

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU

2.4.1. Phương pháp thu thập dữ liệu từ thực nghiệm;

2.4.2. Phương pháp thu thập dữ liệu phi thực nghiệm.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

- Trong phương pháp này, số liệu được thực hiện bằng cách **quan sát, theo dõi, đo đạc** qua các thí nghiệm.
- Để thu thập số liệu, nhà NC cần **đặt ra các biến** để quan sát và đo đạc (thu thập số liệu). Các TN thường được **lặp lại nhiều lần** để làm giảm sai số trong thu thập số liệu.
- **Ví dụ**: Người NC muốn xem xét những mức độ phân bón nào đó để làm tăng năng suất, trong cách bố trí TN thì mỗi mức độ phân bón thường được lặp lại nhiều lần. Kết quả TN là các số liệu được đo từ các chỉ tiêu sinh trưởng và năng suất ở những mức độ phân bón khác nhau.
- PPKH trong thực nghiệm gồm các bước: **lập giả thuyết, xác định biến** (độc lập hay phụ thuộc), **bố trí TN, thu thập số liệu**, kiểm chứng giả thuyết.

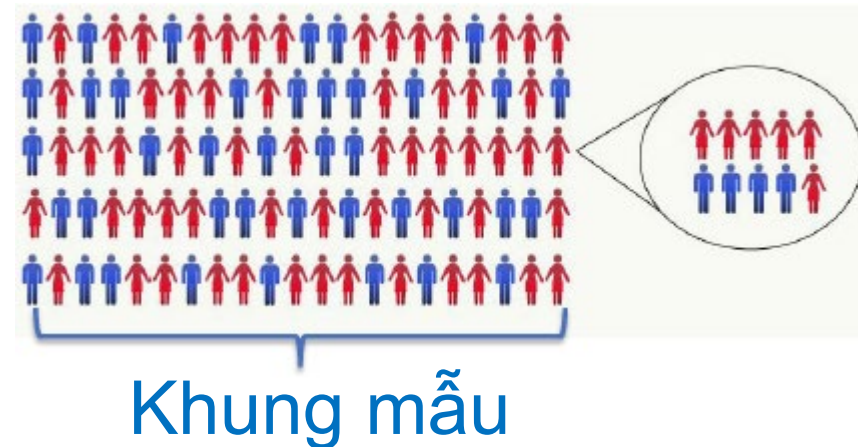
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## a. Khung mẫu (sample frame)

Là danh sách liệt kê dữ liệu cần thiết của toàn bộ khách thể NC (là vật mang đối tượng NC)

### Ví dụ:

Nếu chúng ta muốn khảo sát sự hài lòng của SV quốc tế về điều kiện học tập khi học tại Đại học X thì cần phải xác định **danh sách của tất cả các SV** quốc tế tại Đại học X.



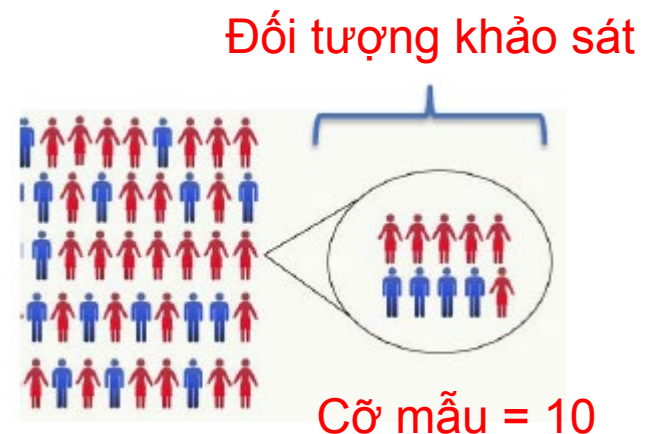
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## b. Đối tượng khảo sát

Là một bộ phận đủ đại diện của khách thể NC (là vật mang đối tượng NC) trong khung mẫu được người NC lựa chọn để xem xét. Hiếm khi người NC có đủ quỹ thời gian và kinh phí để khảo sát trên toàn bộ khách thể. Đối tượng khảo sát thường được chia làm hai nhóm:

a) Nhóm khảo sát: đối tượng được đặt ra trong giả thuyết.

b) Nhóm đối chứng: so sánh với nhóm khảo sát.



# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## c. Xác định cỡ mẫu:

Cỡ mẫu là **số lượng** khách thể NC cần thiết trong **nhóm đối tượng khảo sát** được người NC lựa chọn để xem xét.

Mục đích của việc xác định cỡ mẫu là để giảm đi công lao động và chi phí làm thí nghiệm và điều quan trọng là chọn cỡ mẫu sao cho không làm mất đi các **đặc tính** của mẫu và **độ tin cậy** của số liệu đại diện cho quần thể.

Người NC cần xác định một kích cỡ mẫu **tối thiểu** mà vẫn đánh giá được tương đối chính xác quần thể. Chọn cỡ mẫu quá lớn hoặc lớn hơn mức tối thiểu thì tốn kém còn chọn cỡ mẫu dưới mức tối thiểu lại ít chính xác.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## c. Xác định cỡ mẫu:

Trước khi xác định cỡ mẫu, ta cần xác định dạng phân phối của quần thể. Để đơn giản, ở đây ta thừa nhận quần thể có phân phối **normal distribution**.

Để xác định cỡ mẫu tối thiểu, ta cần phải đánh giá **trung bình mẫu** và **trung bình quần thể  $\mu$** . Thực tế, hai giá trị trung bình này thường khác nhau.

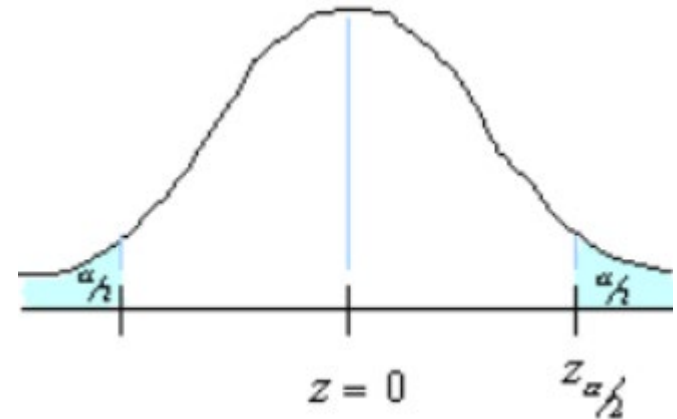
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## c. Xác định cỡ mẫu:

**Sai số biên** (the margin of error)  $d$  thể hiện sự khác nhau giữa trung bình mẫu quan sát và giá trị trung bình của quần thể  $\mu$  được tính như sau:

$$d = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

trong đó  $d$  là sai số biên mong muốn;  $Z_{\alpha/2}$  là giá trị ngưỡng (ứng với khoảng tin cậy  $1 - \alpha$ ) của phân bố chuẩn như chỉ trong Hình 1;  $n$  là cỡ mẫu; và  $\sigma$  là độ lệch chuẩn quần thể



Hình 1: Phân bố  
thống kê của mẫu

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## c. Xác định cỡ mẫu:

Sau đó chúng ta có thể tính cỡ mẫu cần thiết dựa trên khoảng tin cậy  $(1 - \alpha)$  và sai số biên. Cỡ mẫu được tính qua chuyển đổi công thức trên là

$$n = \left( Z_{\alpha/2} \frac{\sigma}{d} \right)^2$$

Để tính được  $n$  thì phải biết  $\sigma$ , sai số biên mong muốn, xác định khoảng tin cậy  $(1 - \alpha)$  để tra giá trị  $Z_{\alpha/2}$  như chỉ trong Bảng

$1 - \alpha$	0.8	0.85	0.9	0.95	0.99
$Z_{\alpha/2}$	1.28	1.44	1.645	1.96	2.85



# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## c. Xác định cỡ mẫu:

**Ví dụ:** Một người NC muốn đánh giá hàm lượng trung bình của phosphorus trong một ao hồ. Một NC trong nhiều năm trước đây có một độ lệch chuẩn quần thể  $\sigma$  có giá trị là 1,5 gram/lít. Bao nhiêu mẫu nước cần thiết nên được lấy để đo hàm lượng phosphorus chính xác mà 95% mẫu (khoảng tin cậy  $(1 - \alpha)$ ) có sai số không vượt quá 0,1 gram?

Áp dụng công thức tính cỡ mẫu: Sai số biên  $d = 0.1$

$$n = \left( Z_{\alpha/2} \frac{\sigma}{d} \right)^2 \Rightarrow n = \left( 1.96 \frac{1.5}{0.1} \right)^2 = 864.36$$

Như vậy, người NC cần lấy **865** mẫu nước để phân tích hàm lượng trung bình của phosphorus trong ao hồ.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## c. Xác định cỡ mẫu:

**Ví dụ:** Một người NC muốn đánh giá sản phẩm hư hỏng trung bình trong một nhà máy sản xuất bánh kẹo. Một NC trong nhiều năm trước đây có một độ lệch chuẩn trong nhà máy  $\sigma = 2 \text{ kg/tấn}$ . Bao nhiêu mẫu bánh kẹo cần thiết nên được lấy để đo mức độ sản phẩm hư hỏng chính xác mà 95% mẫu có sai số  $d$  không vượt quá 0,2 kg ?

Áp dụng công thức tính cỡ mẫu:

$$n = \left( Z_{\alpha/2} \frac{\sigma}{d} \right)^2 \Rightarrow n = \left( 1.96 \frac{2.0}{0.2} \right)^2 = 384.16$$

Như vậy, người NC cần lấy **385** mẫu bánh kẹo.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## d. Phương pháp lấy mẫu

Mục đích của các PP lấy mẫu là đạt được **mẫu đại diện** cho cả quần thể NC.

Khi chọn PP lấy mẫu, cần hiểu rõ các **đặc tính** của quần thể NC để xác định cỡ mẫu quan sát đại diện và để đánh giá tương đối chính xác quần thể.

Trong NC, khi không thể quan sát hết toàn bộ các cá thể trong quần thể, ta sẽ chọn một số lượng đủ các cá thể đại diện hay còn gọi là mẫu TN.

PP chọn mẫu thí nghiệm rất quan trọng, bởi vì có liên quan tới sự biến động hay độ đồng đều của mẫu. Có **2** PP chọn mẫu:

- i. Chọn mẫu **không xác suất** (không chú ý tới độ đồng đều)
- ii. Chọn mẫu **có xác suất** (đề cập tới độ đồng đều)

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## i. Chọn mẫu không có xác suất:

Là cách lấy mẫu trong đó các cá thể của mẫu được chọn **không ngẫu nhiên** hay **không có xác suất lựa chọn giống nhau**. Ví dụ

- Các đơn vị mẫu được tự lựa chọn mà không có PP.
- Các đơn vị mẫu rất dễ dàng đạt được hoặc dễ dàng tiếp cận. VD chọn những hộ trên những con đường dễ đi.
- Các đơn vị mẫu được chọn theo lý do K/Tế, VD trả tiền cho sự tham dự.
- Các đơn vị mẫu được quan tâm bởi người NC trong cách “điển hình” của quần thể mục tiêu. VD: mẫu được chọn nằm trong d/sách “điển hình”
- Các đơn vị mẫu được chọn mà không có sự thiết kế rõ ràng (VD: chọn 50 người đầu tiên đến buổi sáng).

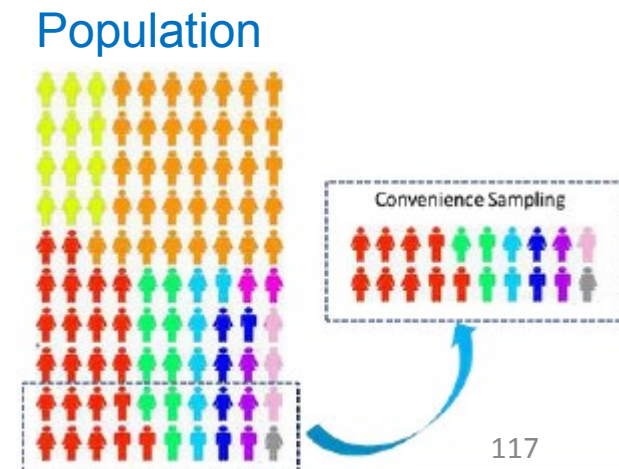
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## i. Chọn mẫu không có xác suất:

PP chọn mẫu không có xác suất thường có **độ tin cậy thấp**.

Mức độ chính xác của cách chọn mẫu không xác suất tùy thuộc vào:

- Sự phán đoán, cách nhìn, kinh nghiệm của người nghiên cứu;
- Sự may mắn hoặc dễ dàng;
- **Không có cơ sở thống kê** trong việc chọn mẫu.



# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

Cơ bản của việc chọn mẫu có xác suất là cách lấy mẫu sao cho việc chọn các cá thể của mẫu sao cho mỗi cá thể có **cơ hội lựa chọn như nhau**.

Nếu như có một số cá thể có cơ hội xuất hiện nhiều hơn thì sự lựa chọn không phải là ngẫu nhiên. Để tối ưu hóa mức độ chính xác trong NC, người nghiên cứu thường sử dụng PP lấy mẫu ngẫu nhiên.

Có **3** dạng chọn mẫu xác suất chính:

- **Chọn mẫu ngẫu nhiên đơn giản** (simple random);
- **Chọn mẫu ngẫu nhiên phân lớp** (stratified random samples);
- **Chọn mẫu ngẫu nhiên hệ thống** (systematic random samples).

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu ngẫu nhiên đơn giản (simple random)

Trong PP này, việc lựa chọn  $n$  các cá thể từ một quần thể dựa trên nguyên tắc **xác suất bằng nhau** nhằm đảm bảo các cá thể có cơ hội được lựa chọn bằng nhau.

**Ví dụ:** Một trường học có 1.000 SV, người NC muốn chọn ra 100 SV để NC về tình trạng sức khỏe trong số 1.000 SV. Theo cách chọn mẫu đơn giản => viết tên 1.000 SV vào mẫu giấy nhỏ => bỏ vào một cái thùng => rút ngẫu nhiên ra 100 mẫu giấy. Như vậy, mỗi SV có một cơ hội lựa chọn như nhau và XS chọn ngẫu nhiên một SV trên dễ dàng được tính.

VD trên ta có quần thể  $N = 1.000$  SV và cỡ mẫu  $n = 100$  SV. Như vậy, SV của trường được chọn trong cách lấy mẫu ngẫu nhiên sẽ có xác suất là  $n/(N \times 100)$  hay  $100/(1000 \times 100) = 10\%$ .

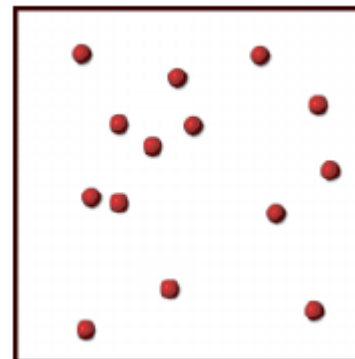
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu ngẫu nhiên đơn giản (simple random)

Một cách chọn mẫu ngẫu nhiên khác là sử dụng bảng số ngẫu nhiên trong sách thống kê phép thí nghiệm hoặc cách chọn số ngẫu nhiên bằng chương trình thống kê trên máy tính.

PP chọn mẫu ngẫu nhiên trong các TN lấy mẫu trong thực tế được thể hiện trong Hình.



Hình: Phương pháp chọn mẫu ngẫu nhiên trong thực tế



# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

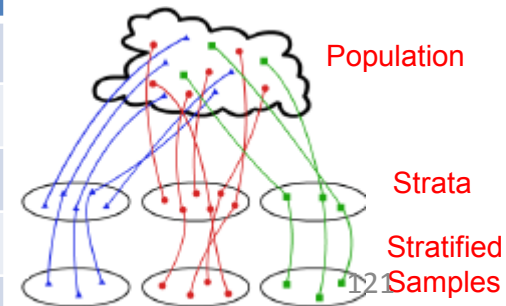
## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu ngẫu nhiên phân lớp (stratified random samples)

Ta chọn mẫu phân lớp khi quần thể mục tiêu chia thành các nhóm hay phân lớp. Trong pp lấy mẫu phân lớp, tổng quần thể ( $N$ ) được chia thành  $L$  lớp các quần thể phụ  $N_1, N_2 \dots N_L$ . Sau đó, người NC sẽ **xác định cỡ mẫu** và **chọn ngẫu nhiên** các cá thể **trong mỗi lớp**.

**VD:** Khi NC về mức độ giàu nghèo của một vùng NC có 4 huyện (4 phân lớp), mỗi huyện có số hộ khác nhau. Người NC muốn thực hiện 200 cuộc PV hộ gia đình trong vùng NC, như vậy cỡ mẫu của mỗi huyện sẽ được tính theo tỷ lệ % như:

Huyện	Số hộ trong mỗi lớp	Tỷ lệ hộ mỗi lớp (%)	Cỡ mẫu PV mỗi lớp
A	250	25	50
B	150	15	30
C	400	40	80
D	200	20	40
	<b>1000</b>	<b>100</b>	<b>200</b>



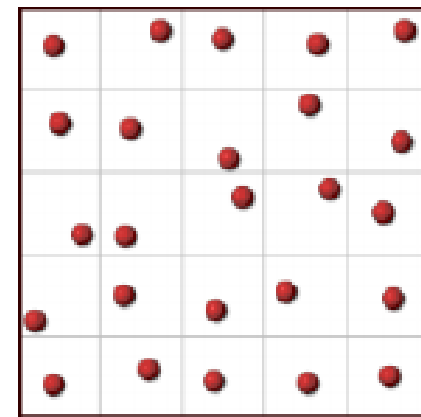
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu ngẫu nhiên phân lớp (stratified random samples)

Trong NC trên, nếu người NC không phân chia các huyện ra thành các lớp, thì pp lấy mẫu ngẫu nhiên PV hộ gia đình trong vùng NC có thể sẽ sai và số liệu thu thập sẽ không đại diện cho vùng NC, do mẫu có thể tập trung ở một huyện nào đó. Một số NC thường được chia lớp trong quần thể mục tiêu gồm:

Phân lớp quần thể mục tiêu là các TP, tỉnh, huyện; vùng sinh thái khác nhau; các hộ gia đình theo mức độ giàu nghèo, trình độ học vấn, ...;



Hình: Phương pháp chọn mẫu phân lớp

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu hệ thống (systematic samples)

Trong chọn mẫu hệ thống, cỡ mẫu  $n$  được chọn từ một quần thể  $N$ , và mẫu được chọn từ các đơn vị mẫu như là **một chuỗi liên tiếp** của các điểm có **khoảng cách bằng nhau**.

Ví dụ: chọn mẫu hệ thống như sau: muốn NC 1 thành viên trong mỗi nhóm có 10 cá thể, quần thể có 10 nhóm (tổng cá thể của quần thể là 100), ta đánh số cá thể từ 1-100. Lúc này nhóm 1 được đánh số từ 1-10; nhóm 2 từ 11-20; nhóm 3 từ 21-30; ...nhóm 10 từ 91-100.

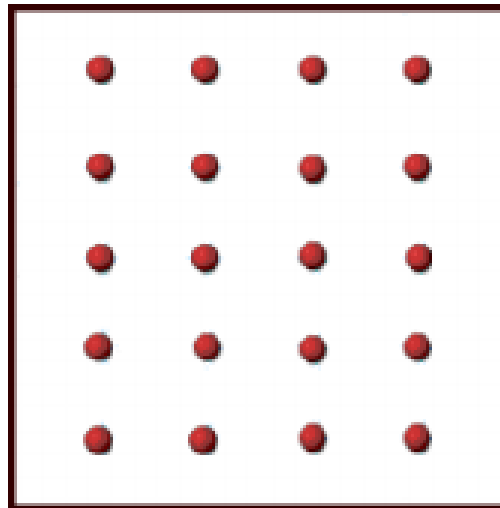
Trước tiên cần sắp xếp thứ tự các đơn vị mẫu. Sau đó chọn điểm đầu tiên bất kỳ có giá trị  $< 10$  (ví dụ 7). Số cá thể tiếp theo sẽ cộng thêm là 10. Như vậy các thành viên được chọn sẽ có số thứ tự là 7, 17, 27, 37, 47,... 97.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu hệ thống (systematic samples)

Phương pháp chọn mẫu hệ thống tạo ra các ô có các điểm có **khoảng cách đều nhau** với các ô có cấu trúc khác nhau như Hình



Hình : Phương pháp  
chọn mẫu hệ thống

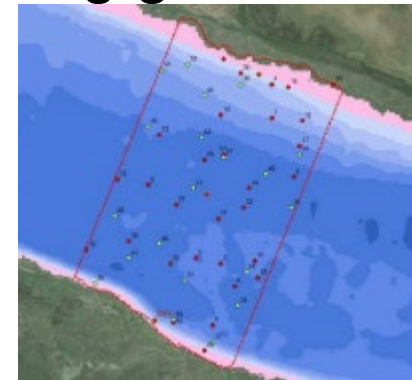
# PHƯƠNG PHÁP THU THẬP DỮ LIỆU TỪ THỰC NGHIỆM

## ii. Chọn mẫu có xác suất:

### ✓ Chọn mẫu không gian (spatial sampling)

Ta có thể sử dụng cách lấy mẫu này khi hiện tượng, sự vật được quan sát có **sự phân bố mẫu** theo không gian.

**Ví dụ:** lấy mẫu nước ở sông, đất ở sườn đồi, hoặc không khí trong phòng. Cách chọn mẫu như vậy thường gặp trong các nghiên cứu sinh học, địa chất, địa lý.



Lấy mẫu theo sự phân bố này yêu cầu có sự giống nhau về không gian qua các PP ngẫu nhiên, hệ thống và phân lớp. Kết quả của một mẫu chọn có thể được biểu diễn như một loạt các điểm trong không gian hai chiều, giống như là bản đồ.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU PHI THỰC NGHIỆM

PP phi thực nghiệm là PP thu thập số liệu dựa trên **sự quan sát** các sự kiện, sự vật đã hay đang tồn tại, từ đó **tìm ra quy luật** của chúng. PP này gồm các loại NC kinh tế và XH, NC nhân chủng học, ...

Loại số liệu thu thập trong PP phi thực nghiệm gồm số liệu được thu thập từ **Bảng câu hỏi** thông qua các **câu hỏi có cấu trúc kín** hoặc các **câu hỏi mở** theo các PP thu thập số liệu.

Cách xác định **Khung mẫu, đối tượng khảo sát, cỡ mẫu** và **cách lấy mẫu** để thu thập số liệu NC thông qua các Bảng hỏi trong PP phi thực nghiệm cũng tương tự như cách bố trí và thu thập số liệu thí nghiệm NC trong PP thực nghiệm.<sup>126</sup>

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU PHI THỰC NGHIỆM

**Một bảng câu hỏi phải thỏa mãn 2 yêu cầu cơ bản sau:**

1. Phải có **đầy đủ câu hỏi** mà nhà NC muốn thu thập dữ liệu
2. Phải tạo điều kiện và kích thích **sự hợp tác** của người trả lời

## **Quy trình thiết kế bảng câu hỏi**

1. Xác định cụ thể dữ liệu cần thu thập
2. Xác định **dạng phỏng vấn** (trực tiếp, điện thoại, gửi thư, internet, ...)
3. Soạn câu hỏi và hình thức trả lời (**mở, đóng, hỗn hợp, thang đo** ...)
4. Xác định cách dùng **thuật ngữ** (đơn giản, ngắn gọn, đơn nghĩa, đơn ý, tránh gợi ý, tránh thang đo không cân bằng, tránh câu hỏi ước đoán không rõ ràng...)
5. Xác định cấu trúc và hình thức (phân theo mục đích, đẹp, )
6. Thử lần 1 => sửa chữa => Bản nháp => Hoàn thiện

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU PHI THỰC NGHIỆM

**Câu hỏi mở:** Là dạng câu hỏi không có sẵn câu trả lời. Người trả lời tự do diễn đạt các ý trả lời. Dữ liệu thu thập được không có cấu trúc và thường không thể số hoá hay mã hoá. **Câu hỏi mở** phù hợp với **NC định tính**.

**Ví dụ:** Câu hỏi về quan điểm chính trị, điều tâm đắc nhất, lý do hài lòng ...

**Câu hỏi đóng:** Là dạng câu hỏi được định trước, có sẵn các câu trả lời và thang đo rõ ràng. Dữ liệu thu thập được có thể được phân tích, số hóa nhưng nó giới hạn sự trả lời. **Câu hỏi đóng** phù hợp với **NC định lượng**.

**Ví dụ:** SV các khóa học được đưa ra các câu hỏi nhận xét về giáo trình, bài giảng, sách, GV ... và được chỉ định trả lời theo thang đánh giá 5 mức độ: (rất hài lòng: +2; hài lòng: +1; bình thường: 0; không hài lòng: -1; rất không hài lòng: -2) để khảo sát sự hài lòng của SV.



# PHƯƠNG PHÁP THU THẬP DỮ LIỆU PHI THỰC NGHIỆM

## Câu hỏi mở:

**Ưu điểm:** Dễ trình bày; Khuyến khích người trả lời trình bày quan điểm dựa trên ngôn ngữ phù hợp với quan điểm của mình; Không bị ràng buộc bởi những câu trả lời cho sẵn; Dữ liệu thu được phong phú và có chiều sâu.

## Nhược điểm:

- ✓ Gây khó khăn cho thu thập và phân tích số liệu (không số hoá được);
- ✓ Tăng chi phí về tiền và thời gian;
- ✓ Buộc người trả lời phải suy nghĩ và kiên nhẫn;
- ✓ Mất nhiều thời gian → phụ thuộc sự nhiệt tình của người trả lời;
- ✓ Trong t/hợp người trả lời ít nói, hoặc câu trả lời tối nghĩa → khó thu thập thông tin;
- ✓ Chất lượng câu trả lời phụ thuộc vào k/n và sự khéo léo của PV viên.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU PHI THỰC NGHIỆM

## Câu hỏi đóng:

**Ưu điểm:** Dễ thực hiện; thời gian trả lời nhanh chóng, không phụ thuộc vào kinh nghiệm và sự khéo léo của PV viên; có thể số hoá số liệu.

## Nhược điểm:

- ✓ Câu trả lời phạm vi hẹp, có thể không phản ánh đúng ý muốn người trả lời;
- ✓ Phương án câu trả lời mang tính chủ quan của người hỏi, có thể mất tính khách quan;
- ✓ Dữ liệu thu thập thường không đủ rộng và sâu;
- ✓ Đòi hỏi phải soạn Bảng hỏi và câu hỏi phải có k/n, kiến thức bao quát vấn đề cần được NC.

# PHƯƠNG PHÁP THU THẬP DỮ LIỆU PHI THỰC NGHIỆM

## Thang đo (Câu hỏi đóng):

Một số thang đo phổ biến:

- ✓ **Thang đo cấp định danh:** dùng để xếp loại, không có ý nghĩa về lượng  
VD: Bạn thích sữa chua Yomilk kg? Thích (1); Không thích (2); Không ý kiến (3)
- ✓ **Thang đo cấp thứ tự:** Số đo dùng để so sánh thứ tự; không có ý nghĩa về lượng  
VD: Bạn vui lòng xếp thứ tự các hãng taxi bạn yêu thích: Vinasun ..., (1) ... (2) ... (3)
- ✓ **Thang đo cấp quãng:** Số đo dùng để chỉ khoảng cách, nhưng gốc 0 không có ý nghĩa  
VD: Thang đo Likert (1932) 5 mức độ: (Rất thích thích bình thường ghét rất ghét)
- ✓ **Thang đo cấp tỷ lệ:** Số đo dùng để đo độ lớn, và gốc 0 có ý nghĩa  
VD: Gia đình bạn có bao nhiêu xe honda; ..... xe

# PHƯƠNG PHÁP TRÌNH BÀY VÀ PHÂN TÍCH DỮ LIỆU

2.5.1. Phương pháp trình bày dữ liệu

2.5.2. Phương pháp phân tích dữ liệu

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

## a. Trình bày dạng bảng

**Cấu trúc bảng số liệu chứa các thành phần:**

- ✓ **Số; Tựa bảng; Tựa cột; Tựa hàng;** Phần thân chính chứa số liệu;
- ✓ **Chú thích** cuối bảng; Các đường ranh giới giữa các phần

**Có 3 đặc trưng cần thể hiện trong một bảng số liệu tốt:**

- ✓ Số liệu thể hiện tính hệ thống, cấu trúc một cách ý nghĩa;
- ✓ Số liệu rõ ràng, chính xác; có thể so sánh và rút ra nhiều kết luận lý thú về số liệu và mối quan hệ giữa các số liệu với nhau.
- ✓ Cần có số liệu nền (gần sát nhau) và số liệu nhấn nổi bật (khác nhau) để nhấn mạnh ưu điểm của NC đang thực hiện.

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

**Bảng 1.** So sánh kết quả của 2 phương pháp tối ưu hoá khi áp dụng đánh giá độ tin cậy

Kết quả	Phương pháp	Năng lượng biến dạng, [N.m]	Góc hướng sợi tối ưu, [độ]	Chiều dày tối ưu, [mm]	CPU, [s]
Khi chưa đánh giá độ tin cậy	SQP	0.001741	135.665796 46.305952	3.178809 2.004429 1.633524	394
	PP Điểm Trong	0.001741	135.149206 45.278144	3.182155 1.996115 1.642719	1220
Khi đánh giá độ tin cậy $\beta$ (99.8%)	SQP	0.002787	135.774375 45.038588	3.175694 2.001571 1.645471	495
	PP Điểm Trong	0.002787	135.128701 45.256593	3.175353 1.999261 1.650310	1962

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

## b. Trình bày dạng Hình

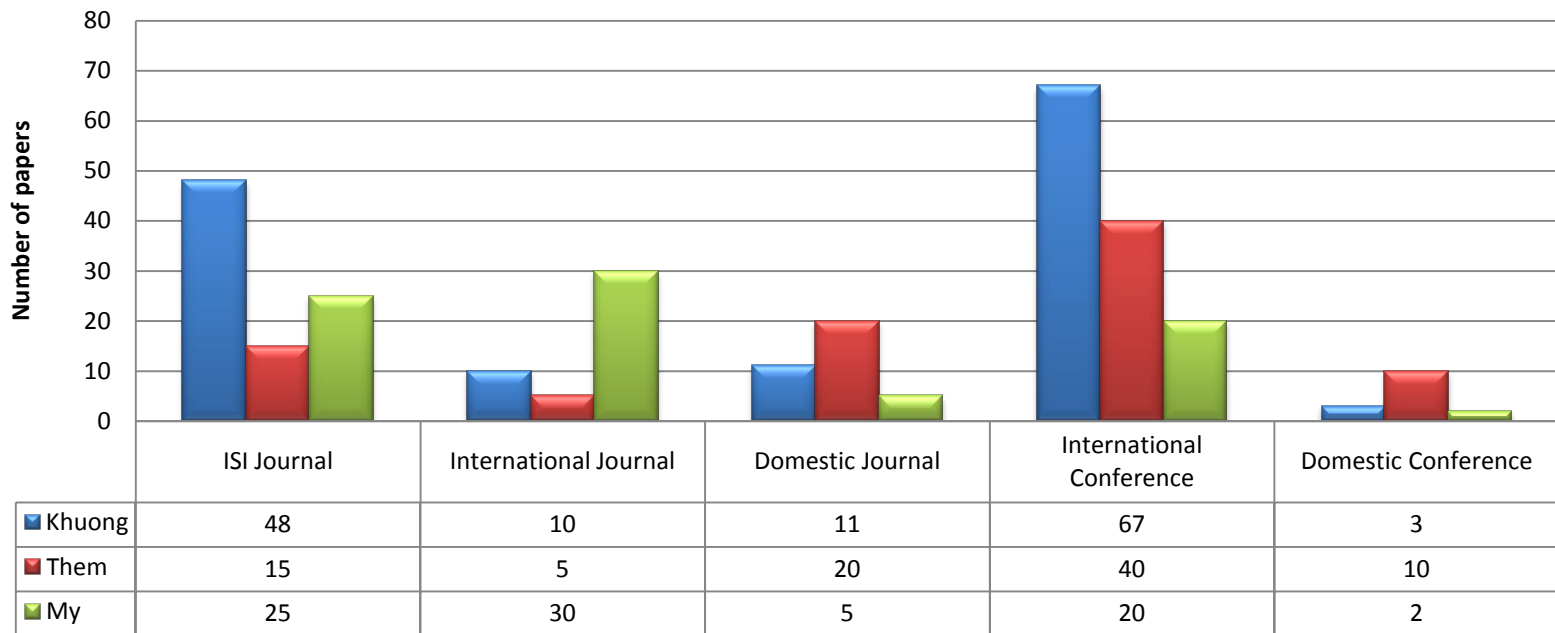
Sử dụng hình nhằm minh họa các **kết quả** và **mối quan hệ** giữa các biến cho độc giả dễ thấy hơn khi trình bày bằng bảng số liệu hoặc chữ. Sử dụng hình có thuận lợi là độc giả hiểu nhanh chóng các số liệu mà không mất nhiều thời gian khi nhìn bảng.

Các dạng hình được sử dụng gồm **biểu đồ cột** (colume chart), **biểu đồ thanh** (bar chart), **biểu đồ tần suất** (frequency histogram), **biểu đồ phân tán** (scatterplot), **biểu đồ đường biểu diễn** (line chart), **biểu đồ hình bánh** (pie chart), ...

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

**Biểu đồ cột và thanh:** So sánh số liệu theo nhóm, hoặc số liệu được phân nhóm, hoặc so sánh phần trăm tổng của nhiều số liệu.

Up to Year 2018

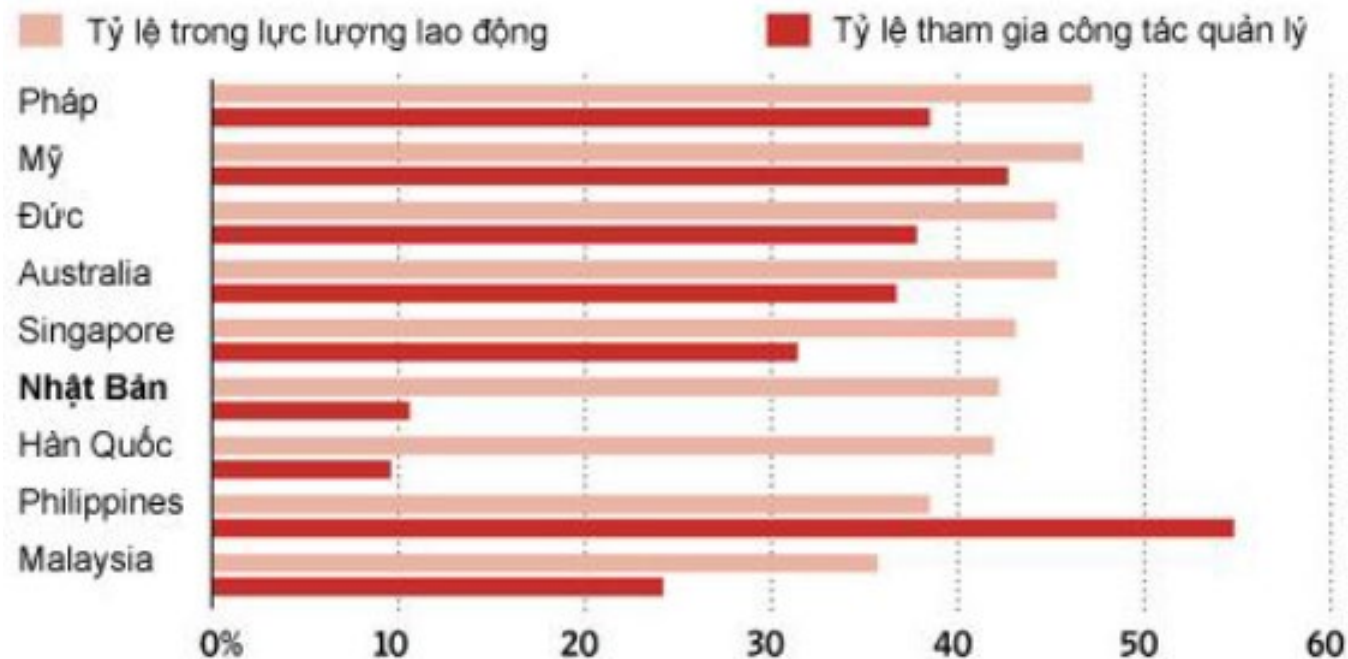


Hình : Ví dụ biểu đồ cột và thanh



# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

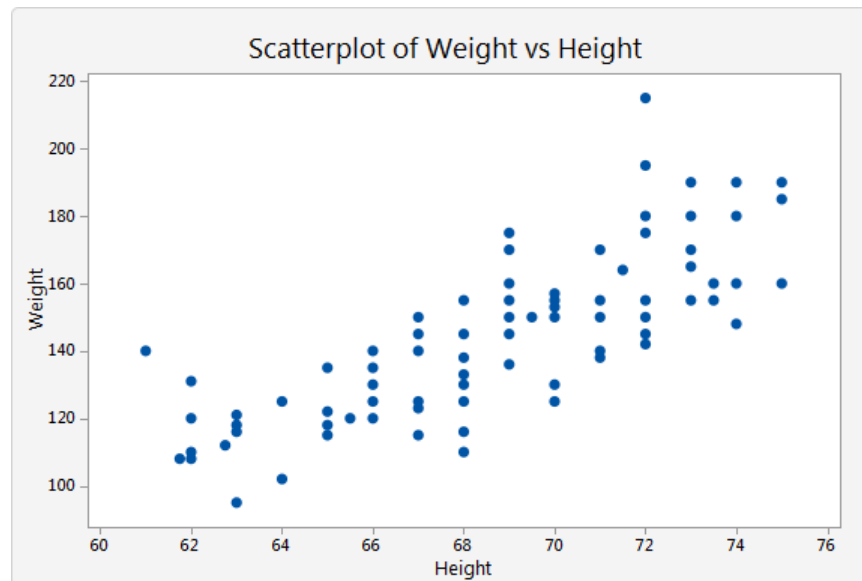
**Biểu đồ cột và thanh:** So sánh số liệu theo nhóm, hoặc số liệu được phân nhóm, hoặc so sánh phần trăm tổng của nhiều số liệu.



Hình : Tỷ lệ lao động nữ tại Nhật Bản và các quốc gia khác

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

**Biểu đồ phân tán:** Trình bày sự phân bố các số liệu và mối quan hệ giữa các số liệu. Các giá trị là các chấm phân bố và mối quan hệ được thể hiện bằng đường hồi quy tương quan. Biến phụ thuộc  $y$  có trục thẳng đứng phụ thuộc vào giá trị của biến độc lập  $x$  là trục nằm ngang.

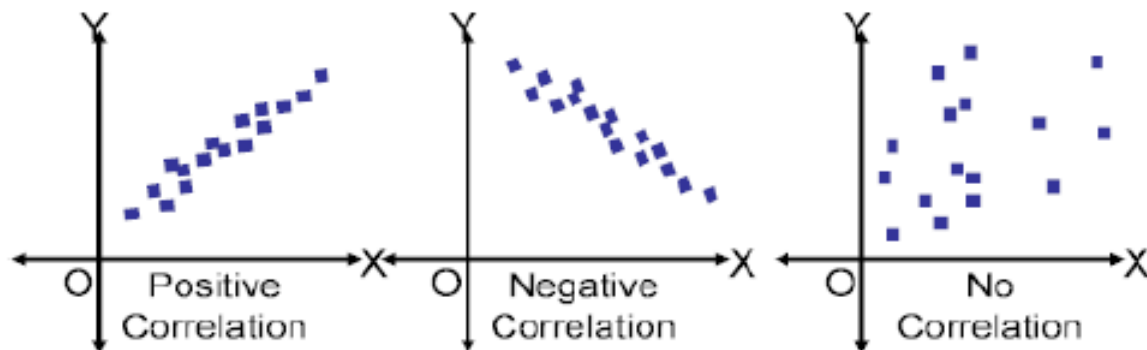


Hình : Ví dụ biểu đồ phân tán

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

## Đặc trưng của biểu đồ phân tán:

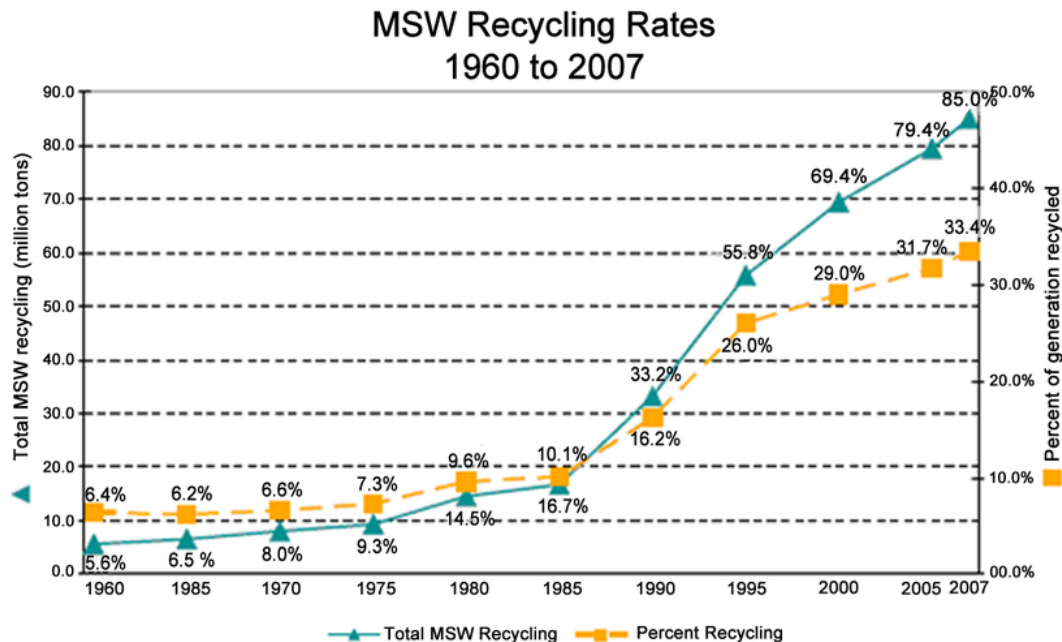
- **Tương quan thuận** (đồng biến) - một biến tăng, biến kia cũng tăng
- **Tương quan nghịch** (nghịch biến) - một biến tăng, biến kia giảm
- **Không tương quan** – Không có ảnh hưởng từ biến này đến biến kia



Hình : Ví dụ 3 dạng biểu đồ phân tán

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

**Biểu đồ đường biểu diễn:** Được trình bày nhằm diễn tả mối quan hệ của chiều hướng biến động và chức năng của các biến. Các giá trị là các điểm nối với nhau bởi đường thẳng hoặc cong. Có thể trình bày nhiều biến phụ thuộc là những đường biểu diễn trên cùng một hình.

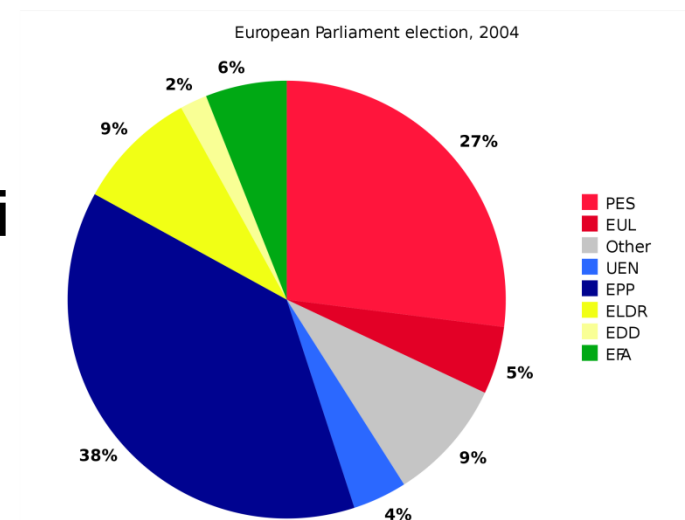


Hình : Ví dụ biểu đồ đường biểu diễn

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

**Biểu đồ hình bánh:** Trình bày, so sánh mối quan hệ tỷ lệ phần trăm tổng của các số liệu khác nhau. Nên tuân theo các quy tắc sau:

- Tổng số các số liệu có giá trị tổng không đổi (thường 100%)
- Các giá trị có sự khác biệt tương đối lớn (có ý nghĩa)
- Mỗi phần chia của hình nên được chú thích
- Số phần chia tương đối nhỏ (thường từ 3-7 phần) và không vượt quá 7



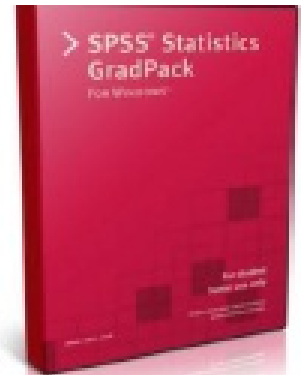
Hình : Ví dụ biểu đồ hình bánh

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

## ✓ Phần mềm hỗ trợ trình bày dữ liệu

Một số phần mềm hỗ trợ phân tích và trình bày dữ liệu thống kê:

- Phần mềm **SPSS** (Statistical Package for the Social Sciences)
- Phần mềm **SAS**
- Phần mềm **R**
- Phần mềm **STATA**
- Phần mềm **Epilnf**
- ...



# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

## ✓ Những điểm lưu ý chính

- Kết quả cần được biểu diễn dưới dạng biểu đồ hay bảng số liệu, và cần diễn giải ngắn gọn; Nếu **số liệu chính xác** là quan trọng thì nên dùng **bảng số liệu**; nếu **xu hướng** quan trọng hơn độ chính xác thì nên dùng **biểu đồ**;
- Khi mô tả bảng số liệu, tránh viết liệt kê. Nên chọn số liệu **nổi trội, quan trọng**, và liên quan **mục tiêu NC**. Tránh lặp lại dữ liệu giống nhau;
- Số liệu được trình bày tuần tự để trả lời trực tiếp các mục đích NC (hay câu hỏi NC); Khi mô tả kết quả NC, cần đề cập đến **xu hướng khác biệt** (directionality) và **mức độ khác biệt** (magnitude);
- Các bảng thống kê, biểu đồ, hình ảnh & ký hiệu cần được chú thích rõ.

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

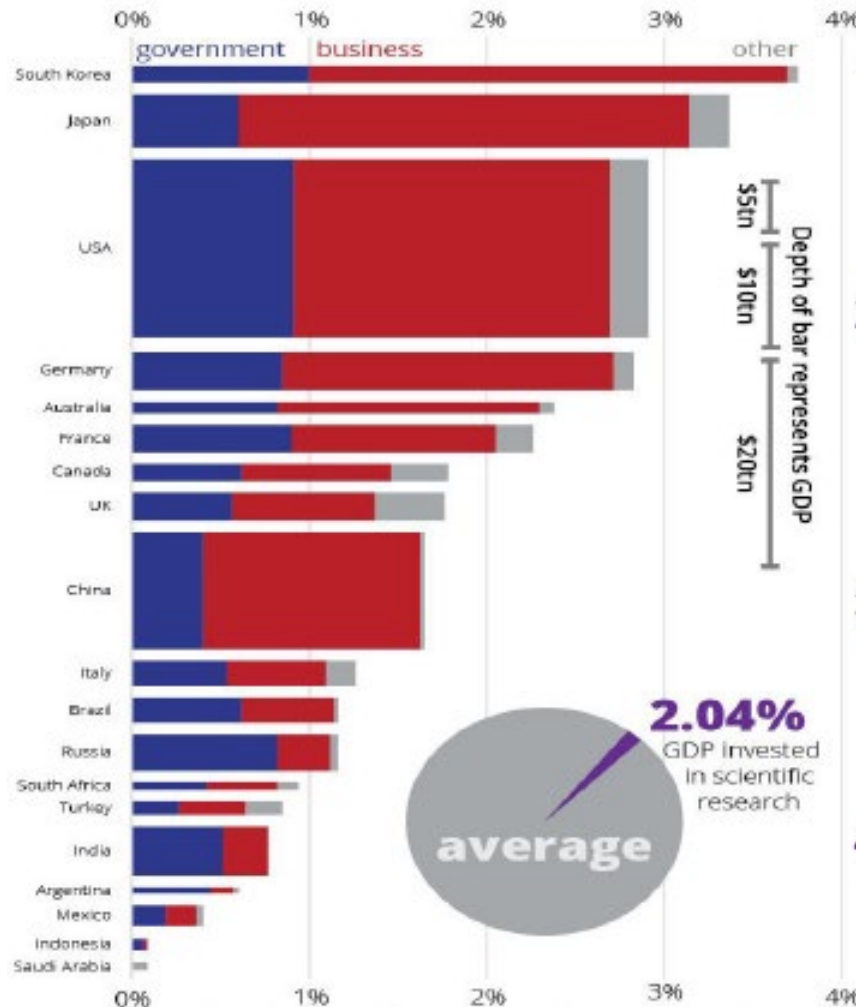
Phần nhận xét, giải thích, phân tích kết quả số liệu thường có 3 nhóm ý:

- Nhóm nhận xét về diễn biến và mối liên hệ giữa các số liệu
  - Nhóm giải thích nguyên nhân của các diễn biến (hoặc mối liên hệ) đó
  - Nhóm phân tích ý nghĩa / hàm ý của các diễn biến (hoặc mối liên hệ) đó
- Với nhóm (1), NCV cần làm rõ sự thay đổi của các giá trị theo thời gian, chỉ ra được khoảng tăng hoặc giảm mang tính chu kỳ, thấy được những thời điểm với những giá trị đặc biệt (sự tăng, giảm đặc biệt), với mức độ khác biệt.
- Với nhóm (2) & (3), NCV cần chú ý đến những yếu tố tác động trực tiếp, gián tiếp vào đối tượng, các mốc thời gian để giải thích/phân tích ý nghĩa hợp lý. NCV cần vận dụng kiến thức tổng hợp để việc giải thích/phân tích sâu sắc, thú vị.



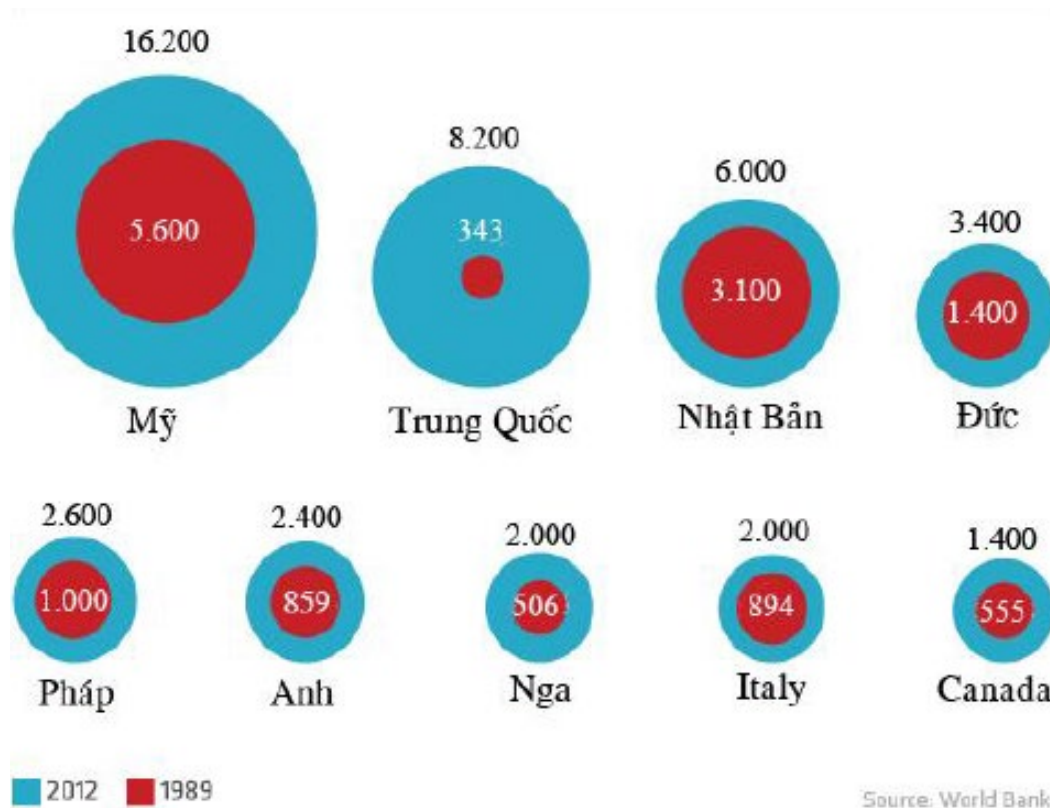
# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

Hình:  
Nguồn  
quỹ đầu  
tư từ nhà  
nước Và  
tư nhân  
vào  
NCKH  
của khối  
G20 tính  
theo %  
GDP



1. % GDP NCKH trung bình tất cả các quốc gia là 2.04%, trong đó khối nhà nước 0.65%, khối tư nhân 1.26%, và 0.13% từ các nguồn khác.
2. % GDP NCKH các quốc gia không đều. Hàn Quốc cao nhất 3.7%. Kế Nhật, Mỹ, Đức... Chú ý, Anh đầu tư dưới trung bình chỉ 1.7% GDP. TQ nỗ lực đầu tư NCKH ngày càng lớn với % GDP gần bằng Anh.
3. Hầu hết quốc gia, khối tư nhân đóng góp lớn nhất vào NCKH. Riêng tại Nga, Ấn Độ, Argentina và Indonesia, nguồn đầu tư chính là ngân sách của chính phủ.
4. Mỹ đầu tư lớn nhất vào NCKH, từ cả phía chính phủ và tư nhân. Tiếp theo Trung Quốc, thứ 3 Nhật Bản

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU



Hình : GDP các nước giai đoạn 1989 – 2012 (tỷ USD)

1. Sau 25 năm, ngôi vị nền KT lớn nhất TG vẫn thuộc về Mỹ. Giai đoạn 1989-2012, GDP nước này tăng gấp 3 lên hơn 16.000 tỷ USD.
2. Tuy nhiên, tốc độ tăng trưởng hai chữ số suốt 3 thập kỷ qua của Trung Quốc đã giúp nước này vượt qua Anh, Đức, Nhật Bản để trở thành nền KT lớn nhì TG từ năm 2010....

# PHƯƠNG PHÁP TRÌNH BÀY DỮ LIỆU

Lịch sử giá dầu thô thế giới từ năm 1861 tới nay



Hình : Lịch sử giá dầu thô thế giới từ 1861 đến nay