

GIÁO TRÌNH

CUSTOMER INSIGHTS & MARKETING STRATEGY

THẤU HIỂU KHÁCH HÀNG VÀ TÁI ĐỊNH VỊ THƯƠNG HIỆU



BAO GỒM CÁC CASE STUDY THỰC TẾ:
UNILEVER, WESTERN SYDNEY
UNIVERSITY (UNLIMITED)



NHÓM DỊCH
TRƯỜNG ĐẠI HỌC PXU



Trợ dịch :
Ms Nguyen Thi Tuyen Nhung

Bản quyền tác phẩm Customer Insights © 2023 của Aila Khan, Munir Hossain và Sabreena Amin được cấp phép theo Giấy phép Quốc tế Creative Commons Attribution-NonCommercial 4.0, trừ khi có ghi chú khác.

Cuốn sách này được xuất bản thông qua Tập thể Tài nguyên Giáo dục Mở của Hội đồng Thủ thư Đại học Australasia (Council of Australasian University Librarians Open Educational Resources Collective). Phiên bản trực tuyến có sẵn tại <https://oercollective.caul.edu.au/customer-insights>

Bìa sách được thiết kế bởi Ms Nguyen Thi Tuyet Nhung .

Tuyên bố miễn trừ trách nhiệm Xin lưu ý rằng các biểu trưng (logos) và dấu hiệu nhận diện thương hiệu (branding) của các doanh nghiệp được loại trừ cụ thể khỏi Giấy phép Creative Commons Attribution Noncommercial của tác phẩm này, và không được phép sao chép dưới bất kỳ hình thức nào nếu không có sự cho phép rõ ràng bằng văn bản từ những người nắm giữ bản quyền.

Bản quyền Tác phẩm này, Customer Insights (tái bản lần 2) là một tác phẩm phái sinh (derivative) từ cuốn Customer Insights của tác giả Aila Khan, được sử dụng theo giấy phép CC BY-NC-SA 4.0.

Tác phẩm Customer Insights (tái bản lần 2) của Aila Khan, Sabreena Amin và Muni Hossain được cấp phép theo giấy phép CC BY-NC 4.0 bởi Western Sydney University.

Định dạng trích dẫn đề xuất: Khan, A., Amin, S. & Hossain, M. (2023). Customer insights (2nd ed.). Western Sydney University.

<https://oercollective.caul.edu.au/customer-insights>

Định dạng ghi nhận tác giả đề xuất (Recommended attribution): Tác phẩm Customer Insights (2nd ed.) của Aila Khan, Sabreena Amin và Muni Hossain được cấp phép theo giấy phép CC BY-NC 4.0 bởi Western Sydney University.

LỜI GIỚI THIỆU TỪ NHÓM DỊCH GIẢ

Trong kỷ nguyên dữ liệu và quản trị hiện đại, việc am hiểu sâu sắc các phương pháp nghiên cứu (Research Methods) và kỹ thuật thiết kế khảo sát (Questionnaire Design) là nền tảng thiết yếu đối với bất kỳ nhà nghiên cứu, sinh viên hay chuyên gia thực hành nào. Nhằm góp phần thu hẹp khoảng cách tiếp cận tri thức chất lượng cao, nhóm dịch giả hân hạnh mang đến cho quý độc giả bản Việt hóa của cuốn giáo trình chuyên sâu này.

Quá trình chuyển ngữ cuốn sách được chúng tôi thực hiện với sự cẩn trọng tối đa và kỷ luật học thuật khắt khe. Điểm nổi bật trong phương pháp dịch thuật của dự án này là nguyên tắc **bảo toàn thuật ngữ chuyên ngành**. Nhận thức được rằng ngôn ngữ khoa học đòi hỏi tính chính xác tuyệt đối và nhu cầu tra cứu tài liệu quốc tế của người học là rất lớn, chúng tôi đã chủ động **giữ lại và đặt trong ngoặc đơn các thuật ngữ chuyên ngành bằng tiếng Anh** bên cạnh phần nghĩa tiếng Việt (ví dụ: *Acquiescence*, *Non-sampling error*, *Research panels*, *Standard Deviation*...). Cách tiếp cận này không chỉ giúp bảo toàn trọn vẹn ngữ nghĩa cốt lõi của tác giả gốc mà còn hỗ trợ người đọc dễ dàng đối chiếu, làm quen với hệ thống từ vựng tiêu chuẩn quốc tế, từ đó tạo đà thuận lợi cho việc hội nhập học thuật toàn cầu.

Bên cạnh cam kết về chất lượng chuyên môn, việc tôn trọng tuyệt đối quyền sở hữu trí tuệ và tính toàn vẹn của tác phẩm gốc là kim chỉ nam của nhóm. Bản dịch này tuân thủ nghiêm ngặt các quy định về giấy phép **Creative Commons (CC BY-NC - Ghi công & Phi thương mại)** được áp dụng trong tài liệu nguyên bản. Cụ thể:

- Mọi cấu trúc lý thuyết, công thức tính toán thống kê, và thiết kế bảng biểu đều được giữ nguyên 100% nội dung.
- Toàn bộ hệ thống nguồn trích dẫn, danh mục tài liệu tham khảo và ghi danh phương tiện truyền thông (Media Attributions) đều được bảo lưu nguyên trạng đường dẫn và thông tin xác thực.

- Tác phẩm phái sinh này được thực hiện hoàn toàn hướng tới mục tiêu phục vụ giáo dục, nghiên cứu khoa học và cộng đồng, không vì mục đích thương mại.

Chúng tôi hi vọng rằng, với sự kết hợp giữa văn phong dịch thuật học thuật chuẩn mực, cấu trúc giữ nguyên bản và hệ thống thuật ngữ song ngữ rõ ràng, cuốn giáo trình này sẽ trở thành một người bạn đồng hành đắc lực, hỗ trợ sinh viên, quý độc giả trên con đường chinh phục các dự án nghiên cứu và khảo sát thực chứng.

Dù đã nỗ lực hết sức, quá trình chuyển ngữ những khái niệm quản trị và thống kê phức tạp khó tránh khỏi những khiếm khuyết nhất định. Nhóm dịch giả rất mong nhận được sự lượng thứ và những ý kiến đóng góp quý báu từ quý độc giả và các chuyên gia trong ngành để bản dịch được hoàn thiện hơn trong các lần tái bản sau.

Trân trọng,

Việt Nam, tháng 05 năm 2026

Nhóm dịch Trường Đại học PXU.

Trợ dịch - Ms. Nguyễn Thị Tuyết Nhung

Nội dung	Trang
Lời giới thiệu từ nhóm dịch giả	3
Lời tri ân vùng đất tổ tiên (Acknowledgement of country)	16
Lời cảm ơn	19
Sự khác biệt giữa nghiên cứu tiếp thị và insight khách hàng	19
Đạo đức nghiên cứu	30
Nghiên cứu thứ cấp	40
Sử dụng dữ liệu điều tra dân số (Siddharth Jain)	47
Nghiên cứu sơ cấp	45

Nghiên cứu định tính và định lượng	66
Các loại thiết kế nghiên cứu	68
Thảo luận nhóm tập trung (Focus groups)	59
Nghiên cứu quan sát	59
Các thang đo hoặc các loại biến số	98
Thiết kế bảng câu hỏi	103
Các phương pháp chọn mẫu	127
Các sai số trong nghiên cứu	147
Nghiên cứu dạng Panel (Research panels)	159

Các phương pháp phân phối khảo sát	169
Thống kê mô tả	173
Sự liên đới giữa các biến số	96
Sự khác biệt giữa các nhóm người trả lời	106
Phân tích cảm xúc (Sentiment analysis)	212
Trí tuệ nhân tạo và thông tin	220
Phân tích mạng xã hội	28
Nghiên cứu các cộng đồng bản địa	241
Truyền đạt insight	276

Đồ họa thông tin (Infographics)	258
Quy trình: từ việc tạo ra đến sử dụng insight khách hàng	267
Nghiên cứu tình huống (Case Study): Sử dụng insight khách hàng để tái định vị Western Sydney University (Jacki Montgomery)	281

Cuốn giáo trình truy cập mở (open-access) này được phát hành lần đầu vào học kỳ Mùa thu năm 2022 tại Western Sydney University. Kể từ đó, đây đã trở thành tài liệu tham khảo được khuyến nghị cho sinh viên khối ngành kinh doanh tại WSU Online, Cơ sở WSU Sydney City và Đại học Kinh tế Thành phố Hồ Chí Minh, Việt Nam. Chúng tôi hoan nghênh mọi ý kiến đóng góp để cải thiện giáo trình tại địa chỉ: a.khan@westernsydney.edu.au.

Bắt nhịp với những bước tiến trong ngành, cùng với các phản hồi nhận được từ sinh viên và giới học giả, lần tái bản thứ hai này có một số điểm điều chỉnh như sau:

- Hiện đây là một cuốn sách có nhiều đồng tác giả, do toàn bộ đội ngũ giảng viên đã cùng nhau chung tay biên soạn cho lần tái bản thứ hai này.
- Chúng tôi đã áp dụng triết lý "sinh viên là đối tác" (students as partners). Các sinh viên đối tác của chúng tôi đã trực tiếp đóng góp vào nội dung sách. Họ cũng cung cấp những phản hồi giá trị, giúp ích rất nhiều cho chúng tôi trong quá trình chỉnh sửa.
- Ấn bản này bao gồm các nội dung "mới" mà nhóm tác giả cảm thấy mang tính thiết yếu và mọi chuyên gia đều cần phải làm quen. Các chương được bổ sung bao gồm "Sử dụng dữ liệu điều tra dân số" và "Phương pháp nghiên cứu Thổ dân bản địa", vốn theo truyền thống không phải là một phần của nghiên cứu tiếp thị.
- Với màn ra mắt mang tính đột phá vào ngày 30 tháng 11 năm 2022 bởi OpenAI – và sự gia tăng chóng mặt lên tới 100 triệu người dùng tính đến tháng 1 năm 2023 – ChatGPT thực sự là một cuộc cách mạng trong không gian AI. Một cuốn giáo trình tiếp thị sẽ không thể hoàn thiện nếu bỏ qua việc ghi nhận sự xuất hiện của chatbot mới mẻ này! Do đó, chương "Trí tuệ nhân tạo và thông tin" của chúng tôi đã được cập nhật và sửa đổi một cách xứng đáng.



Tiến sĩ Aila Khan

Tiến sĩ Aila là Giảng viên Cao cấp tại Trường Kinh doanh thuộc Western Sydney University (WSU). Một số dự án nghiên cứu thú vị nhất của cô – bao gồm robot xã hội (social robots), sức khỏe tinh thần của cá nhân và hoạt động thiện nguyện – đều được khởi xướng bởi chính các sinh viên trong khuôn khổ luận văn tốt nghiệp của họ.



Tiến sĩ Mohammad Munir Hossain

Tiến sĩ Munir có hơn 23 năm kinh nghiệm trong môi trường học thuật. Các lĩnh vực nghiên cứu mà ông quan tâm bao gồm hành vi người tiêu dùng và các nghiên cứu xuyên văn hóa (cross-cultural studies).



Tiến sĩ Sabreena Zoha Amin

Tiến sĩ Sabreena Zoha Amin là một nhà nghiên cứu và học giả tại Trường Kinh doanh thuộc Western Sydney University (WSU), Úc. Lĩnh vực nghiên cứu mà Tiến sĩ Sabreena quan tâm bao trùm các mảng hành vi người tiêu dùng (consumer behaviour), tiếp thị dịch vụ (service marketing), tiếp thị kỹ thuật số (digital marketing) và tiếp thị quốc tế (international marketing).



Tiến sĩ Jacki Montgomery – Tác

giả chương Nghiên cứu tình huống (Case Study): Sử dụng insight khách hàng để tái định vị Western Sydney University

Tiến sĩ Jacki Montgomery là Giảng viên chuyên ngành Quảng cáo tại Western Sydney University. Với nền tảng học thuật về tiếp thị và quảng cáo, cô đặc biệt quan tâm đến sự linh hoạt của ngôn ngữ (dexterity of language) và vai trò sâu sắc mà insight có thể mang lại trong các chiến dịch quảng cáo. Cô thường xuyên diễn

thuyết tại các hội nghị học thuật và đã từng trình bày tại nhiều sự kiện của ngành công nghiệp sáng tạo, bao gồm Semi Permanent và Mumbrella. Tiến sĩ Jacki được công nhận ở cả cấp độ quốc gia và quốc tế như một nhà truyền đạt và giảng viên xuất sắc, và gần đây cô đã vinh dự được trao giải thưởng Giảng viên của Năm (Teacher of the Year) của Western Sydney University.



Tiến sĩ Siddharth Jain

— Tác giả chương Sử dụng dữ liệu điều tra dân số.

Tiến sĩ Siddharth Jain mới nhận bằng Tiến sĩ (Ph.D.) tại WSU, đồng thời sở hữu bằng Thạc sĩ về Kế toán Chuyên nghiệp và Quản trị Kinh doanh. Thầy cũng

là Hội viên liên kết (Associate member) của Hiệp hội Kế toán Công chứng (CPA). Với hơn ba năm kinh nghiệm trong cả công tác giảng dạy và nghiên cứu, Tiến sĩ Siddharth đã rèn luyện các kỹ năng của mình tại Trường Kinh doanh thuộc WSU, tích lũy chuyên môn sâu sắc về kinh tế lượng (econometrics) và phân tích dữ liệu, điều này được minh chứng rõ nét qua hàng loạt công trình nghiên cứu đã được công bố của ông.

Cuốn sách này được ra mắt nhờ vào những nỗ lực đóng góp của các học giả trẻ xuất sắc đến từ chương trình Sinh viên là Đối tác (Students as Partners - SAP) của WSU.

Sanaa Harrass



Sanaa Harrass đã hoàn thành chương trình Cử nhân Khoa học Y khoa tại Western Sydney University. Trước đó, cô cũng đã theo học chương trình cử nhân chuyên ngành Kinh tế và Khoa học Chính trị tại Cairo University. Cô hiện có hai người con gái, 12 và 15 tuổi. Sanaa rất yêu thích việc đọc sách và có niềm đam mê mãnh liệt với các phương pháp chữa bệnh bằng thảo dược.

Amanda Cooper



Amanda Cooper là nghiên cứu sinh Tiến sĩ (Ph.D. candidate) đang công tác tại Trung tâm Nghiên cứu Viết và Xã hội (Writing and Society Research Centre) thuộc Western Sydney University. Sau khi hoàn thành chương trình Cử nhân (Bachelor of Arts) vào năm 2015, cô đã vinh dự được Trường Khoa học Nhân văn và Nghệ thuật Truyền thông (School of Humanities and Communication

Arts) trao tặng Huy chương của Trường khoa dành cho Thành tích Học tập Xuất sắc (Dean's Medal for Academic Excellence). Amanda cũng từng tham gia công tác tại trường với tư cách là trợ giảng (tutor), cố vấn học tập (mentor) và sinh viên đối tác tại trung tâm 21C Partnership Xchange Hub.

Phiên bản trực tuyến của tài liệu này được thiết kế để đáp ứng bộ Tiêu chuẩn Tiếp cận Nội dung Web 2.0 ([Web Content Accessibility Guidelines 2.0](#)).

Hãy cho chúng tôi biết nếu bạn gặp khó khăn trong việc truy cập cuốn sách này

Chúng tôi luôn tìm cách để làm cho các tài nguyên của mình trở nên dễ tiếp cận hơn. Nếu bạn gặp trục trặc khi truy cập tài liệu này, vui lòng [liên hệ](#) và cho chúng tôi biết để chúng tôi có thể khắc phục sự cố.

Vui lòng cung cấp các thông tin sau:

- Vị trí xảy ra sự cố bằng cách cung cấp tiêu đề sách, địa chỉ trang web hoặc mô tả trang.
- Mô tả chi tiết về sự cố.
- Máy tính, phần mềm, trình duyệt và bất kỳ công nghệ hỗ trợ (assistive technology) nào bạn đang sử dụng để giúp chúng tôi chẩn đoán và giải quyết vấn đề của bạn.
 - Ví dụ: Windows 10, Google Chrome (Phiên bản 65.0.3325.181), phần mềm đọc màn hình NVDA (NVDA screenreader).

Ghi nhận bản quyền truyền thông (Media Attributions)

- **Ảnh Aila Khan (đã cắt)** © Sally Tsoutas - Nhiếp ảnh gia WSU, được cấp phép theo giấy phép **CC BY-NC-ND (Attribution NonCommercial NoDerivatives)**.
- **Ảnh Tiến sĩ Sabreena Amin** © Sabreena Amin, được cấp phép theo giấy phép **CC BY-NC-SA (Attribution NonCommercial ShareAlike)**.
- **Ảnh Jacki Montgomery** © Jacki Montgomery, được cấp phép theo giấy phép **CC BY-NC-ND (Attribution NonCommercial NoDerivatives)**.
- **Ảnh Siddharth Jain** © Siddharth Jain, được cấp phép theo giấy phép **CC BY-NC-ND (Attribution NonCommercial NoDerivatives)**.
- **Ảnh Đối tác sinh viên Sanaa Harrass** © Sanaa Harrass, được cấp phép theo giấy phép **CC BY-NC-SA (Attribution NonCommercial ShareAlike)**.

LỜI TRI ÂN VÙNG ĐẤT TỔ TIÊN

Lời tri ân dành cho những chủ nhân truyền thống của vùng đất thuộc Western Sydney University

Western Sydney University ghi nhận và tri ân những người bảo vệ (custodians) vùng đất nơi chúng ta gặp gỡ, làm việc, học tập và giao lưu. Chúng tôi bày tỏ sự tôn kính đối với người dân thuộc các dân tộc Darug, Tharawal, Eora và Wiradjuri, nơi các cơ sở của chúng tôi tọa lạc. Chúng tôi ghi nhận rằng các hoạt động giảng dạy, học tập và nghiên cứu được thực hiện trên khắp các cơ sở của trường là sự tiếp nối của những truyền thống giảng dạy, học tập và nghiên cứu đã diễn ra trên những vùng đất này trong hàng chục ngàn năm qua. Chúng tôi ghi nhận và bày tỏ lòng thành kính đến các bậc Trưởng lão (Elders) trong quá khứ, hiện tại và thế hệ tương lai.



Nguồn: Tác phẩm ‘Matta: Meeting place’ của Jason và Trevor Dalmarri Được đồng sáng tác bởi đội ngũ thuộc Trường Kinh doanh WSU (WSU School of Business) tại Parramatta, tác phẩm nghệ thuật này đại diện cho những vùng đất và cộng đồng nơi

cơ sở Parramatta City được xây dựng. Từ Parramatta bắt nguồn từ từ ngữ của Thổ dân là 'Baramada' hoặc 'Burramatta'. Người Burramatta (trong đó "Burra" nghĩa là địa điểm và "Matta" nghĩa là những con lươn) thuộc dân tộc Dhurug, những người đã sinh sống tại khu vực trù phú thức ăn này từ trước thời kỳ xâm lược của người da trắng.

Câu chuyện đằng sau tác phẩm nghệ thuật này kể về chính vùng đất mà bạn đang làm việc ngày hôm nay, qua đó thể hiện sự đa dạng của khu vực và bày tỏ lòng tôn kính đối với quá khứ.

Những con "Matta" (lươn) đã bơi dọc theo những con sông này, sinh sôi nảy nở mạnh mẽ và dò dẫm dọc theo hai bờ sông Burramatta. Những dải nước màu xanh lá và xanh lam uốn lượn xuyên suốt tác phẩm nghệ thuật, tựa như cách chúng đã từng di chuyển vô cùng duyên dáng dọc theo các con lạch. Tác phẩm khắc họa những bọt nước, dòng chảy và vô số cá tôm bơi qua khu vực mà ngày nay được gọi là Parramatta. Đường bờ biển phác họa ranh giới của đất liền khi các vùng nước hướng ra Sydney Harbour (Cảng Sydney) và đại dương.

Những mảng màu rực rỡ tượng trưng cho tất cả các thị trấn bao quanh Burramatta. Bên trong những thị trấn đó là các vòng tròn thuộc về những gia đình đa văn hóa hiện đang sinh sống trên vùng đất của người Dhurug. Những đường nét đứt và dấu vết vón vón là lối mòn đi bộ của người Dhurug, nay đã biến thành các con đường, tuyến giao thông và tuyến đường sắt.

Những vùng đất này bao trọn sự phát triển của thị trấn nay được gọi là Parramatta nhằm tôn vinh những người Úc đầu tiên (First Australians). Western Sydney University ghi nhận dân tộc Dhurug là những chủ nhân truyền thống của những vùng đất mà chúng ta đang làm việc và sinh sống ngày nay.

Ghi nhận tác quyền hình ảnh (Media Attributions)

Tác phẩm [*Matta: meeting place*](#) © Jason và [Trevor Dalmarri](#) được cấp phép theo giấy phép Đã đăng ký bản quyền ([All Rights Reserved](#)).

LỜI CẢM ƠN

Một số cá nhân tuyệt vời tại Western Sydney Uni (WSU) đã hỗ trợ giám tuyển và biên soạn cuốn sách này. Chúng tôi vô cùng biết ơn đội ngũ nhân viên Thư viện WSU, đặc biệt là Paul Jewell, Karen Sheehy, Lucy Walton, Lauren Tupper và Jeffrey Har, những cộng sự tuyệt vời của chúng tôi.

Frank Hill, Cán bộ Bản quyền của Trường, đã hỗ trợ chúng tôi vô cùng to lớn. Rất nhiều hình ảnh trong cuốn sách này được thực hiện bởi nhiếp ảnh gia chính thức của trường – xin gửi lời cảm ơn đến Sally Tsoutas!

Chúng tôi cũng xin ghi nhận và cảm ơn Tập thể Tài nguyên Giáo dục Mở (Open Educational Resources Collective), một sáng kiến của Hội đồng Thủ thư Đại học Úc (Council of Australian University Librarians - CAUL), vì khoản tài trợ hào phóng giúp cải thiện bản in lần thứ hai của cuốn Customer Insights. CAUL dẫn dắt các sáng kiến trên toàn ngành tập trung vào việc phát triển năng lực và khả năng xuất bản các giáo trình mở cũng như các tài nguyên giáo dục mở khác tại Úc và New Zealand.

Cuối cùng, xin gửi lời tri ân đến các đồng nghiệp học thuật đã đề nghị đánh giá cuốn sách và cung cấp phản hồi trong một thời gian ngắn. Các bạn thật tuyệt vời!

Danh sách các chuyên gia (theo thứ tự bảng chữ cái):

- Tiến sĩ Abdul Babar, Giảng viên thỉnh giảng (Sessional Academic)
- Bruce Cameron, Giảng viên thỉnh giảng (Sessional Academic)
- Fareeza Khan, Giảng viên thỉnh giảng (Sessional Academic), Macquarie University
- Saman Sadiq, Giảng viên thỉnh giảng (Sessional Academic)
- Sana Yousaf, Trợ giảng (Teaching Assistant)

1. SỰ KHÁC BIỆT GIỮA NGHIÊN CỨU TIẾP THỊ VÀ INSIGHT KHÁCH HÀNG

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem chúng trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=31#video-31-1>

MỤC TIÊU HỌC TẬP :

Sau khi học xong chương này, sinh viên phải có khả năng:

- nhận diện sự trỗi dậy của khái niệm insight khách hàng qua các năm
- hiểu rõ những khác biệt then chốt giữa nghiên cứu tiếp thị và insight khách hàng
- xác định các dạng dữ liệu hiện có để phục vụ việc phân tích.

Sự trỗi dậy của insight khách hàng

Theo truyền thống, các nhà tiếp thị thường sử dụng các công cụ như khảo sát, phỏng vấn và thảo luận nhóm tập trung (focus groups) để thực hiện nghiên cứu tiếp thị. Mặc dù các công cụ này vẫn đóng vai trò rất quan trọng, nhưng chúng tồn tại một số hạn chế vốn có. Trước hết, để thu được dữ liệu chất lượng tốt từ các công cụ này, một giả định then chốt là người trả lời sẽ phản hồi chính xác các câu hỏi nghiên cứu. Trong nhiều trường hợp, những người tham gia được kỳ vọng sẽ nhớ lại một trải nghiệm trước đó và báo cáo lại. Ở các tình huống khác, khách hàng được đặt vào các kịch bản giả định và được yêu cầu dự đoán khả năng họ sẽ thực hiện một hành vi nhất định (ví dụ: mua hàng). Do đó, chia khóa để có được dữ liệu chất lượng cao phụ thuộc vào mức độ tương tác, khả năng và sự sẵn lòng cung cấp toàn bộ thông tin cần thiết của người trả lời.

Một hạn chế khác của các phương pháp nghiên cứu tiếp thị truyền thống là sự sai lệch (bias) có thể len lỏi vào thông qua cách dùng từ hoặc cách trình bày câu hỏi. Người ta đều biết rằng câu trả lời cho một câu hỏi có thể thay đổi tùy thuộc vào việc ai là người hỏi, cách hỏi như thế nào, và thậm chí có thể là thời điểm đặt câu hỏi. Hơn nữa, các câu hỏi khảo sát thường được định sẵn và cung cấp rất ít sự linh hoạt để khai thác những ý tưởng mới. Người trả lời thường bị giới hạn trong cách họ có thể phản hồi. Tương tự như vậy, việc tạo ra dữ liệu tốt trong một cuộc phỏng vấn thường đòi hỏi một mối quan hệ tốt đẹp (rapport) giữa người phỏng vấn và người tham gia. Tuy nhiên, không phải tất cả các nhà nghiên cứu đều có kỹ năng tốt trong lĩnh vực này.

Thêm vào đó, các phương pháp thu thập dữ liệu truyền thống – chẳng hạn như khảo sát và thảo luận nhóm tập trung – đòi hỏi rất nhiều thời gian cho việc thiết kế khảo sát, thu thập dữ liệu và phân tích. Các nhà tiếp thị lại thường cần thông tin ngay lập tức, điều mà các công cụ này khó có thể đáp ứng kịp thời. Cũng có khả năng xảy ra "sai lệch do mong muốn xã hội" (social desirability bias). Người trả lời có thể mang một khuynh hướng vô thức muốn làm hài lòng nhà nghiên cứu, đặc biệt là khi nhà

ngiên cứu có mặt trực tiếp. Sự sai lệch như vậy sẽ ảnh hưởng trực tiếp đến câu trả lời của người tham gia.

Cuối cùng, các phương pháp nghiên cứu tiếp thị không hề rẻ. Cần có nhiều nguồn lực – cả về tiền bạc lẫn chuyên môn – để thiết kế một bảng khảo sát, một cuộc phỏng vấn hay một cuộc thảo luận nhóm tập trung (focus group). Nhiều doanh nghiệp vừa và nhỏ thường không có khả năng ứng dụng hiệu quả các công cụ truyền thống để thu thập thông tin.

1. Macdonald, EK, Wilson, HN & Konus, U 2012, *Better customer insight—in real time*, Harvard Business Review, <https://hbr.org/2012/09/better-customer-insight-in-real-time>

Thế giới hiện đại đang tận dụng tối đa các công nghệ kỹ thuật số. Với việc sử dụng ví thông minh, điện thoại thông minh, thiết bị gia dụng thông minh, dịch vụ phát trực tuyến và các nền tảng mạng xã hội, một lượng thông tin tức thời đang được tạo ra liên tục. Điều này hình thành nên một kho lưu trữ ngày càng lớn của "Dữ liệu lớn" (Big Data) – mà về mặt logic – cần được khai thác để đạt được những insight về hành vi người tiêu dùng. Mặc dù các phương pháp nghiên cứu tiếp thị truyền thống vẫn hữu ích, chúng cần được bổ sung bằng các dạng dữ liệu khác đang được tạo ra và lưu trữ thường xuyên.

Rất nhiều tổ chức đã xây dựng "Đội ngũ Insight" (Insights Teams) bao gồm các nhân sự đến từ nhiều nền tảng chuyên môn khác nhau. Trái ngược với một bộ phận nghiên cứu thị trường, nhiệm vụ của Đội ngũ Insight là có khả năng "tổng hợp" (synthesise) lượng dữ liệu khổng lồ đang được tạo ra. Các thành viên trong đội ngũ đóng vai trò là những người tập hợp (aggregators), người diễn giải (interpreters) và người phổ biến thông tin (disseminators). Việc chỉ đơn thuần trình bày "các phát hiện" (findings) cho những câu hỏi khảo sát là không đủ. Với tư cách là một chuyên gia Insight, vai trò then chốt là có khả năng tập hợp các mẫu thông tin khác nhau, diễn giải dữ liệu thông qua việc hợp tác với đồng nghiệp từ các phòng ban khác, và có khả năng đề xuất các điểm hành động cốt lõi (action points) cho sự tăng trưởng trong tương lai.



Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=31#h5p-1>

Ảnh nền hoa © 2022 Western Sydney University do Sally Tsoutas, [Nhiếp ảnh gia của Western Sydney University thực hiện](#), được cấp phép theo giấy phép [Attribution-NonCommercial-NoDerivatives 4.0 International \(Ghi nhận tác giả - Phi thương mại - Không phái sinh 4.0 Quốc tế\)](#).

Nghiên cứu Tiếp thị (Marketing Research)	Insight Khách hàng (Customer Insights)
Thiên về trọng tâm học thuật	Là một thuật ngữ trong thực tiễn ngành
Định nghĩa: Là hoạt động thu thập và phân tích dữ liệu để trả lời các câu hỏi kinh doanh cụ thể.	Insight người tiêu dùng (Consumer Insights) là hoạt động sử dụng dữ liệu sẵn có để đạt được sự thấu hiểu sâu sắc hơn về khách hàng, từ đó giúp đưa ra các quyết định kinh doanh hiệu quả.
Các công cụ được sử dụng: khảo sát, phỏng vấn, thảo luận nhóm tập trung, v.v.	Dữ liệu sẵn có: dữ liệu nội bộ, mạng xã hội, dữ liệu trực tuyến, v.v.
Trọng tâm đặt vào các phát hiện (findings)	Trọng tâm đặt vào các phát hiện và lý do đằng sau chúng; đưa ra các kiến nghị
Tập trung vào câu hỏi ban đầu/mục tiêu nghiên cứu	Tập trung vào việc sử dụng thông tin để hiểu thái độ và hành vi của khách hàng

Phân tích từng luồng dữ liệu một cách riêng lẻ	Tích hợp nhiều luồng dữ liệu lại với nhau
Thông tin/dữ liệu thực tế được tóm tắt và làm cơ sở cho một kiến nghị	Dữ liệu được chuyển tải dưới dạng một câu chuyện hoặc lối tường thuật (narrative)
Trọng tâm vào các phương pháp định lượng, định tính, hỗn hợp	Mang trọng tâm đa ngành (multi-disciplinary)
Bàn giao kết quả dựa trên các mục tiêu nghiên cứu. Được sử dụng để trả lời các câu hỏi cụ thể.	Đưa ra các kiến nghị mang tính đa ngành
Bàn giao dữ liệu cho bộ phận tiếp thị nhưng không tham gia vào các cuộc họp với khách hàng	Tiếp thị là một đối tác kinh doanh. Đội ngũ Insight tham gia trực tiếp vào các cuộc họp nhân sự của khách hàng
Phổ biến hơn ở các nền kinh tế phụ thuộc nhiều vào sản xuất (manufacturing) như Trung Quốc,	Được sử dụng phổ biến hơn ở các nền kinh tế dựa trên dịch vụ (service-based

Đức, Nhật Bản, Đài Loan, Indonesia, Ba Lan và Hàn Quốc	economies) như Mỹ, Brazil, Bermuda, Anh, Hy Lạp, Úc, Singapore
Xây dựng cơ sở dữ liệu nghiên cứu	Cung cấp quyền truy cập vào các bảng điều khiển (dashboards), v.v.

Bảng: Sự khác biệt giữa Nghiên cứu Tiếp thị và Insight Khách hàng là gì?

Sử dụng dữ liệu hiện có để đạt được insight

Khái niệm đằng sau 'insight khách hàng' không phải là thu thập thông tin thông qua các nguồn dữ liệu truyền thống, mà là khai thác các nguồn hiện có đang liên tục tạo ra dữ liệu. Một số nguồn có thể cung cấp các insight hữu ích được thảo luận dưới đây:

2. Purcell, M 2021, *What's the difference between consumer insights and marketing research?*, Greenbook, truy cập ngày 28 tháng 2 năm 2022, <https://www.greenbook.org/mr/market-research-news/whats-the-difference-between-consumer-insights-and-market-research>

Dữ liệu bán hàng nội bộ (Internal sales data)

Các tổ chức sở hữu một nguồn tài nguyên dữ liệu nội bộ dồi dào được thu thập dưới nhiều hình thức khác nhau. Mỗi khi một giao dịch được thực hiện, một bản ghi (record) sẽ được tạo ra. Việc phân tích dữ liệu bán hàng đó, doanh thu trên mỗi khách hàng, thời gian và địa điểm bán hàng có thể tạo ra thông tin hữu ích cho quá trình ra quyết định.

Truy vấn và khiếu nại của khách hàng (Customer queries and complaints)

Mỗi khi khách hàng liên hệ để đưa ra một câu hỏi truy vấn (hoặc một lời khiếu nại), một mảnh dữ liệu lại được tạo ra. Nhiều công ty đã nhận ra rằng bằng cách phân tích các truy vấn và khiếu nại của khách hàng, một lượng lớn thông tin có thể được khai thác. Quá trình này bao gồm việc phân tích các bản ghi nội dung trò chuyện (chat transcripts), email hoặc tin nhắn thoại của khách hàng.

Phân tích trang web (Website analytics)

Trong thế giới trực tuyến ngày nay, mức độ quan tâm của khách hàng đối với các sản phẩm/dịch vụ của công ty có thể được đánh giá thông qua lưu lượng truy cập web (web traffic) và các lượt tương tác. Việc khai thác các số liệu đo lường (metrics) có liên quan có thể giúp giải thích nhiều khía cạnh khác nhau trong hành vi trực tuyến của khách truy cập (ví dụ: thời gian dành cho trang web) cũng như mức độ quan tâm của họ.

Dữ liệu nghiên cứu tiếp thị trước đây (Previous marketing research data)

Mặc dù một số nhà quản trị có thể gạt bỏ các nghiên cứu cũ vì cho rằng chúng đã "lỗi thời" (obsolete), nhưng rất có khả năng một số phát hiện trong đó vẫn chứng minh

được tính hữu ích. Dữ liệu định tính có thể chứa đựng một số insight thú vị và dữ liệu định lượng có thể được "mổ xẻ và phân tách" (sliced and diced) theo nhiều cách khác nhau.

Phân tích mạng xã hội (Social media analytics)

Người ta dự kiến sẽ có hơn 4 tỷ người dùng mạng xã hội vào năm 2025. Con số này chiếm hơn một nửa dân số thế giới ở mức 7,9 tỷ người (năm 2021). Việc phớt lờ mạng xã hội là một rủi ro mà các doanh nghiệp không thể gánh chịu (not affordable). Mặc dù nhiều tổ chức hướng tới việc có sự hiện diện nhất định trên mạng xã hội, các công ty nhỏ với ngân sách hạn hẹp vẫn được khuyến nghị nên theo dõi – ít nhất là – các cuộc hội thoại đang diễn ra trên nhiều nền tảng mạng xã hội khác nhau. Các cộng đồng trực tuyến, mặc dù khác biệt so với Facebook và Instagram, cũng được biết đến là nơi diễn ra các cuộc thảo luận có vai trò sống còn mà các nhà quản trị cần theo dõi.

Công nghệ đeo và thiết bị thông minh (Wearable technologies and smart devices)

Fitbit giúp tạo ra dữ liệu liên quan đến sức khỏe (wellbeing-related data), trong khi các ứng dụng và điện thoại thông minh có thể lưu trữ vô vàn thông tin. Việc tích hợp dữ liệu đó với các mảnh thông tin khác thu thập được đã được chứng minh là rất hữu ích cho các nhà tiếp thị.

Nguồn thông tin từ bên thứ ba (Third-party sources of information)

Một số thông tin được chia sẻ bởi **các bên có quyền lợi liên quan** (stakeholders) trên toàn mạng lưới, trong khi các loại thông tin khác cũng có thể được rao bán. Việc lấy dữ liệu người tiêu dùng, chẳng hạn như từ một công ty thẻ tín dụng, có thể hỗ trợ nhà quản trị có được sự hiểu biết hữu ích về hành vi mua sắm của khách hàng.

Nhân viên (Employees)

Bản thân các nhân viên của công ty thường là điểm tiếp xúc trực tiếp (face-to-face contact) đầu tiên với khách hàng. Nhân viên có thể cung cấp thông tin trực tiếp (firsthand information) về khách hàng. Có thể có một lượng lớn thông tin đã hiện hữu ngay trong nội bộ công ty mà có thể được khai phá để hiểu rõ hơn về khách hàng.



Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem chúng trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=31#oembed-1>

2. ĐẠO ĐỨC NGHIÊN CỨU

Mục tiêu học tập

Sau khi học xong chương này, sinh viên phải:

- hiểu được vai trò và tầm quan trọng của nghiên cứu có đạo đức
- làm quen với việc áp dụng quy tắc ứng xử có đạo đức trong suốt quá trình nghiên cứu
- nhận thức được Bộ Quy tắc Ứng xử Nghề nghiệp (Code of Professional Behaviour) dành cho các nhà nghiên cứu tiếp thị.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem chúng trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=33#oembed-1>

Đạo đức nghiên cứu là gì?

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem chúng trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=33#oembed-2>

Đạo đức được định nghĩa là các nguyên tắc đạo lý điều chỉnh hành vi của một người, đặc biệt là khi cá nhân này là thành viên của một tổ chức nghề nghiệp. Đạo đức dựa trên sự lựa chọn của cá nhân và không bị chi phối bởi pháp luật. Các vấn đề pháp lý liên quan đến luật pháp, và việc không tuân thủ chúng sẽ bị trừng phạt (ví dụ: lái xe khi say rượu). Mặt khác, đạo đức là các nguyên tắc và giá trị của cá nhân, chẳng hạn như việc trình bày thông tin một cách chính xác cho khách hàng.

Mặc dù niềm tin đạo đức được nắm giữ bởi các cá nhân, chúng cũng có thể được phản ánh qua các giá trị, thực tiễn và chính sách định hình nên các lựa chọn do những người ra quyết định thực hiện thay mặt cho tổ chức của họ. Các ngành nghề và tổ chức thường xuyên thiết lập một "Bộ Quy tắc Đạo đức" (Code of Ethics) đóng vai trò hướng dẫn hành vi cho các thành viên của ngành nghề hoặc tổ chức đó.¹

Tại Úc, The Research Society – cơ quan đầu ngành dành cho tất cả những cá nhân tham gia hoặc quan tâm đến nghiên cứu thị trường, xã hội và quan điểm – cung cấp các quy tắc và hướng dẫn liên quan đến hành vi đạo đức trong nghiên cứu. Tất cả các Thành viên của The Research Society cũng như các Đối tác Công ty và Khách hàng đều phải tuân thủ Quy tắc này. Quy tắc này áp dụng cho các hoạt động chuyên môn của các Thành viên và Đối tác. Các tổ chức Đối tác Công ty và Khách hàng được yêu cầu phải đảm bảo rằng tất cả các cá nhân do họ tuyển dụng hoặc hợp tác (bất kể có

phải là Thành viên của The Research Society hay không) đều tuân thủ Quy tắc này. Quy tắc Ứng xử Nghề nghiệp của The Research Society (Quy tắc) được giới thiệu lần đầu vào năm 1955 và đã được xem xét, cập nhật thường xuyên kể từ đó. Quy tắc được thiết kế như một khuôn khổ cho sự tự quản (self-regulation). Mục tiêu của nó là thúc đẩy niềm tin của công chúng vào hoạt động nghiên cứu và chứng minh rằng các hoạt động do các Thành viên thực hiện đều được tiến hành một cách chuyên nghiệp và có đạo đức.

1. Lumen Learning, [*Ethical and legal behaviour*](#), n.d.

Quy tắc Ứng xử Nghề nghiệp trong Nghiên cứu (Code of Professional Behaviour in Research)

Mặc dù các hiệp hội ngành nghề có bộ quy tắc và hướng dẫn riêng, bất kỳ nghiên cứu nào được thực hiện tại các tổ chức học thuật có liên quan đến đối tượng tham gia là con người (hoặc động vật) đều phải nhận được sự chấp thuận về đạo đức (ethical clearance) từ các Hội đồng Đạo đức (Ethics Committees) của trường đại học. Các tổ chức học thuật cũng kỳ vọng các nhà nghiên cứu của mình tuân thủ quy tắc thực hành để đảm bảo tính liêm chính và trách nhiệm trong suốt quá trình nghiên cứu. Một số tiêu chuẩn hành vi cao nhất được các trường đại học tại Úc và tổ chức The Research Society kỳ vọng ở các nhà nghiên cứu. Các điểm cốt lõi được thảo luận dưới đây. Toàn văn tài liệu có sẵn trên trang web của [The Research Society](#)² và Hội đồng Nghiên cứu Y khoa và Sức khỏe Quốc gia ([National Health and Medical Research Council](#))³.

Sự tôn trọng (Respect)

Tôn trọng tất cả những người tham gia nghiên cứu là một giá trị cốt lõi mà các nhà nghiên cứu và tổ chức của họ cần phải tuân thủ.

Chú thích: 2. The Research Society, [Code of professional behaviour, 2022](#). 3. National Health and Medical Research Council, [National statement on ethical conduct of research](#): (NHMRC manual), 2018.

Điều này được thảo luận trong bối cảnh tôn trọng quyền riêng tư, tính bảo mật và sự nhạy cảm đối với các vấn đề văn hóa của người tham gia. Bất kỳ thỏa thuận nào với người tham gia nghiên cứu đều cần phải được thực hiện đầy đủ.

Sự nghiêm ngặt (Rigour)

Khi tiến hành các hoạt động chuyên môn, các nhà nghiên cứu phải đảm bảo rằng công việc của họ được thực hiện tuân theo các nguyên tắc, phương pháp và kỹ thuật nghiên cứu khoa học phù hợp. Không nghi ngờ gì nữa, đây là một yêu cầu then chốt đối với tất cả các dự án tìm kiếm "sự chấp thuận về đạo đức" (ethical approval) tại các trường đại học và viện nghiên cứu. Nghiên cứu phải được thực hiện một cách trung thực, tuân theo các nguyên tắc đã được công nhận về thực hành nghiên cứu. Các phát hiện nghiên cứu (research findings) – dù thuận lợi hay không thuận lợi – đều phải được truyền đạt/phổ biến để chúng có thể đóng góp vào việc nâng cao tri thức và sự hiểu biết.

Thông tin sai lệch (Misleading information)

Các nhà nghiên cứu không được đưa ra những tuyên bố sai sự thật hoặc gây hiểu lầm về kỹ năng, kinh nghiệm hoặc hoạt động của chính họ hay của tổ chức mình. Một nhà nghiên cứu không được đánh lừa người tham gia khi kêu gọi sự hỗ trợ của họ. Cần thừa nhận rằng mọi nghiên cứu đều gây ra một mức độ khó chịu và bất tiện nhất định. Ví dụ, việc trả lời các bảng khảo sát dài dòng có thể gây ra sự bức bối cho một số người trả lời. Trong bất kỳ giai đoạn nào, nhà nghiên cứu cũng không được cố ý "che giấu" thời gian thực tế cần thiết để tham gia nghiên cứu. Tương tự như vậy, một số

câu hỏi nghiên cứu có thể khiến người tham gia cảm thấy lo lắng hoặc đau buồn. Điều này có thể liên quan đến một tình trạng y tế hoặc các quan điểm tôn giáo, v.v. Các nhà nghiên cứu cần lưu tâm đến cách quản lý "rủi ro" như vậy, ngay cả khi họ cảm thấy khả năng xảy ra điều đó là rất nhỏ.

Phân biệt nghiên cứu với các hoạt động khác

Bất kỳ hoạt động tiếp thị nào nhắm vào các cá nhân tham gia, chẳng hạn như xúc tiến bán hàng (sales promotion), tiếp thị trực tiếp (direct marketing), bán hàng trực tiếp (direct selling) và các hoạt động tương tự, đều phải được phân biệt rõ ràng và tách biệt khỏi hoạt động nghiên cứu. Trong tất cả các giai đoạn của quy trình nghiên cứu (phát triển đề xuất, thu thập dữ liệu, phân tích và lập báo cáo), nhà nghiên cứu không bao giờ được tham gia vào các hoạt động nhằm mục đích thao túng, đánh lừa hoặc cưỡng ép các cá nhân....các nhà nghiên cứu và tổ chức của họ cần phải tuân thủ. Điều này được thảo luận trong bối cảnh tôn trọng quyền riêng tư, tính bảo mật và sự nhạy cảm đối với các vấn đề văn hóa của người tham gia. Bất kỳ thỏa thuận nào với người tham gia nghiên cứu đều cần phải được thực hiện đầy đủ.

Sự nghiêm ngặt (Rigour)

Khi tiến hành các hoạt động chuyên môn, các nhà nghiên cứu phải đảm bảo rằng công việc của họ được thực hiện tuân theo các nguyên tắc, phương pháp và kỹ thuật nghiên cứu khoa học phù hợp. Không nghi ngờ gì nữa, đây là một yêu cầu then chốt đối với tất cả các dự án tìm kiếm "sự chấp thuận về đạo đức" (ethical approval) tại các trường đại học và viện nghiên cứu. Nghiên cứu phải được thực hiện một cách trung thực,

tuân theo các nguyên tắc đã được công nhận về thực hành nghiên cứu. Các phát hiện nghiên cứu (research findings) – dù thuận lợi hay không thuận lợi – đều phải được truyền đạt/phổ biến để chúng có thể đóng góp vào việc nâng cao tri thức và sự hiểu biết.

Thông tin sai lệch (Misleading information)

Các nhà nghiên cứu không được đưa ra những tuyên bố sai sự thật hoặc gây hiểu lầm về kỹ năng, kinh nghiệm hoặc hoạt động của chính họ hay của tổ chức mình. Một nhà nghiên cứu không được đánh lừa người tham gia khi kêu gọi sự hỗ trợ của họ. Cần thừa nhận rằng mọi nghiên cứu đều gây ra một mức độ khó chịu và bất tiện nhất định. Ví dụ, việc trả lời các bảng khảo sát dài dòng có thể gây ra sự bức bối cho một số người trả lời. Trong bất kỳ giai đoạn nào, nhà nghiên cứu cũng không được cố ý "che giấu" thời gian thực tế cần thiết để tham gia nghiên cứu. Tương tự như vậy, một số câu hỏi nghiên cứu có thể khiến người tham gia cảm thấy lo lắng hoặc đau buồn. Điều này có thể liên quan đến một tình trạng y tế hoặc các quan điểm tôn giáo, v.v. Các nhà nghiên cứu cần lưu tâm đến cách quản lý "rủi ro" như vậy, ngay cả khi họ cảm thấy khả năng xảy ra điều đó là rất nhỏ.

Phân biệt nghiên cứu với các hoạt động khác

Bất kỳ hoạt động tiếp thị nào nhắm vào các cá nhân tham gia, chẳng hạn như xúc tiến bán hàng (sales promotion), tiếp thị trực tiếp (direct marketing), bán hàng trực tiếp (direct selling) và các hoạt động tương tự, đều phải được phân biệt rõ ràng và tách biệt khỏi hoạt động nghiên cứu. Trong mọi giai đoạn của quy trình nghiên cứu (phát triển đề xuất, thu thập dữ liệu, phân tích và báo cáo), nhà nghiên cứu không bao giờ được

tham gia vào các hoạt động nhằm mục đích thao túng, đánh lừa hoặc cưỡng ép các cá nhân.

Sự đồng ý và tham gia tự nguyện

Nếu các nhà nghiên cứu có ý định tiết lộ thông tin nghiên cứu có thể nhận dạng (identifiable research information) về những người tham gia, họ phải nhận được sự đồng ý của đối tượng. Ví dụ, những người tham gia phải được thông báo trước khi các kỹ thuật quan sát hoặc thiết bị ghi âm, ghi hình được sử dụng cho một dự án. Ở mọi giai đoạn của dự án, sự tham gia của các cá nhân là hoàn toàn tự nguyện và người tham gia phải có khả năng tự nguyện rút lui (opt-out), từ chối cung cấp thông tin tại bất kỳ thời điểm nào trong quá trình thu thập dữ liệu.

Tiết lộ thông tin của người tham gia

Các nhà nghiên cứu phải thông báo cho người tham gia ngay từ khi bắt đầu dự án rằng liệu dự án đó là ẩn danh (anonymous) hay có thể nhận dạng. Nếu là ẩn danh, tính ẩn danh của người tham gia phải được duy trì nghiêm ngặt. Nếu có thể nhận dạng, người tham gia phải được thông báo ở giai đoạn này về mục đích cũng như bên nhận thông tin nghiên cứu có thể nhận dạng đó, trừ khi có những lý do về mặt phương pháp luận (methodological reasons) không cho phép làm như vậy vào thời điểm này. Sau khi dữ liệu đã được thu thập, các nhà nghiên cứu phải thông báo cho người tham gia về mục đích sử dụng và yêu cầu họ xác nhận lại sự đồng ý cho phép tiết lộ thông tin.

Các nhà nghiên cứu phải thông báo cho người tham gia biết ai sẽ là người nhận thông tin nghiên cứu và thông tin đó sẽ được sử dụng cho mục đích gì.

Hành vi bán hàng núp bóng nghiên cứu (*Sugging*) và một thực tiễn có liên quan là gây quỹ núp bóng nghiên cứu (*frugging*) sẽ làm tổn hại nghiêm trọng đến danh tiếng của ngành nghiên cứu thị trường. Hãy thực hiện hoạt động dưới đây để đoán xem hai thuật ngữ này có nghĩa là gì.



Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=33#h5p-2>

Xem xét đối với các nhóm dễ bị tổn thương

Đôi khi, thông tin cần được thu thập từ các nhóm người cụ thể. Những nhóm này có thể bao gồm phụ nữ mang thai, những người cần được chăm sóc y tế liên tục, hoặc những người bị khuyết tật về nhận thức hoặc thể chất. Trong những trường hợp khác, các nhà nghiên cứu có thể cần phỏng vấn những người đã từng tham gia vào các hoạt động bất hợp pháp. Cần có sự xem xét thích đáng trong việc thiết kế và tiến hành nghiên cứu đối với các phân khúc dân số này. Sự an toàn và sức khỏe (well-being) của cả người tham gia và nhà nghiên cứu là điều quan trọng tối thượng.

Khi thu thập thông tin từ trẻ em và thanh thiếu niên, các nhà nghiên cứu phải hết sức thận trọng. Trước khi thu thập thông tin, phải xin phép cha mẹ hoặc người lớn chịu

trách nhiệm giám hộ. Hơn nữa, sự an toàn cũng như an ninh về mặt cảm xúc và tâm lý của trẻ em hoặc thanh thiếu niên phải được đảm bảo.

Trong một số tình huống, một người có thể ở trong một mối quan hệ không bình đẳng hoặc phụ thuộc (ví dụ: một nhân viên của một tổ chức được người sử dụng lao động yêu cầu tham gia vào một nghiên cứu) và có thể cảm thấy bị ép buộc (coerced) phải tham gia vào nghiên cứu đó. Nhà nghiên cứu sẽ cần phải xem xét tất cả các tình huống như vậy và thực hiện các bước nhằm giảm thiểu tác động bất lợi của tình huống đó đối với nghiên cứu cũng như đối với các mối quan hệ của người tham gia.

Lưu trữ và bảo mật dữ liệu

Các nhà nghiên cứu phải đảm bảo an ninh cho tất cả các thông tin liên quan đến dự án. Các nhà nghiên cứu phải thực hiện các biện pháp phòng ngừa hợp lý để bảo vệ dữ liệu nghiên cứu có thể nhận dạng (identifiable research data). Dữ liệu nghiên cứu có thể nhận dạng được thu thập và lưu trữ theo Quy tắc này không được lưu giữ lâu hơn mức cần thiết cho mục đích thu thập hoặc xử lý dữ liệu đó. Khi thông tin nghiên cứu có thể nhận dạng không còn cần thiết nữa, phải thực hiện các bước hợp lý để tiêu hủy nó hoặc đảm bảo rằng nó được phi nhận dạng (de-identified).

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=33#h5p-3>

3. NGHIÊN CỨU THỨ CẤP (SECONDARY RESEARCH)

Mục tiêu học tập

Sau khi học xong chương này, sinh viên phải có khả năng:

- Giải thích khái niệm nghiên cứu thứ cấp
- Làm rõ những lợi ích và hạn chế chính của nghiên cứu thứ cấp
- Đánh giá các nguồn dữ liệu thứ cấp (secondary data) khác nhau

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem chúng trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=37#oembed-1>

Nghiên cứu thứ cấp là gì?

Trong những tình huống mà nhà nghiên cứu không trực tiếp tham gia vào quá trình thu thập dữ liệu (nghiên cứu sơ cấp - *primary research*), họ có thể phải dựa vào các

thông tin và dữ liệu hiện có để đi đến những kết luận hoặc kết quả nghiên cứu cụ thể. Nghiên cứu thứ cấp (*secondary research*), hay còn được gọi là nghiên cứu tại bàn (*desk research*), là một phương pháp nghiên cứu liên quan đến việc sử dụng các thông tin đã được thu thập từ trước cho một mục đích nghiên cứu khác. Trong chương này, chúng tôi sẽ giải thích nghiên cứu thứ cấp là gì, phương pháp này hoạt động như thế nào và chia sẻ một số ví dụ về việc ứng dụng nó trong thực tiễn.



Các nguồn dữ liệu thứ cấp

Hai nguồn dữ liệu thứ cấp chính bao gồm:

- Nguồn nội bộ (Internal sources)
- Nguồn bên ngoài (External sources)

Các nguồn dữ liệu thứ cấp nội bộ tồn tại ngay bên trong tổ chức. Đó có thể là các bản báo cáo, các phát hiện nghiên cứu trước đây, hoặc các tài liệu cũ mà ngày nay vẫn có thể được sử dụng để thấu hiểu một hiện tượng cụ thể. Những thông tin này có thể chỉ được lưu hành nội bộ cho các thành viên của tổ chức và có thể được xem là một tài sản có giá trị.

Các nguồn dữ liệu thứ cấp bên ngoài nằm ngoài phạm vi tổ chức và đề cập đến những thông tin được lưu trữ tại các thư viện công cộng, các cơ quan chính phủ, văn phòng hội đồng, các hiệp hội khác nhau cũng như trên các tờ báo hoặc các bài báo trên tạp chí học thuật (journal articles).

Việc các nhà nghiên cứu tìm kiếm kỹ lưỡng các thông tin thứ cấp trước khi đầu tư thời gian và nguồn lực vào việc thu thập dữ liệu sơ cấp (primary data) là một bước đi hoàn toàn hợp lý. Trong nghiên cứu học thuật, các học giả không được phép chuyển sang giai đoạn tiếp theo cho đến khi họ chứng minh được rằng mình đã thực hiện việc rà soát, tổng quan tất cả các nghiên cứu trước đó. Giả sử một nhà nghiên cứu muốn xem xét các đặc điểm của cộng đồng dân cư nhập cư tại khu vực Tây Sydney (Western Sydney). Những nhóm thông tin sau đây đã có sẵn trong nhiều báo cáo khác nhau được trích xuất từ dữ liệu điều tra dân số của Cục Thống kê Úc (Australian Bureau of Statistics):

- Nơi sinh của cư dân
- Ngôn ngữ được cư dân sử dụng tại nhà
- Quy mô gia đình
- Mức thu nhập
- Trình độ học vấn

Bằng cách tiếp cận nguồn dữ liệu thứ cấp có sẵn như vậy, nhà nghiên cứu có khả năng tiết kiệm đáng kể thời gian, tiền bạc và công sức. Hơn nữa, khi dữ liệu được trích xuất từ một cơ quan có thẩm quyền và uy tín (reputable source), điều này càng làm củng cố thêm độ tin cậy của nhà nghiên cứu trong việc xác định và sử dụng một nguồn thông tin xác đáng.

Đánh giá dữ liệu thứ cấp

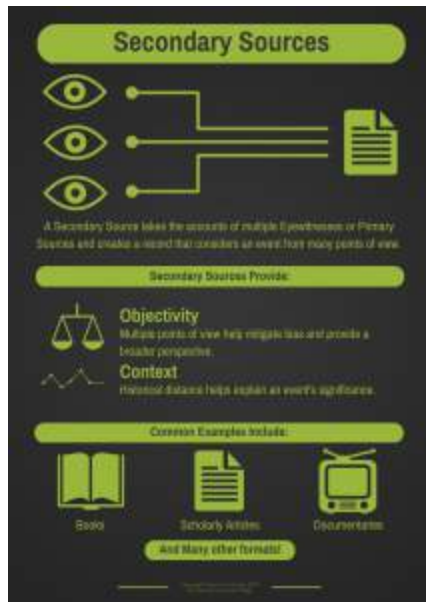
Việc đánh giá dữ liệu thứ cấp đóng vai trò rất quan trọng. Nguồn dữ liệu này không phải lúc nào cũng được cung cấp miễn phí. Cần xem xét các yếu tố sau đây vì chúng liên quan trực tiếp đến độ tin cậy (*reliability*) và tính hợp lệ (*validity*) của các kết quả nghiên cứu, chẳng hạn như liệu:¹

- nguồn cung cấp có đáng tin cậy hay không
- các đặc điểm của mẫu (*sample characteristics*), thời gian thu thập và tỷ lệ phản hồi (*response rate* - nếu có) của dữ liệu có phù hợp hay không
- các phương pháp thu thập dữ liệu có phù hợp và được chấp nhận trong chuyên ngành của bạn hay không
- dữ liệu có được thu thập một cách nhất quán hay không
- mọi hoạt động mã hóa (*data coding*) hoặc điều chỉnh dữ liệu có phù hợp và đầy đủ hay không
- hệ thống tài liệu (*documentation*) của nghiên cứu gốc nơi dữ liệu được thu thập có đủ chi tiết để bạn đánh giá chất lượng của nó hay không
- có đủ thông tin trong siêu dữ liệu (*metadata*) hoặc chính dữ liệu đó để trích dẫn nguồn gốc một cách chuẩn xác hay không.

Bên cạnh những điểm đã đề cập ở trên, một số vấn đề thực tiễn cũng cần được đánh giá, bao gồm chi phí truy cập và khung thời gian liên quan để tiếp cận được dữ liệu có phù hợp hay không.

Chú thích:

1. Griffith University, [Research data: get started, 2022](#).



Nghiên cứu sơ cấp (Primary Research)

- Là nghiên cứu thu thập dữ liệu trực tiếp (first-hand). Quá trình này có thể đòi hỏi rất nhiều thời gian.
 - Tạo ra dữ liệu thô (raw data) thuộc quyền sở hữu của chính nhà nghiên cứu.
 - Có tính liên quan mật thiết và bám sát các mục tiêu của nghiên cứu.
 - Do nhà nghiên cứu trực tiếp tiến hành. Có khả năng chịu ảnh hưởng bởi sự thiên kiến hoặc sai lệch từ phía nhà nghiên cứu (researcher bias).
 - Có thể tiêu tốn nhiều chi phí để thực hiện.
-

Bảng 2: Sự khác biệt giữa Nghiên cứu Sơ cấp và Nghiên cứu Thứ cấp²

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=37#h5p-4>

Ghi nhận tác quyền hình ảnh (Media Attributions)

- Hình ảnh giáo trình tiếp thị © [Sally Tsoutsas - Nhiếp ảnh gia của WSU](#), được cấp phép theo [giấy phép CC BY-NC-ND](#) (Ghi nhận tác giả - Phi thương mại - Không phái sinh / Attribution NonCommercial NoDerivatives).
- Đồ họa thông tin (infographic) về các nguồn tài liệu thứ cấp © [Shonn M. Haren](#), được cấp phép [theo giấy phép CC BY](#) (Ghi nhận tác giả / Attribution).

Chú thích: 2. Qualtrics XM, [Secondary research: definition, methods and examples](#), 2022.

4. SỬ DỤNG DỮ LIỆU ĐIỀU TRA DÂN SỐ

Siddharth Jain

Mục tiêu học tập

- Làm quen với việc truy cập vào một nguồn dữ liệu thứ cấp
- Truy xuất thông tin về dân số Úc

Giới thiệu

Cục Thống kê Úc (Australian Bureau of Statistics - ABS) là cơ quan trung ương của Úc chịu trách nhiệm thiết kế, tiến hành, lưu trữ và thống kê, lập bảng dữ liệu liên quan đến cuộc điều tra dân số quốc gia (national census). Đây là nguồn dữ liệu toàn diện nhất về mọi khía cạnh của đời sống tại Úc, bao gồm dân số, nền kinh tế, môi trường và các ngành công nghiệp.

Dữ liệu điều tra dân số có thể được truy cập thông qua một số công cụ dữ liệu Điều tra dân số, trong đó bao gồm QuickStats. QuickStats là một phương pháp đơn giản để nắm bắt nhanh thông tin về một khu vực. Công cụ này cung cấp một bản tóm tắt về một khu vực, bao gồm các số liệu nhân khẩu học và các bảng biểu về những đặc điểm

chính của người dân sinh sống tại khu vực đó, gia đình của họ và tình trạng nhà ở (dwellings). Dưới đây là một cái nhìn tổng quan nhanh về dân số Úc:

<https://abs.gov.au/census/find-census-data/quickstats/2021/AUS>

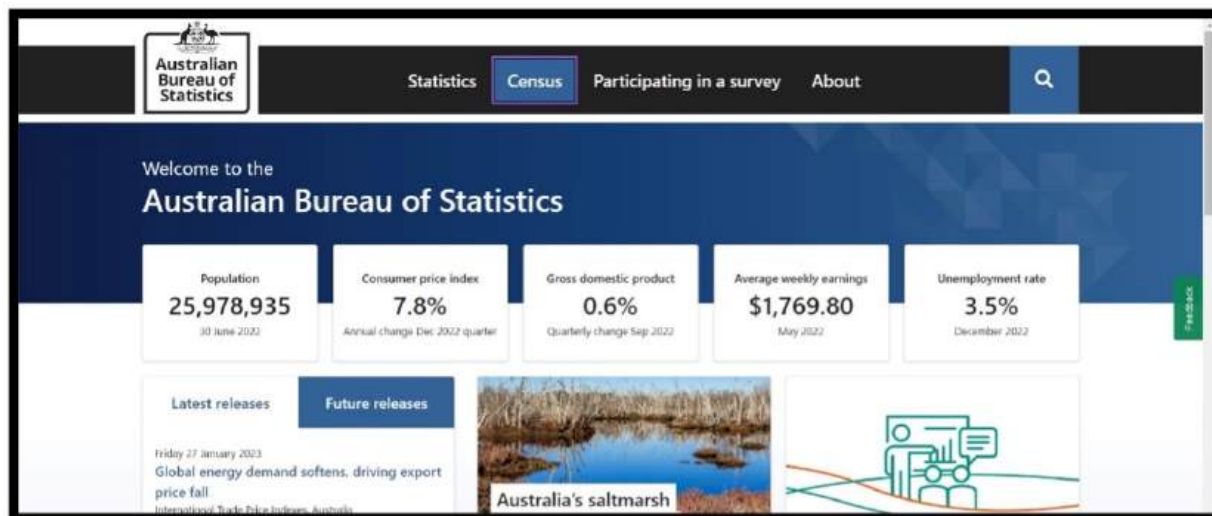
Truy cập QuickStats

QuickStats là một công cụ hữu ích để truy cập các thông tin được tùy chỉnh (customised information) về những người đang sinh sống tại Úc. Có ba loại thông tin hiện có sẵn thông qua QuickStats:

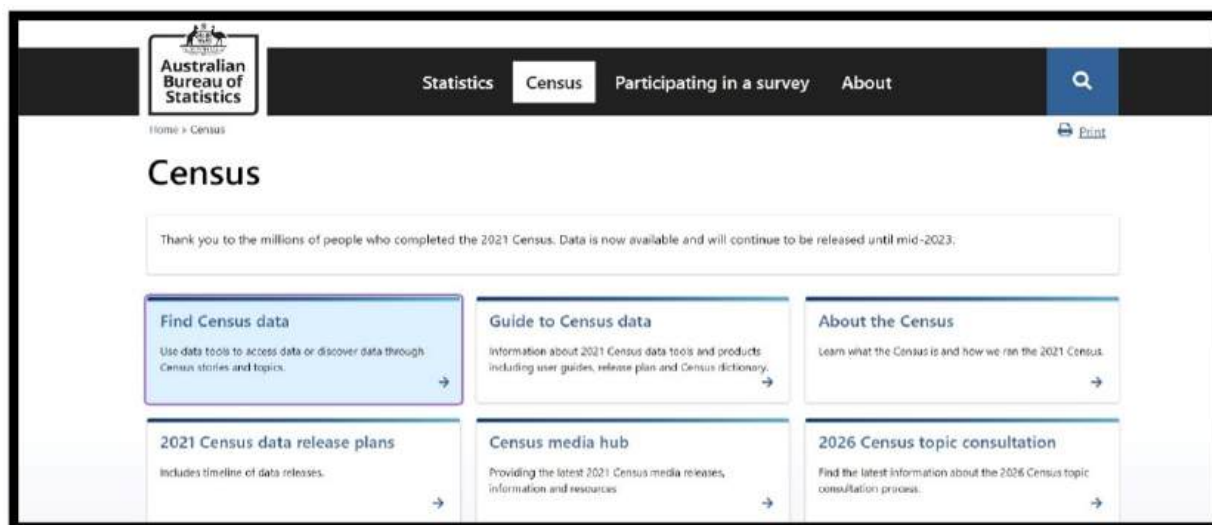
- Tất cả các cá nhân (All persons)
- Người Thổ dân và người dân đảo Torres Strait (Aboriginal and Torres Strait Islander people)
- Quốc gia nơi sinh (Country of Birth)

Giả sử mục tiêu nghiên cứu của bạn là thu thập thông tin về những cá nhân sinh ra tại Nhật Bản đang cư trú tại tiểu bang New South Wales. Bạn có thể thu thập thông tin nhân khẩu học về nhóm đối tượng (cohort) này bằng cách làm theo các hướng dẫn dưới đây:

1. Truy cập vào trang web www.abs.gov.au
2. Nhấp vào mục "Census" (Điều tra dân số) trên thanh công cụ ở phần đầu trang web.



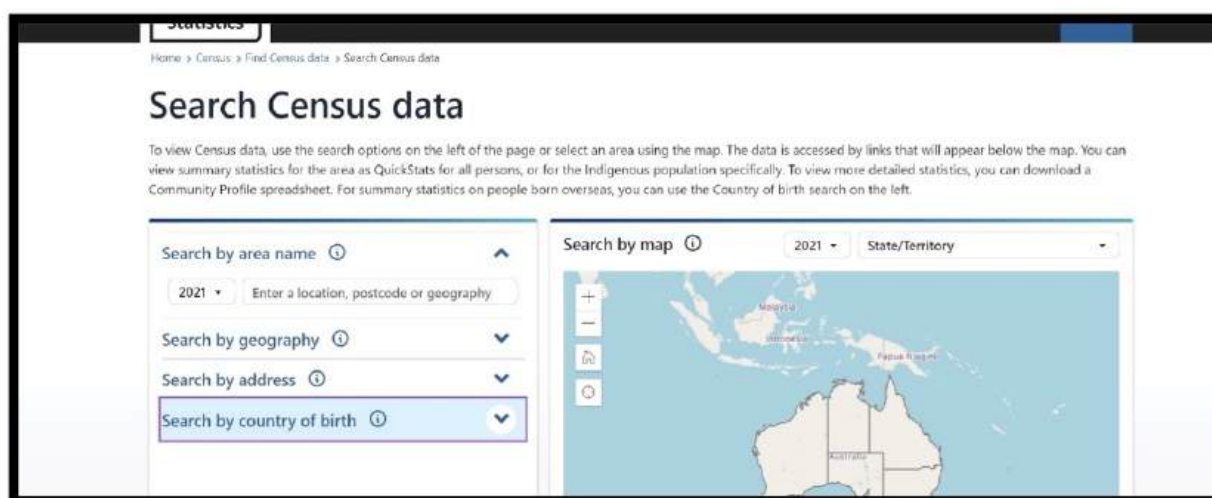
Nhấp vào “Find Census Data” (Tìm dữ liệu Điều tra dân số).



Nhấp vào “Search Census data” (Tìm kiếm dữ liệu Điều tra dân số).



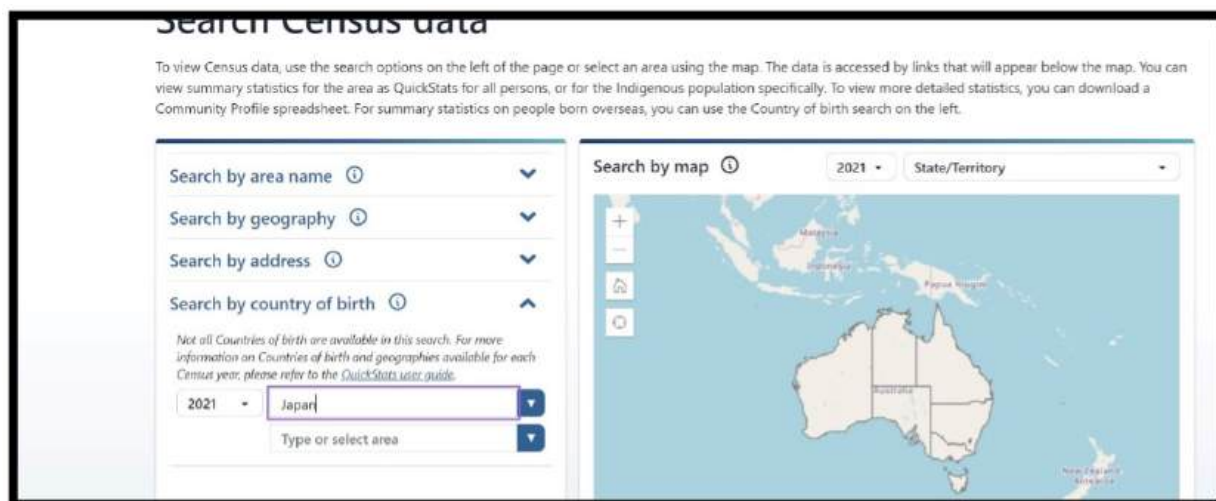
Nhấp vào “Search by country of birth” (Tìm kiếm theo quốc gia nơi sinh).



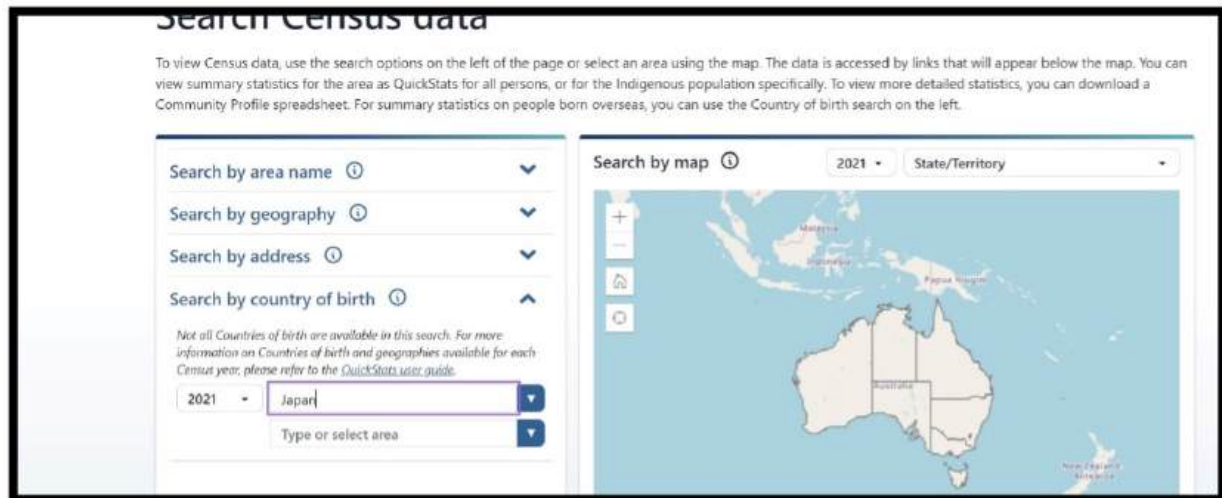
Nhập hoặc chọn “Japan” (Nhật Bản) vào trường dữ liệu.



Nhập hoặc chọn “New South Wales” vào trường dữ liệu.



Nhập vào “**Country of birth**” (Quốc gia nơi sinh) để nhận dữ liệu liên quan.



Hoạt động

Truy cập: <https://www.abs.gov.au/census/find-census-data/search-by-area>. Nhập tên vùng ngoại ô ‘Parramatta’ vào khoảng trống được cung cấp trong mục ‘Search by Area Name’ (Tìm kiếm theo tên khu vực).

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=45#h5p-5>

Về tác giả chương sách – Tiến sĩ Siddharth Jain



Tiến sĩ Siddharth Jain là người vừa tốt nghiệp chương trình Tiến sĩ tại Đại học Western Sydney (WSU), đồng thời sở hữu bằng Thạc sĩ Quản trị Kinh doanh và Kế toán chuyên nghiệp. Siddharth cũng là thành viên liên kết của Hiệp hội Kế toán công chứng (Associate CPA).

Với hơn ba năm kinh nghiệm trong cả lĩnh vực giảng dạy và nghiên cứu, Siddharth đã trau dồi các kỹ năng của mình tại trường kinh doanh thuộc WSU, đạt được chuyên môn sâu trong lĩnh vực kinh tế lượng (econometrics) và phân tích dữ liệu, điều này đã được minh chứng qua nhiều công trình nghiên cứu đã được công bố của ông.

Ghi nhận tác quyền hình ảnh

- Ảnh chụp màn hình trang web <https://www.abs.gov.au/> được cấp phép theo giấy phép CC BY (Ghi nhận tác giả / Attribution).
- Ảnh chụp màn hình trang web <https://www.abs.gov.au/census> được cấp phép theo giấy phép CC BY (Ghi nhận tác giả / Attribution).
- Ảnh chụp màn hình trang web <https://www.abs.gov.au/census/find-census-data> được cấp phép theo giấy phép CC BY (Ghi nhận tác giả / Attribution).
- Ảnh chụp màn hình trang web <https://www.abs.gov.au/census/find-census-data/search-by-area> được cấp phép theo giấy phép CC BY (Ghi nhận tác giả / Attribution).

- Ảnh chụp màn hình trang web <https://www.abs.gov.au/census/find-census-data/search-by-area> được cấp phép theo giấy phép CC BY (Ghi nhận tác giả / Attribution).
- Ảnh chụp màn hình trang web <https://www.abs.gov.au/census/find-census-data/search-by-area> được cấp phép theo giấy phép CC BY (Ghi nhận tác giả / Attribution).
- Chân dung Siddharth Jain © Siddharth Jain được cấp phép theo giấy phép CC BY-NC-ND (Ghi nhận tác giả - Phi thương mại - Không phái sinh / Attribution NonCommercial NoDerivatives).

5. NGHIÊN CỨU SƠ CẤP (PRIMARY RESEARCH)

Mục tiêu học tập

- Định nghĩa về nghiên cứu sơ cấp và xác định được các phương pháp khác nhau để thu thập dữ liệu sơ cấp.
- Giải thích được những ưu điểm và nhược điểm của các phương pháp thu thập dữ liệu (sơ cấp) khác nhau.
- Nhận thức được các lợi thế và bất lợi của nghiên cứu sơ cấp nói chung.
- Hiểu rõ những khác biệt cốt yếu giữa nghiên cứu sơ cấp và nghiên cứu thứ cấp

Nghiên cứu sơ cấp là gì?

Dữ liệu sơ cấp (Primary data) đề cập đến những thông tin được nhà nghiên cứu thu thập riêng biệt cho dự án hiện tại. Phương pháp này đặc biệt hữu ích khi việc chỉ dựa vào dữ liệu thứ cấp là không đủ. Một số câu hỏi nghiên cứu – ví dụ như: *Người tiêu dùng nghĩ gì về thương hiệu nước giải khát mới của chúng ta?* – đòi hỏi phải thu thập dữ liệu sơ cấp trực tiếp (first-hand). Mặc dù việc thu thập dữ liệu sơ cấp tốt và nghiêm ngặt có thể tiêu tốn nhiều thời gian và nguồn lực, nhưng lợi ích mang lại là dữ liệu này hoàn toàn thuộc quyền sở hữu của nhà nghiên cứu hoặc tổ chức của họ.

Dưới đây là một số công cụ chính được sử dụng để thu thập dữ liệu sơ cấp. Cần lưu ý rằng dữ liệu sơ cấp có thể bao gồm cả thông tin định lượng (quantitative) lẫn định tính (qualitative).

Sự khác biệt cốt lõi giữa hai loại dữ liệu sơ cấp:

- **Dữ liệu định lượng:** Tập trung vào các con số, số liệu thống kê để đo lường "bao nhiêu" hoặc "mức độ nào".
- **Dữ liệu định tính:** Tập trung vào việc tìm hiểu lý do "tại sao" và "như thế nào" thông qua các ý kiến, cảm xúc và động lực của khách hàng.

Chú thích:

1. Qualtrics XM, [*Primary Research: Definition, Examples & Methods*](#), 2022

Phỏng vấn (Interviews)



Phỏng vấn bao gồm các phiên làm việc trực tiếp một đối một (1-1), có thể thực hiện qua điện thoại hoặc đối mặt trực tiếp giữa nhà nghiên cứu và người tham gia. Do tác động của đại dịch COVID-19, nhiều cuộc phỏng vấn đã được chuyển sang hình thức trực tuyến và được thực hiện trên các nền tảng họp

trực tuyến khác nhau như Zoom.

Phỏng vấn là phương pháp tuyệt vời để khám phá một chủ đề một cách sâu sắc hơn. Ví dụ, một nhà nghiên cứu có thể muốn tìm hiểu động cơ của mọi người đằng sau việc quyên góp từ thiện cho một tổ chức cụ thể. Cách tiếp cận một đối một sẽ hiệu quả hơn nhiều trong việc dò hỏi (probing) và thấu hiểu những lý do ẩn sau hành động đó.

Phỏng vấn cũng đặc biệt hữu ích khi trao đổi với các chuyên gia. Một chuyên gia trong lĩnh vực cụ thể, chẳng hạn như chuỗi cung ứng, có thể cung cấp những hiểu biết (insights) giá trị về một vấn đề. Với kinh nghiệm tích lũy, những chuyên gia này là sự bổ sung hữu ích vào danh sách đối tượng phỏng vấn của nhà nghiên cứu.

Nhược điểm của phương pháp phỏng vấn:

- **Thời gian:** Tốn nhiều thời gian để thực hiện phỏng vấn với nhiều người khác nhau.
- **Chi phí:** Có thể trở nên tốn kém.
- **Tính chất công việc:** Dễ gây mệt mỏi và tẻ nhạt (tedious), đặc biệt là khi các đối tượng phỏng vấn phân bổ rải rác về mặt địa lý.

Khảo sát (Surveys)

Khảo sát hoặc bảng câu hỏi (questionnaires) là những công cụ có cấu trúc nhằm thu thập dữ liệu định lượng. Khảo sát có thể được tiến hành trực tiếp (ví dụ: tại siêu thị) hoặc qua điện thoại. Hiện nay, nhiều cuộc khảo sát được phân phối qua email. Trong các trường hợp khác, những cá nhân cụ thể (chẳng hạn như nhân viên của một tổ chức) sẽ được mời tham gia khảo sát được lưu trữ trên một trang web nhất định.

Nhìn chung, các cuộc khảo sát sẽ được hoàn thành bởi một số lượng lớn người tham gia để kết quả đạt được **ý nghĩa thống kê** (statistically significant). Mặc dù khảo sát có thể được dùng để thu thập dữ liệu dạng văn bản, nhưng nhiều người thường không cung cấp thông tin chi tiết bằng văn bản viết. Do đó, khảo sát nên được sử dụng cho các **câu hỏi đóng** (closed-ended questions).

Ví dụ về câu hỏi đóng trong khảo sát:

- "Bạn có sở hữu xe ô tô không?"
- "Lần cuối cùng bạn mua tủ lạnh là khi nào?"
- "Trên thang điểm từ 1 đến 5, với 1 là 'hoàn toàn không có khả năng' và 5 là 'rất có khả năng', khả năng bạn mua điện thoại thông minh mới trong năm nay là bao nhiêu?"

Bảng khảo sát cần được thiết kế tốt để có thể thu thập được những thông tin hữu ích. Chúng **không có tính linh hoạt**, vì vậy tất cả các câu hỏi liên quan cần phải được đưa vào trước khi bắt đầu dự án nghiên cứu.

Quan sát (Observations)

Nghiên cứu quan sát không đòi hỏi sự tương tác giữa nhà nghiên cứu và đối tượng nghiên cứu. Cả thông tin định tính và định lượng đều được thu thập trên cơ sở thực hiện các quan sát có hệ thống và ghi chép lại chúng. Mặc dù cách tiếp cận này giúp giảm thiểu khả năng xảy ra thiên kiến của nhà nghiên cứu (researcher bias), nhưng nó cũng có hạn chế là không cung cấp được lời giải thích cho các hành động do cá nhân thực hiện.

Chẳng hạn, một quản lý cửa hàng bán lẻ – với sự trợ giúp của hệ thống camera trong cửa hàng – có thể nhìn thấy cách khách hàng mua sắm giữa các lối đi, nhưng họ có thể không có được lời giải thích cho hành vi này. Tương tự, những "khách hàng bí mật" (mystery shoppers) – những người đóng giả làm người mua sắm và được các siêu thị như Coles sử dụng để đánh giá dịch vụ của chính mình – cũng là những người thực hiện nghiên cứu quan sát.

Thảo luận nhóm tập trung (Focus Groups)

Thảo luận nhóm tập trung là phương pháp hiệu quả để tạo ra một lượng lớn thông tin trong một khoảng thời gian hạn chế. Phương pháp này đặc biệt hữu ích khi sự thảo luận giữa các thành viên tham gia có thể dẫn đến những ý tưởng hoặc giải pháp mới.

Phương pháp nghiên cứu sơ cấp này đòi hỏi chuyên môn của một **người điều phối (facilitator)** giỏi – người có kỹ năng lôi cuốn tất cả các thành viên trong nhóm cùng tham gia. Mặc dù đây là phương pháp yêu thích của một số chuyên gia thực hành, nhưng cần phải cẩn trọng khi vận hành một nhóm thảo luận tập trung. Một số người

có thể không tham gia chút nào, hoặc có thể phản hồi theo một cách nhất định để cố gắng hòa nhập với ý kiến của số đông.

Thiết bị thông minh (Smart devices)

Người tiêu dùng trong năm 2022 sử dụng một loạt các thiết bị thông minh như điện thoại thông minh, công nghệ đeo tay (wearable technologies), và các thiết bị gia dụng thông minh như tủ lạnh và tivi. Vì các thiết bị này được kết nối với internet, chúng không chỉ tạo ra dữ liệu mà còn lưu trữ, xử lý và truyền tải thông tin này. Mặc dù có thể tồn tại nhiều câu hỏi về đạo đức xoay quanh việc loại dữ liệu nào đang bị truy cập và bởi ai, nhưng thực tế là những thông tin này có thể hữu ích cho nhiều doanh nghiệp.



Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=48#oembed-1>

Dữ liệu mạng xã hội: Thứ cấp hay Sơ cấp?

Đây là một điểm thường xuyên được thảo luận trong cộng đồng nghiên cứu. Liệu dữ liệu từ mạng xã hội nên được phân loại là thứ cấp hay sơ cấp? Câu trả lời phụ thuộc vào cách thức dữ liệu trên mạng xã hội được truy xuất.

Nếu một nền tảng mạng xã hội như Twitter được sử dụng để thu thập tất cả các bài đăng (tweets) có gắn thẻ băm (hashtag) về một thương hiệu (ví dụ: #Nike), thì dữ liệu đó phải được xem là dữ liệu sơ cấp. Dữ liệu sơ cấp này đang được thu thập trực tiếp — một cách nguyên bản (first-hand) — bởi nhà nghiên cứu cho một mục đích cụ thể. Một số người có thể lập luận rằng nhà nghiên cứu đã không chủ động yêu cầu các chủ tài khoản đăng bài. Tuy nhiên, tập dữ liệu như vậy về bản chất là sơ cấp vì chưa có ai khác biên soạn thông tin này (với hashtag #Nike) cho một mục tiêu nghiên cứu cụ thể.

Ngược lại, nếu cùng một nhà nghiên cứu đó sử dụng một báo cáo về phân tích mạng xã hội hoặc tải xuống một báo cáo được đăng trên Twitter (ví dụ: một bài báo về việc sử dụng Twitter của những người ở độ tuổi 20), thì những thông tin như vậy có thể được phân loại là dữ liệu "thứ cấp".

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=48#h5p-6>

Ghi nhận tác quyền hình ảnh (Media Attributions)

- Hình ảnh phỏng vấn (Interview) © [Sally Tsoutsas - Nhiếp ảnh gia của WSU](#) (Đại học Western Sydney), được bảo lưu mọi quyền ([All Rights Reserved](#)).

6. NGHIÊN CỨU ĐỊNH TÍNH VỚI NGHIÊN CỨU ĐỊNH LƯỢNG (QUALITATIVE VS QUANTITATIVE RESEARCH)

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Phân biệt giữa dữ liệu định tính và dữ liệu định lượng.
- Lựa chọn phương pháp phù hợp cho các câu hỏi nghiên cứu khác nhau.
- Nhận diện được những hạn chế của nghiên cứu định tính và nghiên cứu định lượng.

Nghiên cứu định tính và nghiên cứu định lượng

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=50#oembed-1>

Một trong những tiêu chí chính được sử dụng để phân loại thông tin là dựa trên việc thông tin đó mang tính định lượng hay định tính. Trong nhiều tình huống, các nhà nghiên cứu kết hợp cả hai phương pháp này để đi đến mô hình "nghiên cứu phương

pháp hỗn hợp" (mixed-methods research). Trong phần thảo luận dưới đây, chúng ta sẽ xem xét sự khác biệt giữa hai loại dữ liệu này cũng như những tình huống mà từng loại dữ liệu có thể được khai thác một cách hữu ích.

Dữ liệu định tính là gì?

Dữ liệu định tính liên quan đến các phán đoán mang tính mô tả, sử dụng các từ ngữ khái niệm thay vì những con số. Giới tính, tên quốc gia, loài động vật, trạng thái cảm xúc, cảm nhận và quan điểm là những ví dụ về thông tin định tính.

Dữ liệu định lượng là gì?

Dữ liệu định lượng liên quan đến một đại lượng có thể đo lường được, nghĩa là các con số được sử dụng. Một số ví dụ bao gồm nhiệt độ, thời gian, số tiền đã chi tiêu, số lượng đơn vị sản phẩm đã mua và số giờ sử dụng máy tính.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=50#h5p-7>

Đặc điểm	Nghiên cứu Định tính	Nghiên cứu Định lượng
Cơ sở thực hiện	Dựa trên thảo luận nhóm tập trung, phỏng vấn, nghiên cứu tình huống, ý kiến chuyên gia.	Dựa trên khảo sát, thăm dò ý kiến.
Dạng câu hỏi	Sử dụng các câu hỏi mở.	Sử dụng các câu hỏi đóng (có/không) và câu hỏi nhiều lựa chọn.
Mục đích	Giúp xây dựng/hình thành một lý thuyết để nghiên cứu.	Kiểm chứng và xác nhận một lý thuyết đã được xây dựng.
Trình bày kết quả	Kết quả được trình bày dưới dạng các chủ đề (văn bản).	Kết quả được trình bày dưới định dạng thống kê (con số, bảng biểu, đồ thị).
Quy mô mẫu	Cần ít đối tượng phản hồi hơn.	Cần nhiều đối tượng phản hồi.

Khả năng lặp lại	Khó lặp lại (để cho ra kết quả tương tự).	Dễ dàng lặp lại.
Tính phù hợp với dữ liệu nhạy cảm	Ít phù hợp hơn: đối tượng phản hồi có thể bị thiên kiến, quá thân thuộc hoặc có xu hướng làm rò rỉ thông tin.	Lý tưởng cho dữ liệu nhạy cảm vì có thể được ẩn danh và bảo mật.

Bảng 2: Sự khác biệt giữa nghiên cứu định tính và nghiên cứu định lượng

Nên thu thập loại dữ liệu nào?

Nhìn chung, việc lựa chọn dữ liệu liên quan trực tiếp đến câu hỏi nghiên cứu đang được xem xét. Dưới đây là một số hướng dẫn chung về loại thông tin được tạo ra với sự hỗ trợ của các phương pháp khác nhau:

Nghiên cứu định tính (Qualitative research)

Thường được sử dụng để thấu hiểu một hiện tượng, chẳng hạn như trải nghiệm của khách hàng tại một cửa hàng bán lẻ.

- Ví dụ: Một nhà nghiên cứu có thể chọn phỏng vấn một số người mua hàng bằng cách đặt các câu hỏi mở, ví dụ: *"Hôm nay quý khách có gặp phải vấn đề gì tại cửa hàng không?"* hoặc *"Quý khách có cảm thấy thoải mái khi mua sắm tại đây không?"* * Mục tiêu: Những câu hỏi này khuyến khích khách hàng trò chuyện thoải mái và nêu bật các vấn đề (ví dụ: khó khăn khi đỗ xe, nhân viên thô lỗ, hoặc hệ thống thông gió kém) mà đôi khi không thể xác định được chỉ qua một khảo sát định lượng.

Nghiên cứu định lượng (Quantitative research)

Hữu ích trong việc ước tính tần suất hoặc con số tổng thể.

- Ví dụ: Có bao nhiêu người đã đến nhà hàng tối nay?
- Mục tiêu: Phương pháp này cũng được sử dụng để kiểm tra hoặc xác nhận một giả thuyết – tức là một dự đoán có cơ sở của người quản lý.
 - *Ví dụ minh họa:* Một giáo viên có thể khảo sát lớp học của mình với câu hỏi: *"Trên thang điểm từ 1 đến 5, bạn cảm thấy buồn chán như thế nào trong lớp này? (với 1 = hoàn toàn không chán và 5 = rất chán)"*. Bằng cách phân tích các câu trả lời bằng số, giáo viên có thể kết luận rằng dự đoán của mình về sự buồn chán của học sinh là chính xác nếu điểm trung bình của cả lớp là 4.5.

Phương pháp hỗn hợp (Mixed method)

Như tên gọi cho thấy, phương pháp này sử dụng cả nghiên cứu định tính và định lượng. Một nhà nghiên cứu có thể sử dụng nghiên cứu định tính để hiểu rõ hơn về một hiện tượng, sau đó sử dụng nghiên cứu định lượng để kiểm định nó về mặt thống kê.

Ví dụ, trong khi phỏng vấn khách hàng, một nhà quản lý có thể nhận thấy rằng một số người trong số họ đang phàn nàn về các vấn đề đỗ xe. Để hiểu mức độ phổ biến của lời phàn nàn này, nhà quản lý có thể bắt đầu một cuộc khảo sát nhanh và ngắn để tìm hiểu xem có bao nhiêu khách hàng có nhu cầu đỗ xe tại khuôn viên của cửa hàng.

Ngược lại, một nhà nghiên cứu có thể tình cờ phát hiện ra một finding (kết quả nghiên cứu) bất ngờ thông qua một cuộc khảo sát (chẳng hạn như điểm số rất thấp cho dịch vụ khách hàng tại một cửa hàng ở Sydney). Các cuộc phỏng vấn có thể được tiến hành để tìm ra nguyên nhân gây ra sự không hài lòng của khách hàng. Quản lý cửa hàng có thể thấy việc bắt đầu các cuộc trò chuyện với khách hàng là hữu ích để hiểu rõ hơn về những điểm số thấp này.

Hạn chế của nghiên cứu định tính

- **Tính chủ quan (Subjectivity):** Dữ liệu định tính tồn tại dưới dạng văn bản viết hoặc lời nói và đòi hỏi sự diễn giải cẩn trọng. Trong một số tình huống, nhà nghiên cứu có thể không quen thuộc với ngôn ngữ của những người tham gia. Việc phân tích dữ liệu có thể đòi hỏi thời gian và sự giúp đỡ từ một người bản địa trong cộng đồng đó, người thông thạo ngôn ngữ liên quan. Ngay cả khi ngôn ngữ không phải là rào cản, việc diễn giải nội dung văn bản vẫn có thể được thực hiện khác nhau bởi các nhà nghiên cứu khác nhau.
- **Số lượng người tham gia ít hơn:** Phỏng vấn và thảo luận nhóm tập trung chỉ có thể được tiến hành với một nhóm nhỏ người tham gia. Trong khi các cuộc khảo sát có thể được gửi đi và hoàn tất bởi hàng trăm (hoặc hàng nghìn) người, thì việc thực hiện phỏng vấn với số lượng người cung cấp thông tin tương tự là điều không khả thi. Do đó, bất kỳ kết quả nào từ nghiên cứu định tính đều cần phải được khái quát hóa (generalize) một cách thận trọng.
- **Thiên kiến (Bias):** Nhà nghiên cứu hoặc nhà phân tích đóng vai trò lớn hơn nhiều trong việc diễn giải thông tin định tính so với việc chỉ đơn thuần chạy một bài kiểm tra thống kê. Đôi khi, người được phỏng vấn có thể đưa ra một điểm rất quan trọng nhưng chỉ đề cập một lần duy nhất – và thậm chí là đề cập một cách ngẫu nhiên. Trong khi một nhà nghiên cứu mới vào nghề dễ dàng bỏ lỡ hoàn toàn thông tin đó, thì một nhà nghiên cứu giàu kinh nghiệm hơn sẽ ghi chú lại nó.

Hạn chế của nghiên cứu định lượng

- **Quy mô mẫu lớn:** Để nghiên cứu định lượng có giá trị về mặt thống kê, một số lượng lớn đối tượng phản hồi phải được đưa vào cuộc khảo sát. Mặc dù không có con số chính xác nào được coi là chuẩn mực, nhưng quy mô mẫu

càng lớn thì độ tin cậy của các kết quả tìm được càng cao. Điều này đòi hỏi nguồn lực lớn về kỹ năng thiết kế dự án nghiên cứu, thời gian và tiền bạc.

- **Các câu hỏi được xác định trước:** Khi một cuộc khảo sát quy mô lớn được thực hiện, tất cả các câu hỏi trong khảo sát đó cần phải được xác định trước. Trong trường hợp nhà nghiên cứu muốn bổ sung thêm bất kỳ câu hỏi nào, họ không thể thực hiện điều đó trong một khoảng thời gian ngắn.

7. CÁC LOẠI THIẾT KẾ NGHIÊN CỨU (TYPES OF RESEARCH DESIGN)

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Giải thích được ba loại thiết kế nghiên cứu được các nhà tiếp thị sử dụng.
- Hiểu rõ việc ứng dụng của các loại thiết kế này.

Thiết kế nghiên cứu

Chương này xem xét các loại thiết kế nghiên cứu được các nhà tiếp thị sử dụng. Một thiết kế nghiên cứu là một kế hoạch hoặc cấu trúc tổng thể cho một dự án nghiên cứu. Một thiết kế nghiên cứu sẽ sử dụng các sự kết hợp khác nhau của dữ liệu sơ cấp, thứ cấp, định tính và định lượng. Tùy thuộc vào các câu hỏi nghiên cứu tổng thể, các thiết kế nghiên cứu trong tiếp thị có thể rơi vào một trong ba danh mục sau:

- **Thiết kế nghiên cứu khám phá (Exploratory research design)**
- **Thiết kế nghiên cứu mô tả (Descriptive research design)**
- **Thiết kế nghiên cứu nhân quả / thực nghiệm (Causal research design - experiments)**

Thiết kế nghiên cứu khám phá mang tính chất không chính thức và ít cấu trúc hơn so với hai loại thiết kế còn lại. Thiết kế khám phá được sử dụng để tìm hiểu một tình huống, đặc biệt là khi nhà nghiên cứu đang ở trong một lĩnh vực chưa quen thuộc. Hầu hết các dự án học thuật đều bắt đầu bằng nghiên cứu khám phá khi các nhà nghiên cứu thực hiện nghiên cứu tại bàn (desk research).

Tương tự, khi các nhà tiếp thị có kế hoạch mạo hiểm vào các thị trường mới, chẳng hạn như Ấn Độ, việc áp dụng thiết kế nghiên cứu khám phá là điều nên làm. Điều này có thể bao gồm việc xem xét các nghiên cứu tình huống, truy cập các báo cáo đã xuất bản (tức là dữ liệu thứ cấp) về thị trường đó, thực hiện các cuộc phỏng vấn kinh nghiệm với các chuyên gia và tiến hành các nhóm thảo luận tập trung – nếu cần thiết.

Thiết kế khám phá hữu ích trong việc thu thập thông tin nền tảng và quyết định các phương pháp nghiên cứu trong tương lai. Nó có thể tạo ra nhiều câu hỏi hơn, những câu hỏi này sẽ cần được giải quyết bằng các loại thiết kế nghiên cứu khác.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=52#h5p-8>

Thiết kế nghiên cứu mô tả (Descriptive research design)

Đúng như tên gọi, thiết kế nghiên cứu mô tả được sử dụng để mô tả các đặc điểm của thị trường hoặc của đối tượng phản hồi. Loại thiết kế nghiên cứu này tạo ra các thông tin định lượng. Do đó, một thiết kế như vậy thường bao gồm các cuộc khảo sát.

Khảo sát rất hữu ích để đo lường các con số mô tả liên quan đến nhóm tuổi, mức thu nhập, mô hình chi tiêu và thậm chí là thái độ của đối tượng phản hồi. Điều này có thể được thực hiện bằng cách sử dụng các thang đo phù hợp, chẳng hạn như: “Trên thang điểm từ 1–5, bạn hài lòng như thế nào với dịch vụ của tổ chức này?”.

Các phép đo sinh lý, chẳng hạn như những phản ứng không tự nguyện của con người (như nhịp tim, các thay đổi về da và chuyển động mắt) đối với các kích thích tiếp thị, ví dụ như một quảng cáo, cũng có thể được xếp vào nhóm nghiên cứu mô tả. Mặc dù việc ghi lại những thông tin này đòi hỏi các thiết bị đặc dụng, nhưng đây vẫn là một loại thông tin mô tả có thể hữu ích cho các nhà tiếp thị.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=52#h5p-9>

Một thiết kế nghiên cứu mô tả có thể là nghiên cứu cắt lớp (*cross-sectional study*) được thực hiện tại một thời điểm nhất định (chẳng hạn như vào tháng 3 năm 2022) hoặc một nghiên cứu theo chiều dọc (*longitudinal study*) được thực hiện với cùng một nhóm đối tượng phản hồi – lặp đi lặp lại – trong một khoảng thời gian nhất định. Do đó, một cuộc khảo sát theo chiều dọc có thể được thực hiện với các sinh viên WSU vào tháng 4, tháng 5 và tháng 6 để theo dõi mức độ hài lòng của sinh viên đối với một môn học trong suốt một học kỳ.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=52#oembed-1>

Thiết kế nghiên cứu nhân quả (Causal research design)

Thiết kế nghiên cứu nhân quả (còn được gọi là thiết kế nghiên cứu thực nghiệm) xem xét các mối quan hệ nguyên nhân và kết quả. Một thí nghiệm được thiết kế tốt là cách tốt nhất để hiểu một biến số (ví dụ: quảng cáo) có thể ảnh hưởng đến một biến số khác (ví dụ: doanh số bán sản phẩm) như thế nào. Một thí nghiệm bao gồm một hoặc nhiều **biến độc lập** (ví dụ: mức giá và các tính năng của sản phẩm), những biến này được tác động (thao túng) để xác định xem chúng có thể ảnh hưởng như thế nào đến một hoặc nhiều **biến phụ thuộc** (chẳng hạn như sở thích của khách hàng hoặc sự hài lòng của khách hàng). Nhà nghiên cứu có thể "thao túng" các biến độc lập, trong khi các biến phụ thuộc là những biến bị ảnh hưởng do những thay đổi của các biến độc lập đó.

Các thí nghiệm có thể được phân loại thành thực nghiệm hiện trường (field experiments) hoặc thực nghiệm trong phòng thí nghiệm (lab experiments). Khi một thí nghiệm được tiến hành trong môi trường tự nhiên – chẳng hạn như tại một cửa hàng bán lẻ – nó được gọi là **thực nghiệm hiện trường**. Mặt khác, các thí nghiệm được tiến hành trong văn phòng của nhà nghiên cứu hoặc lớp học đại học là các **thực nghiệm trong phòng thí nghiệm**. Một thực nghiệm trong phòng thí nghiệm có thể được thực hiện để đo lường tác động của một quảng cáo đối với thái độ của đối tượng nghiên cứu.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=52#h5p-10>

Thực nghiệm hiện trường thường được xem là cung cấp cho các nhà nghiên cứu những kết quả đáng tin cậy vì nó được thực hiện trong môi trường thực tế. Tuy nhiên, việc triển khai một thực nghiệm hiện trường không hề dễ dàng. Rất khó để kiểm soát các yếu tố khác – trong môi trường thực – vốn cũng có thể tác động đến kết quả cuối cùng. Hơn nữa, việc vận hành một thí nghiệm đòi hỏi chuyên môn trong thiết kế thực nghiệm, cũng như yêu cầu các nguồn lực về tài chính và thời gian.

Lấy ví dụ, McDonald's có thể muốn xem liệu việc giảm giá bánh cheeseburger có dẫn đến doanh số bán hàng cao hơn hay không. Nhà hàng thức ăn nhanh này có thể giảm giá tại một địa phương cụ thể, chẳng hạn như Fairfield. Họ cũng sẽ phải đo lường doanh số một cách tỉ mỉ xung quanh thời điểm giảm giá. Khi diễn giải kết quả, họ cần

đảm bảo rằng bất kỳ thay đổi nào trong số liệu doanh số đều có thể được quy cho sự thay đổi giá – chứ không phải do các yếu tố môi trường khác, chẳng hạn như việc tăng giá tại Hungry Jacks hoặc lượng khách hàng tăng tổng thể do một trận bóng đá địa phương.

Vì vậy, việc kiểm soát các **biến ngoại lai (extraneous variables)** như vậy – tức là tất cả những biến số ngoài các biến độc lập đã được xác định nhưng cũng có thể ảnh hưởng đến biến phụ thuộc – là vô cùng quan trọng.

Ví dụ về Thiết kế thực nghiệm

Một trong những thiết kế nghiên cứu thực nghiệm phổ biến là thiết kế Kiểm chứng "Trước-Sau" (Before-After Testing). Đúng như tên gọi, mô hình này đo lường biến phụ thuộc trước và sau khi có sự thay đổi trong biến độc lập. Ví dụ dưới đây sẽ giúp làm rõ khái niệm này:

Mô hình Thực nghiệm

Nhóm thực nghiệm (Experimental Group): $(R) \quad O_1 \quad X \quad O_2$

Nhóm đối chứng (Control Group): $(R) \quad O_3 \quad O_4$

Hiệu ứng Thực nghiệm (Experimental Effect - E): $E = (O_2 - O_1) - (O_4 - O_3)$

Trong đó:

- **R (Random allocation):** Phân bổ ngẫu nhiên các đối tượng nghiên cứu vào nhóm thực nghiệm hoặc nhóm đối chứng.
- **O (Observation):** Quan sát hoặc đo lường biến phụ thuộc.

- **X (Treatment):** Sự tác động hoặc thao túng biến độc lập (ví dụ: thay đổi giá, chạy quảng cáo mới).

Giải thích cho ví dụ trên:

1. Các đối tượng hoặc người tham gia nghiên cứu được phân bổ ngẫu nhiên vào một trong các nhóm.
2. Nhóm thực nghiệm (experimental group) là nhóm mà tại đó biến độc lập (ví dụ: giá cả) được thao túng/điều chỉnh (ví dụ: giảm xuống).
3. Nhóm kiểm soát (control group) không phải chịu bất kỳ sự thay đổi nào của biến độc lập.
4. Các phép đo lường (tức là O) đối với biến phụ thuộc (ví dụ: doanh số bán hàng) – cho cả hai nhóm – được thực hiện trước khi có sự thay đổi về giá (X) – ĐO LƯỜNG TRƯỚC TÁC ĐỘNG (PRE-TEST).
5. Các phép đo lường (tức là O) đối với biến phụ thuộc được thực hiện sau khi có sự thay đổi về giá – ĐO LƯỜNG SAU TÁC ĐỘNG (POST-TEST).
6. Sự chênh lệch giữa O2 và O1 chứng minh sự thay đổi của biến phụ thuộc là do sự thay đổi của X (biến độc lập).
7. Sự chênh lệch giữa O4 và O3 chứng minh bất kỳ sự thay đổi nào của biến phụ thuộc là do các yếu tố khác ngoài X gây ra.
8. Do đó, sự chênh lệch giữa hai nhóm (O2 – O1) – (O4 – O3) chứng minh tác động thực sự của biến độc lập vì nó đã loại bỏ được mọi ảnh hưởng của các biến ngoại lai (extraneous variables).

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=52#h5p-11>

8. THẢO LUẬN NHÓM TẬP TRUNG (FOCUS GROUPS)

Chương này được điều chỉnh từ [Trung tâm Community Tool Box](#) và được sử dụng theo [Giấy phép Creative Commons Attribution-Noncommercial-Share Alike 3.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Định nghĩa và áp dụng khái niệm thảo luận nhóm tập trung vào một dự án nghiên cứu.
- Giải thích các giai đoạn khác nhau trong việc tổ chức một buổi thảo luận nhóm tập trung.
- Đánh giá nhu cầu thực hiện thảo luận thông qua các nhóm tập trung.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=55#oembed-1>

Thảo luận nhóm tập trung là gì?

Thảo luận nhóm tập trung (focus group) là một cuộc thảo luận nhóm nhỏ được dẫn dắt bởi một người điều phối (facilitator) đã qua đào tạo. Phương pháp này được sử dụng để tìm hiểu ý kiến của những người tham gia về một chủ đề được chỉ định, và để định hướng cho các hành động trong tương lai.

Các ví dụ:

- Một nhóm thảo luận tập trung gồm phụ huynh có con ở độ tuổi mẫu giáo họp mặt để thảo luận về nhu cầu chăm sóc trẻ em. Các phụ huynh chia sẻ quan điểm của họ về các chương trình chăm sóc trẻ tại địa phương và những gì có thể được thực hiện để cải thiện chúng.
- Một nhóm thảo luận tập trung gồm các công dân cao tuổi họp mặt tại trung tâm người cao tuổi mới. họ nghĩ gì về các chương trình đang được cung cấp? Những gợi ý và ý tưởng của riêng họ là gì?
- Một cơ quan muốn mở một nhà lưu trú tập thể (group home) cho những người trưởng thành bị khuyết tật về phát triển tại một khu dân cư yên tĩnh. Cơ quan này triệu tập một nhóm các láng giềng tiềm năng. Những mối quan ngại của họ là gì? Liệu dự án này có khả thi không?



Thảo luận nhóm tập trung khác với các nhóm thông thường như thế nào?

Một nhóm thảo luận tập trung khác với bất kỳ nhóm chung nào ở ba khía cạnh cơ bản sau:

- **Sự khác biệt chính** là nhóm có một chủ đề thảo luận cụ thể và tập trung.
- Nhóm có một **người dẫn dắt hoặc người điều phối (facilitator)** đã qua đào tạo.
- Thành phần của nhóm và cuộc thảo luận nhóm được lập kế hoạch cẩn thận để tạo ra một **môi trường không bị đe dọa**, nơi mọi người tự do nói chuyện một cách cởi mở. Các thành viên được khuyến khích tích cực bày tỏ quan điểm của mình.

Tại sao thảo luận nhóm tập trung được sử dụng?

Các phản hồi trong một nhóm thảo luận tập trung thường tồn tại dưới dạng lời nói, có tính chất mở, tương đối rộng và mang tính định tính. Chúng có **độ sâu, sự sắc thái và tính đa dạng** cao hơn.

Thông qua phương pháp này, nhà nghiên cứu còn có thể quan sát được các **giao tiếp phi ngôn ngữ** (cử chỉ, điệu bộ) và **sự tương tác giữa các thành viên** trong nhóm. Do đó, thảo luận nhóm tập trung có thể tiếp cận gần hơn với những gì mọi người thực sự đang suy nghĩ và cảm nhận, mặc dù các câu trả lời của họ có thể khó định lượng hoặc khó chấm điểm trên một thang đo cụ thể.

Khi nào nên sử dụng thảo luận nhóm tập trung?

- **Khi bạn đang cân nhắc việc giới thiệu một chương trình hoặc dịch vụ mới:** Giúp thu thập phản hồi sớm để điều chỉnh sản phẩm trước khi tung ra thị trường.
- **Khi bạn muốn đặt những câu hỏi khó có thể hỏi hoặc trả lời dễ dàng trên một bản khảo sát viết:** Những vấn đề phức tạp đòi hỏi sự giải thích và đào sâu suy nghĩ.
- **Khi bạn muốn bổ sung cho những kiến thức thu được từ các cuộc khảo sát viết:** Giúp giải mã lý do "tại sao" đằng sau các con số định lượng từ khảo sát.
- **Khi bạn quen biết hoặc có thể tìm được một người điều phối nhóm có kinh nghiệm và kỹ năng:** Kỹ năng của người dẫn dắt là yếu tố then chốt quyết định sự thành công của buổi thảo luận.
- **Khi bạn có đủ thời gian, kiến thức và nguồn lực để tuyển chọn một nhóm người tham gia sẵn lòng đóng góp ý kiến:** Việc tìm đúng đối tượng mục tiêu là một phần quan trọng trong thiết kế nghiên cứu.

Ưu và nhược điểm của thảo luận nhóm

Ưu điểm: Một lợi thế của thảo luận nhóm tập trung là **độ sâu và tính phức hợp** của các câu trả lời như đã đề cập trước đó. Các thành viên trong nhóm thường có thể kích thích những suy nghĩ mới cho nhau – điều mà có thể sẽ không xảy ra nếu thực hiện phỏng vấn riêng lẻ. Nhà nghiên cứu có khả năng thu thập một lượng lớn thông tin từ một nhóm người trong một khoảng thời gian tương đối ngắn.

Hạn chế: Tuy nhiên, việc triển khai các nhóm tập trung cũng có một số hạn chế:

- **Thời gian:** Phương pháp này thường tiêu tốn nhiều thời gian trên mỗi đối tượng phản hồi hơn so với khảo sát cá nhân — bởi việc tuyển chọn nhóm và tổ chức buổi thảo luận mất khá nhiều công sức.
- **Tâm lý e ngại:** Một số thành viên có thể cảm thấy do dự khi nói chuyện cởi mở trước đám đông.
- **Yêu cầu chuyên môn:** Người điều phối (moderator) cần phải am hiểu về động lực nhóm (group dynamics). Việc thuê những người điều phối giỏi thường rất tốn kém.
- **Tính nhạy cảm:** Rất quan trọng là các chủ đề nhạy cảm (ví dụ: tôn giáo, xu hướng tính dục hoặc tình trạng bệnh lý cá nhân) có thể không thu được dữ liệu tốt nếu thảo luận thông qua nhóm tập trung.

Dưới đây là một ví dụ về bảng câu hỏi sàng lọc (screening questionnaire):

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=55#h5p-12>

Quy trình thực hiện thảo luận nhóm tập trung

Phần A: Trước khi bắt đầu

Thành phần nhóm và nội dung thảo luận cần được lập kế hoạch cẩn thận nhằm tạo ra một môi trường không gây áp lực, giúp những người tham gia cảm thấy thoải mái để nói chuyện cởi mở và đưa ra những ý kiến trung thực.

Xác định mục tiêu của bạn

Để bắt đầu, bạn cần trả lời rõ ràng các câu hỏi cốt lõi sau:

- **"Tại sao tôi muốn thực hiện buổi thảo luận nhóm tập trung này?"** (Xác định lý do lựa chọn phương pháp này thay vì khảo sát hay phỏng vấn sâu).
- **"Tại sao tôi lại làm việc này?"** (Xác định mục đích kinh doanh hoặc học thuật đằng sau dự án).
- **"Tôi hy vọng sẽ học hỏi/thu thập được điều gì?"** (Xác định các kết quả cụ thể hoặc thông tin chi tiết cần khám phá).

Xem xét các phương pháp khác

Bạn có kế hoạch sử dụng thêm các phương pháp khác để tìm hiểu về ý kiến của đối tượng nghiên cứu hay không? Nói cách khác, cho đến thời điểm này: **Hãy suy nghĩ trước khi bắt đầu, cân nhắc kỹ trước khi hành động.**

- **Nếu có:** Bạn sẽ sử dụng những phương pháp nào và tại sao? (Ví dụ: Kết hợp khảo sát định lượng để kiểm chứng các phát hiện định tính).
- **Nếu không:** Đây có phải là phương pháp duy nhất và tốt nhất để tìm ra điều bạn muốn không?

Tìm kiếm một người dẫn dắt / người điều phối giỏi (Moderator)

Một người điều phối nhóm tập trung giỏi thường hội đủ các phẩm chất sau:

- **Có kinh nghiệm điều phối các nhóm:** Khả năng quản lý động lực nhóm và xử lý các tình huống phát sinh.
- **Có hiểu biết về chủ đề nghiên cứu HOẶC sẵn lòng nghiên cứu về chủ đề đó:** Để có thể đặt những câu hỏi đào sâu (probing questions) một cách thông minh.

- **Có khả năng kết nối tốt với những người tham gia:** Tạo dựng niềm tin và sự thoải mái để họ chia sẻ thật lòng.
- **Sẵn sàng hợp tác chặt chẽ với bạn:** Để đảm bảo mang lại những kết quả đầu ra đúng như mục tiêu mong muốn.

Tìm người ghi chép và bổ nhiệm quan sát viên

Bạn phải đảm bảo rằng cuộc thảo luận được ghi lại một cách đầy đủ. Cần có người ghi chép lại những gì được nói, tương tự như việc ghi biên bản cuộc họp. Hãy sắp xếp việc này trước.

Ngoài ra, bạn có thể ghi âm cuộc thảo luận với sự cho phép của nhóm. Cách này sẽ tốn nhiều thời gian hơn để chuyển soạn (transcribe) từ âm thanh sang văn bản và diễn giải bản chuyển soạn đó, nhưng bạn sẽ có một bản ghi chép đầy đủ, chính xác và lưu trữ lâu dài hơn. Trong một số phiên thảo luận, việc kết hợp cả quan sát viên và thiết bị ghi âm là rất hữu ích.

Quyết định đối tượng mời tham gia (Người tham gia nhóm tập trung)

Lý tưởng nhất, những người được mời nên là một **mẫu đại diện (representative sample)** cho nhóm đối tượng mà bạn đang quan tâm đến ý kiến của họ.

Giả sử bạn quan tâm đến ý kiến của các chủ cửa hàng trên Phố Chính. Hãy lập một danh sách đầy đủ. Sau đó, chọn ra một nhóm đại diện, chẳng hạn như phân loại theo quy mô, loại hình kinh doanh, hoặc liệu họ là chủ sở hữu địa phương hay từ nơi khác đến. Có lẽ bạn muốn lắng nghe ý kiến từ tất cả các loại hình kinh doanh; vì vậy hãy đảm bảo bạn thực hiện được điều đó.

Một **bảng câu hỏi sàng lọc (screening questionnaire)** có thể hỗ trợ nhà nghiên cứu trong việc đảm bảo rằng chỉ những người đáp ứng đủ các **tiêu chuẩn lựa chọn (inclusion criteria)** mới được tuyển dụng.

Quyết định về các hình thức khuyến khích (Incentives)

Bạn có nên đưa ra một hình thức khuyến khích nào đó để mọi người tham gia không? Có thể mọi người đến chỉ vì họ muốn giúp đỡ, hoặc vì họ nghĩ rằng họ sẽ được gặp gỡ những người thú vị khác, học hỏi được điều gì đó, hoặc đơn giản là để chung vui. Đôi khi sự mới lạ của chính trải nghiệm này đã là một động lực. Và có thể tất cả những lý do này đều đúng.

Nhưng có lẽ những lý do đó là chưa đủ và cần đến một hình thức khuyến khích khác. **Tiền mặt** là một phương án; đôi khi các thành viên nhóm tập trung được trả phí, dù chỉ là một khoản nhỏ. Nếu ngân sách cho phép, hãy cân nhắc điều này. Nếu không, hãy nghĩ đến các hình thức khuyến khích khác: đồ ăn và thức uống (nhiều hơn là chỉ khoai tây chiên và nước ngọt); sự công nhận công khai; một món quà mang về; hoặc một cơ hội đào tạo sau này.

Quyết định các chi tiết cụ thể của cuộc họp

Cụ thể, hãy chốt các thông tin này trước khi bạn bắt đầu mời mọi người đăng ký:

- Ngày nào?
- Địa điểm nào?
- Thời gian nào?
- Kéo dài bao lâu?
- Có bao nhiêu nhóm?

Chuẩn bị bộ câu hỏi của bạn

Hãy tham gia với một sự chuẩn bị kỹ lưỡng. Viết ra trước một danh sách các chủ đề và câu hỏi mà bạn muốn hỏi. Điều này sẽ đóng vai trò là kim chỉ nam cho bạn. Dưới đây là một số ví dụ về các câu hỏi chung. Chúng chủ yếu áp dụng cho các nhóm đang thảo luận về một chương trình hoặc dịch vụ hiện tại, nhưng cũng có thể được điều chỉnh cho các

chương trình dự kiến, cũng như cho các nhóm đang giải quyết những vấn đề khác. Ngôn từ chính xác và thứ tự trình bày sẽ phụ thuộc vào chủ đề và nhóm của bạn, nhưng một số câu hỏi trong số này có thể được tùy chỉnh để phù hợp với nhu cầu riêng của bạn.

- "Bạn có suy nghĩ gì về những điều đang diễn ra hiện nay?"
- "Bạn có thể nói rằng mình hài lòng với tình hình hiện tại, với cách mọi việc đang diễn ra không?"
- (Nếu có) "Bạn hài lòng về điều gì? Tại sao lại như vậy?" (Hoặc, "Điều gì đang diễn ra tốt đẹp...?")
- "Có điều gì khiến bạn không hài lòng và muốn thấy sự thay đổi không?" (Hoặc, "Điều gì đang diễn ra không tốt...?")
- (Nếu có) "Đó là những điều gì? Tại sao lại như vậy? Chúng nên được thay đổi như thế nào? Bạn muốn thấy những điều gì sẽ diễn ra?"
- "Còn về khía cạnh cụ thể này (của chủ đề) thì sao. Bạn nghĩ gì về điều đó?"
- Lặp lại đối với các khía cạnh khác nhau của chủ đề, với sự biến tấu trong phong cách. Ví dụ: nếu chủ đề chính của nhóm tập trung là "trị an cộng đồng" (community policing), một số khía cạnh then chốt cần đề cập có thể là sự hiện diện/độ nhận diện, tính nhạy cảm, sự tương tác, sự tôn trọng, v.v.
- "Một số người nói rằng một cách để cải thiện X là làm Y.
- Bạn có đồng ý với điều này không?" (Hoặc, "Bạn cảm thấy thế nào về điều đó?")
- "Bạn có bất kỳ đề xuất nào khác, hoặc những gợi ý nào bạn muốn đưa ra không?"
- "Có điều gì khác bạn muốn chia sẻ trước khi chúng ta kết thúc (wind up) không?"
- Một số câu hỏi "đào sâu/thăm dò" (probes) hoặc "tiếp nối" (follow-ups) được thiết kế để khai thác thêm thông tin cho một câu hỏi nhất định:
 - "Bạn có thể nói thêm về điều đó không?"
 - "Bạn có thể đưa ra một ví dụ không?"
 - "Jane nói X. Còn những người khác thì sao. Các bạn nghĩ gì?"
 - "Còn bạn thì sao, Joe. [Hoặc, "các bạn ở góc đằng kia...."] Các bạn có suy nghĩ gì về điều này không?"

- "Có ai khác có suy nghĩ gì về điều đó không?"

Tuyển dụng thành viên (Recruit your members)

Hãy gọi điện, gửi email hoặc trực tiếp tìm kiếm họ. Hiện nay có rất nhiều bên thứ ba (các công ty nghiên cứu thị trường) sẵn sàng hỗ trợ các nhà nghiên cứu trong việc xác định và kết nối với đúng nhóm đối tượng người tham gia mục tiêu.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=55#h5p-13>

Ví dụ về một bảng câu hỏi sàng lọc để lựa chọn người tham gia:

Phần B: Trong buổi thảo luận

Điều hành nhóm

Một trình tự sự kiện phổ biến cho nhiều buổi thảo luận nhóm tập trung thường diễn ra như sau (Người dẫn dắt thường chịu trách nhiệm thực hiện các bước này):

- **Cảm ơn mọi người đã đến tham dự.**
- **Điểm lại mục đích của nhóm và mục tiêu của buổi họp:** Giới thiệu bối cảnh để người tham gia hiểu rõ vai trò của họ.

- **Thông báo về lộ trình buổi họp:** Cách thức tiến hành và cách các thành viên có thể đóng góp. Thiết lập các **quy tắc chung (ground rules)** và khuyến khích sự tham gia cởi mở.
- **Thiết lập tông điệu (Set the tone):** Điều này rất quan trọng vì có thể ít thành viên nào từng tham gia nhóm tập trung trước đây. Hãy tạo sự thoải mái và tin cậy.
- **Đặt câu hỏi mở đầu:** Đây có thể là một câu hỏi rất chung chung (“Suy nghĩ tổng quát của anh/chị về X là gì?”) hoặc một câu hỏi cụ thể hơn. Cả hai lựa chọn đều hợp lý và cả hai loại câu hỏi này đều có thể được hỏi trước khi buổi thảo luận kết thúc.
- **Đảm bảo mọi ý kiến về câu hỏi đó đều có cơ hội được lắng nghe.**

Một số kỹ thuật phổ biến để khơi gợi ý kiến:

- Tóm tắt những gì bạn nghĩ mình đã nghe được và hỏi xem nhóm có đồng ý hay không.
- Diễn đạt cùng một câu hỏi theo một cách khác.
- Hỏi xem còn ai có ý kiến gì về câu hỏi đó không.
- Đặt một câu hỏi đào sâu (follow-up).
- Quan sát quanh phòng và thiết lập tương tác bằng mắt nhanh, đặc biệt là với những người chưa phát biểu.

Lưu ý quan trọng: Nhiệm vụ của người dẫn dắt là khơi gợi ý kiến chứ không phải phán xét ý kiến đó. Mọi quan điểm đều cần được tôn trọng.

- **Đặt câu hỏi tiếp theo** và tiến hành với các câu hỏi khác theo cách tương tự. Việc diễn đạt câu hỏi, các câu hỏi đào sâu, thứ tự câu hỏi và thời gian dành cho

mỗi câu là những điểm mà người dẫn dắt sẽ phải quyết định — đôi khi là ngay tại chỗ. Một người dẫn dắt giàu kinh nghiệm sẽ làm tốt việc này. Đó là lý do tại sao bạn cần dành thời gian tìm kiếm họ!

- **Khi tất cả các câu hỏi đã được đặt ra**, và trước khi nhóm kết thúc, hãy hỏi xem còn ai có ý kiến nào khác không. Đây là cách hiệu quả để thu thập những quan điểm chưa được nêu ra trước đó.
 - **Thông báo cho các thành viên về các bước tiếp theo** và những gì họ có thể mong đợi sau buổi thảo luận này.
 - **Đừng quên cảm ơn nhóm một lần nữa vì đã dành thời gian tham dự!**
-

Phần C: Sau buổi thảo luận

Xem xét dữ liệu

Nếu bạn có ghi âm, hãy thực hiện chuyển soạn thành văn bản (transcript). Nếu không, hãy lập một bản tóm tắt viết tay từ các ghi chú của nhóm. Trong bất kỳ trường hợp nào, hãy xem xét kỹ lưỡng thông tin bạn đã thu thập được.

Trong một số trường hợp, bạn có thể thiết lập và sử dụng một hệ thống mã hóa (coding system) để "chấm điểm" dữ liệu và đếm số lần một chủ đề cụ thể được nhắc đến. Kinh nghiệm sẽ giúp ích rất nhiều ở bước này. Nhưng dù bạn có làm điều này hay không, hãy cố gắng có ít nhất hai người xem xét kết quả một cách độc lập (bởi ngay cả những người giỏi nhất cũng có những thiên kiến riêng). Sau đó, hãy ngồi lại với nhau để so sánh các cách diễn giải và kết luận của mình.

- Những hình mẫu (patterns) nào xuất hiện?
- Những chủ đề chung là gì?
- Những câu hỏi mới nào nảy sinh?
- Những kết luận nào có vẻ xác thực?

Chia sẻ kết quả với nhóm

Họ đã dành thời gian cho bạn. Điều tối thiểu bạn có thể làm là gửi cho họ một số phản hồi — đây là nghĩa vụ của bạn. Việc này có thể được thực hiện qua thư, điện thoại hoặc email. Đôi khi, bạn thậm chí có thể mời nhóm quay lại phiên thảo luận thứ hai để xem xét kết quả, xác minh tính chính xác và/hoặc khám phá các chủ đề khác.

Lưu ý: Có lẽ hiện tại các thành viên đã trở nên quan tâm hơn đến vấn đề này và muốn tham gia sâu hơn. Hãy cân nhắc việc trao cho họ cơ hội để thực hiện điều đó. Một nhóm thảo luận tập trung, một cách gián tiếp, có thể trở thành một công cụ tuyển dụng (nhân sự hoặc cộng tác viên).

Sử dụng kết quả

Thu thập thông tin hữu ích chính là lý do khiến bạn muốn thực hiện thảo luận nhóm tập trung ngay từ đầu. Bây giờ bạn có cơ hội, và có lẽ là cả trách nhiệm, để đưa nó vào sử dụng. Bạn có thể cải thiện tình huống vốn đã thúc đẩy và khiến bạn nghĩ về việc tổ chức nhóm tập trung ngay từ lúc bắt đầu.

Ghi danh truyền thông (Media Attributions):

- Ảnh Nhóm sinh viên © [Sally Tsoutsas - Nhiếp ảnh gia WSU](#), được cấp phép theo giấy phép [CC BY-NC-ND \(Ghi công - Phi thương mại - Không phái sinh\)](#).

9. NGHIÊN CỨU QUAN SÁT (OBSERVATIONAL RESEARCH)

Chương này được điều chỉnh từ phần [6.5 Nghiên cứu Quan sát](#) của Paul C. Price, Rajiv Jhangiani, I-Chant A. Chiang, Dana C. Leighton, & Carrie Cuttler, được cấp phép theo Giấy phép Quốc tế [Creative Commons Attribution-NonCommercial-ShareAlike 4.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Giải thích các loại phương pháp nghiên cứu quan sát khác nhau.
- Lựa chọn loại phương pháp phù hợp cho một tình huống nghiên cứu cụ thể.



Nghiên cứu quan sát là gì?

Thuật ngữ nghiên cứu quan sát được dùng để chỉ một số loại nghiên cứu phi thực nghiệm khác nhau, trong đó hành vi được quan sát và ghi lại một cách hệ thống. Mục tiêu của nghiên cứu quan sát là mô tả một biến số hoặc một tập hợp các biến số. Nói một cách tổng quát hơn, mục tiêu là để có được một "bức ảnh chụp nhanh" (snapshot) về các đặc điểm cụ thể của một cá nhân, một nhóm hoặc một bối cảnh.

Như đã mô tả trước đó, nghiên cứu quan sát có tính chất **phi thực nghiệm** vì không có yếu tố nào bị thao túng hoặc kiểm soát; do đó, chúng ta không thể đưa ra các kết luận nhân quả khi sử dụng phương pháp này. Dữ liệu thu thập được thường mang tính định tính, nhưng cũng có thể là định lượng hoặc cả hai (phương pháp hỗn hợp).

Quan sát tự nhiên so với Quan sát dàn dựng

1. Quan sát tự nhiên (Naturalistic Observation)

Đây là phương pháp quan sát hành vi của mọi người trong môi trường mà hành vi đó thường xuyên xảy ra. Do đó, quan sát tự nhiên là một loại **nghiên cứu hiện trường** (ngược lại với nghiên cứu trong phòng thí nghiệm).

- **Ví dụ kinh điển:** Nghiên cứu nổi tiếng của Jane Goodall về loài tinh tinh. Bà đã dành ba thập kỷ quan sát chúng trong môi trường tự nhiên ở Đông Phi để nghiên cứu cấu trúc xã hội, mô hình giao phối và vai trò giới.
- **Ví dụ trong kinh doanh:** Quan sát người mua hàng trong cửa hàng tạp hóa, trẻ em trên sân chơi trường học, hoặc bệnh nhân tâm thần trong khoa điều trị.

Quan sát tự nhiên ngụy trang (Disguised naturalistic observation): Các nhà nghiên cứu thường thực hiện quan sát một cách kín đáo nhất có thể để người tham gia không biết họ đang bị nghiên cứu.

- **Tính đạo đức:** Phương pháp này được coi là chấp nhận được nếu người tham gia được giữ ẩn danh và hành vi diễn ra ở nơi công cộng – nơi mọi người thường không kỳ vọng về quyền riêng tư (ví dụ: cho hàng hóa vào giỏ siêu thị).
- **Tranh cãi:** Việc quan sát "hành vi trong nhà vệ sinh" bị coi là vi phạm đạo đức vì mọi người có kỳ vọng hợp lý về quyền riêng tư ở đó, ngay cả khi đó là nhà vệ sinh công cộng.

2. Quan sát dàn dựng (Contrived Observation)

Hình thức này diễn ra trong một môi trường nhân tạo, chẳng hạn như trong phòng thí nghiệm.

- **Ví dụ:** Nhà nghiên cứu muốn đo lường phản ứng sinh lý của mọi người đối với một quảng cáo đang được trình chiếu trong lớp học của trường đại học.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=60#oembed-1>

Quan sát không công khai và quan sát công khai (Disguised vs undisguised observation)

Trong những trường hợp việc tiến hành quan sát tự nhiên không công khai (disguised naturalistic observation) là không hợp đạo đức hoặc không khả thi, các nhà nghiên cứu có thể tiến hành quan sát tự nhiên công khai (undisguised naturalistic observation), nơi những người tham gia được thông báo về sự hiện diện của nhà nghiên cứu và việc hành vi của họ đang bị theo dõi. Tuy nhiên, một mối lo ngại đối với quan sát tự nhiên công khai là tính phản ứng (reactivity). Tính phản ứng đề cập đến việc một phép đo lường làm thay đổi hành vi của người tham gia. Trong trường hợp quan sát tự nhiên công khai, mối lo ngại về tính phản ứng là khi mọi người biết họ đang bị quan sát và nghiên cứu, họ có thể hành xử khác với cách họ vẫn làm bình thường. Ví dụ, bạn có thể hành xử rất khác trong một quán bar nếu bạn biết rằng có ai đó đang quan sát và ghi lại các hành vi của bạn, và điều này sẽ làm mất đi độ giá trị (invalidate) của nghiên cứu. Vì vậy, quan sát không công khai ít xảy ra tính phản ứng hơn và do đó có thể đạt được độ giá trị (validity) cao hơn vì mọi người không nhận thức được rằng các hành vi của họ đang bị quan sát và ghi lại. Tuy nhiên, hiện nay chúng ta biết rằng mọi người thường trở nên quen với việc bị quan sát và theo thời gian, họ bắt đầu hành xử một cách tự nhiên trước sự hiện diện của nhà nghiên cứu. Nói cách khác, theo thời gian, mọi người dần thích nghi (habituate) với việc bị quan sát. Hãy nghĩ về các chương trình truyền hình thực tế như *Big Brother* hoặc *Survivor*, nơi mọi người liên tục bị quan sát và ghi hình. Mặc dù ban đầu họ có thể thể hiện những hành vi chuẩn mực nhất, nhưng chỉ trong một khoảng thời gian khá ngắn, họ bắt đầu tán tỉnh, quan hệ tình dục, ăn mặc hở hang, la hét vào mặt nhau, và đôi khi hành xử như những kẻ ngốc thực sự trước khán giả toàn quốc.

Quan sát tham gia so với Quan sát cơ học

Quan sát có sự tham gia và quan sát bằng máy móc (Participant vs mechanical observation)

Một phương pháp thu thập dữ liệu khác trong nghiên cứu quan sát (observational research) là quan sát có sự tham gia (participant observation). Trong quan sát có sự tham gia, các nhà nghiên cứu trở thành những người tham gia tích cực vào nhóm hoặc tình huống mà họ đang nghiên cứu. Quan sát có sự tham gia rất giống với quan sát tự nhiên (naturalistic observation) ở chỗ nó bao gồm việc quan sát hành vi của con người trong môi trường mà hành vi đó thường xuyên diễn ra. Giống như quan sát tự nhiên, dữ liệu thu thập được có thể bao gồm các cuộc phỏng vấn (thường là không có cấu trúc - unstructured), các ghi chú dựa trên sự quan sát và tương tác của họ, tài liệu, hình ảnh, và các hiện vật (artifacts) khác. Điểm khác biệt duy nhất giữa quan sát tự nhiên và quan sát có sự tham gia là các nhà nghiên cứu tiến hành quan sát có sự tham gia sẽ trở thành các thành viên tích cực của nhóm hoặc tình huống mà họ đang nghiên cứu. Cơ sở lý luận cơ bản cho quan sát có sự tham gia là có thể có những thông tin quan trọng mà chỉ những người tham gia tích cực vào nhóm hoặc tình huống đó mới có thể tiếp cận hoặc giải thích được. Giống như quan sát tự nhiên, quan sát có sự tham gia có thể là không công khai (disguised) hoặc công khai (undisguised).

Trong quan sát có sự tham gia không công khai (disguised participant observation), các nhà nghiên cứu giả vờ là thành viên của nhóm xã hội mà họ đang quan sát và che giấu danh tính thực sự là nhà nghiên cứu của mình. Ngược lại, đối với quan sát có sự tham gia công khai (undisguised participant observation), các nhà nghiên cứu trở thành một phần của nhóm họ đang nghiên cứu và họ tiết lộ danh tính thực sự của mình là nhà nghiên cứu cho nhóm đang được khảo sát. Một lần nữa, có những vấn đề đạo đức (ethical issues) quan trọng cần xem xét đối với phương pháp quan sát có sự tham gia không công khai. Thứ nhất, không thể có được sự đồng thuận có hiểu biết (informed consent) và thứ hai là sự lừa dối thụ động (passive deception) đang được sử dụng. Nhà nghiên cứu đang thụ

động lừa dối những người tham gia bằng cách cố ý giữ kín thông tin về động cơ của họ khi trở thành một phần của nhóm xã hội mà họ đang nghiên cứu. Nhưng đôi khi sự tham gia không công khai là cách duy nhất để tiếp cận một nhóm khép kín/cần sự bảo vệ (chẳng hạn như một giáo phái - a cult). Hơn nữa, quan sát có sự tham gia không công khai ít có xu hướng xảy ra tính phản ứng (reactivity) hơn so với quan sát có sự tham gia công khai.

Một trong những lợi ích chính của quan sát có sự tham gia (participant observation) là nhà nghiên cứu ở một vị thế tốt hơn nhiều để thấu hiểu quan điểm và trải nghiệm của những người mà họ đang nghiên cứu khi bản thân họ trở thành một phần của nhóm xã hội đó. Hạn chế cốt lõi của phương pháp tiếp cận này là chỉ riêng sự hiện diện của người quan sát cũng có thể ảnh hưởng đến hành vi của những người đang bị quan sát. Mặc dù đây cũng là một mối lo ngại đối với quan sát tự nhiên, nhưng khi các nhà nghiên cứu trở thành các thành viên tích cực của nhóm xã hội mà họ đang nghiên cứu, sẽ nảy sinh thêm những lo ngại rằng họ có thể làm thay đổi động lực học xã hội (social dynamics) và/hoặc ảnh hưởng đến hành vi của những người mà họ đang nghiên cứu. Tương tự, nếu nhà nghiên cứu đóng vai trò là người quan sát có sự tham gia (participant-observer), có thể xuất hiện những lo ngại về các thiên kiến (biases) phát sinh từ việc xây dựng các mối quan hệ với những người tham gia. Cụ thể, nhà nghiên cứu có thể trở nên kém khách quan hơn, dẫn đến nhiều thiên kiến của người thực nghiệm (experimenter bias) hơn.

Hình 1: Các loại thiết bị khác nhau được sử dụng để nghiên cứu các phản ứng sinh lý/thần kinh ở người.



Trong nhiều tình huống, các phương tiện quan sát là máy móc (mechanical) chứ không phải con người. Điều này bao gồm máy quay video, máy đếm lưu lượng giao thông (traffic counters), máy quét tại quầy thanh toán (checkout scanners), điện thoại thông minh và một loạt các thiết bị đo lường các phản ứng sinh lý (physiological responses). Các thiết bị này bao gồm màn hình theo dõi chuyển động mắt (eye-tracking monitors), máy đo đồng tử (pupilometers), máy đo điện tâm lý (psychogalvanometers), máy phân tích cao độ giọng nói (voice pitch analyzers) và các công cụ để đo lường tín hiệu thần kinh (neurological signals).

Quan sát có cấu trúc so với Quan sát không cấu trúc

Trong quan sát có cấu trúc (structured observation), trọng tâm được đặt vào việc thu thập dữ liệu định lượng (quantitative data) thay vì dữ liệu định tính (qualitative data). Các nhà

nhà nghiên cứu sử dụng phương pháp tiếp cận này chỉ quan tâm đến một tập hợp giới hạn các hành vi. Điều này cho phép họ định lượng hóa các hành vi mà họ đang quan sát. Nói cách khác, quan sát có cấu trúc ít mang tính bao quát toàn diện (less global) hơn so với quan sát tự nhiên (naturalistic observation) và quan sát có sự tham gia (participant observation), bởi vì nhà nghiên cứu thực hiện các quan sát có cấu trúc chỉ quan tâm đến một số ít các hành vi cụ thể. Do đó, thay vì ghi nhận lại mọi thứ xảy ra, nhà nghiên cứu chỉ tập trung vào các hành vi trọng tâm rất cụ thể. Ví dụ, một nhà tiếp thị (marketer) có thể quan tâm đến số lượng người bước vào một trung tâm thương mại, hoặc số lần mọi người dừng lại để xem một quảng cáo được trưng bày.

Mặt khác, quan sát phi cấu trúc (unstructured observation) lại linh hoạt và ít mang tính công thức (informal) hơn. Không có một danh mục kiểm tra (checklist) nào để nhà nghiên cứu phải tuân theo. Nhà nghiên cứu có thể quan sát mọi khía cạnh của một hiện tượng và sau đó cung cấp các chi tiết về những yếu tố mà họ nhận thấy là có liên quan trong việc thấu hiểu một tình huống. Kỹ thuật này có thể mang tính chủ quan (subjective) cao hơn so với phương pháp tiếp cận có cấu trúc.

Khi các quan sát đòi hỏi sự phán đoán (judgment) từ phía người quan sát (ví dụ: khách hàng trông có vẻ vui vẻ khi mua sắm trong cửa hàng không?), quá trình này thường được gọi là mã hóa (coding). Việc mã hóa nhìn chung đòi hỏi phải định nghĩa rõ ràng một tập hợp các hành vi mục tiêu (target behaviours). Sau đó, những người quan sát sẽ phân loại từng người tham gia dựa trên việc họ đã thực hiện hành vi nào và số lần họ thực hiện mỗi hành vi đó. Những người quan sát thậm chí có thể ghi lại thời lượng (duration) của từng hành vi. Các hành vi mục tiêu phải được định nghĩa theo cách sao cho các người quan sát khác nhau đều có thể mã hóa chúng theo cùng một cách. Ví dụ, trong một nghiên cứu, các nhà nghiên cứu đã ghi hình lại một nhóm nhỏ các phản ứng của người tham gia và để hai người quan sát tiến hành mã hóa chúng một cách độc lập. Hai người quan sát cho thấy họ đồng thuận về các phản ứng được thể hiện trong 97% thời gian, cho thấy độ tin cậy giữa những người đánh giá (interrater reliability) đạt mức tốt.

Một trong những lợi ích chính của quan sát có cấu trúc (structured observation) là nó hiệu quả hơn nhiều so với quan sát tự nhiên (naturalistic observation) và quan sát có sự tham gia (participant observation). Vì các nhà nghiên cứu chỉ tập trung vào các hành vi cụ thể, điều này giúp tiết kiệm thời gian và chi phí. Ngoài ra, thông thường môi trường được cấu trúc hóa để khuyến khích các hành vi đang được quan tâm; điều này một lần nữa đồng nghĩa với việc các nhà nghiên cứu không phải đầu tư quá nhiều thời gian vào việc chờ đợi các hành vi được quan tâm diễn ra một cách tự nhiên. Cuối cùng, các nhà nghiên cứu sử dụng phương pháp tiếp cận này rõ ràng có thể thực thi sự kiểm soát chặt chẽ hơn đối với môi trường. Tuy nhiên, khi các nhà nghiên cứu gia tăng sự kiểm soát đối với môi trường, nó có thể làm cho môi trường trở nên kém tự nhiên hơn, từ đó làm giảm độ giá trị ngoại suy (external validity). Ví dụ, không có gì đảm bảo rõ ràng rằng những quan sát có cấu trúc được thực hiện trong môi trường phòng thí nghiệm có thể khái quát hóa (generalize) sang một môi trường thực tế (real-world environment) hay không. Hơn nữa, vì các nhà nghiên cứu tham gia vào quan sát có cấu trúc thường tiến hành một cách công khai (not disguised), nên có thể nảy sinh nhiều lo ngại hơn về tính phản ứng (reactivity).

Quan sát bằng chứng vật lý (Observation of physical evidence)

Các nhà nghiên cứu thị trường có khả năng thu thập dữ liệu thông qua việc quan sát các bằng chứng vật lý (physical evidence), từ đó có thể cung cấp những **insight** then chốt. Các hiện vật (artefacts) – chẳng hạn như vỏ hộp thức ăn trong thùng rác và việc kiểm đếm hàng tồn kho thực tế (physical inventories) – là những phương pháp then chốt có thể được sử dụng. Việc phân tích nội dung (content analysis) đối với các quảng cáo, báo chí, hình ảnh, hoặc các bản ghi nhớ (memos) có thể giúp các nhà nghiên cứu thấu hiểu những khía cạnh nhất định của một hiện tượng.

Phần này đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây: <https://westernsydney.pressbooks.pub/customerinsights2/?p=60#h5p-14>

Ghi nhận nguồn truyền thông

- **Thiết bị theo dõi ánh mắt (Eye tracking equipment)** © [Sally Tsoutas - Nhiếp ảnh gia WSU](#), được bảo lưu mọi quyền ([All Rights Reserved](#)).
- **Thiết bị camera mô phỏng thần kinh (Neuromorphic camera equipment)** © [Sally Tsoutas - Nhiếp ảnh gia WSU](#), được cấp phép theo giấy phép [CC BY-NC-ND \(Ghi công - Phi thương mại - Không phái sinh\)](#).
- **Thiết bị cảm biến (Sensor equipment)** © [Sally Tsoutas - Nhiếp ảnh gia WSU](#), được cấp phép theo giấy phép [CC BY-NC-ND \(Ghi công - Phi thương mại - Không phái sinh\)](#).

10. CÁC THANG ĐO HOẶC CÁC LOẠI BIẾN SỐ (MEASURES OR TYPES OF VARIABLES)

Chương này được điều chỉnh từ nội dung [UniSkills của Đại học Curtin](#) và được sử dụng theo [Giấy phép Quốc tế Creative Commons Attribution ShareAlike 4.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Xác định các loại dữ liệu/thang đo khác nhau được sử dụng trong phép đo.
- Hiểu rõ cách áp dụng các thang đo khác nhau trong các bảng khảo sát.

Các thang đo hoặc Các loại biến số

Các loại Dữ liệu và Biến số

Chương này trình bày chi tiết một số khái niệm quan trọng sẽ được đề cập trong các trang tiếp theo, bao gồm dữ liệu và biến số là gì, và cách phân biệt giữa các loại khác nhau.

Dữ liệu là gì?

Từ "dữ liệu" (data) đề cập đến những quan sát và đo lường được thu thập bằng một phương thức nào đó, thường là thông qua nghiên cứu.

Dữ liệu được ghi nhận dưới dạng các con số (và do đó đo lường số lượng) là dữ liệu định lượng (quantitative data), trong khi dữ liệu được ghi nhận dưới dạng văn bản (và do đó ghi nhận các đặc tính) là dữ liệu định tính (qualitative data).

Dữ liệu định lượng có thể được phân tích bằng phương pháp thống kê; điều này cũng áp dụng cho dữ liệu định tính ghi nhận các đặc tính theo các danh mục (categories) khác nhau (ví dụ: một người có màu tóc gì, sinh ra ở quốc gia nào, tình trạng hôn nhân của họ ra sao, v.v.), trái ngược với loại dữ liệu ghi nhận các đặc tính dưới dạng suy nghĩ, cảm nhận và ý kiến.

Chính hai loại dữ liệu được đề cập trước đó là đối tượng mà bạn sẽ làm việc trong học phần này, và ngay sau đây, bạn sẽ được giới thiệu một số thuật ngữ khác thường được sử dụng trong thống kê để mô tả dữ liệu có bản chất như vậy.

Biến số là gì?

Biến số (Variables) là các đặc điểm hoặc thuộc tính mà bạn đang quan sát, đo lường và ghi nhận dữ liệu – một số ví dụ có thể kể đến như chiều cao, cân nặng, màu mắt, giống chó, khí hậu, độ dẫn điện, mức độ hài lòng về dịch vụ khách hàng và sĩ số tham dự lớp học, đây chỉ là một vài ví dụ điển hình.

Đúng như tên gọi của nó, giá trị của một biến số sẽ biến thiên/thay đổi (varies) từ đối tượng nghiên cứu (subject - tức là con người, địa điểm hoặc sự vật) này sang đối tượng khác. Ví dụ, biến số chiều cao có thể có giá trị là 170cm đối với người này, 163cm đối với người khác và 154cm đối với một người khác nữa; biến số khí hậu có thể có giá trị là khô hạn đối với một thành phố, nhiệt đới đối với thành phố khác và Địa Trung Hải đối

với một thành phố khác nữa; và biến số sĩ số tham dự lớp học có thể có giá trị là 17 đối với một lớp học, 25 đối với lớp học khác và 32 đối với một lớp khác nữa, v.v.

Dữ liệu phân loại và Dữ liệu Liên tục

Việc lựa chọn đại lượng thống kê (statistic) hoặc kiểm định thống kê (statistical test) chính xác để phân tích dữ liệu của bạn phụ thuộc vào loại dữ liệu, và do đó phụ thuộc vào (các) loại biến số, vì vậy việc có thể phân biệt giữa chúng là rất quan trọng. Trong hầu hết các trường hợp, bạn sẽ chỉ cần phân loại dữ liệu (và do đó là các biến số) của mình thành dạng phân loại (categorical) hoặc liên tục (continuous), nhưng mỗi loại này cũng có thể được phân loại chi tiết hơn (sub-classified). Định nghĩa và các phân loại chi tiết cho từng loại như sau:

Dữ liệu Định loại (Categorical Data)

Dữ liệu phân loại (Categorical data) là dữ liệu được phân nhóm thành các danh mục, chẳng hạn như dữ liệu cho biến số 'giới tính' hoặc 'tình trạng hút thuốc'. Dữ liệu phân loại có thể được phân loại chi tiết hơn thành:

- **định danh (nominal)** khi các danh mục không có thứ tự, chẳng hạn như đối với biến số 'tình trạng hôn nhân' (hơn nữa, nếu chỉ có hai danh mục thì các thuật ngữ **nhị phân (binary)** và/hoặc **lưỡng phân (dichotomous)** đôi khi được sử dụng); hoặc
-

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=62#h5p-15>

thứ bậc (ordinal) khi các danh mục có thứ tự rõ ràng, chẳng hạn như đối với chiếc xe 'tốt nhất' (dẫn đầu), chiếc xe 'á quân' (về nhì) và chiếc xe xếp 'thứ ba' trong danh sách các phương tiện được yêu thích nhất tại Úc.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=62#h5p-16>

Dữ liệu Liên tục (Continuous Data)

Dữ liệu liên tục (Continuous data) là dữ liệu được đo lường trên một thang đo số liên tục và có thể nhận một số lượng lớn các giá trị có thể có, chẳng hạn như dữ liệu cho biến số 'cân nặng' hoặc 'khoảng cách'. Dữ liệu liên tục có thể được phân loại chi tiết hơn thành:

- **thang đo khoảng (interval)** khi nó không có điểm 0 tuyệt đối (absolute zero), và các số âm cũng mang ý nghĩa, chẳng hạn như đối với biến số 'nhiệt độ tính bằng độ C'. Trong tiếp thị (marketing), bất kỳ biến số nào được đo lường trên thang đo

Likert đều được xem là một biến khoảng (interval variable). Dưới đây là một ví dụ:

Ví dụ: Trên thang điểm từ 1 đến 5 (trong đó 1 = cực kỳ không hài lòng và 5 = cực kỳ hài lòng), bạn sẽ đánh giá nhà hàng như thế nào theo các tiêu chí sau:

Sự sạch sẽ (Cleanliness)	1	2	3	4	5
Cách trình bày món ăn (Food presentation)	1	2	3	4	5
Món ăn - hương vị (Food - taste)	1	2	3	4	5
Phục vụ nhanh chóng (Prompt service)	1	2	3	4	5
Nhân viên lịch sự (Courteous employees)	1	2	3	4	5

Đánh giá của khách hàng về các tiêu chí trên có thể được tính trung bình để đưa ra điểm số cho 'mức độ hài lòng đối với nhà hàng'.

thang đo tỷ lệ (ratio) khi nó có một điểm 0 tuyệt đối (absolute zero), và các số âm không mang ý nghĩa, chẳng hạn như 'mức chi tiêu' cho hàng tạp hóa hoặc 'thời gian tiêu tốn' trong nhà bếp.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=62#h5p-17>

Một điểm cuối cùng cần lưu ý về chủ đề này là bất kỳ dữ liệu liên tục (continuous data) nào cũng luôn có thể được chuyển đổi thành dữ liệu phân loại (categorical data), đơn giản bằng cách tạo ra các danh mục từ nó. Ví dụ, dữ liệu liên tục cho biến số 'cân nặng' có thể được chuyển đổi thành dữ liệu phân loại bằng cách tạo ra các danh mục như 50–59kg, 60–69kg, 70–79kg, v.v., và điều này có thể hữu ích nếu bạn muốn phân tích dữ liệu liên tục của mình bằng cách sử dụng các đại lượng thống kê (statistics) và các kiểm định thống kê (statistical tests) được thiết kế riêng cho dữ liệu phân loại. Tuy nhiên, bạn không thể làm theo chiều ngược lại và chuyển đổi dữ liệu phân loại thành dữ liệu liên tục, vì vậy, nếu có quyền lựa chọn, để đạt được sự linh hoạt tối đa

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=62#oembed-1>

11. THIẾT KẾ BẢNG CÂU HỎI (QUESTIONNAIRE DESIGN)

Chương này được điều chỉnh từ nội dung của [Cục Thống kê Úc \(Australian Bureau of Statistics\)](#), được cấp phép theo [Giấy phép Quốc tế Creative Commons Attribution 4.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Nêu bật các bước trong quy trình thiết kế bảng câu hỏi.
- Làm quen với các nguyên tắc để thiết kế một bảng câu hỏi tốt.
- Xác định các lỗi thường gặp trong bảng câu hỏi.

Thiết kế Bảng câu hỏi

Bảng câu hỏi có thể ảnh hưởng đến tỷ lệ phản hồi (response rate) đạt được trong cuộc khảo sát, chất lượng của các câu trả lời thu thập được, và hệ quả là các kết luận được rút ra từ kết quả khảo sát. Việc thiết kế bảng câu hỏi nên được bắt đầu bằng cách xem xét các mục tiêu của cuộc khảo sát và đầu ra (output) được yêu cầu, sau đó xây dựng một danh sách các câu hỏi để thu thập thông tin này một cách chính xác. Cần cân nhắc kỹ lưỡng một số yếu tố bao gồm các loại câu hỏi sẽ được sử dụng, cách diễn đạt từ ngữ của bảng câu hỏi (questionnaire wording), cấu trúc và thiết kế của bảng câu hỏi, cũng như việc thử nghiệm (testing) bảng câu hỏi để đảm bảo thu thập được dữ liệu chất lượng. Một bảng câu hỏi được thiết kế kém có thể là nguồn gốc lớn nhất gây ra sai số phi lấy mẫu (non-sampling error) (dù là trực tiếp hay gián tiếp).

Phần này thảo luận về giai đoạn xây dựng của quá trình thiết kế bảng câu hỏi, cách diễn đạt câu hỏi, các loại câu hỏi khác nhau, cùng với cấu trúc và thiết kế của bảng câu hỏi.

Việc thử nghiệm bảng câu hỏi nháp (draft questionnaire) cũng sẽ được đề cập trong chương này.

Phát triển Bảng câu hỏi (Questionnaire Development)

Phần này liệt kê các bước khác nhau liên quan đến quá trình xây dựng một bảng câu hỏi. Giai đoạn xây dựng này cần đảm bảo rằng các hạng mục dữ liệu (data items) trong bảng câu hỏi cuối cùng là phù hợp và có thể được thu thập một cách chính xác. Điều quan trọng nữa là phải xem xét từ sớm việc mã hóa (coding) các câu trả lời của bảng câu hỏi và loại hệ thống xử lý (processing system) sẽ được sử dụng để xử lý bảng câu hỏi.

Bước 1: Xác định mục tiêu với Khách hàng

Trước khi bắt đầu thiết kế bảng câu hỏi, trước tiên phải dành thời gian làm rõ các mục tiêu của cuộc khảo sát với khách hàng. Điều quan trọng là lý do thu thập thông tin phải được trình bày một cách rõ ràng và rành mạch.

Việc xác định mục tiêu với khách hàng bao gồm các công việc sau:

- **Làm rõ mục tiêu của cuộc khảo sát:** Mục đích cuối cùng của dự án này là gì?
- **Giải trình về việc thu thập dữ liệu dựa trên lợi ích:** Việc khảo sát này sẽ mang lại giá trị gì cho doanh nghiệp hoặc tổ chức?
- **Xác định phạm vi khảo sát:** Cụ thể là ai sẽ được khảo sát? (Đối tượng mục tiêu).
- **Xác định kết quả đầu ra mong muốn:** Chỉ định các bảng biểu/dữ liệu thống kê dựa theo mục tiêu và mức độ chính xác yêu cầu.

- **Chuẩn bị danh mục nội dung:** Liệt kê các thông tin cần thiết dựa trên mục tiêu. Có thể cần nhiều câu hỏi khác nhau để thu thập được một mẫu thông tin mong muốn.
- **Định nghĩa nội dung:** Xác định rõ ràng từng mẫu thông tin cần thu thập có ý nghĩa gì.
- **Giải trình cho nội dung:** Đảm bảo tính cần thiết của mọi thông tin được thu thập. Liệu thông tin này có thực sự phục vụ cho mục tiêu không?
- **Thiết lập thứ tự ưu tiên cho từng hạng mục dữ liệu:** Điều này rất quan trọng để đảm bảo những dữ liệu quan trọng nhất luôn được thu thập. Nó cũng giúp việc loại bỏ các câu hỏi ít quan trọng trở nên dễ dàng hơn nếu ngân sách khảo sát bị cắt giảm.

Bước 2: Nghiên cứu Chủ đề

Điều rất quan trọng là nhà nghiên cứu chuẩn bị thiết kế bảng câu hỏi phải tiến hành nghiên cứu nền tảng (background research) về chủ đề đang được xem xét. Về mặt thiết kế bảng câu hỏi, nghiên cứu nên nhằm mục đích:

- định nghĩa rõ ràng các khái niệm và định nghĩa sẽ được sử dụng trong cuộc khảo sát. Việc sử dụng các định nghĩa và hệ thống phân loại chuẩn (standard definitions and classifications) sẽ làm tăng khả năng so sánh của dữ liệu (data comparability); và
- xác định chính xác các đặc điểm của tổng thể mục tiêu (target population characteristics). Điều này cần được thực hiện để các câu hỏi có thể được đặt ra ở một mức độ/trình độ phù hợp.

Việc xem xét các cuộc khảo sát trước đây về cùng chủ đề để học hỏi từ những kinh nghiệm trong quá khứ cũng rất hữu ích. Những dữ liệu thu thập (collections) như vậy có thể chứng minh được tính hữu ích vì nó giúp tránh được những sai lầm có thể xảy ra. Điều quan trọng là phải xem xét các khái niệm, định nghĩa và cách diễn đạt câu hỏi

(question wording) của các lần thu thập trước đó nếu muốn tạo ra một chuỗi thời gian (time series), bởi vì những thay đổi trong các yếu tố này THIẾT KẾ BẢNG CÂU HỎI | 115 sẽ dẫn đến những thay đổi trong các câu trả lời. Ví dụ, một câu hỏi trong Cuộc tổng điều tra dân số (Population Census) năm 1981, "Bạn có đang thế chấp/vay cầm cố (mortgage) không?" đã nhận được các câu trả lời rất khác so với những câu trả lời thu được trong Cuộc tổng điều tra năm 1986 khi câu hỏi được đổi thành "Bạn đã thanh toán xong tiền nhà chưa?". Nhiều người có khoản thế chấp thứ hai đã tính cả điều này vào trong Cuộc tổng điều tra năm 1981 nhưng lại không đưa vào Cuộc tổng điều tra năm 1986.

Bước 3: Quyết định các câu hỏi sàng lọc (để chọn lọc người trả lời)

Các câu hỏi sàng lọc (còn được gọi là "screeners") được sử dụng để xác nhận đủ điều kiện (qualify) hoặc loại bỏ (disqualify) người trả lời khỏi việc tham gia khảo sát — tùy thuộc vào cách họ trả lời. Để lập ra một danh sách toàn diện các câu hỏi sàng lọc, các nhà nghiên cứu cần xây dựng một tập hợp rõ ràng các tiêu chí lựa chọn (inclusion criteria) và tiêu chí loại trừ (exclusion criteria). Các tiêu chí lựa chọn có thể dựa trên nhân khẩu học (ví dụ: nữ giới, sinh sống tại Sydney, trong độ tuổi từ 18-25 và đang theo học cử nhân Kinh doanh), thái độ (chẳng hạn như ưu tiên mua nông sản trồng tại địa phương), phong cách sống (dành những ngày cuối tuần trên bãi biển), và/hoặc hành vi (người thường xuyên mua cà phê Nescafé).

Hãy tưởng tượng bạn đang làm việc cho Nestlé. Bạn muốn có phản hồi về một loại cà phê hòa tan phối trộn (blend) mới. Tuy nhiên, bạn chỉ muốn thu thập thông tin từ nhóm khách hàng sử dụng với tần suất cao (heavy users) tại Sydney. Bảng câu hỏi sàng lọc của bạn có thể bao gồm các câu hỏi như:

- Bạn sống ở đâu? (LƯU Ý: Dừng khảo sát nếu câu trả lời là bất kỳ nơi nào khác ngoài Sydney)
- Bạn có thích cà phê hòa tan không? (Nếu Không, – hoặc không chắc chắn hoặc không áp dụng – thì không tiếp tục với các câu hỏi sau)
- Trung bình, bạn uống bao nhiêu cà phê hòa tan trong một ngày? (Bất kỳ ai uống dưới 4 tách sẽ bị loại ra)

Bước 4: Hoàn thiện các câu hỏi và các loại thang đo (Type of Scales)

Việc quyết định nội dung của các câu hỏi và thang đo được sử dụng để đo lường chúng sẽ có tác động đến quá trình phân tích dữ liệu (data analysis). Toàn bộ nội dung câu hỏi cần được liên kết chặt chẽ với các mục tiêu ban đầu đã được thống nhất trong quá trình tham vấn (consultation) với khách hàng.

Câu hỏi Mở so với Câu hỏi Đóng

Các câu hỏi nhìn chung có thể được phân loại thành một trong hai loại – câu hỏi mở (open) hoặc câu hỏi đóng (closed) – tùy thuộc vào lượng thông tin có thể được cung cấp. Các câu hỏi mở cho phép người trả lời (respondents) trả lời câu hỏi bằng ngôn từ của

chính họ. Một ví dụ là ‘Nghề nghiệp của bạn là gì?’ Ưu điểm của những loại câu hỏi này là chúng cho phép có nhiều đáp án có thể xảy ra và có thể thu thập dữ liệu chính xác từ nhiều loại phản hồi (responses) khác nhau. Tuy nhiên, chúng đòi hỏi nhiều công sức hơn so với câu hỏi đóng, cả trong việc trả lời lẫn xử lý. Các câu hỏi mở thường được sử dụng trong các cuộc thử nghiệm thí điểm (pilot tests) để xác định phạm vi của các câu trả lời có khả năng xảy ra.

Các câu hỏi đóng cung cấp cho người trả lời một loạt các đáp án có khả năng xảy ra nhất để họ lựa chọn. Những câu hỏi này phù hợp khi nhà nghiên cứu có thể dự đoán (anticipate) được hầu hết các câu trả lời và khi không cần đến các giá trị chính xác tuyệt đối. Tuy nhiên, chúng đòi hỏi nhiều nỗ lực hơn so với các câu hỏi mở trong các giai đoạn xây dựng (development) và thử nghiệm (testing). Thời gian xử lý của các câu trả lời đóng ít hơn nhiều so với các câu trả lời dạng mở (open-ended responses).

Bảng: Các loại Câu hỏi Đóng

Các loại câu hỏi đóng (Types of Closed Questions)	Định dạng câu trả lời (Answer Format)
Lựa chọn giới hạn (Limited Choice)	Có/Không (Yes/No)
Nhiều lựa chọn (Multiple Choice)	Chọn từ một số các câu trả lời (Choose from a number of responses)
Danh mục kiểm tra (Checklist)	Chọn nhiều hơn một câu trả lời từ một danh sách cho sẵn (Choose more than one of the responses from a given list)

Đóng một phần (Partially Closed)	Lựa chọn cuối cùng ghi ‘Khác, vui lòng ghi rõ’ (Last alternative states ‘Other, please specify’)
----------------------------------	--

Bảng: Các loại câu hỏi cho các loại thông tin khác nhau

Loại câu hỏi	Thông tin cần tìm
Câu hỏi Sự thật (Factual)	Thông tin khách quan, ví dụ: Anh/Chị có bằng lái xe không?
Câu hỏi Ý kiến (Opinion)	Ý kiến cá nhân của người trả lời đối với một vấn đề, ví dụ: Thái độ đối với nước giải khát.
Câu hỏi Hành vi (Behavioural)	Hành động hoặc việc thực hiện một hoạt động của người trả lời, ví dụ: Anh/Chị có đi du lịch nước ngoài trong năm nay không?
Câu hỏi Giả định (Hypothetical)	Câu trả lời của người phản hồi cho các tình huống giả định, ví dụ: Anh/Chị sẽ làm gì nếu...

Câu hỏi Nhân khẩu học (Demographic)	Thông tin cá nhân của người trả lời như giới tính, tuổi tác, nghề nghiệp, mã bưu điện; Dùng để phân đoạn đối tượng phản hồi.
Câu hỏi Kiến thức (Knowledge)	Dùng để kiểm tra kiến thức của người trả lời về các vấn đề nhất định, ví dụ: Ai là Thủ tướng hiện tại của Úc?

Trong khi loại câu hỏi liên quan đến những thông tin cần thu thập từ người trả lời, thì cách thức nó được ‘đo lường’ (với các thang đo phù hợp) sẽ tạo điều kiện khả thi để tiến hành các loại hình phân tích thống kê khác nhau. Dưới đây là một ví dụ minh họa cách thức bốn loại thang đo có thể được sử dụng để đo lường cùng một khái niệm.

Giả sử, một nhà tiếp thị quan tâm đến việc đo lường thái độ và hành vi mua sắm của mọi người liên quan đến kem.

Dưới đây là một số cách thức mà các câu hỏi/thang đo có thể được thiết kế:

A. “Bạn có mua một sản phẩm kem nào trong vài ngày qua không?”

Có ___ Không ___

(Thang đo định danh - Nominal scale)

B. “Giả sử bạn đang lập một danh sách mua sắm. Hãy xếp hạng các sản phẩm sau từ 1 đến 4, trong đó 1 = ưu tiên cao nhất và 4 = ít ưu tiên nhất trong danh sách” (Thang đo thứ bậc - Ordinal Scale)

Sản phẩm: Cà phê, Nước ngọt, Kem, Nước ép, Sô-cô-la

1.

2.

3.

4.

C. “Trên thang điểm từ 1 đến 5, trong đó 1 = Hoàn toàn không đồng ý và 5 = Hoàn toàn đồng ý, vui lòng đánh giá mức độ đồng ý của bạn đối với các nội dung sau

(Thang đo khoảng - Interval Scale)

	Rất không đồng ý				Rất đồng ý
Tôi yêu kem	1	2	3	4	5
Tôi cảm thấy hạnh phúc khi ăn kem	1	2	3	4	5

Tôi sẽ không ngại khó khăn để tìm mua loại kem yêu thích của mình	1	2	3	4	5
---	---	---	---	---	---

D. Anh/Chị đã chi bao nhiêu tiền cho mặt hàng kem trong tuần vừa qua? \$ _____
(Thang đo tỷ lệ - Ratio scale)

Bước 5: Kiểm tra cách diễn đạt câu hỏi và các danh mục trả lời

Có một số yếu tố cần cân nhắc khi thiết kế câu hỏi để đảm bảo thu được các câu trả lời phù hợp. Một số khía cạnh của việc thiết kế câu hỏi có thể dẫn đến sai số (errors), cụ thể là:

Ngôn ngữ

Các câu hỏi sử dụng ngôn ngữ phức tạp, mang tính kỹ thuật hoặc các thuật ngữ chuyên ngành (jargon) có thể gây bối rối hoặc khó chịu cho người trả lời. Đối với các cuộc khảo sát có phỏng vấn viên (interviewer-based surveys), những người trả lời không hiểu câu hỏi có thể e ngại việc bị xem là thiếu hiểu biết nên sẽ không yêu cầu phỏng vấn viên giải thích câu hỏi. Khi đó, người trả lời có thể từ chối trả lời hoặc đưa ra một phản hồi không chính xác.

Ngôn ngữ kỹ thuật hoặc thuật ngữ chuyên ngành chỉ nên được sử dụng trong những trường hợp chúng là một phần trong ngôn ngữ giao tiếp thông thường của tổng thể mục tiêu (target population) của cuộc khảo sát. Một ví dụ cho trường hợp này là một cuộc

khảo sát dành cho các chuyên gia công nghệ thông tin: cuộc khảo sát sẽ cần sử dụng ngôn ngữ mang tính chất ‘biệt ngữ/thuật ngữ chuyên ngành’ đối với người thiết kế khảo sát, nhưng lại hoàn toàn phù hợp với người trả lời.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây: <https://westernsydney.pressbooks.pub/customerinsights2/?p=64#h5p-18>

Một nguyên tắc chung cần lưu ý là cách diễn đạt các hạng mục trong bảng câu hỏi (questionnaire items) phải cụ thể, rõ ràng, nhất quán, ngắn gọn, đơn giản và dễ hiểu (tự nó đã rõ nghĩa).

Tính mơ hồ (Ambiguity)

Nếu các từ hoặc cụm từ mơ hồ được đưa vào trong câu hỏi, ý nghĩa của chúng có thể bị diễn giải khác nhau bởi những người khác nhau. Điều này sẽ gây ra sai số cho dữ liệu vì những người trả lời khác nhau thực chất đang trả lời cho những câu hỏi khác nhau.

Ví dụ, hãy xem xét câu hỏi: *"Mức sống của Anh/Chị có bị sụt giảm đáng kể do sự gia tăng đột ngột của các khoản thanh toán thế chấp hàng tháng không?"*. Một câu trả lời "Không" có thể hàm chứa một trong nhiều nghĩa khác nhau – chẳng hạn như: "Không, mức sống của tôi không giảm vì các khoản thanh toán tăng lên" hoặc "Không, các khoản thanh toán của tôi không hề tăng".

Một câu hỏi cũng có thể trông có vẻ thẳng thắn nhưng lại cho phép nhiều loại câu trả lời khác nhau. Việc đưa vào đơn vị đo lường mà bạn yêu cầu ở bất cứ nơi nào có thể áp dụng là rất quan trọng, ví dụ: đô la, ngày, lít.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=64#h5p-19>

Câu hỏi kép (Nhiều khái niệm trong cùng một câu hỏi)

Đây là những câu hỏi bề ngoài có vẻ là câu hỏi đơn nhưng thực chất lại lồng ghép hai câu hỏi khác nhau. Ví dụ: *"Anh/Chị có dự định nghỉ việc và quay lại học toàn thời gian trong năm nay không?"*. Một người có thể dự định nghỉ việc nhưng không quay lại học tập, và ngược lại. Khi các phần khác nhau của câu hỏi đòi hỏi các câu trả lời khác nhau hoặc có phần không liên quan, người trả lời có thể không chắc chắn về cách trả lời. Khi cố gắng diễn giải câu trả lời cho những câu hỏi như vậy, nhà nghiên cứu sẽ không rõ câu trả lời đó tương ứng với phần nào của câu hỏi.

Câu hỏi Dẫn dắt (Leading Questions)

Sai số sẽ xuất hiện nếu câu hỏi dẫn dắt người trả lời hướng tới một phản hồi cụ thể. Ví dụ, câu hỏi *"Anh/Chị đã làm việc bao nhiêu ngày trong tuần trước?"*, nếu được hỏi mà không xác định trước liệu người trả lời có thực sự làm việc trong tuần đó hay không, thì đây là một câu hỏi dẫn dắt. Nó ám chỉ rằng người đó lẽ ra đã hoặc nên đi

làm. Người trả lời có thể trả lời sai để tránh việc phải nói với người phỏng vấn rằng họ đã không đi làm.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=64#h5p-20>

Câu hỏi Thiếu cân bằng (Unbalanced Questions)

Một dạng khác của câu hỏi dẫn dắt là câu hỏi thiếu cân bằng. Ví dụ, câu hỏi *"Anh/Chị có ủng hộ việc kiểm soát súng đạn không?"* chỉ đưa ra một phương án duy nhất để xem xét. Câu hỏi nên được diễn đạt lại thành một dạng như: *"Anh/Chị ủng hộ hay phản đối việc kiểm soát súng đạn?"*, nhằm cung cấp cho người trả lời nhiều hơn một lựa chọn.

Các phương án trả lời của một câu hỏi cũng có thể thiếu tính cân bằng. Ví dụ, một người phản hồi có thể được hỏi một cách trung lập rằng: *"Vui lòng đánh giá tình trạng sức khỏe tổng quát của Anh/Chị"*, nhưng lại bị yêu cầu lựa chọn trong các phương án trả lời gồm: "Kém", "Tốt" và "Xuất sắc".

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=64#h5p-21>

Sai số do Trí nhớ/Hồi tưởng (Recall/Memory Error)

Một mức độ sai số đáng kể có thể xuất hiện trong các câu hỏi yêu cầu người trả lời hồi tưởng lại các sự kiện, chi phí, v.v., đặc biệt nếu các chi tiết được tìm kiếm trong một khoảng thời gian dài. Chất lượng dữ liệu thu thập từ các câu hỏi hồi tưởng bị ảnh hưởng bởi tầm quan trọng của sự kiện đối với người trả lời và khoảng thời gian kể từ khi sự kiện đó diễn ra. Người trả lời cũng có xu hướng ghi nhớ những gì *nên làm* hơn là những gì *đã thực hiện*.

Những chủ đề có tầm quan trọng hoặc mức độ quan tâm lớn hơn đối với người trả lời, hoặc những sự kiện xảy ra không thường xuyên, sẽ được ghi nhớ trong thời gian dài hơn và chính xác hơn. Khi có thể (ví dụ: với thông tin tài chính), các câu hỏi nên được thiết lập sao cho người trả lời có thể tham chiếu đến hồ sơ cá nhân của họ, điều này giúp nâng cao tính chính xác của báo cáo. Việc giảm thiểu giai đoạn hồi tưởng (recall period) cũng giúp giảm bớt thiên kiến trí nhớ.

Một dạng sai số trí nhớ cụ thể là **hiệu ứng viễn kính (telescoping)**. Điều này xảy ra khi người trả lời báo cáo các sự kiện xảy ra sớm hơn hoặc muộn hơn so với thực tế, đưa các sự kiện vào giai đoạn tham chiếu một cách sai lệch. Ảnh hưởng này có thể được giảm bớt phần nào bằng cách quy định cực kỳ cụ thể thời điểm bắt đầu và kết thúc của giai đoạn tham chiếu, ví dụ: sử dụng cụm từ "tuần kết thúc vào Thứ Bảy ngày 1 tháng 9" thay vì "tuần trước".

Câu hỏi xâm phạm quyền riêng tư (nhạy cảm)

Các câu hỏi về những chủ đề mà người trả lời có thể coi là gây bối rối hoặc cực kỳ nhạy cảm có thể tạo ra những câu trả lời không chính xác. Người phản hồi có thể từ chối cung cấp thông tin về các vấn đề cá nhân như chi tiết về sức khỏe hoặc thu nhập.

Nếu người trả lời bị yêu cầu trả lời những câu hỏi với thông tin có vẻ không được xã hội ủng hộ (socially undesirable), họ có thể đưa ra những phản hồi mà họ tin là "được

chấp nhận" hơn cho người phỏng vấn. Trong những trường hợp này, phương pháp tốt hơn thường là cung cấp cho người trả lời một bảng câu hỏi tự điền (self-administered questionnaire) mà người phỏng vấn không nhìn thấy.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=64#h5p-22>

Các đối tượng phản hồi trong khảo sát doanh nghiệp cũng có thể coi một số chủ đề là nhạy cảm, chẳng hạn như vi phạm an ninh CNTT hoặc quyên góp từ thiện, cũng như không muốn tiết lộ các thông tin bảo mật thương mại về doanh nghiệp của họ. Các cuộc khảo sát doanh nghiệp bổ sung thêm một khía cạnh mới trong việc thu thập dữ liệu nhạy cảm, vì thường cần nhiều người phản hồi khác nhau, đôi khi từ các bộ phận khác nhau, để hoàn thành các phần của biểu mẫu và phê duyệt nội dung của nó. Một số người phản hồi có thể không muốn người khác nhìn thấy các câu trả lời cụ thể.

Tác động tiêu cực của các câu hỏi nhạy cảm có thể trở nên nghiêm trọng hơn nếu chúng được đặt ở đầu bảng câu hỏi, do đó có thể dẫn đến việc không phản hồi (non-response) nếu người trả lời không sẵn lòng tiếp tục các câu hỏi còn lại. Nếu một câu hỏi nhạy cảm được đặt sâu hơn trong biểu mẫu, người trả lời đã có cam kết hoàn thành cao hơn, và nếu họ từ chối tiếp tục, thì phản hồi đã thực hiện vẫn hữu ích hơn. Các cách để vượt qua khó khăn liên quan đến câu hỏi nhạy cảm có thể bao gồm

việc trấn an người trả lời rằng thông tin họ cung cấp là bảo mật và không yêu cầu người trả lời viết tên của họ ở bất kỳ đâu trên mẫu khảo sát.

Xu hướng đồng thuận (Acquiescence)

Tình huống này phát sinh khi người trả lời có xu hướng chung là đồng ý thay vì phản đối với bất kỳ điều gì. Nó xảy ra khi người trả lời được hỏi liệu họ đồng ý hay không đồng ý với một nhận định, đặc biệt là khi các nhận định đưa ra được trình bày như những quan điểm chung hợp lý. Xu hướng này cũng có thể xuất hiện đối với các câu hỏi yêu cầu trả lời "Có" hoặc "Không".

Khuynh hướng này có thể do sự kết hợp của nhiều yếu tố, chẳng hạn như tính cách và trình độ học vấn của người trả lời, cũng như các điều kiện phỏng vấn hoặc thiết kế của bảng câu hỏi tự điền. Người trả lời thường sẽ đồng ý khi câu hỏi mơ hồ hoặc khó trả lời. Hiệu ứng này có thể bị phóng đại khi người trả lời mệt mỏi hoặc phải trả lời một chuỗi dài các câu hỏi có cùng các danh mục phản hồi. Một hiệu ứng liên quan là **sự thỏa hiệp (satisficing)**, khi đó người trả lời chọn câu trả lời hợp lý đầu tiên thay vì nỗ lực tìm kiếm hoặc ghi nhớ câu trả lời tốt nhất.

Các danh mục phản hồi đầy đủ (Adequate Response Categories)

Việc đảm bảo các danh mục phản hồi đầy đủ và bao quát mọi phương án trả lời khả thi là rất quan trọng.

- **Ví dụ về sự thiếu sót:** Độ tuổi: 15-19, 21-25. Điều này gây khó khăn cho những người phản hồi đang ở độ tuổi 20.
- **Ví dụ về sự chồng lấn:** Độ tuổi: 15-20, 20-25. Các khoảng giá trị luôn phải **loại trừ lẫn nhau (mutually exclusive)**. Ví dụ này gây rắc rối cho người 20 tuổi vì họ có thể trả lời vào một trong hai hoặc cả hai danh mục.

Các danh mục phản hồi cũng cần được dùng từ cẩn thận, vì người trả lời sẽ dựa vào đó để làm rõ hoặc mở rộng ý nghĩa của câu hỏi. Ví dụ, nếu một câu hỏi sử dụng thang đo tần suất 5 điểm: "Không bao giờ", "Hiếm khi", "Trung bình", "Thường xuyên" và "Rất thường xuyên", người trả lời có thể giả định sai rằng thang đo này đại diện cho sự phân bố của toàn bộ quần thể. Nếu họ tự coi mình là bình thường hoặc cực đoan so với quần thể đối với hoạt động đang được quan tâm, câu trả lời của họ sẽ khác đi bất kể tần suất thực tế mà họ thực hiện hoạt động đó.

Số lượng các lựa chọn phản hồi

Số lượng danh mục phản hồi có thể ảnh hưởng đến chất lượng dữ liệu, vì quá ít hoặc quá nhiều danh mục đều có thể gây ra sai số. Quá nhiều danh mục có thể gây ra sự mệt mỏi và thiếu tập trung cho người trả lời, dẫn đến các câu trả lời thiếu suy nghĩ. Nếu có quá ít danh mục, người trả lời có thể gặp khó khăn trong việc tìm thấy phương án mô tả chính xác tình huống của họ.

Danh mục "Không biết" (Don't Know Category)

Việc quyết định có nên đưa vào hay loại bỏ tùy chọn "Không biết" phụ thuộc phần lớn vào nội dung khảo sát. Thông thường, nếu loại bỏ danh mục này, các phản hồi còn lại sẽ được phân bổ đều cho các phía tiêu cực và tích cực của thang đo, mặc dù điều này phụ thuộc nhiều vào bản chất câu hỏi. Tuy nhiên, loại bỏ tùy chọn này có thể không phải là ý hay, vì người trả lời có thể bị buộc phải đưa ra câu trả lời khi họ thực sự không biết thái độ của mình là gì đối với một chủ đề cụ thể, hoặc không biết câu trả lời cho một câu hỏi thực tế. Khi người trả lời bị buộc phải hình thành thái độ ngay tại chỗ, thái độ đó sẽ cực kỳ thiếu tin cậy.

Sắc thái (Tone)

Một sự thay đổi trong cách dùng từ có thể dẫn đến sự thay đổi trong phản hồi. Ví dụ, hai câu hỏi sau đây có thể thu được các phản hồi khác nhau:

- *"Anh/Chị có nghĩ rằng việc sở hữu súng nên bị cấm (forbidden) không?" hoặc*
- *"Anh/Chị có nghĩ rằng việc sở hữu súng không nên được cho phép (not be allowed) không?"*

Những thay đổi nhỏ trong cách dùng từ cũng có thể có tác động đáng kể đến câu trả lời. Do đó, cần thận trọng khi xem xét các cách diễn đạt thay thế. Nên tránh sử dụng các từ phủ định như "không" (not) trong câu hỏi vì người trả lời dễ bỏ sót. Ngoài ra, việc sử dụng "không" trong một thang đo như "Hài lòng", "Trung lập" và "Không hài lòng" (Not satisfied) không tạo ra một sự đối lập thực sự. "Bất mãn" (Dissatisfied) sẽ là một lựa chọn thay thế tốt hơn, tuy nhiên "Chưa hài lòng" (Unsatisfied) cũng có thể được sử dụng và sẽ mang ý nghĩa hơi khác một chút đối với người trả lời.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=64#oembed-1>

Bước 6: Hoàn thiện Cấu trúc và Bố cục của Bảng câu hỏi

Không chỉ cách dùng từ trong câu hỏi cần được chú trọng đến từng chi tiết, mà "diện mạo" của bảng câu hỏi cũng rất quan trọng. Những bảng câu hỏi được thiết kế kém (ví dụ: văn bản khó đọc) không chỉ làm nản lòng người trả lời trong việc hoàn thành bảng khảo sát mà còn có thể dẫn đến việc người trả lời mắc lỗi. Một số yếu tố quan trọng trong cấu trúc và bố cục của bảng câu hỏi được trình bày dưới đây.

Trình tự (Sequencing)

Các câu hỏi trong một biểu mẫu nên tuân theo một trình tự (sequence) hợp lý đối với người trả lời. Bất kể phương pháp nào được sử dụng để tiến hành bảng câu hỏi (administer the questionnaire), trình tự nên được diễn ra trôi chảy từ câu hỏi này sang câu hỏi tiếp theo. Sự tiến triển trôi chảy qua các câu hỏi đặc biệt quan trọng nếu bảng câu hỏi được trả lời trong những hoàn cảnh khó khăn (ví dụ: một người mẹ đang cố gắng điền vào bảng câu hỏi trong khi những đứa trẻ của cô ấy đang đòi hỏi sự chú ý). Một ý tưởng hay là bắt đầu bảng câu hỏi bằng những câu hỏi dễ chịu và đơn giản để thúc đẩy sự quan tâm đến cuộc khảo sát, và mang lại cho người trả lời sự tự tin vào khả năng trả lời các câu hỏi còn lại của họ. Đặc biệt, các câu hỏi mở đầu (opening questions) nên xác định được rằng người trả lời là một thành viên của tổng thể khảo sát (survey population).

Các câu hỏi còn lại nên được cấu trúc một cách hợp lý để phỏng vấn viên hoặc người trả lời không cần phải lật qua lật lại giữa các trang của bảng câu hỏi. Ví dụ, bất kỳ ghi chú giải thích nào cũng nên được trình bày như một phần của câu hỏi mà chúng đề cập đến, chứ không phải trên một trang riêng biệt.

Những câu hỏi có thể nhạy cảm đối với người trả lời nhìn chung không nên được đặt ở phần đầu của bảng câu hỏi. Thay vào đó, chúng nên được đặt ở một phần của biểu mẫu nơi chúng có ý nghĩa nhất đối với bối cảnh của các câu hỏi khác. Bằng cách này, định dạng của bảng câu hỏi có thể đóng vai trò như một vùng đệm để giúp người trả lời cảm thấy thoải mái hơn với các câu hỏi nhạy cảm sau khi đã thiết lập được sự thiện cảm (rapport).

Trong các bảng câu hỏi tự điền (self-enumeration questionnaires), để đảm bảo rằng người trả lời chỉ trả lời những phần liên quan của bảng câu hỏi, các câu hỏi sàng lọc (filter questions) có thể được sử dụng để hướng dẫn người trả lời bỏ qua (skip) những câu hỏi không áp dụng cho họ. Các câu hỏi sàng lọc cũng được sử dụng trong các cuộc khảo sát có phỏng vấn viên (interviewer-based surveys) để hướng dẫn phỏng vấn viên tuân theo một chuỗi các câu hỏi dựa trên các câu trả lời do người trả lời cung cấp. Các câu hỏi sàng lọc cần được sử dụng một cách cẩn thận vì người trả lời (và các phỏng vấn viên) cần có

đủ thông tin về điều kiện bỏ qua (skip condition) để đánh giá xem người trả lời có nên bỏ qua hay không. Nhìn chung, cũng nên tránh sử dụng các câu hỏi sàng lọc cho các chủ đề nhạy cảm vì người trả lời sẽ có xu hướng đưa ra câu trả lời nhằm tránh việc phải trả lời các câu hỏi nhạy cảm đó.

Nếu các hướng dẫn không rõ ràng và dễ hiểu, phỏng vấn viên hoặc người trả lời có thể làm theo một trình tự không chính xác hoặc bỏ sót câu hỏi. Nhìn chung, chỉ nên đặt một hoặc hai điều kiện trong mỗi hướng dẫn trình tự. Việc phỏng vấn với sự trợ giúp của máy tính (computer-assisted interviewing) và các biểu mẫu điện tử tự điền có thể làm cho việc thiết lập trình tự phức tạp trở nên dễ dàng hơn nhiều.

Các câu hỏi sàng lọc cũng giúp xác định các tổng thể phụ (sub-populations).

Ví dụ: Câu 7 ‘Bạn có sinh ra ở nước ngoài không?’

Nếu ‘Có’ chuyển đến Câu 8, nếu ‘Không’ chuyển đến Câu 12.

Thứ tự các câu hỏi (Order of Questions)

Thứ tự xuất hiện của các câu hỏi có thể ảnh hưởng đến các câu trả lời được đưa ra cho những câu hỏi cụ thể. Các câu trả lời được đưa ra cho những câu hỏi trước đó có thể ảnh hưởng đến câu trả lời cho những câu hỏi sau đó. Ví dụ, nếu một câu hỏi hỏi người tham gia (participants) xem họ có tin rằng các tổ chức công đoàn (trade unions) gây ra sự xáo trộn trong cộng đồng hay không, và sau đó một câu hỏi tiếp theo hỏi về các vấn đề trong nền công nghiệp Úc, thì tác động tiêu cực của các tổ chức công đoàn đối với nền công nghiệp có thể nhận được sự chú ý lớn hơn nhiều so với trường hợp bình thường (khi không có câu hỏi dẫn dắt trước đó).

Thứ tự các phương án trả lời (Order of Response Options)

Thứ tự thực tế của các phương án trả lời cũng có thể gây ra sai số. Các phương án xuất hiện đầu tiên có thể được chọn vì chúng tạo ra tác động ban đầu đối với người

phản hồi, hoặc vì người phản hồi mất tập trung và không nghe hoặc đọc các phương án còn lại. Tương tự, các phương án cuối cùng có thể được chọn vì chúng dễ hỏi trông hơn, đặc biệt nếu người phản hồi được cung cấp một danh sách dài các phương án. Do đó, thứ tự các phương án trả lời có ảnh hưởng lớn hơn đến chất lượng dữ liệu khi một câu hỏi bao gồm số lượng lớn các phương án lựa chọn.

Nếu có thể, các phương án nên được trình bày theo một thứ tự có ý nghĩa. Nếu một số phương án được xã hội mong muốn hơn những phương án khác, chúng nên được đặt cuối cùng để giảm thiểu sai số. Ví dụ, một câu hỏi về giáo dục nên trình bày các bằng cấp theo thứ tự từ thấp đến cao. Đối với một số danh sách tự điền, thứ tự bảng chữ cái là phù hợp nhất để giúp người phản hồi tìm thấy phương án họ muốn, ví dụ như khi người phản hồi phải chọn loại cây trồng mà họ sản xuất.

Các phương án trả lời và Khó khăn của người phản hồi

Khi cuộc khảo sát có người phỏng vấn, các phương án trả lời có thể được trình bày bằng lời nói hoặc trên một thẻ gợi ý (prompt card). Thẻ gợi ý là một danh sách các câu trả lời khả thi cho một câu hỏi được người phỏng vấn đưa ra để hỗ trợ người phản hồi. Điều này giúp giảm thiểu các lỗi do người phản hồi không thể nhớ hết tất cả các phương án được đọc cho họ. Tuy nhiên, những người phản hồi có thị lực kém, người di cư hạn chế về tiếng Anh, hoặc người trưởng thành gặp vấn đề về đọc viết sẽ gặp khó khăn trong việc trả lời chính xác.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=64#h5p-23>

Độ dài (Length)

Độ dài của một bảng câu hỏi có thể được mô tả theo nhiều cách khác nhau. Các nhà thiết kế khảo sát thường bận tâm về số lượng trang, trong khi số lượng câu hỏi (đặc biệt là những câu hỏi bắt buộc) và thời gian cần thiết để hoàn thành thường quan trọng hơn. Bao lâu là quá dài sẽ thay đổi tùy thuộc vào phương thức thu thập dữ liệu (mode of data collection), độ dễ và mức độ thú vị của chủ đề cũng như thiết kế của bảng câu hỏi. Về cuối của một bảng câu hỏi dài, người trả lời có thể ít suy nghĩ hơn cho các câu trả lời của mình và ít tập trung hơn vào các hướng dẫn cũng như câu hỏi, qua đó làm giảm độ chính xác của các thông tin mà họ cung cấp. Nếu một người trả lời được thông báo rằng một cuộc phỏng vấn sẽ kéo dài trong vài giờ, hoặc họ nhận được một bảng câu hỏi dày nhiều trang, điều đó có thể dẫn đến việc người trả lời từ chối tham gia hoàn toàn.

Bố cục Bảng câu hỏi (Questionnaire Layout)

Người trả lời hoặc người phỏng vấn khi sử dụng một bảng câu hỏi có bố cục kém có thể bỏ sót câu hỏi, đi sai trình tự hoặc nhập câu trả lời vào nhầm ô phản hồi, dẫn đến việc thiếu dữ liệu hoặc dữ liệu sai lệch. Bố cục kém cũng góp phần gây ra lỗi ở giai đoạn xử lý. Ví dụ, nếu các ô trả lời không được căn chỉnh thẳng hàng, một số câu trả lời có thể bị bỏ sót hoàn toàn trong quá trình nhập liệu.

Đặc biệt đối với những bảng câu hỏi do người trả lời tự điền, thiết kế có thể gây ra lỗi do:

1. **Khả năng đọc kém** (ví dụ: bản in không rõ ràng hoặc văn bản quá nhỏ);
2. **Vì phạm luồng đọc thông thường** của người trả lời (người đọc tiếng Anh có thói quen đọc từ trên cùng bên trái sang dưới cùng bên phải; ví dụ: các tiêu đề lớn và hình ảnh rục rở ở giữa trang sẽ phá vỡ luồng này);
3. **Hướng dẫn dễ bị bỏ qua** (ví dụ: những hướng dẫn không được phân biệt rõ ràng với câu hỏi hoặc không được đặt gần phần liên quan của câu hỏi); và
4. **Không đủ khoảng trống** để viết câu trả lời.

Thiết kế Vật lý (Physical Design)

Bảng câu hỏi nên được trình bày về mặt vật lý sao cho giảm thiểu tối đa thời gian cần thiết để phỏng vấn, trả lời và xử lý kết quả. Cụ thể, cần xem xét đến cấu trúc, đồ họa và bố cục của biểu mẫu. Bố cục kém dẫn đến sai lầm trong việc hiểu câu hỏi và ghi chép câu trả lời. Nhìn chung, bảng câu hỏi:

1. **Cần phải được hiểu bởi người trả lời, người phỏng vấn và người xử lý dữ liệu.** Điều này được thực hiện bằng cách cung cấp hướng dẫn rõ ràng. Đối với người trả lời, điều này có nghĩa là bố cục câu hỏi và danh mục phản hồi phải đầy đủ với trình tự phù hợp. Người phỏng vấn cần các gợi ý (prompts) để hiểu và người xử lý cần các mã phản hồi (response codes) đầy đủ;
2. **Nên có hình thức đẹp** vì nó có thể ảnh hưởng đến phản hồi, tức là bảng câu hỏi nên được thiết kế và trình bày tốt, từ đó giúp việc trả lời trở nên dễ dàng;
3. **Nên xác định rõ ràng ngày tháng, tiêu đề và tổ chức thực hiện;**
4. **Nên phác thảo rõ ràng mục đích của cuộc khảo sát;**
5. **Nên đảm bảo với người trả lời về tính bảo mật** của thông tin mà họ cung cấp;
6. **Nên cung cấp số điện thoại liên hệ** để người trả lời có thể nhận được sự giúp đỡ nếu cần và nêu rõ hạn chót nộp phiếu; và
7. **Nên có các trang được đánh số liên tiếp** với hệ thống đánh số đơn giản.

Bước 7: Kiểm thử Bảng câu hỏi

Sau khi sự hiểu biết về các khái niệm và định nghĩa đã được khảo sát thông qua các nhóm thảo luận tập trung (focus groups), một bảng câu hỏi sơ bộ có thể được lập ra và kiểm thử một cách không chính thức trên một nhóm nhỏ người, chẳng hạn như đồng nghiệp tại nơi làm việc.

Việc kiểm thử như vậy không nhằm mục đích thu được kết quả mang tính đại diện, mà nhằm tìm ra các lỗi lớn trong câu hỏi, ví dụ như cách dùng từ vụng về. Quá trình này được thiết kế để loại bỏ các "góc cạnh thô ráp" (những thiếu sót cơ bản) của một bảng câu hỏi. Sử dụng các câu hỏi mở trong giai đoạn này là một ý hay để tìm ra các phản hồi tiềm năng. Sau đó, các câu hỏi có thể được tái cấu trúc và phát triển thành một bản dự thảo bảng câu hỏi để sử dụng trong các vòng tiền kiểm thử (pre-testing) không chính thức và kiểm thử thử nghiệm (pilot testing) sau này.

Kết luận

Thiết kế bảng câu hỏi bắt đầu bằng việc làm rõ các mục tiêu của cuộc khảo sát, xác định dữ liệu cần được tạo ra và thiết lập một danh sách các câu hỏi để thu thập dữ liệu này.

Cần cân nhắc kỹ lưỡng một số yếu tố, bao gồm loại câu hỏi được sử dụng, trình tự logic và cách dùng từ của câu hỏi, cũng như thiết kế vật lý của biểu mẫu. Việc kiểm thử từng khía cạnh này của thiết kế bảng câu hỏi với một nhóm đối tượng phản hồi trước khi hoàn thiện là rất quan trọng. Nếu cần thiết, biểu mẫu có thể được sửa đổi và kiểm thử lại cho đến khi người trả lời có thể hoàn thành chính xác và nhanh chóng với mức sai sót tối thiểu.

12. CÁC PHƯƠNG PHÁP CHỌN MẪU

Chương này được điều chỉnh từ [Cục Thống kê Úc](#), được cấp phép theo [giấy phép quốc tế Creative Commons Attribution 4.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu các bước trong một kế hoạch chọn mẫu.
- Phân biệt được các loại phương pháp chọn mẫu khác nhau.

Giới thiệu

Trong hầu hết các tình huống, nhà nghiên cứu không khả thi để thu thập thông tin từ tất cả mọi người trong một quần thể (thường được gọi là điều tra tổng điều tra - census). Việc có một kế hoạch về cách lựa chọn một "mẫu" gồm các cá nhân có thể khai thác để thu thập dữ liệu là rất quan trọng. Các quyết định chính cần đưa ra cho một "kế hoạch chọn mẫu" bao gồm:

- Xác định quần thể và các đơn vị chọn mẫu;
- Xác định khung chọn mẫu (sampling frame);
- Ước tính quy mô mẫu;
- Xác định hệ phương pháp chọn mẫu; và
- Lựa chọn kỹ thuật chọn mẫu.

Chương này sẽ xem xét chi tiết tất cả các chủ đề này.

Xác định quần thể và các đơn vị chọn mẫu

Quần thể (population) là tập hợp hoặc tập hợp các đơn vị (ví dụ: cá nhân, hộ gia đình, trường học, bệnh viện, doanh nghiệp) mà cuộc khảo sát sẽ được thực hiện. Quần thể mục tiêu (target population) còn được gọi là phạm vi của cuộc khảo sát. Đó chính là "quần thể" mà cuộc khảo sát hướng tới.

Một thành phần then chốt của việc chọn mẫu là xác định chính xác quần thể mà nhà nghiên cứu quan tâm khảo sát. Điều này đồng nghĩa với việc đảm bảo rằng định nghĩa bao gồm các đặc điểm chính của các đơn vị quần thể, giúp việc xác định chính xác nhóm cá nhân cần nhắm mục tiêu trở nên dễ dàng hơn.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=67#h5p-24>

Xác định khung chọn mẫu (Sampling frame)

Khung chọn mẫu là danh sách các đơn vị (ví dụ: cá nhân, hộ gia đình, doanh nghiệp, v.v.) trong quần thể khảo sát. Vì việc lựa chọn mẫu dựa trực tiếp trên danh sách này, khung chọn mẫu là một trong những công cụ quan trọng nhất trong thiết kế khảo sát. Nó quyết định mức độ bao phủ của quần thể mục tiêu và ảnh hưởng đến việc lựa chọn

phương pháp thu thập dữ liệu. Ngoài ra, khung chọn mẫu nên chứa các thông tin hỗ trợ về các đơn vị để có thể phát triển một kế hoạch chọn mẫu hiệu quả hơn. Khung chọn mẫu cần bao gồm các điểm liên lạc cho mỗi đơn vị được liệt kê để có thể tiếp cận quần thể. Điều này có nghĩa là đối với khảo sát qua thư, khung chọn mẫu phải có địa chỉ bưu điện; đối với khảo sát có người phỏng vấn, khung chọn mẫu phải có địa chỉ nhà; và đối với khảo sát qua điện thoại, khung chọn mẫu phải có số điện thoại.

Một khung chọn mẫu tốt là khung được cập nhật mới nhất, không bỏ sót đơn vị nào, chỉ bao gồm các đơn vị liên quan, không có sự trùng lặp, người dùng có thể tiếp cận được và chứa đủ thông tin để định danh duy nhất cũng như liên lạc được với từng đơn vị.

Xác định quy mô mẫu (Sample size determination)

Một khía cạnh quan trọng của thiết kế chọn mẫu là quyết định quy mô mẫu dựa trên các mục tiêu và hạn chế hiện có. Vì mỗi cuộc khảo sát đều khác nhau, không có quy tắc cố định nào để xác định quy mô mẫu. Tuy nhiên, các yếu tố cần xem xét bao gồm:

- Quy mô quần thể và mức độ biến thiên (variability) trong quần thể;
- Nguồn lực (thời gian, tiền bạc và nhân lực);
- Mức độ chính xác yêu cầu đối với kết quả;
- Mức độ chi tiết yêu cầu trong kết quả;
- Tỷ lệ không phản hồi (non-response) có khả năng xảy ra; và
- Tầm quan trọng tương đối của các biến số đang được quan tâm.

Khi các vấn đề này được giải quyết, bạn sẽ ở vị thế tốt hơn để quyết định quy mô của mẫu.

Mức độ biến thiên (Variability)

Tổng thể càng có tính biến thiên cao, thì kích thước mẫu (sample size) yêu cầu càng lớn để đạt được các mức độ chính xác cụ thể. Tuy nhiên, tính biến thiên thực tế của tổng thể thường không được biết trước; thông tin từ một cuộc khảo sát trước đó hoặc một cuộc thử nghiệm thí điểm (pilot test) có thể được sử dụng để đưa ra các chỉ báo về tính biến thiên của tổng thể.

Khi đặc điểm đang được đo lường tương đối hiếm gặp, kích thước mẫu lớn hơn sẽ được yêu cầu để đảm bảo rằng có đủ các đơn vị có đặc điểm đó được bao hàm trong mẫu nghiên cứu.

Quy mô quần thể (Population size)

Một khía cạnh ảnh hưởng đến quy mô mẫu cần thiết là quy mô quần thể. Khi quy mô quần thể nhỏ, cần cân nhắc kỹ lưỡng trong việc xác định quy mô mẫu; nhưng khi quy mô quần thể lớn, nó ít ảnh hưởng đến quy mô mẫu. Sự gia tăng độ chính xác từ việc tăng quy mô mẫu hoàn toàn không tỷ lệ thuận với quy mô quần thể.

Nguồn lực và độ chính xác (Resources and accuracy)

Như đã thảo luận trước đó, các ước tính được thu thập từ một mẫu thay vì từ tổng điều tra, do đó các con số ước tính sẽ có sự khác biệt so với giá trị thực của quần thể. Thước đo độ chính xác của ước tính chính là **sai số chuẩn (standard error)**. Một mẫu lớn có nhiều khả năng có sai số chuẩn nhỏ hơn hoặc có độ chính xác cao hơn so với một mẫu nhỏ.

Khi lập kế hoạch khảo sát, bạn có thể muốn tối thiểu hóa quy mô của sai số chuẩn để tối đa hóa độ chính xác của các ước tính. Điều này có thể thực hiện bằng cách chọn một mẫu lớn nhất trong phạm vi nguồn lực cho phép. Một cách khác là bạn có thể ấn định quy mô sai số chuẩn cần đạt được và chọn một quy mô mẫu được thiết kế để đạt tới mức đó. Trong một số trường hợp, chi phí để lấy một quy mô mẫu cần thiết nhằm đạt được mức độ chính xác nhất định sẽ quá cao. Khi đó, cần phải đưa ra quyết định

về việc nói lỏng mức độ chính xác, giảm yêu cầu về dữ liệu, tăng ngân sách hoặc cắt giảm chi phí ở các khâu khác trong quy trình khảo sát.

Mức độ chi tiết yêu cầu (Level of detail required)

Nếu chúng ta chia quần thể thành các nhóm phụ (phân tầng - strata) và chọn mẫu từ mỗi tầng này, thì cần có một quy mô mẫu đủ lớn trong mỗi nhóm phụ để đảm bảo các ước tính ở cấp độ này có độ tin cậy. Quy mô mẫu tổng thể sẽ bằng tổng quy mô mẫu của các nhóm phụ. Một phương pháp tốt là vẽ một bảng trống hiển thị tất cả các đặc điểm cần phân loại chéo (cross-classified). Bảng càng có nhiều ô (cells), quy mô mẫu cần thiết càng lớn để đảm bảo các ước tính đáng tin cậy.

Tỷ lệ không phản hồi có khả năng xảy ra (Likely level of non-response)

Tình trạng không phản hồi có thể gây ra vấn đề cho nhà nghiên cứu theo hai cách. Với một quy mô mẫu ban đầu cố định, tỷ lệ không phản hồi càng cao thì sai số chuẩn sẽ càng lớn. Điều này có thể được bù đắp bằng cách ấn định một quy mô mẫu lớn hơn dựa trên tỷ lệ phản hồi dự kiến, hoặc bằng cách sử dụng phương pháp chọn mẫu theo định mức (quota sampling).

Vấn đề thứ hai đối với những người không phản hồi là đặc điểm của họ có thể khác biệt rõ rệt so với những người có phản hồi. Kết quả khảo sát vẫn sẽ bị sai lệch ngay cả khi tăng quy mô mẫu (tức là việc tăng quy mô mẫu sẽ không có tác dụng đối với **thiên kiến không phản hồi - non-response bias**). Tỷ lệ phản hồi càng thấp, mẫu cuối cùng càng ít tính đại diện cho tổng quần thể và sai lệch của các ước tính mẫu càng lớn. Thiên kiến không phản hồi đôi khi có thể được giảm bớt bằng phương pháp hậu phân tầng (post-stratification) cũng như thông qua việc theo dõi sát sao những người không phản hồi, đặc biệt là ở các tầng có tỷ lệ phản hồi thấp.

Tầm quan trọng tương đối của các biến số quan tâm

Thông thường, các cuộc khảo sát được sử dụng để thu thập một dải dữ liệu về nhiều biến số quan tâm khác nhau. Một quy mô mẫu có thể cho ra thông tin đủ chính xác cho biến số này nhưng lại không đủ chính xác cho biến số khác. Thông thường, việc lựa chọn một mẫu đủ lớn để bao phủ tất cả các biến số ở mức độ chính xác mong muốn là không khả thi. Do đó, trong thực tế, tầm quan trọng tương đối của các biến số quan tâm sẽ được xem xét, các ưu tiên được thiết lập và quy mô mẫu phù hợp sẽ được xác định tương ứng.



Hệ phương pháp chọn mẫu: Xác suất và Phi xác suất

Mẫu xác suất (Probability sample) là mẫu mà trong đó mọi đơn vị của quần thể đều có xác suất được lựa chọn đã biết và khác không, đồng thời được lựa chọn một cách ngẫu nhiên. Mẫu xác suất cho phép thực hiện các suy diễn thống kê về quần thể mục tiêu. Bằng cách biết trước xác suất lựa chọn của mỗi đơn vị, nhà nghiên cứu có thể

thực hiện các lựa chọn khách quan để tạo ra một mẫu có tính đại diện cao hơn. Xác suất đã biết cũng cho phép đo lường độ chính xác của các ước tính khảo sát thông qua sai số chuẩn và khoảng tin cậy.

Nếu xác suất lựa chọn của mỗi đơn vị là không xác định hoặc không thể tính toán được, mẫu đó được gọi là **mẫu phi xác suất (Non-probability sample)**. Các mẫu phi xác suất thường ít tốn kém hơn, dễ thực hiện hơn và không yêu cầu khung chọn mẫu. Tuy nhiên, không thể đánh giá chính xác độ chính xác (tức là sự tiệm cận của các ước tính khi lặp lại việc chọn mẫu cùng quy mô) của các ước tính từ mẫu phi xác suất vì không có sự kiểm soát đối với tính đại diện của mẫu. Nếu một mẫu phi xác suất được thực hiện cẩn thận, sai lệch trong kết quả có thể được giảm thiểu.

Lựa chọn giữa chọn mẫu xác suất và phi xác suất

Việc lựa chọn giữa phương pháp tiếp cận chọn mẫu xác suất hay phi xác suất phụ thuộc vào nhiều yếu tố:

- Mục tiêu và phạm vi của cuộc khảo sát;
- Phương pháp thu thập dữ liệu phù hợp với các mục tiêu đó;
- Độ chính xác yêu cầu của kết quả và liệu độ chính xác đó có cần phải đo lường được hay không;
- Sự sẵn có của khung chọn mẫu;
- Nguồn lực cần thiết để duy trì khung chọn mẫu; và
- Sự sẵn có của thông tin bổ sung về các đơn vị trong quần thể.

Chọn mẫu xác suất thường được ưu tiên khi thực hiện các cuộc khảo sát lớn, đặc biệt là khi có sẵn khung quần thể đảm bảo rằng chúng ta có thể lựa chọn và liên lạc với từng đơn vị trong (khung) quần thể đó. Tuy nhiên, khi các hạn chế về thời gian và tài chính khiến việc chọn mẫu xác suất không khả thi, hoặc khi việc biết mức độ chính xác của kết quả không phải là yếu tố quan trọng, các mẫu phi xác suất vẫn có vai trò nhất định vì chúng rẻ, dễ vận hành và không cần khung chọn mẫu. Vì lý do này, khi

thực hiện nghiên cứu định tính (mang tính điều tra) thay vì nghiên cứu định lượng, các mẫu và kỹ thuật phi xác suất (như nghiên cứu trường hợp - case studies) thường ưu việt hơn so với mẫu xác suất và ước tính định lượng. Chọn mẫu phi xác suất cũng có thể hữu ích khi thực hiện các khảo sát thử nghiệm (piloting).

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=67#h5p-25>

Các kỹ thuật chọn mẫu – Mẫu phi xác suất

Các loại mẫu phi xác suất khác nhau được thảo luận dưới đây. Trong một mẫu phi xác suất, tất cả các thành viên của một quần thể không có cơ hội được lựa chọn ngang nhau vào mẫu. Tuy nhiên, nếu một mẫu phi xác suất được thực hiện cẩn thận, sai lệch trong kết quả có thể được giảm thiểu.

Chọn mẫu theo định mức (Quota sampling)

Để chọn mẫu theo định mức, người phỏng vấn lựa chọn những người phản hồi cho đến khi khảo sát đủ số lượng đã xác định trước trong các danh mục nhất định (ví dụ: người phỏng vấn có thể chọn mẫu để đạt được một cơ cấu độ tuổi/giới tính nhất định phản ánh đúng quần thể mục tiêu).

Đây là phương pháp chọn mẫu thường được sử dụng bởi các nhà nghiên cứu thị trường và những người thăm dò ý kiến chính trị vì nó có thể đưa ra các ước tính khá tốt nếu được thực hiện đúng cách. Khi các đơn vị bổ sung được lựa chọn ngẫu nhiên để lấp đầy định mức và nhà nghiên cứu không sử dụng yếu tố phán đoán nào để chọn đơn vị, phương pháp này rất giống với chọn mẫu xác suất. Tuy nhiên, khi tỷ lệ không phản hồi đáng kể (điều hầu như luôn xảy ra đối với các khảo sát tự nguyện), chọn mẫu theo định mức có thể phản ánh không đầy đủ những phần quần thể không sẵn lòng trả lời hoặc khó liên lạc.

Chọn mẫu thuận tiện và ngẫu nhiên (Convenience and haphazard sampling)

Phòng vấn tại góc phố, bảng câu hỏi trên tạp chí và báo giấy, và các cuộc thăm dò qua điện thoại đều là những ví dụ về mẫu thuận tiện hoặc ngẫu nhiên. Các loại khảo sát này dễ bị sai lệch hoặc mẫu không mang tính đại diện vì chỉ những người có quan điểm mạnh mẽ về chủ đề mới phản hồi. Những khảo sát này cũng có xu hướng đặt những câu hỏi mang tính định hướng hoặc có cách dùng từ thiên kiến.

Phòng vấn tại góc phố có thể bị sai lệch tùy thuộc vào thời gian và vị trí của người phỏng vấn. Không có sự kiểm soát nào đối với việc lựa chọn mẫu người phản hồi trong bất kỳ phương pháp nào nêu trên, tuy nhiên, chúng rất rẻ và dễ quản lý.

Chọn mẫu phán đoán hoặc chọn mẫu có mục đích (Judgment or purposive sampling)

Chọn mẫu phán đoán là phương pháp mà một mẫu "đại diện" được lựa chọn bởi một chuyên gia trong lĩnh vực nghiên cứu. Chọn mẫu phán đoán dễ chịu tác động của các sai lệch không xác định nhưng có thể được chấp nhận đối với các mẫu rất nhỏ. Hình thức chọn mẫu này có thể được sử dụng để chọn mẫu cho một bài kiểm tra thử nghiệm (pilot test) của một cuộc khảo sát xác suất, nhưng không nên đưa ra các suy

diễn về quần thể từ các mẫu phán đoán. Chọn mẫu phán đoán còn được gọi là chọn mẫu có mục đích.

Chọn mẫu quả cầu tuyết (hoặc chọn mẫu giới thiệu dây chuyền) (Snowball sampling)

Theo kỹ thuật chọn mẫu này, nhà nghiên cứu tìm kiếm sự hỗ trợ từ các đối tượng tham gia nghiên cứu để giúp xác định và tuyển dụng các đối tượng tham gia khác cho nghiên cứu. Điều này có thể đặc biệt hữu ích trong tình huống quần thể mục tiêu khó tìm kiếm. Ví dụ bao gồm các nhóm dân tộc thiểu số, những người mắc một tình trạng bệnh lý nhất định, hoặc những người có một hệ tư tưởng cụ thể.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=67#oembed-1>

Các kỹ thuật chọn mẫu – Mẫu xác suất

Chọn mẫu ngẫu nhiên đơn giản (Simple random sampling)

Chọn mẫu ngẫu nhiên đơn giản (SRS) là một cơ chế lựa chọn theo xác suất, trong đó mỗi đơn vị trong quần thể được cung cấp một xác suất lựa chọn ngang nhau; do đó, mọi mẫu khả thi với cùng một quy mô nhất định đều có xác suất được chọn như nhau. Một phương pháp khả thi để chọn mẫu ngẫu nhiên đơn giản là đánh số thứ tự từng đơn vị trên khung chọn mẫu và thực hiện lựa chọn bằng cách tạo ra các "số lựa chọn" từ bảng số ngẫu nhiên hoặc từ một dạng công cụ tạo số ngẫu nhiên nào đó.

Bảng: Ưu điểm và Nhược điểm – Chọn mẫu ngẫu nhiên đơn giản

Ưu điểm	Nhược điểm
Đơn giản; Dễ sử dụng	Yêu cầu phải có danh sách đầy đủ của tất cả các đơn vị trong quần thể
Lựa chọn ngẫu nhiên tuyệt đối nên không có thiên kiến	Chi phí cao; không khả thi để thực hiện

Trong thực tế, chọn mẫu ngẫu nhiên đơn giản hiếm khi được sử dụng vì hầu như luôn có những phương pháp thiết kế mẫu hiệu quả hơn (xét về khía cạnh tạo ra kết quả chính xác với một mức chi phí nhất định). Tuy nhiên, chọn mẫu ngẫu nhiên đơn giản vẫn là nền tảng cho nhiều phương pháp thiết kế mẫu phức tạp hơn và được sử dụng làm chuẩn mực để so sánh với các thiết kế khác.

Ví dụ về tình huống có thể áp dụng kỹ thuật chọn mẫu ngẫu nhiên đơn giản:

Để chọn người tham gia vào một cuộc thảo luận nhóm tập trung (focus group), một giáo viên có thể sử dụng máy tính để tạo ngẫu nhiên tên của 5 học sinh từ danh sách đầy đủ gồm 20 học sinh của cả lớp.

Chọn mẫu hệ thống (Systematic sampling)

Chọn mẫu hệ thống cung cấp một phương pháp đơn giản để lựa chọn mẫu khi khung chọn mẫu tồn tại dưới dạng một danh sách rõ ràng. Trong trường hợp khung chọn

mẫu có chứa các thông tin hỗ trợ, các đơn vị trong khung sẽ được sắp xếp theo thứ tự liên quan đến dữ liệu hỗ trợ đó (ví dụ: quy mô nhân sự của một doanh nghiệp). Một khoảng cách cố định (được gọi là **khoảng nhảy - skip**) sau đó sẽ được sử dụng để chọn các đơn vị từ khung chọn mẫu. Chọn mẫu hệ thống được giải thích tốt nhất bằng cách mô tả cách thức thực hiện việc lựa chọn mẫu.

Bảng: Ưu điểm và Nhược điểm – Chọn mẫu hệ thống

Ưu điểm	Nhược điểm
Dễ sử dụng	Yêu cầu phải có danh sách đầy đủ của tất cả các đơn vị trong quần thể
Kết quả chính xác hơn	Thiên kiến định kỳ (Periodicity Bias): Ví dụ doanh số bán hàng hàng ngày tại siêu thị dự kiến sẽ đạt đỉnh vào cuối tuần. Nếu khoảng nhảy (skip) được tính là 7, một sai lệch sẽ xuất hiện tạo ra các mẫu không đại diện cho quần thể.

Ví dụ về tình huống có thể áp dụng kỹ thuật chọn mẫu hệ thống: Khảo sát cứ mỗi hộ gia đình thứ năm trên một con phố.

Chọn mẫu phân tầng (Stratified sampling)

Chọn mẫu phân tầng là một kỹ thuật sử dụng thông tin bổ trợ (được gọi là các biến phân tầng) để tăng hiệu quả của thiết kế mẫu. Các biến phân tầng có thể mang tính địa lý (ví dụ: tiểu bang, nông thôn/thành thị) hoặc phi địa lý (ví dụ: tuổi tác, giới tính, số lượng nhân viên).

Chọn mẫu phân tầng bao gồm:

- Chia hoặc phân tầng quần thể thành các nhóm đồng nhất (tương tự nhau) được gọi là các **tầng (strata)**; và
- Lựa chọn mẫu bằng phương pháp chọn mẫu ngẫu nhiên đơn giản (SRS) hoặc chọn mẫu hệ thống bên trong mỗi tầng và độc lập với các tầng khác.

Việc phân tầng hầu như luôn cải thiện độ chính xác của các ước tính. Điều này là do sự biến thiên của quần thể có thể được coi là bao gồm các thành phần **nội tầng (within strata)** và **giữa các tầng (between strata)**. Bằng cách chọn mẫu độc lập trong mỗi tầng, chúng ta đảm bảo mỗi tầng được phản ánh một cách phù hợp trong mẫu, từ đó sự biến thiên giữa các tầng bị loại bỏ và chúng ta chỉ còn lại thành phần biến thiên nội tầng.

Với yếu tố này, chúng ta thấy rằng cách hiệu quả nhất để phân tầng là tạo ra các tầng khác biệt nhau càng nhiều càng tốt (để tối đa hóa phương sai bị loại bỏ) trong khi nội bộ mỗi tầng đồng nhất nhất có thể (để tối thiểu hóa phương sai còn lại).

Các cân nhắc thực tiễn

Khi lập kế hoạch chọn mẫu phân tầng, cần lưu ý một số vấn đề thực tiễn sau:

- Các tầng phải được thiết kế sao cho tập hợp của chúng bao hàm tất cả các thành viên của quần thể mục tiêu;

- Mỗi thành viên chỉ được xuất hiện trong duy nhất một tầng, tức là các tầng không được chồng lấn lên nhau; và
- Các định nghĩa về ranh giới của các tầng phải chính xác và không mơ hồ.

Ví dụ về phân tầng: Nếu chúng ta quan tâm đến nền tảng giáo dục của các thành viên trong một khoa Khoa học tại một trường đại học, chúng ta có thể chọn một mẫu từ toàn bộ khoa hoặc chọn các mẫu độc lập từ mỗi bộ môn trong khoa, chẳng hạn như toán học, vật lý, hóa học, v.v. Phương pháp sau sẽ đảm bảo rằng mỗi bộ môn đều được đại diện một cách đầy đủ (điều mà không nhất thiết sẽ xảy ra nếu chọn mẫu chung), và sẽ giúp tăng độ chính xác của ước tính tổng thể.

Mặt khác, nếu chúng ta quan tâm đến trình độ học vấn (Tiến sĩ, Thạc sĩ, Cử nhân) thay vì nền tảng chuyên môn, chúng ta nên phân tầng khoa theo cấp bậc (Giáo sư, Giảng viên cao cấp, Giảng viên) thay vì theo bộ môn. Sử dụng cách phân tầng này, chúng ta có nhiều khả năng tìm thấy sự đồng nhất về tiêu chuẩn giáo dục trong cùng một cấp bậc hơn là trong cùng một lĩnh vực công tác, và chúng ta cũng dễ dàng phân tách nhóm có trình độ cao hơn với nhóm có trình độ thấp hơn.

Bảng: Ưu điểm và Nhược điểm của Chọn mẫu Phân tầng

Ưu điểm	Nhược điểm
Đảm bảo các nhóm thiểu số được đưa vào: Giúp các nhóm nhỏ trong quần thể không bị bỏ sót.	Tăng chi phí: Việc quản lý và thu thập dữ liệu theo từng tầng riêng biệt đòi hỏi nhiều ngân sách hơn.
Kết quả chính xác hơn: Giảm sai số chọn mẫu bằng cách tận dụng sự đồng nhất trong nội bộ từng tầng.	Nguy cơ phân tầng quá chi tiết: Nếu chia tầng quá nhỏ sẽ làm phức tạp hóa quá trình nghiên cứu và giảm hiệu quả thống kê.
Phân tích được trên nhiều nhóm khác nhau: Cho phép so sánh và đối chiếu dữ liệu giữa các phân đoạn quần thể.	

Số lượng tầng (Number of strata)

Không có quy tắc cố định về việc nên chia quần thể thành bao nhiêu tầng. Điều này phụ thuộc vào quy mô, tính đồng nhất của quần thể và định dạng đầu ra yêu cầu. Nếu kết quả cần được báo cáo cho một số nhóm phụ nhất định của quần thể, thì các nhóm phụ này phải được xem xét như các tầng riêng biệt.

Các cuộc khảo sát của ABS

Tất cả các cuộc khảo sát do Cục Thống kê Úc (ABS) thực hiện đều áp dụng phương pháp phân tầng. Các cuộc khảo sát hộ gia đình (như Khảo sát Quần thể hàng tháng và Khảo sát Chỉ tiêu hộ gia đình) sử dụng các tầng địa lý. Các cuộc khảo sát doanh nghiệp sử dụng các biến số như tầng bang và tầng ngành, đồng thời sử dụng một số thước đo quy mô (ví dụ: số lượng nhân viên) để hình thành các tầng quy mô.

Chọn mẫu cụm và chọn mẫu nhiều giai đoạn (Cluster and multi-stage sampling)

Cho đến nay, chúng ta đã xem xét một số cách thức để lựa chọn một mẫu từ các đơn vị quần thể và ước tính các đặc điểm của quần thể dựa trên mẫu này. Trong phần này, chúng ta sẽ xem xét một cơ chế chọn mẫu mà việc lựa chọn các đơn vị quần thể được thực hiện bằng cách chọn các nhóm cụ thể (hoặc các cụm) của các đơn vị đó, sau đó chọn tất cả hoặc một số đơn vị quần thể bên trong các nhóm đã chọn để đưa vào mẫu.

Chọn mẫu cụm (Cluster sampling)

Chọn mẫu cụm bao gồm việc chọn mẫu qua một số giai đoạn (thường là hai). Các đơn vị trong quần thể được nhóm thành các cụm thuận tiện, thường là các cụm phát sinh tự nhiên. Các cụm này không chồng lấn, là các nhóm được xác định rõ ràng và thường đại diện cho các khu vực địa lý. Ở giai đoạn lựa chọn thứ nhất, một số lượng cụm nhất định sẽ được chọn. Ở giai đoạn thứ hai, tất cả các đơn vị trong các cụm đã chọn được lựa chọn để hình thành mẫu.

Các cân nhắc thực tiễn

- Các cụm phải được thiết kế sao cho tập hợp của chúng bao hàm tất cả các thành viên của quần thể mục tiêu;
- Mỗi thành viên phải xuất hiện trong một và chỉ một cụm duy nhất;

- Các định nghĩa hoặc ranh giới của các cụm phải chính xác và không mơ hồ; trong trường hợp các cụm địa lý, các ranh giới tự nhiên hoặc nhân tạo như sông ngòi và đường sá thường được sử dụng để phân định ranh giới cụm.

Bảng: Ưu điểm và Nhược điểm của Chọn mẫu Cụm

Ưu điểm	Nhược điểm
Chi phí thấp: Vì không phải tất cả các đơn vị của quần thể đều được đưa vào khảo sát, giúp tiết kiệm chi phí di chuyển và quản lý.	Độ chính xác thấp hơn: Kỹ thuật này thường kém chính xác hơn so với các kỹ thuật chọn mẫu xác suất khác (do tính đồng nhất trong cụm).

Nếu chúng ta lấy một mẫu ngẫu nhiên đơn giản gồm 10.000 hộ gia đình trên khắp nước Úc, chúng ta có nhiều khả năng bao phủ quần thể một cách đồng đều hơn, nhưng chi phí sẽ đắt hơn nhiều so với việc chọn mẫu 50 cụm (mỗi cụm gồm 200 hộ gia đình).

Chọn mẫu nhiều giai đoạn (Multi-stage sampling)

Chọn mẫu nhiều giai đoạn bao gồm việc chọn mẫu trong ít nhất hai giai đoạn. Ở giai đoạn đầu tiên, các nhóm hoặc cụm lớn của các đơn vị quần thể được chọn. Các cụm này được thiết kế để chứa nhiều đơn vị hơn mức cần thiết cho một mẫu cuối cùng.

Ở giai đoạn thứ hai, các đơn vị được lấy mẫu từ các cụm đã chọn để thu được mẫu cuối cùng. Nếu sử dụng nhiều hơn hai giai đoạn, quá trình lựa chọn các "cụm phụ" (sub-clusters) bên trong các cụm sẽ tiếp tục cho đến khi đạt được mẫu cuối cùng.

Các cân nhắc thực tiễn áp dụng cho chọn mẫu nhiều giai đoạn cũng tương tự như chọn mẫu cụm.

Dưới đây là ví dụ về các giai đoạn lựa chọn có thể được sử dụng trong một cuộc khảo sát hộ gia đình ba giai đoạn:

- **Giai đoạn 1: Các Phân khu Bầu cử.** Các phân khu bầu cử (các cụm) được lấy mẫu từ một thành phố hoặc tiểu bang.
- **Giai đoạn 2: Các khu phố (Blocks).** Các khu phố được chọn từ bên trong các phân khu bầu cử.
- **Giai đoạn 3: Các hộ gia đình.** Các hộ gia đình được chọn từ bên trong các khu phố đã chọn.

Ứng dụng của chọn mẫu nhiều giai đoạn

Chọn mẫu nhiều giai đoạn thường được sử dụng khi việc lập danh sách tất cả các đơn vị trong quần thể mục tiêu quá tốn kém hoặc bất khả thi. Thông thường, một mẫu nhiều giai đoạn cho ra các ước tính kém chính xác hơn so với một mẫu ngẫu nhiên đơn giản có cùng quy mô. Tuy nhiên, một mẫu nhiều giai đoạn thường chính xác hơn một mẫu ngẫu nhiên đơn giản **có cùng mức chi phí**, và chính vì lý do này mà phương pháp này được áp dụng.

Ưu điểm và nhược điểm

Những ưu điểm và nhược điểm của chọn mẫu nhiều giai đoạn tương tự như đối với chọn mẫu cụm. Tuy nhiên, để bù đắp cho độ chính xác thấp hơn, bạn cần phải thực hiện một trong hai cách:

1. Số lượng các cụm được chọn ở giai đoạn đầu tiên phải tương đối lớn (nhưng điều này sẽ làm tăng chi phí khảo sát); hoặc
2. Phân số chọn mẫu (sampling fraction) cho các giai đoạn sau phải cao (tức là phải chọn một tỷ lệ phần trăm lớn các đơn vị trong mỗi cụm).

Chọn mẫu River và chọn mẫu Panel (River sampling and panel sampling)

(Phần này được trích xuất từ tài liệu của Lehdonvirta, Oksanen, Rasanen và Blank (2020) [1], được cấp phép theo giấy phép [Creative Commons 4.0 Quốc tế - CC BY 4.0](#))

Có hai cách tiếp cận phổ biến được sử dụng trong các khảo sát trực tuyến phi xác suất. Cách tiếp cận phi xác suất đơn giản nhất để tuyển dụng người phản hồi trực tuyến là **chọn mẫu "river" (dòng chảy)**, còn được gọi là chọn mẫu đánh chặn (intercept sampling) hoặc chọn mẫu thời gian thực (real-time sampling) (Olivier, 2011; Walsh và cộng sự, 1992; Yun & Trumbo, 2000, trích dẫn trong [Lehdonvirta, Oksanen, Rasanen và Blank 2020](#)). Cách tiếp cận thứ hai được gọi là **chọn mẫu panel**.

Chọn mẫu River có nghĩa là tuyển dụng người phản hồi bằng cách mời họ nhấp vào một liên kết dẫn đến một cuộc khảo sát được đặt trên một trang web, email hoặc một nơi nào đó có khả năng được các thành viên của quần thể mục tiêu chú ý tới.

Tên gọi này bắt nguồn từ ý tưởng các nhà nghiên cứu "nhúng tay" vào lưu lượng truy cập (traffic flow) của một trang web, "bắt" lấy một số người dùng đang lướt qua.

Phương pháp này tương tự như phương pháp chọn mẫu thuận tiện hoặc ngẫu nhiên đã được mô tả ở trên.

Một vấn đề cơ bản của phương pháp chọn mẫu River là **thiên kiến bao phủ (coverage bias)**. Thiên kiến bao phủ xảy ra do không phải mọi nhóm dân cư phụ (subpopulation) đều được đại diện một cách tương xứng, hoặc thậm chí là không hề xuất hiện trên các phương tiện truyền thông kỹ thuật số (Räsänen, 2006, trích dẫn trong [Lehdonvirta, Oksanen, Rasanen và Blank 2020](#)).

Chọn mẫu Panel đề cập đến việc tuyển dụng người trả lời khảo sát thường thông qua sự hỗ trợ của các nhà cung cấp panel trực tuyến phi xác suất mang tính thương mại. Ngày nay có rất nhiều nhà cung cấp như vậy, từ các công ty khởi nghiệp mới đến các tổ chức truyền thông và nghiên cứu lâu đời như Ipsos Mori, Qualtrics và Survey Sampling International (SSI) (Callegaro và cộng sự, 2014a, trích dẫn trong [Lehdonvirta, Oksanen, Rasanen và Blank 2020](#)).

Trái ngược với các panel trực tuyến dựa trên xác suất, chiến lược tuyển dụng điển hình của các panel trực tuyến phi xác suất bao gồm việc đặt quảng cáo trên các trang web hoặc phương tiện truyền thông xã hội và phân phối lời mời tới các nhóm tin tức và danh sách gửi thư. Những người dùng quan tâm sẽ tự nguyện đăng ký (opt-in) để trở thành thành viên của panel, và người dùng từ nhiều nguồn khác nhau thường được hòa trộn vào một panel duy nhất (Lorch, Cavallaro, & van Ossenbruggen, 2014, trích dẫn trong [Lehdonvirta, Oksanen, Rasanen và Blank 2020](#)).

Nguồn tham khảo / Chú thích: [1] Lehdonvirta, V, Oksanen, A, Rasanen, P and Blank, G, '[Social media, web and panel surveys: using non-probability samples in social and policy research](#)', *Policy and Internet journal*, tập 13, số 1, trang 134 – 155.

Một điểm khác biệt chính so với chọn mẫu River là các nhà cung cấp panel cam kết quản lý cấu trúc nhân khẩu học của các tệp người phản hồi của họ. Họ cố gắng khắc phục các sai lệch xuất phát từ mức độ bao phủ không đồng đều của phương tiện truyền thông kỹ thuật số, nhằm đảm bảo các panel có đặc điểm nhân khẩu học tương tự như quần thể quốc gia.

Các khảo sát panel trực tuyến được sử dụng rộng rãi bởi các nhà khoa học xã hội cũng như các chuyên gia tư vấn chính sách và các viện nghiên cứu (think tanks). Tuy nhiên, các nghiên cứu trước đây chỉ ra rằng các khảo sát panel vẫn có tính hệ thống trong việc đại diện vượt mức (over-represent) cho một số nhóm và đại diện thiếu (under-represent) cho các nhóm khác (Willems và cộng sự, 2006, trích dẫn trong [Lehdonvirta, Oksanen, Rasanen và Blank 2020](#)). Ví dụ, các khảo sát panel trực tuyến của Hoa Kỳ thường đại diện vượt mức cho nhóm người da trắng, có học vấn cao hơn và là những người sử dụng Internet tích cực (Chang & Krosnick, 2009; Dever, Rafferty, & Valliant, 2008; Malhotra & Krosnick, 2007, trích dẫn trong [Lehdonvirta, Oksanen, Rasanen và Blank 2020](#)).

Ghi danh các phương tiện truyền thông (Media Attributions)

- *Population sample* © Nghị viện Úc ([Parliament of Australia](#)) được cấp phép theo giấy phép CC BY-NC (Ghi công - Phi thương mại).

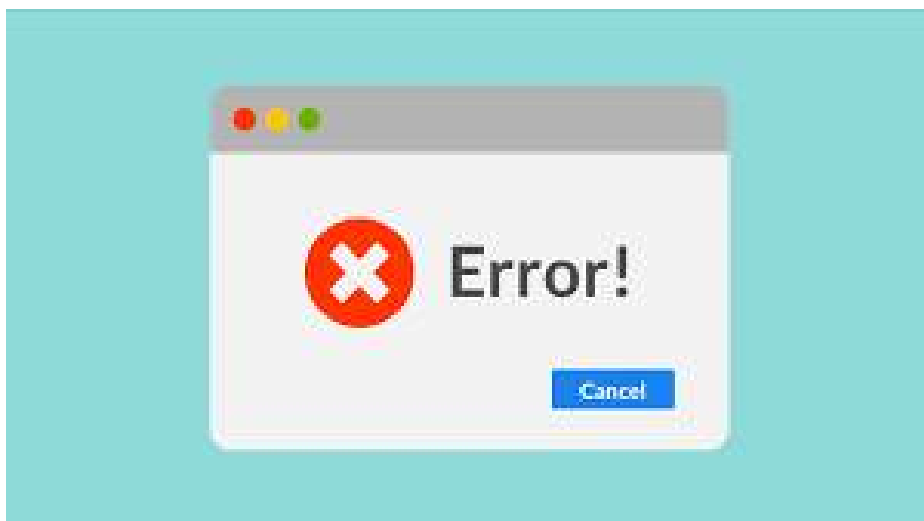
13. SAI SỐ TRONG NGHIÊN CỨU

Chương này được điều chỉnh từ [Cục Thống kê Úc](#), được cấp phép theo giấy phép quốc tế [Creative Commons Attribution 4.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Giải thích được sai số chọn mẫu (sampling errors) và sai số ngoài chọn mẫu (non-sampling errors) trong nghiên cứu;
- Hiểu được các nguồn gốc dẫn đến những sai số đó;
- Thể hiện được sự hiểu biết về các chiến lược nhằm giảm thiểu các sai số đó.



Giới thiệu

Độ chính xác của một ước lượng khảo sát đề cập đến mức độ tiệm cận của ước lượng đó so với giá trị thực của quần thể. Khi có sự sai lệch giữa giá trị của ước lượng khảo sát và giá trị thực của quần thể, sự khác biệt giữa hai giá trị này được gọi là **sai số của ước lượng khảo sát (error of the survey estimate)**. Tổng sai số của ước lượng khảo sát bắt nguồn từ hai loại sai số:

- **Sai số chọn mẫu (Sampling error):** Phát sinh khi chỉ một phần của quần thể được sử dụng để đại diện cho toàn bộ quần thể. Sai số chọn mẫu có thể được đo lường bằng các phương pháp toán học.
- **Sai số ngoài chọn mẫu (Non-sampling error):** Có thể xảy ra ở bất kỳ giai đoạn nào của một cuộc khảo sát chọn mẫu.

Việc nhà nghiên cứu nhận thức được những sai số này, đặc biệt là sai số ngoài chọn mẫu, là rất quan trọng để có thể giảm thiểu hoặc loại bỏ chúng khỏi cuộc khảo sát. Phần giới thiệu chi tiết về sai số chọn mẫu và sai số ngoài chọn mẫu sẽ được trình bày trong các phần tiếp theo.

Sai số chọn mẫu (Sampling error)

Sai số chọn mẫu phản ánh sự khác biệt giữa một ước lượng (ví dụ: giá trị trung bình) thu được từ một cuộc khảo sát chọn mẫu và "giá trị thực" (tức là trung bình thực tế của quần thể) có được nếu toàn bộ quần thể khảo sát được điều tra.

Sai số này có thể được đo lường từ các giá trị của quần thể, nhưng vì những giá trị này thường không xác định (hoặc rất khó tính toán), nó cũng có thể được ước tính từ dữ liệu mẫu. Việc xem xét sai số chọn mẫu khi công bố kết quả khảo sát là rất quan trọng vì nó cung cấp một chỉ báo về độ chính xác của ước lượng, và do đó phản ánh mức độ tin cậy có thể đặt vào các diễn giải kết quả.

Nếu các nguyên tắc chọn mẫu được áp dụng cẩn thận trong giới hạn của các nguồn lực sẵn có, sai số chọn mẫu có thể được đo lường chính xác và được giữ ở mức tối thiểu.

Các yếu tố ảnh hưởng đến Sai số chọn mẫu

Sai số chọn mẫu bị ảnh hưởng bởi một số yếu tố bao gồm: **quy mô mẫu**, **thiết kế mẫu**, **phân số chọn mẫu (sampling fraction)**, và **sự biến thiên trong quần thể**.

- **Quy mô mẫu (Sample size):** Nhìn chung, quy mô mẫu lớn hơn sẽ làm giảm sai số chọn mẫu; tuy nhiên, sự sụt giảm này không tỷ lệ thuận. Theo một quy tắc kinh nghiệm căn bản (rule of thumb), bạn cần *tăng quy mô mẫu lên gấp bốn lần để giảm một nửa sai số chọn mẫu*.
- **Phân số chọn mẫu (Sampling fraction):** Phân số chọn mẫu (tỷ lệ của quy mô mẫu trên quy mô quần thể) có mức độ ảnh hưởng ít hơn nhiều. Tuy nhiên, khi quy mô mẫu tăng lên và chiếm tỷ lệ lớn hơn trong quần thể, sai số chọn mẫu sẽ giảm xuống.
- **Sự biến thiên của quần thể (Population variability):** Sự biến thiên cũng ảnh hưởng đến sai số chọn mẫu. Các quần thể có mức độ biến thiên cao hơn sẽ gây ra sai số lớn hơn, vì các mẫu hoặc các ước lượng được tính toán từ các mẫu khác nhau có nhiều khả năng dao động lớn hơn. Tác động của sự biến thiên trong nội bộ quần thể có thể được giảm bớt bằng cách tăng quy mô mẫu để mẫu mang tính đại diện cao hơn cho quần thể khảo sát.
- **Thiết kế mẫu (Sample design):** Các lựa chọn thiết kế mẫu khác nhau cũng ảnh hưởng đến độ lớn của sai số chọn mẫu. Ví dụ, chọn mẫu phân tầng giúp làm giảm sai số chọn mẫu, trong khi chọn mẫu cụm lại có xu hướng làm tăng sai số này.

Sai số chọn mẫu có thể được ước tính về mặt thống kê và được sử dụng trong quá trình phân tích, diễn giải các kết quả thống kê.

Video tham khảo (Tiếng Anh): <https://www.youtube.com/watch?v=XE7QDfdaQ68>

Sai số phi lấy mẫu (Non-sampling error)

Sai số phi lấy mẫu bao gồm tất cả các loại sai số khác trong các ước lượng (ví dụ: số trung bình mẫu). Những sai số này bao gồm toàn bộ các sai sót xảy ra do những nguyên nhân không phải từ kế hoạch lấy mẫu hay kích thước mẫu. Một số ví dụ về nguyên nhân gây ra sai số phi lấy mẫu là tỷ lệ phản hồi bảng câu hỏi thấp, bảng câu hỏi được thiết kế kém, thiên lệch từ phía người trả lời và các sai số trong quá trình xử lý. Sai số phi lấy mẫu có thể xảy ra ở bất kỳ giai đoạn nào của quy trình. Những sai số này có thể tìm thấy trong cả các cuộc tổng điều tra toàn bộ (censuses) và các cuộc điều tra chọn mẫu (sample surveys).

Các nguồn gây ra sai số phi lấy mẫu sẽ được thảo luận dưới đây.

Thiên kiến Không phản hồi (Non-Response Bias)

Không phản hồi đề cập đến tình huống khi người được hỏi không trả lời bất kỳ câu hỏi khảo sát nào (tức là **không phản hồi toàn bộ - total non-response**) hoặc không trả lời một số câu hỏi khảo sát do câu hỏi quá nhạy cảm, vấn đề về hội tượng, không có khả năng trả lời, v.v. (**không phản hồi một phần - partial non-response**).

Để cải thiện tỷ lệ phản hồi, cần chú trọng đến việc thiết kế bảng câu hỏi, đào tạo người phỏng vấn, đảm bảo tính bảo mật cho người phản hồi và gọi lại vào các thời điểm khác nhau nếu gặp khó khăn khi liên lạc. Việc "gọi lại" (call-backs) rất hiệu quả trong việc giảm thiểu tình trạng không phản hồi nhưng có thể tốn kém đối với các cuộc phỏng vấn trực tiếp. Một email nhắc nhở nhẹ nhàng đối với các khảo sát trực tuyến cũng được sử dụng như một công cụ hữu ích để cải thiện tỷ lệ phản hồi.

Các vấn đề về Bảng câu hỏi (Questionnaire problems)

Nội dung và cách diễn đạt của bảng câu hỏi có thể gây hiểu lầm và bố cục của bảng câu hỏi có thể gây khó khăn cho việc ghi nhận chính xác các câu trả lời. Các câu hỏi không nên mang tính dẫn dắt (loaded), hỏi nhiều vấn đề cùng lúc (double-barrelled), gây hiểu lầm hoặc mơ hồ, và phải liên quan trực tiếp đến các mục tiêu của cuộc khảo sát.

Việc kiểm thử bảng câu hỏi trên một mẫu người phản hồi trước khi hoàn thiện là điều tối quan trọng nhằm xác định các vấn đề về mạch khảo sát (questionnaire flow) và cách dùng từ, đồng thời dự trù đủ thời gian để thực hiện các cải tiến cho bảng câu hỏi. Sau đó, bảng câu hỏi nên được kiểm thử lại để đảm bảo những thay đổi vừa thực hiện không tạo ra các vấn đề mới.

Thiên kiến của người phản hồi (Respondent Bias)

Đôi khi, người phản hồi có thể cung cấp thông tin không chính xác vì họ tin rằng họ đang bảo vệ lợi ích cá nhân và sự toàn vẹn của mình. Việc thiết kế bảng câu hỏi cẩn thận và kiểm thử bảng câu hỏi hiệu quả có thể khắc phục những vấn đề này ở một mức độ nào đó. Dưới đây là hai loại tình huống có thể tránh được thông qua việc thiết kế và thực hiện khảo sát tốt hơn.

Nguồn tham khảo / Chú thích: [1] Chumney, 2016, *Introduction to error sources in survey research: sampling error*, 14 tháng 8, video trực tuyến, truy cập ngày 4 tháng 4 năm 2022.

- **Sự nhạy cảm (Sensitivity)** Nếu người phản hồi phải đối mặt với một câu hỏi mà họ thấy bối rối, họ có thể từ chối trả lời, hoặc chọn một phương án trả lời giúp họ không phải tiếp tục với các câu hỏi đó. Ví dụ, nếu được hỏi: *"Bạn có đang sử dụng thuốc tránh thai đường uống vì bất kỳ lý do gì không?"*, và biết rằng nếu trả lời "Có", họ sẽ bị hỏi thêm chi tiết, thì những người phản hồi cảm thấy bối rối với câu hỏi này có khả năng sẽ trả lời "Không", ngay cả khi điều đó không đúng.
 - **Sự mệt mỏi (Fatigue)** Sự mệt mỏi có thể là một vấn đề trong các cuộc khảo sát đòi hỏi mức độ cam kết cao từ người phản hồi. Ví dụ, các cuộc khảo sát dạng nhật ký (diary surveys), nơi người phản hồi phải ghi lại tất cả các khoản chi tiêu trong khoảng thời gian hai tuần. Trong các loại khảo sát này, mức độ chính xác và chi tiết được cung cấp có thể giảm xuống khi người phản hồi trở nên mệt mỏi với việc phải ghi chép lại mọi khoản chi tiêu.
-

Lỗi xử lý dữ liệu (Processing Errors)

Có bốn giai đoạn trong quá trình xử lý dữ liệu mà lỗi có thể xảy ra: tinh chỉnh dữ liệu (data grooming), nhập liệu (data capture), chỉnh sửa (editing) và ước lượng (estimation).

Việc tinh chỉnh dữ liệu bao gồm kiểm tra sơ bộ trước khi nhập dữ liệu vào hệ thống xử lý trong giai đoạn nhập liệu. Việc kiểm tra và quản lý chất lượng không đầy đủ ở giai đoạn này có thể dẫn đến **mất mát dữ liệu** (khi dữ liệu không được nhập vào hệ thống) và **trùng lặp dữ liệu** (khi cùng một dữ liệu được nhập vào hệ thống nhiều lần). Các quy tắc kiểm tra chỉnh sửa không phù hợp và việc gán trọng số không chính xác trong quy trình ước lượng cũng có thể tạo ra lỗi cho dữ liệu. Để giảm thiểu những lỗi

này, nhân viên xử lý dữ liệu cần được đào tạo đầy đủ và giao một khối lượng công việc thực tế.

Diễn giải sai kết quả (Misinterpretation of Results)

Điều này có thể xảy ra nếu nhà nghiên cứu không nhận thức được các yếu tố nhất định ảnh hưởng đến các đặc điểm đang được điều tra. Một nhà nghiên cứu hoặc bất kỳ người dùng nào khác không tham gia vào giai đoạn thu thập dữ liệu có thể không nhận thức được các xu hướng được hình thành trong dữ liệu do bản chất của việc thu thập, chẳng hạn như phạm vi (scope) của nó.

(Ví dụ: một cuộc khảo sát thu thập dữ liệu về thu nhập với phạm vi bao phủ là tất cả những người trưởng thành (tức là 18 tuổi trở lên), dự kiến sẽ tạo ra một mức ước lượng khác với mức ước lượng của Khảo sát Thu nhập Hàng tuần Trung bình (AWE) của ABS, đơn giản vì AWE bao gồm cả những người 16 và 17 tuổi trong phạm vi của nó). Các nhà nghiên cứu nên tìm hiểu cẩn thận về hệ phương pháp được sử dụng trong bất kỳ cuộc khảo sát cụ thể nào.

Thiên kiến về khoảng thời gian (Time Period Bias)

Điều này xảy ra khi một cuộc khảo sát được tiến hành trong một khoảng thời gian không mang tính đại diện. Ví dụ, nếu một cuộc khảo sát nhằm mục đích thu thập thông tin chi tiết về doanh số bán kem, nhưng lại chỉ thu thập dữ liệu của một tuần trong khoảng thời gian nóng nhất của mùa hè, thì dữ liệu đó khó có thể đại diện cho doanh số bán kem trung bình hàng tuần trong suốt cả năm.

Video tham khảo (Tiếng Anh): <https://www.youtube.com/watch?v=zF37RvnNHnk>

Nguồn tham khảo / Chú thích: [2] Chumney, 2016, *Introduction to error sources in survey research: measurement errors*, 14 tháng 8, video trực tuyến, truy cập ngày 4 tháng 4 năm 2022.

Giảm thiểu sai số ngoài chọn mẫu (Minimising non-sampling error)

Sai số ngoài chọn mẫu có thể rất khó để đo lường chính xác, nhưng có thể được giảm thiểu thông qua:

- Lựa chọn cẩn thận thời điểm tiến hành khảo sát;
 - Sử dụng một khung chọn mẫu được cập nhật và chính xác;
 - Có kế hoạch theo dõi (follow-up) những người không phản hồi;
 - Thiết kế bảng câu hỏi một cách cẩn trọng;
 - Cung cấp đào tạo kỹ lưỡng cho đội ngũ người phỏng vấn và nhân viên xử lý dữ liệu; và
 - Nhận thức được tất cả các yếu tố ảnh hưởng đến chủ đề đang được nghiên cứu.
-

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=70#h5p-26>

Vì nhiều cuộc khảo sát phải đối mặt với tình trạng tỷ lệ phản hồi kém, chúng ta sẽ đặc biệt thảo luận về các cách để giảm thiểu việc không phản hồi từ những đối tượng tiềm năng.

Giảm thiểu tình trạng Không phản hồi (Minimising Non-Response)

Tỷ lệ phản hồi có thể được cải thiện thông qua việc thiết kế khảo sát tốt với các câu hỏi ngắn gọn, đơn giản, các kỹ thuật thiết kế biểu mẫu tốt và giải thích rõ mục đích cũng như tính ứng dụng của cuộc khảo sát.

Những cam kết về tính bảo mật là cực kỳ quan trọng vì nhiều người không sẵn lòng trả lời do lo ngại sự thiếu riêng tư. Việc theo dõi có mục tiêu (targeted follow-ups) đối với những người không thể liên lạc được hoặc những người ban đầu không thể trả lời có thể làm tăng tỷ lệ phản hồi một cách đáng kể. Dưới đây là một số mẹo về cách giảm thiểu sự từ chối trong một cuộc phỏng vấn trực tiếp hoặc qua điện thoại:

- **Tìm hiểu lý do từ chối** và cố gắng trò chuyện để giải quyết những khúc mắc đó.
- **Sử dụng ngôn ngữ tích cực.**
- **Nêu rõ cách thức và những gì bạn dự định làm** để giúp họ hoàn thành bảng câu hỏi.
- **Nhấn mạnh tầm quan trọng** của cuộc khảo sát.
- **Giải thích tầm quan trọng trong câu trả lời của họ** với tư cách là người đại diện cho các đơn vị khác (trong mẫu).
- **Nhấn mạnh những lợi ích từ kết quả khảo sát**, giải thích cách họ có thể nhận được bản tóm tắt kết quả.
- **Đưa ra cam kết về tính bảo mật** cho các câu trả lời.

Các biện pháp khác có thể cải thiện sự hợp tác của người trả lời và tối đa hóa tỷ lệ phản hồi bao gồm:

- **Các hoạt động nâng cao nhận thức cộng đồng** bao gồm thảo luận với các tổ chức và nhóm lợi ích trọng điểm, phát hành thông cáo báo chí, phỏng vấn trên phương tiện truyền thông và các bài báo. Điều này nhằm mục đích thông báo

cho cộng đồng về cuộc khảo sát, xác định các vấn đề đáng quan tâm và giải đáp chúng.

- **Gửi thư thông báo cho các đơn vị được chọn**, cho họ biết trước, giải thích các mục đích của cuộc khảo sát và cách thức khảo sát sẽ được tiến hành.

Trong trường hợp khảo sát qua thư, hầu hết các điểm trên có thể được nêu rõ trong một bức thư giới thiệu (introductory letter) hoặc thông qua một chiến dịch truyền thông quảng bá.

Dự trù cho việc Không phản hồi (Allowing for Non-Response)

Trong trường hợp tỷ lệ phản hồi vẫn ở mức thấp sau khi đã thực hiện mọi nỗ lực theo dõi (follow-up) hợp lý, bạn có thể giảm thiểu sai lệch bằng cách sử dụng các điểm chuẩn của quần thể (population benchmarks) để **hậu phân tầng (post-stratify)** mẫu, theo dõi chuyên sâu một mẫu phụ của những người không phản hồi, hoặc **gán dữ liệu (imputation)** cho tình trạng không phản hồi từng phần (không trả lời một câu hỏi cụ thể).

Mục đích chính của việc gán dữ liệu là tạo ra dữ liệu nhất quán mà không cần quay lại gặp người phản hồi để lấy các giá trị chính xác, từ đó giảm bớt gánh nặng cho người phản hồi cũng như các chi phí liên quan đến cuộc khảo sát. Nói một cách khái quát, các phương pháp gán dữ liệu được chia thành ba nhóm:

- Giá trị được gán được suy ra từ các thông tin khác do chính đơn vị đó cung cấp;
- Các giá trị của các đơn vị khác có thể được sử dụng để tính ra giá trị cho người không phản hồi (ví dụ: gán bằng giá trị trung bình);
- Giá trị chính xác của một đơn vị khác (được gọi là đơn vị cho - *donor*) được sử dụng làm giá trị cho người không phản hồi (được gọi là đơn vị nhận - *recipient*).

Khi quyết định chọn phương pháp gán dữ liệu, cần phải biết việc gán dữ liệu sẽ có tác động như thế nào đến các ước lượng cuối cùng. Nếu thực hiện gán dữ liệu với số lượng lớn, kết quả có thể gây hiểu lầm, đặc biệt là nếu phương pháp gán được sử dụng làm biến dạng sự phân bố của dữ liệu.

Nếu ở giai đoạn lập kế hoạch, nhà nghiên cứu tin rằng có khả năng xảy ra tỷ lệ không phản hồi cao, thì quy mô mẫu có thể được tăng lên để dự trù cho điều này. Tuy nhiên, thiên kiến không phản hồi (non-response bias) sẽ không được khắc phục chỉ bằng cách tăng quy mô mẫu, đặc biệt là nếu các đơn vị không phản hồi có những đặc điểm khác với các đơn vị có phản hồi. Việc hậu phân tầng và gán dữ liệu cũng không thể loại bỏ hoàn toàn thiên kiến không phản hồi khỏi các kết quả.

Mẫu: Tác động của việc Không phản hồi

Giả sử một cuộc khảo sát qua đường bưu điện đối với 3.421 người trồng cây ăn quả được thực hiện để ước tính số lượng cây ăn quả trung bình trên một trang trại. Ban đầu có một khoảng thời gian chờ phản hồi, và do tỷ lệ phản hồi thấp, hai đợt gửi thư nhắc nhở (follow-up reminders) đã được thực hiện. Lượng phản hồi và kết quả thu được như sau:

	Lượn phản hồi	Số cây trung bình
Phản hồi ban đầu	300	456
Phản hồi thêm sau 1 lần nhắc nhở	543	382

Phản hồi thêm sau 2 lần nhắc nhở	434	340
Tổng phản hồi	1277	

Sau hai lần nhắc nhở, tỷ lệ phản hồi vẫn chỉ đạt 37%. Từ các nguồn thông tin khác, người ta biết rằng mức trung bình tổng thể thực sự là 329. Kết quả ước lượng diễn tiến qua các đợt của cuộc khảo sát này như sau:

	Phản hồi lũy kế	Trung bình gộp
Phản hồi ban đầu	300	456
Lũy kế sau 1 lần nhắc nhở	843	408
Lũy kế sau 2 lần nhắc nhở	1277	385

Nếu kết quả được công bố mà không có bất kỳ sự nhắc nhở nào, số lượng cây trung bình sẽ bị ước lượng quá cao vì các trang trại có số lượng cây lớn hơn đường như có xu hướng phản hồi nhanh chóng và sẵn sàng hơn. Nhờ có việc theo dõi nhắc nhở, các

trang trại nhỏ hơn đã gửi lại biểu mẫu khảo sát và mức ước lượng đã dịch chuyển sát với giá trị thực tế hơn.

Nguồn trích dẫn Phương tiện (Media Attributions)

- Thông báo lỗi (**Error message**) © [istockphoto](#) được cấp phép theo giấy phép [CC BY-NC \(Ghi công - Phi thương mại / Attribution NonCommercial\)](#).

14. CÁC PANEL NGHIÊN CỨU

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu khái niệm về panel nghiên cứu;
- Nhận thức được những chỉ trích đối với việc sử dụng các panel nghiên cứu;
- Giải thích quy trình liên quan đến việc quản lý các panel nghiên cứu.



Panel Surveys

Image by kvector/shutterstock.com 1

Nhóm nghiên cứu là gì? (What is a research panel?)

Một nhóm nghiên cứu (Panel) là một nhóm người đã đồng ý tham gia vào các nghiên cứu trong tương lai sau khi được sàng lọc và lập hồ sơ từ trước (pre-screened and pre-profiled). Họ cũng được phổ biến rõ ràng về những kỳ vọng đối với các loại hình nghiên cứu sẽ được tiến hành, tần suất tham gia, cũng như các hình thức và mức thù lao hoặc ưu đãi sẽ được cung cấp.

Ngày nay, hầu hết các panel nghiên cứu đều hoạt động trực tuyến. Thuật ngữ "**Panel độc quyền**" (**Proprietary panels**) đề cập đến các panel nghiên cứu thuộc quyền sở hữu của một nhà cung cấp panel. Tổ chức này chịu trách nhiệm xây dựng và duy trì panel. Chủ sở hữu panel chịu trách nhiệm về "chất lượng" của panel, đảm bảo rằng panel đó được xác thực và "làm mới" (refreshed) thường xuyên.

Lợi ích chính của việc sử dụng panel nghiên cứu là sự dễ dàng và tốc độ trong việc tiếp cận các đối tượng trả lời khảo sát tiềm năng. Tổng chi phí tính trên mỗi đầu người tham gia khảo sát có thể thấp hơn so với các phương pháp khác. Hơn nữa, các nhà nghiên cứu có khả năng tuyển dụng người phản hồi tiềm năng từ khắp nơi trên toàn thế giới. Điều này có thể đặc biệt hữu ích khi các nhà tiếp thị đang lên kế hoạch mở rộng hoạt động sang các khu vực mới.

Cuối cùng, các panel trực tuyến mang lại sự riêng tư và bảo mật cao hơn cho người phản hồi. Do đó, sẽ dễ dàng hơn để thảo luận về các chủ đề nhạy cảm (ví dụ: một tình trạng bệnh lý) hoặc tuyển dụng các nhóm dân số ẩn (hidden populations) (chẳng hạn như các thành viên của cộng đồng LGBTQ+). [2]

Nguồn tham khảo / Chú thích: [1]

<https://www.shutterstock.com/image-vector/vector-emotion-feedback-scale-on-white-1090628780>

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=72#h5p-27>

Những thách thức trong việc sử dụng các nhóm nghiên cứu³

Có bốn vấn đề chính đã được các nhà nghiên cứu và những người thực hành (practitioners) xác định khi sử dụng các **nhóm nghiên cứu (research panels)**.

Vấn đề đầu tiên đề cập đến **“những người trả lời khảo sát chuyên nghiệp” (professional survey-takers)** – những người cố tình tham gia khảo sát một cách thường xuyên. Những người trả lời như vậy chắc chắn sẽ phản hồi các câu hỏi khảo sát khác biệt so với các nhóm **đối tượng phản hồi (respondents)** khác. Các đơn vị cung cấp hội đồng (panel providers) được khuyến khích [công khai] tỷ lệ tham gia trong hội đồng của mình. Tương tự, tỷ lệ bồi hoàn (reimbursement rate) (chẳng hạn như vé xem phim, 10 đô la hoặc khoản quyên góp cho tổ chức từ thiện yêu thích) cho các thành viên hội đồng cần phải được cung cấp minh bạch để khách hàng có thể đánh giá **động lực (incentive)** đằng sau sự tham gia của đối tượng phản hồi.

Nguồn tham khảo / Chú thích:

[2] Greenbook directory 2022, *75 Top online research panel companies: what is a proprietary panel?*, truy cập ngày 11 tháng 4 năm 2022.

[3] Porter, COLH, Outlaw, R, Gale, JP & Cho, TS 2018, ‘The use of online panel data in management research: a review and recommendations’, *Journal of Management*, tập 45, số 1, trang 319-344, truy cập ngày 11 tháng 4 năm 2022, cơ sở dữ liệu SAGE Journal, DOI: 10.1177/0149206318811569.

Vấn đề thứ hai liên quan đến **“tính đại diện” (representativeness)** của một hội đồng trực tuyến. Mặc dù đây từng là một vấn đề trong những ngày đầu của internet, việc xem xét nó vẫn rất hữu ích. Với sự tiếp cận rộng rãi dữ liệu trực tuyến, **khoảng cách kỹ thuật số (digital divide)** đang dần thu hẹp. Tại Úc, 99% người trưởng thành có quyền truy cập internet. Các nhà nghiên cứu hiện nay tự tin hơn trong việc tìm kiếm tính đại diện cao hơn của **tổng thể chung (overall population)** trong các hội đồng trực tuyến, vì việc truy cập internet không còn bị ràng buộc bởi độ tuổi, giáo dục, thu nhập hay nền tảng văn hóa/ngôn ngữ. Tuy nhiên, các nhà nghiên cứu vẫn có thể muốn kiểm tra xem liệu một hội đồng trực tuyến có đủ tính đại diện cho một **tổng thể cụ thể (specific population)** hay không. Ví dụ, nếu một nghiên cứu yêu cầu tuyển dụng “các bà nội trợ sống tại Sydney VÀ sinh ra tại Sudan” thì có thể cần phải thảo luận với đơn vị cung cấp hội đồng.

Điểm tiếp theo liên quan đến **“sức khỏe” (health)** của hội đồng trực tuyến. Tương tự như tất cả các cơ sở dữ liệu, việc chỉ đơn giản sở hữu một danh sách những người sẵn sàng phản hồi các khảo sát và câu hỏi phỏng vấn là chưa đủ. Cần phải có sự xem xét và giám sát liên tục đối với hội đồng. Nếu không thực hiện việc này, sẽ có rủi ro sở hữu một hội đồng với thông tin thành viên đã lỗi thời. **Hồ sơ thành viên (member profiles)** thay đổi liên tục. Mọi người chuyển nơi ở, đổi việc làm, kết hôn, sinh con, hình thành những quan điểm mới và thay đổi các niềm tin trước đó. Trừ khi có một cơ chế để làm mới và cập nhật thông tin về các thành viên, nếu không sẽ có khả năng cao xảy ra tình trạng **“không phản hồi” (non-response)**, vì hội đồng sẽ không còn phù hợp với **hồ sơ đối tượng phản hồi (respondent profile)** yêu cầu. Có một số người thực hành cảm thấy việc tuyển dụng thành viên mới thường xuyên có thể giải quyết vấn đề này. Mặc dù việc tăng quy mô hội đồng là một điểm bán hàng (selling point) tốt, nhưng việc cập nhật thông tin của các thành viên hiện tại thường mang lại hiệu quả chi phí cao hơn là tuyển dụng mới.

Cuối cùng, có một mối quan ngại về **“chất lượng dữ liệu” (quality of data)** được tạo ra từ các hội đồng trực tuyến, vì các thành viên hội đồng được tuyển chọn có thể không hoàn toàn tập trung vào nội dung khảo sát, không nỗ lực suy nghĩ về các câu hỏi hoặc có

sự thiếu tập trung nói chung. Mặc dù đây là một vấn đề quan trọng, việc nhận diện những người trả lời sai lệch (errant survey takers) này cũng rất dễ dàng. Đơn vị cung cấp hội đồng có thể thảo luận tiêu chí với khách hàng và xóa bỏ phản hồi từ các “**speeders**” (tức là những đối tượng phản hồi hoàn thành khảo sát trong một thời gian ngắn tới mức phi lý) và “**straight-liners**” (tức là những đối tượng phản hồi liên tục chọn cùng một phương án cho tất cả các câu hỏi khảo sát). Điều này đặc biệt phổ biến đối với khảo sát sử dụng các câu hỏi dạng **thang đo Likert (Likert-scale)**. Bất kỳ bên mua hội đồng trực tuyến nào cũng muốn được đảm bảo rằng **tính chính trực (integrity)** của hội đồng được duy trì bằng cách loại bỏ bất kỳ thành viên nào bị phát hiện có hành vi như vậy lặp đi lặp lại.

Quản lý các Panel Nghiên cứu Trực tuyến (Management of Online Research Panels)

Bốn giai đoạn của quy trình quản lý panel bao gồm [6]:

- Tuyển dụng thành viên panel (Recruitment of panellists)
- Lựa chọn mẫu (Selection of a sample)
- Giám sát và duy trì panel (Panel monitoring and maintenance)
- Gắn kết panel / Tương tác với panel (Panel engagement)

4. Cơ quan Truyền thông và Truyền thông Úc (Australian Communications and Media Authority) 2021, *Truyền thông và phương tiện tại Úc: cách chúng ta sử dụng internet*, xem ngày 11 tháng 5 năm 2022, [<https://www.acma.gov.au/publications/2021-12/report/communications-andmedia-australia-how-we-useinternet>]([https://www.acma.gov.au/publications/2021-12/report/communications-andmedia-australia-how-we-useinternet#:~:text=Nearly%20all%20Australian%20adults%20\(99,prior%20to%20C\)](https://www.acma.gov.au/publications/2021-12/report/communications-andmedia-australia-how-we-useinternet#:~:text=Nearly%20all%20Australian%20adults%20(99,prior%20to%20C)))

5. Data Decisions Group 2022, *Giữ cho hội đồng nghiên cứu trực tuyến tùy chỉnh của bạn luôn được làm mới (Keeping your custom online research panel refreshed)*, xem ngày 11 tháng 4 năm 2022.

Tuyển chọn thành viên nhóm nghiên cứu (Recruitment of panellists)

Phương pháp tuyển dụng chính được sử dụng để xây dựng một panel được gọi là **"phương pháp chọn tham gia" (opt-in method)** (Goritz 2007, trích dẫn trong Kahn 2012). Với các nhóm nghiên cứu (panel) "chọn tham gia" hoặc "tình nguyện", thường không có hạn chế về đối tượng có thể tham gia.

Lý tưởng nhất, một panel nên được xây dựng bằng cách sử dụng nhiều phương pháp khác nhau (cả trực tuyến và ngoại tuyến) để tuyển dụng người tham gia. Thông thường, những cá nhân quan tâm sẽ được điều hướng đến trang web của tổ chức quản lý panel, nơi họ có thể tìm thấy các điều khoản và điều kiện liên quan đến panel.

Người phản hồi tiềm năng được yêu cầu điền vào một biểu mẫu đăng ký, qua đó hệ thống sẽ tự động tạo ra một cơ sở dữ liệu về nhân khẩu học - xã hội (socio-demographic database). Bằng cách áp dụng nhiều phương pháp tuyển dụng khác nhau, có thể thông qua việc hợp tác với đa dạng các đối tác liên kết, panel được tạo ra chắc chắn sẽ phong phú và có tính đại diện cao hơn so với panel chỉ dựa trên một nguồn tuyển dụng hạn chế.

Lựa chọn mẫu

Một mẫu được lựa chọn theo các yêu cầu của dự án nghiên cứu. Việc thảo luận trước về hồ sơ người phản hồi với nhà cung cấp panel luôn mang lại lợi ích. Ví dụ: nếu nhà nghiên cứu đặc biệt muốn khảo sát *"những bà mẹ sinh ra ở Indonesia, hiện sống tại Sydney và có con đang trong độ tuổi đến trường"*, không phải tất cả các panel trực tuyến đều có sẵn nhóm dân số này.

6 Khan, AM 2012, *Sự ảnh hưởng của hoạt động tài trợ doanh nghiệp đối với nhân viên của bên tài trợ (The influence of corporate sponsorship on the sponsor's employees)*, luận án tiến sĩ (thesis), xem ngày 29 tháng 3 năm 2022.

Phương pháp **chọn mẫu theo định mức (Quota sampling)** thường được áp dụng khi lựa chọn một mẫu từ panel trực tuyến. Nhà nghiên cứu có thể quyết định rằng họ muốn tỷ lệ phân bổ giữa nam và nữ là bằng nhau. Ngay cả khi có sẵn nhiều thành viên panel là nữ hơn sẵn sàng trả lời khảo sát, họ cũng sẽ không được chọn thêm vào nếu giới hạn định mức của nhóm nữ đã đạt đủ.

Giám sát và duy trì panel

Một trong những dấu hiệu của một nhà cung cấp panel chất lượng là các chiến lược của họ liên quan đến việc giám sát và duy trì panel trực tuyến. Để đảm bảo chất lượng của panel, cần phải chắc chắn rằng các thành viên vẫn đang "hoạt động" (active).

Hơn nữa, bất kỳ thành viên nào bị phát hiện là **"người làm khảo sát quá nhanh" (speeders)** hoặc **"người đánh lụi theo đường thẳng" (straight liners)** sẽ bị cảnh báo và bị xóa khỏi panel nếu hành vi này vẫn tiếp diễn. Nhà cung cấp panel cũng phải có các chính sách để ngăn chặn việc "sử dụng quá mức" (over-use) các thành viên panel. Ví dụ: không có người phản hồi nào phải làm khảo sát nhiều hơn một lần một tháng. Tương tự, nếu một thành viên panel đang làm một cuộc khảo sát, họ không nên bị yêu cầu thực hiện đồng thời một cuộc khảo sát khác.

Theo thời gian, các thành viên panel có thể trở nên không hoạt động (inactive) do mất hứng thú, cảm thấy khối lượng công việc không được đánh giá xứng đáng, hoặc do sự thay đổi trong hoàn cảnh gia đình. Các lý do khác như tỷ lệ tử vong tự nhiên, địa chỉ email không hợp lệ và những lo ngại của thành viên về bảo mật dữ liệu cũng có thể góp phần làm giảm mức độ hoạt động. Việc nhà cung cấp dịch vụ giám sát panel để phát hiện bất kỳ dấu hiệu nào như vậy và khám phá các phương án nhằm giữ cho panel luôn "sống" và "khỏe mạnh" là rất hữu ích.

Gắn kết panel (Panel engagement)

Tương tự như cách các công ty duy trì sự gắn kết với nhân viên của họ, việc tương tác và gắn kết với các thành viên panel cũng rất quan trọng. Mặc dù những người phản hồi thường được cung cấp một khoản ưu đãi (bằng tài chính hoặc hình thức khác), người ta thừa nhận rằng sự hợp tác mạnh mẽ của người phản hồi – yếu tố thiết yếu cho chất lượng dữ liệu – là thứ không thể dùng tiền để mua được.

Các nhà cung cấp panel trực tuyến cần quản lý panel của họ bằng cách xây dựng lòng tin và sự cam kết thông qua các cấp độ giao tiếp đa kênh, chẳng hạn như qua liên hệ cá nhân, bản tin email (newsletters) và các cuộc gọi điện thoại. Đây cũng là cơ hội để "đào tạo" người phản hồi về các thực hành tốt bằng cách truyền đạt rõ ràng những kỳ vọng của nhà nghiên cứu và tầm quan trọng của việc tạo ra dữ liệu có chất lượng cao.

Các Nguyên tắc của ESOMAR dành cho Panel Nghiên cứu

ESOMAR (Hiệp hội Nghiên cứu Thị trường và Dư luận Châu Âu - *European Society for Opinion and Marketing Research*) là một tổ chức thành viên phi lợi nhuận, được thành lập vào năm 1948. ESOMAR có trụ sở chính tại Amsterdam, Hà Lan và hiện diện tại hơn 130 quốc gia.

Các thành viên của ESOMAR tạo thành một cộng đồng nhờ có chung mối quan tâm đến phân tích dữ liệu, nghiên cứu và sự thấu hiểu (insights) nhằm giúp cải thiện cuộc sống của các cá nhân trong xã hội cũng như trong các tổ chức.

Là tổ chức hàng đầu đại diện cho các nhà nghiên cứu trên toàn thế giới, ESOMAR đã xuất bản một bản hướng dẫn về cách tuyển dụng mẫu trực tuyến bằng cách sử dụng các panel. Bản hướng dẫn này có sẵn tại trang web **Esomar.org**.

Video: Ví dụ về Panel Người tiêu dùng (Consumer Panel)

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=72#oembed-1>

15. CÁC PHƯƠNG PHÁP PHÂN PHỐI KHẢO SÁT

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Phân biệt được các phương pháp khác nhau hiện có để thu thập dữ liệu khảo sát;
- Hiểu được ưu và nhược điểm của từng phương pháp phân phối khảo sát.

Nguồn tham khảo / Chú thích: [7] Camden BRI, 2011, *Consumer panel information* (Thông tin về panel người tiêu dùng), ngày 20 tháng 4, video trực tuyến, truy cập ngày 4 tháng 4 năm 2022.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=74#oembed-1>

Phân phối Khảo sát để Thu thập Dữ liệu (Survey Distribution for Data Collection)

Có nhiều cách khác nhau để phân phối khảo sát nhằm thu thập dữ liệu. Dưới đây là các phương pháp được sử dụng phổ biến nhất.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=74#h5p-28>

Nguồn tham khảo / Chú thích: [1] Elon University Poll, 2014, *Methods of collecting survey data* (Các phương pháp thu thập dữ liệu khảo sát), ngày 27 tháng 9, video trực tuyến, truy cập ngày 4 tháng 4 năm 2022

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=74#h5p-29>

Lựa chọn Phương pháp Thu thập Dữ liệu (Choosing a Data Collection Method)

Một số yếu tố có thể đóng vai trò quyết định trong việc lựa chọn (các) phương pháp thu thập dữ liệu phù hợp nhất. Những yếu tố này được thảo luận dưới đây:

Thời gian (Time)

Các tổ chức thường có những thời hạn (deadlines) chặt chẽ cần phải đáp ứng. Các nhà nghiên cứu chịu áp lực phải cung cấp thông tin kịp thời. Điều này thường có nghĩa là chỉ những phương pháp có thời gian xoay vòng nhanh (quick turnaround) như khảo sát trực tuyến hoặc phỏng vấn qua điện thoại mới có thể được sử dụng để thu thập dữ liệu. Các phương pháp truyền thống hơn, chẳng hạn như khảo sát tại nhà (in-home surveys), sẽ không phù hợp trong tình huống này.

Chi phí (Cost)

Nhà nghiên cứu phải luôn lưu ý đến nguồn kinh phí hiện có. Trong trường hợp có ngân sách lớn cho nghiên cứu, việc thuê nhân viên thực địa (field workers) để thu thập dữ liệu tại các trung tâm thương mại, văn phòng hoặc tại nhà là hoàn toàn khả thi. Ngược lại, nếu ngân sách hạn hẹp, nhà nghiên cứu có thể phải dựa vào các hình thức phân phối khảo sát tiết kiệm chi phí, chẳng hạn như đăng tải khảo sát trên nền tảng mạng xã hội của công ty.

Khả năng tiếp cận các Nhóm Quần thể (Access to Population Groups)

Phương pháp thu thập dữ liệu nhanh nhất và tiết kiệm chi phí nhất cũng sẽ trở nên vô dụng nếu không thể tiếp cận đúng nhóm quần thể mục tiêu. Đây là một trong những điểm then chốt cần được xem xét khi lựa chọn phương pháp phân phối bảng câu hỏi khảo sát. Ví dụ, một nhóm sinh viên đại học ở Iran có thể dễ dàng tiếp cận bằng các phương pháp trực tiếp, trong khi ở Úc, các chiến lược thay thế khác có thể sẽ cần được áp dụng.

Sự phù hợp với Mục tiêu Nghiên cứu (Suitability for Research Aim)

Nếu một cuộc khảo sát được thiết kế đòi hỏi người phản hồi phải xem các hình ảnh nhất định, thì phỏng vấn qua điện thoại có thể không phải là một phương pháp phù hợp. Tương tự, nếu một bảng câu hỏi yêu cầu người phản hồi truy cập vào một liên kết hoặc một trang web, điều quan trọng là phải áp dụng đúng phương pháp thu thập dữ liệu để tạo điều kiện thuận lợi nhất cho họ.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=74#h5p-30>

16. THỐNG KÊ MÔ TẢ (DESCRIPTIVE STATISTICS)

Chương này được phỏng theo Chương 12.1: "Mô tả các Biến đơn" ([Describing Single Variables](#)) trong tài liệu của Thư viện Đại học Minnesota năm 2010 [1] và được sử dụng theo giấy phép [Creative Commons Ghi công - Phi thương mại - Chia sẻ tương tự \(Attribution-NonCommercial-ShareAlike\)](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Giải thích được thế nào là thống kê mô tả;
- Hiểu được các loại thước đo được sử dụng để mô tả một mẫu về mặt thống kê;
- Tính toán và diễn giải các thước đo xu hướng tập trung (measures of central tendency).

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=76#oembed-1>

Nguồn: Amour Learning 2

Thống kê mô tả đề cập đến một tập hợp các kỹ thuật nhằm tóm tắt và trình bày dữ liệu (của một mẫu). Ở đây, chúng ta giả định rằng dữ liệu có tính định lượng và bao gồm các điểm số trên một hoặc nhiều biến của từng đối tượng tham gia nghiên cứu. Chúng ta sẽ bắt đầu bằng cách xem xét một số kỹ thuật phổ biến nhất để mô tả các biến đơn (single variables).

Phân phối của một biến (The Distribution of a Variable)

Mỗi biến đều có một phân phối, đó là cách các điểm số (giá trị) được phân bổ trên các mức độ của biến đó. Ví dụ, trong một mẫu gồm 100 sinh viên đại học, phân phối của biến "số lượng anh chị em" có thể là: 10 người không có anh chị em nào, 30 người có một, 40 người có hai, v.v. Trong cùng mẫu đó, phân phối của biến "giới tính" có thể là 44 người mang giá trị "nam" và 56 người mang giá trị "nữ".

Nguồn tham khảo / Chú thích: [2] Amour Learning, 2019, *What are descriptive and inferential statistics?* (Thống kê mô tả và thống kê suy luận là gì?), ngày 28 tháng 1, video trực tuyến, truy cập ngày 1 tháng 5 năm 2022.

Nguồn tham khảo / Chú thích: [1] Thư viện Đại học Minnesota 2010, 'Describing single variables' trong *Research Methods in Psychology*, Nhà xuất bản Thư viện Đại học Minnesota.

Bảng tần số (Frequency Tables)

Một cách để trình bày phân phối của một biến là sử dụng bảng tần số. Ví dụ, Bảng 12.1 *"Bảng tần số thể hiện phân phối giả định của các điểm số trên Thang đo Lòng tự trọng Rosenberg"*, là một bảng tần số cho thấy phân phối giả định về điểm số trên Thang đo Lòng tự trọng Rosenberg đối với một mẫu gồm 40 sinh viên đại học.

Cột đầu tiên liệt kê các giá trị của biến — các điểm số có thể đạt được trên thang đo Rosenberg — và cột thứ hai liệt kê tần số (frequency) của từng điểm số. Bảng này cho thấy có ba sinh viên đạt điểm tự trọng là 24, năm sinh viên đạt điểm 23, v.v.

Từ một bảng tần số như thế này, người ta có thể nhanh chóng nhận ra một số khía cạnh quan trọng của phân phối, bao gồm: khoảng điểm (range) dao động từ 15 đến 24; các điểm số phổ biến nhất và ít phổ biến nhất (tương ứng là 22 và 17); cũng như bất kỳ điểm số cực trị (extreme scores) nào khác biệt hẳn so với phần còn lại.

BẢNG 12.1 BẢNG TẦN SỐ THỂ HIỆN PHÂN PHỐI GIẢ ĐỊNH CỦA CÁC ĐIỂM SỐ TRÊN THANG ĐO LÒNG TỰ TRỌNG ROSENBERG

Lòng tự trọng (Self-esteem)	Tần số (Frequency)
24	3
23	5
22	10
21	8
20	5
19	3
18	3
17	0

16	2
15	1

Có một vài điểm khác đáng chú ý về các bảng tần số. Đầu tiên, các mức độ (levels) được liệt kê ở cột thứ nhất thường đi từ giá trị cao nhất ở trên cùng xuống thấp nhất ở dưới cùng, và chúng thường không mở rộng vượt quá điểm số cao nhất và thấp nhất có trong dữ liệu. Ví dụ, mặc dù điểm số trên thang đo Rosenberg có thể dao động từ mức cao nhất là 30 đến mức thấp nhất là 0, Bảng 12.1 "*Bảng tần số thể hiện phân phối giả định của các điểm số trên Thang đo Lòng tự trọng Rosenberg*" chỉ bao gồm các mức độ từ 24 đến 15, vì khoảng này đã bao trọn tất cả các điểm số trong tập dữ liệu cụ thể này.

Thứ hai, khi có nhiều điểm số khác nhau trải dài trên một khoảng giá trị rộng, thường sẽ tốt hơn nếu tạo một **bảng tần số ghép lớp (grouped frequency table)**, trong đó cột đầu tiên liệt kê các khoảng giá trị và cột thứ hai liệt kê tần số của các điểm số rơi vào mỗi khoảng đó.

Ví dụ, Bảng 12.2 "*Bảng tần số ghép lớp thể hiện phân phối giả định của Thời gian phản ứng*" (được nhắc tới ở phần sau), là một bảng tần số ghép lớp thể hiện phân phối giả định về thời gian phản ứng đơn giản của một mẫu gồm 20 người tham gia. Trong một bảng tần số ghép lớp, các khoảng (ranges) phải có độ rộng bằng nhau và thường có từ 5 đến 15 khoảng.

Cuối cùng, các bảng tần số cũng có thể được sử dụng cho các biến phân loại (categorical variables), trong trường hợp này các mức độ chính là các nhãn phân loại (category labels). Thứ tự của các nhãn phân loại có phần tùy ý, nhưng chúng thường

được sắp xếp theo thứ tự từ tần số cao nhất ở trên cùng đến tần số thấp nhất ở dưới cùng.

BẢNG 12.2 BẢNG TẦN SỐ GHÉP LỚP THỂ HIỆN PHÂN PHỐI GIẢ ĐỊNH CỦA THỜI GIAN PHẢN ỨNG

Thời gian phản ứng (ms)	Tần số (Frequency)
241–260	1
221–240	2
201–220	2
181–200	9
161–180	4
141–160	2

BIỂU ĐỒ HISTOGRAM (HISTOGRAMS)

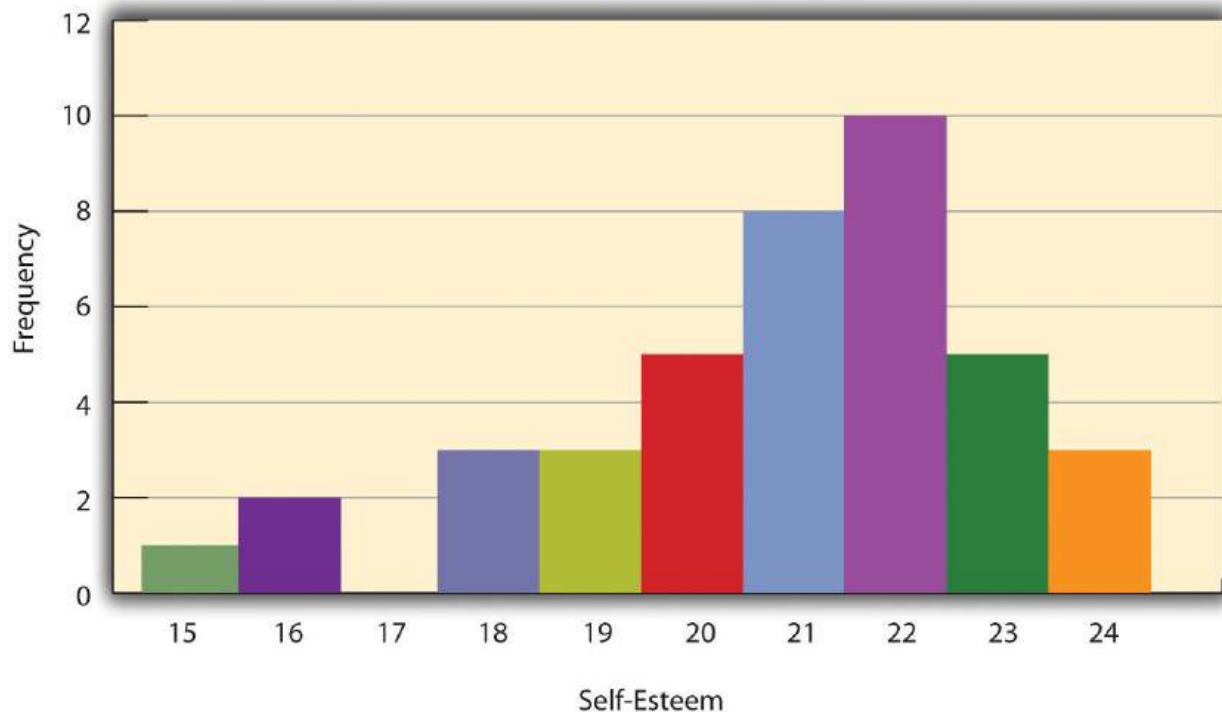
Biểu đồ histogram là một hình thức trình bày dạng đồ thị của một phân phối. Nó trình bày những thông tin giống hệt như một bảng tần số nhưng theo một cách giúp người xem nắm bắt nhanh chóng và dễ dàng hơn nhiều.

Biểu đồ histogram trong Hình 12.1 *"Biểu đồ Histogram thể hiện Phân phối của các Điểm số Lòng tự trọng được trình bày trong Bảng 12.1"* thể hiện phân phối của các điểm số lòng tự trọng trong Bảng 12.1 *"Bảng tần số thể hiện phân phối giả định của các điểm số trên Thang đo Lòng tự trọng Rosenberg"*.

Trục hoành (trục x) của biểu đồ histogram đại diện cho biến số và trục tung (trục y) đại diện cho tần số. Phía trên mỗi mức độ của biến số trên trục hoành là một thanh dọc (cột) đại diện cho số lượng cá nhân đạt được điểm số đó.

Khi biến số là định lượng (quantitative), như trong ví dụ này, thường sẽ không có khoảng trống giữa các thanh cột. Tuy nhiên, khi biến số là phân loại (categorical), thường sẽ có một khoảng trống nhỏ giữa chúng. *(Khoảng trống tại mốc 17 trong biểu đồ histogram này phản ánh thực tế là không có điểm số 17 nào xuất hiện trong tập dữ liệu này).*

HÌNH 12.1 BIỂU ĐỒ HISTOGRAM THỂ HIỆN PHÂN PHỐI CỦA CÁC ĐIỂM SỐ LÒNG TỰ TRỌNG ĐƯỢC TRÌNH BÀY TRONG BẢNG 12.1 "BẢNG TẦN SỐ THỂ HIỆN PHÂN PHỐI GIẢ ĐỊNH CỦA CÁC ĐIỂM SỐ TRÊN THANG ĐO LÒNG TỰ TRỌNG ROSENBERG"



HÌNH DÁNG CỦA PHÂN PHỐI (DISTRIBUTION SHAPES)

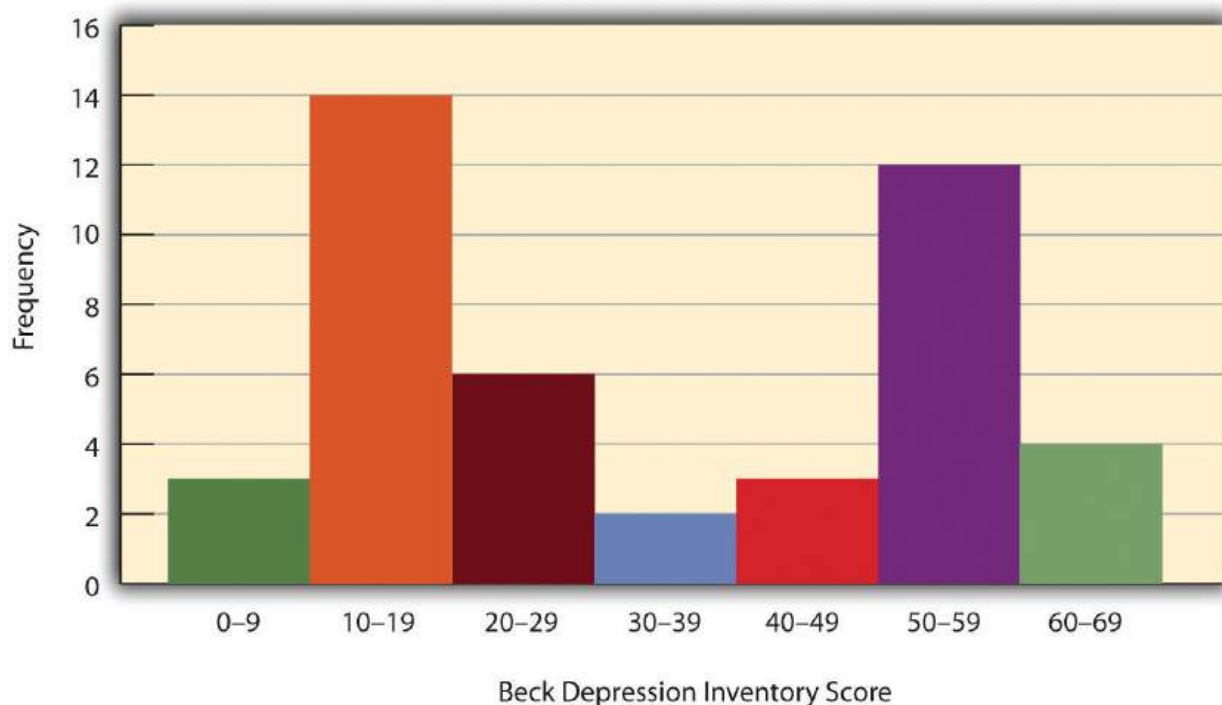
Khi phân phối của một biến định lượng được trình bày trên biểu đồ histogram, nó sẽ có một hình dáng nhất định. Hình dáng của phân phối điểm số lòng tự trọng trong Hình 12.1 là một ví dụ điển hình. Có một đỉnh nằm đâu đó gần giữa phân phối và các "phần đuôi" (tails) vuốt nhọn dần về cả hai hướng tính từ đỉnh.

Phân phối trong Hình 12.1 là **phân phối đơn đỉnh (unimodal)**, nghĩa là nó có một đỉnh rõ rệt. Tuy nhiên, các phân phối cũng có thể là **phân phối lưỡng đỉnh (bimodal)**, nghĩa là chúng có hai đỉnh rõ rệt.

Ví dụ, Hình 12.2 "*Biểu đồ Histogram thể hiện Phân phối lưỡng đỉnh giả định của các điểm số trên Bảng kiểm tra Trầm cảm Beck*", cho thấy một phân phối lưỡng đỉnh giả định của các điểm số trên Bảng kiểm tra Trầm cảm Beck (Beck Depression Inventory). Các phân phối cũng có thể có nhiều hơn hai đỉnh rõ rệt (**đa đỉnh** -

multimodal), nhưng những trường hợp này tương đối hiếm trong nghiên cứu tâm lý học.

HÌNH 12.2 BIỂU ĐỒ HISTOGRAM THỂ HIỆN PHÂN PHỐI LŨNG ĐỈNH GIẢ ĐỊNH CỦA CÁC ĐIỂM SỐ TRÊN BẢNG KIỂM TRA TRẦM CẢM BECK (BECK DEPRESSION INVENTORY)



Một đặc điểm khác về hình dáng của một phân phối là liệu nó có tính **đối xứng (symmetrical)** hay **bị lệch (skewed)**.

Phân phối nằm ở giữa trong Hình 12.3 "*Các biểu đồ Histogram thể hiện Phân phối Lệch âm, Đối xứng và Lệch dương*" là một phân phối đối xứng. Các nửa bên trái và bên phải của nó là hình ảnh phản chiếu của nhau qua gương.

- **Phân phối lệch âm (Negatively skewed):** Là phân phối ở bên trái, với đỉnh của nó bị lệch về phía đầu cao của khoảng giá trị và có một "đuôi âm" (đuôi bên trái) tương đối dài.
- **Phân phối lệch dương (Positively skewed):** Là phân phối ở bên phải, với đỉnh của nó nằm về phía đầu thấp của khoảng giá trị và có một "đuôi dương" (đuôi bên phải) tương đối dài.

HÌNH 12.3 CÁC BIỂU ĐỒ HISTOGRAM THỂ HIỆN PHÂN PHỐI LỆCH ÂM, ĐỐI XỨNG VÀ LỆCH DƯƠNG



Giá trị ngoại lệ (Outlier) là một điểm số cực trị, cao hơn hoặc thấp hơn rất nhiều so với phần còn lại của các điểm số trong phân phối. Đôi khi, các giá trị ngoại lệ đại diện cho những mức độ cực hạn thực sự của biến số đang được nghiên cứu. Ví dụ, trên Bảng kiểm tra Trầm cảm Beck, một người bị trầm cảm lâm sàng duy nhất có thể là một giá trị ngoại lệ trong một mẫu gồm những người bạn cùng lứa tuổi đang hạnh phúc và sinh hoạt bình thường.

Tuy nhiên, các giá trị ngoại lệ cũng có thể là kết quả của các sai sót hoặc hiểu lầm từ phía nhà nghiên cứu hoặc người tham gia, lỗi thiết bị, hoặc các vấn đề tương tự. Chúng ta sẽ thảo luận kỹ hơn về cách diễn giải các giá trị ngoại lệ và cách xử lý chúng ở phần sau của chương này.

CÁC THƯỚC ĐO XU HƯỚNG TẬP TRUNG VÀ ĐỘ BIẾN THIÊN

Việc có thể mô tả các đặc điểm của phân phối một cách chính xác hơn là điều rất hữu ích. Tại đây, chúng ta xem xét cách thực hiện điều này thông qua hai đặc điểm quan trọng: xu hướng tập trung và độ biến thiên của chúng.

XU HƯỚNG TẬP TRUNG

Xu hướng tập trung của một phân phối là điểm trung tâm của nó—điểm mà các điểm số trong phân phối có xu hướng tập trung xung quanh. (Một thuật ngữ khác của xu hướng tập trung là giá trị trung bình). Nhìn lại Hình 12.1 “Biểu đồ Histogram thể hiện phân phối điểm số lòng tự trọng được trình bày trong...”, ví dụ, chúng ta có thể thấy rằng các điểm số lòng tự trọng có xu hướng tập trung quanh các giá trị từ 20 đến 22. Tại đây, chúng ta sẽ xem xét ba thước đo xu hướng tập trung phổ biến nhất: số trung bình (mean), số trung vị (median), và yếu vị (mode).

Số trung bình của một phân phối (ký hiệu là M) là tổng các điểm số chia cho số lượng các điểm số đó. Dưới dạng công thức, nó trông như sau:

$$M = \frac{\sum X}{N}$$

Trong công thức này, biểu tượng Σ (chữ cái Hy Lạp sigma) là dấu tổng và có nghĩa là tính tổng các giá trị của biến X . N đại diện cho số lượng điểm số. Số trung bình cho đến nay là thước đo xu hướng tập trung phổ biến nhất, và có một vài lý do chính đáng cho việc này. Nó thường cung cấp một chỉ báo tốt về xu hướng tập trung của một phân phối, và nó dễ hiểu đối với hầu hết mọi người. Ngoài ra, số trung bình có các thuộc tính thống kê khiến nó đặc biệt hữu ích trong việc thực hiện thống kê suy luận.

Một lựa chọn thay thế cho số trung bình là số trung vị. Số trung vị là điểm số nằm ở giữa theo nghĩa là một nửa số điểm trong phân phối nhỏ hơn nó và một nửa số điểm lớn hơn nó. Cách đơn giản nhất để tìm số trung vị là sắp xếp các điểm số từ thấp nhất

đến cao nhất và xác định điểm số ở chính giữa. Ví dụ, hãy xem xét tập hợp bảy điểm số sau:

8,4,12,14,3,2,3

Để tìm số trung vị, chỉ cần sắp xếp lại các điểm số từ thấp nhất đến cao nhất và xác định điểm ở giữa.

2,3,3,4,8,12,14

Trong trường hợp này, số trung vị là 4 vì có ba điểm số thấp hơn 4 và ba điểm số cao hơn 4. Khi có một số lượng điểm số chẵn, sẽ có hai điểm số nằm ở giữa phân phối, trong trường hợp đó số trung vị là giá trị nằm chính giữa hai điểm đó. Ví dụ, nếu chúng ta thêm điểm số 15 vào tập dữ liệu trước đó, sẽ có hai điểm số (cả 4 và 8) ở giữa phân phối, và số trung vị sẽ là giá trị ở giữa chúng (6).

Thước đo cuối cùng của xu hướng tập trung là yếu vị. Yếu vị là điểm số có tần suất xuất hiện nhiều nhất trong một phân phối. Trong phân phối lòng tự trọng được trình bày trong Bảng 12.1 và Hình 12.1, ví dụ, yếu vị là 22. Số lượng sinh viên có điểm số đó nhiều hơn bất kỳ điểm số nào khác. Yếu vị là thước đo xu hướng tập trung duy nhất cũng có thể được sử dụng cho các biến phân loại.

Trong một phân phối vừa đơn đỉnh vừa đối xứng, số trung bình, trung vị và yếu vị sẽ rất gần nhau tại đỉnh của phân phối. Trong một phân phối lưỡng đỉnh hoặc không đối xứng, số trung bình, trung vị và yếu vị có thể khá khác nhau. Trong một phân phối lưỡng đỉnh, số trung bình và trung vị sẽ có xu hướng nằm giữa các đỉnh, trong khi yếu vị sẽ nằm ở đỉnh cao nhất. Trong một phân phối lệch, số trung bình sẽ khác với số trung vị theo hướng của độ lệch (tức là hướng của phần đuôi dài hơn).

Đối với các phân phối bị lệch mạnh, số trung bình có thể bị kéo đi quá xa theo hướng lệch đến mức nó không còn là một thước đo tốt cho xu hướng tập trung của phân phối đó nữa. Hãy tưởng tượng, ví dụ, một tập hợp gồm bốn thời gian phản ứng đơn giản là

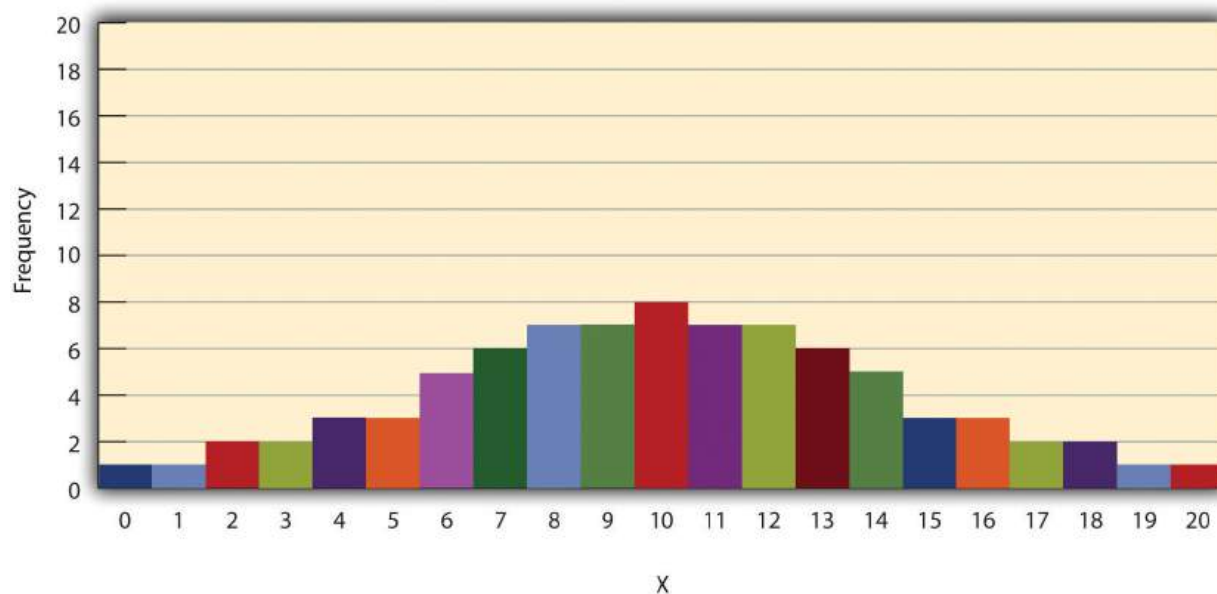
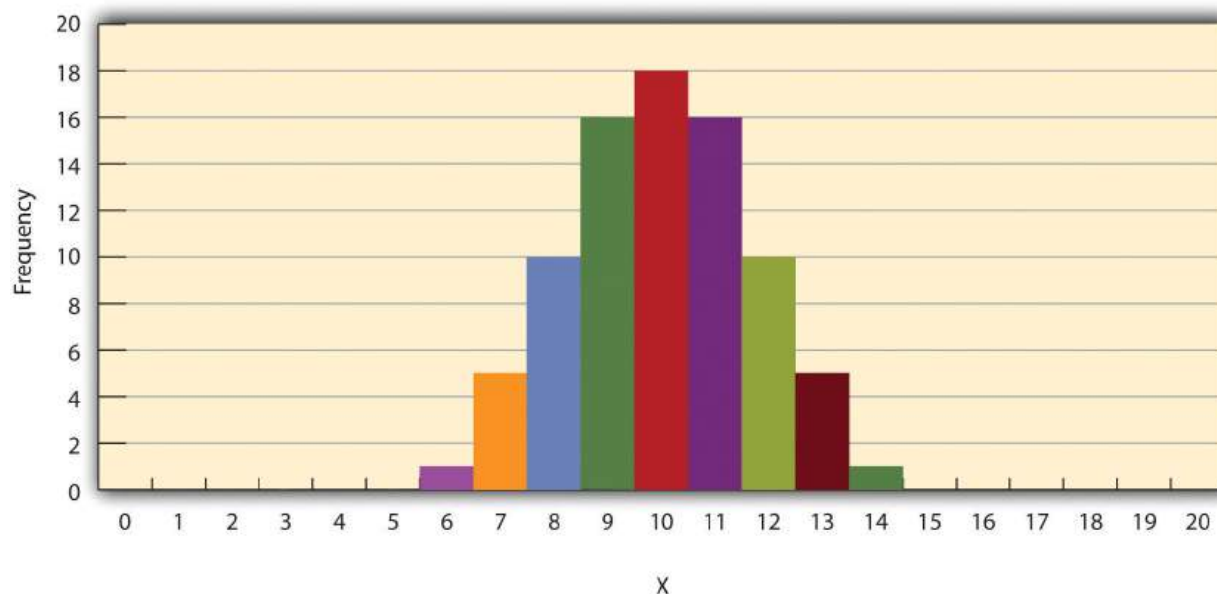
200, 250, 280 và 250 mili giây (ms). Số trung bình là 245 ms. Nhưng việc thêm một điểm số nữa là 5.000 ms—có lẽ vì người tham gia không chú ý—sẽ nâng số trung bình lên thành 1.445 ms. Thước đo xu hướng tập trung này không chỉ lớn hơn 80% các điểm số trong phân phối, mà nó còn dường như không đại diện tốt cho hành vi của bất kỳ ai trong phân phối đó. Đây là lý do tại sao các nhà nghiên cứu thường ưu tiên số trung vị cho các phân phối bị lệch mạnh (chẳng hạn như phân phối thời gian phản ứng).

Tuy nhiên, hãy nhớ rằng bạn không bắt buộc phải chọn một thước đo xu hướng tập trung duy nhất khi phân tích dữ liệu của mình. Mỗi thước đo cung cấp thông tin hơi khác nhau và tất cả chúng đều có thể hữu ích.

CÁC THƯỚC ĐO ĐỘ BIẾN THIÊN

Độ biến thiên của một phân phối là mức độ mà các điểm số thay đổi xung quanh xu hướng tập trung của chúng. Hãy xem xét hai phân phối trong Hình 12.4 “Biểu đồ Histogram thể hiện các phân phối giả định có cùng số trung bình, trung vị và yếu vị (10) nhưng có độ biến thiên thấp (Trên) và độ biến thiên cao (Dưới)”, cả hai đều có cùng xu hướng tập trung. Số trung bình, trung vị và yếu vị của mỗi phân phối đều là 10. Tuy nhiên, hãy lưu ý rằng hai phân phối này khác nhau về độ biến thiên của chúng. Biểu đồ phía trên có độ biến thiên tương đối thấp, với tất cả các điểm số tương đối gần với điểm trung tâm. Biểu đồ phía dưới có độ biến thiên tương đối cao, với các điểm số trải rộng trên một khoảng lớn hơn nhiều.

HÌNH 12.4 CÁC BIỂU ĐỒ HISTOGRAM THỂ HIỆN CÁC PHÂN PHỐI GIẢ ĐỊNH CÓ CÙNG SỐ TRUNG BÌNH, TRUNG VỊ VÀ YẾU VỊ (10) NHƯNG CÓ ĐỘ BIẾN THIÊN THẤP (TRÊN) VÀ ĐỘ BIẾN THIÊN CAO (DƯỚI)



Một thước đo đơn giản về độ biến thiên là **khoảng biến thiên (range)**, đơn giản là hiệu số giữa điểm số cao nhất và thấp nhất trong phân phối. Ví dụ, khoảng biến thiên của các điểm số lòng tự trọng trong Bảng 12.1 “Bảng tần số thể hiện phân phối giả định của các điểm số trên Thang đo Lòng tự trọng Rosenberg” là hiệu số giữa điểm số cao nhất (24) và điểm số thấp nhất (15). Nghĩa là, khoảng biến thiên là $24 - 15 = 9$. Mặc dù khoảng biến thiên dễ tính toán và dễ hiểu, nhưng nó có thể gây hiểu lầm khi

có các giá trị ngoại lệ. Hãy tưởng tượng, ví dụ, một kỳ thi mà tất cả sinh viên đều đạt điểm từ 90 đến 100. Nó có khoảng biến thiên là 10. Nhưng nếu có một sinh viên duy nhất đạt điểm 20, khoảng biến thiên sẽ tăng lên thành 80 — tạo ra ấn tượng rằng các điểm số biến động rất mạnh trong khi thực tế chỉ có một sinh viên khác biệt đáng kể so với những người còn lại.

Cho đến nay, thước đo độ biến thiên phổ biến nhất là **độ lệch chuẩn (standard deviation)**. Nói một cách đại khái, độ lệch chuẩn của một phân phối là khoảng cách trung bình giữa các điểm số và số trung bình. Ví dụ, độ lệch chuẩn của các phân phối trong Hình 12.4 “Biểu đồ Histogram thể hiện các phân phối giả định có cùng số trung bình, trung vị và yếu vị (10) nhưng có độ biến thiên thấp (Trên) và độ biến thiên cao (Dưới)” là 1,69 cho phân phối phía trên và 4,30 cho phân phối phía dưới. Nghĩa là, trong khi các điểm số ở phân phối phía trên khác với số trung bình khoảng 1,69 đơn vị tính theo trung bình, thì các điểm số ở phân phối phía dưới khác với số trung bình khoảng 4,30 đơn vị tính theo trung bình.

Các phép tính độ lệch chuẩn được minh họa cho một tập dữ liệu nhỏ trong Bảng 12.3 “Các phép tính cho độ lệch chuẩn”. Cột đầu tiên là một tập hợp gồm tám điểm số có số trung bình là 5. Cột thứ hai là hiệu số giữa mỗi điểm số và số trung bình. Cột thứ ba là bình phương của mỗi hiệu số này. Lưu ý rằng mặc dù các hiệu số có thể âm, nhưng các bình phương của hiệu số luôn dương — nghĩa là độ lệch chuẩn luôn luôn dương. Ở dưới cùng của cột thứ ba là trung bình của các bình phương hiệu số, giá trị này còn được gọi là **phương sai (ký hiệu là SD^2)**. Mặc dù bản thân phương sai là một thước đo độ biến thiên, nhưng nhìn chung nó đóng vai trò lớn hơn trong thống kê suy luận so với thống kê mô tả. Cuối cùng, dưới phương sai là căn bậc hai của phương sai, đó chính là độ lệch chuẩn.

BẢNG 12.3 CÁC PHÉP TÍNH CHO ĐỘ LỆCH CHUẨN

X	X-M	(X-M)²
3	-2	4
5	0	0
4	-1	1
2	-3	9
7	2	4
6	1	1
5	0	0
8	3	9
M = 5		Tổng = 28

Phương sai: $SD^2 = 28/8 = 3.50$

Độ lệch chuẩn: $SD = \sqrt{3.50} = 1.87$

17. MỐI LIÊN HỆ GIỮA CÁC BIẾN

Nội dung sau đây được phỏng theo [SPSS eTutor của Dee Britton \[1\]](#) và được sử dụng theo giấy phép [Creative Commons Ghi công - Phi thương mại - Chia sẻ tương tự 3.0](#).

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu khái niệm về mối liên hệ (hoặc mối quan hệ) giữa các biến;
- Phân biệt được các phép kiểm định mối liên hệ giữa các biến phân loại (categorical variables) và các biến liên tục (continuous variables);
- Diễn giải các kết quả từ các phép kiểm định thống kê.

Nguồn tham khảo / Chú thích: [1] Britton, D 2020, *SPSS eTutor*, Suny Empire State College, truy cập ngày 2 tháng 5 năm 2022.

Trong khi phân tích dữ liệu, các nhà nghiên cứu quan tâm đến việc tìm hiểu mối liên hệ (hoặc mối quan hệ) giữa các loại biến khác nhau. Có hai loại mối liên hệ được đặc biệt quan tâm:

- **Mối liên hệ giữa các biến phân loại (categorical) / định danh (nominal) / thứ bậc (ordinal).**
 - **Mối liên hệ giữa các biến khoảng (interval) / liên tục (continuous).**
-

Mối liên hệ giữa các biến phân loại

Đôi khi nhà nghiên cứu muốn hiểu mối quan hệ giữa các biến phân loại (categorical variables), vốn còn được gọi là các biến định danh (nominal) hoặc biến thứ bậc (ordinal). Ví dụ về các biến này bao gồm giới tính, các nhóm độ tuổi, nghề nghiệp và sở thích thương hiệu. Các mối liên hệ như vậy có thể được khám phá thông qua:

a) Bảng chéo (cross-tabulation) và/hoặc, b) Phép kiểm định Chi-bình phương (chi-square test)

a) Bảng chéo (Cross-tabulation)

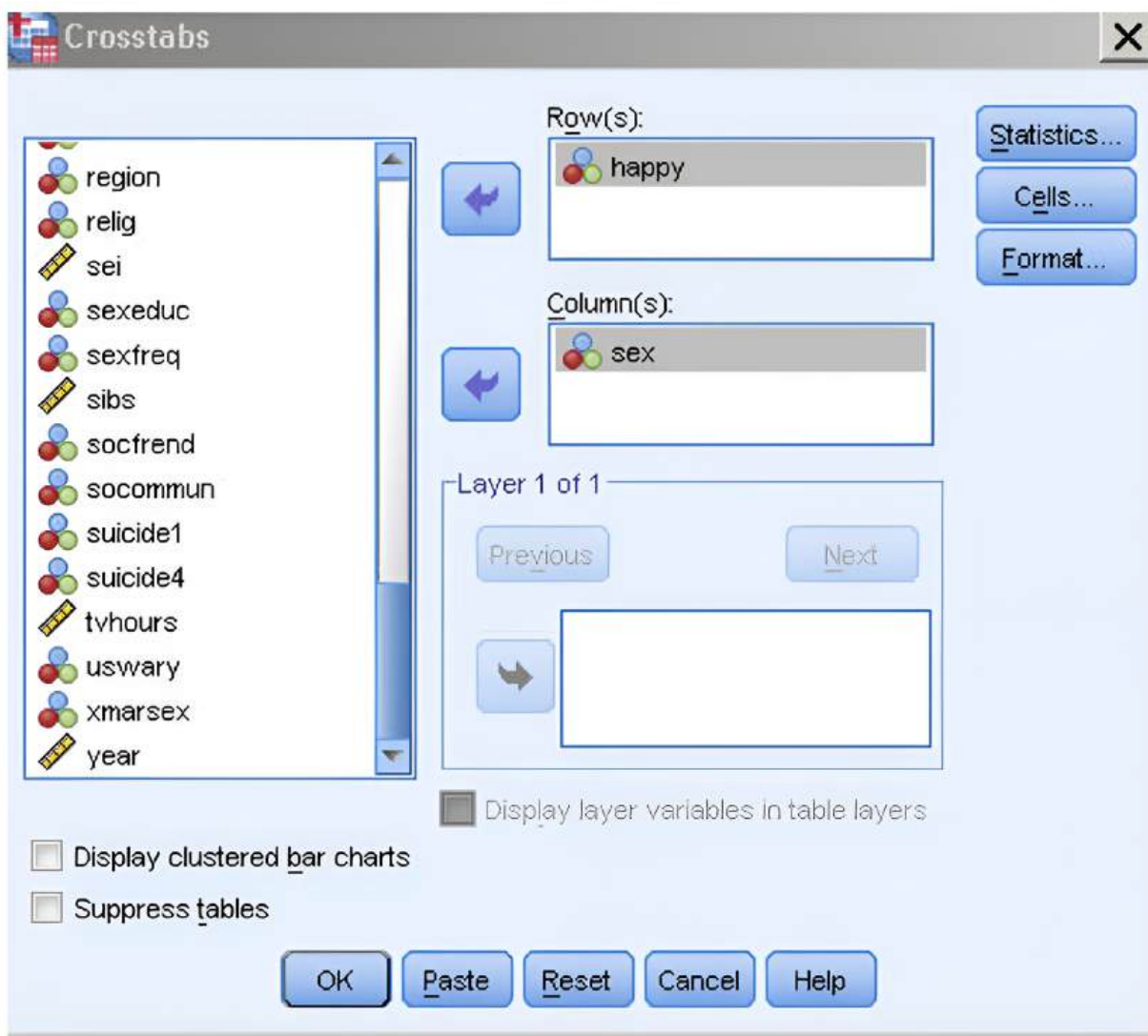
Bảng chéo đơn giản là một phân phối tần số cho hai biến cùng một lúc. Một bảng chéo là một ma trận thể hiện phân phối của một biến ứng với mỗi danh mục của biến thứ hai. [2]

Để chạy bảng chéo trên SPSS, hãy truy cập: **Analyze** (Phân tích), **Descriptive Statistics** (Thống kê mô tả), **Crosstabs** (Bảng chéo).

- Chọn biến phụ thuộc (**dependent variable**), nhấn vào mũi tên hướng về ô **Row** (Hàng).

- Chọn biến độc lập (**independent variable**), nhấn vào mũi tên hướng về ô **Column** (Cột).

Màn hình của bạn sẽ trông như thế này:



Nguồn tham khảo / Chú thích: [2] Britton, D 2020, SPSS eTutor: cross tabulations, Suny Empire State College, truy cập ngày 2 tháng 5 năm 2022.

Nhấn vào **Cells** (Ô) và sau đó nhấn vào ô **Column** percentage (Phần trăm theo cột):

Crosstabs: Cell Display

Counts

- ☒ Observed
- ☐ Expected
- ☐ Hide small counts
Less than

z-test

- ☐ Compare column proportions
- ☐ Aadjust p-values (Bonferroni method)

Percentages

- ☐ Row
- ☒ Column
- ☐ Total

Residuals

- ☐ Unstandardized
- ☐ Standardized
- ☐ Aadjusted standardized

Noninteger Weights

- ☐ Round cell counts
- ☐ Round case weights
- ☐ Truncate cell counts
- ☒ Truncate case weights
- ☐ No aadjustments

Nhấn **Continue** (Tiếp tục), sau đó nhấn **OK**. Kết quả đầu ra (output) của bạn sẽ trông như thế này:

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
GENERAL HAPPINESS * RESPONDENTS SEX	1495	99.7%	5	.3%	1500	100.0%

GENERAL HAPPINESS * RESPONDENTS SEX Crosstabulation

			RESPONDENTS SEX		Total
			MALE	FEMALE	
GENERAL HAPPINESS	VERY HAPPY	Count	198	237	435
		% within RESPONDENTS SEX	28.7%	29.4%	29.1%
	PRETTY HAPPY	Count	384	439	823
		% within RESPONDENTS SEX	55.7%	54.5%	55.1%
	NOT TOO HAPPY	Count	107	130	237
		% within RESPONDENTS SEX	15.5%	16.1%	15.9%
Total	Count	689	806	1495	
	% within RESPONDENTS SEX	100.0%	100.0%	100.0%	

Bảng chéo này có ý nghĩa gì? Kết quả này tương đối dễ diễn giải. 28,7% nam giới và 29,4% nữ giới khẳng định rằng họ đang rất hạnh phúc. Nhưng đối với một bảng chéo có nhiều thuộc tính thì sao? Ví dụ, nếu bạn muốn phân tích mối quan hệ giữa số lượng con cái bạn có và mức độ hạnh phúc chung của bạn? Dưới đây là kết quả đầu ra của phân tích đó:

Việc diễn giải bảng này trở nên khó khăn hơn một chút vì có nhiều ô dữ liệu hơn. Tuy nhiên, chúng ta vẫn tập trung vào **phần trăm theo cột (column percentages)** để so sánh các nhóm. Chẳng hạn, bạn có thể xem xét liệu những người có từ 4 con trở lên có tỷ lệ "rất hạnh phúc" cao hơn hay thấp hơn so với những người không có con.

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
GENERAL HAPPINESS * NUMBER OF CHILDREN	1493	99.5%	7	.5%	1500	100.0%

GENERAL HAPPINESS * NUMBER OF CHILDREN Crosstabulation									
			NUMBER OF CHILDREN						
			0	1	2	3	4	5	6
GENERAL HAPPINESS	VERY HAPPY	Count	97	69	108	84	44	18	
		% of Total	6.5%	4.6%	7.2%	5.6%	2.9%	1.2%	.4
	PRETTY HAPPY	Count	230	140	232	116	55	22	
		% of Total	15.4%	9.4%	15.5%	7.8%	3.7%	1.5%	.8
	NOT TOO HAPPY	Count	65	38	49	44	17	12	
		% of Total	4.4%	2.5%	3.3%	2.9%	1.1%	.8%	.2
	Total	Count	392	247	389	244	116	52	
		% of Total	26.3%	16.5%	26.1%	16.3%	7.8%	3.5%	1.4

b) Phép kiểm định Chi-bình phương (Chi-square test)

Trong khi bảng chéo đơn thuần là một bảng tần số/phần trăm cho hai biến, phép kiểm định Chi-bình phương được thực hiện để xem xét liệu kết quả của bảng chéo đó có ý nghĩa thống kê (statistically significant) hay không. Phép kiểm định Chi-bình phương là một phép kiểm định phi tham số được sử dụng để xác định xem liệu có mối liên hệ có ý nghĩa thống kê giữa hai biến phân loại hay không. [3]

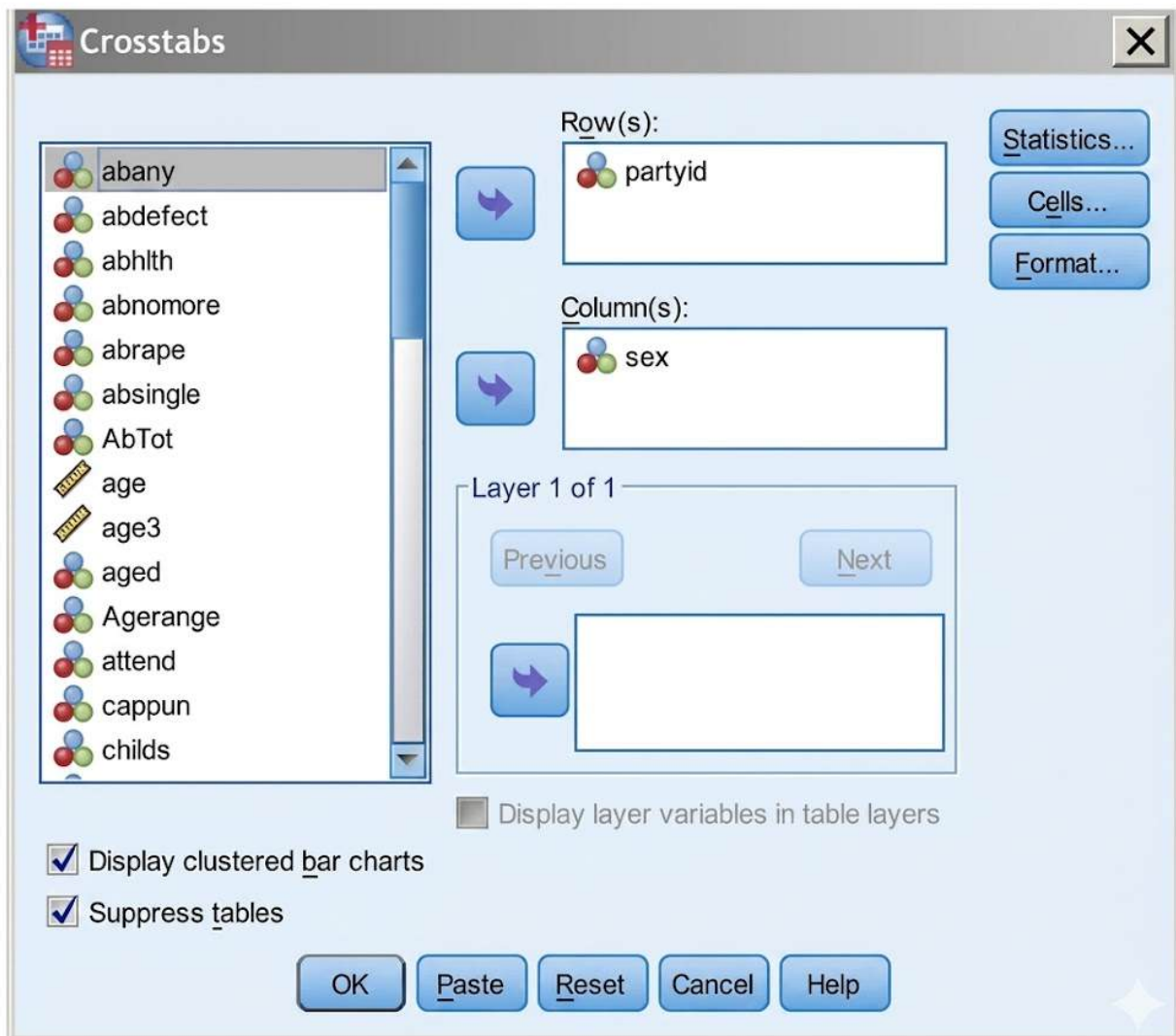
Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=78#oembed-1>

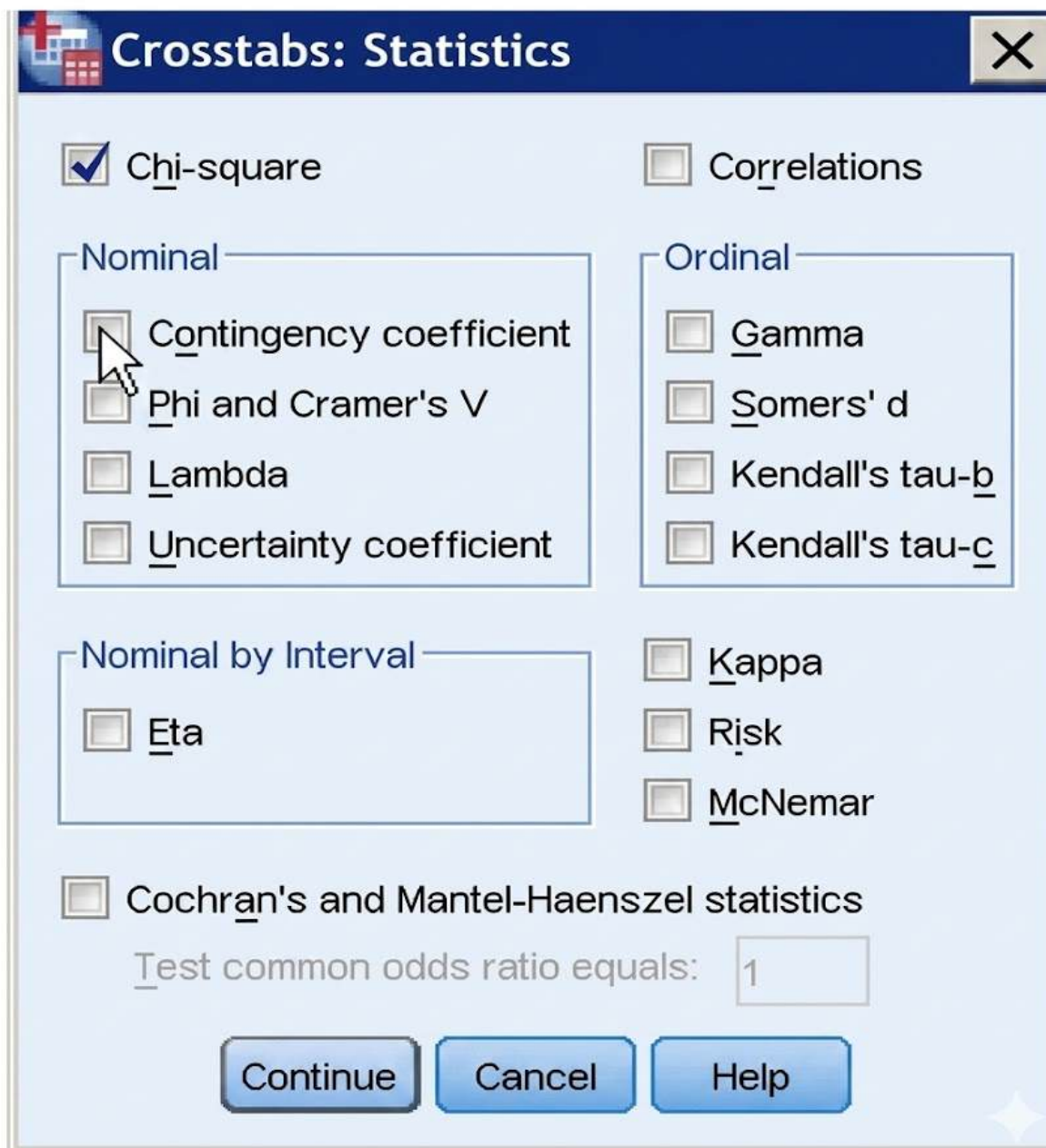
Nguồn tham khảo / Chú thích: [3] Britton, D 2020, SPSS eTutor: chi-square test of independence, Suny Empire State College, truy cập ngày 2 tháng 5 năm 2022.

Nhấn vào **Analyze** (Phân tích), **Descriptive Statistics** (Thống kê mô tả), **Crosstabs** (Bảng chéo). Nhấn vào tên biến phụ thuộc của bạn và đưa nó vào ô “Row” (Hàng), sau đó chọn biến độc lập và đưa vào ô “Column” (Cột). Trong ví dụ này, biến độc lập của chúng ta là **SEX** (Giới tính) và biến phụ thuộc là **PARTYID** (Nhận diện đảng phái).

Màn hình của bạn sẽ trông như thế này:



Bây giờ, hãy nhấn vào nút **Statistics** (Thống kê) và chọn **Chi-square**.



Nguồn tham khảo / Chú thích: [4] EZPSS (hoặc EXPSS), 2019, *Easy SPSS tutorial: chi-square test in SPSS, including interpretation* (Hướng dẫn SPSS dễ hiểu: phép kiểm định Chi-bình phương trong SPSS, bao gồm cả cách diễn giải), ngày 29 tháng 4, video trực tuyến, truy cập ngày 1 tháng 5 năm 2022.

Nhấn **OK**. Sau đó nhấn vào **Cells** (Ô). Hãy kiểm tra để đảm bảo rằng mục **Observed** (Quan sát) đã được chọn trong ô “count” (tần số) và các ô **Row** (Hàng), **Column** (Cột) và **Total** (Tổng cộng) đều được tích chọn trong ô “percentage” (phần trăm).

Crosstabs: Cell Display

Counts

- ☒ Observed
- ☐ Expected
- ☐ Hide small counts
Less than

z-test

- ☐ Compare column proportions
- ☐ Adjust p-values (Bonferroni method)

Percentages

- ☒ Row
- ☒ Column
- ☒ Total

Residuals

- ☐ Unstandardized
- ☐ Standardized
- ☐ Addjusted standardized

Noninteger Weights

- ☒ Round cell counts
- ☒ Round case weights
- ☐ Truncate cell counts
- ☐ Truncate case weights
- ☐ No adjustments

Nhấn **OK**. Kết quả đầu ra (output) của bạn sẽ trông như thế này:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
POLITICAL PARTY AFFILIATION * RESPONDENTS SEX	1488	99.2%	12	.8%	1500	100.0%

			RESPONDENTS SEX		Total
			MALE	FEMALE	
POLITICAL PARTY AFFILIATION	DEMOCRAT	Count	225	311	536
		% within POLITICAL PARTY AFFILIATION	42.0%	58.0%	100.0%
		% within RESPONDENTS SEX	32.6%	39.0%	36.0%
		% of Total	15.1%	20.9%	36.0%
	INDEPENDENT	Count	270	277	547
		% within POLITICAL PARTY AFFILIATION	49.4%	50.6%	100.0%
		% within RESPONDENTS SEX	39.1%	34.7%	36.8%
		% of Total	18.1%	18.6%	36.8%
	REPUBLICAN	Count	177	205	382
		% within POLITICAL PARTY AFFILIATION	46.3%	53.7%	100.0%
		% within RESPONDENTS SEX	25.7%	25.7%	25.7%
		% of Total	11.9%	13.8%	25.7%
	OTHER PARTY	Count	18	5	23
		% within POLITICAL PARTY AFFILIATION	78.3%	21.7%	100.0%
		% within RESPONDENTS SEX	2.6%	.6%	1.5%
		% of Total	1.2%	.3%	1.5%
Total	Count	690	798	1488	
	% within POLITICAL PARTY AFFILIATION	46.4%	53.6%	100.0%	
	% within RESPONDENTS SEX	100.0%	100.0%	100.0%	
	% of Total	46.4%	53.6%	100.0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	15.531 ^a	3	.001
Likelihood Ratio	15.956	3	.001
Linear-by-Linear Association	5.815	1	.016
N of Valid Cases	1488		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.67.

Symmetric Measures

	Value	Approx. Sig.
Nominal by Nominal Phi	.102	.001
Cramer's V	.102	.001
N of Valid Cases	1488	

Diễn giải phép kiểm định Chi-bình phương về tính độc lập

Một trong những yêu cầu của phép kiểm định Chi-bình phương là mỗi và mọi ô (cell) phải có tần số từ 5 trở lên. Trước tiên, bạn cần kiểm tra xem dữ liệu trong bảng của mình có đáp ứng yêu cầu này hay không. Hãy nhìn vào ghi chú dưới bảng **Chi-square Tests**. Kết quả đầu ra của chúng ta bao gồm thông tin này ở ghi chú ‘a’; không có ô nào của chúng ta có tần số nhỏ hơn 5, và do đó chúng ta không vi phạm giả định này của Chi-bình phương.

Bây giờ hãy nhìn vào giá trị **“Pearson Chi-Square Asymp. Sig (2 sided)”**. Vì phép kiểm định Chi-bình phương đang kiểm tra giả thuyết không (null hypothesis), giá trị

Sig. phải từ 0,05 trở xuống để mỗi quan hệ giữa các biến có ý nghĩa thống kê. Trong ví dụ này, giá trị Sig. là 0,001, vì vậy mỗi quan hệ giữa giới tính và nhận diện đảng phái chính trị có ý nghĩa thống kê rất mạnh.

Mối liên hệ giữa các biến khoảng / liên tục

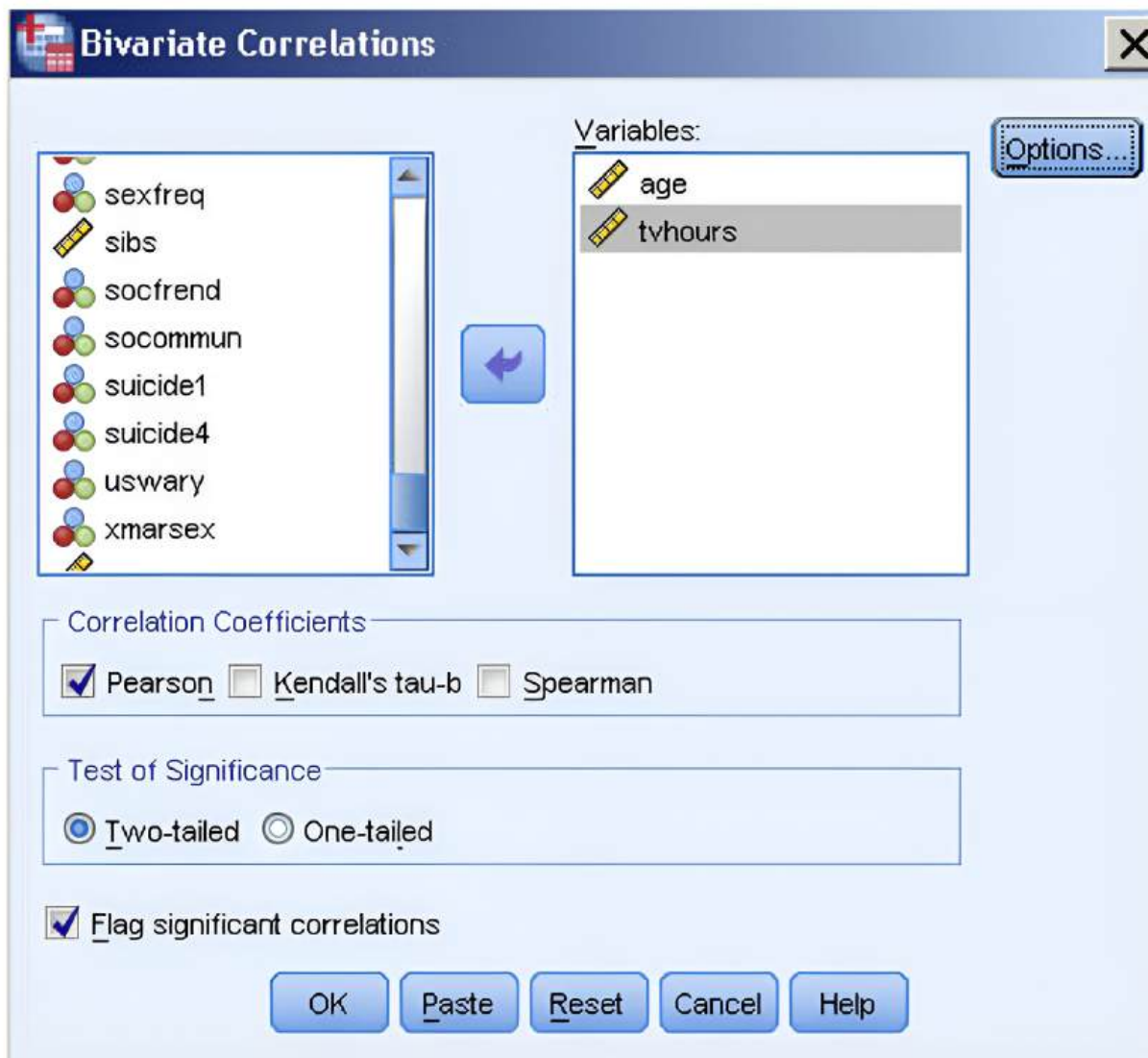
Khi mục tiêu là khám phá mối quan hệ giữa hai biến được đo lường bằng thang đo khoảng/tỷ lệ (interval/ratio scale), các nhà nghiên cứu sẽ sử dụng phép kiểm định tương quan.

Hệ số tương quan Pearson (Pearson's correlation)

Một phép kiểm định giả thuyết về hệ số tương quan Pearson được sử dụng để xác định xem liệu có mối tương quan tuyến tính có ý nghĩa thống kê giữa hai biến liên tục hay không (ví dụ: số tiền chi cho thực phẩm và độ tuổi của người tiêu dùng tính theo năm). [5]

Để tính toán hệ số r của Pearson, hãy truy cập: **Analyze** (Phân tích), **Correlate** (Tương quan), **Bivariate** (Hai biến). Nhập hai biến của bạn vào. Ví dụ, chúng ta có thể xem xét mối tương quan giữa hai biến liên tục là “Age” (Độ tuổi) và “TVhours” (số giờ xem TV mỗi ngày).

Màn hình của bạn sẽ trông như thế này:



Nguồn tham khảo / Chú thích:

[5] Britton, D 2020, SPSS eTutor: descriptive statistics, Suny Empire State College, truy cập ngày 2 tháng 5 năm 2022.

Nhấn **OK**. Tìm nội dung sau đây trong kết quả đầu ra (output) của bạn:

Correlations

		AGE OF RESPONDENT	HOURS PER DAY WATCHING TV
AGE OF RESPONDENT	Pearson Correlation	1	.139**
	Sig. (2-tailed)		.000
	N	1491	966
HOURS PER DAY WATCHING TV	Pearson Correlation	.139**	1
	Sig. (2-tailed)	.000	
	N	966	973

** . Correlation is significant at the 0.01 level (2-tailed).

Lưu ý rằng giá trị “r” của Pearson khi so sánh biến độ tuổi (age) với chính nó là 1, cho thấy mối tương quan hoàn hảo. Nếu bạn suy xét về điều này, nó hoàn toàn hợp lý về mặt logic. Điều bạn thực sự quan tâm là xem xét giá trị \$r\$ của Pearson giữa hai biến khác nhau (trong trường hợp này, giá trị là 0,139). Điều này gợi ý rằng khi một người già đi, họ xem tivi nhiều hơn.

Dưới đây là các hướng dẫn để diễn giải độ mạnh của mối liên hệ đối với Lambda, Gamma và hệ số r của Pearson (hãy nhớ rằng, Lambda chỉ có thể nhận giá trị dương):

Độ mạnh của mối liên hệ	Giá trị của Lambda, Gamma, P
Không có	0,00

Mối liên hệ yếu	+ 0,01 – 0,09
Mối liên hệ trung bình	+ 0,10 – 0,29
Dấu hiệu của mối liên hệ mạnh	+ 0,30 – 0,99
Mối liên hệ hoàn hảo, mạnh nhất có thể	+ 1,00

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=78#oembed-2>

Nguồn : How2stats⁶

Nguồn tham khảo / Chú thích: [6] How2stats 2011, *Pearson Correlation – SPSS EXPSS* (Tương quan Pearson – SPSS EXPSS), ngày 1 tháng 9, video trực tuyến, truy cập ngày 1 tháng 5 năm 2022.

18. SỰ KHÁC BIỆT GIỮA CÁC NHÓM ĐỐI TƯỢNG PHẢN HỒI

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu được sự khác biệt giữa kiểm định t-test cho hai mẫu độc lập (independent samples t-test) và kiểm định ANOVA;
- Diễn giải được kết quả đầu ra từ các phép kiểm định này..

Nội dung sau đây được phỏng theo tài liệu của Đại học Curtin [1] và được sử dụng theo giấy phép Creative Commons Ghi công - Chia sẻ tương tự 4.0.

Kiểm định t-test cho một mẫu (One sample t-test)

Kiểm định t-test cho một mẫu được sử dụng để kiểm tra xem liệu số trung bình của một mẫu thuộc biến liên tục có khác biệt có ý nghĩa so với một "giá trị kiểm định" (một giá trị giả định nào đó) hay không.

Ví dụ, bạn sẽ sử dụng nó nếu bạn có một mẫu điểm thi cuối kỳ của sinh viên và bạn muốn kiểm tra xem liệu chúng có đến từ một quần thể có điểm trung bình cuối kỳ bằng với mức trung bình của năm trước là 70 hay không. Trong trường hợp này, các giả thuyết sẽ là:

- **H0:** Mẫu đến từ một quần thể có điểm trung bình cuối kỳ là 70 ($\mu_{\text{điểm_cuối_kỳ}} = 70$)

- **HA:** Mẫu không đến từ một quần thể có điểm trung bình cuối kỳ là

$$70 (\mu_{\text{điểm_cuối_kỳ}} \neq 70)$$

Trước khi tiến hành kiểm định t-test cho một mẫu, bạn cần kiểm tra xem các giả định sau đây có giá trị hay không:

- **Giả định 1:** Mẫu là một mẫu ngẫu nhiên có tính đại diện cho quần thể.
- **Giả định 2:** Các quan sát là độc lập, nghĩa là các phép đo của một đối tượng không ảnh hưởng đến các phép đo của bất kỳ đối tượng nào khác.
- **Giả định 3:** Biến số có phân phối chuẩn, hoặc quy mô mẫu đủ lớn để đảm bảo tính phân phối chuẩn của phân phối chọn mẫu.

Nếu giả định cuối cùng về tính phân phối chuẩn bị vi phạm, hoặc nếu bạn có một biến thứ bậc thay vì một biến liên tục (chẳng hạn như xếp loại cuối kỳ F, 5, 6, 7, 8, 9, 10), thì phép kiểm định **Wilcoxon signed-rank** cho một mẫu nên được sử dụng thay thế.

Giả sử các giả định cho phép kiểm định t-test một mẫu đã được đáp ứng và phép kiểm định được thực hiện bằng phần mềm thống kê (ví dụ: SPSS trong ví dụ này), kết quả thu được sẽ trông giống như sau:

Nguồn tham khảo / Chú thích:

[1] Curtin University 2021, *Inferential statistics*, truy cập ngày 11 tháng 5 năm 2022.

One-Sample Statistics						
	N	Mean	Std. Deviation	Std. Error Mean		
Final_mark	40	73.1250	8.76528	1.38591		

One-Sample Test						
Test Value = 70						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Final_mark	2.255	39	.030	3.12500	.3217	5.9283

Lưu ý rằng bảng đầu tiên trong số các bảng này hiển thị các số liệu thống kê mô tả, bạn nên quan sát bảng này trước để nắm bắt được những gì đang diễn ra trong mẫu nghiên cứu. Ví dụ, số trung bình của mẫu là 73,125 so với giá trị kiểm định (test value) là 70, tạo ra một mức chênh lệch là 3,125 (giá trị này có thể được tự tính toán hoặc cũng được hiển thị trong bảng thứ hai).

Tuy nhiên, để kiểm tra xem sự khác biệt này có ý nghĩa thống kê hay không, chúng ta cần bảng thứ hai, đặc biệt là **giá trị p** (trong bảng này được liệt kê là '**Sig. (2-tailed)**') và khoảng tin cậy của sự khác biệt. Về giá trị p:

- **Nếu $p \leq 0,05$:** Chúng ta bác bỏ giả thuyết H_0 , nghĩa là mẫu đến từ một quần thể có số trung bình khác biệt đáng kể so với giá trị kiểm định.
- **Nếu $p > 0,05$:** Chúng ta không bác bỏ giả thuyết H_0 , nghĩa là mẫu đến từ một quần thể có số trung bình không khác biệt đáng kể so với giá trị kiểm định.

Trong trường hợp này, giá trị $p = 0,03$ của chúng ta cho thấy sự khác biệt này **có ý nghĩa thống kê**.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=80#h5p-31>

Điều này được xác nhận bởi khoảng tin cậy từ \$(0,3217; 5,9283)\$ cho sự sai khác giữa số trung bình quần thể và giá trị kiểm định của chúng ta (70). Bởi vì khoảng tin cậy này không chứa giá trị 0, nó một lần nữa cho thấy sự khác biệt có ý nghĩa thống kê. Trên thực tế, chúng ta tin tưởng 95% rằng số trung bình thực sự của quần thể cao hơn giá trị kiểm định của chúng ta từ 0,3217 đến 5,9283 điểm.

Lưu ý rằng mặc dù giá trị kiểm định (\$t\$) và bậc tự do (\$df\$) thường phải được báo cáo như một phần của kết quả nghiên cứu, bạn không cần phải diễn giải các chỉ số này khi đánh giá mức độ ý nghĩa của sự khác biệt.

Nếu bạn muốn thực hành diễn giải kết quả của phép kiểm định t-test cho một mẫu về ý nghĩa thống kê, hãy thử thực hiện một hoặc cả hai hoạt động sau đây:

Kiểm định t-test cho hai mẫu độc lập (Independent samples t-test)

Kiểm định t-test cho hai mẫu độc lập được sử dụng để kiểm tra xem liệu có sự khác biệt có ý nghĩa giữa các số trung bình mẫu của một biến liên tục đối với hai nhóm độc lập hay không. Ví dụ, bạn sẽ sử dụng nó nếu bạn thu thập dữ liệu về số giờ xem TV mỗi tuần và muốn kiểm tra xem có sự khác biệt đáng kể về số giờ trung bình giữa nam và nữ hay không. Trong trường hợp này, các giả thuyết sẽ là:

- H_0 : Không có sự khác biệt có ý nghĩa về số giờ xem TV mỗi tuần giữa nam và nữ ($\mu_{nam} = \mu_{nữ}$, hoặc $\mu_{nam} - \mu_{nữ} = 0$).
- H_A : Có sự khác biệt có ý nghĩa về số giờ xem TV mỗi tuần giữa nam và nữ ($\mu_{nam} \neq \mu_{nữ}$, hoặc $\mu_{nam} - \mu_{nữ} \neq 0$).

Trước khi tiến hành kiểm định t-test cho hai mẫu độc lập, bạn cần kiểm tra xem các giả định sau đây có giá trị hay không:

- **Giả định 1:** Mẫu là một mẫu ngẫu nhiên có tính đại diện cho quần thể.
- **Giả định 2:** Các quan sát là độc lập, nghĩa là các phép đo của một đối tượng không ảnh hưởng đến các phép đo của bất kỳ đối tượng nào khác.
- **Giả định 3:** Biến số có phân phối chuẩn cho cả hai nhóm, hoặc quy mô mẫu đủ lớn để đảm bảo tính phân phối chuẩn của phân phối chọn mẫu.

Nếu giả định cuối cùng về tính phân phối chuẩn bị vi phạm, hoặc nếu bạn có một biến thứ bậc thay vì một biến liên tục (chẳng hạn như số giờ được ghi nhận theo các khoảng), thì phép kiểm định **Mann-Whitney U** nên được sử dụng thay thế.

Giả sử các giả định cho phép kiểm định t-test hai mẫu độc lập đã được đáp ứng và phép kiểm định được thực hiện bằng phần mềm thống kê (ví dụ: SPSS trong ví dụ này), kết quả thu được sẽ trông giống như sau:

Lưu ý rằng bảng đầu tiên trong số các bảng này hiển thị các số liệu thống kê mô tả, bạn nên quan sát bảng này trước để nắm bắt được những gì đang diễn ra trong mẫu nghiên cứu. Ví dụ, sự khác biệt giữa các số trung bình mẫu của nam và nữ là 1,70833 (giá trị này có thể được tính toán từ các số trung bình ở bảng đầu tiên, và cũng được hiển thị ở bảng thứ hai – việc nó mang dấu âm đơn giản là vì nữ giới xem TV nhiều hơn nam giới).

Tiếp theo, lưu ý rằng thực sự có ba giá trị p (và hai khoảng tin cậy) trong bảng thứ hai; một giá trị được sử dụng trong tình huống phương sai của hai nhóm xấp xỉ bằng nhau (nằm ở cột '**Sig. (2-tailed)**') ở hàng trên cùng của bảng), một giá trị được sử dụng trong tình huống phương sai của hai nhóm không xấp xỉ bằng nhau (nằm ở cột '**Sig. (2-tailed)**') ở hàng dưới cùng của bảng), và giá trị còn lại (giá trị đầu tiên trong bảng, ở cột '**Sig.**') được sử dụng để xác định xem chúng ta đang ở trong tình huống nào.

Chúng ta cần diễn giải giá trị thứ ba này trước; đây là giá trị p cho **Kiểm định Levene về tính đồng nhất của phương sai (Levene's Test for Equality of Variances)**, trong đó giả thuyết không là các phương sai bằng nhau. Do đó:

- Nếu $p \leq 0,05$, đó là bằng chứng để bác bỏ giả thuyết không này và giả định rằng các phương sai không bằng nhau.
- Nếu $p > 0,05$, đó là bằng chứng để không bác bỏ giả thuyết không này và giả định rằng các phương sai bằng nhau.

Tùy thuộc vào trường hợp nào xảy ra, điều đó sẽ quyết định bạn nên diễn giải giá trị p (và khoảng tin cậy) nào trong số các giá trị còn lại. Trong trường hợp này, giá trị p bằng 0,966 có nghĩa là chúng ta có thể giả định các phương sai bằng nhau và nên diễn giải giá trị p ở hàng trên cùng của phần còn lại trong bảng.

Một lần nữa, cách diễn giải giá trị p này là:

- Nếu $p \leq 0,05$, chúng ta bác bỏ H_0 , nghĩa là số trung bình của hai nhóm khác biệt có ý nghĩa.
- Nếu $p > 0,05$, chúng ta không bác bỏ H_0 , nghĩa là số trung bình của hai nhóm không khác biệt có ý nghĩa.

Trong trường hợp này, giá trị p của chúng ta là 0,564 (ở hàng trên cùng của phần còn lại trong bảng) cho thấy sự khác biệt giữa các số trung bình **không có ý nghĩa thống kê**.

Điều này được xác nhận bởi khoảng tin cậy từ $(-7,654; 4,238)$ cho sự khác biệt giữa các số trung bình. Bởi vì khoảng tin cậy này có chứa giá trị 0, nó một lần nữa có nghĩa là sự khác biệt không có ý nghĩa thống kê. Chúng ta tin tưởng 95% rằng sự khác biệt về số giờ trung bình dành cho việc xem TV mỗi tuần giữa nam và nữ nằm trong khoảng từ -7,654 đến 4,238 giờ.

Lưu ý rằng mặc dù giá trị kiểm định (t) và bậc tự do (df) thường phải được báo cáo như một phần của kết quả nghiên cứu, bạn không cần phải diễn giải các chỉ số này khi đánh giá mức độ ý nghĩa của sự khác biệt.

Nếu bạn muốn thực hành diễn giải kết quả của phép kiểm định t-test cho hai mẫu độc lập về ý nghĩa thống kê, hãy thử thực hiện một hoặc cả hai hoạt động sau đây:

Phân tích phương sai một yếu tố (One-way ANOVA)

Phân tích phương sai một yếu tố (One-way ANOVA) tương tự như kiểm định t-test cho hai mẫu độc lập nhưng được sử dụng khi so sánh từ ba nhóm trở lên. Ví dụ: so sánh cân nặng trung bình của những người không tập thể dục, những người tập thể dục vừa phải và những người tập thể dục nhiều.

Giả thuyết không (null hypothesis) cho phân tích ANOVA một yếu tố tuyên bố rằng tất cả các số trung bình quần thể đều bằng nhau, trong khi giả thuyết đối (alternative hypothesis) tuyên bố rằng có ít nhất một trong số chúng là khác biệt.

Trước khi tiến hành phân tích ANOVA một yếu tố, bạn cần kiểm tra xem các giả định sau đây có giá trị hay không:

- **Giả định 1:** Mẫu là một mẫu ngẫu nhiên có tính đại diện cho quần thể.

- **Giả định 2:** Các quan sát là độc lập, nghĩa là các phép đo của một đối tượng không ảnh hưởng đến các phép đo của bất kỳ đối tượng nào khác.
- **Giả định 3:** Biến số có phân phối chuẩn cho từng nhóm, hoặc quy mô mẫu đủ lớn để đảm bảo tính phân phối chuẩn của phân phối chọn mẫu.
- **Giả định 4:** Các quần thể được so sánh có phương sai bằng nhau.

Nếu một phân tích ANOVA một yếu tố được tiến hành và kết quả cho thấy có ít nhất một trong các số trung bình là khác biệt, bạn sẽ cần phải điều tra thêm để xác định xem sự khác biệt nằm ở đâu bằng cách sử dụng các phép kiểm định hậu định (post hoc tests), ví dụ như kiểm định Tukey's HSD.

19. PHÂN TÍCH CẢM XÚC

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu được khái niệm về phân tích cảm xúc (sentiment analysis);
- Xác định được những ưu điểm và hạn chế của phân tích cảm xúc;
- Nắm được các bước khác nhau trong việc thực hiện phân tích cảm xúc.



Những chú chuột túi (kangaroo) trong khuôn viên trường Đại học Western Sydney © 2022 Đại học Western Sydney, được chụp bởi [Nhiếp ảnh gia Sally Tsoutas của Đại học Western Sydney](#), được cấp phép theo giấy phép [Ghi công - Phi thương mại - Không phái sinh 4.0 Quốc tế \(Attribution-NonCommercial-No Derivatives 4.0 International\)](#).

Phân tích cảm xúc là gì?

Phân tích cảm xúc (Sentiment analysis) trong tiếp thị là quá trình xác định ý kiến của mọi người về một sản phẩm hoặc một công ty bằng cách phân tích các bình luận của công chúng trên nhiều nền tảng trực tuyến khác nhau. Bên cạnh việc để lại bình luận trên trang web riêng của một tổ chức, người dùng trực tuyến cũng có thể cung cấp phản hồi và tham gia vào các cuộc thảo luận, trò chuyện trên nhiều nền tảng khác nhau. Các ví dụ về những nền tảng như vậy bao gồm các trang web truyền thông xã hội (ví dụ: Instagram), nền tảng kiến thức (như Quora), nền tảng chia sẻ phương tiện (ví dụ: YouTube, Vimeo, Spotify), nền tảng cung cấp dịch vụ (ví dụ: Uber, Airbnb) và một số trang web mua sắm (ví dụ: Amazon). Trước đây, khái niệm này được biết đến với tên gọi là ‘đánh giá sản phẩm’ (product reviews). Tuy nhiên, ngày nay với lượng thông tin được tạo ra, các nhà phân tích có khả năng ‘cào’ (crawl) web và khai thác dữ liệu liên quan. Các tuyên bố trực tuyến hoặc bình luận của khách hàng được chuyển thành dữ liệu phân loại (như “tích cực”, “tiêu cực” hoặc “trung tính”) và được tóm tắt lại để cung cấp cho người quản lý một cái nhìn tổng quan về cách công chúng đang phản hồi với thương hiệu hoặc sản phẩm của họ.

Mặc dù cảm xúc cũng có thể được bộc lộ trong các cuộc trò chuyện của khách hàng với trung tâm cuộc gọi của một tổ chức, phần lớn nghiên cứu trong lĩnh vực này đã tập trung vào việc phân tích các bình luận bằng văn bản trực tuyến của mọi người.

Những Ưu điểm khi Sử dụng Phân tích Cảm xúc

Việc đo lường cảm xúc của công chúng chưa bao giờ dễ dàng hơn, nhanh chóng hơn, rẻ hơn và ít bị thiên lệch hơn. Các công cụ tinh vi đã làm cho phân tích cảm xúc trở nên dễ tiếp cận đối với nhiều doanh nghiệp. Các nhà tiếp thị có khả năng thực hiện những thay đổi nhanh chóng đối với các chiến dịch của họ (ví dụ: thay đổi nhạc nền trong một quảng cáo của Expedia) dựa trên ‘cảm xúc’ được thể hiện. Khi so sánh,

việc sử dụng các công cụ nghiên cứu truyền thống sẽ cực kỳ tốn kém và mất thời gian. Hơn nữa, các kết quả cũng sẽ dễ mắc phải một mức độ sai sót nhất định do con người.

Phân tích cảm xúc có thể giúp tiết lộ những ‘người có ảnh hưởng’ (influencers) cho một thương hiệu hoặc một sản phẩm. Mặc dù nhiều công ty có thể ‘chiêu mộ’ những người có ảnh hưởng trên mạng xã hội để tiếp thị thương hiệu của họ, nhưng rất có thể một thương hiệu đã có sẵn một nhóm những người ủng hộ trung thành. Đây là những người đóng vai trò là những người ủng hộ (advocates) thương hiệu và có khả năng ảnh hưởng đến người khác thông qua các bình luận của họ. Mặc dù những người có ảnh hưởng trên mạng xã hội thường được xem là những người có lượng người theo dõi khổng lồ, hiện nay người ta đã nhận ra rằng những ‘người có ảnh hưởng vi mô’ (micro-influencers) cũng rất quan trọng. Số lượng người theo dõi của những micro-influencer có thể tương đối nhỏ, nhưng sự gắn kết của họ với những người theo dõi lại rất mạnh mẽ. Điều này có thể cực kỳ quan trọng khi mọi người đang tìm kiếm lời khuyên chân thành về những vấn đề quan trọng đối với họ.

Tương tự như điểm trên, việc có thể xác định ngôn từ kích động thù địch (hate speech) trực tuyến cũng quan trọng không kém, đặc biệt là khi nó liên quan đến thương hiệu hoặc bất kỳ giá trị cốt lõi nào của thương hiệu. Sự thù ghét thương hiệu (Brand hate) khác với những cảm xúc tiêu cực khác như sự thất vọng, không hài lòng hoặc bức bối đơn thuần. Có thể có những cấp độ thù ghét thương hiệu khác nhau với mức độ cực đoan nhất dẫn đến bắt nạt thương hiệu (brand bullying). Những người thể hiện cảm xúc tiêu cực mạnh mẽ như vậy cũng có khả năng tham gia vào truyền miệng trực tuyến tiêu cực (negative online word of mouth), phàn nàn công khai trực tuyến và thúc đẩy các chiến dịch tẩy chay. Phân tích cảm xúc rất hữu ích trong việc làm nổi bật một xu hướng như vậy, ngay cả khi nó chỉ giới hạn ở một nhóm nhất định. Các chiến lược để định hướng vượt qua sự tiêu cực như vậy của người tiêu dùng trong thế giới kỹ thuật số là cực kỳ quan trọng.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=82#h5p-32>

Các bước trong phân tích cảm xúc

Dưới đây là một quy trình đơn giản để có thể thực hiện phân tích cảm xúc:

1. Phát hiện cảm xúc (Detecting sentiment) Đây là mức độ cơ bản nhất của việc thực hiện phân tích cảm xúc. Nó đơn giản chỉ cung cấp một cái nhìn tổng quan nhanh chóng về ý kiến chung của khách hàng. Điều này có nghĩa là duyệt qua các bình luận trực tuyến và trích xuất dữ liệu mang tính quan điểm (chẳng hạn như “Tôi yêu thích điều này!”). Ngược lại, có thể có những loại dữ liệu khác như dữ liệu khách quan (ví dụ: “nhà hàng nằm ở trung tâm thành phố”), vốn có thể không thể hiện bất kỳ cảm xúc nào của người tiêu dùng.

2. Phân loại cảm xúc và xác định cường độ (Categorising sentiment and identifying the intensity) Bước này liên quan đến việc phát hiện xem cảm xúc là tích cực, tiêu cực hay trung tính. Tùy thuộc vào loại phần mềm được sử dụng, các nhà quản lý cũng có thể thêm trọng số cho các danh mục này, ví dụ:

- Rất tích cực (very positive)
- Hơi tích cực (somewhat positive)
- Trung tính (neutral)
- Hơi tiêu cực (somewhat negative)
- Rất tiêu cực (very negative)

Các danh mục này thể hiện cường độ của cảm xúc.

3. Hàm ý hỗn hợp (Mixed Connotation) Đôi khi, văn bản chứa đựng những ý kiến hỗn hợp hoặc mâu thuẫn, ví dụ: “nhân viên rất thân thiện nhưng chúng tôi phải đợi quá lâu mới được phục vụ”. Người ta sẽ cần phải diễn giải riêng biệt các câu nói như vậy, điều mà một chương trình đơn giản dựa trên máy móc có thể gặp khó khăn trong việc mã hóa và phân tích.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=82#oembed-1>

Nguồn: Fun Robotics [1]

Những hạn chế của phân tích cảm xúc

Dưới đây là một vài hạn chế chính của các công cụ phân tích cảm xúc đã được công nhận trong ngành:

- Phân tích cảm xúc phụ thuộc vào máy móc dựa trên cách mọi người sử dụng ngôn ngữ để bày tỏ ý kiến của họ. Ý nghĩa của một từ trong từ điển có thể rất khác so với cách mọi người sử dụng nó trong các cuộc trò chuyện hàng ngày. Phân tích cảm xúc có thể gặp trục trặc khi người dùng trực tuyến sử dụng các cụm từ để thể hiện sự mỉa mai (sarcasm).
- Người ta vẫn cho rằng phân tích cảm xúc đạt hiệu quả cao nhất khi được sử dụng với các tập dữ liệu lớn và phong phú. Các doanh nghiệp nhỏ có thể nhận thấy rằng không có nhiều dữ liệu sẵn có về sản phẩm/dịch vụ của họ để có thể phân tích một cách hiệu quả.
- Hơn nữa, phân tích cảm xúc không phải là một hoạt động chỉ thực hiện một lần. Nó phải được tích hợp vào chiến lược thu thập thông tin của một công ty. Để thu được giá trị thực sự từ các công cụ phân tích cảm xúc, người ta cần phân tích một lượng lớn dữ liệu văn bản một cách thường xuyên.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=82#h5p-33>

Nguồn tham khảo / Chú thích: [1] Fun Robotics, 2021, What is sentiment analysis? (Phân tích cảm xúc là gì?), ngày 1 tháng 4, video trực tuyến, truy cập ngày 10 tháng 5 năm 2022.

Bảng 2: Các ví dụ về các loại cảm xúc khác nhau, được phân loại thành cảm xúc tích cực và tiêu cực

Loại cảm xúc	Câu phát biểu mẫu (Ví dụ minh họa)
Vui vẻ (Joy)	"Thật thư giãn khi dùng bữa tại nhà hàng này"
Bực bội (Frustration)	"Đây là lần thứ ba tôi cố gọi cho bạn để đặt lịch hẹn"
Tin tưởng (Trust)	"Tôi sẽ chỉ đến Ecco để mua giày"
Tức giận (Anger)	"Lũ \$#_#@U*%() này không chỉ gửi sai hóa đơn cho tôi mà còn phạt tôi vì thanh toán trễ"
Sợ hãi (Fear)	"Tôi không nghĩ mình sẽ mua sắm ở Westfield sau tất cả các ca nhiễm COVID ngày hôm qua"

Buồn bã (Sadness)	"McCain không còn làm món lasagna chay nữa 😞"
Ngạc nhiên (Surprise)	"Tôi không nhận ra là Coles có nướng bánh mì tươi!"

20. TRÍ TUỆ NHÂN TẠO VÀ THÔNG TIN

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Mô tả vai trò của Trí tuệ Nhân tạo (Artificial Intelligence) trong Nghiên cứu Tiếp thị (Marketing Research);
- Hiểu được những lợi ích của Trí tuệ Nhân tạo trong Nghiên cứu;
- Giải thích những hạn chế của công nghệ Trí tuệ Nhân tạo.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=84#oembed-1>

Nguồn: Digital Marketing Made Easy [1] Tài liệu sau đây được phỏng theo Huang và Rust [2] và được sử dụng theo giấy phép Creative Commons Ghi công - Chia sẻ tương tự 4.0 ([Creative Commons Attribution ShareAlike 4.0 Licence](https://creativecommons.org/licenses/by-sa/4.0/)).

Trí tuệ nhân tạo (AI) trong tiếp thị hiện đang ngày càng trở nên quan trọng, nhờ vào sức mạnh tính toán ngày càng tăng, chi phí tính toán giảm, sự sẵn có của dữ liệu lớn (big data) và sự tiến bộ của các thuật toán cũng như mô hình học máy (machine learning). Chúng ta thấy các ứng dụng rộng rãi của AI trong nhiều lĩnh vực tiếp thị khác nhau. Ví dụ: Prime Air của Amazon.com sử dụng máy bay không người lái (drones) để tự động hóa việc vận chuyển và giao hàng. Domino's pizza đang thử nghiệm ô tô tự lái và robot giao hàng để giao pizza đến tận cửa nhà khách hàng.

Trí tuệ nhân tạo (AI) đề cập đến trí thông minh được thể hiện bởi máy móc, trái ngược với trí thông minh tự nhiên được thể hiện bởi con người và động vật. Trí tuệ nhân tạo là việc máy móc hành xử một cách thông minh, thực hiện giải quyết vấn đề hoặc hoàn thành các nhiệm vụ. Đây là một chủ đề rất rộng bao gồm khoa học máy tính, khoa học nhận thức, toán học, triết học, khoa học thần kinh, ngôn ngữ học, v.v.

Trí tuệ Nhân tạo (AI) và Nghiên cứu Tiếp thị

AI có thể được sử dụng theo ba cách khác nhau trong nghiên cứu tiếp thị:

a) thu thập dữ liệu;

b) phân tích thị trường;

và c) thấu hiểu khách hàng và phát triển những thông tin chi tiết (insights) quan trọng.

AI và Thu thập Dữ liệu

AI có thể tự động hóa việc thu thập dữ liệu về thị trường, môi trường, công ty, đối thủ cạnh tranh và khách hàng. Trong một thế giới kết nối kỹ thuật số, dữ liệu thị trường có thể được theo dõi và giám sát một cách dễ dàng. Việc cảm biến, theo dõi và thu thập dữ liệu là những công việc lặp đi lặp lại, mang tính thường quy có thể dễ dàng được tự động hóa thông qua AI.

Các nghiên cứu hiện tại đã chỉ ra nhiều cách khác nhau để sử dụng AI cho việc thu thập dữ liệu. Ví dụ:

- Thông tin tình báo về khách hàng (customer intelligence), bao gồm dữ liệu về người tiêu dùng, các hoạt động và môi trường của họ, có thể được thu thập nếu họ sử dụng các thiết bị được kết nối (ví dụ: thiết bị đeo thông minh, tủ lạnh thông minh và trợ lý kỹ thuật số như Siri).
- Quá trình sử dụng sản phẩm và trải nghiệm tiêu dùng có thể được trực quan hóa bằng Internet Vạn vật (Internet of Things - IoT).
- Các công nghệ và phân tích tiên tiến khác nhau có thể nắm bắt dữ liệu hoạt động tiếp thị phi cấu trúc.
- Các cảm biến trong xe hơi có thể theo dõi hành vi lái xe để xác định phí bảo hiểm.
- Các công nghệ bán lẻ, chẳng hạn như bản đồ nhiệt (heat maps), camera giám sát (video surveillance) và đèn hiệu (beacons), có thể được sử dụng để lập hồ sơ và nhận diện người mua sắm tại cửa hàng.

Các nghiên cứu trước đây cho thấy rằng, với tính chất lặp đi lặp lại, thường quy nhưng khối lượng lớn của dữ liệu thị trường, AI có thể thu thập dữ liệu một cách hiệu quả ở quy mô lớn. Khả năng thu thập dữ liệu của AI không chỉ giới hạn ở dữ liệu hành vi có thể quan sát được; nó cũng có thể được sử dụng để tạo điều kiện thuận lợi cho việc thu thập dữ liệu khảo sát hoặc thực nghiệm nhằm nắm bắt đặc điểm tâm lý, ý kiến và thái độ của người tiêu dùng. Ví dụ, việc con người quản lý và giám sát các cuộc khảo sát đang diễn ra không còn cần thiết nữa và có thể được tự động hóa.

SurveyMonkey và SurveyCake là hai nền tảng khảo sát thương mại giúp tự động hóa việc thiết kế khảo sát và thu thập dữ liệu. Bằng cách kết hợp các nguyên tắc của trò chơi hóa (gaming) và AI, dữ liệu chất lượng tốt hơn đang được thu thập trực tuyến. Các cuộc khảo sát trực tuyến đã trở nên thú vị và hấp dẫn hơn. Những người tham gia khảo sát cũng có nhiều khả năng hoàn thành các cuộc khảo sát này hơn và giới thiệu

chúng cho người khác. Trên thực tế, bằng cách sử dụng công nghệ AI, các nhà nghiên cứu tin rằng trong vài năm tới, gần 1 trong 4 cuộc khảo sát sẽ được thực hiện bằng giọng nói với một trợ lý kỹ thuật số (Qualtrics 2022).

AI và Phân tích Dữ liệu

AI có thể được sử dụng để phân tích thông tin. Dưới đây là một vài ví dụ về cách các nhà nghiên cứu đã sử dụng AI để thực hiện phân tích:

- **Phân tích văn bản tự động** (Automated text analysis) được áp dụng để có được thông tin chi tiết về hành vi của người tiêu dùng (ví dụ: bằng cách phân tích các đánh giá của khách hàng trên một trang web).
- **Các thuật toán học máy** (Machine learning algorithms) và phân loại văn bản dựa trên từ vựng (lexicon-based text classification) được sử dụng để phân tích các tập dữ liệu mạng xã hội khác nhau.

Nguồn tham khảo / Chú thích: [1] Digital Marketing Made Easy 2021, *Artificial intelligence explained in 3 minutes: 3 applications in marketing*, ngày 14 tháng 5, video trực tuyến, truy cập ngày 10 tháng 5 năm 2022. https://www.youtube.com/watch?v=jBaH_g4QKAc

[2] Huang, MH & Rust, RT 2020, 'A strategic framework for artificial intelligence in marketing', *Journal of the Academy of Marketing Science*, tập 49, tr. 30-50, truy cập ngày 11 tháng 5 năm 2022. <https://link.springer.com/article/10.1007/s11747-020-00749-9>

- **Phân tích tiếp thị dữ liệu lớn** (Big data marketing analytics) là một cách tiếp cận chủ đạo để tạo ra những hiểu biết sâu sắc (insights) về tiếp thị. Các ứng dụng cụ thể bao gồm lập bản đồ cấu trúc thị trường cho các chủng loại hàng hóa bán lẻ lớn bằng cách sử dụng mô hình ngôn ngữ mạng nơ-ron (neural network language model), thông qua việc phân tích sự xuất hiện đồng thời của các sản phẩm trong giỏ hàng; phát hiện các ứng dụng di động sao chép bằng phương pháp phát hiện sao chép của học máy; và hỗ trợ kỹ thuật nội dung mạng xã hội bằng cách sử dụng các thuật toán xử lý ngôn ngữ tự nhiên (NLP) để khám phá mối liên hệ giữa nội dung tiếp thị trên mạng xã hội và mức độ tương tác của người dùng.

AI và Cải thiện Sự thấu hiểu Khách hàng

AI có thể được sử dụng để thấu hiểu các nhu cầu và mong muốn của khách hàng hiện tại và tiềm năng, ví dụ như họ là ai, họ muốn gì và các giải pháp hiện tại của họ là gì. Sự khác biệt chính giữa phân tích thị trường và thấu hiểu khách hàng là yếu tố sau thường liên quan đến dữ liệu cảm xúc về tình cảm, cảm giác, sở thích và thái độ của khách hàng. Do đó, AI không chỉ đưa ra các con số một cách nhanh chóng mà còn có khả năng phân tích dữ liệu cảm xúc. Ví dụ, Affectiva đã hợp tác với Ford để tạo ra công nghệ phân tích cảm xúc AutoEmotive, nhằm cố gắng nhận biết trạng thái cảm xúc của người lái xe.

Đối với những khách hàng tiềm năng, các nhà tiếp thị có thể sử dụng AI cảm xúc (feeling AI) để hiểu những gì họ muốn và tại sao họ lại hài lòng với các đối thủ cạnh tranh hoặc các lựa chọn bên ngoài. Một số nhà tiếp thị có thể cảm thấy rằng nhu cầu và mong muốn của khách hàng tiềm năng khó dự đoán hơn vì dữ liệu cảm xúc của họ ít có sẵn hơn. Tuy nhiên, Unilever đã chứng minh điều ngược lại. Với sự trợ giúp của các trung tâm dữ liệu AI, Unilever đã tổng hợp các thông tin chi tiết thông qua quá trình lắng nghe mạng xã hội (social listening) và CRM, đồng thời phát hiện ra mối liên hệ giữa kem và bữa sáng. Unilever phát hiện ra rằng có ít nhất 50 bài hát có cụm

từ ‘kem cho bữa sáng’ ('ice cream for breakfast') trong lời bài hát (mà mọi người đang nghe), cũng như việc Dunkin Donuts bán kem vào buổi sáng. Với thông tin chi tiết này, Unilever đã phát triển một thương hiệu mới Ben & Jerry's với một loạt các loại kem có hương vị ngũ cốc dùng cho bữa sáng.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây: <https://westernsydney.pressbooks.pub/customerinsights2/?p=84#h5p-5>

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây: <https://westernsydney.pressbooks.pub>

ChatGPT

Việc ra mắt chatbot ChatGPT gần đây vào tháng 11 năm 2022 đã tạo ra sự quan tâm đáng kể từ công chúng. Về bản chất, [ChatGPT là một chatbot được hỗ trợ bởi AI](#), cho phép người dùng mô phỏng các cuộc trò chuyện giống con người với một trí tuệ nhân tạo.

GPT là viết tắt của **Generative Pre-trained Transformer** (Mô hình Transformer tạo sinh được huấn luyện trước), [một mô hình xử lý ngôn ngữ do công ty trí tuệ nhân tạo OpenAI của Mỹ phát triển](#). Mô hình ngôn ngữ GPT sử dụng [học sâu](#) (deep learning) để tạo ra các phản hồi giống con người. Học sâu là một nhánh của học máy (machine

learning) liên quan đến việc huấn luyện các mạng nơ-ron nhân tạo (artificial neural networks) bắt chước sự phức tạp của bộ não con người, nhằm tạo ra các phản hồi giống con người.

[ChatGPT](#) có một giao diện thân thiện với người dùng sử dụng công nghệ này, cho phép người dùng tương tác với nó theo cách trò chuyện. Trước công nghệ mới này, cả doanh nghiệp và người tiêu dùng đều thể hiện sự quan tâm lớn đối với việc một sự đổi mới như vậy có thể cách mạng hóa các chiến lược tiếp thị và trải nghiệm khách hàng như thế nào.

Phần về ChatGPT ở trên được trích từ [Fares 2023](#) [3] và được sử dụng theo giấy phép [Creative Commons Ghi công - Phi thương mại - Không phái sinh 2.0 Chung \(CC BY-NC-ND 2.0\)](#).

Khi người dùng đưa ra một truy vấn, công cụ tìm kiếm phổ biến nhất thế giới Google – sử dụng một lộ trình truyền thống – sẽ quét hàng tỷ trang web và trình bày các kết quả phù hợp nhất. Mặt khác, ChatGPT phục vụ một mục đích tương tự lại đưa ra một phản hồi tập trung hơn, giống con người hơn. Thay vì cung cấp cho người dùng một danh sách các 'kết quả tìm kiếm', ChatGPT hành xử như một chuyên gia. Ví dụ, khi được yêu cầu tìm một công thức luộc trứng, Google cung cấp cho chúng ta các liên kết web khác nhau và để người dùng chọn kết quả phù hợp nhất. Khi so sánh, ChatGPT cung cấp cho người dùng một phản hồi được thiết kế riêng biệt bằng ngôn ngữ giống con người [4].

Phần "Điều gì làm cho ChatGPT trở nên đặc biệt?" ở trên được trích từ [Fares 2023](#) [5] và được sử dụng theo giấy phép [Creative Commons Ghi công - Phi thương mại - Không phái sinh 2.0 Chung \(CC BY-NC-ND 2.0\)](#).

Những tương tác như vậy với người dùng được xem là 'có ý nghĩa' hơn từ góc nhìn của người dùng. Trong khi đó, ChatGPT đủ thông minh để không đưa ra bất kỳ câu trả lời nào cho những câu hỏi không phù hợp về mặt tình dục hoặc có thể làm tổn hại đến tình cảm của các chủng tộc và tôn giáo trên toàn cầu.

Điều gì làm cho ChatGPT trở nên đặc biệt?

Điều làm cho ChatGPT trở nên khác biệt so với các chatbot khác chính là quy mô tập dữ liệu của nó. Các chatbot thông thường được huấn luyện trên một tập dữ liệu nhỏ hơn theo cách dựa trên quy tắc (rule-based), được thiết kế để trả lời các câu hỏi cụ thể và thực hiện một số tác vụ nhất định.

Mặt khác, ChatGPT được huấn luyện trên một tập dữ liệu khổng lồ — **175 tỷ tham số** (parameters) và **570 gigabyte** — và có khả năng thực hiện hàng loạt các tác vụ trong các lĩnh vực và ngành nghề khác nhau. Dành cho mục đích so sánh, 570GB tương đương với hơn **385 triệu trang** trên Microsoft Word.

Với khối lượng dữ liệu này, ChatGPT có thể thực hiện nhiều hoạt động liên quan đến ngôn ngữ khác nhau, bao gồm trả lời các câu hỏi trong nhiều lĩnh vực và ban ngành, cung cấp câu trả lời bằng nhiều ngôn ngữ khác nhau và tạo ra nội dung.

Đăng ký ChatGPT

Đăng ký tài khoản ChatGPT tại đây: <https://chat.openai.com/auth/login>

Khi thông tin của bạn đã được xác minh, hãy bắt đầu thử nghiệm với nó. ChatGPT có thể được sử dụng để dịch văn bản từ ngôn ngữ này sang ngôn ngữ khác, tóm tắt thông

tin, thực hiện viết sáng tạo (chẳng hạn như làm thơ), động não tìm ý tưởng (brainstorm) và cung cấp các phản hồi cho những thắc mắc và khiếu nại của khách hàng.

Hình ảnh do AI tạo ra

Có rất nhiều công cụ tạo hình ảnh bằng AI (AI image generators) có sẵn cho người dùng. OpenAI hiện có một phiên bản beta (thử nghiệm) tại đây: <https://labs.openai.com/> (ứng dụng xem trước DALL.E). Mô hình DALL.E tạo ra một hình ảnh chỉ trong vòng vài giây từ bất kỳ câu lệnh văn bản (text prompt) nào.

Một công cụ như vậy mang đến cơ hội để tất cả mọi người được thỏa sức sáng tạo. Người dùng có thể nhập vào một mô tả chi tiết (dài tối đa 400 ký tự) và một hình ảnh sẽ được tạo ra trong vòng vài giây. DALL.E 2 có khả năng xử lý nhiều loại chất liệu nghệ thuật khác nhau, vì vậy một hình ảnh có thể được tạo ra dưới dạng một bức tranh vẽ, tranh kính màu (stained glass), nghệ thuật kỹ thuật số (digital art), v.v.

Các hình ảnh sau đây được tạo ra từ mô tả: *"các sinh viên đại học phong cách siêu nhân đang ăn kem và học về thống kê từ một con mèo"* ("superman style uni students eating ice cream and learning about statistics from a cat").

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=84#h5p-35>

Nếu người dùng không hài lòng với các hình ảnh được tạo ra, họ có thể tinh chỉnh phần mô tả. Ví dụ, mô tả *"bức tranh màu nước vẽ các sinh viên đại học phong cách*

siêu nhân đang ăn kem và học về thống kê từ một con mèo" tạo ra hình ảnh này cùng với các hình ảnh khác:



AI watercolour painting

Nguồn tham khảo / Chú thích: [3] Fares, OH 2023, *ChatGPT could be a game-changer for marketers, but it won't replace humans any time soon*, The Conversation. [4] Singh, S, 2022, *ChatGPT Vs Google: the ultimate comparison of 2023*, <https://www.demandsage.com/chatgpt-vs-google/> [5] Fares, OH 2023, *ChatGPT could be a game-changer for marketers, but it won't replace humans any time soon*, The Conversation.

Những hạn chế của Trí tuệ Nhân tạo



Trí tuệ cảm xúc (Emotional intelligence):

ChatGPT cung cấp nội dung và phản hồi giống con người một cách tối tân. Tuy nhiên, điều quan trọng là phải nhận thức được rằng công cụ này chỉ *giống* con người. [Tương tự như những thách thức truyền thống đối với các chatbot](#), mức độ giống con người sẽ đóng vai trò thiết yếu đối với việc cải thiện quy trình và sáng tạo nội dung.

Các nhà tiếp thị có thể sử dụng ChatGPT để nâng cao trải nghiệm khách hàng, nhưng nếu không có con người để mang lại sự phù hợp, tính cách, kinh nghiệm và [sự kết nối cá nhân](#), sẽ rất khó để tận dụng tối đa ChatGPT. Việc chỉ dựa vào ChatGPT để xây dựng các kết nối và sự gắn kết của khách hàng mà không có sự tham gia của con người có thể làm suy giảm sự kết nối có ý nghĩa với khách hàng thay vì nâng cao nó.

Phần "Trí tuệ cảm xúc" ở trên được trích từ [Fares 2023](#) [6] và được sử dụng theo giấy phép [Creative Commons Ghi công - Phi thương mại - Không phái sinh 2.0 Chung \(CC BY-NC-ND 2.0\)](#).

Chi phí Tài chính và Nguồn nhân lực:

Việc sử dụng AI đòi hỏi các nguồn lực về tài chính và con người. Chi phí ban đầu để thiết lập cơ sở hạ tầng AI luôn tiêu tốn nhiều nguồn lực. Do đó, nó có thể không khả thi đối với tất cả các tổ chức. Đầu tư vào AI có thể đồng nghĩa với những khoản chi tiêu khổng lồ cho việc thu thập dữ liệu, thiết bị tính toán và lưu trữ, cũng như chi phí tuyển dụng nhân sự có liên quan và đào tạo họ. Giống như mọi khoản đầu tư khác, lợi nhuận có thể không đến ngay lập tức. Có thể mất một thời gian để nhận thấy được những lợi ích của việc sử dụng AI.

Các hệ thống AI không nên được xem là sự thay thế cho nguồn vốn nhân lực (human capital). Trên thực tế, các hệ thống AI chỉ hoạt động tốt nhất khi được tích hợp một cách tối ưu với chuyên môn của tổ chức. Việc để khách hàng hoặc nhà cung cấp phó mặc hoàn toàn cho một ‘trợ lý kỹ thuật số’ là điều không nên làm. Có rất nhiều nghiên cứu tình huống (case studies) đã chứng minh những hạn chế của cách tiếp cận như vậy.

Những Kỳ vọng Phi thực tế từ AI / Sự đe dọa từ AI:

Đôi khi, các nhà quản lý có thể có những kỳ vọng phi thực tế từ các công nghệ AI. Không phải mọi khía cạnh của AI đều sẽ gia tăng giá trị cho tất cả các hoạt động kinh doanh. Các nhà quản lý cần phân tích những gì là khả thi để đạt được các mục tiêu cụ thể. Tác giả của cuốn sách này đã từng làm việc trong lĩnh vực robot xã hội (social robotics).

Nguồn tham khảo / Chú thích: [6] Fares, OH 2023, *ChatGPT could be a game-changer for marketers, but it won't replace humans any time soon...*

Việc triển khai một robot xã hội trong môi trường kinh doanh thường đặt ra nhiều thách thức. Hoặc là kỳ vọng của các nhà quản lý doanh nghiệp quá cao đối với robot xã hội (ví dụ: “nó có thể trả lời điện thoại được không?”) hoặc – trong một số trường hợp – nó không giúp ích gì cho các mục tiêu giao tiếp tổng thể của họ. Trong một vài trường hợp khác, nhân viên cảm thấy bị đe dọa bởi sự xuất hiện của một robot xã hội vì họ sợ mất việc làm. Trong nhiều trường hợp, việc đánh giá tổng thể về công nghệ lại được thực hiện bởi những cá nhân có thể đã có sẵn định kiến (bias) đối với việc sử dụng máy móc.

Một yếu tố tương tác H5P đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=84#h5p-36>

21. PHÂN TÍCH MẠNG XÃ HỘI

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Xác định các loại hình phân tích mạng xã hội (social media analytics) khác nhau và chức năng của chúng;

- Liệt kê các bước thực hiện đánh giá hiệu suất mạng xã hội của một công ty bằng cách sử dụng các phân tích.
-

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=86#oembed-1>

Nguồn: Netbase Quid [1] Tài liệu sau đây được phỏng theo Brosius và Cless [2] và được sử dụng theo giấy phép [Creative Commons Ghi công - Chia sẻ tương tự 4.0 \(Creative Commons Attribution ShareAlike 4.0 Licence\)](#).

Việc ứng dụng mạng xã hội trong các chương trình đã phát triển theo cấp số nhân trong thập kỷ qua và đã trở thành một cách thức phổ biến để các chương trình thu hút và tiếp cận các bên liên quan (stakeholders) của họ cũng như cung cấp thông tin cho các nỗ lực tương tác. Do đó, các tổ chức đang sử dụng phân tích dữ liệu từ các nền tảng mạng xã hội như một cách để đo lường mức độ tác động. Những dữ liệu này có thể giúp các nhà quản lý hiểu được cách thức mà các mục tiêu, tiến độ và kết quả của chương trình được phổ biến và sử dụng (ví dụ: thông qua các cuộc thảo luận, lượt xem nội dung, theo dõi các trang mạng xã hội của chương trình).

Mạng xã hội cho phép các chương trình:

- Tiếp cận những đối tượng khán giả rộng lớn và đa dạng;
- Thúc đẩy giao tiếp cởi mở và sự hợp tác;
- Nhận được phản hồi tức thì;
- Dự đoán các tác động trong tương lai – “Việc dự báo dựa trên mạng xã hội đã chứng minh được hiệu quả đáng kinh ngạc trong nhiều lĩnh vực khác nhau, bao gồm dự đoán giá cổ phiếu, kết quả bầu cử và doanh thu phòng vé phim.” (Priem 2014) [3].

Nguồn tham khảo / Chú thích: [1] Netbase Quid 2016, *The importance of social media analytics for business*, ngày 28 tháng 6, video trực tuyến, truy cập ngày 1 tháng 5 năm 2022.

<https://www.youtube.com/watch?v=ocK8o9ROkhE>

[2]. Brosius, L & Cless, A 2019, *Utilizing social media analytics to demonstrate program impact*, viewed 11 May 2022, https://evalu-ate.org/blog/brosius_cless_nov2019

[3] Priem, J 2014, ‘Altmetrics’, trong B. Cronin & C. R. Sugimoto (Chủ biên), *Beyond bibliometrics: harnessing multidimensional indicators of scholarly impact*, The MIT Press, Cambridge, tr. 263-288. https://evalu-ate.org/blog/brosius_cless_nov2019/



Thực tế ảo xã hội (Social Virtual Reality - hay Social VR) gần đây đã chứng kiến sự gia tăng về số lượng người dùng. Ảnh: [Sally Tsoutas © 2022 Đại học Western Sydney](#), được chụp bởi Nhiếp ảnh gia Sally Tsoutas của Đại học Western Sydney, được cấp phép theo giấy phép [Ghi công - Phi thương mại - Không phái sinh 4.0 Quốc tế \(Attribution-NonCommercial-NoDerivatives 4.0 International\)](#).

Phân tích Mạng xã hội là gì?

Phân tích mạng xã hội (Social media analytics) là khả năng thu thập và tìm ra ý nghĩa từ dữ liệu thu thập được từ các kênh xã hội để hỗ trợ các quyết định kinh doanh — và đo lường hiệu suất của các hành động dựa trên những quyết định đó thông qua mạng xã hội. Phân tích mạng xã hội có ý nghĩa rộng hơn các chỉ số (metrics) như lượt thích

(likes), lượt theo dõi (follows), lượt đăng lại (retweets), lượt xem trước (previews), lượt nhấp (clicks) và lượt hiển thị (impressions) thu thập được từ các kênh riêng lẻ.

Phân tích mạng xã hội sử dụng các nền tảng phần mềm được thiết kế đặc biệt (ví dụ: Twitonomy) hoạt động tương tự như các công cụ tìm kiếm trên web. Dữ liệu về các từ khóa hoặc chủ đề được truy xuất thông qua các truy vấn tìm kiếm hoặc các trình 'thu thập dữ liệu' web (web 'crawlers') trải rộng trên nhiều kênh. Các đoạn văn bản được trả về, tải vào cơ sở dữ liệu, phân loại và phân tích để rút ra những thông tin chi tiết (insights) có ý nghĩa.

Có nên tiến hành phân tích mạng xã hội không?

Các chỉ số có sẵn cho mạng xã hội rất phong phú và không phải tất cả đều hữu ích để xác định tác động từ những nỗ lực trên mạng xã hội của một chương trình. Như Sterne (2010) [4] giải thích, các chỉ số mạng xã hội cần phải có ý nghĩa bởi vì “đo lường chỉ vì mục đích đo lường là một việc làm vô ích”; “không có bối cảnh, các phép đo của bạn là vô nghĩa”; và “không có các mục tiêu kinh doanh cụ thể, các chỉ số của bạn là vô nghĩa.” Do đó, điều quan trọng là phải xem xét các mục tiêu cụ thể của chương trình và những chỉ số nào (các chỉ số hiệu suất chính - KPIs) là trọng tâm để đánh giá tiến độ và sự thành công của những mục tiêu này.

Ngoài ra, cũng cần nhận ra rằng các nền tảng mạng xã hội phổ biến luôn thay đổi, việc phân loại các nền tảng mạng xã hội khác nhau rất khó khăn và các chỉ số được sử dụng bởi các nền tảng khác nhau cũng có sự khác biệt.

Để mang lại nhiều ý nghĩa hơn cho các phân tích mạng xã hội của một chương trình, có thể sẽ hữu ích khi xem xét sử dụng một khuôn khổ (framework) để cung cấp cấu trúc cho việc điều chỉnh các chỉ số mạng xã hội phù hợp với mục tiêu của chương trình, đồng thời hỗ trợ chứng minh tiến độ và sự thành công đối với các mục tiêu đó.

Biết khi nào và liệu có nên tiến hành phân tích mạng xã hội hay không là một khái niệm quan trọng cần xem xét. Chỉ vì một phân tích mạng xã hội *có thể* được tiến hành không có nghĩa là nó *nên* được tiến hành. Do đó, trước khi bắt đầu, điều quan trọng là phải dành thời gian để xác định một vài điều:

1. Các mục tiêu cụ thể nào sẽ được giải đáp thông qua mạng xã hội?
2. Các mục tiêu này sẽ được đo lường như thế nào bằng cách sử dụng mạng xã hội?
3. Nền tảng nào sẽ có giá trị/hữu ích nhất trong việc tiếp cận đối tượng mục tiêu?

Các Chỉ số Mạng xã hội trên các nền tảng khác nhau

Bảng và các tài liệu sau đây được phỏng theo Forsyth, E và Doherty, A 2015 [5] và được sử dụng theo giấy phép [Creative Commons Ghi công - Phi thương mại - Chia sẻ tương tự 4.0 \(Creative Commons Attribution Non-Commercial ShareAlike 4.0 Licence\)](#).

Các chỉ số mạng xã hội được sử dụng để đo lường xem liệu các mục tiêu truyền thông cụ thể có đang đạt được hay không. Hiệu suất của một tổ chức trên mạng xã hội không nên chỉ được đo lường thông qua phân tích định lượng. Việc phân tích định tính các bình luận và bài đăng cũng quan trọng không kém.

Nguồn tham khảo / Chú thích: [4] Sterne, J 2010, *Social media metrics: how to measure and optimize your marketing investment*, John Wiley, Hoboken NJ. [5] Forsyth, E và Doherty, A 2015, [Social media analytics in an imperfect world](#), Australian Library and Information Association.

Bảng: Các Chỉ số Mạng xã hội và Các Mục tiêu Liên quan

Chỉ số Mạng xã hội(Social Media Metric)	Chỉ số Hiệu suất(Performance Indicator)	Mục tiêu Tiếp thị/Truyền thông(Marketing/Communication Objective)
• Số lượng và tần suất bài đăng	Sự hiện diện trên mạng xã hội <i>(Social Media presence)</i>	Sự hiện diện cơ bản trên bất kỳ nền tảng mạng xã hội nào.
• Số lượng người theo dõi/người hâm mộ • Số lượt xem ('bài đăng') • Số lượng người theo dõi 'hủy đăng ký' • Nhân khẩu học của người theo dõi	Phạm vi tiếp cận <i>(Reach)</i>	Số lượng và hồ sơ (profile) của những người được tiếp cận.

• Số lượng 'bình luận' • Số lượt 'chia sẻ' hoặc 'đăng lại' • Số lượt 'thích' • Số lượt 'tải xuống' • Số lượt 'nhắc đến' hoặc 'gắn thẻ' • Khối lượng thảo luận được tạo ra	Mức độ tương tác <i>(Engagement)</i>	Số lượng người 'tương tác' với nội dung.
• Tỷ lệ bình luận tích cực so với tổng số bình luận • Tỷ lệ bình luận tiêu cực so với tổng số bình luận	Sự hài lòng <i>(Satisfaction)</i>	Số lượng người cảm thấy hài lòng.
• Số lượt nhấp (click-throughs) đến trang đặt hàng	Hành động được thực hiện	Số lượng người đặt hàng (ví dụ), hoặc đăng ký tham gia một sự kiện, hoặc

• Số lượng đơn hàng được đặt	<i>(Action taken)</i>	đăng ký để nhận thông tin cập nhật.
------------------------------	-----------------------	-------------------------------------

Sau khi thu thập dữ liệu từ các nền tảng mạng xã hội của một công ty, việc phân tích dữ liệu là rất quan trọng. Kết quả từ quá trình phân tích có thể được tổ chức lại để bao gồm các thông tin sau:

- Mức độ mà thông điệp của công ty được xem trên các nền tảng mạng xã hội khác nhau
- Loại đối tượng khán giả đang xem/tương tác với nội dung
- Sự tương tác và sở thích của khán giả đối với nội dung được đăng tải trên các nền tảng mạng xã hội khác nhau
- Những lĩnh vực trọng tâm tiềm năng cho các nỗ lực trên mạng xã hội trong tương lai của công ty

Cách theo dõi hiệu suất mạng xã hội bằng các công cụ phân tích mạng xã hội

Dưới đây là một quy trình có hệ thống để theo dõi hiệu suất mạng xã hội của một công ty:

1. Xác định các mục tiêu cho chiến dịch mạng xã hội của công ty (ví dụ: đạt được **10.000** người theo dõi).
2. Chọn các nền tảng mạng xã hội để sử dụng.
3. Chọn các chỉ số (metrics) liên quan để đo lường hiệu suất so với từng mục tiêu.
4. Thiết lập các điểm chuẩn (benchmarks) / tốc độ tăng trưởng cho từng chỉ số để chứng minh tiến độ.

5. Tạo một báo cáo phân tích mạng xã hội (xác định các xu hướng mới nổi / thông tin chi tiết về hành vi của khách hàng).
 6. Thực hiện hành động để cải thiện hiệu suất mạng xã hội của bạn (ví dụ: tạo điều kiện dễ dàng để khách hàng gửi truy vấn trên Facebook).
 7. Sửa đổi và xem xét lại (các chỉ số mạng xã hội) một cách thường xuyên.
-

22. NGHIÊN CỨU CÁC CỘNG ĐỒNG BẢN ĐỊA

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu rõ sự phân bố của dân số Bản địa trên khắp nước Úc;
- Nhận thức được lịch sử giữa các cộng đồng Bản địa ở Úc và các nhà nghiên cứu;
- Nhận thức được quy tắc đạo đức mà tất cả các nhà nghiên cứu phải tuân thủ khi thực hiện nghiên cứu với các cộng đồng Bản địa;
- Làm quen với các phương pháp nghiên cứu đặc thù được sử dụng với các cộng đồng Bản địa.

Giới thiệu

Tại Úc, có 812.000 người được xác định là người Thổ dân (Aboriginal) và/hoặc người dân đảo Torres Strait trong Cuộc Điều tra Dân số và Nhà ở năm 2021. Người Thổ dân và người dân đảo Torres Strait chiếm 3,2% dân số. Con số này tăng từ 2,8% vào năm 2016 và 2,5% vào năm 2011.

Năm 2021, tỷ lệ lớn nhất của dân số Thổ dân và người dân đảo Torres Strait sống ở New South Wales (34,2%) và Queensland (29,2%). Mặc dù chỉ có 7,5% dân số Thổ dân và người dân đảo Torres Strait của Úc sống ở Lãnh thổ Bắc Úc (Northern Territory), nhưng họ chiếm tới hơn một phần tư (26,3%) dân số của Lãnh thổ này. Con số này cao hơn nhiều so với các tiểu bang và vùng lãnh thổ khác [1].

Bối cảnh Lịch sử của Nghiên cứu với người Úc Bản địa:

Việc thực hiện nghiên cứu có đạo đức với các cộng đồng Thổ dân và người dân đảo Torres Strait là một chủ đề quan trọng. Từ những thời kỳ đầu của quá trình thực dân hóa, các quan điểm về văn hóa và tổ chức xã hội của Thổ dân và người dân đảo Torres Strait (bao gồm cả các giá trị và tập tục của họ) đều dựa trên những nhận thức và giả định thiếu hiểu biết. Những nhận thức này nảy sinh từ những sự so sánh không phù hợp giữa người Thổ dân và người dân đảo Torres Strait với các quan điểm về tinh thần, xã hội, chính trị và kinh tế của những kẻ thực dân châu Âu.

Những kẻ thực dân đã nhìn nhận văn hóa và xã hội của Thổ dân và người dân đảo Torres Strait thông qua lăng kính văn hóa của riêng họ và đánh giá những thuộc tính này dựa trên mức độ mà họ cho là phù hợp với các phong tục và chuẩn mực của châu Âu. Không có gì đáng ngạc nhiên khi những kẻ thực dân thời kỳ đầu không biết gì về các Dân tộc Thổ dân và người dân đảo Torres Strait cũng như văn hóa của họ. Những sai lầm nghiêm trọng trong đánh giá và những quan niệm sai lầm kéo theo sau đó đã tác động đáng kể đến các Dân tộc Thổ dân và người dân đảo Torres Strait kể từ đó cho đến nay [2].

Có những lo ngại cho rằng người Úc Bản địa đã bị "nghiên cứu quá mức" (over-researched) mà không có những cải thiện tương xứng, đặc biệt là liên quan đến các kết quả về sức khỏe. Các lý do chính khiến việc chuyển hóa các phát hiện nghiên cứu thành các chỉ số về thay đổi xã hội hoặc lợi ích diễn ra kém cỏi phần lớn là do nghiên cứu đã bị kiểm soát bởi những người không phải là người Bản địa, được tiến

hành 'trên' người Bản địa; và thiên vị mạnh mẽ đối với các động lực của xã hội thực dân.

Mặc dù có một số cải cách gắn liền với các phong trào chính trị và môi trường Bản địa theo thời gian, một sự ngờ vực sâu xa đối với các nhà nghiên cứu và các hoạt động nghiên cứu liên quan vẫn tồn tại trong các quần thể dân cư Bản địa. Theo giai thoại, các dự án và các cuộc tham vấn của chính phủ về việc thực hiện chính sách công được thực hiện kém hiệu quả cũng đã khiến các cộng đồng người Úc Bản địa phân loại các hoạt động này là nghiên cứu; điều này thường làm tăng thêm nhận thức rằng nghiên cứu đã không mang lại lợi ích cho các cộng đồng [3].

Hành động của một số nhà nghiên cứu cũng đã làm tổn hại đến niềm tin của các cộng đồng Thổ dân vào các hoạt động nghiên cứu. Năm 1948, Frank Setzler, người phụ trách chính về nhân chủng học tại Viện Smithsonian ở Washington, là thành viên của một đoàn thám hiểm khoa học Mỹ - Úc. Các ghi chú thực địa của chính nhà nghiên cứu này đã ghi lại các quá trình bí mật mà ông ta đã áp dụng để vận chuyển hàng trăm 'khúc xương' của những người đã khuất sang Mỹ. Những bộ xương này đã trở thành một phần của một trong những bộ sưu tập giải phẫu người lớn nhất thế giới. Các bộ sưu tập này vẫn nằm lại tại Viện Smithsonian trong hơn 60 năm. Cuối cùng, sau quá trình vận động hành lang căng thẳng, những bộ xương này đã được trả về Úc vào năm 2011 [4].

Cuối cùng, các nhà nghiên cứu tiếp thị có thể thấy khó khăn trong việc xây dựng mối quan hệ với phân khúc này của dân số Úc. Việc thực hiện nghiên cứu đòi hỏi sự tin tưởng, điều mà một phần đã bị xói mòn do cách đối xử của giới doanh nghiệp Úc đối với người tiêu dùng và các cộng đồng Bản địa. Sự việc [Rio Tinto phá hủy hẻm núi Juukan Gorge](#) là một trong những hành động không thể tha thứ gần đây nhất.

[Woolworths đã buộc phải từ bỏ kế hoạch xây dựng một kho chứa rượu gần ba cộng đồng Thổ dân cấm rượu](#) (dry communities). Sau đó, vào năm 2021, [Telstra đã bị tòa](#)

[án liên bang phạt 50 triệu đô la](#) vì bị phát hiện đã bóc lột các khách hàng Bản địa bằng cách đăng ký cho họ các hợp đồng điện thoại mà họ không thể hiểu hoặc không có khả năng chi trả [5].

Cuối cùng, các trò lừa đảo tiếp thị cũng ảnh hưởng đến người Úc Bản địa. Năm 2020, Scamwatch đã nhận được 3.455 báo cáo với thiệt hại hơn 2 triệu đô la từ người Úc Bản địa. Mặc dù mức thiệt hại thấp hơn 4% so với năm 2019, nhưng số lượng báo cáo lại tăng gần 25%. Các trò lừa đảo gây thiệt hại tài chính nặng nề nhất cho các cộng đồng Bản địa là lừa đảo hẹn hò và tình cảm, tiếp theo là lừa đảo đầu tư và lừa đảo mua sắm trực tuyến.

Nguồn tham khảo / Chú thích:

[1] Cục Thống kê Úc (Australian Bureau of Statistics), 2022, *Aboriginal and Torres Strait Islander population summary*.

<https://www.abs.gov.au/articles/australia-aboriginal-andtorres-strait-islander-population-summary>

[2] Hội đồng Nghiên cứu Y tế và Sức khỏe Quốc gia (National Health and Medical Research Council), 2018, *Ethical guidelines for research with Aboriginal and Torres Strait Islander peoples*.

<https://www.nhmrc.gov.au/research-policy/ethics/ethical-guidelines-researchaboriginal-and-torres-strait-islander-peoples>

[3] Bainbridge, T et al. 2015, 'No one's discussing the elephant in the room: contemplating questions of research impact and benefit in Aboriginal and Torres Strait Islander Australian health research', *BMC Public Health*, tập 15, số 696.

[4] Dickie, M 2018, *The American anthro who stole our bones in broad daylight*, NITimes.

<https://www.nit.com.au/the-american-anthro-who-stole-our-bones-in-broad-daylight/>.

[broad-daylight/](#).

[5] Thông tin về vụ phạt Telstra từ quyết định của tòa án liên bang (2021). Remeikis, A 2021, *Australian corporations' treatment of Indigenous customers to be investigated by inquiry*.

<https://www.theguardian.com/australia-news/2021/nov/09/australian-corporations-treatment-of-indigenous-customers-to-be-investigated-by-inquiry>

Quy tắc Đạo đức trong Nghiên cứu với các Dân tộc Thổ dân và Người dân đảo Torres Strait

Trước những lo ngại nêu trên, AIATSIS (Viện Nghiên cứu Thổ dân và Người dân đảo Torres Strait của Úc) đã lần đầu tiên công bố các hướng dẫn về đạo đức vào năm 1999 [6]. Bộ Quy tắc này đại diện cho một cách tiếp cận mới bằng cách định vị người Bản địa là đối tác nghiên cứu, chứ không đơn thuần là đối tượng nghiên cứu. Bộ Quy tắc thiết lập các tiêu chuẩn quốc gia cho việc thực hiện có đạo đức và trách nhiệm đối với tất cả các nghiên cứu về Thổ dân và người dân đảo Torres Strait, trên tất cả các lĩnh vực và phương pháp luận. Quy tắc này dành cho những người thực hiện nghiên cứu, đánh giá nghiên cứu hoặc tài trợ nghiên cứu, bao gồm các cá nhân, trường đại học, chính phủ, doanh nghiệp và các tổ chức cộng đồng.

Nghiên cứu về Thổ dân và người dân đảo Torres Strait bao gồm tất cả các nghiên cứu có tác động hoặc có ý nghĩa đặc biệt đối với các dân tộc Thổ dân và người dân đảo Torres Strait, bao gồm việc lập kế hoạch, thu thập, phân tích và phổ biến thông tin hoặc kiến thức, dưới bất kỳ định dạng hoặc phương tiện nào, có nội dung về hoặc có thể ảnh hưởng đến các dân tộc Bản địa, dù là tập thể hay cá nhân.

Chi tiết về bốn nguyên tắc hình thành nên khuôn khổ đạo đức được khuyến nghị có thể được tìm thấy tại đây:

<https://aiatsis.gov.au/sites/default/files/2020-10/aiatsis-code-ethics.pdf>

Nguyên tắc 1: Quyền tự quyết của người Bản địa Việc công nhận và tôn trọng quyền tự quyết của các dân tộc Thổ dân và người dân đảo Torres Strait là nền tảng cho tất cả các nghiên cứu được tiến hành tại Úc. Bộ Quy tắc này công nhận mối liên hệ độc đáo của các dân tộc Thổ dân và người dân đảo Torres Strait với đất đai và vùng nước của lục địa này từ thời xa xưa. Nghiên cứu tại Úc nên thừa nhận thực tế rằng các dân tộc Thổ dân và người dân đảo Torres Strait đã tồn tại liên tục như những xã hội riêng biệt, với các luật lệ, văn hóa, kiến thức và thể giới quan đa dạng và độc đáo;

những yếu tố này có thể cung cấp thông tin cho nghiên cứu trên nhiều lĩnh vực bao gồm khoa học tự nhiên, khoa học xã hội và nhân văn.

Nguyên tắc 2: Sự lãnh đạo của người Bản địa Để chứng minh giá trị, nghiên cứu về Thổ dân và người dân đảo Torres Strait nên được dẫn dắt bởi người Bản địa. Nghiên cứu được coi là do người Bản địa dẫn dắt tại Úc khi người Thổ dân và/hoặc người dân đảo Torres Strait có trách nhiệm ra quyết định thực sự và nghiên cứu được định hình bởi các ưu tiên, giá trị, quan điểm và tiếng nói của người Bản địa. Sự lãnh đạo của người Bản địa phải được thể hiện rõ ràng trong cả lý do (tại sao) cũng như cách thức (như thế nào) thực hiện nghiên cứu, từ khâu hình thành ý tưởng đến khâu truyền thông kết quả nghiên cứu.

Nguyên tắc 3: Tác động và giá trị Nghiên cứu với các dân tộc Bản địa phải nhằm mục đích mang lại lợi ích cho chính họ. Hơn nữa, để sự đồng ý dựa trên thông tin có hiệu lực, cần phải có sự hiểu biết lẫn nhau về các lợi ích và rủi ro.

Nguyên tắc 4: Tính bền vững và trách nhiệm giải trình Sự tôn trọng đối với các dân tộc Thổ dân, người dân đảo Torres Strait và thế giới quan của họ trong việc tiến hành nghiên cứu đòi hỏi các nhà nghiên cứu và các tổ chức phải có trách nhiệm giải trình trong dài hạn đối với những tác động từ hành động của mình.

Ngoài các nguyên tắc này, cũng cần đề cập đến vai trò của các Trưởng lão Bản địa (Elders) trong bất kỳ dự án nào liên quan đến nghiên cứu. Các Trưởng lão trong cộng đồng được tôn trọng vì những câu chuyện, nghệ thuật, bài hát và ngôn ngữ của họ. Trong khi nhiều người trong số họ đang chăm sóc cháu chắt, họ cũng được tham vấn về các chương trình và dịch vụ. Các Trưởng lão còn được biết đến với vai trò hỗ trợ tại tòa án địa phương. Những người đàn ông và phụ nữ đang đối mặt với các cáo buộc sẽ ngồi cùng các "Dì" và "Chú" (Aunts and Uncles) để trò chuyện (yarn) về hành trình của họ và tìm hướng đi tiếp theo. Các Trưởng lão không chỉ đóng vai trò thiết yếu trong việc bảo tồn kiến thức truyền thống mà còn là then chốt trong việc giúp giải

quyết các vấn đề rộng lớn hơn của cộng đồng. Bất kỳ công trình nghiên cứu nào nhằm tìm hiểu các cộng đồng này đều cần nhận thức được tầm quan trọng của việc tham vấn các Trưởng lão [7].

Nguồn trích dẫn trong văn bản:

5. Remeikis, A 2021, *Australian corporations' treatment of Indigenous customers to be investigated by inquiry*,
<https://www.theguardian.com/australia-news/2021/mar/01/australian-corporations-treatment-of-indigenous-customers-to-be-investigated-by-inquiry>.
6. Australian Institute of Aboriginal and Torres Strait Islander Studies [AIATSIS] 2020, *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander research*,
<https://aiatsis.gov.au/research/ethical-research/code-ethics>.
7. Mroelli, L 2017, *What role does an Elder have in Indigenous Communities?*, NITV,
<https://www.sbs.com.au/nitv/article/2017/02/22/what-role-does-elder-have-indigenous-communities>.

Các phương pháp nghiên cứu Bản địa

Trong hầu hết các trường hợp, các phương pháp nghiên cứu được phát triển bởi các học giả và những người thực hành phương Tây. Mặc dù các phương pháp này có thể hữu ích trong việc thu thập dữ liệu trong phạm vi cuộc sống của các nhà nghiên cứu, chúng có thể không hoạt động tốt với các cộng đồng Bản địa. Do đó, cần phải "giải thực dân hóa" (decolonise) quá trình nghiên cứu để làm cho nó phù hợp hơn với việc nghiên cứu các quần thể Thổ dân và người dân đảo Torres Strait.

Phòng vấn và thảo luận nhóm tập trung — những công cụ chính để tương tác với người tham gia — chỉ có thể cung cấp dữ liệu chất lượng phong phú nếu mối quan hệ giữa nhà nghiên cứu và người tham gia dựa trên sự tin tưởng và tính xác thực. Sự tương tác giữa hai bên sẽ chỉ dừng lại ở mức bề mặt nếu người tham gia nhận thấy mục tiêu nghiên cứu thuần túy chỉ dựa trên việc thu thập dữ liệu. Để củng cố mối quan hệ, phải có một mục tiêu chân thành là giúp người tham gia cải thiện cuộc sống của chính họ.

Chúng ta sẽ thảo luận về hai phương pháp đã nhận được sự chú ý từ các học giả:

Yarning [8] (Trò chuyện chân tình)

Yarning, bắt nguồn từ từ "yarn" (một sợi len, bông, v.v.), về mặt ngôn ngữ học có thể đề cập đến hành động sử dụng sợi để đan hoặc dệt. Tuy nhiên, khi được sử dụng trong tiếng Anh giao tiếp, nó hàm ý việc kể một "câu chuyện" hoặc một câu chuyện giải trí.

‘Yarning’ trong nghiên cứu: Yarning là một truyền thống được thực hiện hàng ngàn năm bởi nhiều người dân thuộc các Quốc gia Đầu tiên (First Nations) tại Úc. Đây là một phần không thể thiếu trong cách học tập và chia sẻ của người Bản địa.

Hoạt động này thường được thực hiện bởi những người Thổ dân tụ họp lại với nhau một cách không chính thức để thư giãn, hoặc theo những cách trang trọng hơn như

thảo luận về các vấn đề cộng đồng hoặc văn hóa. Kể chuyện là một phần quan trọng của *yarning*, cho phép suy ngẫm về lịch sử gần đây hoặc quá khứ, những trải nghiệm sống và chia sẻ kiến thức.

Các nhà nghiên cứu có thể tham gia vào "yarning" bằng cách trò chuyện với những người thuộc các Quốc gia Đầu tiên về việc mỗi người đến từ đâu, những người họ cùng biết, và mối liên hệ của họ với nơi mà họ gặp gỡ, chỉ để nêu một vài ví dụ.

Đối với các nhà nghiên cứu không phải người Thổ dân, việc thiết lập mối quan hệ với người dân thuộc các Quốc gia Đầu tiên là điều thiết yếu khi tiến hành nghiên cứu trong cộng đồng của họ. Đây là nơi các phương pháp nghiên cứu như "yarning" có thể mang lại cơ hội để thiết lập mối quan hệ với các cộng đồng này.

Một khi các nhà nghiên cứu thiết lập được sự kết nối với người dân tại nơi họ mong muốn thực hiện nghiên cứu, một mối quan hệ đôi bên cùng có lợi và hòa nhập có thể được hình thành. Điều này rất quan trọng để đảm bảo những người tham gia nghiên cứu thuộc các Quốc gia Đầu tiên được tính là một phần của nghiên cứu, chứ không bị xem là đối tượng nghiên cứu.

Khả năng xây dựng mối quan hệ là yếu tố sống còn để đảm bảo cuộc sống của người dân thuộc các Quốc gia Đầu tiên được mô tả và ghi lại một cách chính xác, người tham gia không bị lợi dụng và cộng đồng có thể hưởng lợi từ nghiên cứu.

Các phương pháp dựa trên nghệ thuật (Art-based methods)

Các phương pháp dựa trên nghệ thuật nắm giữ nhiều cách thức nhận thức vốn thường ít có không gian trong diễn ngôn khoa học. Các phương pháp dựa trên nghệ thuật (ABMs) bao gồm một loạt các hình thức như sân khấu, nghệ thuật thị giác, hội họa, nhiếp ảnh, âm nhạc và khiêu vũ.

Người Thổ dân truyền tải những câu chuyện văn hóa quan trọng của họ qua các thể hệ thông qua các biểu tượng/hình tượng trong tác phẩm nghệ thuật của họ. Một trong

nhiều phương pháp dựa trên nghệ thuật là *Photovoice*, theo đó mọi người xác định, đại diện và nâng cao cộng đồng của mình thông qua nhiếp ảnh. Trong phương pháp nghiên cứu này, nhà nghiên cứu cung cấp máy ảnh cho các cá nhân, giúp họ đóng vai trò là người ghi chép và là chất xúc tác tiềm năng cho các hành động và thay đổi xã hội trong chính cộng đồng của họ. Người tham gia chụp những bức ảnh giúp kể câu chuyện của họ về một mối quan tâm cụ thể. Câu chuyện của cá nhân thường đi kèm với các bức ảnh để thúc đẩy một phương tiện chia sẻ chuyên môn mang tính tham gia và hiệu quả.

LƯU Ý: Chúng tôi đã dựa vào các đồng nghiệp và các nhà lãnh đạo bản địa trong Trường Kinh doanh để xem xét tài liệu này và cung cấp phản hồi cho chúng tôi.

Nguồn trích dẫn:

8. Ober, R, Oliver, R & Dovchin, S 2022, 'You can't just show up and start asking questions': why researchers need to understand the importance of yarning for First Nations', Australian Council of Graduate Research, <https://www.acgr.edu.au/you-cant-just-show-up-and-start-asking-questions-why-researchers-need-to-understand-the-importance-of-yarning-for-first-nations/>.

23. TRUYỀN TẢI CÁC THÔNG TIN CHI TIẾT (INSIGHTS)

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Làm quen với các cách thức khác nhau để trình bày kết quả nghiên cứu;
- Có khả năng áp dụng các loại biểu đồ khác nhau để trình bày kết quả.

Chương này được phỏng theo [Cục Thống kê Úc](#) (Australian Bureau of Statistics) [1]; được cấp phép theo giấy phép quốc tế [Creative Commons Ghi công 4.0](#) (Creative Commons Attribution 4.0 International).

Trình bày kết quả theo một định dạng rõ ràng và logic cho khách hàng là một trong những nhiệm vụ quan trọng nhất đối với người quản lý khảo sát. Khi trình bày kết quả, hình thức thuyết trình nên được thiết kế riêng để giải quyết các mục đích, mục tiêu của cuộc khảo sát và làm hài lòng những người sẽ sử dụng kết quả đó. Cần cân nhắc đến mức độ hiểu biết về thống kê của khách hàng và người dùng, đặc biệt là đối với các thuật ngữ thống kê. Việc trình bày cần phải hiệu quả, dễ hiểu và truyền tải được các đặc điểm chính của dữ liệu.

Báo cáo bằng văn bản

Khi trình bày dữ liệu, một hình thức báo cáo bằng văn bản là điều thiết yếu. Báo cáo nên truyền tải các đặc điểm chính một cách rõ ràng và tuân theo một tiến trình logic, sử dụng ít thuật ngữ chuyên môn nhất có thể, cung cấp thông tin chi tiết (insight) về

dữ liệu và làm cho các kết quả trở nên thú vị nhất có thể. Với tư cách là tác giả của báo cáo, bạn nên truyền tải các thông điệp cụ thể thay vì thông tin khái quát hóa. Bạn cũng nên đưa ra các giả thuyết để giải thích các phát hiện và khuyến khích người dùng khám phá thêm dữ liệu. Nếu phù hợp, bạn có thể cần bao gồm các kiến nghị.

Nội dung của báo cáo và sự cân bằng giữa câu chữ, bảng biểu và biểu đồ sẽ phụ thuộc tự nhiên vào chủ đề, kết quả và đối tượng độc giả tiềm năng. Biểu đồ và bảng biểu truyền tải thông tin phức tạp một cách rõ ràng và có thể được sử dụng để tạo sự đa dạng. Trong mọi trường hợp, cần nhớ rằng báo cáo được viết ra để đọc, vì vậy nó cần phải có ý nghĩa và dễ hiểu.

Một báo cáo khảo sát thường bao gồm:

Giới thiệu (Introduction) Phần giới thiệu nêu rõ mục đích, mục tiêu của cuộc khảo sát và mục tiêu của báo cáo; đưa ra bối cảnh nghiên cứu; định nghĩa các thuật ngữ và khái niệm; và nêu rõ liệu cuộc khảo sát là để kiểm chứng một giả thuyết hay mang tính khám phá.

Phương pháp luận (Methodology) Phần phương pháp luận mô tả phương pháp lấy mẫu và thông tin về quần thể khảo sát, cũng như cách thức dữ liệu được phân tích và các quy trình thống kê đã được sử dụng.

Phát hiện và Phân tích (Findings and Analysis) Phần phát hiện và phân tích là phần chính của báo cáo, giải quyết chi tiết về số lượng mẫu, tỷ lệ phản hồi, kết quả và việc diễn giải các bảng biểu, v.v., đồng thời thảo luận về các hướng hành động khả thi.

Kết luận và Kiến nghị (Conclusions and Recommendations) Phần kết luận tóm tắt các phát hiện chính của báo cáo và trả lời các câu hỏi đã đặt ra ở phần

giới thiệu. Các kiến nghị phác thảo những hành động cần thực hiện dựa trên cơ sở các kết luận.

Phụ lục và Tài liệu tham khảo (Appendices and References) Phần phụ lục bao gồm các mục có thể hữu ích cho người đọc (ví dụ: bảng câu hỏi) nhưng không thiết yếu đối với báo cáo. Tài liệu tham khảo liệt kê các sách, tạp chí và bài báo được trích dẫn trong nghiên cứu.

Cần nhớ rằng báo cáo có khả năng phục vụ như một cơ sở để thảo luận, vì vậy một số lưu ý quan trọng khác là tiêu đề, cách sử dụng các tiêu đề chính và tiêu đề phụ, màu sắc và thiết kế của trang bìa cũng như diện mạo tổng thể của báo cáo – nó cần phải kích thích sự quan tâm của người đọc.

Nguồn trích dẫn:

1. Australian Bureau of Statistics n.d., *Presentation of results*, truy cập ngày 31 tháng 3 năm 2022,
<https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Language+-+Presentation+of+Results>.

Trình bày Thống kê

Cách thức trình bày kết quả sẽ phụ thuộc vào dữ liệu và các loại thống kê mô tả (descriptive statistics) được yêu cầu. Bảng biểu và biểu đồ là các hình thức trình bày phổ biến nhất, nhưng cũng có những hình thức khác. Nhìn chung, bảng biểu mang tính chi tiết hơn – hiển thị các giá trị thực tế, trong khi biểu đồ hữu ích hơn trong việc hiển thị các mối quan hệ – tập trung vào hình thái, hình dạng và sự chuyển động của dữ liệu. Biểu đồ đặc biệt hữu ích trong việc trình bày sự thay đổi giá trị của một hạng mục dữ liệu trong một khoảng thời gian. Thông thường, chỉ cần một biểu đồ hoặc một bảng biểu là đủ để đại diện cho một tập dữ liệu.

Bảng biểu (Tables)

Bảng biểu là hình thức trình bày thống kê phổ biến nhất. Bảng biểu trình bày các thông tin bổ sung mà văn bản thông thường không thể hiển thị được. Một bảng biểu tốt là bảng biểu mà trong đó các quy luật và các trường hợp ngoại lệ nổi bật hẳn lên khi nhìn vào, theo sau là một đoạn văn ngắn bình luận về bảng đó. Bảng biểu thường chỉ được sử dụng để trình bày một vài giá trị, vì chúng sẽ trở nên khó hiểu nếu bao gồm quá nhiều con số.

Bảng biểu nên có tiêu đề rõ ràng. Một số hướng dẫn hữu ích khác bao gồm:

- Làm tròn dữ liệu đến 4 hoặc 5 chữ số có ý nghĩa (significant figures) để giúp dữ liệu dễ nhìn và dễ xử lý hơn;
- Đọc theo cột (từ trên xuống) dễ dàng hơn đọc theo hàng (ngang), đặc biệt là đối với số lượng hạng mục lớn;
- Các giá trị trung bình hoặc tỷ lệ phần trăm của hàng và cột có thể giúp người đọc diễn giải dữ liệu – nên tránh các cột có khoảng cách quá rộng vì chúng khó so sánh; và
- Việc tính tổng các hàng và/hoặc cột thường rất hữu ích.

Biểu đồ (Graphs)

Chúng ta cũng có thể sử dụng một chuỗi các biểu đồ đơn giản để kể một câu chuyện về dữ liệu và đưa ra cái nhìn sâu sắc về các phát hiện. Điều này có thể được hỗ trợ bởi một đoạn văn giải thích.

Biểu đồ là một cách trình bày dữ liệu hấp dẫn về mặt thị giác. Chúng duy trì sự quan tâm của người đọc và khuyến khích họ suy nghĩ về dữ liệu. Mặc dù lượng thông tin có thể trình bày là hữu hạn, biểu đồ thường tiết lộ hoặc làm nổi bật những sự thật "ẩn giấu" trong dữ liệu phức tạp. Các xu hướng và mối quan hệ được nắm bắt rõ ràng hơn từ biểu đồ so với bảng biểu hay văn bản, và do đó người đọc sẽ nhớ lâu hơn. Tuy nhiên, biểu đồ không phải lúc nào cũng cần thiết và chỉ nên được sử dụng khi phù hợp. Hãy lưu ý rằng người đọc có thể không quen thuộc với dữ liệu, do đó bạn cần đưa vào biểu đồ đầy đủ thông tin để giải thích dữ liệu đó.

Xây dựng Biểu đồ

- **Tiêu đề (Titles):** Tiêu đề là thiết yếu và tốt nhất nên đặt ở trên cùng. Nó nên chỉ ra "cái gì, ở đâu và khi nào" của biểu đồ một cách ngắn gọn nhất có thể và nên có kích thước chữ lớn hơn phần chữ trên chính biểu đồ đó.
- **Tỷ lệ/Thang đo (Scales):** Trục ngang thường đo đơn vị thời gian (trong trường hợp biểu đồ biểu diễn theo thời gian), và trục dọc đo biến số đang được xem xét. Nếu có thể, trục dọc nên bắt đầu từ số 0.
- Có thể thêm **đổ bóng và màu sắc** để tạo điểm nhấn thị giác;
- Cả hai trục nên được đánh dấu và đặt tên cùng với thang đo và đơn vị;
- Tất cả các tham số và biến số nên được định nghĩa ngắn gọn nhất có thể;
- Một **chú giải (legend)** có thể cần thiết, đặc biệt đối với các biểu đồ chi tiết có đổ bóng và/hoặc màu sắc;
- Nên sử dụng **nhãn (labels)** nếu có nhiều hơn một biến số được vẽ trên cùng một biểu đồ; và

- **Chú thích (footnotes)** có thể được sử dụng để giải thích các đặc điểm bất thường, chẳng hạn như sự đứt đoạn trong chuỗi dữ liệu.

Các loại Biểu đồ

Việc quyết định sử dụng loại biểu đồ nào phụ thuộc vào loại dữ liệu đang được trình bày. Dưới đây là một số loại biểu đồ có thể được sử dụng:

- **Biểu đồ Đường hoặc Đồ thị Đường (Line or Curve Graphs)** Những biểu đồ này được sử dụng khi cần nhấn mạnh vào sự chuyển động hơn là kích thước của hạng mục dữ liệu, hoặc khi so sánh nhiều chuỗi dữ liệu với nhau. Các loại biểu đồ này hiển thị sự biến thiên của dữ liệu được vẽ theo một khoảng thời gian.
- **Biểu đồ Thanh hoặc Biểu đồ Cột (Bar or Column Graph)** Loại biểu đồ này mô tả các giá trị số trên một biến số nhất định. Giá trị được thể hiện bằng chiều cao của cột. Loại biểu đồ này đặc biệt hiệu quả trong việc hiển thị những thay đổi lớn từ giai đoạn này sang giai đoạn tiếp theo.
- **Biểu đồ Cột/Thanh Nhóm (Grouped Columns/Bars)** Các biểu đồ này so sánh các danh mục khác nhau trên cùng một biểu đồ. Các nhóm cột, với mỗi cột trong nhóm đại diện cho một danh mục khác nhau, được vẽ trên cùng một trục.
- **Biểu đồ Tròn (Pie Chart)** Những biểu đồ này được sử dụng để so sánh các thành phần (theo tỷ lệ phần trăm). Các thành phần có thể được so sánh với nhau và so với đóng góp của chúng vào tổng thể. Những biểu đồ này đạt hiệu quả tốt nhất khi có số lượng danh mục ít.
- **Bản đồ (Map)** Bản đồ được sử dụng để hiển thị ranh giới của các khu vực. Các cách đồ bóng khác nhau cho các vùng khác nhau có thể cho thấy giá trị của biến số thay đổi như thế nào giữa các địa điểm.

- **Các loại Biểu đồ khác** Một số biểu đồ khác có sẵn bao gồm: Biểu đồ Chỉ số (Index), Biểu đồ Hình ảnh (Pictogram), Biểu đồ Bề mặt (Surface Chart), Đường cong Tích lũy (Cumulative Curve), Đường Sai lệch (Deviation Line), Biểu đồ Thanh trượt (Sliding Bar Chart), Biểu đồ Điểm (Dot Chart), Tháp Dân số (Population Pyramid), Bản đồ Biểu đồ (Map Chart), Biểu đồ Thanh Chồng (Stacked Bar) và Biểu đồ Phân tán (Scatter Diagram).

Các hình thức trình bày khác

Thuyết trình trực tiếp (Oral Presentations)

Tùy thuộc vào hoàn cảnh (người dùng, loại dữ liệu, kết quả), một bản báo cáo bằng văn bản có thể là chưa đủ hoặc cần được bổ sung thêm. Việc thuyết trình trực tiếp các kết quả khảo sát thường bị xem nhẹ dù đây là một phương tiện truyền tải thông tin quan trọng. Trong khi báo cáo văn bản cung cấp chi tiết cụ thể với phạm vi kết quả rộng, thì một buổi thuyết trình trực tiếp chỉ có thể nhấn mạnh vào một vài điểm chính. Tuy nhiên, điều này thường lại phù hợp nhất tùy thuộc vào đối tượng khán giả. Tương tự như báo cáo văn bản, việc thuyết trình kém có thể khiến các kết quả khảo sát bị bác bỏ.

Lời nói và các phương tiện hỗ trợ hình ảnh có thể có tác động lớn đến khán giả. Do đó, người thuyết trình nên nhận thức được bài thuyết trình của mình hướng đến ai và nắm rõ những kết quả khảo sát nào có thể đi ngược lại với các quan điểm hiện hữu.

Áp phích (Posters)

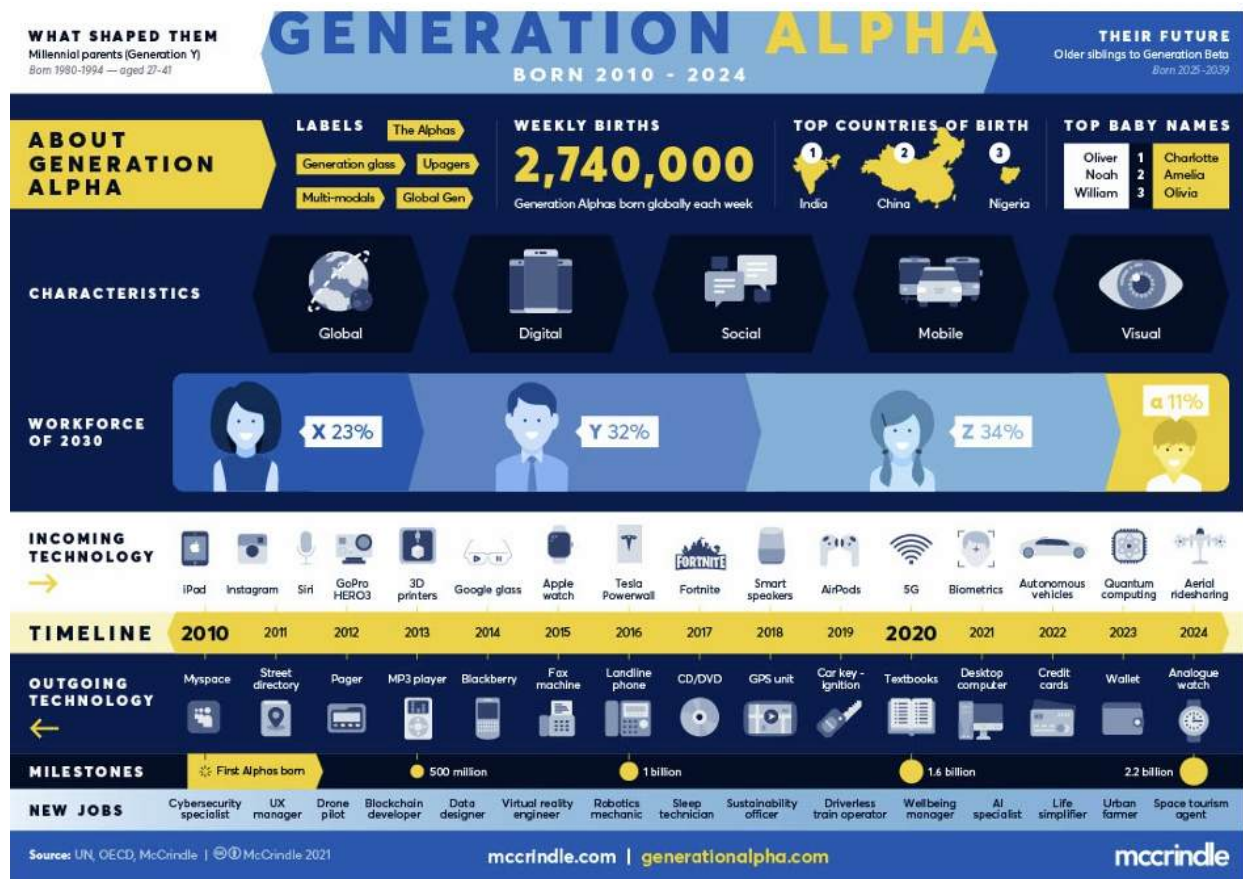
Áp phích là một cách để thu hút sự chú ý, nhưng chỉ nên đưa ra một tuyên bố trực tiếp duy nhất. Thông điệp phải dễ nhận thấy chỉ trong nháy mắt và bản thân áp phích phải hấp dẫn để khuyến khích những người dùng tiềm năng xem xét nó.

Triển lãm bảng thông tin (Panel Exhibits)

Triển lãm bảng thông tin là một dạng mở rộng của việc trình bày bằng áp phích. Loại hình trình bày này cung cấp nhiều chi tiết hơn và mở rộng một vài ý tưởng chính. Một lần nữa, điều quan trọng là các bảng thông tin phải đầy màu sắc và hấp dẫn.

Đồ họa thông tin (Charts - thường được gọi là Infographic)

Một biểu đồ tổng hợp có thể phù hợp tùy thuộc vào dữ liệu. Đây là một bản đồ phác thảo có thể chứa ảnh chụp, chữ viết, số liệu, biểu đồ, sơ đồ, bản đồ, biểu tượng, hình minh họa, v.v. (ví dụ: một biểu đồ thời tiết).



Đồ họa thông tin (Infographic) "Thế hệ Alpha" (Generation Alpha) do McCrindle 2021 tạo ra được cấp phép theo Giấy phép Quốc tế Creative Commons Ghi công 4.0 [2].

Video

Việc sử dụng video và truyền hình có thể cung cấp thêm một phương tiện truyền thông bổ sung cho các kết quả khảo sát.

Kết luận

Việc trình bày dữ liệu tốt là vô cùng quan trọng. Nếu chúng ta đã bỏ ra công sức, thời gian và chi phí để thực hiện một nghiên cứu, thì chúng ta phải đảm bảo rằng mình trình bày dữ liệu đó một cách rõ ràng và logic. Chúng ta phải cân nhắc việc sử dụng văn bản, bảng biểu và sơ đồ để xây dựng một tài liệu hữu ích và dễ hiểu. Mục đích cuối cùng của bất kỳ bài trình bày nào là thông báo cho người dùng về kết quả nghiên cứu theo cách có thể được sử dụng để thúc đẩy việc ra quyết định dựa trên thông tin đầy đủ.

Nguồn trích dẫn:

2. McCrindle 2021, *Generation alpha*, truy cập ngày 28 tháng 2 năm 2022, <https://mccrindle.com.au/resource/infographic-generation-alpha/>.

24. ĐỒ HỌA THÔNG TIN (INFOGRAPHICS)

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Hiểu được tầm quan trọng của việc sử dụng đồ họa thông tin;
- Nắm rõ các yếu tố then chốt để tạo nên một đồ họa thông tin hiệu quả.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=92#oembed-1>

Nguồn: Visme [1]

Đồ họa thông tin là gì?

Đồ họa thông tin (Infographics) truyền tải dữ liệu và thông tin phức tạp một cách trực quan dưới định dạng đồ họa dễ đọc và dễ hiểu. Chúng thường kết hợp các hình minh họa, biểu tượng (icons) và trực quan hóa dữ liệu để thu hút người đọc. Chúng được sử dụng rất hiệu quả trong môi trường trực tuyến để truyền đạt thông tin trong khi sử dụng tối thiểu văn bản.

Đồ họa thông tin là một công cụ quan trọng vì những lý do được giải thích dưới đây:

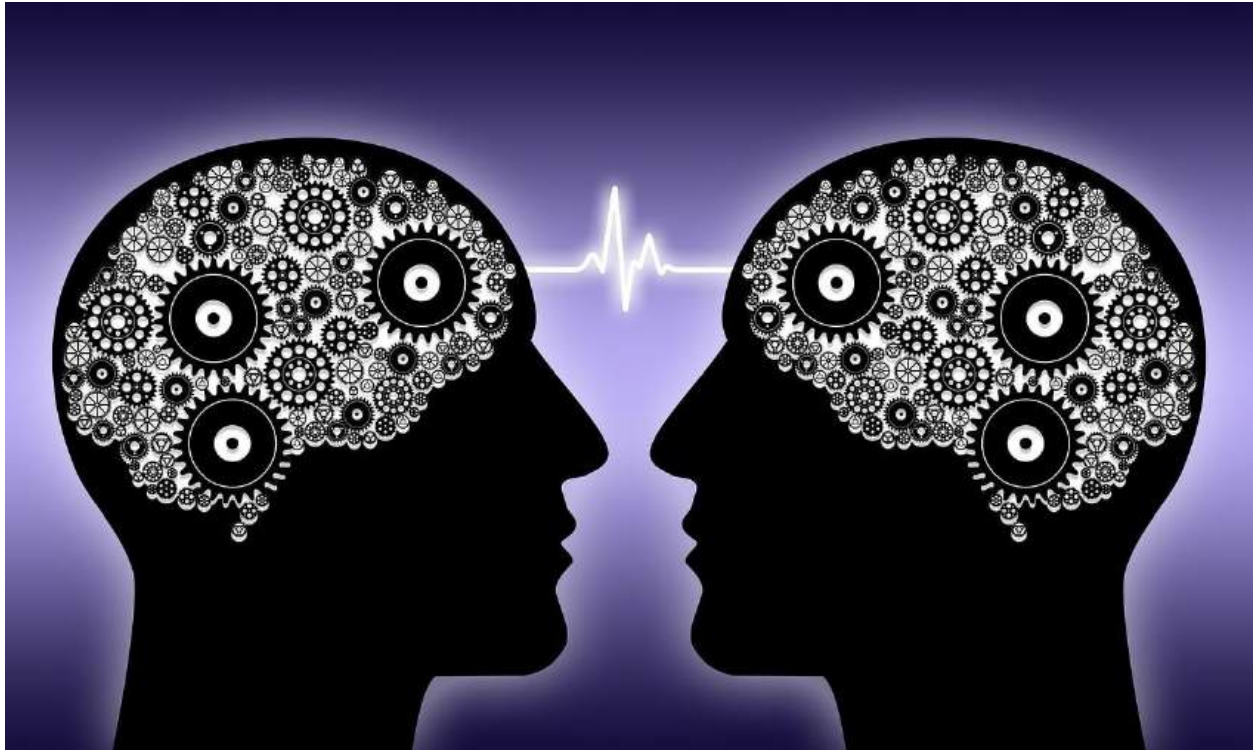
Xử lý Thông tin



Theo lý thuyết về **hiệu ứng ưu thế hình ảnh** (picture superiority effect), con người có khả năng giải mã thông tin hình ảnh nhanh hơn **60.000 lần** so với văn bản. Bộ não

con người được cấu tạo theo cách giúp việc xử lý thông tin hình ảnh trở nên dễ dàng hơn. Kết quả là, hình ảnh có khả năng được ghi nhớ tốt hơn so với ngôn từ.

Quá tải thông tin (Information Overload)



Trong kỷ nguyên quá tải thông tin ngày nay, thông điệp của một nhà tiếp thị cần phải nổi bật giữa bối cảnh cạnh tranh. Nó cần phải dễ nhớ. Thế giới ngoài kia là một "khu rừng thông tin", và việc bắt gặp một hình ảnh đồ họa dễ hiểu là một sự nhẹ nhõm (như đã được kể lại trong một bài diễn thuyết TedTalk của David McCandless).

Nguồn trích dẫn:

1. Visme 2019, *The 14 dos and don'ts to design beautiful and effective infographics*, ngày 8 tháng 10, video trực tuyến, truy cập ngày 30 tháng 5 năm 2022, https://www.youtube.com/watch?v=n8VtcviJ_Yc.

Giao tiếp với Thế hệ Bản địa Kỹ thuật số (Digital Natives)



Thế hệ Millennials và các nhóm độ tuổi tương tự đã lớn lên cùng với công nghệ. Họ ưu tiên cách thức giao tiếp bao gồm các thông điệp ngắn gọn và tức thời. Họ chuyển đổi sự chú ý sang các nguồn thông tin khác rất nhanh chóng. Những bài thuyết giảng dài dòng hay những văn bản viết đồ sộ có thể sẽ không hiệu quả với một thế hệ có mức độ chịu đựng sự nhàm chán rất thấp.

Dễ dàng chia sẻ (Easily Shareable)



Đồ họa thông tin (Infographics) dễ dàng được chia sẻ trên mạng xã hội và các nền tảng khác hơn. Có nhiều khả năng là thông điệp then chốt của bạn trong một đồ họa thông tin sẽ tiếp cận được lượng khán giả lớn hơn so với một bài báo dài dòng.

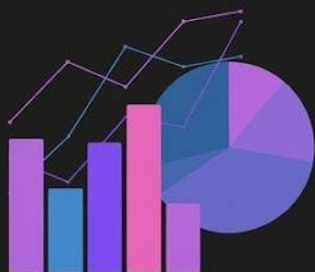
Các yếu tố của một Đồ họa thông tin (Infographic)

Tài liệu sau đây được phỏng theo Đại học Curtin [2] và được sử dụng theo giấy phép Creative Commons Ghi công - Chia sẻ tương tự 4.0 (Creative Commons Attribution ShareAlike 4.0 Licence).

Nguồn trích dẫn:

2. Curtin University 2020, *Visual communication*, truy cập ngày 11 tháng 5 năm 2022, <https://libguides.library.curtin.edu.au/visual-communication/infographics>.

Elements of Infographics



Data Visualisations

Use as an accessible way to show people patterns and trends in data.

Images and Symbols

Images should be engaging and grab the attention of the viewers. They should also relate to the topic of the infographic.



Small Amounts of Text

Try not to overwhelm the viewer with too much text. Instead try to make your point with as few words as possible.

Colour

Colour can be used to draw attention to particular sections of the infographic as well as draw attention to the infographic as a whole.



Tạo lập một Đồ họa thông tin (Infographic)

Các bước sau đây được thực hiện trong quá trình tạo lập một đồ họa thông tin:

Bước 1 – Chuẩn bị (Preparation): Mục tiêu chính đằng sau việc tạo đồ họa thông tin là gì? Thông điệp then chốt cần được truyền tải là gì?

Bước 2 – Lập kế hoạch (Planning): Bước này bao gồm việc chuẩn bị một bản phác thảo cho đồ họa thông tin. Quyết định nội dung chính nào nên được đưa vào. Nội dung này phải giải quyết được các mục đích/mục tiêu đã được nêu rõ ở Bước 1.

Bước 3 – Phát triển (Development): Sử dụng phần mềm phù hợp nhất để xây dựng đồ họa thông tin.

Bước 4 – Đánh giá (Evaluation): Một số tiêu chí chính được thảo luận trong phần bên dưới. Sẽ rất hữu ích nếu thực hiện đối sánh (benchmark) đồ họa thông tin của bạn với một đồ họa khác trong cùng ngành.

Bước 5 – Xuất bản/Phổ biến (Publication/Dissemination): Một đồ họa thông tin chỉ hữu ích nếu nó tiếp cận được đối tượng mục tiêu. Việc đảm bảo rằng nó có thể được chia sẻ dễ dàng cũng giúp ích trong việc phổ biến thông điệp.

Đánh giá một Đồ họa thông tin:

Tài liệu sau đây được phỏng theo Đại học Guelph 2022 [3] và được sử dụng theo giấy phép Creative Commons Ghi công - Chia sẻ tương tự 4.0 ([Creative Commons Attribution ShareAlike 4.0 Licence](https://creativecommons.org/licenses/by-sa/4.0/)).

Khi đánh giá một đồ họa thông tin, hãy sử dụng các tiêu chí sau:

- **Tiêu chí 1:** Nó có kể một câu chuyện hay không?
 - **Tiêu chí 2:** Nó có sử dụng dữ liệu/thông tin tốt và hữu ích không?
 - **Tiêu chí 3:** Phong cách hình ảnh giúp ích hay gây cản trở việc hiểu câu chuyện?
-

Nguồn trích dẫn:

3. University of Guelph 2022, *Infographics*, truy cập ngày 11 tháng 5 năm 2022, <https://guides.lib.uoguelph.ca/Infographics>.

25. QUY TRÌNH: TỪ TẠO LẬP ĐẾN SỬ DỤNG THÔNG TIN CHI TIẾT VỀ KHÁCH HÀNG (CUSTOMER INSIGHTS)

Mục tiêu học tập

Sau khi hoàn thành chương này, sinh viên phải có khả năng:

- Xác định các bước trong việc tạo lập và sử dụng thông tin chi tiết về khách hàng (customer insights);
- Phân biệt giữa các kết quả nghiên cứu truyền thống và khái niệm về thông tin chi tiết có thể hành động (actionable insights).

Quy trình nghiên cứu tiếp thị truyền thống dựa trên bước đầu tiên là "xác định vấn đề" hoặc nêu rõ câu hỏi nghiên cứu. Trong một số ít trường hợp hiếm hoi, các học giả cũng nhấn mạnh việc thực hiện phân tích chi phí - lợi ích để xem liệu có thực sự cần thiết phải thực hiện nghiên cứu tiếp thị tốn kém hay không.

Ngược lại, các nhà quản lý tin tưởng vào khái niệm thông tin chi tiết về khách hàng (customer insights) khẳng định rằng việc khám phá dữ liệu có sẵn trên thực tế sẽ giúp kích hoạt các câu hỏi và những vấn đề liên quan [1]. Cần phải phân tích và diễn giải thêm dữ liệu từ nhiều nguồn khác nhau để nêu rõ mục tiêu trong việc khai thác dữ liệu. Chỉ khi các hành động có ý nghĩa được thực hiện dựa trên kết quả của một thông tin chi tiết mới xuất hiện, thì các vấn đề mới thực sự được giải quyết.

Chương này trình bày quy trình 8 bước trong việc phát triển một chiến lược xoay quanh các thông tin chi tiết về khách hàng. Quy trình này bao gồm việc lập kế hoạch

để tạo lập và sau đó sử dụng các thông tin chi tiết về khách hàng nhằm đáp ứng các mục tiêu của tổ chức. Các bước này đã được tổng hợp sau khi xem xét kinh nghiệm của những người thực hành và các đề xuất của giới học thuật. Mỗi bước được thảo luận chi tiết dưới đây.

Bước 1: Công nhận và cam kết với khái niệm "Thông tin chi tiết về khách hàng" (Customer Insights)

Giống như nhiều chiến lược tổ chức khác, cơ hội thành công của một chương trình thông tin chi tiết về khách hàng phụ thuộc vào ban quản lý. Nếu ban quản lý cấp cao nhận ra giá trị của việc phát triển chương trình này — và sau đó cam kết dành nguồn lực cho nó — sáng kiến mới có thể được khởi động. Theo truyền thống, các nhà quản lý thường quen thuộc hơn với khái niệm nghiên cứu tiếp thị. Tuy nhiên, nếu các nhà lãnh đạo doanh nghiệp dành thời gian (và năng lượng) để tìm hiểu về lợi ích của phương pháp tiếp cận thông tin chi tiết, và hướng dẫn tổ chức coi dữ liệu — đặc biệt là dữ liệu sẵn có — như một tài sản quý giá, thì bước đầu tiên trong việc triển khai chương trình thông tin chi tiết về khách hàng đã được thực hiện. Quan trọng hơn, bộ phận thông tin chi tiết không phải là một đơn vị hoạt động độc lập trong tổ chức. Trên thực tế, các công ty có hiệu suất cao sẽ đảm bảo rằng những người lãnh đạo từ nhóm thông tin chi tiết tham gia vào tất cả các giai đoạn lập kế hoạch chiến lược của công ty.

Nguồn trích dẫn:

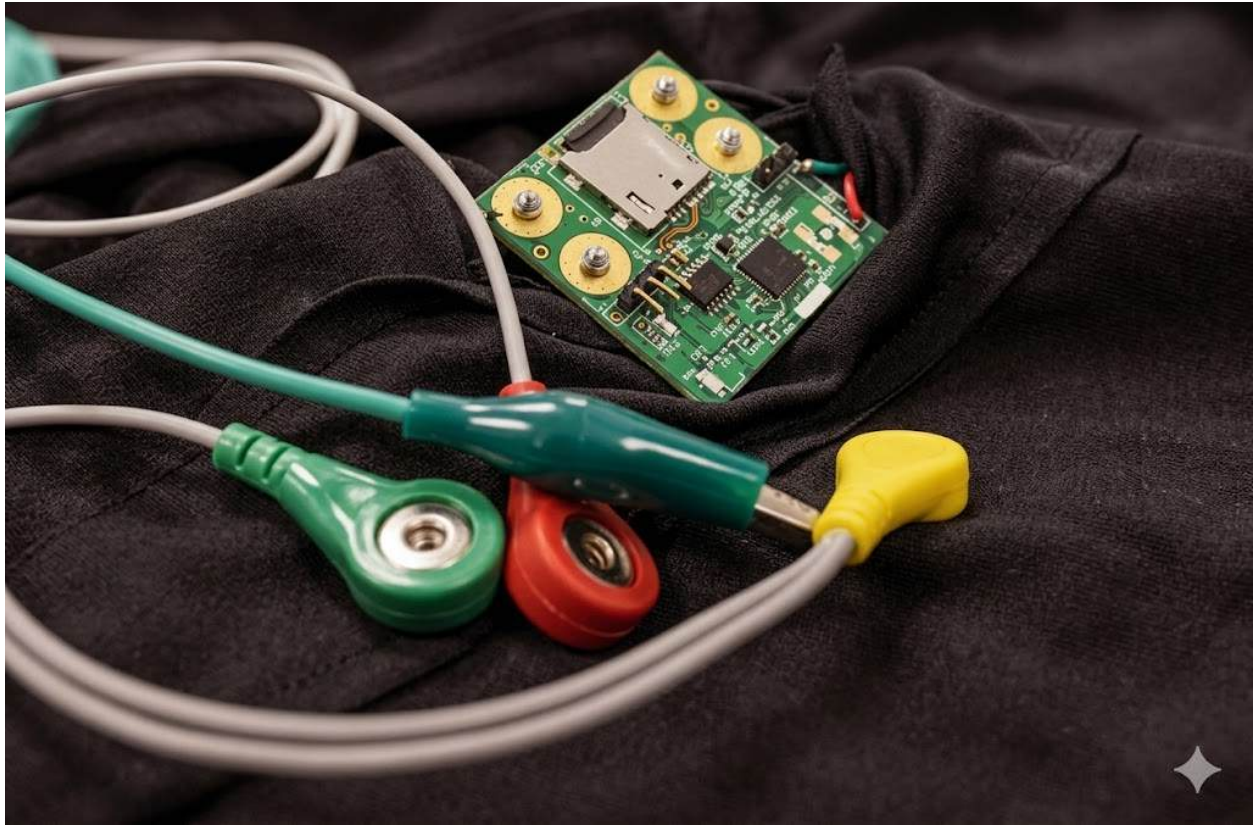
1. Rogers, C 2017, *Marketing insight teaming up to drive better business*, Marketing Week, truy cập ngày 21 tháng 3 năm 2022, <https://www.marketingweek.com/marketing-insight-teaming-up-to-drive-better-business/>.

Bước 2: Xác định các "nguồn dữ liệu" khác nhau – bên trong và bên ngoài doanh nghiệp

Một tổ chức tạo ra dữ liệu liên tục — tuy nhiên, các nhà quản lý không phải lúc nào cũng nhận thức được tiện ích của chúng. Một hồ sơ được tạo ra mỗi khi có giao dịch được thực hiện hoặc nhận được thắc mắc của khách hàng. Nếu một quản lý cấp cơ sở muốn đánh giá hiệu quả hoạt động dịch vụ khách hàng của nhóm mình, họ hoàn toàn có thể khai thác các nguồn dữ liệu khác nhau đang được tạo ra và lưu trữ nội bộ. Một cái nhìn 360° có thể cho thấy dữ liệu từ camera/cảm biến tại cửa hàng, các truy vấn trực tuyến, các cuộc gọi điện thoại và việc trả lại sản phẩm đều hữu ích. Hơn nữa, đội ngũ nhân viên tuyến đầu của công ty — những người trực tiếp tương tác với khách hàng — có thể bổ sung dữ liệu bằng những trải nghiệm thực tế của chính họ [2].

Nguồn trích dẫn:

2. Barton, D & Court, D 2012, 'Making advanced analytics work for you', *Harvard Business Review*, tập 90, số 10, tr. 78-83, truy cập ngày 22 tháng 3 năm 2022, <https://hbr.org/2012/10/making-advanced-analytics-work-for-you>.



Thiết bị công nghệ đeo (Wearables technology) – Thiết bị theo dõi chứng ngưng thở khi ngủ.

Ảnh: Sally Tsoutas © 2022 Đại học Western Sydney, được chụp bởi Nhiếp ảnh gia [Sally Tsoutas của Đại học Western Sydney](#), được cấp phép theo giấy phép Ghi công - Phi thương mại - Không phái sinh 4.0 Quốc tế ([Attribution-NonCommercial-NoDerivatives 4.0 International](#)).

Tương tự như vậy, các nhà quản lý có thể tư duy sáng tạo về các nguồn thông tin sẵn có bên ngoài. Các công nghệ đeo (wearable technologies) có thể ghi lại video người tiêu dùng sử dụng sản phẩm trong thời gian thực (tham khảo: [Unilever's path to innovative insights](#)). Hơn nữa, dữ liệu từ mạng xã hội mang lại kho thông tin phong phú dưới định dạng văn bản, video và hình ảnh. Những thông tin này dựa trên các cuộc trò chuyện tự nhiên giữa các chủ tài khoản, và do đó được xem là "đáng tin cậy" hơn so với các câu trả lời mang tính đối phó trong một cuộc khảo sát chính thức.

Ngoài ra, thông tin từ Cục Thống kê Úc (ABS), Cục Khí tượng (BoM), và các báo cáo định kỳ từ các công ty hàng đầu như ANZ-Roy Morgan (ví dụ: Chỉ số Niềm tin Người tiêu dùng - Consumer Confidence Index) có thể hỗ trợ các nhà quản lý doanh nghiệp trong việc phân tích tình hình.

Một lời cảnh báo: Các nhà quản lý có thể kết thúc với những chồng dữ liệu chất đống trên bàn làm việc mà không có ích lợi gì. Việc một chuyên gia thống kê thực hiện hàng tá kiểm định giả thuyết và đưa ra kết quả cho các nhà quản lý "xem" cũng không giúp ích được gì nhiều. Bí quyết nằm ở việc thiết kế một chiến lược thu thập - đối soát dữ liệu, trong đó các thông tin quan trọng có thể được sàng lọc — từ hàng tấn tài liệu — để cung cấp cho nhân viên và cuối cùng được sử dụng để triển khai các chiến lược cụ thể.

Bước 3: Xây dựng một "Nhóm Thông tin chi tiết" (Insights Team) với các kỹ năng và năng lực khác nhau [3]

Một nhóm thông tin chi tiết về khách hàng phải bao gồm các thành viên với những kỹ năng và năng lực khác nhau. Ví dụ: một giám đốc sản phẩm, một nhà phân tích kinh doanh, một chuyên gia thống kê, một nhân viên IT, một kiến trúc sư/nhà khoa học dữ liệu và một nhà thiết kế đồ họa có thể tạo thành một nhóm làm việc cho một dự án cụ thể.

Cách tiếp cận như vậy để thực hiện một hoạt động "nghiên cứu" có thể đòi hỏi sự thay đổi về văn hóa và tư duy. Đây có thể là một vùng lãnh thổ mới đối với nhiều người theo trường phái truyền thống. Việc giới thiệu nhân viên tiếp cận kỹ thuật đổi mới này có thể được thực hiện bằng cách tiếp cận đa diện — đào tạo, các nhà lãnh đạo làm gương, các chế độ đãi ngộ và các chỉ số đo lường để củng cố hành vi.

Xây dựng đúng đội ngũ là một bước quan trọng trong việc thiết kế chiến lược thông tin chi tiết về khách hàng. Dữ liệu lớn (Big data), thông tin và công nghệ là những thứ

sẽ tồn tại lâu dài. Thay vì thực hiện một cuộc cải tổ lớn, các tổ chức được gợi ý nên tiếp tục xây dựng văn hóa hướng tới sự linh hoạt và nhanh nhẹn. Khi có thêm nhiều nhân viên được đào tạo và tiếp cận với lợi ích của việc sử dụng dữ liệu hiện có, các tổ chức có thể xây dựng những năng lực vượt trội, điều này sẽ chứng minh là một tài sản cạnh tranh mang tính quyết định.

Bước 4: Sử dụng các công cụ hiện có / phát triển các công cụ riêng để quản lý dữ liệu

Với lượng thông tin kỹ thuật số khổng lồ hiện nay, các công ty cần những công cụ đáng tin cậy để đảm bảo rằng các nhà quản lý có thể tiếp cận được thông tin liên quan. Có một số công cụ miễn phí hoặc có giá cả hợp lý mà ngay cả các doanh nghiệp nhỏ cũng có thể dễ dàng áp dụng. Chúng bao gồm YouTube Analytics, Google Analytics, Facebook Audience Insights, Twitter Analytics và các nền tảng khác để đo lường hoạt động trên mạng xã hội. Tất nhiên, còn rất nhiều phần mềm và nền tảng khác có sẵn để các nhà quản lý lựa chọn. Quyết định cuối cùng sẽ phụ thuộc vào nhu cầu của doanh nghiệp cũng như chi phí liên quan đến việc sử dụng các công cụ này.

Tùy thuộc vào năng lực CNTT của công ty, việc thiết kế một nền tảng dữ liệu riêng có thể là khả thi. Mục tiêu của một nền tảng như vậy — được thiết kế và xây dựng nội bộ — là trao quyền cho từng nhân viên truy cập dữ liệu, tự phân tích và sau đó sử dụng dữ liệu này để đưa ra quyết định. Một nền tảng như vậy phải dễ dàng truy cập và sử dụng. Nếu có quá nhiều rào cản (chẳng hạn như quyền truy cập kỹ thuật số hoặc yêu cầu bộ kỹ năng trình độ cao), thì khả năng sử dụng của nền tảng đó sẽ giảm đi. Một nền tảng quản lý dữ liệu nên thân thiện với người dùng (đối với nhân viên) cũng như lấy người dùng làm trung tâm.

Nguồn trích dẫn: Stobierski, P 2021, *How to structure your data analytics team*, Harvard Business School Online, truy cập ngày 22 tháng 3 năm 2022, <https://online.hbs.edu/blog/post/how-to-structure-a-data-analytics-team>.

Bước 5: Diễn giải dữ liệu – Nhà phân tích dữ liệu hay là...?



Phân tích dữ liệu là việc quá quan trọng để có thể phó mặc cho một người – hay một bộ phận duy nhất. Ảnh: Sally Tsoutas © 2022 Đại học Western Sydney, được chụp bởi Nhiếp ảnh gia [Sally Tsoutas của Đại học Western Sydney](#), được cấp phép theo giấy phép Ghi công - Phi thương mại - Không phái sinh 4.0 Quốc tế ([Attribution-NonCommercial-NoDerivatives 4.0 International](#)).

Trong khi các doanh nghiệp thường xuyên tìm kiếm những nhà phân tích dữ liệu lành nghề, có một khái niệm khác đang chứng minh được tính hữu dụng của nó. Phân tích dữ liệu là việc quá quan trọng để có thể phó mặc cho chỉ một người – hay một nhóm người có kỹ năng kỹ thuật. Mọi thành viên trong nhóm đều phải được đào tạo đầy đủ để có thể hiểu được dữ liệu. Các giám đốc điều hành doanh nghiệp đồng ý rằng việc thiếu kỹ năng phân tích dữ liệu ngăn cản họ đưa ra các quyết định then chốt dựa trên

phân tích dữ liệu. Đó là lý do tại sao việc thực hiện kiểm tra (audit) mức độ kỹ năng cơ bản của các thành viên hiện tại trong nhóm là rất quan trọng, từ đó cung cấp cho họ sự đào tạo và nguồn lực để họ có thể tự tin phân tích các tập dữ liệu mà không cần phải dựa dẫm vào việc ai đó tính toán con số cho mình [4].

Bước 6: Xác định các thông tin chi tiết (insights) có ý nghĩa

Mặc dù vẫn có một số người khẳng định rằng các kết quả nghiên cứu (findings) và thông tin chi tiết (insights) là một, nhưng ngày càng có nhiều người thực hành làm nổi bật sự khác biệt giữa hai khái niệm này. Bảng dưới đây nhấn mạnh những điểm khác biệt giữa kết quả nghiên cứu và thông tin chi tiết.

Bảng: Sự khác biệt giữa "Kết quả nghiên cứu" và "Thông tin chi tiết về khách hàng" [5]

Nguồn trích dẫn:

4. Think with Google 2017, *How to make everyone on your team a data-savvy marketer*, truy cập ngày 22 tháng 3 năm 2022, <https://www.thinkwithgoogle.com/intl/en-gb/future-of-marketing/creativity/how-make-everyone-your-team-data-savvy-marketer/>.
5. Doing Data n.d., *Data, information or insight?*, truy cập ngày 22 tháng 3 năm 2022, <https://doingdata.org/blog/data-information-or-insight>.

Bảng: Sự khác biệt giữa ‘Kết quả nghiên cứu’ và ‘Thông tin chi tiết khách hàng’

Đặc điểm	Kết quả nghiên cứu (Research Findings)	Thông tin chi tiết khách hàng (Customer Insights)
Định nghĩa	“Một mẫu thông tin được phát hiện trong quá trình kiểm tra chính thức về một vấn đề hoặc tình huống” <i>(Từ điển Cambridge)</i>	“Một sự thấu hiểu rõ ràng, sâu sắc và đôi khi là bất ngờ về một vấn đề hoặc tình huống phức tạp” <i>(Từ điển Cambridge)</i>
Nguồn dữ liệu	Thường dựa trên một nguồn dữ liệu duy nhất, chẳng hạn như một cuộc khảo sát.	Dựa trên nhiều nguồn dữ liệu, ví dụ: hồ sơ nội bộ, mạng xã hội, v.v.
Bản chất	Những sự thật khách quan (Hard facts).	Một câu chuyện có ý nghĩa dựa trên sự thật.
Khả năng quan sát	Hiển thị rõ ràng (ví dụ: hệ số tương quan là \$0,7\$).	Không quá lộ liễu; đòi hỏi sự khám phá chuyên sâu; là những quy luật ẩn giấu trong dữ liệu.

Kết quả đầu ra	Mọi quá trình phân tích đều tạo ra "kết quả".	Không phải mọi quá trình tổng hợp/khám phá dữ liệu đều dẫn đến "thông tin chi tiết".
Trọng tâm	Tập trung vào phân tích dữ liệu và kiểm định giả thuyết.	Tập trung vào việc tích hợp cả tư duy sáng tạo và tư duy phân tích.
Giá trị sử dụng	Có thể thú vị nhưng vụn vặt (trivial).	Giàu thông tin, có khả năng thực thi (actionable).
Mức độ cần thiết	Biết thì tốt (Nice-to-know).	Bắt buộc phải biết (Need-to-know).
	Chủ yếu dễ hiểu đối với các chuyên gia có kiến thức về thống kê	Dễ hiểu đối với nhân viên

Ví dụ: [Unilever – Đưa tiếng nói của người tiêu dùng vào mọi hoạt động của doanh nghiệp.](#)

Vậy, một "thông tin chi tiết có thể hành động" (actionable insight) có nghĩa là gì? Dưới đây là một ví dụ: Vào năm 2018, Unilever đã khởi động chương trình [Tiếng nói Kỹ thuật số của Người tiêu dùng \(Digital Voice of the Consumer - DVOC\)](#) để hiểu rõ hơn "chất lượng" có ý nghĩa như thế nào đối với người tiêu dùng. DVOC giúp thu thập các đánh giá và phản hồi mà mọi người chia sẻ với các Trung tâm Tương tác Khách hàng của Unilever cũng như trên các nền tảng công cộng khác. Các bảng điều khiển (dashboards) của DVOC hợp nhất và xử lý dữ liệu từ mạng xã hội, tin tức, diễn đàn, trung tâm gọi khách hàng, các trang thương mại điện tử và ý kiến từ các chuyên gia (thông qua một ứng dụng dashboard).

Giá trị to lớn từ DVOC đã phát huy tác dụng trong đợt bùng phát virus corona trên toàn cầu. Unilever cần thực hiện nghiên cứu thị trường nhanh chóng trên toàn cầu để hiểu những khó khăn mới mà người tiêu dùng của họ phải đối mặt. Một phương pháp nghiên cứu thị trường truyền thống với các cuộc khảo sát và thảo luận nhóm tập trung (focus groups) sẽ mất rất nhiều thời gian, và có lẽ không mang lại thông tin chính xác.

Bằng cách phân tích kỹ lưỡng các cuộc trò chuyện trên mạng xã hội, nhóm thương hiệu nước rửa tay đã nhận thấy rằng người tiêu dùng đang bị bối rối bởi các mô tả khác nhau về nồng độ cồn trên nhãn sản phẩm. Nhóm thương hiệu đã có thể xác định mẫu thông tin này là một **"thông tin chi tiết có thể hành động" (actionable insight)**. Họ nhanh chóng đưa ra các hướng dẫn cho các nhóm hỗ trợ và giúp điều chỉnh thông điệp để làm rõ thông tin tốt hơn cho người tiêu dùng, nhằm giúp giảm thiểu bất kỳ sự nhầm lẫn nào về thương hiệu.

Bước 7: Truyền tải thông qua trực quan hóa dữ liệu (Data Visualisation) [6][7]

Trực quan hóa dữ liệu có nghĩa là vẽ các hiển thị đồ họa để trình bày dữ liệu. Trước đây, đây là một kỹ năng được coi là "biết thì tốt" (nice-to-know). Hiện nay, nó đã trở thành một kỹ năng "phải có" (must-have) vì việc giao tiếp với các đối tượng khán giả khác nhau đòi hỏi phải sử dụng các hình ảnh trực quan dễ hiểu.

Một hình ảnh trực quan, ở mức độ tốt nhất, là một sự giải thích bằng hình ảnh. Nó sử dụng hệ thống tri giác thị giác của não bộ để giúp chúng ta hiểu một tập dữ liệu lớn, một tập hợp các mối quan hệ hoặc các mối quan hệ không gian. Trực quan hóa dữ liệu cho phép chúng ta:

- Phân tích và thấu hiểu dữ liệu tốt hơn;
- Vượt qua những danh sách phẳng hai chiều nhàm chán;
- Khuyến khích việc khám phá sự phức tạp của dữ liệu theo những cách lôi cuốn, thú vị và thuyết phục;
- Minh họa những câu chuyện đằng sau dữ liệu.

Phần cứng tốt hơn đồng nghĩa với việc tái tạo hình ảnh chính xác hơn, màu sắc tốt hơn (bao gồm cả kỹ thuật pha trộn alpha - alpha-blending) và tốc độ vẽ nhanh hơn. Phần mềm tốt hơn đồng nghĩa với việc vẽ dễ dàng và linh hoạt hơn, các chủ đề nhất quán và tiêu chuẩn cao hơn. Các nhà khoa học máy tính đã tham gia sâu hơn nhiều, cả về mặt kỹ thuật lẫn việc giới thiệu các phương pháp tiếp cận mới.

Việc đưa trực quan hóa dữ liệu vào các bộ sưu tập của bạn có tiềm năng tăng khả năng tái sử dụng, khám phá và tính kết nối cho dữ liệu nghiên cứu của bạn. Một hình ảnh trực quan tốt nên:

- Hiển thị được dữ liệu;
- Khiến người xem suy nghĩ về nội dung cốt lõi thay vì về phương pháp luận;

- Tránh bóp méo những gì dữ liệu muốn nói;
- Trình bày được nhiều con số trong một không gian nhỏ;
- Làm cho các tập dữ liệu lớn trở nên mạch lạc;
- Tiết lộ dữ liệu ở nhiều mức độ chi tiết khác nhau, từ cái nhìn tổng quan rộng lớn đến cấu trúc tinh vi bên trong.

Bước 8: Kết hợp các thông tin chi tiết "có thể hành động" vào việc phát triển chiến lược

Nhiều người ủng hộ lĩnh vực "thông tin chi tiết về khách hàng" (customer insights) khẳng định rằng quy trình này phải dẫn đến sự xuất hiện của các thông tin chi tiết "có thể hành động" (actionable insights). **Thông tin chi tiết có thể hành động là những kết luận được rút ra từ dữ liệu mà có thể chuyển hóa trực tiếp thành một hành động hoặc một phản hồi.** Thông tin chi tiết có thể hành động là kết quả của quá trình phân tích sâu rộng dựa trên dữ liệu và thông tin, cho phép một tổ chức và nhân viên của mình đưa ra các quyết định sáng suốt (informed decisions).

Nguồn trích dẫn:

6. Unwin, A 2020, *Why is data visualization important? What is important in data visualisation?*, truy cập ngày 16 tháng 3 năm 2022, <https://hdr.mitpress.mit.edu/pub/9vsgp7z0/release/4>.
7. Australian National Data Service n.d., *Data visualisation*, truy cập ngày 16 tháng 3 năm 2022, <https://www.ands.org.au/working-with-data/data-visualisation>.

Phần lớn dữ liệu có thể không mang lại thông tin chi tiết hoặc chỉ là thông tin chung chung. Đây có thể là những thông tin đã được xác lập và công nhận rộng rãi. Tuy nhiên, chính những "mẫu vàng" giá trị ẩn giấu trong các tập dữ liệu — khi được khai thác, chia sẻ và thấu hiểu trong toàn tổ chức — có thể tiết lộ những "thông tin chi tiết" nhất định thúc đẩy hành động. Các thông tin chi tiết có thể hành động sẽ chuyển hóa thành các hoạt động cụ thể, dẫn đến sự sửa đổi và thực thi.

Nguồn: Digital Garage [8]

Tài liệu tham khảo:

Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard business review*, 90(10), 78-83.

<https://online.hbs.edu/blog/post/analytics-team-structure>

<https://www.thinkwithgoogle.com/intl/en-ca/future-of-marketing/digital-transformation/marketing-data-analysis-infographic/>

<https://www.warc.com/newsandopinion/news/unilevers-path-to-innovative-insights/36906>

<https://www-warc-com.ezproxy.uws.edu.au/content/article/event-reports/unilevers-eight-tips-for-method-innovation/107930>

<https://www.unilever.com/news/news-search/2020/how-your-product-reviews-are-shaping-our-brands>

<https://www-warc-com.ezproxy.uws.edu.au/content/article/unilever-putting-the-consumers-voice-through-out-the-business/135005>

Nguồn trích dẫn bổ sung:

8. Digital Garage 2013, *Breaking down your data for insights*, truy cập ngày 22 tháng 3 năm 2022, https://www.youtube.com/watch?v=0_u6YQ_n6-4.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=94#oembed-1>

26. NGHIÊN CỨU TÌNH HUỐNG (CASE STUDY):

SỬ DỤNG THÔNG TIN CHI TIẾT KHÁCH HÀNG ĐỂ TÁI ĐỊNH VỊ ĐẠI HỌC WESTERN SYDNEY

Tác giả: Jacki Montgomery

Giới thiệu

Đại học Western Sydney (Western Sydney University) được coi là một trường đại học "trẻ", được thành lập vào năm 1989. Trường nổi tiếng toàn cầu về sự xuất sắc trong giảng dạy và nghiên cứu [1]. Tuy nhiên, chỉ mới một thập kỷ trước, bản sắc địa phương của trường đã từng là một thách thức lớn. Khi đó, trường có tên là "University of Western Sydney" (Đại học Miền Tây Sydney), và các nghiên cứu cho thấy phụ huynh cũng như sinh viên tương lai không thể thực sự định nghĩa được trường đại học này đại diện cho điều gì, và họ cũng không coi đây là lựa chọn hàng đầu của mình.

Vào năm 2015, đội ngũ tiếp thị đã đặt mục tiêu thách thức những nhận thức này về trường và thực hiện một chiến dịch tái cấu trúc thương hiệu (rebranding) đầy táo bạo. Đó là một thành công chưa từng có – dựa trên nền tảng sâu sắc từ những thông tin chi tiết về khách hàng (consumer insights).

Xác định phạm vi của thách thức



Với tư cách là Giám đốc Tiếp thị tại Đại học Western Sydney, một trong những trách nhiệm chính của Angelo Kourtis là tăng số lượng sinh viên nhập học. Mặc dù tỉ lệ nhập học vẫn duy trì ở mức ổn định, ông tin rằng vẫn còn cơ hội đáng kể để gia tăng các con số này. Ông và các đồng nghiệp nhận thức được rằng nhiều sinh viên trong khu vực địa phương đã lựa chọn đi qua Đại học Western Sydney để đến học tại các trường đại học xa nhà hơn nhiều. Các chiến

dịch quảng cáo trước đây của Đại học Western Sydney thường quảng bá về chất lượng và vị trí của trường nhằm thu hút sinh viên. Thay vì tiếp tục phương pháp này, Kourtis tin rằng đội ngũ cần thực hiện nghiên cứu chuyên sâu và khám phá lý do tại sao sinh viên địa phương lại chọn nhập học tại các trường đại học khác.

Ban đầu, Kourtis cảm thấy định vị của trường đại học chưa rõ ràng; rằng nó thực sự không đại diện cho bất cứ điều gì có thể tạo ra sự cộng hưởng với khu vực hoặc rộng hơn là tại Sydney. Ông đã thuê một Công ty Nghiên cứu Thị trường, và họ đã tiến hành nghiên cứu định tính (phần nghiên cứu này sau đó đã cung cấp thông tin cho việc thiết kế nghiên cứu định lượng nhằm xác minh các phát hiện ban đầu). Nhóm đã tổ chức các buổi thảo luận nhóm tập trung (focus groups) với những người tham gia được coi là các bên liên quan quan trọng trong quá trình ra quyết định, và/hoặc những người đã từng trải qua quá trình chọn trường đại học: sinh viên tiềm năng, sinh viên hiện tại, thành viên cộng đồng, nhà tuyển dụng, giáo viên trung học và cố vấn hướng nghiệp. Nghiên cứu đã tạo ra hai thông tin chi tiết (insights) then chốt.

Insign1: Nhận thức về Đại học Western Sydney (so với các trường đại học khác)

Các nhà nghiên cứu đã yêu cầu những người tham gia phân nhóm một số trường đại học tại Sydney thành các phân khúc, sau đó yêu cầu họ mô tả các phân khúc đó. Phân khúc đầu tiên được gọi là **"Các trường đại học khát vọng" (Aspirational Universities)**; trong nhóm này, những người tham gia đã xếp các trường đại học lâu đời (trường "đá nâu" - brownstone) vốn thường yêu cầu điểm số nhập học cao, và mô tả chúng như những biểu tượng của địa vị.

Phân khúc thứ hai được gọi là **"Các trường đại học chuyên gia ngành" (Discipline Expert Universities)**, được người tham gia công nhận là những trường đại học nổi tiếng về thế mạnh trong các ngành học cụ thể – được xem là quan trọng nếu bạn muốn đảm bảo sự nghiệp trong một trong những lĩnh vực chuyên ngành đó.

Phân khúc thứ ba được gán nhãn là **"Các trường đại học trong tầm tay" (Achievable Universities)** – nghĩa là "các trường đại học lưới an toàn", hoặc là lựa chọn mà bạn đưa ra khi các trường đại học khác không chấp nhận đơn đăng ký của bạn. Đại học Western Sydney thường xuyên nhất bị xếp vào phân khúc thứ ba này.

Kourtis cảm thấy rằng điều này là không chính xác, và trong một thời gian quá dài, trường đại học đã để bản thân bị định nghĩa bởi những người khác. Điều này đã củng cố quyết tâm của ông trong việc thay đổi nhận thức của người tiêu dùng và đưa trường đại học ra khỏi phân khúc "Trong tầm tay".

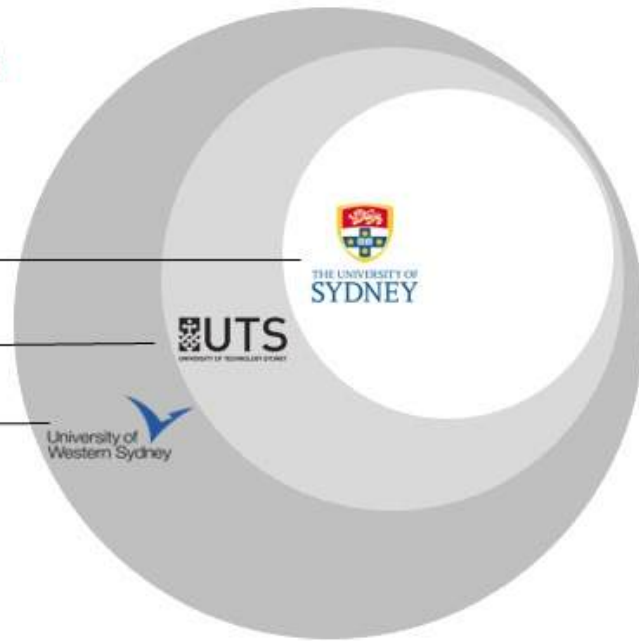
1. *Times Higher Education 2023, World university rankings, truy cập năm 2023, <https://www.timeshighereducation.com/world-university-rankings/western-sydney-university>.*

How universities are perceived

1. Aspirational

2. Discipline Experts

3. Achievable



Insight 2: Nhận thức về khu vực Western Sydney và tâm lý muốn tách biệt khỏi địa phương

Một **insight** then chốt khác nảy sinh từ nghiên cứu là tác động của những nhận thức tiêu cực về khu vực Western Sydney đối với quyết định chọn trường của sinh viên. Mặc dù nhiều đối tượng tham gia nghiên cứu đang sinh sống tại Western Sydney, họ lại không muốn định danh bản thân gắn liền với khu vực này do những định kiến (stigma) tiêu cực bủa vây. Các cuộc thảo luận công khai và thông tin trên truyền thông tính đến thời điểm đó phần lớn đều xoáy sâu vào góc nhìn khiếm khuyết (deficit perspective) về Western Sydney; do đó, giới trẻ đặc biệt mong muốn tạo khoảng cách với khu vực này.

Vì vậy, thông thường, quyết định lựa chọn trường đại học của họ không nằm ở bản thân ngôi trường đó, mà mang tính chất như một lời tuyên bố nhằm thoát ly khỏi tầm ảnh hưởng của địa phương. Mặc dù khu vực Western Sydney quả thực thường phải

chịu nhiều bất lợi về kinh tế - xã hội hơn so với các khu vực khác của Sydney [2], nhưng **insight** về việc những nhận thức này lại ảnh hưởng trọng yếu đến lựa chọn trường đại học của sinh viên là một kết quả đầy bất ngờ.

Ý tưởng chiến dịch được định hình bởi các Insight



Thay vì né tránh những **insight** này, Angelo Kourtis và đội ngũ của mình đã đối mặt trực diện với những tiêu cực đó. Ý tưởng chiến dịch, được hình thành từ các **insight** then chốt, đã hướng đến việc khơi dậy niềm tự hào về khu vực và sự "làm chủ" bản sắc Western Sydney (bao gồm cả việc đổi tên trường thành Western Sydney University). Chiến dịch cũng thách thức các sinh viên tương lai và các bên liên quan khác phải thay đổi cách tư duy về trường Đại học này.

Fortini-Campbell [3] mô tả việc tạo ra sự kết nối giữa **consumer insight** và **brand insight** trong quảng cáo là việc "chạm đúng điểm chạm vàng" (hitting the sweet spot). Đó là việc truyền tải một ý tưởng mạnh mẽ đến mức người tiêu dùng cảm thấy rằng "họ thực sự thấu hiểu mình".

Thông điệp từ Western Sydney University là: việc bạn đến từ đâu không quan trọng, điều quan trọng là nơi bạn đang hướng tới. Nghĩa là, tiềm năng của bạn không bị giới hạn bởi lịch sử cá nhân, hay bởi cách người khác nhìn nhận và định nghĩa về bạn. Và quan trọng hơn hết, Western Sydney University sẽ giúp bạn hiện thực hóa tiềm năng đó.

Thực thi chiến dịch (The campaign execution)

Để minh chứng cho cách mà Western Sydney University khai mở tiềm năng, Kourtis và Agency Sáng tạo đã tập trung vào hành trình cá nhân của các cựu sinh viên – những người gói gọn tinh thần "unlimited" (không giới hạn). Một trong những hành trình đó là của Deng Adut, một cựu binh lính trẻ em người Sudan. Chạy trốn khỏi cuộc chiến tranh tàn khốc tại Sudan, chàng thiếu niên đã được một gia đình ở Western Sydney nhận nuôi. Dù vốn tiếng Anh ít ỏi, anh đã làm thêm nhiều công việc, học tập chăm chỉ và giành được học bổng ngành Luật tại Western Sydney University. Hiện nay, Adut là một luật sư bào chữa thành công và là nhà vận động vì quyền người tị nạn.

Nguồn trích dẫn:

2. Australian Bureau of Statistics 2016, *Socio-economic indexes for areas (SEIFA)*, <https://www.abs.gov.au/ausstats/abs@.nsf/mf/2033.0.55.001>
3. Fortini-Campbell, L. (2001). *Hitting the Sweet Spot: How Consumer Insights Can Inspire Better Marketing and Advertising*. <https://www.amazon.com/Hitting-Sweet-Spot-Advertising-Account/dp/1887229080>

Câu chuyện của anh minh họa cho ý tưởng rằng: bạn có thể đến từ bất cứ đâu và thành công theo cách riêng của mình; rằng tiềm năng của bạn là không giới hạn (**unlimited**). Bộ phim dài 90 giây, bao gồm các cảnh tái hiện chiến tranh và kịch tính nơi phòng xử án, đã được ra mắt trên mạng xã hội và thu hút 3 triệu lượt xem trong vòng 7 ngày. Bộ phim đã được xem tại hơn 56 quốc gia và được các đài truyền hình trong nước cũng như quốc tế tiếp sóng – nâng tổng số lượt xem lên hơn 5 triệu.

Một hoặc nhiều yếu tố tương tác đã bị loại bỏ khỏi phiên bản văn bản này. Bạn có thể xem trực tuyến tại đây:

<https://westernsydney.pressbooks.pub/customerinsights2/?p=99#oembed-1>

Nguồn trích dẫn:

3. Fortini-Campbell, L 2017, *Hitting the sweet spot: how consumer insights can inspire better marketing and advertising*, The Copy Workshop.

Thành công sau chiến dịch (Post-campaign success)

Chiến dịch không chỉ nhận được những lời khen ngợi và giải thưởng toàn cầu, mà còn thay đổi hiệu quả những nhận thức tiêu cực trước đó. Những **insights** được rút ra từ nghiên cứu đã đóng vai trò then chốt trong việc định hình một chiến dịch làm thay đổi cục diện cách nhìn nhận của sinh viên tương lai về trường đại học. Đây là một sự từ chối táo bạo, không để người khác định nghĩa về khu vực và tổ chức của mình.

Đại học Western Sydney đã ghi nhận sự gia tăng ngay lập tức trong số lượng sinh viên nhập học, và các cuộc khảo sát thái độ đã phản ánh những chuyển biến tích cực trong nhận thức của các bên liên quan trọng yếu. Trường đại học hiện đang nắm giữ một **positioning** (định vị) mạnh mẽ là "unlimited" (không giới hạn) trong lĩnh vực giáo dục, và tiếp tục củng cố ý tưởng rằng: nếu bạn sẵn sàng làm việc chăm chỉ, táo bạo và đầy tham vọng, Đại học Western Sydney sẽ trao quyền để bạn đạt được tiềm năng vô hạn của mình.

Câu hỏi thảo luận

1. **"Tiềm năng không giới hạn" (unlimited potential) có ý nghĩa như thế nào đối với bạn?** Thông điệp này có tạo ra sự cộng hưởng (resonate) với bạn không? Bạn có thấy nó truyền cảm hứng không? Tại sao có hoặc tại sao không? Hãy thảo luận theo nhóm. Bạn có thể khám phá ra một số **insights** về bạn bè đồng trang lứa và cách họ nhìn nhận giá trị của giáo dục đại học.
2. **Hãy nghĩ về một chiến dịch mang lại cho bạn cảm giác: "họ thực sự thấu hiểu mình"** (có thể là một thương hiệu chăm sóc da, một thương hiệu xe hơi hoặc một thương hiệu công nghệ). Tại sao nó lại khiến bạn cảm thấy như vậy? Đó có phải là sự kết hợp giữa **consumer insight** và **brand insight** không?

Về tác giả chương này



Tiến sĩ Jacki Montgomery là Giảng viên chuyên ngành Quảng cáo tại Đại học Western Sydney. Với nền tảng học thuật chuyên sâu về tiếp thị và quảng cáo, bà đặc biệt quan tâm đến sự linh hoạt của ngôn ngữ và vai trò sâu sắc mà **insight** có thể đảm nhận trong các chiến dịch quảng cáo. Bà thường xuyên tham gia diễn thuyết tại các hội nghị học thuật và đã trình bày tại các sự kiện công nghiệp sáng tạo uy tín, bao gồm Semi Permanent và Mumbrella.

Jacki được công nhận cả trong nước và quốc tế là một chuyên gia truyền thông và giáo viên xuất sắc. Gần đây, bà đã được trao tặng

danh hiệu "Giảng viên của năm" (Teacher of the Year) tại Đại học Western Sydney.

Về Angelo Kourtis



Angelo Kourtis là thành viên của cộng đồng Đại học Western Sydney từ năm 1986 với các vai trò là sinh viên, người tốt nghiệp và nhân viên của trường. Ông từng nắm giữ nhiều vị trí quản lý cấp cao về tiếp thị trước khi được bổ nhiệm vào chức vụ Phó Hiệu trưởng phụ trách Sinh viên (Pro Vice-Chancellor - Students) vào tháng 10 năm 2011, và Phó Chủ tịch

phụ trách Nhân sự và Phát triển (Vice-President, People and Advancement) vào tháng 3 năm 2015. Ông lãnh đạo các bộ phận: Trải nghiệm Sinh viên và Tiếp thị; Nhân sự; Đối ngoại và Phát triển; Kế hoạch Sinh viên và Đại học – bộ phận sau cùng bao gồm cả mảng Phân tích và Tình báo Cạnh tranh (Competitive Intelligence and Analytics) và Quản trị Sinh viên.

Angelo đã khởi xướng và thiết lập chiến lược vòng đời sinh viên (student lifecycle strategy) tại trường, hé lộ những **insight** vô giá về "hành trình" mà sinh viên dần thân khi khao khát học vấn bậc cao và xa hơn thế nữa. Ông là người tiên phong trong khái niệm "Học giả Công dân" (Citizen Scholar) và lãnh đạo việc thành lập "Học viện" (The Academy) tại trường. Ông cũng là kiến trúc sư trưởng cho chiến lược tái định vị thương hiệu của Đại học Western Sydney – chiến dịch đã thu hút sự tán thưởng nồng nhiệt từ giới chuyên môn trong nước và quốc tế, đồng thời thúc đẩy một loạt các chương trình chuyển đổi trong các lĩnh vực kỹ thuật số, phân tích (analytics) và trải nghiệm sinh viên.

Ghi nhận bản quyền truyền thông (Media Attributions)

- **Ảnh Angelo Kourtis** © Western Sydney University, được cấp phép theo [giấy phép CC BY-NC-ND](#) (Ghi công - Phi thương mại - Không phái sinh).
- **Hình ảnh về Nhận thức (Perceptions) trong nghiên cứu tình huống** © [Jacki Montgomery](#), được bảo lưu [Mọi quyền](#) (All Rights Reserved).
- **Ảnh Deng Adut** © Western Sydney University, được bảo lưu [Mọi quyền](#) (All Rights Reserved).
- **Ảnh Jacki Montgomery** © [Jacki Montgomery](#), được cấp phép theo [giấy phép CC BY-NC-ND](#) (Ghi công - Phi thương mại - Không phái sinh).