

## Bài báo nghiên cứu

# ĐÁNH GIÁ SỰ PHÙ HỢP CỦA DỮ LIỆU TRONG DẠY HỌC THỐNG KÊ Ở CẤP TRUNG HỌC CƠ SỞ: MINH HỌA QUA HAI MÔ HÌNH GEOGEBRA

*Tôn Thất Tú\*, Lê Vũ Khoa, Lê Bùi Quỳnh Chi, Lê Thành Nhân*

*Trường Đại học Sư phạm – Đại học Đà Nẵng, Việt Nam*

*\*Tác giả liên hệ: Tôn Thất Tú – Email: [ttu@ued.udn.vn](mailto:ttu@ued.udn.vn)*

*Ngày nhận bài: 02-02-2025; Ngày nhận bài sửa: 10-9-2025; Ngày duyệt đăng: 08-10-2025*

## TÓM TẮT

*Thống kê và xác suất là một trong ba mạch kiến thức chính trong chương trình giáo dục phổ thông môn Toán 2018, được dạy từ lớp 2 đến lớp 12, với các khái niệm và kỹ năng từ cơ bản đến nâng cao. Ở cấp trung học cơ sở, phần thống kê chủ yếu tập trung vào việc thu thập, phân loại và biểu diễn dữ liệu qua bảng và biểu đồ. Việc thu thập dữ liệu từ nhiều nguồn khác nhau có thể ảnh hưởng đến chất lượng của dữ liệu, từ đó tác động đến việc phân tích và ra quyết định. Do đó, việc đánh giá sự phù hợp của dữ liệu trở thành yêu cầu quan trọng và đã được thể hiện trong nội dung yêu cầu cần đạt của chương trình. Bài báo này phân tích lý do cần đánh giá sự phù hợp của dữ liệu, sự cần thiết của việc sử dụng các tiêu chí đánh giá, đồng thời giới thiệu một số tiêu chí đơn giản phù hợp với cấp THCS. Ngoài ra, hai mô hình động trên GeoGebra cùng với hai tình huống dạy học được xây dựng để minh họa và hỗ trợ giảng dạy.*

**Từ khóa:** Thống kê và Xác suất; Geogebra; mô hình động; đánh giá sự phù hợp; Chương trình giáo dục phổ thông 2018

## 1. Giới thiệu

Môn Toán không chỉ mang tính logic và trừu tượng, mà còn đóng vai trò quan trọng trong việc phát triển năng lực chung và năng lực toán học cho học sinh (HS). Chương trình giáo dục phổ thông 2018 đã chú trọng vào việc tạo môi trường học tập thực tế, giúp HS tiếp cận kiến thức có tính ứng dụng cao, từ đó nâng cao hiệu quả dạy và học. Trong bối cảnh dữ liệu lớn và công nghệ thông tin phát triển nhanh chóng, kỹ năng thu thập, phân tích và diễn giải dữ liệu trở nên thiết yếu. Thống kê và xác suất, vì vậy, đã được công nhận là những mạch kiến thức quan trọng, giúp HS hiểu và áp dụng thông tin trong cuộc sống hàng ngày. Việc tích hợp thống kê và xác suất vào chương trình không chỉ nâng cao chất lượng học tập mà còn trang bị cho HS những kỹ năng cần thiết để thích ứng và thành công trong môi trường số hiện đại. Học sinh sẽ trở thành những công dân có khả năng phân tích thông tin, ra quyết định dựa trên dữ liệu, và tự tin hơn trong việc đối mặt với thách thức của thế giới ngày nay.

---

**Cite this article as:** Ton, T. T., Le, V. K., Le, B. Q. C., & Ly, T. N. (2026). Evaluating the appropriateness of data in teaching statistics at the lower secondary school level: illustrations using two Geogebra models. *Ho Chi Minh City University of Education Journal of Science*, 23(1), 176-188. [https://doi.org/10.54607/hcmue.js.23.1.4705\(2026\)](https://doi.org/10.54607/hcmue.js.23.1.4705(2026))

Nhằm nhấn mạnh vai trò của mạch kiến thức này, trong nội dung Chương trình giáo dục phổ thông môn Toán 2018 (Ministry of Education and Training, 2018, p.16), chỉ ra rằng:

Thống kê và Xác suất là một thành phần bắt buộc của giáo dục toán học trong nhà trường, góp phần tăng cường tính ứng dụng và giá trị thiết thực của giáo dục toán học. Thống kê và Xác suất tạo cho HS khả năng nhận thức và phân tích các thông tin được thể hiện dưới nhiều hình thức khác nhau, hiểu bản chất xác suất của nhiều sự phụ thuộc trong thực tế, hình thành sự hiểu biết về vai trò của thống kê như là một nguồn thông tin quan trọng về mặt xã hội, biết áp dụng tư duy thống kê để phân tích dữ liệu. Từ đó, nâng cao sự hiểu biết và phương pháp nghiên cứu thế giới hiện đại cho HS.

Mạch kiến thức về xác suất và thống kê được giới thiệu từ lớp 2 với những nội dung cơ bản, sau đó dần được nâng cao về mức độ khó hơn ở các lớp học tiếp theo. Trong chương trình Toán ở cấp trung học cơ sở, mạch kiến thức này tập trung vào các kỹ năng như thu thập dữ liệu, lập bảng tần số và tần suất (cho dữ liệu ghép lớp và không ghép lớp), vẽ biểu đồ thống kê, cũng như tính toán xác suất dựa trên số liệu thực nghiệm hoặc theo quan điểm cổ điển đối với các thí nghiệm có số hữu hạn các kết quả đồng khả năng. Nhiều nhà nghiên cứu đã dành sự chú ý và thực hiện các công trình liên quan đến nội dung, cách tiếp cận, phương pháp giảng dạy, cũng như các trở ngại và thách thức trong việc học và dạy mạch kiến thức này. Nguyen và Ho (2023) gợi ý sử dụng các hoạt động trải nghiệm để giảng dạy nội dung thống kê trong chương trình lớp 7. Trong khi đó, Nguyen và Le (2020) đề xuất các cách thức lựa chọn biểu diễn dữ liệu hiệu quả trên biểu đồ thống kê. Le và cộng sự (2024) đã phân tích các yếu tố ảnh hưởng đến niềm tin và thái độ của giáo viên toán trung học cơ sở về việc giảng dạy thống kê, qua đó nhận thấy nhiều giáo viên gặp khó khăn trong việc nắm bắt một số khái niệm và mong muốn được đào tạo thêm. Bên cạnh đó, nghiên cứu của Jamie (2007) về giáo viên toán tại bang Alabama, Mỹ, cho thấy rằng thống kê không chỉ là một môn học phức tạp và khó tiếp cận đối với học sinh mà còn đặt ra nhiều thách thức trong việc giảng dạy ở cấp trung học. Nguyen và Hoang (2021) cũng nhận định rằng việc trình bày mạch thống kê và xác suất gặp nhiều khó khăn, đặc biệt khi mạch này trở thành một trong ba nội dung trọng tâm của chương trình toán phổ thông. Ngoài ra, nghiên cứu của Garfield và Ben-Zvi (2007) đã đưa ra các nguyên tắc cơ bản trong giảng dạy thống kê và xác suất, nhấn mạnh sự cần thiết của việc hỗ trợ học sinh vượt qua trở ngại trong việc hiểu các khái niệm và sửa chữa những lỗi sai trong lập luận thống kê.

Sự phát triển công nghệ thông tin và các phần mềm toán học đã hỗ trợ hiệu quả cho giáo viên đổi mới phương pháp giảng dạy, tạo bài giảng sinh động, trực quan, khơi dậy hứng thú học tập cho HS. Trong nghiên cứu của Pratt et al. (2011), các tác giả đã nêu bật một số vai trò quan trọng của công nghệ thông tin (CNTT) trong việc giảng dạy và học tập thống kê cũng như xác suất. Các vai trò này bao gồm: hỗ trợ thực hiện các tính toán, khám phá dữ liệu thông qua biểu đồ, trực quan hóa các khái niệm phức tạp, tạo lập các thí nghiệm mô phỏng, hỗ trợ điều tra và thu thập dữ liệu, đồng thời cung cấp các môi trường đa dạng nhằm thúc đẩy sự tham gia và hợp tác giữa người học. Geogebra, một phần mềm toán học miễn phí với nhiều tính năng mạnh mẽ như vẽ hình, tính toán đại số, giải tích và xây dựng mô hình động, ngày càng được ưa chuộng bởi các giáo viên toán tại Việt Nam. Các sách giáo khoa toán phổ thông cũng đã dành một số trang để giới thiệu phần mềm này trong việc vẽ hình và tính toán cơ bản. Nhiều nghiên cứu đã trình bày ứng dụng Geogebra trong giảng dạy, đặc biệt ở chủ đề xác suất và thống kê, chẳng hạn như: Pham và Tran (2023); Nguyen và Quach (2023); Ton et al. (2024).

Trong thời đại số với sự bùng nổ thông tin, việc chú trọng đến dữ liệu và các phương pháp thống kê nhằm phân tích dữ liệu không chỉ giúp học sinh hiểu rõ hơn về thế giới xung quanh mà còn trang bị cho các em kỹ năng cần thiết để xử lý thông tin một cách hiệu quả. Đặc biệt, đánh giá tính hợp lý của dữ liệu ngày càng trở nên quan trọng, giúp đảm bảo chất lượng thông tin mà học sinh tiếp nhận. Giáo dục học sinh cách phân tích và đánh giá độ tin cậy của dữ liệu không chỉ nâng cao tư duy phản biện mà còn cải thiện khả năng ra quyết định của các em. Tuy nhiên, nội dung này vẫn còn khá mới mẻ đối với cả học sinh và giáo viên, được lồng ghép trong mạch kiến thức xác suất và thống kê ở bậc trung học cơ sở theo chương trình Giáo dục phổ thông 2018. Việc đánh giá sự phù hợp của dữ liệu trong nội dung thống kê trở thành một yêu cầu quan trọng trong chương trình môn Toán, đòi hỏi phải được nghiên cứu.

Bài báo này tập trung phân tích và tổng hợp những vấn đề liên quan đến việc đánh giá sự phù hợp của dữ liệu trong nội dung thống kê ở cấp Trung học cơ sở. Trên cơ sở đó, nhóm tác giả đề xuất xây dựng hai mô hình minh họa bằng phần mềm Geogebra để hỗ trợ dạy và học nội dung này, góp phần nâng cao năng lực phân tích dữ liệu của học sinh. Mô hình đầu tiên tập trung vào việc nhận diện sự tương thích giữa đơn vị đo, đại lượng cần đo và giá trị đo. Mô hình thứ hai hướng dẫn học sinh vận dụng các công cụ thống kê mô tả cơ bản để phát hiện các điểm dữ liệu bất thường và không hợp lý. Để minh họa cho tính khả thi, bài báo đồng thời xây dựng hai tình huống dạy học giả định nhằm cụ thể hóa phương án vận dụng các mô hình trên vào thực tiễn giảng dạy.

## 2. Nội dung

### 2.1. Vì sao phải đánh giá sự phù hợp của dữ liệu?

Việc trang bị cho HS kiến thức và kỹ năng đánh giá tính hợp lý của dữ liệu là một nhiệm vụ vô cùng quan trọng trong bối cảnh kỷ nguyên số hiện nay. Đây là thời đại mà dữ liệu không chỉ tràn ngập trong mọi lĩnh vực của đời sống mà còn được tiếp cận một cách dễ dàng qua các nền tảng trực tuyến và công cụ công nghệ. Một số lí do nổi bật có thể là:

- *Phát triển tư duy phản biện* (Leighton et al., 2021): HS cần biết không chỉ chấp nhận thông tin một chiều một cách mù quáng mà còn biết tự đặt câu hỏi, tìm kiếm các luận cứ và bằng chứng hỗ trợ. Kiến thức này giúp HS phân tích và đánh giá thông tin một cách cẩn thận, từ đó rèn luyện khả năng tư duy phản biện và ra quyết định hợp lý.

- *Nhận biết và hạn chế ảnh hưởng của thông tin sai lệch* (Zrnec et al., 2022): Trong thời đại công nghệ thông tin, sự bùng nổ thông tin dễ dẫn đến việc phát tán thông tin sai lệch. Học sinh cần biết cách xác định nguồn thông tin đáng tin cậy và phân tích tính chính xác của dữ liệu.

- *Rèn luyện kỹ năng giải quyết vấn đề* (Kjelvik & Schultheis, 2019): Khả năng đánh giá dữ liệu giúp HS tìm ra giải pháp cho các vấn đề phức tạp trong học tập và cuộc sống. Khi HS có khả năng phân tích và đánh giá thông tin, họ sẽ dễ dàng hơn trong việc tìm ra các phương án khả thi.

- *Chuẩn bị kỹ năng nghề nghiệp tương lai* (Taş, 2024): Nhiều lĩnh vực nghề nghiệp hiện nay, đặc biệt các ngành nghề liên quan đến khoa học dữ liệu, tài chính và y tế, yêu cầu khả năng phân tích và đánh giá dữ liệu. Việc trang bị kiến thức này từ sớm sẽ giúp HS dễ dàng thích ứng với các yêu cầu trong môi trường làm việc sau này.

- *Đưa ra các quyết định thông minh* (Khong et al., 2023): Việc có khả năng đánh giá dữ liệu giúp HS đưa ra quyết định sáng suốt hơn trong nhiều khía cạnh của cuộc sống, từ học tập, lựa chọn nghề nghiệp cho đến các vấn đề cá nhân. Người học sẽ không dễ dàng bị tác động bởi cảm xúc hay áp lực xã hội mà sẽ dựa trên các dữ liệu cụ thể và tính logic của các vấn đề.

## 2.2. Vấn đề đánh giá sự phù hợp của dữ liệu trong chương trình GDPT 2018

Phần kiến thức thống kê trong chương trình bậc trung học cơ sở không chỉ giúp HS tiếp cận với các khái niệm cơ bản mà còn trang bị cho các em những kỹ năng cần thiết để xử lý và phân tích dữ liệu trong thực tế. Các chủ đề về nội dung cụ thể thống kê ở bậc trung học cơ sở trong chương trình giáo dục phổ thông 2018 môn Toán bao gồm:

- *Thu thập, phân loại, biểu diễn dữ liệu theo các tiêu chí cho trước*: Học sinh sẽ được hướng dẫn cách thu thập dữ liệu từ nhiều nguồn khác nhau, bao gồm văn bản, bảng biểu, kiến thức trong các lĩnh vực giáo dục khác, phỏng vấn, internet... Việc phân loại dữ liệu theo các tiêu chí như đối tượng, thời gian, hoặc địa điểm giúp HS hiểu rõ hơn về cách tổ chức thông tin. Điều này không chỉ rèn luyện kỹ năng tư duy hệ thống mà còn giúp các em nhận diện các yếu tố có thể ảnh hưởng đến kết quả nghiên cứu.

- *Mô tả và biểu diễn dữ liệu trên các bảng, biểu đồ*: Biểu diễn dữ liệu một cách trực quan là một phần quan trọng trong thống kê. Học sinh sẽ học cách xây dựng các bảng tần số và tần suất để tổ chức thông tin một cách rõ ràng, cũng như vẽ một số loại biểu đồ thống kê như biểu đồ cột, biểu đồ tròn và biểu đồ đường. Qua đó, HS sẽ dần dần nắm vững cách thức tóm tắt nhanh chóng các thông tin từ dữ liệu thành các bảng biểu, biểu đồ trực quan, từ đó dễ dàng hơn trong việc phân tích và so sánh.

- *Hình thành và giải quyết vấn đề đơn giản xuất hiện từ các số liệu và biểu đồ thống kê đã có*: Từ kết quả phân tích dữ liệu thu được HS có thể phát hiện ra các vấn đề, các quy luật đơn giản hay các mối quan hệ thống kê giữa các yếu tố. Học sinh sẽ được khuyến khích tư duy phản biện bằng cách đặt câu hỏi và giải quyết các vấn đề thực tiễn từ các số liệu và biểu đồ. Ví dụ, khi phân tích một biểu đồ, HS có thể đặt ra các giả thuyết hoặc tìm kiếm nguyên nhân cho các xu hướng đã quan sát. Điều này không chỉ giúp các em phát triển kỹ năng phân tích mà còn tạo cơ hội cho việc thảo luận và hợp tác trong lớp học.

- *Bảng tần số, tần số tương đối và biểu đồ tương ứng*: HS sẽ được học cách lập bảng tần số và tần số tương đối. Việc sử dụng các biểu đồ sẽ giúp các em hình dung rõ ràng hơn về mối quan hệ giữa các đại lượng thông qua các giá trị được thu thập trong dữ liệu. Kỹ năng này đặc biệt quan trọng trong việc giúp HS phát hiện được sự phân bố cũng như các xu hướng trong một tập dữ liệu lớn.

Dữ liệu ngày nay được thu thập từ nhiều nguồn khác nhau dẫn đến việc xuất hiện các vấn đề về tính chính xác và độ tin cậy. Những sai sót trong quá trình thu thập và nhập liệu không chỉ ảnh hưởng đến chất lượng dữ liệu mà còn có thể dẫn đến những kết luận sai lệch trong các phân tích sau này. Do đó, việc giáo dục HS về cách nhận biết và đánh giá sự phù hợp của dữ liệu trở nên cực kỳ quan trọng. Điều này không chỉ giúp các em tránh được các lỗi sai trong phân tích mà còn phát triển tư duy phản biện và khả năng ra quyết định dựa trên thông tin chính xác. Khi HS được hướng dẫn để hiểu rõ cách phân loại và biểu diễn dữ liệu, các em sẽ có khả năng xác định được những điểm bất hợp lý hoặc những dữ liệu không đáng tin cậy, từ đó nâng cao chất lượng các phân tích và báo cáo.

Vấn đề về nhận biết, phát hiện và giải thích về sự hợp lý của dữ liệu được thể hiện trong yêu cầu cần đạt đối với nội dung “thu thập, phân loại và biểu diễn dữ liệu theo các tiêu chí cho trước” ở lớp 6, 7 và 8. Đối với lớp 9, yêu cầu này thể hiện thông qua nội dung về “mô tả và biểu diễn dữ liệu trên các bảng, biểu đồ”. Cụ thể như sau:

**Bảng 1. Yêu cầu cần đạt liên quan tính hợp lý của dữ liệu**

Lớp	Nội dung	Yêu cầu cần đạt liên quan
6	Thu thập,	Nhận biết được tính hợp lý của dữ liệu theo các tiêu chí đơn giản.
7	phân loại, biểu diễn dữ liệu theo các tiêu chí cho trước	Giải thích được tính hợp lý của dữ liệu theo các tiêu chí toán học đơn giản (ví dụ: tính hợp lý, tính đại diện của một kết luận trong phỏng vấn; tính hợp lý của các quảng cáo...).
8		Chứng tỏ được tính hợp lý của dữ liệu theo các tiêu chí toán học đơn giản (ví dụ: tính hợp lý trong các số liệu điều tra; tính hợp lý của các quảng cáo...).
9	Mô tả và biểu diễn dữ liệu trên các bảng, biểu đồ	Phát hiện và lí giải được số liệu không chính xác dựa trên mối liên hệ toán học đơn giản giữa các số liệu đã được biểu diễn trong những ví dụ đơn giản.

### 2.3. Các tiêu chí đơn giản đánh giá sự phù hợp

#### 2.3.1. Vì sao phải sử dụng tiêu chí để đánh giá?

Theo nghiên cứu nền tảng của Wang và Strong (1996) về các khía cạnh chất lượng dữ liệu, cũng như phương pháp luận được đề xuất bởi Pipino và cộng sự (2002), việc sử dụng một hệ thống các tiêu chí (hay còn gọi là các thứ nguyên chất lượng – quality dimensions) là bước không thể thiếu trong quy trình đánh giá chất lượng dữ liệu. Các học giả này nhất quán nhấn mạnh rằng khái niệm "chất lượng" phải được hiểu là "sự phù hợp cho mục đích sử dụng" (fitness for use). Để xác định được điều này, một bộ khung tiêu chí rõ ràng đóng vai trò như một thước đo khách quan. Điều này không chỉ giúp chúng ta đảm bảo tính chính xác của dữ liệu mà còn tăng cường tính khách quan, đảm bảo tính nhất quán, cải thiện chất lượng dữ liệu và tạo cơ sở cho phân tích chuyên sâu.

- *Đảm bảo tính chính xác*: Tiêu chí đánh giá dữ liệu giúp xác định xem dữ liệu có chính xác và đáng tin cậy hay không. Việc đánh giá theo các tiêu chí cụ thể sẽ giảm thiểu khả năng dựa vào các nguồn dữ liệu kém chất lượng hoặc sai lệch. Tính chính xác là yếu tố rất quan trọng trong bất kì nghiên cứu hoặc phân tích nào, vì nếu dữ liệu không chính xác, tất cả các quyết định được đưa ra từ dữ liệu đó cũng sẽ không đáng tin cậy.

- *Tăng cường tính khách quan*: Việc sử dụng các tiêu chí rõ ràng giúp quá trình đánh giá trở nên khách quan hơn. Điều này rất quan trọng vì việc đánh giá chủ quan hoặc thiên lệch có thể dẫn đến những kết luận không chính xác.

- *Đảm bảo tính nhất quán*: Tiêu chí đánh giá cũng cho phép so sánh dữ liệu từ các nguồn khác nhau một cách nhất quán. Khi các tiêu chí được thiết lập và áp dụng một cách nhất quán, các nhà phân tích có thể đưa ra các quyết định dựa trên dữ liệu một cách công bằng hơn, từ đó tạo ra sự tin tưởng trong kết quả phân tích.

- *Cải thiện chất lượng dữ liệu*: Khi có tiêu chí đánh giá rõ ràng, các tổ chức và cá nhân có thể cải thiện quy trình thu thập và xử lý dữ liệu. Việc này không chỉ giúp nâng cao chất lượng dữ liệu mà còn tối ưu hóa thời gian và nguồn lực. Các tiêu chí đánh giá thường dẫn đến việc phát hiện các điểm yếu trong quy trình thu thập dữ liệu, từ đó tạo điều kiện cho các biện pháp cải thiện phù hợp. Điều này góp phần tạo ra một hệ thống dữ liệu mạnh mẽ và đáng tin cậy hơn.

- *Tạo cơ sở cho phân tích chuyên sâu*: Việc đánh giá dữ liệu theo tiêu chí sẽ tạo ra một nền tảng vững chắc cho các bước phân tích chuyên sâu tiếp theo. Khi dữ liệu đã được kiểm

tra và đánh giá theo các tiêu chí rõ ràng, các nhà phân tích có thể tập trung vào việc rút ra các kết luận chính xác và có giá trị.

2.3.2. Một số tiêu chí đánh giá

Trong nghiên cứu của Pipino và cộng sự (2002), các tác giả đã chỉ ra 16 tiêu chí để đánh giá chất lượng dữ liệu. Dựa vào việc phân tích các ví dụ trong sách toán cấp THCS, chúng tôi đề xuất một số tiêu chí đơn giản giúp định hướng HS trong việc đánh giá sự phù hợp và tính hợp lí của dữ liệu thu được:

- *Độ chính xác*: Sự chính xác được thể hiện thông qua một số yếu tố như đơn vị đo hoặc định dạng phù hợp, giá trị đo nằm trong miền xác định hoặc phản ánh đúng thực tế (không quá lớn hoặc quá nhỏ)... Nếu thông tin sai lệch, nó có thể dẫn đến những quyết định không đúng. Chẳng hạn, khi cho một HS thu thập thông tin về tuổi của bố mẹ các bạn trong lớp nhưng có một phụ huynh được ghi 157 tuổi là không chính xác.

- *Tính đầy đủ*: Dữ liệu phải có đầy đủ thông tin cần thiết để trở thành mẫu đại diện cho tổng thể tham gia vào quá trình phân tích. Việc thiếu sót có thể gây ra hiểu nhầm. Chẳng hạn, trong lớp có 30 HS và một HS muốn biết môn học được các bạn yêu thích nhất nên đã khảo sát 20 HS để tìm hiểu. Kết luận của HS này có thể không chính xác vì còn thiếu 10 chưa khảo sát.

- *Sự nhất quán*: Dữ liệu từ các nguồn khác nhau phải không mâu thuẫn với nhau. Chẳng hạn, khi bạn đưa cho HS một số tài liệu để viết về lịch sử của nhà trường nhưng có 2 tài liệu ghi năm thành lập của nhà trường khác nhau. Lúc đó, HS cần phát hiện ra tính không hợp lí của dữ liệu vì thiếu sự nhất quán.

- *Tính cập nhật*: Dữ liệu cần phải được cập nhật để phù hợp với yêu cầu đặt ra trong vấn đề nghiên cứu. Chẳng hạn, khi lớp trưởng báo cáo về tình hình học tập của các bạn trong lớp trong tuần này ở tiết sinh hoạt cuối tuần nhưng lại sử dụng số liệu từ tuần trước thì sẽ không hợp lí.

- *Tính hợp lệ*: Dữ liệu phải tuân theo các quy tắc hoặc tiêu chuẩn đã được thiết lập. Chẳng hạn, khi giáo viên yêu cầu mỗi HS hãy ghi số đo 3 góc của một tam giác thì có bạn ghi  $30^0$ ,  $70^0$  và  $90^0$  là không hợp lí vì điều này trái với tính chất đã biết “Tổng 3 góc trong một tam giác bằng  $180^0$ ”.

Trong các tiêu chí nói trên, tiêu chí về “độ chính xác” và “tính hợp lệ” nhìn chung HS dễ dàng nắm bắt và tiếp cận hơn qua các ví dụ minh họa đơn giản về sự phù hợp của dữ liệu dựa vào các quy chuẩn và tính chất đã biết. Thông qua việc phân tích các kiến thức và ví dụ được trình bày trong các bộ sách Cánh diều, Chân trời sáng tạo và Kết nối tri thức, một số tiêu chí đơn giản trong đánh giá sự phù hợp của dữ liệu được tổng hợp lại như sau:

**Bảng 2.** Một số tình huống đánh giá về tính hợp lí của dữ liệu ở các bộ sách

Lớp	Tiêu chí	Nội dung thể hiện	Ví dụ
6	Độ chính xác	- Xếp các đối tượng khác loại vào cùng một hạng mục. - Giá trị không thực tế (quá lớn hoặc quá nhỏ)	- Liệt kê một số món ăn yêu thích: mì xào bò, gà quay, mực nướng, rượu vang, lẩu hải sản. - Số liệu thống kê số lượng HS của 5 lớp ở khối 6: 32, 30, 37, 87, 95.
	Tính hợp lệ	- Không phù hợp với tiêu chuẩn	- Đo nhiệt độ của một số bệnh nhân, ta được các số liệu (đơn vị độ C): 38, 39, 35, 80, -7.
7	Tính đầy đủ	- Dữ liệu được khảo sát phải có tính đại diện	- Đề khảo sát sự hài lòng của khách hàng trên chuyến bay nhưng hãng hàng không chỉ khảo sát những người đi ở hạng thương gia.
	Độ chính xác	- Giá trị thu được không thuộc miền xác định	- Lớp trưởng thông kê trong lớp có 40 bạn tham gia chạy việt dã trong khi sĩ số của lớp chỉ có 32 HS.

	- Đúng định dạng	- Thu thập dữ liệu email: vanduong08@gmail.com, thuhang_chu.vn, maiphuongpt@yahoo.com.
8	<p>Độ chính xác</p> <p>- Giá trị thuộc miền xác định</p> <p>- Giá trị không thực tế (quá lớn hoặc quá nhỏ)</p> <p>Tính hợp lệ</p> <p>- Tổng các giá trị thành phần phải bằng giá trị của toàn thể</p> <p>Tính đầy đủ</p> <p>- Dữ liệu khảo sát phải có tính đại diện</p>	<p>- Kết quả 5 bài kiểm tra môn Toán của bạn Dũng: 8, -6, 7, 5, 9.</p> <p>- Khảo sát 3 bạn trong lớp về số tivi có trong nhà của bạn, ta được kết quả: 3; 52; 2.</p> <p>- Lớp 3C có 30 HS được chia làm 3 tổ với số HS tương ứng mỗi tổ là 10, 10, 9.</p> <p>- Công ty may đang thiết kế đồng phục cho học sinh THCS nên đã khảo sát ý kiến của 60 HS lớp 6.</p>
9	<p>Độ chính xác</p> <p>- Các giá trị đo của một tiêu chí có đơn vị khác nhau</p> <p>Tính hợp lệ</p> <p>- Tổng các tỉ lệ xuất hiện các giá trị trong mẫu phải bằng 1</p>	<p>- Thống kê tuổi thọ trung bình của người dân Indonesia, Myanmar, Thái Lan, Việt Nam lần lượt: 71,3 năm; 69,1 năm; 932,4 tháng; 75,4 năm. Hãy vẽ biểu đồ cột biểu diễn.</p> <p>- Vẽ biểu đồ hình tròn với các giá trị tỉ lệ ghi ở phần chú thích trên biểu đồ có tổng nhỏ hơn 1.</p>

Nhìn chung, việc đánh giá sự hợp lí của dữ liệu ở sách giáo khoa lớp 6, 7 và 8 được thể hiện rõ nét thông qua các ví dụ và bài tập. Tuy nhiên, đối với lớp 9 phần kiến thức này còn khá mờ nhạt, đặc biệt chưa khai thác và ứng dụng các kiến thức đã học như bảng biểu (bảng tần số, tần suất) hay biểu đồ để hỗ trợ đánh giá sự hợp lí của dữ liệu.

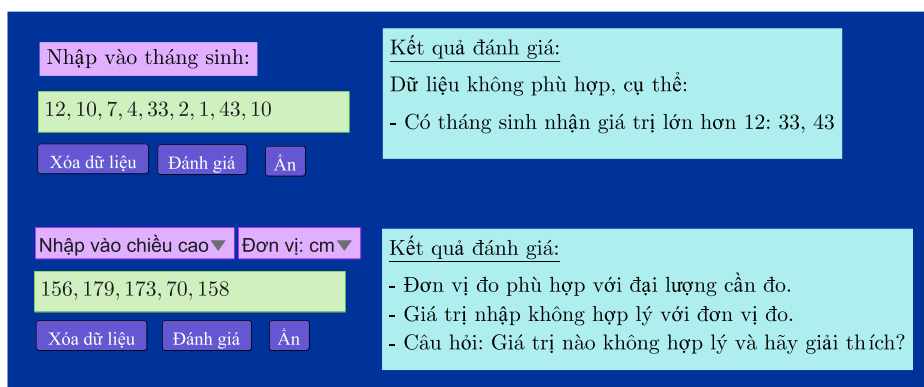
#### 2.4. Các mô hình đề xuất

Tiếp theo, chúng tôi xây dựng một số mô hình động trên phần mềm Geogebra nhằm hỗ trợ cho các giáo viên dạy học về nội dung sự phù hợp của dữ liệu trong phần kiến thức thống kê ở bậc trung học cơ sở. Thông qua việc áp dụng Geogebra, giáo viên có thể tạo ra các tình huống học tập phong phú, cho phép HS tương tác trực tiếp với các mô hình, quan sát sự thay đổi và đưa ra các đánh giá. Điều này không chỉ giúp HS dễ dàng tiếp thu các tiêu chí đơn giản về sự phù hợp của dữ liệu mà còn phát triển kĩ năng phân tích, đánh giá và giải quyết vấn đề. Đồng thời, mô hình động giúp giảm bớt sự khô khan trong việc học lí thuyết, tạo ra một môi trường học tập sinh động và đầy hứng khởi.

##### 2.4.1. Mô hình thứ nhất

Trong mô hình đầu tiên này, chúng tôi tập trung vào việc đánh giá sự phù hợp của dữ liệu qua các yếu tố chính: tính hợp lí của tháng sinh, của chiều cao và cân nặng so với đơn vị đo và các giá trị đo được. Mô hình cho phép người dùng nhập dữ liệu vào một giao diện tương tác và tùy chỉnh, từ đó có thể tạo ra các tình huống thực tế khác nhau để kiểm tra tính hợp lí của các giá trị dữ liệu. Việc này giúp người học và giáo viên dễ dàng thử nghiệm và phân tích các tình huống thống kê trong những bối cảnh cụ thể, đồng thời nâng cao khả năng tư duy phản biện và phân tích dữ liệu.

Khi dữ liệu được nhập vào mô hình, người dùng có thể nhấn nút “Đánh giá” để hệ thống tự động kiểm tra tính hợp lí của các giá trị dữ liệu. Với phần ví dụ về tháng sinh, mô hình sẽ đưa ra kết quả đánh giá rất rõ ràng và chi tiết. Nếu dữ liệu tháng sinh không hợp lí, chẳng hạn như một giá trị tháng sinh vượt quá phạm vi từ 1 đến 12, mô hình sẽ cung cấp lời giải thích rõ ràng về lí do tại sao giá trị đó không hợp lệ. Điều này giúp HS dễ dàng nhận diện lỗi và hiểu được nguyên lí đằng sau việc xác định sự phù hợp của dữ liệu theo các tiêu chuẩn cơ bản.

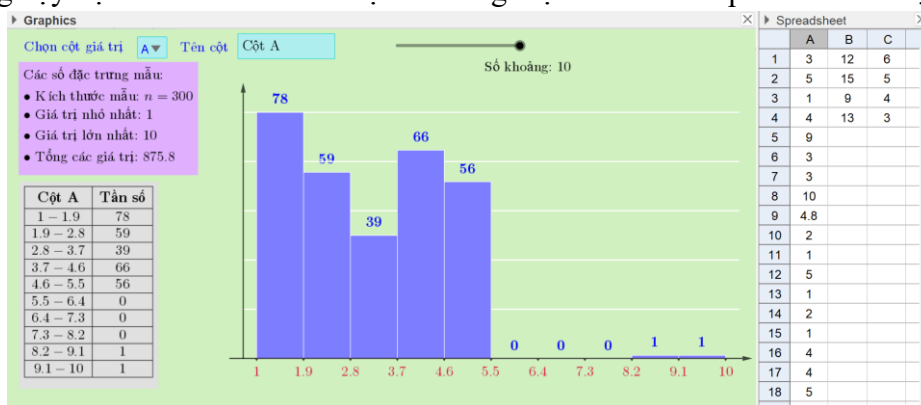


**Hình 1.** Mô hình về tháng sinh, chiều cao và cân nặng

Tuy nhiên, đối với phần đánh giá sự phù hợp của chiều cao và cân nặng, kết quả đánh giá là chung chung và một câu hỏi sẽ được đưa ra yêu cầu người dùng tự xác định giá trị nào có thể không phù hợp, đồng thời yêu cầu giải thích tại sao giá trị đó có thể không chính xác. Chẳng hạn, nếu một giá trị chiều cao hay cân nặng có sự khác biệt quá lớn so với mức trung bình hoặc không khớp với đơn vị đo chuẩn. Điều này không chỉ giúp người học kiểm tra lại tính hợp lý của các giá trị đo lường mà còn khuyến khích các em phát triển khả năng tự phân tích và đánh giá dữ liệu.

**2.4.2. Mô hình thứ hai**

Đối với mô hình này, giáo viên có thể nhập giá trị vào các cột A, B hoặc C trong bảng tính (Spreadsheet), và trong phần giao diện, giáo viên được phép chọn lấy dữ liệu từ cột A, B, hay C tương ứng. Các thông tin về các số đặc trưng mẫu đơn giản như kích thước mẫu, giá trị nhỏ nhất và lớn nhất sẽ được liệt kê để giúp giáo viên và HS có cái nhìn tổng quát về dữ liệu. Ngoài ra, bảng tần số dạng ghép lớp cũng sẽ được hiển thị, cho phép giáo viên dễ dàng phân tích phân bố của dữ liệu. Tên của các cột giá trị cũng được phép thay đổi linh hoạt, sao cho phù hợp với dữ liệu nhập vào, nhằm tối ưu hóa việc sử dụng và trình bày thông tin. Việc cho phép lấy dữ liệu từ 3 cột cũng như thay đổi số khoảng (bảng thanh trượt) trong bảng tần số ghép lớp sẽ giúp tăng tính linh hoạt của mô hình, tạo điều kiện để giáo viên có thể tạo ra nhiều tình huống dạy học khác nhau và hỗ trợ HS trong việc tìm hiểu và phân tích dữ liệu.



**Hình 2.** Đánh giá sự phù hợp thông qua thông tin từ thống kê mô tả

Mô hình này được xây dựng với mục đích hỗ trợ HS có thể phát hiện ra các điểm không phù hợp trong dữ liệu thông qua các tiêu chí đơn giản. Đặc biệt, khi dữ liệu tương đối nhiều,

việc rà soát từng giá trị để đánh giá sự phù hợp là điều rất khó thực hiện một cách thủ công. Do đó, việc sử dụng các phần mềm có chức năng tính toán thống kê để tóm tắt dữ liệu thông qua các thống kê mô tả là điều cần thiết, giúp hỗ trợ quá trình đánh giá dữ liệu. Ở cấp trung học cơ sở, HS đã được học về bảng tần số, tần suất và một số biểu đồ thống kê cơ bản. Vì vậy, trong mô hình này, chúng tôi lựa chọn sử dụng các công cụ này (bảng tần số và biểu đồ cột) đi kèm với một số chỉ số thống kê (giá trị lớn nhất, nhỏ nhất và tổng các giá trị) để giúp HS đánh giá dữ liệu một cách dễ dàng. Giáo viên có thể nhập dữ liệu và tạo ra các tình huống học tập cụ thể, sau đó yêu cầu HS quan sát, phân tích và đưa ra nhận xét dựa trên những kết quả tóm tắt và biểu đồ thu được, từ đó đánh giá sự phù hợp của dữ liệu nhập vào.

### 2.5. Một số tình huống dạy học

Sau đây là một số tình huống dạy học mà giáo viên có thể thiết kế với sự hỗ trợ của phần mềm Geogebra và hai mô hình động được xây dựng ở trên. Các mô hình này có thể chạy trực tiếp trên máy tính cá nhân của giáo viên hoặc giáo viên có thể thực hiện online bằng cách truy cập vào địa chỉ: <https://www.geogebra.org/m/x34v3wgr>.

#### 2.5.1. Tình huống 1

Giáo viên cung cấp một bảng dữ liệu của 20 HS về tháng sinh, chiều cao (cm) và cân nặng (kg), ngoài ra dữ liệu này chứa một vài giá trị không hợp lý (chẳng hạn, tháng sinh là 15, chiều cao là 250cm). Sau đó, GV hướng dẫn HS nhập dữ liệu vào mô hình trên Geogebra, thực hiện đánh giá tính hợp lý rồi đưa ra giải thích cho kết quả thu được.

**Bảng 3. Số liệu tháng sinh, chiều cao (cm), cân nặng (kg) của 20 HS**

Tháng sinh	11	12	4	5	6	2	4	2	1	7
	8	9	10	10	8	3	11	15	11	3
Chiều cao (cm)	145	160	151	165	150	151	160	155	160	155
	152	172	160	154	148	145	250	158	153	163
Cân nặng (kg)	57	52	65	55	47	35	537	35	56	62
	57	70	46	41	-57	58	35	40	59	55

**Mục tiêu:** HS bước đầu làm quen với việc đánh giá tính hợp lý của dữ liệu, giải thích lý do nếu dữ liệu không hợp lý.

GV: Mở phần mềm Geogebra, hướng dẫn và hỗ trợ HS nhập các dữ liệu khảo sát được lên phần mềm Geogebra cho phần dữ liệu tháng sinh và chiều cao (với đơn vị đo: cm).

GV: Hãy nhấn nút “Đánh giá” và cho biết dữ liệu về tháng sinh có hợp lý hay không và vì sao?

**Nhập vào tháng sinh:**

11, 12, 4, 5, 6, 2, 4, 2, 1, 7, 8, 9, 10, 1

Xóa dữ liệu   Đánh giá   Ấn

**Kết quả đánh giá:**

Dữ liệu không phù hợp, cụ thể:

- Có tháng sinh nhận giá trị lớn hơn 12: 15

**Nhập vào chiều cao**   Đơn vị: cm

145, 160, 151, 165, 150, 151, 160, 1

Xóa dữ liệu   Đánh giá   Ấn

**Kết quả đánh giá:**

- Đơn vị đo phù hợp với đại lượng cần đo.
- Giá trị nhập không hợp lý với đơn vị đo.
- Câu hỏi: Giá trị nào không hợp lý và hãy giải thích?

**Hình 3. Đánh giá tính hợp lý của dữ liệu về 20 học sinh**

HS (Câu trả lời mong đợi): Trong dữ liệu tháng sinh có giá trị 15 là không hợp lí vì giá trị này lớn hơn 12.

GV: Kết quả đánh giá về dữ liệu chiều cao trong mô hình là không hợp lí. Vậy, nhìn lại bảng dữ liệu đã cho, giá trị nào không hợp lí và vì sao?

HS (Câu trả lời mong đợi): Giá trị 250cm là không hợp lí vì chiều cao này là quá cao với HS nói riêng và con người nói chung.

GV: Tiếp theo, GV yêu cầu HS nhập dữ liệu cân nặng (với đơn vị đo: kg) vào mô hình và yêu cầu HS giải đáp câu hỏi ở phần đánh giá kết quả trong mô hình.

GV: Nếu ta có bảng số liệu cho cả nghìn HS của trường thì việc đánh giá tính hợp lí của dữ liệu bằng cách quan sát từng giá trị trong bảng sẽ gặp khó khăn gì?

HS (Câu trả lời mong đợi): Việc làm này sẽ mất thời gian và có thể vô tình bị bỏ sót giá trị cần đánh giá.

GV: Đúng như vậy. Ngoài việc mất thời gian, nếu quan sát bằng mắt thường, chúng ta đôi lúc sẽ bị sót các giá trị không hợp lí, dẫn đến dữ liệu không được đánh giá chính xác. Do đó, để hỗ trợ trong việc đánh giá tính hợp lí ta có thể sử dụng các thông tin bổ sung từ các chỉ số thống kê đơn giản và các biểu diễn của dữ liệu (bảng và biểu đồ).

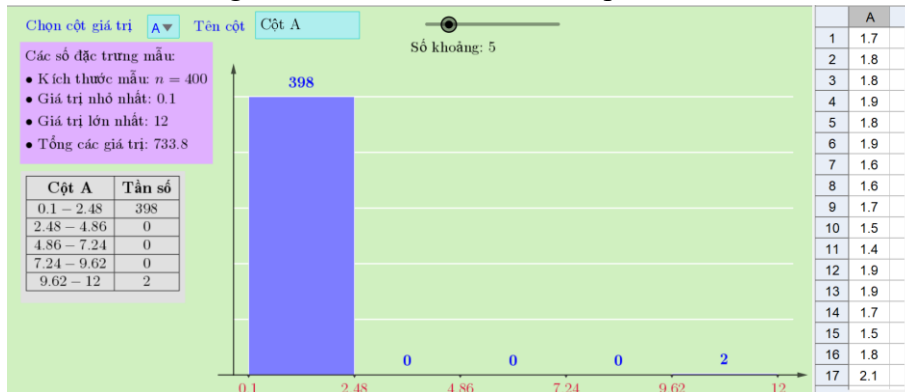
2.5.2. Tình huống 2

Giáo viên chuẩn bị một bộ dữ liệu về lượng nước (lít) mà 400 HS uống mỗi ngày, trong dữ liệu có một số giá trị bất hợp lí (chẳng hạn, có 4 số liệu là 0.1 lít, 0.2 lít, 10 lít và 12 lít). Dữ liệu này chứa trong file Excel ở cột A. Sau đó HS nhập dữ liệu vào mô hình Geogebra và mô hình sẽ hiển thị bảng tần số rồi quan sát và biểu đồ cột cùng một số chỉ số đặc trưng thống kê đơn giản. GV yêu cầu HS đánh giá sự hợp lí và giải thích nếu phát hiện dữ liệu không hợp lí.

**Bảng 4.** Số liệu về lượng nước (lít) uống mỗi ngày của 50 học sinh đầu tiên

1.7	1.8	1.8	1.9	1.8	1.9	1.6	1.6	1.7	1.5
1.4	1.9	1.9	1.7	1.5	1.8	2.1	1.8	2	1.5
1.9	1.6	1.9	1.9	2	1.9	2	1.8	1.8	2.2
1.6	1.8	2	1.9	1.8	1.5	1.8	2	1.8	1.9
1.7	1.9	1.9	1.6	1.7	1.5	1.7	1.8	1.6	1.8

**Mục tiêu.** Học sinh biết cách sử dụng các chỉ số thống kê, bảng tần số và biểu đồ để phát hiện các điều bất thường và từ đó kết luận về tính hợp lí của dữ liệu.



**Hình 4.** Phân tích thống kê mô tả với số khoảng chia bằng 5

GV: Mở phần mềm Geogebra, hướng dẫn và hỗ trợ HS sao chép dữ liệu từ file Excel đã chuẩn bị và dán dữ liệu vào cột A của cửa sổ Spreadsheet trong mô hình trên Geogebra (xem Hình 4).

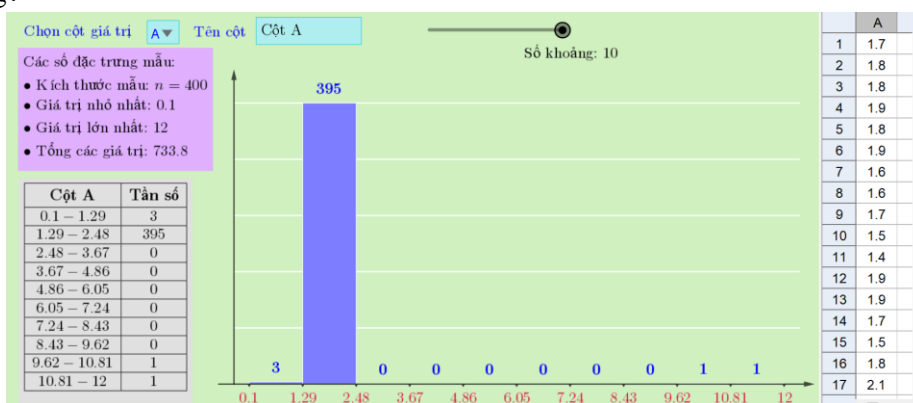
GV: Kéo số khoảng chia bằng 5, rồi đưa câu hỏi: “Từ biểu đồ, các em có phát hiện được điều gì bất thường hay không?”

HS (Câu trả lời mong đợi): Có hai giá trị rơi vào khoảng từ 9.62 đến 12. Điều này là không hợp lí vì một HS không thể uống nhiều nước đến vậy trong một ngày.

GV: Nếu nhìn vào bảng tần số ghép lớp, các em có phát hiện được điều gì bất thường hay không?

HS (Câu trả lời mong đợi): Phát hiện tình huống tương tự như khi nhìn biểu đồ.

GV: Kéo số khoảng chia bằng 10, rồi đưa câu hỏi: “Bây giờ, các em có phát hiện gì mới hơn không?”



Hình 5. Phân tích thống kê mô tả với số khoảng chia bằng 10

HS (Câu trả lời mong đợi): Có hai giá trị lớn bất thường (một giá trị rơi vào khoảng từ 9.62 đến 10.81, một giá trị rơi vào khoảng từ 10.81 đến 12). Ngoài ra, có 3 giá trị rơi vào khoảng từ 0.1 đến 1.29.

GV: Ba giá trị rơi vào khoảng từ 0.1 đến 1.29 có bất thường không?

HS (Câu trả lời mong đợi): Khó kết luận được.

GV: Bảng tần số ghép lớp và biểu đồ ở trên đều có thể cho ta các thông tin như các bạn đã phát hiện, tuy nhiên công cụ nào giúp ta dễ phát hiện điều bất thường hơn trong tình huống này? Vì sao?

HS (Câu trả lời mong đợi): Biểu đồ, vì nó trực quan hơn nên dễ nhận biết.

GV: Nếu nhìn thêm vào các số đặc trưng mẫu, các em có phát hiện thêm được điều gì bất thường không? Vì sao?

HS (Câu trả lời mong đợi): Giá trị nhỏ nhất (0.1 lít) và giá trị lớn nhất (12 lít) trong mẫu là không hợp lí. Lượng nước uống chỉ có 0.1 lít, không đủ để đáp ứng nhu cầu cơ bản của con người. Lượng nước uống là 12 lít vượt quá mức bình thường và không thực tế. Như thế, trong 3 giá trị trong khoảng 0.1 đến 1.29 có ít nhất một giá trị bất thường.

GV: Việc kết hợp với các công cụ của thống kê mô tả (như các số đặc trưng, bảng tần số hay biểu đồ) có thể hỗ trợ chúng ta trong việc phát hiện các bất thường, từ đó giúp ta ra quyết định về sự hợp lí của dữ liệu hiệu quả hơn.

### 3. Kết luận và kiến nghị

Bài báo này tập trung phân tích và đánh giá tính phù hợp của dữ liệu trong các bài toán thống kê ở chương trình trung học cơ sở, đồng thời xây dựng hai mô hình minh họa trên phần mềm Geogebra để hỗ trợ giảng dạy. Các mô hình được thiết kế theo nguyên tắc “mở” với giao diện thân thiện giúp giáo viên dễ dàng tạo ra các tình huống giảng dạy phong phú. Những mô

hình này được nhóm tác giả chia sẻ miễn phí, giáo viên có thể tải về để chạy trên máy tính cá nhân, chạy trực tiếp trên trình duyệt web hoặc thậm chí trên điện thoại di động. Việc sử dụng thống kê mô tả để phát hiện dấu hiệu bất thường và đánh giá tính hợp lý của dữ liệu giúp học sinh nhận thấy giá trị của các công cụ toán học, đặc biệt là thống kê, trong việc ra quyết định. Công nghệ thông tin, đặc biệt các phần mềm như Geogebra, đóng vai trò quan trọng trong việc hỗ trợ giáo viên thiết kế và trình bày mô hình trực quan, sinh động. Những công cụ này không chỉ giúp tạo ra các tình huống học tập đa dạng nhanh chóng mà còn giúp học sinh tiếp cận kiến thức dễ dàng, từ đó nâng cao khả năng phân tích và nhận diện dữ liệu bất thường. Việc kết hợp công nghệ vào giảng dạy kì vọng khơi dậy hứng thú học toán ở học sinh, đồng thời phát triển tư duy phân tích và kĩ năng xử lí dữ liệu hiệu quả.

❖ **Tuyên bố về quyền lợi:** Các tác giả xác nhận hoàn toàn không có xung đột về quyền lợi.

#### TÀI LIỆU THAM KHẢO

- Garfield, J., & Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, 75(3), 372-396. <https://doi.org/10.1111/j.1751-5823.2007.00029.x>
- Jamie, D. M. (2007). Teacher perceptions and attitudes about teaching statistics in P-12 education. *Educational Research Quarterly*, 30(4).
- Khong, I., Yusuf, N. A., Nuriman, A., & Yadila, A. B. (2023). Exploring the impact of data quality on decision-making processes in information intensive organizations. *APTISI Transactions on Management (ATM)*, 7(3), 246–253. <https://doi.org/10.33050/atm.v7i3.2138>
- Kjelvik, M. K., & Schultheis, E. H. (2019). Getting Messy with Authentic Data: Exploring the Potential of Using Data from Scientific Research to Support Student Data Literacy. *CBE—Life Sciences Education*, 18(2). <https://doi.org/10.1187/cbe.18-02-0023>
- Le, T. B. T. T., Tang, M. D., & Tran, M. M. (2024). Some factors of teacher beliefs and attitudes of middle school mathematics teachers in Bac Lieu province about teaching statistics. *Dong Thap University Journal of Science*, 13(02S), 52-63. <https://doi.org/10.52714/dthu.13.02S.2024.1344>
- Leighton, J. P., Cui, Y., & Cutumisu, M. (2021). Key Information Processes for Thinking Critically in Data-Rich Environments. *Frontiers in Education*, 6. <https://doi.org/10.3389/educ.2021.561847>
- Ministry of Education and Training. (2018). *Chuong trình giao duc pho thong mon Toan [General education program in Mathematics]*. Hanoi.
- Nguyen, C.T., & Hoang, L.M. (2021). Nhung dinh huong co ban ve day hoc noi dung thong ke cap trung hoc pho thong trong chuong trình mon toan 2018 [Basic directions on teaching the content of statistics at high school in the mathematics education curriculum 2018]. *Vinh University Journal of Science*, 50(4B), 67-75.
- Nguyen, P. T., & Ho, N. N. L. (2023). Day hoc noi dung “Thong ke” (Toan 7) thong qua hoat dong trai nghiem [Teaching the content "Statistics" (Grade 7 Math) through experiential activities]. *Vietnam Journal of Education*, 23(21), 1–6.
- Nguyen, T. T. T., & Quach, T. S. (2023). Mot so bien phap day hoc chu de Thong ke va Xac suat cho hoc sinh lop 6 voi su ho tro cua cong nghe thong tin [Some methods of teaching statistics and probability topics for 6th grade students with the support of information technology]. *The Vietnam Institute of Educational Sciences*, 19, 27-34.

- Pham, T. Q., & Tran, T. (2023). Ung dung cong nghe thong tin trong mot so tinh huong day hoc mach kien thuc thong ke va xac suat o trung hoc pho thong [Application of information technology in some situations of teaching statistics and probability in high school]. *Vietnam Journal of Education*, 23(5), 7-11.
- Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211-218.
- Pratt, D., Davies, N., & Connor, D. (2011). *The role of technology in teaching and learning statistics*. Teaching statistics in school mathematics-challenges for teaching and teacher education: A Joint ICMI/IASE Study: The 18th ICMI Study (pp. 97-107): Springer. [http://dx.doi.org/10.1007/978-94-007-1131-0\\_13](http://dx.doi.org/10.1007/978-94-007-1131-0_13)
- Taş, E. (2024). Data literacy education through university-industry collaboration. *Information and Learning Sciences*, 125(5-6), 389–405. <https://doi.org/10.1108/ILS-06-2023-0077>
- Ton, T.T., Hoang, T.T.T & Nguyen, D.N. (2024). Mot so mo hinh ho tro day hoc khai niem ngau nhien va y tuong do luong xac suat [Some models support teaching the concept of randomness and measurement of probability]. *Ho Chi Minh City University of Education Journal of Science*, 21(2), 245-255.
- Wang, R. Y., & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), 5–33. <https://doi.org/10.1080/07421222.1996.11518099>
- Zrnec, A., Poženel, M., & Lavbič, D. (2022). Users' ability to perceive misinformation: An information quality assessment approach. *Information Processing & Management*, 59(1). <https://doi.org/10.1016/j.ipm.2021.102739>

**EVALUATING THE APPROPRIATENESS OF DATA IN TEACHING STATISTICS  
AT THE LOWER SECONDARY SCHOOL LEVEL:  
ILLUSTRATIONS USING TWO GEOGEBRA MODELS**

**Ton That Tu\*, Le Vu Khoa, Le Bui Quynh Chi, Ly Thanh Nhan**

*The University of Danang – University of Science and Education, Vietnam*

*\*Corresponding author: Ton That Tu – Email: ttu@ued.udn.vn*

*Received: February 02, 2025; Revised: September 10, 2025; Accepted: October 08, 2025*

**ABSTRACT**

*Statistics and Probability is one of the three core areas of knowledge in the 2018 General Education Curriculum for Mathematics, taught from grade 2 to grade 12, with concepts and skills progressing from basic to advanced levels. At the secondary school level, the focus of statistics is primarily on data collection, classification, and representation through tables and charts. Collecting data from various sources may affect the quality of the data, which in turn influences analysis and decision-making. Therefore, evaluating the appropriateness of data has become an important requirement, which is reflected in the curriculum's learning objectives. This paper analyzes the reasons for evaluating data appropriateness and the necessity of using assessment criteria, while also introducing some simple criteria suitable for the lower secondary level. In addition, two dynamic GeoGebra models and two teaching situations are developed to illustrate and support instruction.*

**Keywords:** Statistics and Probability; Geogebra; Dynamic Models; Data Suitability Assessment; 2018 General Education Curriculum