



การบ้านปฏิบัติการ 11 Extra Credit

1D Lists and Tuples

ข้อกำหนด

- **Extra Credit** เป็นงานเดี่ยว ไม่อนุญาตให้ขอคำปรึกษาจากผู้อื่น หากมีการทุจริตจะได้คะแนน ติดลบ เทียบเท่าคะแนน Extra Credit ทั้งสองฝ่าย

(HW11_EX_5XXXXXXXXX.py) พิจารณาโปรแกรมภาษา python เพื่อทำหน้าที่คำนวณค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐานของคะแนนของนักศึกษา โดยมีการทำงานแบ่งเป็น 3 ขั้นตอน (Step) ดังแสดงในฟังก์ชัน main() ด้านล่าง

```
06 def main():
07     # STEP1: รับ input แล้วแยกเป็น list ของรหัสนักศึกษา, ชื่อ-นามสกุล, และคะแนน
08     id_list, name_list, score_list = input_data()
09     print("ID List = ", id_list)
10     print("name List = ", name_list)
11     print("score List = ", score_list)
12
13     # STEP2: คำนวณค่าเฉลี่ย (mean) ของคะแนน และสร้าง list ของรหัสนักศึกษาที่ได้คะแนนต่ำกว่าค่าเฉลี่ย พร้อมคืนค่า
14     mean, below_mean_list = id_below_mean(id_list, score_list)
15     print("\nMean =", mean)
16     print("ID of students with below average scores:")
17     print(below_mean_list)
18
19     # STEP3: คำนวณค่าเบี่ยงเบนมาตรฐาน (SD) สร้าง list ชื่อ (ไม่รวมนามสกุล)
20     # ของนักศึกษาที่ได้คะแนนในช่วง mean - sd ถึง mean + sd
21     sd, names_in_range_list = name_in_SD_range(name_list, score_list, mean)
22     print("\nSD. =", sd)
23     print("First names of students with scores in range (%f, %f): " %
24           (mean - sd, mean + sd))
25     print(names_in_range_list)
```

ตัวอย่างการ Run

```
$ python T111_Q2_1_5XXXXXXXXX.py < scores.txt
ID List = [580510001, 580510002, 580510003, 580510004, 580510005]
name List = ['Jacky Chang', 'Billy Smith', 'Lisa Mark', 'Maria Brook', 'Johny Walker']
score List = [61.12, 85.25, 75.48, 41.96, 55.38]

Mean = 63.838
ID of students with below average scores:
[580510001, 580510004, 580510005]

SD. = 15.176179229305378
First names of students with scores in range (48.661821, 79.014179):
['Jacky', 'Lisa', 'Johny']
```

คำสั่ง

- ให้ Download ไฟล์ที่มีฟังก์ชัน main() จาก <https://goo.gl/NUorlj>
- เขียนฟังก์ชันต่อไป่นี้ เพื่อให้โปรแกรมทำงานสมบูรณ์
 - `input_data()`
 - `id_below_mean(id_list, score_list)`
 - `name_in_SD_range(name_list, score_list, mean)`

1) (STEP1: 35 คะแนน) ฟังก์ชัน `input_data()` มีหน้าที่รับชุดข้อมูลด้วยวิธี **input redirection** จากไฟล์ข้อมูล

Input

ไฟล์ข้อมูล ซึ่งแต่ละบรรทัดจะอยู่ในรูป

```
student_id;first_name last_name;score
```

ตามลำดับโดยข้อมูลดังกล่าวจะคั่นด้วย เครื่องหมาย **Semicolon** `' ; '`

ตัวอย่างข้อมูล ในรูปของ **text file** ชื่อ `scores.txt`

```
580510001;Jacky Chang;61.12
580510002;Billy Smith;85.25
580510003;Lisa Mark;75.48
580510004;Maria Brook;41.96
580510005;Johny Walker;55.38
```

- ข้อมูลตัวแรกในแต่ละแถว คือ รหัสนักศึกษา
- ข้อมูลตัวที่สองในแต่ละแถว คือ ชื่อ-นามสกุล คั่นด้วยวรรค `' '`
- ข้อมูลตัวที่สามในแต่ละแถว คือ คะแนน

Output

ฟังก์ชันจะคืนค่าในรูปแบบดังต่อไปนี้

```
id_list, name_list, score_list
```

- `id_list` คือ list ของข้อมูลรหัสนักศึกษา (int)
- `name_list` คือ list ของข้อมูล ชื่อ-นามสกุล (str)
- `score_list` คือ list ของข้อมูล คะแนน (float)

2) (STEP2: 35 คะแนน) ฟังก์ชัน `id_below_mean(id_list, score_list)` เพื่อตรวจสอบว่านักศึกษารหัสใดบ้างที่ได้คะแนนต่ำกว่าค่าเฉลี่ย (mean) ของคะแนนทั้งหมด

Input

ฟังก์ชันพารามิเตอร์แต่ละตัวคือ

- `id_list` คือ list ของข้อมูลรหัสนักศึกษา ที่ได้จากฟังก์ชัน `input_data()`
- `score_list` คือ list ของข้อมูล คะแนน ที่ได้จากฟังก์ชัน `input_data()`

Output

ฟังก์ชันจะคืนค่าในรูปแบบดังต่อไปนี้

```
mean, below_mean_list
```

- `mean` คือ ค่าเฉลี่ยของคะแนนทั้งหมด
- `below_mean_list` คือ list ของรหัสนักศึกษา ที่ได้คะแนนต่ำกว่า mean

3) (STEP3: 30 คะแนน) ฟังก์ชัน `name_in_SD_range(name_list, score_list, mean)` เพื่อหาชื่อของนักศึกษา (ไม่รวมนามสกุล) ที่ได้คะแนนในช่วง $(\text{mean} - \text{SD})$ ถึง $(\text{mean} + \text{SD})$ โดยที่ SD ย่อมาจาก Standard Deviation คือ ค่าเบี่ยงเบนมาตรฐาน

Input

ฟังก์ชันพารามิเตอร์แต่ละตัวคือ

- `name_list` คือ list ของข้อมูลชื่อ-นามสกุล ที่ได้จากฟังก์ชัน `input_data()`
- `score_list` คือ list ของข้อมูลคะแนน ที่ได้จากฟังก์ชัน `input_data()`
- `mean` คือ ค่าเฉลี่ยของคะแนนทั้งหมดที่ได้จากฟังก์ชัน `id_below_mean()`

Output

ฟังก์ชันจะคืนค่าในรูปแบบดังต่อไปนี้

`sd, names_in_range_list`

- `sd` คือค่าเบี่ยงเบนมาตรฐาน
- `names_in_range_list` คือ list ชื่อแรกของนักศึกษาที่ได้คะแนนอยู่ในช่วง $(\text{mean} - \text{SD})$ ถึง $(\text{mean} + \text{SD})$

วิธีคำนวณหาค่า SD จากตัวอย่างข้อมูลในข้อ STEP1

$$SD_N = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$$

$$SD_N = \sqrt{\frac{(61.12-63.84)^2 + (85.25-63.84)^2 + (75.48-63.84)^2 + (41.96-63.84)^2 + (55.38-63.84)^2}{5}}$$

$$= 15.176179$$

หมายเหตุ ค่าทศนิยม 2 ตำแหน่งที่แสดงในการคำนวณในสูตรด้านบน เป็นค่าประมาณเพื่อใช้ในการแสดงผลตัวอย่างเท่านั้น (ไม่ต้องทำการปัดทศนิยม)

ดังนั้นช่วงค่าคะแนนที่สนใจคือ

$$\begin{aligned} (\text{mean} - \text{SD}) &\leq \text{score} \leq (\text{mean} + \text{SD}) \\ (63.838000 - 15.176179) &\leq \text{score} \leq (63.838000 + 15.176179) \\ 48.661821 &\leq \text{score} \leq 79.014179 \end{aligned}$$