BÀI GIẢNG: THU THẬP VÀ PHÂN LOẠI DỮ LIỆU
CHUYÊN ĐỀ: DỮ LIỆU VÀ BIỂU ĐỒ
MÔN: TOÁN 8 (KẾT NỐI TRI THỨC)
GIÁO VIÊN: NGUYỄN THANH HẢI

MỤC TIÊU

- ✓ *Nắm được khái niệm về thu thập dữ liệu và cách thu thập dữ liệu.*
- ✓ *Nắm được khái niệm về phân loại dữ liệu và cách phân loại dữ liệu.*

I. THU THẬP DỮ LIỆU

1 Thu thập dữ liệu trực tiếp và gián tiếp

- Thu thập dữ liệu trực tiếp là việc thu thập thông qua quan sát, làm thí nghiệm, lập bảng hỏi, phỏng vấn.
- Thu thập dữ liệu gián tiếp là việc thu thập dữ liệu từ những nguồn có sẵn như sách báo, mạng internet.

2 Số liệu liên tục và số liệu rời rạc

Là việc thu thập dữ liệu từ những nguồn có sẵn như: sách, báo, mạng Internet,…

Để có thể đưa ra kết luận hợp lí, dữ liệu thu nhập được phải đảm bài tính đại diện cho toàn bộ đối tượng đang được quan tâm.

Bài 1. Để thu thập dữ liệu sau, ta nên làm thế nào? Đó là thu thập dữ liệu trực tiếp hay gián tieeos?

a) Dữ liệu về xếp hạng FIFA của bóng đá nam Việt Nam trong thời gian gần đây.

b) Số liệu về sự phát triển chiều cao của một giống cây mới theo thời gian.

Cách giải:

a) Cách tốt nhất ta vào website của Liên đoàn Bóng đá Thế giới (FIFA) tại địa chỉ fifa.com/fifa-world-ranking/vie để thu thập. Đây là phương pháp thu thập dữ liệu gián tiếp.

b) Để có số liệu về sự phát triển chiều cao của một giống cây mới theo thời gian, ta trồng cây và định kì đo chiều cao, ghi lại kết quả. Đây là phương pháp thu thập dữ liệu trực tiếp.

Bài 2. Hãy cho biết phương pháp thu thập dữ liệu trong mỗi trường hợp sau là trực tiếp hay gián tiếp.

a) Nam vào website của Cục Thống kê và ghi lại số xã/phường của các tỉnh/thành phố thuộc đồng bằng Bắc Bộ.

b) Thầy giáo dạy Giáo dục thể chất đã đo và ghi lại thời gian chạy cự li 1000 m của các bạn học sinh khối 8

Cách giải:

a) Đây là phương pháp thu thập dữ liệu gián tiếp

b) Đây là phương pháp thu thập dữ liệu trực tiếp

Bài 3. Nên sử dụng phương pháp thu thập dữ liệu nào để thu được mỗi dữ liệu sau:

a) Tên của 10 quốc gia có diện tích lớn nhất.

b) Ý kiến của các bạn về địa điểm đi tham quan trong tuần tới.

c) Chiều cao của các cây giống sau 6 tháng trồng.

Cách giải:

a) Phương pháp thu thập dữ liệu gián tiếp

b) Phương pháp thu thập dữ liệu trực tiếp

c) Phương pháp thu thập dữ liệu trực tiếp

II. PHÂN LOẠI DỮ LIỆU

1 Ví dụ

Cho hai dãy dữ liệu sau về 5 học sinh

(A) Chiều cao (đơn vị: cm): 128,1; 132,9; 125,7; 131,3; 133,6

(B) Số môn thể thao học sinh biết chơi: 2; 1; 5; 2; 3

a) Hai dãy dữ liệu này có phải là số liệu không?

b) Đo chiều cao (kí hiệu là h) một học sinh khác và hỏi về số môn thể thao (kí hiệu là n) mà em đó biết chơi.

+ h có thể nhận gía trị bất kì lớn hơn 120 cm và nhỏ hơn 150 cm được không?

+ n có thể nhận được giá trị lớn hơn 3 và nhỏ hơn 4 được không?

Cách giải:

a) Hai dãy dự liệu dạng số được gọi là số liệu,

b) h có thể nhận gía trị bất kì lớn hơn 120 cm và nhỏ hơn 150 cm

n không thể nhận được giá trị lớn hơn 3 và nhỏ hơn 4.

2 Số liệu liên tục và số liệu rời rạc

a. Số liệu liên tục

- Số liệu có thể nhận giá trị tùy ý trong một khoảng nào đó được gọi là số liệu liên tục.

b. Số liệu rời rạc

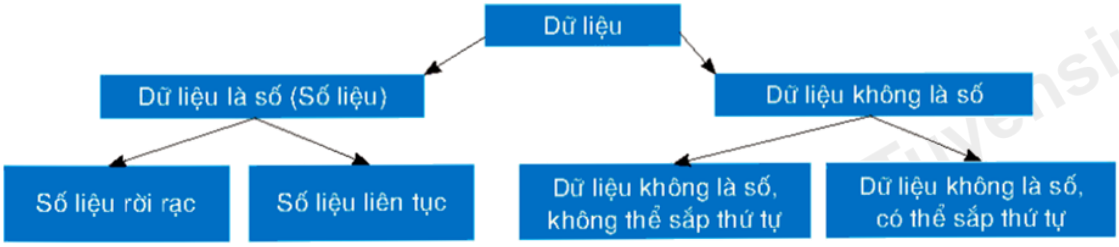
- Số liệu không phải là số liệu liên tục được gọi là số liệu rời rạc.

Chú ý:

- Dạng hay gặp của số liệu liên tục là số liệu thu được từ các phép đo như chiều cao, cân nặng, nhiệt độ,…

- Dạng hay gặp của số liệu rời rạc là số liệu đếm số phần tử của một tập hợp nào đó, chẳng hạn số học sinh trong lớp học, số sản phẩm một công nhân làm được trong ngày…

3 Số đồ phân loại dữ liệu



Bài 4. Với mỗi câu hỏi sau, Quỳnh đã hỏi ba bạn và ghi lại câu trả lời.

a) Nhà bạn có bao nhiêu chiếc ti vi? Kết quả: 5; 52; 2

b) Bạn mất bao nhiêu thời gian (đơn vị: giờ) để hoàn thành bài tập về nhà? Kết quả: 1,5; 2,3; 1,9

Mỗi dãy dữ liệu trên thuộc loại nào? Chỉ ra giá trị không hợp kí nếu có

Cách giải:

a) Dữ liệu thu được là số liệu rời rạc. Giá trị 52 không hợp lí

b) Dữ liệu thu được là số liệu liên tục.

Bài 5. Với mỗi câu hỏi sau, An đã hỏi 5 bạn và ghi lại câu trả lời.

a) Bạn nặng bao nhiêu kilogam? Kết quả: 48; 51; 46; 145; 48

b) Tên bạn có bao nhiêu chữ cái? Kết quả: 4; 5; 6; 3; 5.

Mỗi dãy dữ liệu trên thuộc loại nào? Chỉ ra giá trị không hợp lí nếu có.

Cách giải:

a) Dữ liệu thu được là số liệu liên tục. Giá trị 145 không hợp lí.

b) Dữ liệu thu được là số liệu rời rạc.

Bài 6. Dữ liệu thu được trong mỗi câu hỏi sau thuộc loại nào?

a) Bạn cao bao nhiêu?

b) Mạng điện thoại bạn đang dung là gì?

c) Gia đình bạn có bao nhiêu người dưới 18 tuổi?

Cách giải:

a) Dữ liệu thu được là số liệu liên tục.

b) Dữ liệu thu được không phải là số liệu

c) Dữ liệu thu được là số liệu rời rạc.

Bài 7. Ghép cặp cho phù hợp

a. Số liệu rời rạc.	A. Kết quả đánh giá của 5 bạn về để kiểm tra học kì I môn Toán: Khó, Rất khó, Trung bình, Dễ, Khó.
b. Số liệu liên tục.	B. Nhiệt độ (°C) tại Nha Trang trong 5 ngày đầu tháng 6 là: 23,2; 25,7; 31,4; 27,3; 28,6.
c. Dữ liệu không là số, không thể sắp thứ tự.	C. Số hoạt động hè mà các bạn trong tổ tham gia: 2; 1; 3; 0; 4.
d. Dữ liệu không là số, có thể sắp thứ tự.	D. Tên các môn thể thao mà các bạn yêu thích nhất: Bóng đá, Cầu lông, Cờ vua, Võ thuật, Bóng bàn.

Cách giải:

a – C; b – B; c – D; d – A