Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 11, No. 2, 2023

Automatic Pricing and Replenishment Decision Model for Vegetable Products Based on Optimization Algorithm

Xiaolei Chen *

Electronic and Information Engineering, Guangdong Polytechnic Normal University, Guangzhou, China
* **Corresponding author Email:** 1933593729@qq.com

Abstract: Due to the short shelf life of vegetable commodities, vegetable superstores need to make replenishment and pricing decisions on a daily basis based on the historical sales and demand of each commodity. Therefore, this paper establishes an optimization model to help superstores filter and analyze data and optimize replenishment and pricing decisions. First of all, we obtained the sales volume distribution data and images of each category over time, analyzed the interrelationship of each vegetable category, studied the "cost-plus pricing" method on this basis, and analyzed the correlation between the total sales volume and cost-plus pricing. Finally, the objective function of maximizing the revenue of the superstore is chosen, and a linear programming model is established under the constraints to obtain the total daily replenishment quantity and pricing strategy of each vegetable category in the coming week, which can maximize the revenue of the superstore.

Keywords: Benefit Maximization; Pearson Coefficient Analysis; Linear Programming Models.

1. Introduction

Vegetable commodities in fresh produce superstores have a short freshness period and deteriorate in quality after increased sales time, and most varieties cannot be re-sold the next day. Supermarkets are replenished daily based on historical sales and demand for the commodities [1]. Due to the many varieties of vegetables and their different origins, merchants need to make replenishment decisions without knowing exactly what the specific individual products and purchase prices are [2-3]. Vegetables are generally priced using the "cost-plus pricing" method, and supermarkets often sell damaged and poor-quality items at a discount [4]. From both the demand side and the supply side, the limitation of sales space in supermarkets makes the reasonable sales mix extremely important. In this paper, we analyze the distribution pattern and interrelationship of the sales volume of each vegetable category and individual product, analyze the relationship between the total sales volume of each vegetable category and the cost-plus pricing, and determine the total daily replenishment volume and pricing strategy for each vegetable category in the coming week to maximize the revenue of the superstore.

2. Analysis of the distribution pattern and correlation of various categories of vegetables in supermarkets

2.1. Correlation analysis based on Pearson coefficient

According to the results obtained from the descriptive statistics of different vegetable categories, we can initially know that there is a certain degree of interrelationship within them, in order to further explore the correlation, through the category sales volume over time change law, the use of Pearson correlation coefficient to analyze the correlation of the data within the data, in order to measure the degree of linear correlation between the two variables [5]. By calculating the Pearson correlation coefficient, we can know the degree of correlation between the sales volume of different categories, the correlation coefficient is between -1 and 1, the Pearson coefficient of the two variables is expressed as the quotient of their covariance and variance. The following is the calculation procedure:

Calculate the sample covariance:

$$\mathrm{Cov}(x,y)=\frac{\sum_{i=1}^{n}(X_I-X)(Y_i-Y)}{n-1} \tag{1}$$

Calculate the sample standard deviation:

$$S_x=\sqrt{\frac{\sum_{i=1}^{n}(X_i-X)^2}{n-1}} \tag{2}$$

$$S_y=\sqrt{\frac{\sum_{i=1}^{n}(Y_i-Y)^2}{n-1}} \tag{3}$$

Then the sample Pearson correlation coefficient:

$$r_{xy}=\frac{\mathrm{Cov}(X,Y)}{S_X S_Y} \tag{4}$$

Refer to the relationship between Pearson's correlation coefficient and the degree of correlation: the closer to 1 or -1, the larger the absolute value of the coefficient, the stronger the degree of correlation; and the closer to 0, the weaker the degree of correlation. The degree of correlation between the variables can be determined by the range of values shown in Table 1.

Table 1. Scale of correlation strength r-values

Numerical range	degree of relevance
0-0.2	Very weak correlation or no correlation
0.2-0.4	weak correlation
0.4-0.6	Moderately relevant
0.6-0.8	strong correlation
0.8-1.0	Highly relevant

2.2. Correlation analysis based on Pearson coefficient

Considering that both the total sales and the cost-plus ratio show seasonality in the graph with time as the horizontal axis, the relationship between the total sales and the cost-plus ratio is analyzed with the help of Spearman's correlation analysis in Spspro. Spearman's correlation analysis calculates the correlation coefficients (the degree of correlation) between the two data.

Firstly, the existence of a statistically significant relationship between XY was tested to determine whether the P-value presented significance ($P<0.05$), and it was found that there was no significance in any of the six categories, and then the correlation coefficients were analyzed for their positive and negative directions, as well as the degree of correlation. Variables X and Y were ranked in order from smallest to largest and were represented by the ranks R_X and R_Y. When sorting, the phenomenon of equal data resulting in the same rank is called holding, and the average of these is taken as the rank of each data. Remember that the Spearman correlation coefficient is η, then:

$$\eta = \frac{\sum R_X R_Y - \frac{(\sum R_X)(\sum R_Y)}{n}}{\sqrt{(\sum R_X^2 - \frac{(\sum R_X)^2}{n})}\sqrt{(\sum R_Y^2 - \frac{(\sum R_Y)^2}{n})}} \tag{5}$$

The Spearman correlation coefficient is similar to the Pearson correlation coefficient in that it provides an objective statistical measure of the strength of the relationship between the variables. The Spearman correlation coefficient ranges from -1 to 1, with -1 indicating a perfect negative correlation and 1 indicating a perfect positive correlation. Spearman's correlation coefficient ranges from -1 to 1, with -1 indicating a perfect negative correlation, 1 indicating a perfect positive correlation, and 0 indicating no linear relationship between the two variables. The final table of correlation coefficients for each category is obtained (the more "+", the greater the correlation with the cost-plus ratio).

2.3. Cost-plus pricing strategy analysis

The cost-plus pricing method means that the price of a product compensates for the costs of production and distribution and provides a reasonable return. The theory assumes that the superstore has a dominant role in determining the price and that the buyer can influence the cost-plus rate. The "cost-plus pricing" method determines the selling price of the product being sold by setting the appropriate markup rate on the wholesale price $Q_1 Q_2$, so the cost-plus rate ω is a direct reflection of the level of cost-plus pricing, and the next step is to consider using the cost-plus rate ω to measure the level of cost-plus pricing for different vegetable categories.

Q_1 The correlation between, Q_2, ω is as follows:

$$Q_2 = (1+\omega)Q_1 \tag{6}$$

Then it can be obtained:

$$\omega = \frac{Q_2}{Q_1} - 1 \tag{7}$$

2.4. Macro total sales volume and category sales volume analysis

Firstly, quantitative statistics were conducted from the vegetable categories, counting the total sales of individual items to derive the total sales of vegetable categories over the past three years, and exploring the relationship between them through comparative analysis to help understand the sales situation among different vegetable categories. It was concluded that the total sales of flowering and leafy vegetables far exceeded that of other vegetables, while the total sales of eggplants were generally low. The statistical graph is shown in Figure 2.

Considering the vegetable category sales volume index has obvious time, so consider analyzing the relationship between category sales volume change with time, firstly, carry out data screening processing, and then data categorization, respectively, with the number of days as the x-axis, accumulating and analyzing the sales volume data of each category in the same day, drawing a line graph of sales volume change with the number of days of the same day of the categories, and making a comparative analysis. The line graph is shown in Figure 2. Through the chart, it can be found that each vegetable category has a certain seasonality.

Spsspro was used to import the sales metrics of each vegetable category over time, perform Pearson correlation analysis, and represent the final Pearson coefficient results in a heat map to visualize the structure of the relationship between the categories. The heat map is shown in Figure 3.

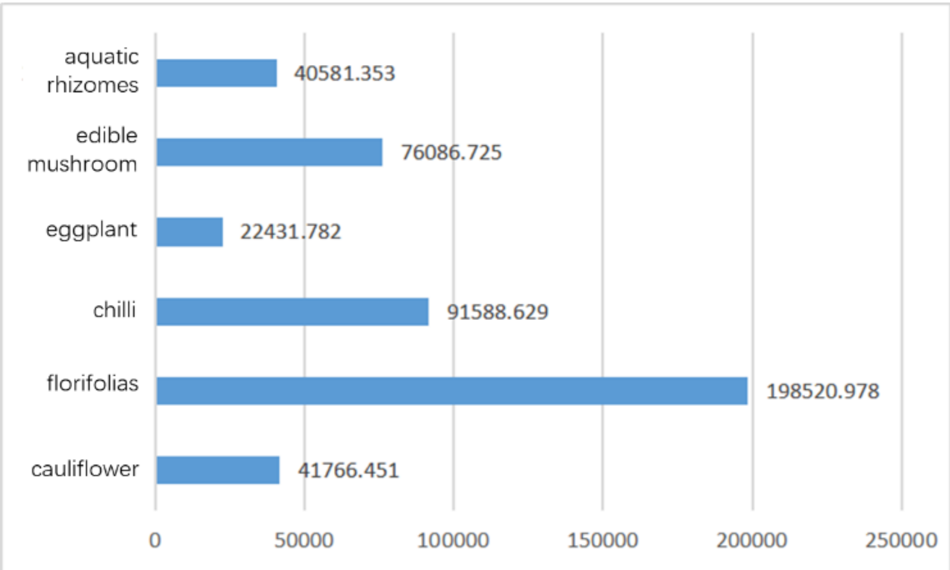

Figure 1. Histogram of the distribution of total sales volume of vegetables by category (unit: kg)

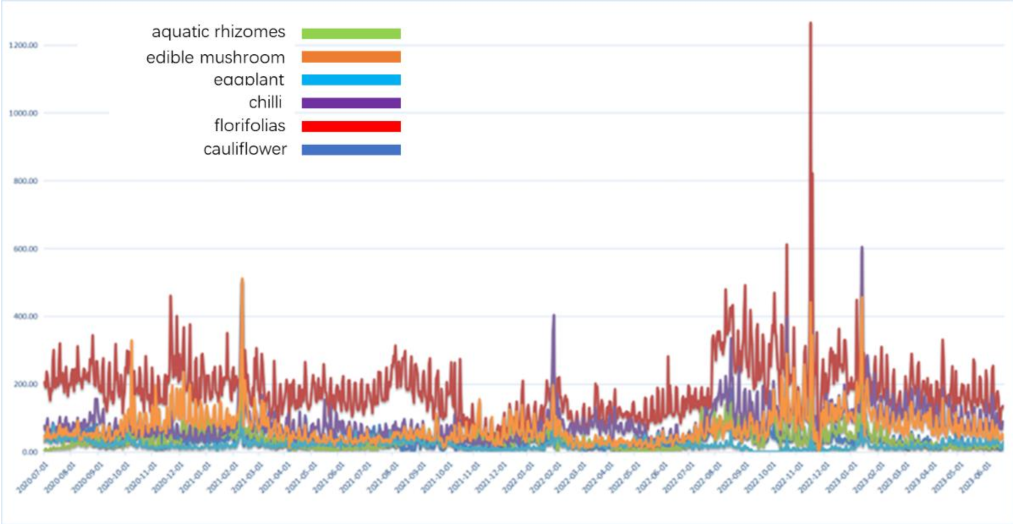

Figure 2. Curve of sales volume by category over time

	Florifolias	cauliflower	aquatic rhizomes	eggplant	chilli	edible mushroom
Florifolias	1	0. 627	0. 561	0. 257	0. 659	0. 631
cauliflower	0. 627	1	0. 542	0. 312	0. 551	0. 523
aquatic rhizomes	0. 561	0. 542	1	0. 074	0. 614	0. 67
eggplant	0. 257	0. 312	0. 074	1	0. 273	0. 12
chilli	0. 659	0. 551	0. 614	0. 273	1	0. 687
edible mushroom	0. 631	0. 523	0. 67	0. 12	0. 687	1

Figure 3. Heat map of Pearson correlation coefficient

Based on the structure of the relationships presented in the heat map, it can be concluded that among them, the floral and foliage category is more strongly correlated with the pepper category, the aquatic rhizomes category and the edible mushrooms category, whereas the deep-rooted category is weakly correlated with the eggplant category. Observing the correlation coefficient of eggplant category, it is found that its Pearson coefficient with other vegetable categories is generally low.

3. A model for automated replenishment and pricing of superstore vegetable items

3.1. Relationship between total sales and cost-plus pricing

Using the wholesale price (i.e., cost) as the vertical axis and time as the horizontal axis, a line graph of various types of costs over time is plotted as shown in Figure 4. According to the image it can be found that the fluctuation of each type of cost has a certain seasonality.

With the sales price Q_2 as the vertical axis, time as the horizontal axis to draw the sales price of each category with the time change of the line graph shown in Figure 5, according to the image can be found in the sales price of each category fluctuations there is also a certain seasonality.

Where let the unit price of sales u and sales t be there (Summing the selling prices of individual items):

$$Q_2 = \sum ut \quad (8)$$

Then, according to the above formula $\omega = \frac{Q_2}{Q_1} - 1$, the cost markup rate of each category can be calculated, and plotted a line graph of cost markup rates over time for each category as shown in Figure 6.

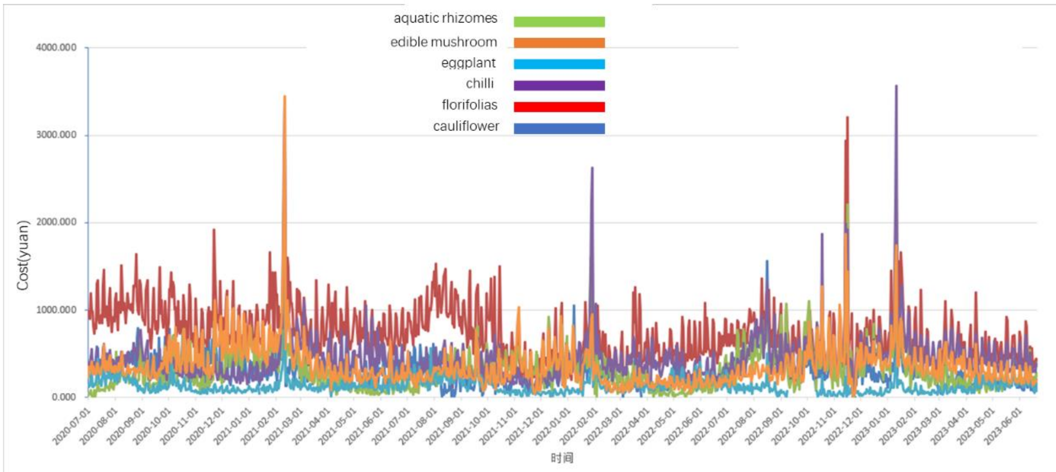

Figure 4. Line graph of cost changes over time by category

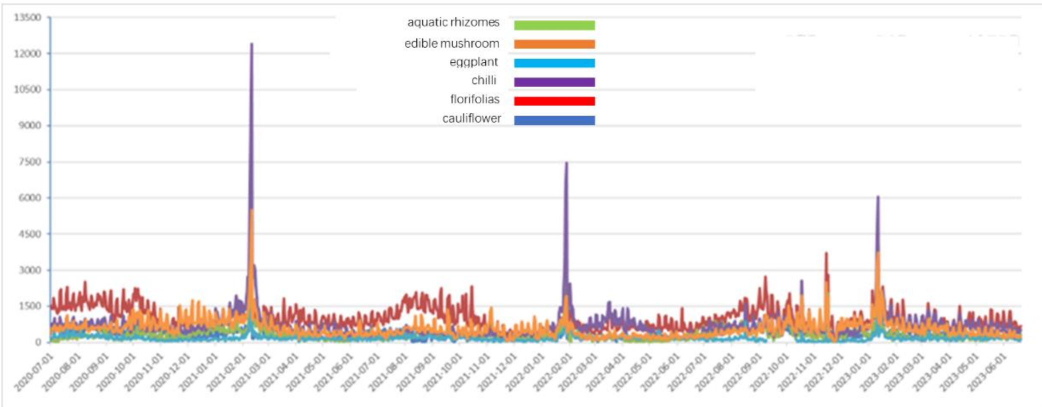

Figure 5. Line graph of sales price changes over time by category

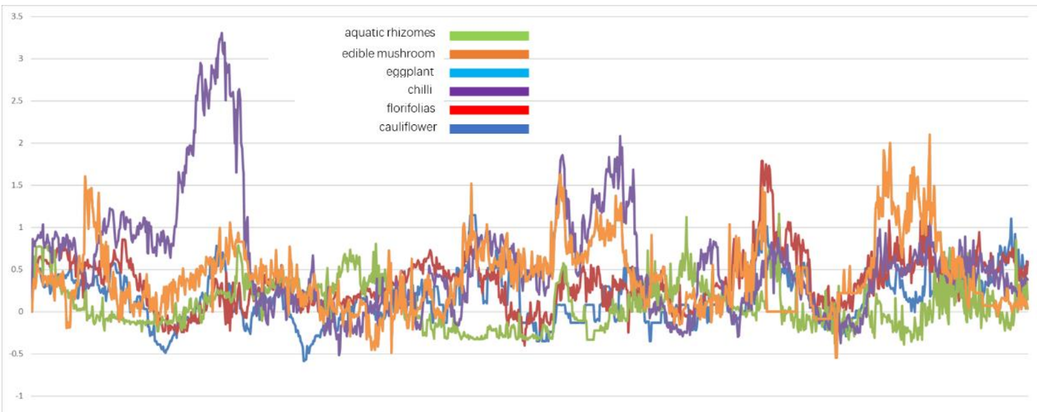

Figure 6. Line graph of cost-plus rates over time by category

The categories can be seen to fluctuate in a roughly seasonal pattern.

Table 2. Table of correlation coefficients by category

cauliflower (Brassica oleracea var. botrytis)	philodendron	Aquatic rhizomes	capsicum	eggplant	edible mushrooms
-0.152	0.103	-0.64	-0.428	-0.528	0.1
+	+	+++	++	++	+

From Table 2, it can be concluded that total sales of aquatic roots and tubers, peppers and eggplants are weakly and negatively correlated with the cost-plus rate. While cauliflower, leafy vegetables and edible mushrooms generally have low correlation and almost no correlation.

3.2. Total replenishment and pricing strategy

For the total daily replenishment and pricing strategy of each vegetable category for the coming week, a linear programming model is considered to be used to solve the problem, taking into account the cost-plus pricing strategy of the superstore as well as the linear relationship of the constraints of the topic.

First define the decision variable x_{ij} ,which denotes the total sales of the vegetable category numbered j on day i, where, $i = 1,2,3,\cdots,7, j = 1,2,3,\cdots,6$, each vegetable category is numbered as shown in Table 3.

Table 3. Number of each vegetable category

cauliflower (Brassica oleracea var. botrytis)	philodendron	Aquatic rhizomes	capsicum	eggplant	edible fungi
1	2	3	4	5	6

Similarly, define Q_{ij} to denote the final selling price of the vegetable category on the day of i with the number j, define ω_{ij} to denote the cost plus rate of the vegetable category on the day of i with the number j, and n_{ij} to denote the wholesale price of the vegetable category on the day of i with the number j, and thus Q_{ij}, ω_{ij}, and n_{ij} are related to each other:

$$Q_{ij} = (1 + \omega_{ij})n_{ij} \tag{9}$$

Define P_{ij} as the total profit of the vegetable category numbered j on day i , then:

$$P_{ij} = \omega_{ij}x_{ij} \tag{10}$$

For a week the total profit of the superstore P_t can be expressed as:

$$P_t = \sum_{i=1}^{7}\sum_{j=1}^{6} P_{ij} \tag{11}$$

In summary the following optimization model can be determined:

$$\max \ P_t = \sum_{i=1}^{7}\sum_{j=1}^{6} P_{ij} \tag{12}$$

$$s.t.\begin{cases} Q_{ij} = a_{j0} + a_{j1}x_{ij} \\ Q_{ij} = (1+\omega_{ij})n_{ij} \\ P_{ij} = \omega_{ij}x_{ij} \\ x_{ij}, P_{ij}, Q_{ij} \geq 0 \\ \omega_{ij} \in R \end{cases} \tag{13}$$

Finally, the total daily replenishment and pricing strategy for the coming week are obtained.

4. Conclusion

This paper establishes an optimization model to help superstores filter and analyze data and make better replenishment and pricing decisions. Firstly, we obtain the sales volume of individual vegetable products and categories, and then we further analyze the interrelationships, using the processed sales volume distribution data and images of each category over time, and using Pearson coefficient correlation analysis to study the interrelationships of each vegetable category. On this basis, the "cost-plus pricing" method is studied, and the mathematical method is used to describe the cost-plus pricing strategy with the help of the cost-plus rate, and the correlation analysis of Spearman's coefficient is used to analyze the correlation between the total volume of sales and the cost-plus pricing, which is more suitable. Finally, the objective function of maximizing the revenue of the superstore is selected, and a linear programming model is established under the constraints to obtain the total daily replenishment and pricing strategy of each vegetable category in the coming week that can maximize the revenue of the superstore. Overall, the automatic pricing and replenishment model established in this paper is based on data screening and analysis, correlation analysis between variables and optimization model establishment, which can be used to solve the problems of replenishment and pricing decisions of superstores with objectivity, standardization, and comprehensiveness, and it can help superstores rationally allocate resources in combination with the reality of various indicators, and adopt better strategies to maximize the sales profit. It is objective, standardized and comprehensive.

References

[1] Yang Haoxu. Correlation analysis between vegetable price and sales: taking oil wheat as an example [J]. Food Safety Guide,2018(21):179-181.

[2] Pan Xiaofei, Xie Zhiheng, Wang Shuyun, Considering the optimization decision of fresh food stores and pricing supermarket considering loss avoidance[J].Highway traffic technology,2022,39(06):177-185+190.

[3] Yang Shuai, Huang Xiangmeng, Wang Junbin. Research on joint optimization strategy of shelf allocation and pricing of fresh food [J].Supply chain Supply chain management, 2022,3(08):49-59.

[4] Xu Weichao. Review of correlation coefficients[J].

[5] Wang Zengyan, Zhou Tianzhi, Wang Dadong, Accounting for Accounts Receivable Financing[J]. National Excellent Economic Journal of Finance and Accounting Monthly, 2012:22-36.