GROUPBY FUNCTION

In [1]:

```
import pandas as pd
var=pd.DataFrame({'name':['pooja','shekhar','mausmi','huma','sam','pooja','sam','tilak','huma','huma','aj
```

In [2]:

```
var
```

Out[2]:

	name	gender	sub_1	sub_2
0	pooja	F	12	45
1	shekhar	M	14	78
2	mausmi	F	12	65
3	huma	F	59	34
4	sam	M	30	86
5	pooja	F	45	29
6	sam	M	25	47
7	tilak	M	47	58
8	huma	F	76	65
9	huma	F	32	68
10	ajit	M	39	39

In [3]:

```
var1=var.groupby('name')
```

In [17]:

```
var1
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x000001EF55951120>
```

In [5]:

```
for x,y in var1:
    print(x)
    print("--------------------------------------------")
    print(y)
    print("--------------------------------------------")
```

```
ajit
--------------------------------------------
    name gender  sub_1  sub_2
10  ajit      M     39     39
--------------------------------------------
huma
--------------------------------------------
   name gender  sub_1  sub_2
3  huma      F     59     34
8  huma      F     76     65
9  huma      F     32     68
--------------------------------------------
mausmi
--------------------------------------------
     name gender  sub_1  sub_2
2  mausmi      F     12     65
--------------------------------------------
pooja
--------------------------------------------
    name gender  sub_1  sub_2
0  pooja      F     12     45
5  pooja      F     45     29
--------------------------------------------
sam
--------------------------------------------
  name gender  sub_1  sub_2
4  sam      M     30     86
6  sam      M     25     47
--------------------------------------------
shekhar
--------------------------------------------
      name gender  sub_1  sub_2
1  shekhar      M     14     78
--------------------------------------------
tilak
--------------------------------------------
    name gender  sub_1  sub_2
7  tilak      M     47     58
--------------------------------------------
```

In [6]:

```
var1.get_group('huma')
```

Out[6]:

	name	gender	sub_1	sub_2
3	huma	F	59	34
8	huma	F	76	65
9	huma	F	32	68

In [7]:

```
var1.min()
```

Out[7]:

	gender	sub_1	sub_2
name			
ajit	M	39	39
huma	F	32	34
mausmi	F	12	65
pooja	F	12	29
sam	M	25	47
shekhar	M	14	78
tilak	M	47	58

In [8]:

```
var1.max()
```

Out[8]:

	gender	sub_1	sub_2
name			
ajit	M	39	39
huma	F	76	68
mausmi	F	12	65
pooja	F	45	45
sam	M	30	86
shekhar	M	14	78
tilak	M	47	58

In [9]:

```
var1.mean()
```

```
C:\Users\Asus\AppData\Local\Temp\ipykernel_3252\772936438.py:1: FutureWarning: The default value of nume
ric_only in DataFrameGroupBy.mean is deprecated. In a future version, numeric_only will default to Fals
e. Either specify numeric_only or select only columns which should be valid for the function.
  var1.mean()
```

Out[9]:

	sub_1	sub_2
name		
ajit	39.000000	39.000000
huma	55.666667	55.666667
mausmi	12.000000	65.000000
pooja	28.500000	37.000000
sam	27.500000	66.500000
shekhar	14.000000	78.000000
tilak	47.000000	58.000000

In [10]:

```
li=list(var1) #convert this data frame into list
```

In [11]:

```
li
```

Out[11]:

```
[('ajit',
      name gender  sub_1  sub_2
  10  ajit      M     39     39),
 ('huma',
     name gender  sub_1  sub_2
  3  huma      F     59     34
  8  huma      F     76     65
  9  huma      F     32     68),
 ('mausmi',
       name gender  sub_1  sub_2
  2  mausmi      F     12     65),
 ('pooja',
      name gender  sub_1  sub_2
  0  pooja      F     12     45
  5  pooja      F     45     29),
 ('sam',
    name gender  sub_1  sub_2
  4  sam      M     30     86
  6  sam      M     25     47),
 ('shekhar',
        name gender  sub_1  sub_2
  1  shekhar      M     14     78),
 ('tilak',
      name gender  sub_1  sub_2
  7  tilak      M     47     58)]
```

In [12]:

```
var.groupby(['name'])['name'].agg('count')
```

Out[12]:

```
name
ajit       1
huma       3
mausmi     1
pooja      2
sam        2
shekhar    1
tilak      1
Name: name, dtype: int64
```

In [13]:

```
var.groupby(['gender'])['sub_1'].agg({'mean'})
```

Out[13]:

	mean
gender	
F	39.333333
M	31.000000

In [14]:

```
var.groupby(['gender']).agg({'sub_1':'count','sub_2':'count'})
```

Out[14]:

	sub_1	**sub_2**
gender		
F	6	6
M	5	5

In [15]:

```
var.groupby(['gender']).agg({'sub_1':'mean','sub_2':'mean'})
```

Out[15]:

	sub_1	**sub_2**
gender		
F	39.333333	51.0
M	31.000000	61.6

In [ ]: