```
import pandas as pd
import numpy as np

df = pd.read_csv("Downloads/Employee_Salary_Dataset.csv")

df.head(10)
```

	ID	Experience_Years	Age	Gender	Salary
0	1	5	28	Female	250000
1	2	1	21	Male	50000
2	3	3	23	Female	170000
3	4	2	22	Male	25000
4	5	1	17	Male	10000
5	6	25	62	Male	5001000
6	7	19	54	Female	800000
7	8	2	21	Female	9000
8	9	10	36	Female	61500
9	10	15	54	Female	650000

```
df.tail(5)
```

	ID	Experience_Years	Age	Gender	Salary
30	31	10	34	Male	80000
31	32	15	54	Male	900000
32	33	20	55	Female	1540000
33	34	19	53	Female	9300000
34	35	16	49	Male	7600000

```
df.mean()
```

```
ID                  1.800000e+01
Experience_Years    9.200000e+00
Age                 3.548571e+01
Salary              2.059147e+06
dtype: float64
```

```
df.loc[:,'Age'].mean()
```

```
35.48571428571429
```

```
df.loc[:,'Salary'].mean()
```

```
2059147.142857143
```

```
df.median()
```

```
ID                     18.0
Experience_Years        6.0
Age                    29.0
Salary              250000.0
dtype: float64
```

```
df.mode()
```

```
    ID  Experience_Years   Age  Gender    Salary
0    1               2.0  54.0  Female   25000.0
1    2               NaN   NaN     NaN  250000.0
2    3               NaN   NaN     NaN       NaN
3    4               NaN   NaN     NaN       NaN
4    5               NaN   NaN     NaN       NaN
5    6               NaN   NaN     NaN       NaN
6    7               NaN   NaN     NaN       NaN
7    8               NaN   NaN     NaN       NaN
8    9               NaN   NaN     NaN       NaN
9   10               NaN   NaN     NaN       NaN
10  11               NaN   NaN     NaN       NaN
11  12               NaN   NaN     NaN       NaN
12  13               NaN   NaN     NaN       NaN
13  14               NaN   NaN     NaN       NaN
14  15               NaN   NaN     NaN       NaN
15  16               NaN   NaN     NaN       NaN
16  17               NaN   NaN     NaN       NaN
17  18               NaN   NaN     NaN       NaN
18  19               NaN   NaN     NaN       NaN
19  20               NaN   NaN     NaN       NaN
20  21               NaN   NaN     NaN       NaN
21  22               NaN   NaN     NaN       NaN
22  23               NaN   NaN     NaN       NaN
23  24               NaN   NaN     NaN       NaN
24  25               NaN   NaN     NaN       NaN
25  26               NaN   NaN     NaN       NaN
26  27               NaN   NaN     NaN       NaN
27  28               NaN   NaN     NaN       NaN
28  29               NaN   NaN     NaN       NaN
29  30               NaN   NaN     NaN       NaN
30  31               NaN   NaN     NaN       NaN
31  32               NaN   NaN     NaN       NaN
32  33               NaN   NaN     NaN       NaN
33  34               NaN   NaN     NaN       NaN
34  35               NaN   NaN     NaN       NaN
```

```
df.mean(axis=1)[0:4]
```

```
0    62508.50
1    12506.00
2    42507.25
3     6257.00
dtype: float64
```

```
df.mean(axis = 1)[0:4]
```

```
0    62508.50
1    12506.00
2    42507.25
```

```
3     6257.00
dtype: float64
```

```
df.loc[:,'Age'].median()
```

```
29.0
```

```
df.loc[:,'Age'].mode()
```

```
0    54
dtype: int64
```

```
df.min()
```

```
ID                         1
Experience_Years           1
Age                       17
Gender                Female
Salary                  3000
dtype: object
```

```
df.max()
```

```
ID                           35
Experience_Years             27
Age                          62
Gender                     Male
Salary                 10000000
dtype: object
```

```
df.loc[:,'Age'].min()
```

```
17
```

```
df.loc[:,'Age'].max()
```

```
62
```

```
df.loc[:,'Age'].min(skipna = False)
```

```
17
```

```
df.loc[:,'Age'].min(skipna = True)
```

```
17
```

```
df.std()
```

```
ID                   1.024695e+01
Experience_Years     7.552950e+00
Age                  1.464355e+01
Salary               3.170124e+06
dtype: float64
```

```
df.loc[:,'Age'].std()
```

```
14.643551940884361
```

```
df.std(axis=1)[0:4]
```

```
0    124994.333900
1     24996.001694
2     84995.167190
3     12495.336570
dtype: float64
```

```
df.groupby(['Salary'])['Age'].mean()
```

```
Salary
3000        18
6000        21
6100        21
7500        23
8900        23
9000        21
10000       17
15000       21
20000       22
25000       24
50000       21
61500       36
80000       34
87000       27
170000      23
220100      40
250000      27
330000      36
650000      54
800000      54
900000      54
930000      34
1400000     29
1540000     55
5000000     54
5001000     62
6000050     39
6570000     54
6845000     29
7600000     49
7900000     54
9300000     53
10000000    62
Name: Age, dtype: int64
```

```
df.groupby(['Gender'])['Age'].mean()
```

```
Gender
Female    37.111111
Male      33.764706
Name: Age, dtype: float64
```

```
df_u=df.rename(columns= {'Salary':'Income'},inplace=True)
```

```
df.head()
```

```
   ID  Experience_Years  Age  Gender  Income
0   1                 5   28  Female  250000
1   2                 1   21    Male   50000
2   3                 3   23  Female  170000
3   4                 2   22    Male   25000
4   5                 1   17    Male   10000
```

```
csv_url =
'https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.d
ata'
```

```
col_names =
['Sepal_Length','Sepal_Width','Petal_Length','Petal_Width','Species']
```

```
iris = pd.read_csv(csv_url, names = col_names)
```

```
irisSet = (iris['Species']== 'Iris-setosa')
print('Iris-setosa')
print(iris[irisSet].describe())
```

```
Iris-setosa
       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
```

```
irisVer = (iris['Species']== 'Iris-versicolor')
```

```
print('Iris-versicolor')
print(iris[irisVer].describe())
```

```
Iris-versicolor
       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
```

```
print('Iris-versicolor')
```

```
Iris-versicolor
```

```
print(iris[irisVer].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
```

```
irisVir = (iris['Species']== 'Iris-virginica')
```

```
print('Iris-virginica')
```

```
Iris-virginica
```

```
print(iris[irisVir].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
```