```
import pandas as pd
```

```
s=pd.Series([1,2,3,4,5,6,7,8])
```

```
type(s)
```

```
pandas.core.series.Series
```

```
print(s)
```

```
0    1
1    2
2    3
3    4
4    5
5    6
6    7
7    8
dtype: int64
```

```
print(s)
```

```
0    1
1    2
2    3
3    4
4    5
5    6
6    7
7    8
dtype: int64
```

```
for i in s:
    print(i)
```

```
1
2
3
4
5
6
7
8
```

```
for i in range(0,len(s)):
   print("at",i,"we have",s[i])
```

```
at 0 we have 1
at 1 we have 2
at 2 we have 3
at 3 we have 4
at 4 we have 5
at 5 we have 6
```

```
at 6 we have 7
at 7 we have 8
```

```
l=[1,2,3,4]
l*2
```

```
[1, 2, 3, 4, 1, 2, 3, 4]
```

```
s
```

```
0    1
1    2
2    3
3    4
4    5
5    6
6    7
7    8
dtype: int64
```

```
s-2
```

```
0   -1
1    0
2    1
3    2
4    3
5    4
6    5
7    6
dtype: int64
```

```
s=s*10
```

```
print(s)
```

```
0    10
1    20
2    30
3    40
4    50
5    60
6    70
7    80
dtype: int64
```

```
s[-1]#not supported
```

```
len(s)
```

```
8
```

```
min(s)
```

```
10
```

```
max(s)
```

```
80
```

```
sum(s)
```

```
360
```

```
s=pd.Series([1,2,3,4])
s2=pd.Series([10,20,30,40,50])
s3=s+s2
```

```
print(s3)
```

```
0    11.0
1    22.0
2    33.0
3    44.0
4     NaN
dtype: float64
```

```
s[0:4]
```

```
0    1
1    2
2    3
3    4
dtype: int64
```

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

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving mycsv.csv to mycsv.csv

```
import pandas as pd
import io
df=pd.read_csv(io.BytesIO(uploaded['mycsv.csv']))
print(df)
```

```
    id   name  gender   salary
0    1   aaaa    male    10000
1    2   bbbb  female     8000
2    3  ccccc    male   120000
3    4   dddd    male    45000
4    5   eeee  female   334567
5    6   ffff  female    21234
6    7   gggg  female     2345
7    8   hhhh  female    23456
8    9   iiii    male     7654
```

```
9   10    jjj   male   3456578
10  11   kkkk    male    45678
11  12   llll    male      345
12  13   mmmm  female     9876
13  14  nnnnn    male    34567
14  15   jjjj    male    87654
15  16   iiii  female    34567
16  17   mmmm  female     1234
17  18  nnnnn  female    23456
18  19   oooo  female     5555
19  20   pppp    male     3455
```

```
df
```

| | id | name | gender | salary |
|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 |
| **1** | 2 | bbbb | female | 8000 |
| **2** | 3 | ccccc | male | 120000 |
| **3** | 4 | dddd | male | 45000 |
| **4** | 5 | eeee | female | 334567 |
| **5** | 6 | ffff | female | 21234 |
| **6** | 7 | gggg | female | 2345 |
| **7** | 8 | hhhh | female | 23456 |
| **8** | 9 | iiii | male | 7654 |
| **9** | 10 | jjj | male | 3456578 |
| **10** | 11 | kkkk | male | 45678 |
| **11** | 12 | llll | male | 345 |
| **12** | 13 | mmmm | female | 9876 |
| **13** | 14 | nnnnn | male | 34567 |
| **14** | 15 | jjjj | male | 87654 |
| **15** | 16 | iiii | female | 34567 |
| **16** | 17 | mmmm | female | 1234 |
| **17** | 18 | nnnnn | female | 23456 |
| **18** | 19 | oooo | female | 5555 |
| **19** | 20 | pppp | male | 3455 |

```
df.name
```

```
0      aaaa
1      bbbb
2     ccccc
3      dddd
4      eeee
```

```
5      ffff
6      gggg
7      hhhh
8      iiii
9       jjj
10     kkkk
11     llll
12     mmmm
13    nnnnn
14     jjjj
15     iiii
16     mmmm
17    nnnnn
18     oooo
19     pppp
Name: name, dtype: object
```

```
print(df.name,df.salary)
```

```
0      aaaa
1      bbbb
2     ccccc
3      dddd
4      eeee
5      ffff
6      gggg
7      hhhh
8      iiii
9       jjj
10     kkkk
11     llll
12     mmmm
13    nnnnn
14     jjjj
15     iiii
16     mmmm
17    nnnnn
18     oooo
19     pppp
Name: name, dtype: object 0       10000
1        8000
2      120000
3       45000
4      334567
5       21234
6        2345
7       23456
8        7654
9     3456578
10      45678
11        345
12       9876
13      34567
14      87654
15      34567
16       1234
17      23456
18       5555
19       3455
Name: salary, dtype: int64
```

```
df[['salary','name']]
```

| | salary | name |
|---|---|---|
| **0** | 10000 | aaaa |
| **1** | 8000 | bbbb |
| **2** | 120000 | ccccc |
| **3** | 45000 | dddd |
| **4** | 334567 | eeee |
| **5** | 21234 | ffff |
| **6** | 2345 | gggg |
| **7** | 23456 | hhhh |
| **8** | 7654 | iiii |
| **9** | 3456578 | jjj |
| **10** | 45678 | kkkk |
| **11** | 345 | llll |
| **12** | 9876 | mmmm |
| **13** | 34567 | nnnnn |
| **14** | 87654 | jjjj |
| **15** | 34567 | iiii |
| **16** | 1234 | mmmm |
| **17** | 23456 | nnnnn |
| **18** | 5555 | oooo |
| **19** | 3455 | pppp |

```
df.salary*0.10
```

```
0       1000.0
1        800.0
2      12000.0
3       4500.0
4      33456.7
5       2123.4
6        234.5
7       2345.6
8        765.4
9     345657.8
10      4567.8
11        34.5
12       987.6
13      3456.7
14      8765.4
15      3456.7
```

```
16       123.4
17      2345.6
18       555.5
19       345.5
Name: salary, dtype: float64
```

```
df['tax']=df.salary*0.10
```

```
df
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 | 1000.0 |
| **1** | 2 | bbbb | female | 8000 | 800.0 |
| **2** | 3 | ccccc | male | 120000 | 12000.0 |
| **3** | 4 | dddd | male | 45000 | 4500.0 |
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **5** | 6 | ffff | female | 21234 | 2123.4 |
| **6** | 7 | gggg | female | 2345 | 234.5 |
| **7** | 8 | hhhh | female | 23456 | 2345.6 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |
| **10** | 11 | kkkk | male | 45678 | 4567.8 |
| **11** | 12 | llll | male | 345 | 34.5 |
| **12** | 13 | mmmm | female | 9876 | 987.6 |
| **13** | 14 | nnnnn | male | 34567 | 3456.7 |
| **14** | 15 | jjjj | male | 87654 | 8765.4 |
| **15** | 16 | iiii | female | 34567 | 3456.7 |
| **16** | 17 | mmmm | female | 1234 | 123.4 |
| **17** | 18 | nnnnn | female | 23456 | 2345.6 |
| **18** | 19 | oooo | female | 5555 | 555.5 |
| **19** | 20 | pppp | male | 3455 | 345.5 |

```
df.loc[4]
```

```
id             5
name        eeee
gender    female
salary    334567
tax      33456.7
Name: 4, dtype: object
```

```
for i in range(0,len(df)):
    print(df.loc[i])
```

```
id             1
name        aaaa
gender      male
salary     10000
tax       1000.0
Name: 0, dtype: object
id             2
name        bbbb
gender    female
salary      8000
tax        800.0
Name: 1, dtype: object
id             3
name        ccccc
gender       male
salary     120000
tax       12000.0
Name: 2, dtype: object
id             4
name        dddd
gender      male
salary     45000
tax       4500.0
Name: 3, dtype: object
id             5
name         eeee
gender     female
salary     334567
tax       33456.7
Name: 4, dtype: object
id             6
name        ffff
gender    female
salary     21234
tax       2123.4
Name: 5, dtype: object
id             7
name        gggg
gender    female
salary      2345
tax        234.5
Name: 6, dtype: object
id             8
name        hhhh
gender    female
salary     23456
tax       2345.6
Name: 7, dtype: object
id            9
name       iiii
gender     male
salary     7654
tax       765.4
Name: 8, dtype: object
id            10
name         jjj
gender      male
salary    3456578
```

```
df.to_csv("final.csv")
```

```
df.head(10)#default 5 else n given
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 | 1000.0 |
| **1** | 2 | bbbb | female | 8000 | 800.0 |
| **2** | 3 | ccccc | male | 120000 | 12000.0 |
| **3** | 4 | dddd | male | 45000 | 4500.0 |
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **5** | 6 | ffff | female | 21234 | 2123.4 |
| **6** | 7 | gggg | female | 2345 | 234.5 |
| **7** | 8 | hhhh | female | 23456 | 2345.6 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |

```
df.tail(10)
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **10** | 11 | kkkk | male | 45678 | 4567.8 |
| **11** | 12 | llll | male | 345 | 34.5 |
| **12** | 13 | mmmm | female | 9876 | 987.6 |
| **13** | 14 | nnnnn | male | 34567 | 3456.7 |
| **14** | 15 | jjjj | male | 87654 | 8765.4 |
| **15** | 16 | iiii | female | 34567 | 3456.7 |
| **16** | 17 | mmmm | female | 1234 | 123.4 |
| **17** | 18 | nnnnn | female | 23456 | 2345.6 |
| **18** | 19 | oooo | female | 5555 | 555.5 |
| **19** | 20 | pppp | male | 3455 | 345.5 |

```
df[4:15]
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **5** | 6 | ffff | female | 21234 | 2123.4 |
| **6** | 7 | gggg | female | 2345 | 234.5 |
| **7** | 8 | hhhh | female | 23456 | 2345.6 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |
| **10** | 11 | kkkk | male | 45678 | 4567.8 |

```
1 df
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 | 1000.0 |
| **1** | 2 | bbbb | female | 8000 | 800.0 |
| **2** | 3 | ccccc | male | 120000 | 12000.0 |
| **3** | 4 | dddd | male | 45000 | 4500.0 |
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **5** | 6 | ffff | female | 21234 | 2123.4 |
| **6** | 7 | gggg | female | 2345 | 234.5 |
| **7** | 8 | hhhh | female | 23456 | 2345.6 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |
| **10** | 11 | kkkk | male | 45678 | 4567.8 |
| **11** | 12 | llll | male | 345 | 34.5 |
| **12** | 13 | mmmm | female | 9876 | 987.6 |
| **13** | 14 | nnnnn | male | 34567 | 3456.7 |
| **14** | 15 | jjjj | male | 87654 | 8765.4 |
| **15** | 16 | iiii | female | 34567 | 3456.7 |
| **16** | 17 | mmmm | female | 1234 | 123.4 |
| **17** | 18 | nnnnn | female | 23456 | 2345.6 |
| **18** | 19 | oooo | female | 5555 | 555.5 |
| **19** | 20 | pppp | male | 3455 | 345.5 |

```
1 df['gender']=="male"#only gives True False--filter
```

```
0      True
1     False
```

```
2      True
3      True
4     False
5     False
6     False
7     False
8      True
9     False
10     True
11     True
12    False
13     True
14     True
15    False
16    False
17    False
18    False
19     True
Name: gender, dtype: bool
```

```
df[df['gender']=="male"]
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 | 1000.0 |
| **2** | 3 | ccccc | male | 120000 | 12000.0 |
| **3** | 4 | dddd | male | 45000 | 4500.0 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **10** | 11 | kkkk | male | 45678 | 4567.8 |
| **11** | 12 | llll | male | 345 | 34.5 |
| **13** | 14 | nnnnn | male | 34567 | 3456.7 |
| **14** | 15 | jjjj | male | 87654 | 8765.4 |
| **19** | 20 | pppp | male | 3455 | 345.5 |

```
df1=df[df['gender']=="male"]
df1.to_csv("Male.csv")
df2=df[df['gender']=="female"]
df2.to_csv("Female.csv")
```

```
df[(df['gender']=="female")&(df['salary']>=30000)]
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **15** | 16 | iiii | female | 34567 | 3456.7 |

```
max(df.salary)
```

```
3456578
```

```
avg=sum(df.salary)/len(df)
print("Average salary for employee is:",avg)
df[df['salary']>avg]
```

```
Average salary for employee is: 213761.05
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |

```
df.sort_values(by="salary")[-2:-1]

```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **4** | 5 | eeee | female | 334567 | 33456.7 |

```
df
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|

```
df.drop(labels=["tax","salary"],axis=1)
```

| | id | name | gender |
|---|---|---|---|
| **0** | 1 | aaaa | male |
| **1** | 2 | bbbb | female |
| **2** | 3 | ccccc | male |
| **3** | 4 | dddd | male |
| **4** | 5 | eeee | female |
| **5** | 6 | ffff | female |
| **6** | 7 | gggg | female |
| **7** | 8 | hhhh | female |
| **8** | 9 | iiii | male |
| **9** | 10 | jjj | male |
| **10** | 11 | kkkk | male |
| **11** | 12 | llll | male |
| **12** | 13 | mmmm | female |
| **13** | 14 | nnnnn | male |
| **14** | 15 | jjjj | male |
| **15** | 16 | iiii | female |
| **16** | 17 | mmmm | female |
| **17** | 18 | nnnnn | female |
| **18** | 19 | oooo | female |
| **19** | 20 | pppp | male |

```
df
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 | 1000.0 |
| **1** | 2 | bbbb | female | 8000 | 800.0 |
| **2** | 3 | ccccc | male | 120000 | 12000.0 |
| **3** | 4 | dddd | male | 45000 | 4500.0 |
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **5** | 6 | ffff | female | 21234 | 2123.4 |
| **6** | 7 | gggg | female | 2345 | 234.5 |
| **7** | 8 | hhhh | female | 23456 | 2345.6 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |
| **10** | 11 | kkkk | male | 45678 | 4567.8 |
| **11** | 12 | llll | male | 345 | 34.5 |

```
1 df.drop([1,5,10])
```

| | id | name | gender | salary | tax |
|---|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 | 1000.0 |
| **2** | 3 | ccccc | male | 120000 | 12000.0 |
| **3** | 4 | dddd | male | 45000 | 4500.0 |
| **4** | 5 | eeee | female | 334567 | 33456.7 |
| **6** | 7 | gggg | female | 2345 | 234.5 |
| **7** | 8 | hhhh | female | 23456 | 2345.6 |
| **8** | 9 | iiii | male | 7654 | 765.4 |
| **9** | 10 | jjj | male | 3456578 | 345657.8 |
| **11** | 12 | llll | male | 345 | 34.5 |
| **12** | 13 | mmmm | female | 9876 | 987.6 |
| **13** | 14 | nnnnn | male | 34567 | 3456.7 |
| **14** | 15 | jjjj | male | 87654 | 8765.4 |
| **15** | 16 | iiii | female | 34567 | 3456.7 |
| **16** | 17 | mmmm | female | 1234 | 123.4 |
| **17** | 18 | nnnnn | female | 23456 | 2345.6 |
| **18** | 19 | oooo | female | 5555 | 555.5 |
| **19** | 20 | pppp | male | 3455 | 345.5 |

```
df.drop("gender",axis=1)
```

| | id | name | salary | tax |
|---|---|---|---|---|
| **0** | 1 | aaaa | 10000 | 1000.0 |
| **1** | 2 | bbbb | 8000 | 800.0 |
| **2** | 3 | ccccc | 120000 | 12000.0 |
| **3** | 4 | dddd | 45000 | 4500.0 |
| **4** | 5 | eeee | 334567 | 33456.7 |
| **5** | 6 | ffff | 21234 | 2123.4 |
| **6** | 7 | gggg | 2345 | 234.5 |
| **7** | 8 | hhhh | 23456 | 2345.6 |
| **8** | 9 | iiii | 7654 | 765.4 |
| **9** | 10 | jjj | 3456578 | 345657.8 |
| **10** | 11 | kkkk | 45678 | 4567.8 |
| **11** | 12 | llll | 345 | 34.5 |
| **12** | 13 | mmmm | 9876 | 987.6 |
| **13** | 14 | nnnnn | 34567 | 3456.7 |
| **14** | 15 | jjjj | 87654 | 8765.4 |
| **15** | 16 | iiii | 34567 | 3456.7 |
| **16** | 17 | mmmm | 1234 | 123.4 |
| **17** | 18 | nnnnn | 23456 | 2345.6 |
| **18** | 19 | oooo | 5555 | 555.5 |
| **19** | 20 | pppp | 3455 | 345.5 |

```
df.drop(df[df['gender']=="female"].index)
```

| | id | name | gender | salary |
|---|---|---|---|---|
| **0** | 1 | aaaa | male | 10000 |

```
1
```

| | | | | |
|---|---|---|---|---|
| **3** | 4 | dddd | male | 45000 |
| **8** | 9 | iiii | male | 7654 |
| **9** | 10 | jjj | male | 3456578 |
| **10** | 11 | kkkk | male | 45678 |
| **11** | 12 | llll | male | 345 |
| **13** | 14 | nnnnn | male | 34567 |
| **14** | 15 | jjjj | male | 87654 |
| **19** | 20 | pppp | male | 3455 |