```
import pandas as pd
```

Making a DataFrame

```
a=pd.DataFrame({"neha":19,"farwa":19,"arooj":20},index=["A","B"])
a
```

	neha	farwa	arooj
A	19	19	20
B	19	19	20

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

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

Working on DataSet from Seaborn Library

```
import seaborn as sns
df=sns.load_dataset("titanic")
df
```

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

891 rows × 15 columns

Checking Information about Data

```
df.info()
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_male	deck	e
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	S
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	S
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	S
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...	

Checking first 5 Entries

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

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

891 rows × 15 columns

Double-click (or enter) to edit

Checking last 5 Entries

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

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	e
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	S
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	S
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	S
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...	
886	0	2	male	27.0	0	0	13.0000	S	Second	man	True	NaN	S
887	1	1	female	19.0	0	0	30.0000	S	First	woman	False	B	S
888	0	3	female	NaN	1	2	23.4500	S	Third	woman	False	NaN	S

Summary Statics

```
df.describe()
df
```

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

891 rows × 15 columns

Checking No. OF rOWS ANS cOLUMS

```
df.shape[0]
```

```
891
```

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

```
15
```

```
rows="the number of rows are", df.shape[0]
columns="the number of colums are",df.shape[1]
print(rows)
print(columns)
```

```
('the number of rows are', 891)
('the number of colums are', 15)
```

checking columns name

```
df.columns
```

```
Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
       'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
```

```
       'alive', 'alone'],
      dtype='object')
```

checking row heading

```
df.index
```

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

removing specific columns

```
df1=df.drop(["deck","alone"],axis =1)
df1
```

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

891 rows × 13 columns

Double-click (or enter) to edit

Checking missing values

```
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
```

checking unique values

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

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

```
df.embark_town.unique()
```

```
array(['Southampton', 'Cherbourg', 'Queenstown', nan], dtype=object)
```

✓ 0s completed at 11:20 AM