In [41]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [42]:
```
data = pd.read_csv('data.csv')
print(data)
```

```
       age     workclass  fnlwgt     education  educational-num  \
0       25       Private  226802          11th                7
1       38       Private   89814       HS-grad                9
2       28     Local-gov  336951    Assoc-acdm               12
3       44       Private  160323  Some-college               10
4       18             ?  103497  Some-college               10
...    ...           ...     ...           ...              ...
48837   27       Private  257302    Assoc-acdm               12
48838   40       Private  154374       HS-grad                9
48839   58       Private  151910       HS-grad                9
48840   22       Private  201490       HS-grad                9
48841   52  Self-emp-inc  287927       HS-grad                9

           marital-status         occupation relationship   race  gender  \
0           Never-married  Machine-op-inspct    Own-child  Black    Male
1      Married-civ-spouse    Farming-fishing      Husband  White    Male
2      Married-civ-spouse    Protective-serv      Husband  White    Male
3      Married-civ-spouse  Machine-op-inspct      Husband  Black    Male
4           Never-married                  ?    Own-child  White  Female
...                   ...                ...          ...    ...     ...
48837  Married-civ-spouse       Tech-support         Wife  White  Female
48838  Married-civ-spouse  Machine-op-inspct      Husband  White    Male
48839             Widowed       Adm-clerical    Unmarried  White  Female
48840       Never-married       Adm-clerical    Own-child  White    Male
48841  Married-civ-spouse    Exec-managerial         Wife  White  Female

       capital-gain  capital-loss  hours-per-week native-country income
0                 0             0              40  United-States  <=50K
1                 0             0              50  United-States  <=50K
2                 0             0              40  United-States   >50K
3              7688             0              40  United-States   >50K
4                 0             0              30  United-States  <=50K
...             ...           ...             ...            ...    ...
48837             0             0              38  United-States  <=50K
48838             0             0              40  United-States   >50K
48839             0             0              40  United-States  <=50K
48840             0             0              20  United-States  <=50K
48841         15024             0              40  United-States   >50K

[48842 rows x 15 columns]
```

In [43]:
```
print(data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 48842 entries, 0 to 48841
Data columns (total 15 columns):
 #   Column           Non-Null Count  Dtype 
---  ------           --------------  ----- 
 0   age              48842 non-null  int64 
 1   workclass        48842 non-null  object
 2   fnlwgt           48842 non-null  int64 
 3   education        48842 non-null  object
 4   educational-num  48842 non-null  int64 
 5   marital-status   48842 non-null  object
 6   occupation       48842 non-null  object
 7   relationship     48842 non-null  object
 8   race             48842 non-null  object
 9   gender           48842 non-null  object
 10  capital-gain     48842 non-null  int64 
 11  capital-loss     48842 non-null  int64 
 12  hours-per-week   48842 non-null  int64 
 13  native-country   48842 non-null  object
 14  income           48842 non-null  object
dtypes: int64(6), object(9)
memory usage: 5.6+ MB
None
```

In [44]:
```
print(data.describe())
```

```
                age        fnlwgt  educational-num  capital-gain  \
count  48842.000000  4.884200e+04     48842.000000  48842.000000   
mean      38.643585  1.896641e+05        10.078089   1079.067626   
std       13.710510  1.056040e+05         2.570973   7452.019058   
min       17.000000  1.228500e+04         1.000000      0.000000   
25%       28.000000  1.175505e+05         9.000000      0.000000   
50%       37.000000  1.781445e+05        10.000000      0.000000   
75%       48.000000  2.376420e+05        12.000000      0.000000   
max       90.000000  1.490400e+06        16.000000  99999.000000   

       capital-loss  hours-per-week  
count  48842.000000    48842.000000  
mean      87.502314       40.422382  
std      403.004552       12.391444  
min        0.000000        1.000000  
25%        0.000000       40.000000  
50%        0.000000       40.000000  
75%        0.000000       45.000000  
max     4356.000000       99.000000  
```

In [45]:
```
print(data['age'].mean())
```

```
38.64358543876172
```

In [46]:
```
print(data['age'].min())
```

```
17
```

In [47]:
```
print(data['age'].max())
```

```
90
```

In [48]:
```
print(data['age'].median())
```

```
37.0
```

In [49]:
```
print(data['age'].std())
```

```
13.710509934443557
```

In [50]:
```
print(data['age'].quantile(0.25))
```

```
28.0
```

In [51]:
```
print(data['age'].quantile(0.75))
```

```
48.0
```

In [52]:
```
print(data['income'].nunique())
```

```
2
```

In [53]:
```
print(data.groupby(['income','age']).count())
```

```
              workclass  fnlwgt  education  educational-num  marital-status  \
income age
<=50K  17           595     595        595              595             595
       18           862     862        862              862             862
       19          1050    1050       1050             1050            1050
       20          1112    1112       1112             1112            1112
       21          1090    1090       1090             1090            1090
...                 ...     ...        ...              ...             ...
>50K   83             2       2          2                2               2
       84             1       1          1                1               1
       85             1       1          1                1               1
       88             1       1          1                1               1
       90            13      13         13               13              13

              occupation  relationship  race  gender  capital-gain  \
income age
<=50K  17            595           595   595     595           595
       18            862           862   862     862           862
       19           1050          1050  1050    1050          1050
       20           1112          1112  1112    1112          1112
       21           1090          1090  1090    1090          1090
...                  ...           ...   ...     ...           ...
>50K   83              2             2     2       2             2
       84              1             1     1       1             1
       85              1             1     1       1             1
       88              1             1     1       1             1
       90             13            13    13      13            13

              capital-loss  hours-per-week  native-country
income age
<=50K  17             595             595             595
       18             862             862             862
       19            1050            1050            1050
       20            1112            1112            1112
       21            1090            1090            1090
...                   ...             ...             ...
>50K   83               2               2               2
       84               1               1               1
       85               1               1               1
       88               1               1               1
       90              13              13              13

[142 rows x 13 columns]
```

In [54]:
```
print(data.groupby(['income','age']).max())
```

```
                    workclass   fnlwgt     education  educational-num  \
income age
<=50K  17          State-gov   806316  Some-college               10
       18          State-gov   761006  Some-college               14
       19       Without-pay  1047822  Some-college               13
       20          State-gov   745817  Some-college               14
       21       Without-pay   811615  Some-college               14
...                     ...      ...           ...              ...
>50K   83       Self-emp-inc   240150     Bachelors               13
       84       Self-emp-inc   172907  Some-college               10
       85       Self-emp-inc   155981     Bachelors               13
       88   Self-emp-not-inc   263569          11th                7
       90   Self-emp-not-inc   313986   Prof-school               15

                 marital-status        occupation   relationship   race gender  \
income age
<=50K  17              Widowed  Transport-moving      Unmarried  White   Male
       18              Widowed  Transport-moving           Wife  White   Male
       19            Separated  Transport-moving           Wife  White   Male
       20            Separated  Transport-moving           Wife  White   Male
       21              Widowed  Transport-moving           Wife  White   Male
...                        ...               ...            ...    ...    ...
>50K   83   Married-civ-spouse   Farming-fishing        Husband  White   Male
       84   Married-civ-spouse             Sales        Husband  White   Male
       85              Widowed   Exec-managerial  Not-in-family  White   Male
       88   Married-civ-spouse   Farming-fishing        Husband  White   Male
       90        Never-married             Sales           Wife  White   Male

             capital-gain  capital-loss  hours-per-week native-country
income age
<=50K  17           34095          1721              50  United-States
       18           34095          1721              72  United-States
       19           34095          2129              99     Yugoslavia
       20           34095          2258              84     Yugoslavia
       21            4865          2603              98        Vietnam
...                   ...           ...             ...            ...
>50K   83           20051          2392              55  United-States
       84               0             0              35  United-States
       85               0             0              40  United-States
       88            6418             0              40  United-States
       90           20051          1825              72  United-States

[142 rows x 13 columns]
```

In [55]:
```
print(data.groupby('income')['age'].min())
```

```
income
<=50K    17
>50K     19
Name: age, dtype: int64
```

In [56]:
```
print(data.groupby('income')['age'].mean())
```

```
income
<=50K    36.872184
>50K     44.275178
Name: age, dtype: float64
```

In [57]:
```
print(data.groupby('income')['age'].std())
```

```
income
<=50K    14.104118
>50K     10.558983
Name: age, dtype: float64
```

In [58]:
```
print(data.groupby('income')['age'].median())
```

```
income
<=50K    34.0
>50K     43.0
Name: age, dtype: float64
```

In [59]:
```
def calc_mean(data) :
    if len(data)==0:
        return None
    else:
        return (sum(data)/len(data))


def calc_std(data, mean):
    if len(data)==0:
        return None
    else:
        squareDiff = sum((x - mean)**2 for x in data)
        return (squareDiff/(len(data)-1))**0.5

def calc_percentile(data, percentile):
    sorted_data = sorted(data)
    index = int(percentile*len(data))
    return sorted_data[index]

def display_stat(df) :

    print("Mean :")
    mea = calc_mean(df)
    print(mea)

    print("\nStandard Deviation :")
    std = calc_std(df, mea)
    print(std)

    percent = [0.25, 0.5, 0.75]
    # print(f'\n{col} : ')
    for percentile in percent:
        print(f'\n{percentile} percentile: {calc_percentile(df, percentile)}')

data_age = data['age']
display_stat(data_age)
```

```
Mean :
38.64358543876172

Standard Deviation :
13.710509934443557

0.25 percentile: 28

0.5 percentile: 37

0.75 percentile: 48
```

In [60]:
```
iris_data = sns.load_dataset('iris')
print(iris_data)

print(iris_data.describe())

print(iris_data.groupby('species').count())

setosa_data = iris_data[iris_data['species']=='setosa']
print("**********setosa_data**********\n\n",setosa_data)

versicolor_data = iris_data[iris_data['species']=='versicolor']
print("\n\n\n**********versicolor_data**********\n\n",versi

virginica_data = iris_data[iris_data['species']=='virginica']
print("\n\n\n**********virginica_data**********\n\n",virgin

print("**********setosa_data**********\n\n",setosa_data.des
print("\n\n\n**********versicolor_data**********\n\n",versi
print("\n\n\n**********virginica_data**********\n\n",virgin
```

```
     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 x 5 columns]
       sepal_length  sepal_width  petal_length  petal_width
count    150.000000   150.000000    150.000000   150.000000
mean       5.843333     3.057333      3.758000     1.199333
std        0.828066     0.435866      1.765298     0.762238
min        4.300000     2.000000      1.000000     0.100000
25%        5.100000     2.800000      1.600000     0.300000
50%        5.800000     3.000000      4.350000     1.300000
75%        6.400000     3.300000      5.100000     1.800000
max        7.900000     4.400000      6.900000     2.500000
            sepal_length  sepal_width  petal_length  petal_width
species
setosa                50           50            50           50
versicolor            50           50            50           50
virginica             50           50            50           50
**********************setosa_data**********************

     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
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

**********versicolor_data**********

	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
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

**********virginica_data**********

	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
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

**********setosa_data**********

	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

**********versicolor_data**********

	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

**********virginica_data**********

	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
```

In [61]:

```
# Using built-in functions

def display_stat(species_data, species_name) :
    nc = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width']
    print(f"Statistics for {species_name}\n")
    print("Mean :")
    print(species_data[nc].mean())
    print("Min :")
    print(species_data[nc].min())
    print("Max :")
    print(species_data[nc].max())
    print("Standard Deviation :")
    print(species_data[nc].std())
    print("25% :")
    print(species_data[nc].quantile(0.25))
    print("50% :")
    print(species_data[nc].quantile(0.50))
    print("75% :")
    print(species_data[nc].quantile(0.75))


# Filtering the data
setosa_data = iris_data[iris_data['species']=='setosa']
display_stat(setosa_data, 'Iris-setosa')
print("\n\n")

versicolor_data = iris_data[iris_data['species']=='versicolor']
display_stat(versicolor_data, 'Iris-versicolor')
print("\n\n")

virginica_data = iris_data[iris_data['species']=='virginica']
display_stat(virginica_data, 'Iris-virginica')
```

```
Statistics for Iris-setosa

Mean :
sepal_length    5.006
sepal_width     3.428
petal_length    1.462
petal_width     0.246
dtype: float64
Min :
sepal_length    4.3
sepal_width     2.3
petal_length    1.0
petal_width     0.1
dtype: float64
Max :
sepal_length    5.8
sepal_width     4.4
petal_length    1.9
petal_width     0.6
dtype: float64
Standard Deviation :
sepal_length    0.352490
sepal_width     0.379064
petal_length    0.173664
petal_width     0.105386
dtype: float64
25% :
sepal_length    4.8
sepal_width     3.2
petal_length    1.4
petal_width     0.2
Name: 0.25, dtype: float64
50% :
sepal_length    5.0
sepal_width     3.4
petal_length    1.5
petal_width     0.2
Name: 0.5, dtype: float64
75% :
sepal_length    5.200
sepal_width     3.675
petal_length    1.575
petal_width     0.300
Name: 0.75, dtype: float64



Statistics for Iris-versicolor

Mean :
sepal_length    5.936
sepal_width     2.770
petal_length    4.260
petal_width     1.326
dtype: float64
Min :
```

```
sepal_length    4.9
sepal_width     2.0
petal_length    3.0
petal_width     1.0
dtype: float64
Max :
sepal_length    7.0
sepal_width     3.4
petal_length    5.1
petal_width     1.8
dtype: float64
Standard Deviation :
sepal_length    0.516171
sepal_width     0.313798
petal_length    0.469911
petal_width     0.197753
dtype: float64
25% :
sepal_length    5.600
sepal_width     2.525
petal_length    4.000
petal_width     1.200
Name: 0.25, dtype: float64
50% :
sepal_length    5.90
sepal_width     2.80
petal_length    4.35
petal_width     1.30
Name: 0.5, dtype: float64
75% :
sepal_length    6.3
sepal_width     3.0
petal_length    4.6
petal_width     1.5
Name: 0.75, dtype: float64



Statistics for Iris-virginica

Mean :
sepal_length    6.588
sepal_width     2.974
petal_length    5.552
petal_width     2.026
dtype: float64
Min :
sepal_length    4.9
sepal_width     2.2
petal_length    4.5
petal_width     1.4
dtype: float64
Max :
sepal_length    7.9
sepal_width     3.8
petal_length    6.9
```

```
petal_width     2.5
dtype: float64
Standard Deviation :
sepal_length    0.635880
sepal_width     0.322497
petal_length    0.551895
petal_width     0.274650
dtype: float64
25% :
sepal_length    6.225
sepal_width     2.800
petal_length    5.100
petal_width     1.800
Name: 0.25, dtype: float64
50% :
sepal_length    6.50
sepal_width     3.00
petal_length    5.55
petal_width     2.00
Name: 0.5, dtype: float64
75% :
sepal_length    6.900
sepal_width     3.175
petal_length    5.875
petal_width     2.300
Name: 0.75, dtype: float64
```

In [62]:

```
# Using user-defined functions

def calc_mean(data) :
    if len(data)==0:
        return None
    else:
        return (sum(data)/len(data))


def calc_std(data, mean):
    if len(data)==0:
        return None
    else:
        squareDiff = sum((x - mean)**2 for x in data)
        return (squareDiff/(len(data)-1))**0.5

def calc_percentile(data, percentile):
    sorted_data = sorted(data)
    index = int(percentile*len(data))
    return sorted_data[index]

def display_stat(species_data, species_name) :
    nc = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width']
    print(f"********Statistics for {species_name}********\n")

    print("Mean :\n")
    mea = [calc_mean(species_data[col]) for col in nc]
    print(pd.Series(mea,index=nc))
```

```
    print("\nStandard Deviation :\n")
    std = [calc_std(species_data[col], mea[i]) for i,col in enumerate(nc)]
    print(pd.Series(std,index=nc))

    percent = [0.25, 0.5, 0.75]
    for col in nc:
        print(f'\n{col} : ')
        for percentile in percent:
            print(f'{percentile} percentile: {calc_percentile(species_data[col], pe


display_stat(setosa_data, 'Iris-setosa')
print("\n")

display_stat(versicolor_data, 'Iris-versicolor')
print("\n")

display_stat(virginica_data, 'Iris-virginica')
```

```
********Statistics for Iris-setosa********

Mean :

sepal_length    5.006
sepal_width     3.428
petal_length    1.462
petal_width     0.246
dtype: float64

Standard Deviation :

sepal_length    0.352490
sepal_width     0.379064
petal_length    0.173664
petal_width     0.105386
dtype: float64

sepal_length :
0.25 percentile: 4.8
0.5 percentile: 5.0
0.75 percentile: 5.2

sepal_width :
0.25 percentile: 3.2
0.5 percentile: 3.4
0.75 percentile: 3.7

petal_length :
0.25 percentile: 1.4
0.5 percentile: 1.5
0.75 percentile: 1.6

petal_width :
0.25 percentile: 0.2
0.5 percentile: 0.2
0.75 percentile: 0.3


********Statistics for Iris-versicolor********

Mean :

sepal_length    5.936
sepal_width     2.770
petal_length    4.260
petal_width     1.326
dtype: float64

Standard Deviation :

sepal_length    0.516171
sepal_width     0.313798
petal_length    0.469911
petal_width     0.197753
dtype: float64
```

```
sepal_length :
0.25 percentile: 5.6
0.5 percentile: 5.9
0.75 percentile: 6.3

sepal_width :
0.25 percentile: 2.5
0.5 percentile: 2.8
0.75 percentile: 3.0

petal_length :
0.25 percentile: 4.0
0.5 percentile: 4.4
0.75 percentile: 4.6

petal_width :
0.25 percentile: 1.2
0.5 percentile: 1.3
0.75 percentile: 1.5


********Statistics for Iris-virginica********

Mean :

sepal_length    6.588
sepal_width     2.974
petal_length    5.552
petal_width     2.026
dtype: float64

Standard Deviation :

sepal_length    0.635880
sepal_width     0.322497
petal_length    0.551895
petal_width     0.274650
dtype: float64

sepal_length :
0.25 percentile: 6.2
0.5 percentile: 6.5
0.75 percentile: 6.9

sepal_width :
0.25 percentile: 2.8
0.5 percentile: 3.0
0.75 percentile: 3.2

petal_length :
0.25 percentile: 5.1
0.5 percentile: 5.6
0.75 percentile: 5.9

petal_width :
```

```
0.25 percentile: 1.8
0.5 percentile: 2.0
0.75 percentile: 2.3
```