	4.500000	1.40000
25%	6.22500	2.800000	5.100000	1.80000
50%	6.50000	3.000000	5.550000	2.00000
75%	6.90000	3.175000	5.875000	2.30000
max	7.90000	3.800000	6.900000	2.50000

135 7.7 3.0 6.1 2.3 Iris-virginica

```
import math
```

```
df
```

	gender	math score	reading score	writing score	Placement Score	placement offer count	Region
0	female	72	72	74.0	78.0	1	Pune
1	female	69	90	88.0	NaN	2	na
2	female	90	95	93.0	74.0	2	Nashik
3	male	47	57	NaN	78.0	1	Na
4	male	na	78	75.0	81.0	3	Pune
5	female	71	Na	78.0	70.0	4	na
6	male	12	44	52.0	12.0	2	Nashik
7	male	NaN	65	67.0	49.0	1	Pune
8	male	5	77	89.0	55.0	0	NaN

140 5.9 3.0 5.1 1.8 Iris-virginica

149	5.9	3.0	5.1	1.8	Iris-virginica

```
sum_values = df['writing score'].sum()
```

```
count_values = df['writing score'].count()
```

```
mean = sum_values/count_values
```

```
mean
```

```
77.0
```

```
np.sort(df['writing score'])
```

```
array([52., 67., 74., 75., 78., 88., 89., 93., nan])
```

```
middle_index = math.floor(count_values/2)
```

```
median = np.sort(df['writing score']) [middle_index]
```

```
median
```

```
78.0
```

```
max_value = np.sort(df['writing score']) [-1]
```

```
max_value
```

```
nan
```

```
min_value = np.sort(df['writing score']) [0]
```

```
min_value
```

```
52.0
```

```
def max_value(variable) :
  return np.sort(df[variable])[-1]
```

```
max_value('writing score')
```

```
     nan
```

```
def min_value(variable) :
  return np.sort(df[variable])[0]
```

```
min_value('writing score')
```

```
     52.0
```

```
df['writing score'].std()
```

```
     13.416407864998739
```

```
def std_dev (variable):
  lst=[]
  for i in range (count_values):
    sigma = (df[variable] [i]-mean) **2
    lst.append(sigma)
  return np.sqrt(sum(lst)/count_values)
```

```
std_dev('writing score')
```

```
     nan
```

```
def variance (data, ddof=0):
  n = len (data)
  mean = sum (data) / n
  return sum ((x - mean) ** 2 for x in data)/(n-ddof)
```

```
variance(df['writing score'])
```

```
nan
```

```
df.describe()
```

	writing score	Placement Score	placement offer count
count	8.000000	8.000000	9.000000
mean	77.000000	62.125000	1.777778
std	13.416408	23.295846	1.201850
min	52.000000	12.000000	0.000000
25%	72.250000	53.500000	1.000000
50%	76.500000	72.000000	2.000000
75%	88.250000	78.000000	2.000000
max	93.000000	81.000000	4.000000

✓ 0s completed at 12:51 AM