

```
1 import pandas as pd
```

```
1 ### a series
```

```
1 a=pd.Series([1,2,3,4,5],index=["A","B","C","D","E"])
```

```
2 a
```

```
A    1
B    2
C    3
D    4
E    5
dtype: int64
```

```
1 ###MAKING a dataframe
```

```
1 b=pd.DataFrame({"Umar":22,"Umair":21,"Asad":24},index=["A","B","C"])
```

```
2 b
```

|          | Umar | Umair | Asad |
|----------|------|-------|------|
| <b>A</b> | 22   | 21    | 24   |
| <b>B</b> | 22   | 21    | 24   |
| <b>C</b> | 22   | 21    | 24   |

```
1 ### Working on DataSet from seaborn Llibrary
```

```
1 import seaborn as sns
```

```
2 df=sns.load_dataset ("titanic")
```

```
3 df
```

|            | survived | pclass | sex    | age  | sibsp | parch | fare    | embarked | class  | who   | adult |
|------------|----------|--------|--------|------|-------|-------|---------|----------|--------|-------|-------|
| <b>0</b>   | 0        | 3      | male   | 22.0 | 1     | 0     | 7.2500  | S        | Third  | man   |       |
| <b>1</b>   | 1        | 1      | female | 38.0 | 1     | 0     | 71.2833 | C        | First  | woman |       |
| <b>2</b>   | 1        | 3      | female | 26.0 | 0     | 0     | 7.9250  | S        | Third  | woman |       |
| <b>3</b>   | 1        | 1      | female | 35.0 | 1     | 0     | 53.1000 | S        | First  | woman |       |
| <b>4</b>   | 0        | 3      | male   | 35.0 | 0     | 0     | 8.0500  | S        | Third  | man   |       |
| ...        | ...      | ...    | ...    | ...  | ...   | ...   | ...     | ...      | ...    | ...   |       |
| <b>886</b> | 0        | 2      | male   | 27.0 | 0     | 0     | 13.0000 | S        | Second | man   |       |
| <b>887</b> | 1        | 1      | female | 19.0 | 0     | 0     | 30.0000 | S        | First  | woman |       |
| <b>888</b> | 0        | 3      | female | NaN  | 1     | 2     | 23.4500 | S        | Third  | woman |       |
| <b>889</b> | 1        | 1      | male   | 26.0 | 0     | 0     | 30.0000 | C        | First  | man   |       |
| <b>890</b> | 0        | 3      | male   | 32.0 | 0     | 0     | 7.7500  | Q        | Third  | man   |       |

891 rows × 15 columns



```
1 ### Checking information about Data
2 df.info()
3 df
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 15 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   survived    891 non-null    int64
1   pclass      891 non-null    int64
2   sex         891 non-null    object
3   age         714 non-null    float64
4   sibsp       891 non-null    int64
5   parch       891 non-null    int64
6   fare        891 non-null    float64
7   embarked    889 non-null    object
8   class       891 non-null    category
9   who         891 non-null    object
10  adult_male  891 non-null    bool
11  deck        203 non-null    category
12  embark_town 889 non-null    object
13  alive       891 non-null    object
14  alone       891 non-null    bool
dtypes: bool(2), category(2), float64(2), int64(4), object(5)
memory usage: 80.7+ KB
```

|     | survived | pclass | sex    | age  | sibsp | parch | fare    | embarked | class  | who   | adult |
|-----|----------|--------|--------|------|-------|-------|---------|----------|--------|-------|-------|
| 0   | 0        | 3      | male   | 22.0 | 1     | 0     | 7.2500  | S        | Third  | man   |       |
| 1   | 1        | 1      | female | 38.0 | 1     | 0     | 71.2833 | C        | First  | woman |       |
| 2   | 1        | 3      | female | 26.0 | 0     | 0     | 7.9250  | S        | Third  | woman |       |
| 3   | 1        | 1      | female | 35.0 | 1     | 0     | 53.1000 | S        | First  | woman |       |
| 4   | 0        | 3      | male   | 35.0 | 0     | 0     | 8.0500  | S        | Third  | man   |       |
| ... | ...      | ...    | ...    | ...  | ...   | ...   | ...     | ...      | ...    | ...   | ...   |
| 886 | 0        | 2      | male   | 27.0 | 0     | 0     | 13.0000 | S        | Second | man   |       |
| 887 | 1        | 1      | female | 19.0 | 0     | 0     | 30.0000 | S        | First  | woman |       |
| 888 | 0        | 3      | female | NaN  | 1     | 2     | 23.4500 | S        | Third  | woman |       |
| 889 | 1        | 1      | male   | 26.0 | 0     | 0     | 30.0000 | C        | First  | man   |       |
| 890 | 0        | 3      | male   | 32.0 | 0     | 0     | 7.7500  | Q        | Third  | man   |       |

891 rows × 15 columns



```
1 ### checking first 5 Entries
2 df.head()
```

```
1 ### checking last 5 Entires
2 df.tail()
3
```

|     | survived | pclass | sex    | age  | sibsp | parch | fare  | embarked | class  | who   | adult_ |
|-----|----------|--------|--------|------|-------|-------|-------|----------|--------|-------|--------|
| 886 | 0        | 2      | male   | 27.0 | 0     | 0     | 13.00 | S        | Second | man   |        |
| 887 | 1        | 1      | female | 19.0 | 0     | 0     | 30.00 | S        | First  | woman | I      |
| 888 | 0        | 3      | female | NaN  | 1     | 2     | 23.45 | S        | Third  | woman | I      |
| 889 | 1        | 1      | male   | 26.0 | 0     | 0     | 30.00 | C        | First  | man   |        |
| 890 | 0        | 3      | male   | 32.0 | 0     | 0     | 7.75  | Q        | Third  | man   |        |

```
1 ### Summary statistics
2 df.describe()
```

|       | survived   | pclass     | age        | sibsp      | parch      | fare       |
|-------|------------|------------|------------|------------|------------|------------|
| count | 891.000000 | 891.000000 | 714.000000 | 891.000000 | 891.000000 | 891.000000 |
| mean  | 0.383838   | 2.308642   | 29.699118  | 0.523008   | 0.381594   | 32.204208  |
| std   | 0.486592   | 0.836071   | 14.526497  | 1.102743   | 0.806057   | 49.693429  |
| min   | 0.000000   | 1.000000   | 0.420000   | 0.000000   | 0.000000   | 0.000000   |
| 25%   | 0.000000   | 2.000000   | 20.125000  | 0.000000   | 0.000000   | 7.910400   |
| 50%   | 0.000000   | 3.000000   | 28.000000  | 0.000000   | 0.000000   | 14.454200  |
| 75%   | 1.000000   | 3.000000   | 38.000000  | 1.000000   | 0.000000   | 31.000000  |
| max   | 1.000000   | 3.000000   | 80.000000  | 8.000000   | 6.000000   | 512.329200 |

```
1 ### Checking number of row and columns
2 df.shape
```

```
(891, 15)
```

```
1 ### Checking number of row and columns
2 df.shape[0]
```

```
891
```

```
1 ### checking column name
2 df.columns
```

```
Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
       'alive', 'alone'],
      dtype='object')
```

```
1 ### checking row heading
2 df.index
```

```
RangeIndex(start=0, stop=891, step=1)
```

```
1 df.shape[1]
```

```
1 ### Remove Soectfic column
2 df.drop(["deck","alone"],axis=1)
3 df
```

|     | survived | pclass | sex    | age  | sibsp | parch | fare    | embarked | class  | who   | adul |
|-----|----------|--------|--------|------|-------|-------|---------|----------|--------|-------|------|
| 0   | 0        | 3      | male   | 22.0 | 1     | 0     | 7.2500  | S        | Third  | man   |      |
| 1   | 1        | 1      | female | 38.0 | 1     | 0     | 71.2833 | C        | First  | woman |      |
| 2   | 1        | 3      | female | 26.0 | 0     | 0     | 7.9250  | S        | Third  | woman |      |
| 3   | 1        | 1      | female | 35.0 | 1     | 0     | 53.1000 | S        | First  | woman |      |
| 4   | 0        | 3      | male   | 35.0 | 0     | 0     | 8.0500  | S        | Third  | man   |      |
| ... | ...      | ...    | ...    | ...  | ...   | ...   | ...     | ...      | ...    | ...   | ...  |
| 886 | 0        | 2      | male   | 27.0 | 0     | 0     | 13.0000 | S        | Second | man   |      |
| 887 | 1        | 1      | female | 19.0 | 0     | 0     | 30.0000 | S        | First  | woman |      |
| 888 | 0        | 3      | female | NaN  | 1     | 2     | 23.4500 | S        | Third  | woman |      |
| 889 | 1        | 1      | male   | 26.0 | 0     | 0     | 30.0000 | C        | First  | man   |      |
| 890 | 0        | 3      | male   | 32.0 | 0     | 0     | 7.7500  | Q        | Third  | man   |      |

891 rows × 15 columns



```
1 ### checking missing values
2 df.isnull().sum()
```

```
survived      0
pclass        0
sex           0
age          177
sibsp         0
parch         0
fare          0
embarked      2
class         0
who           0
adult_male    0
deck         688
embark_town   2
alive         0
alone         0
dtype: int64
```

```
1 ### checking missing values
2 df.isnull()
```

|     | survived | pclass | sex   | age   | sibsp | parch | fare  | embarked | class | who   | adult_male |
|-----|----------|--------|-------|-------|-------|-------|-------|----------|-------|-------|------------|
| 0   | False    | False  | False | False | False | False | False | False    | False | False | False      |
| 1   | False    | False  | False | False | False | False | False | False    | False | False | False      |
| 2   | False    | False  | False | False | False | False | False | False    | False | False | False      |
| 3   | False    | False  | False | False | False | False | False | False    | False | False | False      |
| 4   | False    | False  | False | False | False | False | False | False    | False | False | False      |
| ... | ...      | ...    | ...   | ...   | ...   | ...   | ...   | ...      | ...   | ...   | ...        |
| 886 | False    | False  | False | False | False | False | False | False    | False | False | False      |

```
1 ### checking unique values
```

```
2 df.who.unique()
```

```
array(['man', 'woman', 'child'], dtype=object)
```

```
1 df.age.unique()
```

```
array([22. , 38. , 26. , 35. , nan, 54. , 2. , 27. , 14. ,
       4. , 58. , 20. , 39. , 55. , 31. , 34. , 15. , 28. ,
       8. , 19. , 40. , 66. , 42. , 21. , 18. , 3. , 7. ,
       49. , 29. , 65. , 28.5 , 5. , 11. , 45. , 17. , 32. ,
       16. , 25. , 0.83, 30. , 33. , 23. , 24. , 46. , 59. ,
       71. , 37. , 47. , 14.5 , 70.5 , 32.5 , 12. , 9. , 36.5 ,
       51. , 55.5 , 40.5 , 44. , 1. , 61. , 56. , 50. , 36. ,
       45.5 , 20.5 , 62. , 41. , 52. , 63. , 23.5 , 0.92, 43. ,
       60. , 10. , 64. , 13. , 48. , 0.75, 53. , 57. , 80. ,
       70. , 24.5 , 6. , 0.67, 30.5 , 0.42, 34.5 , 74. ])
```