

Problem-Based Learning

Python 3

Lecture: 10-Pandas (2)

Online Python

- <https://www.programiz.com/python-programming/online-compiler/>

Lecture Notes

- <https://web.phy.ntnu.edu.tw/~hongyi/?url=notes>

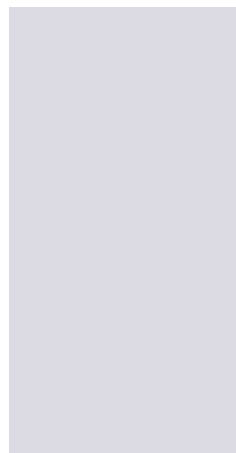


串列

- 五個學生國英數成績如下，計算每個學生的平均成績，並附加上原表格。

	國文	英文	數學
學生1	96	65	73
學生2	88	76	82
學生3	92	84	89
學生4	82	73	64
學生5	70	83	68

+



矩陣(Numpy) 、串列(List) 、表格(Pandas)

- 串列List

```
[[96, 65, 73], [88, 76, 82], [92, 84, 89]]
```

- 矩陣Numpy array

```
[[96 65 73]  
 [88 76 82]  
 [92 84 89]]
```

- 表格Pandas DataFrame

	A	B	C ← columns
a	96	65	73
b	88	76	82
c	92	84	89

↑
index

Python list → Pandas DataFrame

- `import pandas as pd`
- `a = [[1, 2, 3],
 [6, 7, 8],
 [2, 3, 4]]`
- `ind = ['a', 'b', 'c']`
- `col = ['A', 'B', 'C']`
- `d = pd.DataFrame(a, index=ind, columns=col)`
- `print(d)`

	A	B	C
a	1	6	2
b	2	7	3
c	3	8	4

Pandas DataFrame：增加直行單筆資料

- `b = d`
- `c = [7, 8, 9]`
- `b['D'] = c`
- `print(b)`

	A	B	C	D
a	1	6	2	7
b	2	7	3	8
c	3	8	4	9

Pandas DataFrame：增加直行多筆資料

- `b = d`
- `a2 = [[7,3],
 [8,4],
 [9,5]]`
- `col2 = ['D', 'E']`
- `d2 = pd.DataFrame(a2, index=ind, columns=col2)`
- `b = pd.concat([b,d2], axis=1)`
- `print(b)`

	A	B	C	D	E	← axis=1
a	1	6	2	7	3	
b	2	7	3	8	4	
c	3	8	4	9	5	

↑
axis=0

Pandas DataFrame：增加橫列單筆資料

- `b = d`
- `c = [4, 9, 5]`
- `b.loc['d'] = c`
- `print(b)`

	A	B	C
a	1	6	2
b	2	7	3
c	3	8	4
d	4	9	5

Pandas DataFrame：增加橫列單筆資料

- `b = d`
- `c = [4, 9, 5]`
- `d2 = pd.DataFrame([c], index=['d'], columns=col)`
- `b = pd.concat([b,d2], axis=0)`
- `print(b)`

	A	B	C
a	1	6	2
b	2	7	3
c	3	8	4
d	4	9	5

Pandas DataFrame：增加橫列多筆資料

- `b = d`
- `a2 = [[4,9,5],`
 `[5,1,6]]`
- `ind2 = ['d', 'e']`
- `d2 = pd.DataFrame(a2, index=ind2, columns=col)`
- `b = pd.concat([b,d2], axis=0)`
- `print(b)`

	A	B	C ← axis=1
a	1	6	2
b	2	7	3
c	3	8	4
d	4	9	5
e	5	1	6

↑
axis=0

Pandas DataFrame：計算 mean

- `b = d`
- `c = b.mean(axis=1)`
- `print(c)`
- `print(c.index)`
- `print(c.values)`

	A	B	C
a	1	6	2
b	2	7	3
c	3	8	4

← axis=1

↑ axis=0

a	3.0
b	4.0
c	5.0

↑ index ↑ values

Pandas series to Pandas DataFrame

- `c = pd.DataFrame({"平均":c.values}, index=c.index)`
- `print(c)`

Problem

- 計算每個學生的平均成績，並附加到原表格。
- 找出每一個學生的最高分和最低分 and 其所對應的學科，並附加到原表格。

	國文	英文	數學
學生1	96	65	73
學生2	88	76	82
學生3	92	84	89
學生4	82	73	64
學生5	70	83	68

+

平均

+

最高分

+

最低分

提示：
b.min(axis=?)
b.max(axis=?)
b.mean(axis=?)

b.idxmin(axis=?)
b.idxmax(axis=?)