

## Chapter 3 - Exercise 2: Đọc và chuyển dữ liệu, sau đó tính BMI theo điều kiện, truy xuất dữ liệu

Dữ liệu được trích xuất từ

[http://wiki.stat.ucla.edu/socr/index.php/SOCR\\_Data\\_MLB\\_HeightsWeights](http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data_MLB_HeightsWeights)

*Ghi chú:* Major League Baseball (MLB) là giải đấu bóng chày chuyên nghiệp. Major League Baseball có tổng cộng 30 đội bóng đến từ nhiều bang khác nhau của Mỹ và Canada (29 đội từ Mỹ và 1 đội từ Canada). MLB luôn được sự quan tâm lớn của hầu hết fan bóng chày trên toàn thế giới, và cũng được xem là giải đấu nổi tiếng và uy tín nhất, tập hợp những cầu thủ có trình độ cao nhất trong bộ môn này. Dữ liệu **heights** (tính theo inches) và **weights** (tính theo pounds) là chiều cao và cân nặng của các cầu thủ có tham gia 1 số giải của MLB.

Cho tập tin dữ liệu heights\_1.txt, weights\_1.txt

### ▼ Các kiến thức sử dụng trong bài tập:

1. Sử dụng các phép toán số học trên mảng
2. Truy xuất phần tử của mảng thông qua chỉ số

### ▼ Thực hiện các yêu cầu sau, và đối chiếu với kết quả cho trước:

```
import numpy as np

# Đọc file
with open('data\heights_1.txt', 'r') as f:
    x = f.read()
x=x.replace('[', '')
x=x.replace(']', '')
height = x.split(sep=',')
height
arr_height = np.asarray(height, dtype=int)
arr_height

array([74, 74, 72, ..., 75, 75, 73])

# Chép dữ liệu từ tập tin heights_1.txt vào list height
height = [74, 74, 72, 72, 73, 69, 69, 71, 76, 71, 73, 73, 74, 74, 69, 70, 73, 75, 78, 79,
# Chép dữ liệu từ tập tin weights_1.txt vào list weight
weight = [180, 215, 210, 210, 188, 176, 209, 200, 231, 180, 188, 180, 185, 160, 180, 185,
```

```
print(len(height))
print(len(weight))
```

```
1015
1015
```

```
# Câu 1: Tạo numpy array arr_height từ list height
arr_height = np.asarray(height)
# In danh sách các phần tử của arr_height
print(arr_height)
# Cho biết kích thước (shape) của arr_height
print('Kích thước arr_height = ', arr_height.shape)
```

```
[74 74 72 ... 75 75 73]
Kích thước arr_height = (1015,)
```

► Nhấn vào đây để xem kết quả !

```
# Câu 2: Tạo numpy array arr_weight từ list weight
arr_weight = np.asarray(weight)
# In danh sách các phần tử của arr_weight
print(arr_weight)
# Cho biết kích thước (shape) của arr_weight
print(arr_weight.shape)
```

```
[180 215 210 ... 205 190 195]
(1015,)
```

► Nhấn vào đây để xem kết quả !

```
# Câu 3: Cho hệ số quy đổi từ inch sang m là 0.0254
# Tạo array arr_height_m dựa trên công thức: arr_height * hệ số quy đổi
arr_height_m = arr_height * .0254
# In danh sách các phần tử của arr_height_m
print(arr_height_m)
```

```
[1.8796 1.8796 1.8288 ... 1.905 1.905 1.8542]
```

► Nhấn vào đây để xem kết quả !

```
# Câu 4: Cho hệ số quy đổi từ pound sang kg là 0.453592
# Tạo array arr_weight_kg dựa trên công thức: arr_weight * hệ số quy đổi
arr_weight_kg = arr_weight * 0.453592
# In danh sách các phần tử của arr_weight_kg
print(arr_weight_kg)
```

```
[81.64656 97.52228 95.25432 ... 92.98636 86.18248 88.45044]
```

► Nhấn vào đây để xem kết quả !

```
# Câu 5: Tính giá trị BMI (Body Mass Index) của arr_height_m và arr_weight_kg và lưu vào a
# Gợi ý: Tính theo công thức BMI = Cân nặng / (Chiều cao * Chiều cao)
arr_bmi = arr_weight_kg/(arr_height_m**2)
# In ra danh sách các phần tử của arr_bmi
print(arr_bmi)

[23.11037639 27.60406069 28.48080465 ... 25.62295933 23.74810865
 25.72686361]
```

► Nhấn vào đây để xem kết quả !

```
# Câu 6: Cho biết giá trị cân nặng ở vị trí index = 50 trong arr_weight_kg
print(arr_weight_kg[50])

90.7184
```

► Nhấn vào đây để xem kết quả !

```
# Câu 7: Tạo array arr_height_m_100 bao gồm các phần tử có vị trí index từ 100 đến 110 (lấy)
arr_height_m_100 = arr_height_m[100:111]
print(arr_height_m_100)

[1.8542 1.8796 1.8288 1.8542 1.7526 1.8288 1.8542 1.905 1.905 1.8542
 1.8288]
```

► Nhấn vào đây để xem kết quả !

```
# Câu 8: Tạo và in kết quả của biểu thức điều kiện dùng để lấy ra các cầu thủ bóng chày có bmi_less_21 = arr_bmi < 21
print(bmi_less_21)
# Áp dụng biểu thức điều kiện đã tạo để in ra các cầu thủ bóng chày có bmi < 21 trong arr_
print(arr_bmi[bmi_less_21])

[False False False ... False False False]
[20.54255679 20.54255679 20.69282047 20.69282047 20.34343189 20.34343189
 20.69282047 20.15883472 19.4984471 20.69282047 20.9205219 ]
```

► Nhấn vào đây để xem kết quả !

```
# Câu 9: Cho biết chiều cao trung bình và cân nặng trung bình của các cầu thủ
avg_height = np.mean(arr_height_m)
print('Chiều cao trung bình:', avg_height)

avg_weight = np.mean(arr_weight_kg)
print('Cân nặng trung bình: ', avg_weight)

Chiều cao trung bình: 1.8717172413793102
Cân nặng trung bình: 91.33019058916256
```

# Câu 10: Cho biết chiều cao và cân nặng lớn nhất của các cầu thủ

```
max_height = np.max(arr_height_m)
print('Chiều cao lớn nhất: ', max_height)
```

```
max_weight = np.max(arr_weight_kg)
print('Cân nặng lớn nhất: ', max_weight)
```

```
Chiều cao lớn nhất:  2.1082
Cân nặng lớn nhất:  131.54167999999999
```

# Câu 11: Cho biết chiều cao và cân nặng nhỏ nhất của các cầu thủ

```
min_height = np.min(arr_height_m)
print('Chiều cao thấp nhất: ', min_height)
```

```
min_weight = np.min(arr_weight_kg)
print('Cân nặng bé nhất: ', min_weight)
```

```
Chiều cao thấp nhất:  1.7018
Cân nặng bé nhất:  68.0388
```