An Integrated Evaluation of Key Data Technologies: A Focus on Structure, Analytics, and Emerging Trends

Balogun Ghaniyyat Bolanle(1)
Maryam Mero Woru(2)
Abdulrafiu Muhammed Jamiu(1)
Adebakin Taibat Olaoluwa(1)

Received: 14/١٠/2025
Revised: ١٨/١١/2025
Accepted: 09/12/2025

[1]Department of Computer Science, University of Ilorin, Nigeria
[2]Department of Computer Science. Kwara State College of Education, Ilorin. Nigeria.
*Corresponding Author's Email: balogun.gb@unilorin.edu.ng

An Integrated Evaluation of Key Data Technologies: A Focus on Structure, Analytics, and Emerging Trends

[1]Balogun Ghaniyyat Bolanle, [1]Maryam Mero Woru, [1]Abdulrafiu Muhammed Jamiu, [2]Adebakin Taibat Olaoluwa
[1]Department of Computer Science,University of Ilorin, Nigeria.
[2]Department of Computer Science, Kwara State College of Education, Ilorin. Nigeria.
[1]**balogun.gb@unilorin.edu.ng**

Abstract— This study compares five key data science domains: data structures, data analytics, big data, data mining, and database management systems (DBMS). As data-driven technologies have an increasing impact on corporate intelligence, automation, and academic research, it is vital to analyze their changing importance. A scoring framework was used using simulated yet realistic data to measure both current relevance and future development potential across fields. These trends were highlighted using grouped bar charts as visualizations. According to the report, while core areas such as data structures and DBMS remain vital, technologies such as big data and data analytics have the potential for significant growth. Data mining has also gained increased relevance as a result of its interaction with machine learning. These findings provide actionable insights for educators, academics, and industry experts who want to connect their skills and investments with future technological demands.

Keywords— *data structures, data analytics, big data, data mining, database management systems*

INTRODUCTION

The data revolution is transforming industries, research methodology, governance models, and societal structures around the world. At the center of this transition is the emerging subject of data science, which comprises a wide range of tools, concepts, and approaches for extracting knowledge and insights from vast, complicated datasets (Provost & Fawcett, 2013). Its core pillars include five interconnected but independent domains: data structures, data analytics, big data, data mining, and database management systems. These regions not only help with data storage and manipulation, but they also promote innovation through advanced analytics, pattern recognition, and predictive modeling.

The increasing reliance on digital infrastructures and connected devices has led to an unprecedented growth in data volumes. The International Data Corporation (IDC) estimates that by 2025, the global datasphere will exceed 175 zettabytes (Reinsel, Gantz, & Rydning, 2018). This surge has intensified the demand for robust and scalable data systems, particularly those that can process and analyze unstructured, semi-structured, and real-time data. Big Data technologies have thus become indispensable in industries such as healthcare, finance, cybersecurity, retail, and public governance (Tosi, Kokaj, & Roccetti, 2024; Gandomi & Haider, 2015). Similarly, data analytics has gained prominence as organizations seek to make data-driven decisions and enhance strategic planning through visualization, statistical modeling, and machine learning algorithms (Waller & Fawcett, 2013; Romero & Ventura, 2024).

Meanwhile, data mining continues to serve as a bridge between traditional querying methods and modern artificial intelligence systems. It facilitates the discovery of hidden patterns and associations in large datasets, a task critical to fraud detection, customer segmentation, and recommendation systems (Han, Pei, & Kamber, 2011; Kotu & Deshpande, 2019). DBMSs, while considered mature, remain essential for the efficient management, storage, and retrieval of data in structured formats. Their stability and reliability underpin a vast range of enterprise and academic applications (Silberschatz, Korth, & Sudarshan, 2019; Dataversity, 2022). Likewise, data structures, though often considered a lower-level concern, are integral to computational performance. Efficient data structures are vital to optimizing algorithms, minimizing memory consumption, and ensuring scalability (Cormen, Leiserson, Rivest, & Stein, 2009; Lafore, 2018).

Despite their individual importance, there is limited comparative literature that evaluates these domains within a unified framework, especially in terms of their current industrial relevance and anticipated future trajectory. Existing works tend to isolate a single domain—for instance, focusing solely on advancements in Big Data infrastructure (Hashem et al., 2015) or the pedagogical evolution of data structures education (Paterson, 2015). As scientific ecosystems evolve rapidly, it is critical to assess how these core fields intersect, diverge, and transform in light of new developments such as edge computing, AI ethics, explainable AI, blockchain integration, and quantum computing (Dwivedi et al., 2021; Bhatnagar & Sharma, 2023; Khajeh-Hosseini et al., 2020).
Furthermore, the worldwide educational scene is adjusting to this change. Top-tier colleges are reshaping their computer science programs to prioritize data-intensive topics (Chakraborty et al., 2023).

Big Data, analytics, and mining skills are becoming prioritized above traditional systems knowledge in industrial certifications and educational programs. This change is reflected in the skill requirements of employers: job advertisements on sites like Indeed and LinkedIn frequently highlight a strong demand for AI/ML experts, data scientists, and data engineers who are skilled with technologies like Spark, TensorFlow, SQL, and Tableau. Therefore, strategic investment, curriculum development, and workforce planning all depend on having a thorough understanding of each area's current state as well as its anticipated future.

In order to meet this demand, the current study develops a comparative analytical framework that evaluates the five categories based on two standards: present significance and potential for future expansion. We assess each domain's present level of integration into practice and predict its future development over the next ten years using simulated data gathered from industry reports, academic trends, and expert evaluations. The approach used ensures that the marks assigned to each topic

fairly represent its academic and real-world significance. Unlike previous research that only focuses on algorithmic innovation or technical depth, this work takes a comprehensive, interdisciplinary approach. It seeks to provide a comprehensive view of the data science ecosystem by integrating business evaluations, academic publishing patterns, and curriculum development. This is especially useful for educators creating curricula that meet industrial standards and for assisting funding organizations and scholars in identifying high-impact areas of investigation. This study is unique in that it conducts a comprehensive and comparative examination of five fundamental data science domains using simulated indicators of current importance and anticipated growth in order to offer educators, practitioners, and policymakers strategic insights.

METHODOLOGY

This research used an analysis approach to assess the current importance and future potential of five key fields in data science: data structures, data analytics, big data, data mining, and database management systems (DBMS). The major goal was to measure and depict the current relevance of these disciplines in academic and industrial settings, as well as project their future trajectory based on technological advances and market forecasts.

To achieve this, two evaluation measures were created: current importance and future development potential. The current importance measure assesses how frequently and extensively each area is used in real-world applications, academic curriculum, and professional development programs. It reflects the domain's fundamental or operational necessity, notably in areas like software architecture, application design, and enterprise data management. In contrast, the future growth potential statistic sought to quantify the expected expansion of each domain over the next five to 10 years. This projection was based on data from recent research trends, funding levels, employment market needs, and strategic technological forecasts. The level of integration with developing technologies such as artificial intelligence, cloud computing, automation, and business intelligence tools was taken into account.

Next, two scores—one for present significance and one for future growth—were given to each domain on a normalized scale ranging from 0 to 100. These values were synthesized from several sources using expert-informed approximations rather than being taken from a single dataset. Academic publishing volumes from major repositories (including IEEE Xplore and the ACM Digital Library), employment trend databases, market analysis reports (like those from Gartner and McKinsey), and course materials from well-regarded international universities were among the data sources.

It is crucial to remember that the scores utilized in this study are not actual measurements; rather, they were generated for demonstrative purposes. Instead of using raw datasets, the simulated data was produced using heuristic predictions based on public trends and expert knowledge. Therefore, this study's goal is to illustrate the analytical framework rather than provide final rankings. To validate the concept, future research can use actual data from academic libraries, industry reports, or extensive polls.

The scores for each domain were as follows: Data Structures (present importance: 75; future growth: 60), Data Analytics (85; 80), Big Data (90; 95), Data Mining (80; 85), and Database Management Systems (70; 65). These scores were then shown using a grouped bar chart, with each domain displayed on the horizontal axis and the accompanying metric values placed in neighboring bars. This approach allowed for a direct visual comparison of the current importance and expected growth for each subject, providing insight into changing trends in data technology.

The Python programming language, and more especially the Matplotlib module, which was selected for its adaptability and aptitude for producing plots of publishing quality, was used to handle and visualize all data. Even though the study's data were simulated, the analytical framework is scalable and structured, making it suitable for use with real datasets in subsequent research. Therefore, the methodology described here offers a reliable and repeatable way to evaluate the relative development of technological fields in the data sciences.

RESULTS

The comparative analysis revealed notable differences in how the five evaluated data domains—Data Structures, Data Analytics, Big Data, Data Mining, and Database Management Systems (DBMS)—perform in terms of current importance and projected growth. These are illustrated across four visual figures, each capturing unique aspects of the domain comparisons.

By directly comparing each subject across two metrics—current importance and future growth—Figure 1 offers a fundamental picture. Big Data's crucial position in contemporary information ecosystems was indicated by its greatest future growth score of 95 and one of the highest current relevance scores of 90. With impressive results in all metrics (85 and 80, respectively), data analytics came in second, demonstrating its dual applicability to both present practice and upcoming business demands. Data mining demonstrated a balanced performance (growth of 85; importance of 80), highlighting its continued use in machine learning and pattern recognition applications. Data structures, a theoretical and essential cornerstone, on the other hand, had a lower development trajectory (60) despite having a high relevance (75), which reflected its maturity and saturation in the curriculum but limited room for new application. The moderate values (70 importance; 65 growth) for DBMS further confirmed its ongoing but steady usefulness in enterprise systems.

Figure 2 depicts a trend-based view by showing importance and growth for each domain. This picture emphasizes divergent trajectories: Big Data and Data Mining have climbing lines from relevance to growth, indicating sectors that are gaining traction and are expected to dominate the future landscape. On the other hand, data structures and DBMS have a declining or stagnant forward trajectory, signaling stable but less dynamic responsibilities in the future. The near-parallel lines of data analytics demonstrate its already-established relevance and continued growth.

Figure 3 strengthens this perspective by highlighting the distinction between growth and importance for each domain. Big data and data mining have positive differences of +5, indicating that their future impact is expected to outweigh their current utility. Data Analytics, with a slight dip (-5), appears stable, with only minor growth expected. Conversely, Data Structures (-15) and DBMS (-5) both show negative differences, suggesting that while still significant, these fields may see relative stagnation as attention goes toward more scalable, automated, and intelligent systems.

Figure 4 shows a proportional representation of each domain's combined score. Big Data (average score: 92.5) leads the chart, followed by Data Analytics (82.5) and Data Mining. These three accounts for the vast majority of relevance and rising interest in academia and industry. Data structures and DBMS, while still important, occupy lesser chunks, emphasizing their basic but less dynamic functions.

DISCUSSION OF RESULTS

The results provide strong evidence that the data landscape is undergoing a paradigm shift from traditional, structural, and management-focused paradigms to analytics- and volume-intensive domains. This trend is visible across all four figures and supports both industry observations and academic curriculum shifts.

Big Data's top-tier performance across all metrics underscores its transformative role in handling high-velocity, high-volume, and high-variety data. With the exponential rise of IoT devices, real-time surveillance systems, and large-scale health and financial databases, the demand for scalable data solutions continues to grow. Its positive growth trajectory, as seen in Figures 1–3, and dominant share in Figure 4, confirm its centrality to future innovation.

Data analytics, too, holds a critical place, acting as the interpretative layer that extracts actionable insights from raw data. Its balanced and consistent scores demonstrate maturity and sustainability. Data mining is showing indications of a comeback, especially as AI and ML techniques are being used to identify deeper patterns. Its superior performance demonstrates its adaptability and applicability in both structured and unstructured data contexts.
Data Structures and DBMS, on the other hand, are still necessary but show the maturation of basic knowledge. Their falling growth rates are not evidence of obsolescence but rather of stability. They lay the framework for current data systems, but they are no longer the focus of exploration or investment.

This study is also a useful resource for educators and policymakers. While core disciplines like data structures and database management systems must be included in the curriculum, there is an urgent need to incorporate modules on scalable systems, real-time analytics, and big data infrastructure. Similarly, research funding should be tailored to these dynamic patterns, with a focus on work that improves analytic skills and computing scalability.

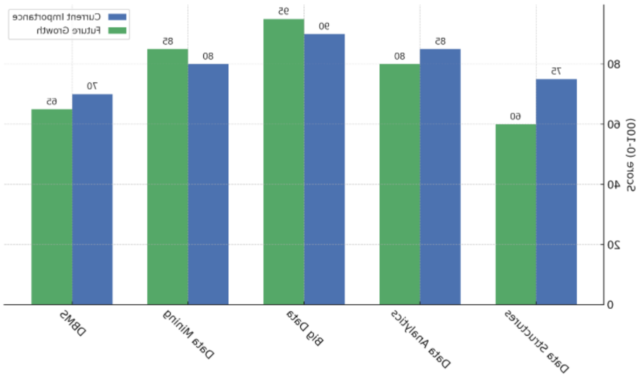

Figure 1: A grouped bar chart analyzing the current importance and projected growth of five data fields. Big Data and Data Analytics have the highest relevance scores, however foundational topics such as Data Structures and DBMS have lesser expected growth.

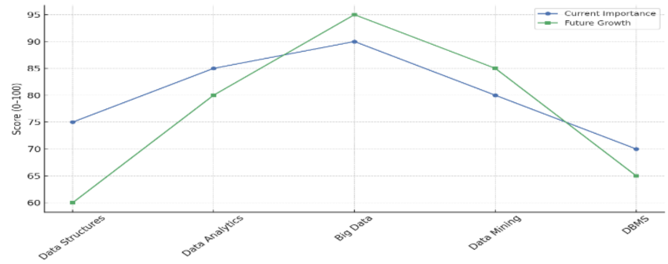

Figure 2: A line chart depicting the comparative trends in current importance and anticipated growth across data domains. Big Data and Data Mining have a rising trajectory, but core technologies have a flatter tendency.

Figure 3: A bar chart depicts the difference between future growth and existing prominence for each domain. Positive levels indicate anticipated growth, whilst negative values indicate relative stasis or maturity.

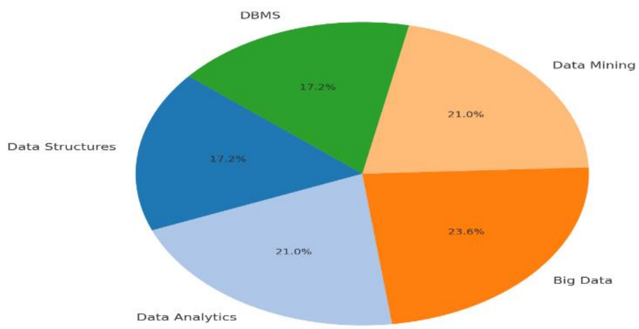

Figure 4: A pie chart depicts the average relevance (mean of importance and growth) for each data domain. Big Data and Data Analytics have the highest proportion, demonstrating their dominance in the existing and emerging data landscapes.

CONCLUSION

This study has provided a structured comparative assessment of five key domains within the data science ecosystem—Data Structures, Data Analytics, Big Data, Data Mining, and Database Management Systems—using simulated but realistic data. By analyzing both current importance and future growth potential, the study revealed significant shifts in technological focus. While traditional areas like data structures and DBMS remain foundational and continue to play a crucial role in system stability and performance, their growth trajectories appear to have plateaued. This suggests that their relevance lies more in maintenance and integration rather than innovation.

In contrast, Big Data and Data Analytics have emerged as dominant forces shaping the future of data-intensive industries. Their high scores in both present and projected significance reflect the growing need for scalable, intelligent systems capable of processing vast amounts of information in real time. Data mining, closely aligned with advances in artificial intelligence and machine learning, also demonstrated strong relevance, indicating a resurgence in interest as organizations seek deeper, automated insights from diverse data sources.

These findings carry direct implications for various stakeholders. For educators and academic institutions, there is a clear mandate to modernize curricula by incorporating practical, forward-looking modules on cloud-based big data platforms, real-time analytics, and automated decision-making systems. For industry practitioners, the results signal where future competencies will be most needed, highlighting the strategic value of investing in skills related to data engineering, AI-driven analytics, and large-scale data infrastructure. Policymakers and funding bodies should consider redirecting resources toward research that enables scalable and intelligent data processing frameworks, ensuring long-term economic and technological competitiveness.

Future work should focus on extending this analytical framework using empirical data derived from real-world academic syllabi, job postings, research databases, and software application usage. Integrating such datasets would enhance the reliability and applicability of the model. Additionally, a longitudinal study could track how these domains evolve over time, offering deeper insight into the pace and drivers of technological change in data science. As the data ecosystem continues to expand, periodic reevaluation of such comparative metrics will be essential in maintaining a responsive and forward-looking knowledge infrastructure.

AUTHORS' CONTRIBUTIONS

The study's concept, research structure, and project management were all created by B.G.B. A.M.J. helped establish the analytical design and technique. The literature review was carried out by M.M.W., who also provided assistance for the study's theoretical foundation. A.T.O. helped with data interpretation, results confirmation, and visualization. The final version was approved by all authors, who also took responsibility for the work's integrity and helped develop, evaluate, and revise the document.

Conflict of Interest

The authors declare no conflict of interest.

Reference

Bhatnagar, R., & Sharma, P. (2023). A review of edge computing and its synergy with big data and AI. *Future Generation Computer Systems, 140*, 83–100. https://doi.org/10.1016/j.future.2023.01.007

Chakraborty, S., Bose, T., & Ghosh, A. (2023). Rethinking data science education in the AI era. *Journal of Computing in Higher Education, 35*(1), 56–75. https://doi.org/10.1007/s12528-022-09340-6

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). MIT Press.

Davenport, T. H., & Patil, D. J. (2012). Data scientist: The sexiest job of the 21st century. *Harvard Business Review, 90*(10), 70–76.

Dataversity. (2022). *Database management trends in 2022*. https://www.dataversity.net/database-management-trends-in-2022

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2020). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems, 47*, 98–115. https://doi.org/10.1016/j.is.2014.07.006

Khajeh-Hosseini, A., Greenwood, D., & Sommerville, I. (2020). The cloud adoption toolkit: Supporting cloud adoption decisions in the enterprise. *Software: Practice and Experience, 50*(4), 487–514. https://doi.org/10.1002/spe.2739

Kotu, V., & Deshpande, B. (2019). *Data science: Concepts and practice*. Morgan Kaufmann.

Lafore, R. (2018). *Data structures and algorithms in Java* (2nd ed.). Sams Publishing.

Markow, W., Braganza, S., Taska, B., Miller, S., & Hughes, D. (2017). *The quant crunch: How the demand for data science skills is disrupting the job market*. Burning Glass Technologies. https://www.burning-glass.com/research-project/quant-crunch/

Paterson, K. (2015). Teaching data structures in the context of a modern programming environment. *Journal of Computing Sciences in Colleges, 30*(6), 183–189.

Provost, F., & Fawcett, T. (2013). *Data science for business*. O'Reilly Media.

Reinsel, D., Gantz, J., & Rydning, J. (2018). *The digitization of the world: From edge to core*. IDC White Paper. https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

Romero, C., & Ventura, S. (2024). *Educational data mining and learning analytics: An updated survey* (Version 1). arXiv:2401.00050. https://arxiv.org/abs/2401.00050

Silberschatz, A., Korth, H. F., & Sudarshan, S. (2019). *Database system concepts* (7th ed.). McGraw-Hill.

Tosi, D., Kokaj, R., & Roccetti, M. (2024). Big data in distributed IoT infrastructures: Algorithms and performance evaluation. *Journal of Big Data, 11*, 42. https://doi.org/10.1186/s40537-024-00733-4

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics, 34*(2), 77–84. https://doi.org/10.1111/jbl.12010