

THỐNG KÊ ỨNG DỤNG

ĐỖ LÂN

dolan@tlu.edu.vn
Đại học Thủy Lợi

Ngày 26 tháng 3 năm 2019

Nội dung môn học

- ➊ Tổng quan về Thống kê
- ➋ **Thu thập dữ liệu**
- ➌ Tóm tắt và trình bày dữ liệu bằng bảng và đồ thị
- ➍ Tóm tắt dữ liệu bằng các đại lượng thống kê mô tả
- ➎ Xác suất căn bản và biến ngẫu nhiên
- ➏ Phân phối của tham số mẫu và ước lượng tham số tổng thể
- ➐ Kiểm định giả thuyết về tham số một tổng thể
- ➑ Kiểm định giả thuyết về tham số hai tổng thể
- ➒ Phân tích phương sai
- ➓ Kiểm định chi - bình phương
- ➔ Hồi quy đơn biến
- ➕ Hồi quy đa biến

Phần III

Thu thập dữ liệu

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Khi thu thập dữ liệu, trước tiên dựa vào vấn đề nghiên cứu và mục tiêu nghiên cứu mà:

- xác định dữ liệu cần thu thập,
- thứ tự ưu tiên của các dữ liệu này.
- xác định rõ giới hạn, phạm vi dữ liệu.

Ví dụ

Khi nghiên cứu về vấn đề điều kiện sinh hoạt có ảnh hưởng đến kết quả học tập của sinh viên hay không, ta có hai nhóm dữ liệu chính cần phải thu thập là: điều kiện sinh hoạt và kết quả học tập. Với nhóm điều kiện sinh hoạt, có thể thu thập những dữ liệu sau:

- *Ở nhà với cha mẹ, người thân, ở trọ hay ở kí túc?*
- *Có phòng riêng hay không?*
- *Chỗ ở có ồn ào, chật trội, nóng bức không?*
- *Có phải làm thêm, phụ giúp bố mẹ không? Nếu có thời gian làm khoảng bao nhiêu giờ mỗi ngày?*
- ...

Ví dụ

Tuy nhiên những thông tin như sau là không cần thiết:

- *Bàn học màu gì? Cũ hay mới?*
- *Nhà được xây năm nào?*
- *Nhà vệ sinh có đẹp hay không?*
- ...

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Bài toán thực tế

Lớp trưởng 59HT muốn tìm hiểu điểm thi và điểm quá trình môn Thống kê ứng dụng để tư vấn các bạn.

Bài toán thực tế

Lớp trưởng 59HT muốn tìm hiểu điểm thi và điểm quá trình môn Thống kê ứng dụng để tư vấn các bạn.

- 1 Bạn có thể lấy được nguồn dữ liệu điểm bằng những cách nào?

Bài toán thực tế

Lớp trưởng 59HT muốn tìm hiểu điểm thi và điểm quá trình môn Thống kê ứng dụng để tư vấn các bạn.

- 1 Bạn có thể lấy được nguồn dữ liệu điểm bằng những cách nào?
- 2 So sánh ưu nhược điểm của từng cách thức bạn vừa đề cập ở trên.

Dữ liệu thứ cấp:

Dữ liệu thứ cấp: là dữ liệu thu thập từ những nguồn có sẵn, thường là đã qua tổng hợp, xử lí.

Nguồn:

- các báo cáo về sản xuất, tài chính, nhân sự,...;
- từ các cơ quan thống kê nhà nước
- từ báo, tạp chí;
- Các tổ chức, hiệp hội, viện nghiên cứu, ...

Dữ liệu thứ cấp:

Dữ liệu thứ cấp: là dữ liệu thu thập từ những nguồn có sẵn, thường là đã qua tổng hợp, xử lí.

Dữ liệu thứ cấp:

Dữ liệu thứ cấp: là dữ liệu thu thập từ những nguồn có sẵn, thường là đã qua tổng hợp, xử lí.

Nguồn:

- các báo cáo về sản xuất, tài chính, nhân sự,...;
- từ các cơ quan thống kê nhà nước
- từ báo, tạp chí;
- Các tổ chức, hiệp hội, viện nghiên cứu, ...

Ví dụ

Thông tin trên Thời báo kinh tế Sài Gòn: Theo khảo sát đầu năm 2015 của Nielsen phỏng vấn hơn 30.000 người tại 60 quốc gia ở tất cả các châu lục, báo cáo cho thấy:

- *77% người được khảo sát dùng tiền nhàn dỗi để tiết kiệm.*
- *18% dùng tiền để đầu tư cổ phiếu và rót vào các quỹ đầu tư.*

Dữ liệu sơ cấp: dữ liệu thu thập trực tiếp từ đối tượng nghiên cứu. Dữ liệu sơ cấp thường được thu thập theo một qui trình bài bản tùy theo nghiên cứu thống kê là nghiên cứu thử nghiệm hay nghiên cứu quan sát. Nguồn:

- thử nghiệm,
- điều tra phỏng vấn,
- thuê dịch vụ, ...

Dữ liệu sơ cấp: dữ liệu thu thập trực tiếp từ đối tượng nghiên cứu. Dữ liệu sơ cấp thường được thu thập theo một qui trình bài bản tùy theo nghiên cứu thống kê là nghiên cứu thử nghiệm hay nghiên cứu quan sát.

Dữ liệu sơ cấp: dữ liệu thu thập trực tiếp từ đối tượng nghiên cứu. Dữ liệu sơ cấp thường được thu thập theo một qui trình bài bản tùy theo nghiên cứu thống kê là nghiên cứu thử nghiệm hay nghiên cứu quan sát. Nguồn:

- thử nghiệm,
- điều tra phỏng vấn,
- thuê dịch vụ, ...

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Vì sao phải chọn mẫu

- ① Khi nghiên cứu một tổng thể lớn.
- ② Quá trình nghiên cứu phải phá hủy.
- ③ ...

Vì sao phải chọn mẫu

- ❶ Khi nghiên cứu một tổng thể lớn.
- ❷ Quá trình nghiên cứu phải phá hủy.
- ❸ ...

Do vậy, ta chỉ thu thập được một mẫu của tổng thể. Từ mẫu này sử dụng các phương pháp thống kê ta suy diễn cho tổng thể.

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Một số khái niệm về chọn mẫu

- Chọn mẫu ngẫu nhiên là quá trình lựa chọn sao cho trong mỗi lần chọn mỗi đơn vị tổng thể có cơ hội được chọn vào mẫu như nhau.

Một số khái niệm về chọn mẫu

- Chọn mẫu ngẫu nhiên là quá trình lựa chọn sao cho trong mỗi lần chọn mỗi đơn vị tổng thể có cơ hội được chọn vào mẫu như nhau.
- Quá trình chọn mẫu không đáp ứng được yêu cầu trên thì gọi là chọn mẫu không ngẫu nhiên.

Một số khái niệm về chọn mẫu

- Chọn mẫu ngẫu nhiên là quá trình lựa chọn sao cho trong mỗi lần chọn mỗi đơn vị tổng thể có cơ hội được chọn vào mẫu như nhau.
- Quá trình chọn mẫu không đáp ứng được yêu cầu trên thì gọi là chọn mẫu không ngẫu nhiên.
- Chọn mẫu có hoàn lại là cách chọn mà khi mỗi đơn vị tổng thể được chọn ra lại đặt trả lại tổng thể.

Một số khái niệm về chọn mẫu

- Chọn mẫu ngẫu nhiên là quá trình lựa chọn sao cho trong mỗi lần chọn mỗi đơn vị tổng thể có cơ hội được chọn vào mẫu như nhau.
- Quá trình chọn mẫu không đáp ứng được yêu cầu trên thì gọi là chọn mẫu không ngẫu nhiên.
- Chọn mẫu có hoàn lại là cách chọn mà khi mỗi đơn vị tổng thể được chọn ra lại đặt trả lại tổng thể.
- Chọn mẫu không hoàn lại là cách chọn mà khi mỗi đơn vị tổng thể được chọn ra không được đặt trả lại tổng thể.

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Chọn mẫu ngẫu nhiên đơn giản

- Khi số lượng đơn vị tổng thể ít (vài chục hoặc vài trăm), ta có thể chọn các đơn vị của mẫu bằng cách bốc thăm hay quay số.
- Khi cỡ là lớn và có danh sách thứ tự, ta có thể dùng phần mềm để chọn ngẫu nhiên.

Chọn mẫu ngẫu nhiên đơn giản

- Khi số lượng đơn vị tổng thể ít (vài chục hoặc vài trăm), ta có thể chọn các đơn vị của mẫu bằng cách bốc thăm hay quay số.
- Khi cỡ là lớn và có danh sách thứ tự, ta có thể dùng phần mềm để chọn ngẫu nhiên.

Ví dụ

- *Chọn ngẫu nhiên một khách hàng trong ngày để trao thưởng.*

Chọn mẫu ngẫu nhiên đơn giản

- Khi số lượng đơn vị tổng thể ít (vài chục hoặc vài trăm), ta có thể chọn các đơn vị của mẫu bằng cách bốc thăm hay quay số.
- Khi cỡ là lớn và có danh sách thứ tự, ta có thể dùng phần mềm để chọn ngẫu nhiên.

Ví dụ

- *Chọn ngẫu nhiên một khách hàng trong ngày để trao thưởng.*
- *Chọn ngẫu nhiên một khách hàng trong tháng để trao thưởng.*

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Lấy mẫu hệ thống



Ví dụ

Giả sử cần so sánh mức độ chăm học của K58 và K59 TLU. Ta có thể chọn mẫu gồm 300 sinh viên K58, 200 sinh viên K59 theo cách sau:

- ① Lấy số liệu từ phòng Đào tạo về danh sách sinh viên, giả sử có 3000 sinh viên K58, 2000 sinh viên K59.*
- ② Lấy mẫu cho K58: Thực hiện chia khoảng $3000/300 = 10$; Từ 1 đến 10 lấy một số ngẫu nhiên, chẳng hạn 5. Sau đó lấy tiếp: 15; 25; ... cho đến khi đủ 300 sinh viên.*
- ③ Làm tương tự cho việc chọn mẫu cho K59.*

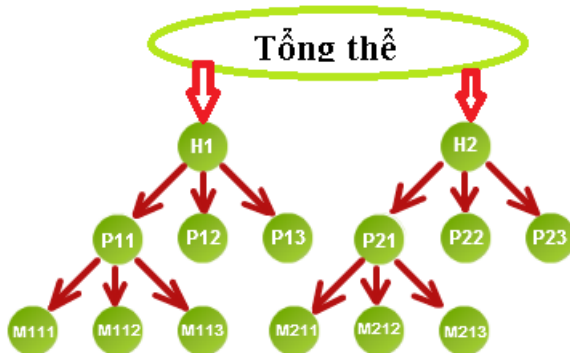
1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn



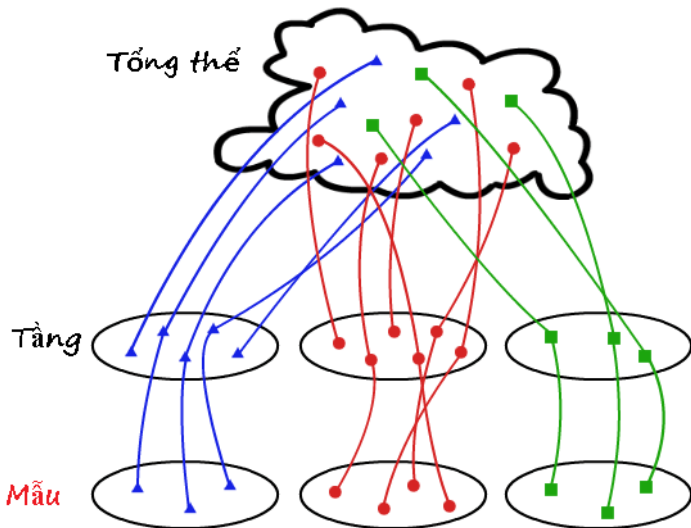
1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- **Lấy mẫu phân tầng**
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Lấy mẫu phân tầng



Ví dụ về lấy mẫu phân tầng

Ví dụ

Giả sử cần so sánh mức độ chăm học của K58 và K59 TLU. Ta có thể chọn mẫu gồm 300 sinh viên K58, 200 sinh viên K59 theo cách sau:

- ① Vì các ngành khác nhau ở chương trình đào tạo, mức độ yêu cầu về nhiệm vụ và thiên hướng yêu thích việc học của sinh viên, nên ta chia theo các ngành, mỗi ngành coi như một tầng.*
- ② Lấy số liệu từ phòng Đào tạo về danh sách sinh viên mỗi ngành học.*
- ③ Với K58, từ danh sách sinh viên ở mỗi ngành lấy một mẫu theo cách chọn hệ thống. Gộp các mẫu con lại để được mẫu đại diện cho K58. (nếu các ngành khác nhau có chênh lệch lớn về số lượng sinh viên thì cần phân bổ tỉ lệ hợp lí)*
- ④ Làm tương tự cho việc chọn mẫu cho K59.*

1 Xác định dữ liệu cần thu thập

2 Dữ liệu thứ cấp và dữ liệu sơ cấp

3 Các kĩ thuật lấy mẫu

- Một số khái niệm về chọn mẫu
- Chọn mẫu ngẫu nhiên
- Lấy mẫu hệ thống
- Lấy mẫu theo khối và lấy mẫu nhiều giai đoạn
- Lấy mẫu phân tầng
- Một số lỗi mắc phải khi chọn mẫu và thu thập dữ liệu

Lỗi do chọn mẫu và không do chọn mẫu

- Lỗi không do chọn mẫu:

Lỗi do chọn mẫu và không do chọn mẫu

- Lỗi không do chọn mẫu: mất dữ liệu, ghi chép sai, phân tích sai,

Lỗi do chọn mẫu và không do chọn mẫu

- Lỗi không do chọn mẫu: mất dữ liệu, ghi chép sai, phân tích sai, người trả lời phỏng vấn không hiểu, không nói hoặc phóng đại vấn đề được hỏi, ... Những lỗi này gọi là lỗi không phải chọn mẫu.

Lỗi do chọn mẫu và không do chọn mẫu

- Lỗi không do chọn mẫu: mất dữ liệu, ghi chép sai, phân tích sai, người trả lời phỏng vấn không hiểu, không nói hoặc phóng đại vấn đề được hỏi, ... Những lỗi này gọi là lỗi không phải chọn mẫu.
→ Không có phương pháp thống kê nào đo hoặc điều khiển được những lỗi kiểu này.

Lỗi do chọn mẫu và không do chọn mẫu

- Lỗi không do chọn mẫu: mất dữ liệu, ghi chép sai, phân tích sai, người trả lời phỏng vấn không hiểu, không nói hoặc phóng đại vấn đề được hỏi, ... Những lỗi này gọi là lỗi không phải chọn mẫu.
→ Không có phương pháp thống kê nào đo hoặc điều khiển được những lỗi kiểu này.
- Sai lầm do chọn mẫu: mẫu không đại diện được cho tổng thể. Khi mẫu không đại diện được cho tổng thể thì các tính toán trên mẫu không thể suy diễn cho tổng thể được.

Các nguyên nhân của lỗi do chọn mẫu

- 1 Nguyên nhân đầu tiên có thể do cách chọn mẫu, chẳng hạn khi tạo mẫu ngẫu nhiên, tổng thể được liệt kê thiếu một nhóm các phần tử mà lớp các phần tử này có đặc tính khác biệt với phần còn lại của tổng thể.

Các nguyên nhân của lỗi do chọn mẫu

- 1 Nguyên nhân đầu tiên có thể do cách chọn mẫu, chẳng hạn khi tạo mẫu ngẫu nhiên, tổng thể được liệt kê thiếu một nhóm các phần tử mà lớp các phần tử này có đặc tính khác biệt với phần còn lại của tổng thể.
- 2 Nguyên nhân thứ hai đến từ việc không trả lời (nonresponse) của ứng viên đã được chọn vào mẫu ngẫu nhiên. Điều này xuất hiện khi điều tra viên không thể liên lạc được với ứng viên hoặc ứng viên từ chối tham gia trả lời.

Các nguyên nhân của lỗi do chọn mẫu

- 1 Nguyên nhân đầu tiên có thể do cách chọn mẫu, chẳng hạn khi tạo mẫu ngẫu nhiên, tổng thể được liệt kê thiếu một nhóm các phần tử mà lớp các phần tử này có đặc tính khác biệt với phần còn lại của tổng thể.
- 2 Nguyên nhân thứ hai đến từ việc không trả lời (nonresponse) của ứng viên đã được chọn vào mẫu ngẫu nhiên. Điều này xuất hiện khi điều tra viên không thể liên lạc được với ứng viên hoặc ứng viên từ chối tham gia trả lời.
- 3 Nguyên nhân thứ ba là người được hỏi trả lời không trung thực các câu hỏi mang tính nhạy cảm ảnh hưởng đến lợi ích của họ.

Các nguyên nhân của lỗi do chọn mẫu

- 1 Nguyên nhân đầu tiên có thể do cách chọn mẫu, chẳng hạn khi tạo mẫu ngẫu nhiên, tổng thể được liệt kê thiếu một nhóm các phần tử mà lớp các phần tử này có đặc tính khác biệt với phần còn lại của tổng thể.
- 2 Nguyên nhân thứ hai đến từ việc không trả lời (nonresponse) của ứng viên đã được chọn vào mẫu ngẫu nhiên. Điều này xuất hiện khi điều tra viên không thể liên lạc được với ứng viên hoặc ứng viên từ chối tham gia trả lời.
- 3 Nguyên nhân thứ ba là người được hỏi trả lời không trung thực các câu hỏi mang tính nhạy cảm ảnh hưởng đến lợi ích của họ.
- 4 Nguyên nhân thứ tư đến từ bản thân câu hỏi lái người được hỏi trả lời theo dụng ý người khảo sát.

Ví dụ

Trong cuộc bầu cử tổng thống của Mỹ năm 1936 có hai ứng cử viên sáng giá cho chức vụ tổng thống là: Franklin D. Roosevelt (Đảng Dân chủ) và Alf Landon (Đảng Cộng hòa).

Ví dụ

Trong cuộc bầu cử tổng thống của Mỹ năm 1936 có hai ứng cử viên sáng giá cho chức vụ tổng thống là: Franklin D.Roosevelt (Đảng Dân chủ) và Alf Landon (Đảng Cộng hòa).

- *Tạp chí Literary Digest đã chọn một mẫu gồm 2.4 triệu người chọn từ danh bạ điện thoại và chọn từ tên các thành viên của các câu lạc bộ để điều tra. Kết quả điều tra: Alf Landon sẽ thắng Franklin D.Roosevelt với tỉ lệ là 57% và 43%.*

Ví dụ

Trong cuộc bầu cử tổng thống của Mỹ năm 1936 có hai ứng cử viên sáng giá cho chức vụ tổng thống là: Franklin D.Roosevelt (Đảng Dân chủ) và Alf Landon (Đảng Cộng hòa).

- *Tạp chí Literary Digest đã chọn một mẫu gồm 2.4 triệu người chọn từ danh bạ điện thoại và chọn từ tên các thành viên của các câu lạc bộ để điều tra. Kết quả điều tra: Alf Landon sẽ thắng Franklin D.Roosevelt với tỉ lệ là 57% và 43%.*
- *Cùng lúc đó, nhóm điều tra do George Gallup sáng lập đã chọn một mẫu gồm 5000 người theo một phương pháp chọn mẫu kiểu hệ thống và dự đoán Roosevelt sẽ dành chiến thắng trước đối thủ Alf Landon.*

Ví dụ

Thực tế Roosevelt đã dành chiến thắng lớn. Như vậy nhóm George Gallup với mẫu nhỏ hơn rất nhiều đã dự đoán đúng còn tạp chí Literary Digest đã dự đoán sai. Nguyên nhân vì sao?

Ví dụ

Thực tế Roosevelt đã dành chiến thắng lớn. Như vậy nhóm George Gallup với mẫu nhỏ hơn rất nhiều đã dự đoán đúng còn tạp chí Literary Digest đã dự đoán sai. Nguyên nhân vì sao?

Sai lầm của Literary Digest:

- *Chỉ chọn những người để điều tra từ danh bạ điện thoại và tên của các thành viên của các câu lạc bộ.*

Ví dụ

Thực tế Roosevelt đã dành chiến thắng lớn. Như vậy nhóm George Gallup với mẫu nhỏ hơn rất nhiều đã dự đoán đúng còn tạp chí Literary Digest đã dự đoán sai. Nguyên nhân vì sao?

Sai lầm của Literary Digest:

- *Chỉ chọn những người để điều tra từ danh bạ điện thoại và tên của các thành viên của các câu lạc bộ.*
- *Năm 1936, Mỹ chưa phục hồi từ cuộc đại suy thoái nên rất nhiều người thất nghiệp và có thu nhập thấp không có điện thoại và cũng không tham gia vào câu lạc bộ nào. Chương trình chọn mẫu của Literary Digest đã bỏ qua những người này, mà chính họ đã bỏ phiếu ủng hộ Roosevelt.*

Bài tập nhóm

Mỗi nhóm gồm khoảng 6-8 sinh viên thực hiện 2 nhiệm vụ:

Bài tập 1

Tìm hiểu và trình bày về ít nhất 3 ứng dụng của thống kê trong cuộc sống.

Bài tập 2

Thiết kế một nghiên cứu nhỏ.

- 1 Lên ý tưởng một chủ đề nghiên cứu.
- 2 Xác định dữ liệu cần thu thập gồm những biến nào? Thang đo của mỗi biến là gì?
- 3 Lập bảng hỏi hoặc cách thức thu thập dữ liệu.
- 4 Tổng thể cần nghiên cứu là gì?
- 5 Lựa chọn cách chọn mẫu nào? Trình bày cách thực hiện lấy mẫu.

Bài tập ôn luyện

Bài toán: Nghịch lý Simpson

Một người nghiên cứu muốn xác định xem giữa hai loại thuốc dùng để chữa cùng một loại bệnh, loại nào tốt hơn. Kết quả thống kê về về lượng người chữa được khỏi bệnh, phân biệt theo giới tính được cho bởi bảng dưới đây:

<i>Giới tính: Nữ</i>	<i>Thuốc I</i>	<i>Thuốc II</i>
<i>Chữa được</i>	150	15
<i>Không chữa được</i>	850	285
<i>Giới tính: Nam</i>	<i>Thuốc I</i>	<i>Thuốc II</i>
<i>Chữa được</i>	190	720
<i>Không chữa được</i>	10	180

Hỏi thuốc nào tốt hơn?

Dựa vào bảng thống kê này, có hai câu trả lời trái ngược nhau cho câu hỏi thuộc nào tốt hơn như sau:

Câu trả lời 1

Dựa vào bảng thống kê này, có hai câu trả lời trái ngược nhau cho câu hỏi thuốc nào tốt hơn như sau:

Câu trả lời 1

Thuốc I đem cho 1200 người dùng, chữa được cho 340 người.
Thuốc II đem cho 1200 người dùng, chữa được 735 người.
Như vậy thuốc II tốt hơn.

Câu trả lời 2

Dựa vào bảng thống kê này, có hai câu trả lời trái ngược nhau cho câu hỏi thuốc nào tốt hơn như sau:

Câu trả lời 1

Thuốc I đem cho 1200 người dùng, chữa được cho 340 người.
Thuốc II đem cho 1200 người dùng, chữa được 735 người.
Như vậy thuốc II tốt hơn.

Câu trả lời 2

Đối với nữ, tỉ lệ chữa được bệnh của thuốc I là 15, của thuốc II là 5.
Đối với nam, tỉ lệ chữa được bệnh của thuốc I là 95, của thuốc II là 80.
Trong cả 2 trường hợp thì tỉ lệ chữa khỏi bệnh của thuốc I đều cao hơn,
vậy thuốc I tốt hơn.

Vậy, câu trả lời nào là đúng?

Nội dung chính của chương

- Giới thiệu về khái niệm, cách thức thu thập hai kiểu dữ liệu: Sơ cấp và thứ cấp.

Nội dung chính của chương

- Giới thiệu về khái niệm, cách thức thu thập hai kiểu dữ liệu: Sơ cấp và thứ cấp.
- Trình bày lí do vì sao phải chọn mẫu và một số khái niệm về chọn mẫu.

Nội dung chính của chương

- Giới thiệu về khái niệm, cách thức thu thập hai kiểu dữ liệu: Sơ cấp và thứ cấp.
- Trình bày lí do vì sao phải chọn mẫu và một số khái niệm về chọn mẫu.
- Giới thiệu các kiểu lấy mẫu: Lấy mẫu ngẫu nhiên, lấy mẫu không ngẫu nhiên.

Nội dung chính của chương

- Giới thiệu về khái niệm, cách thức thu thập hai kiểu dữ liệu: Sơ cấp và thứ cấp.
- Trình bày lí do vì sao phải chọn mẫu và một số khái niệm về chọn mẫu.
- Giới thiệu các kiểu lấy mẫu: Lấy mẫu ngẫu nhiên, lấy mẫu không ngẫu nhiên.
- Một số lỗi trong quá trình chọn mẫu.

Những kiến thức cơ bản cần hiểu trong chương

- 1 Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.
- 2 Hiểu vì sao phải điều tra chọn mẫu mà không tổng điều tra.
- 3 Nắm được một số khái niệm về chọn mẫu: Chọn mẫu có hoàn lại, không hoàn lại, chọn mẫu ngẫu nhiên, không ngẫu nhiên.
- 4 Nắm được một số kĩ thuật chọn mẫu ngẫu nhiên: Chọn mẫu ngẫu nhiên đơn giản, chọn mẫu hệ thống, lấy mẫu phân tầng, lấy mẫu theo khối, lấy mẫu nhiều giai đoạn.
- 5 Nắm được một số kĩ thuật chọn mẫu không ngẫu nhiên: Chọn mẫu thuận tiện, chọn mẫu định mức, chọn mẫu phán đoán.
- 6 Hiểu những lỗi có thể gặp trong quá trình chọn mẫu.

Những kiến thức cơ bản cần hiểu trong chương

- 1 Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.

Những kiến thức cơ bản cần hiểu trong chương

- ① Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.
- ② Hiểu vì sao phải điều tra chọn mẫu mà không tổng điều tra.

Những kiến thức cơ bản cần hiểu trong chương

- ① Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.
- ② Hiểu vì sao phải điều tra chọn mẫu mà không tổng điều tra.
- ③ Nắm được một số khái niệm về chọn mẫu: Chọn mẫu có hoàn lại, không hoàn lại, chọn mẫu ngẫu nhiên, không ngẫu nhiên.

Những kiến thức cơ bản cần hiểu trong chương

- 1 Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.
- 2 Hiểu vì sao phải điều tra chọn mẫu mà không tổng điều tra.
- 3 Nắm được một số khái niệm về chọn mẫu: Chọn mẫu có hoàn lại, không hoàn lại, chọn mẫu ngẫu nhiên, không ngẫu nhiên.
- 4 Nắm được một số kĩ thuật chọn mẫu ngẫu nhiên: Chọn mẫu ngẫu nhiên đơn giản, chọn mẫu hệ thống, lấy mẫu phân tầng, lấy mẫu theo khối, lấy mẫu nhiều giai đoạn.

Những kiến thức cơ bản cần hiểu trong chương

- 1 Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.
- 2 Hiểu vì sao phải điều tra chọn mẫu mà không tổng điều tra.
- 3 Nắm được một số khái niệm về chọn mẫu: Chọn mẫu có hoàn lại, không hoàn lại, chọn mẫu ngẫu nhiên, không ngẫu nhiên.
- 4 Nắm được một số kĩ thuật chọn mẫu ngẫu nhiên: Chọn mẫu ngẫu nhiên đơn giản, chọn mẫu hệ thống, lấy mẫu phân tầng, lấy mẫu theo khối, lấy mẫu nhiều giai đoạn.
- 5 Nắm được một số kĩ thuật chọn mẫu không ngẫu nhiên: Chọn mẫu thuận tiện, chọn mẫu định mức, chọn mẫu phán đoán.

Những kiến thức cơ bản cần hiểu trong chương

- 1 Phân biệt dữ liệu sơ cấp, thứ cấp. Hiểu được ưu, nhược điểm của từng loại dữ liệu và cách thức thu thập từng loại dữ liệu.
- 2 Hiểu vì sao phải điều tra chọn mẫu mà không tổng điều tra.
- 3 Nắm được một số khái niệm về chọn mẫu: Chọn mẫu có hoàn lại, không hoàn lại, chọn mẫu ngẫu nhiên, không ngẫu nhiên.
- 4 Nắm được một số kĩ thuật chọn mẫu ngẫu nhiên: Chọn mẫu ngẫu nhiên đơn giản, chọn mẫu hệ thống, lấy mẫu phân tầng, lấy mẫu theo khối, lấy mẫu nhiều giai đoạn.
- 5 Nắm được một số kĩ thuật chọn mẫu không ngẫu nhiên: Chọn mẫu thuận tiện, chọn mẫu định mức, chọn mẫu phán đoán.
- 6 Hiểu những lỗi có thể gặp trong quá trình chọn mẫu.