

Automatic Pricing and Replenishment Decision for Vegetable Products based on Planning Model and Regression Algorithm

Guanyu Fu *

School of Finance, Jilin University of Finance and Economics, Changchun, Jilin, China

* Corresponding author Email: gy1207604626@outlook.com

Abstract: Based on the distribution data of this superstore, this study establishes a planning model to help the superstore make replenishment and pricing decisions for vegetable goods to obtain higher profits. First, the mapping function is used to correspond the vegetable number to the vegetable name, wholesale price, and wastage rate, and the sales data are treated as outliers. Pearson correlation analysis was utilized to explore the interrelationship between the sales volume of each vegetable category and a single product, indicating a correlation between the total sales volume of different categories of vegetables. Then, based on the regression analysis to study the relationship between the total sales volume of each vegetable category and cost-plus pricing, a planning model was established to solve the total daily replenishment and pricing strategy of each vegetable category. Finally, with profit maximization as the objective function, a planning model is developed to solve the single-item replenishment quantity and pricing strategy for a single day. This study develops better replenishment and pricing strategies for superstores.

Keywords: Superstore Pricing; Replenishment Strategy; Pearson Correlation; Planning Models.

1. Introduction

In fresh food superstores, the freshness period of general vegetable products is relatively short. The quality of the product deteriorates with the increase in sales time, and most of the varieties, if not sold on the same day, cannot be sold on the next day. Therefore, supermarkets usually replenish their stocks daily according to the historical sales and demand of each product. Because supermarkets sell many different varieties of vegetables from different origins, merchants must make replenishment decisions for each vegetable category on the same day without knowing exactly what the individual products and purchase prices will be. Vegetables are generally priced using the "cost-plus pricing" method, and supermarkets usually sell at a discount for shipping losses and poor-quality goods. Reliable market demand analysis is important for replenishment and pricing decisions. The purpose of this study is to forecast the market demand for vegetables, to guide the future demand for vegetable types and quantities, and to support the replenishment and pricing decisions of fresh food superstores. Therefore, this paper

collects three years of vegetable sales data from a superstore and establishes a relevant model to optimize the replenishment plan and pricing strategy of vegetable items, which helps the superstore to improve operational efficiency and profitability [1].

2. Modeling and Solving

2.1. Relationship between Vegetable Categories and Sales Volume Per Item

To make better replenishment and pricing decisions for vegetable items, we first analyzed the sales volume of each category and individual items of vegetables. First, use Pearson correlation analysis to explore the interrelationship between the sales volume of various categories and individual products of vegetables, and get the correlation coefficients of the sales volume of different vegetable categories and individual products. The sales volume data of different categories of vegetables and different individual products were counted [2], and the correlation graphs were plotted, as shown in Figure 1.

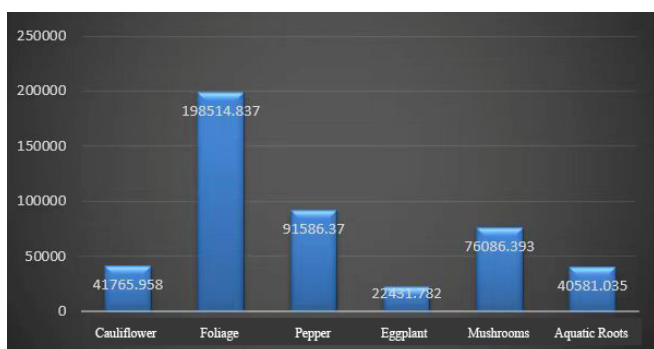
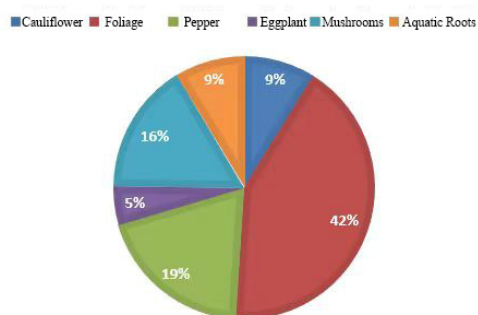


Figure 1. Total sales of vegetables by category



Pearson's correlation analysis was performed based on SPSS for sales of different categories and individual items.

Using quarters to categorize the data in terms of time, using pivot tables to sum the sales for each quarter using each

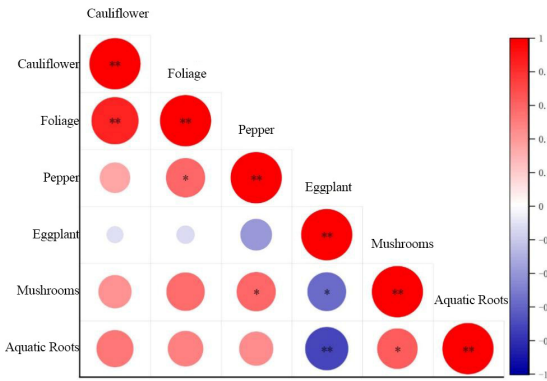
category and each item as rows, respectively, and then tabulating these quarterly sales tables to obtain a sales table with categories as rows and quarterly columns (category sales table) and a sales table with individual categories as rows and quarterly columns (individual item sales table) [3].

The Pearson correlation coefficient can be used to indicate the correlation (linear correlation) between two variables X and Y , with a value between -1 and 1.

$$r = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N \sum x_i^2 - (\sum x_i)^2} \sqrt{N \sum y_i^2 - (\sum y_i)^2}} \quad (1)$$

When $r > 0$, it indicates that the two variables are positively correlated; conversely, when $r < 0$, it indicates that the two variables are negatively correlated. The larger the absolute value of the correlation coefficient, the stronger the correlation; the closer the correlation coefficient is to 0, the weaker the correlation. Usually, you can judge the strength of the correlation between variables by its range of values.

Firstly, the Pearson correlation analysis of the total sales of each vegetable category was carried out using SPSS, and the results obtained are shown in Figure 2.



* $p < 0.05$ ** $p < 0.01$

Figure 2. Results of correlation analysis of total sales volume of vegetables by category

It can be concluded that there is a correlation between the total sales of different categories of vegetables, with a highly significant positive correlation between the total sales of vegetables in the cauliflower and foliar categories, and a highly significant negative correlation between the total sales of vegetables in the aquatic rootstocks and eggplant categories. There is a significant positive correlation between the total sales volume of vegetables between cauliflower and chili, a significant positive correlation between the total sales volume of edibles and chili and aquatic roots and tubers, and a significant negative correlation between the total sales volume of edibles and eggplant.

2.2. Relationship between Total Sales and Cost-plus Pricing for Each Vegetable Category

Firstly, the total sales volume and average price of different vegetables are analyzed in terms of categories, and the relationship between the total sales volume and cost-plus pricing of each vegetable category is studied based on regression analysis. Then, with profit maximization as the objective function and vegetable selling price and sales volume meeting the basic requirements as the constraints, a

planning model is established to solve the total daily replenishment and pricing strategy of each vegetable category in the coming week [4].

Taking quarter as the statistical unit, the average value of selling price in each quarter and the total sales of vegetables in each quarter were obtained by using a pivot table, and regression analysis based on SPSS was used to study the relationship between sales and selling price of different categories of vegetables, and the results obtained are shown in Table 1.

Table 1. Relationship between total sales volume and cost-plus pricing of different vegetable categories

Vegetable category	Regression results	R^2	Significance (P)
Foliage	$y=6641.387-1628.304x$	0.466	0.000
Cauliflower	$y=3702.895-140.822x$	0.611	0.004
Aquatic Roots	$y=4495.651-269.121x$	0.480	0.001
Eggplant	$y=1039.492-10.922x$	0.490	0.000
Edible fungi	$y=-673.860-1538.585x$	0.680	0.000
Capsicum	$y=3139.355-49.520x$	0.740	0.001

To solve the total daily replenishment and pricing strategy for each vegetable category in the coming week, a planning model is established with profit maximization as the objective function, satisfying the basic requirements of sales price and sales volume [5].

Objective function:

$$\text{Max } Y = \sum_{j=1}^n \sum_{i=1}^6 t_{ij} (x_{ij} - c_{ij}) \quad (2)$$

Constraints:

$$\min t \leq t_{ij} \leq \max t \quad (3)$$

$$\min x \leq x_{ij} \leq \max x \quad (4)$$

Where, Y denotes the superstore profit, $i=1, 2, \dots, 6$ denotes the six categories of vegetables, which are foliage, cauliflower, aquatic roots and tubers, eggplant, edible mushrooms, and chili peppers, $j=1, 2, \dots, n$ denotes a single specific vegetable, x is the matrix of vegetable selling price, c is the matrix of vegetable wholesale price, and t is the matrix of vegetable under sales volume.

By solving the model, we can find the sales unit price and sales volume when the revenue of the superstore is maximized, where the sales volume is used as the daily replenishment and the sales unit price is used as the pricing strategy, and the six categories are solved so that we can finally get the replenishment volume of the superstore and the pricing strategy according to the sales volume of each vegetable category.

2.3. Replenishment and Pricing Strategies for Individual Vegetable Items

First, according to the data related to the vegetable varieties that can be sold, to maximize the profit as the objective function, assuming that the total number of sellable varieties and single products that can only be sold is between 24-33, each vegetable category is sold, the vegetable selling price, the sales volume to meet the basic constraints, the establishment of the planning model, as follows:

$$\text{Max } Y = \sum_{j=1}^n \sum_{i=1}^6 t_{ij} (x_{ij} - c_{ij}) a_{ij} \quad (5)$$

Constraints:

$$\min t \leq t_{ij} \leq \max t \quad (6)$$

$$\min x \leq x_{ij} \leq \max x \quad (7)$$

If $t_{ij} > 0$, then $b_{ij} = 1, t_{ij} > 2.5$

$$24 \leq \sum_{j=1}^n \sum_{i=1}^6 b_{ij} \leq 30 \quad (8)$$

Where Y denotes the profit of the superstore, i denotes the six categories of vegetables, which are leafy flowers, cauliflower, aquatic roots and tubers, eggplant, edible mushrooms, and chili peppers, j denotes a single specific vegetable, x denotes the matrix of selling price of vegetables, c denotes the matrix of wholesale price of vegetables, and t denotes the matrix of the volume of vegetables sold. a denotes the matrix of whether the vegetables are available for sale, and if the vegetables are available for sale from June 24-30, $a_{ij} = 1$, and if the vegetables are not sellable on June 24-30, then $a_{ij} = 0$. The planning model is solved to derive the replenishment quantity and pricing strategy for a single item on a single day.

2.4. Other Factors Affecting Replenishment and Pricing Decisions for Vegetable Commodities

2.4.1. Requirements for Vegetable Commodities in Their Own Right

(1) Data on suitable temperature and humidity conditions

In the process of transportation, vegetable commodities are more sensitive to temperature, which will lead to different degrees of loss of vegetable commodities under an unsuitable temperature environment. It is necessary to adjust the environment during transportation according to the different properties of different commodities. When the outside temperature is high, often vegetable commodities cannot maintain a good selling point, which greatly reduces the sales volume and profitability of the superstore. Therefore, supermarkets should collect the appropriate temperature and environmental conditions of various types of vegetables, this kind of data has a significant role in reducing the loss of supermarkets. In storage, different vegetable commodities have different suitable storage conditions, it is easy because of the unsuitable temperature, and humidity conditions under the deterioration. Therefore, in inventory management, it is more important to know the suitable storage conditions of vegetable products.

(2) Supply side: seasonal timing of vegetable commodities

For certain seasonal commodities, superstores often need to determine the launch date, launch cycle and replenishment time based on the previous sales of the same period.

2.4.2. Objective Conditions

(1) Weather conditions on the day

Since the public tends not to travel and shop on rainy or cloudy days according to consumer psychology, the amount of sales for the day is predicted based on the weather conditions and replenished accordingly. In addition, when the weather is rainy, snowy, etc., and the price of goods remains unchanged, consumers tend to reduce the number of trips and the time and frequency of shopping. At this time, if you want to increase sales to achieve greater profitability, you need to change the pricing and adjust the price downward, to attract more consumers to the store; In addition, during the holidays, consumer desire and demand for consumption increases, this time the super should increase the supply of vegetable products for the super to bring higher economic returns.

(2) Demand side: types and quantities of vegetable commodities demanded

The demand for different types of vegetables varies in different regions and cities. Supermarkets in a certain region should select the varieties and quantities of vegetables on the shelves according to the characteristics of the region; in different seasons, the demand for different types of vegetables and at different times also varies. Therefore, supermarkets should restock according to the different needs of consumers; for consumers of different age groups, the demand for different types of vegetables is not the same. For example, the elderly are more inclined to eat vegetables that move when bitten, if the residents around the supermarket are mostly elderly, the supermarket should increase the supply of vegetables that the elderly love to eat.

2.4.3. Macroeconomic Level

(1) Local GDP

For different regions, the local GDP affects the sales of vegetable goods. For different economic development of different regions, the economic consumption level of each region is also different, the level of purchasing power is different, so it will lead to different sales of commodities. Therefore, the demand for vegetables in different regions is judged according to the level of economic development of each region [6].

(2) Size of the local population and percentage of each group

The amount of consumption varies from region to region in areas with different amounts of population. Undoubtedly, under the condition that the economic development of a certain place is smooth, the larger the population is, the consumption capacity and quantity will increase. Therefore, for supermarkets, it is important to evaluate and calculate the quantity to be replenished during a certain period of time and make pricing decisions based on the population size and density of a certain area and the number of corresponding supermarkets, and to replenish the corresponding proportion of vegetable products based on the proportion of the age group of the residents around the supermarket and the preference of each group of people.

3. Conclusion

The mathematical model constructed in this paper shows significant advantages in analyzing the sales volume of vegetable commodities and pricing strategy, closely integrated with the actual, using Pearson correlation analysis to explore the interrelationship between the sales volume of each vegetable category and single product, to ensure that the pricing replenishment strategy of the high efficiency, stability, and high degree of fit. At the same time, Pearson correlation analysis was conducted to analyze the total sales volume of each vegetable category, as well as the relationship between the sales volume of individual items. However, the model was deficient in accounting for external factors such as vegetable availability, climate, and precipitation, and had limited adaptability to long-term sales and price trends. Therefore, although the model is valid under the current conditions, it needs to be further improved and adapted in future applications to adapt to a wider range of market and environmental changes, to more comprehensively improve the sales efficiency and profitability of vegetable commodities in supermarkets.

References

- [1] Du Heng, Lu Ke. Retailer pricing strategy considering consumption heterogeneity under supply shortage[J/OL]. China Management Science, 1-17[2023-12-19] <https://doi.org/10.16381/j.cnki.issn1003-207x.2021.2453>.
- [2] War Shuai. Research on quality influencing factors of public welfare fundraising platform based on Pearson correlation analysis [J]. International Public Relations,2022, (19): 71-73. DOI: 10.16645/j.cnki.cn11-5281/c.2022.19.057.
- [3] Su Pingping, Lv Fuqiang. Research on the design of intelligent creation of community vegetable farms based on SPSS statistical analysis--Taking Hohhot racecourse railroad district as an example[J]. Footwear Craft and Design,2023,3(22):157-159.
- [4] JIA Yongqing, GE Song, XI Bingjie. Logistics-based regression analysis of risk factors for respiratory failure in patients with chronic obstructive pulmonary disease[J]. Journal of Rare Diseases,2023,30(12):42-44.
- [5] HUANG Sen, XU Xiangdong. Reliable path planning models and algorithms for transportation networks considering random travel time heterogeneity of road segments[J/OL]. Journal of Transportation Engineering, 1-16[2023-12-19] <http://kns.cnki.net/kcms/detail/61.1369.U.20231206.1449.002.html>.
- [6] SUN Ying, FAN Jin, JIA Weiguo. Research on the Stages and Effective Paths to Achieve Carbon Peak - Based on the Perspective of Green GDP Accounting[J]. Industrial Technology and Economics,2023,42(12):95-104.