

# Customer Segmentation in the Retail Market

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Jeen Mary John , Olamilekan Shobayo , Bayode Ogunleye

Peoples' awareness of online purchases has significantly risen. This has given rise to online retail platforms and the need for a better understanding of customer purchasing behaviour. Retail companies are pressed with the need to deal with a high volume of customer purchases, which requires sophisticated approaches to perform more accurate and efficient customer segmentation. Customer segmentation is a marketing analytical tool that aids customer-centric service and thus enhances profitability.

customer segmentation

DBSCAN

Gaussian mixture model

K-means clustering

BIRCH

RFM framework

## 1. Introduction

In the era of Big Data, decision making based on data is becoming increasingly important for organisations, especially the retail industry. Big Data Analytics (BDA) is used to analyse massive datasets and identify patterns in data, which can be used to make informed decisions <sup>[1]</sup>. Segmentation is a well-established marketing strategy that involves dividing customers into distinct groups, known as market segments, and focusing marketing efforts towards the most favourable segment <sup>[2]</sup>. Gaining insight into consumer behaviour and the decision-making process is crucial in formulating efficient strategies to deliver a customer-centric service and enhance profitability. Business Intelligence (BI) and Big Data Analytics (BDA) are employed in the customer segmentation framework to discern the prospective customer profile and market both new and established products. However, there is a lack of contextual framework behind practice development, specifically regarding the influence on decision marketing strategic issues.

Online customer data could be used to generate personas that represent distinct groups of individuals <sup>[3][4]</sup>. Business Intelligence (BI) and Big Data Analytics (BDA) could also be leveraged for the creation of a framework aimed at discerning the prospective customer profile for the purpose of marketing both new and established products <sup>[4]</sup>.

## 2. Customer Segmentation in the Retail Market

To assist the economy, client loyalty research is essential. According to a study conducted in the UK, consumers are more devoted to specific retailers and brands when purchasing is convenient. The cost of the goods is not as significant as service operations and customer pleasure. To successfully control consumer purchasing behaviour

and boost retail sales, retail companies must strengthen service operations skills [5]. By allowing personalised shopping experiences, predicting patterns, and taking wise choices based on market information, BD has transformed the retail industry [6].

During the era of Big Data (BD), decision making driven by data is common, irrespective of the scale of the business or the industry, who will be able to tackle complex business problems by taking actions based on data-driven insights [7]. According to Oussous et al. [8], BD, in contrast to traditional data, refers to vast, expanding data collections that encompass many different diverse forms, including un-structured, semi-structured, and structured data.

A case study was conducted by Jin and Kim on a typical courier company's sorting and logistics processing [9]. According to the study, BD can enhance business efficiency by transforming raw data into valuable information. Identifying the type and scope of data is crucial for achieving sustainable growth and competitiveness. Integrated use of BI, BD, and BDA in management decision support systems may aid businesses in achieving time- and cost-effectiveness. This case study provides useful insights for future company plans to reduce trial-and-error iterations. The authors of [10] developed a generic framework for organisational excellence by integrating Baldrige and BI frameworks. They formed respective matrices and adapted the Baldrige and BI frameworks, incorporating knowledge management and BI frameworks with specific key performance indicator (KPI) parameters. The framework integrates KPIs from the Baldrige framework, customer management, workforce engagement, knowledge management, operations focus, strategic planning, and academic accreditation. The dashboard is designed as an infographic mechanism, incorporating business, content, analytics, and continuous intelligence.

A K-means clustering (KC)-based consumer segmentation model was proposed in [11] and implemented using RStudio. The CSV-formatted dataset was analysed using various R functions. Visualisations were developed to understand client demographics. To identify the clusters, the elbow technique, average silhouette method, and gap statistics were used.

To examine the use of BDA in supply chain demand forecasting, the authors of [12] employed neural networks (NNs) and a regression analysis approach. The research findings are based on closed-loop supply chains (CLSCs) and provide suggestions for future research.

Customer segmentation is crucial in marketing as it helps organisations understand and meet their consumers' demands by breaking down an intended market into groups based on shared traits. Machine-learning-based customer segmentation models, such as k-means clustering, density-based spatial clustering of applications with noise (DBSCAN), and balanced iterative reducing and clustering using hierarchies (BIRCH), have presented potential insights in analysing customer data for effective decision making in the marketing sector. Ushakovato and Mikhaylov [13] developed a Gaussian-based framework for analysing smart meter data for predicting tasks. The experimental study analysed and compared different clustering techniques for predicting time series data and cluster assignments without additional customer information. Fontanini and Abreu [14] used the BIRCH algorithm to determine common load forms in neighbourhoods. The global clustering measure was developed by solving issues

generated during optimisation of load forms. Lorbeer et al. [15] found that BIRCH requires the maximum number of clusters for a better clustering quality performance. This technique extracts load forms from large databases, clusters high, moderate, and minimal load types using cost functions, and determines the right number of clusters for global grouping. It is suitable for urban-scale loading assessments and real-time online education. Hicham and Karim [16] proposed a clustering ensemble method which consists of DBSCAN, k-means, MiniBatch k-means, and the mean shift algorithm for customer segmentation. They applied their clustering ensemble method to 35,000 records and achieved a Silhouette Score of 0.72. The authors of [17] developed a customer segmentation model (RFM+B) using indicators of RFM and balance (B) to enhance marketing decisions. This model classifies client savings using transactional patterns and current balances using the RFM and B qualities. The model achieved an accuracy of 77.85% using the K-means clustering technique. Hossain [18] segmented customer data based on spending patterns through k-means clustering and the DBSCAN algorithm, determining clients with out-of-the-ordinary spending patterns. This study suggested incorporating neural-network-based clustering techniques for customer segmentation on large datasets. Punhani et al. [19] applied k-means clustering to 25,000 online customer records obtained from Kaggle repository. They used the Davies–Bouldin Index (BDI) to evaluate their optimal number of K and thus segmented the customer records into four segments. Turkmen [20] compared DBSCAN, k-means clustering, agglomerative clustering, and the RFM framework with 541,909 retail customer datasets. They employed random forest for selection of variables fitted into the algorithms. They showed k-means clustering achieved the best result with a Silhouette Score of 0.6. In summary, literature review findings suggest there is limited comparative studies of sophisticated unsupervised learning approaches for customer segmentation in the retail marketing context.

## References

1. Lekhwar, S.; Yadav, S.; Singh, A. Big data analytics in retail. In *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2018, Volume 2*; Springer: Singapore, 2019; pp. 469–477.
2. Gwozdz, W.; Steensen Nielsen, K.; Müller, T. An environmental perspective on clothing consumption: Consumer segments and their behavioral patterns. *Sustainability* 2017, 9, 762.
3. An, J.; Kwak, H.; Jung, S.G.; Salminen, J.; Jansen, B.J. Customer segmentation using online platforms: Isolating behavioral and demographic segments for persona creation via aggregated user data. *Soc. Netw. Anal. Min.* 2018, 8, 54.
4. Fotaki, G.; Spruit, M.; Brinkkemper, S.; Meijer, D. Exploring big data opportunities for online customer segmentation. *Int. J. Bus. Intell. Res. (IJBIR)* 2014, 5, 58–75.
5. Ramanathan, U.; Subramanian, N.; Yu, W.; Vijaygopal, R. Impact of customer loyalty and service operations on customer behaviour and firm performance: Empirical evidence from UK retail sector. *Prod. Plan. Control* 2017, 28, 478–488.

6. Li, C.; Chen, Y.; Shang, Y. A review of industrial big data for decision making in intelligent manufacturing. *Eng. Sci. Technol. Int. J.* 2022, 29, 101021.
7. Arunachalam, D.; Kumar, N. Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Syst. Appl.* 2018, 111, 11–34.
8. Oussous, A.; Benjelloun, F.Z.; Lahcen, A.A.; Belfkih, S. Big Data technologies: A survey. *J. King Saud Univ.-Comput. Inf. Sci.* 2018, 30, 431–448.
9. Jin, D.H.; Kim, H.J. Integrated understanding of big data, big data analysis, and business intelligence: A case study of logistics. *Sustainability* 2018, 10, 3778.
10. Jayakrishnan, M.; Mohamad, A.K.; Yusof, M.M. Understanding big data analytics (BDA) and business intelligence (BI) towards establishing organisational performance diagnostics framework. *Int. J. Recent Technol. Eng.* 2019, 8, 128–132.
11. Mathew, A.; Scholar, P.G.; Jobin, T.J. Role of Big Data Analysis and Machine Learning in Ecommerce-Customer Segmentation. In *Proceedings of the National Conference on Emerging Computer Applications (NCECA)*, Online, 17 June 2021; p. 189.
12. Seyedan, M.; Mafakheri, F. Predictive big data analytics for supply chain demand forecasting: Methods, applications, and research opportunities. *J. Big Data* 2020, 7, 1–22.
13. Ushakova, A.; Mikhaylov, S.J. Big data to the rescue? Challenges in analysing granular household electricity consumption in the United Kingdom. *Energy Res. Soc. Sci.* 2020, 64, 101428.
14. Fontanini, A.D.; Abreu, J. A data-driven BIRCH clustering method for extracting typical load profiles for big data. In *Proceedings of the 2018 IEEE Power & Energy Society General Meeting (PESGM)*, Portland, OR, USA, 5–10 August 2018; pp. 1–5.
15. Lorbeer, B.; Kosareva, A.; Deva, B.; Softić, D.; Ruppel, P.; Küpper, A. Variations on the clustering algorithm BIRCH. *Big Data Res.* 2018, 11, 44–53.
16. Hicham, N.; Karim, S. Analysis of Unsupervised Machine Learning Techniques for an Efficient Customer Segmentation using Clustering Ensemble and Spectral Clustering. *Int. J. Adv. Comput. Sci. Appl.* 2022, 13, 122–130.
17. Firdaus, U.; Utama, D. Development of bank's customer segmentation model based on rfm+ b approach. *Int. J. Innov. Comput. Inf. Cont.* 2021, 12, 17–26.
18. Hossain, A.S. Customer segmentation using centroid based and density based clustering algorithms. In *Proceedings of the 2017 3rd International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, Bangladesh, 7–9 December 2017; pp. 1–6.

19. Punhani, R.; Arora, V.S.; Sabitha, S.; Shukla, V.K. Application of clustering algorithm for effective customer segmentation in E-commerce. In Proceedings of the 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 17–18 March 2021; pp. 149–154.
20. Turkmen, B. Customer Segmentation with Machine Learning for Online Retail Industry. *Eur. J. Soc. Behav. Sci.* 2022, 31, 111–136.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/114405>