International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 5685-5702

https://doi.org/10.70517/ijhsa463490

# Quantitative Research on Changes in Residents' Consumption Behavior and the Path of Consumption Upgrading Based on Big Data Analysis

**Jing Li**[1,*]
[1] School of Economic, Shandong University, Jinan, Shandong, 250100, China
Corresponding authors: (e-mail: lijingsdu@hotmail.com).

**Abstract** With the development of the digital economy, residents' consumption behaviors have become increasingly diverse. This study employs big data analytics to explore the patterns of change in consumer behavior and the pathways of consumption upgrading. Building upon the traditional RFMT model, the study introduces the "T" indicator for recommendation traffic to construct an enhanced RFMT segmentation model. Additionally, the SOM neural network model is improved through optimized learning rate design to enhance training stability and clustering accuracy. Based on the consumption data of 4,158 households, clustering analysis identifies four typical consumer groups: core type, habitual type, supportive type, and general type. By incorporating five individual factors such as city tier, the study finds that the core type represents a growth engine composed of high-net-worth individuals, the habitual type reflects a pragmatic group with stable repurchase behaviors, the supportive type includes high-potential scenario-driven consumers, and the general type consists of price-sensitive long-tail users. Further analysis using eight indicators, including digital technology usage, identifies four consumption behavior types: technology-empowered consumption, interest-driven consumption, socially embedded consumption, and resource-constrained consumption. The findings reveal that residents' consumption behaviors are influenced by a combination of factors, with significant differences among consumer groups. The study recommends designing differentiated consumption upgrading strategies tailored to the needs of each group to expand the consumer market and promote high-quality development of the consumption economy.

**Index Terms** residents' consumption behaviour, big data analysis, RFMT model, SOM neural network, consumption upgrade path, quantitative research

## I. Introduction

Residents' consumption behaviour is influenced by their physiological, psychological, social, economic, cultural and other factors [1]. Observation of reality found that the relationship between residents' income growth and consumption growth is intricate, and the practice of simply stimulating farmers to increase income to boost consumption growth is questionable [2]. In fact, behind the explicit factors such as income, expectation, uncertainty, etc., the trajectory of changes in residents' consumption behaviour has been marked by institutional change [3]. By arranging and adjusting the contradictions between farmers' productivity and production relations, consumption capacity and consumption relations, the system and its changes cause changes in residents' consumption psychology and consumption behaviour, which are ultimately reflected in the changes in residents' consumption level and consumption structure [4]-[6]. Therefore, exploring the influencing factors and changing rules of residents' consumption behaviour is not only urgent, but also has strong theoretical value and practical application value [7].

The issue of changes in residents' consumption behaviour and consumption upgrading has been a hot issue in academic research at home and abroad, and scholars at home and abroad have published a large number of research results around the connotation, influencing factors, and mechanism of residents' consumption behaviour changes and consumption upgrading [8]-[10].Chen, D and Guo, X found that under the synergistic effect of the digital economy and financial development, China significantly promotes residents' consumption upgrading and Chen, D and Guo, X found that under the synergistic effect of digital economy and financial development, China has significantly promoted the upgrading of residents' consumption and the transformation of consumption structure, the common mechanism of which is the optimisation of industrial structure, which exhibits regional and urban-rural heterogeneity as well as the non-linear characteristics of new urbanisation [11].Feng, X and Du, G pointed out that the level of financial literacy has a significant contribution to residents' consumption upgrading and that this effect is more pronounced for rural households [12].Shen, A et al. found that new infrastructure has a significant heterogeneous effect on the expansion and promotion of China's residents' consumption upgrading. Chinese residents' consumption upgrading has a significant heterogeneous effect, in which the optimisation of industrial structure is the key mechanism [13].Le, X et al. study the relationship between urbanisation, digital economy and rural residents' consumption upgrading in China, and find that urbanisation mediated by digital economy significantly promotes rural consumption upgrading, and that urbanisation has a double-threshold effect on rural consumption upgrading [14].Zhang, P and Gao , J. found that lowering the burden of medical costs can stabilise or increase individuals' expectations of future income levels, and high-quality public health insurance can promote the upgrading of China's personal consumption structure [15].

The emergence of big data has invariably changed people's way of thinking and lifestyle, and the difficulty of data processing will be greatly increased. Deep mining and analysis through big data can help to help study the trend of changes in residents' behaviour and provide reliable and accurate information for follow-up work [16].Theodorakopoulos, L et al. use big data to analyse consumer behaviour, mining individual consumer characteristics from massive data to understand consumer preferences, needs and behavioural data [17]. Applying big data technology to accurately grasp the changes and patterns of consumer behaviour through data collection, data analysis and processing can develop a more precise strategy for consumption upgrading in order to facilitate a more flexible response to consumer demand and effectively improve economic efficiency [18], [19].

Residents' consumption behaviour is influenced by their physiological, psychological, social, economic and cultural factors. Realistic observation shows that the relationship between residents' income growth and consumption growth is intricate and complex, and the practice of simply stimulating farmers' incomes to boost consumption growth needs to be further explored. In fact, behind the obvious factors such as income, expectation and uncertainty, the trajectory of changes in residents' consumption behaviour has been marked by institutional changes. By arranging and adjusting the contradictions between farmers' productive forces and production relations, and between consumption capacity and consumption relations, the system and its changes have caused changes in residents' consumption psychology and behaviour, which are ultimately reflected in the changes in residents' consumption level and consumption structure. Therefore, it is of great theoretical value and practical application value to explore the influencing factors and changing rules of residents' consumption behaviour.

Changes in residents' consumption behaviour and consumption upgrading have been a hot issue in academic research at home and abroad, and scholars at home and abroad have published a large number of research results on the connotation, influencing factors, and mechanism of residents' consumption behaviour changes and consumption upgrading. Chen and Guo found that the synergistic effect of China's digital economy and financial development significantly promotes the upgrading of residents' consumption and the transformation of the consumption structure, and the common mechanism is the industrial structure, which is the result of the synergistic effect of the digital economy and financial development. Feng and Du pointed out that the level of financial literacy has a significant promotion effect on residents' consumption upgrading, and this effect is more significant for rural households.Shen et al. found that the new infrastructure has a significant heterogeneous effect on expanding and enhancing the consumption upgrading of China's residents, with the optimisation of the industrial structure as the key mechanism.Le et al. investigated the relationship between urbanisation, the digital economy and the consumption upgrading of China's rural residents Le et al. found that urbanisation mediated by digital economy significantly promotes rural consumption upgrading, and that urbanisation has a double threshold effect on rural consumption upgrading. Zhang and Gao found that lowering the burden of healthcare costs can stabilise or raise individuals' expectations of future income levels, and that high-quality public health insurance can promote the upgrading of China's individual consumption structure.

The emergence of big data has invariably changed people's way of thinking and lifestyle, and the difficulty of data processing will be greatly increased. Deep mining and analysis through big data can help to help study the changing trends of consumer behaviour and provide reliable and accurate information for the follow-up work.Theodorakopoulos et al. analyse consumer behaviour through big data, mining individual consumer consumption characteristics from massive data to understand their preferences, needs and behaviours and other data. Applying big data technology to accurately grasp the changes and patterns of consumer behaviour through data collection, data analysis and processing can provide a more precise strategy for consumption upgrading, thus responding more flexibly to consumer demand and effectively improving economic efficiency.

Starting from the perspective of big data, this study firstly introduces T indicators on the basis of the traditional RFM model, constructs the RFMT segmentation model, and determines the weights of each indicator through hierarchical analysis; secondly, improves the SOM neural network model, optimises the design of the learning rate function, and improves the clustering accuracy; then, conducts a cluster analysis of the residents' consumption data, and identifies the characteristics of different types of consumption groups; finally, adopts the group analysis method to explore the influencing factors of residents' consumption behaviour and the path of consumption upgrading.

Through quantitative research, the rule of change of residents' consumption behaviour is deeply analysed to provide data support and decision-making reference for promoting the development of consumer economy.

# II. Analysis of the consumer behaviour of the population

## *II. A. Modelling RFMT Segmentation*

### II. A. 1) Determination of RFMT model metrics

On the basis of the traditional RFM model, combined with the importance of consumer group fission in the context of the traffic era, this paper will introduce the T (referral flow) indicator to construct the RFMT model.The introduction of the T indicator can more comprehensively measure the ability of the consumer group to be converted into traffic through fission and other means, and also identify the potential value of the consumer group, which makes the RFMT model more in line with the characteristics of the contemporary era, and provides a theoretical basis for the individual segmentation provides a theoretical basis.

where R (recency) denotes the time interval between the most recent consumption and the analysed time node, the larger the R value indicates that the longer the most recent consumption is from the current time interval, the lower the potential value of the individual F (frequency) denotes the number of times an individual consumes within the analysed time node, the larger the F value indicates that the higher the number of consumptions, the higher the potential value [20].M (average monetary amount) denotes the average amount of consumption within the analysed time node, the larger the M value indicates that the higher the consumption amount, the higher the value. M (average amount of money) indicates the average amount of consumption within the analysed time node, the larger the M value indicates the higher the amount of consumption and the higher the value. t (recommended traffic) indicates that the individual is converted into traffic through the consumer group cleavage within the analysed time node, the larger the T value indicates that the stronger the word-of-mouth recommendation ability is, the more new consumer groups are generated, and the higher the value of the individual is.

### II. A. 2) Value classification of consumer groups based on RFMT model

From the classification of consumer groups based on RFMT model, consumer groups can be classified into four categories, and the value classification of consumer groups is shown in Fig. 1, which can be classified into four categories: Core type, Habitual type, Supportive type and General type.

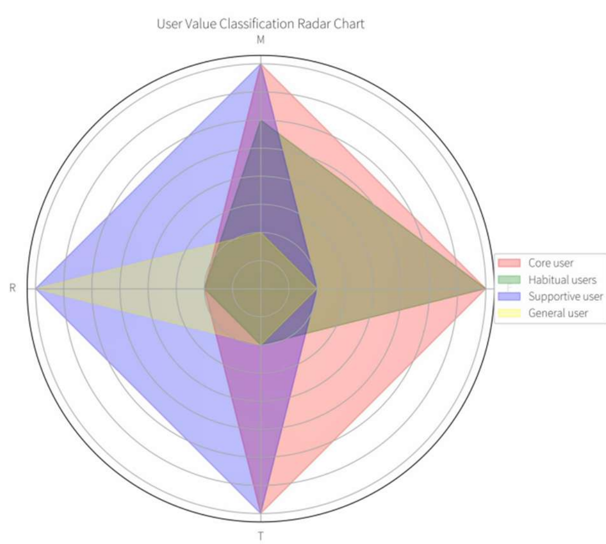


Figure 1: Classification of type values

As shown in Figure 1, the core type is set to high F, high M, low R, high T, indicating active and high value; the habitual type is set to high F, medium M, low R, low T, indicating habitual purchases but no promotion; the supportive type is set to low F, high M, high R, high T, belonging to the type of occasional large-dollar spending and active promotion; and the general type is set to low F, low M, high R, low T, indicating inactive and low value.

### *II. B. Construction of the methodological system for consumer group segmentation*

#### II. B. 1) Data pre-processing

(1) Data collection

The data used for the study were collected to create a panel dataset. Generally come to resource a nationally representative household level financial database, covering a more complete range of provinces, autonomous regions and municipalities to ensure the representativeness of the results. Also note that the time interval is within a reasonable range, generally defaulting to 10 years.

(2) Data pre-processing

In order to reduce the waste of resources and ensure the accuracy of the data mining results, the data should be pre-processed before analysing the raw data, which mainly includes the processing of noisy data and missing values, as well as standardised calculation of the raw data.

(a) Noise data processing

Noisy data is often anomalous data with errors or deviations from expected values in the raw data, which can be pre-processed using methods such as regression, split-box and outlier analysis.

(b) Missing value processing

In the treatment of missing values in the raw data, filling in the missing areas of the raw data is mostly done by using indicator-centred metrics and methods such as decision trees and regression.

(c) Data standardisation

Before analysing the raw data, it is necessary to standardise the data and then analyse the standardised data. Commonly used data standardisation methods include "deviation standardisation (min-max standardisation)", "standard deviation standardisation (Z-score standardisation)" and "normalisation standardisation". The specific application of each method is described below:

Min-max normalisation, also known as deviation normalisation, is a method that is suitable for cases where the maximum and minimum values of the sample are known. A linear transformation is applied to the original data and then the result is mapped to the interval [0, 1]. The specific formula is as follows

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

Z-value standardisation, also known as standard deviation standardisation, this method is applicable to the case where the processed data meets the standard normal distribution and the maximum and minimum values of the original data are unknown. The specific formula is as follows

$$y_i = \frac{x_i - \overline{x}}{s}, (1 \le i \le n) \tag{2}$$

where $\bar{x}$ is the mean of the original data and $s$ is the standard deviation of the original data. After normalisation by Z-value, 0 and 1 are the mean and standard deviation of the indicator respectively, which are pure quantities without any units.

The normalisation method is used to determine the weights of indicators, and the new sequence $y_i$ takes the value range of [0, 1] and satisfies $\sum_{1}^{n} y_i = 1$ . The specific formula is as follows:

$$y_i = \frac{X_i}{\sum_{1}^{n} X_i}, (1 \le i \le n) \tag{3}$$

#### II. B. 2) Calculation of indicator weights

The determination of the weight of each influencing factor affecting customer value is different, so it is necessary to combine the context of the times and the development of the industry to give each indicator a different weight. The specific calculation process of the hierarchical analysis method is as follows:

(1) Hierarchical modelling

Decision objectives, considerations and decision objects are classified in their interrelationship into a top, middle and bottom level, thus forming a hierarchical structure. The top level represents the purpose of the decision or the problem to be addressed, the middle level represents the factors to be considered or the decision criteria, and the bottom level represents the options for decision-making.

(2) Constructing a judgement matrix

The judgement matrix represents the comparison of the relative importance of all the decision factors in this level with the decision factors in the upper level. Where the elements of the judgement matrix $a_{ij}$ are given by Santy's 1-9 scale method.

According to the table, the relative importance scale of each element is determined by comparing each decision element two by two. The relative importance scales of the two decision elements are then quoted, resulting in the ratio of the relative importance scales of the two decision elements, thus constructing a judgement matrix A. The specific formula for the judgement matrix is as follows

$$A=(a_{ij})_{n*n}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \tag{4}$$

(3) Single ranking of hierarchical importance

A is normalised and its eigenvector, which corresponds to the largest eigenvalue of A, is calculated:

$$W=(w_1,w_2,\cdots,w_i,\cdots,w_n)^T \tag{5}$$

$$w_i=\frac{\sum_{j=1,i=1,2,\cdots,n}^{n} a_{ij}}{n\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}} \tag{6}$$

Find the maximum eigenvalue : $\lambda\max$

$$\lambda_{\max}=\sum_{i=1}^{n}\frac{\sum_{i=1}^{n} a_{ij}w_j}{nw_i} \tag{7}$$

(4) Consistency testing

Firstly, the consistency index CI is calculated with the following formula:

$$CI=\frac{\lambda_{\max}-n}{n-1} \tag{8}$$

Considering the random nature of consistency bias, when testing whether the judgement matrix satisfies consistency, it is also necessary to compare the CI value with the average random consistency index RI to derive the test coefficient, which is given by

$$CR=\frac{CI}{RI} \tag{9}$$

CR is the consistency ratio of the judgement matrix, when CR is smaller, it means the reliability of the indicator weights is higher. When CR=1, the judgement can be considered fully consistent. When CR<0.1, it can be regarded as satisfactory. When CR>0.1, the judgement matrix should be corrected [21].

After determining the weights of each indicator, the weight coefficients were substituted into the RFMT composite score formula, and the values were divided according to the individual's RFMT composite score ranking, which was specified as follows:

$$RFMT=w_R R'+w_F F'+w_M M'+w_T T' \tag{10}$$

where $w_R, w_F, w_M, w_T$ is the weight coefficient of each indicator and $R', F', M', T'$ is the pre-processed indicator value.

### *II. C.SOM Neural Network Model Analysis and Improvement*

The SOM neural network is capable of forming input signals on an array of processing units (one or two dimensions). The self-organising mapping function of the model is achieved by simulating the neural network of the human brain.The SOM neural network mainly consists of an input layer and a competing layer.The number of vectors in the input network determines the number of neurons in the input layer. When external signals are fed into the network, the competing layer analyses and compares the input variables to find patterns and classify them [22].The structure of the SOM neural network is shown in Fig. 2.

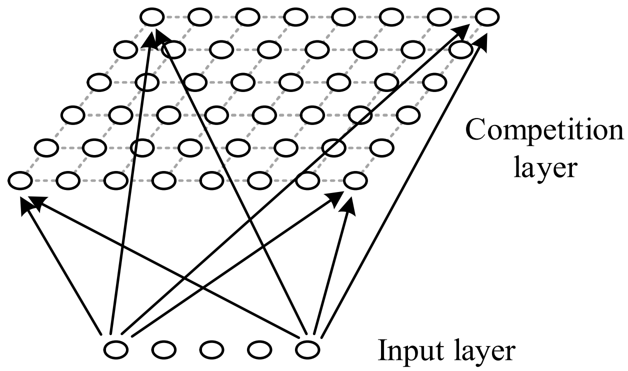


Figure 2: Structure of SOM neural network

(1) Initialisation: randomly initialise the parameters of each node. The number of parameters for each node is consistent with the dimension of the input data.

(2) Find the node that best matches each input data. Suppose the dimension of the input data is $D$ , i.e., $X=\{x_i\}$ , where $i=1,2,\cdots,D$ , the discriminant function is the Euclidean distance as shown in Eq. (11).

$$d_j(x)=\sum_{i=1}^{D}(x_i-w_{ij})^2 \tag{11}$$

(3) Update the nodes adjacent to the found activation node $I(x)$ . Let $S_{i,j}$ denote the Euclidean distance between nodes $i$ and $j$ and assign update weights to the nodes neighbouring $I(x)$ as shown in equation (12).

$$W_{j,I(x)}=\exp(S_{j,I(x)}^2/2\sigma^2) \tag{12}$$

That is, the update degree of neighbouring nodes is affected by their distance from the active node $I(x)$ .

(4) Update the node parameters according to the gradient descent method as shown in Eq. (13).

$$\nabla w_{ji}=\eta(t)\cdot W_{j,I(x)}(t)\cdot(x_i-w_{ji}) \tag{13}$$

Iterate until convergence.

How to choose the appropriate learning rate has always been the difficulty in SOM neural network training. When the learning rate is large, the weight vectors will be repeatedly oscillated and updated, leading to a decrease in training stability. When the learning rate is close to 0, although the stability of learning is improved, the convergence efficiency of the network is reduced [23]. Therefore, the learning rate is set as a monotonically decreasing function over time $t$ , which ensures that the model learns at a faster rate at the beginning of training, and the learning rate converges to 0 towards the end of training, ensuring the stability of model training. The learning rate function $\lambda(t)$ is shown in equation (14).

$$\lambda(t)=\lambda(0)\exp(-\frac{t}{5T}) \tag{14}$$

where $\lambda(0)$ is the initial learning rate of the model and $T$ is the learning step size. This method can improve the learning rate of the SOM, but also ensure the convergence of the training process and the stability of the SOM neural network.

The flow of the improved SOM neural network algorithm is shown in Fig. 3. In this paper, the following two algorithms are used as comparison algorithms of SOM neural network algorithm:

(1) Method 1: Traditional SOM neural network algorithm.

(2) Method II: Setting the learning efficiency as

$$\lambda(t)=\lambda(0)\exp(-\frac{t}{3T}) \tag{15}$$

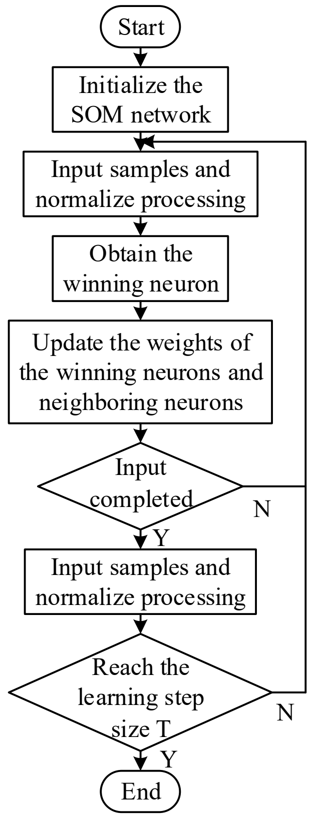


Figure 3: Training process of the improved SOM neural network model

## III. Cluster analysis of the behaviour of the population's consumption activities

### *III. A. Raw data cleansing*

The data used in this study come from the China Household Finance Survey (CHFS) conducted by the China Household Finance Survey and Research Centre of Southwestern University of Finance and Economics in 2013, 2015, 2017 and 2019 (swufe.edu.cn, accessed on 18 January 2024). The CHFS has created the first nationally representative household-level financial database, covering 29 provinces (including autonomous regions and municipalities directly under the central government) in 2013. In subsequent years, the survey sample included both newly included participants (the new sample) and previously surveyed households (the old sample). Tracking data spanning four periods (2013-2019), totalling a sample of 4,158 tracked households, are retained for this study. The sample size of the balanced panel data is 16,332.The specific cleaning process is as follows (Table 1).

Table 1: Description of cleaning objectives and raw indicators

| norm | define | Cleaning priorities |
|---|---|---|
| Average time of use per day | Average daily minutes of platform usage | Range calibration (0 to 1440 minutes) |
| Average amount spent | Total annual consumption ÷ Number of valid consumption | Non-negative value checking, extreme value filtering |
| Average trading interval | Average number of days between two adjacent purchases | Logic check (interval ≥ 0) |
| Number of transactions per year | Backcasting through intervals: 365 ÷ average number of days between intervals | Special handling when the interval is 0 |

Note: Each family x category x year = 1 row

For outlier handling, the reasonable range of the length of use is 0~1440 minutes, and the exceeding value is set as NA; the reasonable range of the consumption amount is 0.1~50000 RMB, and the exceeding sample is deleted; the reasonable range of the transaction interval is 0~365 days, and the interval=0 is regarded as a high-frequency consumption group. Cross-check the deviation rate between the number of annual transactions and the original number of consumption, and allow ±10% error. Examples of cleaned data are shown below:

Table 2: Example of cleaned data

| Family ID | Average time of use per day (minutes) | Average amount spent ($) | Average trading interval (days) | Number of transactions per year | data state |
|---|---|---|---|---|---|
| F001 | 86.3 | 420.5 | 5.2 | 70.2 | validity |
| F002 | 152.1 | 1250.8 | 15 | 24.3 | efficiently |
| F003 | 25 | 630.2 | - | removing | interval missing |
| F004 | 1500 | - | 30 | removing | excessive usage time |

The cleaned data contains household ID, RFMT metrics, and number of annual transactions, with an original sample size of 4,158 households, an effective sample size of ≥4,100 households (cleaning loss rate <8%), and an outlier accounting for <5% of the total number of outliers after controlling for them through boundary rules.

### *III. B. SOM Clustering*

In this study, SOM cluster analysis was carried out, and the algorithm was used to cluster the empowered dataset to obtain four groups of residents, and the mean values of the four indicators for each group and the number of individuals in each category were calculated. The results of clustering of consumer groups are shown in Table 3 (N denotes the number of samples, P denotes the proportion of samples).

Table 3 Results of SOM cluster analysis

| form | Core type | Habitual type | Supportive type | General type |
|---|---|---|---|---|
| R | 0.018 | 0.016 | 0.032 | 0.028 |
| F | 0.21 | 0.18 | 0.07 | 0.05 |
| M | 0.19 | 0.11 | 0.15 | 0.08 |
| T | 0.18 | 0.05 | 0.2 | 0.04 |
| N | 762 | 1211 | 1074 | 1113 |
| P | 18.31 per cent | 29.11 per cent | 25.82 per cent | 26.76 per cent |

Combined with the clustering results in Table 3SOM, this paper plots the RFMT probability distributions of the four categories of residents, as shown in Figures 4-Figure 7, respectively:

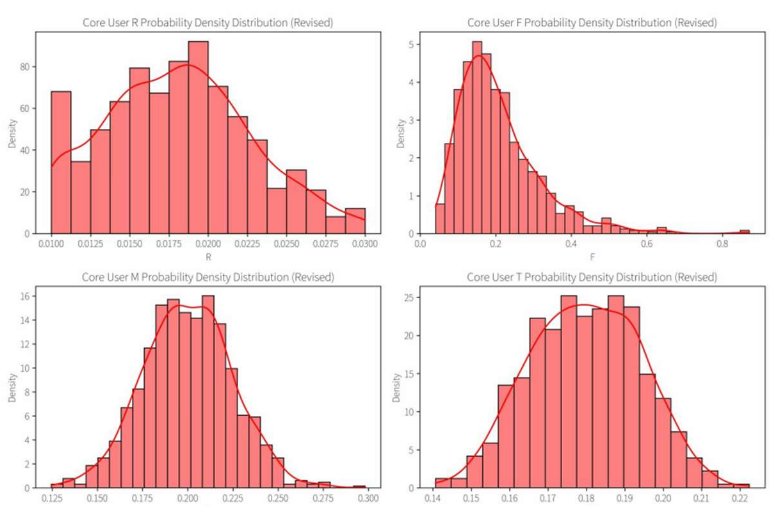


Figure 4: Probability density distribution of RFMT for Core type

As shown in Table 2 and Figure 4, Core type accounts for 17.5%, which is a value benchmarking consumer group. It has the highest activity (R-value: 0.0179), the highest purchase frequency (F-value: 0.2036), the strongest consumption ability (M-value: 0.1852) and outstanding communication influence (T-value: 0.1924). Typical characteristics are high purchase frequency, high consumption amount and high promotion degree, which belong to the core supporters of the brand and the main force of word-of-mouth communication. From the perspective of probability distribution, the R value of Core type is highly concentrated in the 0.01-0.03 range (mean value: 0.0179), indicating that the consumption time is close; the distribution of F value is right-skewed (peak value: 0.20), with significant high-frequency consumption characteristics; the wide-peak distribution of M value (0.15-0.25) reflects the diversity of high unit price; the bimodal distribution near the T value of 0.18 combines both stable promotion and explosive communication ability. explosive dissemination ability.

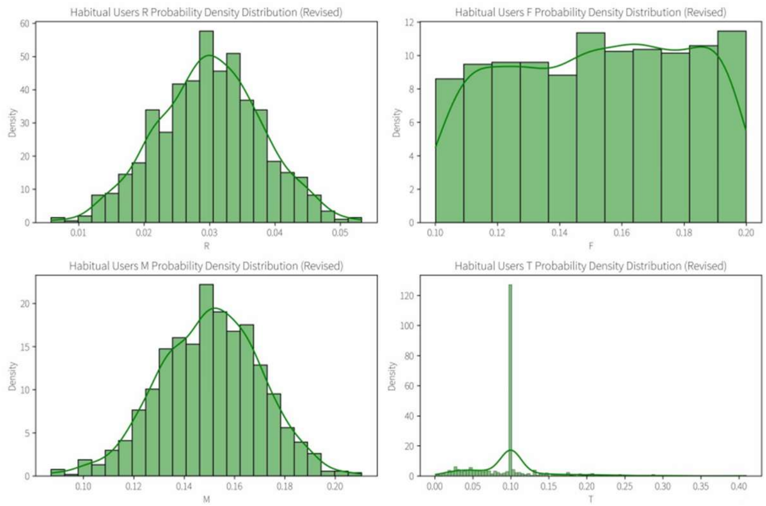


Figure 5: Habitual type RFMT probability density distribution

As shown in Table 2 and Figure 5, Habitual type accounted for 25.8%, which is a stable consumer group. It is a stable consumer group with long-term repurchase behaviour (low R-value: 0.0326, high F-value: 0.1528), but medium consumption power (M-value: 0.1327) and weak willingness to actively disseminate (T-value: 0.1046). This type of consumer group relies on consumption inertia, and although it contributes stable income, it needs to be stimulated by personalised services to stimulate its social promotion potential. In terms of probability distribution, the R value is a dense single peak near 0.03, which is slightly more dispersed than the core type; the F value is a narrow peak in the range of 0.10-0.20, which shows stable repurchase behaviours; the M value is a symmetric bell-shaped distribution, with small fluctuations in the amount of consumption; and the T value is a long-tail with a peak at 0.10, which shows occasional promotional behaviours.

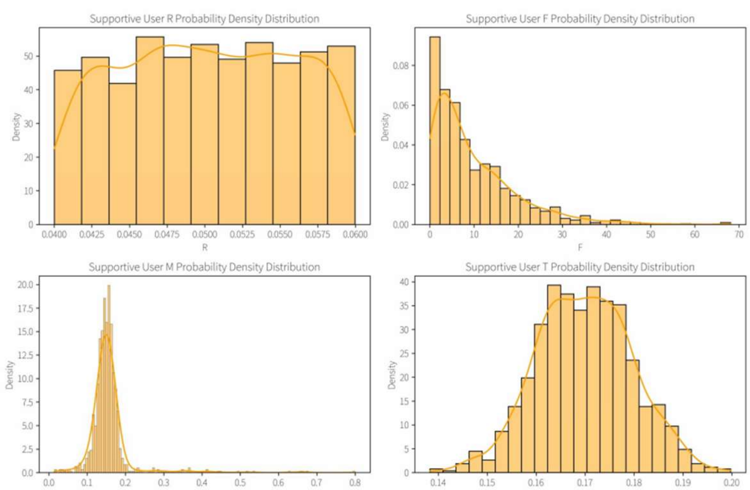


Figure 6: Supportive type RFMT probability density distribution

As shown in Table 2 and Figure 6, the Supportive type, accounting for 21.7% of the total, is the high spreading power group. It has the strongest promotional ability (highest T-value: 0.1763) and high single consumption (M-value: 0.1539, second only to the core type). Characterised by episodic large-value consumption (e.g. limited edition products or holiday consumption) with high referral rate, they are able to radiate more potential consumer groups through social influence. From the perspective of probability distribution, there is a wide peak of R value at 0.04-0.06, indicating medium consumption activity; F value shows exponential decay in the 0-0.15 interval, indicating its low-frequency consumption characteristics; M value at 0.15 reflects the peak + long-tail characteristics, indicating the existence of episodic high-value consumption behaviours; and T value shows a peak near 0.17, reflecting a very strong concentration of promotions.

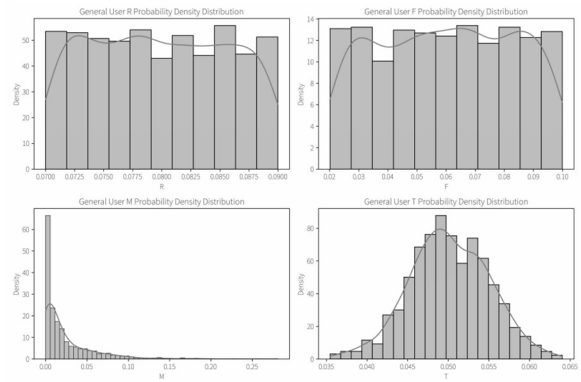


Figure 7: General type RFMT probability density distribution

As shown in Table 2 and Figure 7, the General type is a low-value group with a proportion of 35.0%. They have the lowest activity (R value: 0.0784, F value: 0.0621), weak consumption contribution (M value: 0.0718) and almost no communication behaviour (T value: 0.0562). It is a low-frequency and low-value consumption, which needs to be improved through activation strategies (such as promotion or experience optimisation) to enhance its conversion possibility. From the perspective of probability distribution, the R value is dense in the interval of 0.07-0.09, indicating that the consumption interval is the longest; the F value is gently distributed in the interval of 0.02-0.10, indicating that there is no consumption pattern; the M value exhibits an obvious near-zero-valued single peak, which is reflected in the lowest level of consumption; and the T value shows a narrow peak at 0.05, indicating that this type of consumer group has almost no promotional behaviours.

### *III. C. Characterisation of residents of different value types*

In order to further analyse the characteristics of residents of different value types, this paper rearranged the data and found that the distribution of residents of different value types is most obvious for five types of indicators, namely, city_level, per capita income (perinc), Internet use (internet), education level (edu), Engel's coefficient (engel = foodfee/total_consump), etc. Therefore, the above five indicators are set as prerequisites to analyse the characteristics of the four types of consumer groups with different values, respectively:

#### III. C. 1) Core: a growth engine for HNWIs

The results show that the core type group belongs to, first-tier cities, high-knowledge groups, all-dimensional consumption leaders, with the following characteristics: concentrated in first-tier cities, higher per capita income, frequent use of the Internet, higher education, and a lower Engel's coefficient. These characteristics together reflect the high value, high activity and high loyalty of the core type.

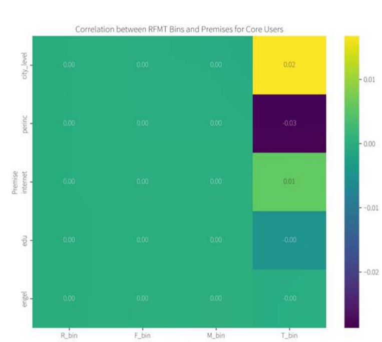


Figure 8: Characteristic heat map of the core type

As shown in Figure 8, the core type is concentrated in first-tier cities (accounting for more than 70% of the total), and its consumption behaviour exhibits significant "high-frequency and high customer order" characteristics: the R value (≤30 days) is strongly correlated with the M value (≥90th percentile) (r>0.6), reflecting a high degree of unity between its consumption activity and payment ability. This group is typically labelled with high income (per capita monthly consumption up to 3 times that of the general consumer group) and high education (80% above bachelor's degree), forming a self-reinforcing closed loop of consumption upgrading - disposable income released by a low Engel's coefficient (≤0.25) is highly invested in quality consumption of recreation and entertainment, tourism and health care. The depth of its digital penetration is particularly outstanding: 90% rely on online channels for repurchase, and the correlation coefficient between Internet usage and F/M value reaches 0.52, highlighting the reliance on all-link digital consumption.

#### III. C. 2) Habitual: a pragmatic group of steady state repurchasers

The results show that the habitual type group usually exhibits regular purchasing characteristics dominated by middle-class, rigid demand in second- and third-tier cities, and the correlation between their RFMT indicators and the five types of indicators is generally weaker compared to the core type. This reflects that the behavioural patterns of habitual types are relatively stable, and the differences in different indicators are less obvious than those of core types.

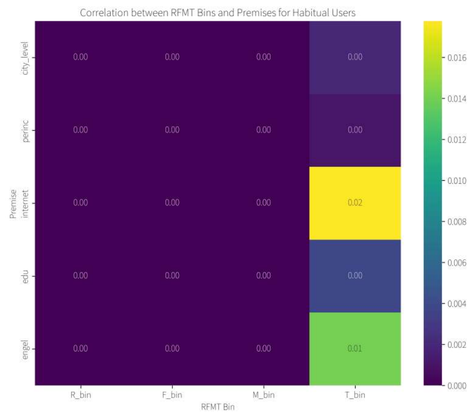


Figure 9: Characteristic heat map of habitual type

As shown in Figure 9, the habitual type is mainly settled in second- and third-tier cities (60 per cent of the total), and its consumption structure shows a distinct rigid demand orientation: the proportion of spending on necessities (food/daily necessities) accounts for 45 per cent of the total, with a median Engel coefficient of 0.35, which significantly squeezes out the space for optional consumption. Compared with the core type, this group is less sensitive to the urban tier (r=0.21), and income growth has not been synchronised to translate into consumption upgrading - per capita income is only weakly correlated with M-value (r=0.19), reflecting a pragmatic view of consumption. Their Internet use is instrumental: 68% of the consumer group uses price comparison tools with high frequency, but the correlation between online behaviour and F/M value (r=0.16) is much lower than that of brand loyalty, reflecting the behavioural logic of "giving priority to channel value over emotional connection". For this group, subscription-based stocking discounts and points multiplier programmes can effectively increase the frequency of repurchase, while the expansion of category breadth needs to be tied to family scenarios (e.g., mother and baby products bundled with daily consumables).

#### III. C. 3) Supportive: high-potential scenario triggers

Compared with the core and habitual types, the supportive group is characterised by low-frequency, high-communication, scene-triggered consumption, and the correlation between the RFMT indicators and the five types of indicators is generally weaker. This reflects the relatively stable behavioural patterns of the supportive types and the differences in different indicators are less pronounced than those of the core and habitual types.

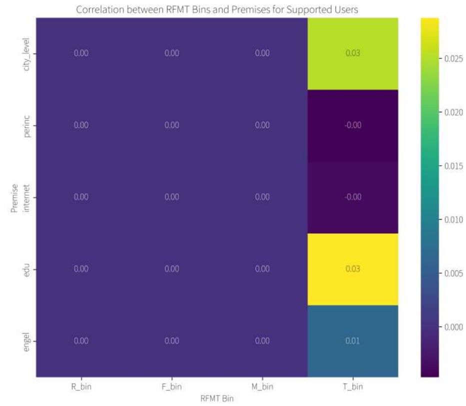

Figure 10: Characteristic heat map of the support type

As shown in Figure 10, the core of the value of the support type lies in the discontinuous high-impact consumption characteristics: its T-value (trust index) reaches 0.7, but the F-value is ≤2 times/quarter, its consumption behaviours are highly concentrated in the social scenarios such as festival gifts and family purchases, and the fluctuation range of the single consumption amount is up to 300%. This group breaks the logic of conventional geographical distribution - the T value of the consumer group in the fourth-tier cities exceeds that of the first-tier cities by 15%, and there is a significant deviation between educational background and loyalty (37% of the consumer group with low education and high dissemination). Their fission value is realised through the strong communication attributes of the Internet: the conversion rate of sharing behaviour is 2.1 times higher than the average value, which is especially suitable for social incentive strategies such as cashback for group sharing. To activate this type of consumer group, we need to focus on scenario-based explosive product design (e.g. holiday limited edition gift boxes), and use the "woolgathering mentality" to create the perception of taking advantage of a bargain, such as a limited-time free gift combined with the KOC grass-raising commission mechanism.

**III. C. 4) General type: price-sensitive long-tail flows**

Compared to Core, Habitual and Support, the General type group is a low-activity, low-loyalty, bottom-of-the-funnel segment with the weakest correlation between RFMT indicators and the five categories. This reflects the fact that the General type has the most stable behavioural patterns and the differences in the different indicators are not as pronounced as in the other consumer group segments.

Figure 11: Characteristic heat map of the general type

As shown in Figure 11, the consumption trajectory of the general type exhibits irregular randomness: the correlation coefficients between each indicator and RFMT are all ≤0.15, and 60% of consumption is triggered by promotion-driven or emergency demand. Survival cost squeeze is its core constraint - the median Engel's coefficient is 0.42, more than half of the consumer group has average monthly consumption <2,000 RMB, and rigid expenditure severely inhibits optional consumption upgrading. This group is trapped in the digital divide: the Internet usage rate is only 41%, online transactions account for less than 30%, and traditional channels have significant barriers to reach. The operation strategy needs to be extremely lightweight: activate the bottom of the funnel traffic through the 9.9 yuan high discount diversion products, and at the same time layout of offline community touchpoints (such as convenient service stations) to reduce the cost of customer acquisition; for the consumer group with no conversion intention, the natural elimination mechanism can be activated to reduce the loss of resources.

# IV. Exploration of factors influencing residents' consumption behaviour and the path of upgrading

## IV. A. Analysis of the need for individual conditions

The necessity test for individual conditions was a necessary prerequisite for subgroup analyses, and a conditional variable was considered necessary to lead to an outcome if it existed with a concordance greater than 0.9. Indicators for individual conditions are summarised in Table 4:

Table 4: Summary of indicators for individual conditions

| dimension (math.) | Directly related indicators | Indirect/derived indicators | The significance of economic behaviour |
|---|---|---|---|
| Use of digital technology | wanggou, sphone, internet | c8001c | Digital life penetration and dependency |
| Product Interests | wenyu, cloth, lux | travel | High value-added product sensitivity |
| Access to information | tongxin, tongxun | internet | Intensity of investment in information channels |
| affinities | renqing | child, happiness | Social trust-building tendencies |
| Policy implications | yanglaoxian, yiliao_shehui | yiliao_shangye | Policy Compliance and Risk Avoidance |
| self-improvement | jiaopei, tisheng | edu | Sustainability of investment in human capital |
| capital resources | total_asset, total_debt | fangchan, car | Resilience and asset structure |

After that, in accordance with the consistency test criteria of the AHP method in 2.2.2, 30 experts were invited to rate the relative importance of each indicator, and the invited experts covered the fields of sociology, public management, and residents' consumption behaviours, etc., to test the above individual condition indicators. Considering the space reason, this paper only lists the final results of the consistency test, such as Table 5.

Table 5: Results of AHP consistency test for individual condition indicators

| dimension (math.) | Judgement matrix order(n) | Maximum eigenvalue ($\lambda_{max}$) | CI value | RI value | CR value | Pass or fail the test (CR<0.1) |
|---|---|---|---|---|---|---|
| Use of digital technology | 3 | 3.024 | 0.012 | 0.58 | 0.021 | be |
| Product Interests | 3 | 3.086 | 0.043 | 0.58 | 0.074 | be |
| Access to information | 2 | 2 | 0 | 0 | 0.000 | be |
| affinities | 3 | 3.105 | 0.052 | 0.58 | 0.090 | Yes (critical adoption) |
| Policy implications | 3 | 3.172 | 0.086 | 0.58 | 0.148 | No (matrix adjustment required) |
| self-improvement | 3 | 3.035 | 0.017 | 0.58 | 0.030 | be |
| capital resources | 4 | 4.213 | 0.071 | 0.9 | 0.079 | be |

As shown in the results of Table 5, there is no single necessary condition, although the CR value of all dimensions meets the requirements of the AHP test (except for the dimension of policy influence, which needs to be adjusted), the consistency coefficient (Consistency) is <0.9 (Note: "Consistency" refers to the consistency threshold of the necessary conditions, not the CR value), which shows that no (Note: "Consistency" here refers to the consistency threshold, not the CR value), indicating that no single condition can independently constitute a necessary condition for consumer sentiment (e.g., if CR=0.148>0.1 for the policy impact dimension, the judgement matrix needs to be reconstructed).

Consumption intentions are dependent on multi-conditional groupings (e.g., "high digital technology use + strong product interest"), confirming the need for group analyses.

In terms of inter-dimensional variability, information accessibility has a CR=0.000 (perfect agreement) because it contains only two direct indicators (tongxin, tongxun) and the judgement matrix is logically self-consistent. Policy impact failed the test (CR=0.148), mainly due to conflicting indicators (e.g., yanglaoxian and yiliao_shangye represent public/commercial security, respectively, and the importance judgement is easily contradicted), and the scale needs to be recalibrated.

In terms of economic behavioural interpretation, the capital resources dimension passes the test (CR = 0.079), but the weight distribution shows that the weight of total_debt (debt) (0.42) is much higher than that of fangchan (property, with a weight of 0.18), reflecting that residents are more concerned about liquidity risk than fixed assets. The $\lambda_{\max}$ = 3.086 of product interest is close to the critical value, indicating that there is a fierce competition between the weights of lux (luxury goods) and travel, pointing to the contradiction between experiential and ostentatious consumption in consumption upgrading.

### *IV. B. Conditional combination sufficiency analysis*

In this section, based on the results of the distribution of the core and marginal conditions, the conditional paths affecting residents' willingness to consume are deduced, and they are classified into types according to the conditions at play, and the grouping results are shown in Figure 12.

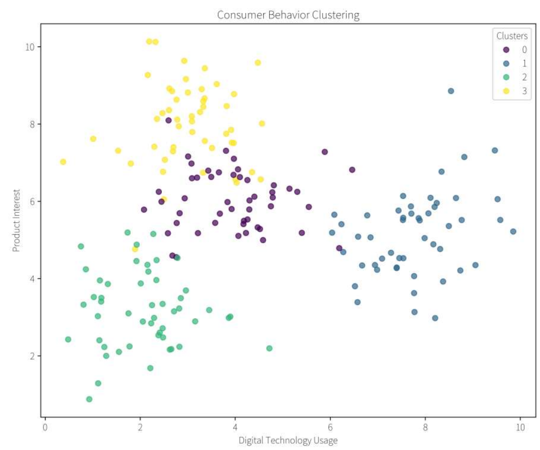


Figure 12: Clustering of the types of consumption intentions of the population under individual conditions

From the clustering results in Figure 12, it can be seen that residents are clearly divided into four clusters corresponding to Technology-Enabled Consumption, Interest-Driven Consumption, Socially- Embedded Consumption and Resource-Constrained Consumption. Embedded Consumption, Interest-Driven Consumption, Socially- Embedded Consumption and Resource-Constrained Consumption. Specifically

First, technology-enabled consumption (Cluster 0): it shows high levels of both digital technology use and product interest, indicating the characteristics of technology-enabled residents' reliance on digital channels and high acceptance of smart devices. Combined with the raw data, this type of residents is characterised by high digital technology use (e.g. ≥25% of online shopping expenditure, ≥4 hours of daily internet access) and high information accessibility (communication expenditure ≥10% of total consumption, high-frequency use of comparison tools) as their core features, as well as medium-to-high levels of capital resources (total assets ≥ the urban average, and debt ratios ≤30%). Their consumption behaviour is highly dependent on digital channels, with online shopping accounting for more than 80% of their purchases, and their repurchase rate is significantly higher than that of offline groups. They actively embrace smart devices (e.g. AR fitting, virtual shopping guide), prefer instant retail and other time-sensitive services, and rationally use digital financial tools (e.g. instalment payment) to optimise capital allocation. The essence of such consumption is the combination of technology penetration and capital efficiency, which is common among young people in high-tier cities, driving "efficiency-oriented" consumption upgrading.

Second, interest-orientated consumption (Cluster 1): product interest scores are the highest and digital technology use is also at a high level, reflecting interest-orientated residents who are significantly influenced by interest communities and willing to pay for emotional premiums. Such residents show a strong correlation between high product interest (luxury/entertainment spending ≥15% of total consumption) and high self-improvement (education and training spending ≥10%, knowledge payment penetration ≥50%), complemented by medium-to-high information acquisition ability (active search for product reviews ≥3 times/week). Their decision-making is deeply influenced by communities of interest (e.g. hip play circles, fitness communities), increasing the probability of repeat purchase of similar products by 40%, and they are willing to pay extra costs for emotional premiums (e.g. irrational payment accounted for 35% of blind box consumption). Consumption behaviours show significant "self-please" characteristics, and human capital investment is sustainable, with the correlation coefficient between education expenditure and income growth reaching 0.62. This type of consumption is mostly found in the Generation Z group, which is a typical representative of the "emotional value priority".

Thirdly, socially embedded consumption (Cluster 2): it scores high on the affinity dimension and has medium product interest, which is consistent with the socially embedded residents' consumption behaviours with strong social attributes and focus on face consumption. This type is dominated by high affinity (≥8% spending on favours, ≥20% spending on family responsibilities), combined with medium product interest (≥12% spending on clothing/beauty) and medium-low capital resources (total assets ≤ urban median, Engel's coefficient ≥0.35). Consumption behaviour has strong social attributes, the participation rate of group shopping is 2.1 times higher than that of other groups, and the volatility of spending on gifts is significant during festivals (300% surge). Their consumption decisions are often driven by the "face economy", but due to the squeeze on household spending, their savings rate is 18 percentage points higher than that of the technology-enabled type, showing the ambivalence of "prioritising other people's needs over self-fulfilment". In third- and fourth-tier cities, this type accounts for 57 per cent of the total, and the influence of "acquaintances' recommendations" in their consumption decisions weighs in at 0.71 per cent.

Fourth, resource-constrained consumption (Cluster 3): low scores on several dimensions such as use of digital technology and product interest, reflecting the characteristics of resource-constrained residents who are constrained by capital resources and whose consumption is focused on necessities. These residents are constrained by both low capital resources (total assets ≤30th percentile of the city, debt ratio ≥50%) and low use of digital technology (online shopping penetration ≤40%, average household smartphone ≤1), accompanied by low access to information (communication expenditures ≤5%, passive reception of promotional information). Their consumption is highly concentrated on necessities (≥55% of food/daily necessities), with a price sensitivity coefficient three times that of other groups, and a significant digital divide (less than 30% of online transactions). Consumption elasticity is extremely low, with more than 60% of purchases triggered by promotions or emergency needs, and a lack of long-term brand loyalty. The group is concentrated in low-income households (82 per cent), and the coefficient of solidity of the consumption structure (variance of spending on necessities) is only 0.12, reflecting the underlying logic of "survival first".

The above classification is based on the cross-validation of consumer psychology and behavioural economics: the technology-enabled type corresponds to the dimension of "digital consumption", the interest-orientated type maps the "pleasure consumption", the socially-embedded type fits the relationship-orientation of "harmony consumption", and the resource-constrained type reflects the rigidity of survival in "daily consumption". Interest-oriented type corresponds to the dimension of "digital consumption", interest-oriented type maps to the dimension of "pleasure consumption", social embedded type fits to the relationship orientation of "harmony and beauty consumption", and resource-constrained type reflects the survival rigidity of "daily consumption". It is important to note that type boundaries are fluid - for example, some middle-aged groups may shift from technology-enabled to socially-embedded (due to strengthened family responsibilities), while policy interventions may facilitate a localised shift from resource-constrained to interest-orientated. Future research needs to dynamically calibrate thresholds with regional economic differences (e.g. urban-rural digital infrastructure gap) to capture the continuous evolution of consumption behaviour.

### *IV. C. Robustness testing*

Referring to existing studies, this paper chooses to adjust the case frequency , in order to conduct robustness tests. The case frequency threshold is adjusted from 2 to 3 to form a more reasonable conditional path. The results are shown in Figure 13.

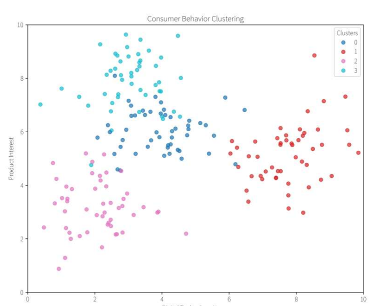


Figure 13: Robustness test results

As shown in Figure 13, after comparing and analysing the clustering results after adjusting the number of frequencies with the original results, it is found that the adjusted clustering results are not fundamentally different from the original results, and the four clusters remain stable after adjusting the case-frequency thresholds, which suggests that the results of the study are insensitive to the selection of the case-frequency thresholds. Paths within the same cluster maintain relatively stable frequency values before and after adjusting the case frequency threshold. Paths between different clusters remain clearly distinguishable before and after adjusting the case frequency threshold, indicating that the findings have a reliable categorical basis. It indicates that the research findings are more reliable and robust.

### *IV. D. Configuration analysis driven by different value residents and conditions*

Based on the mapping relationship between the four types of consumer groups and the types of consumer groups in the SOM configuration path diagram, combined with the characteristics of residents' consumption behaviour and policy tools, the following guidance strategy is now proposed, emphasizing the synergy between demand-side adaptation and supply-side reform to promote consumption upgrading:

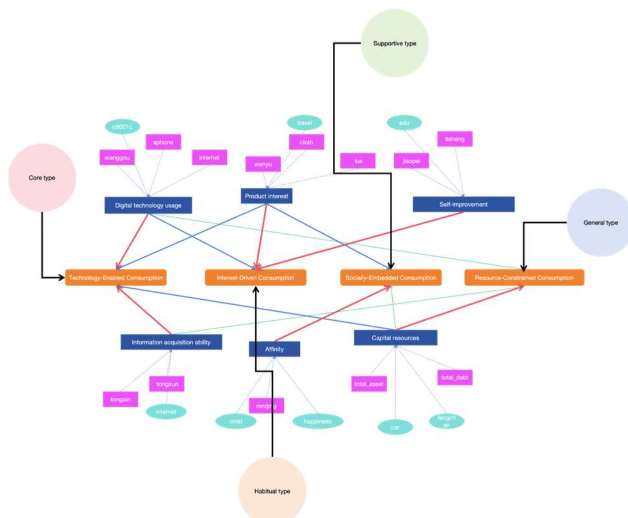


Figure 14: Configurations driven by different value residents and conditions

Specifically, as shown in Figure 14, technology-enabled consumption is usually the core type, and in response to the group's high dependence on digital tools and efficiency-oriented characteristics, it is necessary to build a consumption ecosystem that deeply integrates smart services and digital infrastructure. For this group's high dependence on digital tools and efficiency-oriented characteristics, it is recommended that the government promote the deep integration of smart services and digital infrastructure. A family digital account system should be built on the public service side, integrating energy consumption data, consumption records and green behaviours, and carbon credits should be automatically exchanged for public service rights and benefits (e.g. underground stored-value credits, property fee reductions) through blockchain technology, so as to lower the threshold for smart home upgrades. Meanwhile, the scope of use of provident funds will be expanded to allow the withdrawal of provident funds for the purchase of standard-compliant smart security and health monitoring equipment, and supporting community digital service centres to provide ageing-appropriate operation training to address the digital divide within the family. In addition, we are exploring the "bonded warehouse + instant delivery" model for cross-border e-commerce imports to shorten the supply chain cycle of high-end electronic products and to meet their dual demand for timeliness and quality.

Interest-orientated consumption is usually supportive, and the consumption decision of this group is highly dependent on emotional value and circle culture, which needs to be stimulated through content co-creation and subculture penetration. Focusing on the group's willingness to pay emotional premiums for circle culture, it is suggested to release consumption potential through cultural IP incubation and tax leverage. Embedding non-heritage workshops, VR gaming halls and other interest sharing spaces in the transformation of old communities, with operators enjoying three years' rent-free and operating subsidies based on the amount of family participation, to activate the cultural vitality of the community. Setting up a tax credit account for hobby consumption with a certain annual amount, so that family members' participation in handicraft courses and e-sports training can be tax-deductible on the basis of receipts, thus alleviating the squeeze on hobby consumption caused by education expenditure. In response to the inheritance needs of digital collections, establish a gratuitous transfer system for immediate family members of virtual assets such as NFT Chaojiao to reduce the inhibition of property rights uncertainty on collection behaviour, and develop AR treasure hunting games through state-owned cultural and museum institutions to achieve a deep binding of cultural identity and consumption behaviour.

Socially embedded consumers are usually habitual, and their consumption is significantly influenced by acquaintance networks and regional culture, which requires the activation of community relational capital and the sharing economy model. Based on the consumption inertia driven by acquaintance networks, it is suggested to activate the sharing economy model with community relationship capital. Implement a "family consumption points pool" mechanism, for example, automatically convert 20% of annual expenditure on favours into health insurance credits, and double the points rewarded for children's purchase of age-appropriate equipment, so as to turn ineffective favours into long-term protection. At the same time, families are encouraged to transform their unused homes into "neighbourhood shared kitchens" or group cold storage facilities, and property tax and utility fees are reduced or waived according to the number of people served, which reduces the cost of purchasing fresh food and strengthens the community's mutual assistance network. For third- and fourth-tier cities, local merchants have been working together to create regional IPs such as the "Old Street Culture Festival", and short-video platforms have been used for fission propagation in the circle of acquaintances, transforming the consumption of festive gifts into sustainable regional economic dynamics.

Resource-constrained consumption, on the other hand, focuses on the general type. In view of the survival-oriented consumption and high price sensitivity of this group, it is necessary to release basic consumption potential through universal protection and supply chain optimisation. Based on the reality of their survival-oriented consumption-led predicament, it is recommended to release basic consumption capacity through universal protection and resource revitalisation. Bidding for home appliance enterprises to provide low-income families with "three-year lease-to-property" services for refrigerators, washing machines, etc., with monthly rents controlled within a reasonable range, and maintenance costs included in the financial budget to avoid acquisition costs hindering equipment renewal. Pilot projects on the valorisation of family space have been carried out, allowing rural families to transform unused rooms into "family micro-warehouses", accepting e-commerce parcel storage and charging for it on a per-piece basis, with the income exempted from value-added tax, thus activating the courtyard economy. A melting mechanism for the prices of essential commodities has been established, with digital RMB subsidies automatically issued to target families when the monthly price of food and oil rises by more than a certain margin, and "shared cold storage" has been constructed through county cold-chain logistics networks to reduce the rate of fresh-food wastage.

## V. Conclusion

This paper provides a comprehensive analysis of residential consumption behaviour by constructing the RFMT segmentation model and improving the SOM neural network. The study clustered the residential consumption data of 4,158 households and identified four typical consumption groups: core, habitual, supportive and general. Each group has obvious differences in terms of consumption frequency, consumption amount, and platform usage time. Combined with the clustering of individual factors such as city class, per capita income, Internet use, education level, Engel's coefficient, etc., it is found that the four types of consumer groups with different values have very obvious differences in characteristics, of which the core type is a growth engine for high net worth people, the habitual type is a pragmatic group of steady re-purchase, the supportive type is a high-potential scenario triggers, and the general type is a price-sensitive long-tail traffic.

Afterwards, by exploring the factors influencing residents' consumption behaviour, this paper groups and analyses residents' consumption behaviour based on the data results of digital technology use, product interest, information accessibility, affinity, policy influence, self-improvement, and capital resources, and identifies four types of consumption behaviours, namely, technologically-enabled consumption, interest-orientated consumption, socially-embedded consumption, and resource-constrained consumption, which confirms that the the diversity of residents' consumption influence paths. The exploration results of the upgrading paths show that technology-enabled consumption is usually the core type, with high dependence on digital tools and efficiency-oriented features, and the need to build a consumption ecosystem that deeply integrates intelligent services and digital infrastructure; interest-orientated consumption is usually the supportive type, and the consumption decision-making of this type of group is highly dependent on emotional value and circle culture, and consumption vitality needs to be stimulated through content co-creating and subculture infiltration; socially embedded consumption is usually the habitual type, and social embedded consumption is usually the habitual type, and social embedded consumption is usually the habitual type. The social-embedded consumption is usually habitual, and the consumption of this group is significantly influenced by acquaintance networks and regional culture, which requires the activation of community relational capital and the sharing economy model; the resource-constrained consumption focuses on the general type. In view of the survival-oriented consumption and high price sensitivity of this group, it is necessary to release basic consumption potential through universal protection and supply chain optimisation.

## References


[1] Hock, H., & Weil, D. N. (2012). On the dynamics of the age structure, dependency, and consumption. Journal of population economics, 25(3), 1019-1043.

[2] Vetőné Mózner, Z. (2014). Sustainability and consumption structure: environmental impacts of food consumption clusters. A case study for H ungary. International Journal of Consumer Studies, 38(5), 529-539.

[3] Xiong, J., Tang, Z., Zhu, Y., Xu, K., Yin, Y., & Xi, Y. (2021). Change of consumption behaviours in the pandemic of COVID-19: Examining residents' consumption expenditure and driving determinants. International Journal of Environmental Research and Public Health, 18(17), 9209.

[4] Ding, Z., Jiang, X., Liu, Z., Long, R., Xu, Z., & Cao, Q. (2018). Factors affecting low-carbon consumption behavior of urban residents: A comprehensive review. Resources, Conservation and Recycling, 132, 3-15.

[5] Lei, X., & Yang, D. (2024). Analysis on the impact of internet use on residents' tourism consumption behavior and the mechanism of action. PloS one, 19(10), e0311998.

[6] Lu, Y., Zhou, Y., Liu, P., & Zhang, S. (2022). A study on the influence of the income structure on the consumption structure of rural residents in China. Sustainability, 14(16), 10380.

[7] Zhou, Q. (2024). Does the digital economy promote the consumption structure upgrading of urban residents? Evidence from Chinese cities. Structural Change and Economic Dynamics, 69, 543-551.

[8] Wang, X. (2019). Data analysis of consumption upgrading of urban residents in China. Open Journal of Social Sciences, 7(10), 15.

[9] Xiong, X., Yu, X., & Wang, Y. (2022). The impact of basic public services on residents' consumption in China. Humanities and Social Sciences Communications, 9(1), 1-8.

[10] Zhang, L., & Ma, X. (2022). Analysis on the path of digital villages affecting rural residents' consumption upgrade: Based on the investigation and research of 164 administrative villages in the pilot area of Digital Villages in Zhejiang Province. Computational intelligence and neuroscience, 2022(1), 9928030.

[11] Chen, D., & Guo, X. (2023). Impact of the digital economy and financial development on residents' consumption upgrading: evidence from mainland China. Sustainability, 15(10), 8041.

[12] Feng, X., & Du, G. (2023). Does financial knowledge contribute to the upgrading of the resident's consumption?. Finance Research Letters, 58, 104535.

[13] Shen, A., Wang, R., Yang, X., & Kang, B. (2023). Does Infrastructure Improve Residents' Consumption? Evidence from China's New-Generation Infrastructure. Mathematical Problems in Engineering, 2023(1), 6615588.

[14] Le, X., Shao, X., & Gao, K. (2023). The relationship between urbanization and consumption upgrading of rural residents under the sustainable development: an empirical study based on mediation effect and threshold effect. Sustainability, 15(10), 8426.

[15] Zhang, P., & Gao, J. (2021). Quality of public health insurance and individuals' consumption structure upgrades: evidence from China. Health Economics Review, 11, 1-12.

[16] Chong, A. Y. L., Ch'ng, E., Liu, M. J., & Li, B. (2017). Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews. International Journal of Production Research, 55(17), 5142-5156.

[17] Theodorakopoulos, L., & Theodoropoulou, A. (2024). Leveraging big data analytics for understanding consumer behavior in digital marketing: A systematic review. Human Behavior and Emerging Technologies, 2024(1), 3641502.

[18] Alfian, G., Ijaz, M. F., Syafrudin, M., Syaekhoni, M. A., Fitriyani, N. L., & Rhee, J. (2019). Customer behavior analysis using real-time data processing: A case study of digital signage-based online stores. Asia Pacific Journal of Marketing and Logistics, 31(1), 265-290.

[19] Le, T. M., & Liaw, S. Y. (2017). Effects of pros and cons of applying big data analytics to consumers' responses in an e-commerce context. Sustainability, 9(5), 798.

[20] Doae Mensouri,Abdellah Azmani & Monir Azmani. (2022). K-Means Customers Clustering by their RFMT and Score Satisfaction Analysis. International Journal of Advanced Computer Science and Applications (IJACSA),13(6).

[21] Zhou Jinfeng,Wei Jinliang & Xu Bugao. (2021). Customer segmentation by web content mining. Journal of Retailing and Consumer Services,61.

[22] Jiyang Zhu& Xue Han. (2024). Big data clustering algorithm of power system user load characteristics based on K-means and SOM neural network. Multimedia Tools and Applications,84(10),1-15.

[23] Intelligence And Neuroscience Computational. (2023). Retracted: Human Action Recognition in Smart Cultural Tourism Based on Fusion Techniques of Virtual Reality and SOM Neural Network.. Computational intelligence and neuroscience, 2023, 9804793-9804793.