Chương 2
THU THẬP DỮ LIỆU

Ths. Nguyễn Tiến Dũng
Viện Kinh tế và Quản lý, Trường ĐH Bách khoa Hà Nội
Email: dung.nguyentien3@hust.edu.vn

MỤC TIÊU CỦA CHƯƠNG

- Sau khi kết thúc chương này, người học có thể:
  - Phân biệt được DL thứ cấp và DL sơ cấp
  - Kể tên và nêu được đặc điểm, phạm vi ứng dụng của các kỹ thuật lấy mẫu ngẫu nhiên và không ngẫu nhiên
  - Phân biệt được DL định tính và DL định lượng và ý nghĩa của hai loại DL này trong phân tích TK.

CÁC NỘI DUNG CHÍNH

2.1 XÁC ĐỊNH DL CẦN THU THẬP

2.2 DỮ LIỆU THỨ CẤP VÀ DL SƠ CẤP

2.3 CÁC PP THU THẬP DL SƠ CẤP

2.4 CÁC KỸ THUẬT LẤY MẪU

2.5 DL ĐỊNH TÍNH VÀ DL ĐỊNH LƯỢNG

2.1 XÁC ĐỊNH DL CẦN THU THẬP

- Tiến trình NC
  - Vấn đề quản trị >> Mục tiêu NC >> Kế hoạch NC >> Thu thập DL >> Phân tích DL …
- TD: NC mức độ yêu thích chuyên ngành của SV và kết quả học tập
  - Mức độ yêu thích chuyên ngành
  - Kết quả học tập
  - DL khác:
    - Về SV: khoa, trường, học năm thứ mấy
    - Về môi trường học tập: đặc điểm nhà trường, xã hội …
- 2 loại DL:
  - DL thứ cấp
  - DL sơ cấp

2.2 DỮ LIỆU THỨ CẤP

- DL không có sẵn
- Tiếp cận nguồn DL
- Nguồn bên trong DN
- Nguồn bên ngoài DN
  - Các tổ chức chuyên TK của nhà nước: TCTK (gso.gov.vn), Cục TK …
  - Các tổ chức không chuyên TK
  - Các tổ chức dịch vụ cung cấp thông tin: TNS, A.C. Nielsen …

General Statistical Office in Vietnam: www.gso.gov.vn

English | FAQ | Sơ đồ website

TỔNG CỤC THỐNG KÊ

6B Hoàng Diệu, Ba Đình, Hà Nội

Trang chủ | Số liệu thống kê | Phương pháp thống kê | Tin tức | Ấn phẩm thống kê | Danh mục | Văn bản pháp lý | Giới thiệu

THÔNG TIN MỚI

Trị giá xuất, nhập khẩu phân theo

TIN TỨC - SỰ KIỆN

Hội nghị tổng kết công tác thống kê Bộ, ngành 2011 - 2013 và phương hướng nhiệm vụ 2013 – 2015 01/08/2013

- Lễ trao tặng Kỷ niệm chương "Vì sự nghiệp Báo chí Việt Nam" (31/07/2013)
- Hội nghị sơ kết công tác Công đoàn 6 tháng đầu năm, triển khai kế hoạch công tác 6 tháng cuối năm (30/07/2013)
- Khảo sát thực tế của Lớp Bồi dưỡng kiến thức Quản lý Nhà nước (22/07/2013)
- Hội nghị công tác tổ chức, cán bộ và đào tạo của Tổng cục Thống kê (15/07/2013)

Một số chỉ tiêu chủ yếu 7 tháng đầu năm 2013 (tăng/giảm) so với cùng kỳ năm 2012 (%)

Chỉ số sản xuất công nghiệp	+5,2
Tổng mức bán lẻ hàng hoá và doanh thu dịch vụ tiêu dùng	+12,0
Tổng kim ngạch xuất khẩu	+14,3
Tổng kim ngạch nhập khẩu	+15,0
Khách quốc tế đến Việt Nam	+5,9
Vốn đầu tư ngân sách nhà nước thực hiện so với kế hoạch năm	54,4

ĐIỀU TRA THỐNG KÊ

- Chương trình điều tra thống kê
- Kế hoạch điều tra hàng năm
- Kết quả điều tra

DANH MỤC

2.3 DỮ LIỆU SƠ CẤP

- DL không có sẵn
- Tự làm hay thuê ngoài
- Các phương pháp thu thập DL sơ cấp
  - Quan sát
  - Nhật ký
  - Phỏng vấn sâu
  - Thảo luận nhóm
  - Điều tra chọn mẫu: PV trực tiếp, qua điện thoại, qua thư

2.4 CÁC KỸ THUẬT (PHƯƠNG PHÁP) LẤY MẪU

Yêu cầu đối với lấy mẫu

- Sai số
- Chi phí
- Thời gian

Các phương pháp lấy mẫu ngẫu nhiên

- Lấy mẫu ngẫu nhiên đơn giản
- Lấy mẫu hệ thống
- Lấy mẫu phân tầng
- Lấy mẫu theo cụm

Các phương pháp lấy mẫu không ngẫu nhiên

- Lẫy mẫu thuận tiện
- Lẫy mẫu theo định mức
- Lẫy mẫu bằng phán đoán
- Lấy mẫu theo giới thiệu

2.4.1 Các kỹ thuật lấy mẫu ngẫu nhiên

- 2.4.1.1 Lấy mẫu ngẫu nhiên đơn giản
  - Tra bảng số ngẫu nhiên
  - Rút thăm ngẫu nhiên
  - Khởi tạo số ngẫu nhiên:
    - Hàm RANDBETWEEN(a;b)

2.4.1.2 Lấy mẫu hệ thống

- N = 49; n = 5
- Cách 1:
  - Tính bước nhảy k = [N/n] = 9
  - Chọn ngẫu nhiên p.tử đầu tiên trong [1;9] → 6
  - Tính ra các p.tử tiếp theo: 15, 24, 33, 42
- Cách 2:
  - Tính bước nhảy k = N/n = 9,8 → làm tròn là 10.
  - Chọn ngẫu nhiên p.tử đầu tiên trong [1;10] → 7
    - Các phần tử tiếp theo: 17, 27,37,47 (đủ)
  - Chọn ngẫu nhiên p.tử đầu tiên: [1;10] → 10
    - Các phần tử tiếp theo: 20, 30, 40, 50
    - 50 – 49 = 1 → lấy mẫu quay vòng

2.4.1.3 Lấy mẫu phân tầng

- Tổng thể gồm nhiều nhóm (gọi là tầng), có đặc điểm khác nhau.
- Yêu cầu: lấy ra một số lượng hay tỷ lệ nhất định theo từng nhóm (tầng).
- TD: Có DS gồm N = 20 khách nam + 20 khách nữ. Cần chọn ra n = 4 khách nam và 4 khách nữ để phỏng vấn (rút mẫu đều .
- Theo mục đích, tỷ lệ rút mẫu trong các tầng có thể không đều

2.4.1.4 Lấy mẫu theo cụm (cluster sampling)

- Tổng thể gồm nhiều nhóm (gọi là cụm), có đặc điểm khác nhau, mỗi cụm lại có thể bao gồm các cụm con.
- Lấy mẫu theo cụm 1 giai đoạn
  - Chọn ngẫu nhiên một số cụm
  - Trong các cụm đã chọn, chọn ngẫu nhiên một số p.tử
- Lấy mẫu theo cụm 2 giai đoạn
  - Chọn ngẫu nhiên một số cụm
  - Trong các cụm đã chọn, lấy ngẫu nhiên một số cụm con
  - Trong các cụm con đã chọn, lấy ngẫu nhiên một số p.tử
- TD: N = 20 nhà x 10 tầng x 10 hộ; n = 100

2.4.2 Các kỹ thuật lấy mẫu không ngẫu nhiên (phi xác suất)

- Lấy mẫu thuận tiện (convenience sampling)
- Lấy mẫu định mức (quota sampling)
- Lấy mẫu theo phán đoán (judgemental sampling)
- Lấy mẫu theo giới thiệu (snowball sampling)

2.5 DỮ LIỆU ĐỊNH TÍNH VÀ ĐỊNH LƯỢNG

- Dữ liệu định tính
  - DL định danh
  - DL thứ bậc
- Dữ liệu định lượng
  - DL khoảng
  - DL tỷ lệ
- Không thể tính TB và PS của DL định tính
- Thang đo → Dữ liệu → Phương pháp PT DL