Enhancing Bank Customer Segmentation in Electronic Service Adoption: An Integrated Analytic Hierarchy Process - Agglomerative Hierarchical Clustering Approach within the Recency, Frequency, and Monetary value Framework

Author 1: *Arash Moheimani* [1]
Author 2: *Seyed Mohammad Hassan Hosseini* [2]
Author 3: *Alessio Ishizaka* [3]

Highlights

- Dual AHP usage improves the accuracy of customer clustering and prioritization.
- First application of RFM for electronic banking adoption, offering a data-driven approach with transactional data, reducing reliance on subjective survey methods.
- Cluster demographics used to provide strategic recommendations for enhancing e-service engagement.

Abstract

Clustering techniques, invaluable for classifying data based on feature similarities, often face challenges with hierarchical structures and varying submetrics. This study integrates the Analytic Hierarchy Process (AHP) and Agglomerative Hierarchical Clustering (AHC) to address these limitations, prioritizing criteria and enhancing customer segmentation in electronic banking service adoption through internet and mobile bank portals. We introduce a unique framework based on the Recency, Frequency, Monetary value (RFM) model to assess customer engagement with concrete data. Data were collected from an Iranian bank in the first half of 2022 to test our proposed method. The research findings, validated by the Silhouette Coefficient (SC), reveal a four-cluster solution representing varying levels of customer engagement with e-services, distributed as 32.5%, 55.2%, 12.1%, and 0.2%, from the least to the most engaged clusters. A subsequent AHP analysis prioritized these clusters, informing targeted marketing strategies. Finally, our study offers practical, demographic-based recommendations for bank directors, aimed at enhancing engagement with electronic banking services and improving operational efficiency.

Keywords: RFM, agglomerative hierarchical clustering (AHC), analytic hierarchy process (AHP), decision support

1. Introduction

In recent years, clustering has become a widely used machine-learning technique for grouping similar data points based on specific features. This approach is utilized across many domains, including market segmentation, customer clustering, and electronic banking (Hosseini et al., 2022). However, traditional clustering methods, such as Agglomerative Hierarchical Clustering (AHC), often rely solely on measuring similarities based on distances between objects' characteristics, without accounting for hierarchical structures or priority relations between features, thereby limiting their ability to fully capture complex patterns. Similarly, the results of clustering lack preferential order, meaning the clusters are formed solely based on data characteristics without any inherent ranking or prioritization. In the banking industry, this oversight is particularly critical, as understanding customer behaviors, preferences, and demographics is essential for effective segmentation and targeted marketing strategies (Hosseini et al., 2022). To address these limitations, this paper proposes the integration of the Analytic Hierarchy Process (AHP) with AHC to enhance customer segmentation in electronic service adoption. Additionally, the Recency, Frequency, and Monetary (RFM) model serves as the conceptual framework for analyzing customer engagement with digital banking portals. By combining AHP with AHC and leveraging RFM data, this study introduces a

[1] Arash Moheimani. Mr, PhD researcher, Department of Information Systems, Supply Chain & Decision Support, NEOMA Business School, 1 rue du Maréchal Juin - BP 215, Mont-Saint-Aignan Cedex, 76825, France, arash.moheimani@neoma-bs.fr (ORCID:0000-0002-8257-2496)
[2] Seyed Mohammad Hassan Hosseini, Mr, Dr, Associate Professor, Department of Industrial Engineering and Management, Shahrood University of Technology, Shahrood, Iran, sh.hosseini51@gmail.com (ORCID: 0000-0002-6164-3179)
[3] Alessio Ishizaka, Mr, Distinguished Professor, Department of Information Systems, Supply Chain & Decision Support, NEOMA Business School, 1 rue du Maréchal Juin - BP 215, Mont-Saint-Aignan Cedex, 76825, France, Alessio.ishizaka@neoma-bs.fr (ORCID:0000-0002-2531-292X)

more comprehensive clustering method that prioritizes customer attributes and better reflects the nuances of customer behaviors.

2. Literature Review

2.1. Electronic Banking Adoption

Research on electronic banking adoption has highlighted several key factors influencing its uptake, such as time and cost savings, freedom of location, and customer satisfaction. Despite these advantages, banks, especially in developing countries, face significant challenges in encouraging the use of digital banking platforms (Sharma et al., 2020). Prior studies, primarily using Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) frameworks, have focused on customer perceptions, intentions, and subjective data (Kwateng et al., 2018). However, there has been a lack of emphasis on transactional behavior, which is a key aspect in understanding actual service adoption.

2.2. RFM-based Customer Segmentation

The RFM model has been widely used for customer segmentation in marketing, particularly for analyzing transactional behaviors like the recency, frequency, and monetary value of customer interactions (Khajvand & Tarokh, 2011). In the banking sector, studies have shown that RFM analysis provides a concrete and tangible measure of customer engagement. Despite its effectiveness, the RFM model is underutilized in electronic banking adoption research, where subjective models currently dominate. This study aims to extend the application of the RFM model to the electronic banking context, offering more objective and data driven segmentation of customer engagement based on their actual transactions.

2.3. Background of AHP-AHC Method

The integration of AHP with AHC has been explored in various fields, offering more structured clustering by accounting for the hierarchical nature of features. AHP assigns weights to features before clustering, addressing the gap in conventional methods that often overlook the priority relations between customer attributes. While previous studies have demonstrated the effectiveness of this integration (Hillerman et al., 2017), its application in the banking sector remains novel, especially for customer segmentation where transaction attributes vary in importance. Our study extends this method by reapplying AHP after clustering to prioritize the resulting clusters, enhancing its effectiveness in electronic banking adoption analysis.

3. Objectives

This study aims to develop a robust decision-making model that integrates AHP with AHC to enhance customer segmentation in electronic banking, using customer engagement levels in internet and mobile banking portals as indicators of electronic banking adoption. The specific objectives of the study are:

Objective 1. To propose a novel framework that integrates AHP and AHC for customer segmentation based on the RFM model and apply this framework to a real-world dataset from an Iranian bank.

Objective 2. To provide actionable managerial insights based on demographic analysis and prioritized clusters, assisting bank managers in designing targeted strategies for customer engagement and retention.

4. Methodology

This study's methodological process is illustrated in Fig. 1 and comprises data collection and preprocessing, RFM reorganization of data, weighting RFM elements by AHP, clustering by AHC, cluster analysis, cluster prioritization, demographic analysis, and providing managerial recommendations using the outcomes of the analyses.

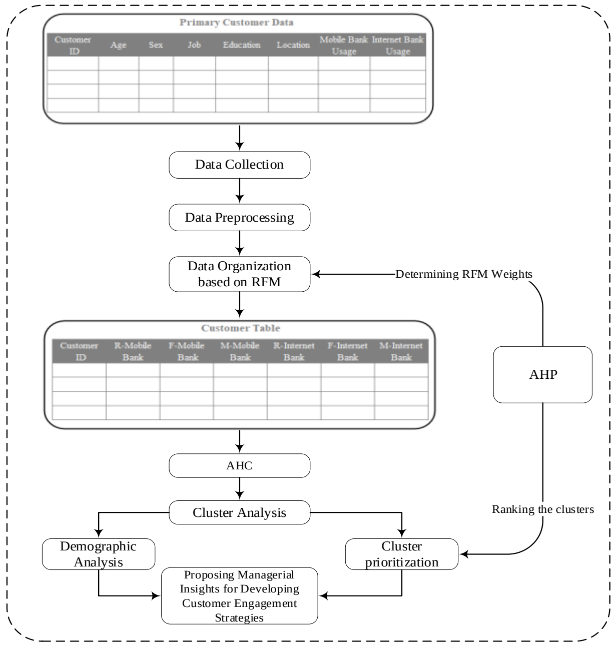


Fig. 1. The methodological procedure of the present study

4.1. An RFM-based AHP-AHC Approach

Data was collected from the bank's records through established connections, covering multiple branches. The preprocessing phase involved two key steps:

1. Data Integration: Consolidating records from various branches into a unified dataset.
2. Data Cleaning: Addressing missing values, removing duplicates, and resolving inconsistencies.

The cleaned data was structured using the RFM model. AHP was applied to assign weights to the RFM variables and the electronic banking portals (Internet and Mobile Banking). The decision-making hierarchy, representing the RFM variables and electronic portals, is shown in Fig. 2.

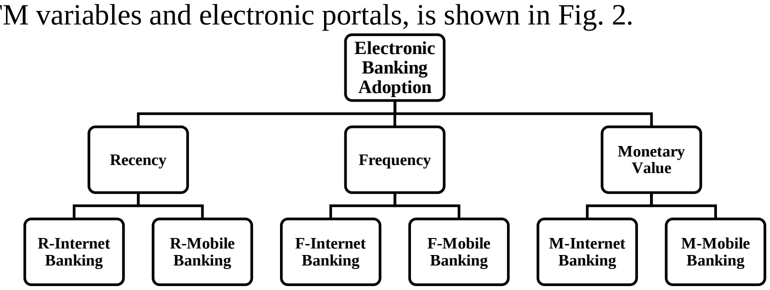


Fig. 2. The hierarchy of the RFM model and the electronic banking portals portrayed by AHP

4.2. Weights Assignment and Aggregation

The AHP process involved ten experts: four senior bank managers and six experienced banking agents. Each expert provided pairwise comparisons, and their judgments were aggregated using the geometric mean method to ensure a fair representation of all opinions. Table 1 summarizes the assigned weights for the main RFM variables and their respective sub-features. To ensure the consistency of expert judgments, we calculated the Consistency Ratio (CR) for the AHP matrix. A CR value of 0.003, well below the acceptable threshold of 0.10, confirmed the consistency and reliability of the expert evaluations.

Table 1. The weights of the features and sub-features used in the clustering

Features	*Recency*		*Frequency*		*Monetary value*	
Weights	0.221		0.527		0.252	
Sub-features	*R-Mobile Bank*	*R-Internet Bank*	*F-Mobile Bank*	*F- Internet Bank*	*M-Mobile Bank*	*M-Internet Bank*

Weights	0.75	0.25	0.6	0.4	0.27	0.73

4.3. AHC

AHC is a bottom-up clustering approach that merges data points or clusters based on similarity, forming a hierarchical structure. At each step, the algorithm calculates the distance between clusters, progressively merging the two closest ones to build a hierarchy. The choice of metric and linkage methods significantly influences the clustering results.

- **Metric Methods**: These define how distances between data points are measured. In this study, we used both the *Euclidean Distance* and *Manhattan Distance* to evaluate similarity.
- **Linkage Methods**: These determine how distances between clusters are calculated during the merging process. We employed several linkage methods, including:
    - Single Linkage: Merges clusters based on the smallest distance between any two points.
    - Complete Linkage: Merges clusters based on the largest distance between points.
    - Average Linkage: Merges based on the average distance between all pairs of points in the clusters.
    - Ward Linkage: Minimizes the increase in within-cluster variance during merging.

To determine the optimal number of clusters, we used both the Silhouette Coefficient (SC) and the dendrogram. The SC quantitatively evaluates cluster separation and compactness, with values closer to 1 indicating well-defined clusters. Simultaneously, the dendrogram provided a visual representation of the hierarchical clustering process, allowing for an intuitive assessment of the cluster structure. Together, these methods helped guide the selection of the most appropriate number of clusters.

5. Case Study

This study focuses on a private bank in Iran to analyze customer behavior and segment users based on electronic banking adoption. The dataset consists of demographic and usage data from **14,000 customers** with active accounts during the first six months of 2022. After preprocessing to handle incomplete records, the data was used to evaluate customer adoption of internet and mobile banking services. This large sample size allows the findings to be generalized to similar banks within the country.

6. Results

6.1. Cluster Analysis

The clustering analysis was conducted using both Manhattan and Euclidean distance metrics, which yielded identical results, indicating that the choice of metric had little impact on the final clustering. However, different linkage methods (Single, Average, Ward, Complete) provided varying cluster formations, emphasizing the importance of careful selection when determining the optimal number of clusters.

To determine the optimal number of clusters, we calculated the SC score for each method, from two to five clusters. As shown in Table 2, the Ward method suggested three clusters, while the Complete method pointed to four clusters. The Average and Single methods both indicated two clusters. Based on the geometric mean SC scores, four clusters had the highest average score, providing the best solution.

Table 2. Comparison of SC scores for different numbers of clusters using various linkage methods

Number of Clusters / Linkage methods	Ward	Complete	Average	Single	**Geomean Score**
2 clusters	0.809723727	0.620098437	**0.903757626**	**0.903757626**	0.800249
3 clusters	**0.870409387**	0.594449701	0.778862888	0.852049311	0.765493
4 clusters	0.750442177	**0.879855255**	0.847332232	0.822576792	**0.823644**
5 clusters	0.7431672	0.85885619	0.828125822	0.800298516	0.806471

During discussions with the bank managers, they acknowledged that two clusters would not offer sufficient insight into strategic decision-making. After reviewing all methods and results, they agreed that four clusters provide a more balanced and meaningful segmentation. Finally, a 3D visualization of the RFM variables (Fig. 4) revealed clear distinctions in customer engagement levels across the four clusters, from the least engaged (Cluster 1) to the most engaged (Cluster 4). This segmentation allows for more targeted decision-making and aligns with the bank's strategic priorities.

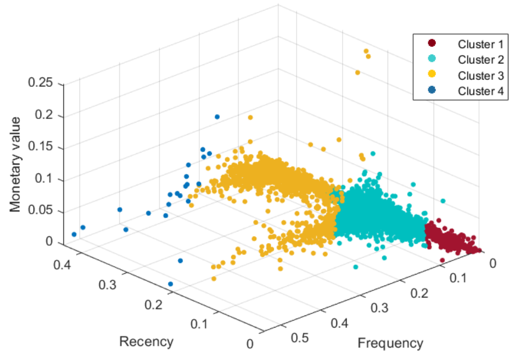


Fig. 4. Three-Dimensional illustration of the data points in terms of RFM variables

6.2. Demographic Analysis

This section provides a demographic analysis of the four customer clusters based on their engagement with electronic banking services. We examined five key demographic elements: occupation, gender, location, education level, and age, to understand customer profiles and provide insights for marketing strategies. Comprehensive statistical methods, including T-tests and ANOVA, were employed to validate the significance of observed differences across demographic dimensions

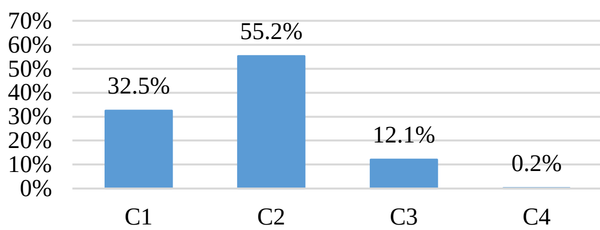
Fig. 5, with six distinct charts from (a) to (f), illustrates the demographic breakdown across clusters:

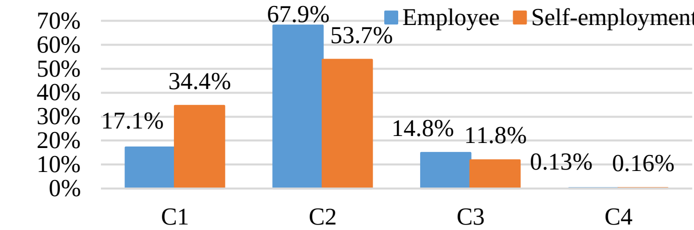
- **Cluster distribution (Fig. 5a)**: Most customers fall into C1 and C2, representing lower and moderate engagement levels, respectively, while C3 and C4, which show progressively higher interaction, contain significantly fewer customers. This pattern aligns with trends in developing countries, where growth in electronic service adoption remains an opportunity (Sharma et al., 2020).
- **Occupation (Fig. 5b)**: Employed individuals show higher engagement across C2 and C3, while self-employed customers show a higher percentage in C1. This may be because employees, restricted by office hours, often turn to electronic services out of necessity. In contrast, many self-employed individuals rely on in-person transactions, where reputation and personal negotiation play a crucial role.

Gender (Fig. 5c): Men are slightly more skewed toward higher engagement clusters, while women are more concentrated in lower engagement clusters. This difference reflects broader sociocultural factors and varying levels of digital access.

- **Location (Fig. 5d)**: Customers from larger cities are more engaged with electronic services, predominantly in C2, C3 and C4, while those from smaller cities cluster around C1. The urban-rural divide can be influenced by digital infrastructure and lifestyle differences.
- **Education (Fig. 5e)**:C1 has a notably high proportion of customers without a diploma, while C2 and C3 show a significant decrease in this group, with a sharp rise in customers holding higher education levels, suggesting that technological literacy plays an important role in electronic banking adoption. As for C4, its low representation of any educational background suggests that engagement in this cluster may be driven more by financial means than by formal education.
- **Age (Fig. 5f)**: The highest engagement is seen in C2 and C3 for adults aged 21 to 50, reflecting their comfort with technology. C4 has a consistent, though lower, presence across all age groups, indicating that other factors, such as professional needs, might drive their usage.

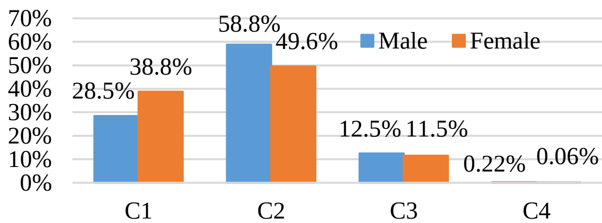
Across all demographic categories, C2 is the most represented cluster, followed by C1, with C3 and C4 showing higher engagement levels. These patterns highlight a general trend of moderate electronic banking adoption in Iran, with considerable opportunities for growth in the more engaged clusters.

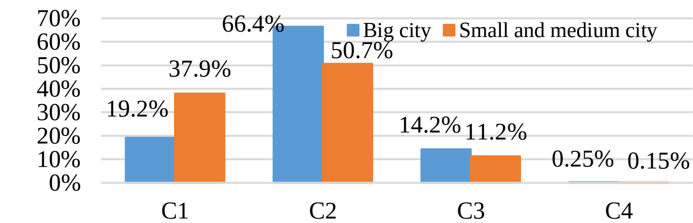




(a) Percentages of banking customers clustered from C1 to C4 regarding accepting electronic services

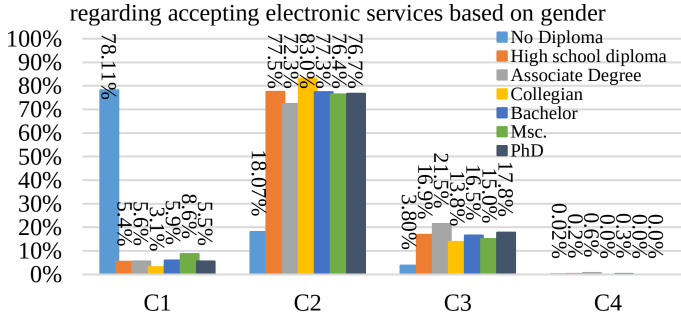
(b) Percentages of banking customers clustered from C1 to C4 regarding electronic service engagement based on occupation

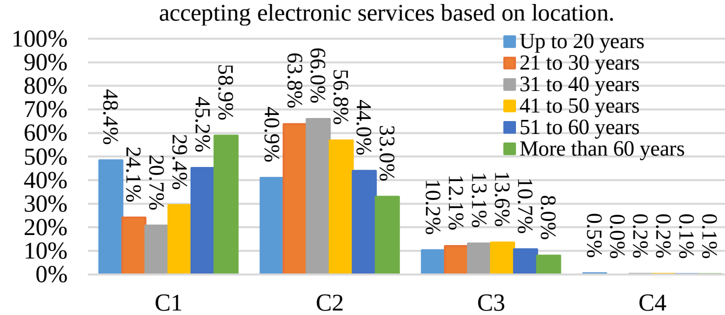




(c) Percentages of banking customers clustered as C1 to C4 regarding accepting electronic services based on gender

(d) Percentages of banking customers clustered from C1 to C4 regarding accepting electronic services based on location.





(e) Percentages of banking customers clustered from C1 to C4 regarding electronic service engagement based on education

(f) Percentages of banking customers clustered from C1 to C4 regarding electronic service engagement based on age

Fig. 5. Clusters' demographic details

6.3. Cluster Prioritization

Aiming to derive marketing prioritization, AHP was reused to rank the clusters. First, experts assigned names to the clusters to enhance comprehension. Subsequently, three new criteria were added to the RFM basis to help determine the most important clusters:

a) **Cluster Volume (CV)**: Reflects the size of the customer base.
b) **Average Balance (AB)**: Indicates the financial potential within each cluster.
c) **Average Non-Electronic Usage (ANEU)**: Captures reliance on non-electronic services, highlighting an opportunity to shift these customers toward digital banking.

Table 4 shows the weights of these new criteria, derived through AHP, with a CR value of 0.007, and Fig. 6 presents the AHP decision tree of the relationships between the criteria and the clusters.

Table 4. Weights of the features used for cluster ranking

Features	RFM	CV	AB	ANEU
Weights	0.252	0.302	0.223	0.223

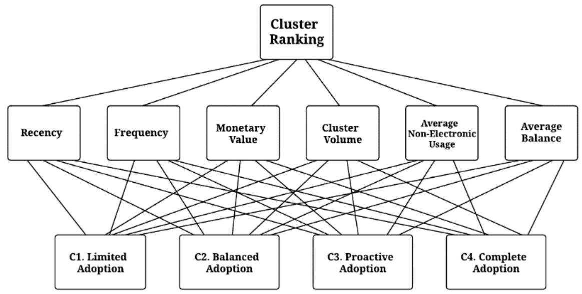


Fig. 6. Hierarchy of the criteria used to rank the clusters portrayed by AHP

Finally, the ranking was done using AHP, Cluster C3 ("Proactive Adoption") ranks first due to its high engagement potential, followed by Cluster C2 ("Balanced Adoption"), which has a large customer base. Cluster C4 ("Complete Adoption") ranks third as these customers are already highly engaged, requiring less marketing focus. Cluster C1 ("Limited Adoption") ranks last, indicating minimal engagement and lower strategic importance.

6.4. Managerial Insights

In this section, our attention shifts to the two top-ranked clusters, C3 and C2, which hold the most potential for banking growth strategies. Fig. 7 offers a deep demographic dive into these clusters, highlighting the participation of different demographic groups in C3 and C2. This detailed analysis aims to furnish bank strategists with granular managerial insights that can be pivotal in crafting tailored customer engagement, retention, and acquisition strategies. Fig. 7 is organized into six distinct charts from (a) to (f):

Overall distribution among all clusters (Fig. 7a): This pie chart visually encapsulates the customer distribution across the four clusters: C1 (32.5%), C2 (55.2%), C3 (12.1%), and C4 (0.2%), underscoring the prominence of clusters C2 and C3.

Occupation distribution in C2 and C3 (Fig. 7b): This bar chart delineates the employment landscape of clusters C2 and C3. Notably, in C2, 15.5% are employees, while 39.7% identify as self-employed. In C3, the self-employed group still maintains a considerable share with 8.4%, compared to employees at 3.7%.

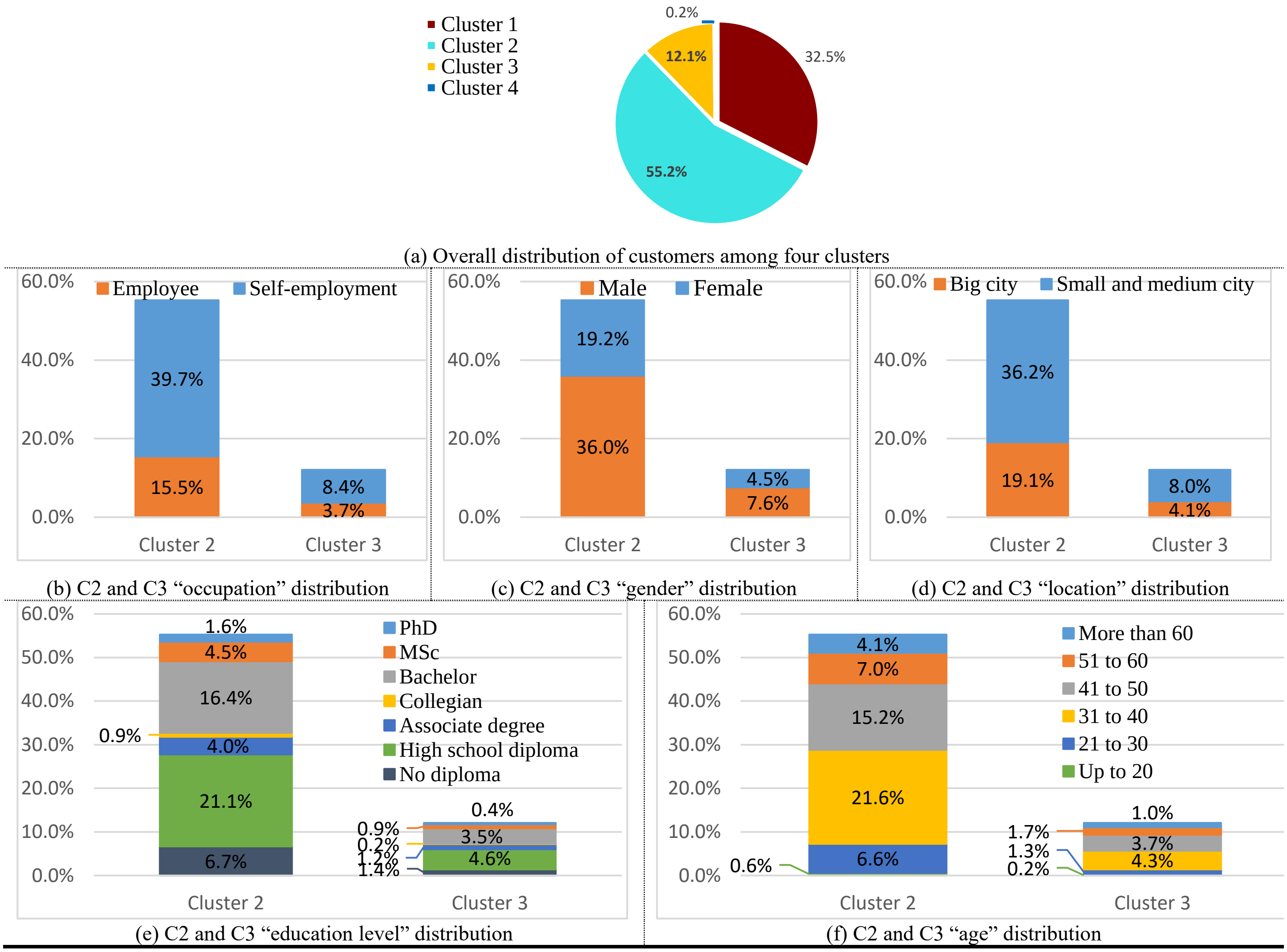
(Gender distribution in C2 and C3 (Fig. 7c): Gender dynamics in the two clusters reveal that males comprise 36% and 7.6% in C2 and C3, respectively. Females, on the other hand, constitute 19.2% of C2 and 4.5% of C3.

Location distribution in C2 and C3 (Fig. 7d): Examining urbanization levels, customers from larger cities form 19.1% of C2 and 4.1% of C3. Those from small and medium cities represent 36.2% of C2 and 8% of C3.

Education level distribution in C2 and C3 (Fig. 7e): The education bar chart shows the spread across seven levels, ranging from no diploma to PhDs. For instance, in C2, individuals with a high school diploma make up 21.1%, while in C3, they represent 4.6%.

Age distribution in C2 and C3 (Fig. 7f): The age demographics span six brackets. In C2, the age group of 31 to 40 has the highest representation at 21.6%. This trend continues somewhat in C3, with the same age group constituting 4.3%.

While the percentages across demographic groups within C2 and C3 vary, it is pivotal to recognize that all these groups hold potential revenue avenues. As the bank directs attention to these two clusters, these demographic breakdowns serve as foundational data in shaping targeted and effective marketing strategies.



(a) Overall distribution of customers among four clusters

(b) C2 and C3 "occupation" distribution (c) C2 and C3 "gender" distribution (d) C2 and C3 "location" distribution

(e) C2 and C3 "education level" distribution (f) C2 and C3 "age" distribution

Fig. 7. Demographic breakdown of customers in clusters C2 and C3

As can be seen in the demographic profiles from Fig. 7, our analysis explores potential strategies for clusters C3 and C2 based on each demographic element. This detailed segmentation aims to use these insights to create customized engagement plans for electronic banking users, guiding the bank's future marketing efforts. Cluster C3, with its engaged user base, receives advanced recommendations to further enhance relationships with proactive users. Meanwhile, strategies for C2 are foundational, catering to newcomers or those yet to fully use electronic banking. These strategic recommendations are provided in Table 5.

Table 5. Suggestions for enhancing customer engagement with electronic banking services

Elements	C3 – Proactive Adoption	C2 – Balanced Adoption
Occupation	Employees are highly engaged; consider loyalty programs with premium features. For self-employed individuals, enhance engagement through business tools and dedicated services.	Features that streamline banking, like automatic bill payments and personal finance tools, could benefit employees. Self-employed individuals may appreciate incentives and tools for business management, addressing their specific concerns.
Gender	Gender distribution is balanced. Enhance male engagement through investment options; for females, segment-specific marketing and financial education focusing on personal goals could deepen engagement.	Tailor marketing to women's interests, emphasizing financial planning and security to boost their banking activities.
Location	Urban customers could enjoy exclusive events and local business collaborations. In smaller cities, community-based initiatives could engage customers and encourage ambassadorship.	Big-city dwellers may value urban lifestyle benefits, while emphasizing convenience and advanced tools could increase engagement in smaller cities.

Education	A diverse educational background suggests tiered rewards and cumulative discounts as effective strategies. Specialized programs for highly engaged users could incentivize further interaction.	Simplified applications and a straightforward discount system may cater well to varying educational levels, encouraging a shift towards more engaged clusters.
Age	Target the 31-50 age group with tools for financial management and career peak challenges, adding value with personalized offers and insights.	Cater to the financial planning needs of those 31-50 and introduce engaging financial literacy for the younger demographic to foster early banking habits.

7. Conclusion

This study makes notable contributions to both theory and practice by advancing the AHP-AHC methodology and applying it to electronic banking adoption, using RFM as its framework. The key contributions are:

1. Extended AHP-AHC methodology: This study introduces a dual application of AHP in the integrated AHP-AHC method, allowing for a more refined and entangled decision-making process. The dual usage enhances the theoretical framework of AHP-AHC, offering a more detailed prioritization of customer segments, which leads to better-targeted marketing strategies and more effective decision-making in service-based industries.
2. RFM as a framework for electronic banking adoption: This study is the first to apply the RFM model to electronic banking adoption, providing a more tangible and data-driven approach to customer segmentation. Using actual transactional data offers a clearer and more reliable analysis of customer engagement, moving away from subjective survey data and contributing to more accurate strategic planning.

These contributions enhance the theoretical understanding of multi-step AHP applications and provide practical tools for improving customer segmentation. The model's robustness is further confirmed by CR values of 0.003 and 0.007, along with the experts' approval. Future research could explore this approach's applicability in other service sectors like e-commerce or healthcare to test its broader generalizability.

8. Limitations

This study has several limitations. First, it uses data from a single Iranian bank, limiting the generalizability of the findings to other regions or industries with different economic or cultural contexts. Second, while the RFM framework captures transactional data effectively, it may overlook factors like customer satisfaction or digital literacy, which could offer a more comprehensive view of customer behavior. Future research should address these limitations by applying the methodology across diverse contexts and incorporating a wider range of customer behaviors to enhance generalizability and objectivity.

9. Key References

Hillerman, T., Souza, J. C. F., Reis, A. C. B., & Carvalho, R. N. (2017). Applying clustering and AHP methods for evaluating suspect healthcare claims. *Journal of Computational Science*, 19, 97–111. https://doi.org/10.1016/J.JOCS.2017.02.007

Hosseini, M., Abdolvand, N., & Harandi, S. R. (2022). Two-dimensional analysis of customer behavior in traditional and electronic banking. *Digital Business*, 2(2), 100030. https://doi.org/10.1016/J.DIGBUS.2022.100030

Khajvand, M., & Tarokh, M. J. (2011). Estimating customer future value of different customer segments based on adapted RFM model in retail banking context. *Procedia Computer Science*, 3, 1327–1332. https://doi.org/10.1016/J.PROCS.2011.01.011

Kwateng, K. O., Atiemo, K. A. O., & Appiah, C. (2018). Acceptance and use of mobile banking: an application of UTAUT2. *Journal of enterprise information management*, 32(1), 118-151. https://doi.org/10.1108/JEIM-03-2018-0055

Sharma, R., Singh, G., & Sharma, S. (2020). Modelling internet banking adoption in Fiji: A developing country perspective. *International Journal of Information Management*, 53, 102116. https://doi.org/10.1016/J.IJINFOMGT.2020.102116