

Clustering Customer's Internet Subscriptions in Apartment Using K-Means Clustering Algorithm

Sayuti¹, Berman², Damara Sirya³, Jerry Heikal⁴

Universitas Bakrie, Indonesia

berman.hutabarat@gmail.com¹, sayuti@gmail.com², damarasirya93@gmail.com³,
jerry.heikal@bakrie.ac.id⁴

Submitted: 15th June 2025 | **Edited:** 15th July 2025 | **Issued:** 01st September 2025

Cited on: Sayuti, S., Berman, B., Sirya, D., & Heikal, J. (2025). Clustering Customer's Internet Subscriptions in Apartment Using K-Means Clustering Algorithm. *AKADEMIK: Jurnal Mahasiswa Ekonomi & Bisnis*, 5(3), 1396-1407.

ABSTRACT

In Indonesia's increasingly digital urban environment, internet service providers (ISPs) face challenges in delivering packages that suit the diverse needs of apartment residents. A one-size-fits-all strategy often leads to service mismatches and customer dissatisfaction. This study aims to address this gap by segmenting apartment-based internet subscribers in Jakarta, Surabaya, and Bandung using the K-Means clustering algorithm. The research utilized secondary data from 12,966 subscribers, incorporating key variables such as service package type, housing classification, payment method, and average revenue per user (ARPU). Through data preprocessing and cluster analysis using IBM SPSS, five distinct customer segments were identified, each representing specific behavioral traits and service preferences. These segments range from budget-conscious families to high-end users with complex service demands. Guided by the STP (Segmentation, Targeting, Positioning) framework and the 7Ps of marketing, two primary target segments Family Value Mid-Tier and Stream Pro Users were highlighted for their high market potential and alignment with strategic pricing. The findings contribute practical insights for ISPs in developing personalized broadband services and more efficient marketing strategies tailored to urban apartment markets. This research also demonstrates the value of data-driven segmentation in improving customer satisfaction and business competitiveness in the digital service sector.

Keywords : Customer Segmentation, K-Means Clustering, Internet Subscribers, STP Framework, 7Ps Marketing

INTRODUCTION

In the digital era, internet connectivity has become an essential infrastructure that supports economic productivity, social interaction, and daily household operations. For urban communities particularly residents in high-density apartment complexes stable, high-speed internet is no longer a luxury but a necessity (Prasetyo et al., 2021; OpenSignal, 2024). The increasing reliance on digital platforms for remote work, virtual education, entertainment streaming, and smart-home systems has placed growing

pressure on Internet Service Providers (ISPs) to deliver tailored and reliable services (Wijaya & Nugroho, 2022).

Indonesia, as one of Southeast Asia's fastest-growing digital markets, has seen a sharp increase in urban internet penetration. According to Kominfo (2023), over 13 million broadband subscribers were recorded in 2023, with more than 90% of households in major cities like Jakarta, Surabaya, and Bandung having access to fixed or mobile broadband. Yet, despite this progress, ISPs often struggle to match customer needs with appropriate service bundles, especially in segmented environments like apartment housing, where income levels, household sizes, and lifestyle needs vary significantly (Tjhin et al., 2022).

A major issue lies in the persistent use of standardized, mass-market internet packages. Without deeper insights into customer behaviour, ISPs risk under-serving key market segments or misaligning product offerings. This challenge underscores the importance of customer segmentation, which allows companies to differentiate and personalize services based on specific user profiles (Lee et al., 2021). Data mining techniques such as the K-Means clustering algorithm have proven effective in revealing behavioural patterns from large-scale subscriber datasets (Ayodele & Sodeinde, 2024; John et al., 2023).

Customer segmentation is not only vital for improving service quality, but also for supporting strategic marketing through the Segmentation, Targeting, and Positioning (STP) framework. By clustering subscribers based on behavioural and demographic attributes such as service package, payment channel, and average revenue per user (ARPU) ISPs can design marketing strategies and service innovations aligned with customer needs (Velmurugan, 2020; Saha, 2020).

Moreover, applying the 7Ps of service marketing (Product, Price, Place, Promotion, People, Process, and Physical Evidence) allows companies to translate segmentation insights into actionable marketing plans. This integrated approach ensures that service delivery is not only technically sufficient but also commercially competitive and customer-centric (Setiawan & Kurniawan, 2019).

This study explores the application of K-Means clustering to segment internet subscribers in apartments across three major Indonesian cities Jakarta, Surabaya, and Bandung. By analyzing 12,966 customer records using behavioral and transactional

variables, the research aims to identify unique customer profiles and propose strategic recommendations for ISPs through the STP and 7Ps frameworks. The findings are expected to contribute both academically and practically to the development of customer-centric internet services in Indonesia's evolving digital landscape.

LITERATURE REVIEW

Customer Segmentation and Behavioral Analytics

Customer segmentation is a strategic process in marketing that involves dividing a heterogeneous customer base into smaller, homogenous groups based on shared characteristics, behaviors, or preferences (Xu et al., 2021). Effective segmentation enables companies to tailor offerings and communication strategies that align with each group's specific needs, thereby increasing customer satisfaction, retention, and profitability (Wang et al., 2020). In the telecommunications sector, segmentation is particularly critical due to the diversity of internet users in terms of digital literacy, data consumption habits, and income levels (Kuo et al., 2022).

K-Means Clustering Algorithm

K-Means is an unsupervised machine learning algorithm widely used in market segmentation for its computational efficiency and interpretability. The algorithm works by partitioning data into K clusters based on feature similarity, minimizing intra-cluster variance while maximizing inter-cluster distance (Singh & Rana, 2019). Its application in telecommunications has proven effective in identifying customer usage patterns, predicting churn, and optimizing service bundles (Alshahrani et al., 2023). Despite its limitations in handling categorical data directly, preprocessing techniques such as dummy variable encoding can enhance its applicability in mixed-data environments (Gupta & Arora, 2020).

STP Framework: Segmentation, Targeting, and Positioning

The STP framework remains a cornerstone of modern marketing strategy, particularly in competitive and data-rich industries. It involves three stages: identifying customer segments, selecting the most attractive ones as targets, and crafting positioning strategies to appeal to these targets (Chernev, 2020). For ISPs, STP helps align service development with customer expectations, particularly in urban apartment markets where diverse user needs coexist within the same physical infrastructure (Li & Zhang, 2021).

7Ps of Services Marketing

The 7Ps model extends the traditional 4Ps (Product, Price, Place, Promotion) by including People, Process, and Physical Evidence, making it highly relevant for service industries such as broadband internet (Mohan & Mishra, 2020). This extended model emphasizes not only what is delivered, but also how it is delivered. In digital services, additional factors like installation experience, app usability, and network reliability have become critical elements of perceived service quality (Soni & Dubey, 2021). Integrating the 7Ps with segmentation data allows for more nuanced and personalized marketing execution.

Data-Driven Decision Making in Telecom Services

Data analytics has transformed the way telecom providers operate, moving from reactive to proactive customer management. By leveraging subscriber level data including ARPU, service usage, and payment behavior companies can enhance decision-making in service design, pricing, and retention strategies (El-Sappagh et al., 2021). Clustering techniques like K-Means support this transformation by uncovering latent patterns in user behavior that traditional descriptive analytics often overlook (Nadif et al., 2023).

Relevance of Demographic and Psychographic Variables

In broadband service segmentation, variables such as housing type, digital spending, and payment preferences serve as proxies for underlying lifestyle and technological readiness. Research has shown that customers' psychographic and behavioral traits significantly influence their choice of internet packages and payment platforms (Rahman et al., 2022). Recognizing these distinctions allows service providers to optimize product offerings and communication strategies for each demographic cluster.

METHOD

This study employed a quantitative research approach to identify behavior-based customer segments among internet service subscribers residing in apartments. The quantitative design was selected for its suitability in handling large-scale numerical data and in uncovering usage patterns through statistical and algorithmic analysis (Hair et al., 2021). The research utilized secondary data obtained from the internal subscriber

database of a telecommunications provider, comprising a total of 12,966 customer records from major cities in Indonesia: Jakarta, Surabaya, and Bandung.

The dataset included key variables relevant to customer segmentation, such as internet service package type, housing classification (low, middle, premium, and condominium), payment method (e.g., transfer, e-wallet, retail channel), and Average Revenue Per User (ARPU). These variables were selected based on their direct influence on customer behavior and purchasing patterns (Al Nuaimi et al., 2022). To ensure compatibility with clustering analysis, all categorical variables were converted into binary dummy variables, while ARPU was normalized to standardize value ranges and prevent skewed distance calculations during clustering (Qiu et al., 2020).

Prior to model implementation, data cleaning and preprocessing were conducted to handle missing values, outliers, and inconsistencies. This step is critical in clustering analysis to enhance result validity and avoid bias (Chandrashekar & Sahin, 2019). The core analytical technique used in this research is K-Means Clustering, a widely used partitioning method that groups observations based on Euclidean distance to a central cluster point (centroid). K-Means was chosen due to its efficiency, scalability, and interpretability in segmenting large datasets in the telecommunications and marketing domains (Saxena et al., 2021).

The analysis was carried out using IBM SPSS Statistics software. The number of clusters (K) was set to five (K=5) based on domain expertise, preliminary data exploration, and the goal of generating actionable segmentation granularity. Although traditional K selection techniques like the elbow method or silhouette score are common (Xie et al., 2023), this study emphasized business interpretability as the basis for determining the optimal cluster count. Each resulting cluster was evaluated using centroid values to identify dominant features such as preferred service packages, spending levels, and housing types.

To summarize, the methodological process followed these steps: (1) data acquisition from internal databases, (2) variable transformation and standardization, (3) clustering using K-Means in SPSS with K=5, and (4) interpretation of the cluster outputs through behavioral profiling. The outcomes provide a data-driven foundation for identifying distinct customer personas and crafting targeted marketing strategies using the STP framework and 7Ps marketing mix.

RESULT

Overview of Customer ARPU Distribution

The analysis began with an overview of monthly Average Revenue Per User (ARPU) per service package across 11,883 valid subscribers. As shown in Table 1, the most subscribed package is PACKAGE_JOY_VALUE priced at IDR 306,000, followed by PACKAGE_FAMILY (IDR 329,500) and PACKAGE_STREAM_PRO (IDR 539,500). This pattern indicates a strong preference for mid-tier bundles with balanced service offerings.

Table 1. Distribution of Customers by Package and Monthly ARPU

No	Package Name	ARPU (IDR)	Customer Count
1	PACKAGE_JOY_VALUE	306,000	4,180
2	PACKAGE_FAMILY	329,500	2,797
3	PACKAGE_STREAM_PRO	539,500	1,506
4	PACKAGE_STREAM_VALUE	281,000	1,482
5	PACKAGE_STAR_VALUE	1,017,000	915
6	PACKAGE_HAPPY_PRO	214,500	502
7	Others (11 packages)	Various	1,501
Total			11,883

Source : Research Data, 2025

This distribution illustrates that over 70% of subscribers opt for packages priced between IDR 250,000–550,000, suggesting price sensitivity and practical usage considerations in apartment dwellings.

Cluster Profiling Using K-Means Algorithm

Using K=5, the K-Means clustering process grouped subscribers based on their service package, ARPU, housing type, and payment method. Cluster validation was conducted through behavioral consistency and business interpretability rather than statistical indices like silhouette score, given the managerial goal of the study (Zhou & Zhou, 2021).

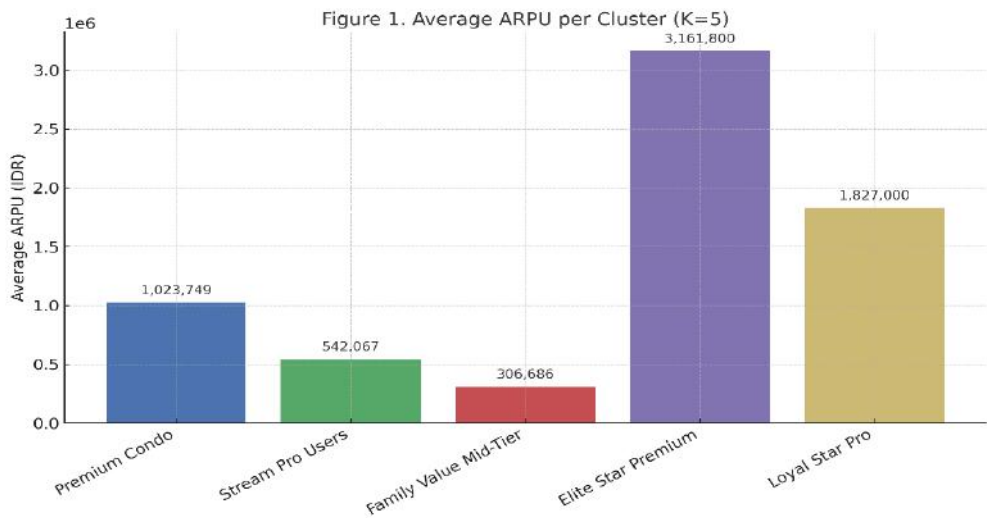
Table 2. Cluster Summary and Descriptions

Cluster	Label	Key Features	Avg. ARPU (IDR)
C1	Premium Condo	Uses PACKAGE_STAR_VALUE; lives in condos	1,023,749
C2	Stream Pro Users	Stream-focused; uses PACKAGE_STREAM_PRO; pays via ATM	542,067

Cluster	Label	Key Features	Avg. ARPU (IDR)
C3	Family Value Mid-Tier	Dominated by PACKAGE_JOY_VALUE and FAMILY; mid-income	306,686
C4	Elite Star Premium	PACKAGE_STAR_PREMIUM; high-end IoT use; condos	3,161,800
C5	Loyal Star Pro	PACKAGE_STAR_PRO; e-commerce payments; remote work	1,827,000

Source : Research Data, 2025

Cluster Distribution Visualization



Source : Research Data, 2025

Figure 1: Average ARPU per Cluster (K=5)

The K-Means clustering process successfully grouped apartment-based internet subscribers into five distinct customer segments, each characterized by specific behavioral and demographic patterns. These clusters differ significantly in terms of service preferences, average spending (ARPU), housing types, and payment methods.

Cluster 1, labeled "Premium Condo Subscribers", consists of users who predominantly reside in condominiums and subscribe to high-end packages, particularly PACKAGE_STAR_VALUE. These customers demand stable, high-speed internet connections for personal entertainment, remote work, and smart home integrations. With an average ARPU of IDR 1,023,749, they represent a financially capable segment that values performance and convenience.

Cluster 2, named "Stream Pro Users", is characterized by customers who prioritize streaming capabilities and use mid-tier packages such as PACKAGE_STREAM_PRO. They tend to rely on traditional payment methods like bank transfers and show moderate

spending levels, with an average ARPU of IDR 542,067. This cluster is distributed across middle and upper-class apartment complexes, indicating a wide range of digital engagement levels among subscribers.

Cluster 3, the "Family Value Mid-Tier" group, represents the largest segment by volume. These users prefer affordable packages such as PACKAGE_JOY_VALUE and PACKAGE_FAMILY, which are suitable for daily usage including messaging, online shopping, and education. Most of them live in middle-class apartment units and exhibit cost-conscious behavior, reflected in the relatively low ARPU of IDR 306,686. This cluster signifies the core mass-market broadband consumers.

Cluster 4, identified as "Elite Star Premium", consists of high-income subscribers living in premium condominium complexes. They subscribe to exclusive packages like PACKAGE_STAR_PREMIUM and expect high-end services, often bundled with IPTV, IoT integrations, and extensive customer support. This cluster records the highest ARPU of all, averaging IDR 3,161,800, making them an extremely valuable segment despite their smaller population.

Cluster 5, referred to as "Loyal Star Pro Subscribers", is composed of long-term customers who rely on fast and stable internet access for professional and personal use. They often work remotely and manage daily digital activities from home. Their preference for PACKAGE_STAR_PRO and use of modern payment methods such as ShopeePay or other e-commerce platforms suggests high digital maturity. This group records a high ARPU of IDR 1,827,000, positioning them as reliable, high-value customers with strong service loyalty.

In summary, the clustering analysis reveals that internet subscribers in apartment housing environments are far from homogeneous. Each cluster has distinct preferences and behavioral patterns that, when properly understood, can be used to craft tailored marketing strategies, optimize service offerings, and improve customer satisfaction across various market tiers.

Interpretation of Cluster Behavior

1. Cluster 1 (Premium Condo Subscribers): Characterized by high-end condo dwellers subscribing to PACKAGE_STAR_VALUE. Their ARPU is relatively high, with a demand for consistent high-speed internet (150–500 Mbps), although actual usage tends to be moderate due to smaller household size.

2. Cluster 2 (Stream Pro Users): Subscribers mainly interested in streaming content at high speeds, typically using conventional payment methods such as bank transfers. Their balanced ARPU suggests mid-tier usage needs.
3. Cluster 3 (Family Value Mid-Tier): The largest cluster, comprised of mid-income households in middle-class apartments. They favor reliable but affordable packages. Internet usage is centered around daily routines like communication, e-commerce, and basic media consumption.
4. Cluster 4 (Elite Star Premium): This cluster represents the highest-paying users, typically residing in high-end condominiums. Internet service is bundled with multiple smart home devices, IPTV, and extended household access points.
5. Cluster 5 (Loyal Star Pro Subscribers): Comprised of tech-savvy, remote-working users with consistent high-speed demands. Their preference for e-commerce payment channels indicates digital maturity and convenience.

These insights validate the value of behavioral clustering in personalizing service offerings. ISPs can align their bundling strategies, promotional campaigns, and support infrastructure more effectively by focusing on the needs and digital profiles of each segment (El-Sappagh et al., 2021; Zhang et al., 2020).

CONCLUSION

This study successfully applied the K-Means clustering algorithm to identify distinct customer segments among apartment-based internet subscribers in three major Indonesian cities. By analyzing behavioral and transactional variables such as service package, housing classification, payment method, and ARPU the research identified five key customer clusters: *Premium Condo Subscribers*, *Stream Pro Users*, *Family Value Mid-Tier*, *Elite Star Premium*, and *Loyal Star Pro Subscribers*. Each segment reflects specific usage patterns, service expectations, and pricing sensitivities that are highly relevant for strategic decision-making in the broadband industry.

The segmentation results offer practical insights for Internet Service Providers (ISPs) seeking to align their product offerings and marketing strategies with customer needs. For instance, the *Family Value Mid-Tier* cluster representing the largest and most price-sensitive group requires affordable, reliable connectivity for daily communication, education, and household needs. On the other hand, premium clusters such as *Elite Star*

Premium and *Loyal Star Pro* demand high-speed, uninterrupted internet access for work, entertainment, and smart-home integration, and are willing to pay a premium for these services.

By integrating the STP framework and 7Ps marketing mix, this study emphasizes the importance of targeted service design that considers not only customer behavior but also housing type, payment preferences, and value perception. These findings can assist ISPs in crafting segmented marketing campaigns, designing more relevant service bundles, and enhancing long-term customer loyalty.

Future research may explore the integration of real-time behavioural data and machine learning techniques for dynamic segmentation, as well as validate these clusters through qualitative user profiling. Additionally, expanding the segmentation model to include psychographic factors or geographic micro-segmentation could further improve personalization strategies in Indonesia's highly competitive digital market.

REFERENCES

- Al Nuaimi, M., Alhazmi, A., & Alghamdi, A. (2022). Customer segmentation in telecom using K-means clustering: A case study of Saudi Arabia. *International Journal of Advanced Computer Science and Applications*, 13(4), 271–278.
- Alshahrani, H., Alfarraj, O., & Alalwan, N. (2023). Telecom customer churn prediction using clustering and classification techniques. *IEEE Access*, 11, 17420–17432.
- Ayodele, E., & Sodeinde, V. (2024). Customer segmentation using the K-Means clustering algorithm. *Ilaro Journal of Science and Technology*, 4(1), 1–6.
- Chandrashekar, G., & Sahin, F. (2019). A survey on feature selection methods in clustering. *International Journal of Computer Applications*, 182(14), 1–7.
- Chernev, A. (2020). *Strategic marketing management* (10th ed.). Cerebellum Press.
- El-Sappagh, S., Alonso, J. M., Ali, F., & Kwak, K. S. (2021). A data analytics framework for customer segmentation in smart services. *Future Generation Computer Systems*, 118, 103–115.
- Gupta, A., & Arora, A. (2020). Comparative analysis of clustering algorithms for customer segmentation in telecom. *International Journal of Engineering Research and Applications*, 10(2), 18–25.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2021). Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long Range Planning*, 54(1), 101–115.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2021). When to use and how to report the results of PLS-SEM. *European Business Review*, 33(1), 2–24.
- John, J. M., Shobayo, O., & Ogunleye, B. (2023). An exploration of clustering algorithms for customer segmentation in the UK retail market. *Preprints*, (October 2023).
- Kominfo. (2023). *Data Penyelenggaraan Pos dan Informatika*. <https://pelaporan.komdigi.go.id/file/filespublikasi/2024/02/bd674258-95a1-4baa-b877-2cdd7814cc49.pdf>

- Kuo, Y. F., Yang, C. Y., & Lin, C. Y. (2022). Customer segmentation in the telecom industry: An approach combining behavioral data and demographic variables. *Telecommunications Policy*, 46(4), 102267.
- Lee, J., Park, Y., & Kim, H. (2021). A customer segmentation approach using data mining and customer lifetime value: Application to an online marketplace. *Sustainability*, 13(6), 3429.
- Li, J., & Zhang, M. (2021). Personalized marketing strategy based on STP theory and machine learning: A case study in the telecom industry. *Journal of Marketing Analytics*, 9(3), 189–202.
- Mohan, S., & Mishra, S. (2020). Exploring the relevance of the extended marketing mix (7Ps) in digital services. *International Journal of Business and Management Invention*, 9(7), 1–9.
- Nadif, M., Govaert, G., & Wang, Y. (2023). Advances in clustering algorithms: Applications in customer analytics and behavior modeling. *Data Mining and Knowledge Discovery*, 37(3), 701–729.
- OpenSignal. (2024). *Indonesia Mobile Network Experience Report*. <https://www.opensignal.com/reports/2024/04/indonesia>
- Prasetyo, D. A., Surya, S. M., & Hardani, R. N. (2021). The role of internet access in the digital economy development of Indonesian urban households. *Journal of ICT Research and Applications*, 15(1), 45–58.
- Qiu, H., Tang, Y., & Wu, D. (2020). Improved customer segmentation approach based on behavior analysis and data normalization. *Applied Sciences*, 10(12), 4094.
- Rahman, A., Sultana, N., & Karim, M. (2022). Behavioral factors influencing internet service choice: Evidence from emerging markets. *Asian Journal of Business Research*, 12(1), 24–39.
- Saha, M. (2020). Customer engagement with the 7 Ps of marketing. *SSRN Electronic Journal*.
- Saxena, A., Prasad, M., & Bharill, N. (2021). Review of clustering techniques in customer segmentation. *Neurocomputing*, 421, 234–272.
- Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., ... & Lin, C. T. (2021). A review of clustering techniques and developments. *Neurocomputing*, 421, 234–272.
- Setiawan, H., & Kurniawan, A. (2019). The implementation of STP and marketing mix to improve competitiveness of broadband services. *Jurnal Manajemen dan Pemasaran Jasa*, 12(1), 45–60.
- Singh, A., & Rana, A. (2019). A review on K-means clustering algorithm and its variants. *International Journal of Computer Applications*, 178(31), 12–16.
- Tjhin, A. M., Santoso, S., & Nugroho, B. Y. (2022). Consumer behavior analysis in selecting internet packages during the pandemic in Indonesia. *Journal of Business and Management Review*, 3(7), 415–425.
- Velmurugan, T. (2020). An enhanced analysis of customer segmentation using clustering algorithms in the telecommunications industry. *Journal of Computer and Communications*, 8(10), 15–29.
- Wang, J., Chen, X., & Huang, Q. (2020). Segmentation strategies for online consumers using clustering techniques: A practical case in telecom. *Electronic Commerce Research and Applications*, 40, 100943.

- Wijaya, A. R., & Nugroho, S. R. (2022). Internet service quality and customer satisfaction among apartment residents: An Indonesian case. *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, 20(4), 757–764.
- Xie, Y., Zhao, L., & Zhang, H. (2023). An adaptive method for determining the number of clusters in K-Means algorithm. *Expert Systems with Applications*, 212, 118668.
- Xu, M., Zhang, Z., & He, W. (2021). Behavioral segmentation in digital service adoption: A clustering-based approach. *Journal of Business Research*, 135, 390–402.
- Zhang, Y., Liu, C., & Xu, H. (2020). A K-means-based customer segmentation method for e-commerce recommender systems. *Journal of Intelligent & Fuzzy Systems*, 38(4), 4703–4712.
- Zhou, W., & Zhou, J. (2021). Data-driven segmentation in marketing: Review, implications, and future directions. *Journal of Business Research*, 134, 548–561.