

```
In [ ]: Pandas -- > Pandas is very powerful tool for data manipulation and analysis.
```

Or

Pandas is fast, powerful flexiable and easy to use open source data analysis and why Use Pandas?

- 1- Easy data manipulation and cleaning
- 2- Powerful capabilities for data analysis
- 3- it is handled different different data formats(csv,excels,json, SQL etc)

History :

in, 2008, wes mackenny started the deploying the pandas for high performanace

- 1- Pandas is the Open source
- 2- easy to use by anyone
- 3- it is data analysis tool

```
In [1]: # installation
```

```
!pip install pandas
```

```
Requirement already satisfied: pandas in c:\users\jitud\appdata\local\programs\python\python313\lib\site-packages (2.2.3)
Requirement already satisfied: numpy>=1.26.0 in c:\users\jitud\appdata\local\programs\python\python313\lib\site-packages (from pandas) (2.2.6)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\jitud\appdata\local\programs\python\python313\lib\site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in c:\users\jitud\appdata\local\programs\python\python313\lib\site-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in c:\users\jitud\appdata\local\programs\python\python313\lib\site-packages (from pandas) (2025.2)
Requirement already satisfied: six>=1.5 in c:\users\jitud\appdata\local\programs\python\python313\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.17.0)
```

```
[notice] A new release of pip is available: 25.1.1 -> 25.2
```

```
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
In [2]: # Import pandas packages
import pandas as pd
```

```
In [4]: # to check the pandas version
print(pd.__version__)
```

```
2.2.3
```

```
In [ ]: # Pandas has two components
```

```
1- Series
```

```
2- DataFrame
```

1. Series

Series is one dimensional array like structure with homogeneous data

eg. - [23,12,3,45,6,67,67,567,567,5,756,756,756,765]

```
In [5]: import pandas as pd
```

```
In [6]: # create a empty object
```

```
s = pd.Series()
print(s)
```

```
Series([], dtype: object)
```

```
In [11]: # from a list
```

```
lst = [2,3,45,6,6]
#print(type(lst))

series1 = pd.Series(lst)
#print(type(series1))
series1
```

```
Out[11]: 0    2
         1    3
         2   45
         3    6
         4    6
         dtype: int64
```

```
In [12]: # from a tuple
```

```
series2 = pd.Series((2,3,45,6,6))
series2
```

```
Out[12]: 0    2
         1    3
         2   45
         3    6
         4    6
         dtype: int64
```

```
In [13]: # from a dictionary
```

```
d = {"name": "john", "age" : 20}
series2 = pd.Series(d)
series2
```

```
Out[13]: name    john
         age      20
         dtype: object
```

```
In [20]: # from a array
```

```
import numpy as np
```

```
arr = np.array([1,2,3,45,5])
#print(type(arr))

arr_series = pd.Series(arr)
arr_series
```

```
Out[20]: 0      1
        1      2
        2      3
        3     45
        4      5
        dtype: int64
```

```
In [21]: # with custom index

series2 = pd.Series([12,14,16,18], index=[100,200,300,400])
series2
```

```
Out[21]: 100     12
        200     14
        300     16
        400     18
        dtype: int64
```

DataFrame

Dataframe is two dimensional array with heterogeneous data .

```
In [23]: # empty dataframe

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

```
Empty DataFrame
Columns: []
Index: []
```

```
In [26]: # creating a dataframe from dictionary

data = {
    "Name" : ["John", "Bob", "Cat"],
    "Age" : [20,30,25]
}

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

```
Out[26]:
```

	Name	Age
0	John	20
1	Bob	30
2	Cat	25

```
In [ ]:
```