

Exploring the Link Between Education, Governance Quality, and Academic Integrity: A Cross-Country Clustering Analysis

Artem Artyukhov,  ORCID: <https://orcid.org/0000-0003-1112-6891>

Dr of Economics, Research Institute of Trade and Sustainable Business, Faculty of Commerce, University of Economics in Bratislava, Slovakia

Olena Churikanova,  ORCID: <https://orcid.org/0000-0001-5703-2271>

Dr of Economics, Professor, TUBAF Digital, TU Bergakademie Freiberg, Germany

Yuliia Yehorova,  ORCID: <https://orcid.org/0000-0002-8756-4073>

Ph.D., Research Institute of Trade and Sustainable Business, Faculty of Commerce, University of Economics in Bratislava, Slovakia

Corresponding author: Artem Artyukhov, artem.artiukhov@euba.sk

Type of manuscript: research paper

Abstract: *Given the increasing global concerns over academic dishonesty and its implications for institutional credibility, the relationship between education, governance quality, and academic integrity is a crucial study area. This study aims to examine how education levels and governance quality impact academic integrity by clustering countries based on key indicators such as expected years of schooling, mean years of schooling, rule of law index, and document similarity percentage (a proxy for plagiarism). The analysis includes data from over 100 countries, spanning both developing and developed nations, with indicators drawn from sources such as the UNDP Human Development Reports, the World Bank's Rule of Law Index for the year 2022, and aggregated similarity scores from Plagiarism Checker X for the 01.02.2025. These indicators were selected to represent educational attainment, institutional governance quality, and observed academic originality, allowing for a multidimensional statistical clustering approach. The k-means clustering algorithm was applied to identify natural groupings of countries, with the optimal number of clusters determined using the elbow method, silhouette scores, and Davies-Bouldin index to ensure robustness. Using k-means clustering, the research identifies two distinct country groups and analyses their characteristics to explore systemic influences on academic behaviours. The findings reveal that nations with higher education levels and more substantial governance structures exhibit lower similarity percentages, indicating more extraordinary originality and stronger academic integrity practices. Conversely, countries with lower education levels and weaker governance tend to have higher similarity percentages, suggesting challenges in upholding academic integrity. These results highlight the need for targeted interventions to improve governance and educational policies to strengthen academic integrity globally.*

Keywords: academic integrity, clustering analysis, education index, plagiarism, rule of law.

JEL Classification: I25, I28, D73, K42, O57, C38.

Received: 28 November 2024

Accepted: 01 February 2025

Published: 31 March 2025

Funding: This research was funded by the European Union grant “NextGenerationEU through the Recovery and Resilience Plan for Slovakia” (No. 09I03-03-V01-00130) and project VEGA – 1/0392/23 “Changes in the approach to the creation of companies’ distribution management concepts influenced by the effects of social and economic crises caused by the global pandemic and increased security risks”.

Publisher: Academic Research and Publishing UG (i.G.) (Germany).

Founder: Academic Research and Publishing UG (i.G.) (Germany).

Cite as: Artyukhov, A., Churikanova, O., & Yehorova, Yu. (2025). Exploring the Link Between Education, Governance Quality, and Academic Integrity: A Cross-Country Clustering Analysis. *Business Ethics and Leadership*, 9(1), 143–153. [http://doi.org/10.61093/bel.9\(1\).143–153.2025](http://doi.org/10.61093/bel.9(1).143–153.2025).



Copyright: © 2025 by the author. Licensee: Academic Research and Publishing UG (i.G.) (Germany). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

INTRODUCTION

The link between education, governance quality, and academic integrity is significant contemporary relevance, driven by increased global concern about academic dishonesty and its implications for educational quality and institutional credibility. As a cornerstone of quality education, academic integrity directly influences the reliability and authenticity of research outputs, which are critical in informing policymaking and fostering societal trust. Recent studies and reports have highlighted a significant global increase in plagiarism and academic dishonesty, intensified by widespread digital technology use. The International Centre for Academic Integrity has specifically emphasized this rise, noting that such misconduct directly undermines the authenticity and credibility of research outputs.

Consequently, this poses a considerable threat to institutional credibility and can erode public trust in educational and governmental institutions globally (ICAI, n.d.). Leading international organizations, including UNESCO and the OECD, have increasingly emphasized the importance of academic integrity in maintaining quality education and reliable research. UNESCO's initiatives, such as the Global Ethics Observatory, aim to promote ethical standards in science and technology, highlighting integrity and transparency as essential components of effective governance and credible policymaking (UNESCO, n.d.-a). Similarly, the OECD underscores the significance of building a robust evidence base in education to enhance the quality and efficiency of education systems, thereby reinforcing the critical role of academic integrity in fostering societal trust and advancement (OECD, n.d.).

In an era of rapid information exchange, digital accessibility, and cross-border educational collaboration, understanding how governance structures and education systems contribute to or mitigate plagiarism and other integrity-related issues becomes vital. By linking governance quality – institutional effectiveness, transparency, and the rule of law – to educational outcomes, policymakers and educators can better grasp how systemic factors affect academic behaviors. UNESCO has actively promoted academic integrity by supporting the development of ethical guidelines and governance frameworks within educational institutions. For instance, the International Institute for Educational Planning (IIEP, n.d.) has assisted universities in creating charters of ethics, emphasizing the importance of involving faculty and students in their development and implementation. These initiatives highlight how structured governance and clear ethical policies can foster a culture of integrity in education.

Instances of governance lapses have led to significant educational disruptions. For example, Brindabella Christian College in Canberra entered voluntary administration, owing \$8 million to the Australian Taxation Office. Despite receiving substantial taxpayer funding, financial mismanagement resulted in unpaid staff wages and operational instability (Jackson, 2025). This case underscores how inadequate governance can adversely affect educational institutions' integrity and functionality.

The proliferation of generative artificial intelligence (GenAI) tools has significantly impacted academic practices. A recent survey revealed that 92% of UK students have utilized GenAI tools like ChatGPT, with 88% employing them for assessments. This surge has prompted universities to reevaluate assessment methods to uphold academic integrity (Weale, 2025). These developments highlight the critical need for effective governance in educational institutions to uphold academic integrity, especially in the face of emerging technological challenges.

LITERATURE REVIEW

Exploring the interconnectedness between education, governance quality, and academic integrity has become critical in contemporary scholarly research. Existing literature has examined the factors contributing to academic integrity violations, underscoring their implications for institutional governance and educational excellence.

Ahmad and Fauzi (2024) employed a bibliometric analysis to examine trends in plagiarism within academic writing in higher education institutions. Their findings indicate a significant correlation between institutional governance quality and the incidence of plagiarism, suggesting that effective governance structures and robust educational policies are essential for minimizing academic misconduct. Ahsan et al. (2021) conducted a systematic review addressing contract cheating in higher education, identifying governance inadequacies, deficient ethical guidelines, and insufficient institutional accountability as primary contributors to academic integrity breaches. They advocate for establishing clear governance frameworks and heightened ethical standards to mitigate contract cheating and associated integrity violations.

Further contributing to the discussion, Artyukhov et al. (2021) examined Ukrainian universities, highlighting that the quality of scientific activity, technology transfer, and research integrity significantly depends on institutional governance mechanisms. Their research emphasizes the necessity of transparent and

accountable governance practices in sustaining high academic standards. Additionally, Artyukhova et al. (2024) argued that transparency in university leadership, particularly in communication with external stakeholders, substantially influences institutional credibility and academic integrity. They stress that effective governance and leadership transparency are crucial for building stakeholder trust and enhancing institutional integrity. In the context of technological advancements, Balalle and Pannilage (2025) conducted a systematic literature review addressing the impact of artificial intelligence on academic integrity. Their findings underscore the urgent need for adaptive governance structures capable of responding to emerging integrity challenges posed by AI technologies.

Guraya (2018) and Kennet and Shkodkina (2018) presented cross-cultural perspectives on academic integrity. Guraya's comparative analysis of Saudi and UK medical schools revealed significant disparities in recommended sanctions for integrity breaches attributed to distinct governance approaches and cultural contexts. Similarly, Kennet and Shkodkina identified variances in academic integrity violations between the United States and Ukraine, linking these differences to institutional governance and cultural factors. A cross-cultural study by Orosz et al. (2018) examined the relationship between academic cheating and national corruption levels. The findings revealed a strong correlation: countries with higher levels of perceived corruption also reported higher academic dishonesty among students. This suggests that the broader governance environment influences individual ethical behaviors within educational settings. Hammerschmidt (2013) explored Chinese graduate students' understanding of academic integrity within U.S. institutions, demonstrating that governance practices, transparency in expectations, and educational interventions play critical roles in shaping international students' adherence to integrity standards.

Technological interventions to uphold academic integrity were explored by Ison and Szathmary (2016), who demonstrated the efficacy of SafeAssign software in strengthening institutional governance mechanisms through improved plagiarism detection capabilities. The rise of artificial intelligence in education has introduced new dimensions to the discourse on academic integrity. A study by Ateeq et al. (2024) examines the implications of AI on academic integrity and the shift toward holistic assessment methods. The research reveals a strong link between the educational impact of AI and academic success, highlighting that while AI offers personalized learning experiences, it also necessitates the development of policies and ethical guidelines to ensure responsible use.

The influence of governance on academic outcomes and student behavior was further supported by Mohamed Nafuri et al. (2022), whose clustering analysis linked governance quality directly to student academic performance and integrity practices. Nguyen and Tuamsuk (2024) similarly emphasized governance quality's impact on scientific integrity within scholarly publications, using bibliometric methods to assess governance effectiveness. Recent studies highlight the significance of institutional governance in promoting academic integrity.

For instance, a chapter by Andrews and Glendinning (2024) emphasizes the essential role governance bodies within universities play in ensuring the quality of academic activities. The chapter delves into how effective governance frameworks can mitigate issues related to academic dishonesty, thereby enhancing the overall educational experience.

Shaghghi (2021) examined macro-social factors contributing to plagiarism in Iranian universities, identifying governance quality, policy efficacy, and cultural influences as pivotal determinants of academic dishonesty. The study advocates comprehensive governance reforms to address integrity issues effectively. Lastly, Zhao et al. (2024) conducted a meta-analysis highlighting the interaction between cultural values, achievement orientations, and governance quality, finding that effective governance significantly mediates educational aspirations and the prevalence of academic cheating.

These studies collectively emphasize that effective governance practices, transparent leadership, robust policy frameworks, and culturally informed strategies are indispensable in promoting academic integrity. Addressing these interrelated factors is critical for establishing and sustaining ethical standards, accountability, and excellence within higher education institutions globally.

This study aims to examine the relationship between education, governance quality, and the prevalence of document similarity (a proxy for plagiarism) across different countries. The research intends explicitly to classify countries into meaningful clusters based on indicators such as expected years of schooling, mean years of schooling, rule of law index, and document similarity percentage.

Using *k*-means clustering, the study identifies distinct groups of countries, analyzes their characteristics, and explores how governance and education levels impact academic integrity and originality. The ultimate objective is to offer insights for educators, researchers, and policymakers to understand global academic integrity issues better and suggest targeted strategies for improving education, governance, and academic standards across nations.

METHODOLOGY

Data Collection

The study utilizes a dataset containing country-level indicators related to education and governance. The dataset was compiled from publicly available sources such as the World Bank, the United Nations Development Program (UNDP), and academic integrity reports. Data preprocessing was conducted in R Studio to clean missing values, normalize variables, and prepare them for clustering analysis. The primary variables include:

- Expected years of schooling (x_1) measures the number of years a child is expected to spend in the education system (UNDP, n.d.); the recent available datasets are related to 2022.
- Mean years of schooling (x_2) represents the average number of years of schooling completed by individuals aged 25 and above (UNDP, n.d.); the recent available datasets are related to 2022.
- Rule of law index (x_3) assesses governance quality by evaluating adherence to laws, institutional stability, and public trust (World Bank, n.d.); similar to the previous indicator, the data from 2022 were chosen.
- Similarity percentage (y) represents the extent of document similarity (Plagiarism Checker X, n.d.). The Global Plagiarism Live Statistics feature aggregates data from users worldwide who utilize Plagiarism Checker X to scan their documents. By analyzing this anonymized data, the platform provides real-time insights into the prevalence of text similarity across different countries and regions. This information can be valuable for educators, researchers, and policymakers aiming to understand and address global academic integrity issues. The datasets were downloaded on 01.02.2025.

Clustering Technique

The k -means clustering algorithm was employed to identify meaningful country groupings based on education and governance quality. k -means was chosen because it efficiently partitioned datasets into distinct, non-overlapping clusters based on Euclidean distances.

Optimal Cluster Selection

To determine the optimal number of clusters (k), the following methods were applied:

- The elbow method plots the within-cluster sum of squares (WCSS) against different values of k to identify the “elbow point”, where adding more clusters does not significantly reduce variance.
- The silhouette method measures the silhouette width to evaluate how well-separated clusters are, with higher values indicating better clustering quality.
- The Davies-Bouldin Index (DBI) is calculated to assess cluster compactness and separation. A lower DBI indicates better clustering performance. These methods determine the optimal number of clusters, providing the best balance between intra-cluster cohesion and inter-cluster separation.

Cluster Interpretation

After identifying the number and list of the countries included in clusters, their characteristics are analyzed using descriptive statistics. To visualize differences between clusters, boxplots are generated to compare key variables across groups.

Validation and Robustness Checks

To quantify reproducibility, the Adjusted Rand Index (ARI) is used; it measures how similar two clustering results are. If $ARI = 1$, that means perfect agreement between two runs; $ARI \approx 0$ means no agreement (random clustering); $ARI < 0$ means worse than random clustering. If ARI is high (above 0.8), it indicates strong reproducibility. If it is low (below 0.5), clustering may not be stable.

Software and Tools

All data processing, clustering analysis, and visualizations were conducted using R Studio. This methodology ensures that the clustering approach is robust and interpretable and provides meaningful insights into the relationship between education, governance, and document similarity across countries.

RESULTS

The elbow method plot helps determine the optimal number of clusters (k) for k -means clustering (Figure 1). The y-axis represents the total within-cluster sum of squares (WCSS), which measures the variance within each cluster. The x-axis represents the number of clusters (k). The “elbow point” is where the WCSS decreases slower (forming an elbow shape).

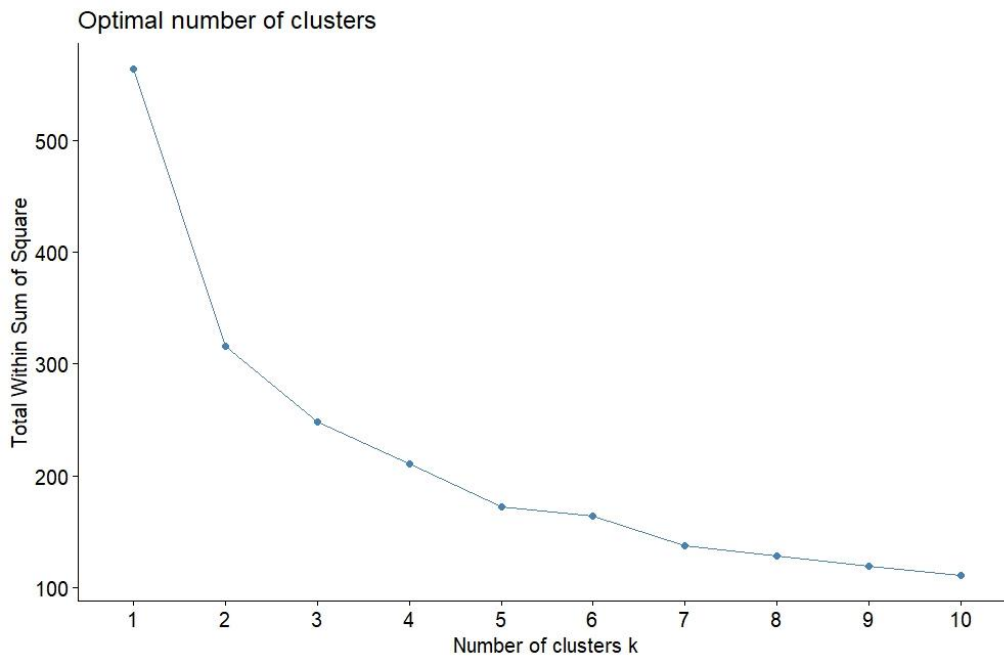


Figure 1. The Elbow Method Plot

Source: authors' calculation in R Studio.

From Figure 1, the elbow appears to be at $k = 3$ or $k = 4$ because the drop in WCSS is significant from $k = 1$ to $k = 3$, and after $k = 3$ or $k = 4$, the reduction in WCSS slows down, indicating diminishing returns.

The silhouette method plot (Figure 2) helps determine the optimal number of clusters (k) based on the average silhouette width, which measures how well-separated the clusters are. The y-axis represents the average silhouette width, which measures clustering quality (higher values are better). The x-axis represents the number of clusters (k). The optimal k is the number of clusters that results in the highest silhouette score.

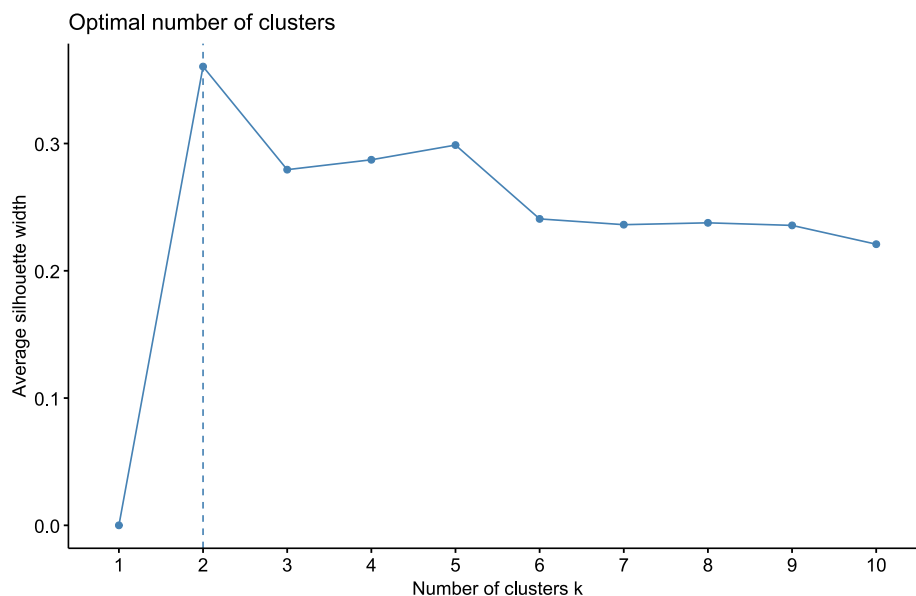


Figure 2. The Silhouette Method Plot

Source: authors' calculation in R Studio.

From Figure 2, the highest silhouette width is at $k = 2$. The silhouette width decreases after $k = 2$, indicating that adding more clusters might not improve the clustering quality.

The elbow and the silhouette methods are commonly used to determine the optimal number of clusters in k -means, but they evaluate clustering from different perspectives. If the goal is to minimize within-cluster variance (better representation of natural groupings), the elbow method is preferred. The silhouette method is more useful if the goal is to maximize inter-cluster separation (ensuring clusters are distinct). If they suggest different k values, it is best to use additional validation methods like the Davies-Bouldin index (Table 1).

Table 1. DBI Values for Different Numbers of Clusters

<i>k</i> (Number of Clusters)	DBI Value
<i>k</i> = 2	1.1243 (Lowest)
<i>k</i> = 3	1.346
<i>k</i> = 4	1.242

Note: DBI = Davies-Bouldin Index; *k* = cluster.

Source: authors' calculation in R Studio.

Since the Davies-Bouldin Index (DBI) is lowest for *k* = 2 (1.124), this suggests that 2 clusters provide the best separation and compactness for data (see Table 1).

The cluster plot (Figure 3) shows the two clusters found using *k*-means (*k* = 2), based on the optimal Davies-Bouldin index (DBI) and silhouette method.

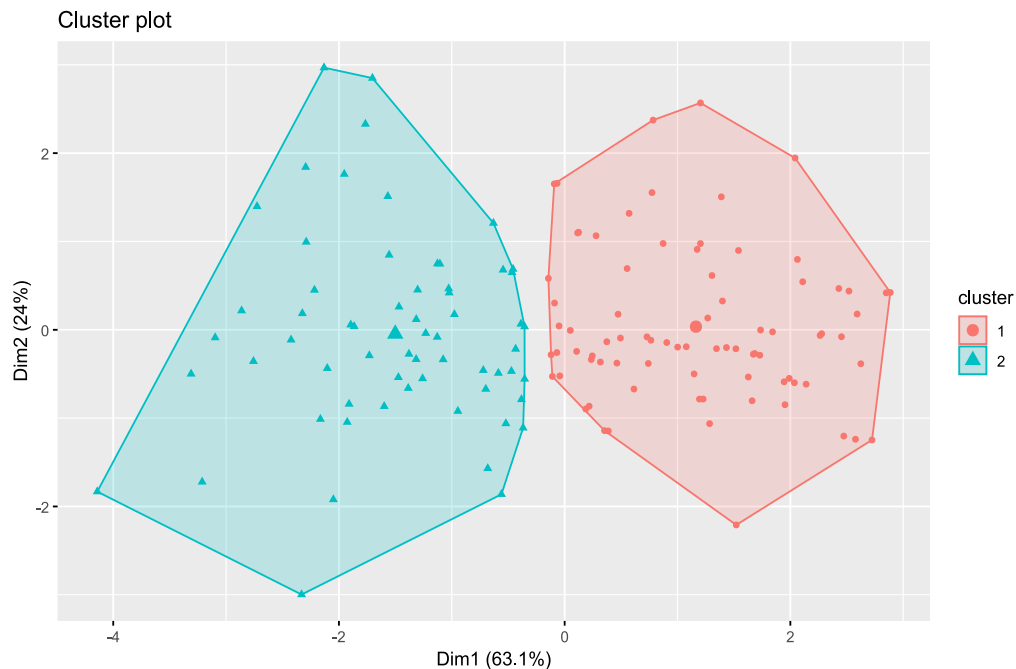


Figure 3. The Cluster Plot

Source: authors' calculation in R Studio.

The clustering analysis has revealed two distinct groups of countries based on their education levels and governance quality. The results clearly link a nation's educational development and governance strength (Table 2).

Table 2. Final Cluster Assignment Summary

Cluster	Description	Countries
1 (Red)	Developing Nations with Educational Challenges	Afghanistan, Algeria, Angola, Bangladesh, Belize, Benin, Burkina Faso, Cambodia, Cameroon, Congo, Egypt, El Salvador, Eritrea, Ethiopia, Gambia, Ghana, Guatemala, Guinea, Guyana, Honduras, India, Indonesia, Iran, Iraq, Jamaica, Kenya, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Maldives, Mali, Mexico, Morocco, Mozambique, Myanmar, Namibia, Nepal, Niger, Nigeria, Pakistan, Palestine, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Rwanda, Senegal, Sierra Leone, Somalia, Sudan, Syria, Tanzania, Uganda, Ukraine, Venezuela, Vietnam, Zambia, Zimbabwe
2 (Green)	High-Education, Strong-Governance	Albania, Andorra, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahrain, Barbados, Belgium, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czechia, Denmark, Dominican Republic, Ecuador, Estonia, Fiji, Finland, France, Georgia, Germany, Greece, Hong Kong, Hungary, Iceland, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Kuwait, Latvia, Lithuania, Luxembourg, Malaysia, Malta, Mauritius, Montenegro, Netherlands, New Zealand, North Macedonia, Norway, Oman, Poland, Portugal, Qatar, Romania, Russia, Saudi Arabia, Serbia, Seychelles, Singapore, Slovakia, Slovenia, South Africa, South Korea, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Trinidad And Tobago, Tunisia, Turkey, United Arab Emirates, United Kingdom, United States, Uruguay

Source: authors' calculation in R Studio.

As demonstrated in Table 2, the clustering analysis highlights a strong correlation between education levels, governance quality, and the similarity percentage in checked documents across countries. Nations with higher education levels and stronger governance tend to have a lower similarity percentage, indicating a greater emphasis on originality, research integrity, and academic rigor. In contrast, countries with lower education levels and weaker governance often exhibit higher similarity percentages, suggesting challenges in academic standards, institutional policies, and enforcement mechanisms to ensure originality in educational and governmental practices.

Countries in the high-education, strong-governance cluster invest in education and institutional stability and foster a culture of independent thinking and originality, reflected in lower similarity percentages in document checks. This suggests that stronger governance and well-developed education systems contribute to higher academic integrity and stricter policies against plagiarism. On the other hand, the developing nations with educational challenges cluster show higher similarity percentages, indicating that weaker education systems and governance structures may struggle with plagiarism detection, enforcement of academic integrity policies, and promoting original research.

The clusters show distinct differences in education and governance quality (Table 3). Cluster 1 represents countries with higher education levels and stronger governance. These nations have an average of 15.9 expected years of schooling, 11.5 mean years of schooling, and a higher rule of law score (0.629). This suggests a well-established educational system and stable governance. A suitable name for this group could be “High-Education, Strong-Governance Nations”.

Table 3. Cluster Interpretation Based on Mean Values

Cluster	Similarity, % (y)	Expected Years of Schooling, (x1)	Mean Years of Schooling, (x2)	Rule of Law (x3)
Cluster 1	0.132	15.9 years	11.5 years	0.629 (Stronger governance)
Cluster 2	0.151	11.3 years	6.24 years	0.343 (Weaker governance)

Source: authors' calculation in R Studio.

As shown in Table 3, cluster 2 consists of countries with lower education levels and weaker governance. The average expected years of schooling is 11.3, while the mean years of schooling is only 6.24. The rule of law score is significantly lower at 0.343, indicating weaker institutions and governance structures. This cluster represents developing nations with educational and governance challenges, making “Developing Nations with Educational Challenges” a fitting name.

The boxplot (Figure 4a) illustrates the differences in expected years of schooling (x1) between the two clusters. Countries in the high-education and strong-governance cluster have a higher median expected schooling duration, averaging around 15 to 16 years, with some countries exceeding 20 years. In contrast, the developing nations with educational challenges cluster shows significantly lower expected schooling, with a median of approximately 11 years and a broader spread, including some countries falling below 10 years. This difference highlights the strong link between governance quality and educational opportunities. Countries with better governance tend to provide more years of education, while those with governance challenges often struggle with lower schooling expectations.

The boxplot (Figure 4b) compares the mean years of schooling (x2) between the two clusters. Countries in high-education, strong-governance cluster have higher median mean years of schooling, typically around 10 to 12 years, indicating a well-established education system where most of the population completes secondary education. In contrast, the developing nations with educational challenges cluster exhibits a lower median of around 5 to 6 years, with some countries falling below this level, suggesting significant educational access and retention issues. The disparity between these clusters reinforces the connection between governance quality and actual educational attainment, where countries with stronger governance provide more excellent educational opportunities.

The boxplot (Figure 4c) visually compares the rule of law (x3) between the two identified clusters. Countries in the high-education, strong-governance cluster generally have higher rule of law scores, with most values ranging between 0.5 and 0.8. The median score is above 0.6, indicating stable institutions and stronger governance. In contrast, the developing nations with educational challenges cluster has lower rule of law scores, mostly between 0.2 and 0.5, with a median of around 0.35. This suggests weaker institutions and governance structures. Some outliers appear below 0.1, representing countries with extreme governance challenges. The clear separation between clusters highlights the strong relationship between education levels and governance quality. Countries with better education systems tend to have stronger governance, while those with limited schooling often face governance issues.

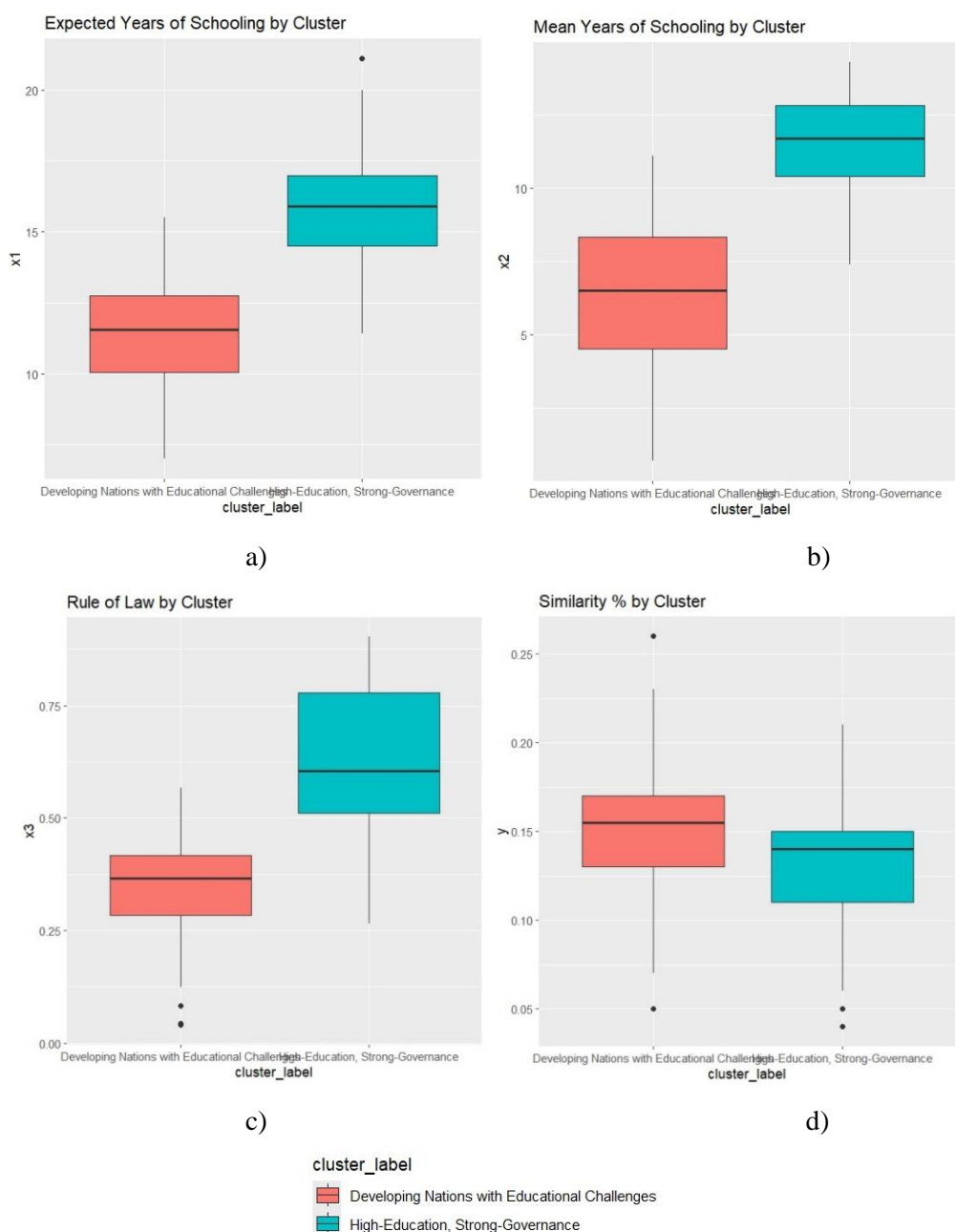


Figure 4. The Boxplot Compares the Different Variables Between the Two Clusters

Source: authors' calculation in R Studio.

The boxplot (Figure 4d) compares the two clusters' similarity percentages (y). The developing nations with educational challenges cluster has a slightly higher median similarity percentage than the high-education, strong-governance cluster. The interquartile range for both clusters is relatively close, but the developing nations show a wider spread, with several outliers both above and below the main range. This suggests that while the similarity percentage varies across both groups, countries in the developing cluster tend to have more significant variation, possibly due to differing governance and education quality. The more developed cluster appears more consistent in similarity percentage, indicating more uniform governance and education structures.

Addressing these disparities requires a multi-faceted approach. Developing nations should focus on improving education infrastructure, strengthening institutional policies on academic integrity, and enhancing governance mechanisms to regulate originality in research and documentation. Investing in advanced plagiarism detection tools, teacher training, and strict enforcement policies can help reduce similarity percentages, fostering a more original and ethical academic environment. Nations with strong governance and education systems should continue supporting innovation, upholding strict academic integrity policies, and collaborating internationally to maintain high research and documentation standards.

From a global perspective, tackling educational and governance disparities alongside similarity percentage concerns is crucial for long-term development. Countries with higher similarity percentages should implement targeted interventions, including curriculum improvements, awareness programs on research ethics, and technological advancements in plagiarism detection. More developed nations should play a role in supporting initiatives that encourage academic originality and research transparency worldwide.

Future research should explore how similarity percentage correlates with academic performance, research output, and governance efficiency, providing further insights into the broader implications of originality and institutional integrity in education and governance systems. The Adjusted Rand Index (ARI) score is 1, which means there is a perfect agreement between the two clustering results. This indicates that *k*-means clustering is highly stable and reproducible, meaning that the same cluster assignments are produced even when the random seed is changed.

DISCUSSION

The clustering analysis in this study identified two distinct groups of countries based on education levels, governance quality, and document similarity percentage. The first cluster, labelled “High-Education, Strong-Governance Nations”, comprises countries with higher expected years of schooling, stronger governance (higher rule of law index scores), and lower document similarity percentages. In contrast, the second cluster, “Developing Nations with Educational Challenges”, consists of countries with lower schooling levels, weaker governance, and higher similarity percentages, suggesting a greater prevalence of academic dishonesty. These findings align with and expand upon previous research exploring the intersection of governance, education, and academic integrity.

The results confirm Ahmad and Fauzi (2024), who demonstrated a strong correlation between institutional governance quality and the prevalence of plagiarism. This study similarly finds that countries with weaker governance structures have a higher document similarity percentage, reinforcing the argument that effective governance mechanisms are essential in reducing academic misconduct. The clustering analysis in this study provides additional empirical support by showing a clear division between well-governed nations with low similarity percentages and poorly governed nations with high similarity percentages.

The findings are also consistent with Ahsan et al. (2021), who highlighted governance inadequacies and the lack of institutional accountability as primary contributors to contract cheating. Cluster 2 (developing nations with educational challenges) exhibits lower rule of law index scores in this study, reflecting weaker enforcement mechanisms and governance deficiencies, which may explain the higher similarity percentages. This supports the notion that institutional weaknesses directly contribute to breaches in academic integrity.

Moreover, Orosz et al. (2018) linked higher levels of national corruption with increased academic dishonesty. These results reinforce this finding, as countries with lower rule of law index scores (an indicator of corruption and weak governance) tend to have higher document similarity percentages. This suggests that a nation’s broader governance environment influences individual ethical behaviors within educational settings.

This study further complements the work of Artyukhov et al. (2021), who emphasized the role of transparent and accountable governance practices in sustaining high academic standards. The results suggest that countries in Cluster 1 (high-education, strong-governance nations), which demonstrate stronger governance and transparency, also tend to have lower document similarity percentages. This reinforces the claim that academic integrity is closely tied to governance stability. Another key comparison arises from Shaghghi (2021), who investigated macro-social factors contributing to plagiarism in Iranian universities. The study identified policy inefficacy and cultural influences as significant determinants of academic dishonesty. While this study does not explicitly address cultural factors, it does support the argument that policy effectiveness (as reflected in governance quality) significantly impacts plagiarism prevalence.

The rise of Generative AI (GenAI) has introduced new challenges in academic integrity, as Balalle and Pannilage (2025) noted. Their study underscores the urgent need for adaptive governance structures to address emerging integrity concerns. However, this study’s clustering analysis suggests that governance quality and educational development remain fundamental determinants of academic honesty, irrespective of AI advancements. Countries with weaker governance structures and lower education levels are likely to struggle more with AI-related academic dishonesty, reinforcing the need for stronger institutional regulations and technological oversight in these regions.

CONCLUSIONS

This study aimed to investigate the relationship between educational achievement, governance quality, and document similarity (indicative of plagiarism prevalence) across different nations. The research

specifically aimed to classify countries into meaningful clusters to understand global patterns in academic integrity better.

Employing a robust methodology, the study utilized secondary data sourced from international organizations, such as the World Bank and UNDP, complemented by aggregated similarity percentage data from Plagiarism Checker X. The *k*-means clustering algorithm was applied, with optimal clusters determined through the elbow, silhouette, and Davies-Bouldin index methods, ensuring a comprehensive and statistically rigorous analysis. The main findings revealed two distinct clusters. The first cluster labelled “High-Education, Strong-Governance”, featured countries with higher education levels, more substantial governance structures, and lower similarity percentages, indicating greater academic integrity and originality. Conversely, the second cluster, “Developing Nations with Educational Challenges”, exhibited lower education levels, weaker governance, and higher document similarity percentages, reflecting potential deficiencies in educational quality and governance practices.

These findings offer valuable outputs for regulations and governance, highlighting the need for targeted interventions. For developing nations, enhancing educational infrastructure, strengthening institutional governance, and adopting stringent academic integrity policies are crucial steps. Meanwhile, countries identified within the high-education and strong-governance cluster should continue reinforcing existing standards, fostering international collaboration, and supporting global initiatives to universally improve educational and governance outcomes. Ultimately, the results underscore the importance of integrating education quality improvement and governance enhancement policies to cultivate originality and academic integrity on a global scale.

Limitation

This research has several limitations that should be considered when interpreting the findings. Firstly, the clustering analysis relies on available secondary data from international organizations such as the World Bank and UNDP, which might not fully capture country-specific nuances or real-time changes in educational and governance conditions. While widely accepted, the indicators used may not entirely represent the complexities inherent in education systems and governance quality. Secondly, the measure of document similarity, derived from aggregated global user data via Plagiarism Checker X, may contain biases related to the self-selection of users and the unequal distribution of software use across countries. Consequently, similarity percentages might not perfectly reflect nationwide plagiarism levels or academic integrity practices.

Additionally, the choice of clustering method (*k*-means) imposes assumptions such as spherical-shaped clusters and equal cluster variances, potentially overlooking other meaningful groupings. Alternative clustering methods could yield different insights or highlight different relationships among variables. Lastly, the cross-sectional nature of the dataset does not allow for causal inferences or exploration of trends over time. Future research should employ longitudinal analyses, incorporate qualitative data for deeper insights, and potentially explore additional variables influencing educational and governance outcomes.

Author Contributions

Conceptualization: A. A., O. C., Y. Y.; data curation: A. A., O. C.; formal analysis: A. A., Y. Y.; funding acquisition: Y. Y.; investigation: A. A., O. C.; methodology: A. A.; project administration: A. A.; resources: O. C., Y. Y.; software: Y. Y.; supervision: A. A.; validation: A. A., O. C.; visualization: Y. Y.; writing – original draft: A. A., O. C., Y. Y.; writing – review & editing: A. A., O. C., Y. Y.

Conflicts of Interest

No conflict to declare.

Data Availability Statement

Not applicable.

Informed Consent Statement

Not applicable.

References

1. Ahmad, H., & Fauzi, M. A. (2024). Plagiarism in Academic Writing in Higher Education Institutions: A Bibliometric Analysis. *International Journal on Social and Education Sciences*, 6(1), 64–84. [\[CrossRef\]](#)
2. Ahsan, K., Akbar, S., & Kam, B. (2021). Contract cheating in higher education: a systematic literature review and future research agenda. *Assessment & Evaluation in Higher Education*, 47(4), 523–539. [\[CrossRef\]](#)

3. Andrews, S., & Glendinning, I. (2024). Governing Academic Integrity: Conceptualizing the Assurance and Efficacy of Strategies and Outcomes. In Eaton, S.E. (eds) *Second Handbook of Academic Integrity*. Springer International Handbooks of Education. Springer, Cham. [\[CrossRef\]](#)
4. Artyukhov, A., Wołowiec, T., Artyukhova, N., Bogacki, S., & Vasylieva, T. (2024). SDG 4, Academic Integrity and Artificial Intelligence: Clash or Win-Win Cooperation? *Sustainability*, 16(19), 8483. [\[CrossRef\]](#)
5. Artyukhova, N., Vorontsova, A., Artyukhov, A., Yehorova, Y., Vasić, S., Rubanov, P., & Vasylieva, T. (2024). Analysis of trends in the structure of higher education market of European countries. *Knowledge and Performance Management*, 8(2), 91–108. [\[CrossRef\]](#)
6. Ateeq, A., Alzoraiki, M., Milhem, M., & Ateeq, R. A. (2024). Artificial intelligence in education: Implications for academic integrity and the shift toward holistic assessment. *Frontiers in Education*, 9. [\[CrossRef\]](#)
7. Balalle, H., & Pannilage, S. (2025). Reassessing academic integrity in the age of AI: A systematic literature review on AI and academic integrity. *Social Sciences & Humanities Open*, 11, 101299. [\[CrossRef\]](#)
8. Guraya, S. Y. (2018). Comparing recommended sanctions for lapses of academic integrity measured by Dundee Polyprofessionalism Inventory I: Academic integrity from a Saudi and a UK medical school. *Journal of the Chinese Medical Association*, 81(9), 787–795. [\[CrossRef\]](#)
9. Hammerschmidt, J. E. (2013). *An Investigation of Chinese Graduate Student Understanding of Academic Integrity in U's Higher Education*. Dissertations. 519. [\[Link\]](#)
10. International Institute for Educational Planning Unesco (IIEP). (n.d.). *Platform on ethics and corruption in education*. IIEP Unesco – Etico. [\[Link\]](#)
11. International Center for Academic Integrity (ICAI). (n.d.). *Facts & Statistics*. ICAI. [\[Link\]](#)
12. Ison, D. C., & Szathmary, K. J. (2016). Assessing Academic Integrity Using SafeAssign Plagiarism Detection Software. *Collegiate Aviation Review International*, 34(1). [\[CrossRef\]](#)
13. Jackson, B. (2025). *Private Christian school dodges bankruptcy by entering voluntary administration*. News. [\[Link\]](#)
14. Kennet, A. M., & Shkodkina, Yu. (2018). A Review of the Factors behind Academic Integrity Violations: Comparing the United States and Ukraine. *Business Ethics and Leadership*, 2(2), 84–96. [\[CrossRef\]](#)
15. Mohamed Nafuri, A. F., Sani, N. S., Zainudin, N. F. A., Rahman, A. H. A., & Aliff, M. (2022). Clustering Analysis for Classifying Student Academic Performance in Higher Education. *Applied Sciences*, 12(19), 9467. [\[CrossRef\]](#)
16. Nguyen, L. T., & Tuamsuk, K. (2024). Unveiling scientific integrity in scholarly publications: a bibliometric approach. *International Journal for Educational Integrity*, 20(1). [\[CrossRef\]](#)
17. OECD. (n.d.). *Education research*. [\[Link\]](#)
18. Orosz, G., Tóth-Király, I., Bóthe, B., Paskuj, B., Berkics, M., Fülöp, M., & Roland-Lévy, C. (2018). Linking cheating in school and corruption. *European Review of Applied Psychology*, 68(2), 89–97. [\[CrossRef\]](#)
19. Plagiarism Checker X. (n.d.). *Global Plagiarism Live statistics*. Plagiarism Checker X – Text Similarity Detector. [\[Link\]](#)
20. Shaghghi, M. (2021). Interactional Effects of Macro Scale Social Factors in the Plagiarism Phenomenon: A Case of Iran. *Iranian Journal of Information Processing and Management*, 37(2), 411–438. [\[CrossRef\]](#)
21. Sharkey, J. (2025). *AI cheating surges at universities*. The Times & The Sunday Times. [\[Link\]](#)
22. UNESCO. (n.d.-a). *Ethics of Artificial Intelligence*. UNESCO.org. [\[Link\]](#)
23. UNDP. (n.d.-b). *Human Development Index*. UNDP | Human Development Reports. [\[Link\]](#)
24. Weale, S. (2025). *UK universities warned to 'stress-test' assessments as 92% of students use AI*. The Guardian. [\[Link\]](#)
25. World Bank. (n.d.). *World Bank Open Data*. World Bank Open Data. [\[Link\]](#)
26. Zhao, L., Yang, X., Yu, X., Zheng, J., Mao, H., Fu, G., Fang, F., & Lee, K. (2024). Academic Cheating, Achievement Orientations, and Culture Values: A Meta-Analysis. *Review of Educational Research*, 0(0). [\[CrossRef\]](#)