

Problem-Based Learning

Python 3

Lecture: 08-list-example

Online Python

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

Lecture Notes

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



Python list, Numpy array, Pandas DataFrame

✓ Python list:

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

✓ Numpy array:

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

✓ Pandas DataFrame:

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

Numpy Array → Pandas DataFrame

- `import pandas as pd`
- `import numpy as np`
- `a = np.array(
 [[96, 65, 73],
 [88, 76, 82],
 [92, 84, 89],
 [82, 73, 64],
 [70, 83, 68]]
)`
- `d = pd.DataFrame(a)`

Python List → Pandas DataFrame

- `import pandas as pd`
- `a = [[96, 65, 73],
[88, 76, 82],
[92, 84, 89],
[82, 73, 64],
[70, 83, 68]]`
- `d = pd.DataFrame(a)`

Problem

- 五個學生國英數成績如下：
- 按照平均成績的大小**排序**
- 計算各科的平均值和標準差，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

- 使用Pandas
- 僅使用for迴圈來計算平均成績和標準差
- 僅使用for迴圈來進行大小的排序
- 使用iloc[i][j]來存取特定位置的元素

Solution (1)

- `import numpy as np`
- `import pandas as pd`

Solution (1)

- `a = [[96, 65, 73],
[88, 76, 82],
[92, 84, 89],
[82, 73, 64],
[70, 83, 68]]`
- `ind = ["學生1", "學生2", "學生3", "學生4", "學生5"]`
- `col = ["國文", "英文", "數學"]`
- `d = pd.DataFrame(a, index=ind, columns=col)`
- `print(d)`

Solution (1)

- `a2 = np.zeros(5)`
- `for i in range(d.shape[0]):`
- `for j in range(d.shape[1]):`
- `a2[i] = a2[i] + d.iloc[i,j]`
- `a2[i] = a2[i] / d.shape[1]`
- `print(a2)`
- `d2 = pd.DataFrame(a2, index=ind, columns=['平均'])`
- `print(d2)`
- `d = pd.concat([d,d2], axis=1)`
- `print(d)`

Solution (1)

- `for i in range(d.shape[0]):`
- `for j in range(i):`
- `if d.iloc[i,3] < d.iloc[j,3]:`
- `for k in range(d.shape[1]):`
- `d.iloc[i,k], d.iloc[j,k] = d.iloc[j,k],`
 `d.iloc[i,k]`
- `ind[i], ind[j] = ind[j], ind[i]`
- `d.index = ind`
- `print(d)`

Solution (1)

- `a3 = np.zeros(3)`
- `for j in range(d.shape[1]):`
- `for i in range(d.shape[0]):`
- `a3[j] = a3[j] + d.iloc[i,j]`
- `a3[j] = a3[j] / d.shape[0]`
- `print(a3)`
- `d.loc['平均'] = a3`
- `print(d)`

Solution (1)

- `a4 = np.zeros(3)`
- `for j in range(d.shape[1]):`
 - `s = 0`
 - `for i in range(d.shape[0]):`
 - `s = s + (d.iloc[i,j] - a3[j])**2`
 - `a4[j] = np.sqrt(s / d.shape[0])`
- `print(a4)`
- `d.loc['標準差'] = a4`
- `print(d)`