

Phương pháp lấy mẫu nghiên cứu

Tháng 10 năm 2008

nielsen

The Nielsen logo is displayed in a serif font. The letter 'n' is blue, while the remaining letters 'ielsen' are grey. Below the text, there is a horizontal row of nine grey dots, each centered under a letter: 'n', 'i', 'e', 'l', 's', 'e', 'n', and two additional dots at the end.

Nội dung

1. Tiêu chuẩn của Nielsen
2. Phương pháp Door-to-door
 - a. Khung lấy mẫu
 - b. Thiết kế mẫu
 - c. Việc lấy mẫu trong thực tế.
3. Các phương pháp lấy mẫu khác

1. Tiêu chuẩn của Nielsen

Tiêu chuẩn của Nielsen về thiết kế và phân tích nghiên cứu

- Công ty Nielsen đã biên soạn một tài liệu quy định về những tiêu chuẩn trong thiết kế và phân tích nghiên cứu được áp dụng trên phạm vi toàn cầu, tài liệu đó có tên là WatchBuilder.
- Bộ phận Msci được kiểm tra hàng năm để đảm bảo việc thực thi nghiêm ngặt những tiêu chuẩn có trong WatchBuilder.
- Việc kiểm tra hàng năm còn có tác dụng cập nhật, phát triển quy trình nghiên cứu của từng nước cho phù hợp với cả khu vực.

2. Phương pháp Door-to-door

- i. Khung lấy mẫu
- ii. Thiết kế mẫu
- iii. Việc chọn mẫu trong thực tế

Khung lấy mẫu

Cấu trúc quản lý hành chính của Việt Nam

Dân số 2009: 85,789,573 người

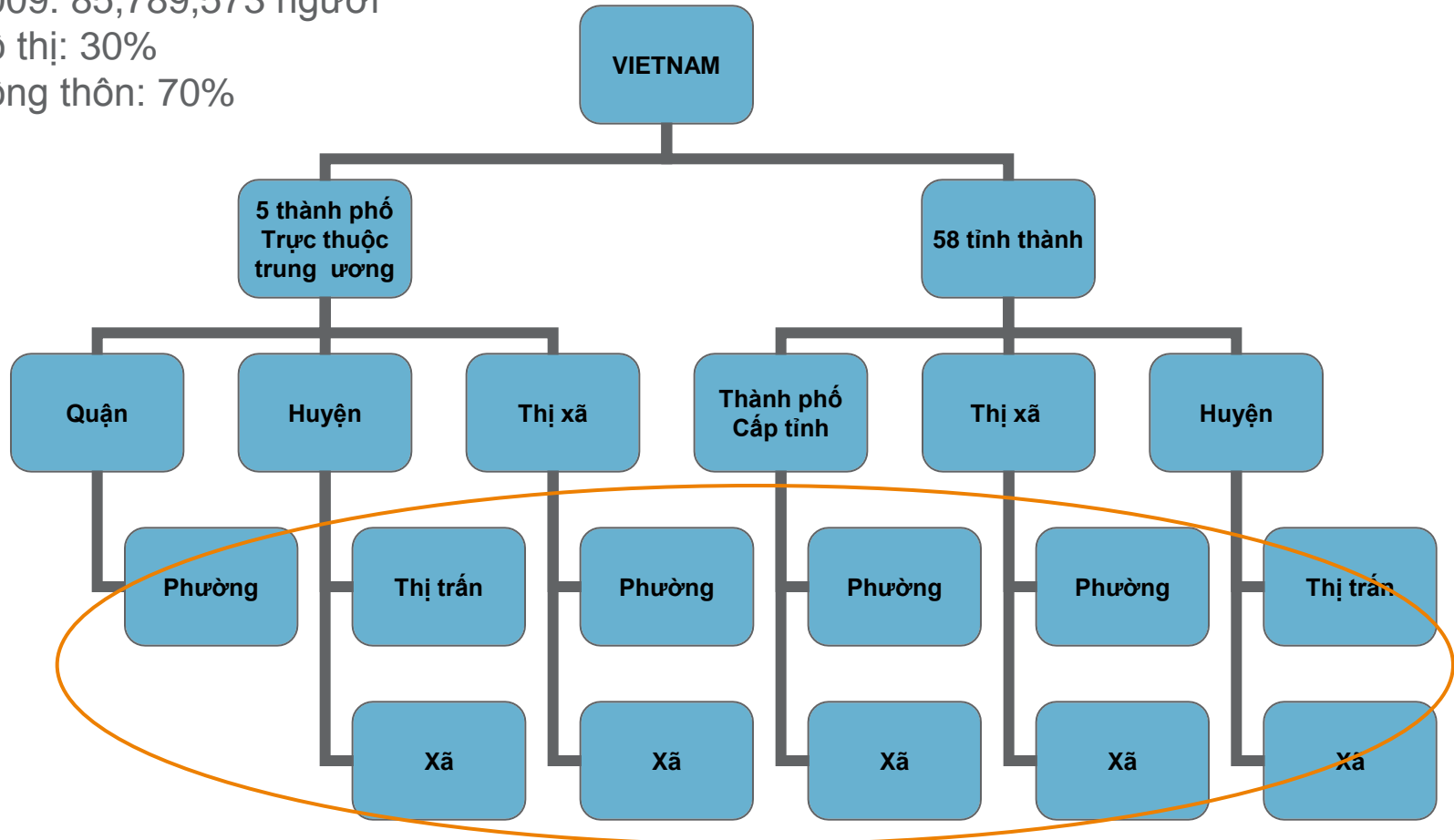
Dân số đô thị: 30%

Dân số nông thôn: 70%

Cấp tỉnh/thành phố

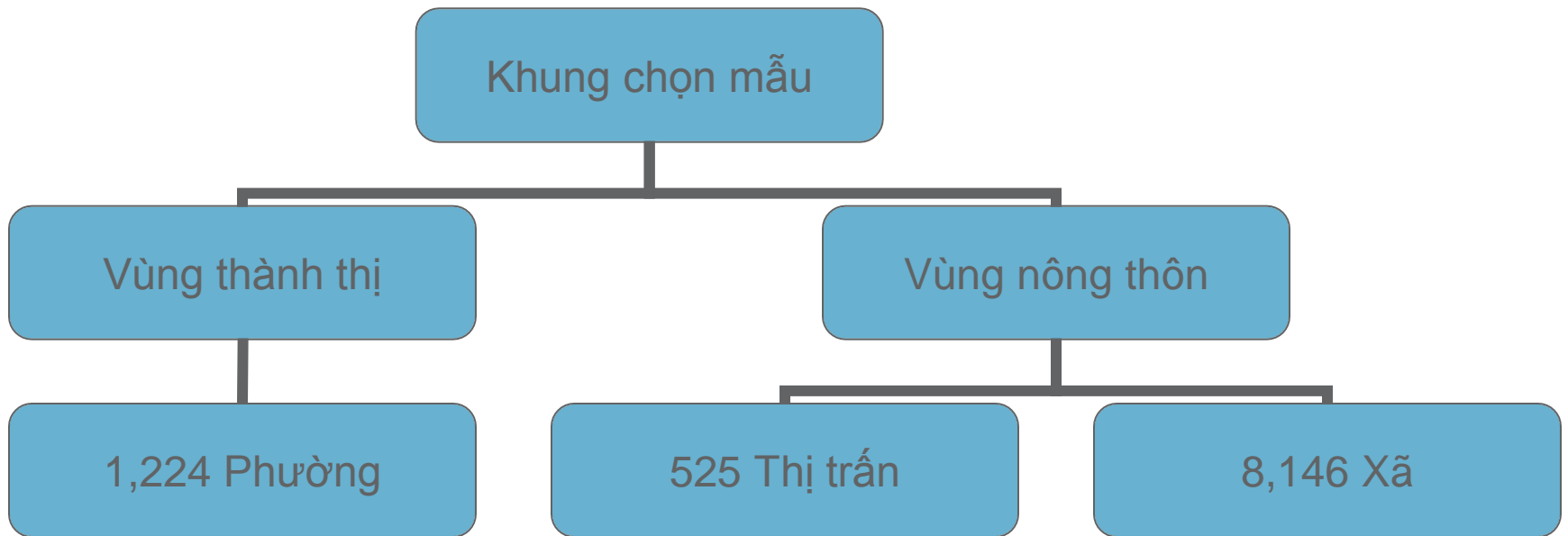
Cấp quận/huyện
Thị xã

Cấp phường



Khung chọn mẫu

Khung chọn mẫu



Đơn vị chọn mẫu cơ sở (PSU)

- Ổn định
- Dễ nhận diện theo quản lý hành chính.
- Có mật độ dân số khả thi.
- Phân tầng theo khu vực địa lý
- Cập nhật hàng năm.

Khung chọn mẫu chính

- Khung chọn mẫu chính của Nielsen Việt Nam được tạo ra dựa trên danh sách các đơn vị hành chính địa giới của tổng cục thống kê Việt nam.
- Khung chọn mẫu chính bao gồm 2 vùng:
 - Vùng thành thị: 1,224 phường;
 - Vùng nông thôn: 525 thị trấn và 8,146 xã (không tính 66 thị trấn và 932 xã vùng núi hoặc hải đảo).

Đơn vị chọn mẫu cơ sở.

- Nielsen Việt nam chọn cấp độ phường/xã/thị trấn làm đơn vị chọn mẫu cơ sở vì những lý do như sau:
 - Các đơn vị này có tính ổn định theo việc quản lý hành của chính phủ.
 - Các đơn vị này được phân chia địa giới rõ ràng .
 - Có dân cư sinh sống và dân số được cập nhật hàng năm.
- Các đơn vị chọn mẫu cơ sở được chọn dựa vào số dân số mà chúng có.

Phân nhóm các đơn vị chọn mẫu cơ sở.

- Có 3 cấp độ phân nhóm:
 - Thành thị/nông thôn;
 - Thành phố/ tỉnh;
 - Quận / Huyện.
- Định nghĩa về thành thị/nông thôn
 - Theo nghị định 72/2001 của quốc hội Việt Nam (ND 72/2001 QH), vùng thành thị phải đảm bảo các yếu tố sau:
 - Phần trăm lao động phi nông nghiệp tối thiểu là 65%.
 - Dân số tối thiểu là 4000 người.
 - Cơ sở hạ tầng đô thị đầy đủ.
- Ví dụ
 - **Thành thị** bao gồm các phường (thuộc quận và thị xã).
 - **Bán thành thị** bao gồm các thị trấn (thuộc huyện)
 - **Nông thôn** bao gồm các xã (thuộc huyện)

Hệ thống bản đồ số

- Nielsen Việt Nam đã có hệ thống bản đồ số cho 5 thành phố trực thuộc trung ương (Hanoi, Haiphong, Danang, HCMC and Cantho), theo các cấp độ:
 - Quận;
 - Phường;
 - Block.
- Đối với các tỉnh thành khác, Nielsen sử dụng “Atlas hành chính địa giới Việt Nam” do cục bản đồ xuất bản để xác định địa giới các phường, thị trấn, xã.
- Lý do của việc ưu tiên phát triển cơ sở dữ liệu bản đồ số ở 5 thành phố chính là do các thành phố này tập trung nhiều dân cư có nhu cầu tiêu dùng cao.
- Bộ phận Reference Data (RD) chịu trách nhiệm quản lý dữ liệu bản đồ số này.

Thiết kế mẫu

Quy trình chọn mẫu

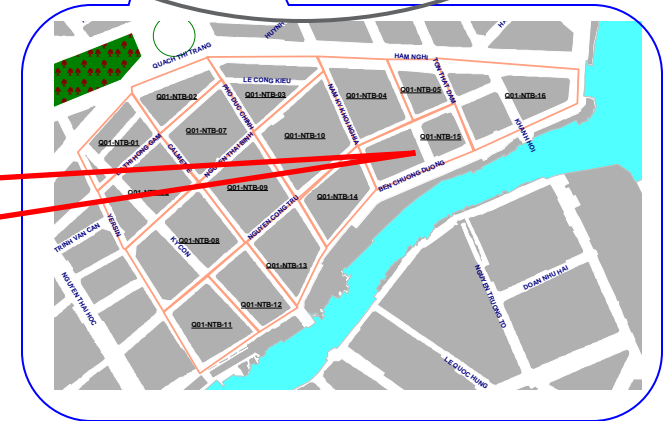
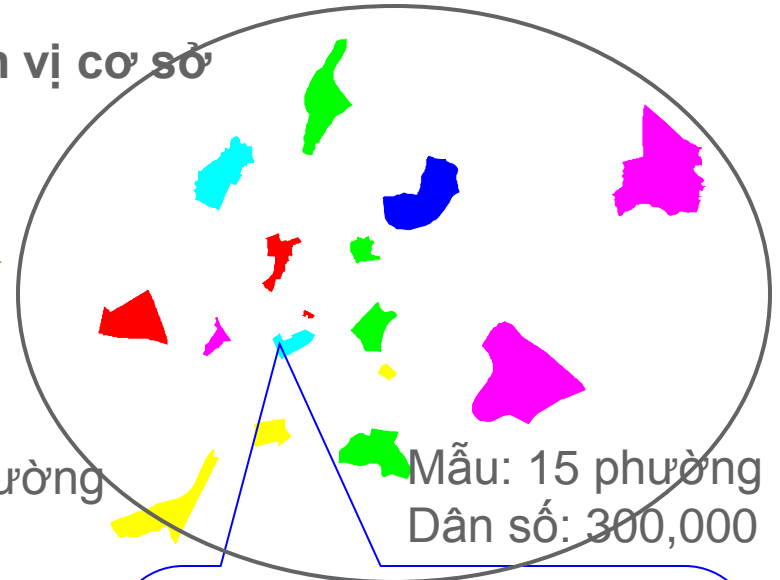
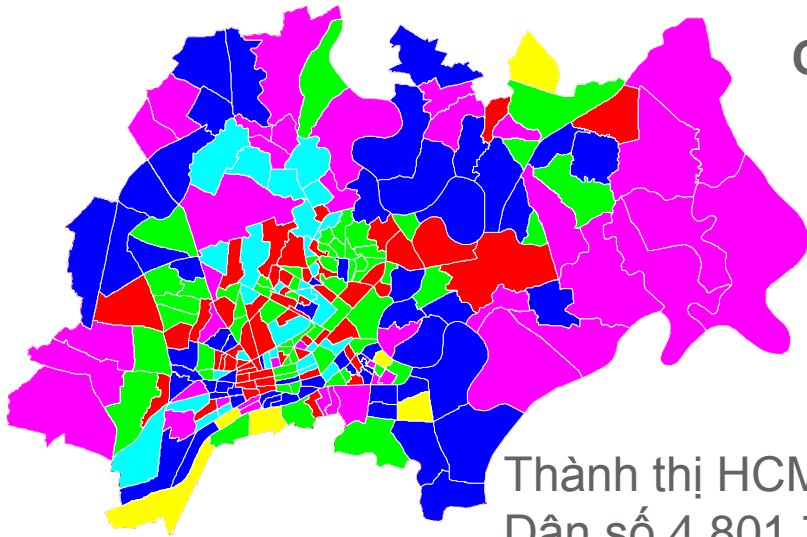
- Việc lấy mẫu được thực hiện tuần tự trên 4 giai đoạn
 - Giai đoạn 1 – Chọn đơn vị lấy mẫu cơ bản (Phường/ xã/ thị trấn)
 - Giai đoạn 2 - Chọn Block.
 - Giai đoạn 3 – Chọn hộ gia đình.
 - Giai đoạn 4 – Chọn đáp viên.

Bước 1 chọn đơn vị lấy mẫu cơ sở

- Một số đơn vị chọn mẫu cơ sở (phường/xã/thị trấn) sẽ được chọn ra theo phương pháp PPS (probability proportional to size) để làm các đơn vị đại diện cho tỉnh/thành phố.
- Mỗi đơn vị chọn mẫu cơ sở có một số lượng block nhất định được xác định và quản lý bởi bộ phận RD. Block bắt đầu của mỗi đơn vị cơ sở được chọn ngẫu nhiên bởi bộ phận Msci sau đó, điểm bắt đầu của mỗi block sẽ được chọn bởi bộ phận RD trên tờ bản đồ.
- Trung bình, số đáp viên được chọn ra trong mỗi đơn vị cơ sở không được vượt quá 20.
- Level 3.2 không chọn ngẫu nhiên đơn vị chọn mẫu cơ sở bản.

Ví dụ

Chọn đơn vị cơ sở



Chọn ngẫu nhiên block bắt đầu Chọn block trong mỗi phường

Bước 2 – Chọn hộ gia đình

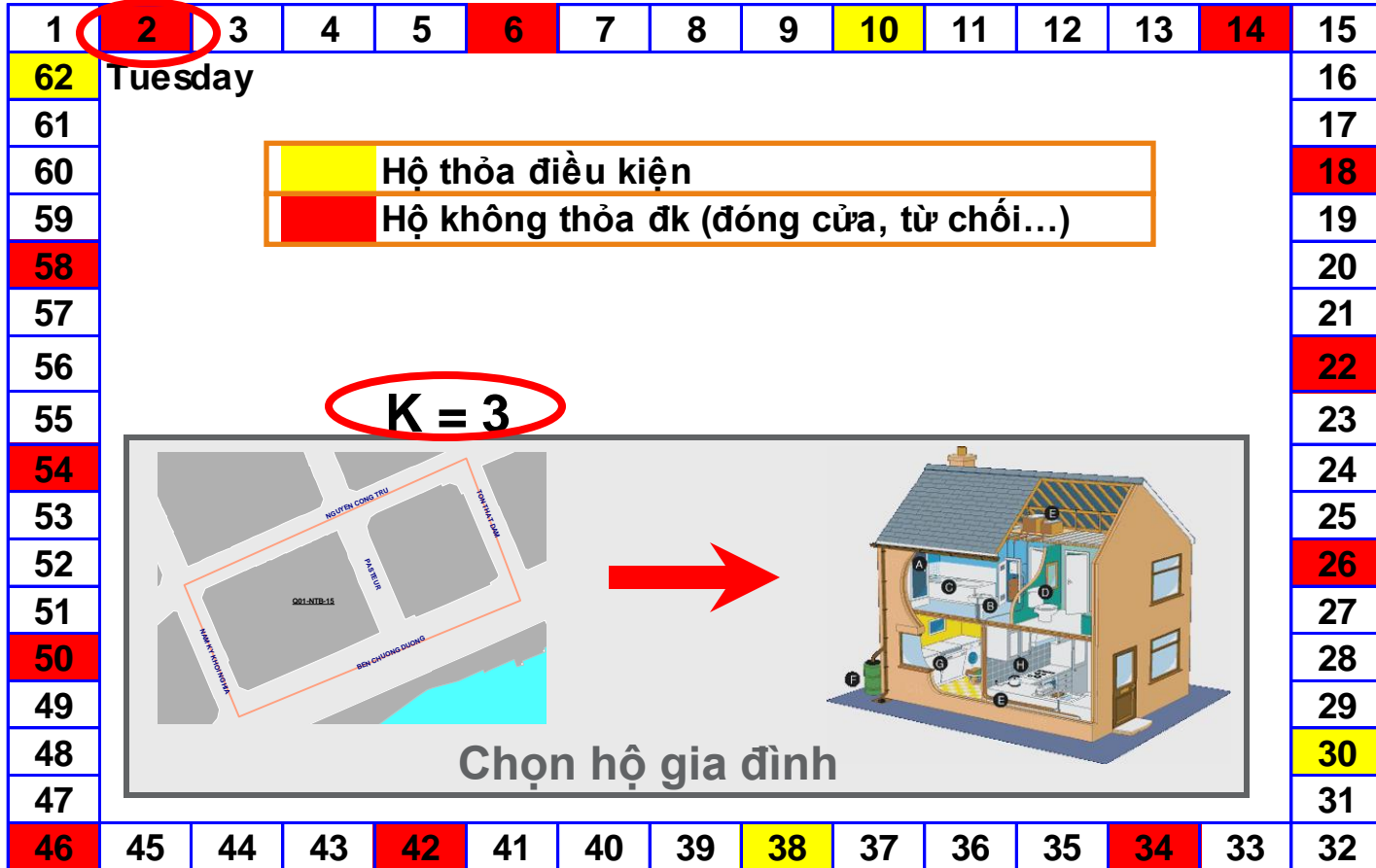
- Trong mỗi block, các hộ sẽ được chọn ngẫu nhiên theo nguyên tắc bàn tay phải, điểm bắt đầu ngẫu nhiên và bước nhảy K. Hoặc các hộ được chọn thuận tiện tùy theo yêu cầu của dự án.
 - Điểm bắt đầu được chọn bởi bộ phận RD.
 - Khoảng cách giữa hai hộ tiếp xúc được xác định bởi Msci.
- Trường hợp các dự án đi level 3, hộ gia đình sẽ được PVV chọn thuận tiện.

$$k = \frac{\text{BlockUniverseSize} * \text{ECR}}{\text{SampleSize per PSU}} - 1$$

ECR	<5%	5%-20%	20%->30%	30%->40%	40%->50%	>50%
K	Level 3 K>=0	Level 3 K>=1	Random K=2 or K=3	Random K=3 or K=4 or K=5	Random K=5 or K=6	Random K=6 or higher

Ví dụ

Điểm bắt đầu



Trường hợp gặp phải chung cư



- Mỗi toàn nhà chung cư được xem như một block.
- Phỏng vấn viên sẽ bắt đầu đi từ tầng cao nhất xuống tầng thấp nhất để chọn hộ gia đình và đáp viên.
- Đối với một số dự án (ví dụ như CLT) mỗi tầng chỉ chọn 1 đáp viên và chọn không quá 12 đáp viên trong một tòa nhà chung cư.

Các cấp độ lấy mẫu

- Level 1

- Chọn **ngẫu nhiên** các đơn vị lấy mẫu cơ bản (phường/xã).
- Trong mỗi đơn vị lấy mẫu cơ bản, chọn **ngẫu nhiên** các block.
- Trong mỗi block chọn **ngẫu nhiên** các hộ gia đình.
- Trong mỗi hộ gia đình chọn **ngẫu nhiên** ra 1 đáp viên (Kish grid/ call back 3 times)

- Level 2

- Chọn **ngẫu nhiên** các đơn vị lấy mẫu cơ bản (phường/xã).
- Trong mỗi đơn vị lấy mẫu cơ bản, chọn **ngẫu nhiên** các block.
- Trong mỗi block chọn **ngẫu nhiên** các hộ gia đình.
- Trong mỗi hộ gia đình chọn ra 1 đáp viên theo **quota** cho trước (Độ tuổi, Giới tính, tầng lớp kinh tế, Tình trạng làm việc...) hoặc chọn **thuận tiện** ra một người đại diện cho cả hộ gia đình (ví dụ: người quyết định mua các sản phẩm nấu nướng cho gia đình...)

Các cấp độ lấy mẫu

- Level 3.1

- Chọn **ngẫu nhiên** các đơn vị lấy mẫu cơ bản (phường/xã).
- Trong mỗi đơn vị lấy mẫu cơ bản, chọn **thuận tiện** các block.
- Trong mỗi block chọn **thuận tiện** các hộ gia đình.
- Trong mỗi hộ gia đình chọn ra 1 đáp viên theo **quota/ thuận tiện**.

- Level 3.2: snow balling

- Chọn **thuận tiện** các đơn vị lấy mẫu cơ bản (phường/xã).
- Trong mỗi đơn vị lấy mẫu cơ bản, chọn **thuận tiện** các block.
- Trong mỗi block chọn **thuận tiện** các hộ gia đình.
- Trong mỗi hộ gia đình chọn ra 1 đáp viên **theo quota/ thuận tiện**.

So sánh các cấp độ lấy mẫu

	Bước 1: Msci sử dụng máy tính để chọn		Bước 2: FW thực hành	Bước 3: FW thực hành
	GĐ 1: Chọn phường/xã	GĐ 2: Chọn Block	GĐ 3: Chọn hộ gia đình	GĐ 4: Chọn đáp viên
Level 1	Ngẫu nhiên	Ngẫu nhiên	Ngẫu nhiên	Ngẫu nhiên
Level 2	Ngẫu nhiên	Ngẫu nhiên	Ngẫu nhiên	Quota/ Thuận tiện
Level 3.1	Ngẫu nhiên	Thuận tiện	Thuận tiện	Quota/ Thuận tiện
Level 3.2	Thuận tiện	Thuận tiện	Thuận tiện	Quota/ Thuận tiện

So sánh các cấp độ lấy mẫu

- Level 1: Ngẫu nhiên hoàn toàn, mẫu có tính đại diện cao nhất nhưng cũng khó thực hiện và chi phí thực hiện cũng cao nhất. Cấp độ lấy mẫu này phù hợp cho những nhóm ngành tiêu dùng cá nhân phổ biến-những mặt hàng mà người tiêu dùng tự có khả năng tự quyết định và tự mua để sử dụng cho bản thân. Ví dụ như: các sản phẩm chăm sóc cá nhân (kem, phấn, sữa rửa mặt), dịch vụ điện thoại di động, thuốc lá, các loại đồ uống có cồn/gas ...
- Level 2: chỉ khác Level 1 ở giai đoạn 4. Level 2 không chọn đáp viên 1 cách ngẫu nhiên mà theo quota/thuận tiện và không áp dụng kish grid hoặc call back 3 lần. Level 2 thực sự phù hợp cho những sản phẩm tiêu dùng mang cho cả hộ gia đình-những sản phẩm mà mỗi gia đình có một người quyết định mua sản phẩm để tiêu dùng cho cả hộ gia đình (ví dụ như: các mặt hàng lương thực/thực phẩm, bột giặt, sữa (cho em bé/bà mẹ), nước chấm, gia vị, dịch vụ tài chính gia đình... Level 2 cũng là một giải pháp tiết kiệm chi phí cho những dự án điều tra về các mặt hàng tiêu dùng cá nhân (như level 1) nhưng trong trường hợp này, phải đảm bảo chúng ta đã có trước những thông số của thị trường để tính khung quota cho mẫu (theo độ tuổi, giới tính, SEC...) có như vậy mẫu mới mang tính đại diện tốt nhất với chi phí lại phù hợp nhất.

So sánh các cấp độ lấy mẫu

- Level 3.1: Khác hẳn với level 1 và 2, Level 3.1 chỉ ngẫu nhiên ở bước chọn đơn vị chọn mẫu cơ bản (phường) còn lại việc chọn block, hộ gia đình, đáp viên là thuận tiện. Level 3 phù hợp với những dự án mà xác suất phỏng vấn thành công thấp ($ECR < 20\%$) Ví dụ như những dự án product test, CLT, những dự án có điều kiện đáp viên quá khắt khe (chỉ lấy class A, chỉ lấy 1 khoảng tuổi hẹp nào đó hay chỉ chọn PV những bà mẹ mang thai lần đầu...)
- Level 3.2: Thuận tiện hoàn toàn, phỏng vấn viên sẽ tự đi tìm đáp viên (theo sự giới thiệu của người khác hoặc theo mối quan hệ của đáp viên). Level 3.2 thường được áp dụng cho những dự án “siêu khó” với tỉ lệ $ECR < 5\%$. Ví dụ như những dự án về nhóm ngành xe hơi, kính sát trùng, du lịch nước ngoài
- * Trong thực tế, có một số dự án mặc dù ECR thấp nhưng khách hàng lại muốn đi random để tính các thông số của thị trường. Chúng ta có thể đề xuất giải pháp kết hợp. Một nửa giai đoạn đi theo cấp độ này và một nửa giai đoạn thi theo cấp độ khác (xem thêm booster sampling-Mục 3). Trường hợp này CSG nên xin lời khuyên từ những người có kinh nghiệm hơn.

**Thực tế của việc lấy
mẫu.**

Chương trình lấy mẫu tự động

- Msci sẽ sử dụng một chương trình lấy mẫu tự động để lên một kế hoạch lấy mẫu và gửi cho bộ phận thu thập dữ liệu hiện trường. Chương trình này hoạt động dựa trên nguyên tắc PPS và các hàm ngẫu nhiên.
- Khung mẫu của 5 thành phố chính không chỉ được phân nhóm ra thành các khu vực địa lý (phường/quận) khác nhau mà còn được phân nhóm theo tầng lớp kinh tế của dân cư (SEC), Điều này sẽ giúp cho mẫu mang tính đại diện cao cho thị trường.
- Khung mẫu của các tỉnh thành còn lại chỉ được phân nhóm theo khu vực địa lý (phường/xã/thị trấn).

3. Phương pháp khác

Phương pháp chọn mẫu ngẫu nhiên theo danh sách.

- Khung mẫu được sử dụng cho phương pháp lấy mẫu đó là danh sách tổng thể. Danh sách tổng thể này cần cần phải được cập nhật thường xuyên.
- Phương pháp chọn mẫu hệ thống sẽ được áp dụng để chọn ra một tập hợp các đối tượng mẫu từ danh sách tổng thể này.
- Quota sẽ được áp dụng để đảm bảo cho việc chọn mẫu này mang tính đại diện cao nhất.

Chọn mẫu boosters

- Mẫu Booster được chọn trong trường hợp đối tượng của nghiên cứu là hiếm và khó gặp. Mẫu boosters được thực thi sau khi giai đoạn đi lấy mẫu ngẫu nhiên (mức 1, mức 2) hoàn tất mà vẫn không lấy đủ được số mẫu tối thiểu.
- Mặc dù các đối tượng của chọn mẫu booster không được chọn từ một phương pháp lấy mẫu ngẫu nhiên, Nielsen sẽ vẫn đảm bảo rằng các đối tượng này được lựa chọn trên cơ sở bao phủ về mặt địa lý cũng như quota (nếu có) để đảm bảo tính đại diện cao nhất của mẫu được chọn ra trong nghiên cứu.

Cám ơn!

nielsen
• • • • • • • • •