

In [1]:

```
#import pandas
import pandas as pd
```

In [31]:

```
#To get the current version of pandas
print(pd.__version__)
```

1.1.3

In [32]:

```
#Create a series from a list
list1=[10,20,30,40,50]
s1=pd.Series(list1)
print(s1)
```

```
0    10
1    20
2    30
3    40
4    50
dtype: int64
```

In [33]:

```
#Customize the index
list2=['a','b','c','d','e']
s2=pd.Series(list1,index=list2)
print(s2)
```

```
a    10
b    20
c    30
d    40
e    50
dtype: int64
```

In [34]:

```
#Create a series from random numbers
#To get random numbers by invoking randn()
import numpy as np
x=np.random.randn(10)
x
```

Out[34]:

```
array([ 0.14639237, -0.7489306 ,  0.52563224,  0.29672407, -1.25144376,
        0.07001811, -0.46579607,  0.78406185,  0.76350546,  1.06214105])
```

In [35]:

```
#Then create a Series by taking random numbers
s3=pd.Series(x)
s3
```

Out[35]:

```
0    0.146392
1   -0.748931
2    0.525632
3    0.296724
4   -1.251444
5    0.070018
6   -0.465796
7    0.784062
8    0.763505
9    1.062141
dtype: float64
```

In [36]:

```
#Create a series from a dictionary
dict1={'a':100, 'b':200, 'c':300}
s4=pd.Series(dict1)
s4
```

Out[36]:

```
a    100
b    200
c    300
dtype: int64
```

In [37]:

```
print(s1)
```

```
0    10
1    20
2    30
3    40
4    50
dtype: int64
```

In [38]:

```
#To get the minimum value
s1.min()
```

Out[38]:

```
10
```

In [39]:

```
#To get the max value  
s1.max()
```

Out[39]:

50

In [40]:

```
#To get mean value  
s1.mean()
```

Out[40]:

30.0

In [41]:

```
#To get the median value  
s1.median()
```

Out[41]:

30.0

In [42]:

```
#Created another series s5  
s5=pd.Series([10,20,30])  
s5
```

Out[42]:

```
0    10  
1    20  
2    30  
dtype: int64
```

In [43]:

```
#Add two series  
s1.add(s5)
```

Out[43]:

```
0    20.0  
1    40.0  
2    60.0  
3     NaN  
4     NaN  
dtype: float64
```

In [44]:

```
#subtract two series  
s1.sub(s5)
```

Out[44]:

```
0    0.0  
1    0.0  
2    0.0  
3    NaN  
4    NaN  
dtype: float64
```

In [45]:

```
#multiply two series  
s1.mul(s5)
```

Out[45]:

```
0    100.0  
1    400.0  
2    900.0  
3     NaN  
4     NaN  
dtype: float64
```

In [46]:

```
#Division of two series  
s1.div(s5)
```

Out[46]:

```
0    1.0  
1    1.0  
2    1.0  
3    NaN  
4    NaN  
dtype: float64
```

In [2]:

```
#Create a DataFrame from a List  
list2=['java','python','php','ruby']  
df1=pd.DataFrame(list2)  
print(df1)
```

```
0  
0    java  
1    python  
2     php  
3     ruby
```

In [3]:

```
#Create a DataFrame from a dictionary
employee_detail={'eid':[101,102,103], 'ename':['Rahul', 'Sachin', 'Sourav'], 'esal':[10000,20000,30000]}
df2=pd.DataFrame(employee_detail)
print(df2)
```

	eid	ename	esal
0	101	Rahul	10000
1	102	Sachin	20000
2	103	Sourav	30000

In [6]:

```
#Create a DataFrame from a csv file
df3=pd.read_csv("E:\cipet\salarinfo.csv")
df3.head()
```

Out[6]:

	YearsExperience	Salary
0	1.1	39343
1	1.3	46205
2	1.5	37731
3	2.0	43525
4	2.2	39891

In [7]:

```
df3.tail()
```

Out[7]:

	YearsExperience	Salary
25	9.0	105582
26	9.5	116969
27	9.6	112635
28	10.3	122391
29	10.5	121872

In [8]:

```
df3.shape
```

Out[8]:

```
(30, 2)
```

In [9]:

df3.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  -
0   YearsExperience  30 non-null    float64
1   Salary          30 non-null    int64
dtypes: float64(1), int64(1)
memory usage: 608.0 bytes
```

In [10]:

df3.describe()

Out[10]:

	YearsExperience	Salary
count	30.000000	30.000000
mean	5.313333	76003.000000
std	2.837888	27414.429785
min	1.100000	37731.000000
25%	3.200000	56720.750000
50%	4.700000	65237.000000
75%	7.700000	100544.750000
max	10.500000	122391.000000

In [25]:

```
#Concatenating two data frames
dict1={'regdno':[101,102,103], 'name':['Sachin','Sourav','Rahul'], 'branch':['CSE','IT','EEE']}
df4=pd.DataFrame(dict1,index=[0,1,2])
df4
```

Out[25]:

	regdno	name	branch
0	101	Sachin	CSE
1	102	Sourav	IT
2	103	Rahul	EEE

In [26]:

```
dict2={'regdno':[104,105,106], 'name':['Akshay', 'Ajay', 'Abhay'], 'branch':['ETC', 'Civil', 'Mechanical']}  
df5=pd.DataFrame(dict2, index=[3,4,5])  
df5
```

Out[26]:

	regdno	name	branch
3	104	Akshay	ETC
4	105	Ajay	Civil
5	106	Abhay	Mechanical

In [27]:

```
#concatenate two data frame  
result=pd.concat([df4,df5])  
print(result)
```

	regdno	name	branch
0	101	Sachin	CSE
1	102	Sourav	IT
2	103	Rahul	EEE
3	104	Akshay	ETC
4	105	Ajay	Civil
5	106	Abhay	Mechanical

In [28]:

```
#merging two data frame  
dict3={'regdno':[107,108,109], 'name':['Lelin', 'Nirmal', 'Kalandi'], 'branch':['ETC', 'Civil', 'Mechanical']}  
df6=pd.DataFrame(dict3)  
df6
```

Out[28]:

	regdno	name	branch
0	107	Lelin	ETC
1	108	Nirmal	Civil
2	109	Kalandi	Mechanical

In [29]:

```
dict4={'regdno':[107,108,109], 'address':['BBSR', 'Cuttack', 'Paradeep'], 'age':[22,24,26]}  
df7=pd.DataFrame(dict4)  
df7
```

Out[29]:

	regdno	address	age
0	107	BBSR	22
1	108	Cuttack	24
2	109	Paradeep	26

In [30]:

```
result1=pd.merge(df6,df7,on='regdno')  
print(result1)
```

	regdno	name	branch	address	age
0	107	Lelin	ETC	BBSR	22
1	108	Nirmal	Civil	Cuttack	24
2	109	Kalandi	Mechanical	Paradeep	26

In []: