

# Optimizing Vegetable Commodity Pricing and Replenishment: An Integration of ARIMA and Nonlinear Programming Approaches

Haoyu Liu<sup>1, \*</sup>, Rui Hu<sup>1</sup>, Guocan Qiu<sup>2</sup>

<sup>1</sup>School of Civil Engineering, Xiamen University Tan Kah Kee College, Zhangzhou, China, 363105

<sup>2</sup>School of accounting and finance, Xiamen University Tan Kah Kee College, Zhangzhou, China, 363105

\* Corresponding Author Email: QUS21029@gmail.com

**Abstract.** Automated pricing and replenishment decision-making for vegetable items in supermarkets has attracted the attention of many operators and become a trend, and has become an indispensable part of supermarkets for many operators. In order to develop reliable automatic pricing and replenishment decisions for vegetable products, this paper firstly uses logistic regression analysis to obtain the regression equations of total sales and cost-plus pricing for each vegetable category, and then uses the ARIMA time-series forecasting model to forecast the average wholesale prices of vegetable categories for the next 7 days. After selecting the average wholesale price and pricing-sales volume model for each vegetable category, a nonlinear programming model is constructed to solve the total daily replenishment and pricing strategy for each vegetable category in the coming week. Considering the constraints of limited merchandising categories, this paper introduces three constraints: the number of saleable items, the minimum display volume and the selling space, and estimates the future wholesale price by using the historical average to make a prediction. The optimal replenishment quantity and price are derived by combining the nonlinear programming model and guided by the constraint function.

**Keywords:** Autoregressive Differential Moving Average Model, Constraint Function, Optimisation Model

## 1. Introduction

As China's agricultural sector advances and living standards elevate, consumers' appetite for premium fresh products at reasonable prices has surged.[1] Concurrently, market demand is increasingly volatile, influenced by the ever-shifting consumer behaviors and the progressively shorter product life cycles.[2] Given the perishable nature of most vegetables, items unsold within a day typically become difficult to market the next day. Consequently, supermarkets frequently restock based on historical sales and demand data. Supermarkets, serving as the cornerstone for daily life necessities, hold pivotal roles in multiple facets of life. However, they are simultaneously faced with stringent criteria when it comes to automated pricing and replenishment decisions for vegetable commodities. As such, automating these decisions has captivated many operators, evolving into a modern trend and an integral component of effective supermarket management. Bolstered by the swift progression of Internet technologies, the realm of automated pricing and replenishment for vegetables has transformed into a state-of-the-art business strategy. This strategy integrates computational capabilities and data modeling, aiming to harness potential benefits in the most efficient and streamlined manner.

Current automated strategies for vegetable commodity pricing and replenishment predominantly emphasize historical sales and daily demand, often neglecting the nuanced market dynamics. External market variables can considerably influence sales profitability, rendering vegetable commodity forecasting and market capture increasingly complex. Extracting salient indicators from extensive market data to devise robust automated pricing and replenishment mechanisms presents a significant challenge.

Earlier studies have delved into the integrated decision-making framework of investment, dynamic pricing, and replenishment for eco-friendly distribution technologies catering to perishables without stock-outs. Spearheaded by online retailers, this initiative focuses on leveraging green value-added efforts to stimulate consumer demand. By instituting a unified decision-making model centered on maximizing online retailers' profit over time, and drawing upon Pontryagin's Maximum Principle, an optimal distribution paradigm for perishables has been developed. This model elucidates optimal strategies for green technology investments, dynamic pricing, and perishable replenishments. Numerical validations and sensitivity analyses pertaining to key system parameters further substantiate these findings.[3]

We conducted an analysis on the sales data of vegetable items spanning from 1 July 2020 to 30 June 2023, sourced from a supermarket located in an urban setting. To optimize the revenue stream for this establishment, we evaluated two distinct scenarios, aiming to formulate a robust automated pricing and replenishment framework for vegetable commodities.

In an endeavor to meet market demand, the supermarket is postulated to strategize its replenishment on a categorical basis, constraining the overall count of salable items and ensuring a minimum display quantity to fulfill the requisites of each commodity. We assessed the correlation between the aggregate sales volume of each vegetable category and its respective cost-plus pricing. Subsequent strategies for daily replenishment and pricing across vegetable categories for the week spanning 1-7 July 2023 were devised to maximize potential revenue. A regression analysis facilitated the derivation of a linear relationship between cumulative sales and cost-plus pricing. For the latter, recognizing the time frame as short-term, we employed the ARIMA time-series forecasting model to anticipate the average wholesale prices of vegetable categories over the impending week. A detailed analysis of sales metrics and wholesale pricing for vegetable commodities within an urban supermarket was undertaken. This exploration permitted the aggregation and processing of sales and pricing data to determine the cumulative sales for each vegetable category. The synthesized data is presented in Figure 8.

On 1 July 2023, the replenishment quantity and pricing strategy for individual vegetable items were determined, considering the confined sales space within the supermarket and regulating the total number of marketable items, all while adhering to a minimum display volume threshold of 2.5 kg. Drawing inspiration from the optimization problem presented in [4], historical data was harnessed to forecast future sales volumes for individual vegetable items. This involved an in-depth analysis of parameters like sales history, pricing, and wastage rates, with the overarching aim of revenue maximization for the supermarket. The strategy is further constrained by three parameters: the total count of marketable individual items, the stipulated minimum display quantity, and the available sales space. The crux of this research lies in optimizing both replenishment quantity and pricing strategy to augment revenue while satisfying these constraints. The culmination of this process involved the synthesis of a nonlinear programming model, steered by an underlying constraint function, to ascertain the optimal replenishment quantities and corresponding prices.

## **2. Modeling and Strategy for Weekly Vegetable Pricing and Replenishment (1-7 July 2023) in a City Supermarket**

From an in-depth analysis of the sales flow data, a pronounced cyclicity within the dataset is evident. Our projection for the daily replenishment of a supermarket in a specific city, spanning 1-7 July 2023, categorizes this as a short-term trend rather than a long-term one. For this forecast, we predominantly employ the Autoregressive Integrated Moving Average model (ARIMA) [5].

The ARIMA model, a staple in time series analysis, excels in delineating and anticipating trends within time series datasets. This model's foundation rests on the premise that temporal sequences exhibit discernible trends and periodicities, which can be comprehensively represented through autoregressive, differencing, and moving average computations. The ARIMA model encompasses three primary components:

Autoregressive (AR): The autoregressive component captures the influence of previous time points on current values within a time series. It involves a linear relationship where the present value is a combination of its recent predecessors, accounting for specific time lags.

Integrated (I): Differencing seeks to stabilize the time series by eliminating trend and seasonality. Through this operation, an initially non-stationary series can be transformed into a stationary one, facilitating better predictability.

Moving Average (MA): The moving average component accounts for the impact of past errors on the current value in the time series. It models the present error as a linear combination of past error terms across specified time lags. The parameters of the ARIMA model are usually denoted as  $p$ ,  $d$  and  $q$ , which correspond to the orders of autoregression, difference and moving average, respectively. Choosing the appropriate values of  $p$ ,  $d$  and  $q$  is a crucial step in building an ARIMA model.

To formulate the optimal daily replenishment and pricing strategy for each vegetable category in the forthcoming week—aiming to maximize the superstore's revenue—a preliminary analysis must be undertaken. This entails understanding the relationship between the total sales volume of each vegetable category and the cost-plus pricing, essential to discern the influence of sales pricing on overall sales volume.

Given that the evaluation is category-centric, it necessitates the computation of the average daily sales price for each vegetable category. This derived value will supplant the conventional cost-plus pricing. The computation for the same is articulated as follows:

$$S_{avg_n} = \frac{\sum_{men} \cdot S_{m, d} * SN_{m, d}}{\sum_{men} \cdot SN_{m, d}} \quad (1)$$

The formula for calculating the total daily sales for each vegetable category is as follows:

$$SN_{n, d} = \sum_{men} \cdot SN_{m, d} \quad (2)$$

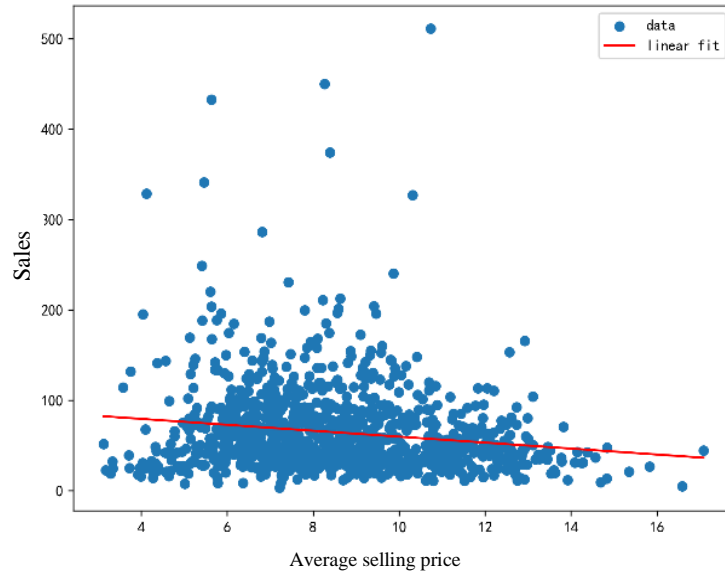
### 3. Solving and analysing the results of automatic pricing and replenishment decisions for vegetable items

#### 3.1. Solving and analysing the results of total daily replenishment and pricing strategies for each vegetable category in a supermarket in a city for the coming week (1-7 July 2023)

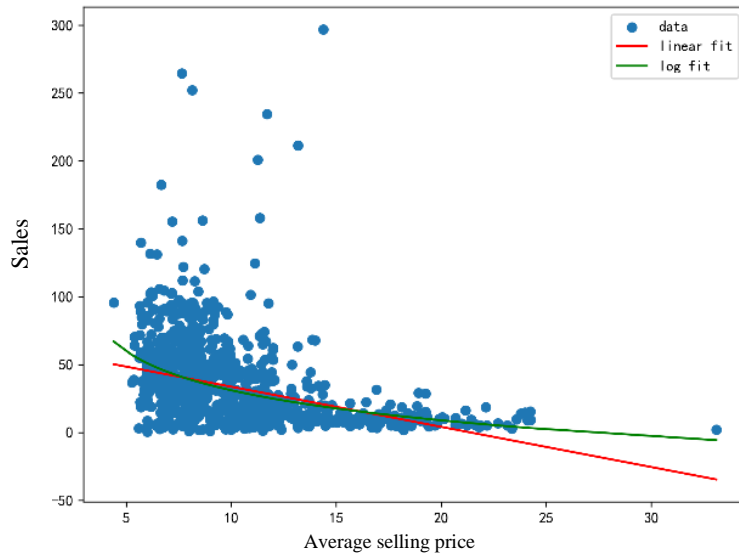
To capture the relationship between average sales pricing and total sales across various vegetable categories, we explore several functional forms, including linear, power, and logarithmic functions. The model that exhibits the best fit is subsequently chosen to represent the pricing-sales relationship [6]. Figures 1 through 6 illustrate the fit achieved by these functional forms for six distinct vegetable varieties. The parameters for the chosen models are detailed in Table 1:

**Table 1.** Plot of fitted model parameters

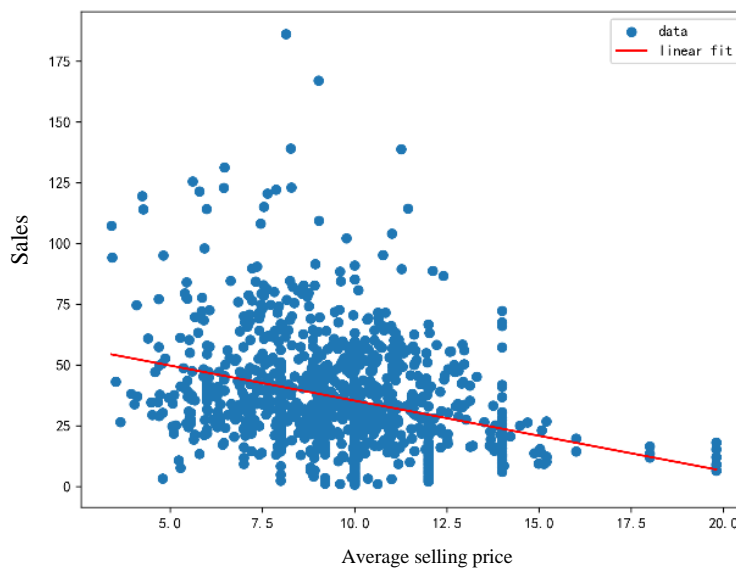
Classification name	Model name	model parameter
Aquatic rhizomes	logarithmic function	[-26.11035516, 2.51727164, 83.60511972]
philodendron	logarithmic function	[-36.00775817, 2.35435122, 209.756959]
cauliflower (Brassica oleracea var. botrytis)	linear function	[-2.89165146, 64.2021206]
eggplant	logarithmic function	[-8.16177241, 1.62367199, 36.4701182]
chilli	logarithmic function	[-17.07560033, 3.10816607, 105.16379865]
edible mushroom	linear function	[-3.28878025, 92.53732638]



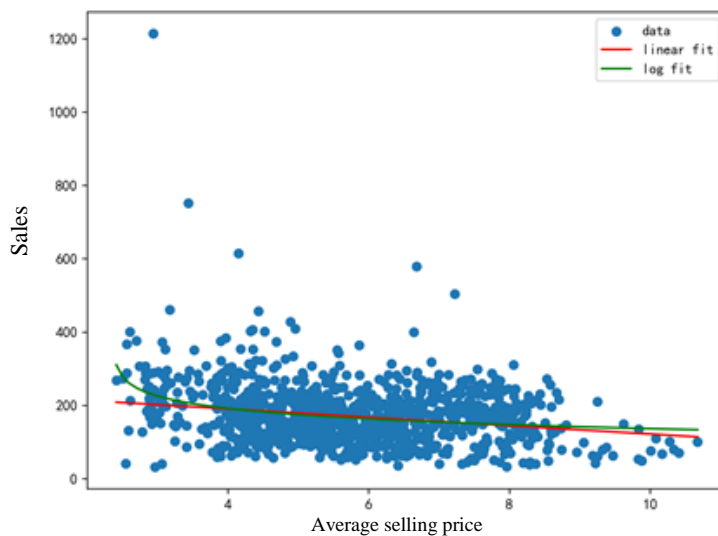
**Fig. 1** Price vs. Volume for Aquatic Roots & Tubers



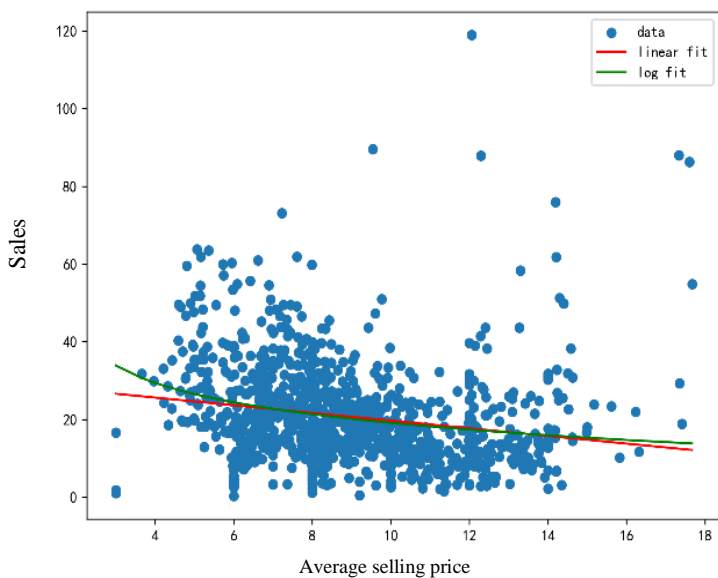
**Fig. 2** Price vs. Volume for Edible Mushrooms



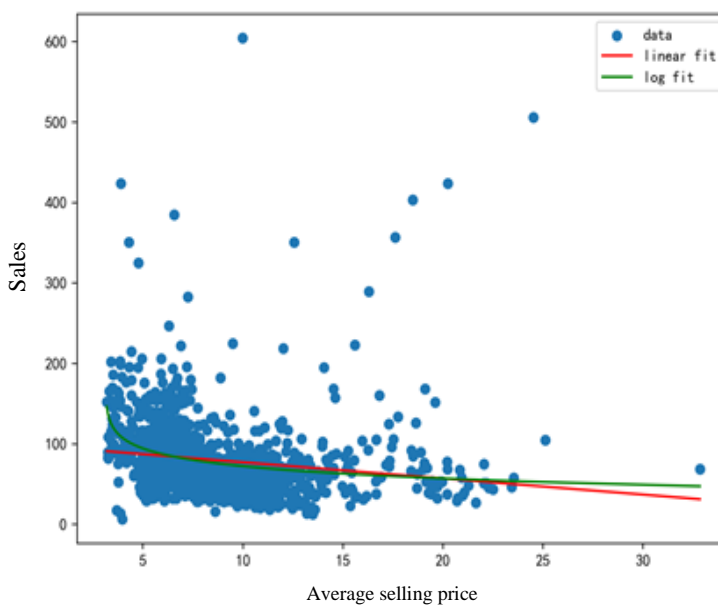
**Fig. 3** Price vs. Volume for Cauliflower



**Fig. 4** Price vs. Volume for Cauliflower & Leaves



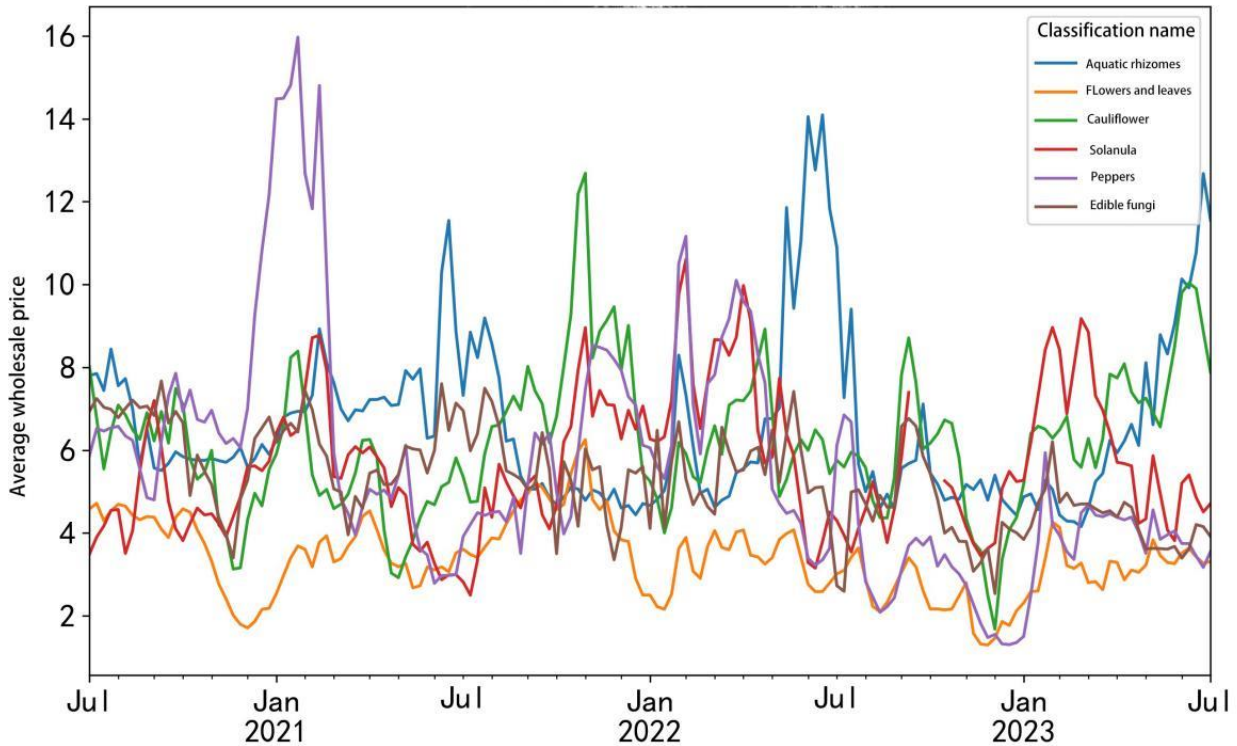
**Fig. 5** Price vs. Volume for Tomatoes



**Fig. 6** Price vs. Volume for Chillies

In order for the superstore to maximise the benefits of vegetable products, it is also necessary to forecast the average wholesale price (cost) of each vegetable category, and the formula for the average wholesale price of vegetable categories is calculated as follows:

$$W_{avg} - S_{n, d} = \frac{\sum_{men} (1 + A_n) * SN_{m, d} * WS_{m, d}}{\sum_{men} (1 + A_n) * SN_{m, d}} \quad (3)$$



**Figure 7.** Change in average wholesale prices of vegetable commodities

As illustrated in Figure 7, the trend in the average wholesale prices of various vegetable categories over the past three years, based on our investigation of the supermarket, is presented. Initially, the ARIMA model is employed to forecast the average wholesale price for each vegetable category over the forthcoming seven-day period. The model's parameters were optimized using the Akaike Information Criterion (AIC) [7], resulting in a parameter set of (2,2,2) to best predict the average wholesale price across the vegetable categories for the subsequent week. Upon acquiring the average wholesale price and the corresponding pricing-sales volume model for each category, a nonlinear programming model was formulated. This was undertaken to determine the ideal total daily replenishment and pricing strategy. The specified nonlinear programming model is detailed as follows [8]:

$$objective: \max (R_{n, d} = S_{n, d} * SN_{n, d} - Y_{n, d} * WS_{n, d}) \quad (4)$$

$$s.t.: Y_{n,d} > SN_{n,d} + A_n * Y_{n,d} \quad (5)$$

$$S_{n, d} > WS_{n,d} \quad (6)$$

$$S_{n, d}, Y_{n, d} > 0 \quad (7)$$

$$SN_{n, d} = func (S_{n, d}) \quad (8)$$

Where the objective function is to maximise the benefits of the superstore on the first  $d$  days in the vegetable category  $n$ , meaning total sales price - total wholesale price,  $Y_{n, d}$  is the replenishment volume of commodity  $n$  on day  $d$ ,  $S_{n, d}$  is the sales volume of commodity  $n$  on day  $d$ ; constraint function (4) is the replenishment volume is greater than the sales volume plus wastage, constraint function (5) is the sales price is greater than the cost of the wholesale price, constraint function (6) is the replenishment volume and the sales pricing is greater than 0. Constraint function (7) is the Pricing - sales volume function, through the function [9] can be determined by the sales pricing sales volume, substitution of data, solve to obtain the superstore the next 7 days of the vegetable category daily replenishment and pricing strategy as shown in Table 2 to Table 7

**Table 2.** Quantity and pricing of aquatic roots and tubers stocked over 7 days

dates	Incoming quantity (kg)	Pricing (yuan/kg)
1	10.95334	19.62878
2	10.94375	19.6342
3	10.82813	19.69978
4	10.80803	19.71121
5	10.79247	19.72005
6	10.78517	19.72421
7	10.77935	19.72752

**Table 3.** 7-day intake and pricing of foliage categories

dates	Incoming quantity (kg)	Pricing (yuan/kg)
1	159.7414	9.440909
2	159.9985	9.396931
3	160.0699	9.384782
4	159.9614	9.403269
5	160.1486	9.371391
6	160.0068	9.395525
7	160.2002	9.362627

**Table 4.** 7-day intake and pricing of cauliflower category

dates	Incoming quantity (kg)	Pricing (yuan/kg)
1	22.07802	15.75169
2	22.1002	15.74521
3	22.08577	15.74943
4	22.09992	15.74529
5	22.08922	15.74842
6	22.10044	15.74514
7	22.09229	15.74752

**Table 5.** Quantity and Pricing of Tomato Stocked in 7 Days

dates	Incoming quantity (kg)	Pricing (yuan/kg)
1	17.07341	14.00598
2	17.08081	13.99511
3	17.06499	14.0179
4	17.06977	14.01114
5	17.0663	14.01606
6	17.06447	14.01864
7	17.06293	14.02083

**Table 6.** 7-day stocking and pricing of chilli types

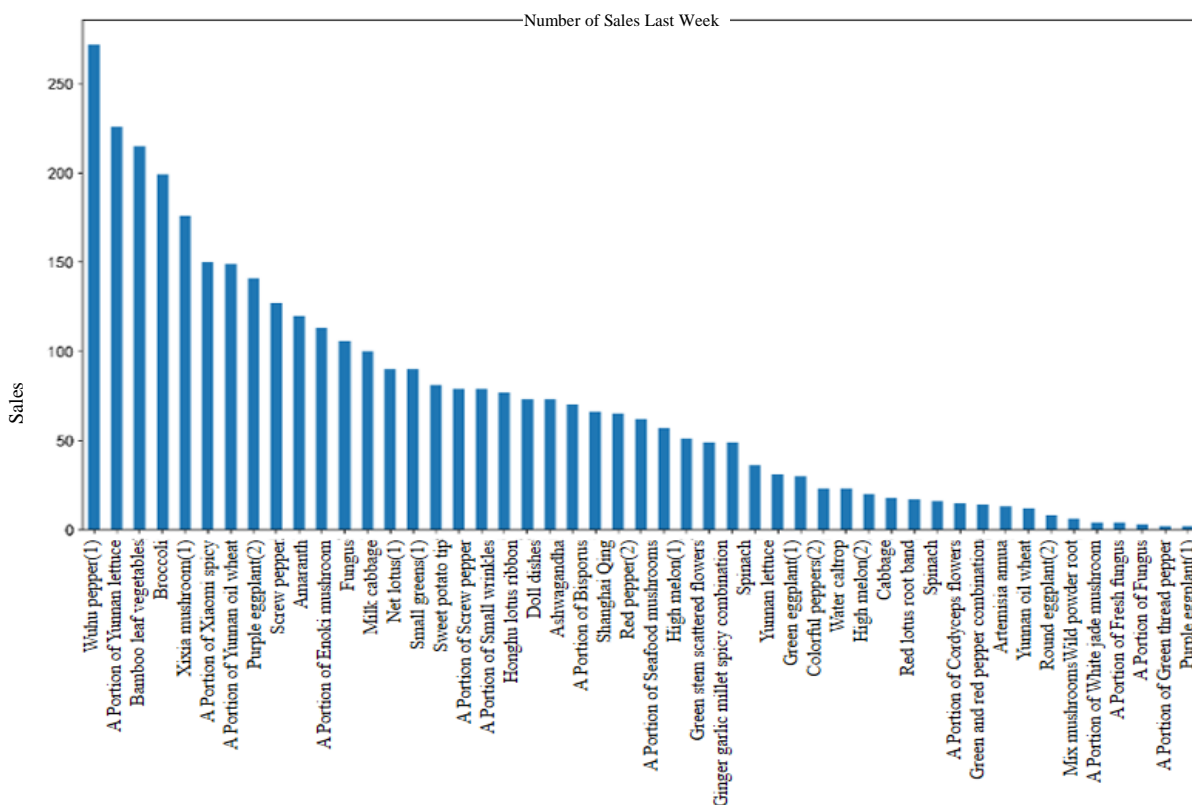
dates	Incoming quantity (kg)	Pricing (yuan/kg)
1	76.95735	11.01945
2	76.7876	11.09116
3	76.96742	11.01522
4	76.81373	11.08008
5	76.98788	11.00663
6	76.8376	11.06997
7	77.00889	10.99781

**Table 7.** Edible mushroom intake and pricing over 7 days

dates	Incoming quantity (kg)	Pricing (yuan/kg)
1	54.24866	13.20099
2	52.53689	13.67229
3	51.98256	13.82492
4	53.4264	13.42738
5	51.83244	13.86625
6	53.38939	13.43758
7	52.05502	13.80497

Through a comparative assessment across vegetable categories, Figure 8 illustrates the proposed daily replenishment and pricing strategy for each category over the upcoming week (1-7 July 2023) for the superstore.

**3.2. Solving and analysing the results of total replenishment and pricing strategy for vegetables of a supermarket in a city for the total replenishment of a single item on 1 July 2023**



**Figure 8.** Vegetable sales volume of individual items in supermarkets

Figure 8 reflects how the number of supermarket vegetable items sold in the last week. Firstly, the wholesale price of each individual product is predicted for the coming day, where the historical mean prediction method [10] is used to get the future wholesale price  $WS_{m, d}$ . The nonlinear programming model is defined as follows:

$$\text{objective: } \max (R_d = \text{selected} * SN_d - \text{selected} * Y_d * WS_d) \quad (9)$$

$$\text{s.t.: } \text{selected} * Y_{m, d} > \text{selected} * SN_{m, d} + \text{selected} * A_n * Y_{m, d}, \text{ where } \text{men} \quad (10)$$

$$S_{m, d} > WS_{m, d} \quad (11)$$

$$\text{selected} * Y_{m, d} > \text{selected} * 2.5 \quad (12)$$

$$27 \leq \text{sum}(\text{selected}) \leq 33 \quad (13)$$

$$\text{selected} * Y_{m, d} = Y_{m, d} \quad (14)$$

$$\text{selected} * S_{m, d} = S_{m, d} \quad (15)$$

$$SN_{m, d} = \text{func} (S_{m, d}), \text{ where } \text{men} \quad (16)$$

$$\text{num} (n) = 6 \quad (17)$$

$$\text{selected} = [\text{random.randint}(0, 1) \text{ for\_in\_range}(49)] \quad (18)$$

Where the objective function is to maximise the sales profit of the superstore, as the sum of the sales profit of each vegetable item, here we use a selection list *selected* to control the quantity of vegetable items purchased, which is a list with length 49 and value 0 or 1, 0 means no purchasing, and 1 means purchasing (constraint function (16)).  $Y_d$  is a list of replenishment quantities, with a length of 49 and a value of the quantity of vegetable items purchased;  $S_d$  is a list of selling prices, with a length of 49 and a value of the unit price of the vegetable items sold.

The constraints applied in our model are meticulously defined to cater to various operational and market-driven factors. Specifically:

Constraint (8) ensures that the replenishment quantity surpasses the sum of sales and wastage quantities.

Constraint (9) stipulates that the selling price must exceed the wholesale price.

Constraint (10) is imposed to honor the minimum display quantity requirements.

Constraint (11) addresses the permissible number of individual product selections.

Constraint (12) guarantees consistency between the replenishment quantity list and the selection list in terms of zero-valued positions.

Constraint (13) ensures the sales price list and the selection list maintain congruency with respect to zero-valued positions.

Constraint (14) draws upon the vegetable category pricing-sales volume function acquired from Problem 2, translating individual product categories and pricing to projected sales volumes for each product.

Constraint (15) acknowledges market demand for varied vegetable categories, stipulating that the replenished quantity of individual vegetables belongs to the respective category, with a total of six distinct categories in our study.

The restocking and pricing strategies for vegetable items over the forthcoming day are detailed in Table 8:

**Table 8.** Quantity of Vegetable Commodities Stocked and Selling Prices in 1 Day

Item Name	quantity of imported goods	sales price
amaranth greens (genus Amaranthus)	0	0
Chinese flowering cabbage	0	0
broccoli	4.760075759	14
Spinach (portions)	4.223043276	5.510204082
water caltrop or water chestnut (genus Trapa)	29.86845451	14
Cordyceps flowers (portions)	9.693139396	3.6125
king cobra or ghost chili (Naga jolokia)	16.07547691	11.29133858
Screw peppers (portions)	0	0
Crabmeat Mushroom & White Jade Mushroom Duo (Box)	0	0
broccoli	19.3910901	12.4080402
Xixia Flower Mushroom (1)	22.80611168	24
Golden Needle Mushroom (box)	19.10032914	1.879646018
long term eggplant	0	0
Green and Red Hanging Pepper Combo (Servings)	29.03063805	5.493333333
Green Thread Pepper (portions)	22.70106748	4.3
Green aubergine (1)	4.141759794	6
High melon (1)	0	0
Yunnan oilseed rape (Brassica napus)	25.66818909	7.2
Yunnan lettuce	14.28796708	9.2
High melon (2)	0	0
White Jade Mushroom (bag)	0	0
Fresh fungus (portions)	0	0
Mullein (portion)	0	0
Purple aubergine (1)	26.75966417	9
Wild Lotus Root	23.37169162	26
Item Name	quantity of imported goods	sales price
Colourful Peppers (2)	28.3699084	18.98666667
Shanghai Qingdao	3.542098523	8
Yunnan oilseed rape (portion)	20.77066984	4.159060403
Yunnan lettuce (portions)	14.16476615	4.461504425
Net root (1)	0	0
Agaricus bisporus (box)	0	0
Round Eggplant (2)	28.58223041	6.133333333
Chrysanthemum coronarium	0	0
cabbage (round cabbage most commonly found in Chinese medicine)	8.539312758	4.792
Ginger, Garlic & Millet Pepper Combo Pack (small portion)	20.98586099	4.72244898
baby Chinese cabbage (mini-sized variety)	5.602326454	6.54109589
Small wrinkled skin (portions)	0	0
Millet peppers (portions)	14.09529155	5.769333333
Baby bok choy (1)	9.498576104	5.2
snow fungus (Tremella fuciformis)	15.98918396	5.332075472
Zijiang Qingdian Scattered Flowers	0	0
Honghu Lotus Roots	0	0
Seafood Mushroom (Packet)	24.68027962	2.748387097
Brussels sprouts (Brassica oleracea var. botrytis)	12.49620544	3.773953488
Purple Eggplant (2)	0	0
Red pepper (2)	5.627657919	18.89230769
Red Lotus Root Scallops	0	0
sweet potato tip	0	0
Wuhu green pepper (1)	7.728847498	5.2

## 4. Conclusion

This study embarked on the challenge of optimizing the automatic pricing and replenishment decision-making processes for vegetable categories, with the overarching aim of maximizing superstore revenues. Central to this endeavor was the synthesis of the Pearson correlation coefficient method, the Autoregressive Integrated Moving Average (ARIMA), and a suite of regression analysis techniques. The key achievements from our investigation can be summarized as:

1. Regression Insights: Through logistic regression analysis, we were able to formulate regression equations correlating the total sales volume and cost-plus pricing for individual vegetable categories.

2. Forecasting Dynamics: The ARIMA model, refined using the Akaike Information Criterion to an optimal configuration of (2,2,2), effectively predicted the average wholesale prices for vegetable categories over a prospective 7-day timeframe.

3. Modeling & Replenishment: Upon deriving the average wholesale prices and establishing a pricing-sales volume model, we developed a nonlinear planning model. This facilitated the calculation of daily replenishment and pricing strategies for vegetable categories in the week of 1-7 July 2023. Detailed outputs of this endeavor are cataloged in Tables 2 to 7, offering a pathway towards maximum revenue realization for superstores.

4. Constraints & Optimization: The optimization model incorporated three pivotal constraints: a controlled tally of saleable items (27-33), a minimum display quantity threshold of 2.5kg, and limitations on selling space. Using historical averages as a forecasting anchor, future wholesale prices were estimated. The union of these constraints with the nonlinear programming model ensured the effectiveness of the replenishment strategies, as illustrated in Table 8.

In reflection, the empirical results underscore the validity and potential of the proposed model in real-world scenarios, promising a more efficient and profit-maximizing approach to vegetable commodity management in superstores.

## References

- [1] Yu Chao. A comprehensive study of fresh produce inventory based on freshness[D]. Tsinghua University,2010.
- [2] Zhang Fan. Research on Demand Forecasting and Inventory Management Mode of Retail Supply Chain Based on Commodity Categories [D]. University of Electronic Science and Technology, 2011.
- [3] Yang Tianshan,Yuan Gonglin.Research on dynamic pricing and replenishment strategy of perishables considering green distribution technology investment[J].China Management Science:1-13[2023-09-20].
- [4] Xiao Yixin,Liu Sanyang. Locust optimisation algorithm for fusion pattern search and its application[J/OL]. Journal of Xi'an Electronic Science and Technology University:1-20[2023-09-21].
- [5] Zhang Qi-Ping,Ji Cong-Hua,Lu Jin-Qi et al. Study on the effect of applying ARIMA product seasonal model on the prediction of epidemic trend of carbapenem-resistant *Klebsiella pneumoniae*[J]. China Hospital Statistics,2023,30(03):173-178.
- [6] Yang Haoxu. Research on the correlation analysis of vegetable price and sales volume--Taking oil wheat vegetable as an example[J]
- [7] Hu Hu-Jiang,Li Li-Ping,Jin Hao et al. First-arrival pickup method of microvibration signals from sudden water surges based on high-order statistical skewness and red pool information criterion[J]
- [8] Zhang Xiaoping,Cao Qingsong,Gao Xiaolin et al. Nonlinear planning of electric vehicle charging station layout based on simplified model[J]. Automotive Practical Technology,2023,48(11):15-21.
- [9] Liu Jian,Zhang Zhihong. An improved Drosophila visual evolutionary neural network for solving constraint function optimisation[J]. Microelectronics and Computers,2022,39(04):41-48.
- [10] Chen Baolin,Zhang Bincai,Wu Jing et al. Historical mean method for MODIS image pixel cloud compensation--A case study in Gansu Province[J]. Remote Sensing of Land Resources,2021,33(02):85-92.