

Chapter 3

Chapter 3 - Exercise 1: Thực hiện các yêu cầu sau:

Phần 1: Các thao tác cơ bản trên mảng (1)

1. Tạo numpy array arr có giá trị từ 0-9. Hiển thị các phần tử có trong arr, xem kiểu dữ liệu và kích thước của arr.
2. Từ array arr ở câu 1 => tạo arr_odd và arr_even.
3. Từ array arr ở câu 1 => tạo arr_update_1 với các phần tử chẵn giữ nguyên, các phần tử lẻ thay bằng 100.

Phần 2: Các thao tác cơ bản trên mảng (2)

1. Cho 2 array arr_a = [1,2,3,2,3,4,3,4,5,6] và arr_b = [7,2,10,2,7,4,9,4,9,8], tạo array arr_c chỉ lấy duy nhất các phần tử xuất hiện ở cả arr_a và arr_b.
2. Từ arr_a và arr_b ở câu 1 => Tạo arr_d chứa các phần tử chỉ xuất hiện ở arr_a.
3. Cho arr_e = np.array([2, 6, 1, 9, 10, 3, 27, 8, 6, 25, 16]), hãy tạo arr_f chỉ chứa các phần tử có giá trị từ 5 đến 10 của arr_e.

Phần 3: Các thao tác cơ bản trên mảng (3)

1. Tạo arr_zeros có 10 phần tử 0, cập nhật phần tử ở vị trí thứ 5 là 1.
2. Tạo arr_h có giá trị từ 10 đến 24. In danh sách các phần tử theo thứ tự đảo ngược của arr_h.
3. Cho arr_k = np.array([1, 2, 0, 8, 2, 0, 1, 3, 0, 5, 0]), tạo arr_l từ arr_k với các phần tử khác 0.
4. Từ arr_l của câu 3, thêm 2 phần tử có giá trị là 10 và 20 vào cuối array.
5. Từ array của câu 4, thêm phần tử có giá trị 100 vào vị trí có index = 5.
6. Từ array của câu 5, xóa các phần tử tại vị trí có index = 0, 1, 2.

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

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* => hãy chép dữ liệu từ tập tin này vào list là *height*, *weight*

1. Tạo numpy array *arr_height* từ list *height*.
2. Tạo numpy array *arr_weight* từ list *weight*.
3. Cho hệ số quy đổi từ inch sang m là 0.0254, tạo *arr_height_m* dựa trên công thức: *arr_height* * hệ số quy đổi.
4. Cho hệ số quy đổi từ pound sang kg là 0.453592, tạo *arr_weight_kg* dựa trên công thức: *arr_weight* * hệ số quy đổi.
5. Tính BMI của *arr_height_m* và *arr_weight_kg* theo công thức BMI = Cân nặng / (Chiều cao * Chiều cao), và lưu vào *arr_bmi*.
6. Cho biết giá trị cân nặng ở vị trí *index = 50* trong *arr_weight_kg*
7. Tạo *arr_height_m_100* gồm các phần tử có vị trí *index* từ 100 đến 110 (lấy cả *index 110*) trong *arr_height_m*
8. Cho biết các cầu thủ bóng chày có *bmi < 21* trong *arr_bmi*
9. Cho biết chiều cao trung bình và cân nặng trung bình của các cầu thủ
10. Cho biết chiều cao và cân nặng lớn nhất của các cầu thủ
11. Cho biết chiều cao và cân nặng nhỏ nhất của các cầu thủ

Chapter 3 - Exercise 3: Thao tác trên mảng nhiều chiều

1. Tạo array *arr* có kích thước 3x3 với các giá trị True
2. Cho *arr_1D = np.array([0 1 2 3 4 5 6 7 8])*. Tạo array 2 chiều có kích thước 3x3 từ *arr_1D*, và lưu vào *arr_2D* Trong *arr_2D*, chuyển cột 1 sang cột 3 và ngược lại.
3. Từ *arr_2D* của câu 2 (sau khi đổi thứ tự cột), chuyển dòng 1 sang dòng 2 và ngược lại.
4. Từ *arr_2D* của câu 3, đảo ngược các dòng của *arr_2D*.
5. Từ *arr_2D* của câu 4, đảo ngược các cột của *arr_2D*.

6. Cho `arr_2D_null = np.array([[1, 2, 3], [np.NaN, 5, 6], [7, np.NaN, 9], [4, 5, 6]])`,
Kiểm tra trong array có giá trị rỗng không?
7. Từ `arr_2D_null` của câu 6, thay thế giá trị null bằng 0.

Chapter 3 - Exercise 4: Thao tác dữ liệu mảng dữ liệu baseball

Cho tập tin `baseball_2D.txt` => chép dữ liệu từ tập tin vào list là `baseball`

Dữ liệu baseball cho biết chiều cao (cột 1) tính theo inch và cân nặng (cột 2) tính theo pounds của các cầu thủ

1. Tạo 2D numpy array tên `np_baseball` từ `baseball`. Xem kiểu dữ liệu và kích thước của `np_baseball`
2. In các giá trị của dòng thứ 50 trong `np_baseball`
3. Tạo numpy array `np_weight` với dữ liệu được lấy từ cột hai của `np_baseball`.
4. Cho biết chiều cao của vận động viên thứ 124.
5. Cho biết chiều cao trung bình, cân nặng trung bình của các cầu thủ
6. Bạn nhận xét gì về mối tương quan giữa chiều cao và cân nặng của các cầu thủ: có/ không có tương quan, tương quan thuận/ngịch

Chapter 3 - Exercise 5: Tính median của chiều cao (height) dựa vào vị trí (position)

Cho 2 tập tin `heights.txt` và `positions.txt` => chép dữ liệu từ 2 tập tin vào 2 list là `heights` và `positions`

'GK' (goalkeeper), 'M' (midfield), 'A' (attack) and 'D' (defense)

1. a) Tạo numpy array `np_positions` từ list `positions`. Xem kiểu dữ liệu của `np_positions`
b) Tạo numpy array `np_heights` từ list `heights`. Xem kiểu dữ liệu của `np_heights`
2. Tính chiều cao trung bình của các GK.
3. Tính chiều cao trung bình của những vị trí khác (Không phải là GK).
4. Tạo mảng dữ liệu có cấu trúc tự định nghĩa `players` gồm 'position' kiểu văn bản (U5) và 'height' kiểu 'float'
5. Sắp mảng `players` theo height, cho biết vị trí có chiều cao cao nhất và chiều cao thấp nhất