

In [1]:

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

Data Structures

Pandas Series

Pandas Data Frames

In [2]:

```
s=pd.Series([7,15,4,4,7,4],index=['r1','r2','r3','r4','r5','r6'])
s
```

Out[2]:

```
r1    7
r2   15
r3    4
r4    4
r5    7
r6    4
dtype: int64
```

In [3]:

```
s.r2
```

Out[3]:

```
15
```

In [4]:

```
s[1]
```

Out[4]:

```
15
```

In [5]:

```
s.values
```

Out[5]:

```
array([ 7, 15,  4,  4,  7,  4], dtype=int64)
```

In [6]:

```
s.index
```

Out[6]:

```
Index(['r1', 'r2', 'r3', 'r4', 'r5', 'r6'], dtype='object')
```

In [7]:

```
s.value_counts()
```

Out[7]:

```
4    3
7    2
15    1
dtype: int64
```

In [8]:

```
#DataFrame
```

In [9]:

```
df = pd.DataFrame([[1,8],[3,4],[5,9],[7,2],[2,6]], index= ['R1','R2','R3','R4','R5'], columns=['C1','C2'])
df
```

Out[9]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
R4   7   2
R5   2   6
```

In [10]:

```
df[:3]
```

Out[10]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
```

In [11]:

```
df['C2']
```

Out[11]:

```
R1    8
R2    4
R3    9
R4    2
R5    6
Name: C2, dtype: int64
```

In [12]:

```
df
```

Out[12]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
R4   7   2
R5   2   6
```

In [13]:

```
df.iloc[2,1]
```

Out[13]:

```
9
```

In [14]:

```
df.values
```

Out[14]:

```
array([[1, 8],
       [3, 4],
       [5, 9],
       [7, 2],
       [2, 6]], dtype=int64)
```

In [15]:

```
df[df<5]
```

Out[15]:

```
   C1  C2
R1  1.0 NaN
R2  3.0  4.0
R3  NaN NaN
R4  NaN  2.0
R5  2.0 NaN
```

In [16]:

```
df<5
```

Out[16]:

```
   C1  C2
R1  True False
R2  True  True
R3 False False
R4 False  True
R5  True False
```

In [17]:

```
df.head(3)
```

Out[17]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
```

In [18]:

```
df.tail(2)
```

Out[18]:

```
   C1  C2
R4   7   2
R5   2   6
```

In [19]:

```
df
```

Out[19]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
R4   7   2
R5   2   6
```

In [20]:

```
df.index
```

Out[20]:

```
Index(['R1', 'R2', 'R3', 'R4', 'R5'], dtype='object')
```

In [21]:

```
df.columns
```

Out[21]:

```
Index(['C1', 'C2'], dtype='object')
```

In [22]:

```
df # (5, 2)
```

Out[22]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
R4   7   2
R5   2   6
```

In [23]:

```
df.T # (2,5)
```

Out[23]:

```
   R1  R2  R3  R4  R5
C1   1   3   5   7   2
C2   8   4   9   2   6
```

In [24]:

```
df.drop('R3',axis=0)
```

Out[24]:

```
   C1  C2
R1   1   8
R2   3   4
R4   7   2
R5   2   6
```

In [25]:

```
df.drop('C2',axis=1)
```

Out[25]:

```
   C1
R1   1
R2   3
R3   5
R4   7
R5   2
```

In [26]:

```
df
```

Out[26]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
R4   7   2
R5   2   6
```

In [27]:

```
df.describe()
```

Out[27]:

```
   C1  C2
count  5.000000  5.000000
mean   3.600000  5.800000
std    2.408319  2.863564
min    1.000000  2.000000
25%    2.000000  4.000000
50%    3.000000  6.000000
75%    5.000000  8.000000
max    7.000000  9.000000
```

In [28]:

```
df
```

Out[28]:

```
   C1  C2
R1   1   8
R2   3   4
R3   5   9
R4   7   2
R5   2   6
```

In [29]:

```
#other example
```

In [30]:

```
df = pd.DataFrame( { 'x1':[1,2], 'x2':[3,4] , 'x3':[5,6] } )
df
```

Out[30]:

```
   x1  x2  x3
0    1   3   5
1    2   4   6
```

In [31]:

```
data = {'Name': ["Ali", "Sara", "Reza", "Taha"],
        'Location': ["yazd", "Theran", "shiraz", "Hamedan"],
        'Age' : [35, 28, 55, 5]
}
```

In [32]:

```
df = pd.DataFrame(data)
df
```

Out[32]:

```
   Name Location Age
0    Ali   yazd   35
1  Sara  Theran   28
2  Reza  shiraz   55
3  Taha  Hamedan    5
```

In [33]:

```
df.sort_index(axis=1)
```

Out[33]:

```
   Age Location Name
0   35   yazd   Ali
1   28  Theran  Sara
2   55  shiraz  Reza
3    5  Hamedan  Taha
```

In [34]:

```
df.sort_index(axis=0,ascending=False)
```

Out[34]:

```
   Name Location Age
3  Taha  Hamedan    5
2  Reza  shiraz   55
1  Sara  Theran   28
0   Ali   yazd   35
```

In [35]:

```
df.sort_index(axis=0,ascending=True)
```

Out[35]:

```
   Name Location Age
0   Ali   yazd   35
1  Sara  Theran   28
2  Reza  shiraz   55
3  Taha  Hamedan    5
```

In [36]:

```
df.sort_values(by='Age',ascending=False)
```

Out[36]:

```
   Name Location Age
2  Reza  shiraz   55
0   Ali   yazd   35
1  Sara  Theran   28
3  Taha  Hamedan    5
```

In [37]:

```
df[df.Age <= 30]
```

Out[37]:

```
   Name Location Age
1  Sara  Theran   28
3  Taha  Hamedan    5
```

In [38]:

```
df
```

Out[38]:

```
   Name Location Age
0   Ali   yazd   35
1  Sara  Theran   28
2  Reza  shiraz   55
3  Taha  Hamedan    5
```

In [39]:

```
df.iloc[2][1]
```

Out[39]:

```
'shiraz'
```

In [40]:

```
df.iloc[1]['Age']
```

Out[40]:

```
28
```