

Customer Segmentation and Purchase Behavior Analysis

Athiseshan A Student, Department of Management Studies, School of management studies,

Vels Institute Of Science Technology & Advanced Studies Email: aathiseshanarumugam@gmail.com

Dr. G. Madhumita, Professor, School of Management Studies,

Vels Institute of Science Technology & Advanced Studies Email: madhu.sms@vistas.ac.in

Abstract

This study examines consumer purchase behavior and develops a data-driven customer segmentation framework based on behavioral characteristics. The study uses primary data collected from 119 respondents through a structured online questionnaire covering demographic characteristics, shopping behavior and seventeen purchase-behavior statements. The behavioral statements were organized into five composite constructs—Price Sensitivity, Brand Loyalty, Deliberation, Impulsiveness and Service Satisfaction—and K-Means cluster analysis was used to identify distinct customer segments. The study identified four interpretable segments: Disengaged/Low-Involvement Shoppers, Price-Sensitive Deal Seekers, Impulsive Brand-Conscious Shoppers, and Value-Driven Loyal Customers. The largest segment was Impulsive Brand-Conscious Shoppers (35.3%), followed by Value-Driven Loyal Customers (30.3%), Price-Sensitive Deal Seekers (21.0%), and Disengaged/Low-Involvement Shoppers (13.4%). Chi-square analysis showed that segment membership was not significantly associated with occupation. One-way ANOVA indicated significant differences in overall purchase behavior across income groups and in brand loyalty across occupation and income groups. Pearson correlation showed a weak positive relationship between brand loyalty and service satisfaction ($r = 0.251$, $p = 0.006$). The findings demonstrate that behavioral segmentation can reveal meaningful customer differences that are not captured by demographic profiling alone and can support more targeted marketing, customer relationship and sales strategies.

Keywords: Customer segmentation, purchase behavior, behavioral segmentation, K-Means clustering, price sensitivity, brand loyalty, impulsiveness, customer satisfaction

1. Introduction

Every business ultimately depends on understanding its customers. Consumers differ in their needs, motivations, spending capacity, brand attitudes and purchasing habits. A uniform marketing approach can therefore overlook important differences in how customers evaluate products, respond to offers and decide when and where to buy.

Customer segmentation is the process of dividing a heterogeneous customer base into smaller and relatively homogeneous groups on the basis of shared characteristics. Segmentation may be demographic, geographic, psychographic or behavioral. Behavioral segmentation focuses on what customers actually do, including purchase frequency, spending patterns, price sensitivity, brand loyalty, information search and responsiveness to promotions.

Purchase behavior is shaped by both rational and psychological influences. Some consumers compare alternatives and read reviews before purchasing, while others respond quickly to convenience, promotions or emotional triggers. Some customers are strongly price-sensitive, whereas others attach greater importance to trusted brands, quality or service. Understanding these patterns enables marketers to develop differentiated communication and customer strategies.

The increasing availability of customer data and analytical techniques has made data-driven segmentation more accessible. Cluster analysis, particularly K-Means clustering, can group respondents according to similarities across multiple behavioral dimensions without requiring the researcher to assign the segment labels in advance.

The present study focuses on general cross-category consumer purchase behavior and uses a multi-dimensional behavioral framework to identify actionable customer segments. The study also examines whether the identified behavioral segments correspond with demographic characteristics and selected statistical relationships.

2. Review of Literature

Sharma and Verma (2021) examined the evolution of market segmentation from traditional demographic and geographic bases towards behavioral and data-driven approaches. Their work emphasized the usefulness of indicators such as purchase frequency, spending and promotional responsiveness and supported combining behavioral segmentation with demographic profiling.

Iyer and Kapoor (2022) applied K-Means clustering using recency, frequency and monetary value together with attitudinal responses. Their findings supported the use of cluster-based approaches for identifying groups with different purchase and loyalty characteristics.

Banerjee (2023) examined behavioral segmentation in financial-services distribution and argued that behavioral and attitudinal indicators can complement income and occupation when organizations assess customer receptiveness and cross-selling potential.

Chatterjee and Nair (2024) compared K-Means, hierarchical clustering and latent class analysis and reported that broad behavioral archetypes can remain consistent even when the exact cluster boundaries vary across methods. This supports careful selection of behavioral variables as an important part of segmentation design.

Reddy and Srinivasan (2020) studied price sensitivity and promotional offers among urban Indian consumers. Their findings highlighted discount-seeking behavior and category switching, supporting the treatment of price sensitivity as a separate behavioral dimension.

Fernandes and Pinto (2021) examined brand trust, brand reputation and repurchase intention and reported that brand reputation and customer service quality were important predictors of repurchase intention and recommendation behavior.

Krishnan (2022) distinguished value-consciousness from simple price sensitivity and noted that consumers can simultaneously value discounts, quality and brand reputation. This supports analyzing price sensitivity and brand loyalty as separate constructs.

Gupta and Malhotra (2021) examined information search, online reviews and social media influence. Their findings suggested that deliberation-oriented consumers represent a distinct group requiring detailed and information-rich marketing communication.

Desai and Shah (2023) found that deliberate information search and impulsive buying are related but distinct behaviors. This supports retaining deliberation and impulsiveness as separate segmentation variables.

Agarwal and Bose (2020) examined antecedents of impulsive buying and identified promotions, availability, peer influence and mood as important triggers. Menon (2023) further reported higher impulsiveness among younger consumers and students.

Varghese and Thomas (2021) examined satisfaction, service quality and repurchase intention and highlighted the importance of considering customer satisfaction together with brand loyalty. Joseph and Pillai (2022) similarly identified product availability and service responsiveness as important factors in repurchase behavior.

Ramesh and Kaur (2020) examined occupation, income, shopping frequency and channel preference. Philip and D'Souza (2023) examined family size and residence type in relation to purchase motivations. Rao and Ahluwalia (2022) and Sen and Bhattacharyya (2024) discussed the application of behavioral segmentation to financial-product distribution.

The literature indicates that behavioral segmentation is increasingly important, but there remains scope for Indian studies that combine multiple purchase-behavior dimensions in a single cross-category consumer dataset and then

connect the resulting segments with demographic and statistical analysis. The present study addresses this gap through a five-construct behavioral framework and K-Means clustering.

3. Research Methodology

A descriptive-cum-analytical research design was adopted. The descriptive component was used to understand the demographic and shopping profile of respondents, while the analytical component was used to test relationships and identify behaviorally distinct customer groups.

Primary data were collected from 119 respondents through a structured online questionnaire. Convenience sampling was used because respondents were selected based on accessibility and willingness to participate. The questionnaire contained three broad sections: demographic profile, shopping behavior and purchase-behavior statements.

The demographic section covered age, gender, marital status, educational qualification, occupation, monthly income, place of residence and family size. Shopping-behavior questions covered shopping frequency, preferred channel, monthly shopping expenditure, frequently purchased product categories and main reason for purchase.

Seventeen purchase-behavior statements were measured on a five-point Likert scale from Strongly Disagree (1) to Strongly Agree (5). These statements covered price sensitivity, brand orientation and loyalty, deliberation, impulsiveness, social and digital influence, and satisfaction with service and product availability.

Five composite behavioral constructs were created: Price Sensitivity, Brand Loyalty, Deliberation, Impulsiveness and Service Satisfaction. K-Means cluster analysis was then applied to standardized composite scores to identify behaviorally distinct customer segments.

Additional analysis included percentage analysis, weighted mean and agreement-ratio analysis, Chi-Square tests of independence, One-Way ANOVA and Karl Pearson's correlation. Statistical computation was carried out using Python-based analytical tools.

The study has limitations. The sample is restricted to 119 respondents and uses convenience sampling. The data are self-reported, and the questionnaire captures general cross-category consumer behavior rather than behavior from a single product category. The four-cluster solution is one interpretable partition of the sample, and other clustering methods or numbers of clusters could produce somewhat different boundaries.

4. Data Analysis and Interpretation

4.1 Respondent Profile

The sample was dominated by young respondents. The 18–25 age group represented 42.0% of respondents, followed by 26–35 years at 27.7%, below 18 at 15.1%, 36–45 at 14.3% and above 45 at 0.8%. Males represented 67.2%, females 29.4%, and 3.4% preferred not to disclose gender. Single respondents accounted for 59.7%, while 33.6% were married.

Postgraduates represented 39.5% of the sample and undergraduates 31.1%. Students formed the largest occupational group at 40.3%, followed by employees at 31.9%. In terms of income, 45.8% reported monthly income below ₹20,000 and 29.7% reported ₹20,001–₹40,000.

Respondents were distributed across urban (40.7%), semi-urban (31.4%) and rural (28.0%) locations. A family size of 3–4 members was most common at 42.4%, followed by 1–2 members at 40.7%.

4.2 Shopping Behaviour

Weekly shopping was the most common pattern at 39.8%, followed by monthly shopping at 33.1%, daily shopping at 16.9% and rarely at 10.2%. More than half of respondents (51.7%) used both online and offline channels, while 27.6% preferred online and 20.7% preferred offline shopping.

Monthly shopping expenditure was ₹2,000–₹5,000 for 39.5% of respondents, followed by ₹5,001–₹10,000 for 26.1%, below ₹2,000 for 23.5% and above ₹10,000 for 10.9%. Groceries (37.0%) and clothing (29.4%) were the most frequently purchased categories. Need (28.6%) and price (26.1%) were the leading stated reasons for purchase, followed by quality (22.7%) and brand (16.8%).

4.3 Overall Purchase-Behavior Responses

All seventeen purchase-behavior statements recorded weighted mean scores above the scale midpoint of 3.0. Agreement ratios ranged from 2.13:1 to 7.82:1, indicating an overall favorable orientation towards active purchase behavior. The statement relating to personal preferences in purchasing recorded the highest mean score of 4.02, while preference for buying from trusted brands recorded the lowest mean of 3.45.

4.4 Customer Segmentation Using K-Means

The seventeen statements were consolidated into five composite constructs. Price Sensitivity represented responsiveness to discounts, seasonal offers and value considerations. Brand Loyalty captured trust in known brands, reputation, recommendation and continued purchase intention. Deliberation covered product comparison, review reading, recommendations and social-media information. Impulsiveness represented spontaneous purchase tendency. Service Satisfaction captured product availability, customer service and satisfaction with recent purchases.

K-Means clustering identified four segments. Segment 1, Disengaged/Low-Involvement Shoppers, contained 16 respondents (13.4%) and recorded the lowest scores across the behavioral dimensions: Price Sensitivity 2.75, Brand Loyalty 2.86, Deliberation 2.44, Impulsiveness 2.25 and Service Satisfaction 3.27.

Segment 2, Price-Sensitive Deal Seekers, contained 25 respondents (21.0%). This segment recorded Price Sensitivity of 4.13 and relatively low Brand Loyalty of 3.11. These consumers showed strong responsiveness to discounts, seasonal offers and value-for-money considerations and relatively low attachment to particular brands.

Segment 3, Impulsive Brand-Conscious Shoppers, contained 42 respondents (35.3%) and was the largest segment. It recorded the highest Impulsiveness score at 4.21, together with Brand Loyalty of 3.93 and Deliberation of 3.80. The profile combines spontaneous purchasing with meaningful brand preference and product comparison.

Segment 4, Value-Driven Loyal Customers, contained 36 respondents (30.3%). It recorded the highest scores on four of the five dimensions: Price Sensitivity 4.35, Brand Loyalty 4.18, Deliberation 3.87 and Service Satisfaction 4.30, with Impulsiveness also high at 3.94. This segment combines value consciousness, brand loyalty, information search and strong service satisfaction.

Across all four segments, the combination of online and offline channels was the most common shopping preference. The Value-Driven Loyal segment showed the highest offline-only preference among the segments.

4.5 Statistical Testing

Chi-square analysis found no statistically significant association between occupation and customer segment. Other tested relationships—gender with shopping frequency, age with shopping channel, and income with main purchase reason—were also not statistically significant at the 5% level. These results indicate that behavioural segments are not simply a re-labelling of a single demographic variable.

One-Way ANOVA found no significant difference in the Overall Purchase Behaviour Index across occupation groups ($F = 0.777$, $p = 0.5684$). However, the index differed significantly across income groups ($F = 4.18$, $p = 0.0076$). Brand Loyalty also differed significantly across occupation groups ($F = 3.199$, $p = 0.0097$) and income groups ($F = 4.44$, $p = 0.0055$).

Karl Pearson's correlation between Brand Loyalty and Service Satisfaction was $r = 0.251$ with $p = 0.0058$, indicating a statistically significant but weak positive relationship. The result supports treating the two constructs as related but distinct behavioral dimensions.

5. Findings

The study found that the respondent base was relatively young, with students forming the largest occupational group and a substantial share of respondents reporting lower monthly income.

Weekly shopping and combined online-and-offline shopping were common patterns, indicating recurring and omni-channel purchase behavior within the sample.

Groceries and clothing were the most frequently purchased product categories, while need and price were the two most commonly cited purchase motivations.

All seventeen behavioral statements had means above the midpoint, indicating generally positive and active purchase orientations across the sample.

K-Means analysis identified four distinct behavioral segments: Disengaged/Low-Involvement Shoppers (13.4%), Price-Sensitive Deal Seekers (21.0%), Impulsive Brand-Conscious Shoppers (35.3%), and Value-Driven Loyal Customers (30.3%).

The Impulsive Brand-Conscious segment was the largest and combined high impulsiveness with above-average brand loyalty and deliberation.

The Value-Driven Loyal segment showed the strongest scores across four behavioral dimensions and combined price sensitivity with brand loyalty and service satisfaction.

Segment membership was not significantly associated with occupation, supporting the idea that behavioral segmentation captures differences beyond conventional demographic categories.

Income was significantly associated with differences in overall purchase-behavior engagement, while brand loyalty varied significantly across both occupation and income groups.

Brand Loyalty and Service Satisfaction showed a weak but statistically significant positive relationship, indicating that satisfaction and loyalty should be monitored as separate but connected constructs.

6. Suggestions

Organizations should combine behavioral segmentation with demographic profiling rather than relying on demographic variables alone.

For Disengaged/Low-Involvement Shoppers, communication can focus on awareness, basic product relevance and simple value propositions rather than complex promotional programmers.

For Price-Sensitive Deal Seekers, discount-led offers, value comparisons and transparent pricing can be used, while recognizing that this group may be responsive to competing offers.

For Impulsive Brand-Conscious Shoppers, time-bound promotions and visually engaging communication can be combined with credible brand information and easy purchase processes.

For Value-Driven Loyal Customers, relationship-building, loyalty initiatives, personalized recommendations and service-quality reinforcement can be used to maintain engagement.

Because online-and-offline shopping was common across segments, organizations should provide a consistent customer experience across channels.

Marketing performance should be evaluated using behavioral and outcome measures, not only broad demographic counts. Customer response, repeat purchase, satisfaction and segment movement can be monitored over time.

Future research can use larger and more diverse samples, actual transaction data, longitudinal tracking and alternative clustering methods to validate and refine the behavioral segments.

7. Conclusion

This study examined customer segmentation through purchase behavior using a structured survey and a data-driven clustering approach. The results show that customers differ meaningfully in price sensitivity, brand loyalty, deliberation, impulsiveness and service satisfaction, and that these differences can be converted into interpretable behavioral segments.

The four-segment solution provides a practical framework for understanding customers beyond demographic labels. The segments differ in their intensity of involvement, response to price and promotions, attachment to brands, tendency to deliberate and level of service satisfaction. The statistical results further indicate that segment membership

is not simply determined by occupation, while income and brand loyalty show significant variation in selected analyses.

The findings support the use of behavioral data as an important input for targeted marketing and customer relationship decisions. At the same time, the results should be interpreted within the limitations of the sample and study design. Future research using larger samples and actual customer transaction data can strengthen the validation and commercial applicability of the segmentation framework.

References

- Sharma, S., & Verma, R. (2021). Evolution of Market Segmentation Practice in Indian Retail and FMCG Firms.
- Iyer, A., & Kapoor, R. (2022). Cluster-Based Customer Segmentation Using RFM and Attitudinal Data in Organized Retail.
- Banerjee, P. (2023). Behavioral Segmentation Practices among Financial Services Distributors in India.
- Chatterjee, S., & Nair, P. (2024). A Comparative Study of Clustering Algorithms for Consumer Behavioral Segmentation.
- Reddy, R., & Srinivasan, S. (2020). Price Sensitivity, Promotional Offers and Purchase Decisions among Urban Indian Consumers.
- Fernandes, J., & Pinto, M. (2021). Brand Trust, Brand Reputation and Repurchase Intention among Indian Consumers.
- Krishnan, R. (2022). Value-Consciousness as a Consumer Behavioral Trait.
- Gupta, A., & Malhotra, S. (2021). Information Search, Online Reviews and Social Media Influence on Purchase Decisions.
- Desai, P., & Shah, R. (2023). Information Search Behavior and Impulsive Buying Tendencies.
- Agarwal, N., & Bose, S. (2020). Antecedents of Impulsive Buying Behavior among Indian Retail Consumers.
- Menon, A. (2023). Generational Differences in Impulsive Buying Behavior.
- Varghese, M., & Thomas, J. (2021). Customer Satisfaction, Service Quality and Repurchase Intention.
- Joseph, P., & Pillai, R. (2022). Product Availability, Customer Service Responsiveness and Repurchase Behavior.
- Ramesh, K., & Kaur, P. (2020). Occupation, Income, Shopping Frequency and Channel Preference among Indian Consumers.
- Philip, J., & D'Souza, M. (2023). Family Size, Residence Type and Purchase Motivation.
- Rao, S., & Ahluwalia, N. (2022). Customer Segmentation Practices in Indian NBFCs and Loan Distribution Companies.
- Sen, A., & Bhattacharyya, D. (2024). Behavioural Segmentation and Financial-Product Distribution

Table 1. Respondent and Shopping Profile

Variable	Leading category	Percentage
Age	18–25 years	42.0%
Gender	Male	67.2%
Occupation	Student	40.3%
Monthly income	Below ₹20,000	45.8%
Shopping frequency	Weekly	39.8%

Shopping channel	Both online and offline	51.7%
Monthly expenditure	₹2,000–₹5,000	39.5%
Product category	Groceries	37.0%
Main purchase reason	Need	28.6%

Table 2. Behavioural Customer Segments

Segment	Respondents	Share	Key behavioral profile
Disengaged / Low-Involvement Shoppers	16	13.4%	Low across behavioral dimensions
Price-Sensitive Deal Seekers	25	21.0%	High price sensitivity; lower brand loyalty
Impulsive Brand-Conscious Shoppers	42	35.3%	Highest impulsiveness; above-average loyalty and deliberation
Value-Driven Loyal Customers	36	30.3%	High price sensitivity, loyalty, deliberation and service satisfaction

Table 3. K-Means Cluster Centroids

Segment	Price Sensitivity	Brand Loyalty	Deliberation	Impulsiveness	Service Satisfaction
Disengaged / Low-Involvement	2.75	2.86	2.44	2.25	3.27
Price-Sensitive Deal Seekers	4.13	3.11	—	—	—
Impulsive Brand-Conscious	—	3.93	3.80	4.21	—
Value-Driven Loyal	4.35	4.18	3.87	3.94	4.30