The Research on Optimal Replenishment Decision-making strategies for Vegetable Markets in the Context of Big Data Analysis

Lu Zheng, Weilun Jing, Jinxu Lou, Zhuo Qi *

Liaoning Institute of Science and Engineering, Jinzhou, China, 121000

* Corresponding Author Email: qizhuo198504@163.com

Abstract. In consideration of the limited shelf life of vegetable commodities, this study analysed the correlation between sales data for said commodities. The analysis was then used to propose pricing and replenishment decisions for vegetable commodities. Firstly, correlations between sales volume and each medium vegetable category were analysed using bubble charts. the results showed that leafy vegetables and cauliflower, edibles and aquatic rootstocks exhibited significant positive correlations (Pearson $r > 0$), while aubergines exhibited negative correlations with edibles and chillies. Secondly, a seasonal index model was developed and seasonal indices and other influential parameters were calculated, the analysis also enabled the prediction of replenishment and pricing for the following week. Finally, the restocking volume of the six categories in the coming week was predicted and the pricing strategy was formulated and this followed the dynamic pricing pattern of "high early-week→mid-week discounts→weekend rebound". The present study puts forward a series of recommendations for vegetable markets regarding their replenishment decision-making strategies, which has broad practical value and promotional significance in guiding intelligent decision-making and optimizing procurement and replenishment strategies.

Keywords: Seasonal index model, Pricing and replenishment decision, Vegetable commodities.

1. Introduction

In the fresh produce market, particularly in the speciality vegetable market where fresh produce has a short shelf life (typically 1-2 days), the consequences of overstocking or out-of-stocking are well-documented [1]. In addition, rising consumer expectations for freshness and convenience are forcing supermarkets to synchronise rapid replenishment with dynamic pricing strategies. This dual need to reduce waste and maintain profitability also makes business decisions in the fresh produce market more challenging [2].

To mitigate these risks, big data analytics has emerged as a disruptive enabler. Pivotal studies have demonstrated its effectiveness: Herbon A et al (2017) Optimal dynamic pricing and ordering of perishables under the additive effects of price and time on demand, indicate the extent to which retailers can benefit from implementing a dynamic pricing strategy [3]. Dellino G et al (2018) also studied the pricing of goods based on machine learning and operations research [4]. Kayikci Y et al (2022) have innovatively proposed a dynamic pricing model based on real-time Internet of Things (IoT) sensor data to provide retailers with pricing decision options for each stage of the sales cycle [5]. Li Z's (2024) research focused on optimal replenishment and pricing strategies for different supply and demand scenarios, as evidenced by their study of sales data [6].

While the evidence presented here indicates the transformative potential of data-driven approaches, it is clear that further research is required into the adaptability of vegetable markets to seasonal and demand fluctuation patterns.

Despite advances in demand forecasting and inventory optimization, the exclusion of external variables such as weather variability and supplier reliability ignores the realities of multi-level replenishment [7-9].

In this paper, the total sales volume of each vegetable category was enumerated, and the correlation between the sales volume of each category and the influence of season on sales volume was derived. Secondly, a seasonal index model was established, and the seasonal index and other parameters were calculated. Lastly, replenishment and pricing in the vegetable market were predicted based on the simulation results.

2. Seasonality and sales relationships between products

2.1. Relationship between sales volume and seasons

This paper uses a variety of vegetable sales data from supermarkets over the course of their operations to forecast demand for goods and thus further optimize the replenishment and inventory management of supermarkets. Developing replenishment and pricing strategies is an important aspect of running a supermarket. Seasonality is a key factor that cannot be ignored as it will have a significant impact on sales [10]. With this in mind, the supermarket counted the total sales of each vegetable category on a monthly basis, and the relevant data is shown in Figure 1.

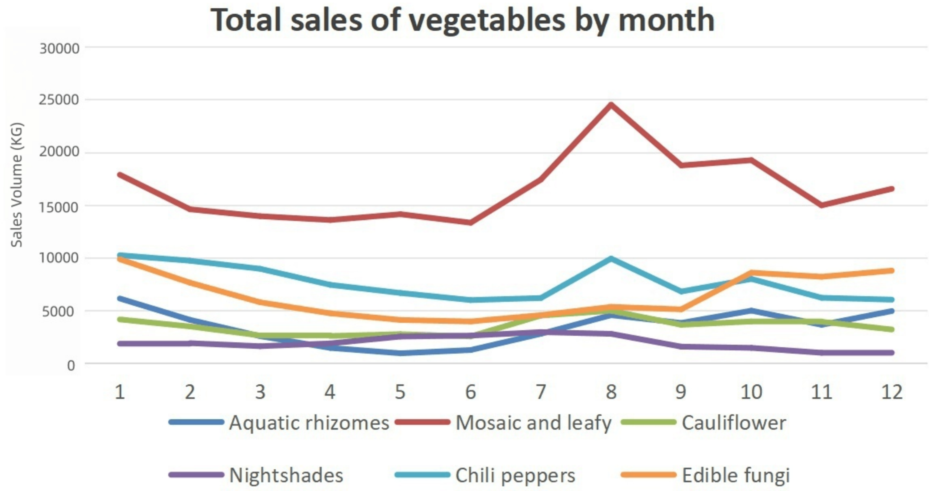

Figure 1. Total sales of vegetables by month

As demonstrated in Figure 1, the sales volume of each vegetable category is seasonally influenced. For instance, cauliflower and leafy vegetables demonstrate high sales volumes on a monthly basis. In addition, chillies, edible mushrooms, aquatic roots and tubers exhibit the highest sales volumes in January, while cauliflower demonstrates the highest sales volume in August. Notably, nightshades achieve the highest sales volume in July. It is imperative for supermarkets to formulate a scientific and reasonable replenishment strategy based on the relationship between seasons and vegetable sales, in order to optimise their operations and management. In advance of the peak season, supermarkets can arrange the replenishment volume according to the sales volume of each type of vegetable.

2.2. Correlation of sales volume of individual commodities

In the context of analysis the relationship between sales of different products, the use of correlation analysis bubble charts provides a means to observe the correlation of sales in the distribution of scattered dots, with the size of the bubbles representing the magnitude of the correlation coefficient between the corresponding categories. Specifically, points with larger blue bubbles indicate higher correlation, with larger points corresponding to higher correlation coefficients (Pearson r>0). Conversely, smaller bubbles indicate a correlation coefficient of approximately 0, whereas larger red bubbles indicate negative correlation and larger points indicate a lower correlation coefficient.

An investigation was conducted to ascertain the correlation between various categories of vegetables. The correlation between the sales volume of different individual items in different vegetable categories is also an important influencing factor in the formulation of replenishment and pricing strategies for vegetable individual items, the following Figure 2 illustrates the correlation between the various categories of vegetables.

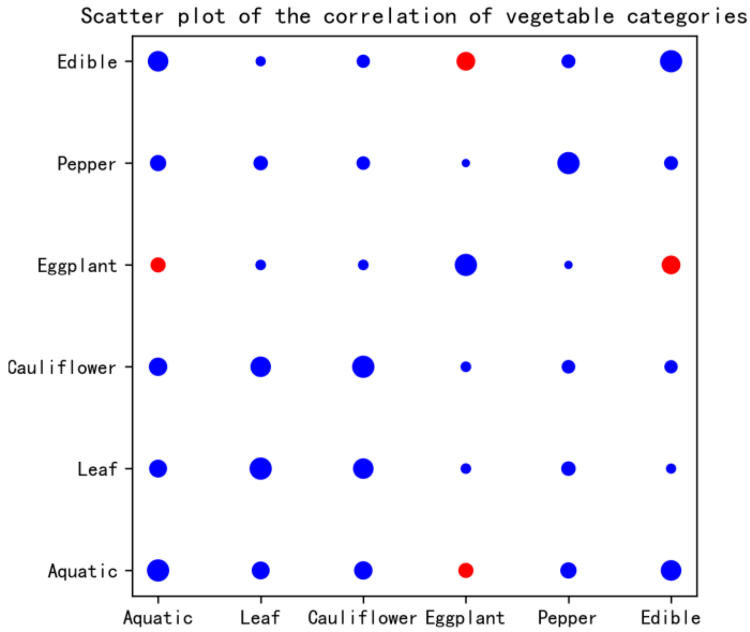

Figure 2. Scatterplot of correlation of vegetable categories

The study revealed a relatively weak correlation between peppers and eggplant, as well as a weak correlation between eggplant and edible fungi. However, a strong correlation was observed between leafy vegetables and cauliflower, as well as a strong correlation between edible fungi and aquatic rhizomes.

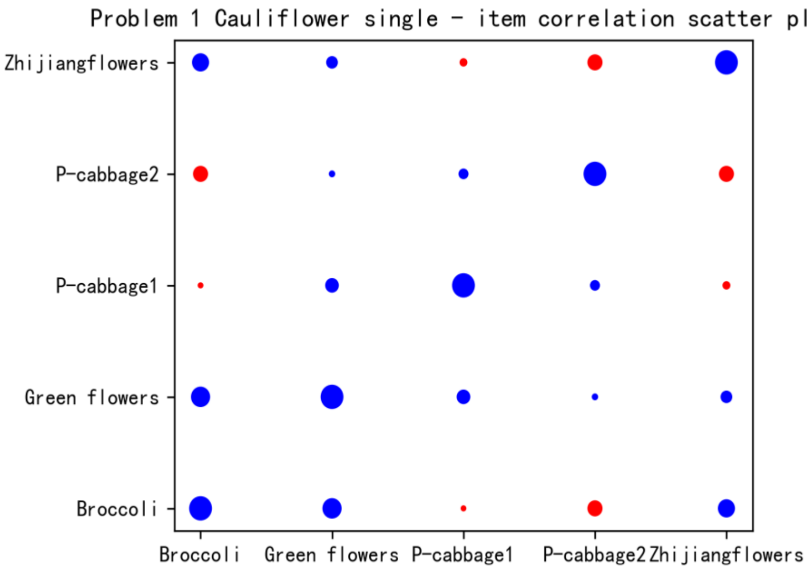

Figure 3. Cauliflower single - item correlation scatter plot

For the cauliflower category, the correlation coefficients between broccoli and clear-stemmed loose flowers show a high correlation, and the correlation coefficients of purple cabbage and green-stemmed loose flowers are negative as shown in Figure 3.

3. Replenishment-ricing integrated optimization

Given that the sales volume and order quantity of commodities will be subject to both temporal and spatial constraints, supermarkets must forecast the sales volume and order quantity of commodities and its trend in the forthcoming period, and devise the most effective plan for supermarket replenishment to maximise revenue and minimise loss due to the substandard quality of vegetables resulting from an inadequate replenishment plan within the constraints of a specified spatial capacity.

To this end, this paper will utilise existing data to predict the total daily replenishment for the next seven days and provide the final pricing strategy.

This section determines the optimal replenishment strategy by modelling the seasonality index and the associated constraints, and provides the pricing strategy for each product.The seasonality index model adopts the time series decomposition method to divide the historical sales data into trend, seasonal and stochastic terms, where the seasonal term is calculated by the moving average ratio method to reflect the demand fluctuation characteristics in different months. The product categories encompassed in this study are diverse. the wastage rates and average wholesale prices of these products during the marketing process are elucidated in Figure 4 and Figure5.

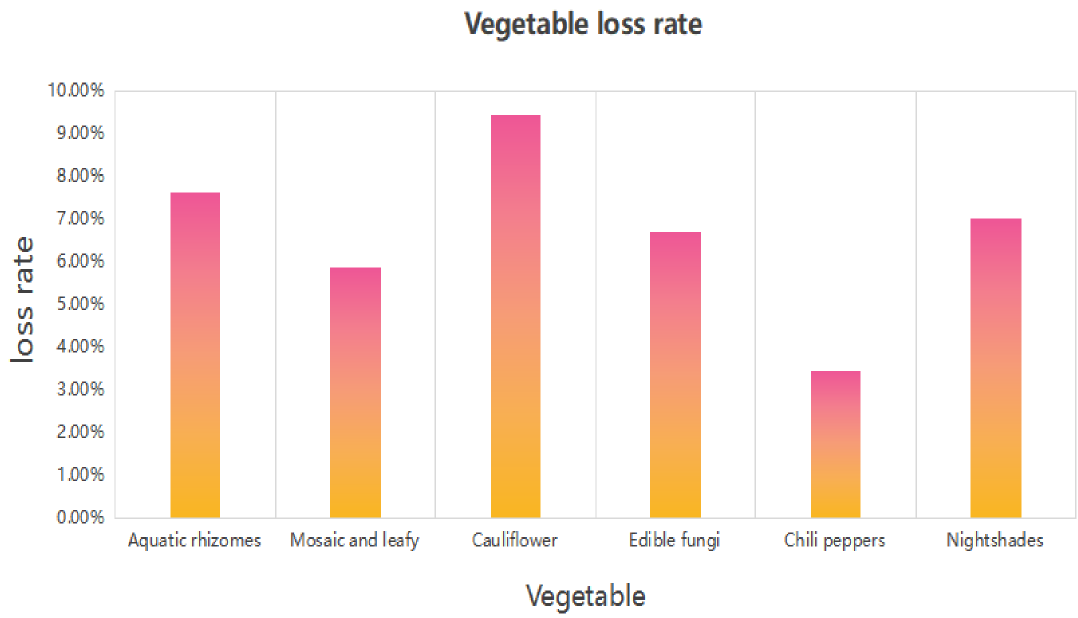

Figure 4. Vegetable loss rate

Figure 5. Average Wholesale Price

As demonstrated in Figure 4, cauliflower exhibits the highest rate of wastage, with a percentage of approximately 3.5%. In contrast, chilli demonstrates a comparatively low rate of wastage, with an approximate figure of 3.5%. The figure also reveals that edible fungi are associated with the highest unit price of sale, while cauliflower and chilli are positioned at the second and third highest prices, respectively. Nightshades, conversely, is associated with the lowest unit price of sale.

The total sales volume of vegetable commodities for the past three weeks are shown in Figure 6.

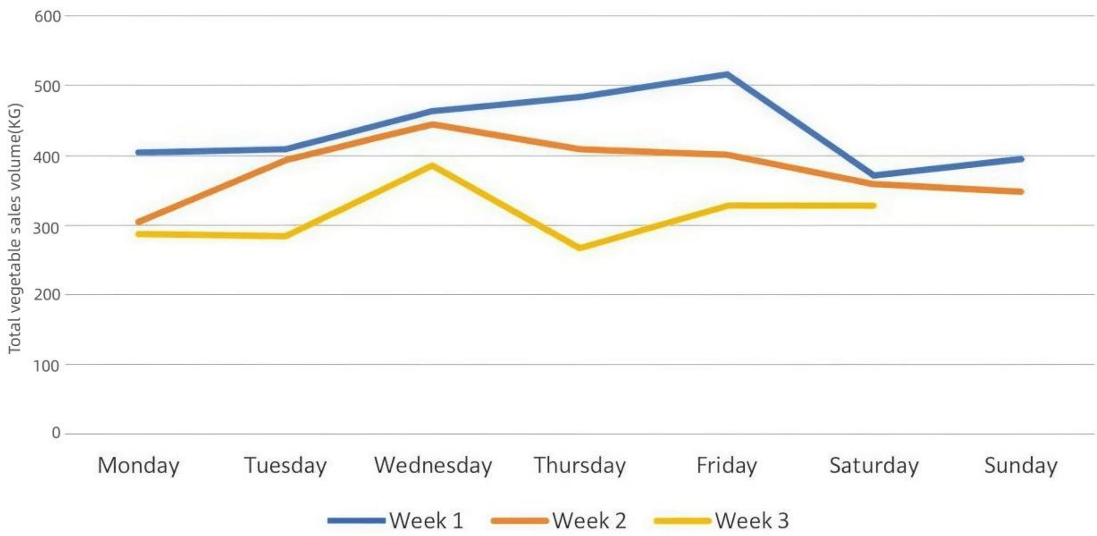

Figure 6. Total sales of vegetable products

It is evident from further analysis and summarisation of the available data that the total sales volume of vegetable items exhibits a cyclical distribution. In the context of cyclical data, the formulation of a strategy for the subsequent week's replenishment and pricing is contingent upon the analysis of the relationship between the total sales volume and the cost-plus pricing. This is due to the seasonal nature of the commodity price. The employment of a time series model for forecasting is therefore recommended. The present paper aims to develop a seasonal index forecasting model.

Modelling seasonal indices generally involves the following steps: data pre-processing, calculation of seasonal indices, building a forecasting model, and model evaluation and adjustment. For time series that contain both a linear trend component and a seasonal component, their components need to be decomposed, and this decomposition is based on the following multiplicative model:

$$Y_t = T_t \times S_t \times I_t \tag{1}$$

Where Y_t denotes the sales forecast for month t, T_t denotes the trend component for month t, S_t denotes the seasonal component for month t and I_t denotes the irregular component. Because of the unpredictability of the irregular component, the forecast can be expressed as the product of the trend component and the seasonal component.

The total amount of daily vegetable sales data from the shop over the past three weeks was taken as the actual value of the total amount of daily purchases in the shop. A scatter plot of the average weekly total amount of purchases over the past three weeks was then created, and the regression function was labelled.

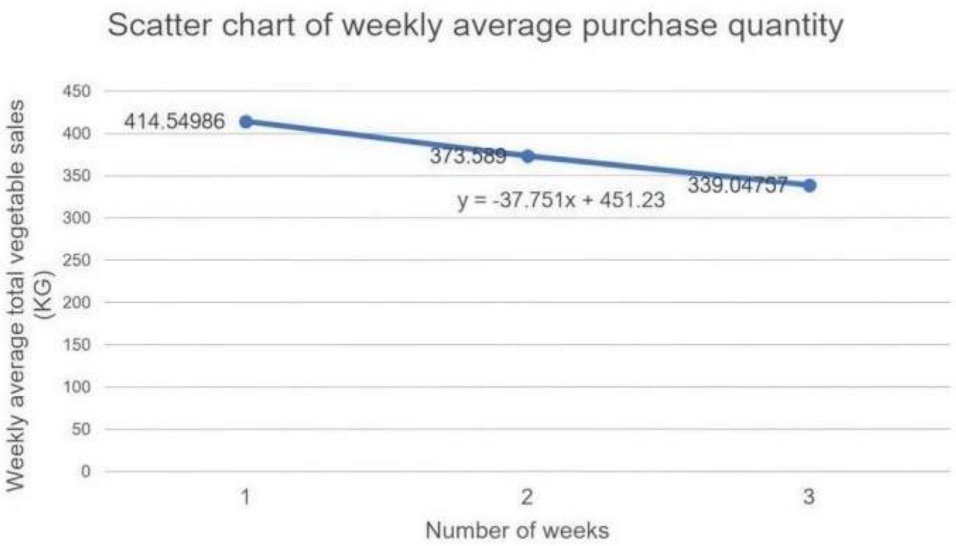

Figure 7. Scatter chart of weekly average purchase quantity

As illustrated in Figure 7, a seasonal index model for the long-term trend is established, and the seasonal index is obtained in order to predict the latest vegetable sales quantity. Through the integration and processing of the aforementioned data, the sales volume of vegetable commodities in the shop is extracted over the past three weeks. A table of the actual sales volume data is established, and a scatter plot is profitably plotted to mark the trend line and the linear function, with the functional expression:

$$y = -37.751x + 451.23 \tag{2}$$

Where x denotes the number of weeks, y denotes average weekly sales in kg of the vegetable.

The daily data from the previous three weeks must be symmetrized proportionally in order to calculate the trend sales volume of all vegetable commodities from the same period. The mean of these values must then be determined.

At this point, S_t = actual average / T_t, and the trend index corresponding to each day is then obtained. The trend value of daily sales volume of the shop over the past three weeks is then multiplied by the trend index to obtain the specific value of predicted sales volume data for all vegetable categories in the coming week. The weekly average of the total vegetables' volume over the past three weeks can be calculated as shown in Table 1:

Table 1. The total volume of vegetables

weeks	first week	second week	third week
on average (kg)	414.5498571	373.589	339.0475714

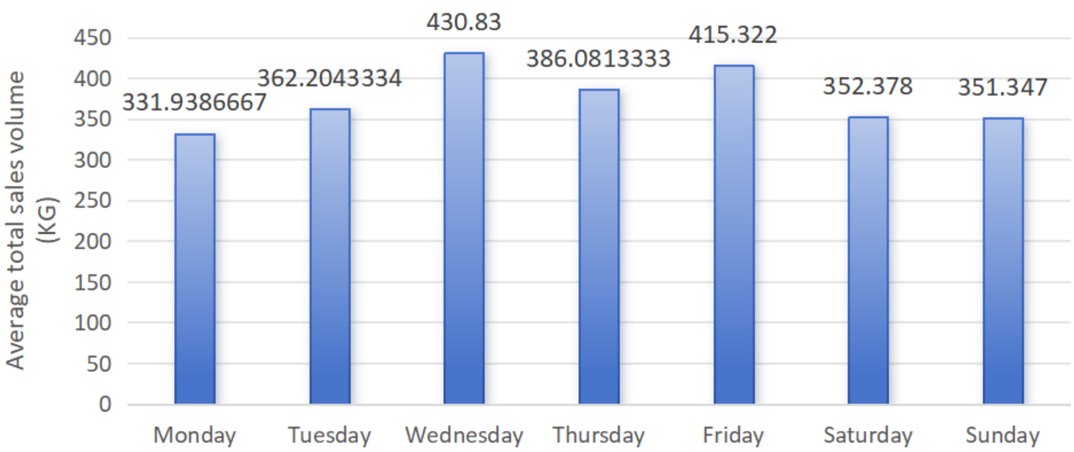

Figure 8. Average weekly sales of various vegetables with the total quantity

As demonstrated shown in Figure 8, the trend value of total vegetable merchandising can be calculated. It shows the average weekly sales of various vegetables with the total quantity is high in wedensday andfriday.

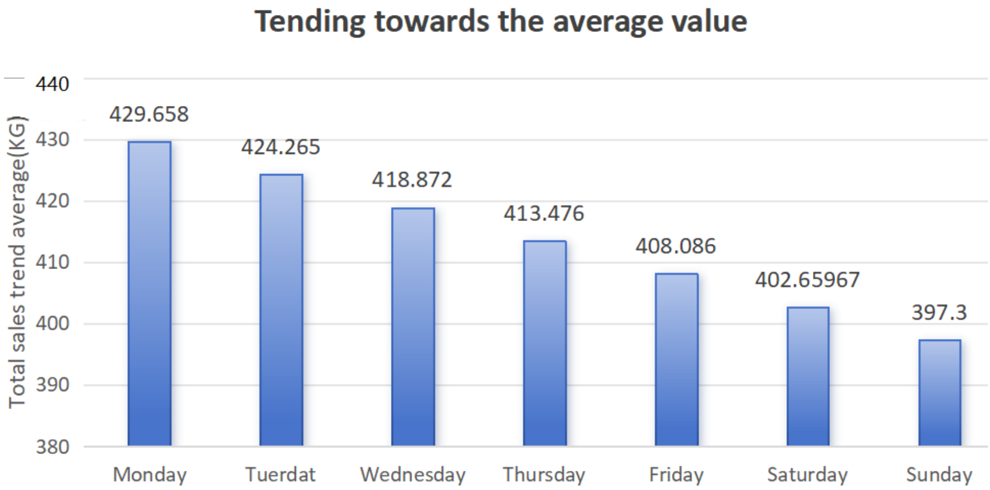

Figure 9. Tending towards the average value

As shown in Figure 9, the graph illustrates the step in the seasonal index model employed for the purpose of extracting the trend and eliminating seasonality. This is achieved by smoothing out the short-term fluctuations in the original data and converging to the mean.

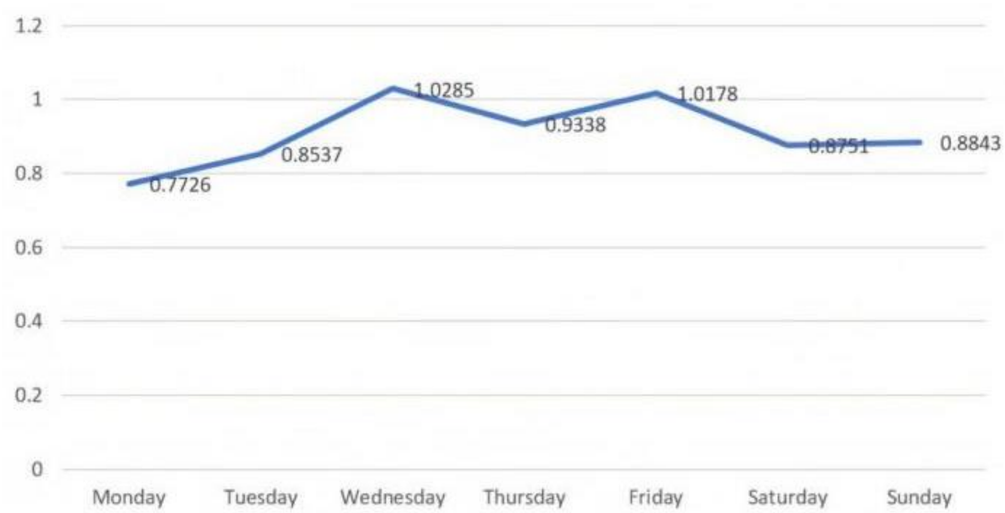

Figure 10. Trend seasonal index

As demonstrated in the Figure 10, the trend-adjusted seasonal index of total vegetable merchandising can be calculated. Utilising this data, the following linear scatter trend graph can be plotted, which illustrates the trend value of total vegetable merchandising. As illustrated in Figure 11, the functional expression for the given graph is as follows:

$$y = -5.3924x + 472.79 \tag{3}$$

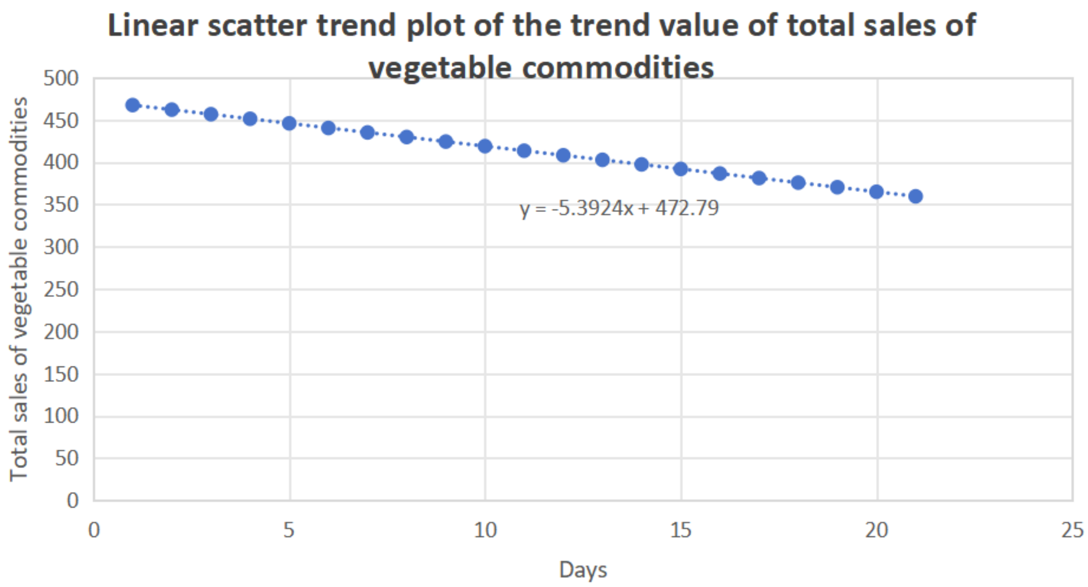

Figure 11. Linear Scatter Trend plot of the Trend value of sales of vegetable commodities

The total amount of vegetables sold per day is predicted to be:

$$\begin{gathered}(-5.3924x_i + 472.79)\times S_t \\ x_i = 21, 22, \ldots, 27\end{gathered} \tag{4}$$

This model is then used to predict the total daily sales of vegetables in shops for the following week, as illustrated in Figure 12.

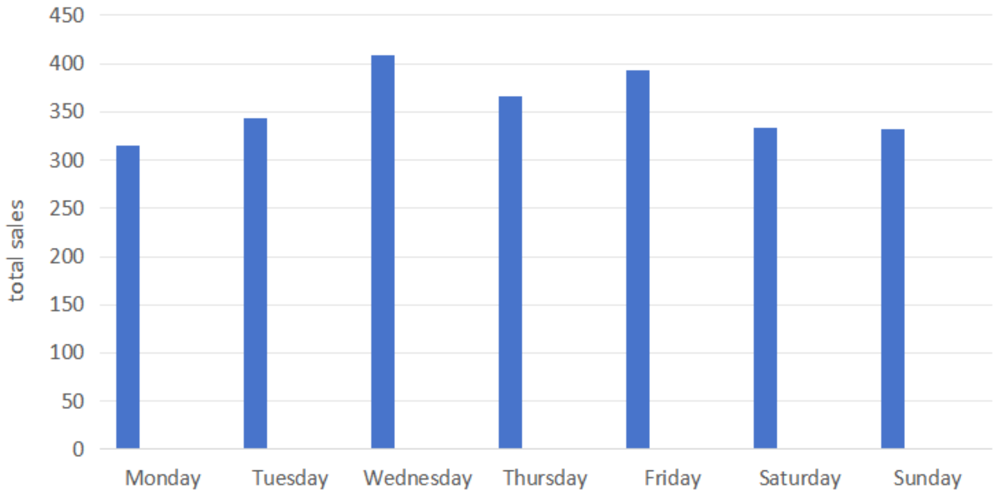

Figure 12. Forecast of total sales

The weights of each vegetable type will be converted into a class based on its wastage rate, in order to ascertain the sales volume of each type over the next seven days of the week. As the wastage rate is a cost-effective attribute, it can be assumed that a greater wastage rate will result in greater weights. The weights of the six categories of vegetables are calculated as follows:

$$\omega = [0.1316、0.1706、0.1061、0.1493、0.2915、0.1429] \tag{5}$$

The predicted values of the total daily sales of vegetable commodities above are multiplied by their respective weights to arrive at the predicted values of the daily sales of each category of vegetable commodities for the next seven days, as illustrated in the Figure 13.

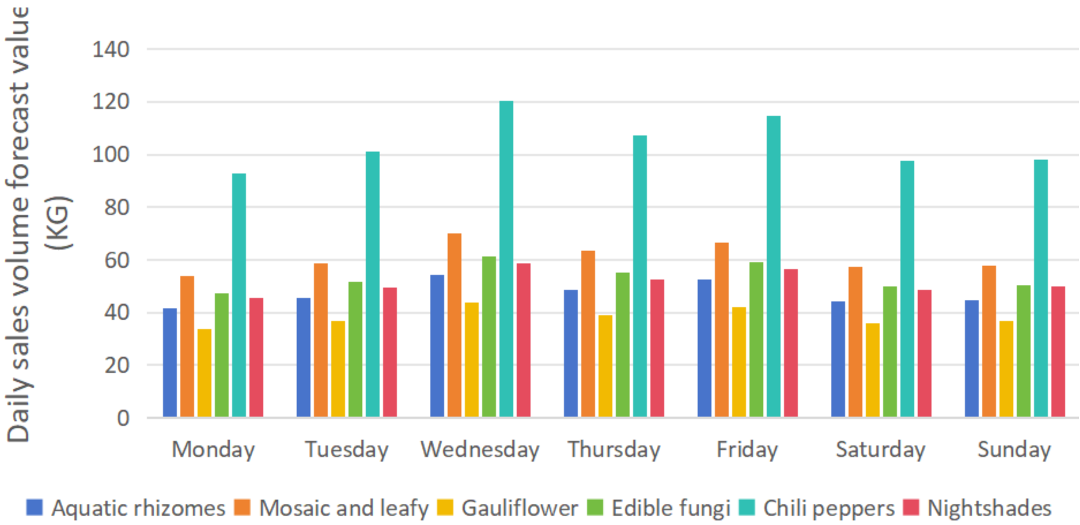

Figure 13. Forecast value for each vegetable product

Table 2. Replenishment volume for 6 categories in the next week

Categories (kg)	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Aquatic rhizomes	41.79	45.58	54.21	48.54	52.78	44.29	44.61
Mosaic and leafy	54.09	58.70	70.15	63.46	66.61	57.29	57.62
Cauliflower	33.75	36.81	43.79	39.21	42.24	35.79	36.79
Edible fungi	47.37	51.67	61.45	55.02	59.23	49.88	50.20
Chili peppers	92.93	101.21	120.18	107.06	114.81	97.62	97.94
Nightshades	45.36	49.47	58.83	52.68	56.72	48.67	48.99

As shown in Table 2, the data show variations in daily supply: aquatic rhizomes (41.79-54.21 kg), mosaic and leafy vegetables (54.09-70.15 kg), cauliflower (33.75-43.79 kg), edible fungi (47.37-61.45 kg), chilli peppers (92.93-120.18 kg, which is the highest demand) and nightshades (45.36-58.83 kg).

Pricing of vegetable commodities is generally based on the "cost-plus pricing" method.

$$\text{cost} = \text{wholesale price} * (1 + \text{wastage rate}/100 \text{ per cent}) \tag{6}$$

This is a system of calculation whereby the cost of each category is determined by the wholesale price, the wastage rate and the forecasted sales volume. The initial step involves the amalgamation and compilation of the wholesale price, wastage rate and forecasted sales volume. Subsequent to this, the cost of each category is calculated. The cost of each category is illustrated in Figure 14.

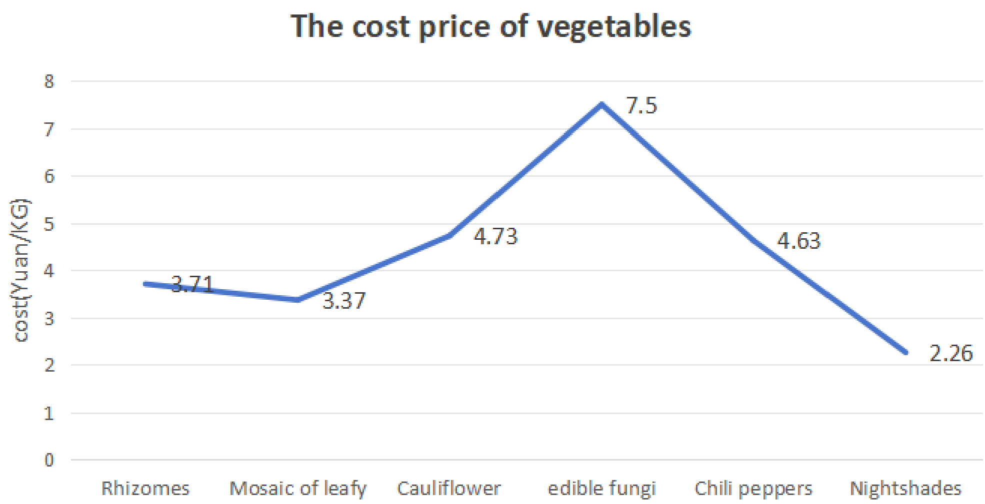

Figure 14. The cost price of vegetables

Based on the expected sales volume of different types of vegetables, they were compared to determine the mark-up rates for different total sales volumes:

(1) If the expected sales of vegetables is less than 50kg , then the mark-up rate = 145 per cent

(2) If the expected sales of vegetables is exceeding 50kg , then the rate of mark-up = 123 per cent

Since the level of sales volume of an item will indirectly affect the sales rate and sales heat of a single item, the rate of return of the superstore can be balanced by adjusting the mark-up rate of the item in time, so that it tends to maximise the rate of return. By calculating the selling price of each category of vegetables (selling price = cost * mark-up rate), finally the pricing strategy of the different categories of vegetables in the superstore for the coming week is derived as shown in Figure 15.

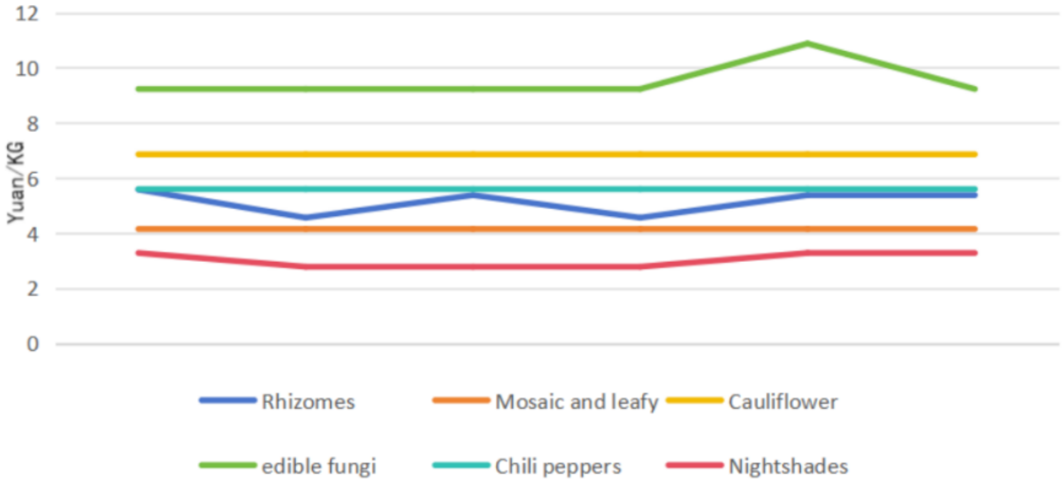

Figure 15. The pricing strategy for different categories of vegetables for the coming week

Table 3. Pricing for 6 categories in the next week(Yuan/kg)

Categories	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Aquatic rhizomes	5.58	5.58	4.56	5.38	4.56	5.38	5.38
Mosaic and leafy	4.15	4.15	4.15	4.15	4.15	4.15	4.15
Cauliflower	6.86	6.86	6.86	6.86	6.86	6.86	6.86
Edible fungi	10.88	9.23	9.23	9.23	9.23	10.88	9.23
Chili peppers	5.6	5.6	5.6	5.6	5.6	5.6	5.6
Nightshades	3.28	3.28	2.78	2.78	2.78	3.28	3.28

As shown in Table 3, pricing strategies vary: aquatic rhizomes peak at ¥5.58 early in the week before fluctuating mid-week (¥4.56-5.38); edible fungi reach ¥10.88 on Mondays/Saturdays (¥9.23 otherwise); nightshades drop to ¥2.78 mid-week but recover to ¥3.28 at the weekend, while mosaic and leafy (¥4.15), cauliflower (¥6.86) and chillies (¥5.6) remain stable. 60% of categories follow a dynamic pricing pattern of "high early-week→mid-week discounts→weekend rebound", providing supermarkets with data-driven strategies to balance inventory turns and profits.

The ability to make effective replenishment decisions necessitates a multi-dimensional analysis framework.

Firstly, demand forecasting integrates historical sales data decomposition (the trend of the historical sales data, seasonality, and cyclical patterns) with real-time market signals. Secondly, inventory management balances stock levels through dynamic safety stock adjustments. Thirdly, effective supply chain coordination involves the evaluation of supplier reliability metrics.

Crucially, external risks such as weather-induced freshness decay are quantified through regression models. Finally, customer segmentation guides strategy refinement: time-series analysis of demographic purchasing patterns enables targeted replenishment aligned with forecasted demand clusters.

4. Conclusions and Implications

This project focuses on the vegetable market and aims to explore optimal replenishment and pricing strategies.

In the inquiry process, data analysis is used to develop optimal replenishment and pricing strategies for vegetable commodities. To arrive at an accurate and effective strategy, a number of influencing factors must be fully considered. Among them, the influence of seasonal factors on the sales volume of commodities should not be underestimated. People's demand for vegetables varies greatly from season to season, for example, in summer, cool melon vegetables tend to have higher sales, while in winter, cold-resistant root vegetables are more favoured by consumers. At the same time, there is also a complex interaction between commodities and commodities.

Based on this, the sales of various vegetable commodities are listed on a monthly basis. This initiative enables a detailed look at the fluctuations in vegetable sales from month to month, thus the data characteristics of different seasons were analysed in depth, as if searching for the codes of vegetable sales in different seasons. Subsequently, we drew a scatterplot for correlation analysis of different kinds of vegetables. Through this scatterplot, we clearly derived the correlation between different types of vegetables and between different individual items of the same type of vegetables. For example, we found that the sales volume of each category of vegetables was strongly related to the season, The correlation between these items has a non-negligible impact on replenishment and pricing.

In examining the optimal solution to replenishment and pricing, the significant impact of vegetable wastage rates in the vegetable sales process. Firstly, an evaluation is carried out on the known data, which can accurately determine the validity and stability of the model. On this basis, the optimal replenishment and pricing strategies for each category of vegetable goods were then carefully calculated and given by applying cost-plus pricing and weighting method.

Finally, to further enhance the usefulness of this project, we have also listed a series of relevant factors that will have an impact on the vegetable replenishment and pricing situation in actual sales. These factors cover a wide range of aspects such as changes in market supply and demand, shifts in consumer preferences, and fluctuations in transport costs. By considering these factors, the results of the project can be better adapted to the complex and changing market environment, and provide more valuable references for vegetable market operators.

References

[1] Zhang J, Zhou X Y, Chi L, et al. Vegetables Market Situation Analysis 2024 and Outlook 2025 [J]. China Vegetables, 2025, (01): 1 - 6.

[2] Bayir B, Charles A, Sekhari A, et al. Issues and challenges in short food supply chains: A systematic literature review [J]. Sustainability, 2022, 14 (5): 3029.

[3] Herbon A, Khmelnitsky E. Optimal dynamic pricing and ordering of a perishable product under additive effects of price and time on demand [J]. European Journal of Operational Research, 2017, 260 (2): 546 - 556.

[4] Dellino G, Laudadio T, Mari R, et al. A reliable decision support system for fresh food supply chain management [J]. International Journal of Production Research, 2018, 56 (4): 1458 - 1485.

[5] Kayikci Y, Demir S, Mangla S K, et al. Data-driven optimal dynamic pricing strategy for reducing perishable food waste at retailers [J]. Journal of Cleaner Production, 2022, 344: 131068.

[6] Li Z. Pricing and Replenishment of Vegetable Commodities Based on Machine Learning and Operations Research [J]. Advances in Computer and Communication, 2024, 5 (2).

[7] Li J, Xu X. Optimality of (s, S)-policies in a dynamic pricing model with replenishment opportunities: A sample path approach [J]. Operations Research Letters, 2013, 41 (6): 719 - 725.

[8] Chen X, Simchi-Levi D. Coordinating inventory control and pricing strategies: The continuous review model [J]. Operations Research Letters, 2006, 34 (3): 323 - 332.

[9] Chen Xin, Hu Peng, Hu Zhenyu, et al. Efficient Algorithms for the Dynamic Pricing Problem with Reference Price Effect [J]. Management Science, 2017, 63 (12).

[10] Withycombe R. Forecasting with combined seasonal indices [J]. International Journal of Forecasting, 1989, 5 (4): 547 - 552.