

Phân nhóm doanh nghiệp ngành bất động sản bằng phân tích dữ liệu đa hợp dựa trên mô hình DuPont

Trịnh Thị Hương¹, Nguyễn Anh Tuấn²

Trường Đại học Thương mại, Việt Nam¹, Trường Đại học Hà Nội, Việt Nam²

Ngày nhận: 18/09/2025

Ngày nhận bản sửa: 02/12/2025

Ngày duyệt đăng: 18/12/2025

Tóm tắt: Nghiên cứu này tiến hành phân tích phân nhóm các doanh nghiệp bất động sản niêm yết tại Việt Nam giai đoạn 2012- 2023 bằng phương pháp phân tích dữ liệu đa hợp (CoDa) dựa trên mô hình DuPont. Bộ dữ liệu gồm 63 doanh nghiệp với 621 quan sát được thu thập từ ba sàn HOSE, HNX và UPCOM. Kết quả chỉ ra ba nhóm doanh nghiệp có đặc trưng tài chính khác biệt. Nhóm 1 chiếm đa số, duy trì tỷ trọng nợ dài hạn khoảng 45% tổng nợ và đạt hiệu quả sinh lời ổn định. Nhóm 2 có mức sinh lời tương tự nhưng gần như không sử dụng nợ dài hạn, phản ánh chiến lược thận trọng và ít phổ biến. Nhóm 3, mặc dù sử dụng đòn bẩy tài chính cao nhất, lại có ROA và ROE thấp, thể hiện rủi ro khi lạm dụng nợ. Kết quả nghiên cứu nhấn mạnh vai trò của cơ cấu vốn, đặc biệt là nợ dài hạn, trong việc duy trì khả năng sinh lời và nâng

Cluster analysis of real estate enterprises using compositional data analysis method based on DuPont model

Abstract: This study conducts a compositional data analysis (CoDa)- based clustering of listed real estate firms in Vietnam over the period 2012- 2023, using the DuPont model as the analytical framework. The dataset comprises 63 firms and 621 firm-year observations collected from the HOSE, HNX, and UPCOM exchanges. The findings reveal three distinct clusters characterized by heterogeneous financial profiles. Cluster 1, which represents the majority of the sample, maintains a long-term debt ratio of approximately 45% of total liabilities and exhibits stable profitability. Cluster 2 shows a comparable level of profitability but relies almost entirely on short-term financing, suggesting a conservative and less prevalent capital structure strategy. Cluster 3 employs the highest degree of financial leverage yet reports the lowest ROA and ROE, indicating heightened financial risk associated with excessive debt utilisation. The results highlight the critical role of capital structure-particularly long-term debt-in sustaining profitability and strengthening the competitive position of real estate firms. The study provides meaningful implications for managers, investors, and policymakers in assessing corporate financial health, formulating sustainable development strategies, and making investment decisions aligned with market conditions.

Keywords: Real estate enterprises, Cluster analysis, Compositional data analysis, DuPont model

Link Doi: <https://doi.org/10.59276/JELB.2026.04.3070>

Trinh Thi Huong¹, Nguyen Anh Tuan²

Email: trinthithuong@tmu.edu.vn¹, tuanna@hanu.edu.vn²

Organization: Thuongmai University, Viet Nam¹, Hanoi University, Viet Nam²

cao năng lực cạnh tranh của doanh nghiệp bất động sản; cung cấp cơ sở tham khảo cho các nhà quản trị, nhà đầu tư và nhà hoạch định chính sách, nhằm đánh giá sức khỏe tài chính của doanh nghiệp, xây dựng chiến lược phát triển bền vững và đưa ra quyết định đầu tư phù hợp với điều kiện thị trường.

Từ khóa: *Doanh nghiệp bất động sản, Phân nhóm, Dữ liệu đa hợp, Mô hình Dupont*

Trích dẫn: Trịnh Thị Hương, & Nguyễn Anh Tuấn. (2026). Phân nhóm doanh nghiệp ngành bất động sản. bằng phân tích dữ liệu đa hợp dựa trên mô hình DuPont. *Tạp chí Kinh tế - Luật và Ngân hàng*, 28(4), 143-156. <https://doi.org/10.59276/JELB.2026.04.3070>

1. Giới thiệu

Thị trường bất động sản được xem là một trong những trụ cột của nền kinh tế, giúp tập trung các nguồn lực, tạo ra tài sản cố định cho quốc gia. Sự phát triển của thị trường bất động sản tạo ra động lực tăng trưởng cho các ngành nghề liên quan, đáp ứng nhu cầu về chỗ ở và phát triển đô thị, du lịch. Theo Tổng cục Thống kê, đóng góp trung bình của ngành xây dựng và bất động sản vào tổng GDP cả nước các năm gần đây chiếm khoảng 10% (Phạm Quốc Duẩn, 2023). Trong đó, ngành bất động sản trực tiếp chiếm khoảng 3,5%, đóng góp trung bình khoảng 0,5 điểm phần trăm vào tăng trưởng GDP.

Phân nhóm (phân cụm) doanh nghiệp là hoạt động phân tách các doanh nghiệp thành những nhóm khác nhau dựa trên các đặc điểm như quy mô vốn, hiệu quả hoạt động, loại hình sở hữu, cũng như các yếu tố về đổi mới sáng tạo (Akhmetshin và cộng sự, 2017; Truong, 2015). Trong giai đoạn 2012- 2023, thị trường bất động sản Việt Nam đã trải qua sự biến động mạnh mẽ với tốc độ mở rộng nhanh của khu vực tư nhân, quá trình cổ phần hóa doanh nghiệp nhà nước và sự gia tăng của dòng vốn đầu tư nước ngoài (Trần Kim Chung, 2023). Những thay đổi này tạo ra sự phân hóa rõ rệt về năng lực tài chính và hiệu quả hoạt động giữa các doanh nghiệp. Từ đó, việc

phân nhóm doanh nghiệp đã trở thành một vấn đề cấp thiết nhằm nhận diện cấu trúc thị trường và đánh giá khả năng cạnh tranh trong toàn ngành. Đối với các nhà đầu tư, việc phân nhóm doanh nghiệp đã trở thành công cụ không thể thiếu trong việc đánh giá cơ hội và rủi ro đầu tư. Đối với doanh nghiệp, việc xác định nhóm tương đồng giúp họ đánh giá vị thế tài chính tương đối, nhận diện khoảng cách hiệu quả so với các đối thủ và xây dựng chiến lược cải thiện hiệu quả sử dụng vốn. Từ góc độ quản lý nhà nước, phân nhóm doanh nghiệp bất động sản theo năng lực tài chính cung cấp công cụ hữu ích cho việc hoạch định và giám sát chính sách, đặc biệt trong bối cảnh kiểm soát tín dụng, cơ cấu vốn và phát triển thị trường bền vững.

Nghiên cứu này tập trung phân nhóm doanh nghiệp bất động sản thông qua các tỷ số tài chính theo mô hình Dupont (Coenders, 2022; Saus-Sala và cộng sự, 2021). Tỷ số tài chính là công cụ hữu ích trong đánh giá hiệu quả hoạt động của doanh nghiệp, tỷ số tài chính sẽ là biến đầu vào của các phương pháp định lượng, để đo lường, phân tích và nhận định sức khỏe tài chính và hiệu suất vận hành của một doanh nghiệp. Ngoài ra, tỷ số tài chính còn được sử dụng để so sánh hiệu quả hoạt động giữa các doanh nghiệp trong cùng ngành hoặc đánh giá sự thay đổi về hiệu suất tài chính theo thời gian. Dữ liệu tài chính có đặc

điểm tự nhiên của dữ liệu đa hợp (Linares-Mustarós và cộng sự, 2022). Chúng tôi sử dụng phương pháp phân nhóm dựa trên dữ liệu đa hợp (Compositional Data Analysis, CoDa)- phương pháp phân nhóm dữ liệu hiệu quả được đề xuất bởi Coender và cộng sự (Ferrer-Rosell & Coenders, 2018; Jofre-Campuzano & Coenders, 2022; Linares-Mustarós và cộng sự, 2018). Một khoảng trống quan trọng trong nghiên cứu học thuật là chưa có nghiên cứu định lượng nào phân nhóm các doanh nghiệp bất động sản niêm yết tại Việt Nam. Nghiên cứu này giải quyết vấn đề phương pháp luận bằng cách áp dụng CoDa kết hợp mô hình DuPont. CoDa vượt trội hơn các tiếp cận truyền thống nhờ khả năng xử lý đối xứng các tỷ số tài chính, giảm ảnh hưởng trị ngoại biên, cho phép sử dụng nhiều tỷ số và không yêu cầu phân phối chuẩn.

Từ dữ liệu của 63 doanh nghiệp trong giai đoạn 2012- 2023, nghiên cứu xác định ba nhóm có đặc trưng tài chính khác biệt. Nhóm 1 chiếm đa số, duy trì tỷ trọng nợ dài hạn khoảng 45% và đạt mức sinh lời ổn định (ROA 4,5%, ROE 11,6%). Nhóm 2 đạt mức sinh lời tương tự (ROA 4,9%, ROE 10,7%) nhưng hầu như không dùng nợ dài hạn (0,01), thể hiện chiến lược thận trọng và ít phổ biến. Ngược lại, Nhóm 3 sử dụng đòn bẩy cao nhất (4,78) nhưng có ROA và ROE thấp nhất (3,5% và 6,8%), cho thấy rủi ro của việc lạm dụng nợ khi hiệu quả hoạt động yếu. Kết quả cho thấy CoDa kết hợp DuPont phân tách doanh nghiệp theo cấu trúc tài chính, đồng thời cung cấp bằng chứng thực nghiệm cho nhà đầu tư và nhà hoạch định chính sách trong đánh giá tính bền vững tài chính và rủi ro hệ thống của ngành.

Bài viết được cấu trúc gồm năm phần: Giới thiệu; Tổng quan nghiên cứu; Phương pháp nghiên cứu; Kết quả và thảo luận; Kết luận và khuyến nghị.

2. Tổng quan nghiên cứu

2.1. Mô hình Dupont

Mô hình DuPont, được giới thiệu lần đầu năm 1914 bởi F. Donaldson Brown (Soliman, 2008), là một phương pháp phân tích lợi nhuận, bao gồm:

Khả năng sinh lời (Profitability): Đo lường bằng biên lợi nhuận (profit margin), định nghĩa là thu nhập ròng (net income)/tổng doanh thu (total revenue). Tỷ số này giúp các nhà phân tích có thể xác định tỷ lệ phần trăm chảy vào “lợi nhuận ròng” trên mỗi đơn vị doanh thu.

Hiệu suất (Efficiency): Đo lường bằng tỷ lệ vòng quay tổng tài sản (turnover), được định nghĩa là tổng doanh thu/tổng tài sản.

Cơ cấu vốn (Capital structure). Cơ cấu vốn được xem xét dựa trên tài sản dài hạn, tài sản ngắn hạn, nợ dài hạn và nợ ngắn hạn. Việc sử dụng nợ có khả năng khuếch đại tỷ lệ lợi nhuận trên vốn chủ sở hữu (ROE) theo cả chiều hướng tích cực và tiêu cực.

Khả năng thanh khoản (Liquidity) của công ty thường được đo lường thông qua tỷ lệ giữa tài sản ngắn hạn và nợ dài hạn. Phân tích lợi nhuận và khả năng thanh khoản là hai yếu tố quan trọng trong đánh giá hiệu suất của bất kỳ tổ chức nào để đưa ra quyết định đúng đắn (Hasidi và cộng sự, 2024).

Bên cạnh mô hình Dupont, mô hình Altman cũng được sử dụng trong phân tích tài chính của doanh nghiệp (Kasilingam & Jayabal, 2012). Mô hình Dupont tập trung nhiều hơn vào phân tách nguồn gốc hiệu quả sinh lời của doanh nghiệp, do đó phù hợp hơn với mục tiêu phân nhóm doanh nghiệp, còn mô hình Altman dự báo rủi ro phá sản của doanh nghiệp thông qua kết hợp các tỷ số tài chính (Altman và cộng sự, 2017). Mô hình Dupont có ưu điểm là các biến tài chính trong mô hình Dupont thu được trực tiếp từ báo cáo tài chính của doanh nghiệp.

2.2. Phân nhóm doanh nghiệp

Phân nhóm đóng vai trò quan trọng trong nghiên cứu và quản trị doanh nghiệp. Thông qua phương pháp phân nhóm, các doanh nghiệp có đặc điểm tương đồng sẽ được gom lại (Feng và cộng sự, 2025). Có nhiều tiêu chí về phân nhóm doanh nghiệp: ngành nghề, địa lý hay hiệu quả hoạt động. Phân nhóm doanh nghiệp là một yếu tố quan trọng trong quá trình phát triển của các doanh nghiệp, từ đó thúc đẩy sự hợp tác, chia sẻ kiến thức và đổi mới (Akhmetshin và cộng sự, 2017). Trong tài chính và kinh doanh, phân nhóm doanh nghiệp là một công cụ hiệu quả để đánh giá hoạt động của doanh nghiệp và dự báo kết quả kinh doanh (Sáenz và cộng sự, 2023). Nhờ vậy, các nhà đầu tư có thể dùng phương pháp này để đánh giá sức khỏe của doanh nghiệp, để có thể đưa ra các quyết định đầu tư sáng suốt và phân bổ nguồn lực hợp lý. Chủ sở hữu và các nhà quản lý doanh nghiệp có thể so sánh hiệu suất công ty của họ với nhóm doanh nghiệp gần nhất thay vì mức trung bình toàn ngành.

Có nhiều phương pháp phân nhóm khác nhau như phân cụm phân cấp (hierarchical), K-means, phân cụm theo mô hình (model-based), phân cụm theo mật độ (density-based) và liên kết Ward (Ward linkage). Các phương pháp có sự khác nhau về giả định và cách hình thành cụm. Đối với dữ liệu đa hợp, Greenacre (2018) cho thấy phương pháp Ward phù hợp và hiệu quả.

2.3. Phân nhóm doanh nghiệp bằng phương pháp CoDa

Phân nhóm doanh nghiệp dựa trên các hiệu suất tài chính được sử dụng từ giữa thế kỷ 19 trong nhiều lĩnh vực khác nhau (Ayaba và cộng sự, 2023; Gupta & Huefner, 1972; Jensen, 1971). Năm 2021, Germà Coenders

và cộng sự đề xuất sử dụng phương pháp phân tích dữ liệu đa hợp (Compositional Data Analysis, CoDa) để trực quan hóa phân tích DuPont và phân nhóm doanh nghiệp (Coenders, 2022). Hướng tiếp cận này dựa trên dữ liệu cấu trúc vốn, thường được công bố trong các báo cáo tài chính của công ty (Coenders, 2022), có tính tương đồng và tự nhiên của dữ liệu CoDa. Dữ liệu tài chính đa hợp của các công ty là vectơ đa hợp 6 thành phần (tài sản cố định, tài sản lưu động, nợ dài hạn, nợ ngắn hạn, doanh thu tổng và chi phí hoạt động) (Creixans-Tenas và cộng sự, 2019). Phương pháp CoDa giúp xử lý các tỷ lệ tài chính một cách đối xứng sao cho kết quả không phụ thuộc vào tử số và mẫu số. So với hướng tiếp cận thông qua phân nhóm trực tiếp trên biến tài chính, phân tích CoDa cũng giúp làm giảm thiểu các vấn đề kỹ thuật và cho kết quả phân nhóm tốt hơn. Cụ thể, phương pháp Coda ít ảnh hưởng bởi trị ngoại biên, có thể sử dụng nhiều tỷ số tài chính và dữ liệu không phân phối chuẩn (Arimany-Serrat và cộng sự, 2022, 2023; Creixans-Tenas và cộng sự, 2019; Linares-Mustarós và cộng sự, 2022). Phương pháp tiếp cận mới đã được cải tiến và áp dụng phân nhóm doanh nghiệp hiệu quả trong nhiều ngành và bối cảnh khác nhau (Coenders & Serrat, 2023).

Các ứng dụng của CoDa trong phân tích tài chính bao gồm các báo cáo kế toán của các công ty ngành rượu vang tại Tây Ban Nha (Coenders, 2022). Thông qua các tỷ số tài chính của mô hình Dupont, các tác giả cho thấy phương pháp CoDa là hữu ích khi thực hiện các phân nhóm doanh nghiệp, ví dụ doanh nghiệp trong ngành du lịch và nông nghiệp (Jofre-Campuzano & Coenders, 2022; Saus-Sala và cộng sự, 2021). Trong đó, một vài nghiên cứu đã ứng dụng phân tích CoDa trong phân nhóm hồ sơ doanh nghiệp (Jofre-Campuzano & Coenders, 2022). Phân tích dữ liệu đa hợp, khi so

sánh với việc sử dụng các tỷ lệ tài chính căn bản, đem lại các kết quả hoàn toàn khác biệt (Jofre-Campuzano & Coenders, 2022; Saus-Sala và cộng sự, 2021). Do đó, phương pháp này có khả năng mang lại kết quả phân nhóm doanh nghiệp hiệu quả khi áp dụng vào các nghiên cứu tại một quốc gia khác hoặc một ngành khác.

Tại Việt Nam, phương pháp tiếp cận này lần đầu được sử dụng trong phân nhóm các doanh nghiệp thủy sản và sản xuất thực phẩm tại Việt Nam trong bối cảnh đại dịch Covid-19. Các doanh nghiệp này được phân loại thành ba nhóm với các đặc điểm riêng biệt về khả năng thanh khoản và lợi nhuận dựa trên phân tích DuPont. Kết quả phân nhóm cũng cho thấy sự dịch chuyển từ nhóm này sang nhóm khác trong đại dịch COVID-19 (Dao và cộng sự, 2024). Dựa trên dữ liệu báo cáo tài chính của 300 doanh nghiệp lớn tại Thái Lan, Indonesia và Việt Nam, Đào Thị Thanh Bình và cộng sự (2025) đã phân nhóm doanh nghiệp và phân tích đặc tính cũng như sự dịch chuyển giữa các nhóm giai đoạn 2018- 2022. Kết quả xác định ba nhóm theo hiệu quả kinh doanh và rủi ro tài chính. Trong bối cảnh COVID-19, các doanh nghiệp Việt Nam thể hiện hiệu suất tích cực và duy trì tỷ lệ lớn trong nhóm có hoạt động tốt nhất với mức an toàn thanh khoản cao.

Theo tổng quan của nhóm tác giả, các nghiên cứu định lượng về phân nhóm doanh nghiệp tại Việt Nam, đặc biệt trong lĩnh vực bất động sản, vẫn còn hạn chế dù phương pháp này phổ biến quốc tế nhờ dữ liệu tài chính sẵn có. Trong bối cảnh thị trường bất động sản Việt Nam biến động mạnh, việc phân tích sâu về đặc điểm tài chính- kinh doanh của doanh nghiệp cung cấp bằng chứng hữu ích cho nhà hoạch định chính sách, lãnh đạo doanh nghiệp và nhà đầu tư. Đồng thời, việc ứng dụng phân tích dữ liệu đa hợp dựa trên các tỷ lệ tài

chính từ mô hình DuPont góp phần bổ sung khoảng trống về phương pháp luận.

3. Phương pháp nghiên cứu

3.1. Dữ liệu nghiên cứu

Dữ liệu nghiên cứu được thu thập từ báo cáo tài chính của các doanh nghiệp bất động sản tại Việt Nam giai đoạn 2012-2023, tương ứng với một chu kỳ kinh tế biến động mạnh. Giai đoạn này khởi đầu sau khủng hoảng 2012, phục hồi vào 2013-2014, bùng nổ trong 2015- 2018, suy giảm từ cuối 2018 đến đầu 2022 dưới tác động của COVID-19, và tiếp tục phát triển nhưng đối mặt thách thức vĩ mô trong năm 2023. Doanh nghiệp được lựa chọn theo phân ngành GICS và hệ thống phân loại của HOSE, HNX và UPCOM. Danh sách được lấy từ cơ sở dữ liệu Refinitiv. Tiêu chí bao gồm: (1) niêm yết liên tục trên ba sàn trong giai đoạn 2012- 2023; (2) có đầy đủ báo cáo tài chính kiểm toán năm; (3) không bị cảnh báo, kiểm soát hoặc đình chỉ giao dịch; và (4) không thay đổi ngành nghề kinh doanh chính. Mẫu cuối cùng gồm 63 doanh nghiệp với 621 quan sát.

Dữ liệu bao gồm các biến tài chính từ báo cáo tài chính hợp nhất (tổng tài sản, tài sản ngắn hạn và dài hạn, tổng nợ, cơ cấu nợ, doanh thu, chi phí, lợi nhuận sau thuế) và các biến phi tài chính như thời gian niêm yết, loại hình sở hữu và số năm hoạt động. Trong 621 quan sát, có 455 quan sát từ HOSE, 119 từ HNX và 47 từ UPCOM.

3.2. Đo lường tỷ số tài chính từ mô hình DuPont

Mô hình DuPont, được giới thiệu lần đầu năm 1914 bởi F. Donaldson Brown, là một phương pháp phân tích lợi nhuận. Mô hình Dupont được tính trực tiếp từ các biến tài

Phân nhóm doanh nghiệp ngành bất động sản
bằng phân tích dữ liệu đa hợp dựa trên mô hình DuPont

chính trong báo cáo tài chính. Cụ thể, ký hiệu:

- x_1 : Tài sản dài hạn (Non current assets)
- x_2 : Tài sản ngắn hạn (Current assets)
- x_3 : Nợ dài hạn (Non current liabilities)
- x_4 : Nợ ngắn hạn (Current liabilities)
- x_5 : Doanh thu (Net sales)
- x_6 : Chi phí (Costs)

Từ đó, tổng tài sản (Total assets) của doanh nghiệp là $x_1 + x_2$ và tổng nợ (Total liabilities) là $x_3 + x_4$. Các tỷ số về khả năng thanh khoản được quan tâm là: Khả năng thanh khoản hiện hành, Tỷ số nợ dài hạn trên tổng nợ và cấu trúc vốn (Bảng 1). Theo mô hình Dupont, tỷ số tài chính được quan tâm nhất là tỷ suất lợi nhuận trên vốn chủ sở hữu Bảng 1. Mô hình Dupont cho phép phân tích khả năng sinh lời như sau:

$$ROA = \text{Turnover} \times \text{Margin}$$

$$ROE = ROA \times \text{Leverage}$$

3.3. Phân nhóm dữ liệu đa hợp đối với biến tài chính

Một vectơ đa hợp gồm D thành phần dương, ký hiệu $x = (x_1, x_2, \dots, x_D)$ và thông tin liên quan giữa các thành phần trong vectơ (tỷ số giữa các phần đó) là mối quan tâm (Egozcue & Pawlowsky-Glahn, 2019).

Một đặc trưng của dữ liệu đa hợp là các tính toán thông thường (như khoảng cách) đối với vectơ Euclide không thực hiện được (Pawlowsky-Glahn & Buccianti, 2011). Do đó, các tiếp cận đối với phân tích CoDa là sử dụng phép chuyển đổi để đưa một vectơ đa hợp về vectơ trong không gian Euclide thông thường.

Đối với các biến tài chính x_1, x_2, \dots, x_D đã giới thiệu ở trên đều không âm, tỷ số giữa các thành phần (tức là tỷ số tài chính dựa trên mô hình Dupont) có ý nghĩa quan trọng nên biến tài chính này được xem là một vectơ đa hợp gồm 6 thành phần. Để phù hợp và hiệu quả phân tích nhóm và dữ liệu đa hợp, Coender và cộng sự (Coenders, 2022; Jofre-Campuzano & Coenders, 2022) đề xuất sử dụng biến đổi logarit tỉ lệ trung tâm có trọng số (Weighted CLR): với $j=1, \dots, 6$; $D=6$ ta có:

$$wCLR_j = \sqrt{w_j} \log \left(\frac{x_j}{x_1^{w_1} x_2^{w_2} \dots x_D^{w_D}} \right),$$

Trong đó, trọng số w_j được chọn sao cho tổng $\sum_j w_j = 1$ và w_j tỷ lệ với trung bình mẫu \bar{x}_j của thành phần thứ j (Greenacre & Lewi, 2009; Jofre-Campuzano & Coenders,

Bảng 1. Tỷ số tài chính dựa trên mô hình Dupont

Tỷ số tài chính		Công thức
Tên tiếng Việt	Tên tiếng Anh	
Khả năng thanh khoản hiện hành	Liquidity	x_2/x_3
Tỷ số nợ dài hạn trên tổng nợ	Debt-maturity	$x_3/(x_3+x_4)$
Tỷ số nợ	Indebtedness	$(x_3+x_4)/(x_1+x_2)$
Tỷ lệ vòng quay tổng tài sản	Turnover	$x_5/(x_1+x_2)$
Biên lợi nhuận ròng	Margin	$(x_5-x_6)/x_5$
Tỷ suất lợi nhuận trên tổng tài sản	Return On Assets (ROA)	$(x_5-x_6)/(x_1+x_2)$
Hệ số đòn bẩy tài chính	Leverage	$(x_1+x_2)/((x_1+x_2)-(x_3+x_4))$
Tỷ suất lợi nhuận trên vốn chủ sở hữu	Return On Equity (ROE)	$(x_5-x_6)/((x_1+x_2)-(x_3+x_4))$

Nguồn: Tác giả tổng hợp dựa trên (Đào Thị Thanh Bình và cộng sự, 2025; Jofre-Campuzano & Coenders, 2022).

2022)- trọng số tối ưu theo Greenacre (2018) và cũng là trọng số mặc định của hàm phân tích trong gói easyCODA trên phần mềm R. Biến đổi wCLR phân nhóm hiệu quả hơn đối với các tỉ số tài chính (Jofre-Campuzano & Coenders, 2022).

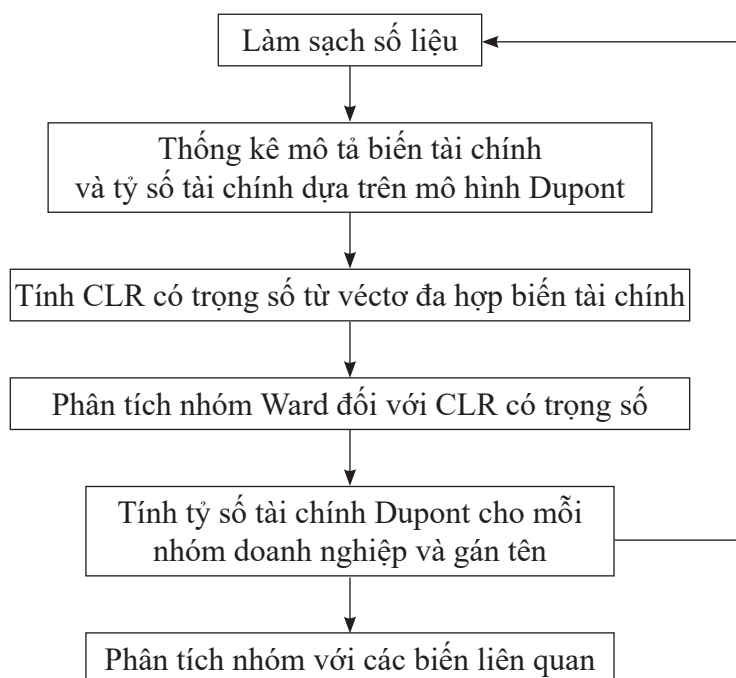
Khi đó, các phương pháp phân nhóm thông thường như K-mean, Ward được sử dụng đối với dữ liệu ($wCLR_1, wCLR_2, \dots, wCLR_D$). Về mặt kỹ thuật thống kê, số nhóm tối ưu được lựa chọn theo nhiều chỉ số và tiêu chí khác nhau bằng hàm NbClust của gói NbClust trong phần mềm R (Charrad và cộng sự, 2014). Thuật toán xác định số nhóm tối ưu dựa trên 30 chỉ số, đồng thời sử dụng biểu đồ dendrogram và biểu đồ khuỷu tay để hỗ trợ lựa chọn trực quan. Về mặt thực tiễn, tiêu chí phân nhóm dựa trên mô hình DuPont, trong đó các tỷ số hiệu quả hoạt động càng cao càng tốt, còn các tỷ số nợ (tỷ trọng nợ dài hạn, tỷ số nợ, hệ số đòn bẩy) càng thấp càng tốt.

Hình 1 minh họa quy trình 06 bước của phân tích dữ liệu đa hợp dựa trên mô hình DuPont: làm sạch dữ liệu, thống kê mô tả các biến và tỷ số tài chính, tính CLR có trọng số, phân nhóm Ward, tính tỷ số tài chính cho từng nhóm và đặt tên nhóm, cuối cùng phân tích nhóm theo các biến quan sát bổ sung. Quy trình này có thể được lặp lại nhiều lần để đạt được các nhóm doanh nghiệp phù hợp và có ý nghĩa.

4. Kết quả nghiên cứu và thảo luận

4.1. Thống kê mô tả

Bảng 2 trình bày cái nhìn khái quát về đặc điểm của bộ dữ liệu nghiên cứu, bao gồm thống kê mô tả đối với 63 doanh nghiệp ngành bất động sản niêm yết trên ba sàn chứng khoán tại Việt Nam (HOSE, HNX và UPCOM), với tổng số 621 quan sát trong giai đoạn 2012- 2023. Bộ dữ liệu ban



Nguồn: Tác giả đề xuất dựa trên Dao và cộng sự (2024).

Hình 1. Quy trình phân tích dữ liệu đa hợp dựa trên mô hình DuPont

Phân nhóm doanh nghiệp ngành bất động sản
bằng phân tích dữ liệu đa hợp dựa trên mô hình DuPont

Bảng 2. Thống kê mô tả dữ liệu (đơn vị: tỉ VND)

Năm	Tài sản dài hạn	Tài sản ngắn hạn	Nợ dài hạn	Nợ ngắn hạn	Doanh thu	Chi phí
2012	1150,75 (3802,24)	2331,45 (5210,43)	1075,29 (3424,05)	1532,8 (4566,84)	462,11 (1376,49)	374,54 (1053,08)
2013	1683,19 (5320,65)	3492,34 (7631,89)	1583,23 (5109,6)	2205,4 (5347,66)	1079,44 (3118,39)	804,64 (2003,72)
2014	2505,34 (7900,19)	3505,95 (7081,46)	2144,41 (6289,68)	2102,68 (5184,71)	1312,2 (4267,44)	1159,39 (3703,18)
2015	3061,75 (10618,93)	4280,18 (10776,94)	2154,42 (6759,79)	3091,91 (9908,1)	1559,01 (4975,14)	1386,35 (4735,1)
2016	3394,33 (12346,35)	5289,68 (14072,99)	1999,93 (5892,93)	3898,54 (14132,13)	2323,28 (8043,36)	1999,1 (7370,63)
2017	3156,78 (14572,57)	5555,51 (15163,7)	1684,55 (5319,51)	4247,45 (16986,79)	2857,76 (11955,75)	2508,69 (11176,38)
2018	4665,99 (19682,08)	7700,11 (21634,32)	3054,3 (11121,61)	4692,72 (15412,07)	4096,77 (16388,05)	3425,25 (15147,78)
2019	5953,77 (26447,36)	10083,71 (31231,82)	3449,53 (14102,63)	7244,58 (27367,34)	4597,37 (17577,14)	3694,81 (15795,42)
2020	8198,95 (34865,24)	10869,59 (28810,26)	5168,76 (18672,11)	7693,99 (25830,45)	4694,09 (16995,08)	3804,43 (14831,92)
2021	8806,88 (37568,51)	11805,51 (30874,99)	5778,62 (21382,96)	7059,88 (21884,65)	5511,62 (19506,74)	4598,12 (18204,79)
2022	9939,67 (41918,58)	13580,48 (45253,07)	4553,57 (19334,4)	11363,55 (46393,81)	4482,11 (15726,14)	3664,37 (13926,74)
2023	12598,21 (49450,94)	20771,81 (62601,05)	7068,56 (22983,2)	16580,27 (63336,1)	6657,95 (26416,13)	5663,82 (24034,93)

Chú thích: Giá trị trình bày trong bảng là giá trị trung bình và độ lệch chuẩn (trong ngoặc) khi biến quan tâm là liên tục; giá trị là tần suất (dạng phần trăm) khi biến quan tâm là rời rạc.

Nguồn: Tính toán của tác giả trên phần mềm R bằng gói lệnh tidyverse.

đầu gồm 654 quan sát, trong đó 33 quan sát bị loại theo quy trình ở Hình 1 (6 quan sát có toàn bộ biến tài chính rỗng và 10 quan sát có chi phí âm). Với các quan sát có biến tài chính bằng 0, thuật toán EM được áp dụng thông qua hàm lrEM của gói zCompositions để thay thế bằng giá trị khác 0. Quy trình phân cụm được lặp lại bốn lần và loại bỏ thêm 17 quan sát khi xuất hiện các cụm riêng lẻ có quá ít quan sát (1- 5 doanh nghiệp).

Bảng 2 trình bày thống kê mô tả của sáu biến đầu vào (x_1 - x_6), gồm giá trị trung bình và độ lệch chuẩn trong ngoặc đơn. Độ lệch chuẩn cao, nhiều trường hợp gấp đôi giá trị trung bình, phản ánh mức biến động lớn và sự khác biệt đáng kể về quy mô, quản trị và hiệu quả hoạt động giữa các doanh nghiệp.

Bảng 3 trình bày các biến được tính từ tỷ số tài chính DuPont dựa trên dữ liệu đầu vào. Giai đoạn 2012- 2023 cho thấy trung bình và độ lệch chuẩn của khả năng thanh khoản và đòn bẩy tài chính tăng trong nửa đầu giai đoạn, sau đó giảm trong thời kỳ COVID-19. Biên lợi nhuận ròng có xu hướng giảm liên tục, đồng thời độ lệch chuẩn cũng giảm. Tỷ suất sinh lời trên tài sản (ROA) duy trì ở mức thấp, tăng nhẹ giữa giai đoạn ròi giảm dần, thậm chí âm vào năm 2023. Tỷ suất sinh lời trên vốn chủ sở hữu (ROE) nhìn chung ổn định trong toàn bộ giai đoạn nghiên cứu.

4.2. Kết quả phân nhóm

Kết quả phân nhóm dữ liệu đa hợp đối với dữ liệu ($wCLR_1, wCLR_2, \dots, wCLR_D$)

Bảng 3. Thống kê mô tả các biến phân tích Dupont

Năm	Tỷ số nợ dài hạn trên tổng nợ	Cơ cấu vốn	Hệ số đòn bẩy tài chính	Tỷ lệ vòng quay tổng tài sản	Biên lợi nhuận ròng	Khả năng thanh khoản hiện hành	Tỷ suất lợi nhuận trên tổng tài sản	Tỷ suất lợi nhuận trên vốn chủ sở hữu
2012	0,56 (0,21)	3,15 (2,48)	0,2 (0,2)	0,15 (0,17)	3,07 (2,92)	0,02 (0,03)	0,07 (0,11)	1,66 (0,75)
2013	0,58 (0,18)	2,86 (1,4)	0,26 (0,18)	0,13 (0,22)	2,92 (3,29)	0,03 (0,05)	0,09 (0,15)	1,56 (0,72)
2014	0,57 (0,19)	3,03 (2,08)	0,24 (0,19)	0,13 (0,18)	3,65 (4,44)	0,02 (0,04)	0,04 (0,22)	1,45 (0,66)
2015	0,56 (0,22)	3,38 (3,66)	0,26 (0,18)	0,18 (0,17)	3,02 (2,64)	0,04 (0,04)	0,06 (0,35)	1,4 (0,64)
2016	0,55 (0,22)	3,22 (3,2)	0,28 (0,18)	0,19 (0,19)	3,54 (3,81)	0,05 (0,06)	0,12 (0,16)	1,47 (0,67)
2017	0,57 (0,19)	3,05 (1,92)	0,33 (0,29)	0,2 (0,24)	2,4 (1,77)	0,05 (0,08)	0,13 (0,27)	1,51 (0,71)
2018	0,56 (0,2)	3,19 (2,88)	0,33 (0,24)	0,2 (0,19)	2,28 (1,58)	0,06 (0,06)	0,12 (0,34)	1,47 (0,7)
2019	0,56 (0,19)	2,96 (2,11)	0,32 (0,25)	0,24 (0,26)	2,19 (1,85)	0,06 (0,05)	0,16 (0,16)	1,6 (0,8)
2020	0,59 (0,17)	3,34 (2,82)	0,28 (0,2)	0,19 (0,21)	1,99 (0,96)	0,05 (0,05)	0,13 (0,15)	1,6 (0,79)
2021	0,56 (0,19)	3,88 (8,45)	0,27 (0,22)	0,25 (0,2)	2,29 (1,45)	0,05 (0,04)	0,15 (0,2)	1,56 (0,77)
2022	0,59 (0,19)	3,55 (3,58)	0,22 (0,16)	0,16 (0,21)	2,27 (1,47)	0,03 (0,04)	0,11 (0,21)	1,59 (0,78)
2023	0,56 (0,19)	4,32 (8,57)	0,16 (0,12)	0,11 (0,24)	2,42 (1,63)	0,02 (0,04)	-0,01 (0,52)	1,55 (0,7)

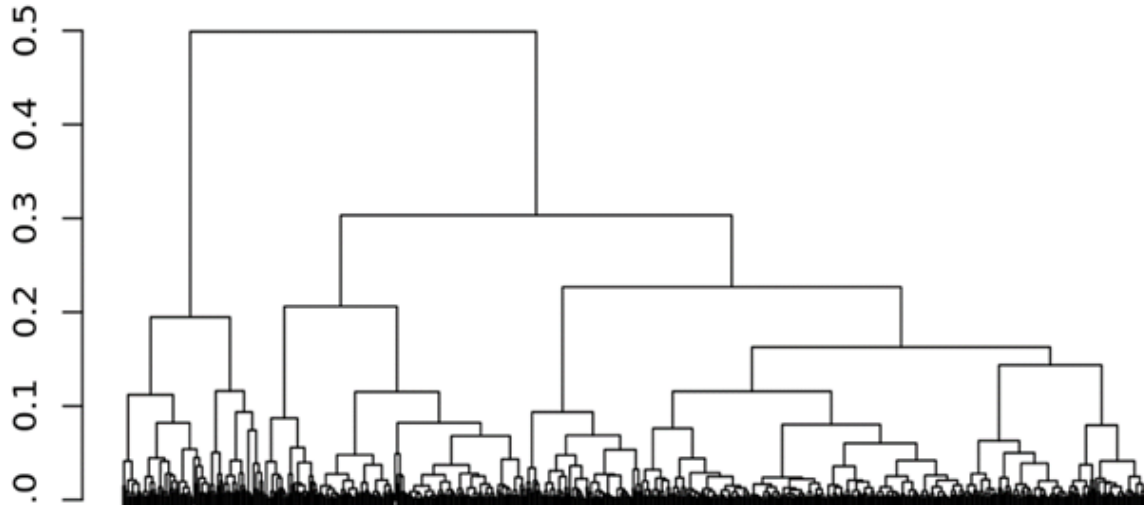
Chú thích: Giá trị trình bày trong bảng là giá trị trung bình và độ lệch chuẩn (trong ngoặc) khi biến quan tâm là liên tục.

Người: Tính toán của tác giả trên phần mềm R bằng gói lệnh tidyverse.

thông qua hàm NbClust cho thấy 8 chỉ số phân cụm đề xuất 3 nhóm là tối ưu (Charrad và cộng sự, 2014). Đồ thị Dendrogram cho thấy 3 cụm là một điểm cắt phù hợp (Hình 2). Hàm NbClust cũng cho thấy 3 là số nhóm tối ưu đối với các phương pháp phân cụm bằng Kmeans và phương pháp phân cụm phân cấp. Số lượng quan sát của 3 nhóm từ dữ liệu đa hợp là 376, 86, 159. Số lượng quan sát của 3 cụm từ phương pháp Kmeans là 620, 12, 6. Số lượng quan sát của 3 nhóm từ phương pháp phân cụm phân cấp là 630, 6, 2. Số lượng quan sát trong các nhóm như trên là không hợp lý do quá chênh lệch số quan sát. Từ kết quả gợi ý phân nhóm đối với dữ liệu đa hợp, chúng ta cần xem xét đặc trưng của các nhóm- tức các nhóm doanh nghiệp bất động sản (Bảng 4).

Sau khi phân nhóm bằng phương pháp phân tích, kết quả được trình bày trong Bảng 3. Các nhóm được sắp xếp từ tốt nhất đến yếu nhất. Kiểm định giữa các nhóm cho thấy khác biệt ở mức ý nghĩa 1% với hầu hết tỷ số tài chính, trừ biên lợi nhuận ròng, phản ánh sự khác biệt chiến lược tài chính và hiệu quả hoạt động của doanh nghiệp bất

Phân nhóm doanh nghiệp ngành bất động sản
bằng phân tích dữ liệu đa hợp dựa trên mô hình DuPont



Nguồn: Tính toán của tác giả trên phần mềm R bằng hàm WARD trong gói lệnh easyCODA.

Hình 2. Biểu đồ Dendrogram

động sản niêm yết.

Nhóm 1 và nhóm 2 có tỷ suất lợi nhuận tương đương, nhưng nhóm 1 cao hơn. Khác biệt chính nằm ở tỷ lệ nợ dài hạn: nhóm 1 duy trì mức khoảng 45%, trong khi nhóm 2 hầu như không dùng nợ dài hạn. Nhóm 1 cũng chiếm đa số, còn nhóm 2 chỉ gồm ít doanh nghiệp. Điều này phản ánh xu hướng ưu tiên nợ dài hạn để tăng ổn định tài chính và giảm rủi ro thanh khoản, dù chi phí vốn cao hơn.

Nhóm 3 có hiệu quả thấp hơn rõ rệt, với ROA và ROE đều thấp. Dù sử dụng đòn bẩy cao nhất, vòng quay tài sản thấp khiến nhóm không đạt mức sinh lời kỳ vọng, cho thấy rủi ro khi lạm dụng đòn bẩy trong điều kiện hiệu quả kinh doanh kém. Điều này nhấn mạnh đòn bẩy chỉ hữu ích khi doanh nghiệp duy trì hiệu quả ổn định.

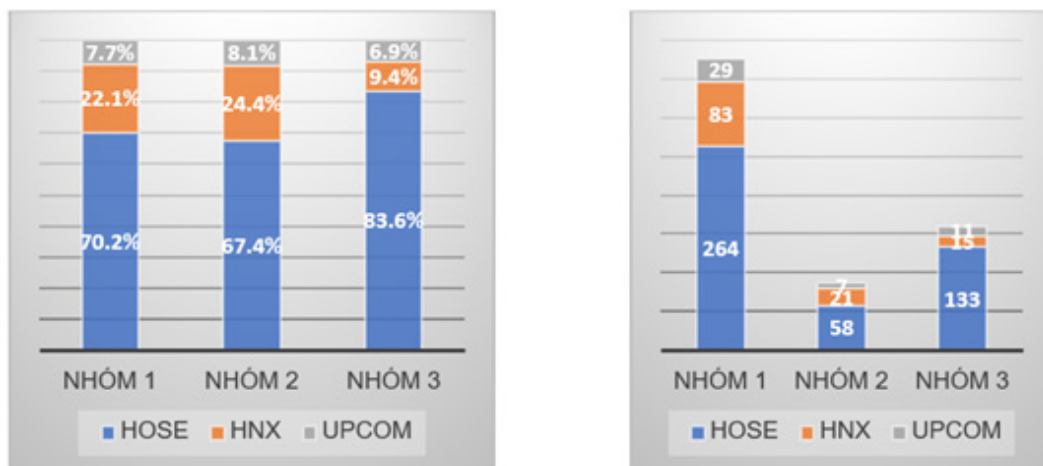
Xét theo sàn giao dịch, UPCOM có số quan sát hạn chế, trong khi HOSE chiếm khoảng 2/3 mẫu. Tỷ lệ này phân bố tương đối đồng

Bảng 4. Trung bình các tỷ số tài chính theo kết quả phân nhóm

Các tỷ số tài chính	Nhóm 1	Nhóm 2	Nhóm 3	p-giá trị
	N = 376	N = 86	N = 159	
Khả năng thanh khoản hiện hành	2,68 (2,61)	2,21 (2,95)	2,66 (1,67)	<0,001
Tỷ số nợ dài hạn trên tổng nợ	0,43 (0,25)	0,01 (0,01)	0,33 (0,19)	<0,001
Hệ số đòn bẩy tài chính	2,97 (2,26)	2,34 (1,59)	4,78 (7,58)	<0,001
Tỷ lệ vòng quay tổng tài sản	0,29 (0,21)	0,31 (0,27)	0,20 (0,16)	<0,001
Biên lợi nhuận ròng (%)	19,90 (17,47)	18,16 (20,28)	20,12 (20,14)	0,2
Tỷ suất lợi nhuận trên tổng tài sản (%)	4,83 (4,19)	5,16 (6,79)	3,76 (4,71)	<0,001
Tỷ suất lợi nhuận trên vốn chủ sở hữu (%)	13,19 (14,13)	11,27 (18,64)	11,36 (16,60)	0,001

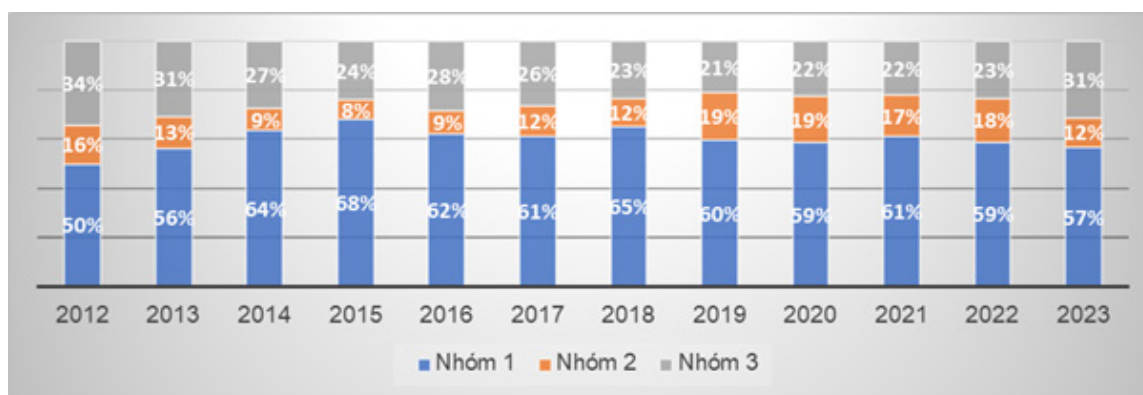
Chú thích: Giá trị trung bình và độ lệch chuẩn (trong ngoặc). P-giá trị của kiểm định Kruskal-Wallis về so sánh các nhóm.

Nguồn: Tính toán trên phần mềm R bằng hàm tbl_summary trong gói lệnh gtsummary.



Nguồn: Tính toán của tác giả trên phần mềm R bằng gói lệnh tidyverse và Excel.

Hình 3. Tỷ lệ niêm yết theo sàn trong 3 nhóm (trái) và số lượng quan sát phân loại theo sàn niêm yết và trong các nhóm (phải)



Nguồn: Tính toán của tác giả trên phần mềm R bằng gói lệnh tidyverse và Excel.

Hình 4. Tỷ lệ phân nhóm theo năm giai đoạn 2012- 2023 (Đơn vị: %)

đều ở cả ba nhóm, kể cả nhóm yếu nhất, cho thấy không thể kết luận doanh nghiệp trên sàn nào hoạt động tốt hơn do HOSE vốn tập trung nhiều doanh nghiệp quy mô lớn. Hình 4 thể hiện sự dịch chuyển giữa các nhóm doanh nghiệp trong giai đoạn 2012-2023, được gọi là ma trận dịch chuyển. Nhóm 3, tức nhóm các công ty có kết quả kinh doanh yếu nhất, duy trì trạng thái khá ổn định trong toàn bộ giai đoạn (chiếm tỷ lệ khoảng 25%). Điều này phản ánh rằng, trong ngành bất động sản, một bộ phận đáng kể các doanh nghiệp thường xuyên gặp khó khăn về hiệu quả hoạt động và khó

thoát ra khỏi tình trạng kém hiệu quả. Sự ổn định của Nhóm 3, mặc dù ở mức thấp, cũng cho thấy rằng khả năng chuyển dịch từ Nhóm 3 sang Nhóm 2 hoặc Nhóm 1 là khá hạn chế, và do đó, các chính sách cải thiện hiệu quả hoạt động trong ngành vẫn chưa tạo ra sự tác động mạnh mẽ đến nhóm yếu kém này. Ngược lại, sự biến động chủ yếu về cơ cấu có thể quan sát ở Nhóm 2-nhóm các doanh nghiệp có kết quả hoạt động tương đối tốt, tuy nhiên đặc điểm nổi bật của nhóm này là phụ thuộc nhiều vào nguồn vốn vay ngắn hạn, trong khi nợ dài hạn hầu như không được sử dụng. Tỷ lệ

của Nhóm 2 trong toàn ngành dao động từ 9% đến 20%. Nhóm 2 không đại diện cho một chiến lược đặc thù hay khác biệt, bởi mặc dù các doanh nghiệp này hoạt động trong lĩnh vực bất động sản- vốn đòi hỏi nhu cầu vốn dài hạn rất lớn- nhưng lại lựa chọn chiến lược huy động vốn thiên về ngắn hạn.

Hình 4 cho thấy tỉ lệ các công ty Nhóm 1 thấp nhất và Nhóm 3 cao nhất năm 2012- tương ứng là năm cuối trong giai đoạn khủng hoảng của thị trường bất động sản trước khi bước vào giai đoạn phục hồi từ 2013- 2014 với nhiều chính sách hỗ trợ tháo gỡ khó khăn và xử lý nợ xấu. Giai đoạn 2013- 2014 cho thấy sự dịch chuyển từ công ty Nhóm 2 sang công ty thuộc Nhóm 1. Đặc biệt, Hình 4 cho thấy sự gia tăng của các công ty Nhóm 1 trong giai đoạn 2015- 2018, tương ứng với thị trường bùng nổ mạnh mẽ ở hầu hết các phân khúc và các địa phương trong giai đoạn này. Giai đoạn cuối năm 2018 đến đầu 2022, thị trường giảm sút cùng tác động của COVID-19 khiến nhiều dự án đình trệ nhưng nhu cầu đầu tư vẫn cao, dẫn đến giá bất động sản cao cấp tăng nhanh (Nguyễn Thị Minh Phương và cộng sự, 2024). Từ năm 2023, thị trường bất động sản tiếp tục phát triển trở lại nhưng với nhiều thách thức hơn do tác động từ vĩ mô và cả quốc tế (Trần Kim Chung, 2023). Theo dự đoán, các doanh nghiệp bất động sản tiếp tục có tiềm năng phát triển trong giai đoạn tiếp theo, tuy nhiên, các doanh nghiệp này cũng tiếp đối mặt với nhiều khó khăn.

5. Kết luận và một số khuyến nghị

Nghiên cứu sử dụng phương pháp phân tích dữ liệu đa hợp (CoDa) kết hợp mô hình DuPont để phân nhóm các doanh nghiệp bất động sản niêm yết giai đoạn 2012- 2023. Với dữ liệu 63 doanh nghiệp

và 621 quan sát, kết quả cho thấy ba nhóm với đặc trưng tài chính khác biệt, phản ánh sự đa dạng về chiến lược vốn và hiệu quả hoạt động.

Phần lớn doanh nghiệp thuộc nhóm có hiệu quả tương đối cao, duy trì tỷ lệ nợ dài hạn ổn định, hỗ trợ đầu tư dài hạn và giảm rủi ro thanh khoản. Một nhóm nhỏ có hiệu quả tương đương nhưng hạn chế sử dụng nợ dài hạn, thể hiện chiến lược thận trọng nhưng ít phổ biến. Nhóm yếu nhất có khả năng sinh lời thấp dù sử dụng đòn bẩy cao, cho thấy đòn bẩy chỉ hiệu quả khi doanh nghiệp duy trì hiệu quả hoạt động ổn định, nếu không sẽ làm tăng rủi ro và giảm lợi nhuận.

Bên cạnh đó, kết quả nghiên cứu cũng cho thấy hầu hết các doanh nghiệp bất động sản niêm yết trên thị trường chứng khoán Việt Nam đều tập trung ở HOSE, do đó, vai trò trung tâm của HOSE trong việc phản ánh sự dịch chuyển giữa các nhóm là rất rõ ràng. Kết quả phân tích sự dịch chuyển qua các sản phẩm giao dịch trong giai đoạn 2012- 2023 không mang lại nhiều kết quả rõ rệt về khác biệt chất lượng doanh nghiệp giữa các sàn. Điều này cho thấy việc lựa chọn niêm yết trên sàn nào chủ yếu mang tính chiến lược thị trường và quy mô doanh nghiệp, chứ không phản ánh sự khác biệt về hiệu quả hoạt động. Điểm đáng chú ý duy nhất là, trong Nhóm 3- nhóm yếu kém nhất- hầu như không có doanh nghiệp nào thuộc về HNX hay UpCOM. Đặc biệt, ma trận dịch chuyển trong giai đoạn nghiên cứu đã thể hiện được chu kỳ kinh tế đầy biến động và khó khăn của các doanh nghiệp bất động sản.

Từ các phát hiện trên, có thể thấy cơ cấu vốn- đặc biệt là tỷ trọng nợ dài hạn- đóng vai trò quan trọng trong việc duy trì khả năng sinh lời ổn định và củng cố vị thế cạnh tranh của doanh nghiệp bất động sản niêm yết. Đồng thời, kết quả nghiên cứu cũng góp phần bổ sung bằng chứng thực nghiệm về hiệu quả của phương pháp CoDa trong

phân tích tài chính doanh nghiệp, khi cho phép xử lý dữ liệu tỷ số tài chính một cách linh hoạt và khắc phục được nhiều hạn chế của các cách tiếp cận truyền thống.

Nhìn chung, nghiên cứu mang lại giá trị cả về mặt học thuật lẫn thực tiễn. Về mặt lý thuyết, nghiên cứu mở rộng ứng dụng của phương pháp phân tích dữ liệu đa hợp trong lĩnh vực tài chính doanh nghiệp tại Việt Nam, đặc biệt trong bối cảnh thị trường bất động sản có nhiều biến động. Trong khi các biến tài chính có sự dao động lớn, thông qua tiếp cận CoDa, chúng ta có thể tìm thấy sự tương đồng của các nhóm. Về mặt thực tiễn, kết quả nghiên cứu cung cấp cơ sở tham khảo cho các nhà quản trị, nhà đầu tư và nhà hoạch định chính sách trong việc đánh giá sức khỏe tài chính của doanh nghiệp, xây dựng chiến lược phát triển bền vững và đưa ra quyết định đầu tư phù hợp với điều kiện thị trường.

Tài liệu tham khảo

- Akhmetshin, E. M., Kolkpak, E. P., Sulimova, E. A., Kireev, V. S., Samarina, E. A., & Solodilova, N. Z. (2017). Clustering as a criterion for the success of modern industrial enterprises. *International Journal of Applied Business and Economic Research*, 15(23), 221–231. https://www.researchgate.net/publication/322294489_Clustering_as_a_criterion_for_the_success_of_modern_industrial_enterprises
- Altman, E. I., Iwanicz-Drozowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of international financial management & accounting*, 28(2), 131-171. <https://doi.org/10.1111/jifm.12053>
- Arimany-Serrat, N., Farreras-Noguer, M. À., & Coenders, G. (2023). Financial resilience of Spanish wineries during the Covid-19 lockdown. *International Journal of Wine Business Research*, 35(2), 346–364. <https://doi.org/10.1108/IJWBR-03-2022-0012>
- Arimany-Serrat, N., Farreras Noguer, M., & Coenders, G. (2022). New developments in financial statement analysis: liquidity in the winery sector. *Accounting*, 8(3), 355–366. <http://dx.doi.org/10.5267/j.ac.2021.10.002>
- Ayaba, M. M., Amo-Bediako, E., & Takawira, O. (2023). Evaluating the financial performance of listed REIT firms in South Africa: A 7-step DuPont model technique. *International Journal of Applied Economics, Finance and Accounting*, 17(2), 267–278. <https://doi.org/10.33094/ijaefa.v17i2.1146>
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. *Journal of statistical software*, 61, 1-36. <https://doi.org/10.18637/jss.v061.i06>
- Coenders, G. (2022). *Accounting statements as compositional data*. https://www.researchgate.net/publication/351845833_Accounting_Statements_as_Compositional_Data
- Coenders, G., & Serrat, N. A. (2023). *Accounting statement analysis at industry level. A gentle introduction to the compositional approach*. <https://arxiv.org/abs/2305.16842>
- Creixans-Tenas, J., Coenders, G., & Arimany-Serrat, N. (2019). Corporate social responsibility and financial profile of Spanish private hospitals. *Heliyon*, 5(10). <https://doi.org/10.1016/j.heliyon.2019.e02623>
- Dao, B. T. T., Coenders, G., Lai, P. H., Dam, T. T. T., & Trinh, H. T. (2024). An empirical examination of financial performance and distress profiles during COVID-19: The case of fishery and food production firms in Vietnam. *Journal of Financial Reporting and Accounting*. <https://doi.org/10.1108/JFRA-09-2023-0509>
- Đào Thị Thanh Bình, Nguyễn Anh Tuấn, Lê Thanh Bình, & Phạm Văn Hùng. (2025). Phân tích phân cụm các doanh nghiệp ở Việt Nam, Thái Lan và Indonesia trong giai đoạn 2018-2022: Ứng dụng phương pháp phân tích dữ liệu đa hợp. *Tạp Chí Nghiên Cứu Kinh Tế và Kinh Doanh Châu Á*, 36(3), 61-78. <https://doi.org/10.24311/jabes/2025.36.3.04>
- Egozcue, J. J., & Pawlowsky-Glahn, V. (2019). Compositional data: the sample space and its structure. *Test*, 28(3), 599-

Nghiên cứu có một số hạn chế như mẫu dữ liệu chỉ hạn chế trên các doanh nghiệp niêm yết trên sàn chứng khoán. Do phân nhóm doanh nghiệp không có suy luận nhân quả nên chưa tìm được yếu tố quyết định của sự dịch chuyển doanh nghiệp giữa các nhóm (như quy mô, tuổi doanh nghiệp, loại hình sở hữu). Kết quả phân nhóm doanh nghiệp có thể kết hợp với các phương pháp phân tích khác (như phương pháp hồi quy) để có kết quả khoa học toàn diện hơn. ■

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Bộ Giáo dục và Đào tạo Việt Nam, trong khuôn khổ đề tài nghiên cứu khoa học cấp Bộ, Mã số B2024-NHF-04, theo Quyết định số 1393/QĐ-BGDĐT ngày 16 tháng 5 năm 2023.

Phân nhóm doanh nghiệp ngành bất động sản
bằng phân tích dữ liệu đa hợp dựa trên mô hình DuPont

638. <https://doi.org/10.1007/s11749-019-00670-6>
- Feng, L., Zheng, L., & Jin, Y. (2025). Identification of financial distress in companies based on improved K-means clustering algorithm. *2025 International Conference on Intelligent Systems and Computational Networks (ICISCN)*, 1–5. <https://doi.org/10.1109/ICISCN64258.2025.10934452>
- Ferrer-Rosell, B., & Coenders, G. (2018). Destinations and crisis. Profiling tourists' budget share from 2006 to 2012. *Journal of Destination Marketing & Management*, 7, 26–35. <https://doi.org/10.1016/j.jdmm.2016.07.002>
- Greenacre, M., & Lewi, P. (2009). Distributional equivalence and subcompositional coherence in the analysis of compositional data, contingency tables and ratio-scale measurements. *Journal of Classification*, 26, 29–54. <https://doi.org/10.1007/s00357-009-9027-y>
- Greenacre, M. (2018). *Compositional data analysis in practice*. Chapman and Hall/CRC, <https://doi.org/10.1201/9780429455537>
- Gupta, M. C., & Huefner, R. J. (1972). A cluster analysis study of financial ratios and industry characteristics. *Journal of Accounting Research*, 10(1), 77–95. <https://doi.org/10.2307/2490219>
- Hasidi, M., Baheri, J., & Hajar, K. (2024). Financial performance evaluation using profitability and liquidity ratio analysis. *Jurnal Ilmiah Manajemen Kesatuan*, 12(4), 1347–1358. <https://doi.org/10.37641/jimkes.v12i4.2742>
- Jensen, R. E. (1971). A cluster analysis study of financial performance of selected business firms. *The Accounting Review*, 46(1), 36–56. Retrieved September 18, 2025, from <https://www.jstor.org/stable/243887>
- Jofre-Campuzano, P., & Coenders, G. (2022). Compositional classification of financial statement profiles: The weighted case. *Journal of Risk and Financial Management*, 15(12), 546. <https://doi.org/10.3390/jrfm15120546>
- Kasilingam, R., & Jayabal, G. (2012). Profitability and solvency analysis of a manufacturing company using Dupont and Altman Model. *BVIMR Management Edge*, 5(2). Retrieved September 18, 2025, from https://scholar.google.com/citations?view_op=view_citation&hl=en&user=sbzb-fsAAAAJ&citation_for_view=sbzb-fsAAAAJ:aqVkm33-oC
- Linares-Mustarós, S., Coenders, G., & Vives-Mestres, M. (2018). Financial performance and distress profiles. From classification according to financial ratios to compositional classification. *Advances in Accounting*, 40, 1–10. <https://doi.org/10.1016/j.adiac.2017.10.003>
- Linares-Mustarós, S., Farreras-Noguer, M. À., Arimany-Serrat, N., & Coenders, G. (2022). New financial ratios based on the compositional data methodology. *Axioms*, 11(12), 694. <https://doi.org/10.3390/axioms11120694>
- Nguyễn Thị Minh Phương, Lê Anh Đức, Nguyễn Lê Khánh Duy, Dương Thị Phương Trang, Phạm Linh Chi, & Đỗ Thị Việt Huệ. (2024). Thị trường kinh doanh bất động sản Việt Nam: Thực trạng và giải pháp. Hội thảo Khoa học quốc gia “Thị trường bất động sản Việt Nam trong bối cảnh sửa đổi Luật Đất đai, Nhà ở và Kinh doanh Bất động sản”, 39–58. Truy cập ngày 20/10/2025 từ <https://khoa-hoc.neu.edu.vn/vi/2024-2253/hoi-thao-khoa-hoc-quoc-gia-thi-truong-bat-dong-san-viet-nam-trong-boi-can-h-sua-doi-luat-dat-dai-nha-o-va-kinh-doanh-bat-dong-san>
- Pawlowsky-Glahn, V., & Buccianti, A. (2011). *Compositional data analysis*. Wiley Online Library. <https://doi.org/10.1002/9781119976462>
- Phạm Quốc Duẩn. (2023). Vai trò của thị trường bất động sản trong nền kinh tế và giải pháp ổn định thị trường tại Việt Nam. *Tạp chí Kinh tế và Dự Báo*, 10(4). Truy cập ngày 18/9/2025 từ <https://kinh-te-va-du-bao.vn/vai-tro-cua-thi-truong-bat-dong-san-trong-nen-kinh-te-va-giai-phap-on-dinh-thi-truong-tai-viet-nam-27533.html>
- Sáenz, J. V., Quiroga, F. M., & Bariviera, A. F. (2023). Data vs. information: Using clustering techniques to enhance stock returns forecasting. *International Review of Financial Analysis*, 88, 102657. <https://doi.org/10.1016/j.irfa.2023.102657>
- Saus-Sala, E., Farreras-Noguer, À., Arimany-Serrat, N., & Coenders, G. (2021). Compositional DuPont analysis. A visual tool for strategic financial performance assessment. In P. Filzmoser, K. Hron, J. A. Martín-Fernández, & J. Palarea-Albaladejo (Eds.). *Advances in Compositional Data Analysis: Festschrift in Honour of Vera Pawlowsky-Glahn* (pp. 189–206). Cham: Springer. https://doi.org/10.1007/978-3-030-71175-7_10
- Soliman, M. T. (2008). The use of DuPont analysis by market participants. *The Accounting Review*, 83(3), 823–853. <https://doi.org/10.2308/accr.2008.83.3.823>
- Trần Kim Chung. (2023). Thị trường bất động sản Việt Nam 2013-2023: Mười năm nhìn lại và định hướng 2023- 2030. *Báo Việt Nam Hội Nhập*. Truy cập ngày 18/9/2025 từ <https://vietnamhoinhap.vn/vi/thi-truong-bat-dong-san-viet-nam-2013-2023--muoi-nam-nhin-lai-va-dinh-huong-2023--2030-44286.htm>
- Truong, C. M. (2015). Small and medium sized enterprises: clustering and owner manager's typology. *VNUHCM Journal of Science and Technology Development*, 18(4), 25-33. <https://vjol.info.vn/index.php/JSTD/article/view/23717>