

# Phần MỘT SỐ YẾU TỐ THỐNG KÊ VÀ XÁC SUẤT

Chương 4.

## MỘT SỐ YẾU TỐ THỐNG KÊ

### Bài 1. THU THẬP VÀ PHÂN LOẠI DỮ LIỆU

#### A. KIẾN THỨC CẦN NHỚ

- Những thông tin thu thập được như: số, chữ, hình ảnh, ... được gọi là **dữ liệu**. Dữ liệu dưới dạng số được gọi là **số liệu**. Việc thu thập, phân loại, tổ chức và trình bày dữ liệu là những hoạt động **thống kê**.
- Thông tin rất đa dạng và phong phú. Việc sắp xếp thông tin theo những tiêu chí nhất định gọi là **phân loại dữ liệu**.
- Để đánh giá tính hợp lý của dữ liệu, ta cần đưa ra các **tiêu chí đánh giá**, chẳng hạn như dữ liệu phải:
  - Đúng định dạng.
  - Nằm trong phạm vi dự kiến.

#### B. BÀI TẬP MẪU

**Bài 1.** Bạn Cúc muốn điều tra về các môn thể thao được yêu thích của các bạn trong lớp. Em hãy giúp Cúc hoàn thành công việc.

Môn thể thao	Kiểm đếm	Số bạn chọn
Bóng đá		
Cầu lông		
Bóng chuyền		

Đá cầu		
Bóng bàn		
Bóng rổ		

*Giải*

Môn thể thao	Kiểm đếm	Số bạn chọn
Bóng đá		12
Cầu lông		8
Bóng chuyền		4
Đá cầu		7
Bóng bàn		4
Bóng rổ		5

**Bài 2.** Từ kết quả kiểm đếm của bạn Cúc ở bài 1, em hãy cho biết:

- a) Cúc đang thống kê về vấn đề gì?
- b) Bạn ấy thu thập được các loại dữ liệu gì?
- c) Môn thể thao nào được các bạn trong lớp của Cúc yêu thích nhất?

*Giải*

- a) Cúc đang thống kê về các môn thể thao được yêu thích của các bạn trong lớp.
- b) Các loại dữ liệu mà Cúc thu thập được là: Danh sách các môn thể thao được các bạn trong lớp yêu thích và số lượng các bạn yêu thích từng môn thể thao trong danh sách đó.
- c) Môn bóng đá được các bạn trong lớp của Cúc yêu thích nhất.

**Bài 3.** Tìm kiếm các thông tin không hợp lý của bảng dữ liệu sau đây:

**Số học sinh vắng trong ngày của các lớp khối 6 trường THCS Quang Trung**

6A1	6A2	6A3	6A4	6A5	6A6	6A7	6A8
1	2	0	K	0	-1	0,5	2

*Giải*

Các thông tin không hợp lý trong bảng dữ liệu trên là:

K ; -1; 0,5.

**Bài 4.** Các thông tin không hợp lý ở bài 3 vi phạm những tiêu chí nào? Hãy giải thích.

*Giải*

Thông tin không hợp lý	Tiêu chí vi phạm	Giải thích
K	Dữ liệu phải đúng định dạng.	Phải dùng định dạng số.
-1	Dữ liệu phải nằm trong phạm vi dự kiến.	Số học sinh vắng phải là số tự nhiên.
0,5	Dữ liệu phải nằm trong phạm vi dự kiến.	Số học sinh vắng phải là số tự nhiên.

### C. BÀI TẬP

1. Tìm các dữ liệu trong bảng thống kê sau:

Tốc độ chạy trung bình của một số động vật	
Con vật	Tốc độ (km/h)
Chó sói	69
Ngựa vằn	64
Sơn dương	98
Thỏ	56
Hươu cao cổ	51
Báo gấm	112

2. Em hãy thu thập và phân loại dữ liệu từ đoạn văn bản lịch sử (theo Viện Sử học) sau đây:

Nhà Ngô: 939 – 965;  
 Nhà Đinh: 968 – 980;  
 Nhà Tiền Lê: 980 – 1009;  
 Nhà Lý: 1009 – 1225;  
 Nhà Trần: 1226 – 1400;  
 Nhà Hồ: 1400 – 1407;  
 Nhà Hậu Lê: 1428 – 1788;  
 Nhà Tây Sơn: 1788 – 1802;  
 Nhà Nguyễn: 1802 – 1945.

Trình bày thông tin thu thập được theo mẫu sau:

Các triều đại phong kiến Việt Nam	
Triều đại	Thời gian tồn tại (năm)
Nhà Ngô	27
Nhà Đinh	...
...	...

3. Hãy làm thống kê trong lớp của em về một chủ đề mà lớp em quan tâm. (Gợi ý một số chủ đề: Các loại trái cây, thể loại nhạc, loài hoa, ... mà các bạn yêu thích).
4. Thân nhiệt (độ C) của bệnh nhân B trong 12 tiếng theo dõi được ghi lại trong bảng sau:

38	39	Nóng	40	41	45
38	37	36	Lạnh	37	37

- a) Em hãy tìm kiếm các thông tin không hợp lý của bảng dữ liệu trên.  
 b) Các thông tin không hợp lý đó đã vi phạm những tiêu chí nào? Hãy giải thích.