INTRO TO PANDAS

LECTURE # 1

PANDAS

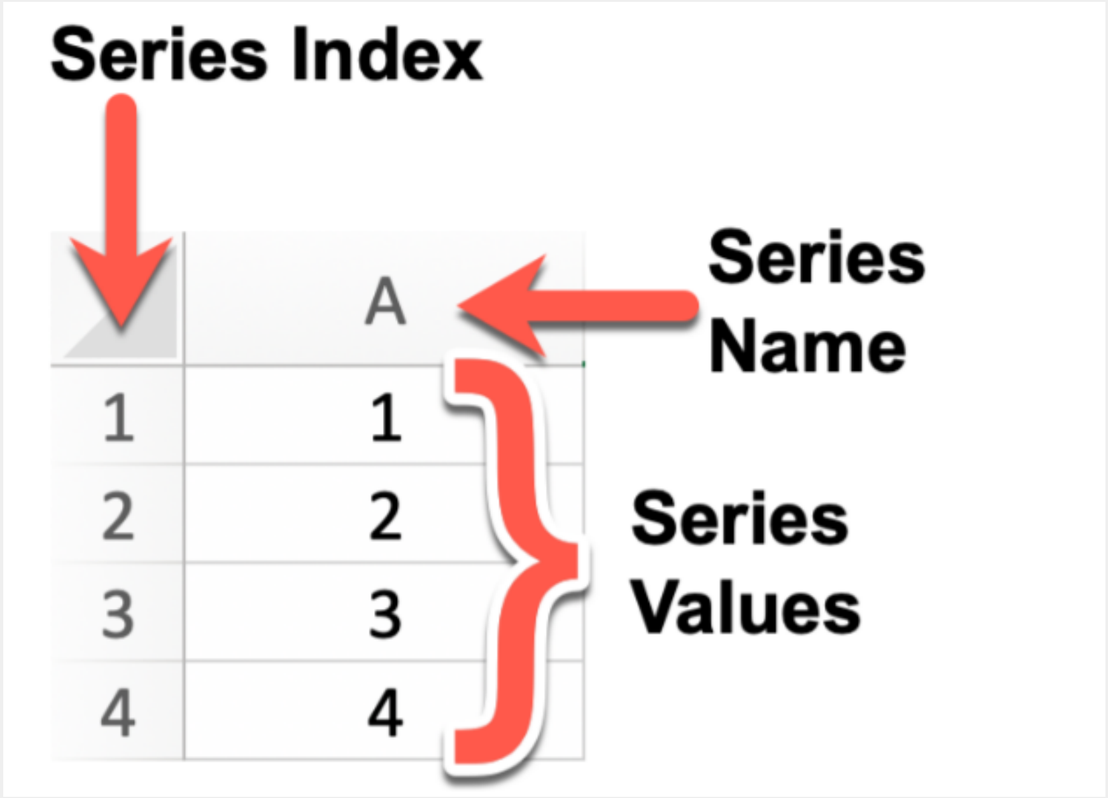
Library for data manipulation and analysis.

pip install pandas

import pandas as pd

SERIES

It is a one-dimensional array holding data of any type.



SERIES

It is a one-dimensional array holding data of any type.

1

```
d = {'a': 1, 'b': 2, 'c': 3}
ser = pd.Series(d)
ser
```

✓ 0.0s

```
a    1
b    2
c    3
dtype: int64
```

2

```
d = {'a': [1,10], 'b': {2,20}, 'c': (3,30)}
ser = pd.Series(d)
ser
```

✓ 0.0s

```
a     [1, 10]
b     {2, 20}
c     (3, 30)
dtype: object
```

FUNC IN SERIES

1

```
courses=["ICA","Accounts","Islamiat", 25]
var=pd.Series(courses)
print(var)
```

✓ 0.0s

```
0         ICA
1    Accounts
2    Islamiat
3          25
dtype: object
```

```
var[1]
```

```
'Accounts'
```

2

```
var=pd.Series(courses,index=['c1','c2','c3'])
print(var)
```

```
c1         ICA
c2    Accounts
c3    Islamiat
dtype: object
```

```
var['c1']
```

```
'ICA'
```

FUNC IN SERIES

pd.Series(<data structure>, index= <list>, name= <name>, dtype= <datatype>)

```
var=pd.Series(courses,index=['c1','c2','c3'],dtype='string',name='Courses')
print(var)
```

⊗ 0.8s

MORE ON SERIES

```
s1=pd.Series(7,index=[1,2,3,4,5])
print(s1)
```

```
1    7
2    7
3    7
4    7
5    7
dtype: int64
```

```
s2=pd.Series(3,index=[1,2,3])
print(s2)
```

```
1    3
2    3
3    3
dtype: int64
```

```
print(s1+s2)
```

```
1    10.0
2    10.0
3    10.0
4     NaN
5     NaN
dtype: float64
```

DATA FRAME

Series is like a column, a DataFrame is the whole table.

```
[27]  marks={"ali":[60,70,90], "sana":[90,30,80]}
      ✓ 0.0s

[28]  S=pd.Series(marks)
      print(S)
      ✓ 0.0s

...   ali     [60, 70, 90]
      sana    [90, 30, 80]
      dtype: object

[29]  D=pd.DataFrame(marks)
      print(D)
      ✓ 0.0s

...      ali  sana
      0   60    90
      1   70    30
      2   90    80
```

DATA FRAME

Series is like a column, a DataFrame is the whole table.

to display = pd.DataFrame(<DataStructure name>, columns = <name>, index = <name>

```
[124]
marks={"ali":[60,70,90], "sana":[90,30,80],"adnan":[78,91,60]}
D=pd.DataFrame(marks,columns=["ali","adnan"],index=["Quiz 1","Quiz 2","Quiz 3"])
print(D)
```

```
...          ali  adnan
Quiz 1   60     78
Quiz 2   70     91
Quiz 3   90     60
```

```
[126]
print(D["ali"]["Quiz 1"])
```

```
...  60
```

to display = <DataFrame name> [<col name>][<row name>]

DATA FRAME

Series is like a column, a DataFrame is the whole table.

Transpose
<DF name> = <DF name>.T

Adding a IF condition

```
D["Pass"]=D["Average"]>=50
```

⊗ 0.1s

To remove a col

```
D= D.drop(columns=["Average", "Minimum"])
```

```
D.pop("Minimum")
print(D)
```

DATA FRAME

Series is like a column, a DataFrame is the whole table.

To add a col

```
D["Average"] = D.mean(axis=1)
D["Minimum"] = D.min(axis=1)
```

```
D.insert(3,"Quiz 4",[70,60,80])
```

Mean of the column
<DataFrame name>**.mean(axis=0)**

DATA FRAME

Series is like a column, a DataFrame is the whole table.

Q1 :
Creating a Series using List of String Data Type. Size of list should be 5. and print the Series

Q2
Repeat Q1 using DataFrame

SERIES

Questions

Q.1- Given the following Series1

A	100
B	200
C	300
D	400
E	500

Write the command to create above Series and then double the value in series and store in another series named Series2

Q.2- State whether True or False

a. A series object is size mutable.

b. A Dataframe object is value mutable

Q.3- Consider a given Series , Series1:

200	700
201	700
202	700
203	700
204	700

Write a program in Python Pandas to create the series and display it.

Q.4- Consider the following Series object, s

IP	95
Physics	89
Chemistry	92
Math	95

i. Write the Python syntax which will display only IP.

ii. Write the Python syntax to increase marks of all subjects by 10.

Q.5- Consider a given series : SQTR

QTR1	50000
QTR2	65890
QTR3	56780
QTR4	89000
QTR5	77900

Write a program in Python Pandas to create and display the series.

Q.6- What will be the output produced by the following programming statements 1 & 2?

```
import pandas as pd
S1=pd.Series(data=[31,41,51])
print(S1>40)          -->Statement1
print(S1[S1>40])      -->Statement2
```

Q.7- Given two series S1 and S2

S1		**S2**	
A	39	A	10
B	41	B	10
C	42	D	10
D	44	F	10

Find the output for following python pandas statements?

a. S1[: 2]*100

b. S1 * S2

c. S2[: : -1]*10

Q.8- Given the following Series S1 and S2:

S1		**S2**	
A	10	A	5
B	20	B	4
C	30	C	6
D	40	D	8

Write the command to find the multiplication of series S1 and S2

Q.9- Consider a given Series , Subject:

ENGLISH	75
HINDI	78
MATHS	82
SCIENCE	86

Write a program in Python Pandas to create this series

DATA FRAME

Write a program in python to find maximum value over index in Data frame.

Ans:

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

# Creating the dataframe
df = pd.DataFrame({"A":[4, 5, 2, 6],
                   "B":[11, 2, 5, 8],
                   "C":[1, 8, 66, 4]})
```

Q5. Create the following dataframe using List of Dictionaries.

	A	B	C
0	1	2	3
1	5	6	8

```
import pandas as _____________
L1 = [["Aman", 45],  ["Ankit", 56],  ["_________",  67]]
DF = pd._____________(L1, _____________=["Name", "Marks"], index=[_________])
print(DF)
```

OUTPUT :

	Name	Marks
1	Aman	45
2	Ankit	56
3	Sunita	67

```
import pandas as pd
L1 = {"Name" : ["Aman", "Ankit", "Sunita"], "Marks" : [45, 56, 67]}
DF = pd.DataFrame(L1, columns = [_____________________], index = [1, 2, 3])
print(DF)
```

OUTPUT :

	Marks	Name
1	45	Aman
2	56	Ankit
3	67	Sunita

DATA FRAME

Series is like a column, a DataFrame is the whole table.

https://www.upgrad.com/blog/pandas-interview-questions-answers-for-freshers-experienced/

http://davburhar.in/File/425/IP%20Class%2012.pdf

https://csiplearninghub.com/pandas-series-class-12-ip-important-questions/

https://www.upgrad.com/blog/pandas-interview-questions-answers-for-freshers-experienced/
http://davburhar.in/File/425/IP%20Class%2012.pdf https://csiplearninghub.com/pandas-series-class-12-ip-important-questions/

https://csiplearninghub.com/important-pandas-dataframe-questions-12-ip/