International Journal for Research Trends in Social Science & Humanities
E-ISSN 2584-2455
Volume 2 Issue 2
Mar-Apr 2024, PP 156-165

# Customer Behavior Analysis on Digital Shopping Platform for Increasing Engagement and Sales

**Dr Anjna Dubey Associate Professor[1], Sumit Sharma Assistant Professor[2], Dr Navneet Kaur Assistant Professor[3]**

Oriental College of Technology

Anjnamba06@gmail.com[1], navneet.khanuja03@yahoo.com[3]

**Abstract.** The fast growth of mobile apps and online shopping sites has made it easier to collect large amounts of data to help people decide what to buy. On these websites, analyzing customer reviews can greatly improve user satisfaction and build loyalty. This paper has finds the latests techniques applied to analysis the visitor behavior, shopping experience, etc. Paper has list some of the indicators that help scholars to understand the behavior of the customer at different level of shopping. As various researcher work in same field of customer behavior analysis, so diferent outcomes and approaches were discussed in the paper. This paper help online platform to evaluate and modify current working environment.

**Index Terms-** Customer Analysis, Marketing, Pattern Finding, Shopping Website Optimization.

## I. Introduction

With the rapid advancement of technology and societal progress, online shopping has become a key component in fulfilling people's everyday consumption needs. The growing trend of utilizing customer behavior and demand data to analyze factors influencing product sales through user-generated content (UGC) is gaining significant attention globally [1], [2]. As highlighted by the 41st Internet Development Statistics Report from the China Internet Network Information Centre (CNNIC) in 2018, by December 2017, the number of online shoppers in China had reached an estimated 533 million, marking a 14.3% increase from 2016. The report further stated that these online shoppers accounted for 69.1% of the total Internet users in the country [3]. This surge can be largely attributed to the convenience offered by the Internet, allowing customers easy access to a wealth of product information. Additionally, the increasing focus on various factors affecting purchase decisions, such as online reviews, is driving this trend [4], [5]. Nowadays, customers give greater importance to aspects like the tone of reviews (review valence) and the number of product images available. This type of information plays a critical role in reducing the uncertainty customers often face before making a purchase [6]. Therefore, efficiently gathering market information and controlling the factors influencing sales can help manufacturers and sellers optimize inventory management and boost competitiveness. Moreover, many e-commerce platforms are now

International Journal for Research Trends in Social Science & Humanities
E-ISSN 2584-2455
Volume 2 Issue 2
Mar-Apr 2024, PP 156-165

manipulating data to curate favorable reviews, thereby influencing product sales [7]. For instance, processes like clustering, screening, and sorting customer reviews have become pivotal in shaping consumer decisions and product success.

However, simply attracting visitors to an e-marketplace app does not ensure profitability, as value is only created when users complete transactions. Actual purchases begin when a user searches for products, adds items to their Wishlist or Favorites list, moves them to the shopping cart, and proceeds through the checkout process. Yet, the transaction can fail at any point before payment is made. The likelihood of failed transactions in e-marketplace apps increases as potential purchases grow, with an average of four failed transactions for every successful one [8].

One of the most common challenges faced by e-commerce businesses is cart abandonment, where users leave items in their shopping cart without completing the purchase [9]. Users can abandon their transactions at various stages, whether it's after product searching, before adding items to a Wishlist or Favorites list, or even during the checkout or payment phase. This behavior results in more failed transactions than successful ones, leading to substantial losses for e-marketplace companies due to the costs and efforts involved in maintaining their services [6]. As a result, e-commerce businesses are actively seeking strategies to increase the number of successful transactions and reduce the rate of cart abandonment.

## II. Related Work

In their study, Alkhawaldeh et al. [10] explored the complex relationship between customer value and sustainable green purchasing decisions, aiming to understand the factors that influence consumer behavior in this context. The research also sought to investigate the mediating role of customer attitudes and behaviors in the relationship between customer value and green purchasing decisions. To collect data, the researchers developed a comprehensive questionnaire based on previous literature in the field, which they then distributed to customers at major shopping malls in Jordan. Out of the collected responses, 145 were deemed valid for statistical analysis. The study utilized structural equation modeling (SEM) to test its hypotheses. The results indicated significant direct and indirect effects between the variables, as evidenced by p-values below 0.002. However, the study found no significant effect of customer attitudes on sustainable green purchasing decisions, with a p-value of 0.659, suggesting that customer attitudes alone may not be a decisive factor in promoting green purchasing behavior.

Zhou et al. [11] conducted a study that combined hierarchical clustering with an extended version of the RFM (Recency, Frequency, Monetary) analysis to improve customer segmentation in a retail use case. In this extension, the researchers introduced the concept of interpurchase time, which is defined as the time gap between two consecutive purchases made at the same location, whether it be a website or a physical store. By incorporating this additional variable, the RFM analysis was expanded to include four distinct features, enhancing the ability to analyze and segment customer behaviors. Following the calculation of these features, the customers were clustered into groups based on their purchasing patterns, helping retailers better understand and target their audience.

Jiang et al. [12] redefined the concept of perceived retailer innovativeness (PRI) within the context of smart retailing and proposed three dimensions of PRI: solution innovativeness, experience innovativeness, and meaning innovativeness. Drawing on the self-determination theory, the study investigated the impact of these three dimensions on customer engagement behavior (CEB). Using a sample of 428 respondents, the empirical analysis revealed that all three dimensions of PRI positively influenced the self-determined satisfaction of customers. This, in turn, was found to have a positive relationship with customer engagement behavior, suggesting that retailers who prioritize innovation in solutions, customer experiences, and the meanings they create for consumers can enhance customer satisfaction and engagement in smart retailing environments.

In the work of Archawaporn et al. [13], the authors analyzed and clustered clickstream data from previous anonymous sessions to develop a predictive model capable of identifying a user's profile after just a few clicks during an online session. This neural network-based model allows e-commerce platforms to anticipate customer preferences and behavior early in the interaction process, which can then be used by the platform's decision-making system to generate personalized recommendations. By tailoring services and recommendations based on the user's predicted profile, e-commerce businesses can improve the customer experience and increase the likelihood of conversions, making the model a valuable tool for online retailers.

Xu et al. [14] conducted an experiment using a dataset of sales event logs generated from an Omnichannel distribution service system, which included multiple sales channels such as marketplaces, web stores, social media, social media shops, and media messengers. To analyze the sales processes, the researchers employed several process mining algorithms, including the Inductive Miner Algorithm, Heuristic Algorithm, Alpha Miner Algorithm, and Fuzzy Miner Algorithm. The goal was to model the sales processes and measure their accuracy using conformance checking techniques. After comparing the models generated by each algorithm, the Fuzzy Miner Algorithm was found to produce the best process model for predicting consumer behavior in Omnichannel distribution services. This model was highly effective in capturing the complexities of consumer behavior across multiple sales channels, providing insights that could help businesses optimize their sales strategies.

In a study by Muangpan et al. [15], the researchers explored the factors and indicators that contribute to customer satisfaction in the transportation and distribution aspects of an online shopping company. Using survey research and exploratory factor analysis (EFA), the study focused on Chinese customers in Thailand who had used the company's services. The results were tested for validity, reliability, and analyzed using the Kaiser-Meyer-Olkin (KMO) statistical approach.

The factor loadings confirmed the validity of the factor explanations, with loadings exceeding the common threshold of 0.7. The study identified three primary factors that influenced customer satisfaction: the responsibility of product and delivery in transportation and distribution, the quality of customer service, and the effectiveness of communication during transportation and distribution. These factors were further broken down into fifteen key indicators, providing a detailed framework for understanding how different aspects of the transportation and distribution process impact customer satisfaction.

Bong et al. [16] conducted research to examine the application of the Technology Acceptance Model (TAM) to consumer behavior in social media-based online commerce. To collect data, a survey was distributed to a sample of 400 respondents, all of whom were part of Generation Z in Indonesia and regularly used social media platforms. The results of the data analysis, conducted using the TAM framework, showed that all six hypotheses proposed in the study were significant and accepted. This indicates that various aspects of customer behavior, such as perceived ease of use and perceived usefulness, have a significant influence on the decisions Generation Z consumers make when shopping online via social media. The study highlights the importance of understanding how user behavior shapes purchasing decisions in the context of social media, which has become an increasingly important platform for online commerce.

Lastly, Vankhede et al. [17] explored the application of predictive analytics to user behavior analysis on e-commerce websites, with a focus on using the random forest method to forecast behavioral patterns. The study utilized a dataset that included variables such as purchase history, browsing habits, and consumer demographics. In addition to predictive modeling, the study also applied exploratory data analysis (EDA) to extract insights from the dataset. By employing the random forest algorithm, which is renowned for its ability to handle complex datasets and generate accurate predictions, the study aimed to improve the understanding and prediction of user behavior on e-commerce websites. The random forest method builds multiple decision trees and combines their predictions using ensemble learning, leading to more reliable and accurate outcomes. The model was used to predict various aspects of user behavior, including browsing patterns, the likelihood of making a purchase, and customer segmentation, providing valuable insights that could help e-commerce platforms optimize their marketing and sales strategies.

## III. Customer Satisfaction Indicators

Shopping websites need to monitor and updates its system to increase good experience of customers. Hence some of indicators one use to monitor supply chain of delivery [18], are mentioned below.

- Product delivery is fast and punctual.
- Product arrival at the scheduled time
- Excha9nged Product on delivery

- Product correct delivery
- Product good packaging delivery
- Customer information safety on delivery
- Arrange Product a delivery time
- Picking up products and delivering door to door
- Convenience of return product delivery
- Carrier personality on delivery
- Informing customers before delivery

Apart from above parameters people can use other set for checking the customer experience based on Customer Care Executive.

- Courtesy of staff
- Knowledge of staff
- Quality of advice & information received
- Ease of finding telephone number
- Speed /efficiency with which query was dealt with
- Speed with which phone was answered

Digital advertising campaigns use a wide array of media channels to deliver their message to the target audience. These channels act as points of contact, where businesses and customers interact. In traditional offline marketing, commonly used channels include television, radio, print media such as magazines, and physical banners or billboards. However, in the digital marketing realm, the scope of channels expands significantly. Marketers leverage social media platforms, search engine advertising (e.g., Google Ads), display ads on various websites, and email marketing campaigns to reach their audience. One of the defining features of digital campaigns is the flexibility to use multiple channels to deliver a cohesive campaign across different platforms, thereby reaching a broader and more diverse audience.

It is important not to confuse a digital campaign with individual advertisements placed on different websites. Instead, a digital campaign is often delivered through a mix of various channels, with the intent to reach the target audience through different touchpoints. These touchpoints are essentially moments when the customer interacts with the brand or its advertising material. These interactions can take different forms, such as clicking on an ad, watching a video, or being exposed to a display banner on their mobile phone, tablet, or desktop. The device used to engage with the ad, along with the nature of the interaction, defines the customer's touchpoint experience.

## IV. Digital Data Analysis Techniques

### 1. Descriptive Analytics

Descriptive analytics serves as the foundational layer of digital data analysis, focusing on summarizing and interpreting historical data to understand what has occurred within a dataset. Techniques such as data visualization and statistical summaries are integral to this process. Data visualization tools like Tableau, Power

International Journal for Research Trends in Social Science & Humanities
E-ISSN 2584-2455
Volume 2 Issue 2
Mar-Apr 2024, PP 156-165

BI, and libraries like matplotlib in Python enable analysts to create graphs and charts that reveal patterns, trends, and correlations within the data, facilitating easier interpretation (Few, 2009) [19]. Additionally, statistical measures such as mean, median, mode, variance, and standard deviation provide quantitative summaries that help in understanding the central tendencies and variability of the data. Exploratory Data Analysis (EDA), as introduced by Tukey (1977) [20], further enhances descriptive analytics by allowing analysts to investigate data distributions and identify anomalies or outliers that may influence subsequent analyses.

**2. Diagnostic Analytics**

Diagnostic analytics delves deeper into data to uncover the reasons behind observed outcomes by examining relationships and causations within the dataset. This involves techniques such as correlation analysis, regression analysis, and Analysis of Variance (ANOVA). Correlation analysis assesses the degree to which two variables move in relation to each other, providing insights into potential associations (Montgomery, 2017) [21]. Regression analysis, including both linear and logistic regression, models the relationship between dependent and independent variables, enabling the prediction of outcomes based on various predictors. ANOVA is employed to compare means across multiple groups, determining whether there are statistically significant differences between them. Additionally, root cause analysis helps in identifying the underlying factors that contribute to specific data patterns or issues, thereby facilitating informed decision-making.

**3. Predictive Analytics**

Predictive analytics focuses on forecasting future events or trends based on historical data by leveraging statistical models and machine learning algorithms. Supervised learning methods, such as decision trees, random forests, and support vector machines, are commonly used to make predictions by learning from labeled datasets. Time series analysis, including models like ARIMA and exponential smoothing, is particularly effective for predicting future values in sequential data (Hastie et al., 2009) [22]. Furthermore, neural networks and deep learning techniques have revolutionized predictive analytics by enabling the modeling of complex, non-linear relationships within large datasets, which are essential for tasks like image and speech recognition. These predictive models not only enhance forecasting accuracy but also support proactive decision-making across various industries.

**4. Prescriptive Analytics**

Prescriptive analytics goes beyond prediction by recommending actions to achieve desired outcomes, utilizing optimization and simulation techniques. Optimization models, such as linear programming and integer programming, are employed to identify the best possible solutions within given constraints. Simulation methods, like Monte Carlo simulations and discrete event simulations, allow analysts to explore different scenarios and assess their potential impacts, providing a dynamic understanding of complex systems (Law & Kelton, 2000) [23]. Decision analysis frameworks, including decision trees and Bayesian networks, support the evaluation of various alternatives under uncertainty, facilitating strategic planning and operational efficiency. By integrating these techniques, prescriptive analytics provides actionable insights that guide organizations in making informed and effective decisions.

**5. Big Data Analytics**

Big data analytics addresses the challenges associated with processing and analyzing extremely large and complex datasets that traditional data processing tools cannot handle efficiently. This field leverages distributed computing frameworks such as Hadoop and Apache Spark to perform parallel processing across multiple machines, enabling the handling of vast amounts of data. NoSQL databases like MongoDB and Cassandra are designed to store unstructured or semi-structured data, offering flexibility and scalability essential for big data applications. Real-time analytics tools, including Apache Storm and Apache Flink, facilitate the immediate processing of streaming data, which is critical for applications requiring timely insights. Big data analytics thus empowers organizations to harness the full potential of their data assets, driving innovation and competitive advantage.

**6. Text and Sentiment Analysis**

Text and sentiment analysis involve the extraction of meaningful information from textual data, enabling the understanding of underlying sentiments and topics. Natural Language Processing (NLP) techniques, such as tokenization, part-of-speech tagging, and named entity recognition, break down text into analyzable components (Jurafsky & Martin, 2020). Sentiment analysis algorithms assess the emotional tone of text, categorizing it as positive, negative, or neutral, which is valuable for gauging public opinion and customer satisfaction (Bird et al., 2009) [24]. Topic modeling techniques, including Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF), identify prevalent themes within large document collections, aiding in content organization and information retrieval. These methods are instrumental in various applications, from social media monitoring to customer feedback analysis, providing deep insights into textual data.

**7. Data Mining**

Data mining encompasses a range of techniques aimed at discovering hidden patterns and knowledge from large datasets. Association rule learning, exemplified by algorithms like Apriori and Eclat, identifies relationships between variables, such as items frequently purchased together in retail settings. Anomaly detection methods pinpoint rare events or outliers that may indicate fraud, system failures, or other significant occurrences. Sequential pattern mining uncovers regular sequences of events, which is useful for understanding user behavior and predicting future actions. Data mining thus facilitates the extraction of actionable intelligence from data, supporting strategic decisions and operational improvements across various domains.

**8. Data Cleaning and Preprocessing**

Data cleaning and preprocessing are critical steps in the data analysis pipeline, ensuring that the data is accurate, consistent, and ready for analysis. This involves tasks such as data integration, where data from multiple sources are combined into a coherent dataset, and data transformation, which includes normalizing or scaling data to facilitate comparison and analysis (Kotsiantis et al., 2006) [25]. Handling missing values is another essential aspect, where techniques like imputation or deletion are used to address incomplete records (Rahm & Do, 2000). Outlier detection methods identify and treat anomalies that could distort analysis results, thereby enhancing data quality and reliability. Effective data cleaning and preprocessing lay the groundwork for robust and meaningful data analysis, minimizing errors and biases in subsequent steps.

**9. Ethical and Privacy Considerations**

Ethical and privacy considerations are paramount in digital data analysis, ensuring that data is used responsibly and that individuals' privacy is protected. Techniques such as anonymization remove personally identifiable information from datasets, mitigating the risk of privacy breaches. Data governance frameworks establish policies and procedures for managing data access, usage, and security, promoting ethical standards and compliance with legal regulations (O'Neil, 2016) [26]. Additionally, addressing ethical issues like bias, fairness, and transparency is crucial for maintaining public trust and ensuring that data-driven decisions do not inadvertently cause harm. As data analysis becomes increasingly pervasive, ensuring ethical and fair use of data is essential for responsible and sustainable analytics practices.

## V. Conclusion

In conclusion, the studies reviewed in this paper collectively underscore the intricate relationship between customer behavior, innovative retail strategies, and the dynamics of online commerce. They highlight the significance of customer value and the role of attitudes in sustainable purchasing decisions while enhancing customer segmentation through integrated approaches to analysis. The research emphasizes the positive impact of perceived retailer innovativeness on customer engagement in smart retailing and demonstrates how predictive modeling from clickstream data can personalize e-commerce experiences. Additionally, effective process modeling in Omnichannel distribution is shown to be crucial for understanding consumer behavior, alongside the identification of critical factors influencing customer satisfaction in online transportation services. Finally, the application of the Technology Acceptance Model illustrates the importance of understanding Generation Z's purchasing behavior on social media, underscoring the need for retailers to adapt their strategies to optimize engagement and meet the evolving demands of the digital marketplace.

## References

1. Y. L. Chong, E. Ch'ng, M. J. Liu, and B. Li, ''Predicting consumer product demands via big data: The roles of online promotional marketing and online reviews,'' Int. J. Prod. Res., vol. 55, no. 17, pp. 5142–5156, Sep. 2017.

2. F. Hou, B. Li, A. Y.-L. Chong, N. Yannopoulou, and M. J. Liu, ‘‘Understanding and predicting what influence online product sales? A neural network approach,’’ Prod. Planning Control, vol. 28, nos. 11–12, pp. 964–975, Sep. 2017.
3. CNNIC. The 41st Internet Development Statistics Report. Accessed: Mar. 5, 2018.
4. T. Tang, E. Fang, and F. Wang, ‘‘Is neutral really neutral? The effects of neutral user–generated content on product sales,’’ J. Marketing, vol. 78, no. 4, pp. 41–58, Jul. 2014.
5. K. Floyd, R. Freling, S. Alhoqail, H. Y. Cho, and T. Freling, ‘‘How online product reviews affect retail sales: A meta-analysis,’’ J. Retailing, vol. 90, no. 2, pp. 217–232, Jun. 2014.
6. W. Shi, X. Gong, Q. Zhang, and L. Wang, ‘‘A comparative study on the first-time online reviews and appended online reviews,’’ J. Manage. Sci., vol. 29, no. 4, pp. 45–58, 2016.
7. Q. Min, L. Qin, and K. Zhang, ‘‘Factors affecting the perceived usefulness of online reviews,’’ Manage. Rev., vol. 29, no. 10, pp. 95–107, 2017.
8. J. Cho, ‘‘Likelihood to abort an online transaction: Influences from cognitive evaluations, attitudes, and behavioral variables,’’ Inf. Manage., vol. 41, no. 7, pp. 827–838, Sep. 2004.
9. N. Sondhi, ‘‘Segmenting & profiling the deflecting customer: Understanding shopping cart abandonment,’’ Procedia Comput. Sci., vol. 122, pp. 392–399, Jan. 2017.
10. Alkhawaldeh, K., Suwan, R. & Qadri, M. The influence of customer value on sustainable green purchasing decision: the mediating role of customer attitudes and behavior. OPSEARCH (2024).
11. Zhou J, Wei J, Xu B (2021) Customer segmentation by web content mining. J Retail Consum Serv.
12. Jiang, K., Zheng, J., Qin, M. et al. Perceived retailer innovativeness and its impact on customer engagement behavior in smart retailing. Serv Bus 18, 255–285 (2024).
13. Laksamon Archawaporn and Adisorn Leelasantitham. "Managing Factors to Stages of the Online Customer Journey Influence on Brand Trust". IEEE Access 12 July 2021.
14. G. Xu, K. Kang and M. Lu, "An Omnichannel Retailing Operation for Solving Joint Inventory Replenishment Control and Dynamic Pricing Problems From the Perspective of Customer Experience," in IEEE Access, vol. 11, pp. 14859-14875, 2023.
15. Muangpan, T. (2022). Customer satisfaction indicators in transportation and distribution factors of online shopping company. Cogent Business & Management, 9(1).
16. S. L. Bong, A. B. Y. Adriawan, G. Kamilah and E. D. Madyatmadja, "Analysis of Consumer Behavior in Online Shopping on Social Media," 2023 International Conference On Cyber Management And Engineering (CyMaEn), Bangkok, Thailand, 2023, pp. 449-453,
17. P. Vankhede and S. Kumar, "Predictive Analytics for Website User Behavior Analysis," 2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2024, pp. 1-6.

18. Muangpan, T. (2022). Customer satisfaction indicators in transportation and distribution factors of online shopping company. Cogent Business & Management, 9(1).
19. Few, S. (2009). Now You See It: Simple Visualization Techniques for Quantitative Analysis. Analytics Press.
20. Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley.
21. Montgomery, D. C. (2017). Design and Analysis of Experiments. Wiley.
22. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. Springer.
23. Law, A. M., & Kelton, W. D. (2000). Simulation Modeling and Analysis. McGraw-Hill.
24. Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.
25. Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). Data Preprocessing for Supervised Learning. International Journal of Computer Science, 1(2), 111–117.
26. O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Publishing Group.