

CHUYÊN ĐỀ 11:

DỮ LIỆU, XÁC SUẤT THỰC NGHIỆM

ÔN HÈ MÔN: TOÁN - LỚP 6



BIÊN SOẠN: BAN CHUYÊN MÔN LOIGIAIHAY.COM

Dạng 1. Phân tích dữ liệu

A. Lý thuyết

I. Thu thập dữ liệu

- 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**.

- Số liệu là một loại dữ liệu nhưng dữ liệu chưa chắc là số liệu.

- Các cách thu thập dữ liệu: Quan sát, lập phiếu điều tra (phiếu hỏi), thu thập từ những nguồn có sẵn (sách, báo, trang web,...)

Có nhiều cách để thu thập dữ liệu như quan sát, làm thí nghiệm, lập phiếu hỏi,... hay thu thập từ những nguồn có sẵn như sách báo, trang web,...

Sau khi thu thập, tổ chức, phân loại, biểu diễn dữ liệu bằng bảng hoặc biểu đồ, ta cần phân tích và xử lý các dữ liệu đó để tìm ra các thông tin hữu ích và rút ra kết luận.

II. Phân loại dữ liệu

1. Khái niệm

- Phân loại dữ liệu là ta đi sắp xếp các thông tin theo những tiêu chí nhất định.

2. Ví dụ

Hình dưới đây là các bình ga của một cửa hàng đang bán.

a) Cửa hàng đang bán tất cả bao nhiêu bình ga?

b) Cửa hàng bán mấy loại bình ga?

Hãy cho biết số lượng bình ga mỗi loại.

Giải:

a) Cửa hàng bán tất cả 8 bình ga.

b) Quan sát hình ta thấy các bình ga có hai màu vàng và hồng.

Có hai kích thước: lớn và nhỏ

Nếu lấy tiêu chí là kích thước thì có hai loại ga là:

- Bình cỡ nhỏ: 6 bình
- Bình cỡ lớn: 2 bình

Nếu lấy tiêu chí là màu sắc để phân loại thì có 2 loại bình là:

- Màu hồng: 6 bình
- Màu vàng: 2 bình.

III. Tính hợp lí của dữ liệu

- Để đánh giá tính hợp lí của dữ liệu, ta cần đưa ra các tiêu chí đánh giá, chẳng hạn như dữ liệu phải:

- + Đúng định dạng: Họ và tên phải là chữ, số tuổi phải là số, email thì phải có kí hiệu @,...
- + Nằm trong phạm vi dự kiến: Số người thì phải là số tự nhiên, cân nặng của người Việt Nam thì phải dưới 200kg, số tuổi của người nguyên dương,...

Ví dụ:

Danh sách đội học sinh dự thi văn nghệ của lớp 6A được ghi lại trong bảng sau:

STT	Họ và tên
1	Hoàng Thu Trang
2	Đỗ Ngọc Hà
3	Phạm Văn Vũ
4	0384888586
5	Trần Nhật Minh
6	Nguyễn Minh Trí

Dữ liệu không hợp lí ở đây là 0384888586 trong cột Họ và tên vì đây không phải là tên người.

B. Bài tập

Bài 1:

Lan đun nước và đo nhiệt độ của nước tại một số thời điểm sau khi bắt đầu đun được kết quả như sau:

Số phút sau khi bắt đầu đun	5	6	7	8	9	10	11
Nhiệt độ ($^{\circ}\text{C}$)	41	76	85	94	97	100	104

- Lan đã thu được dữ liệu trên bằng cách nào: quan sát, làm thí nghiệm hay lập bảng hỏi?
- Tìm các giá trị không hợp lí (nếu có) trong dữ liệu về nhiệt độ của nước mà Lan đo được. Giải thích.

Bài 2:

Đoạn sau được trích từ Thời báo tài chính Việt Nam số ra ngày 26-5-2020:

"Căn cứ báo cáo nhanh của Văn phòng Bộ Công An và Cục Hàng hải Việt Nam, trong tháng 5-2020 (từ ngày đến 14-5-2020), toàn quốc xảy ra 998 vụ tai nạn giao thông, làm chết 529 người và làm bị thương

660 người. So với tháng cùng kì năm 2019 thì tháng 5–2020 đã giảm 328 vụ, giảm 29 người tử vong, giảm 415 người bị thương."

Hãy liệt kê số vụ tai nạn giao thông, số người tử vong, số người bị thương trong tháng 5–2019 trên toàn quốc.

Lời giải chi tiết

Bài 1:

Lan đun nước và đo nhiệt độ của nước tại một số thời điểm sau khi bắt đầu đun được kết quả như sau:

Số phút sau khi bắt đầu đun	5	6	7	8	9	10	11
Nhiệt độ ($^{\circ}\text{C}$)	41	76	85	94	97	100	104

- Lan đã thu được dữ liệu trên bằng cách nào: quan sát, làm thí nghiệm hay lập bảng hỏi?
- Tìm các giá trị không hợp lí (nếu có) trong dữ liệu về nhiệt độ của nước mà Lan đo được. Giải thích.

Phương pháp

- Xác định phương pháp để thu thập dữ liệu bài toán.
- Phân loại những khả năng nhiệt độ phù hợp của nước. Kiểm tra tính hợp lí về nhiệt độ của nước.

Lời giải

- Lan đã làm thí nghiệm để thu được dữ liệu.
- Giá trị 104 là giá trị không hợp lí vì ở điều kiện bình thường nước sôi ở 100 độ C và sẽ bay hơi.

Bài 2:

Đoạn sau được trích từ Thời báo tài chính Việt Nam số ra ngày 26–5–2020:

"Căn cứ báo cáo nhanh của Văn phòng Bộ Công An và Cục Hàng hải Việt Nam, trong tháng 5–2020 (từ ngày đến 14–5–2020), toàn quốc xảy ra 998 vụ tai nạn giao thông, làm chết 529 người và làm bị thương 660 người. So với tháng cùng kì năm 2019 thì tháng 5–2020 đã giảm 328 vụ, giảm 29 người tử vong, giảm 415 người bị thương."

Hãy liệt kê số vụ tai nạn giao thông, số người tử vong, số người bị thương trong tháng 5–2019 trên toàn quốc.

Phương pháp

- + Phân tích dữ liệu trong đoạn trích (Số vụ tai nạn trong tháng 5–2020)
- + Xác định mối quan hệ của từ dữ liệu trong tháng 5–2020 với tháng 5–2019.
- + Tính toán tìm số vụ tai nạn giao thông, số người tử vong, số người bị thương trong tháng 5–2019 trên toàn quốc.

Lời giải

Trong tháng 5–2019:

Số vụ tai nạn giao thông là: $998 + 328 = 1326$ (vụ)

Số người tử vong là: $529 + 29 = 558$ (người)

Số người bị thương là: $660 + 415 = 1075$ (người)