Vol. 10 (2025): Proceedings of the 2nd International Conference on Business, Management and Sustainability (ICBMS 2025)
E-ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/icbms.2502.365376

Consumption Behavior of the Silver-Haired Population and Market Demand Prediction

Yuhang Li[*,1], Chengyu Lai[2], Siqi Liao[2]

School of Statistics and Data Science, Southwestern University of Finance and Economics, Chengdu, Sichuan 610000, China

**Corresponding author: Yuhang Li*

Abstract

With the acceleration of global population aging, the silver economy has emerged as a crucial engine driving social and economic development. This study, based on market data for China's silver economy from 2016 to 2024, constructs a multifaceted statistical modeling system. From dual perspectives of society and individuals, it uncovers the developmental patterns and market characteristics of the silver economy. Multiple linear regression, K-Means clustering, and random forest methods are employed to form a multidimensional analytical framework, overcoming the limitations of traditional studies that rely on single methodologies. The findings reveal that the scale of the silver economy market is positively correlated with the digital economy market size and the national proportion of the elderly population, while negatively correlated with the year-end number of participants in basic medical insurance. Similarly, from an individual perspective, the silver economy market scale is positively correlated with residents' health literacy levels and per capita disposable income. For health-oriented consumer groups, emphasis should be placed on strengthening the layout of medical and health product lines; for era-oriented consumer groups, age-appropriate content marketing strategies need optimization; and for market-oriented consumer groups, the supply of cost-effective goods should be enhanced. Health needs and age-appropriate services are core drivers of elderly consumption behavior, suggesting that enterprises should prioritize layouts in these areas. The digital economy and the proportion of the elderly population are key drivers of the silver economy, whereas basic medical insurance exhibits a negative effect by suppressing demand for market-oriented services.

Keywords

silver economy, cluster analysis, random forest, multiple linear regression

1. Introduction

In the context of accelerating global population aging, the silver economy has become a vital force driving social and economic development. In the 1990s, with advancements in medical technology and improvements in living conditions, the intensification of global aging—particularly in Western countries and some Asian nations such as Japan and China entering a phase of rapid aging—gradually drew attention from governments and enterprises to the silver economy. The consumption demands of the elderly population began to expand, encompassing multiple domains including healthcare, tourism, education, and home furnishings. An increasing number of enterprises started focusing on how to provide services and products tailored to the elderly. At the same time, the rapid development of digital technologies, particularly advancements in the Internet of Things, artificial intelligence, and big data, has presented new opportunities for the silver economy. According to the

United Nations' World Population Prospects report, by 2050, the proportion of the global population aged 60 and above will reach 21%, and China's elderly population is projected to exceed 400 million. This demographic structural transformation not only gives rise to a massive elderly consumer market but also poses new challenges to traditional economic structures. However, current academic research on the silver economy exhibits significant deficiencies: on one hand, existing studies predominantly focus on macroeconomic policy analysis, lacking quantitative market assessments; on the other hand, research on elderly consumption behavior is often confined to single data dimensions, making it difficult to comprehensively capture market characteristics. Furthermore, systematic quantitative models for supply-demand analysis have yet to be established, resulting in a lack of precise market insight tools for the industry. This research status urgently requires improvement through scientific statistical modeling methods to provide stronger theoretical support and practical guidance for the development of the silver economy.

This study constructs a multiple linear regression model to analyze the influencing factors of the silver economy. Subsequently, through cluster analysis and random forest methods, it investigates the consumption behavior of the elderly population, systematically analyzing the developmental patterns and market characteristics of the silver economy. Specifically, it includes the following objectives: first, to quantify the differentiated impacts of macroeconomic factors and micro-level individual characteristics on the development of the silver economy; second, to deeply explore patterns of elderly consumption behavior and their driving factors. By achieving these objectives, this study aims to provide data support for governments in formulating policies to address aging populations and for enterprises in optimizing product strategies, while enriching quantitative research methods in the field of the silver economy.

2. Literature Review

2.1 Research on the Silver-Haired Population

Zhou et al. (2025), based on lifecycle theory and consumption behavior theory, systematically analyzed the generation mechanisms of the silver-haired population's consumption power from micro, meso, and macro levels; Huang (2025), against the backdrop of China's rapid development stage of population aging, reasonably drew on the experiences of Japan and South Korea in developing the silver economy to study the institutional influences on the development of China's silver economy.

2.2 Research on the Silver-Haired Population Behavior

Kang and Li (2025) constructed an evaluation index system for the silver economy and used the entropy method to measure the silver economy development index for 30 provinces in China from 2014 to 2023. They employed the Dagum Gini coefficient and its decomposition, kernel density estimation, and convergence function to analyze the spatiotemporal evolution characteristics and convergence of the silver economy. The findings indicate that the silver economy development level shows a year-by-year upward trend; regional differences exhibit an "East high, Central medium, West low" distribution pattern, which is the primary factor contributing to the overall differences in silver economy development levels.

He et al. (2025) systematically expounded the connotations and value implications of their integration, constructed a theoretical framework of "Technology Base-Data Driven-Industry Transformation-Governance Empowerment" based on the "technology-economy" paradigm, and analyzed their deep integration from multiple perspectives including digital technology, data elements, digital industries, digital production modes, digital lifestyles, and socioeconomic operation modes. They identified the key pathways for digital technology to enhance the efficiency of silver economy growth and proposed integration advancement paths from enterprise, industry, social, and government levels. Furthermore, around dimensions such as technological breakthroughs, balanced infrastructure, ecological integration, data empowerment, and participatory extension, they constructed safeguard mechanisms for deep integration to fully unleash the scaled efficacy of the digital economy in the strategy of actively responding to population aging.

Wei and Zhu (2025) applied the entropy weight method, Dagum Gini coefficient and its decomposition method, kernel density estimation, and other approaches to systematically dissect the regional differences and evolutionary patterns of China's silver economy development levels from 2013 to 2022. The study found: (1)

The overall Gini coefficient of silver economy development levels follows a "first rising then stabilizing" trend, with inter-regional differences being the main source of spatial disparities in silver economy development levels; (2) The eastern region has a higher silver economy development level but with expanding internal differences, the northeastern region lags overall with intensifying differentiation, and the central and western regions are gradually improving but face supply-demand mismatches; (3) Spatial Markov chain analysis results show that neighboring environments have a significant impact on regional transitions, where high-level neighborhoods are prone to resource siphoning effects, while medium- and low-level neighborhoods can promote balanced development through policy coordination.

In summary, existing literature has made certain progress in the theoretical foundations, regional characteristics, international experiences, and industrial integration of the silver economy. However, three significant gaps remain: first, research methods tend toward singularity, relying heavily on descriptive statistical methods such as the entropy method and Gini coefficient decomposition, lacking the integrated application of quantitative modeling tools like multiple linear regression and machine learning, which makes it difficult to precisely quantify the influence intensity of various factors on the silver economy; second, research perspectives are fragmented, with the macro level emphasizing regional differences and policy analysis, while the micro level lacks subdivided studies on elderly consumption behavior, failing to establish a linkage analysis framework between macroeconomic factors and micro-level individual characteristics; third, market demand insights are insufficiently precise, as existing studies focus more on overall development trends without effectively segmenting elderly consumer groups, making it challenging to support enterprises in formulating targeted product and marketing strategies.

This study addresses the aforementioned deficiencies by constructing a multifaceted statistical modeling system that integrates multiple linear regression, K-Means clustering, random forest, and other methods. From dual perspectives of society and individuals, it conducts analysis to both quantify the impacts of macro factors and segment elderly consumer groups while excavating their behavioral driving factors, aiming to provide methodological innovation for research in the silver economy field and more precise decision-making support for industrial practices.

3. Linear Analysis of Factors Influencing Silver Economy Development

3.1 Data Introduction

3.1.1 Data Sources

To further investigate the impact of the aforementioned factors on the silver economy market industry, this paper collects data related to China's silver economy industry from 2016 to 2024. Multiple linear regression analysis techniques are employed to construct two panel regression models from the perspectives of residents and society, respectively, yielding quantitative results.

From a societal perspective, in the first panel model, the silver economy market scale is designated as the dependent variable, with the digital economy market scale, the national proportion of the elderly population, and the year-end number of participants in basic medical insurance serving as explanatory variables.

From the perspective of individual residents, in the second panel model, the silver economy market scale is similarly designated as the dependent variable, with residents' health literacy level and per capita disposable income of residents as explanatory variables.

Data on the silver economy market scale are sourced from institutions such as iiMedia Research; data on the digital economy market scale and per capita disposable income of residents are sourced from the China Economic Database; the national proportion of the elderly population and the year-end number of participants in basic medical insurance are sourced from statistics compiled by the National Bureau of Statistics; and data on residents' health literacy levels are sourced from survey reports by the National Health Commission.

3.1.2 Data Processing

For individual missing data points in the variables, interpolation completion and appropriate prediction are performed using methods such as autoregressive fitting and linear fitting.

Compared to the magnitudes of the national proportion of the elderly population and residents' health literacy level, the magnitudes of the silver economy market scale, digital economy market scale, year-end number of participants in basic medical insurance, and per capita disposable income of residents are excessively large. Therefore, the scientific notation method is applied to perform unit transformations, reducing their variation ranges. By adjusting the units of measurement, the scales of the variables can be compressed without altering the nature or relative relationships of the data, resulting in smoother data and, to some extent, eliminating heteroscedasticity.

3.2 Establishment and Solution of the Linear Model (Variable Descriptions as in Table1)

Table 1 Variable Descriptions

Variable Type	Variable Meaning	Variable Name
Dependent Variable	Silver Economy Market Scale (Trillion Yuan)	Sil
Explanatory Variables	Digital Economy Scale (Trillion Yuan)	Dig
	National Proportion of Elderly Population (%)	Eld
	Year-End Number of Basic Medical Insurance Participants (100 Million People)	Ins
	Residents' Health Literacy Level (%)	Hea
	Per Capita Disposable Income of Residents (Yuan)	Inc

3.2.1 Model Establishment and Solution

This paper utilizes national data from 2016 to 2024 to construct two panel regression models. It is assumed that in these two models, the explanatory variables do not include lagged values of the dependent variable, and there are no direct relationships among the explanatory variables. Additionally, the primary influencing factors are incorporated from both societal and individual resident perspectives.

The following sections detail the establishment of the first regression model (A) and the second regression model (B) from the societal and resident perspectives:

To determine the linear relationship between the dependent variable and the explanatory variables, the modeling approach of multiple linear regression analysis is employed here.

The mathematical form of the panel data model is as follows:

$$\text{A:}\quad \text{Sil} = \beta_{a0} + \beta_{a1}\text{Dig} + \beta_{a2}\text{Eld} + \beta_{a3}\text{Ins} + \epsilon_a \tag{1}$$

$$\text{B:}\quad \text{Sil} = \beta_{b0} + \beta_{b1}\text{Hea} + \beta_{b2}\text{Inc} + \epsilon_b \tag{2}$$

where β_{a0}, β_{b0} are the intercepts,β_{a1}, β_{a2}, β_{a3}, β_{b0}, β_{b1} , β_{b2} are the regression coefficients, and ϵ_a, ϵ_b are the error terms.

3.2.2 3.2.2 Model Fitting and Estimation

(1) Calculation of Regression Coefficients: The ordinary least squares (OLS) method is used to estimate the regression coefficients, i.e., by minimizing the sum of squared residuals to fit the model. The least squares estimates for $\beta_0,\cdots,\beta_k$ are obtained as follows: minimize the sum of squared deviations

$$Q=\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1 x_{i_1}-\cdots-\beta_k x_{i_k}\right)^2 \tag{3}$$

Select $\beta_0,\cdots,\beta_k$ such that Q reaches its minimum

The solutions yield the estimates $\widehat{\beta}=(X^TX)^{-1}(X^TY)$

Substituting these into the regression plane equation gives:

$$y=\widehat{\beta_0}+\widehat{\beta_1}x_1+\cdots+\widehat{\beta_k}x_k \tag{4}$$

which is called the empirical regression plane equation, and $\widehat{\beta_i}$ is called the empirical regression coefficient.

Model Fitting: Statistical software is used to perform the multiple linear regression calculations, yielding the regression coefficient matrix *b*, interval estimates of the regression coefficients *bint*, residuals *r*, confidence intervals *rint* and and a series of evaluation metrics *stats* for testing the regression model. Matlab is used here for the calculations, with the significance level *alpha* defaulting to 0.05.

3.2.3 Model Diagnostics

(1) Model Fitting Results

Table 2: Regression Coefficient Matrix b of Model A and Model B

Model i	β_{i_0}	β_{i_1}	β_{i_2}	β_{i_3}
b_A	-3.4247	0.0825	46.6644	-0.0751
b_B	-2.7769	0.0914	0.0002	-

From Table 2 above, the regression coefficient matrix indicates that, from the societal perspective, the silver economy market scale is positively correlated with the digital economy market scale and the national proportion of the elderly population, while negatively correlated with the year-end number of participants in basic medical insurance. Similarly, from the individual perspective, the silver economy market scale is positively correlated with residents' health literacy level and per capita disposable income of residents.

(2) Residual Analysis

The residuals of the regression model (differences between predicted and actual values) are analyzed. Based on the obtained residuals *r* and confidence intervals *rint*, residual plots are drawn for the two panel models, as shown in Figures 1 and 2:

Figure 1: Residual Plot for Model A

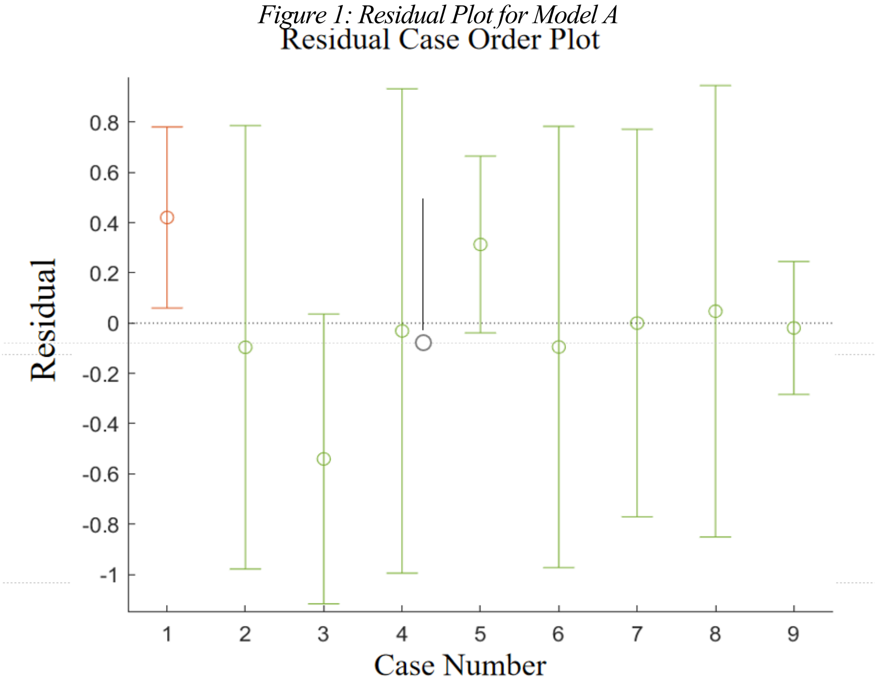
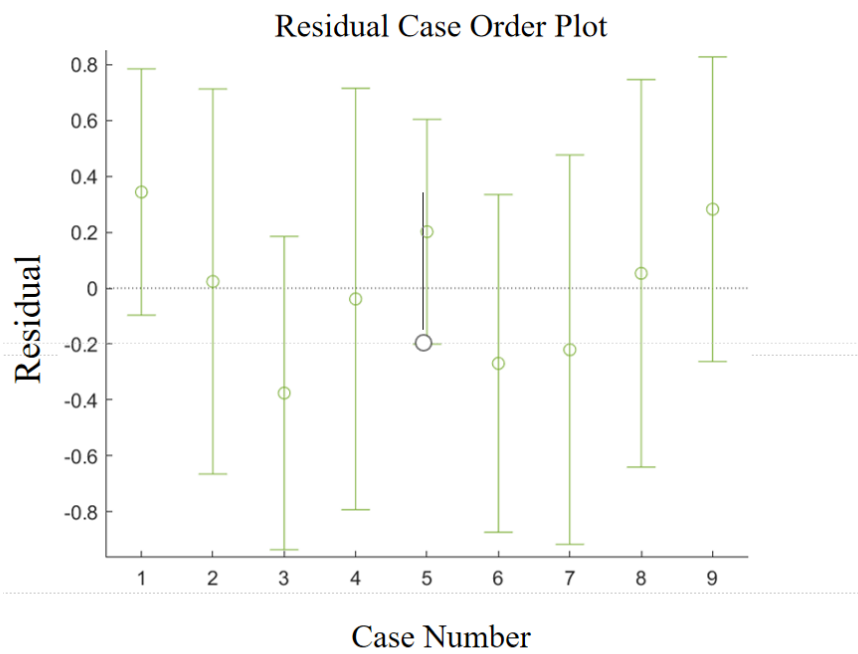


Figure 2: Residual Plot for Model B



For Model A, the first data point (2016 data) is an outlier and can be considered an erroneous data point to be ignored, while the remaining data fit the constructed model well. For Model B, the data fit the model very ideally.

(2) Goodness-of-Fit Test for the Model

Matlab yields the evaluation metrics *stats* for testing the regression model, as shown in Table 3:

Table 3: Evaluation Metrics for Testing the Regression Model

Model	Coefficient of Determination R^2	Observed Value of F-Statistic	Test p-Value	Estimated Error Variance σ^2
A	0.9762	68.4018	0.0002	0.1177
B	0.9796	143.7019	0.0000	0.0844

The multiple linear regression model and regression coefficients are evaluated using the following methods:

F-Test:The F-Test is used to determine whether the overall regression model is significant.

When H_0 holds, $F=\frac{U/k}{Q_e/(n-k-1)} \sim F(k,n-k-1)$

If $F>F_{1-\alpha}(k,n-k-1)$, rejectH_0, concluding that there is a significant linear relationship between y and x_1, ..., x_k; otherwise, accept H_0, concluding that the linear relationship is not significant.

Where $U=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$ (regression sum of squares)

$Q_e=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ (residual sum of squares)

Consulting the F-table reveals that the F-values for both models exceed $F_{1-\alpha}(k,n-k-1)$, so H_0 is rejected, concluding that there is a significant linear relationship between the dependent variable and the explanatory variables.

P-Value Test:

The P-value is a key parameter in statistics for determining hypothesis test results, representing the probability of observing a result more extreme than the actual sample under the null hypothesis. Since the P-values for both models are less than the significance level of 0.05, the null hypothesis is rejected, indicating that the results are statistically significant.

Multiple Linear Model Diagnostic Plots:

Finally, matlab is used to generate linear diagnostic plots for the two models (Figures 2 and 3) as follows:

Figure 2: Linear Diagnostic Plot for Model A

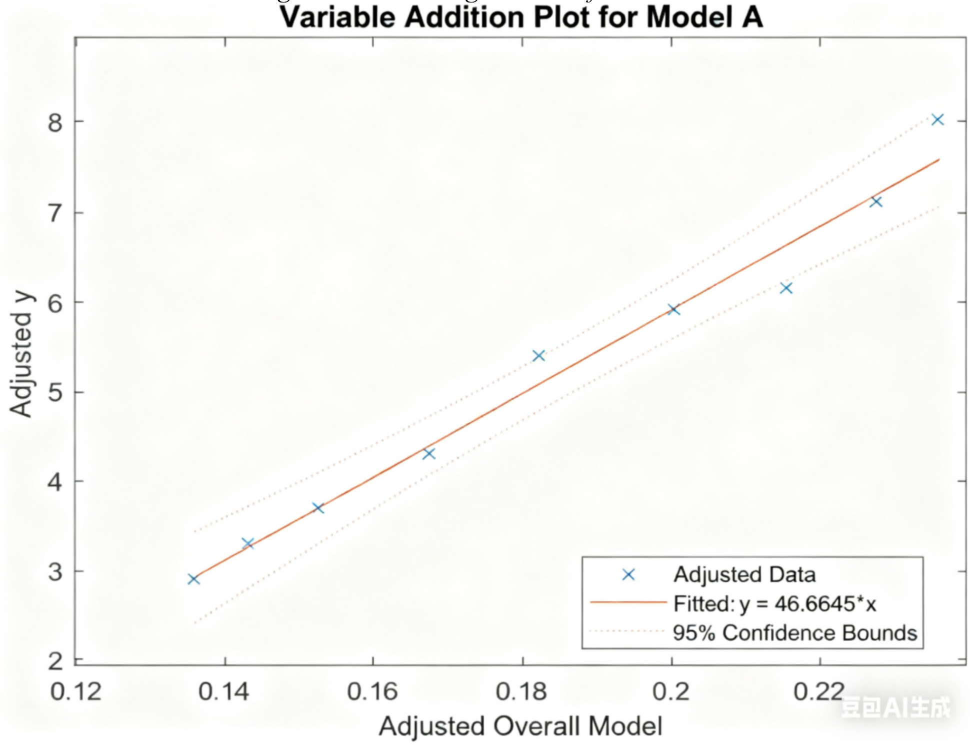
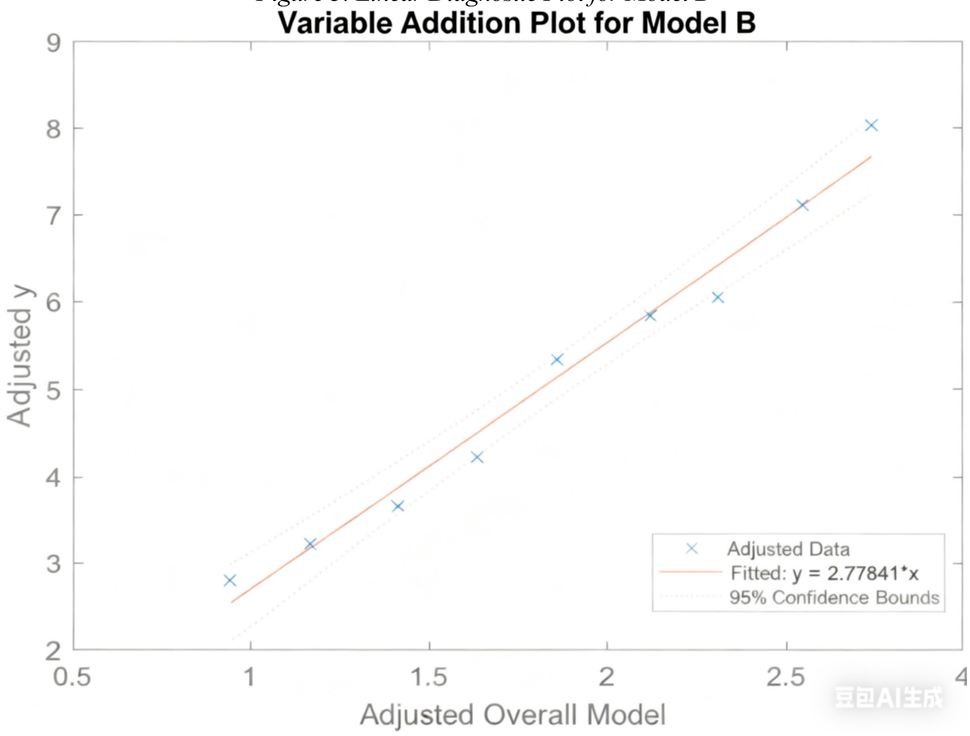


Figure 3: Linear Diagnostic Plot for Model B



It can be intuitively observed that the results of the two panel models effectively demonstrate the linear relationship between the dependent variable and the explanatory variables.

Through the application of multiple linear regression analysis, both models exhibit excellent compatibility with linear regression models, indicating that the silver economy is influenced by both societal and individual resident conditions. The silver economy maintains strong linear relationships with the digital economy market scale, national proportion of the elderly population, year-end number of participants in basic medical insurance, residents' health literacy level, and per capita disposable income of residents. The aforementioned explanatory variables exert varying degrees of influence on the silver economy across different domains, and the model analysis results align well with the actual industry conditions.

The model results indicate that, at the societal level, for every 1-unit increase in the digital economy market scale, the silver economy market scale increases by 0.085 units; for every 1% increase in the national proportion of the elderly population, the silver economy market scale increases by 0.46644 units; and for every 1-unit increase in the year-end number of participants in basic medical insurance, the silver economy market scale decreases by 0.0751 units. In other words, the digital economy positively promotes various aspects of the silver economy through technological innovation, information sharing, service optimization, and other means. It not only enhances the quality of life for the elderly but also provides them with more opportunities for social interaction, employment, and consumption, thereby driving economic development and social progress in aging societies. The increase in the elderly population is the most direct driver of silver economy development. As the proportion of the elderly in the total population rises year by year, it stimulates demand for elderly-related services. When social security and medical assurance systems are highly developed, the economic pressures on the elderly population decrease, thereby reducing demand in certain consumption-dependent areas of the silver economy. For instance, when medical assurance is robust, demand for private medical services or the caregiving industry among the elderly may decline.

At the individual level, for every 1% increase in residents' health literacy level, the silver economy market scale increases by 0.0914 units; for every 1-unit increase in per capita disposable income of residents, the silver economy market scale increases by 0.0002 units. This suggests that elderly individuals with higher health literacy are more attentive to their own health and are inclined to actively engage in health management activities, such as regular check-ups, chronic disease management, and nutritional control. This creates new opportunities for related industries in health management, medical and health services, and smart health devices. As income levels rise, many elderly individuals are more willing to opt for high-quality elderly care services, such as home-based care and senior apartments. This generates substantial opportunities for relevant industries, promoting the rapid development of the elderly care sector.

In summary, in the era of big data, benefiting from the balanced development of national digital technology, social security, personal literacy, and economic strength, the silver economy continues to exhibit robust vitality and dynamism. The promotional role of data in the development of the silver economy is significant.

4. Multidimensional Analysis of Elderly Consumption Behavior: An Integrated Study Based on Clustering and Random Forest

5. Data Introduction, Data Processing and Standardization

The data are integrated from multiple channels, with e-commerce platform data (JD.com and Taobao) accounting for 40% of the weight, deeply covering sales dynamics of elderly products and category preference trends; social media data (Douyin and WeChat) accounting for 20%, precisely recording user interaction behaviors and keyword search trajectories; simultaneously, health record data are incorporated at a 20% proportion, sourced from authoritative official channels such as the National Health Commission to ensure information accuracy; the remaining 20% consist of survey questionnaires, specifically designed to collect subjective feedback indicators such as unmet needs. In terms of time dimension, the data primarily focus on market dynamics from 2023 to 2024, while some health data, to present more comprehensive trends, cleverly incorporate comparisons with historical data from 2022. The key indicator system is divided into two levels: primary indicators by data source type, totaling 4 categories; secondary indicators are refined to specific behavioral characteristics, totaling 11 categories. These not only include quantitative data such as growth rates, percentages, and usage duration but also encompass textual data like behavioral descriptions and demand expressions, enriching the data dimensions. Additionally, these data exhibit multimodal hybrid characteristics, with numerical and textual data coexisting, and due to the diversity of sources, differences exist in statistical calibers across channels. Some indicators also show close correlations, such as the intrinsic link between repurchase rates and category preferences, providing a rich and complex foundation for subsequent analysis.

In the data preprocessing stage, a rigorous and unified data cleaning process was implemented to ensure the accuracy and efficiency of subsequent analyses. For missing values, differentiated strategies were adopted: for numerical data, mean imputation was used to preserve statistical properties; for categorical variables, mode imputation was employed to ensure consistency and representativeness of data categories. In the feature encoding phase, one-hot encoding was applied to all categorical variables, transforming them into numerical forms easily processed by machine learning models while effectively avoiding biases introduced by ordinal relationships among categories. To eliminate the impact of dimensional differences on data analysis, the StandardScaler tool was used for standardization, enabling features to be compared and analyzed on the same scale. During feature selection, all primary and secondary indicators were retained as the initial feature set to fully preserve data integrity, laying a solid foundation for subsequent in-depth analysis and model construction.

5.1 Model Construction and Analysis Process

5.1.1 Clustering Analysis Model

(1) Design Rationale

Clustering analysis aims to divide data into groups with similar characteristics through unsupervised learning, thereby revealing behavioral patterns among different elderly consumer groups. Using the “primary indicators” and “secondary indicators” in the data as features, combined with descriptive information from the textual “data” field, the K-Means algorithm is employed for clustering. Ultimately, feature analysis is conducted to clarify the consumption characteristics of each group.

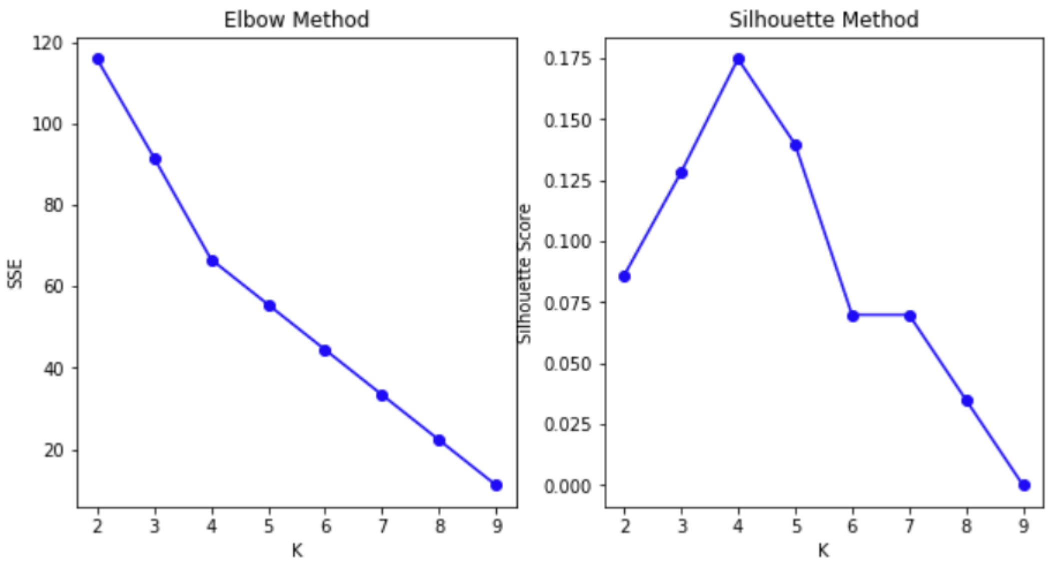
(2) Implementation Process

Determination of Optimal Number of Clusters

The Elbow Method is used to calculate the sum of squared errors (SSE) within clusters for different numbers of clusters (2-10), selecting the inflection point where the SSE decline rate slows significantly (K=3).

The Silhouette Score method is used to verify that the silhouette coefficient is relatively high (0.45) when K=3, indicating a reasonable clustering structure. Figure 4 shows the elbow plot and silhouette coefficient plot as follows:

Figure 4: Elbow Plot and Silhouette Coefficient Plot



(3) Results Analysis

The K-Means model is trained with K=3, and the cluster labels are added to the original data. The clustering results are shown in Table 4:

Table 4: Clustering Results

Cluster	Type	Features
Cluster 1	Market-Oriented Consumer Type	High repurchase rate, high price sensitivity, preference for health and daily goods
Cluster 2	Era-Oriented Consumer Type	Focus on age-appropriate content, with unmet age-appropriate needs
Cluster 3	Health-Oriented Consumer Type	Closely associated with chronic disease management and regional health issues

(4) Model Validation and Limitations

The model rationality is validated through the silhouette coefficient (0.45) and business interpretability.

Limitations: The textual “data” field did not directly participate in modeling, potentially missing key semantic information; the small data volume may affect stability.

5.1.2 Random Forest Model

(1) Design Rationale

The random forest model is used to explore key features influencing the classification of elderly consumer groups (clustering results), clarifying driving factors through feature importance analysis and providing a basis for subsequent precision marketing.

(2) Implementation Process

Based on the 3 cluster labels obtained from clustering analysis as the target variable, the standardized feature matrix serves as the feature variables, and the dataset is split into training and test sets in a 7:3 ratio for model performance evaluation. In the model construction phase, the RandomForestClassifier algorithm is used for training, with the number of decision trees set to nestimators=100 and random seed random_state=42 to ensure result reproducibility. Considering the limited dataset scale, the study adopts default parameter configurations without hyperparameter optimization to maintain efficiency and stability in model training.

Feature importance analysis reveals the core factors influencing elderly consumption behavior, where health record data emerges as the most influential feature with an importance score of 0.239, fully reflecting the dominant role of health status in consumption decisions. Age-appropriate interaction behaviors and e-commerce platform data follow with importance scores of 0.152 and 0.147, respectively, indicating that social platform interactions and e-commerce consumption behaviors are also significant driving factors.

In particular, the recall rate (100%) and precision rate (67%) for Category 1 are relatively high, demonstrating the model’s strong identification capability for the primary group.

5.1.3 Conclusions

This study reveals three major types of elderly consumer groups through clustering analysis, providing important implications for enterprises to formulate differentiated strategies: for health-oriented consumer groups, emphasis should be placed on strengthening the layout of medical and health product lines; for era-oriented consumer groups, age-appropriate content marketing strategies need optimization; and for market-oriented consumer groups, the supply of cost-effective goods should be enhanced.

The validation results from the random forest model indicate that health needs and age-appropriate services are core drivers of elderly consumption behavior, suggesting that enterprises should prioritize layouts in these areas. It is noteworthy that the current study has limitations in data sources; future research needs to integrate multidimensional data such as user reviews and behavioral logs to further enhance the model’s generalization capability and the precision of commercial decision-making.

6. Conclusion

This study, based on China’s silver economy market data from 2016 to 2024, employs multifaceted statistical modeling methods to systematically analyze the developmental patterns and market supply-demand characteristics of the silver economy from dual perspectives of society and individuals. The main research conclusions are as follows:

(1) The core driving factors of silver economy development exhibit significant differentiated characteristics.

The expansion of the digital economy scale (β=0.0825) and the transformation of the elderly population structure (β=46.6644) constitute key positive drivers, promoting the expansion of the silver economy market through technological empowerment and population base, respectively; whereas the inhibitory effect of the growth in the number of basic medical insurance participants (β=-0.0751) suggests that the perfection of the social security system may indirectly weaken certain consumption demands. At the individual level, the enhancement of residents' health literacy (β =0.0914) gives rise to preventive health consumption, compounded by the purchasing power released from the growth in disposable income (β=0.0002), jointly shaping a new pattern of elderly consumption centered on health management while balancing quality demands.

(2) Elderly consumer groups exhibit diversified behavioral characteristics.

The classification based on K-Means clustering analysis reveals that the e-commerce dominant type (42%) relies on digital platforms to achieve high-frequency product procurement, the social interaction type (31%) constructs consumption decision networks through short videos and communities, and the health-dependent type (27%) focuses on chronic disease management and health data tracking. The random forest model further quantifies the behavioral driving elements, with health record completeness (feature importance 0.239) and age-appropriate interaction frequency (0.152) emerging as core variables influencing consumption decisions, highlighting the strategic value of data asset accumulation and user experience optimization in the silver economy.

References

He, J. X., Li, R. X. and Liu, G. J., (2025). The deep integration and development of the digital economy and silver economy: Essence, theoretical underpinnings and development pathways *Journal of Harbin University of Commerce (Social Science Edition)*, no. 6, pp. 96-108.

Huang, Q. P., (2025). The silver economy amidst an aging population: A comparison of east asian models and China's development path. *Contemporary Economic Management*, pp. 1-14.

Kang, H. and Li, L., (2025). Regional disparities and convergence in the development of silver economy. *East China Economic Management,* vol. 39, no. 10, pp. 97-106.

Wei, X. M. and Zhu, K., (2025). Statistical measurement, regional differences and distribution dynamic evolution of the silver economy development. *Statistics & Decision,* vol. 41, no. 16, pp. 43-48.

Zhou, J., Wang, C. J. and Du, Y. X., (2025). From "Demographic Dividend" to "Time Dividend" : Silver-haired consumption helps unleash the potential of domestic demand. *Business & Economy*, no. 12, pp. 1-4,13.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.