

THIẾT KẾ DỰ ÁN HỌC TẬP THỐNG KÊ THEO TIẾP CẬN PHÂN TÍCH DỮ LIỆU KHÁM PHÁ TRONG DẠY HỌC TOÁN Ở TRƯỜNG TRUNG HỌC PHỔ THÔNG

DESIGNING A STATISTICS LEARNING PROJECT THROUGH THE EXPLORATORY DATA ANALYSIS
APPROACH IN HIGH SCHOOL MATHEMATICS TEACHING

Phạm Thế Quân^{1,+},
Nguyễn Thị Thu Huyền²

¹Trường Đại học Sư phạm Hà Nội 2; ²Trường Đại học Thủ đô Hà Nội
+ Tác giả liên hệ • Email: phamthequan@hpu2.edu.vn

Article history

Received: 27/01/2026

Accepted: 23/3/2026

Published: 29/4/2026

Keywords

Statistics teaching,
exploratory data analysis,
project-based learning,
statistical reasoning

ABSTRACT

In a context where data are increasingly prevalent across all areas of life, the ability to read, interpret, and analyze data has become an important requirement for learners. Accordingly, equipping students with statistical knowledge and developing data-related skills is of practical significance in general education. This paper presents the theoretical foundations and proposes a procedure for designing project-based learning activities in the teaching of statistics at the high school level based on the Exploratory Data Analysis (EDA) approach. The EDA approach emphasizes the roles of data exploration, visualization, and interpretation. Based on an analysis of the characteristics of EDA and project-based learning, the paper proposes a five-step procedure for designing learning projects, including: (1) identifying objectives; (2) selecting content and data; (3) designing learning tasks; (4) developing an assessment plan; and (5) piloting and revising. The paper also presents an illustrative example designed for Grade 10 students to demonstrate the feasibility of the proposed procedure. The findings contribute to operationalizing the data exploration approach in teaching activities suited to the current context of mathematics education in Vietnam.

1. Mở đầu

Trong bối cảnh hiện nay, dữ liệu ngày càng xuất hiện phổ biến và có tác động đáng kể đến nhiều lĩnh vực của đời sống xã hội. Khả năng hiểu, phân tích và sử dụng dữ liệu - thường được gọi là năng lực xử lý và phân tích dữ liệu, đang trở thành một yêu cầu quan trọng giúp mỗi cá nhân thích ứng với xã hội hiện đại. Frank và cộng sự (2016) cho rằng đây là năng lực cơ bản, có thể so sánh với khả năng biết đọc biết viết truyền thống. Những phân tích tổng quan gần đây cho thấy xu hướng nghiên cứu về năng lực xử lý và phân tích dữ liệu đang gia tăng, đồng thời gợi ý rằng việc tích hợp các kỹ năng này vào chương trình học phổ thông là cần thiết để chuẩn bị hành trang cho người học (Bowler và Shaw, 2024; Khalid và Mason, 2024).

Trong dạy học thống kê (TK), khả năng làm việc với dữ liệu được thể hiện rõ nét thông qua các hoạt động thu thập, phân tích và diễn giải dữ liệu thực tế. Tiếp cận phân tích dữ liệu khám phá (Exploratory Data Analysis - EDA) được quan tâm nhờ khả năng thúc đẩy quá trình tìm tòi, trực quan hóa và phát hiện quy luật từ dữ liệu. Bên cạnh đó, dạy học theo dự án (DHTDA) được nhiều nghiên cứu trong và ngoài nước ghi nhận là có khả năng góp phần phát triển năng lực, cải thiện kết quả và thái độ học tập của HS đối với môn Toán (Holmes và Hwang, 2016; Kokotsaki và cộng sự, 2016; Trần Thị Hà Phương, 2018; Nguyễn Thị Thu Thảo và Trần Việt Cường, 2019). Tuy nhiên, việc dạy học TK ở trường phổ thông hiện nay vẫn chủ yếu thiên về các thao tác tính toán trên dữ liệu giả định, trong khi các hoạt động khám phá dữ liệu thực tế mang tính mở chưa được triển khai thường xuyên. Mặc dù DHTDA đã được nghiên cứu áp dụng, nhưng các công trình tập trung cụ thể vào chủ đề TK với định hướng EDA cho đối tượng HS THPT còn khá hạn chế. Hiện nay, vẫn thiếu các khung hướng dẫn hoặc quy trình thiết kế dự án học tập TK mang tính hệ thống, phù hợp với đối tượng HS trung học phổ thông. Thực trạng này cho thấy sự tồn tại của một khoảng trống nghiên cứu cần được tiếp tục làm rõ nhằm hỗ trợ GV trong quá trình đổi mới phương pháp dạy học.

Xuất phát từ thực tiễn đó, bài báo này trả lời câu hỏi về cách thiết kế các dự án học tập TK theo tiếp cận EDA một cách hệ thống và khả thi cho HS THPT. Thông qua việc đề xuất các nguyên tắc sư phạm và quy trình thiết kế gồm 05 bước, bài báo góp phần xây dựng khung định hướng kết nối lý thuyết EDA với DHTDA, đồng thời hỗ trợ

GV tổ chức dạy học TK gắn với thực tiễn theo định hướng của Chương trình giáo dục phổ thông 2018. Nghiên cứu được thực hiện theo định hướng nghiên cứu lí luận kết hợp với thiết kế minh họa. Quá trình nghiên cứu bao gồm: (1) phân tích, tổng hợp tài liệu để xây dựng các nguyên tắc và quy trình thiết kế; (2) xây dựng dự án minh họa; (3) thử nghiệm dự án ở quy mô nhỏ. Dữ liệu định tính được thu thập thông qua quan sát giờ học và phân tích sản phẩm học tập của HS nhằm bước đầu nhận diện các khó khăn trong quá trình triển khai, qua đó điều chỉnh và hoàn thiện quy trình thiết kế.

2. Kết quả nghiên cứu

2.1. Một số vấn đề lí luận

2.1.1. Phân tích dữ liệu khám phá

Khái niệm “phân tích dữ liệu khám phá” lần đầu tiên được John Tukey giới thiệu vào năm 1977 như một hướng tiếp cận nhằm tìm hiểu dữ liệu để phát hiện các quy luật chung và hình thành các giả thuyết, thay vì chỉ tập trung vào việc kiểm định các giả thuyết đã được xác định trước. Tukey (1977) lập luận rằng thay vì chỉ dựa vào một số ít các giả thuyết có sẵn và áp dụng các phương pháp TK để kiểm chứng, nhà nghiên cứu nên bắt đầu bằng việc “thăm hiểm” dữ liệu để tìm kiếm những hiểu biết mới. Cách tiếp cận này tạo ra sự phân biệt rõ ràng giữa EDA và phân tích dữ liệu xác nhận (Confirmatory Data Analysis - CDA). Trong khi CDA thiên về tính chặt chẽ của suy diễn và kiểm định, thì EDA lại nhấn mạnh vào tính linh hoạt, trực quan và khả năng phát hiện những điều bất thường hoặc các xu hướng tiềm ẩn trong tập dữ liệu.

Về mặt phương pháp luận, EDA khuyến khích việc sử dụng các biểu đồ và các số đặc trưng TK để tóm tắt dữ liệu. Tukey (1977) đề xuất sử dụng các số sau để tóm tắt dữ liệu: giá trị lớn nhất, nhỏ nhất, trung vị và các tứ phân vị. Các dạng biểu đồ đặc trưng của EDA như biểu đồ hộp, biểu đồ thân và lá hay biểu đồ phân tán. Theo Batanero và Borovenik (2016), EDA thúc đẩy thái độ phân tích tích cực: tìm kiếm các quy luật nội tại, phát hiện các sai lệch so với quy luật đó và diễn giải các phát hiện trong ngữ cảnh cụ thể của dữ liệu. Quá trình này mang tính tương tác cao, trong đó các câu hỏi mới liên tục nảy sinh từ chính bối cảnh dữ liệu.

Trong giáo dục TK, cách tiếp cận EDA có những đặc điểm phù hợp với quá trình nhận thức của người học. Biehler (1994) chỉ ra rằng với EDA, người học thường tập trung vào chính tập dữ liệu đang phân tích mà không nhất thiết phải giả định đó là mẫu đại diện cho một tổng thể, do đó các kết luận đưa ra thường mang tính gợi mở và gắn liền với ngữ cảnh. Điều này giúp giảm bớt gánh nặng về các khái niệm xác suất phức tạp trong giai đoạn đầu, cho phép HS tập trung phát triển tư duy TK thông qua việc tương tác trực tiếp với dữ liệu. Nghiên cứu của Ben-Zvi và Arcavi (2001) cho thấy rằng việc tham gia vào các hoạt động EDA giúp HS xây dựng cái nhìn tổng quan về dữ liệu, chuyển từ việc chỉ quan tâm đến các giá trị cục bộ sang việc nhận biết các xu hướng và mối quan hệ tổng quát. Đồng thời, sự hỗ trợ của các phần mềm giáo dục TK hiện đại cho phép trực quan hóa dữ liệu, qua đó có thể thúc đẩy sự tò mò và khả năng tự đặt câu hỏi của người học - một yếu tố quan trọng trong dạy học dựa trên khám phá.

2.1.2. Dạy học thống kê theo tiếp cận phân tích dữ liệu khám phá ở trường phổ thông

Việc đưa EDA vào chương trình giáo dục phổ thông đã được nhiều tổ chức giáo dục uy tín khuyến nghị như một hướng đi hiệu quả để phát triển tư duy cho HS. Báo cáo GAISE do Hiệp hội Thống kê Hoa Kỳ công bố nhấn mạnh rằng mục tiêu của giáo dục TK không chỉ dừng lại ở trang bị kĩ năng tính toán mà cần hướng tới phát triển suy luận TK và tư duy TK (Franklin và cộng sự, 2007). Trong bối cảnh trường phổ thông, tiếp cận EDA cho phép HS tiếp cận dữ liệu như những “thám tử”, khuyến khích các em đặt câu hỏi về nguồn gốc dữ liệu, tìm kiếm xu hướng và những điểm bất thường thay vì chỉ áp dụng công thức một cách máy móc.

Điểm đặc trưng của dạy học theo hướng tiếp cận này là sự chuyển dịch vai trò của người học từ thụ động tiếp nhận sang chủ động khám phá. Thông qua việc sử dụng các phần mềm giáo dục TK có tính tương tác cao như CODAP hay TinkerPlots, HS có cơ hội tư duy với dữ liệu: từ thu thập, làm sạch, trực quan hóa đến diễn giải kết quả. Theo khung chương trình của New Zealand, HS trung học cần nhận thức được rằng các kết luận TK luôn phụ thuộc vào bối cảnh và chất lượng dữ liệu (New Zealand Ministry of Education, 2007). Quá trình này giúp hình thành thói quen xem xét, đánh giá có cơ sở khoa học và khả năng phản biện - những yếu tố quan trọng của năng lực TK.

Tại Việt Nam, định hướng của Chương trình giáo dục phổ thông môn Toán 2018 đối với mạch kiến thức TK - Xác suất cũng thể hiện sự tương đồng với triết lí của EDA. Các yêu cầu cần đạt như “thu thập, phân loại, biểu diễn, phân tích và xử lí dữ liệu” (Bộ GD-ĐT, 2018) không chỉ đòi hỏi kĩ năng thao tác mà còn nhấn mạnh vào khả năng ra quyết định dựa trên dữ liệu. Việc vận dụng quy trình EDA trong dạy học có thể xem là một giải pháp sư phạm phù hợp để hiện thực hóa các mục tiêu này, giúp HS kết nối kiến thức toán học với các vấn đề thực tiễn, qua đó phát triển năng lực giải quyết vấn đề toán học và năng lực mô hình hóa toán học.

2.1.3. Dạy học theo dự án ở trường phổ thông

DHTDA được hiểu là một mô hình dạy học trong đó người học đóng vai trò trung tâm, thực hiện các nhiệm vụ phức hợp gắn liền với thực tiễn dưới sự hướng dẫn của GV. Theo Meier và Nguyễn Văn Cường (2014), đặc trưng cơ bản của phương pháp này là sự kết hợp chặt chẽ giữa lý thuyết và thực hành, trong đó người học tự lực thực hiện nhiệm vụ - thường dưới hình thức làm việc nhóm - để tạo ra các sản phẩm cụ thể có thể giới thiệu được. Đồng quan điểm này, Nguyễn Văn Tuấn (2022) nhận định rằng trong DHTDA môn Toán, tính tự giác và chủ động của người học được đề cao từ khâu xác định mục đích, lập kế hoạch đến thực hiện và đánh giá sản phẩm. Quá trình này không chỉ giúp người học củng cố kiến thức mà còn tạo điều kiện phát triển các kỹ năng mềm và năng lực giải quyết vấn đề trong môi trường mở.

Các nghiên cứu thực nghiệm đã cung cấp bằng chứng về tác động tích cực của DHTDA đối với người học. Holmes và Hwang (2016) chỉ ra rằng việc áp dụng DHTDA ở trường trung học có thể giúp cải thiện kỹ năng của HS trong các chủ đề đại số và hình học, đồng thời thúc đẩy động lực và sự tham gia vào quá trình học tập. Nghiên cứu của Han và cộng sự (2015) về các dự án STEM gợi ý rằng phương pháp này có xu hướng mang lại lợi ích rõ rệt cho nhóm HS có thành tích thấp, qua đó có thể góp phần thu hẹp khoảng cách về kết quả học tập giữa các nhóm đối tượng khác nhau. Điều này cho thấy DHTDA có tiềm năng trở thành một phương pháp tiếp cận công bằng và hiệu quả, đặc biệt đối với những HS gặp khó khăn với các phương pháp truyền thống.

Khi xem xét mối quan hệ giữa DHTDA và tiếp cận EDA, có thể nhận thấy sự tương thích về mặt sự phạm. Cả hai hướng tiếp cận đều đề cao tính khám phá và giải quyết các vấn đề mở, trong đó dữ liệu thực tế đóng vai trò quan trọng thúc đẩy quá trình tìm tòi. Tuy nhiên, việc triển khai DHTDA trong thực tế vẫn đối mặt với những thách thức nhất định. Haatainen và Aksela (2021) ghi nhận rằng mặc dù GV nhận thức được lợi ích của DHTDA về mặt gắn kết thực tiễn, họ vẫn gặp khó khăn trong công tác tổ chức và quản lý lớp học. Thực tế này gợi ý sự cần thiết phải xây dựng các quy trình thiết kế dự án cụ thể, đặc biệt là trong lĩnh vực TK - nơi mà sự kết hợp giữa kỹ năng toán học, hiểu biết về bối cảnh và công nghệ là đặc biệt quan trọng.

2.2. Nguyên tắc và quy trình thiết kế dự án học tập thông kê theo tiếp cận phân tích dữ liệu khám phá

2.2.1. Nguyên tắc thiết kế dự án theo phân tích dữ liệu khám phá

Để đảm bảo tính khả thi và hiệu quả sự phạm khi tích hợp DHTDA với tiếp cận EDA, việc thiết kế các hoạt động học tập cần tuân thủ một hệ thống các nguyên tắc định hướng. Dựa trên các đặc trưng của EDA và chủ đề TK trong Chương trình môn Toán 2018, bài báo đề xuất năm nguyên tắc cơ bản sau đây:

(1) *Tính khám phá*: Dự án cần tạo điều kiện để HS đặt câu hỏi, khám phá và diễn giải dữ liệu, coi dữ liệu là đối tượng trung tâm của quá trình suy ngẫm và khái quát hóa.

(2) *Tính thực tiễn*: Chủ đề và dữ liệu của dự án nên gắn với các bối cảnh quen thuộc trong đời sống của HS, qua đó giúp HS nhận thấy ý nghĩa của dữ liệu và thúc đẩy nhu cầu tìm hiểu vấn đề thông qua công cụ TK.

(3) *Tính công nghệ*: Hoạt động dự án cần được thiết kế với sự hỗ trợ của các phần mềm giáo dục TK như CODAP, TinkerPlots hoặc Google Sheets nhằm tạo điều kiện để HS trực quan hóa dữ liệu, nhận diện xu hướng và kiểm chứng giả thuyết thông qua tương tác với các dạng biểu diễn dữ liệu khác nhau.

(4) *Tính phát triển*: Hoạt động dự án cần hướng tới việc hình thành, phát triển tư duy và suy luận TK, giúp HS vượt qua mức mô tả dữ liệu để giải thích, dự báo và ra quyết định dựa trên dữ liệu.

(5) *Tính phản biện*: Trong quá trình thực hiện EDA, HS cần được rèn luyện thói quen tự đánh giá và phản biện khoa học. Điều này bao gồm việc xem xét lại chất lượng nguồn dữ liệu, đánh giá tính đại diện của mẫu và nhận biết các giới hạn trong kết luận TK.

2.2.2. Quy trình thiết kế

Trên cơ sở phân tích các đặc trưng của phương pháp DHTDA và định hướng EDA, đồng thời tham khảo và kế thừa các mô hình DHTDA đã được nghiên cứu bởi Meier và Nguyễn Văn Cường (2014), Trần Thị Hà Phương (2018) cũng như Nguyễn Thị Thu Thảo và Trần Việt Cường (2019), bài báo đề xuất quy trình thiết kế dự án học tập TK gồm năm bước như sau:

Bước 1: Xác định mục tiêu dự án. Mục tiêu của dự án cần được xác định trên cơ sở yêu cầu cần đạt của Chương trình môn Toán 2018, đặc biệt là ở chủ đề TK. GV cần cụ thể hóa các năng lực mà dự án hướng tới, chẳng hạn: năng lực thu thập và xử lý dữ liệu, năng lực phân tích và diễn giải kết quả, năng lực sử dụng công cụ TK và công nghệ, năng lực hợp tác và giao tiếp trong giải quyết vấn đề. Việc xác định rõ mục tiêu không chỉ giúp định hướng cho toàn bộ quá trình thiết kế, mà còn là cơ sở để xây dựng các nhiệm vụ học tập và tiêu chí đánh giá phù hợp.

Bước 2: Lựa chọn nội dung và dữ liệu. Trọng tâm của một dự án học tập theo tiếp cận EDA được đặt vào việc lựa chọn và khai thác dữ liệu. Ở bước này, GV cần xác định loại dữ liệu phù hợp với mục tiêu học tập, chẳng hạn dữ

liệu ghép nhóm hoặc không ghép nhóm, đồng thời lựa chọn các bộ dữ liệu có nguồn gốc rõ ràng, bảo đảm tính xác thực. Bên cạnh đó, mức độ phức tạp của dữ liệu cần được cân nhắc để bảo đảm sự phù hợp với trình độ nhận thức của HS cũng như điều kiện hạ tầng công nghệ hiện có. Việc mô tả các biến số, bao gồm loại biến (định lượng hoặc định tính), số lượng giá trị và đơn vị đo lường là cần thiết nhằm dự kiến các dạng biểu đồ và phương pháp phân tích mà HS có thể sử dụng trong quá trình học tập. Ở giai đoạn này, GV có thể khai thác dữ liệu từ các nguồn mở như Kaggle, Google Dataset hoặc thiết kế các hoạt động thu thập dữ liệu trực tiếp trong lớp học.

Bước 3: Thiết kế các nhiệm vụ học tập. GV cần xác định câu hỏi trung tâm của dự án - một vấn đề mở gắn với bối cảnh dữ liệu và có khả năng kích thích tư duy khám phá. Các nhiệm vụ học tập nên được sắp xếp theo tiến trình của EDA, bao gồm: làm quen và mô tả dữ liệu ban đầu; tổ chức và trực quan hóa dữ liệu; phân tích dữ liệu thông qua việc nhận diện xu hướng, mối quan hệ và đặt, kiểm chứng các giả thuyết ban đầu; diễn giải, trình bày kết quả và mở rộng vấn đề. Việc tích hợp công nghệ là yếu tố không thể thiếu trong bước này. GV cần thiết kế các hoạt động cho phép HS thao tác với phần mềm, tạo biểu đồ, kéo-thả biến, lọc dữ liệu và quan sát sự thay đổi tức thời của các đại lượng. Các sản phẩm học tập có thể bao gồm báo cáo TK, bài trình bày, poster hoặc video giới thiệu kết quả phân tích.

Bước 4: Xây dựng kế hoạch đánh giá. Đánh giá cần được thiết kế đa dạng, kết hợp giữa đánh giá quá trình và đánh giá sản phẩm. GV cần xác định các tiêu chí cụ thể để đo lường mức độ đạt được của HS, chẳng hạn: Khả năng mô tả và tổ chức dữ liệu; Mức độ chính xác và sáng tạo trong việc sử dụng biểu đồ, chỉ số TK; Năng lực diễn giải và suy luận từ dữ liệu; Tính hợp tác, giao tiếp và trình bày khoa học. Các công cụ đánh giá có thể bao gồm rubric đánh giá, phiếu quan sát hay bảng kiểm tự đánh giá.

Bước 5: Thử nghiệm và điều chỉnh. GV rà soát lại toàn bộ tiến trình, tham khảo ý kiến đồng nghiệp hoặc tổ chức dạy thử nghiệm ở quy mô nhỏ. Mục đích là để phát hiện các điểm chưa hợp lý trong nội dung dữ liệu, dự báo những khó khăn mà HS có thể gặp phải hoặc sự chưa rõ ràng trong các tiêu chí đánh giá. Từ những phản hồi này, bản thiết kế sẽ được điều chỉnh và tối ưu hóa để đảm bảo tính khả thi và hiệu quả cao nhất.

Điểm khác biệt của quy trình đề xuất so với các mô hình DHTDA truyền thống nằm ở việc tích hợp tiếp cận EDA trong toàn bộ quá trình triển khai. Nếu DHTDA truyền thống thường bắt đầu từ một vấn đề mở, thì quy trình này xem dữ liệu thực tế là điểm khởi nguồn của hoạt động học tập (Bước 2). Tiến trình thực hiện dự án nhấn mạnh hoạt động khám phá dữ liệu, trực quan hóa và hình thành giả thuyết thay vì chủ yếu kiểm chứng giả thuyết có sẵn (Bước 3). Đồng thời, hoạt động đánh giá tập trung vào khả năng suy luận TK của HS thay vì chỉ đánh giá sản phẩm báo cáo cuối cùng (Bước 4). Qua đó, mô hình đề xuất cho thấy một định hướng phù hợp với đặc trưng của tiếp cận EDA.

2.3. Ví dụ minh họa

Để làm rõ tính khả thi của quy trình đề xuất, phần tiếp theo của bài báo trình bày ví dụ thiết kế một dự án học tập dành cho HS lớp 10 với chủ đề “Khám phá các yếu tố ảnh hưởng đến kết quả học tập”. Dự án được thiết kế để thực hiện sau khi HS đã học nội dung về các số đặc trưng của mẫu số liệu không ghép nhóm, với sự hỗ trợ của phần mềm CODAP.

2.3.1. Xác định mục tiêu dự án

Dựa trên yêu cầu cần đạt của chương trình môn Toán lớp 10, dự án được thiết kế nhằm mục tiêu giúp HS không chỉ thành thạo kỹ năng tính toán các số đặc trưng (trung bình, trung vị, tứ phân vị, độ lệch chuẩn) mà còn phát triển khả năng suy luận TK thông qua việc xử lý dữ liệu thực tế. Cụ thể, HS cần đạt được khả năng đọc và diễn giải ý nghĩa của các con số TK trong bối cảnh thực tế, phát hiện các giá trị bất thường và bước đầu đưa ra các nhận định về mối quan hệ giữa các biến số. Bên cạnh đó, dự án cũng hướng đến việc rèn luyện kỹ năng sử dụng phần mềm giáo dục TK (CODAP) để trực quan hóa dữ liệu và kỹ năng làm việc nhóm để giải quyết vấn đề.

2.3.2. Lựa chọn nội dung và dữ liệu

Nội dung dự án được lựa chọn gắn với chủ đề “Phân tích dữ liệu không ghép nhóm”, phù hợp với chương trình lớp 10. Bộ dữ liệu “Phong cách sống hàng ngày và thành tích học tập của HS” (Ishara, 2024) được lựa chọn vì: (1) Có quy mô mẫu lớn với 2000 quan sát và các biến đa dạng, giúp HS có cơ hội khám phá nhiều khía cạnh khác nhau; (2) Các biến có ý nghĩa gần gũi với đời sống HS, từ đó tăng hứng thú học tập và tính thực tiễn của dự án; (3) Dữ liệu có thể sử dụng linh hoạt cho các dạng phân tích mô tả, so sánh nhóm và tìm kiếm mối liên hệ tiềm năng.

2.3.3. Thiết kế các nhiệm vụ học tập

Các hoạt động học tập được tổ chức xoay quanh câu hỏi trung tâm: “Các yếu tố trong phong cách sống hàng ngày (thời gian học tập, vận động, giấc ngủ, giao tiếp xã hội, mức độ căng thẳng) ảnh hưởng như thế nào đến kết quả học tập của HS?”.

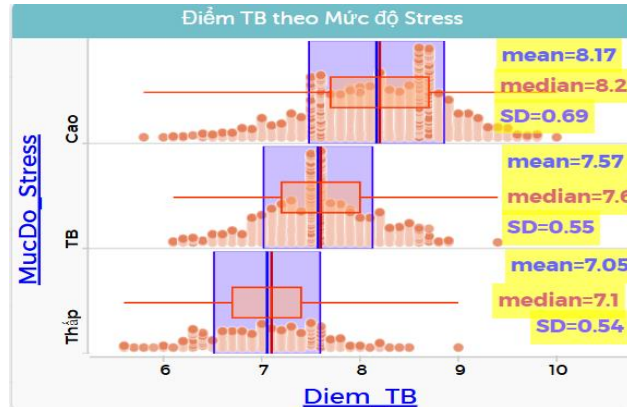
Các nhiệm vụ học tập được chia thành bốn giai đoạn chính, cụ thể như sau:

Giai đoạn 1 - Làm quen và khám phá dữ liệu: HS tìm hiểu nguồn gốc, cấu trúc, các biến và kiểu dữ liệu của bộ dữ liệu. Làm quen với phần mềm CODAP: nhập, sắp xếp, phân loại dữ liệu theo các biến số khác nhau. Thảo luận và đề xuất các câu hỏi khám phá ban đầu (ví dụ: “Liệu HS học nhiều hơn có điểm trung bình cao hơn không?”).

Giai đoạn 2 - Tổ chức và trực quan hóa dữ liệu: HS biểu diễn dữ liệu bằng các hình thức phù hợp: biểu đồ cột, biểu đồ hộp, biểu đồ phân bố tần suất. Thực hiện xác định các số đo TK cơ bản: trung bình, trung vị, độ lệch chuẩn, khoảng tứ phân vị. Quan sát và mô tả đặc điểm phân bố, phát hiện các giá trị bất thường.

Giai đoạn 3 - Phân tích và diễn giải kết quả:

HS phân tích mối quan hệ giữa các biến như: Thời gian học và điểm trung bình; Mức độ căng thẳng và kết quả học tập; Thời gian giao tiếp xã hội và điểm trung bình; Mức độ vận động và trạng thái căng thẳng. Kết quả phân tích được diễn giải bằng ngôn ngữ TK: mô tả xu hướng, so sánh nhóm, đưa ra nhận xét và kết luận. Ví dụ, từ kết quả phân tích, HS có thể nhận thấy: Điểm trung bình tăng dần theo thời gian học; HS có mức căng thẳng “Cao” lại có điểm trung bình cao hơn nhóm “Thấp” và “Trung bình” (hình 1); Việc vận động thường xuyên liên quan đến mức độ căng thẳng thấp hơn. Những kết quả này khuyến khích HS lí giải và thảo luận về nguyên nhân - hệ quả, cũng như các yếu tố có thể ảnh hưởng khác.



Hình 1. Biểu đồ điểm trung bình của các nhóm HS theo mức độ căng thẳng

Giai đoạn 4 - Trình bày, kết luận và mở rộng: HS tổng hợp kết quả, so sánh với giả thuyết ban đầu và trình bày kết quả dưới dạng báo cáo hoặc thuyết trình nhóm. GV hướng dẫn HS diễn giải ý nghĩa TK của các kết quả trong bối cảnh thực tế (ví dụ: “Vì sao căng thẳng cao lại gắn với điểm số cao hơn?”). HS được khuyến khích đặt ra các câu hỏi mở rộng (ví dụ: “Liệu thời gian ngủ hay hoạt động ngoại khóa có ảnh hưởng đến điểm số không?”) và đề xuất hướng thu thập dữ liệu mới tại chính trường học của mình. Các nhiệm vụ học tập được thiết kế theo hướng mở, cho phép HS chủ động khám phá, đề xuất và kiểm chứng, qua đó thể hiện tinh thần của EDA trong dạy học TK.

2.3.4. Xây dựng kế hoạch đánh giá

Kế hoạch đánh giá được cụ thể hóa thông qua rubric gồm năm tiêu chí: (1) Tổ chức và thu gọn dữ liệu; (2) Phân tích và suy luận TK; (3) Sử dụng công nghệ; (4) Hợp tác và giao tiếp; (5) Phản biện và đánh giá. Với mỗi tiêu chí, thang đánh giá được phân chia thành ba mức độ: Mức 1 - Cần cố gắng; Mức 2 - Đạt yêu cầu và Mức 3 - Tốt. Ví dụ với tiêu chí (2) Phân tích và suy luận TK, sự phân hóa dựa trên biểu hiện của HS khi xem xét sự chuyển dịch từ việc chỉ thực hiện thao tác tính toán (mức 1), đến khả năng diễn giải ý nghĩa các con số trong bối cảnh thực tế (mức 2) và đạt mức cao nhất khi HS biết sử dụng dữ liệu làm bằng chứng thuyết phục để bảo vệ các giả thuyết, bác bỏ các suy đoán sai lầm hoặc đánh giá lại độ tin cậy của chính nguồn dữ liệu (mức 3).

2.3.5. Thử nghiệm và điều chỉnh

Dự án được thử nghiệm ở quy mô nhỏ trong một lớp học nhằm thu thập phản hồi từ HS và đồng nghiệp. Kết quả bước đầu cho thấy HS tham gia tích cực vào hoạt động phân tích dữ liệu thực tế và chủ động hơn trong việc diễn giải biểu đồ, chỉ số TK. Về tiêu chí phân tích và suy luận, đa số HS đạt mức 2. Tuy nhiên, quá trình triển khai cũng ghi nhận một số khó khăn thực tiễn. Việc GV đưa ra nhiều câu hỏi định hướng ở giai đoạn đầu có thể làm giảm tính tự chủ của HS trong hoạt động khám phá dữ liệu. Bên cạnh đó, việc chưa quen với các thao tác trên phần mềm khiến một số HS gặp khó khăn trong việc tập trung vào mục tiêu suy luận TK. Từ đó, một số điều chỉnh được đề xuất như: + Giảm số lượng câu hỏi ban đầu để HS có thêm thời gian thảo luận và khám phá dữ liệu; + Bổ sung hướng dẫn chi tiết về cách sử dụng phần mềm CODAP; + Khuyến khích HS mở rộng vấn đề bằng việc tự thu thập dữ liệu nhỏ trong lớp học.

3. Kết luận

Bài báo đã hệ thống hóa cơ sở lí luận và đề xuất quy trình thiết kế dự án học tập trong dạy học TK ở THPT theo tiếp cận EDA. Kết quả nghiên cứu cho thấy việc kết hợp EDA và DHTDA có thể tạo điều kiện để HS chủ động hơn trong việc tìm hiểu, phân tích và diễn giải dữ liệu thực tế, thay vì chủ yếu thực hiện các thao tác tính toán kĩ thuật. Bài báo cũng đề xuất năm nguyên tắc sư phạm và quy trình thiết kế dự án học tập TK gồm năm bước dành cho cấp

THPT. Quy trình này góp phần cụ thể hóa tiếp cận khám phá dữ liệu trong các hoạt động dạy học phù hợp với bối cảnh giáo dục Toán học Việt Nam hiện nay.

Mặc dù vậy, nghiên cứu vẫn tồn tại giới hạn nhất định khi trọng tâm mới dừng ở việc đề xuất mô hình lý thuyết và bước đầu đánh giá tính khả thi trên quy mô hẹp. Vì vậy, các nghiên cứu tiếp theo có thể tập trung đánh giá hiệu quả thực tế thông qua thực nghiệm sư phạm trên diện rộng, đồng thời mở rộng theo hướng xây dựng công cụ đánh giá khả năng suy luận TK của HS và khảo sát các yếu tố ảnh hưởng đến tính khả thi của việc áp dụng tiếp cận EDA trong bối cảnh trường phổ thông tại Việt Nam.

Tài liệu tham khảo

- Batanero, C., & Borovcnik, M. (2016). *Statistics and probability in high school*. Brill.
- Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. *Educational Studies in Mathematics*, 45(1), 35-65. <https://doi.org/10.1023/A:1013809201228>
- Biehler, R. (1994). Probabilistic thinking, statistical reasoning, and the search for causes - Do we need a probabilistic revolution after we have taught data analysis. *Research papers from ICOTS*, 4, 20-37.
- Bowler, L., & Shaw, C. (2024). Trends in data literacy, 2018-2023: a review of the literature. *Information Research an International Electronic Journal*, 29(2), 198-205. <https://doi.org/10.47989/ir292822>
- Bộ GD-ĐT (2018). *Chương trình giáo dục phổ thông môn Toán* (ban hành kèm theo Thông tư số 32/2018/TT-BGDĐT ngày 26/12/2018 của Bộ trưởng Bộ GD-ĐT).
- Frank, M., Walker, J., Attard, J., & Tygel, A. (2016). Data Literacy-What is it and how can we make it happen? *The Journal of Community Informatics*, 12(3), 4-8. <https://doi.org/10.15353/joci.v12i3.3274>
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report*. In: Alexandria, VA: American Statistical Association.
- Haatainen, O., & Aksela, M. (2021). Project-based learning in integrated science education: Active teachers' perceptions and practices. *LUMAT: International Journal on Math, Science and Technology Education*, 9(1), 149-173. <https://doi.org/10.31129/LUMAT.9.1.1392>
- Han, S., Capraro, R., & Capraro, M. M. (2015). How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: The impact of student factors on achievement. *International Journal of Science and Mathematics Education*, 13(5), 1089-1113. <https://doi.org/10.1007/s10763-014-9526-0>
- Holmes, V.-L., & Hwang, Y. (2016). Exploring the effects of project-based learning in secondary mathematics education. *The Journal of Educational Research*, 109(5), 449-463. <https://doi.org/10.1080/00220671.2014.979911>
- Ishara, S. (2024). *Student Stress & Performance Insights*. <https://www.kaggle.com/code/sulaniishara/student-stress-performance-insights/input>
- Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools*, 19(3), 267-277. <https://doi.org/10.1177/1365480216659733>
- Khalid, K., & Mason, J. (2024). *Challenges to Augmenting Literacy in the Digital Environment*. International Conference on Computers in Education.
- Meier, B., Nguyễn Văn Cường (2014). *Lí luận dạy học hiện đại*. NXB Đại học Sư phạm.
- New Zealand Ministry of Education. (2007). *The New Zealand curriculum*. Wellington, NZ.
- Nguyễn Thị Thu Thảo, Trần Việt Cường (2019). Tổ chức dạy học theo dự án chủ đề Tam giác đồng dạng và ứng dụng cho học sinh lớp 8 ở trường trung học cơ sở. *Tạp chí Giáo dục*, 456, 35-41.
- Nguyễn Văn Tuấn (2022). *Tổ chức dạy học theo dự án trong môn Toán cao cấp cho sinh viên đại học khối ngành Kỹ thuật*. Luận án tiến sĩ Khoa học giáo dục, Trường Đại học Sư phạm - Đại học Thái Nguyên.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading/Addison-Wesley, MA.
- Trần Thị Hà Phương (2018). *Dạy học theo dự án một số chủ đề Toán rời rạc cho học sinh chuyên Toán*. Luận án tiến sĩ Khoa học giáo dục, Trường Đại học Sư phạm Hà Nội.