

# **SỬ DỤNG DỮ LIỆU THỰC TẾ KHI GIẢNG DẠY THỐNG KÊ: NHẬN THỨC CỦA SINH VIÊN VỀ CHIẾN LƯỢC NÀY TRONG KHÓA HỌC NHẬP MÔN THỐNG KÊ**

*David L. Neumann, Michelle Hood, Michelle M. Neumann*

## **Tóm tắt:**

*Nhiều giáo viên thống kê khuyến nghị sử dụng dữ liệu thực tế trong các bài học trên lớp. Tuy nhiên, không có nhiều nghiên cứu có hệ thống về hiệu quả của phương pháp giảng dạy đó đối với sự tham gia và học tập của sinh viên. Nghiên cứu này đã nghiên cứu vấn đề nói trên trong khóa học thống kê đại học năm thứ nhất. Các sinh viên ( $n = 38$ ) đã được phỏng vấn và phản ánh của họ về việc sử dụng dữ liệu thực tế trong các lớp học được mã hóa thành các chủ đề. Các chủ đề kết quả là (a) quan điểm liên quan trong học tập, (b) quan tâm, (c) học/ghi nhớ tài liệu, (d) động lực, (e) liên quan/tham gia, và (f) hiểu biết về thống kê. Các kết quả cho thấy cả hai yếu tố nhận thức và tình cảm/động lực có liên quan đến việc sử dụng dữ liệu thực tế để giảng dạy thống kê. Kết quả cũng đưa ra các tính năng trong tập dữ liệu thống kê mà giáo viên cần phải tìm kiếm khi thiết kế bài học của mình.*

## **1. Giới thiệu**

Trong quá khứ, các khóa học thống kê tại trường đại học đã bị chỉ trích vì quá cứng nhắc và trừu tượng do sử dụng các phương pháp dạy học loại bỏ phần lớn việc học hỏi (Hogg, 1991; Willett & Singer, 1992). Các yếu tố đó đặc biệt quan trọng với sinh viên theo học ngành khoa học xã hội vì nhiều người trong số họ có thái độ tiêu cực hoặc lo lắng về thống kê (Onwuegbuzie & Wilson, 2003; Tremblay,

Gardner, & Heipel, 2000). Vì lý do đó, một số phương pháp nhằm cải thiện việc giảng dạy thống kê đã được phát triển. Ví dụ, Moore (1997) đề xuất tích hợp nhiều hơn giữa nội dung khóa học, phương pháp sư phạm và công nghệ. Các chiến lược giảng dạy cụ thể để đạt được điều này có thể bao gồm các hoạt động dựa trên máy tính (ví dụ: Morris, Joiner, & Scanlon, 2002) và đa phương tiện tương tác (ví dụ: González & Birch, 2000).

Tầm quan trọng của việc tạo ra những ảnh hưởng tích cực và tính kiên trì ở sinh viên cũng được ghi nhận (Budé, Van de Wiel, Imbos, Candel, Broers, & Berger, 2007). Ngoài việc nghiên cứu các chiến lược cụ thể, những thay đổi lớn hơn đối với chương trình giảng dạy và phương pháp giảng dạy đã được đề xuất. Bao gồm nhấn mạnh tư duy thống kê, sử dụng ít lý thuyết hơn và nhiều dữ liệu hơn, lấy phân tích dữ liệu làm trung tâm,

## ➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

xây dựng khả năng trực giác, khuyến khích chủ động học tập và sử dụng bối cảnh để phát triển suy luận thống kê (Cobb, 1992; Makar & Ben-Zvi, 2011; Scheaffer, 2001).

Việc áp dụng dữ liệu thực tế trong giảng dạy thống kê là một phương pháp ngày càng được khuyến khích trong lĩnh vực giáo dục thống kê. Dự án Hành động Chương trình giảng dạy của Hiệp hội toán học Mỹ trong những năm 1990 bao gồm trong các hướng dẫn của mình rằng cần có nhiều dữ liệu hơn trong giảng dạy (Cobb, 1992). Sau đó, các tài liệu Dự án về hiểu biết định lượng được phát triển. Những tài liệu này nhằm vào cấp trung học, không giống như Dự án Hành động chương trình giảng dạy của Hiệp hội Toán học Mỹ. Trong mô tả các tài liệu này, Scheaffer (2001) lưu ý rằng dữ liệu thực tế phải được sử dụng khi giảng dạy thống kê, đặc biệt là dữ liệu liên quan và phù hợp với sinh viên. Năm 2005, Hướng dẫn Đánh giá và Hướng dẫn về Giáo dục Thống kê (GAISE) của Hiệp hội Thống kê Hoa Kỳ đề xuất sử dụng dữ liệu thực tế nhằm đạt được các mục tiêu học tập mong muốn (Franklin & Garfield, 2006). Tương tự, các hướng dẫn

chương trình giảng dạy quốc gia về giảng dạy thống kê ở Úc, Anh, Nam Phi và New Zealand đều nhấn mạnh sự cần thiết phải sử dụng dữ liệu thực tế trong giảng dạy (Connor & Davies, 2002). Khái niệm về phát triển một môi trường học tập lý luận thống kê (Garfield & Ben-Zvi, 2009) nhấn mạnh vai trò của dữ liệu thực tế và động lực trong việc phát triển các kỹ năng lập luận thống kê của sinh viên. Sách hướng dẫn này được viết để cung cấp hướng dẫn và chiến lược cho giáo viên thống kê, đồng thời bao gồm đề xuất sử dụng dữ liệu thực tế trong giảng dạy (Dunn, Smith, & Beins, 2007; Garfield & Ben-Zvi, 2008; Hulsizer & Woolf, 2009). Nói chung, nhiều nhà sư phạm thống kê khác đưa ra lý lẽ để ủng hộ việc sử dụng dữ liệu thực tế trong giảng dạy thống kê (xem thêm, Bradstreet, 1996; Diamond & Sztendur, 2002; Holmes, 2002; Neumann, Neumann, & Hood, 2010; Rumsey, 2002; Singer & Willett, 1990).

Từ góc độ lý thuyết, việc áp dụng dữ liệu thực tế trong giảng dạy thống kê hướng tới các lý thuyết về học tập. Trong phạm vi lý thuyết học tập theo xu hướng xây dựng, sinh viên

sẽ xây dựng kiến thức dựa trên kinh nghiệm của mình bằng cách sử dụng các tập dữ liệu thực tế (Cobb, 1992; Garfield & Ben-Zvi, 2009). Kiến thức mới được tích hợp với kiến thức trước đây về giải thích dữ liệu sử dụng số liệu thống kê có liên quan. Cần sử dụng các tập dữ liệu để sinh viên có cơ hội phản ánh công việc của họ bằng dữ liệu, đồng thời giáo viên chỉ cung cấp hướng dẫn đủ để giúp sinh viên xây dựng kiến thức của riêng mình. Các tập dữ liệu thực cũng cung cấp bối cảnh cho một vấn đề thống kê. Bối cảnh có thể dựa trên dữ liệu hoặc có thể dựa trên môi trường học tập tự nhiên và xã hội (Pfannkuch, 2011). Sinh viên có thể xây dựng lý luận thống kê của mình thông qua sự tương tác giữa kiến thức bối cảnh về tập dữ liệu và kiến thức thống kê mới (Dierdorp, Bakker, Eijkelhof, & van Maanen, 2011; Pfannkuch, 2011).

Các tập dữ liệu có thể được sinh viên sử dụng để thực hành tính toán, thu được kinh nghiệm trong việc giải thích kết quả và xây dựng lý luận thống kê của mình về một vấn đề (Garfield & Ben-Zvi, 2009). Giáo viên cũng có thể sử dụng bộ dữ liệu để minh họa

các phương pháp nghiên cứu khác nhau, phương pháp phân tích dữ liệu và ứng dụng lý thuyết thống kê để giải quyết các vấn đề thực tế. Ví dụ, Morgan (2001) đưa ra hướng dẫn về cách thu thập dữ liệu từ các cáo phó được công bố trên một bài báo địa phương đã được sử dụng để dạy cho sinh viên đại học ngành tâm lý học và giáo dục trung học. Tập dữ liệu kết quả được sử dụng để giảng dạy các nguyên tắc về nhận dạng và giải thích bên ngoài, các nguyên tắc xử lý dữ liệu lộn xộn, các nguyên tắc giải thích các mối tương quan và nguyên tắc trình bày dữ liệu trong báo cáo.

Dữ liệu thực tế có thể có nhiều loại khác nhau (Hiệp hội Thống kê Hoa Kỳ, 2005). Khác biệt lớn có thể được thực hiện giữa dữ liệu thực tế (ví dụ: Dữ liệu lưu trữ, dữ liệu được thu thập trong các dự án nghiên cứu, dữ liệu do lớp học tạo ra) và dữ liệu nhân tạo (ví dụ: Dữ liệu giả định hoặc mô phỏng). Việc sử dụng dữ liệu thực tế được cho là có lợi thế đáng kể. Dữ liệu được tạo ra củng cố nhận thức rằng các số liệu thống kê là nhân tạo, tẻ nhạt và không hề thú vị (Singer & Willett, 1990). Trái lại, dữ liệu thực tế có thể là một

công cụ thúc đẩy học tập có ý nghĩa và chuẩn bị cho sinh viên sử dụng các kỹ thuật thống kê trong thế giới thực (Diamond & Sztendur, 2002). Tương tự, các tập dữ liệu thực tế có thể rất quan trọng đối với những sinh viên chưa có kinh nghiệm trong ngành (Bradstreet, 1996). Nhiều tác giả (ví dụ, Garfield & Ben-Zvi, 2009; Scheaffer, 2001; Singer & Willett, 1990) tranh luận kịch liệt về lợi thế của tập dữ liệu thực tế vì đây có thể là một phương tiện hữu ích và hiệu quả cho việc giảng dạy thống kê, giúp sinh viên phát triển các kỹ năng phân tích thông qua các tình huống nghiên cứu thực tế. Hơn nữa, dữ liệu thực tế không chỉ hỗ trợ giáo viên trong việc truyền đạt cách phân tích dữ liệu mà còn giải thích tại sao dữ liệu được phân tích. Các dữ liệu nên mang tính tích cực và có lợi cho sinh viên để cho thấy kỹ thuật thống kê có thể khám phá thông tin có ý nghĩa từ các con số như thế nào (Cobb, 1992; Garfield & Ben-Zvi, 2009). Cobb (1992) chỉ ra một cách đơn giản là "các khái niệm thống kê được học tập tốt nhất trong bối cảnh của các tập dữ liệu thực tế" (tr. 7).

Một cách để giáo viên có được dữ liệu thực tế là để họ

thu thập dữ liệu từ chính sinh viên (ví dụ: Neumann và cộng sự, 2010). Dữ liệu do sinh viên tạo ra cũng mang lại cho sinh viên kinh nghiệm trong thiết kế nghiên cứu, thu thập và phân tích dữ liệu, cũng như tạo sự hấp dẫn tương tác trong lớp học và chiều sâu trong quá trình học tập (Diamond & Sztendur, 2002). Chottiner (1991) đã giải thích việc sử dụng dữ liệu thực tế bao gồm tính cá nhân, thành tích học tập, phân bổ thời gian và dữ liệu khác thu được từ các sinh viên hoàn thành khóa học. Chottiner (1991) đã chỉ ra tính cá nhân của dữ liệu làm tăng sự thú vị đối với việc tìm hiểu về các kỹ thuật thống kê giúp thu được ý nghĩa từ dữ liệu. Mặc dù việc sử dụng dữ liệu thu được trực tiếp từ sinh viên có thể tạo ra sự quan tâm mạnh mẽ, việc thu thập dữ liệu theo cách này có thể không phù hợp với mọi hoàn cảnh và yêu cầu sinh viên và giáo viên đầu tư đủ thời gian cho bài tập.

Một cách khác để tạo tập dữ liệu từ sinh viên là thu thập dữ liệu từ các nguồn xuất bản hoặc kho dữ liệu. Các nguồn tài nguyên trực tuyến như DASL (Thư viện dữ liệu và câu chuyện), Lưu trữ dữ liệu của *Tạp chí*

## ➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

*Thống kê Giáo dục* (JSE), StatLib và website thống kê của chính phủ (ví dụ, Cục Thống kê) là các lựa chọn thay thế dễ dàng và nhanh chóng để thu thập dữ liệu. Cũng có sẵn các nguồn xác định cho bộ dữ liệu ở các bài báo (ví dụ, Willett & Singer, 1992), các trung tâm thông tin trực tuyến (ví dụ: Tài nguyên thống kê của Đại học Cologne) và các sách chứa dữ liệu (ví dụ: Andrews & Herzberg, 1985; Chatterjee, Handcock & Simonoff 1995; Hamilton, 1990; Hand, Daly, Lunn, McConway, & Ostrowski, 1994; Hodges, Krech, & Crutchfield, 1975; Nelson, 1982; Tanner, 1990; Tufte, 1974). Bạn cũng có thể thu được nhiều tập dữ liệu chuyên biệt hơn. Ví dụ, Dierdorp và cộng sự (2011) đã mô tả làm thế nào để dạy cách tương quan và hồi quy để giám sát đề điều trong bối cảnh sử dụng dữ liệu được tạo ra bởi các vệ tinh theo dõi sự biến dạng của bề mặt đất.

Tóm lại, dữ liệu thực tế đóng vai trò trung tâm trong các phương pháp giảng dạy hiện đại nhằm mục đích phát triển kiến thức thống kê và lý luận trong sinh viên. Tuy nhiên, có rất ít nghiên cứu về các kết quả có được từ việc sử dụng dữ liệu thực tế

đối với sự tham gia và học tập của sinh viên. Tiến hành nghiên cứu bằng cách sử dụng phương pháp thử nghiệm khiến việc đưa ra kết luận nhân quả gặp khó khăn trong các cơ sở giáo dục do lo ngại về mặt đạo đức rằng một nhóm sinh viên cần phải được dạy thống kê mà không sử dụng bất kỳ dữ liệu thực tế nào. Tuy nhiên, Gordon (2004) nhấn mạnh rằng thông tin quan trọng có thể thu được thông qua phân tích định tính các báo cáo do các sinh viên thống kê đưa ra về việc học tập của họ. Có ý kiến cho rằng phương pháp này làm tăng sự chú ý vào hành động và mục tiêu của sinh viên trong bối cảnh rộng hơn của các yếu tố thể chế và xã hội. Một phương pháp định tính cũng được áp dụng trong nghiên cứu này. Trong một khóa học thống kê đại học giới thiệu, Neumann và cộng sự (2010) đã đánh giá định tính về việc sử dụng dữ liệu thu thập được từ sinh viên về trải nghiệm học tập của sinh viên. Việc kiểm tra phản hồi của sinh viên chỉ ra rằng việc sử dụng dữ liệu của sinh viên cung cấp một cách thức để kết cấu sự tham gia trong lớp học, được các sinh viên nhận thức là một phương pháp khác, được sinh viên

báo cáo để giúp duy trì sự quan tâm, và sự hữu ích đối với bản thân trong việc xây dựng kiến thức của mình và đánh giá cao sự liên quan của thống kê.

Nghiên cứu này đã đánh giá việc sử dụng dữ liệu thực tế trong một khóa học thống kê năm thứ nhất đại học được giảng dạy trong một chương trình tâm lý học. Không giống như nghiên cứu của Neumann và cộng sự (2010) đánh giá việc sử dụng dữ liệu thu thập được từ sinh viên, nghiên cứu này tập trung vào việc sử dụng dữ liệu thực tế được thu thập từ các nguồn xuất bản hoặc kho dữ liệu trên World Wide Web. Theo ghi nhận của Earley (2007), một điều quan trọng là nghiên cứu không nên chỉ tập trung vào các kết quả thành tích liên quan đến chiến lược giảng dạy mà còn phải tìm hiểu cách chúng ảnh hưởng đến trải nghiệm của sinh viên. Vấn đề nghiên cứu được đề cập trong nghiên cứu này là: Nhận thức của sinh viên về cách sử dụng dữ liệu thực tế ảnh hưởng đến kinh nghiệm của họ trong việc học thống kê là gì? Cuộc điều tra này đã sử dụng phương pháp định tính trong đó sinh viên được phỏng vấn để gợi ra phản ánh của họ về cách sử

dụng dữ liệu thực tế liên quan đến sự tham gia và học tập của họ. Các cuộc phỏng vấn được sao chép và mã hóa để xác định các chủ đề chính được ghi lại trong các câu trả lời. Dựa trên các khuyến nghị trước đây cho thấy lợi ích của việc sử dụng dữ liệu thực tế, giả thuyết rằng sinh viên sẽ khẳng định việc sử dụng nó trong giảng dạy sẽ liên quan đến kinh nghiệm học tập của họ. Tuy nhiên, đặc biệt là các quá trình mà theo đó sinh viên nhận thức rằng việc sử dụng dữ liệu thực tế sẽ có liên quan đến việc học tập và tham gia của họ.

## 2. Phương pháp

### 2.1. Mô tả khóa học và sử dụng dữ liệu thực tế

Khóa học thống kê là khóa học bắt buộc ở trình độ năm nhất trong chương trình tâm lý học tại Đại học Griffith. Chỉ tiêu tuyển sinh thường dao động từ 220 đến 250 sinh viên. Các chủ đề được đề cập trong khóa học bao gồm giới thiệu về các phương pháp nghiên cứu, thống kê mô tả, phân bố chuẩn và điểm  $z$ , xác suất, phân bố mẫu, khoảng tin cậy và kiểm định giả thuyết bằng kiểm định  $t$  (kiểm định  $t$  trung bình 1 mẫu, kiểm định  $t$  trung bình 2 mẫu độc lập).

Khóa học được giảng dạy qua các bài giảng trong hai giờ hàng tuần và một buổi hướng dẫn trong một giờ. Các bài giảng theo định dạng bài giảng truyền thống, mặc dù chúng bao gồm phân tích dữ liệu dựa trên máy tính với phần mềm SPSS, mô phỏng các khái niệm thống kê và video. Các hướng dẫn bao gồm việc thực hành thống kê thông qua tính toán bằng tay và sử dụng phần mềm SPSS.

Các tập dữ liệu thực tế được tích hợp trong các bài giảng và chương trình hướng dẫn. Các tập dữ liệu được thu thập từ nhiều nguồn khác nhau, bao gồm các bài báo xuất bản, kho dữ liệu trên World Wide Web và dữ liệu được thu thập từ các nguồn có sẵn của người hướng dẫn khóa học. Các tập dữ liệu bao gồm các tập dữ liệu có liên quan đến môn tâm lý học và những người có cùng mối quan tâm. Các tập dữ liệu được sử dụng để phát triển các khái niệm thống kê, thay vì chỉ đơn thuần là cung cấp một số các số để tính toán số liệu thống kê. Các bộ dữ liệu và khái niệm thống kê được sử dụng để phát triển là: Danh sách các tỷ phú trên thế giới của Forbes (số liệu thống kê mô tả), thời gian phun trào của mạch nước phun Old

Faithful tại Công viên quốc gia Yellowstone (số liệu thống kê mô tả), lương của vận động viên bóng rổ từ Boston Celtics (số liệu thống kê mô tả), thời gian bơi của nam giới từ Thế vận hội Sydney 2000 (xác định các ngoại lệ), dữ liệu lịch sử về thời gian chạy marathon nam và nữ (tương quan và dự đoán), đo lường lường cô-ca-in dựa trên ghi chép Euro (lấy mẫu), giá trị trung bình và độ lệch chuẩn về hiệu suất của những người chơi thể thao hàng đầu (điểm  $z$ ), chi phí chăm sóc cho những người trong bệnh viện tâm thần (kiểm định  $t$  trung bình 1 mẫu), thời gian để uống một ly bia (khoảng tin cậy của giá trị trung bình), và một tập dữ liệu bao gồm kích thước não và điểm kiểm tra IQ của nam và nữ (tương quan và kiểm định  $t$  cho các nhóm độc lập).

Một hoặc nhiều bộ dữ liệu thực tế đã hình thành cốt lõi của bài học về một chủ đề thống kê đã cho. Ví dụ, số liệu thống kê mô tả được giảng dạy bằng cách sử dụng danh sách các tỷ phú trên thế giới của Forbes. Các sinh viên được giới thiệu đến trang web của Forbes và phương pháp mà dữ liệu được thu thập từ trang web và được nhập vào một tệp

## ➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

dữ liệu SPSS. Định dạng dữ liệu trong tệp dữ liệu SPSS đã được phác thảo (ví dụ: Đặt tên các biến, xác định thang đo). Các số liệu thống kê mô tả đồ họa qua biểu đồ tần số, phân phối tần số, biểu đồ hình tròn, biểu đồ cột và các điểm gốc được tính toán dựa trên dữ liệu. Các biểu đồ kết quả đã được thảo luận về cách chúng trình bày dữ liệu (ví dụ, cách chúng được thành lập) và cho phép diễn giải dữ liệu (ví dụ: Đặc trưng của phân bố dữ liệu). Số liệu thống kê mô tả bằng mode, trung vị, trung bình, tứ phân vị, độ lệch chuẩn, phương sai, phạm vi, vùng phân tán giá trị được tính toán tương tự dựa trên tập dữ liệu. Các giá trị kết quả đã được thảo luận về cách chúng được tính toán (ví dụ: Các mô hình/công thức thống kê) và cách chúng có thể được sử dụng để mô tả tổng thể.

### 2.2. Thu thập và phân tích dữ liệu

**Người tham gia** Nhóm lựa chọn những người tham gia bao gồm các sinh viên theo học trong khóa học nhập môn thống kê 1003PSY Phương pháp nghiên cứu và Thống kê 1 tại Đại học Griffith. Giáo viên của lớp này (tác giả đầu tiên) không tham gia vào việc tuyển sinh

viên và không được cho biết sinh viên nào đã được liên lạc hoặc đã đồng ý tham gia. Sinh viên cũng được thông báo về quá trình này. Như vậy, các thủ tục tuyển chọn và giấu tên sinh viên tham gia cho phép sinh viên cung cấp các đánh giá trung thực. Để chọn lựa những người tham gia, các nhà nghiên cứu ban đầu nhóm sinh viên theo cấp độ khóa học của họ (xuất sắc, giỏi, khá, đạt, trượt). Số lượng người tham gia tiềm năng được lựa chọn ngẫu nhiên từ mỗi nhóm cấp độ tỷ lệ thuận với tổng số sinh viên trong khóa học đã đạt ở mỗi cấp độ. Các sinh viên được tuyển chọn đã được gọi điện thoại và được mời tham gia vào nghiên cứu. Trong số 50 sinh viên được chọn ngẫu nhiên, không thể liên lạc được với 5 sinh viên do số điện thoại không chính xác từ hồ sơ sinh viên, 5 sinh viên từ chối tham gia (1 xếp loại xuất sắc, 1 xếp loại giỏi, 1 xếp loại khá và 2 xếp loại đạt), và 40 sinh viên đồng ý tham gia.

Trong số 40 sinh viên ban đầu đồng ý tham gia, 1 sinh viên xếp loại trượt không được tiếp xúc sau đó và 1 sinh viên khác xếp loại đạt đã rút lui. Mẫu cuối cùng bao gồm 29 nữ và 9 nam ở độ tuổi trung bình là 24,0

tuổi ( $SD = 7.3$ ). Một số người tham gia đã hoàn thành giáo dục sau trung học, bao gồm bằng cử nhân ( $n=1$ ), chứng chỉ ( $n=3$ ), và văn bằng ( $n=4$ ). Những người tham gia còn lại không có giáo dục sau trung học trước đó. Mẫu cuối cùng của những người tham gia báo cáo tham dự hầu hết các bài giảng trong suốt khóa học ( $M = 91.0\%$ ,  $SD = 13.1\%$ ). Tất cả những người tham gia đều được tặng một voucher cà phê trị giá 7 USD để đánh giá cao sự tham gia của họ trong nghiên cứu.

### *Thủ tục phỏng vấn*

Nghiên cứu đã được ban giám định thể chế phê duyệt. Tờ thông tin và mẫu chấp thuận được gửi qua bưu điện cho người tham gia sau khi họ đã đồng ý tham gia trong cuộc điện thoại ban đầu của trợ lý nghiên cứu. Một trợ lý nghiên cứu thứ hai sau đó liên lạc với những người tham gia để tiến hành phỏng vấn. Người phỏng vấn này cũng không tham gia vào khóa học, không tham gia vào việc lựa chọn những người tham gia và không nhìn thấy điểm mà sinh viên đã đạt được. Các cuộc phỏng vấn được tiến hành sau khi khóa học đã được hoàn thành và sinh viên đã biết được xếp loại.

Khi bắt đầu cuộc phỏng vấn, tất cả những người tham gia đã đồng ý cho cuộc gọi được ghi lại để phản hồi có thể được sao chép chính xác trong quá trình mã hóa dữ liệu. Mỗi cuộc phỏng vấn qua điện thoại kéo dài khoảng 20 phút và bắt đầu với các câu hỏi bán cấu trúc hỏi về việc sử dụng các tập dữ liệu thực tế trong tài liệu giảng dạy của khóa học, bao gồm tham chiếu đến một ví dụ cụ thể được sử dụng trong khóa học. Sau đó, người tham gia được hỏi "Suy nghĩ của bạn về việc sử dụng các tập dữ liệu thực tế?" Khi thích hợp, sinh viên được yêu cầu làm rõ hoặc giải thích về câu trả lời. Thông tin bổ sung được người phỏng vấn gợi ý bằng cách sử dụng lời nhắc: "Nó có giúp bạn tương tác với tài liệu không? Như thế nào?", "Nó có giúp thúc đẩy bạn tìm hiểu về thống kê không?", "Một số khía cạnh tích cực với nó là gì?", và "Một số khía cạnh tiêu cực đối với nó là gì?". Mọi người tham gia đều được hỏi những câu hỏi đó. Cuộc phỏng vấn cũng bao gồm các câu hỏi về các sáng kiến giảng dạy khác trong khóa học và những điều này đã được báo cáo ở nơi khác (Neumann, Hood, & Neumann, 2009, 2012;

Neumann, Neumann, & Hood, 2010, 2011). Khi kết thúc cuộc phỏng vấn, những người tham gia được hỏi về các khía cạnh không liên quan của khóa học và một số câu hỏi dựa trên nhân khẩu học.

**Mã hóa dữ liệu** Các câu trả lời từ các cuộc phỏng vấn qua điện thoại đã ghi được nhập trực tiếp vào tài liệu điện tử. Mỗi câu trả lời được kèm theo mã số nhận dạng người tham gia. Tổng cộng, có 124 câu trả lời. Sử dụng quy trình mã hóa do Neuman mô tả (2006), mã hóa mở được sử dụng đầu tiên để nhóm các câu trả lời thành các chủ đề phân tích sơ bộ. Điều này ban đầu đi qua các tập dữ liệu đã được thực hiện bởi một mã hóa viên không tham gia vào việc lựa chọn người tham gia hoặc quá trình phỏng vấn. Mã hóa viên đã nhóm các câu trả lời thành các chủ đề có tính chất khái niệm tương tự. Một danh mục các câu trả lời không thể nhóm thành một chủ đề cũng được tạo ra. Một mã hóa viên độc lập thứ hai đã kiểm tra các câu trả lời và phân bổ độc lập chúng qua các chủ đề đã được tạo bởi bộ mã đầu tiên. Do nhiều câu trả lời bao gồm nhiều chủ đề, tất cả đều có thể được đặt vào các danh mục

khác nhau, hai mã hóa viên được tự do mã hóa toàn bộ câu trả lời vào một chủ đề cụ thể hoặc chia câu trả lời thành các thành phần nhỏ hơn để chúng có thể được mã hóa thành nhiều chủ đề. Hai mã hóa viên tán thành 87% trong các quyết định mã hóa. Sự khác biệt về mã hóa thường tập trung vào mức độ câu trả lời nên được phân chia theo các chủ đề khác nhau. Ví dụ, một mã hóa viên đã phân bổ toàn bộ câu trả lời "Nó giúp bạn nhận ra tầm quan trọng của chủ đề này và mức độ liên quan của nó đối với mọi thứ bạn làm" vào một danh mục *Mức độ liên quan thực tế*, trong khi mã hóa viên khác đặt "Nó giúp bạn nhận ra tầm quan trọng của chủ đề này" vào danh mục *Kiến thức*, và "Mức độ liên quan của nó đối với mọi thứ bạn làm" vào danh mục *Mức độ liên quan thực tế*. Tất cả các khác biệt đã được giải quyết thông qua thảo luận với cá nhân thứ ba.

Theo quy trình mã hóa mở, có 7 chủ đề. Một trong những chủ đề này đã bị xóa khỏi các phân tích sâu hơn vì nó chỉ chứa 8% người tham gia. Chủ đề này đã được đưa ra nhãn dự kiến của các báo cáo *Cần thiết* và phản ánh trong đó sinh viên nhận xét

## ➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

rằng việc sử dụng tập dữ liệu thực tế là cần thiết để tìm hiểu số liệu thống kê. Các câu trả lời đã được mã hóa trong danh mục "Khác" cũng bị bỏ qua. Sáu chủ đề cuối cùng đã được kiểm tra thêm. Nhận được sử dụng để mô tả chủ đề đã được chọn lọc và định nghĩa cho mỗi chủ đề được thuyết minh để phản ánh chính xác bản chất của các câu trả lời trong đó. Các câu trả lời mẫu cũng được chọn cho mỗi chủ đề. Các nhãn, định nghĩa và ví dụ được tạo trong quá trình này đã được thảo luận thêm và sàng lọc giữa các mã hóa viên. Theo sau cuộc

thảo luận, các bản sửa đổi đã được thực hiện thành hai định nghĩa danh mục và hai ví dụ về danh mục.

### 3. Kết quả

Các chủ đề nổi lên từ các cuộc phỏng vấn (xem Bảng 1) bao gồm các tham chiếu đến mức độ liên quan, sở thích, học tập, động lực, sự gắn kết với tài liệu và kiến thức. Chủ đề được trích dẫn nhiều nhất bởi các sinh viên được dán nhãn *Mức độ liên quan thực tế* (63%). Các sinh viên nhận xét rằng dữ liệu thực tế liên quan đến quan điểm rộng hơn và phù hợp hơn về thống kê. Ví dụ,

một sinh viên nhận xét rằng sử dụng dữ liệu "Cung cấp ví dụ thực tế về cách sử dụng số liệu thống kê" và một sinh viên khác ghi chép rằng "Nó giúp bạn nhận ra tầm quan trọng của chủ đề này và mức độ liên quan đến mọi thứ bạn làm". Mức độ liên quan được kết hợp với dữ liệu thực tế có thể đã xuất hiện là kết quả của việc hiển thị cách các số liệu thống kê có thể được áp dụng trong các điều kiện thực tế kết hợp với thực tế là các tình huống đã quen thuộc hoặc quan trọng đối với sinh viên.

*Bảng 1. Chủ đề, định nghĩa và phần trăm đã cung cấp phản hồi cho mỗi chủ đề*

Chủ đề	Định nghĩa	Phần trăm
Mức độ liên quan thực tế	Dữ liệu thực tế áp dụng các khái niệm thống kê cho các tình huống thực tế để cung cấp quan điểm thống kê rộng hơn và phù hợp hơn	63%
Quan tâm	Dữ liệu thực tế liên quan đến sự quan tâm và thích thú khi tìm hiểu về thống kê	58%
Học tập và trí nhớ	Dữ liệu thực tế đã đóng một vai trò trong việc học và ghi nhớ các khái niệm và lý thuyết thống kê	50%
Động lực	Dữ liệu thực tế đã khuyến khích và kích thích khi học về thống kê	37%
Tham gia	Dữ liệu thực tế đã thu hút sinh viên vào việc tìm hiểu về thống kê và khiến họ tham gia và thúc đẩy suy nghĩ xa hơn	32%
Kiến thức	Dữ liệu thực tế liên quan đến việc hiểu số liệu thống kê rõ ràng và toàn diện hơn	24%

Điều này được minh họa bằng các bình luận như "Bởi

vì bạn có thể đưa nó vào quan điểm và thực sự áp

dụng những gì bạn biết" và "Tôi nghĩ chúng là tốt nhất



bởi vì nó tốt để rút lại mọi thứ chứ không phải là ứng dụng thực tế và tất cả đều là những thứ mà mọi người có thể liên quan đến”.

Hai chủ đề đã ghi lại các cách mà các tập dữ liệu thực tế liên quan đến việc học được gắn nhãn *Học tập và trí nhớ* (50%) và *Kiến thức* (24%). Chủ đề *Học tập* đã nắm bắt được cách học và ghi nhớ các khái niệm thống kê và lý thuyết liên quan đến thảo luận về các tập dữ liệu thực tế. Dữ liệu thực tế dường như mang lại lợi ích đặc biệt vì bản chất trừu tượng và toán học của các thống kê như được thể hiện bởi các nhận xét “Nó dễ nhớ hơn vì bạn nhớ ví dụ bằng công thức thống kê” và “Tôi nghĩ rằng các ví dụ thực sự luôn ở trong tâm trí của bạn”. Một số sinh viên cho rằng các bối cảnh sử dụng dữ liệu thực tế như là nguyên tắc trong đánh giá, được thể hiện bởi bình luận “Nó đã giúp kích hoạt những trí nhớ và những thứ khi chúng tôi làm bài kiểm tra”. Trong một số trường hợp, các ví dụ chia tài liệu thành nhiều phần để quản lý hơn, ví dụ: “Giúp học tập tài liệu và chia nhỏ tài liệu”, “Nó khác biệt và chia nhỏ tài liệu

bài giảng” và “Thật tuyệt khi nghe một câu chuyện và không chỉ là số liệu thống kê mà còn biết nhiều hơn về nó”.

Chủ đề *Kiến thức* (24%) được phân bổ cho một chủ đề riêng biệt vào *Học tập và trí nhớ* bởi vì các bình luận đủ đa dạng để yêu cầu hai loại. Chủ đề *Kiến thức* được mã hóa là có liên quan nhiều hơn để hiểu và tạo nên ý nghĩa của tài liệu. Ngược lại, chủ đề *Học tập và trí nhớ* chủ yếu nắm bắt được ý tưởng chỉ đơn thuần là có thể tạo ra những trí nhớ hoặc sau đó lấy tài liệu. Việc sử dụng dữ liệu thực tế có liên quan đến các nhận xét về sự rõ ràng của tài liệu và cách nó đóng một vai trò trong việc xây dựng và nắm bắt các khái niệm thống kê. Ví dụ, một sinh viên nhận thấy rằng “Chúng là một công cụ ngăn để giúp bạn hiểu” và một sinh viên khác nhận thấy “Chúng giúp bạn hiểu, bạn sẽ có được một mô tả rõ ràng về việc nó được giải thích và điều này giúp bạn vì vậy tôi nghĩ nó tốt”.

Dữ liệu thực tế cũng thúc đẩy suy nghĩ thêm về khái niệm thống kê ở chỗ nó “Có ý nghĩa hơn... bởi vì bạn nghĩ, bạn sẽ hiểu nó nhiều hơn”.

Ba chủ đề nổi lên liên quan đến nhiều khía cạnh tình cảm hơn các chủ đề đã thảo luận trước đó. Một số lượng đáng kể sinh viên viện dẫn *Quan tâm* (58%). Ví dụ, có nhận xét rằng “Thú vị hơn là chỉ cung cấp loại lý thuyết của sự vật” và “Có thể làm cho công việc vui vẻ hơn và dễ dàng hơn để tập trung, nó là một cái gì đó không khô khan như các nội dung khác vì vậy nó là tốt”. Một số dữ liệu thực tế khác nhau hoặc đáng ngạc nhiên và điều này có thể đóng một vai trò trong giá trị quan tâm của nó, như được thể hiện bởi các bình luận “Thú vị hơn là loại thống kê cũ. Thú vị hơn vì nó không bình thường” và “Thật thú vị khi tìm ra cách họ đã tìm ra điều đó và nhận được những số liệu thống kê từ đó”.

Chủ đề *Động lực* (37%) cũng xuất hiện cho khoảng một phần ba số sinh viên minh họa rằng dữ liệu thực tế có thể kích thích một số sinh viên tìm hiểu về thống kê. Ví dụ, một sinh viên nhận xét “Chúng là một công cụ động lực tốt”. Một sinh viên khác nhận xét “Nó đã giúp tôi có một lý do để nghiên cứu số liệu thống kê”. Tài liệu có liên quan đến khía cạnh động lực của thống kê

## ➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

học tập có thể có nhiều khả năng khiến sinh viên tìm kiếm và tìm hiểu thêm về các khái niệm, như được thể hiện bởi các nhận xét "Nó chỉ thúc đẩy tôi muốn biết nhiều hơn và biết thêm thông tin về nó" và "Nó làm cho bạn muốn tìm hiểu về nó nhiều hơn một chút".

Chủ đề cuối cùng là *Tham gia* (32%). Chủ đề này được phân biệt với *Quan tâm* và *Động lực* trong đó nó liên quan đến vai trò tiềm năng của tập dữ liệu thực tế trong sự liên quan của sinh viên trong vấn đề này. Một bình luận minh họa là "Tôi nghĩ rằng chúng đã giúp bạn tham gia và trải nghiệm số liệu thống kê". Việc sử dụng dữ liệu thực tế và mối quan hệ của nó với sự tham gia có thể đã giúp thúc đẩy suy nghĩ sâu sắc, như được thể hiện bởi nhận xét "Nó chỉ khiến bạn nghĩ về nó nhiều hơn một chút thay vì chỉ ghi nhớ mọi thứ." Một sinh viên khác đã nhận xét "Bạn có thể tham gia vì đó là điều mà mọi người biết" đề xuất rằng việc sử dụng các biến số và biện pháp quen thuộc là một yếu tố quan trọng của sự tương tác.

### 4. Thảo luận

Phân tích các báo cáo sinh viên trong nghiên cứu này cho thấy việc sử dụng

dữ liệu thực tế liên quan đến các chủ đề nhận thức và tình cảm/ động lực trong các trải nghiệm học tập của sinh viên. Chủ đề được trích dẫn phổ biến nhất là chủ đề liên quan đến mức độ liên quan thực tế. Điều này chỉ ra rằng việc sử dụng các tập dữ liệu thực tế có liên quan đến việc đạt được ý nghĩa từ các số liệu thống kê theo một ý nghĩa được áp dụng hoặc thực tế hơn. Một nửa mẫu cho ý kiến được mã hóa là có liên quan đến học tập và trí nhớ. Những sinh viên này thường tin rằng dữ liệu thực tế đóng một vai trò trong việc học các khái niệm và kỹ thuật mới, cũng như đóng vai trò là một công cụ để ghi nhớ nội dung. Gần một phần tư mẫu báo cáo rằng việc sử dụng dữ liệu thực tế là một yếu tố quan trọng trong việc hiểu tài liệu khóa học dễ dàng và hoàn chỉnh hơn. Ngoài ra, các sinh viên đã báo cáo rằng dữ liệu thực tế có liên quan đến sự quan tâm, động lực và sự tham gia của họ trong việc tìm hiểu về thống kê.

Các chủ đề nổi lên trong nghiên cứu này là phù hợp rộng rãi với lời khuyên sử dụng dữ liệu thực tế trong giảng dạy thống kê. Theo ghi chép của một số tác giả

(ví dụ, Bradstreet, 1996; Cobb, 1992; Garfield & Ben-Zvi, 2008; Scheaffer, 2011), dữ liệu thực tế cho biết thêm giá trị cho quá trình học tập, cũng như kích thích động lực thông qua đối đầu với các vấn đề thực tế. Tương tự, Diamond và Sztendur (2002) đã nói rằng dữ liệu thực tế là một công cụ thúc đẩy cho phép hiểu biết sâu hơn về các phương pháp thống kê, cũng như cung cấp sự thú vị trong quá trình tìm hiểu. Phát hiện của cả hai nhận thức (*Học tập hỗ trợ, Kiến thức*) và chủ đề tình cảm/ động lực (*Quan tâm, động lực, tham gia*) trong các câu trả lời của sinh viên trong nghiên cứu này phù hợp với kết luận này. Chủ đề *Mức độ liên quan thực tế* được tìm thấy trong nghiên cứu này phù hợp với khái niệm rằng kinh nghiệm nghiên cứu nên xác thực và có căn cứ trong bối cảnh (Cobb, 1992; Dierdorp và cộng sự, 2011; Garfield & Ben-Zvi, 2009; Pfannkuch, 2011; Singer & Willett, 1990). Chủ đề này cũng phù hợp với các phương pháp giúp sinh viên xây dựng kiến thức thông qua kinh nghiệm của họ với các tập dữ liệu thực tế (Garfield & Ben-Zvi, 2009) dựa trên lý thuyết học tập theo xu hướng xây dựng.

Tuy nhiên, những kết quả này chỉ liên quan đến kết quả cuối cùng của việc sử dụng dữ liệu thực tế. Các phương pháp được sử dụng không kiểm tra quá trình học tập trong suốt các hoạt động của bài học. Để hiểu rõ hơn về việc xây dựng kiến thức bằng cách sử dụng dữ liệu thực tế, cần phải quan sát sinh viên trước, trong và sau các bài học kết hợp các tập dữ liệu thực tế.

Dựa trên kết quả này và báo cáo của những người khác, một số khuyến nghị được đưa ra giúp giáo viên chọn tập dữ liệu tốt cho buổi học của họ. Để tăng cường lợi ích thực tế trong việc sử dụng dữ liệu, giáo viên nên chọn các tập dữ liệu liên quan đến các chủ đề trong lĩnh vực học tập của sinh viên hoặc quen thuộc với sinh viên trong cuộc sống hàng ngày của họ (ví dụ: Xem Gourgey, 2000; Hulsizer) & Woolf, 2009). Bằng cách này, dữ liệu về một chủ đề thú vị hoặc quan trọng đối với sinh viên để đảm bảo rằng nó tạo ra sự tham gia và động lực của họ. Ví dụ, trong ứng dụng hiện nay, dữ liệu liên quan đến chi phí chăm sóc cho những người trong bệnh viện tâm thần và mối quan hệ giữa

kích thước não và chỉ số IQ có khả năng thu hút mỗi quan tâm chuyên môn của sinh viên tâm lý học. Dữ liệu về thời gian uống bia trong ứng dụng hiện tại có thể chỉ thú vị đối với yếu tố mới lạ hoặc mức độ liên quan đến số lượng giới hạn các sinh viên đại học. Tuy nhiên nên tránh việc lựa chọn dữ liệu thực tế chỉ đơn thuần là cho các hiệu ứng tiềm năng trong việc tăng cường sự tham gia và động lực. Điều quan trọng là các số liệu thống kê được tính toán từ tập dữ liệu sẽ minh họa hiệu quả khái niệm thống kê đang được giảng dạy (Cobb & McClain, 2004; Scheaffer, 2001).

Trong khi việc sử dụng dữ liệu thực tế có thể có một số ưu điểm cho giáo viên và kết quả học tập của sinh viên, thì cũng có thể có những nhược điểm tiềm ẩn. Mặc dù không có chủ đề nào xuất hiện trong nghiên cứu này cho thấy những bất lợi lớn, các nhận xét của sinh viên bị cô lập nêu ra một số cảnh báo. Ví dụ, một sinh viên nhận xét rằng các ví dụ không dựa trên tâm lý học không tốt bằng các ví dụ về tâm lý học. Điều này cho thấy rằng giáo viên sẽ cần dành thời gian để thu thập

các tập dữ liệu có liên quan đến chương trình học của sinh viên (trong trường hợp này là tâm lý học). Mặc dù có sẵn các tập dữ liệu trên Internet, nhưng chỉ có một số lượng dữ liệu hạn chế có thể liên quan đến sinh viên và khái niệm thống kê mà giáo viên muốn thảo luận. Một sinh viên khác nhận xét về việc bị phân tâm trong bài giảng của các sinh viên khác nói về chủ đề mà dữ liệu liên quan đến. Vì vậy, trong khi các tập dữ liệu có thể có ưu điểm cho việc tạo ra sự quan tâm và tham gia, giáo viên cũng nên cẩn thận để giữ cho dữ liệu thực tế không bị mất tập trung.

Mặc dù nghiên cứu này không cho thấy bất kỳ chủ đề nào liên quan đến cân nhắc về đạo đức và bản quyền, nhưng các vấn đề đó vẫn có thể quan trọng khi sử dụng tập dữ liệu thực tế trong giảng dạy (ví dụ: Morgan, 2001; Neumann và cộng sự, 2010). Khuyến cáo rằng dữ liệu thực tế chỉ được sử dụng khi dữ liệu được thu thập một cách có đạo đức và khi những người tham gia đã đồng ý cho người khác sử dụng dữ liệu. Nghiên cứu này không tìm thấy bất kỳ nhận xét nào của sinh viên liên quan đến các vấn đề

## ➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

đạo đức. Giáo viên cũng nên kiểm tra bất kỳ hạn chế bản quyền hoặc sở hữu trí tuệ nào có thể được liên kết với dữ liệu mà họ đang sử dụng. Các tập dữ liệu thực tế cũng có thể "lộn xộn" ở chỗ chúng chứa các bản phân phối lệch hoặc các ngoại lệ. Trong khi điều này có thể trở thành một lợi thế bằng cách dạy sinh viên về kiểm tra phân phối và cách đối phó với các ngoại lệ, thảo luận đó có thể tập trung vào khái niệm thống kê chính mà giáo viên muốn sinh viên học.

### **Hạn chế và nghiên cứu trong tương lai**

Nghiên cứu này dựa trên một số lượng nhỏ sinh viên thuộc một nhóm cụ thể (tức là, những sinh viên đã hoàn thành khóa học nhập môn thống kê là một phần của bằng tâm lý học). Như vậy, kết quả có thể khác nhau đối với sinh viên trong các trường hợp khác. Ví dụ, các đặc điểm của sinh viên như năm học, mức độ và quan tâm về thống kê, cũng như các đặc điểm khóa học như mức độ của khóa học thống kê được hoàn thành có thể ảnh hưởng đến lợi ích nhận thức của dữ liệu thực tế. Cần nghiên cứu sâu hơn xem xét kết quả hiện tại khái quát như thế nào giữa các nhóm

sinh viên và các loại thống kê khác nhau. Người ta cho rằng chủ đề bao quát được tìm thấy trong nghiên cứu này, Mức độ liên quan thực tế, sẽ được tìm thấy sự đa dạng trên nhiều nhóm sinh viên và đặc biệt là giữa các sinh viên đại học đang nghiên cứu thống kê trong các chương trình phi toán học.

Sự thiên vị có thể được đưa vào nghiên cứu thông qua phương pháp được sử dụng. Sự tham gia vô danh của sinh viên và việc sử dụng một người phỏng vấn không liên quan đến khóa học và chương trình cấp bằng của sinh viên được thiết kế để tránh thiên vị của người tham gia và người thực nghiệm. Tuy nhiên, cấu trúc của cuộc phỏng vấn có thể đã thúc đẩy nhiều phản ứng tích cực hơn so với những phản ứng tiêu cực. Điều này là do ba câu hỏi trực tiếp đầu tiên được bố trí một cách tích cực. Đó chỉ là câu hỏi cuối cùng hỏi về những trải nghiệm tiêu cực. Nghiên cứu trong tương lai nên có cùng một số câu hỏi tích cực và tiêu cực và đối trọng với thứ tự của các câu hỏi để kiểm soát sự thiên vị tiềm năng được đưa vào bởi cấu trúc của cuộc phỏng vấn. Ngoài ra, các câu hỏi bổ

sung có thể được hỏi để tăng tính đa dạng và chiều sâu của các câu trả lời của sinh viên. Ví dụ, sinh viên có thể được hỏi "điều gì xảy ra nếu" hoặc các câu hỏi giả định hoặc được yêu cầu phản ánh khi giảng dạy thống kê từ quan điểm của giáo viên.

Các chủ đề *Động lực*, *Tham gia*, và *Kiến thức* đã được xác nhận bởi chỉ một phần nhỏ của mẫu. Một giải thích cho phát hiện này là các sinh viên, trên mức trung bình, không tìm thấy việc sử dụng dữ liệu thực tế để thúc đẩy, thu hút hoặc quan trọng để hiểu về thống kê. Ngoài ra, nó có thể phản ánh bản chất của phương pháp thu thập dữ liệu đã được sử dụng. Các sinh viên được phỏng vấn và trả lời tự do các câu hỏi. Như vậy, các chủ đề nổi lên từ một cá nhân sinh viên có thể là những chủ đề phù hợp nhất với họ và điều đó xuất hiện trong suốt cuộc phỏng vấn. Có thể một số sinh viên có thể đã tìm thấy việc sử dụng dữ liệu thực tế để thúc đẩy, nhưng không đề cập đến thực tế này trong cuộc phỏng vấn của họ. Để có được thông tin kỹ lưỡng hơn về tỷ lệ sinh viên xác nhận mỗi chủ đề, cần thu thập dữ

liệu theo cách cho phép sinh viên xếp hạng hoặc đưa ra nhận xét liên quan đến tất cả các chủ đề.

Phương pháp quan sát được sử dụng trong nghiên cứu này do đó bị giới hạn bởi tính chất mô tả. Phương pháp này rất hữu ích trong các cuộc điều tra ban đầu về các phương pháp giảng dạy sáng tạo nhằm hiểu rõ hơn về bản chất tác động của phương pháp tiếp cận đối với trải nghiệm của sinh viên. Nó cũng cung cấp chi tiết phong phú để làm cơ sở nghiên cứu trong tương lai trong lĩnh vực này và có thể đặt nền móng cho việc tiến hành nghiên cứu định lượng. Các nghiên cứu trong tương lai nên xem xét sáu khía cạnh nổi lên trong nghiên cứu này.

Một giới hạn nữa của nghiên cứu quan sát là nó không cho phép rút ra các mối quan hệ nhân quả. Ví dụ, mặc dù nhiều sinh viên đã tuyên bố rằng việc sử dụng các ví dụ thực tế đã được thu hút, nhưng không thể kết luận rằng việc sử dụng các ví dụ đã làm tăng mức độ tương tác. Kết luận này có thể được rút ra chỉ khi nghiên cứu có sinh viên được đặt trong một nghiên cứu và một nhóm kiểm soát, được lựa chọn bởi một số

quá trình ngẫu nhiên. Điều này có thể đặc biệt mang tính hướng dẫn trong việc chứng minh mối quan hệ nhân quả của việc sử dụng dữ liệu thực tế về kết quả học tập, chẳng hạn như thành tích học tập trong khóa học. Tuy nhiên, luôn luôn có những hạn chế thực tế khi có thể chỉ định ngẫu nhiên sinh viên theo các phương pháp giảng dạy khác nhau.

Trọng tâm của nghiên cứu này là sử dụng các tập dữ liệu thực tế, trái ngược với dữ liệu "được hình thành" hoặc "được tạo ra" với mục đích minh họa một khái niệm thống kê. Trong phương pháp này (xem Lesser, 2011 để biết ví dụ), các kết quả toán học làm nền tảng cho các số liệu thống kê dựa hoàn toàn vào các con số trong tập dữ liệu chứ không dựa vào cách chúng bắt nguồn. Như vậy, sinh viên có thể đạt được nhiều lợi ích nếu họ thực hành tính toán và thao tác các con số bằng bất kỳ loại tập dữ liệu nào. Nó sẽ là hướng dẫn cho nghiên cứu trong tương lai để kiểm tra những lợi ích cụ thể là sử dụng dữ liệu thực tế khi so sánh với dữ liệu chưa được thu thập từ các quan sát thực tế. Dựa trên nghiên cứu

này, người ta dự kiến rằng dữ liệu thực tế sẽ có thể tốt hơn để chứng minh mức độ liên quan thực tế của thống kê. Dữ liệu thực tế cũng được kỳ vọng sẽ tạo ra nhiều sự quan tâm, động lực và sự tham gia của sinh viên hơn là các tập dữ liệu không dựa trên các quan sát thực tế.

**Kết luận** Tóm lại, nghiên cứu này đã chỉ ra rằng một số chủ đề dường như nổi lên khi dữ liệu thực tế liên quan đến việc giảng dạy thống kê. Khi dữ liệu thực tế được sử dụng để minh họa các khái niệm thống kê, tính toán thực hành và thể hiện ứng dụng của thống kê, có một số kết hợp với các loại trải nghiệm học tập cụ thể của sinh viên như mức độ liên quan, kiến thức và học tập, quan tâm, động lực và tương tác. Những kết quả này cho thấy rằng, phù hợp với các khuyến nghị của các tác giả khác (Bradstreet, 1996; Cobb, 1992; Garfield & Benzvi, 2008; Moore, 1997; Scheaffer, 2001) có thể hiệu quả để các giáo viên thống kê kết hợp dữ liệu thực tế vào việc dạy học của mình.

*Lê Hằng, Ngọc Bích (dịch)*

*Nguồn: Bài viết đã được phổ biến tại Hội nghị ICOTS10, <https://icots.info/10>*