**Program no:2**

**Aim: Data frames series using pandas**

**Source Code:**

**1.**

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22])

s = pd.Series(arr)

print(s)

**Output:**

0 10

1 15

2 18

3 22

dtype: int64

**2.**

import pandas as pd

import numpy as np

arr=np.array(['a','b','c','d'])

s=pd.Series(arr,

index=['first','second','third','fourth'])

print(s)

**Output:**

first a

second b

third c

fourth d

dtype: object

**3.**

import pandas as pd

s=pd.Series(50,index=[0,1,2,3,4])

print(s)

**Output:**

0 50

1 50

2 50

3 50

4 50

dtype: int64

**4.**

import pandas as pd

d={'Name':'Hardik','Iplteam':'MI','Runs':1500 }

s=pd.Series(d)

print(s)

**Output:**

Name Hardik

Iplteam MI

Runs 1500

dtype: object

**5.**

import pandas as pd

s=pd.Series([1,2,3,4,5])

print('To multiply all values in a series by 2')

print('------------------------------------------')

print(s\*2)

print('To find the square of all the values in a series')

print('------------------------------------------')

print(s\*\*2)

print('To print all the values in a series that are greater 2')

print('------------------------------------------')

print(s[s>2])

**Output:**

To multiply all values in a series by 2

------------------------------------------

0 2

1 4

2 6

3 8

4 10

dtype: int64

To find the square of all the values in a series

------------------------------------------

0 1

1 4

2 9

3 16

4 25

dtype: int64

To print all the values in a series that are greater 2

------------------------------------------

2 3

3 4

4 5

dtype: int64

**6.**

import pandas as pd

s1=pd.Series([1,2,3,4,5],index=['a','b','c','d','e'])

s2=pd.Series([10,20,30,40,50],index=['a','b','c','d','e'])

s3=pd.Series([5,14,23,32],index=['a','b','c','d'])

print('To add series1 and series2')

print('------------------------------------------')

print(s1+s2)

print('To add series2 and series 3')

print('------------------------------------------')

print(s2+s3)

print('To add series2 and series 3 and filled non matching index with 0')

print('------------------------------------------')

print(s2.add(s3,fill\_value=0))

**Output:**

To add series1 and series2

------------------------------------------

a 11

b 22

c 33

d 44

e 55

dtype: int64

To add series2 and series 3

------------------------------------------

a 15.0

b 34.0

c 53.0

d 72.0

e NaN

dtype: float64

To add series2 and series 3 and filled non matching index with 0

------------------------------------------

a 15.0

b 34.0

c 53.0

d 72.0

e 50.0

dtype: float64

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **7.** |  |  |  |  |  |

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22,55,77,42,48,97])

s=pd.Series(arr)

print(s.head(5))

print(s.tail(5))

**Output:**

0 10

1 15

2 18

3 22

4 55

dtype: int64

4 55

5 77

6 42

7 48

8 97

dtype: int64

**8.**

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22,55,77,42,48,97])

s=pd.Series(arr)

print(s.head(6))

**Output:**

0 10

1 15

2 18

3 22

4 55

5 77

dtype: int64

**9.**

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22,55,77])

s=pd.Series(arr)

print(s)

print(s.loc[:6])

**Output:**

0 10

1 15

2 18

3 22

4 55

5 77

dtype: int64

0 10

1 15

2 18

3 22

4 55

5 77

dtype: int64

**10.**

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22,55,77])

s=pd.Series(arr)

print(s)

print(s.loc[:2])

print(s.loc[3:4])

s.loc[2:3]

**Output:**

0 10

1 15

2 18

3 22

4 55

5 77

dtype: int64

0 10

1 15

2 18

dtype: int64

3 22

4 55

dtype: int64

2 18

3 22

dtype: int64

**11.**

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22,55,77])

s=pd.Series(arr)

print(s)

print(s[1])

print('\n')

print(s.loc[3:4])

s[:3]

**Output:**

0 10

1 15

2 18

3 22

4 55

5 77

dtype: int64

15

3 22

4 55

dtype: int64

0 10

1 15

2 18

dtype: int64

**13.**

import pandas as pd

import numpy as np

arr=np.array(['a','b','c','d'])

s=pd.Series(arr,index=['first','second','third','fourth'])

print(s)

print('\n indexes in Series are:::')

print(s.index)

**Output:**

first a

second b

third c

fourth d

dtype: object

indexes in Series are:::

Index(['first', 'second', 'third', 'fourth'], dtype='object')

**14.**

import pandas as pd

import numpy as np

arr=np.array([10,15,18,22,55,77])

s=pd.Series(arr,index=['A','B','C','D','E','F'])

print(s)

print(s[1:5:2])

print(s[0:6:2])

**Output:**

A 10

B 15

C 18

D 22

E 55

F 77

dtype: int64

B 15

D 22

dtype: int64

A 10

C 18

E 55

dtype: int64

**15.**

import pandas as pd

name=pd.Series(['Hardik','Virat'])

team=pd.Series(['MI','RCB'])

dic={'Name':name,'Team':team}

df=pd.DataFrame(dic)

print(df)

**Output:**

Name Team

0 Hardik MI

1. Virat RCB

**16.**

import pandas as pd

l=[{'Name':'Sachin','SirName':'Bhardwaj'},

   {'Name':'Vinod','SirName':'Verma'},

   {'Name':'Rajesh','SirName':'Mishra'}]

df1=pd.DataFrame(l)

print(df1)

**Output:**

Name SirName

0 Sachin Bhardwaj

1 Vinod Verma

2 Rajesh Mishra

**17.**

import pandas as pd

l=[{'Name':'Sachin','SirName':'Bhardwaj'},

   {'Name':'Vinod','SirName':'Verma'},

   {'Name':'Virat','SirName':'Kohli'},

   {'Name':'Hardik','SirName':'Pandya'},

   {'Name':'Rohit','SirName':'Sharma'},

   {'Name':'KL','SirName':'Rahul'}

   ]

df1=pd.DataFrame(l)

print(df1)

for(row\_index,row\_value)in df1.iterrows():

  print('\nRow index is::',row\_index)

  print('\nRow values is::')

  print(row\_value)

**Output:**

Name SirName

0 Sachin Bhardwaj

1 Vinod Verma

2 Virat Kohli

3 Hardik Pandya

4 Rohit Sharma

5 KL Rahul

Row index is:: 0

Row values is::

Name Sachin

SirName Bhardwaj

Name: 0, dtype: object

Row index is:: 1

Row values is::

Name Vinod

SirName Verma

Name: 1, dtype: object

Row index is:: 2

Row values is::

Name Virat

SirName Kohli

Name: 2, dtype: object

Row index is:: 3

Row values is::

Name Hardik

SirName Pandya

Name: 3, dtype: object

Row index is:: 4

Row values is::

Name Rohit

SirName Sharma

Name: 4, dtype: object

Row index is:: 5

Row values is::

Name KL

SirName Rahul

Name: 5, dtype: object

18.

import pandas as pd

import pandas as pd

l=[{'Name':'Sachin','SirName':'Bhardwaj'},

   {'Name':'Vinod','SirName':'Verma'},

   {'Name':'Virat','SirName':'Kohli'},

   {'Name':'Hardik','SirName':'Pandya'},

   {'Name':'Rohit','SirName':'Sharma'},

   {'Name':'KL','SirName':'Rahul'}

   ]

df1=pd.DataFrame(l)

print(df1)

for(col\_name,col\_value)in df1.iteritems():

  print('\n')

  print('\nColumn index is::',col\_name)

  print('\nColumn values is::')

  print(col\_value)

**Output:**

Name SirName

0 Sachin Bhardwaj

1 Vinod Verma

2 Virat Kohli

3 Hardik Pandya

4 Rohit Sharma

5 KL Rahul

Column index is:: Name

Column values is::

0 Sachin

1 Vinod

2 Virat

3 Hardik

4 Rohit

5 KL

Name: Name, dtype: object

Column index is:: SirName

Column values is::

0 Bhardwaj

1 Verma

2 Kohli

3 Pandya

4 Sharma

5 Rahul

Name: SirName, dtype: object

**19.**

import pandas as pd

empdata={'empid':[101,102,103,104,105,106],

         'ename':['Sachin','Vinod','Lakhbir','Anil','Devinder','Umaselvi'],

         'Doj':['12-01-2012','15-01-2012','05-09-2007','17-01-2012','05-09-2007','16-01-2012']}

df=pd.DataFrame(empdata)

print(df)

**Output:**

empid ename Doj

0 101 Sachin 12-01-2012

1 102 Vinod 15-01-2012

2 103 Lakhbir 05-09-2007

3 104 Anil 17-01-2012

4 105 Devinder 05-09-2007

5 106 Umaselvi 16-01-2012

**20.**

import pandas as pd

s = pd.Series([10,15,18,22])

df=pd.DataFrame(s)

df.columns=['List1']

df['List2']=20

df['List3']=df['List1']+df['List2']

print(df)

**Output:**

List1 List2 List3

0 10 20 30

1 15 20 35

2 18 20 38

3 22 20 42

**21.**

import pandas as pd

s= pd.Series([10,20,30,40])

df=pd.DataFrame(s)

df.columns=['List1']

df['List2']=40

df1=df.drop('List2',axis=1)

df2=df.drop(index=[2,3],axis=0)

print(df)

print(" After deletion::")

print(df1)

print("After row deletion:")

print(df2)

**Output:**

List1 List2

0 10 40

1 20 40

2 30 40

3 40 40

After deletion::

List1

0 10

1 20

2 30

3 40

After row deletion:

List1 List2

0 10 40

1 20 40

**22.**

import pandas as pd

Runs={'TCS':{'Qtr1':2500,'Qtr2':2000,'Qtr3':3000,'Qtr4':2000},

      'WIPRO':{'Qtr1':2800,'Qtr2':2400,'Qtr3':3600,'Qtr4':2400},

      'L&T':{'Qtr1':2100,'Qtr2':5700,'Qtr3':35000,'Qtr4':2100}}

df=pd.DataFrame(Runs)

print(df)

print(df.loc['Qtr3', : ])

print(df.loc['Qtr1':'Qtr3', : ])

**Output:**

TCS WIPRO L&T

Qtr1 2500 2800 2100

Qtr2 2000 2400 5700

Qtr3 3000 3600 35000

Qtr4 2000 2400 2100

TCS 3000

WIPRO 3600

L&T 35000

Name: Qtr3, dtype: int64

TCS WIPRO L&T

Qtr1 2500 2800 2100

Qtr2 2000 2400 5700

Qtr3 3000 3600 35000

23.

import pandas as pd

Runs={'TCS':{'Qtr1':2500,'Qtr2':2000,'Qtr3':3000,'Qtr4':2000},

      'WIPRO':{'Qtr1':2800,'Qtr2':2400,'Qtr3':3600,'Qtr4':2400},

      'L&T':{'Qtr1':2100,'Qtr2':5700,'Qtr3':35000,'Qtr4':2100}}

df=pd.DataFrame(Runs)

print(df)

print(df.loc[ : 'TCS'])

print(df.loc[ : ,'TCS':'WIPRO'])

**Output:**

TCS WIPRO L&T

Qtr1 2500 2800 2100

Qtr2 2000 2400 5700

Qtr3 3000 3600 35000

Qtr4 2000 2400 2100

TCS WIPRO L&T

Qtr1 2500 2800 2100

Qtr2 2000 2400 5700

Qtr3 3000 3600 35000

Qtr4 2000 2400 2100

TCS WIPRO

Qtr1 2500 2800

Qtr2 2000 2400

Qtr3 3000 3600

Qtr4 2000 2400

24.

import pandas as pd

empdata={ 'Doj':['12-01-2012','15-01-2012','05-09-2007',

 '17-01-2012','05-09-2007','16-01-2012'],

 'empid':[101,102,103,104,105,106],

'ename':['Sachin','Vinod','Lakhbir','Anil','Devinder','UmaSelvi']

}

df=pd.DataFrame(empdata)

print(df)

print(df.head(5))

**Output:**

Doj empid ename

0 12-01-2012 101 Sachin

1 15-01-2012 102 Vinod

2 05-09-2007 103 Lakhbir

3 17-01-2012 104 Anil

4 05-09-2007 105 Devinder

5 16-01-2012 106 UmaSelvi

Doj empid ename

0 12-01-2012 101 Sachin

1 15-01-2012 102 Vinod

2 05-09-2007 103 Lakhbir

3 17-01-2012 104 Anil

4 05-09-2007 105 Devinder

**25.**

import pandas as pd

empdata={ 'Doj':['12-01-2012','15-01-2012','05-09-2007',

 '17-01-2012','05-09-2007','16-01-2012'],

 'empid':[101,102,103,104,105,106],

'ename':['Sachin','Vinod','Lakhbir','Anil','Devinder','UmaSelvi']

}

df=pd.DataFrame(empdata)

print(df)

print(df.tail(2))

**Output:**

Doj empid ename

0 12-01-2012 101 Sachin

1 15-01-2012 102 Vinod

2 05-09-2007 103 Lakhbir

3 17-01-2012 104 Anil

4 05-09-2007 105 Devinder

5 16-01-2012 106 UmaSelvi

Doj empid ename

4 05-09-2007 105 Devinder

5 16-01-2012 106 UmaSelvi

**26.**

import pandas as pd

dic={

    'Name':['Sachin Bhardwaj','vinod Verma','rajesh Mishra'],

     'Age':[32,35,40]

}

df=pd.DataFrame(dic,index=[True,False,True])

print(df)

print(df.loc[True])

print()

print('result of the iloc method')

print(df.iloc[1])

**Output:**

Name Age

True Sachin Bhardwaj 32

False vinod Verma 35

True rajesh Mishra 40

Name Age

True Sachin Bhardwaj 32

True rajesh Mishra 40

result of the iloc method

Name vinod Verma

Age 35

dtype: object

**27.**

import pandas as pd

dic1={

    'id':['1','2','3','4','5'],'value1':['A','C','E','G','I'],

    'value2':['B','D','F','H','J']

}

dic2={

    'id':['2','3','6','7','8',],'value1':['K','M','O','Q','S'],

    'value2':['L','N','P','R','T']

}

df1=pd.DataFrame(dic1)

df2=pd.DataFrame(dic2)

df3=pd.concat([df1,df2])

print(df3)

**Output:**

id value1 value2

0 1 A B

1 2 C D

2 3 E F

3 4 G H

4 5 I J

0 2 K L

1 3 M N

2 6 O P

3 7 Q R

4 8 S T

**28.**

import pandas as pd

dic1={

    'id':['1','2','3','4','5'],'value1':['A','C','E','G','I'],

    'value2':['B','D','F','H','J']

}

dic2={

    'id':['2','3','6','7','8',],'value1':['K','M','O','Q','S'],

    'value2':['L','N','P','R','T']

}

dic3={

    'id':['1','2','3','4','5','7','8','9','10','11'],

    'value3':[12,13,14,15,16,17,15,12,13,23]

}

**Output:**

df1=pd.DataFrame(dic1)

df2=pd.DataFrame(dic2)

df3=pd.concat([df1,df2])

df4=pd.DataFrame(dic3)

df5=pd.merge(df3,df4,on='id')

print(df5)

id value1 value2 value3

0 1 A B 12

1 2 C D 13

2 2 K L 13

3 3 E F 14

4 3 M N 14

5 4 G H 15

6 5 I J 16

7 7 Q R 17

8 8 S T 15

**29.**

import pandas as pd

student\_data1={

    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Storey','Bryce Jenson','Ed Bernal','Kwame Morin'],

    'marks':['200','210','190','222','199']

}

student\_data2={

    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Williamson','Dante Morse','Kaiser William','Madeeha Preston'],

    'marks':['201','200','198','219','201']

}

df1=pd.DataFrame(student\_data1)

df1=pd.DataFrame(student\_data2)

df3=pd.merge(df1,df2,on='id',how='outer')

print(df3)

**Output:**

id name marks value1 value2

0 s4 Scarlette Fisher 201 NaN NaN

1 s5 Carla Williamson 200 NaN NaN

2 s6 Dante Morse 198 NaN NaN

3 s7 Kaiser William 219 NaN NaN

4 s8 Madeeha Preston 201 NaN NaN

5 2 NaN NaN K L

6 3 NaN NaN M N

7 6 NaN NaN O P

8 7 NaN NaN Q R

9 8 NaN NaN S T

**30.**

import pandas as pd

student\_data1={

    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Storey','Bryce Jenson','Ed Bernal','Kwame Morin'],

    'marks':['200','210','190','222','199']

}

student\_data2={

    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Williamson','Dante Morse','Kaiser William','Madeeha Preston'],

    'marks':['201','200','198','219','201']

}

df1=pd.DataFrame(student\_data1)

df1=pd.DataFrame(student\_data2)

df3=pd.merge(df1,df2,on='id',how='inner')

print(df3)

**Output:**

Empty DataFrame

Columns: [id, name, marks, value1, value2]

Index: []

**31.**

import pandas as pd

student\_data1={

    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Storey','Bryce Jenson','Ed Bernal','Kwame Morin'],

    'marks':['200','210','190','222','199']

}

student\_data2={

    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Williamson','Dante Morse','Kaiser William','Madeeha Preston'],

    'marks':['201','200','198','219','201']

}

df1=pd.DataFrame(student\_data1)

df1=pd.DataFrame(student\_data2)

df3=pd.merge(df1,df2,on='id',how='right')

print(df3)

**Output:**

id name marks value1 value2

0 2 NaN NaN K L

1 3 NaN NaN M N

2 6 NaN NaN O P

3 7 NaN NaN Q R

4 8 NaN NaN S T

**32.**

import pandas as pd

student\_data1={

    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Storey','Bryce Jenson','Ed Bernal','Kwame Morin'],

    'marks':['200','210','190','222','199']

}

student\_data2={

    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Williamson','Dante Morse','Kaiser William','Madeeha Preston'],

    'marks':['201','200','198','219','201']

}

**Output:**

df1=pd.DataFrame(student\_data1)

df1=pd.DataFrame(student\_data2)

df3=pd.merge(df1,df2,on='id',how='left')

print(df3)

id name marks value1 value2

0 s4 Scarlette Fisher 201 NaN NaN

1 s5 Carla Williamson 200 NaN NaN

2 s6 Dante Morse 198 NaN NaN

3 s7 Kaiser William 219 NaN NaN

4 s8 Madeeha Preston 201 NaN NaN

**33.**

import pandas as pd

student\_data1={

    'id':['s1','s2','s3','s4','s5'],'name':['Danniella Fenton','Ryder Storey','Bryce Jenson','Ed Bernal','Kwame Morin'],

    'marks':['200','210','190','222','199']

}

student\_data2={

    'id':['s4','s5','s6','s7','s8',],'name':['Scarlette Fisher','Carla Williamson','Dante Morse','Kaiser William','Madeeha Preston'],

    'marks':['201','200','198','219','201']

}

df1=pd.DataFrame(student\_data1)

df1=pd.DataFrame(student\_data2)

df3=pd.merge(df1,df2,right\_index=True,left\_index=True)

print(df3)

**Output:**

id\_x name marks id\_y value1 value2

0 s4 Scarlette Fisher 201 2 K L

1 s5 Carla Williamson 200 3 M N

2 s6 Dante Morse 198 6 O P

3 s7 Kaiser William 219 7 Q R

4 s8 Madeeha Preston 201 8 S T