Study on the Development of China's Pet Industry and Strategic Recommendations Based on Data Analysis

Hui Wang [1, *, #], Yupeng Bian [2, #], Zhengan Yan [2, #]

[1] School of Life Sciences, Tiangong University, Tianjin, China, 300387

[2] School of Electrics and Information Engineering, School of Tiangong University, Tianjin, China, 300387

* Corresponding author: 15222716375@139.com

[#]These authors contributed equally.

Abstract. The global pet industry is experiencing significant growth, particularly pronounced in regions such as North America, Europe, and parts of Asia, including China and Japan. In China, the adjustment of family planning policies alongside societal aging has transformed pets into crucial sources of emotional support, thereby invigorating the domestic pet market. Internationally, there is a pronounced demand for high-quality pet products and services in developed countries. Previous studies, including those by Zhang et al., who explored optimization strategies for enhancing pet brand competitiveness through the AARRR model, and You et al., who applied PEST, Porter's Five Forces, and SWOT analyses to investigate smart wearable devices in pet care, have contributed valuable insights. Despite these advancements, challenges persist, notably in achieving sustainable sourcing of pet food ingredients and comprehending the rapidly evolving pet electronics sector. This article aims to provide strategic suggestions for China's pet industry by using principal component analysis and polynomial and fitting linear regression models to deeply analyze the current situation, challenges and future prospects of China's pet industry.

Keywords: Principal Component Analysis, Polynomial Regression Model, Spearman's Rank Correlation Coefficient.

1. Introduction

The global pet industry is undergoing significant expansion, especially prominently in North America, Europe, and specific Asian countries such as China and Japan. The role of pets in families is becoming increasingly crucial and they are gradually becoming indispensable members of the family. In China, the trend towards more open family - planning policies and the emergence of an aging population have promoted pets to become an important source of emotional comfort, thus driving the vigorous development of the domestic pet market. The widespread use of the Internet and social media has accelerated the spread of pet culture and greatly promoted the online sales of pet - related products. In developed economies, the pet industry has become relatively mature, and the demand for high - end pet products and services continues to rise.

Zhang et al. [1] are committed to exploring how to enhance the competitiveness of pet brands through digital transformation; You et al. [2] use a variety of analytical tools to conduct research on smart wearable devices in pet management; Huang et al. [3] introduce a human - pet relationship model; the relevant research in North Carolina [4] integrates pet data into applications to promote the progress of pet technology; Jones et al. [5] predict the consumption trends of pet products in the United States.

However, despite much progress, the pet food ingredient sustainability and the pet electronics field still face many challenges and urgently need in - depth research. This study uses principal component analysis, polynomial and fitted linear regression models to analyze the Chinese pet industry, predicts its future development trends under the background of constantly changing global economic policies, and proposes sustainable development strategies, aiming to provide extremely valuable references for industry stakeholders. And the cited references also highlight the significance of multiple analytical methods in guiding the evolution of the pet industry.

2. Methods

2.1. Principal Component Analysis (PCA)

2.1.1. Algorithm overview

Principal Component Analysis (PCA) is a commonly used data dimensionality reduction technique, whose main purpose is to transform the original data into a new coordinate system through linear transformation, so that in this new coordinate system, the first principal component of the data has the largest variance, the second principal component has the next largest variance, and so on. This preserves the main information in the data while reducing the dimensionality of the data.

2.1.2. Algorithm Principles and Implementation Steps

Fundamentals of the PCA algorithm:

(1) Data standardization: first, the raw data need to be standardized to remove the effects of the scale between different variables. This is usually done by subtracting the mean of each variable and dividing by its standard deviation.[5]

(2) Calculation of covariance matrix: the normalized data matrix can be used to calculate the covariance matrix which describes the correlation between the variables. The covariance matrix is a symmetric matrix that represents the linear relationship between the variables in the data set. [6]

(3) Eigenvalue Decomposition: PCA usually uses Eigenvalue Decomposition (EVD) or Singular Value Decomposition (SVD) for covariance matrices. The eigenvalues of the covariance matrix correspond to the magnitude of the variance of the principal components, while the eigenvectors define the direction of the principal components. By selecting the largest eigenvalue and its corresponding eigenvector, the most important principal components can be extracted. [7]

(4) Determine the number of principal components: based on the cumulative contribution ratio (usually required to be 85%-95%), the first few principal components are selected. These principal components are able to retain most of the information in the original data and are orthogonal to each other. [8] [5]

(5) Reconstructing the data: finally, the original data is projected onto the selected principal components to obtain the dimensionality reduced data. This process can be realized by matrix multiplication, where the projection matrix consists of selected eigenvectors. [7]

The main advantage of PCA is that it can effectively reduce the dimensionality of the data while retaining most of the information in the original data. In addition, since the principal components are independent of each other, subsequent data analysis and visualization can be simplified. [9-10]

Implementation steps of the PCA algorithm:

(1) Data pre-processing

First, all data should be standardized first in order to eliminate the effect of the scale:

(2) Construct the covariance matrix or correlation coefficient matrix

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}, k = 1,2,\ldots,n, j = 1,2,\ldots,p \tag{1}$$

① covariance matrix

In order to describe the linear relationship between the variables, a covariance matrix should be constructed:

$$\mathrm{Cov}(X) = \frac{1}{n-1} X^T X \tag{2}$$

② Correlation coefficient matrix

In order to describe the linear correlation between the variables, a matrix of correlation coefficients should be constructed:

$$\mathrm{Corr}(X) = \frac{\mathrm{Cov}(X)}{\sigma_i \sigma_j} \tag{3}$$

(3) Calculate eigenvalues and eigenvectors

$$Av = \lambda v \tag{4}$$

(4) Selection of principal components

The eigenvectors corresponding to the top k eigenvalues are selected by ranking the eigenvalues from largest to smallest and letting these eigenvectors form the principal components.

(5) Variance contribution

The significance of each principal component was assessed by calculating the variance contribution ratio:

$$VAF = \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j} \tag{5}$$

(6) Calculate principal component scores

The raw data is projected into the principal component space to obtain the principal component scores:

$$F = XV_k \tag{6}$$

2.2. Fitted regression method

2.2.1. Algorithm overview

Fitted regression is an algorithmic model used to analyze and determine quantitative interdependencies between multiple variables. The core idea is to describe the relationship between variables by building a mathematical model and to use the observed data to estimate the parameters of the model, thus enabling prediction and analysis of unknown data.

2.2.2. Algorithm Principles and Implementation Steps

Fundamentals of the fitted regression method:

The best-fit model is found by minimizing the sum of squares of the residuals (RSS). Specifically, given a set of observations (x_i, y_i), where x_i is the independent variable and y_i is the dependent variable, the goal is to find a function $f(x)$ that minimizes the sum of squares of the error term ε in $y = f(x) + \varepsilon$. This process is usually accomplished by least squares, i.e.:

$$RSS = \sum_{i=1}^{n} (y_i - f(x_i))^2 \tag{7}$$

Minimizing the above RSS yields parameter estimates for the regression model. In addition, the fitted regression method can be extended to complex cases such as multiple regression, nonlinear regression, etc., by introducing polynomial terms or nonlinear functions to characterize the relationship between the variables.

Steps to implement the fitted regression method:

(1) Data collection

Relevant data is collected as data points (x_i, y_i), where $i = 1,2,\ldots,n$, n is the number of data points.

(2) Hypothetical model form

It is assumed that all data points can be represented by a linear model, i.e., $y = ax + b$, where a is the slope and b is the intercept.

(3) Define the error function

$$S(a, b) = \sum_{i=1}^{n} (y_i - (ax_i + b))^2 \tag{8}$$

(4) Solve the minimization error function

In order to find the parameters a and b that minimize the error function $S(a, b)$, it is necessary to take the partial derivatives of $S(a, b)$ with respect to a and b, respectively, and make them equal to zero.

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^{n} x_i (y_i - (ax_i + b)) = 0 \tag{9}$$

Take partial derivatives with respect to a:

Take partial derivatives with respect to b:

$$\frac{\partial S}{\partial b} = -2\sum_{i=1}^{n}(y_i - (ax_i + b)) = 0 \tag{10}$$

(5) Solve systems of equations

Join the two above equations with partial derivatives equal to zero and solve this system of equations to obtain the values of a and b:

Group of equations:

$$\begin{gathered}\sum_{i=1}^{n} x_i(y_i - (ax_i + b)) = 0 \\ \sum_{i=1}^{n}(y_i - (ax_i + b)) = 0\end{gathered} \tag{11}$$

The simplified system of equations:

$$\begin{gathered}\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} {x_i}^2 + b\sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} x_i + nb\end{gathered} \tag{12}$$

Solve the system of equations:

By solving the above system of equations, the expressions for a and b can be obtained:

$$\begin{gathered}a = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2} \\ b = \frac{\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i}{n}\end{gathered} \tag{13}$$

(6) Substitution modeling

The fitted data model is obtained by substituting the obtained a and b into the linear model $y = ax + b$.

2.3. Data Source

The data sourse was shown in Table 1, involving specific data websites and basic information for subsequent predictions

Table 1. Data sources cited in this article

Number	Name	Descriptive Citation with Source Webpage
1	Number of Pet Cats and Dogs in China (2019-2023) (in ten thousand heads)	As shown in the "2023-2024 China Pet Industry White Paper", provides significant insights into the trends of pet ownership in China.
2	Market Size of Cats and Dogs in China (2019-2023) (in million yuan)	Revealed by iMedia Data Center, illustrates the growth trend of the domestic pet market.
3	Market Size of Cat and Dog Food in China (2019-2023) (in million yuan)	Provided by Euromonitor International and Asia Pet Institute, showcases the evolution of the pet food market.
4	Market Size of Cat and Dog Nutritional Products in China (2019-2023) (in million yuan)	As documented in the "2023-2024 China Pet Industry White Paper", highlights the expansion of the pet nutrition market.
5	Per Capita Disposable Income of Chinese Residents (2019-2023) (in ten thousand yuan)	Reported by the National Bureau of Statistics and iResearch Consulting Institute, reflects the impact of income levels on the pet market.
6	Scale of Single-Occupancy Households in China (2019-2023) (in million households)	Based on data from the National Bureau of Statistics, analyzes the influence of changing family structures on the pet industry.
7	Aging Trend in China (proportion of population aged 60 and over) (2019-2023)	As indicated by the National Bureau of Statistics, explains the role of societal aging in shaping pet market demands.
8	Number of Pet Cats and Dogs Overseas (2019-2023) (in ten thousand heads)	As presented by Euromonitor International and CITIC Securities Research Department, shows international pet market dynamics.
9	Number of Global Cat and Dog Pet Companies (2019-2023)	Sourced from Euromonitor International, offers insights into the distribution and development trends of the global pet industry.
10	Global Market Size of Pet Food (2019-2023) (in billion US dollars)	As detailed by Euromonitor International and Qianzhan Industry Research Institute, helps understand the scale and trends of the global pet food market.

3. Results

3.1. Analysis of the development of the pet industry in China

3.1.1. Modeling

Using the method of establishing principal component analysis and polynomial regression model as well as using time series analysis, a model is built to forecast the development of China's pet industry in the next three years, and the results of the model forecast are used to provide a scientific basis and reference for the development of the industry.

3.1.2. Analysis of results

(1) Analysis of the development of the pet industry in China:

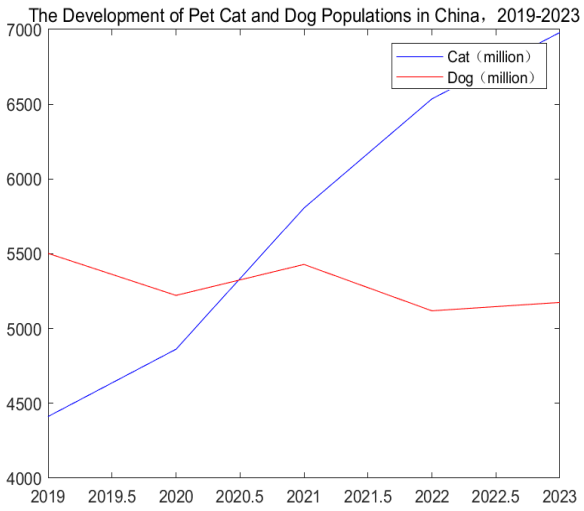

Figure 1. Trends in the number of dogs and cats in China over the past five years

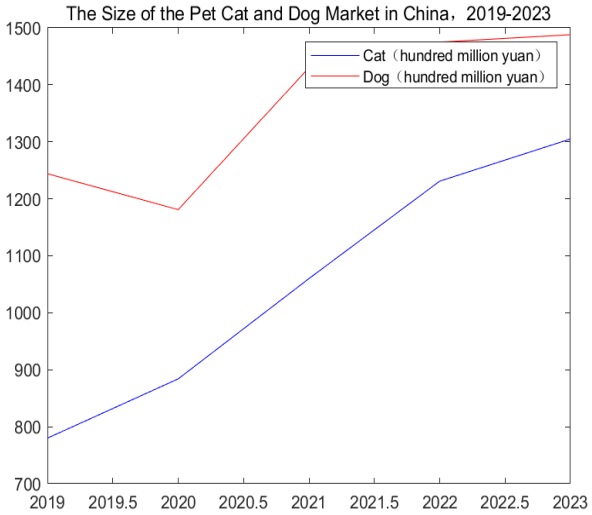

Figure 2. China's dog and cat market size development trend in the past five years

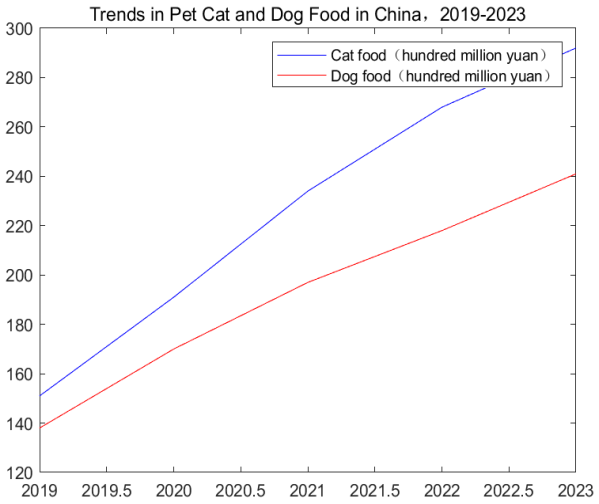

Figure 3. China's dog and cat food development trend in the past five years

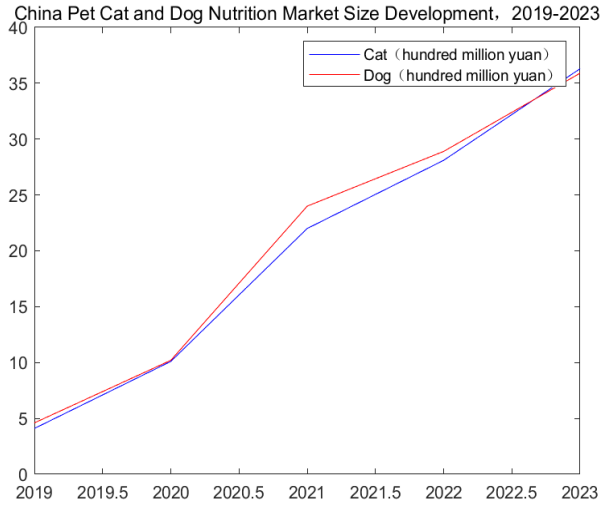

Figure 4. China Dog and Cat Pet Nutrition Market Size and Trend chart for the last five years

The analysis of China's pet industry development reveals several significant trends and factors contributing to its growth. Firstly, as depicted in Figure 1, the number of dogs and cats in China has seen a notable increase over the past five years, with cats showing the most substantial growth. This trend is complemented by Figure 2, which shows the continuous expansion of China's dog and cat market size during the same period. Additionally, Figure 3 highlights the upward trajectory of pet food sales, indicating an increased focus on pet nutrition, further supported by Figure 4, demonstrating strong growth momentum in the pet nutrition market. Collectively, these figures suggest that China's pet industry is experiencing rapid development, transforming into a formidable economic force.

(2) Analysis of the development factors of China's pet industry

Through principal component analysis we get the highest score for the degree of population aging at 0.8852, followed by the factor of the number of people living alone and the increase in disposable income per capita as the second and third in that order.

(3) Analysis of China's pet industry forecast for the next three years

By solving the polynomial regression model, we obtain the following results:

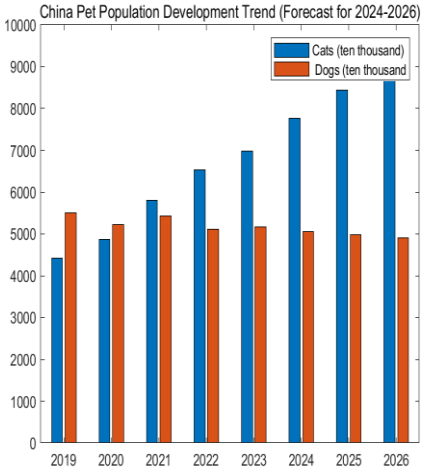

Figure 5. China's forecast of the number of dog and cat pets in the next three years

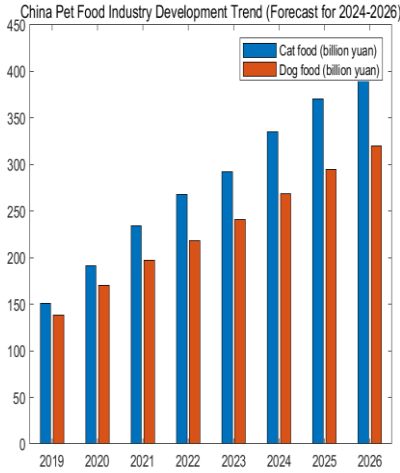

Figure 6. China's pet food development forecast for the next three years

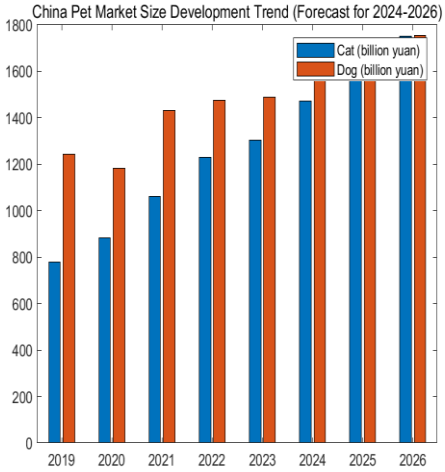

Figure 7. China's Pet Market Size Forecast for the Next Three Years

According to Figure 5 it can be seen that the number of pets in China will continue to grow over the next three years (2024-2026), with the number of cats expected to increase from about 70 million at present to nearly 90 million, while the number of dogs will gradually stabilize. Meanwhile, Figure 6 reveals that the pet food market will see significant expansion, with the cat food market expected to grow from about 30 billion yuan now to nearly 40 billion yuan, and the dog food market likely to grow from about 25 billion yuan to more than 30 billion yuan. Figure 7, on the other hand, reflects that the cat market is expected to reach about $180 billion by 2026, while the dog market will reach about $160 billion, and the gap between the cat and dog market sizes is gradually narrowing. All of the above predicts that the pet industry will continue to thrive in the coming years.

3.2. Analysis of the development of the global pet industry

3.2.1. Modeling

For this problem we use least squares to fit the data points and find the best linear regression model.

3.2.2. Analysis of results

(1) Global Pet Industry Development Analysis

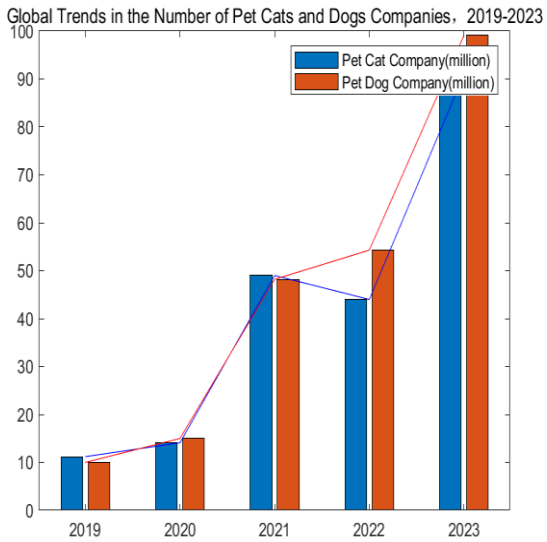

Figure 8. Trends in the number of global pet companies over the past five years

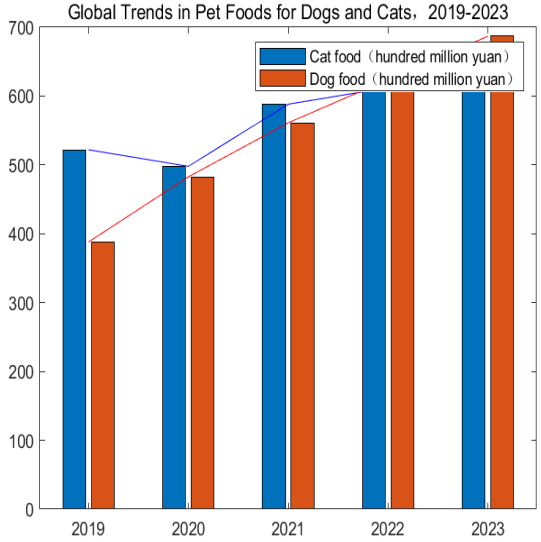

Figure 9. Global Pet Food Size Trends in the Last Five Years

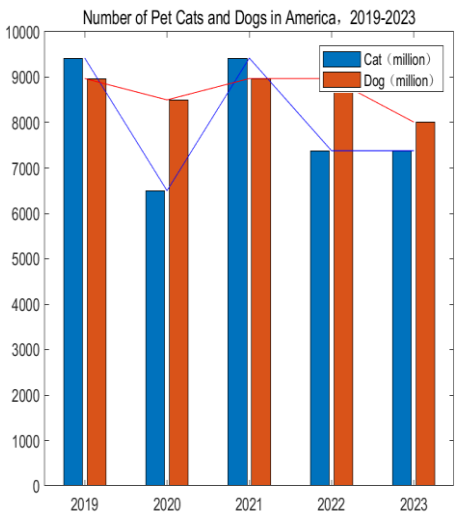

Figure 10. Trends in the number dogs and cats in the United States, 2019-2023

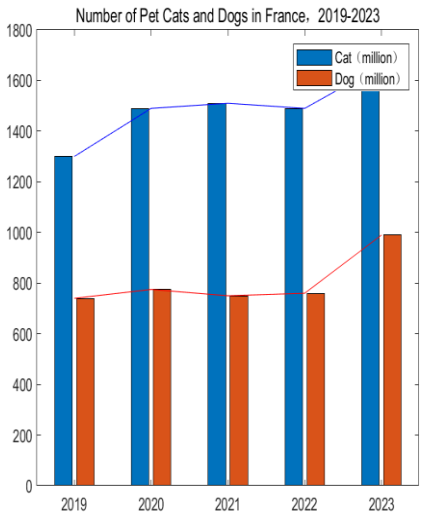

Figure 11. Trends in the number of dogs and cats in France, 2019-2023

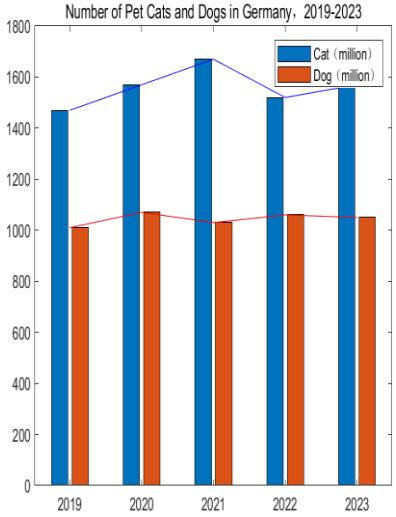

Figure 12. Trends in the number of dogs and cats in Germany, 2019-2023

The global pet industry is experiencing rapid growth, as illustrated by Figure 8, which shows an increasing trend in the number of pet companies over the past five years. This expansion is complemented by the significant rise in the global pet food market size, highlighted in Figure 9, demonstrating strong annual increases in cat and dog food sales, reaching 6.52 million tons and 6.87 million tons respectively in 2023. Figures 10, 11, and 12 depict trends in pet ownership in the United States, France, and Germany from 2019 to 2023, underscoring a growing emphasis on pet nutrition and health. These visuals collectively indicate that the industry's future outlook is promising, with demand expected to continue its upward trajectory in the next three years.

(2) Global Forecast for Pet Food in the Next Three Years

By solving the linear regression model, we can obtain the following results:

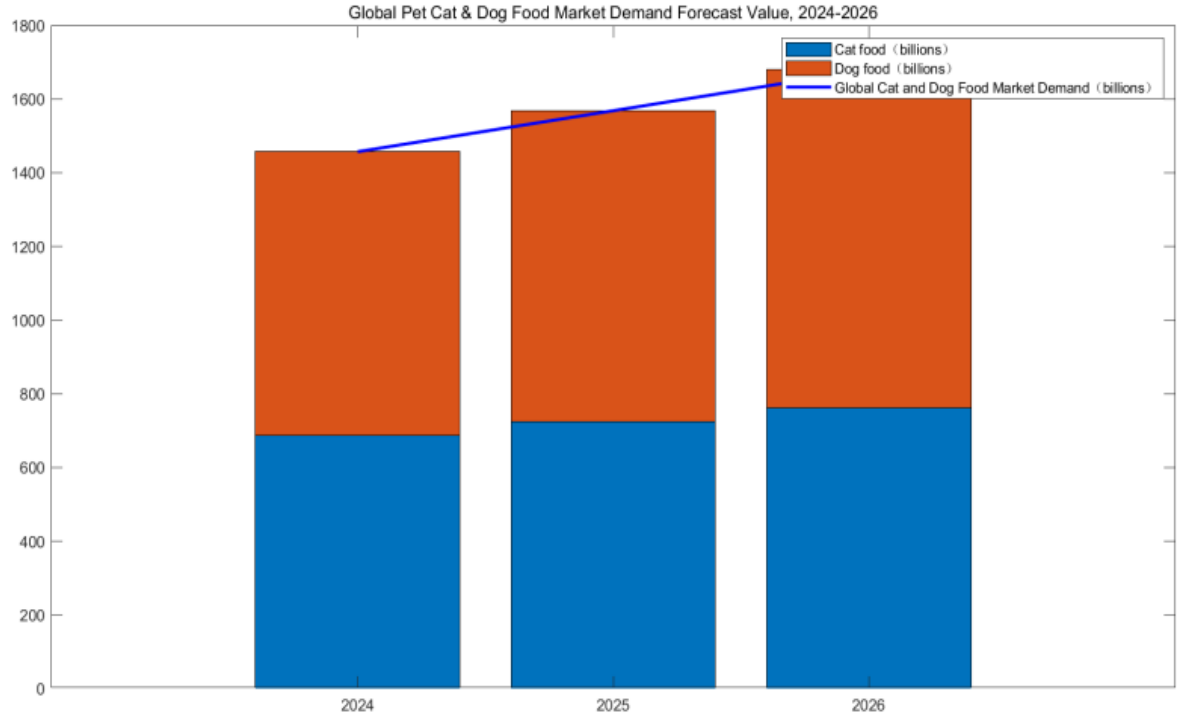

Figure 13. Global demand for pet food over the next three years

Figure 13 reflects the growing trend of global pet cat and dog food market demand between 2024 and 2026. As can be seen from the figure, the market demand for both cat food and dog food is rising year on year. The dog food market, in particular, is growing significantly faster than the cat food

market. Combined with the previous analysis, firstly, this may mean that consumer demand for high-quality pet food is increasing, and also reflects that people are paying more and more attention to the health and nutrition of pets; secondly, with the improvement of living standards and social and economic development, more and more families are raising pets, which also pushes the development and growth of the pet food industry. Therefore, we can foresee that the global pet cat and dog food market will continue to maintain a strong growth momentum in the coming years.

3.3. Analyzing the relationship between China's pet food production and China's pet food exports

3.3.1. Modeling

We modeled the relationship between China's pet food production and China's pet food exports through the following steps were conducted to analyze the relationship between:

(1) Level shifting:

(2) Calculate the sum of the squares of the grade differences:

(3) Calculate the Spearman rank correlation coefficient

We then use the model to derive the production and export volume of pet food in China for the next three years by fitting a linear regression.

3.3.2. Analysis of results

According to the above steps of the algorithm we get the following results:

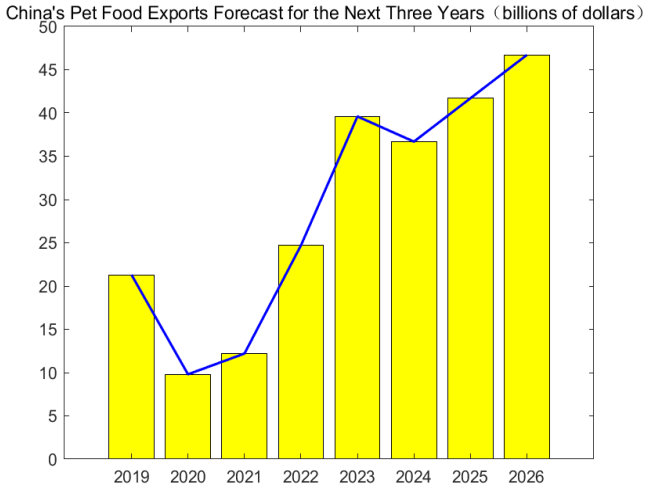

Figure 14. China Pet Food Export Forecast

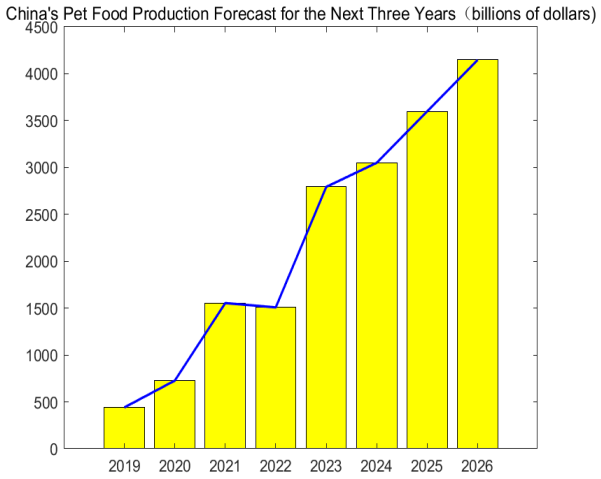

Figure 15. China Pet Food Production Forecast

Figure 14 and Figure15 illustrates although China's pet food exports have fluctuated, they are expected to increase significantly from 2022 - 2026, and the production of pet food from 2019 - 2026 will also increase substantially. This indicates that with the improvement of consumers' awareness and the expansion of the market, the development potential of the industry is huge.

3.4. Analysis and forecast of future trends in the pet market

3.4.1. Modeling

The ask we combine two methods of linear regression forecasting and principal component analysis, aiming to analyze and predict the future trends of the pet market from multiple perspectives. Through PCA, data dimensions can be effectively reduced while retaining key information for further analysis or decision support.

3.4.2. Analysis of results

Based on the above algorithm to solve the model, we finally get to keep the three components, of which the highest score is the market size of cats, followed by per capita disposable income and demographic changes (mainly refers to aging), i.e., in order to minimize the development of the

European and American countries on China's pet industry, we should focus the pet industry on the above three aspects, however, policies such as tariffs and other policies are closely related to each one of us, therefore We should also diversify our market layout, expand the global market as well as strengthen localization.

4. Conclusion

In the context of global economic integration and consumption upgrades, the pet industry has emerged as a rapidly growing sector worldwide, especially in China, where improved living standards, an aging society, and evolving family structures present unique opportunities. Utilizing historical data and statistical methods, including Principal Component Analysis (PCA) and regression models, this study examines China's pet industry to propose strategic recommendations. The findings of this study highlight aging as a key driver for pet market growth, with empirical evidence showing a preference among the elderly for pets as companions. Additionally, shifting consumer preferences towards cats and rising household incomes are bolstering pet-related expenditures.

Despite these insights, PCA and regression models exhibit limitations, particularly in handling large datasets or avoiding overfitting. To mitigate these issues, this article recommends incorporating prior knowledge into PCA and utilizing regularization techniques for regression models. This study aims to guide the development of a modernized pet industry system that aligns with national and international standards, maximizing economic and social benefits. Continuous exploration of optimization pathways is essential for ensuring the effectiveness and adaptability of proposed strategies. Future research should leverage artificial intelligence and big data analytics to better address dynamic market demands.

References

[1] Hidalgo-Fernández A, Moral-Cuadra S, Menor-Campos A, et al. Pet tourism: motivations and assessment in the destinations [J]. Consumer Behavior in Tourism and Hospitality, 2023, 18 (3): 335 - 350.

[2] Lawson, S., Kirman, B., et al. Problematizing upstream technology through speculative design: The case of quantified cats and dogs [C]//Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. Northumbria Research Link, 2015.

[3] Yang, Y., Yuan, Y., et al. Application of digital transformation in pet accessories market [C]//Proceedings of the 2022 2nd International Conference on Enterprise Management and Economic Development (ICEMED 2022), 2022.

[4] Guilherme-Fernandes J, Aires T, Fonseca A J M, et al. Squid meal and shrimp hydrolysate as novel protein sources for dog food [J]. Frontiers in Veterinary Science, 2024, 11: 1360939.

[5] Li B, Li J, Lu J. Research on the coupled risk of key nodes in maritime transport based on improved catastrophe theory [J]. Sustainability, 2019, 11 (17): 4640.

[6] Bruni, A., Leite, V. B. Quantum chemistry and chemometrics applied to conformational analysis [M]. IntechOpen, 2012.

[7] Chin T L, Shih W N. Concisely Indexed Multi-Keyword Rank Search on Encrypted Cloud Documents [J]. Applied Sciences, 2021, 11 (23): 11529.

[8] Han J, Zhou J. A Study on the Construction of Cultivating High-Quality Farmers in Higher Vocational Colleges and Universities under the OBE-PDCA Approach [J]. Applied Mathematics and Nonlinear Sciences, 2023, 9 (1).

[9] Onyeama, J. Credit card fraud detection in the Nigerian financial sector: A comparison of unsupervised TensorFlow-based anomaly detection techniques, autoencoders and PCA algorithm [R]. Enugu State University of Science and Technology, 2024.

[10] Košmelj K. Metoda glavnih komponent: osnove in primer [J]. Acta agriculturae Slovenica, 2007, 89 (1): 159 - 172.