Transactions on Computational and Applied Mathematics (2024) DOI: 10.23977/tracam.2024.040102
Clausius Scientific Press, Canada ISSN 2616-1826 Vol. 4 Num. 1

Research on Vegetable Pricing and Replenishment Based on Exponential Smoothing Model Prediction

Zhenyu Zhou, Yuang Zhang, Jiaxing Zhao

School of Mathematics and Statistics, Hubei Engineering University, Xiaogan, 432000, China

Keywords: Exponential Smoothing Forecasting, Optimization Model, Time Series, Pearson Correlation Analysis

Abstract: Vegetables and fresh food are necessities in daily life. Their quality and price not only affect the interests of businesses, but also have a direct impact on people's livelihood issues. However, this type of commodity has its particularity, its appearance and quality will deteriorate over time, resulting in the need to discount sales or even unable to sell. Therefore, it is a very important task for businesses to seek to establish a good and healthy supply and marketing channel. This article has established an exponential smoothing model to predict the sales volume of its vegetable categories and individual items respectively. It then solves the problem through optimization model algorithms, calculating the optimal daily replenishment total and pricing strategy, thus maximizing the supermarket's revenue.

1. Introduction

With the rapid development of the economy, people's demands for the quality of life and living standards are also increasing. The concept of ordinary people has changed from "having enough to eat" to "eating well". Because of this, freshness becomes particularly important. Generally speaking, the shelf life of vegetable products is relatively short and the appearance of vegetables will deteriorate with time, transportation damage, temperature, humidity, etc [1]. If they are not sold on the same day, they cannot be resold or discounted the next day, which has a certain impact on both customers and merchants. Therefore, supermarkets generally formulate daily replenishment strategies based on the sales history and demand of each commodity. There are many varieties of vegetables with different origins, and the trading time of vegetables is usually between 3 and 4 am, so merchants must make replenishment decisions for various vegetable categories on that day under the condition of uncertainty about specific dishes and purchase prices. Supermarkets usually adopt the "cost-plus pricing" method for vegetables. They will discount vegetables whose appearance deteriorates due to factors such as transportation damage [2]. The sales volume of vegetable products often has a certain relationship with time, so the sales mix, pricing strategy, and replenishment decision-making of supermarkets become particularly important.

2. Pearson correlation coefficient

Through observation of data, it can be found that there are no null values or outliers [3], so this article groups by classification name, calculates the sales volume of each category, then visualizes and statistically analyzes the sales situation of various vegetable categories, and uses Jupyter Notebook for calculation and drawing, resulting in the following Fig 1:

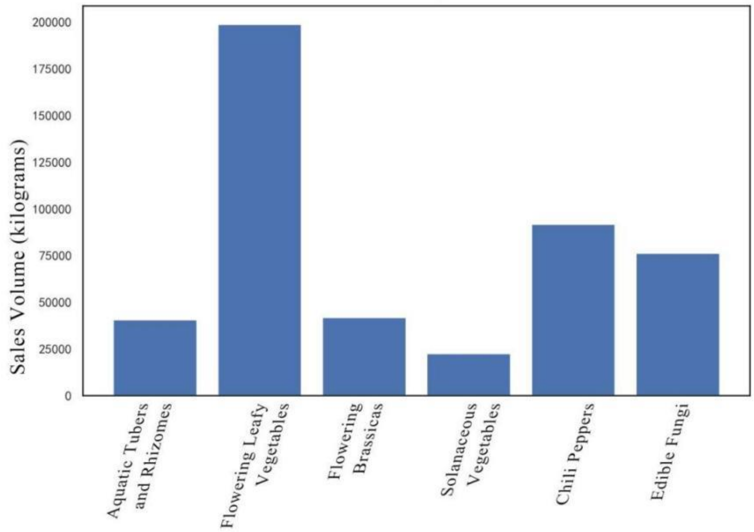

Figure 1: Distribution of sales volume of vegetables by category

Through observation of the bar chart and the results shown in the following Fig, we can see that the total sales volume of different vegetable categories from high to low are: leafy vegetables > peppers > edible fungi > cauliflower > aquatic roots and stems > eggplant. The sales volume of leafy vegetables is the highest among all categories, while the sales volume of eggplant is the lowest. In this paper, the data of sales volume is grouped by category name and quarter, and the sales volume of each category in each quarter is calculated. The trend of changes in the sales volume of different vegetable categories is visualized and presented in Fig 2:

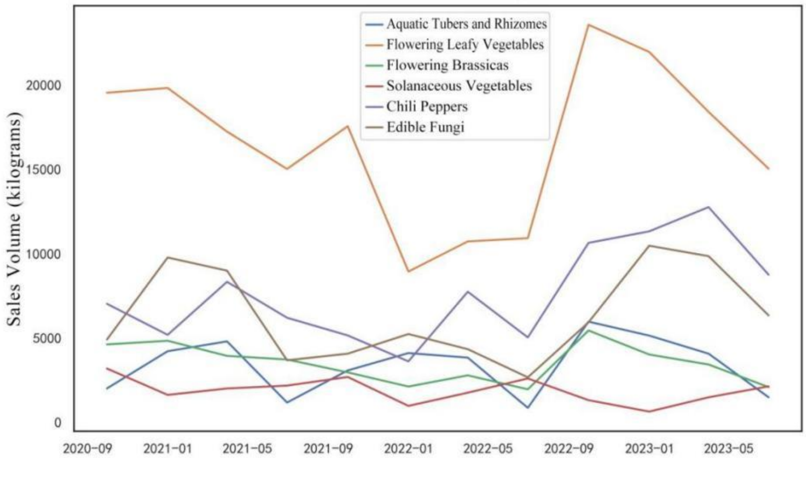

Figure 2: The trend of sales volume of vegetables in various categories

This article group the sales data by category name and quarter, calculate the average sales unit price and total sales price of each category in each quarter, and make a visual chart of the trend of changes in sales unit price and total sales price as shown in Fig 3 and Fig 4:

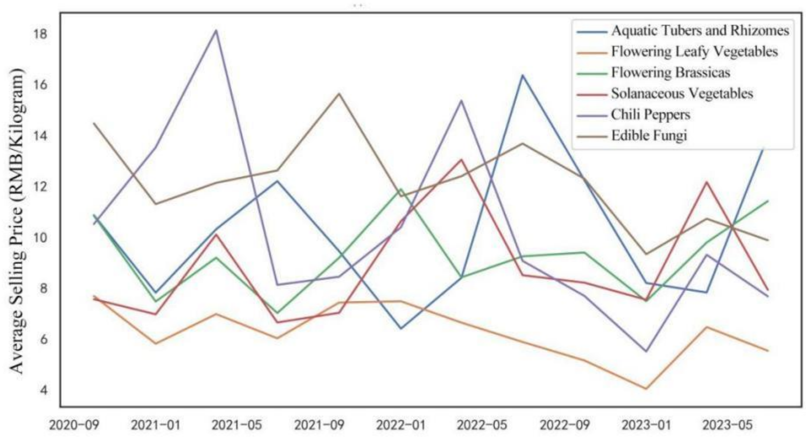

Figure 3: The trend of the average sales unit price of vegetables in various categories

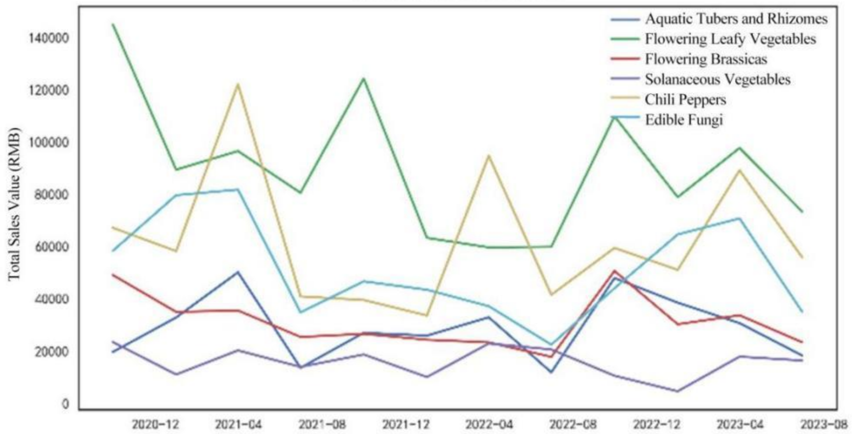

Figure 4: The trend of the total sales price of various categories of vegetables

This article calculate the total sales price, wholesale total price and loss total price. This sales data is grouped by category name and quarter, calculate the total sales price, wholesale total price and loss total price of each category in each quarter, and calculate the profit of each category in each quarter. The visual trend of profit change is demonstrated in Fig 5:

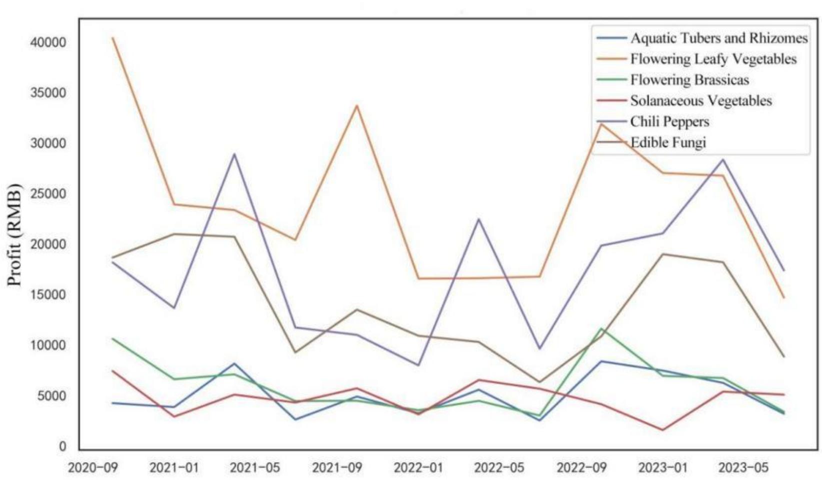

Figure 5: Profit change trend of various vegetable categories

In order to facilitate calculation, this article first converts the data within the sales date into date type and groups them by quarter and category name [4]. This allows us to calculate the sales volume of each category in each quarter. Then, we calculate the correlation coefficient between the sales volume of each category and visualize their relationship as a matrix. If the absolute value of the correlation coefficient is greater than 0.7, it indicates a strong correlation. From the results in the table below, we can see that the correlation coefficient between eggplant and aquatic root vegetables is -0.72, indicating a strong negative correlation, which means they are substitute products for each other. The Pearson correlation coefficient between cauliflower and leafy vegetables is 0.9 [5], indicating a strong positive correlation between cauliflower and leafy vegetables, making them complementary dishes.

Group by category name and calculate the sales volume of each category. Because the number of vegetable varieties is huge, only the top 15 vegetable items with the highest sales volume are listed here. This article creates different colors for each item based on sales quantity, and visualizes the sales trend of the top 15 vegetable items on a quarterly basis. It resamples sales data by quarter, groups sales data by category name and quarter, calculates the sales quantity for each category in each quarter, and obtains the sales change of the top 15 vegetable items. It translates the sales date into date type, groups by quarter and category name, and calculates the sales quantity for each category in each quarter. It translates the $MultiIndex$ into the $DataFrame$, calculates the sales correlation coefficient between various categories, deletes the top 15 vegetable items with the highest sales, removes the correlation coefficients of the top 15 vegetable items with the highest sales from the correlation coefficient matrix, and creates a correlation coefficient matrix. According to the results, the correlation coefficient between Yunnan Oil Lettuce and Yunnan Lettuce (serving) is 0.96, indicating that these two items show a strong correlation, making them complementary items. Similarly, bubble lettuce and baby cabbage combination, Sichuan peppercorn and Yunnan lettuce (serving) also show complementary relationships as seen from the above chart. However, Yunnan oil lettuce and mushroom (box) and Yunnan lettuce (serving) are substitutes for each other.

3. Exponential Smoothing Model

3.1 Exponential Smoothing Forecasting Model

Due to the fact that supermarkets make replenishment plans based on categories, this paper analyzes the relationship between the total sales volume and cost-plus pricing of various vegetable categories, and provides a replenishment plan and pricing strategy for the next week to maximize the supermarket's revenue. This paper uses exponential smoothing method to predict the total sales volume of vegetable categories in the next week, in order to better understand the trend of market demand [6]. By comprehensively considering the total sales volume, cost-plus pricing and market demand, supermarkets can formulate optimal replenishment plans and pricing strategies to achieve maximum revenue.

Initialization:

Initial forecast value (usually the first observation in a time series):

$$\widehat{Y}_i = Y_i \tag{1}$$

The initial smoothing factor (usually a constant between 0 and 1, indicating the weight of new data):α

Predict Future Value:

Prediction value for time step $t(t > 1)$:

$$\widehat{Y_t} = \alpha \cdot Y_t + (1-\alpha) \cdot \widehat{Y_{t-1}} \tag{2}$$

In this formula, $\widehat{Y_t}$ represents the predicted value at time step t, Y_t is the actual observed value, α is the smoothing coefficient.

The core idea of this model is to predict future values based on the exponential weighted average of historical observation values. The smoothing factor α determines the weight of new data. The larger α will react more quickly to new observations, but it may also lead to larger fluctuations. The smaller α has stronger smoothing effect, but its reaction to new data is slower. In order to select an appropriate value more accurately α , this paper evaluates the performance of the model by using methods such as cross-validation and selects the smoothing factor that can produce the best prediction results [7].

Based on the sales forecast of vegetable categories predicted by smooth index, an optimization model for profit has been established. The results of the model are shown below:

$$s.t.\max\ profit = \sum S_{it} * X_{it} - D_{it} * P_i/(1-L_i) \tag{3}$$

$$X_{it} = S_{it} * P_i/(1-L_i) * (1+N_i) \tag{4}$$

$$\sum D_{it} * (1-L_i) \geq \sum S_{it} \tag{5}$$

In which, $profit$ is the total profit in 7 days,Xit is the price for the cost of the category i on day t, Sit is the sales volume of the category i on the day t, Dit is the purchase quantity of the category i on the day t, Li is the loss rate of the category i, Ni is the addition rate for the category i, by iterating to obtain the total replenishment quantity, the supermarket's revenue is maximized. The constraint condition is that the purchase quantity of each category in 7 days is greater than or equal to the sales quantity.

Through the calculation formula, we can calculate the markup rate, cost-plus pricing and average loss rate of vegetable categories. From the results, it can be seen that the markup rate of leafy vegetables is the highest at 0.398703, while the markup rate of aquatic root vegetables is the lowest at 0.308087. Supermarkets can adjust their daily replenishment volume and pricing strategies for the next week based on the markup rates, cost-plus pricing and average loss rates of different categories, combined with market demand and competition, in order to maximize profits. Different types of vegetables may require different strategies to meet market demand and ensure maximum business efficiency of supermarkets [8]. In addition, supermarkets can also adjust these strategies based on historical sales data and trends to cope with changes in the market.

Using the exponential smoothing forecasting model, using the exponential smoothing forecasting model, the vegetable category sales volume for 2023/07/01-2023/07/07 is predicted based on historical data. The results are shown in Table 1:

Table 1: Exponentially smoothed sales forecast results

Time	Aquatic rhizomes	Mosaic and leafy	Cauliflower	Nightshades	Chili peppers	edible fungi
2023/7/1	31.757944	190.784084	29.937935	29.636192	128.784073	82.188835
2023/7/2	26.725007	180.914038	27.963673	30.309923	108.57332	66.56872
2023/7/3	15.972296	116.515183	17.323893	17.503669	65.761534	35.933747
2023/7/4	13.627105	125.811216	17.430301	15.144404	65.079625	32.542406
2023/7/5	16.117473	118.38412	18.292211	16.465227	65.654687	37.569905
2023/7/6	17.901674	117.148984	20.020189	16.384542	73.831031	38.046365
2023/7/7	20.610012	135.318808	22.278354	19.435877	90.994907	54.261552

The evaluation result of the Exponential Smoothing Model is shown in Table 2:

In general, the predictive performance of the model varies among different vegetable categories. For aquatic rhizomes, cauliflowers, and eggplants, the model's predictions are relatively accurate with smaller errors. However, for leafy vegetables and peppers, the model's prediction errors are larger, which may require further optimization or consideration of other prediction methods. In practical applications, supermarkets can develop more refined replenishment and pricing strategies based on the model performance of different categories to improve operational efficiency and meet market demand.

Table 2: Model evaluation results

The name of the vegetable category	MSE	RMSE
Aquatic rhizomes	59.801573	7.733148
Mosaic and leafy	1118.106881	33.438105
Cauliflower	112.832465	10.622263
Nightshades	22.860704	4.781287
Chili peppers	439.208248	20.957296
edible fungi	342.587190	18.509111

The optimal model solution result is a total revenue of 2398.722244 yuan in 7 days.

3.2 Prediction of Exponential Smoothing Model

Due to limited sales space, the replenishment plan for customized items needs to be further optimized. The total number of salable items should be between 27 and 33, with a minimum display quantity of each item at 2.5 kilograms. Based on the available varieties from June 24-30, 2023, please provide the individual item replenishment quantity and pricing strategy for July 1. This should be done while maximizing the supermarket's profits as much as possible, while also meeting the market's requirements for various vegetable products.

In formulating the replenishment plan, factors such as sales volume, inventory, loss rate, and cost of individual items should be considered to avoid overstocking or insufficient inventory. In terms of pricing, this paper formulates a reasonable pricing strategy based on the cost and gross profit margin of each vegetable item, combined with market demand and competition, to maximize the profits of supermarkets [9]. The model is shown as follows:

$$s.t. \max\ profit = \sum S_i * X_i - D_{it} * P_i/(1 - L_i) \tag{6}$$

$$X_{it} = S_i * P_i/(1 - L_i) * (1 + N_i) \tag{7}$$

$$\sum D_{it} * (1 - L_i) \geq \sum S_i \tag{8}$$

$$27 \leq Count(i) \leq 33 \tag{9}$$

$$S_i \geq 2.5 \tag{10}$$

In which, $profit$ is the total profit in 7 days, X_i is the cost pricing for product categories i, S_i is the sales volume of the category i, D_i is the purchase quantity for the category i, L_i is the loss rate of the category i, N_i is the addition rate for the category i, through iterations, the total replenishment quantity is obtained to maximize the supermarket's revenue [10]. The constraint is that the inventory for each category should be greater than or equal to the sales volume within 7 days. The total number of sell items should be controlled between 27 and 33, and each individual item must meet the minimum display quantity requirement of 2.5 kilograms.

Firstly, a total of 48 individual items were counted as available for sale from June 24th to June 30th, 2023. The top 33 individual items with the highest sales volume were selected. The selected individual items are shown in Table 3:

Similarly, the calculation results of the 33 individual items' markup rate are shown in Table 4:

Table 3: The total sales volume of the 33 single items included in the model

The name of the item	Total sales volume	The name of the item	Total sales volume	The name of the item	Total sales volume
Shanghai green	27.029	Fungus vegetables	41.534	water caltrop	12.224
Yunnan oily wheat vegetable (serving)	149	Zhijiang green stalk scattered flowers	25.203	Cordyceps flowers (servings)	16
Yunnan lettuce (serving)	226	Honghu lotus root	28.232	Screw pepper	47.897
Pure Lotus Root(1)	42.184	Seafood mushrooms (pack)	62	Screw pepper (serving)	79
Bisporus mushroom (box)	70	Bamboo leafy vegetables	93.077	broccoli	87.904
Milk cabbage	44.959	Purple Eggplant (2)	76.32	West Gorge Mushroom(1)	32.369
Ginger, garlic and millet pepper combination (small portion)	49	Red Pepper (2)	14.25	Enoki mushroom (box)	113
Baby cabbage	73	Sweet potato tip	31.546	Long-line eggplant	29.501
Small wrinkles (servings)	79	Wuhu Green Pepper(1)	99.634	Green and red Hangzhou pepper combination (portion)	15
Millet pepper (serving)	150	amaranth	62.461	Green Eggplant (1)	18.727
Baby greens(1)	34.306	Spinach (serving)	49	Tall melon (1)	21.031

Table 4: 33 single item bonuses

The name of the item	Bonus rate	The name of the item	Bonus rate	The name of the item	Bonus rate
Shanghai green	0.241426	Fungus vegetables	0.228563	water caltrop	0.347482
Yunnan oily wheat vegetable (serving)	0.37031	Zhijiang green stalk scattered flowers	0.225874	Cordyceps flowers (servings)	0.371056
Yunnan lettuce (serving)	0.340252	Honghu lotus root	0.310422	Screw pepper	0.253664
Pure Lotus Root(1)	0.314326	Seafood mushrooms (pack)	0.229227	Screw pepper (serving)	0.207168
Bisporus mushroom (box)	0.294364	Bamboo leafy vegetables	0.227705	broccoli	0.296912
Milk cabbage	0.204526	Purple Eggplant (2)	0.298046	West Gorge Mushroom(1)	0.256355
Ginger, garlic and millet pepper combination (small portion)	0.254016	Red Pepper (2)	0.378101	Enoki mushroom (box)	0.320617
Baby cabbage	0.270857	Sweet potato tip	0.301809	Long-line eggplant	0.25279
Small wrinkles (servings)	0.259521	Wuhu Green Pepper(1)	0.252224	Green and red Hangzhou pepper combination (portion)	0.227357
Millet pepper (serving)	0.334326	amaranth	0.227518	Green Eggplant (1)	0.320549
Baby greens(1)	0.34825	Spinach (serving)	0.300579	Tall melon (1)	0.262982

The maximum profit of the supermarket on that day is calculated to be 701.4735 yuan through algorithm.

4. Conclusion

The improvement and promotion of models is a continuous process for supermarkets, which can continuously improve operational efficiency, reduce costs, increase customer satisfaction, and competitiveness. Supermarkets can consider adopting more advanced data analysis and prediction models, such as machine learning and artificial intelligence. These models can more accurately predict demand, optimize supply chains, formulate pricing strategies, and automatically adjust replenishment plans. This will help supermarkets respond more flexibly to market changes. Supermarkets can also consider the sustainability and environmental impact of models. Optimizing replenishment plans and supply chains can reduce energy consumption and waste, thereby reducing the impact on the environment. In conclusion, the improvement and promotion of models are key steps for supermarkets to continuously improve operational efficiency, reduce costs, and increase competitiveness. By continuously improving data collection, analysis methods, and decision support systems, supermarkets can better adapt to changing market demands, improve operational efficiency, provide better products and services, and achieve sustainable growth and success. The improvement and promotion of models are a strategic investment that will bring lasting competitive advantages in the future.

References

[1] QI An. Research on Pricing of Science and Technology Novelty Searching Services Based on Cost-Plus Pricing Method [J]. Library Research and Work, 2021(10):25-31+24.

[2] Liyuan Ning. Pricing Strategy and Techniques for Green Vegetables [J]. Guangdong Sericulture, 2020, 54(07):111-112.

[3] Zhenhai Cai. Practical Exploration of Anomaly Data Processing and Analysis Based on Python [J]. Computer Knowledge and Technology, 2023, 19(27):62-65

[4] Shuai Xu, Dandan Liu. Electric Load Forecasting Based on Time Series Combination Model [J]. Electronic Design Engineering, 2023, 31(23): 1-6.

[5] Jinwei Wang, Yuhan Shan, Baohai Shan. Data Analysis of Short Track Speed Skating Competition Based on Pearson Correlation Coefficient [J]. Ice and Snow Sports, 2023, 45(04):9-12+17.

[6] Decai Liu, Zhijie Zhou. Research on the demand forecast of aquatic product cold chain logistics based on combined prediction model - taking Jiangsu Province as an example [J]. Logistics Engineering and Management, 2023, 45(10): 4-8.

[7] Dike Sang, Zhihao Ren. Research on Logistics Optimization Based on Maximizing Benefits of SCCY Company [J]. China Logistics and Purchasing, 2023(15):119-120.

[8] Xiapei Li. Prediction of Agricultural Product Logistics Demand Based on Grey Linear Combination Model [J]. Journal of Beijing Jiaotong University (Social Science Edition), 2017, 16(01): 120-126.

[9] Fang Huang, Rui Yan, Jinzhou Niu. Construction of energy consumption prediction model for cold storage based on time series index smoothing method [J]. Shanghai Energy Conservation, 2023(03):336-341.

[10] Xiaobin Wang, Fujun Lu,Ting Sun, et al. Iterative Computation and Big Data Analysis in Finance [J]. Information and Communication Technology, 2017, 11(04): 47-52.