

```
In [2]: #Series data structure
import pandas as pd
s=pd.Series([1,2,3,4,5,6])
s
```

```
Out[2]: 0    1
        1    2
        2    3
        3    4
        4    5
        5    6
dtype: int64
```

```
In [4]: #Creating a DataFrame by passing a NumPy array, with a datetime index and labeled
dates=pd.date_range("20220101", periods=4)
dates
```

```
Out[4]: DatetimeIndex(['2022-01-01', '2022-01-02', '2022-01-03', '2022-01-04'], dtype
='datetime64[ns]', freq='D')
```

```
In [10]: #Creating a DataFrame by passing a dictionary of objects that can be converted in
import pandas as pd
import numpy as np
df=pd.DataFrame({
    "A":1.0,
    "B":pd.Timestamp("20220101"),
    "C":pd.Series(1,index=list(range(4)),dtype="float32"),
    "D":np.array([2] * 4,dtype="int32"),
    "E":"Egg",
})
df
```

```
Out[10]:
```

	A	B	C	D	E
0	1.0	2022-01-01	1.0	2	Egg
1	1.0	2022-01-01	1.0	2	Egg
2	1.0	2022-01-01	1.0	2	Egg
3	1.0	2022-01-01	1.0	2	Egg

```
In [12]: #DataFrame have different datatypes
df.dtypes
```

```
Out[12]: A    float64
         B    datetime64[ns]
         C    float32
         D    int32
         E    object
dtype: object
```

```
In [13]: #viewing of data  
#HEAD  
df.head()
```

```
Out[13]:
```

	A	B	C	D	E
0	1.0	2022-01-01	1.0	2	Egg
1	1.0	2022-01-01	1.0	2	Egg
2	1.0	2022-01-01	1.0	2	Egg
3	1.0	2022-01-01	1.0	2	Egg

```
In [15]: #TAIL  
df.tail(2)
```

```
Out[15]:
```

	A	B	C	D	E
2	1.0	2022-01-01	1.0	2	Egg
3	1.0	2022-01-01	1.0	2	Egg

```
In [17]: df.index
```

```
Out[17]: Int64Index([0, 1, 2, 3], dtype='int64')
```

```
In [18]: df.columns
```

```
Out[18]: Index(['A', 'B', 'C', 'D', 'E'], dtype='object')
```

```
In [19]: #Dataframe to numpy  
df.to_numpy()
```

```
Out[19]: array([[1.0, Timestamp('2022-01-01 00:00:00'), 1.0, 2, 'Egg'],  
                [1.0, Timestamp('2022-01-01 00:00:00'), 1.0, 2, 'Egg'],  
                [1.0, Timestamp('2022-01-01 00:00:00'), 1.0, 2, 'Egg'],  
                [1.0, Timestamp('2022-01-01 00:00:00'), 1.0, 2, 'Egg']],  
              dtype=object)
```

```
In [20]: #describe() shows a quick statistic summary of your data
df.describe()
```

Out[20]:

	A	C	D
count	4.0	4.0	4.0
mean	1.0	1.0	2.0
std	0.0	0.0	0.0
min	1.0	1.0	2.0
25%	1.0	1.0	2.0
50%	1.0	1.0	2.0
75%	1.0	1.0	2.0
max	1.0	1.0	2.0

```
In [27]: #Sorting
df.sort_index(axis=1,ascending=False)
```

Out[27]:

	E	D	C	B	A
0	Egg	2	1.0	2022-01-01	1.0
1	Egg	2	1.0	2022-01-01	1.0
2	Egg	2	1.0	2022-01-01	1.0
3	Egg	2	1.0	2022-01-01	1.0

```
In [28]: #by values
df.sort_values(by="B")
```

Out[28]:

	A	B	C	D	E
0	1.0	2022-01-01	1.0	2	Egg
1	1.0	2022-01-01	1.0	2	Egg
2	1.0	2022-01-01	1.0	2	Egg
3	1.0	2022-01-01	1.0	2	Egg

```
In [30]: #GETTING
df["A"]
```

Out[30]:

0	1.0
1	1.0
2	1.0
3	1.0

Name: A, dtype: float64

In [31]: `df[0:2]`

Out[31]:

	A	B	C	D	E
0	1.0	2022-01-01	1.0	2	Egg
1	1.0	2022-01-01	1.0	2	Egg

In [33]: *#selection by labels*
`df.loc[0]`

Out[33]:

A	1
B	2022-01-01 00:00:00
C	1
D	2
E	Egg

Name: 0, dtype: object

In [34]: `df.loc[:,["A","B"]]`

Out[34]:

	A	B
0	1.0	2022-01-01
1	1.0	2022-01-01
2	1.0	2022-01-01
3	1.0	2022-01-01

In [37]: *#selection by position*
`df.iloc[2]`

Out[37]:

A	1
B	2022-01-01 00:00:00
C	1
D	2
E	Egg

Name: 2, dtype: object

In [38]: `df.iloc[3:4,0:2]`

Out[38]:

	A	B
3	1.0	2022-01-01

```
In [41]: #random values  
pd.Series(np.random.randn(5),index=['a','b','c','d','e'])
```

```
Out[41]: a    2.522757  
        b   -0.101144  
        c    0.485219  
        d    0.695772  
        e    1.133065  
        dtype: float64
```

```
In [ ]:
```