2 Pandas

Aim:

To install pandas and do the DataFrame operations

Description:

1. Declare Empty Dataframe
2. Declare and print the DataFrame Series
3. Add one column and row.
4. Extract any one column and row based on condition
5. Do the functions like Sum ,square root ,min,max functiion , sort and merge of values .
6. Create series from array , Dictionary
7. Create Series using Scalar value,index.

```
#1

import pandas as pd

df=pd.DataFrame()

print(df)
```

o/p: Empty DataFrame
Columns: []
Index: []

```
#2

e=pd.Series(['A','B','C','D'])

id=pd.Series([102,107,109,114])

f={'E':e,'ID':id}

r=pd.DataFrame(f)

print(r)
```

o/p:

	E	ID
0	A	102
1	B	107
2	C	109
3	D	114

```
#3(a)

r['Age']=pd.Series([21,22,23,24])

print(r)
```

o/p:

	E	ID	Age
0	A	102	21
1	B	107	22
2	C	109	23
3	D	114	24

```
#3(b)

D=pd.DataFrame([['G',123,25],['H',143,30]],columns=['E','ID','Age'])

print(pd.concat([r,D]))
```

o/p:

```
  E   ID  Age
0  A  102   21
1  B  107   22
2  C  109   23
3  D  114   24
0  G  123   25
1  H  143   30
```

#4(a)

```
print(r['E'])
```

o/p:

```
0    A
1    B
2    C
3    D
Name: E, dtype: object
```

#4(b)

```
print(r.loc[2])
```

o/p:

```
E       C
ID    109
Age    23
Name: 2, dtype: object
```

#5

```
import numpy as np
i=pd.DataFrame([[2,7]]*3,columns=['p','q'])
print(i.apply(np.sqrt))
print(i.apply(np.sum,axis=0))
ia=pd.DataFrame([[1,5,7],[2,7,8],[3,6,9]],columns=['x','y','z'])
print(ia.agg(['min','max']))
inf=pd.DataFrame(np.random.randn(5,2),index=[3,2,0,4,1],columns=['col3','col4'])
print(inf.sort_index())
i['y'] = [5, 7, 6]
print(pd.merge(i,ia))
```

o/p:

```
          p         q
0  1.414214  2.645751
1  1.414214  2.645751
2  1.414214  2.645751
```

```
p     6
q    21
dtype: int64
     x  y  z
min  1  5  7
max  3  7  9
       col3      col4
0 -0.972850 -1.660178
1 -0.197431 -0.970238
2  1.008251 -0.955465
3  0.617825 -0.845842
4 -2.339865  1.382442
  p q y x z
0 2 7 5 1 7
1 2 7 7 2 8
2 2 7 6 3 9
```

#6

arr = [10, 20, 30, 40, 50]

ser_arr = pd.Series(arr)

print(ser_arr)

dic = {'a': 10, 'b': 20, 'c': 30, 'd': 40, 'e': 50}

ser_dic = pd.Series(dic)

print(ser_dic)

o/p:

```
0    10
1    20
2    30
3    40
4    50
dtype: int64
a    10
b    20
c    30
d    40
e    50
```

dtype: int64

#7

scalar = 10

ser_scalar = pd.Series(scalar, index=['a', 'b', 'c', 'd', 'e'])

print(ser_scalar)

o/p: a 10
b 10
c 10
d 10
e 10
dtype: int64
