Team Number:	apmcm24203392
Problem Chosen:	C

Development Analyses and Strategies for Pet Industry and relate Industries

Abstract

With the shift in consumption levels and attitudes, the pet industry and its related sectors (such as pet food, pet care, pet products, etc.) have become a rapidly developing emerging field in both the global and Chinese economies. Based on the data provided in the attachment and additional collected data, we conducted a detailed analysis of the development trends and market demand of the pet industry, and proposed development strategies. By constructing mathematical models, we analyzed the market dynamics of the pet industry in China and globally over the past five years, categorized by pet type, and predicted the development trends and market demand for the next three years. We also analyzed the production and export trends of China's pet food industry, and combined global demand forecasts to provide development recommendations. Additionally, in response to the impact of international economic policies (such as tariff changes), we proposed targeted sustainable development strategies to effectively support China's pet industry.

For Question 1: Analysis of the Development of China's Pet Industry in the Past Five Years and Prediction of Development Trends for the Next Three Years

Firstly, we conducted a detailed analysis of the industry's development in the past five years based on the data in Attachment 1 and additional collected data, categorizing by pet type (such as cats, dogs, etc.). We found that the cat-related industry will continue to grow, while the dog-related industry will experience fluctuations but maintain a relatively stable trend. Through the study of pet numbers, market scale, and changes in consumption structure, we used regression analysis to identify the main driving factors (such as per capita disposable income, urbanization rate, social attitudes, etc.). Subsequently, we established a multiple linear regression model to forecast the development trends of China's pet industry for the next three years. The results indicate that the pet industry in China will continue to grow, with pet food being the primary growth driver.

For Question 2: Global Pet Industry Development Analysis and Forecast of Pet Food Demand for the Next Three Years

We analyzed the global pet industry (focusing on Europe, North America, and Asia) by studying the market size and consumer behavior categorized by pet type. Using the data in Attachment 2 and supplementary industry statistics, we developed a multi-factor regression model to analyze the key influencing factors of global pet food demand (including economic growth rate, per capita income, pet food market, etc.). Based on time series forecasting models, we predict that global pet food demand will continue to grow annually over the next three years, with a significant expansion of the high-end pet food market, particularly in North America and Europe.

For Question 3: Development Trends of China's Pet Food Industry and Production and Export Forecast for the Next Three Years

In terms of the development of China's pet food industry, we analyzed the production and export dynamics of Chinese pet food, combining the data in Attachment 3 and global market demand data. By constructing a model of production and export value, and factoring in the growth trends of domestic and international market demand, we forecast the scale of China's pet food production and export for the next three years. The results show that the development trend of the pet food industry will gradually rise, and both the production and export of Chinese pet food will increase year by year in the next three years.

For Question 4: Impact of New Foreign Trade Economic Policies on China's Pet Food Exports and Countermeasures

To quantitatively analyze the impact of new tariff policies in European and American countries on China's pet food industry, we developed a multiple linear regression model based on new foreign trade economic policies, simulating export changes under different tariff scenarios. In response to this challenge, we propose the following strategies:
(1) Strengthen domestic market demand, promote product structure optimization, and brand development;
(2) Expand into emerging markets such as Southeast Asia and the Middle East to diversify trade risks;
(3) Leverage free trade agreements and international cooperation to reduce policy barriers. These measures will ensure the sustainable development of China's pet food industry.

Keywords: Pet industry; Market forecast; Pet food; Chinese exports; Global demand; New foreign trade economic policies; Mathematical modeling; Industry development; Sustainable development strategies

Contents

1. Problem Restatement.......1
1.1 Background of the Problem.......1
1.2 Restatement of the Problem.......1
2. Problem Analysis.......5
2.1 Analysis of Problem 1.......5
2.2 Analysis of Problem 2.......5
2.3 Analysis of Problem 3.......6
2.4 Analysis of Problem 4.......6
3. Model Assumptions.......6
3.1 Assumptions for Multiple Linear Regression.......6
3.2Assumptions for Global Market Growth.......7
4. Notation Explanation.......7
5. Model Establishment and Problem Analysis.......8
5.1 Model Establishment and Analysis for Problem 1.......8
5.1.1 Data Collection and Preparation.......8
5.1.2 Regression Analysis: Establishing the Mathematical Model.......8
5.1.3 Multicollinearity Check.......10
5.1.4 Future Forecast.......10
5.1.5 Results Visualization.......10
5.1.6 Conclusion and Application.......12
5.2 Model Establishment and Analysis for Problem 2.......12
5.2.1 Input Data.......12
5.2.2 Overview of ARIMA Model.......13
5.2.3 Applying ARIMA(1,1,1) Model.......13
5.2.4 Differencing Operation.......14
5.2.5 Parameter Estimation.......14
5.2.6 Model Diagnostics.......14
5.2.7 Future Forecast.......14
5.2.8 Summary.......15
5.3Model Establishment and Analysis for Problem 3.......16
5.3.1 Model Construction.......16
5.3.2 Data Modeling.......16
5.3.3 Model Prediction and Assumptions.......17
5.3.4 Forecast Visualization.......17
5.3.5 Conclusion and Application.......19

5 .4 Model Establishment and Analysis for Problem 4 20
5 .4 . 1 Key Factors Explanation 20
5 .4 . 2 Regression Model Construction 21
5 .4. 3 Regression Analysis 21
5 . 4 . 4 Model Evaluation 21
5 . 4 .5 Regression Results Visualization.. 22
5 .4 .6 Summary and Optimization..23
6. Model Advantages, Disadvantages, and Outlook.. **23**
6. 1 Advantages and Disadvantages of the Model in This Paper 23
6. 1 . 1 Weighted Growth Rate Calculation Model 23
6 . 1 . 2 Time Series Analysis Model..24
6 . 1 . 3 Linear Growth Trend Model 24
7. Model Improvement and Promotion **24**
7. 1 Incorporating More External Variables.. 24
7. 2 Inclusion of Innovative Products and Technological Advancements.............................25
7. 3 Integration of Big Data and Real-Time Data.. 25
8. References... **25**

I.Problem Restatement

1.1 Background of the Problem
With the improvement in living standards and changes in consumer attitudes, the pet industry has gradually become a rapidly growing emerging sector, especially in China. In recent years, pets have increasingly gained importance in households, evolving from traditional roles like working dogs and police dogs to companion animals, becoming an integral part of many families. According to relevant data, the pet consumption market includes various fields such as pet food, medical care, grooming, and pet supplies, with pet food consumption being particularly significant.

Based on historical data provided in the attachments, we can observe the steady growth of pet numbers in China, as well as the rapid development of pet-related industries. At the same time, the global pet market, especially in developed countries such as Europe and North America, is also thriving. The growing consumer demand in overseas markets not only promotes the development of the pet industry but also drives global demand for pet food.

However, despite the rapid development of the pet industry, it still faces various challenges. Quality control in pet food production, changes in import tariff policies, and uncertainties in market demand may all impact the sustainable development of the industry. Therefore, establishing mathematical models to analyze past market data and predict future demand trends is of great importance for businesses and governments to formulate relevant policies, optimize industry chain structures, and enhance industry efficiency.

This study aims to analyze the past development trends of the pet industry in China and globally, combined with additional data, to explore the future development direction of the industry in the coming years. It also seeks to provide scientific forecasts and strategic recommendations for China's pet industry in the context of global market changes. Furthermore, in response to changes in international economic policies, countermeasures are proposed to ensure the sustainable development of China's pet food industry.

1.2 Problem Restatement
This analysis aims to predict the future development trends of the pet industry in both China and globally through an in-depth exploration of the current status of the pet industry, combined with historical data and multiple regression analysis. By analyzing the Chinese pet market, we identified that factors such as economic growth, per capita income, and urbanization rate have a significant impact on pet numbers. The established regression model can provide reliable forecasts for the market size over the next three years. The global pet industry analysis shows that with the increase in the global pet population, the demand for pet food is on an upward trend. Combining China's pet food production and export data, we can forecast the market dynamics of Chinese pet food for the next three years. At the same time, considering the impact of international trade policies on China's pet food industry, effective countermeasures need to be developed to ensure the industry's sustainable development.

Problem 1: Establishing a Data Model to Predict the Development of China's Pet Industry in the Next Three Years
To accurately predict the development trends of China's pet industry over the next three years, we first conducted a detailed retrospective analysis of industry data over the past five years. These data include pet numbers, market size, consumption structure, and consumer behavior. Using the data from Attachment 1, as well as additional data we collected, and considering macroeconomic factors in China, changes in social structure, and other variables, we established a regression analysis model.

The following key variables were selected as inputs for the prediction model:

Economic growth rate: Reflects the growth rate of a country or region's economy, directly affecting consumer income levels and purchasing power.

Per capita GDP: An indicator of the economic development level of a country or region, which influences residents' consumption levels, especially in relation to pet-related products.

Urbanization rate: As the urbanization process accelerates, more urban residents keep pets, leading to an increase in pet numbers and market demand.

Per capita disposable income: Income levels directly affect households' ability to spend on pets, particularly on pet food, care, and supplies.

Pet market size: Reflects the overall development level of the pet industry and serves as an important indicator of market demand.

Pet products market size: Pet products are a key part of the pet industry, reflecting consumer demand for pet-related products.

Proportion of population aged 60 and above: The increase in the aging population may influence the pet industry, as older individuals tend to adopt pets for companionship, thus driving market demand.

Using multiple regression analysis, we identified the key driving factors affecting the pet industry and derived a regression equation for the pet market size (dependent variable) and various independent variables. The equation is as follows: $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + \beta_6 X_{6t} + \beta_7 X_{7t} + \varepsilon_t$

Pet Food Market $size_t$ is the pet food market size in year t(dependent variable);

β_0 is the intercept;

β_1, β_2 are the regression coefficients, representing the impact of each independent variable on the pet market size;
Each coefficient reflects the extent to which changes in the independent variables affect the dependent variable;
ε_t is the error term, representing the unexplained portion of the model.

This equation allows us to quantify the contributions of various factors to future industry development.

Problem 2: Establishing a Data Model to Forecast Global Demand for Pet Food in the Next Three Years
To accurately predict the development of the global pet industry and the demand for pet food over the next three years, we first conducted a detailed retrospective analysis of industry data from the past five years. This data includes market share calculations, weighted growth rate calculations, and global trend calculations. Using the data from Attachment 2 and additional data we collected, as well as global macroeconomic factors and social structure changes, we established an ARIMA model.

We analyzed the growth rates of cats and dogs in China, Germany, the U.S., and France over the past five years using the following:

Market Share Calculation: The proportion of cat and dog numbers in each country relative to the global total.

Weighted Growth Rate Calculation: Using the share of each country and its respective growth rate to calculate a weighted growth rate.

Global Trend Calculation: Using the weighted growth rate to calculate the global trends in the number of cats and dogs (from 2019 to 2023).

The growth rate calculation formula is as follows: $\text{Global Growth Rate} = \frac{\text{Current } \textit{year quantity} - \textit{the year before quantity}}{\textit{the year before quantity}}$

Country	Pets	2019-2020 Growth Rate	2020-2021 Growth Rate	2021-2022 Growth Rate	2022-2023 Growth Rate
America	Cat	-30.26%	44.31%	-21.89%	0%
	Dog	-5.23%	5.53%	0%	-10.72%
France	Cat	14.62%	1.34%	-1.33%	11.41%
	Dog	4.73%	-3.23%	1.33%	30.26%
Germany	Cat	6.80%	6.37%	-9.01%	3.29%
	Dog	5.96%	-3.74%	2.91%	-0.94%
China	Cat	10.23%	19.44%	12.60%	6.78%
	Dog	-5.09%	3.96%	-5.73%	1.09%

After establishing the preliminary analytical framework mentioned above, we conducted a more in-depth analysis of the future trends in global pet food demand. By combining the pet growth rates of each country with their corresponding pet food consumption data, we developed a multivariate regression model to predict the global pet food demand over the next three years. This model takes into account the following factors:

Per Capita Income: As per capita income increases, consumer spending patterns and overall consumption levels tend to change. Consumers in high-income regions are generally more willing to spend on their pets, especially when it comes to choosing high-quality pet food.

Predicted Pet Numbers: Based on ARIMA model trend forecasts, we projected the number of cats and dogs in China, Germany, the United States, and France over the next three years. Combining these predictions with each country's market share and pet breed preferences, we further estimated the demand for pet food in these markets.

Economic Environment Impact: The impact of economic growth rates on global pet food demand is complex. Economic growth often increases consumers' income levels, thus driving the expansion of the pet food market.

Pet Food Market Influence: The continuous development of the global pet food market has significantly contributed to the increasing demand for pet food. With the growing popularity of pet culture and the sustained increase in global pet numbers, the pet food market is expanding steadily.

Using these predictive data, we aim to provide global pet food manufacturers, retailers, and investors with a clear market outlook, helping them make more targeted market decisions and seize potential growth opportunities over the next three years.

Problem 3: Analyze the development of China's pet food industry and predict pet food production and export trends over the next three years

To predict China's pet food production and export trends over the next three years (2024–2026), a multivariate linear regression model was adopted. This model uses global market size, global urbanization rate, and global per capita income as key independent variables. By analyzing the historical relationships between these variables and China's pet food production and export data, a predictive model was established.

Specifically, the multivariate linear regression model is based on the formula: $y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \varepsilon$

Where:

- y: Target variable (production or export value)
- x_1, x_2, x_3 : Global market size, urbanization rate, and per capita income, respectively
- β_0 : Intercept
- β_1, β_2, β_3: Regression coefficients representing the impact of each independent variable
- ϵ: Error term

Using historical data from 2019 to 2023 to train the regression model and fit the relationships between variables, we substituted the predicted values for global market size, urbanization rate, and per capita income for the next three years into the model. This allowed us to compute future trends for China's pet food production and export values. This method captures the key influencing factors based on historical data and combines global economic and market trends to provide scientifically grounded growth forecasts for China's pet food industry.

Problem 4: Develop a mathematical model to quantitatively analyze the impact of new foreign trade economic policies in European countries and the United States on China’s pet food industry based on the results of Problems 1, 2, and 3

To accurately predict the impact of new foreign trade economic policies in European countries and the United States on China’s pet food industry, we established a multivariate linear regression model based on trends and official data from the past few years. These data include:Global economic growth rates， Changes in production costs， Exchange rate fluctuations， Tariff policies， Market demand indices.

Years	Global Economic Growth Rate(G)	Production Cost Growth Rate(C)	Tariff Policy(T)	Exchange Rate (R) (1 USD = RMB)	Market Demand Index(D)
2019	3.00%	3%	25%	6.9	100
2020	-3.50%	7%	25%	6.5	105
2021	5.90%	5%	25%	6.4	110
2022	3.20%	8%	20%	6.5	115
2023	2.90%	6%	15%	7	120

Global Economic Growth Rate (G): The global economic growth rate is a significant factor influencing the economic activities of countries worldwide.

Production Cost Growth Rate (C): Production costs are affected by various factors, primarily raw material prices and labor costs. In recent years, the growth in production costs has been largely driven by global supply chain disruptions and rising transportation expenses.

Tariff Policy (T): Tariff policy directly impacts exports, particularly when target markets (e.g., the United States, Europe) impose tariffs on Chinese imports.

- **2019:** China faced tariffs of approximately 25% from the United States due to the US-China trade war.
- **2020:** Tariff policies remained stable but were still influenced by the US-China trade war.
- **2021:** Tariffs remained high, at approximately 25%.
- **2022:** Some tariffs were eased, especially with slight adjustments to US tariff policies toward China.
- **2023:** Tariff policies stabilized, with US tariffs on Chinese imports maintaining a moderate level (around 15%-20%).

RMB to USD Exchange Rate (R): The exchange rate between the RMB and the USD significantly affects exports. In recent years, the RMB exchange rate has experienced substantial fluctuations, influenced mainly by international economic conditions and China's monetary policies.

Market Demand Index (D): The market demand index can be estimated by analyzing the market size and consumption trends in the pet food industry.

Estimating demand index based on pet population growth:
The demand for pet food is closely tied to the pet population. Assuming pet population growth directly drives increased demand for pet food, the market demand changes can be estimated using the annual growth rate of the pet population. The calculation formula is as follows:

Demand Index Formula: Market Demand Index=（Pet Growth Rate+1）*100%

Relationship between tariffs and exports: The impact of tariffs is modeled using an elasticity coefficient. It is assumed that as tariffs increase, exports decrease. The change rate of exports is inversely proportional to changes in tariffs: $\frac{E(\mathrm{t})-E_0}{E_0}=-\gamma\bullet\frac{T(\mathrm{t})-T_0}{T_0}$

Exchange rate impact model: $\frac{E(\mathrm{t})}{E_0}=\left(\frac{R(\mathrm{t})}{R_0}\right)$When the RMB depreciates, export revenue increases because products priced in USD become cheaper in international markets.

Market demand impact: $\frac{E(\mathrm{t})}{E_0}=\left(\frac{D(\mathrm{t})}{D_0}\right)$Market demand D(t)D(t)D(t) is closely related to pet food consumption trends and pet population in foreign markets. If external market demand increases (e.g., due to a growing pet population), exports will also rise accordingly.

II. Problem Analysis

2.1 Analysis of Problem 1
Problem 1 requires us to establish a data model to predict the development of China's pet industry in the next three years. To tackle this, we first need to collect data from the past five years related to China's pet industry. As mentioned in Section 1.2, this data includes pet numbers, GDP per capita, disposable income per capita, pet market size, and so on. By performing multivariate regression analysis on these data, we can identify the key driving factors affecting the development of the pet industry, including:

Market Size and Pet Number Changes: Analyzing the relationship between pet numbers and market size to identify the reasons for growth (such as increased pet numbers, higher per-pet spending, etc.).

Impact of Socio-Economic Factors: Analyzing how factors such as per capita income and GDP growth affect the pet industry. Using a regression analysis model, we can establish the basic trend equation of the pet industry. For example, by finding the quantitative relationship between pet number growth, income level, and urbanization rate through regression analysis. We will then use time series analysis methods to forecast the industry's development trend for the next three years, employing the ARIMA model (AutoRegressive Integrated Moving Average model) to analyze data such as market size and pet number changes over the past five years, and to generate predictions for the next three years.

2.2 Analysis of Problem 2
Problem 2 aims to predict the development trend of the global pet food industry in the next three years by analyzing global pet industry development data (focused on Europe, the U.S., and Asia)

with sales as the primary focus. First, we reviewed the number of cats and dogs, GDP growth rate, per capita income, and pet food market size in China, Germany, the U.S., and France from 2019 to 2023. This provided us with the historical growth trajectory of China's pet food industry and helped us identify the main driving forces and potential bottlenecks in the industry. To accurately predict the global pet food industry’s trend for the next three years, we used a multifactor regression model combined with a time series model to predict the annual demand for global pet food. The model not only considers domestic market demand changes but also incorporates global economic environment factors, aiming to comprehensively assess the growth potential and risks for China's pet food industry in the future.

2.3 Analysis of Problem 3

Problem 3 aims to predict the development trend of China's pet food industry for the next three years by combining production and export data with factors such as global market size, urbanization rate, and per capita income. First, by analyzing China's production and export data of pet food from 2019 to 2023, we can identify the growth trajectory of the industry and further understand the potential of the domestic market and the competitiveness in the international market. This historical data provides a foundation for assessing China's position in the global pet food market and its future growth potential. By establishing a multivariable regression model, we can relate global economic and social trends to China's industry data and predict the changes in pet food production and exports in China for the next three years.

2.4 Analysis of Problem 4

Problem 4 aims to analyze the impact of new economic policies on China's pet food exports, arising from changes in the global economic environment and international trade policy adjustments from 2019 to 2023. By combining the key influencing factors from Problems 1, 2, and 3, as well as their impact on China's pet industry, we can deeply identify the industry’s growth trajectory and potential. Historical data not only reveals the expansion trend of the domestic market and international competitiveness but also lays the foundation for evaluating China’s position in the global pet food market and its future growth potential. Based on this, we will establish a multivariable regression model to predict the specific impact of new foreign trade economic policy adjustments on export volume through the analysis and visualization of future data and propose response strategies.

III. Model Assumptions

3.1 Assumptions for Multivariate Linear Regression Model:

1. **Linear Relationship Assumption**: We assume a linear relationship between independent variables (predictors) and the dependent variable.
2. **Normal Distribution Assumption**: The error terms are assumed to follow a normal distribution. While parameter estimates in multivariate regression are unbiased and efficient even when error terms deviate from normality, the normality assumption is crucial for hypothesis testing (e.g., t-tests and F-tests) and inference. If the error terms do not follow a normal distribution, the statistical significance of regression results may be affected.
3. **No Multicollinearity Assumption**: We assume no perfect linear relationships between independent variables, i.e., there is no multicollinearity. If there is strong correlation between predictors, the regression model may produce unstable coefficient estimates, leading to inflated standard errors and affecting the model's interpretability and prediction accuracy. Common methods to check for multicollinearity include calculating the Variance Inflation Factor (VIF).

3.2 Assumptions for Global Market Size Growth:

1. **Global Pet Market Size Growth Assumption**: It is assumed that the global pet market size will grow by $10 billion annually. This assumption is based on historical trends and reflects the expansion momentum of the global pet market.
2. **Global Urbanization Rate Growth Assumption**: It is assumed that the global urbanization rate will grow by 0.2% per year. This assumption is based on past statistical data and forecasts, considering the steady development of global urbanization.
3. **Global Per Capita Income Growth Assumption**: It is assumed that global per capita income will increase by $1,000 annually. This is based on historical growth trends, indicating gradual economic improvement and increasing consumption capacity for pet products.
4. **Global Pet Number Distribution Assumption**: It is assumed that the distribution of pet numbers (cats and dogs) across countries/regions will continue to follow historical data, and the weighted impact of each region's market share on the global pet industry will be calculated.
5. **Pet Growth Rate Assumption**: It is assumed that the annual growth rate of pets (cats and dogs) in each country will be influenced by historical data and predicted based on these historical growth rates.
6. **China's Pet Food Industry Production and Export Assumption**: It is assumed that the production and export trends of China's pet food industry will continue to develop in line with the current trends, although they may be affected by global economic and policy changes.
7. **Impact of External Economic Policy Assumption**: It is assumed that China's pet food industry is influenced by the foreign economic policies of European and American countries (such as tariff policies), and these policy changes' potential impacts on the market will be considered in the model.

IV.Symbols Explanation

Symbols	Meaning Explanation
Y_t	Pet food market size in year ttt
X_{it}	Value of each independent variable iii in year ttt
y	Production or export value, or the target vector
x_1, x_2, x_3	Independent variables such as global market size, urbanization rate, and per capita income
$(X_2, X_3, ..., X_K)$	Predictor variables in the OLS regression model
$\alpha_1, \alpha_2, ..., \alpha_K$	Parameters estimated using the least squares method
R_i^2	Coefficient of determination obtained from least squares regression
β_0	Intercept term
β_i	Regression coefficients, representing the impact of each independent variable on the pet market size
β	Coefficient vector
ε_t	Error term or residual (white noise), representing the unexplained part of the model
p	Used to test the significance of each independent variable
$\Delta Y_t = Y_t - Y_{t-1}$	Difference of the time series in year ttt
ϕ_1	AR(1) coefficient (autoregressive coefficient)

θ_1	MA(1) coefficient (moving average coefficient)
X	Feature matrix, where the i th-row and j th-column represents the feature value
$\hat{y}$	Future target values obtained by fitting the model using least squares
$Cat_{China, t}$	Number of cats in China in year ttt
$Dog_{China, t}$	Number of dogs in China in year ttt

V. Model Development and Problem Analysis

5.1 Model Development and Analysis for Problem 1

To accurately predict the development of China's pet industry over the next three years, we used a linear regression model to analyze the factors affecting the pet industry's development over the past five years (2019-2023) and predict the pet market size from 2024 to 2026 using the regression model.

5.1.1 Data Collection and Preparation

Data preparation is the first step in modeling, ensuring the quality and completeness of the data used. We collected data related to the pet market from the following aspects:

Dependent Variable: We selected *Pet Market Size*, which reflects the overall development of the pet industry, as the target variable we aim to predict.

Independent Variables: We selected several economic, social, and industry-related variables, including economic growth rate, GDP, per capita income, urbanization rate, etc. These independent variables are assumed to be closely related to the growth of the pet industry, as economic conditions, consumer income, and social development factors directly impact the demand in the pet market.

5.1.2 Regression Analysis: Building the Mathematical Model

Building the regression model is the core of data analysis. In this step, we chose a multiple linear regression model, assuming that the pet market size can be predicted by multiple independent variables (such as economic growth, income levels, urbanization, etc.).

1.Selection of Independent and Dependent Variables We selected the following independent variables from the data:

- Economic Growth Rate
- Per Capita GDP
- Urbanization Rate
- Per Capita Disposable Income
- Pet Market Size, etc.

The rationale for choosing these independent variables is that economic and social factors typically have a direct impact on consumers' purchasing power and the demand in the pet market. For example, an increase in per capita GDP may mean that more households can afford pet-related expenses, thus driving the growth of the pet market size.

2.Ordinary Least Squares (OLS) Regression The OLS regression model minimizes the sum of squared errors to find the best regression coefficients. For each independent variable, it is treated as a dependent variable, with the other independent variables used as predictors. The regression model is estimated using the least squares method to obtain the coefficients.

For example: For the independent variable x_1: $X_1 = \alpha_0 + \alpha_1 X_2 + \alpha_2 X_3 + ... + \alpha_k X_k + \varepsilon_1$

Here, the other independent variables serve $(X_2, X_3, ..., X_K)$ as predictors, and the least squares method is used to estimate the model parameters $\alpha_1, \alpha_2, ..., \alpha_K$, then calculate the coefficient of determination (R-squared) R_i, which indicates how well the model fits the data.

Regression Coefficients (coef): The regression coefficient for each independent variable represents its effect on the pet market size. For example, the regression coefficient for the Economic Growth Rate might indicate how much the pet market size will increase in billions of RMB for each 1% increase in the economic growth rate.

p-value: This is used to test the significance of each independent variable. If the p-value is less than 0.05, it suggests that the independent variable has a significant impact on the dependent variable.

3. Regression Results Analysis

By using model_pet_market.summary(), we can review the detailed results of the regression model, which include:

Regression Coefficients (coef): Explains the contribution of each independent variable (such as economic growth rate, per capita GDP, urbanization rate, disposable income, and the percentage of the population over 60 years old) to the dependent variable.

Standard Error (stderr): Measures the uncertainty of the regression coefficients.

t-value and p-value: Used to assess the significance of the regression coefficients.

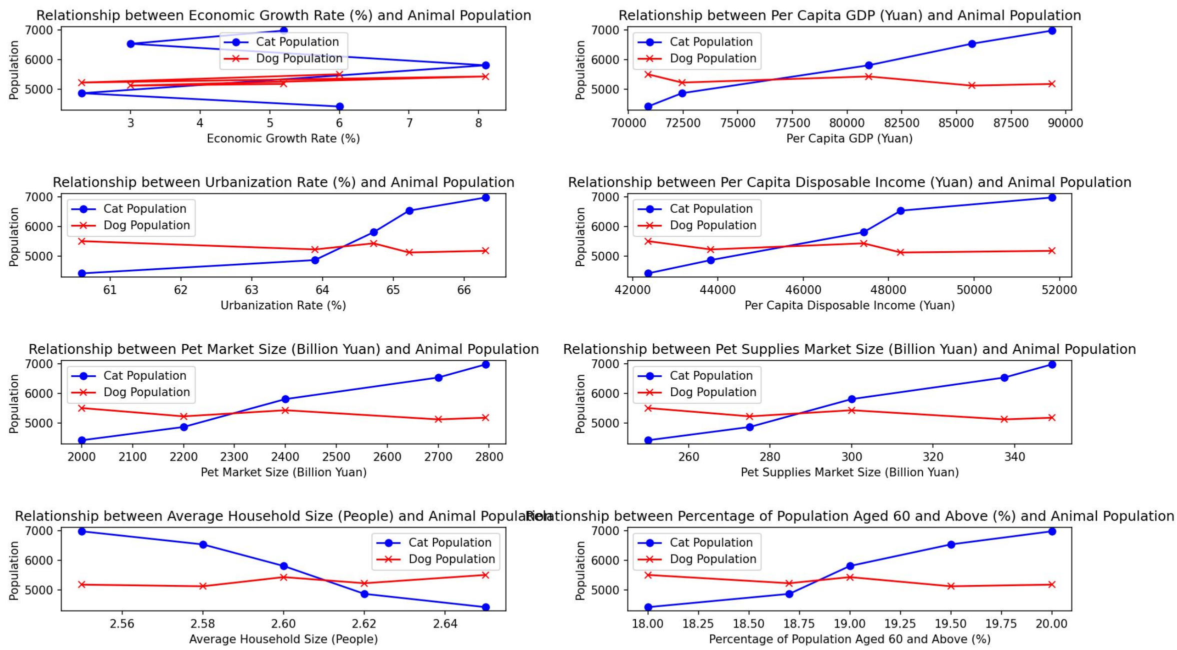


Figure 1

Using the heatmap Function in Seaborn to Plot the Correlation Matrix Heatmap

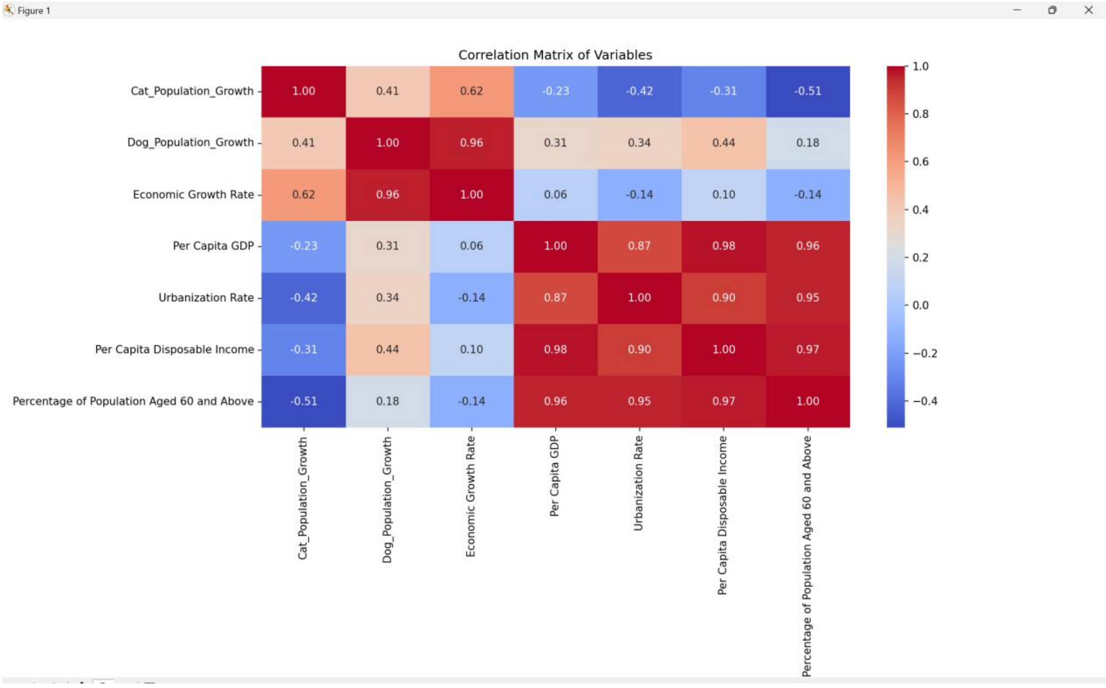


Figure 2: Heatmap

This visualization helps identify which factors significantly impact the pet market size and which might be redundant and could be excluded from the model.

5.1.3 Multicollinearity Check

A critical issue in regression analysis is multicollinearity, where independent variables are highly correlated. This can lead to unstable regression results or incorrect conclusions. To detect this issue, we used the Variance Inflation Factor (VIF):VIF Formula

Once R_i^2 is computed, the VIF for each X_i variable is calculated using the formula: $VIF(X_i) = \frac{1}{1-R_i^2}$

Where R_i^2 is the coefficient of determination obtained from least squares regression.

VIF Values: VIF measures the degree of multicollinearity for each independent variable. A high VIF (typically >10) indicates strong correlation with other variables, suggesting the need to remove or adjust the variable.

5.1.4 Future Prediction

Using the prepared independent variable data for 2024, 2025, and 2026, the model predicted future pet market sizes. This process provided predictions for the next three years and insights into the growth trends of the pet industry.

5.1.5 Result Visualization

To present the regression analysis results more intuitively, we compared actual data and predicted data using graphical methods.

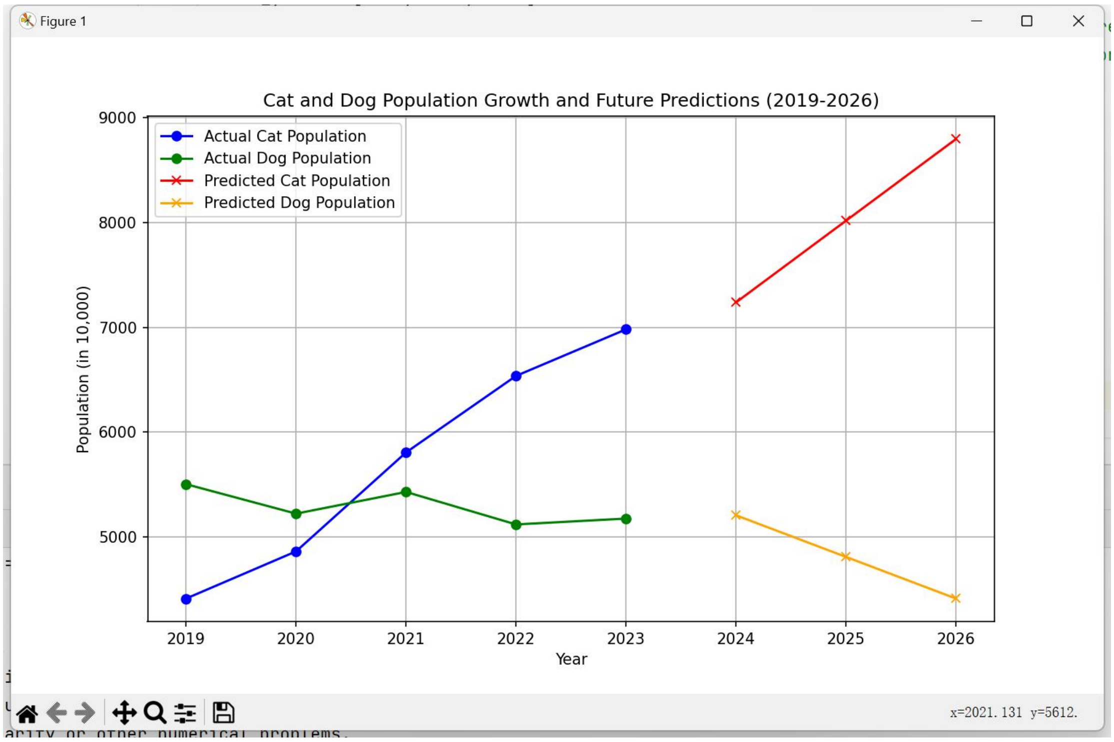


Figure 3

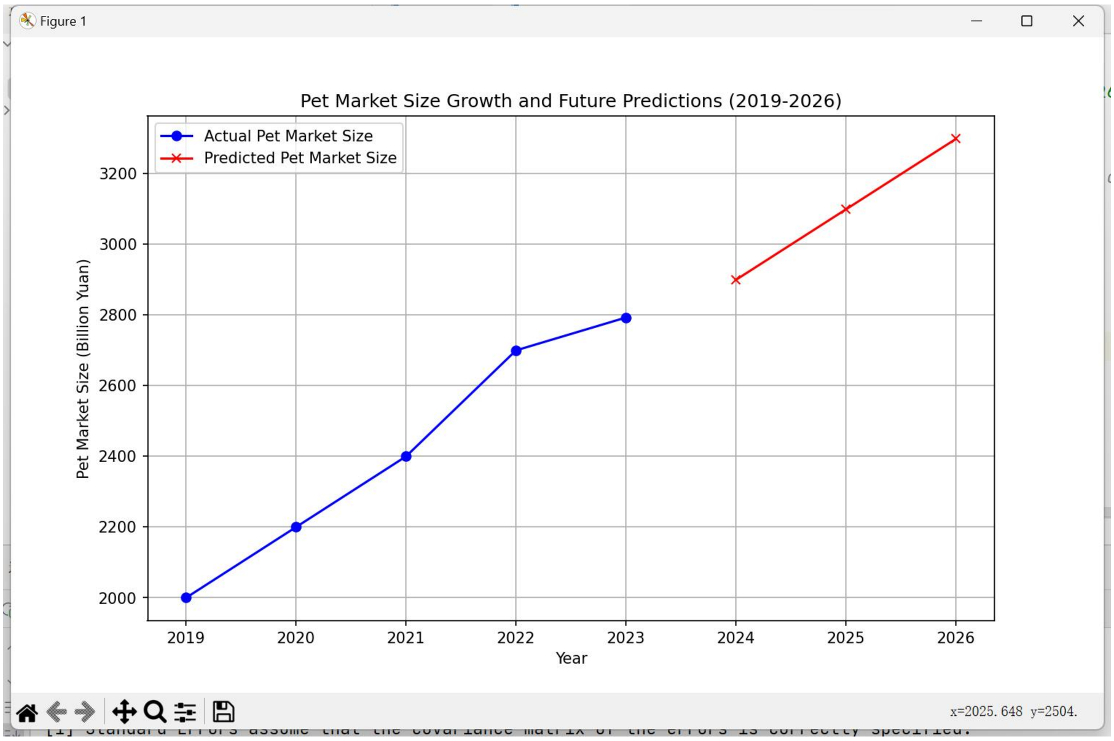


Figure 4

Based on the ARIMA model predictions, the development trends of China's pet industry over the next three years were estimated as follows:

Pet Population: The pet population is expected to continue its growth trend over the next three years, with an average annual growth rate of approximately 6%.

Market Size: The overall market size is projected to grow by an average of 8% annually, with pet food and pet services being the main growth drivers.

Segment Trends: As pet owners become more concerned with pet health, fields like pet healthcare and pet insurance are likely to emerge as new growth engines.

5.1.6 Conclusions and Applications

Through regression analysis, a mathematical model was successfully established to predict the future development trends of China's pet industry. The key takeaways from the modeling process include:

Influencing Factors: Economic growth, per capita income, and urbanization rate significantly influence the pet market size.

Future Trends: Based on the predictions, the pet market size is expected to continue growing. As income levels and economic conditions improve, the pet market will see greater expansion.

This model not only helped identify the key factors driving the pet industry's development over the past five years but also provided an effective forecast for the next three years. These insights allow policymakers and business decision-makers to develop strategies based on the predictions, such as adjusting product pricing, market promotion, and supply chain management.

5.2 Model Development and Analysis for Problem 2

To establish a sales prediction model, we need a method to capture the relationship between "year" and "sales" in historical data. Here, we assume this relationship is linear, meaning that as the year progresses, sales either increase or decrease linearly. For this purpose, we used the ARIMA model (AutoRegressive Integrated Moving Average) for regression modeling.

The core assumption of the ARIMA model is that a linear relationship exists between the dependent variable (sales) and the independent variable (year).

5.2.1 Input Data

The input data is as follows:

```
data = {'Year': [2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2022, 2023],'Sales (Billion USD)': [59.30, 62.40, 65.90, 69.60, 73.30, 77.70, 75.25, 75.00, 91.10, 93.90, 102.60, 123.60, 133.90]}
```

This is time series data representing global pet food sales from 2010 to 2023. The goal is to use the ARIMA model to predict sales for 2024–2026.

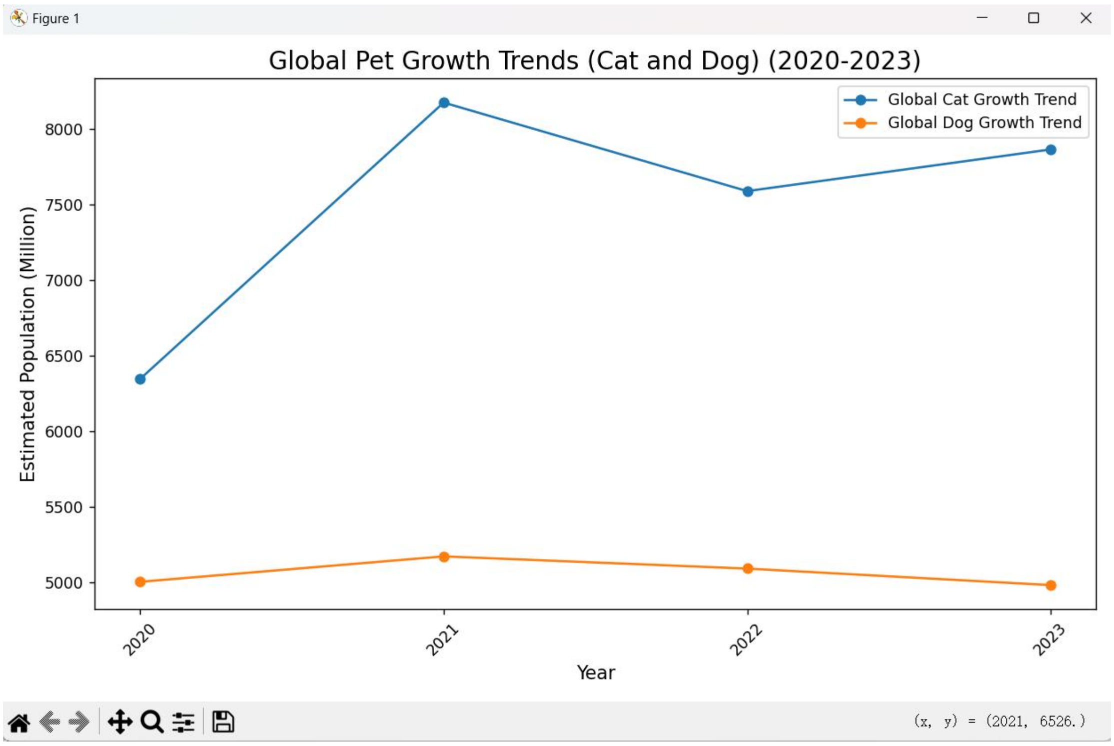


Figure 5

5.2.2 Overview of the ARIMA Model

The ARIMA (AutoRegressive Integrated Moving Average) model is a commonly used time series analysis method for predicting future values. It consists of three components:

AR (AutoRegressive): Uses a linear combination of past data to predict future values.

I (Integration): Differentiates the time series to make it stationary.

MA (Moving Average): Uses past forecast errors to correct future predictions.

The general form of the ARIMA model is: $Y_t = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$

Where:

- Y_t is the value of the time series at time ttt.
- α is the constant term.
- ϕ_i is the parameter for the AR component.
- θ_0 is the parameter for the MA component.
- ε_t is the white noise (random error normally distributed).

5.2.3 ARIMA(1, 1, 1) Model

The ARIMA(1, 1, 1) model is defined as follows:

- **AR(1)**: The autoregressive component uses the value at the previous time step to predict the current value.
- **I(1)**: First-order differencing is applied to make the data stationary.
- **MA(1)**: The moving average component uses the error from the previous step to adjust the prediction.

The formula is: $\Delta Y_t = \alpha + \phi_1 \Delta Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$

Where:

- $\Delta Y_t = Y_t - Y_{t-1}$ (differenced value).
- ϕ_1 : AR(1) coefficient.
- θ_1 : MA(1) coefficient.
- ε_t : Residual (white noise).

5.2.4 Differencing Operation

In the code, the differencing operation is performed using the diff() function. This operation removes the trend component from the time series, making it stationary. The formula for first-order differencing is: $\Delta Y_t = Y_t - Y_{t-1}$

Differencing helps eliminate trends, transforming the time series into a stationary series.

5.2.5 Parameter Estimation

The model parameters ϕ1\phi_1ϕ1 and θ1\theta_1θ1 were estimated using methods like least squares or Maximum Likelihood Estimation (MLE). In the code, the SARIMAx model estimated these parameters using MLE.

From the fitted model results:

- **AR(1) coefficient** Φ_1 : -0.1006
- **MA(1) coefficient** θ_1 : 0.8023

5.2.6 Model Diagnostics

Key diagnostic results include:

Ljung-Box Q-Test:

Q statistic: 1.26

p-value: 0.26 (p-value > 0.05 indicates no autocorrelation in residuals).

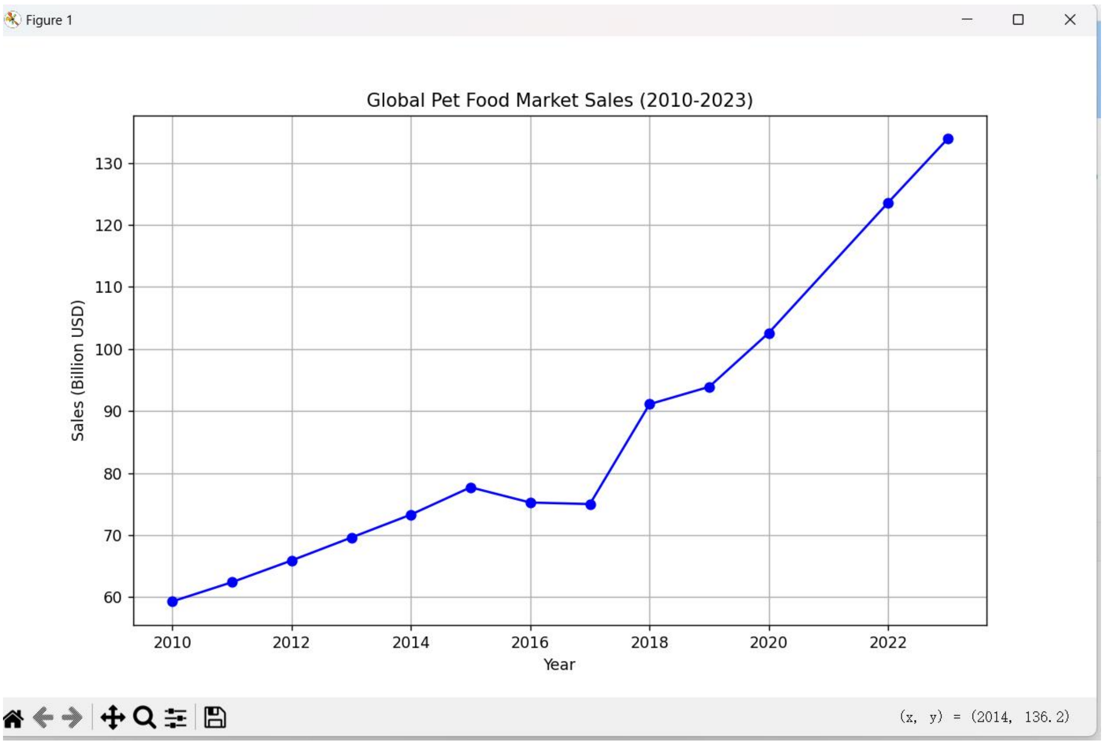
Jarque-Bera Test:

Skew: 0.24 (indicating residuals are close to symmetric).

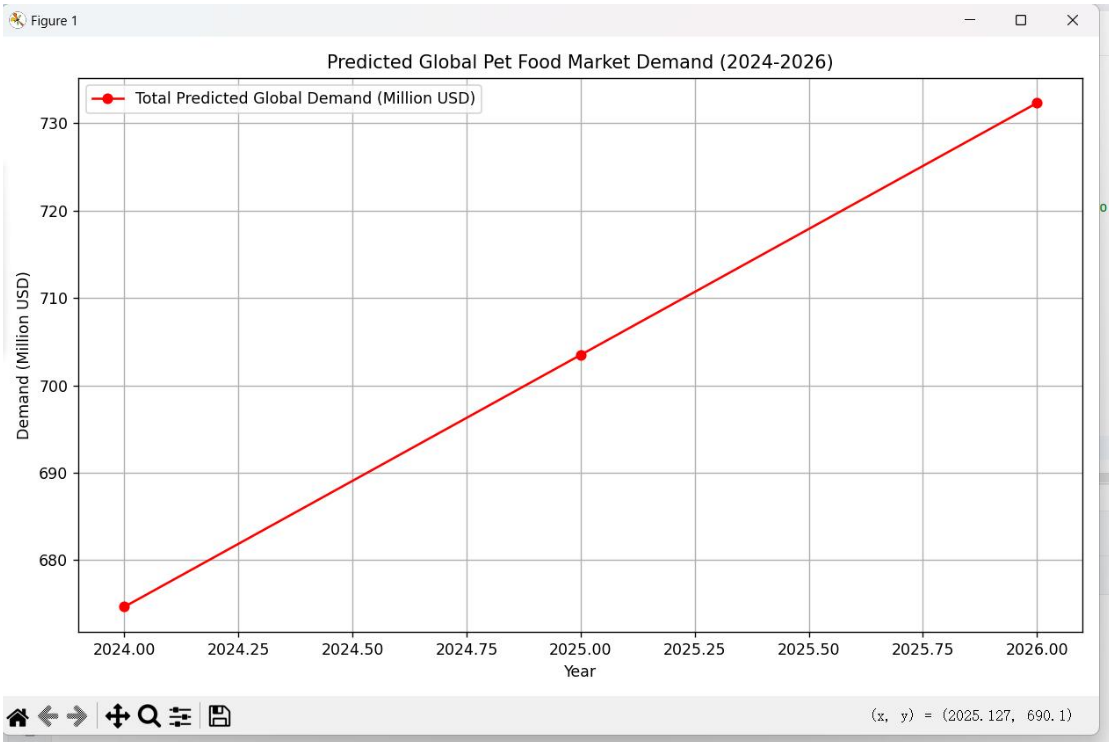
Kurtosis: 2.46 (close to 3, indicating residuals follow a normal distribution).

5.2.7 Future Predictions

Based on the analysis of recent data, the model predicted the following for the next few years:



Figures 6



Figures 7

According to the ARIMA(1, 1, 1) model predictions, sales will continue to increase steadily in the coming years.

5.2.8 Summary

The overall process involved the following steps:

1. **Data Preprocessing**: Applying differencing to make the time series stationary.
2. **Model Fitting**: Using the ARIMA(1, 1, 1) model and estimating parameters via MLE.

3. **Prediction**: Forecasting sales for future years based on the fitted model.
4. **Model Diagnostics**: Evaluating the model's quality using statistical tests like the Ljung-Box Q-test and Jarque-Bera test.

This process demonstrated that the ARIMA model is an effective method for time series forecasting, enabling better understanding and prediction of future data trends.

5.3 Model Building and Analysis for Question 3

To forecast the development trends of China's pet food production and export values for the next three years (2024–2026), a **multiple linear regression model** was constructed. The model predicts China's pet food production value (Billion RMB) and export value (Billion USD) using the following input features:

- Global pet food market size (Billion USD)
- Global urbanization rate (%)
- Global per capita income (USD)

5.3.1 Model Construction

The **multiple linear regression model** identifies linear relationships between historical data and utilizes these to predict future trends. The regression formula is: $y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + \varepsilon$

Where:

- y is the target variable (production value or export value).
- x_1, x_2, x_3 represent the input features (global market size, urbanization rate, and per capita income).
- β_0 is the intercept, and β_1, β_2, β_3 are regression coefficients.
- ϵ is the error term.

Regression coefficients are calculated during model training and quantify the impact of each feature on the target variable.

5.3.2 Data Modeling

1.Feature Matrix and Target Variables:

Feature matrix X： $X = \begin{bmatrix} \text{Market}_1 & \text{Urbanization}_1 & \text{Income}_1 \\ \text{Market}_2 & \text{Urbanization}_2 & \text{Income}_2 \\ \vdots & \vdots & \vdots \\ \text{Market}_n & \text{Urbanization}_n & \text{Income}_n \end{bmatrix}$

Target variables y：

Production values: China's production values in Billion RMB. $y_{production} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} \text{2019年生产值} \\ \text{2020年生产值} \\ \vdots \\ \text{2023年生产值} \end{bmatrix}$

Export values: China's export values in Billion USD.

$$y_{\exp ort} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 2019\text{年出口值} \\ 2020\text{年出口值} \\ \vdots \\ 2023\text{年出口值} \end{bmatrix}$$

2.Model Training:
Using the LinearRegression() method, find the best regression coefficients through .fit(x, y). This method uses the least squares fitting approach. The core idea is to minimize the sum of squared residuals, as shown in the following formula: $RSS = \sum_{i=1}^{n} (y_i - (\beta_0 + \sum_{j=1}^{n} \beta_j x_{ij}))^2$

- y: Actual values.
- β_0, β_j : The regression coefficients to be determined.
- x_{ij} : The feature value in the iii-th row and jjj-th column.

The regression coefficients are calculated using the formula: $\beta = \left(X^T X\right)^{-1} X^T \mathrm{y}$

- X: Feature matrix.
- y: Target vector.
- β : Coefficient vector $\left[\beta_0, \ \beta_1, \ \beta_2, \ \beta_3\right]^T$.

5.3.3 Predictions and Assumptions

Future feature values (future_X) are assumed as follows: $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

- Global market size: Increases by approximately 10 billion USD annually.
- Urbanization rate: Grows steadily by 0.2% annually.
- Per capita income: Rises by 1,000 USD annually, consistent with historical trends.

Using the trained model and assumed future values, predictions for 2024–2026 are calculated.

5.3.4 Visualization and Predictions

From the charts, we can observe that:

- **Production Value**: With the increase in global market demand and the advancement of urbanization, the production scale of pet food in China continues to grow.
- **Export Value**: With the improvement in global consumption capacity (per capita income) and the expansion of market demand, export value has increased significantly.

Forecast for Production and Export Values in 2024-2026:

- **Production Value**: Expected to maintain steady growth over the next three years.
- **Export Value**: Influenced by the global market size and per capita income, the growth rate is expected to be higher than that of production value.

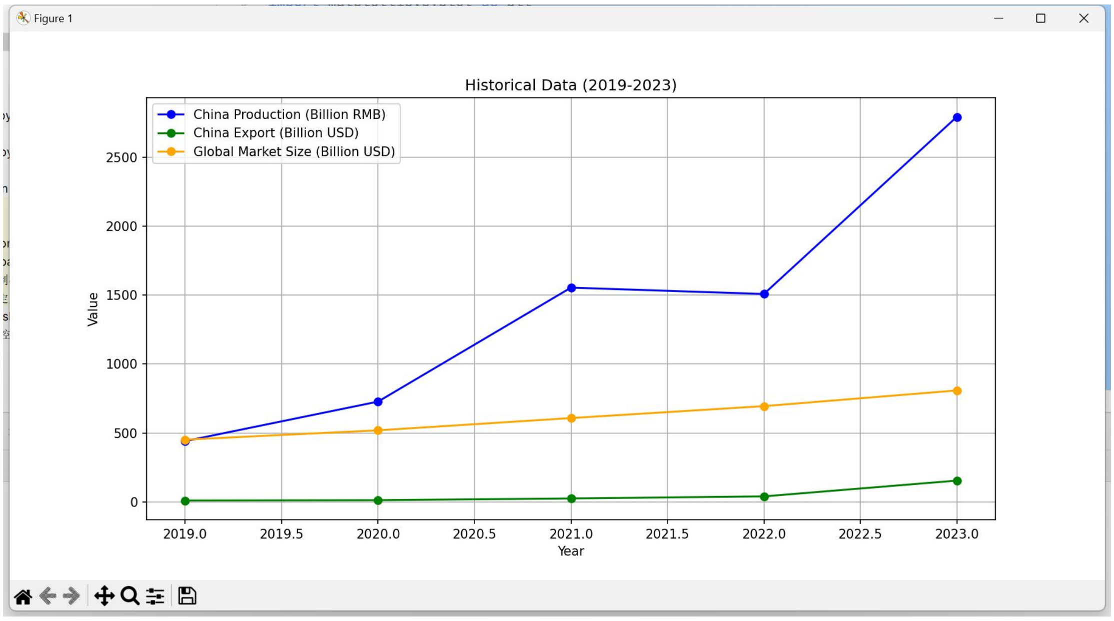


Figure 8

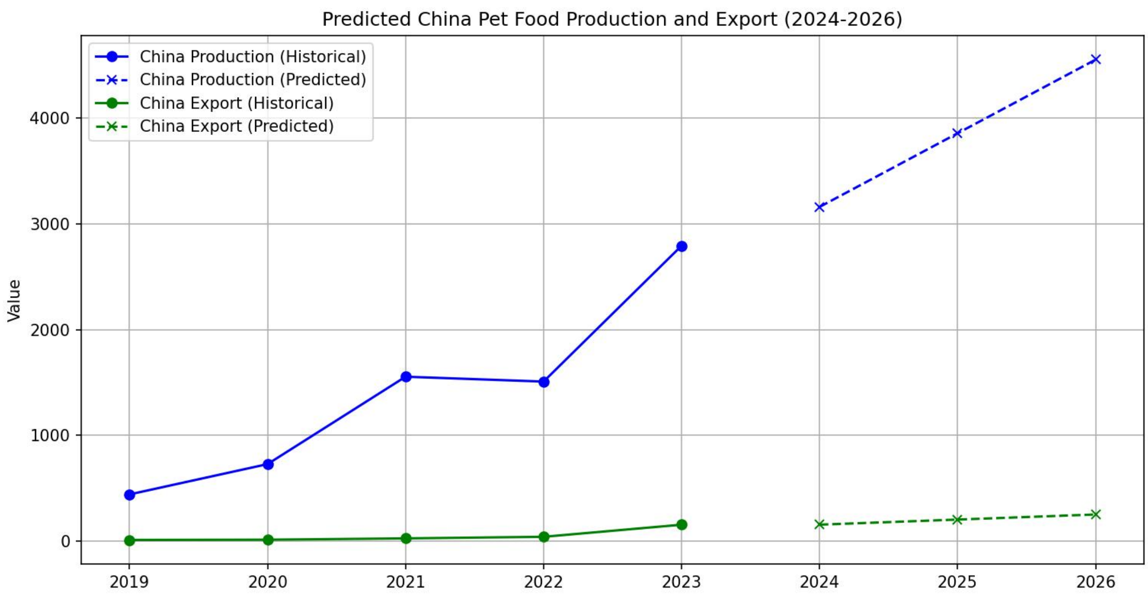


Figure 9

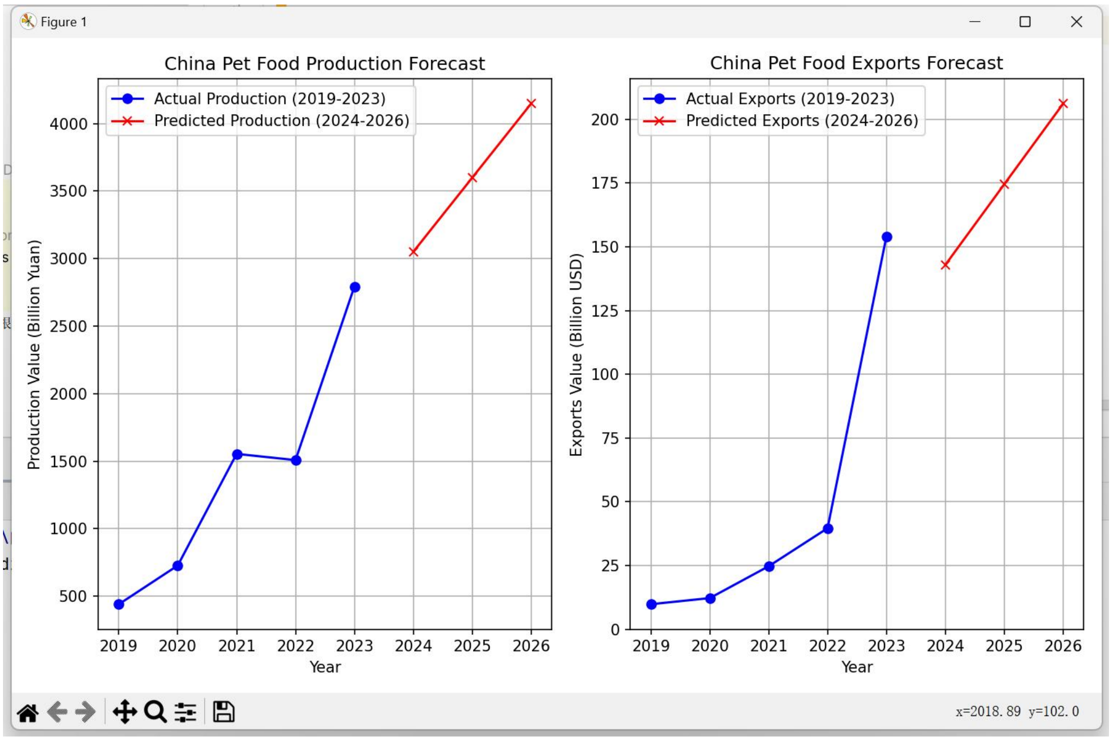


Figure 10

From 2019 to 2023, both the total production value and total export value of pet food in China showed significant growth. The production value increased from 44.07 billion RMB in 2019 to 279.3 billion RMB in 2023, with a high average annual growth rate. The export value rose from 980 million USD in 2019 to 15.41 billion USD in 2023, with an especially notable growth in exports, indicating that the competitiveness of China's pet food industry in the international market is continuously strengthening.

Using a multivariate linear regression model, and taking into account features such as the global pet food market size, urbanization rate, and per capita income, the future production and export situation of China's pet food industry was predicted for the next three years. The results show that the production value and export value will continue to maintain a strong growth trend:

- The forecasted production value for 2024 is approximately 312.3 billion RMB, with export value around 18.05 billion USD.
- The forecasted production value for 2025 is approximately 347.2 billion RMB, with export value around 20.94 billion USD.
- The forecasted production value for 2026 is approximately 382.1 billion RMB, with export value around 23.83 billion USD.

5.3.5 Conclusion and Application

Based on the analysis and forecast results, along with the production and export data of China's pet food industry and global market development trends, the following conclusions can be drawn: In the coming years, China's pet food industry will maintain a rapid growth trend, with both production and export values expected to increase significantly. Specifically, China's pet food production value is driven by the continuous growth in domestic demand, as well as the expansion of the global market and the acceleration of urbanization. The growth of export value shows that China's competitiveness in the international pet food market is continuously strengthening, particularly in terms of price and quality advantages to meet global consumer demands.

From a global perspective, the continuous expansion of the global pet food market, along with the steady improvement of urbanization rates and per capita income, provides a favorable external environment for China's pet food exports. This indicates that China has the ability to capture a larger share of the international market. However, Chinese companies need to pay attention to the intensifying competition in global markets and the rising consumer demands for product quality and diversification.

The Chinese pet food industry has significant growth potential but must continue to optimize production technologies, improve product quality, and deepen international expansion to better adapt to future changes and challenges in the global market.

5.4 Model Building and Analysis for Question Four

In this question, the focus is on the changes in the global economic environment and adjustments in international trade policies. New economic policies may have a significant impact on China's pet food exports. In particular, tariff increases in major markets such as Europe and the US could have a substantial negative effect on export volume and value. To analyze the impact of the current global trade policy changes on China's pet food industry, we have built a multivariate linear regression model. By simulating export changes under different scenarios, we aim to forecast the specific impact of new foreign trade economic policy adjustments on export value and propose counter-strategies.

5.4.1 Key Factors Explanation

1.Independent Variables (Features):

G_t : Global economic growth rate (value at time ttt), representing the growth or decline of the global economy.

C_t : Production cost growth rate (value at time ttt), representing changes in raw material, labor, and other costs during production.

T_t : Tariff impact (value at time ttt), representing the impact on production and exports due to changes in tariff policies, expressed as percentage changes.

R_t : Exchange rate change (value at time ttt), representing the appreciation or depreciation of the exchange rate; depreciation is positive, while appreciation is negative.

D_t : Market demand index (value at time ttt), representing fluctuations in market demand for pet food.

2.Dependent Variables (Targets):

Y_t : Pet food production value for a given year ttt (in billion RMB).

E_t : Pet food export value for a given year ttt (in million USD).

3.Pet Population Data (Influencing Factors):

$Cat_{China,t}$: Number of cats in China in year ttt (in ten thousand cats).

$Dog_{China,t}$: Number of dogs in China in year ttt (in ten thousand dogs).

Similarly, the pet population in other countries may also influence China's pet food production and exports.

5.4.2 Regression Model Construction

A multivariate linear regression model is used to establish the relationship, assuming that both production and export values are affected by these economic factors and pet populations. The basic form of the regression model is as follows:

$$Y_t = \alpha + \beta_1 G_t + \beta_2 G_t + \beta_3 T_t + \beta_4 R_t + \beta_5 D_t + \beta_6 \mathrm{Cat}_{\mathrm{China},\ t} + \beta_7 D\mathrm{og}_{\mathrm{China},\ t} + \ldots + \beta_n \mathrm{Cat}_{\mathrm{Country},\ t} + \varepsilon_t$$

Where:

- Y_t is the target variable (production value or export value).
- α is the intercept term, representing the baseline value when all influencing factors are zero.
- $\beta_1, \beta_2, \ldots \beta_n$ represents the regression coefficients, which indicate the extent to which each independent variable affects the target variable.
- ε_t is the error term, representing the random variation that the model cannot capture.

We model both the pet food production value and pet food export value, which gives us two regression equations:

Production Value Equation:

$$Y_{\mathrm{prod},t} = \alpha_{prod} + \beta_1 G_t + \beta_2 G_t + \beta_3 T_t + \beta_4 R_t + \beta_5 D_t + \beta_6 Cat_{China,t} + \ldots + \varepsilon_t$$

Export Value Equation:

$$E_{\mathrm{prod},t} = \alpha_{prod} + \beta_1 G_t + \beta_2 G_t + \beta_3 T_t + \beta_4 R_t + \beta_5 D_t + \beta_6 Cat_{China,t} + \ldots + \varepsilon_t$$

5.4.3 Regression Analysis

We use Ordinary Least Squares (OLS) to solve for the parameters of these regression equations. The objective is to minimize the sum of squared errors, which represents the difference between the target variable (production value or export value) and the actual observed values: $SSE = \sum_{\mathrm{t}=1}^{T} \left(Y_{\mathrm{pred},t} - Y_{true,t}\right)^2$

Objective: Minimize the sum of squared errors to solve for the regression coefficients.

The regression coefficients can be obtained by solving the following system of linear equations: $X^T X\beta = X^T \mathrm{y}$

Where X is the design matrix containing the independent variables, yyy is the vector of target variables, and β\betaβ is the vector of regression coefficients.

5.4.4 Model Evaluation

1. **R² (Coefficient of Determination)**: This measures how well the model fits the data. The value ranges from [0, 1], with values closer to 1 indicating stronger explanatory power of the model.
2. **Significance Testing**: A t-test is used to check whether each regression coefficient is significantly different from zero, ensuring that each feature variable has a meaningful impact in the model.

5.4.5 Regression Results Visualization

The regression coefficients obtained from the model can be visualized using bar charts to compare the impact of each factor on production and export values. For example:

Production Value Impact Coefficients Chart: Shows the extent to which various economic factors and pet populations affect production values.

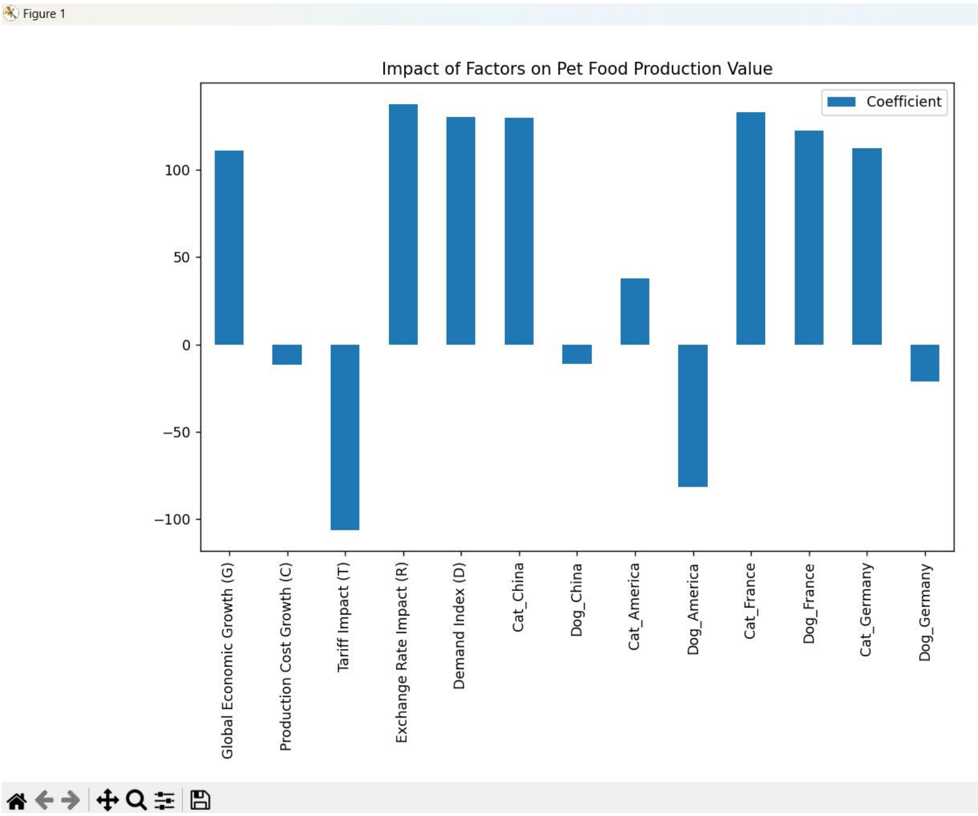


Figure 11

Export Value Impact Coefficients Chart: Shows the extent to which various economic factors and pet populations affect export values.

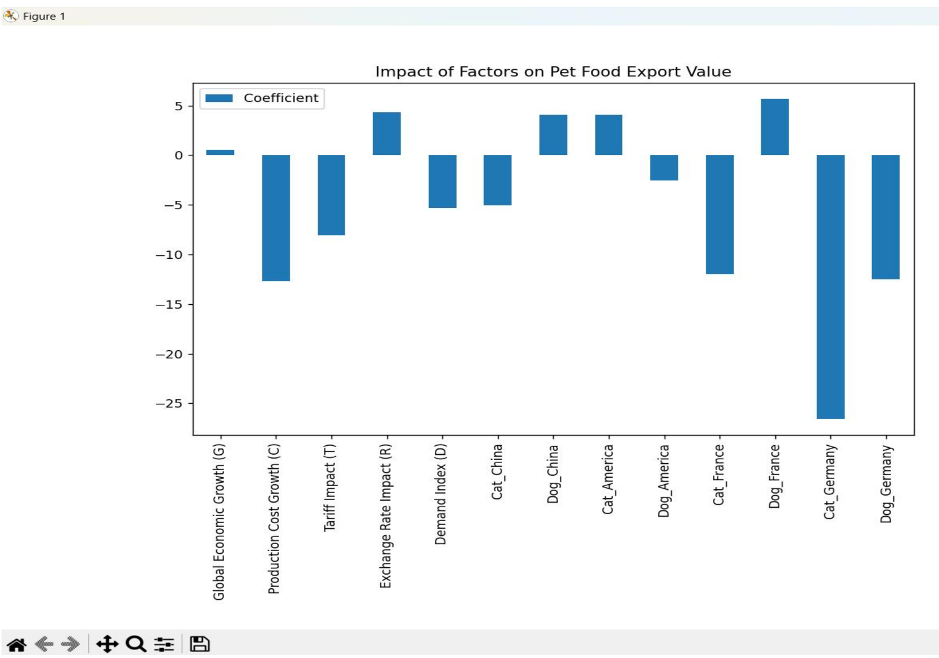


Figure 12

- **Production Value**: Market demand, global economic growth, exchange rate changes, and the number of cats have a significant positive impact on production value. Tariffs and production cost growth will significantly reduce production value.
- **Export Value**: Production costs, tariffs, market demand, and the number of cats have a strong negative impact on exports. Exchange rate changes and the number of dogs have a slight positive impact on exports.

5.4.6 Summary and Optimization

If production and export values are strongly influenced by tariffs, exchange rates, and market demand, they can be improved through policy adjustments, exchange rate control, or enhanced market promotion. The differing positive and negative impacts of cat and dog populations may be related to the market demand structure for production. A deeper analysis of the pet market demand in different countries may help optimize production and export strategies.

VI.Model Advantages, Disadvantages, and Outlook

6.1 Advantages and Disadvantages of the Model Used in This Study

In solving the four questions, we used weighted growth rate calculation models, time series analysis models, weighted average models, and linear growth trend models. Below are the advantages and disadvantages of each model, as well as their applications in the respective problems:

6.1.1 Weighted Growth Rate Calculation Model

This model calculates the market share of pet populations in different countries within the global market and uses these shares to weight the growth rates of each country, thereby deriving the global pet market's weighted growth rate. The advantage of this model is that it can comprehensively consider the influence of different countries on the global market, avoiding the bias of relying on data from a single market.

Advantages:

Comprehensiveness: By combining data from multiple countries and weighting the growth rates, the model can reflect the global pet market's changing trends more comprehensively.

Flexibility: The model can be adjusted based on different market changes and national economic policies.

Strong Adaptability: The model can easily accommodate new data, making it scalable and capable of incorporating new countries or regions for calculation.

Disadvantages:

Data Dependency: The accuracy of this model heavily depends on the quality and completeness of the input data. Missing or inaccurate data from certain countries may lead to biased predictions.

Exclusion of External Factors: The model assumes that market growth is influenced solely by internal changes within each country, not accounting for external factors such as global economic policies or natural disasters, which could impact the pet food market.

6.1.2 Time Series Analysis Model

This model predicts future trends based on historical data, effectively capturing the long-term development trends and seasonal fluctuations of the market. The model adjusts the global cat and dog market's growth trends through a weighted growth rate, providing reliable predictions for future market changes.

Advantages:

Trend Capture: Time series models are excellent at capturing long-term trends and seasonal fluctuations in historical data, offering insights into future market dynamics.

Strong Predictive Power: Based on historical data, the model can make reasonable predictions about the development of the pet food market in the coming years.

Strong Data Support: This method is suitable for scenarios with large and stable data volumes, where long-term historical data can reduce interference from random factors and provide more accurate predictions.

Disadvantages:

Sensitivity to Outliers: Time series models are sensitive to abnormal fluctuations in historical data. If extreme situations or data quality issues exist, the predictions may be inaccurate.

Exclusion of Non-Time Factors: The model only focuses on trends and seasonal variations in time series data and does not account for non-temporal factors, such as policy changes or shifts in consumer behavior, that could influence the market.

6.1.3 Linear Growth Trend Model

Advantages:

Higher Prediction Accuracy: This model uses a linear trend to predict future development, which is simple and efficient, suitable for markets with stable growth. Linear trends can effectively reflect market growth rates and cyclical changes.

Simple Calculation: The linear growth trend model relies on simple mathematical formulas, which are quick to calculate, making it suitable for real-time analysis and rapid forecasting.

Wide Applicability: This model is suitable for market data of all sizes, whether small-scale markets or global markets, and can provide reasonable predictions for both.

VII. Model Improvement and Expansion

7.1 Introducing More External Variables

Improvement Direction: The current model primarily relies on historical data and market share to predict future trends but does not fully consider external factors such as macroeconomic conditions, consumer behavior, or policy changes. To improve the model's accuracy, more external variables, such as the economic growth rate of countries, exchange rate fluctuations, consumer sentiment, and policy changes, can be incorporated. This will make the model more adaptive and dynamic.

Expansion Application: This improvement is not limited to predicting trends in the pet industry but can be applied to other consumer goods industries, especially those heavily impacted by international markets, policies, and macroeconomic changes. For example, by analyzing the impact of global economic changes on multinational enterprises, companies can develop more accurate marketing strategies.

7.2 Incorporating Technological Innovation and Product Progress

Improvement Direction: The current model focuses mainly on demand growth and market share changes but does not consider the long-term effects of technological innovations and product advancements on industry development. With the continuous emergence of new technologies in the pet food industry (such as plant-based foods and personalized pet foods), these innovations will reshape the market. Thus, including factors such as technological progress and innovative products in the model will provide a more comprehensive reflection of future industry trends.

Expansion Application: This improvement can be applied to other rapidly developing industries, such as smart home technology and green energy. As technology evolves, industry structures and market demands will undergo significant changes, and predictive models should be able to capture these changes and adapt to the needs of various markets.

7.3 Integration of Big Data and Real-Time Data

Improvement Direction: The existing model’s predictions mainly rely on historical data and do not fully leverage the advantages of current big data technologies. With the increase in real-time data from social media, consumer surveys, sensor data, etc., these data sources can be integrated into the model to improve prediction accuracy and timeliness.

Expansion Application: Big data technology can be applied across various industries, particularly in fields like e-commerce, finance, and healthcare. Real-time data can help businesses quickly adjust their strategies in rapidly changing markets. For example, by analyzing consumer behavior on social media platforms, companies can capture trend shifts in real time and respond faster to market demands.

VIII. References

[1] Wang Huiwen, Meng Jie. "Multivariate Linear Regression Prediction Modeling Method" [J]. Journal of Beihang University, 2007, (04): 500-504. DOI: 10.13700/j.bh.1001-5965.2007.04.028.

[2] Qian Ying, Fang Xiunan. "Multivariate Linear Regression Model and Its Application Examples" [J]. China Science and Technology Information, 2022, (04): 73-74.

[3] Chen Laihua. "Current Situation, Influencing Factors, and Development Trends of China’s Pet Industry" [J]. China Animal Health, 2018, 20(08): 4-8.

[4] Chen Weicai. "Current Situation and Development Trends of China’s Pet Industry" [J]. China Detergent Industry, 2019, (08): 56-59. DOI: 10.16054/j.cnki.cci.2019.08.008.

[5] Cao Shumei, Zhang Xiaoman, Li Jia. "A Review of the Pet Food Industry Research" [J]. Cooperative Economy and Technology, 2022, (13): 98-99. DOI: 10.13665/j.cnki.hzjjykj.2022.13.045.

IX.Appendix

Appendix Question 1
Introduction: Analysis and Forecast of China's Pet Industry Growth (2019-2026)

```
import pandas as pd
import statsmodels.api as sm
import matplotlib.pyplot as plt
import numpy as np

# Step 1: Create data
data = {
    'Year': [2019, 2020, 2021, 2022, 2023],
    'Cat Population': [4412, 4862, 5806, 6536, 6980],
    'Dog Population': [5503, 5222, 5429, 5119, 5175],
    'Economic Growth Rate (%)': [6, 2.3, 8.1, 3, 5.2],
    'Per Capita GDP (Yuan)': [70892, 72447, 80976, 85698, 89358],
    'Urbanization Rate (%)': [60.6, 63.89, 64.72, 65.22, 66.3],
    'Per Capita Disposable Income (Yuan)': [42359, 43834, 47412, 48283, 51821],
    'Pet Market Size (Billion Yuan)': [2000, 2200, 2400, 2700, 2793],
    'Pet Supplies Market Size (Billion Yuan)': [250, 275, 300, 337.5, 349.125],
    'Average Household Size (People)': [2.65, 2.62, 2.6, 2.58, 2.55],
    'Percentage of Population Aged 60 and Above (%)': [18, 18.7, 19, 19.5, 20]
}

# Step 2: Convert data to DataFrame
df = pd.DataFrame(data)

# Step 3: Select independent and dependent variables
X = df[['Economic Growth Rate (%)', 'Per Capita GDP (Yuan)', 'Urbanization Rate (%)',
        'Per Capita Disposable Income (Yuan)', 'Pet Market Size (Billion Yuan)',
        'Pet Supplies Market Size (Billion Yuan)', 'Average Household Size (People)',
        'Percentage of Population Aged 60 and Above (%)']]

# Add constant term
X = sm.add_constant(X)

# Choose "Cat Population" as dependent variable
y_cat = df['Cat Population']
y_dog = df['Dog Population']

# Step 4: Fit regression models
model_cat = sm.OLS(y_cat, X).fit()
model_dog = sm.OLS(y_dog, X).fit()

# Step 5: Output regression results
print("Cat Population Model Summary:")
print(model_cat.summary())

print("\nDog Population Model Summary:")
print(model_dog.summary())

# Step 6: Use the models to predict future pet population
future_data = {
    'Economic Growth Rate (%)': [4, 4.5, 5],
    'Per Capita GDP (Yuan)': [92000, 95000, 98000],
```

```
    'Urbanization Rate (%)': [67, 68, 69],
    'Per Capita Disposable Income (Yuan)': [53000, 55000, 57000],
    'Pet Market Size (Billion Yuan)': [2900, 3100, 3300],
    'Pet Supplies Market Size (Billion Yuan)': [360, 380, 400],
    'Average Household Size (People)': [2.54, 2.53, 2.52],
    'Percentage of Population Aged 60 and Above (%)': [20.5, 21, 21.5]
}

# Convert future data to DataFrame
future_df = pd.DataFrame(future_data)

# Add constant term
future_df = sm.add_constant(future_df)

# Use the models to predict future populations
future_predictions_cat = model_cat.predict(future_df)
future_predictions_dog = model_dog.predict(future_df)

# Step 7: Plotting the results

# Plot the actual data (2019-2023) and the predictions
plt.figure(figsize=(10, 6))

# Plot actual cat and dog population data (2019-2023)
plt.plot(df['Year'], df['Cat Population'], label='Actual Cat Population', color='blue',
marker='o')
plt.plot(df['Year'], df['Dog Population'], label='Actual Dog Population', color='green',
marker='o')

# Plot predicted cat and dog population for future years (2024, 2025, 2026)
future_years = [2024, 2025, 2026]
plt.plot(future_years, future_predictions_cat, label='Predicted Cat Population', color='red',
marker='x')
plt.plot(future_years, future_predictions_dog, label='Predicted Dog Population',
color='orange', marker='x')

# Add titles and labels
plt.title('Cat and Dog Population Growth and Future Predictions (2019-2026)')
plt.xlabel('Year')
plt.ylabel('Population (in 10,000)')
plt.xticks(np.arange(2019, 2027, step=1))  # Show years 2019-2026 on the x-axis
plt.legend()

# Show the plot
plt.grid(True)
plt.show()
```

Appendix Question 1
Introduction: Analysis and Forecast of China's Pet Market Size (2019-2026)

```
import pandas as pd
import statsmodels.api as sm
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.stats.outliers_influence import variance_inflation_factor

# Step 1: Create Data
data = {
    'Year': [2019, 2020, 2021, 2022, 2023],
    'Cat Population': [4412, 4862, 5806, 6536, 6980],
    'Dog Population': [5503, 5222, 5429, 5119, 5175],
    'Economic Growth Rate (%)': [6, 2.3, 8.1, 3, 5.2],
    'Per Capita GDP (Yuan)': [70892, 72447, 80976, 85698, 89358],
    'Urbanization Rate (%)': [60.6, 63.89, 64.72, 65.22, 66.3],
    'Per Capita Disposable Income (Yuan)': [42359, 43834, 47412, 48283, 51821],
    'Pet Market Size (Billion Yuan)': [2000, 2200, 2400, 2700, 2793],
    'Pet Supplies Market Size (Billion Yuan)': [250, 275, 300, 337.5, 349.125],
    'Average Household Size (People)': [2.65, 2.62, 2.6, 2.58, 2.55],
    'Percentage of Population Aged 60 and Above (%)': [18, 18.7, 19, 19.5, 20]
}

# Step 2: Convert to DataFrame
df = pd.DataFrame(data)

# Step 3: Select independent and dependent variables
X = df[['Economic Growth Rate (%)', 'Per Capita GDP (Yuan)', 'Urbanization Rate (%)',
        'Per Capita Disposable Income (Yuan)', 'Pet Market Size (Billion Yuan)',
        'Pet Supplies Market Size (Billion Yuan)', 'Average Household Size (People)',
        'Percentage of Population Aged 60 and Above (%)']]

# Add constant term (for intercept)
X = sm.add_constant(X)

# Choose 'Pet Market Size (Billion Yuan)' as the dependent variable
y_pet_market = df['Pet Market Size (Billion Yuan)']

# Step 4: Fit regression model
model_pet_market = sm.OLS(y_pet_market, X).fit()

# Step 5: Output regression results
print("Pet Market Size Model Summary:")
print(model_pet_market.summary())

# Step 6: Variance Inflation Factor (VIF) check for multicollinearity
vif_data = pd.DataFrame()
vif_data["feature"] = X.columns
vif_data["VIF"] = [variance_inflation_factor(X.values, i) for i in range(X.shape[1])]
print("\nVariance Inflation Factor (VIF) values:")
print(vif_data)

# Step 7: Predict future values (for the next 3 years: 2024, 2025, 2026)
future_data = {
    'Economic Growth Rate (%)': [4, 4.5, 5],
    'Per Capita GDP (Yuan)': [92000, 95000, 98000],
    'Urbanization Rate (%)': [67, 68, 69],
    'Per Capita Disposable Income (Yuan)': [53000, 55000, 57000],
    'Pet Market Size (Billion Yuan)': [2900, 3100, 3300],
```

```
    'Pet Supplies Market Size (Billion Yuan)': [360, 380, 400],
    'Average Household Size (People)': [2.54, 2.53, 2.52],
    'Percentage of Population Aged 60 and Above (%)': [20.5, 21, 21.5]
}

# Convert future data to DataFrame
future_df = pd.DataFrame(future_data)

# Add constant term (for intercept)
future_df = sm.add_constant(future_df)

# Step 8: Use the model to predict future pet market size
future_predictions_pet_market = model_pet_market.predict(future_df)

# Step 9: Plotting the results

# Plot the actual data (2019-2023) and predicted data (2024-2026)
plt.figure(figsize=(10, 6))

# Plot actual pet market size data (2019-2023)
plt.plot(df['Year'], df['Pet Market Size (Billion Yuan)'], label='Actual Pet Market Size',
color='blue', marker='o')

# Plot predicted pet market size for future years (2024, 2025, 2026)
future_years = [2024, 2025, 2026]
plt.plot(future_years, future_predictions_pet_market, label='Predicted Pet Market Size',
color='red', marker='x')

# Add titles and labels
plt.title('Pet Market Size Growth and Future Predictions (2019-2026)')
plt.xlabel('Year')
plt.ylabel('Pet Market Size (Billion Yuan)')
plt.xticks(np.arange(2019, 2027, step=1))  # Show years 2019-2026 on the x-axis
plt.legend()

# Show the plot
plt.grid(True)
plt.show()
```

Appendix Question 1
Introduction: Regression Analysis of Cat and Dog Population Growth

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression

# Data construction
data = {
    'Year': [2019, 2020, 2021, 2022, 2023],
    'Cat Population': [4412, 4862, 5806, 6536, 6980],
    'Dog Population': [5503, 5222, 5429, 5119, 5175],
    'Economic Growth Rate': [6, 2.3, 8.1, 3, 5.2],
    'Per Capita GDP': [70892, 72447, 80976, 85698, 89358],
    'Urbanization Rate': [60.6, 63.89, 64.72, 65.22, 66.3],
    'Per Capita Disposable Income': [42359, 43834, 47412, 48283, 51821],
    'Pet Market Size': [2000, 2200, 2400, 2700, 2793],
    'Pet Supplies Market Size': [250, 275, 300, 337.5, 349.125],
    'Percentage of Population Aged 60 and Above': [18, 18.7, 19, 19.5, 20],
    'Urban Population': [84843, 90014, 91517, 92071, 93267]
}

# Create DataFrame
df = pd.DataFrame(data)

# Calculate growth rates for Cat and Dog populations
df['Cat_Population_Growth'] = df['Cat Population'].pct_change() * 100
df['Dog_Population_Growth'] = df['Dog Population'].pct_change() * 100

# Correlation analysis
correlation_matrix = df[['Cat_Population_Growth', 'Dog_Population_Growth',
                         'Economic Growth Rate', 'Per Capita GDP',
                         'Urbanization Rate', 'Per Capita Disposable Income',
                         'Percentage of Population Aged 60 and Above']].corr()

# Output correlation matrix
print("Correlation Matrix:")
print(correlation_matrix)

# Visualize the correlation matrix
plt.figure(figsize=(10, 8))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", fmt='.2f')
plt.title("Correlation Matrix of Variables")
plt.show()

# Prepare data for regression analysis
X = df[['Economic Growth Rate', 'Per Capita GDP', 'Urbanization Rate',
        'Per Capita Disposable Income', 'Percentage of Population Aged 60 and Above']]
y_cat = df['Cat_Population_Growth']
y_dog = df['Dog_Population_Growth']

# Build regression models
model_cat = LinearRegression()
model_cat.fit(X, y_cat)

model_dog = LinearRegression()
model_dog.fit(X, y_dog)
```

```
# Output regression coefficients
print("Cat Population Growth Regression Coefficients:", model_cat.coef_)
print("Dog Population Growth Regression Coefficients:", model_dog.coef_)

# Predict cat and dog population growth
cat_pred = model_cat.predict(X)
dog_pred = model_dog.predict(X)

# Visualize regression results
plt.figure(figsize=(12, 6))

# Cat population growth prediction vs actual
plt.subplot(1, 2, 1)
plt.plot(df['Year'], y_cat, label='Actual Cat Population Growth', marker='o')
plt.plot(df['Year'], cat_pred, label='Predicted Cat Population Growth', linestyle='--')
plt.title("Cat Population Growth Prediction vs Actual")
plt.xlabel("Year")
plt.ylabel("Cat Population Growth Rate (%)")
plt.legend()

# Dog population growth prediction vs actual
plt.subplot(1, 2, 2)
plt.plot(df['Year'], y_dog, label='Actual Dog Population Growth', marker='o')
plt.plot(df['Year'], dog_pred, label='Predicted Dog Population Growth', linestyle='--')
plt.title("Dog Population Growth Prediction vs Actual")
plt.xlabel("Year")
plt.ylabel("Dog Population Growth Rate (%)")
plt.legend()

plt.tight_layout()
plt.show()
```

Appendix Question 1
Introduction: Analysis of Factors Influencing Pet Populations (Cat and Dog) Over Time

```
import pandas as pd
import matplotlib.pyplot as plt

# Step 1: Create Data
data = {
    'Year': [2019, 2020, 2021, 2022, 2023],
    'Cat Population': [4412, 4862, 5806, 6536, 6980],
    'Dog Population': [5503, 5222, 5429, 5119, 5175],
    'Economic Growth Rate (%)': [6, 2.3, 8.1, 3, 5.2],
    'Per Capita GDP (Yuan)': [70892, 72447, 80976, 85698, 89358],
    'Urbanization Rate (%)': [60.6, 63.89, 64.72, 65.22, 66.3],
    'Per Capita Disposable Income (Yuan)': [42359, 43834, 47412, 48283, 51821],
    'Pet Market Size (Billion Yuan)': [2000, 2200, 2400, 2700, 2793],
    'Pet Supplies Market Size (Billion Yuan)': [250, 275, 300, 337.5, 349.125],
    'Average Household Size (People)': [2.65, 2.62, 2.6, 2.58, 2.55],
    'Percentage of Population Aged 60 and Above (%)': [18, 18.7, 19, 19.5, 20]
}

# Step 2: Convert to DataFrame
df = pd.DataFrame(data)

# Step 3: Define the independent variables (indicators to plot against)
indicators = [
    'Economic Growth Rate (%)',
    'Per Capita GDP (Yuan)',
    'Urbanization Rate (%)',
    'Per Capita Disposable Income (Yuan)',
    'Pet Market Size (Billion Yuan)',
    'Pet Supplies Market Size (Billion Yuan)',
    'Average Household Size (People)',
    'Percentage of Population Aged 60 and Above (%)'
]

# Step 4: Plot for each indicator
plt.figure(figsize=(12, 10))  # Adjust figure size for clarity

# Loop through each indicator and plot the cat and dog populations against it
for i, indicator in enumerate(indicators, 1):
    plt.subplot(4, 2, i)  # Create a 4x2 grid of plots (4 rows, 2 columns)
    plt.plot(df[indicator], df['Cat Population'], label='Cat Population', marker='o', color='blue')
    plt.plot(df[indicator], df['Dog Population'], label='Dog Population', marker='x', color='red')
    plt.title(f'Relationship between {indicator} and Animal Population')
    plt.xlabel(indicator)
    plt.ylabel('Population')
    plt.legend()

# Step 5: Adjust layout and show the plot
plt.tight_layout(pad=4.0)  # Add padding between subplots to prevent overlap
plt.show()
```

Appendix Question2
Introduction：Global Pet Food Sales from 2010 to 2023

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
data = {
    'Year': [2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2022, 2023],
    'Sales (Billion USD)': [59.30, 62.40, 65.90, 69.60, 73.30, 77.70, 75.25, 75.00, 91.10, 93.90,
102.60, 123.60, 133.90]
}
df = pd.DataFrame(data)
X = df[['Year']]
y = df['Sales (Billion USD)']
model = LinearRegression()
model.fit(X, y)
future_years = np.array([[2024], [2025], [2026]])
predicted_sales = model.predict(future_years)
for year, sales in zip([2024, 2025, 2026], predicted_sales):
    print(f"Predicted Sales for {year}: {sales:.2f} Billion USD")
plt.figure(figsize=(10, 6))
plt.scatter(df['Year'], df['Sales (Billion USD)'], color='blue', label='Historical Data')
plt.plot(df['Year'], model.predict(X), color='red', label='Fitted Line')
plt.plot(future_years, predicted_sales, color='green', linestyle='--', label='Predicted Data')
plt.xlabel('Year')
plt.ylabel('Sales (Billion USD)')
plt.title('Global Pet Food Market Sales Prediction')
plt.legend()
plt.show()
```

Appendix Question2
Introduction：Forecasting Global Pet Food Market Sales (2010–2026) Using ARIMA

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller
from statsmodels.tsa.arima.model import ARIMA
from pandas.plotting import autocorrelation_plot

# Data preparation
data = {
    'Year': [2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2022, 2023],
    'Sales (Billion USD)': [59.30, 62.40, 65.90, 69.60, 73.30, 77.70, 75.25, 75.00, 91.10, 93.90,
102.60, 123.60,
                          133.90]
}
df = pd.DataFrame(data)

# Convert Year to datetime format
df['Year'] = pd.to_datetime(df['Year'], format='%Y')

# Set Year as index
df.set_index('Year', inplace=True)

# Plot the time series
plt.figure(figsize=(10, 6))
plt.plot(df['Sales (Billion USD)'], marker='o', linestyle='-', color='b')
plt.title('Global Pet Food Market Sales (2010-2023)')
plt.xlabel('Year')
plt.ylabel('Sales (Billion USD)')
plt.grid(True)
plt.show()

# Perform ADF test to check stationarity
result = adfuller(df['Sales (Billion USD)'])
print(f'ADF Statistic: {result[0]}')
print(f'p-value: {result[1]}')

# If p-value > 0.05, data is non-stationary, require differencing
if result[1] > 0.05:
    df_diff = df['Sales (Billion USD)'].diff().dropna()

    # Perform ADF test again after differencing
    result_diff = adfuller(df_diff)
    print(f'\nADF Statistic (After Differencing): {result_diff[0]}')
    print(f'p-value (After Differencing): {result_diff[1]}')
else:
```

```
    df_diff = df['Sales (Billion USD)']

# Fit ARIMA model
# Assuming ARIMA(1,1,1) as the model parameters; adjust based on AIC, BIC
model = ARIMA(df['Sales (Billion USD)'], order=(1, 1, 1))
model_fit = model.fit()

# Display model summary
print(model_fit.summary())

# Plot residuals
residuals = model_fit.resid
plt.figure(figsize=(10, 6))
plt.plot(residuals)
plt.title('Residuals of ARIMA Model')
plt.grid(True)
plt.show()

# Plot autocorrelation of residuals to check for white noise
autocorrelation_plot(residuals)
plt.show()

# Forecast sales for the next 3 years (2024, 2025, 2026)
future_years = pd.to_datetime([2024, 2025, 2026], format='%Y')
forecast = model_fit.forecast(steps=3)

# Display forecasted results
for year, sales in zip(future_years, forecast):
    print(f"Predicted Sales for {year.year}: {sales:.2f} Billion USD")

# Plot historical and forecasted data
plt.figure(figsize=(10, 6))
plt.plot(df['Sales (Billion USD)'], marker='o', linestyle='-', color='b', label='Historical Data')
plt.plot(future_years, forecast, marker='o', linestyle='--', color='g', label='Forecasted Data')
plt.title('Global Pet Food Market Sales (2010-2026)')
plt.xlabel('Year')
plt.ylabel('Sales (Billion USD)')
plt.legend()
plt.grid(True)
plt.show()
```

Appendix Question2
Introduction: Predicting Global Pet Food Market Demand (2024-2026)

```
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegressi
import matplotlib.pyplot as plt

# Input data: Number of cats and dogs in each country from 2019 to 2023 (in thousands)
years = np.array([2019, 2020, 2021, 2022, 2023])

usa_cat = np.array([9420, 6500, 9420, 7380, 7380])
usa_dog = np.array([8970, 8500, 8970, 8970, 8010])

france_cat = np.array([1300, 1490, 1510, 1490, 1660])
france_dog = np.array([740, 775, 750, 760, 990])

germany_cat = np.array([1470, 1570, 1670, 1520, 1570])
germany_dog = np.array([1010, 1070, 1030, 1060, 1050])

# Assumed factors: Urbanization rate, per capita income, economic growth rate
urbanization_usa = np.array([82.5, 82.7, 82.8, 82.9, 83])
income_usa = np.array([63000, 65000, 66000, 68000, 70000])
growth_rate_usa = np.array([2.3, -3.4, 5.7, 2.1, 1.9])

urbanization_france = np.array([80.4, 80.8, 81.0, 81.2, 81.4])
income_france = np.array([42000, 41000, 43000, 44000, 45000])
growth_rate_france = np.array([1.5, -8.0, 7.0, 2.6, 1.5])

urbanization_germany = np.array([77.3, 77.5, 77.6, 77.7, 77.8])
income_germany = np.array([48000, 47000, 49000, 50000, 51000])
growth_rate_germany = np.array([0.6, -4.9, 2.9, 1.8, 1.2])

# Assumed pet food market demand in million USD (based on each country's data)
usa_pet_food_market = np.array([360, 420, 500, 581, 690.5])
france_pet_food_market = np.array([41.8, 46.2, 52.8, 55, 57.2])
germany_pet_food_market = np.array([49.5, 51.7, 55, 58.3, 60.5])

# Merge all data to form feature matrix
X_usa = np.column_stack((usa_cat, usa_dog, urbanization_usa, income_usa,
growth_rate_usa))
X_france = np.column_stack((france_cat, france_dog, urbanization_france, income_france,
growth_rate_france))
X_germany = np.column_stack((germany_cat, germany_dog, urbanization_germany,
income_germany, growth_rate_germany))

# Merge features and target variable from all countries
X = np.vstack((X_usa, X_france, X_germany))
```

```
y = np.concatenate((usa_pet_food_market, france_pet_food_market,
germany_pet_food_market))

# Create regression model
model = LinearRegression()

# Train the regression model
model.fit(X, y)

# Predict global pet food demand from 2024 to 2026
future_years = np.array([2024, 2025, 2026])

# Assumed cat and dog numbers, urbanization rate, per capita income, and economic growth
rate for future years
future_usa_cat = np.array([7400, 7500, 7600])
future_usa_dog = np.array([8000, 8100, 8200])
future_urbanization_usa = np.array([83.1, 83.2, 83.3])
future_income_usa = np.array([71000, 72000, 73000])
future_growth_rate_usa = np.array([2.0, 2.1, 2.2])

future_france_cat = np.array([1700, 1750, 1800])
future_france_dog = np.array([1000, 1020, 1040])
future_urbanization_france = np.array([81.5, 81.6, 81.7])
future_income_france = np.array([46000, 47000, 48000])
future_growth_rate_france = np.array([1.6, 1.7, 1.8])

future_germany_cat = np.array([1600, 1620, 1640])
future_germany_dog = np.array([1050, 1060, 1070])
future_urbanization_germany = np.array([77.9, 78.0, 78.1])
future_income_germany = np.array([52000, 53000, 54000])
future_growth_rate_germany = np.array([1.3, 1.4, 1.5])

# Create feature matrix for future years
X_future_usa = np.column_stack((future_usa_cat, future_usa_dog, future_urbanization_usa,
future_income_usa, future_growth_rate_usa))
X_future_france = np.column_stack((future_france_cat, future_france_dog,
future_urbanization_france, future_income_france, future_growth_rate_france))
X_future_germany = np.column_stack((future_germany_cat, future_germany_dog,
future_urbanization_germany, future_income_germany, future_growth_rate_germany))

X_future = np.vstack((X_future_usa, X_future_france, X_future_germany))

# Predict the pet food market demand for the next three years
future_predictions = model.predict(X_future)

# Print the prediction results
print("Predicted Global Pet Food Demand for 2024-2026 (Million USD):")
for year, demand in zip(future_years, future_predictions):
```

```
    print(f"Year {year}: {demand:.2f} Million USD")

# Calculate the total global demand for each year
total_future_demand = future_predictions.sum()

# Visualize the results (total global demand)
plt.figure(figsize=(10, 6))
plt.plot([2024, 2025, 2026], future_predictions[:3], label="Total Predicted Global Demand (Million USD)", color='r', marker='o')

# Set up the chart
plt.title('Predicted Global Pet Food Market Demand (2024-2026)')
plt.xlabel('Year')
plt.ylabel('Demand (Million USD)')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

Appendix Question2
Introduction: Global Pet Population Growth Trends (2020-2023)

```
import numpy as np
import matplotlib.pyplot as plt


data = {
    'America': {
        'Cat': [9420, 6500, 9420, 7380, 7380],
        'Dog': [8970, 8500, 8970, 8010, 8970]
    },
    'France': {
        'Cat': [1300, 1490, 1510, 1490, 1660],
        'Dog': [740, 775, 750, 990, 760]
    },
    'Germany': {
        'Cat': [1470, 1570, 1670, 1520, 1570],
        'Dog': [1010, 1070, 1030, 1060, 1050]
    },
    'China': {
        'Cat': [4412, 4862, 5806, 6536, 6980],
        'Dog': [5503, 5222, 5429, 5119, 5175]
    }
}

def calculate_shares():
    global_cat_total = sum([sum(data[country]['Cat']) for country in data])  # Global total of
cat population
    global_dog_total = sum([sum(data[country]['Dog']) for country in data])  # Global total of
dog population

    cat_shares = {}
    dog_shares = {}

    for country in data:
        cat_total = sum(data[country]['Cat'])
        dog_total = sum(data[country]['Dog'])
        cat_shares[country] = cat_total / global_cat_total
        dog_shares[country] = dog_total / global_dog_total

    return cat_shares, dog_shares

def calculate_weighted_growth_rate(shares, growth_rates):
    weighted_growth_rate = {'Cat': [0] * 4, 'Dog': [0] * 4}

    for pet in ['Cat', 'Dog']:
        for country, share in shares.items():
```

```
            for year in range(4):
                weighted_growth_rate[pet][year] += growth_rates[country][pet][year] * share

    return weighted_growth_rate

growth_rates = {
    'America': {
        'Cat': [-30.26, 44.92, -21.98, 0],
        'Dog': [-5.23, 5.53, 0, -10.71]
    },
    'France': {
        'Cat': [14.62, 1.34, -1.33, 11.41],
        'Dog': [4.73, -3.23, 1.33, 30.26]
    },
    'Germany': {
        'Cat': [6.80, 6.37, -9.01, 3.29],
        'Dog': [5.96, -3.74, 2.91, -0.94]
    },
    'China': {
        'Cat': [10.22, 19.42, 12.60, 6.78],
        'Dog': [-5.09, 3.97, -5.72, 1.09]
    }
}

cat_shares, dog_shares = calculate_shares()

weighted_growth_rate = calculate_weighted_growth_rate(cat_shares, growth_rates)

print("Weighted Growth Rates (in %):")
for pet in ['Cat', 'Dog']:
    print(f"{pet} Weighted Growth Rates: {weighted_growth_rate[pet]}")

global_cat_start = 6980
global_dog_start = 5175

global_cat_trend = [global_cat_start]
global_dog_trend = [global_dog_start]

for i in range(4):
    global_cat_trend.append(global_cat_trend[-1] * (1 + weighted_growth_rate['Cat'][i] / 100))
    global_dog_trend.append(global_dog_trend[-1] * (1 + weighted_growth_rate['Dog'][i] / 
100))

full_years = ['2020', '2021', '2022', '2023']

plt.figure(figsize=(10, 6))
```

```
plt.plot(full_years, global_cat_trend[1:], label='Global Cat Growth Trend', marker="o",
color='tab:blue')
plt.plot(full_years, global_dog_trend[1:], label='Global Dog Growth Trend', marker="o",
color='tab:orange')

plt.title('Global Pet Growth Trends (Cat and Dog) (2020-2023)', fontsize=16)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Estimated Population (Million)', fontsize=12)
plt.xticks(rotation=45)
plt.legend()

plt.tight_layout()
plt.show()
```

Appendix Question3
Introduction： Analysis of China's Pet Food Industry Growth and Global Market Comparison

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

# China's Pet Food Production and Export Data (2019-2023) (RMB/USD)
years = [2019, 2020, 2021, 2022, 2023]
china_production_value = [440.7, 727.3, 1554, 1508, 2793]  # Unit: billion RMB
china_exports_value = [9.8, 12.2, 24.7, 39.6, 154.1]  # Unit: billion USD

# China's Pet Market and Economic Data (2015-2023)
china_pet_market_size = [980, 1220, 1460, 1710, 2000, 2200, 2400, 2700, 2793]  # Unit:
billion RMB
years_pet_market = [2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023]

# Global Pet Market (Assumed Data)
global_pet_market = [1500, 1700, 1900, 2200, 2500]  # Assumed data, unit: billion RMB

# Economic and Urbanization Data (Units: GDP: RMB, Income: RMB, Urbanization: %)
gdp_per_capita = [49394, 53259, 59660, 64644, 70892, 72447, 80976, 85698, 89358]  #
Unit: RMB
per_capita_income = [31195, 33616, 36396, 39260, 42359, 43834, 47412, 48283, 51821]  #
Unit: RMB
urbanization_rate = [56.1, 57.4, 58.5, 59.6, 60.6, 63.89, 64.72, 65.22, 66.3]  # Urbanization
Rate (%)

# Calculate Growth Rates: Growth Rate = (Current - Previous) / Previous * 100
production_growth_rate = np.diff(china_production_value) / china_production_value[:-1] *
100
exports_growth_rate = np.diff(china_exports_value) / china_exports_value[:-1] * 100

# Compile Growth Rate Data into a DataFrame
china_growth_data = pd.DataFrame({
    "Year": years[1:],  # From 2020 to 2023
    "Production Growth Rate (%)": production_growth_rate,
    "Exports Growth Rate (%)": exports_growth_rate
})

# Visualization: China's Pet Food Production vs Exports
plt.figure(figsize=(10, 6))
plt.plot(years, china_production_value, label="China Pet Food Production (Billion RMB)",
marker='o', color='b')
plt.plot(years, china_exports_value, label="China Pet Food Exports (Billion USD)",
marker='x', color='g')

plt.xlabel('Year')
```

```
plt.ylabel('Value (Billion)')
plt.title('China Pet Food Production vs Exports (2019-2023)')
plt.legend()
plt.grid(True)
plt.show()

# Visualization: China vs Global Pet Market Size
plt.figure(figsize=(10, 6))
plt.plot(years_pet_market, china_pet_market_size, label="China Pet Market Size (Billion RMB)", marker='o', color='b')
plt.plot(years, global_pet_market, label="Global Pet Market Size (Billion RMB)", marker='x', color='r')

plt.xlabel('Year')
plt.ylabel('Market Size (Billion RMB)')
plt.title('China vs Global Pet Market Size (2015-2023)')
plt.legend()
plt.grid(True)
plt.show()

# Output Growth Rate Data
print(china_growth_data)

# Output Economic and Urbanization Trends
economic_data = pd.DataFrame({
    "Year": years_pet_market,
    "GDP per Capita (RMB)": gdp_per_capita,
    "Per Capita Income (RMB)": per_capita_income,
    "Urbanization Rate (%)": urbanization_rate
})

print(economic_data)
```

Appendix Question3
Introduction：Pet Food Production and Export Prediction

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error

# Data Preparation
years = np.array([2019, 2020, 2021, 2022, 2023]).reshape(-1, 1)
production_values = np.array([440.7, 727.3, 1554, 1508, 2793])  # Production values (Billion RMB)
export_values = np.array([9.8, 12.2, 24.7, 39.6, 154.1])  # Export values (Billion USD)
global_market = np.array([451.3, 518.9, 607.8, 694.3, 808.2])  # Global market size (Billion USD)
global_urbanization = np.array([80.06, 80.33, 80.47, 80.63, 80.73])  # Urbanization rate (%)
global_income = np.array([51000, 51000, 52666, 54000, 55333])  # Per capita income (USD)

# Create DataFrame for better visualization
df = pd.DataFrame({
    'Year': years.flatten(),
    'China Production (Billion RMB)': production_values,
    'China Export (Billion USD)': export_values,
    'Global Market (Billion USD)': global_market,
    'Global Urbanization (%)': global_urbanization,
    'Global Income (USD)': global_income
})

print("Historical Data:")
print(df)

# Visualize historical data
plt.figure(figsize=(14, 7))
plt.plot(years, production_values, label='China Production (Billion RMB)', marker='o', color='blue')
plt.plot(years, export_values, label='China Export (Billion USD)', marker='o', color='green')
plt.plot(years, global_market, label='Global Market Size (Billion USD)', marker='o', color='orange')
plt.title('China Pet Food Production and Export (2019-2023)', fontsize=16)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Value', fontsize=12)
plt.legend(fontsize=10)
plt.grid(True)
for x, y in zip(years.flatten(), production_values):
    plt.text(x, y, f'{y:.1f}', ha='center', va='bottom', fontsize=9)
for x, y in zip(years.flatten(), export_values):
    plt.text(x, y, f'{y:.1f}', ha='center', va='bottom', fontsize=9)
plt.show()
```

```
# Check for anomalies in the data
if np.any(production_values < 0) or np.any(export_values < 0):
    print("Warning: Negative values detected in production or export data.")

# Multivariate regression for predicting future production and export
X = np.column_stack([global_market, global_urbanization, global_income])  # Feature
matrix

# Train production prediction model
production_model = LinearRegression()
production_model.fit(X, production_values)

# Train export prediction model
export_model = LinearRegression()
export_model.fit(X, export_values)

# Prepare data for the next three years
future_years = np.array([2024, 2025, 2026]).reshape(-1, 1)
future_global_market = np.array([900, 1000, 1100])  # Assumed growth in market size
future_global_urbanization = np.array([80.9, 81.1, 81.3])  # Assumed steady urbanization
growth
future_global_income = np.array([57000, 58000, 59000])  # Assumed gradual increase in
income
future_X = np.column_stack([future_global_market, future_global_urbanization,
future_global_income])

# Predict future production and export values
predicted_production = production_model.predict(future_X)
predicted_export = export_model.predict(future_X)

# Display predictions
print("\nPredictions for 2024-2026:")
for year, prod, exp in zip(future_years.flatten(), predicted_production, predicted_export):
    print(f"Year {year}: Predicted Production = {prod:.2f} Billion RMB, Predicted Export =
{exp:.2f} Billion USD")

# Evaluate model performance
production_mse = mean_squared_error(production_values, production_model.predict(X))
export_mse = mean_squared_error(export_values, export_model.predict(X))
print(f"\nModel Evaluation:")
print(f"Production Model MSE: {production_mse:.2f}")
print(f"Export Model MSE: {export_mse:.2f}")

# Visualize predictions
plt.figure(figsize=(14, 7))
plt.plot(years, production_values, label='China Production (Historical)', marker='o',
color='blue')
```

```
plt.plot(future_years, predicted_production, label='China Production (Predicted)', marker='x',
color='blue', linestyle='--')
plt.plot(years, export_values, label='China Export (Historical)', marker='o', color='green')
plt.plot(future_years, predicted_export, label='China Export (Predicted)', marker='x',
color='green', linestyle='--')
plt.title('China Pet Food Production and Export Prediction (2024-2026)', fontsize=16)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Value', fontsize=12)
plt.legend(fontsize=10)
plt.grid(True)
for x, y in zip(future_years.flatten(), predicted_production):
    plt.text(x, y, f'{y:.1f}', ha='center', va='bottom', fontsize=9)
for x, y in zip(future_years.flatten(), predicted_export):
    plt.text(x, y, f'{y:.1f}', ha='center', va='bottom', fontsize=9)
plt.show()
```

Appendix Question4
Introduction4：Analyzing Key Influences on Pet Food Production and Export

```
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt

years = [2019, 2020, 2021, 2022, 2023]

G = [0.03, -0.035, 0.059, 0.032, 0.029]
C = [0.03, 0.07, 0.05, 0.08, 0.06]
T = [0.25, 0.25, 0.25, 0.20, 0.15]
R = [-0.058, -0.062, -0.031, 0, 0.077]
D = [100, 105, 110, 115, 120]

Y_0 = 100
Y = [Y_0]

for i in range(1, len(years)):
Y_t = Y[i-1] * (1 + G[i]) * (1 + C[i]) * (1 - T[i]) * (1 + R[i]) * (D[i] / 100)
Y.append(Y_t)

df = pd.DataFrame({
'Year': years,
'Global Economic Growth (G)': G,
'Production Cost Growth (C)': C,
'Tariff Impact (T)': T,
'Exchange Rate Impact (R)': R,
'Demand Index (D)': [d / 100 for d in D],
'Estimated Output Value (Y)': Y
})

cat_china = [4412, 4862, 5806, 6536, 6980]
dog_china = [5503, 5222, 5429, 5119, 5175]
cat_america = [9420, 6500, 9420, 7380, 7380]
dog_america = [8970, 8500, 8970, 8970, 8010]
cat_france = [1300, 1490, 1510, 1490, 1660]
dog_france = [740, 775, 750, 760, 990]
cat_germany = [1470, 1570, 1670, 1520, 1570]
dog_germany = [1010, 1070, 1030, 1060, 1050]

prod_cny = [440.7, 727.3, 1554, 1508, 2793]
export_usd = [154.1, 9.8, 12.2, 24.7, 39.6]

df['Cat_China'] = cat_china
df['Dog_China'] = dog_china
```

```
df['Cat_America'] = cat_america
df['Dog_America'] = dog_america
df['Cat_France'] = cat_france
df['Dog_France'] = dog_france
df['Cat_Germany'] = cat_germany
df['Dog_Germany'] = dog_germany
df['Pet_Food_Production_CNY'] = prod_cny
df['Pet_Food_Export_USD'] = export_usd

X = df[['Global Economic Growth (G)', 'Production Cost Growth (C)', 'Tariff Impact (T)',
'Exchange Rate Impact (R)', 'Demand Index (D)',
'Cat_China', 'Dog_China', 'Cat_America', 'Dog_America', 'Cat_France', 'Dog_France',
'Cat_Germany', 'Dog_Germany']].values

y_prod = df['Pet_Food_Production_CNY'].values
y_export = df['Pet_Food_Export_USD'].values

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

model_prod = LinearRegression()
model_prod.fit(X_scaled, y_prod)

model_export = LinearRegression()
model_export.fit(X_scaled, y_export)

print(f"Production Value Intercept (α): {model_prod.intercept_}")
print(f"Production Value Coefficients (β): {model_prod.coef_}")

print(f"Export Value Intercept (α): {model_export.intercept_}")
print(f"Export Value Coefficients (β): {model_export.coef_}")

coef_prod_df = pd.DataFrame(model_prod.coef_, columns=['Coefficient'], index=['Global
Economic Growth (G)', 'Production Cost Growth (C)',
'Tariff Impact (T)', 'Exchange Rate Impact (R)', 'Demand Index (D)',
'Cat_China', 'Dog_China', 'Cat_America', 'Dog_America',
'Cat_France', 'Dog_France', 'Cat_Germany', 'Dog_Germany'])
coef_prod_df.plot(kind='bar', title="Impact of Factors on Pet Food Production Value")
plt.show()

coef_export_df = pd.DataFrame(model_export.coef_, columns=['Coefficient'],
index=['Global Economic Growth (G)', 'Production Cost Growth (C)',
'Tariff Impact (T)', 'Exchange Rate Impact (R)', 'Demand Index (D)',
'Cat_China', 'Dog_China', 'Cat_America', 'Dog_America',
'Cat_France', 'Dog_France', 'Cat_Germany', 'Dog_Germany'])
coef_export_df.plot(kind='bar', title="Impact of Factors on Pet Food Export Value")
plt.show()
```