Pandas : 提供 series
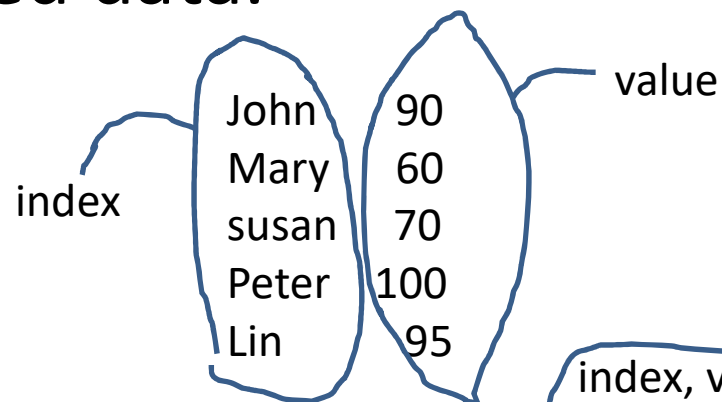
- 建構於NumPy之上的套件(package)
- import pandas as pd
- A Pandas series is a one-dimensional array of indexed data.

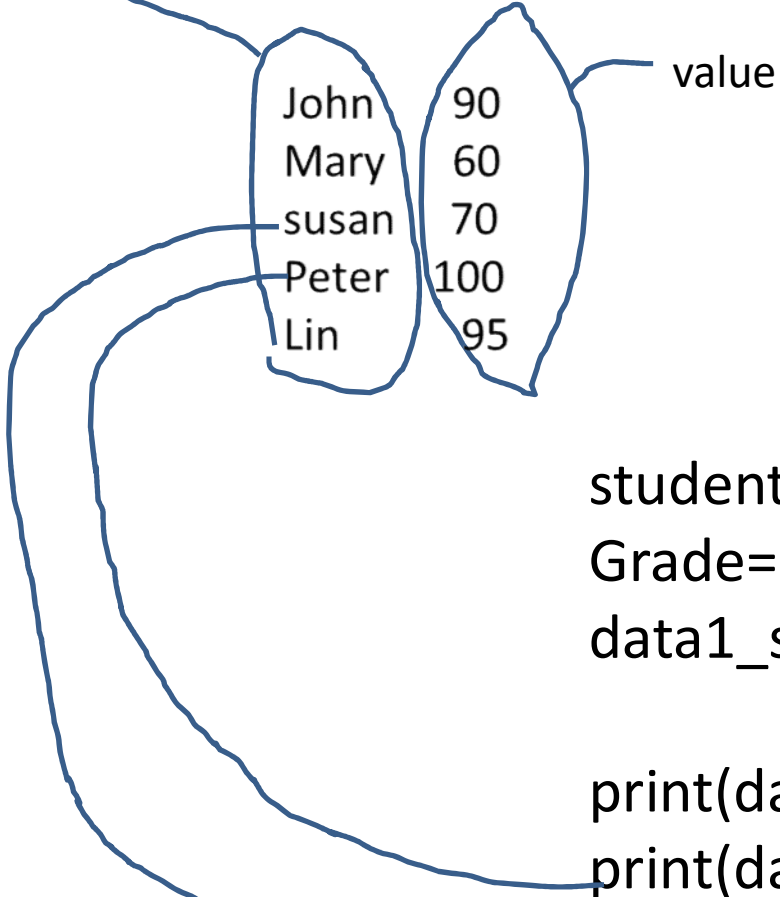| index	value
John	90
Mary	60
susan	70
Peter	100
Lin	95

index, value

The NumPy array has an implicitly defined integer index used to access the values, the Pandas series has explicitly defined index associated with the values.

Series就是有index 的 array

Series: Access the item

The index need not be an integer

John	90
Mary	60
susan	70
Peter	100
Lin	95

value

```
students=['John','Mary','susan','Peter','Lin']
Grade=[90,60,70,100,95]
data1_series = pd.Series(Grade,index=students)
```

創造一個 Series
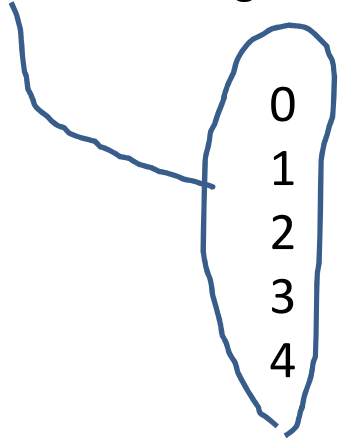
```
print(data1_series[0])
print(data1_series[3])

print(data1_series['susan'])
```

Another Series:

- The series values can be list

- Example:

```
t2=[[1,2,3],[2,3,4],[4,5,6],[0,2,4.5],['go','go','never give up']]
data_5=pd.Series(t2)
print(data_5)
```

The index is an integer

0   [1, 2, 3]

1   [2, 3, 4]

2   [4, 5, 6]

3   [0, 2, 4.5]

4   [go, go, never give up]

Pandas : 提供 DataFrame

- 建構於NumPy之上的套件(package)
- import pandas as pd
- A Pandas DaraFrame is a two-dimensional array of indexed data.

	HK2	Hk1	Midterm
John	A	A	90
Mary	B	B	60
susan	A	A	70
Peter	C	C	100
Lin	A	A	95

創造一個 DataFrame

pd1=pd.DataFrame({'Midterm':midterm_series, 'Hk1':hk1_series, 'HK2':hk2_series})

取值

- 單欄:

- 多欄:

```
col=pd1['Hk1']
print(col)
John    A
Mary    B
susan   A
Peter   C
Lin     A
Name: Hk1, dtype: object
```

```
cols=pd1[['Hk1','HK2']]
print(cols)
       Hk1    HK2
John    A    A
Mary    B    A
susan   A    A
Peter   C    B
Lin     A    A
```

Create the DataFrame from Series

```
students=['John','Mary','susan','Peter','Lin']
Grade=[90,60,70,100,95]
midterm_series = pd.Series(Grade,index=students)
hk1=['A','B','A','C','A']
hk1_series=pd.Series(hk1,students)
hk2=['A','A','A','B','A']
hk2_series=pd.Series(hk2,students)
pd1=pd.DataFrame({'Midterm':midterm_series ,'Hk1':hk1_series ,'HK2':hk2_series })
print(pd1)
```

Series

3 個 Series 組成

	HK2	Hk1	Midterm
John	A	A	90
Mary	A	B	60
Susan	A	A	70
Peter	B	C	100
Lin	A	A	95

JSON

- JSON (**J**ava**S**cript **O**bject **N**otation)
- Example:
- [展覽資訊| 政府資料開放平臺](#)
- {'version': xxx ,

   'UID': xxx ,

   'title': xxx ,

   ... ,

   'showInfo':[{'AA': , 'BB': ,'CC': }]

   }
- 每一筆資料都是一個dict

JSON

- How to create a DataFrame from the given many dicts ?

- import pandas as pd

- df1=pd.DataFrame([{   }, {   }, ...,{   }])

List of dicts

How to append DataFrame ?

Only one dict

- df2=pd.DataFrame([{   }])
- df1.append(df2)
- Unlike append of list, append in Panda does not modify the original object