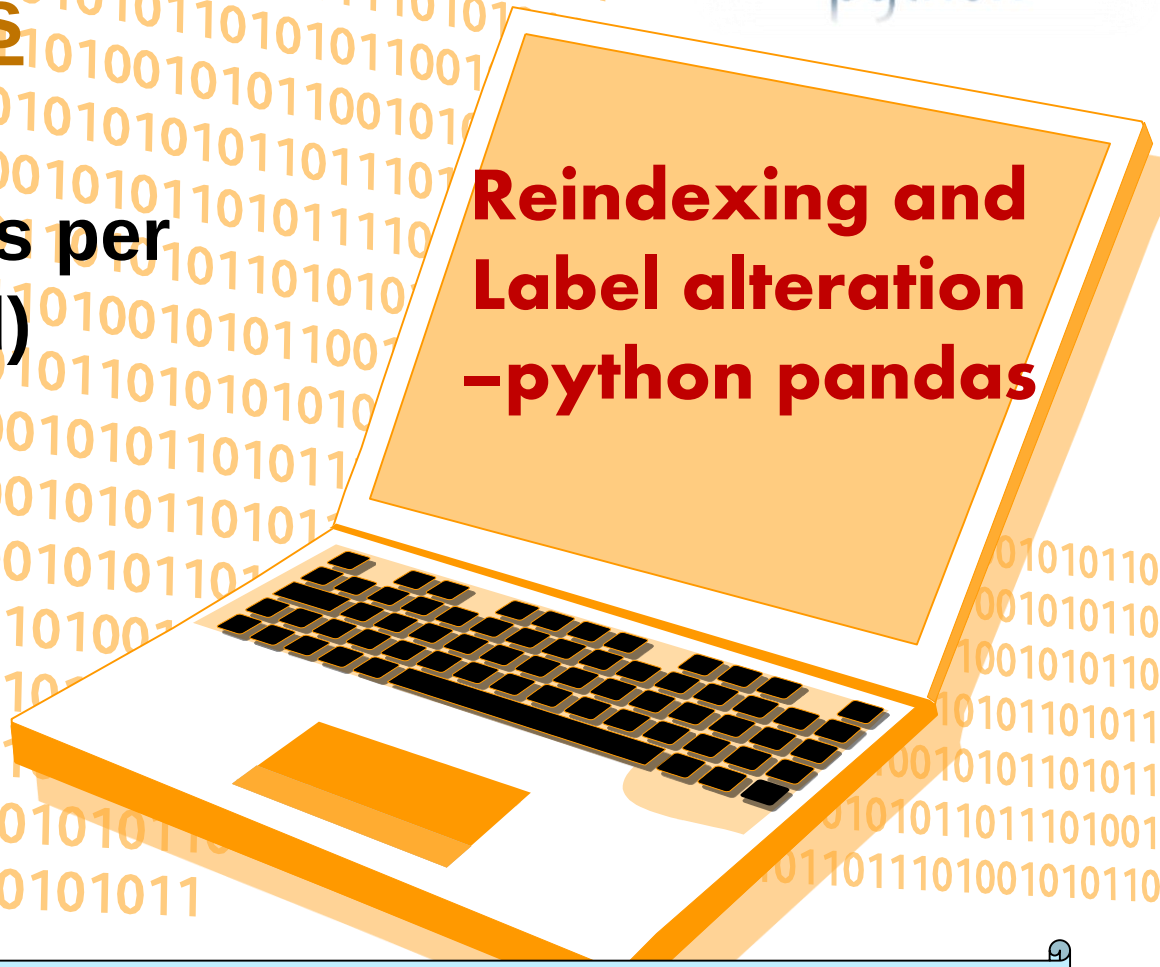
Chapter 4 :

Informatics Practices

Class XII (As per CBSE Board)



Reindexing and Label alteration –python pandas

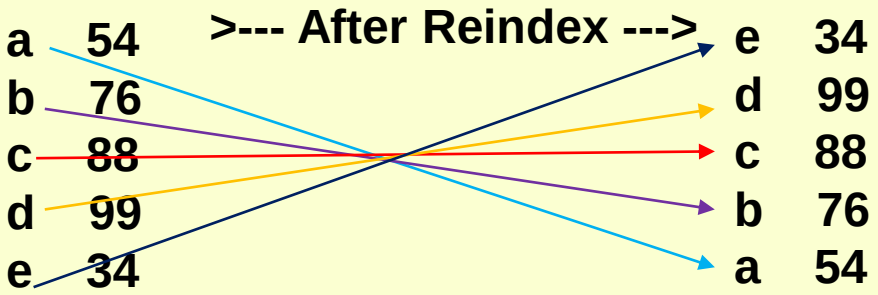
New Syllabus 2019-20

Visit : python.mykvs.in for regular updates

Reindexing – Python Pandas

It is a fundamental operation over pandas series or dataframe.It is a process that makes the data in a Series/data frame conforms to a set of labels. It is used by pandas to perform much of the alignment process. Reindex in python pandas or change the order of the rows and column in python pandas dataframe or change the order of data of series object is possible with the help of reindex() function.

E.g. Given below for series- first column is label(as index) and second column for value



Reindexing – Python Pandas

Reindexing pandas series

The program given below creates a pandas series with some numeric values then index it with a,b,c,d,e labels,then after index is changed to e,d,c,b,a with the help of reindex() function.

e.g. program

```
import pandas as pd
import numpy as np
data = np.array([54,76,88,99,34])
s1 = pd.Series(data,index=['a','b','c','d','e'])
print (s1)
s2=s1.reindex(['e','d','c','b','a'])
print(s2)
```

OUTPUT

```
a    54
b    76
c    88
d    99
e    34
dtype: int32
e    34
d    99
c    88
b    76
a    54
dtype: int32
```

Reindexing – Python Pandas

Reindexing pandas series without label -

Reindex Insert NaN markers where no data exists for a label.In below program f,g are not available as label.

e.g. program

```
import pandas as pd
import numpy as np
data = np.array([54,76,88,99,34])
s1 = pd.Series(data,index=['a','b','c','d','e'])
print (s1)
s2=s1.reindex(['f','g','c','b','a'])
print(s2)
```

OUTPUT

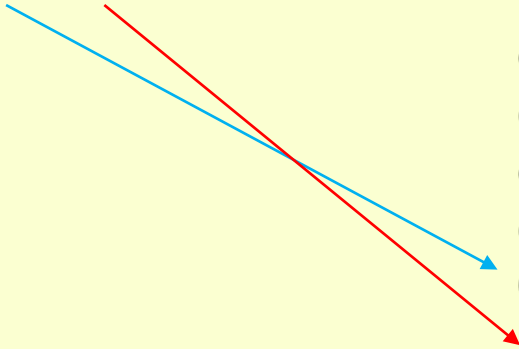
```
a    54
b    76
c    88
d    99
e    34
dtype: int32
f     NaN
g     NaN
c    88.0
b    76.0
a    54.0
dtype: float64
```

Reindexing – Python Pandas

Altering series label –

e.g. program

```
import pandas as pd
import numpy as np
data = np.array([54,76,88,99,34])
s1 = pd.Series(data,index=['a','b','c','d','e'])
print (s1)
s2=s1.rename(index={'a':0,'b':1})
print(s2)
```



```
OUTPUT
a    54
b    76
c    88
d    99
e    34
dtype: int32
0    54
1    76
c    88
d    99
e    34
dtype: int32
```

Reindexing – Python Pandas

Reindexing Rows in pandas Dataframe

e.g.program

```
from collections import OrderedDict
from pandas import DataFrame
import pandas as pd
import numpy as np

table = OrderedDict((
    ("name", ['vishal', 'anil', 'mayur', 'viraj','mahesh']),
    ('age',[15, 16, 15, 17,16]),
    ('weight',  [51, 48, 49, 51,48]),
    ('height',   [5.1, 5.2, 5.1, 5.3,5.1]),
    ('runsscored',   [55,25, 71, 53,51])
))
d = DataFrame(table)
print("DATA OF DATAFRAME")
print(d)
print("DATA OF DATAFRAME AFTER REINDEX")
df=d.reindex([2,1, 0,4,3])
print(df)
```

OUTPUT

DATA OF DATAFRAME

	name	age	weight	height	runsscored
0	vishal	15	51	5.1	55
1	anil	16	48	5.2	25
2	mayur	15	49	5.1	71
3	viraj	17	51	5.3	53
4	mahesh	16	48	5.1	51

DATA OF DATAFRAME AFTER REINDEX

	name	age	weight	height	runsscored
2	mayur	15	49	5.1	71
1	anil	16	48	5.2	25
0	vishal	15	51	5.1	55
4	mahesh	16	48	5.1	51
3	viraj	17	51	5.3	53

Reindexing – Python Pandas

Reindexing columns in pandas Dataframe

e.g.program

```
from collections import OrderedDict
from pandas import DataFrame
import pandas as pd
import numpy as np

table = OrderedDict((
    ("name", ['vishal', 'anil', 'mayur', 'viraj','mahesh']),
    ('age',[15, 16, 15, 17,16]),
    ('weight',  [51, 48, 49, 51,48]),
    ('height',   [5.1, 5.2, 5.1, 5.3,5.1]),
    ('runsscored',   [55,25, 71, 53,51])
))
d = DataFrame(table)
print("DATA OF DATAFRAME")
print(d)
print("DATA OF DATAFRAME AFTER REINDEX")
df=d.reindex(columns=['name','runsscored','age'])
print(df)
```

OUTPUT

DATA OF DATAFRAME

	name	age	weight	height	runsscored
0	vishal	15	51	5.1	55
1	anil	16	48	5.2	25
2	mayur	15	49	5.1	71
3	viraj	17	51	5.3	53
4	mahesh	16	48	5.1	51

DATA OF DATAFRAME AFTER REINDEX

	name	runsscored	age
0	vishal	55	15
1	anil	25	16
2	mayur	71	15
3	viraj	53	17
4	mahesh	51	16

Reindexing – Python Pandas

Altering dataframe labels

e.g.program

```
from collections import OrderedDict
from pandas import DataFrame
import pandas as pd
import numpy as np

table = OrderedDict((
    ("name", ['vishal', 'anil', 'mayur', 'viraj','mahesh']),
    ('age',[15, 16, 15, 17,16]),
    ('weight',  [51, 48, 49, 51,48]),
    ('height',   [5.1, 5.2, 5.1, 5.3,5.1]),
    ('runsscored',   [55,25, 71, 53,51])
))
d = DataFrame(table)
print("DATA OF DATAFRAME")
print(d)
print("DATA OF DATAFRAME AFTER REINDEX")
df=d.rename(index={0:'a',1:'b'})
print(df)
```

OUTPUT

DATA OF DATAFRAME

	name	age	weight	height	runsscored
0	vishal	15	51	5.1	55
1	anil	16	48	5.2	25
2	mayur	15	49	5.1	71
3	viraj	17	51	5.3	53
4	mahesh	16	48	5.1	51

DATA OF DATAFRAME AFTER REINDEX

	name	age	weight	height	runsscored
a	vishal	15	51	5.1	55
b	anil	16	48	5.2	25
2	mayur	15	49	5.1	71
3	viraj	17	51	5.3	53
4	mahesh	16	48	5.1	51