

# BÀI GIẢNG: DỮ LIỆU VÀ THU THẬP DỮ LIỆU

## CHUYÊN ĐỀ: DỮ LIỆU VÀ XÁC SUẤT THỰC NGHIỆM

### MÔN TOÁN LỚP 6

#### THẦY GIÁO: ĐỖ VĂN BẢO



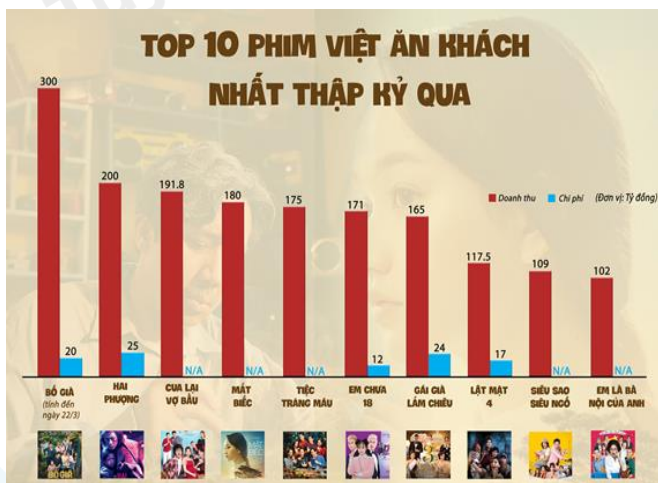
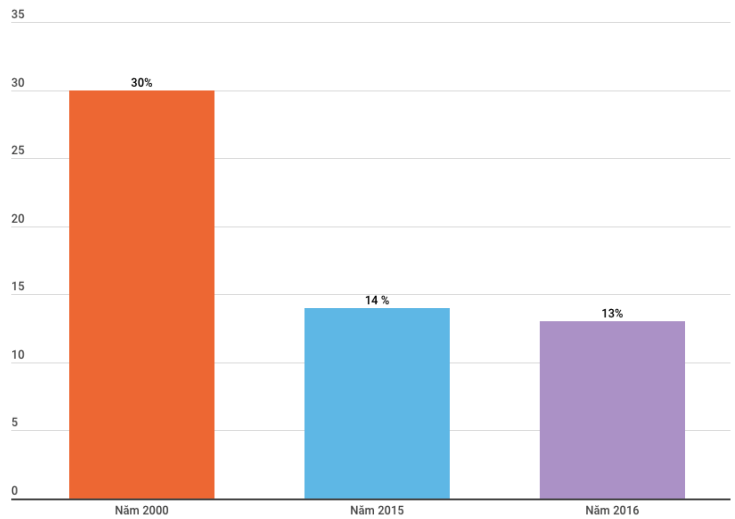
### Thông kê dữ liệu

#### Thành tích của Việt Nam qua 15 lần tham dự SEA Games

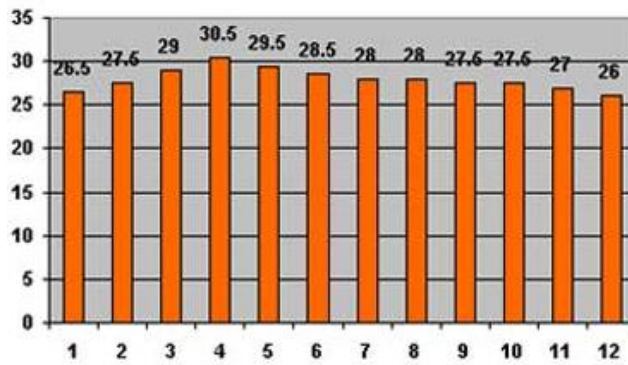
Các kỳ SEA Games VIỆT NAM THAM GIA	VĐV	VÀNG	BẠC	ĐỒNG	XẾP HẠNG
29 (2017) tại Malaysia	476	58	50	60	3/11
28 (2015) tại Singapore	392	73	53	60	3/11
27 (2013) tại Myanmar	519	73	86	86	3/11
26 (2011) tại Indonesia	608	96	92	100	3/11
25 (2009) tại Lào	434	83	75	57	2/11
24 (2007) tại Thái Lan	605	64	58	82	3/11
23 (2005) tại Philippines	528	71	68	89	3/11
22 (2003) tại Việt Nam	690	156	91	93	1/11
21 (2001) tại Malaysia	416	33	35	64	4/10
20 (1999) tại Brunei	170	17	20	27	6/10
19 (1997) tại Indonesia	340	35	48	50	5/10
18 (1995) tại Thái Lan	183	10	18	24	6/10
17 (1993) tại Singapore	139	9	6	19	6/9
16 (1991) tại Philippines	100	7	12	10	7/9
15 (1989) tại Malaysia	46	3	11	5	7/9

Nguồn: Viện Khoa học Thể dục Thể thao  
Infographics.vn

Tỷ lệ trẻ suy dinh dưỡng thể nhẹ cân giảm qua các năm  
(Nguồn: Viện Dinh dưỡng Quốc gia, Đơn vị: %)



Nhiệt độ không khí trung bình các tháng (°C)



### Xác suất thực nghiệm

- Xác suất thực nghiệm là khả năng xảy ra một sự kiện gì đó

- VD: Tung đồng xu, tung xúc xắc, đập bài, trò chơi chiếc nón kỳ diệu, trò bịt mắt đập niêu



### Mục tiêu bài học

- Nhận biết các loại dữ liệu
- Nhận biết tính hợp lý của dữ liệu
- Các phương pháp thu thập dữ liệu



### 1. Dữ liệu thống kê

Dự báo thời tiết 10 ngày tới										
Thời gian	Thứ Năm 11/04/2019	Thứ Sáu 12/04/2019	Thứ Bảy 13/04/2019	Chủ Nhật 14/04/2019	Thứ Hai 15/04/2019	Thứ Ba 16/04/2019	Thứ Tư 17/04/2019	Thứ Năm 18/04/2019	Thứ Sáu 19/04/2019	Thứ Bảy 20/04/2019
Thời tiết	Có mây, có mưa rào và dông	Nhiều mây, có mưa, mưa rào	Có mây, không mưa	Có mây, có mưa rào và dông	Nhiều mây, có mưa, dông	Nhiều mây, không mưa	Nhiều mây, có mưa, mưa rào	Có mây, không mưa	Có mây, không mưa	Có mây, không mưa
Nhiệt độ thấp nhất (°C)	26	24	23	24	23	22	22	23	23	23
Nhiệt độ cao nhất (°C)	34	29	28	31	26	28	26	29	31	28

- Các thông tin thu được ở trên như nhiệt độ thấp nhất, nhiệt độ cao nhất, ngày không mưa,... được gọi là **dữ liệu**. Trong các dữ liệu ấy, có dữ liệu là số (số liệu), có dữ liệu không phải là số.

DỰ BÁO THỜI TIẾT 10 NGÀY TỚI								
Thứ sáu 04/06/2021	Thứ bảy 05/06/2021	Chủ nhật 06/06/2021	Thứ hai 07/06/2021	Thứ ba 08/06/2021	Thứ tư 09/06/2021	Thứ năm 10/06/2021	Thứ sáu 11/06/2021	Thứ bảy 12/06/2021
32°C 27°C	33°C 26°C	32°C 25°C	32°C 24°C	32°C 23°C	34°C 24°C	33°C 23°C	33°C 23°C	33°C 24°C
Nhiều mây, có mưa, mưa rào	Có mây, có mưa rào	Nhiều mây, có mưa, mưa rào	Nhiều mây, có mưa, mưa rào	Nhiều mây, có mưa, mưa rào	Nhiều mây, có mưa, mưa rào	Nhiều mây, có mưa, mưa rào	Có mây, có mưa rào	Có mây, có mưa rào

- Các thông tin thu được từ các bảng biểu, bài báo, câu chuyện, đoạn văn,... được gọi là các dữ liệu  
- Trong các dữ liệu, có dữ liệu là số (được gọi là số liệu), có dữ liệu không phải là số

### Ví dụ 1

An cùng bạn liệt kê tên một số loài thực vật để làm bài tập môn Khoa học tự nhiên, được đây dữ liệu như sau:

Dương xỉ, thông, dừa, rêu, đậu, bưởi, vi khuẩn

a) Dữ liệu trên có phải là số liệu không?

b) Em hãy chỉ ra dữ liệu không hợp lí (nếu có) trong dãy dữ liệu trên

### Lời giải:

a) Dữ liệu trên không phải số liệu

b) Dữ liệu không hợp lí là: vi khuẩn

### Luyện tập 1

Cho hai dãy dữ liệu như sau:

(1) Số học sinh các lớp 6 trong phòng:

35 42 87 38 40 41 38

(2) Tên món ăn yêu thích của các thành viên trong gia đình:

Bánh chưng, pizza, canh cua, gà rán, rau muống luộc, cá kho, rượu vang

a) Trong các dãy dữ liệu trên, dãy nào là dãy số liệu?

b) Hãy tìm dữ liệu không hợp lí (nếu có) trong mỗi dãy dữ liệu trên

**Lời giải:**

a) Dãy (1) là dãy số liệu

b) Dãy (1): dữ liệu không hợp lí là: 87

Dãy (2): dữ liệu không hợp lí là: rượu vang

- Dữ liệu mà có thể thể hiện bằng số lượng hay cân, đong, đo, đếm, tính toán được thì thường là **số liệu**

Hoa điểm tốt Tháng Ba			
		Tổ Một	
Tên	Điểm 8	Điểm 9	Điểm 10
Nam Anh	✓✓	✓	✓
Thanh Hằng	✓✓	✓	✓
Đức Minh	✓	✓✓	



## 2. Thu thập dữ liệu thống kê

- Quân muốn thu thập số liệu về số lượng học sinh đeo kính trong một số lớp học để làm một dự án học tập. Em hãy giúp Quân bằng cách quan sát và đếm xem lớp em có bao nhiêu bạn đeo kính

- Vuông gieo một đồng xu 5 lần và ghi lại kết quả như bảng bên (mỗi gạch ứng với một lần xuất hiện)

Em hãy gieo một đồng xu 10 lần và ghi lại kết quả theo cách tương tự

Mặt sấp	///
Mặt ngửa	//

**VD1:** Lớp em chuẩn bị tổ chức một buổi liên hoan cuối năm. Lớp trưởng muốn thống kê các món ăn mà các thành viên trong lớp thích nhất để chuẩn bị tiệc. Hãy đưa ra một cách thu thập ý kiến của tất cả các thành viên trong lớp

⇒ Lập phiếu hỏi

- Có nhiều cách để thu thập dữ liệu như quan sát, làm thí nghiệm, lập phiếu hỏi,... hay thu thập từ nguồn có sẵn như sách báo, internet, truyền hình,...

- Mỗi cách thu thập có những ưu nhược điểm khác nhau và được ứng dụng tùy theo hoàn cảnh cụ thể

- Người ta thường lập các bảng biểu để thu thập, ghi chép dữ liệu

**VD2:** Thống kê số người tham gia giao thông vượt đèn đỏ

Nên thu thập dữ liệu bằng cách hỏi những người tham gia giao thông xem họ có vượt đèn đỏ không.



Nên đặt máy quay ở ngã tư để thu thập dữ liệu.



### 3. Áp dụng

**9.1:** Trong các dữ liệu sau, dữ liệu nào là số liệu, dữ liệu nào không phải là số liệu?

1. Cân nặng của trẻ sơ sinh (đơn vị tính là gam)
2. Quốc tịch của các học sinh trong một trường quốc tế
3. Chiều cao trung bình của một số loại cây thân gỗ (đơn vị tính là mét)

**9.3:** Bảng sau cho biết số anh chị em ruột trong gia đình của 35 học sinh lớp 6A

Số anh chị em ruột	0	1	2	3
Số học sinh	18	12	15	1

Hãy tìm điểm không hợp lí trong bảng thống kê trên?

**9.4:** Hãy tìm dữ liệu không hợp lí (nếu có) trong dãy dữ liệu sau:

Thủ đô của một số quốc gia châu Á:

Hà Nội      Bắc Kinh      Paris

Tokyo      Đà Nẵng

**9.5:** Để hoàn thành bảng sau, em sẽ sử dụng phương pháp thu thập dữ liệu nào?

Cây	Môi trường sống	Dạng thân	Kiểu lá
Đậu			
Bèo tây			

**9.6:** Hãy lập phiếu hỏi để thu thập dữ liệu về phương tiện đến trường của các thầy, cô giáo trong trường em

**Bài 1:** Trong các dữ liệu sau, dữ liệu nào là số liệu, dữ liệu nào không phải là số liệu?

- Số cuốn sách mỗi bạn học sinh lớp 6A đọc được trong 1 tháng
- Tên sách mỗi bạn học sinh lớp 6A đã đọc được trong 1 tháng
- Chiều cao của mỗi bạn học sinh lớp 6A
- Tên của các bạn học sinh trong lớp 6A
- Số học sinh nghỉ học tuần này của lớp 6A
- Tên học sinh nghỉ học tuần này của lớp 6A
- Lý do học sinh nghỉ học tuần này của lớp 6A

**Bài 2:** COVID – 19 có ảnh hưởng mạnh mẽ đến hệ hô hấp, bao gồm đường hô hấp trên và đường hô hấp dưới.

Các bộ phận trên cơ thể liên quan đến hệ hô hấp là: phế quản, phổi, mũi, họng, chân,...

Hãy tìm dữ liệu không hợp lí nếu có trong thống kê trên

**Bài 3:** Các bộ phận trên cơ thể người: hệ tuần hoàn; hệ hô hấp; hệ thần kinh; hệ vận động; hệ tiêu hóa; hệ nội tiết; hệ bài tiết; hệ sinh sản; hệ quần áo

Hãy tìm dữ liệu không hợp lí nếu có trong thống kê trên

**Bài 4:** Lớp 6A5 có 40 học sinh. Kết thúc học kì 1, lớp trưởng Diệu Ly thu thập dữ liệu về tình hình đánh giá học sinh ở mức Tốt, Khá, Đạt, Chưa Đạt bằng cách hỏi từng bạn và lập thành một bảng như sau:

Đánh giá	Tốt	Khá	Đạt	Chưa Đạt	Tổng số
Số lượng	25	10	6	0	

Hãy tìm điểm không hợp lí (nếu có) trong bảng thống kê trên và giải thích

**Bài 5:** Lớp phó sự kiện Khánh Vy muốn tổ chức sinh nhật chung cho các bạn cùng tháng. Khánh Vy cần phải liệt kê từng tháng có bao nhiêu bạn sinh nhật để lớp chuẩn bị quà thì nên sử dụng phương pháp nào để thu thập dữ liệu trong các phương pháp: quan sát, làm thí nghiệm, lập phiếu hỏi? Tại sao?

**Bài 6:** Để thu thập dữ liệu về các quốc gia đông dân nhất trên thế giới hiện nay, em nên dùng phương pháp thu thập dữ liệu nào?

**Bài 7:** Để ước tính số người tham gia giao thông vào giờ cao điểm tại cổng trường, bạn Hân nên dùng phương pháp thu thập dữ liệu nào trong các phương pháp: làm thí nghiệm, lập phiếu hỏi, quan sát, tìm kiếm trên internet? Tại sao?



#### 4. Hướng dẫn lời giải

**9.1:** Trong các dữ liệu sau, dữ liệu nào là số liệu, dữ liệu nào không phải là số liệu?

- Cân nặng của trẻ sơ sinh (đơn vị tính là gam)
- Quốc tịch của các học sinh trong một trường quốc tế
- Chiều cao trung bình của một số loại cây thân gỗ (đơn vị tính là mét)

**Lời giải:**

Dữ liệu 1, 3 là số liệu

Dữ liệu 2 không phải là số liệu

**9.3:** Bảng sau cho biết số anh chị em ruột trong gia đình của 35 học sinh lớp 6A

Số anh chị em ruột	0	1	2	3
Số học sinh	18	12	15	1

Hãy tìm điểm không hợp lí trong bảng thống kê trên?

**Lời giải:**

Tổng số học sinh:  $18 + 12 + 5 + 1 = 36 \neq 35$

**9.4:** Hãy tìm dữ liệu không hợp lí (nếu có) trong dãy dữ liệu sau:

Thủ đô của một số quốc gia châu Á:

Hà Nội      Bắc Kinh      Paris

Tokyo      Đà Nẵng

**Lời giải:**

Dữ liệu không hợp lí: Đà Nẵng

**9.5:** Để hoàn thành bảng sau, em sẽ sử dụng phương pháp thu thập dữ liệu nào?

Cây	Môi trường sống	Dạng thân	Kiểu lá
Đậu			
Bèo tây			

**Lời giải:**

Có thể áp dụng phương pháp: quan sát, tra cứu dữ liệu qua internet

**9.6:** Hãy lập phiếu hỏi để thu thập dữ liệu về phương tiện đến trường của các thầy, cô giáo trong trường em

**Bài 1:** Trong các dữ liệu sau, dữ liệu nào là số liệu, dữ liệu nào không phải là số liệu?

- Số cuốn sách mỗi bạn học sinh lớp 6A đọc được trong 1 tháng
- Tên sách mỗi bạn học sinh lớp 6A đã đọc được trong 1 tháng
- Chiều cao của mỗi bạn học sinh lớp 6A
- Tên của các bạn học sinh trong lớp 6A
- Số học sinh nghỉ học tuần này của lớp 6A
- Tên học sinh nghỉ học tuần này của lớp 6A
- Lý do học sinh nghỉ học tuần này của lớp 6A

**Lời giải:**

- Số cuốn sách mỗi bạn học sinh lớp 6A đọc được trong 1 tháng: Số liệu
- Tên sách mỗi bạn học sinh lớp 6A đã đọc được trong 1 tháng: Không phải số liệu
- Chiều cao của mỗi bạn học sinh lớp 6A: Số liệu
- Tên của các bạn học sinh trong lớp 6A: Không phải số liệu
- Số học sinh nghỉ học tuần này của lớp 6A: Số liệu
- Tên học sinh nghỉ học tuần này của lớp 6A: Không phải số liệu
- Lý do học sinh nghỉ học tuần này của lớp 6A: Không phải số liệu

**Bài 2:** COVID – 19 có ảnh hưởng mạnh mẽ đến hệ hô hấp, bao gồm đường hô hấp trên và đường hô hấp dưới.

Các bộ phận trên cơ thể liên quan đến hệ hô hấp là: phế quản, phổi, mũi, họng, chân,...

Hãy tìm dữ liệu không hợp lí nếu có trong thống kê trên

**Lời giải:**

Dữ liệu không hợp lí trong thống kê trên là: chân

**Bài 3:** Các bộ phận trên cơ thể người: hệ tuần hoàn; hệ hô hấp; hệ thần kinh; hệ vận động; hệ tiêu hóa; hệ nội tiết; hệ bài tiết; hệ sinh sản; hệ quần áo

Hãy tìm dữ liệu không hợp lí nếu có trong thống kê trên

**Lời giải:**

Dữ liệu không hợp lí trong thống kê trên là: hệ quần áo

**Bài 4:** Lớp 6A5 có 40 học sinh. Kết thúc học kì 1, lớp trưởng Diệu Ly thu thập dữ liệu về tình hình đánh giá học sinh ở mức Tốt, Khá, Đạt, Chưa Đạt bằng cách hỏi từng bạn và lập thành một bảng như sau:

Đánh giá	Tốt	Khá	Đạt	Chưa Đạt	Tổng số
Số lượng	25	10	6	0	

Hãy tìm điểm không hợp lí (nếu có) trong bảng thống kê trên và giải thích

**Lời giải:**

Tổng số học sinh thống kê trong bảng:  $25 + 10 + 6 = 41 > 40$

**Bài 5:** Lớp phó sự kiện Khánh Vy muốn tổ chức sinh nhật chung cho các bạn cùng tháng. Khánh Vy cần phải liệt kê từng tháng có bao nhiêu bạn sinh nhật để lớp chuẩn bị quà thì nên sử dụng phương pháp nào để thu thập dữ liệu trong các phương pháp: quan sát, làm thí nghiệm, lập phiếu hỏi? Tại sao?

**Lời giải:**

Bạn Khánh Vy có thể sử dụng phương pháp lập phiếu hỏi hoặc hỏi trực tiếp

**Bài 6:** Để thu thập dữ liệu về các quốc gia đông dân nhất trên thế giới hiện nay, em nên dùng phương pháp thu thập dữ liệu nào?

**Lời giải:**

Phương pháp thu thập dữ liệu nên dùng: tra cứu dữ liệu qua internet

**Bài 7:** Để ước tính số người tham gia giao thông vào giờ cao điểm tại cổng trường, bạn Hân nên dùng phương pháp thu thập dữ liệu nào trong các phương pháp: làm thí nghiệm, lập phiếu hỏi, quan sát, tìm kiếm trên internet? Tại sao?

**Lời giải:**

Phương pháp thu thập dữ liệu nên dùng: quan sát.