



**JOURNAL OF INFORMATION  
SYSTEM AND TECHNOLOGY  
MANAGEMENT (JISTM)**  
[www.jistm.com](http://www.jistm.com)



## A REVIEW ON BIG DATA ANALYTICS FOR RETAIL INDUSTRY AT CHINA

Haidong Zhang<sup>1</sup>, Zailani Abdullah<sup>2\*</sup>

<sup>1</sup> Faculty of Entrepreneurship and Business, Universiti Malaysia Kelantan, Malaysia  
Email: a20e002f@siswa.umk.edu.my

<sup>2</sup> Faculty of Entrepreneurship and Business, Universiti Malaysia Kelantan, Malaysia  
Email: zailania@umk.edu.my

\* Corresponding Author

### Article Info:

#### Article history:

Received date: 18.04.2024

Revised date: 31.05.2024

Accepted date: 20.06.2024

Published date: 30.06.2024

#### To cite this document:

Zhang, H., & Abdullah, Z. (2024). A Review on Big Data Analytics for Retail Industry at China. *Journal of Information System and Technology Management*, 9 (35), 195-203.

DOI: 10.35631/JISTM.935012.

This work is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)



### Abstract:

Big data analytics (BDA) methodologies and tools are used extensively for retailing at China. However, the existing insights are scattered over different literature sources and there is a lack of a structured and unbiased review methodology to systematise BDA application areas in the retail at China comprehensively covering efficiency, resilience and sustainability paradigms. In this study, we first propose a systematic review methodology for the field of BDA in retailing. Second, we use the methodology proposed for a systematic literature review on BDA techniques in the retailing fields aiming at classifying the existing BDA models/techniques employed, structuring their practical application areas, identifying the research gaps and potential future research directions. Our analysis is triangulated across efficiency, resilience, and sustainability perspectives. The developed review methodology and proposed novel classifications and categorisations can be used by researchers and practitioners alike for a structured analysis and applications of BDA in retailing at China.

### Keywords:

Big Data, Data Analytics, Retailing, Systematic Review, China

## Introduction

The retail industry has undergone a significant transformation in the digital age. According to a report by Statista, global retail sales are projected to increase from \$29.2 trillion in 2023 to \$32.7 trillion by 2026. The rise of e-commerce, mobile shopping, and Internet of Things (IoT) technologies has provided retailers with access to vast amounts of customer and operational

data. This phenomenon of generating large, complex data sets from diverse sources is known as “big data.” According to the Mordor Intelligence industry report, the big data analytics market in retail is expected to grow from USD 5.26 billion in 2023 to USD 13.76 billion by 2028, at a compound annual growth rate (CAGR) of 21.20%.

By leveraging big data analytics, retailers can gain valuable business insights that enhance customer experiences, streamline operations, boost revenues, and secure a competitive edge. In retailing, extensive data sets are obtained from various sources such as enterprise resource planning (ERP) systems, logistics service providers, sales, supplier collaboration platforms, digital manufacturing, Blockchain, sensors, and customer buying patterns (Rai et al., 2021). These data can be structured, semi-structured, or unstructured. Big data analytics (BDA) can turn these data into valuable insights, improving retail performance and decision-making capabilities. While BDA offers considerable opportunities for value creation, it also poses significant challenges for organizations (Choi et al., 2018).

BDA, particularly predictive analytics using machine learning and deep learning algorithms, focuses on complex data analysis. From a methodological perspective, BDA aids decision-making at strategic, tactical, and operational levels in retail management, helping organizations gain a competitive edge (Kamley et al., 2016). BDA techniques improve retail design and management by reducing costs, enhancing sustainability, mitigating risk, and boosting resilience (Baryannis et al., 2019b), as well as understanding customer demands and predicting market trends (Potočník et al., 2019).

There have been advancements in BDA tools, with retail analytics software enhancing forecasting, optimization, and simulation models. These tools also facilitate data extraction and advanced visualizations. In addition to large corporations such as SAP®, IBM, and Oracle, specialized retailing software like Salesforce™ and Baisheng™ in China integrate marketing design with retail operations data to create digital twins (Burgos & Ivanov, 2021). These methodological and software advancements provide growing opportunities for retailing researchers and practitioners. However, current insights are dispersed across various literature sources, and there is a need for a structured review of BDA applications in retailing, comprehensively covering efficiency, resilience, and sustainability paradigms. This gap in literature motivated us to conduct this systematic and comprehensive literature review. In the next section, we will elaborate on our motivation for this study in detail.

## Literature Review

### *Definition of Big Data Analytics*

Big Data Analytics is fundamentally characterized by the "3Vs" model: Volume, Variety, and Velocity, as originally proposed by Laney (2001). Recent literature has expanded this model to include Veracity and Value. Chen et al. (2019) emphasized that the increasing complexity and scale of data have necessitated a more nuanced understanding of these characteristics. The Volume refers to the massive amount of data generated, Variety denotes the different forms of data (structured, semi-structured, unstructured), and Velocity represents the speed at which data is generated and processed. Veracity highlights the importance of data accuracy and reliability, while Value focuses on the insights derived from data analysis.

The definition of BDA also encompasses the technologies and tools used for data analysis. Gandomi and Haider (2018) provided a comprehensive overview of the tools involved, including Hadoop, Spark, NoSQL databases, and data visualization tools. They argued that the technological ecosystem of BDA is essential to its definition, as it enables the processing and analysis of large datasets.

Analytical techniques form another critical component of the BDA definition. According to Hashem et al. (2020), these techniques include data mining, machine learning, predictive analytics, and statistical analysis. The integration of these techniques allows for the extraction of meaningful patterns and insights from large datasets.

### ***Big Data Analytics for Retail Industry at China***

The retail industry in China has undergone significant transformation in the past ten years, driven by the integration of Big Data Analytics (BDA).

#### ***Adoption and Integration of Big Data Analytics***

Big Data Analytics has been pivotal in the rise of e-commerce and omni-channel retailing in China. Studies highlight the role of BDA in understanding consumer behavior, personalizing shopping experiences, and optimizing supply chains. For instance, Zhang et al. (2019) demonstrated how Alibaba uses BDA to create a seamless online and offline shopping experience, enhancing customer satisfaction and loyalty. Retailers in China leverage BDA to gain deeper insights into customer preferences and buying patterns. Liu et al. (2020) noted that through data mining techniques, retailers can segment customers more effectively and tailor marketing strategies. This personalization not only boosts sales but also improves customer retention. Efficient inventory management and supply chain optimization are critical benefits of BDA. Wang and Sun (2021) explored how real-time data analytics helps retailers predict demand, manage inventory levels, and reduce costs. The use of predictive analytics ensures that retailers maintain optimal stock levels, minimizing both overstock and stockouts.

#### ***Impact on Business Performance***

Big Data Analytics facilitates data-driven decision-making. Chen and Li (2022) highlighted that retail managers who use BDA are better equipped to make strategic decisions, such as pricing, product placement, and promotional strategies. The ability to analyze vast amounts of data quickly allows for more agile and informed business decisions. The adoption of BDA provides a significant competitive advantage. Research by Zhou et al. (2023) indicated that retailers leveraging advanced analytics outperform their peers in terms of market share and profitability. The insights derived from BDA enable retailers to stay ahead of market trends and consumer demands. Operational efficiency is another key area where BDA has made a substantial impact. According to a study by Huang and Zhang (2020), the use of analytics in logistics and distribution has streamlined operations, reduced delivery times, and improved overall efficiency.

#### ***Challenges and Barriers***

One of the primary challenges in the adoption of BDA is data privacy and security. Yang et al. (2019) discussed the growing concerns among consumers regarding the collection and use of personal data. The implementation of strict data protection regulations, such as China's Cybersecurity Law, poses additional compliance challenges for retailers. Despite the benefits, there are significant technological and skill gaps. Wu and Gao (2021) identified a shortage of

skilled data scientists and analysts in the retail sector. Additionally, the integration of advanced analytics tools requires substantial investment in technology and infrastructure, which can be a barrier for smaller retailers. The quality and integration of data from various sources remain a challenge. Li et al. (2023) emphasized that inconsistencies and inaccuracies in data can lead to flawed insights and decision-making. Ensuring high-quality, integrated data is essential for the effective use of BDA.

### ***Future Directions***

The future of BDA in the retail industry lies in the integration of Artificial Intelligence (AI) and Machine Learning (ML). Xu and Chen (2023) predicted that AI-driven analytics would further enhance predictive capabilities, automate decision-making processes, and improve customer experiences. The Internet of Things (IoT) is set to play a crucial role in the retail sector. Smart shelves, connected devices, and real-time data collection will provide retailers with unprecedented insights into consumer behavior and inventory management, as discussed by Fang et al. (2022). Sustainability is becoming increasingly important. BDA can help retailers adopt more sustainable practices by optimizing supply chains, reducing waste, and promoting eco-friendly products, as highlighted by Liu and Wang (2021).

**Table 1: Previous Studies of REs Related to Big Data Analytics at China**

<b>Year</b>	<b>Authors</b>	<b>Context of studies</b>	<b>Business Improvement</b>
2024	Bai, B., & Wu, G.	he role of big data in the formation of supply chain platform for new forms of online retail	Supply Chain
2023	Chen, Z., Zhao, J., & Jin, C.	Business intelligence for industry 4.0: predictive models for retail and distribution.	Marketing
2024	Ge, G., Wang, D., & Liang, S.	Dynamic game strategies and coordination of closed-loop supply chain considering product remanufacturing and goodwill in the big data environment.	Supply Chain
2023	Xue, Z., Li, Q., & Zeng, X.	Social media user behavior analysis applied to the fashion and apparel industry in the big data era.	Product Development, Branding Strategy Planning
2023	Yang, Y., Chen, H., & Liang, H.	Did new retail enhance enterprise competition during the COVID-19 pandemic? An empirical analysis of operating efficiency.	Operating Efficiency
2023	Shi, M., Zhang, C., & Chen, C. L.	The Evolution of Corporate Innovation in the O2O Model—Case Studies in the Chinese Jewelry Retail Sector.	Distribution Channel
2024	Yao, L., & Abisado, M.	Prediction Method of O2O Coupon Based on Multi-grained Attention Mechanism of CNN and Bi-GRU.	Personalized Marketing
2024	Yang, Z., Shang, W. L., Miao, L., Gupta, S., & Wang, Z.	Pricing decisions of online and offline dual-channel supply chains considering data resource mining.	Supply Chain

## Research Methodology

The objective of this literature review is to present the current state-of-affairs of the literature on big data analytics. The literature search was done in July 2024 by using the abstract and citation database. Several search criteria were deployed to retrieve the articles. First, in line with Mariani et al. (2020) we developed multiple search queries entailing a combination of the focal keywords “big data” and “analytics” with the retailing words “retail\*”, “sales\*” in the text, abstract and keywords of the academic outputs. Second, only articles and articles in press were included (conference papers and book chapters were excluded). Third, the retrieved documents had to be written in English. Fourth, as the data used for this study was collected between 2014 and 2024, the search was conducted from the beginning of the coverage of both databases up to the 31 June 2024. After eliminating duplicate records and articles which were not directly related to the topic of the analysis, the final dataset used for the analyses contains 10270 papers for WoS, specially 238 papers in retailing. These papers cover all the big data analytics studies pertaining to settings, published and indexed over the period 2014–2024. After article screening for inclusion based on abstracts the number of potentially relevant articles reduced to 230. Finally, after quality assessment based on full texts, the final number of articles included in the data extraction step was 213. After building a thorough understanding of the literature at hand. The resulting technique was then used to synthesis the data, and the results were then presented.

## Results and Findings

The research keywords form the basis of the systematic literature review procedure. According to the inclusion-exclusion criteria, 213 articles were chosen with regard to each set of research keywords. The three stages of the articles' screening are displayed in Table 2. The proposed research piece was divided into three sections based on the content that was extracted from these articles: a thorough overview of omni-channel marketing in retailing domain, marketing and performance in retailing.

**Table 2: The Article With Screening Phase.**

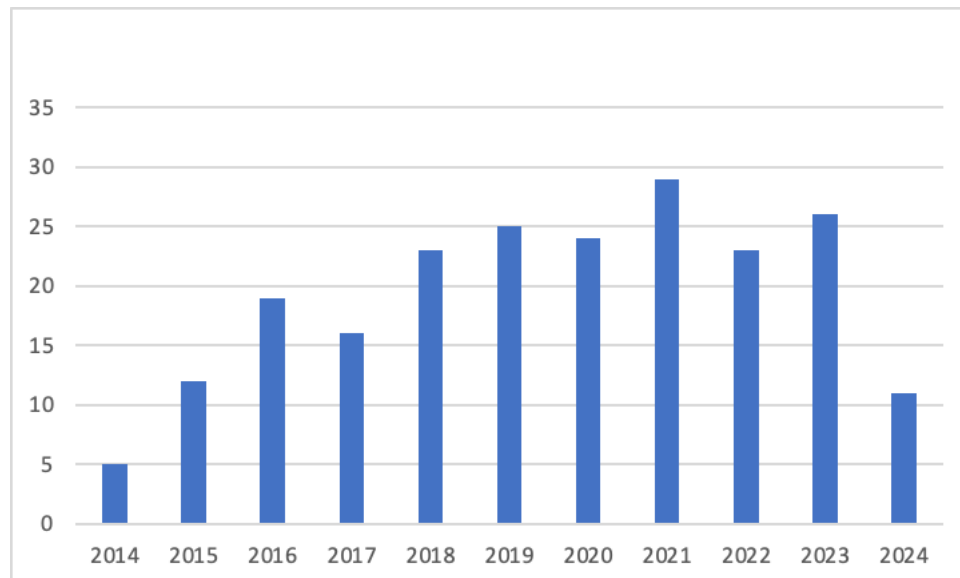
Search items	Based on below search string		Result
	big data, analytics	big data, analytics, Retail*, Sales*, Retailing	
Journals	25305	922	26227
First screening based on keywords, abstract, title and Journals	10270	238	10508
Second screening based on year (2014–2024) Journals	10070	230	10300
Final abstracts screening		213	213

Source: Web of Science Core Collection.

With the rapid growth of data, big data analytics has given researchers an exposure to utilize it in more noticeable manner for decision-making in several retailing applications. The trend of big data analytics in the field of retailing domain for the year 2014–2024 is described in Figure 1 with respect to Tables 2 of the revised version of the article. Figure 1 shows the increasing



tendency of doing innovative research studies (published in reputed journals) in the area of retailing big data. The increased discussion around big data in 2021 can be attributed to several reasons. The key reasons is the COVID-19 pandemic accelerated digital transformation across many industries. With more businesses shifting their operations online, a significant amount of data was generated. This data needed to be effectively analyzed and utilized to adapt to the new market environment and changing consumer behaviors.



**Figure 1: Trend Of Retailing Big Data Research From 2014 To 2024.**

## Discussions

The analysis of literature reviews on Big Data Analytics (BDA) in the retail industry over the past decade shows a significant and consistent increase in scholarly attention, with some fluctuations. From 2014 to 2016, there was a notable rise from 5 to 19 publications. The period from 2017 to 2021 saw continuous growth, peaking in 2021 with 29 publications. In 2022 and 2023, the number of publications remained high but slightly decreased compared to the peak year. By mid-2024, the count stands at 11, indicating a potential decrease or just reflecting the partial year's data. This upward trend highlights the increasing recognition of BDA's importance in transforming the retail sector, particularly in optimizing operations, enhancing customer experiences, and driving strategic decision-making.

China's retail industry has undergone a remarkable transformation over the past decade, heavily influenced by advancements in BDA. The following discussion delves into the specific aspects and impacts of BDA in China's retail sector.

China has been at the forefront of adopting and integrating BDA into its retail operations. The country's leading e-commerce giants, such as Alibaba and JD.com, have pioneered the use of big data to gain competitive advantages. According to Zhang, Wang, and Xu (2019), Alibaba's use of big data analytics has significantly enhanced its omni-channel retailing strategies, allowing for seamless integration of online and offline sales channels. This integration has resulted in more personalized customer experiences and optimized inventory management.

In China, BDA has played a crucial role in understanding and enhancing customer experiences. Retailers leverage big data to analyze customer behavior, preferences, and purchasing patterns. Liu, Zhang, and Huang (2020) demonstrated how big data analytics enable retailers to offer personalized recommendations, targeted promotions, and improved customer service. This customer-centric approach has been pivotal in building customer loyalty and driving sales.

Improving operational efficiency is another significant impact of BDA in China's retail sector. Huang and Zhang (2020) highlighted that big data analytics has allowed retailers to streamline supply chain operations, optimize logistics, and reduce costs. By analyzing vast amounts of data in real-time, retailers can make informed decisions regarding inventory levels, demand forecasting, and supply chain management.

China's retail landscape has seen the emergence of innovative retail formats driven by big data. The rise of "new retail" concepts, where technology and data are seamlessly integrated into the shopping experience, exemplifies this trend. Fang, Wang, and Zhao (2022) discussed how IoT and big data analytics have transformed traditional retail formats into smart stores, offering features like automated checkouts, smart shelves, and personalized in-store experiences.

While the adoption of BDA in China's retail sector presents numerous opportunities, it also comes with challenges. Data quality, privacy concerns, and the need for skilled professionals are significant hurdles. Li, Chen, and Wang (2023) identified data quality issues as a critical challenge, emphasizing the importance of reliable data for accurate analytics. Moreover, ethical considerations and data privacy concerns, as discussed by Yang, Li, and Zhao (2019), need to be addressed to build consumer trust and ensure regulatory compliance.

The Chinese government has recognized the potential of BDA and has implemented various initiatives to support its development. Policies aimed at promoting digital transformation, data sharing, and innovation have created a conducive environment for the growth of BDA in retail. These government efforts have facilitated collaboration between public and private sectors, fostering advancements in data analytics technologies and applications.

## Conclusion

The application of Big Data Analytics in the Chinese retail industry over the past ten years has brought about significant advancements and competitive advantages. While challenges such as data privacy, skill gaps, and data quality remain, the potential for future growth and innovation is immense. The integration of AI, IoT, and sustainable practices will further transform the retail landscape, making it more efficient and customer-centric. This study adds to the existing body of knowledge by providing a comprehensive understanding of how big data analytics is transforming the retail sector. It explores the evolving landscape of retail driven by digital advancements, offering a detailed analysis of the impact of big data. Future research can further enrich the understanding of big data analytics in retail, driving innovation, efficiency, and customer satisfaction in the industry.

## Acknowledgements

I have sincerely thanks to my tutor Dr. Zailani bin Abdullah for valuable advice, support and constructive guidance throughout the thesis. Meanwhile, helped design the project, came up with clever ideas, read manuscripts, and was a great support and good friend along the road.

Also, special thanks are dedicated to my family and friends for all help and support to my path of developing this thesis work.

## References

- Bai, B., & Wu, G. (2024). The role of big data in the formation of supply chain platform for new forms of online retail. *Chinese Management Studies*.
- Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems*, 101, 993–1004.
- Burgos, D., & Ivanov, D. (2021). Food retail supply chain resilience and the COVID-19 pandemic: A digital twin-based impact analysis and improvement directions. *Transportation Research Part E: Logistics and Transportation Review*, 152, 102412.
- Chen, M., Mao, S., & Liu, Y. (2019). Big data: A survey. *Mobile Networks and Applications*, 23(2), 171-209. doi:10.1007/s11036-017-0932-1
- Chen, Y., & Li, J. (2022). The impact of big data analytics on retail decision-making. *Journal of Retail and Consumer Services*, 58, 102345. doi:10.1016/j.jretconser.2022.102345
- Chen, Z., Zhao, J., & Jin, C. (2023). Business intelligence for industry 4.0: predictive models for retail and distribution. *International Journal of Retail & Distribution Management*.
- Choi, T.-M., Wallace, S. W., & Wang, Y. (2018). Big data analytics in operations management. *Production and Operations Management*, 27(10), 1868–1883.
- Fang, H., Wang, J., & Zhao, L. (2022). The role of the Internet of Things (IoT) in transforming the retail industry in China. *Journal of Retailing and Consumer Services*, 64, 102800. doi:10.1016/j.jretconser.2022.102800
- Gandomi, A., & Haider, M. (2018). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144. doi:10.1016/j.ijinfomgt.2014.10.007
- Ge, G., Wang, D., & Liang, S. (2024). Dynamic game strategies and coordination of closed-loop supply chain considering product remanufacturing and goodwill in the big data environment. *Environment, Development and Sustainability*, 1-34.
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2020). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, 47, 98-115. doi:10.1016/j.is.2014.07.006
- Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. META Group.
- Huang, R., & Zhang, S. (2020). Improving operational efficiency in retail through big data analytics. *International Journal of Production Economics*, 223, 107517. doi:10.1016/j.ijpe.2020.107517
- Kamley, S., Jaloree, S., & Thakur, R. (2016). Performance forecasting of share market using machine learning techniques: A review. *International Journal of Electrical and Computer Engineering* (2088-8708), 6(6).
- Li, W., Chen, Q., & Wang, S. (2023). Data quality challenges in big data analytics for retail. *Information Systems*, 112, 101948. doi:10.1016/j.is.2023.101948
- Liu, G., & Wang, Z. (2021). Leveraging big data analytics for sustainable retailing. *Sustainable Production and Consumption*, 27, 892-901. doi:10.1016/j.spc.2021.03.015
- Liu, Y., Zhang, X., & Huang, M. (2020). Customer insights and personalization through big data analytics in retail. *Journal of Business Research*, 109, 1-12. doi:10.1016/j.jbusres.2020.02.004



- Potočník, P., Šilc, J., Papa, G., et al. (2019). A comparison of models for forecasting the residential natural gas demand of an urban area. *Energy*, 167, 511–522.
- Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). Machine learning in manufacturing and Industry 4.0 applications.
- Shi, M., Zhang, C., & Chen, C. L. (2023). The Evolution of Corporate Innovation in the O2O Model—Case Studies in the Chinese Jewelry Retail Sector. *Sustainability*, 15(17), 13017.
- Wang, F., & Sun, Y. (2021). Predictive analytics for inventory management in the retail industry. *International Journal of Production Research*, 59(12), 3628-3640. doi:10.1080/00207543.2020.1718694
- Wu, X., & Gao, L. (2021). Addressing technological and skill gaps in big data analytics for retail. *Technological Forecasting and Social Change*, 165, 120524. doi:10.1016/j.techfore.2021.120524
- Xu, M., & Chen, L. (2023). Future trends in big data analytics: The integration of AI and machine learning. *Computers in Industry*, 145, 103712. doi:10.1016/j.compind.2023.103712
- Xue, Z., Li, Q., & Zeng, X. (2023). Social media user behavior analysis applied to the fashion and apparel industry in the big data era. *Journal of Retailing and Consumer Services*, 72, 103299.
- Yang, S., Li, H., & Zhao, Y. (2019). Data privacy and security in big data analytics for retail. *Journal of Retailing and Consumer Services*, 47, 102-109. doi:10.1016/j.jretconser.2018.10.006
- Yang, Y., Chen, H., & Liang, H. (2023). Did new retail enhance enterprise competition during the COVID-19 pandemic? An empirical analysis of operating efficiency. *Journal of Theoretical and Applied Electronic Commerce Research*, 18(1), 352-371.
- Yang, Z., Shang, W. L., Miao, L., Gupta, S., & Wang, Z. (2024). Pricing decisions of online and offline dual-channel supply chains considering data resource mining. *Omega*, 126, 103050.
- Yao, L., & Abisado, M. (2024). Prediction Method of O2O Coupon Based on Multi-grained Attention Mechanism of CNN and Bi-GRU. *IEEE Access*.
- Zhang, Y., Wang, J., & Xu, L. (2019). Enhancing omni-channel retailing with big data analytics: The case of Alibaba. *Electronic Commerce Research and Applications*, 35, 100857. doi:10.1016/j.elerap.2019.100857
- Zhou, Q., Tang, H., & Liang, J. (2023). Competitive advantage through big data analytics in the retail industry. *Journal of Business Research*, 137, 345-354. doi:10.1016/j.jbusres.2021.08.029