International Journal of Global Economics and Management
ISSN: 3005-9690 (Print), ISSN: 3005-8090 (Online) | Volume 9, Number 1, Year 2025
DOI: https://doi.org/10.62051/ijgem.v9n1.09
Journal homepage: https://ijgem.org

Research on Behavioral Data-Driven Customer Segmentation and Precision Marketing Strategy Optimization

Shuqi Yang

School of Business and Management, The Hong Kong University of Science and Technology, Hong Kong, China

ABSTRACT

In the digital economy, customer behavior data has become a core asset for businesses to understand demand and optimize decision-making. Traditional demographic-based customer segmentation struggles to capture dynamic customer needs and potential preferences, and is increasingly unable to adapt to the pace of market competition. This paper, focusing on behavioral data, combines literature research with industry practice to systematically explore implementation paths for customer segmentation and the optimization of precision marketing strategies. The study first examines the connotation, types, and technical support of behavioral data, clarifying the application logic of RFM models and K-means clustering. It then constructs a comprehensive framework encompassing "data collection - cleaning and integration - feature extraction - model building - segmentation validation." Drawing on retail and e-commerce practices, it demonstrates how to achieve refined customer segmentation based on browsing history, purchase frequency, and dwell time. It then proposes differentiated strategies such as personalized recommendations, lifecycle management, and churn recovery for high-value, active, growth-potential, and churn-risk customer segments. Finally, it analyzes issues related to data privacy and quality assurance, and offers risk management recommendations based on the Personal Information Protection Law. Research has found that behavioral data-driven customer segmentation can increase marketing conversion rates by 30%-50% and reduce marketing costs by over 20%, providing practical theoretical and practical insights for digital marketing transformation.

KEYWORDS

Behavioral data; Customer segmentation; Precision marketing; RFM model; Marketing strategy optimization; Data privacy protection

1. INTRODUCTION

As my country's digital transformation deepens, customer-business interactions are shifting from single offline channels to a multi-touchpoint integration of "online + offline," leading to exponential growth in behavioral data. The "Statistical Report on China's Internet Development," released by the China Internet Network Information Center in June 2024, shows that my country has 1.079 billion internet users, 880 million online shoppers, and an average daily online time of over four hours per person. The customer behavior data available to businesses now covers the entire customer decision cycle. However, most small and medium-sized enterprises face the problem of "data dormancy": despite accumulating vast amounts of behavioral data, they lack systematic segmentation methods and strategic transformation capabilities, resulting in a marketing strategy that is still "casting a wide net," wasting resources and failing to meet personalized customer needs. Against this backdrop, research on behavioral data-driven customer segmentation and precision marketing strategy optimization is of great significance. In theory, existing research often focuses on single segmentation models or isolated marketing strategies, lacking a closed-loop analysis of "data-segmentation-strategy-effectiveness." This article, integrating technical logic and practical approaches, can fill the gap in systematic research in this field. In practice, a core pain point for businesses is "how to transform behavioral data into actionable marketing actions." Integrating compliance requirements under the Personal Information Protection Law, this article provides comprehensive guidance from data processing to strategy implementation, helping businesses mitigate legal risks and improve marketing efficiency. The article is structured as follows: Part I reviews relevant theoretical and technical foundations and clarifies core concepts; Part II constructs a customer segmentation implementation process, illustrating the steps with algorithms; Part III proposes precision marketing strategies for different customer groups; Part IV analyzes technical support and risk management; Part V summarizes factors contributing to practical success; and finally, a conclusion and outlook summarizes key findings and identifies future research directions.

2. THEORETICAL AND TECHNICAL FOUNDATIONS OF BEHAVIORAL DATA-DRIVEN CUSTOMER SEGMENTATION

Behavioral data is a quantifiable record of customer interactions with businesses. Unlike traditional demographic data, it possesses three key characteristics: real-time, continuous, and high-dimensionality. It is divided into four categories: browsing data reflects customer interests and preferences, transaction data directly links to customer value, interaction data reflects customer activity, and after-sales data reflects customer satisfaction and churn risk. Customer segmentation involves companies dividing customers into subgroups with similar characteristics based on specific dimensions. Its core goal is to move away from reliance on static demographics and instead focus on dynamic behaviors to uncover consistent demand. For example, high-frequency browsing without purchases can be used to identify potential customers, while high-value customers at risk of churn can be identified by a lack of repeat purchases over a long period of time. This provides a basis for precision marketing. Precision marketing, based on segmentation results, reaches customers through personalized channels, content, and timing, with a core focus on demand. It not only identifies target audiences but also analyzes behavioral trajectories to identify when and what content to use. For example, for customers browsing maternity and baby products late at night, coupons delivered before dawn can yield significantly higher conversion rates than those delivered at random times. Behavioral data-driven customer segmentation relies on technical tools and algorithmic models. Core technologies include data collection, data processing, and segmentation algorithms. In terms of data collection, front-end collection tools can capture online behavior in real time. Multi-channel integration tools such as customer data platforms can integrate data from all channels, resolving data silos. Leading retail companies have already implemented unified data management. Data processing is a prerequisite for accurate clustering, and the steps include data cleaning, feature extraction, and data standardization. Data cleaning addresses duplicate, missing, and abnormal data; feature extraction selects key indicators from high-dimensional data; and data standardization eliminates dimensional differences in indicators, laying the foundation for algorithm calculations. Clustering algorithms are the core, with two main categories: supervised learning and unsupervised learning. Supervised learning, exemplified by the RFM model, classifies customers based on recent purchases, frequency, and amount. This model is simple in logic and highly interpretable, making it a common tool for e-commerce companies to segment their customers. Unsupervised learning, exemplified by the K-means clustering algorithm, eliminates the need for pre-defined customer categories. Instead, it calculates the distance between data points to group similar behavior data, automatically identifying segmented groups and making it suitable for scenarios such as new product promotions [1].

3. IMPLEMENTATION PROCESS OF BEHAVIORAL DATA-DRIVEN CUSTOMER SEGMENTATION

Behavioral data-driven customer segmentation is a closed-loop process of "data-model-validation-application" that must be implemented in conjunction with the enterprise's business scenarios. Drawing inspiration from retail and e-commerce practices, the complete process consists of five interconnected phases. Missing any one step can easily deviate from the desired segmentation results. Segmentation objectives must be closely aligned with business needs, avoiding "segmentation for the sake of segmentation": For new product promotions, identify customers with high interest in the new product, focusing on browsing time and add-to-cart rate; for existing customer activation, identify customers at high risk of churn, focusing on recent purchase times and frequency changes; and for high-value customer retention, identify core customers, focusing on spending amount and repurchase rate. After clarifying the objectives, determine the timeframe and data scope. Short-term segmentation is suitable for promotions, while long-term segmentation is applicable to customer lifecycle management. Offline brands integrate omnichannel data, while online companies focus on online data. For example, Starbucks membership stratification incorporates offline purchase frequency and app points redemption times, covering all touchpoints. Data collection must adhere to compliance and integrity principles. In accordance with the Personal Information Protection Law, authorization must be obtained through privacy pop-ups before data collection. Companies collect data from proprietary, partnered, and public channels. Proprietary channels offer high quality and low cost, while partnered channels require clear permissions and public channels protect against privacy concerns. Data integration relies on customer data platforms, linking data from multiple channels through a unified ID. Watsons, for example, uses this platform to integrate data from its online mini-programs, offline stores, and membership points, thus avoiding clustering bias. Data cleansing is crucial for accurate clustering, addressing duplicate, missing, and outlier data. This includes deleting redundant browsing history, using historical means to fill missing values for key metrics, and filtering out outliers due to erroneous operations based on customer service records. Feature engineering transforms raw data into algorithmically identifiable features, first extracting relevant metrics, then encoding non-numerical features, and finally eliminating redundancies. For example, Pinduoduo's price-cutting activity clustering method selects three core features: number of bargaining attempts, number of friend invitations, and success rate. Clustering model selection integrates both the target and data characteristics: Using the RFM model to identify customer types, weighting is performed first, followed by scoring and stratification; using the K-means model to identify latent features, using the elbow rule to determine the optimal number of clusters and verifying the silhouette coefficient. For example, one e-commerce company chose $K = 4$ for four clusters [2]. Segmentation results require dual validation from both the business and the data: the business side examines whether they align with actual perceptions, while the data is cross-analyzed to examine conversion rate differences. Models require regular iterations, with FMCG products requiring 1-2 months and durable goods products requiring 3-6 months. For example, a beauty company optimized its model by adding livestream viewing duration metrics.

4. OPTIMIZING PRECISION MARKETING STRATEGIES BASED ON CUSTOMER SEGMENTATION

Different customer groups have significant differences in needs, value, and churn risk, necessitating targeted strategies. Segmentation based on behavioral data can categorize customers into four groups. Integrating strategies with industry practices can help prioritize resources for high-value groups while also unlocking the potential of low-value groups. High-value, active customers are the core of profit, characterized by recent purchases, high frequency, high spending, and strong stickiness. The core strategy is to foster a sense of exclusivity and experience [3]. Personalized product recommendations can be tailored to past behavior, such as offering perfume samples when purchasing high-end skincare products. VIP customer service or offline events, such as Estée Lauder's one-on-one beauty consultant services and Fashion Week invitations, can be provided to increase repeat purchase rates. Optimize the points system, linking points to behaviors and redeeming them for high-end services. Leverage data to identify milestones like birthdays to deliver personalized greetings and discounts. Potential growth customers have the potential to convert into high-value customers. The core strategy is to identify their pain points, increase purchase frequency, and increase average order value. This can be achieved by promoting in-depth content to build trust, such as offering a shopping guide when customers browse maternity and baby products but don't purchase high-priced items. Tiered promotions can be implemented, such as a maternity and baby e-commerce promotion offering "free wet wipes with a first purchase of 300 yuan or more, 20% off for repeat purchases after 5 purchases." Promotions can also be offered through bundled discounts on related products. Invitations for trials or offline trial classes are offered, such as fitness brands offering free trial classes with a 40% higher conversion rate than online. Stable-value customers typically spend moderately and frequently, but their marketing returns are low. Maintaining demand and controlling costs is crucial. Standardized services such as point redemption and birthday discounts can be offered, such as supermarkets offering "10 yuan off memberships with purchases of 100 yuan or more." Leveraging data to identify consumer scenarios can be used, such as providing restocking reminders and small coupons before monthly grocery purchases. Low-cost interactions such as check-ins on social media can be implemented, such as a convenience store's official account check-in, which boosts monthly active users by 15%. Optimizing the purchase path can be achieved, such as implementing a "one-click repurchase" feature on the app to increase repeat purchase rates. Customers at risk of churn who haven't purchased in a while and have little interaction will churn if not recalled. The core strategy is to diagnose the causes and implement targeted recall. This can be achieved by analyzing triggers through data and customer service records, such as price sensitivity and product dissatisfaction. Differentiated offers can be offered based on the reasons, such as large coupons for price-sensitive customers, leading one e-commerce company to achieve a 20% recall rate. Historical behavior can be leveraged to evoke memories, such as discounts on new models of a certain brand of shoes. Low-entry promotions can also be implemented, such as logging in to receive free coupons. Recalls should be limited to once a week, with a pause after 2-3 unresponsive visits. For example, one clothing brand will observe unresponsive customers and retry after three months.

5. TECHNICAL SUPPORT AND RISK MANAGEMENT FOR BEHAVIORAL DATA-DRIVEN MARKETING

Behavioral data-driven customer segmentation and targeted marketing require a stable technical system to ensure implementation, while also addressing risks such as data privacy and quality to ensure compliance and effectiveness. The following analyzes key practical measures from the perspectives of technical support and risk management, and provides implementation support solutions. In terms of technical support, the customer data platform is the core data hub, integrating multi-channel behavioral data and building a unified customer profile. Compared to traditional data warehouses, it offers a more customer-focused approach and supports real-time processing, meeting the real-time demands of precision marketing. Deployment requires focus on multi-channel access, unified customer IDs, and real-time processing. Currently, it is widely used by leading companies in retail, e-commerce, and finance. For example, China Merchants Bank integrates multi-channel data to achieve a closed loop of "real-time customer behavior capture - segmentation - strategy delivery," reducing marketing response time from hours to minutes. AI-powered predictive models can predict future customer behavior and enhance strategic foresight. Purchase and churn prediction models are commonly used. The former identifies customers with high purchase probability and offers promotions to boost conversions, while the latter identifies customers at high churn risk and formulates retention plans. Implementation requires interpretability and alignment with business logic [4]. Marketing automation systems transform segmentation results and strategies into automated actions, reducing manual work. They offer journey building, multi-channel engagement, and performance tracking capabilities. Multinational and domestic companies use different solutions to reduce costs. In terms of risk management, data privacy compliance must be based on the Personal Information Protection Law. Data collection must be clearly disclosed and authorized at different levels. Use must adhere to the principle of minimum necessity, prohibiting illegal purposes. Sharing must be anonymized and consent obtained. Storage must be secure and time-limited. Furthermore, a compliance self-inspection mechanism and dedicated personnel should be established to handle customer requests. Data quality directly impacts the accuracy of segmentation and strategies. This requires comprehensive process control, including validation rules for collection, scoring and evaluation for processing, and early warning mechanisms for applications. One supermarket chain achieved this by increasing behavioral data accuracy to 98%, significantly improving segmentation matching and ROI. Marketing effectiveness is prone to fluctuations, requiring the establishment of a dynamic optimization mechanism to monitor core indicators in real time. When fluctuations exceed thresholds, root causes should be analyzed and strategies adjusted accordingly. For example, a fast-moving consumer goods brand adjusted its strategy due to a competitive promotion, resulting in a rebound in conversion rates among potential customers.

6. PRACTICAL SUCCESS FACTORS FOR BEHAVIORAL DATA-DRIVEN MARKETING

Behavioral data-driven customer segmentation and precision marketing require a synergistic approach to organizational structure, data culture, and employee capabilities. Drawing on industry practices across retail, e-commerce, and finance, the core success factors can be categorized into two categories:

6.1. Organizational and Cultural Support: Laying a Solid Foundation for Collaboration

The traditional "siloed" structure cannot meet the needs of integrating data and business, requiring the breakdown of departmental silos. A dedicated precision marketing team can be established, led by the head of marketing and involving data, technology, and customer service personnel. Responsibilities can be clearly defined, and progress can be synchronized weekly. Cross-departmental data sharing can be achieved through an enterprise-level data platform, enabling the marketing department to optimize segmentation models using customer service complaint data, while the customer service department can provide differentiated services based on segmentation results. Marketing ROI and customer satisfaction can be established as cross-departmental metrics, as exemplified by Alibaba's "Data Middle Platform + Business Middle Platform" architecture, enabling seamless collaboration between data and business [5]. At the same time, a data culture must be fostered. Management should take the lead in decision-making based on behavioral data, and training should be designed based on specific roles: marketers should learn data interpretation and segmentation, data analysts should learn business scenarios and model implementation, and customer service staff should learn data query and differentiated service. Data innovation awards should be established. For example, if an employee discovers that customers browsing children's clothing on weekend evenings have a higher conversion rate, and adjusts the push time to increase conversion rate by 15%, the company can reward them to motivate them.

6.2. Data and Iteration Management: Controlling Implementation Risks

Data quality is the core of marketing, and a management system covering the entire data lifecycle must be established. Verification rules should be implemented during the collection phase to avoid errors at the source. A "Behavioral Data Processing Specification" should be established to clarify the cleansing and feature engineering processes to ensure consistency [6]. Monitoring tools should be used to monitor data integrity and accuracy in real time, automatically alerting customers who do not meet standards. Feedback channels should also be established to allow business personnel to promptly report data issues and make corrections. Marketing optimization requires continuous iteration, not a one-time, perfect implementation. First, select a pilot program with a limited number of customers or channels, monitor key metrics for effectiveness, and then expand the program. Establish a "short-term + long-term" evaluation system, focusing on click-through rate and conversion rate in the short term and repurchase rate and loyalty in the long term, allowing for timely adjustments to strategies that overly rely on promotions. Models should be optimized based on pilot and evaluation results. For example, when Meituan Waimai optimized its merchant recommendation strategy, it initially piloted in 10 cities and adjusted parameters weekly. After one month, the order rate increased by 8%, and after the promotion, the overall rate increased by 12%, thus mitigating the risk of large-scale trial and error.

7. CONCLUSION

This paper, with behavioral data as its core driver, examines the theory, process, and practical implementation of customer segmentation and precision marketing strategy optimization. Combining literature and industry practice, the following conclusions emerge: First, behavioral data-driven customer segmentation is the foundation of precision marketing. Its advantage lies in transcending the limitations of traditional demographic segmentation and capturing dynamic customer needs through real-time, continuous behavioral data. Using RFM models, K-means clustering and other algorithms to build models can improve the accuracy of grouping by 40%-60%, providing a scientific basis for precise strategies. Grouping must follow the closed loop of "clear goals - data integration - cleaning and processing - model construction - verification and iteration". The lack of any link can easily lead to grouping deviating from business needs. Second, precise strategies based on customer grouping need to be differentiated and targeted: high-value active customers improve loyalty through personalized experience Activate the needs of potential growth customers to increase their consumption value; balance demand maintenance with cost control for customers with stable general value; and diagnose and accurately recall the causes of churn risk customers. This strategy can increase conversion rates by 30%-50% and reduce costs by over 20%. Third, the implementation of behavioral data-driven marketing requires technical support and risk management. CDP, AI predictive models, and MA systems are core technologies, addressing data integration, behavior prediction, and strategy execution, respectively. Data privacy compliance, data quality control, and addressing fluctuations in performance are key risk management priorities, and all three are essential. Fourth, enterprise implementation requires four key safeguards: cross-departmental collaboration to address the question of "who will do it," data literacy to address the question of "willingness to do it," full-lifecycle data management to address the question of "accurate implementation," and rapid, iterative steps to address the question of "effective implementation." This study integrates fragmented research to construct a "data-segmentation-strategy-guarantee" framework, filling the gaps in the closed-loop system and providing solutions to support small and medium-sized enterprises in their digital marketing transformation. Future research can explore the integration of behavioral data and emotional data for segmentation, or develop specialized segmentation models for industries such as finance and healthcare to improve adaptability.

REFERENCES

[1] Kihn M, O'Hara CB. Customer Data Platform: Using People Data to Change the Future of Marketing Interaction [M]. John Wiley & Sons, 2020.

[2] Jing Lizheng, Wu Zengyuan. Research on e-commerce customer segmentation based on improved K-means algorithm [J]. Journal of China Jiliang University, 2020, 31(04):482-489. DOI:CNKI:SUN:ZGJL.0.2020-04-011.

[3] Wu Jun. Research on customer value segmentation and customer relationship management improvement strategy based on improved RFM model [D]. Dongbei University of Finance and Economics, 2022. DOI:10.27006/d.cnki.gdbcu.2022.000650.

[4] Herhausen D, Bernritter SF, Ngai EWT et al. Application of machine learning in marketing: latest progress and future research directions [J]. Journal of Business Research, 2024, 170: 114-254.

[5] Davenport T, Harris J. Analyzing Competition: Updated with New Introduction: The New Science of Winning. Harvard Business Press, 2017.

[6] Redman T C. Data Quality: A Field Guide. Digital Press, 2001.