

# 基于超市商品补货策略的分析

## ——以2023年全国数学建模竞赛C题为例

杨若涵, 张莹, 周文可

北京工商大学数学与统计学院, 北京

收稿日期: 2023年12月25日; 录用日期: 2024年1月19日; 发布日期: 2024年1月25日

### 摘要

由于蔬菜类商品保质期较短, 易变质会影响其销售价值, 商超通常会选择每天都补货。蔬菜品种众多, 商家在进货前无法准确得知具体商品及其进货价格。本文将基于已知数据和蔬菜的时令性即商超销售空间的限制, 从销量、进价、售价与商品相互间的关系进行市场需求分析, 通过预测未来商品的销量, 制定合理的补货和定价决策, 确保商超利益最大化。通过系统聚类将辣椒类、花叶类、水生根茎类、食用菌、花菜类、茄类六个品类分成两组, 根据每月销量绘制折线图来判断并描述其分布规律。其次分析不同品类的相互关系, 先绘制散点图初步观察, 再通过计算皮尔逊相关系数来判断不同品类有无相关性, 可知六个品类间花菜类和花叶类相关性最强。其次为预估未来几天的各品类的日补货总量, 借助前一个月数据进行预估, 将品类销量按日汇总后, 先检验时间序列的稳定性, 再分别求得其自相关系数(ACF)和偏自相关系数(PACF), 通过对自相关图和偏自相关图的分析, 得到最佳的阶数 $p$ 、 $q$ , 检验模型, 再利用ARIMA模型进行时间序列预测分析, 得出未来七天的预估销售量, 最后根据补货量与销售量的关系得出日补货总量。

### 关键词

系统聚类, 皮尔逊相关系数, ARIMA模型, 时间序列

# Analysis on Supermarket Product Replenishment Strategy

## —A Case Study of Problem C in the 2023 National Mathematical Modeling Competition

Ruohan Yang, Ying Zhang, Wenke Zhou

School of Mathematics and Statistics, Beijing Technology and Business University, Beijing

Received: Dec. 25<sup>th</sup>, 2023; accepted: Jan. 19<sup>th</sup>, 2024; published: Jan. 25<sup>th</sup>, 2024

## Abstract

Due to the short shelf life and perishable nature of vegetable products, their sales value can be affected, so supermarkets often choose to restock daily. With a wide variety of vegetable varieties, merchants are unable to accurately determine the specific products and their purchase prices before stocking. This article will analyze market demand based on known data and the seasonality of vegetables, as well as the limitations of supermarket sales space. By predicting future product sales, reasonable restocking and pricing decisions can be made to maximize the interests of supermarkets. Using systematic clustering, six categories of vegetables, namely chili peppers, leafy vegetables, aquatic roots and tubers, edible mushrooms, cauliflower, and tomatoes, are divided into two groups. Line charts of monthly sales volume are plotted to determine and describe their distribution patterns. Furthermore, the relationships between different categories are analyzed. Scatter plots are first plotted for initial observation, and then the Pearson correlation coefficient is calculated to determine the correlation between different categories. It is found that cauliflower and leafy vegetables have the strongest correlation among the six categories. Next, the daily total restocking volume for each category in the coming days is estimated using data from the previous month. After aggregating the sales volume by category on a daily basis, the stability of the time series is tested, and the autocorrelation coefficient (ACF) and partial autocorrelation coefficient (PACF) are calculated. Through the analysis of the autocorrelation plot and partial autocorrelation plot, the optimal order of  $p$  and  $q$  is determined. The model is then tested, and the ARIMA model is used for time series forecasting analysis to obtain the estimated sales volume for the next seven days. Finally, the daily total restocking volume is determined based on the relationship between restocking volume and sales volume.

## Keywords

Systematic Clustering, Pearson Correlation Coefficient, ARIMA Model, Time Series

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在生鲜商超中,一般蔬菜类商品的保鲜期都比较短,且品相随销售时间的增加而变差,大部分品种如当日未售出,隔日就无法再售。因此,商超通常会根据各商品的历史销售和需求情况每天进行补货。由于商超销售的蔬菜品种众多、产地不尽相同,而蔬菜的进货交易时间通常在凌晨 3:00~4:00,为此商家须在不确切知道具体单品和进货价格的情况下,做出当日各蔬菜品类的补货决策。蔬菜的定价一般采用“成本加成定价”方法,商超对运损和品相变差的商品通常进行打折销售。可靠的市场需求分析,对补货决策和定价决策尤为重要。

关于补货的研究,已经有一些学者尝试利用销售数据和成本信息来建立模型,以预测蔬菜商品的补货量和定价策略。其中,一些研究者尝试使用随机森林模型来进行预测[1]。通过分析不同时间段内各类别和单品蔬菜商品的销售情况,以及销售量与成本加成定价之间的关系,他们建立了一个随机森林模型,用于预测未来一周的蔬菜商品的补货量和定价策略。这些研究成果为我们提供了宝贵的参考,为我们的研究提供了一定的借鉴意义。陈军[2]等以周期利润最大化为目标,采用变质库存理论,构建需求依赖于

价格和库存水平的零售商双渠道定价与库存补货联合决策模型，分析最优解存在性的性质。基于两种求解思路，设计启发式算法求解出最优保鲜期、销售周期、定价和利润，最后对单位成本、需求率、保鲜期和变质率等参数进行敏感性分析。乔雪[3]以最大化利润为目标，提出了两种鲁棒补货策略：基于库存概率分布的策略和基于库存均值的策略。前者首先给出了估计真实库存概率分布的递推算法，然后提出基于估计库存概率分布的补货策略；后者使用均值近似模型中的随机变量，证明了最优补货策略的零库存性质和目标函数的单峰性，提出了计算最优补货量的方法。

对于本文数据特征，采用时间序列分析和回归方程两种方法来拟合，从而得出关系式进行分析。基于数据总量，以月份为单位，便于绘制和观察折线图，同时进行系统聚类、方差分析来观察整体的分布状况，确定销量最高的时间，进而估计市场需求。关于求相关关系，可先绘制散点图初步观察，对于不同品类，利用皮尔逊相关系数来进行检验，再借助 SPSS 对散点图进行拟合，从而得到不同品类两两之间的具体函数关系。最后通过上述建模来判断未来几天的定价和补货量。

## 2. 模型假设

为了构建更为精确的数学模型，本文根据实际情况做出合理的假设：

- 假设一：商超的蔬菜供应链和进货渠道是稳定可靠的，没有突发的供应问题或运输延迟，亦不受自然灾害等不可抗拒因素影响。
- 假设二：商超的补货决策是基于准确的历史销售和需求数据进行的，没有因为数据收集、处理或分析错误而导致的补货决策失误。
- 假设三：商超的定价决策是基于准确的成本 and 市场需求分析进行的，没有因为定价策略或数据不准确而导致的定价失误。
- 假设四：价格遵循市场规律，即供不应求时价格上升，供大于求时价格下降。

## 3. 符号申明

符号	符号申明
$\rho_{XY}$	样本皮尔森相关系数
$Cov(x, y)$	样本协方差
$S_x$	样本标准差
$\mu$	样本均值

## 4. 方法说明

### 4.1. Pearson 相关性分析

在统计学中，皮尔逊相关系数又称皮尔逊积矩相关系数，是用于度量连个变量  $x$  和  $y$  之间的线性相关性，其值介于-1 到 1 之间。设两个总体分别为  $X$ 、 $Y$ ，其样本分别为  $(x_1, x_2, \dots, x_n)$  和  $(y_1, y_2, \dots, y_n)$ ，则皮尔逊相关系数为：

$$\rho_{XY} = \frac{Cov(x, y)}{S_x S_y}$$

$$\text{其中, } Cov(x, y) = \frac{\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}{n-1}, S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu_x)^2}{n-1}}, S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \mu_y)^2}{n-1}}.$$

皮尔逊相关系数的使用范围:

- 1) 两个变量之间是线性关系, 且是连续数据。
- 2) 两个变量的总体是正态分布, 或接近正太的单峰分布。
- 3) 两个变量的观测值是成对的, 并且每对观测值之间相互独立。

通常可以通过以下取值范围判断变量的相关强度(表 1):

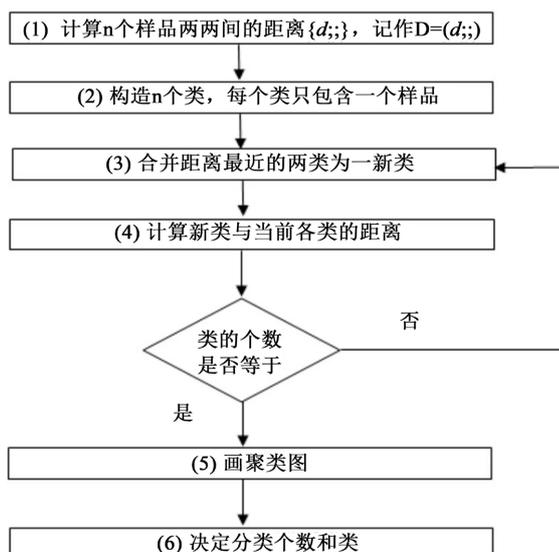
**Table 1.** Pearson correlation coefficient intensity table

**表 1.** 皮尔逊相关系数强度表

范围	强度
0.8~1.0	极强相关
0.6~0.8	强相关
0.4~0.6	中等程度相关
0.2~0.4	弱相关
0.0~0.2	极弱相关或无相关

## 4.2. 系统聚类

系统聚类法在聚类分析诸方法中用的最多。基本思想是按照距离远近, 将距离相近的变量先聚成类, 距离较远的变量后聚成类, 依次进行, 直到每个变量都归入合适的类中。其包含以下步骤:



系统聚类分析的前提是计算和确定类间距离, 因此类间距离的计算方法不同, 系统聚类法也不同。常用的类间距离定义有 7 种, 对应 7 种系统聚类法, 分别为: 最短距离法、最长距离法、中间距离法、重心法、组间连接法、组内连接法、离差平方和法(也称 Ward 法)。具体内容和公式不在本文中叙述, 详情请见周志华机器学习[4]。

## 4.3. 回归分析

对于此次数据, 宜采用一元线性回归方程求解。在统计学中, 线性回归是利用称为线性回归方程的

最小平方函数对一个或多个自变量和因变量之间关系进行建模的一种回归分析。其模型为：

$$\begin{cases} Y = \hat{a}X + \hat{b} \\ \varepsilon \sim N(0, \sigma^2) \end{cases}$$

其中， $\hat{a}$ 、 $\hat{b}$  代表一元线性回归方程参数， $\varepsilon$  为随机误差项。

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

其中， $\bar{x}$  和  $\bar{y}$  都为样本均值。

建立回归模型的步骤如下：

- 1) 选取一元线性回归模型的变量。
- 2) 绘制计算表和拟合散点图。
- 3) 计算变量间的回归系数及其相关的显著性。

模型的检验：

- 1) 回归标准差检验。
- 2) 拟合优度检验。
- 3) 回归系数的显著性检验。

#### 4.4. ARIMA 时间序列分析模型

在生产和科学研究中，对某一个或一组变量  $t$  进行观察测量，将在一系列时刻  $t_1, t_2, \dots, t_n$  ( $t$  为自变量) 按照时间次序排列，并用于解释变量和相互关系的数学表达式。  $t_1 < t_2 < \dots < t_n$  所得到的离散数字组成序列集合  $x(t_1), x(t_2), \dots, x(t_n)$ ，称之为时间序列。时间序列包括多种模型，其中 ARIMA 全称为自回归积分滑动平均模型，是由 Box 和 Jenkins 于 70 年代初提出一著名时间序列预测方法，所以又称为 Box-Jenkins 模型[5]。ARIMA( $p, d, q$ )称为差分自回归移动平均模型，AR 是自回归， $p$  为自回归项；MA 为移动平均， $q$  为移动平均项数， $d$  为时间序列成为平稳时所做的差分次数。其公式为：

$$y_t = \alpha + \sum_{i=1}^p \rho_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

其中  $\alpha$  是常数项， $\rho_i$  是自相关系数， $\theta_i$  是误差项系数， $\varepsilon_t$  是白噪声。自相关函数 ACF 用来确定  $q$  值， $ACF(k) = p_k = \frac{Cov(y_t, y_{t-k})}{Var(y_t)}$ ，反应了同一序列在不同时序的取值之间的相关性，范围为[-1, 1]；偏自相

关函数 PACF 确定  $p$  值，它剔除了中间  $k-1$  个随机变量干扰后的相关程度。

基本程序如下：

- 1) 根据时间序列的散点图、自相关函数和偏自相关函数图以 ADF 单位根检验其方差、趋势及其季节性变化规律，对序列的平稳性进行识别。一般来讲，经济运行的时间序列都不是平稳序列。
- 2) 对非平稳序列进行平稳化处理。如果数据序列是非平稳的，并存在一定的增长或下降趋势，则需要对数据进行差分处理，如果数据存在异方差，则需对数据进行技术处理，直到处理后的数据的自相关函数值和偏相关函数值无显著地异于零。
- 3) 根据时间序列模型的识别规则，建立相应的模型。若平稳序列的偏相关函数是截尾的，而自相关函数是拖尾的，可断定序列适合 AR 模型；若平稳序列的偏相关函数是拖尾的，而自相关函数是截尾的，

则可断定序列适合 MA 模型;若平稳序列的偏相关函数和自相关函数均是拖尾的,则序列适合 ARMA 模型。

- 4) 进行参数估计, 检验是否具有统计意义。
- 5) 进行假设检验, 诊断残差序列是否为白噪声。
- 6) 利用已通过检验的模型进行预测分析。

## 5. 模型建立与求解

### 5.1. 品类关系分析

#### 5.1.1. 分布规律

基于已知数据, 首先将三年内共 1094 天各个菜品类进行汇总, 考察各品类之间的销售量的关系。如图 1, 做出柱状图之后, 花叶类的销售总量最高, 达到了几乎二十万千克, 所占比例最高。相比于花叶类, 其他五个菜品类就相对较少, 其中茄类的总销量最小, 只在两万五千千克左右。由此可以得出, 市民更偏向于花叶类的单品, 可暂时以其作为主要补货对象。

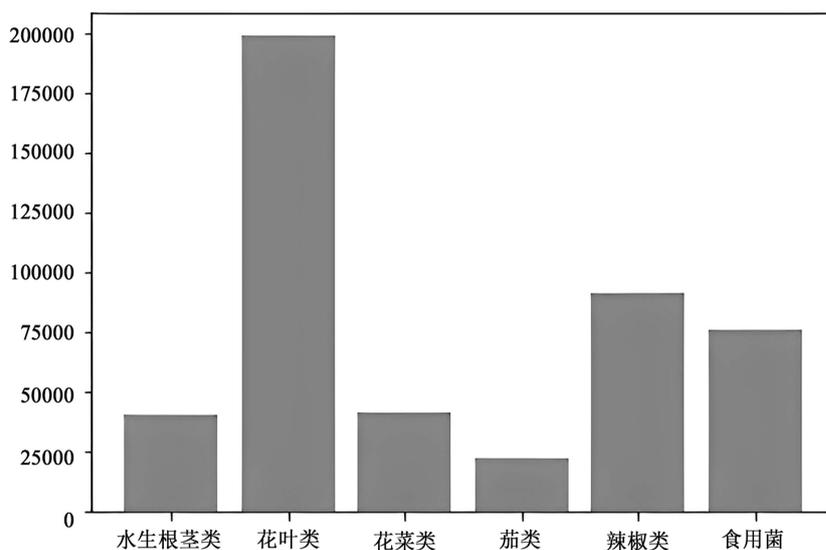


Figure 1. Chart of total sales by category

图 1. 各品类总销量图

由于数据过于庞大, 且考虑到销售量是一个随时间变化而变化的量, 在保证使得分布规律尽可能准确的情况下, 把各菜品类按照月份汇总, 并绘制多重折线图。折线图的横坐标为月份, 纵坐标为销量情况, 图中有六条线, 分别对应了六个菜品类。由图 2 所呈现的分布规律可知, 各品类销售量都呈现出了显著的季节性。花叶类在三年的销售量最多, 该类随月份和季节变化的波动也较大。辣椒类和食用菌在这些月份的总销量占比处于中间, 表明这两类随月份、季节变化的波动呈中性。水生根茎类、花菜类和茄类的总销量较低, 波动幅度最小。

为了增强准确性, 不妨对六个菜品类进行系统聚类。计算出样本之间的欧式距离(即  $d = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ )后进行系统聚类, 将六个品种归为了三类, 如下图所示。由图 3 可知, 总共分为了两个大类, 其中一个大类又分为了两个小类, 与上述分析相同, 以此可以确定按季节分为三大类: 第一大类有水生根茎类、花菜类和茄类, 称之为弱季节性类; 第二大类有辣椒类和食用菌, 称之为中季节性类; 第三大类只有花叶类, 称之为强季节性类。

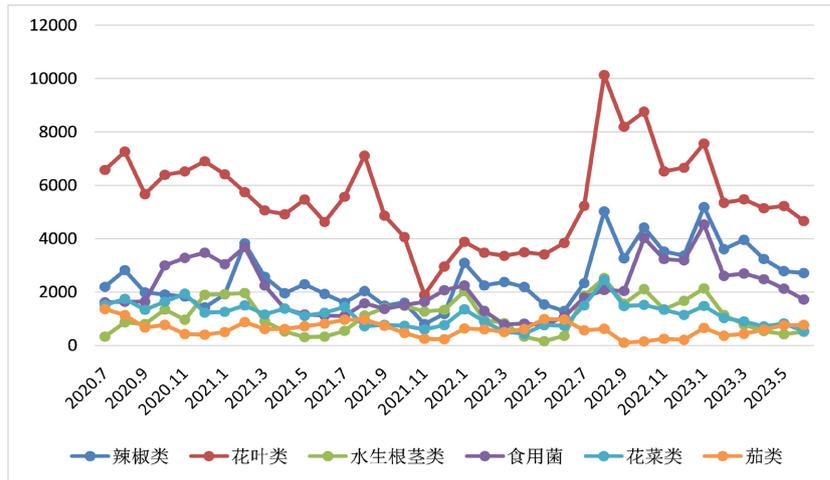


Figure 2. Line chart of total sales volume of different categories for each month from July 1, 2020 to June 30, 2023

图 2. 不同品类在 2020 年 7 月 1 日至 2023 年 6 月 30 日每个月份的总销量折线图

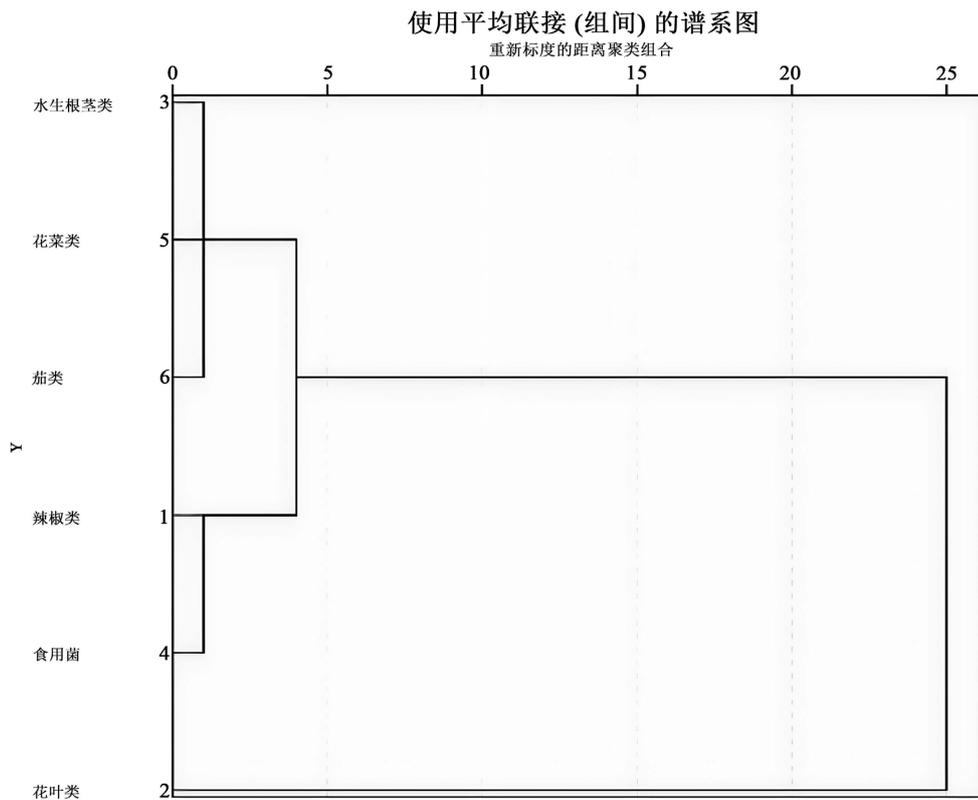


Figure 3. Cluster pedigree diagram of each category system

图 3. 各品类系统聚类谱系图

### 5.1.2. 相关性分析

把各品类按月汇总销量后，每个品类之间依次绘制二维散点图，得到如下的散点图矩阵。从矩阵可以很明显地看出，如花叶类和花菜类、水生根茎类和食用菌的散点较为密集，而如水生根茎类和茄类的散点较为稀疏。随后计算各品类之间的 Pearson 相关系数，得到如下表的相关性矩阵(图 4，表 2)。

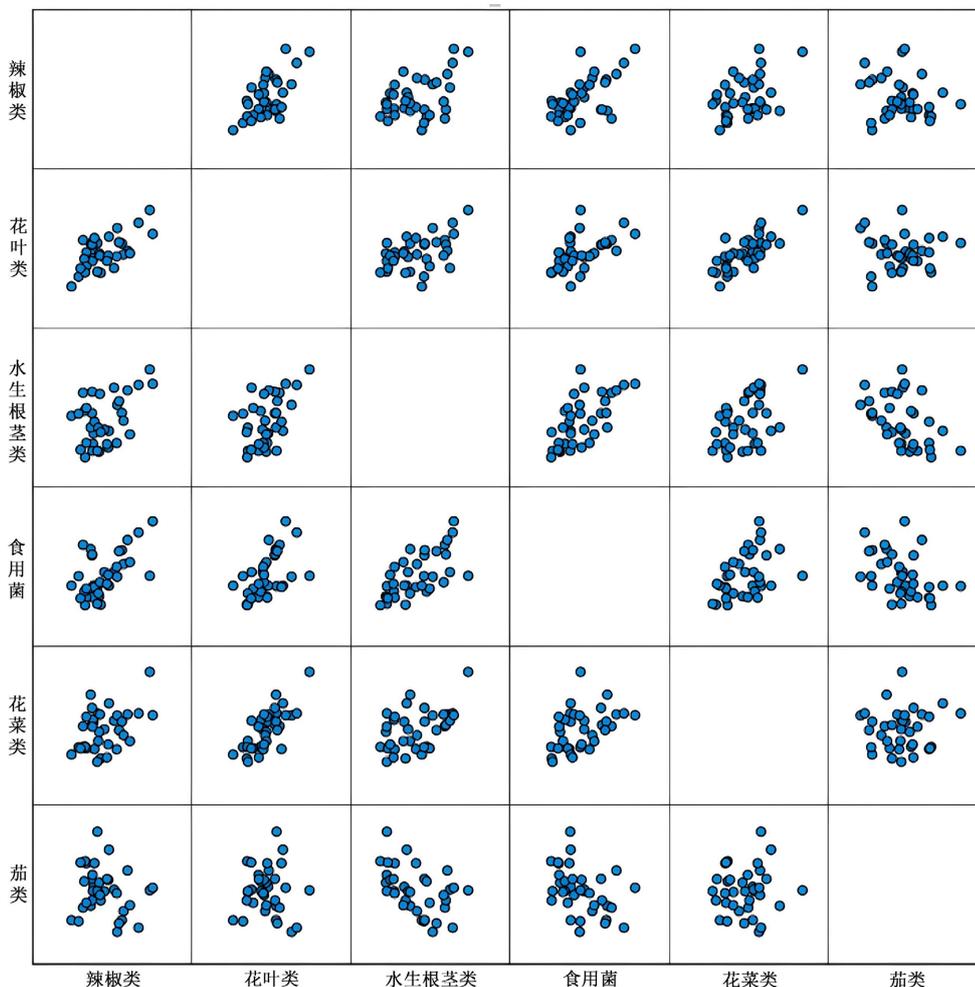


Figure 4. Scatter diagram matrix between categories  
 图 4. 各品类间散点图矩阵

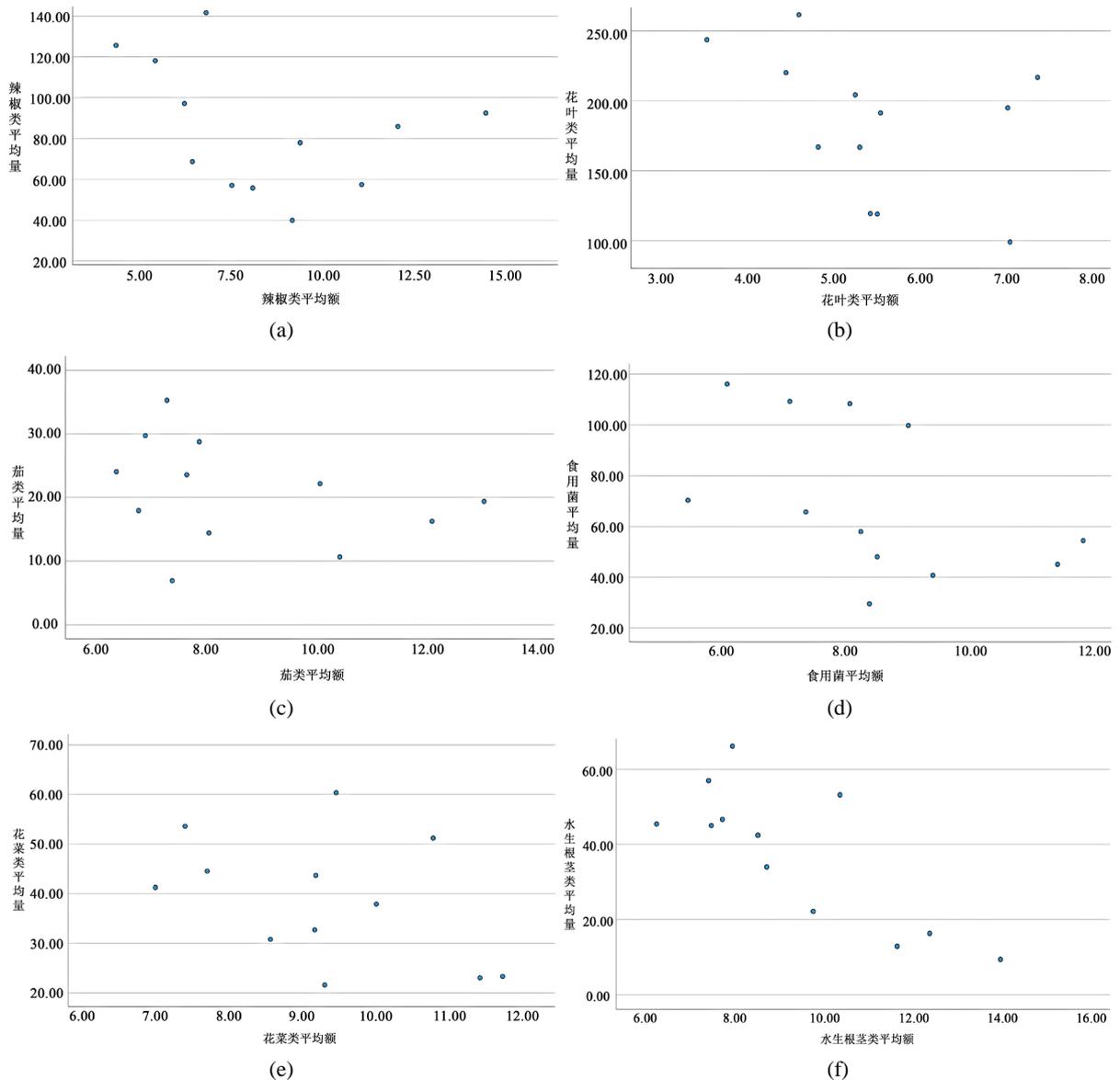
Table 2. Correlation matrix between different categories  
 表 2. 不同品类间相关性矩阵

	辣椒类	花叶类	水生根茎类	食用菌	花菜类	茄类
辣椒类	1.000	0.625	0.459	0.576	0.425	-0.191
花叶类	0.625	1.000	0.486	0.548	0.748	-0.038
水生根茎类	0.459	0.486	1.000	0.653	0.472	-0.466
食用菌	0.576	0.548	0.653	1.000	0.429	-0.409
花菜类	0.425	0.748	0.472	0.429	1.000	0.058
茄类	-0.191	-0.038	-0.466	-0.409	0.058	1.000

由表可知散点图符合线性相关性，其中花叶类和花菜类有正相关性最大，说明花叶类的需求带动了花菜类的需求；水生根茎类和茄类有负相关性最大，说明当其中一类的库存不够时，会使市民更偏向于购买另一类菜品。

### 5.2. 基于回归模型的销量与定价分析

考虑到时令性蔬菜受季节影响较大，在蔬菜的不同品类按月汇总销量的基础上，继续将销量按季度汇总，以此来合并数据，简化求解过程。由于每一个品类有诸多不同单品，且每个单品在每天的销售单价并不一样，所以由 SPSS 进行加权求平均处理，并绘制散点图，以便分析相关性。其中横坐标为平均销售额，纵坐标为平均销售量(图 5)。



**Figure 5.** Scatter chart of quarterly average sales volume and average sales volume of each category  
**图 5.** 各品类季度平均销售额与平均销售量散点图

计算其相关系数，得下表 3:

由此得出结论：销售量和销售额之间呈现弱相关性，相关程度较低。

分别建立回归方程，得到具体表达式：

Table 3. Correlation coefficient

表 3. 相关系数

花叶类	辣椒类	花菜类	水生根茎类	食用菌	茄类
0.168	0.139	0.167	0.245	0.239	0.100

$$\text{辣椒类: } y = -3.94x + 119$$

$$\text{花叶类: } y = -18.23x + 285$$

$$\text{茄类: } y = -1.17x + 31.03$$

$$\text{食用菌: } y = -7.82x + 137$$

$$\text{花菜类: } y = -3.43x + 70.14$$

$$\text{水生根茎类: } y = -6.27x + 96.88$$

### 5.3. 基于时间预测模型的未来补货决策制定

#### 5.3.1. 数据预处理

由于销量与定价的相关性较低,若想预测未来一周内的补货量和定价,便可以只用 2023 年 6 月份每天的数据来估算。为此,先绘制各菜品类在该月每天的销售总量折线图,观察其分布情况。由图 6 可以看出,各品类的峰值基本都在第 17 天左右,月末时波动基本比较平稳。

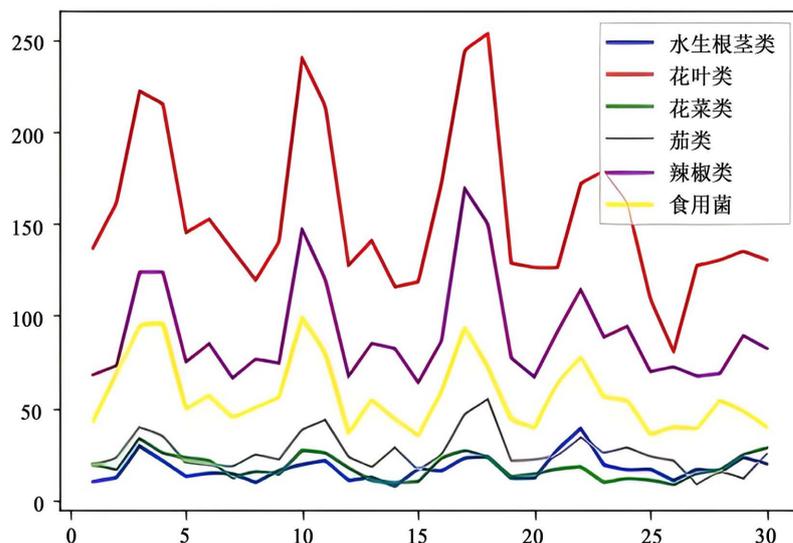


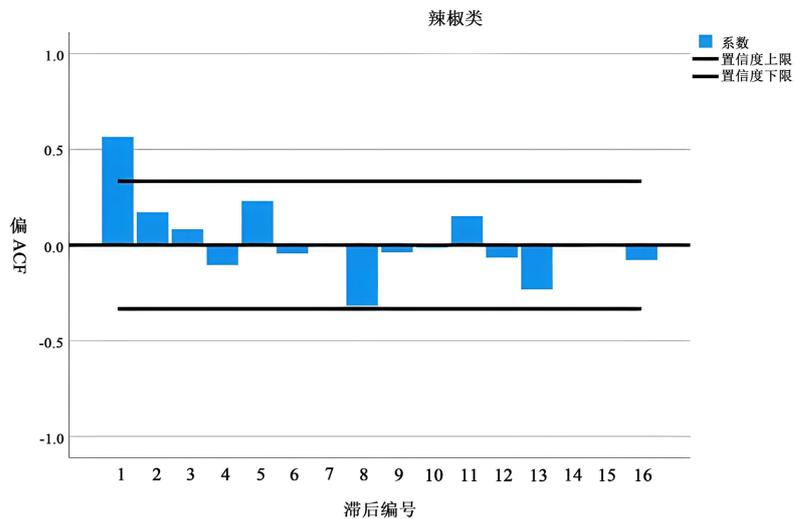
Figure 6. Sales distribution of each category in June 2023

图 6. 各品类在 2023 年 6 月销量分布

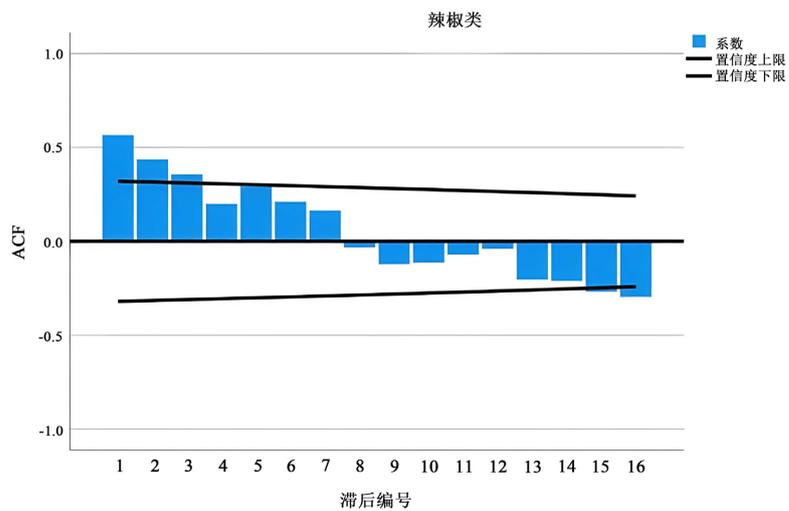
根据补货量 = 销售量 / (1 - 损耗率),再借助 ARIMA 模型,便可计算出未来七天的预测补货量,为此先计算每个品类的自相关系数和偏自相关系数。具体如下图 7,这里只展示两种菜品类,其余四种做法一样。

#### 5.3.2. ARIMA 模型时间预测

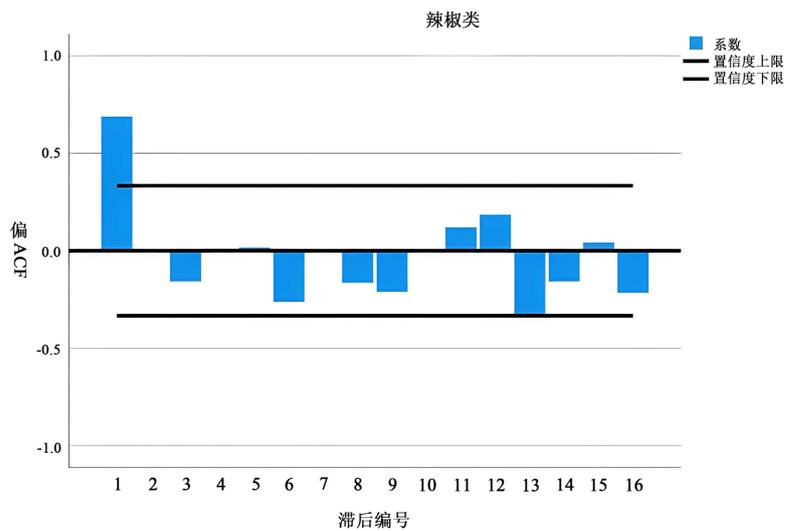
基于已知公式和计算出的自相关系数与偏自相关系数,便可以做出时间序列预测图像(置信区间为 95%),大致预测未来七天各品类补货的走向(图 8)。



(a)



(b)



(c)

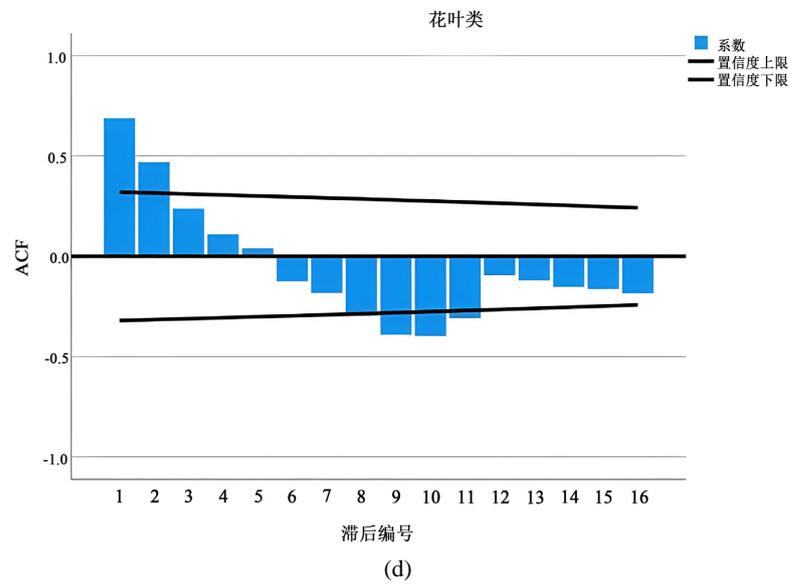
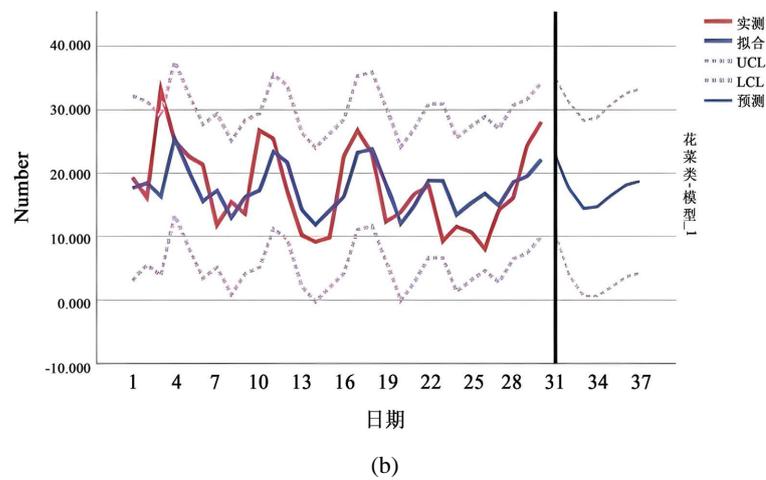
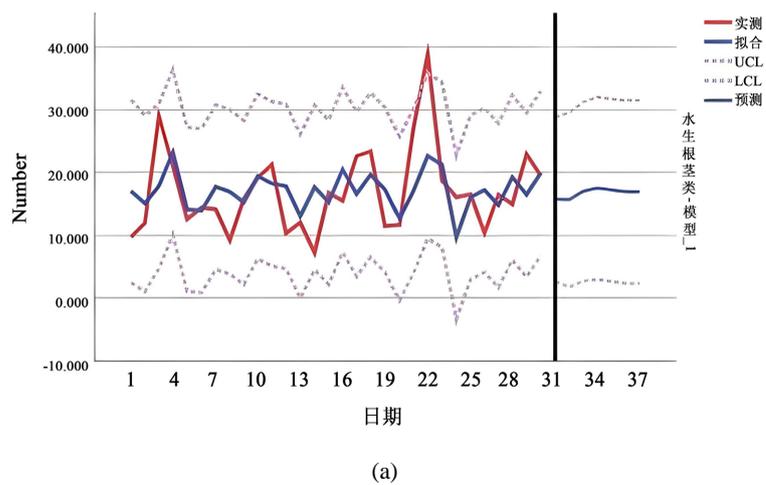


Figure 7. Partial ACF and ACF sequence diagram  
图7. 偏 ACF 和 ACF 序列图



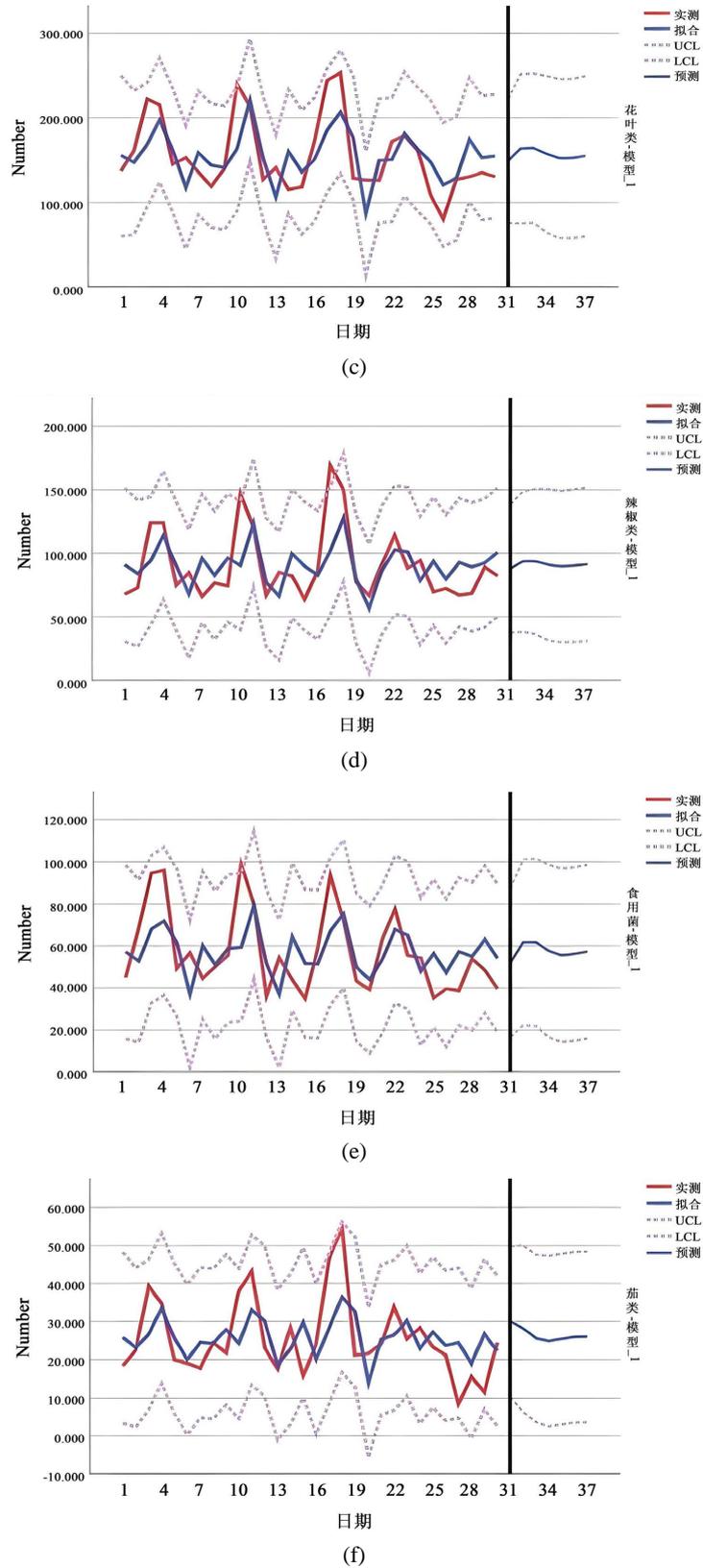


Figure 8. Time series images of each category  
图 8. 各品类时间序列图像

根据这些图像得出最终结果(表 4):

**Table 4.** Replenishment quantity of each category in the next seven days (unit: kg)

**表 4.** 各品类未来七天补货量(单位: 千克)

	辣椒类	花叶类	水生根茎类	食用菌	花菜类	茄类
2023.7.1	96.35	166.64	19.00	52.80	25.11	32.01
2023.7.2	102.66	183.16	18.93	62.85	19.35	30.16
2023.7.3	102.55	183.82	20.51	61.84	15.96	27.32
2023.7.4	99.48	175.88	21.12	58.69	16.27	26.56
2023.7.5	98.16	170.19	20.76	56.57	18.21	27.16
2023.7.6	99.04	170.39	20.41	57.20	20.02	27.69
2023.7.7	100.07	173.41	19.42	58.40	20.68	27.75

而通过回归拟合模型, 得到未来七天的定价策略(表 5):

**Table 5.** Pricing strategy of each category in the next seven days (unit: yuan, with two decimal places)

**表 5.** 各品类未来七天定价策略(单位: 元, 保留两位小数)

	辣椒类	花叶类	水生根茎类	食用菌	花菜类	茄类
2023.7.1	6.07	17.32	8.48	7.40	10.18	8.28
2023.7.2	7.34	19.44	8.69	7.45	11.15	9.27
2023.7.3	7.22	21.52	9.05	7.69	12.00	9.36
2023.7.4	8.75	20.95	9.34	8.00	11.77	9.22
2023.7.5	6.87	23.66	8.72	7.68	10.98	10.02
2023.7.6	7.03	22.17	9.37	7.84	11.13	8.59
2023.7.7	6.56	18.94	8.99	7.86	11.54	9.48

若想预测未来更多天的补货和定价策略, 可以继续沿用此方法。如果预测未来的天数较多, 可以根据历史数据适当地往前调, 比如预测未来一个月, 可以用历史的前两个月或三个月的数据进行拟合, 使得预测的数据更加准确。

## 6. 补充

### 6.1. 基于优化模型的未来一天预测

因蔬菜类商品的销售空间有限, 实际的补货量肯定还会有偏差。根据实际情况, 若要求可售品类总数控制在  $a \sim b$ 。根据已知的一周的可售品种, 给出未来一天的菜品类补货量和定价策略, 在尽量满足市场对各品类蔬菜商品需求的前提下, 使得商超收益最大。

为此可以建立如下规划模型, 其中  $x_i$  为每个品类的可售量,  $q_i$  为补货量,  $c_i$  为预测当天可售品类的平均定价,  $p_i$  为每个品类补货的成本,  $l_i$  为  $x_i$  对应的平均损耗率,  $D$  为该周所有菜品类的平均总销量之和,  $Z$  为利润函数。

目标函数确定：总收入减去总成本，再与可售量相乘，是当天的利润，即：

$$\text{Max } Z = \sum_i c_i x_i - \sum_i p_i x_i / (1 - l_i)$$

其中  $x_i / (1 - l_i) = q_i$

$i=1,2,3,4,5,6$ ，分别对应六个菜品类。

约束条件确定：每个品类可售量均大于等于  $c$  千克，且总个数不少于  $a$ ，不大于  $b$ ，可售量乘以非损耗率为平均销量和，即：

约束条件： $a \leq x_i \leq b$

$$\sum_i (1 - l_i) x_i = D$$

## 6.2. 建议

为了更好地制定蔬菜商品的补货和定价决策，商超还需要采集市场需求数据、竞争对手数据、成本数据、天气和节假日数据、商超库存状况数据、客流量数据及消费者意见数据，这些数据对解决蔬菜商品的补货和定价问题非常有帮助。以下为意见所对应的理由：

- 1) 通过分析市场需求数据，商超可以更准确地预测未来的需求量，从而合理安排补货量。
- 2) 竞争对手数据可以帮助商超了解市场竞争情况，分析自身的优势劣势，制定有竞争力的定价策略，还可以学习改进促销方案，在处理滞销商品的同时尽可能的提高收益。
- 3) 成本数据可以帮助商超在定价决策中考虑利润和成本的平衡，确保商超持续盈利的能力。
- 4) 天气和节假日数据一定程度上影响消费者出门购物的行为。如暴雨天气导致积水 and 交通问题，阻碍消费者出行。节假日的空余时间以及促销活动则可以刺激消费者的购买欲。
- 5) 商超库存情况数据有利于及时了解菜品数目，方便预测哪些蔬菜库存量高需要打折销售，减少因滞销造成的蔬菜变质，以及哪些库存量低需要补货或是抬高价格。
- 6) 客流量数据方便了解不同时间段的顾客数量以及不同年龄段顾客分布情况，可以针对性的设置举办促销活动的时间以及菜品摆放的位置。
- 7) 消费者意见数据对于商超这类接近于服务产业很有价值，针对消费者提出的建议和问题，考量后做出相应采纳和修正。例如商品的质量和价格等方面的反馈，还可以定期搞活动让消费者填问卷，统计菜品间常见的搭配关系。

## 7. 小结

本文主要围绕着蔬菜类商品的自动定价与补货决策展开研究。首先建立并解决了蔬菜类别和销售分布之间的皮尔逊相关模型，得出了每个类别之间的相关关系。随后，建立了 ARIMA 预测模型，用于预测未来一周内新鲜食品超市中每个蔬菜类别的补货数量和定价策略。最终，利用线性规划对 ARIMA 预测模型进行优化，以在市场需求下获得最大收入的情况下解决蔬菜商品的自动定价和补货决策。在本文中，运用了多种数据分析方法，提出了一套有效的定价和补货决策方案，为新鲜食品超市的经营管理提供了重要的参考依据，希望能够为相关领域的研究和实际应用提供有益的借鉴和启发。

## 参考文献

- [1] Wang, K.L., Su, K.M. and Li, H. (2023) Random Forest-Based Restocking and Pricing Prediction for Vegetable Items. *Frontiers in Business, Economics and Management*, **11**, 352-356. <https://doi.org/10.54097/fbem.v11i2.12638>
- [2] 陈军, 康莎, 等. 面向双渠道销售的农产品定价与库存补货联合决策[J]. *工业工程*, 2023, 26(3): 39-46.

- [3] 乔雪. 考虑销售损失的生鲜产品的联合补货定价策略[D]: [博士学位论文]. 南京: 东南大学, 2023.
- [4] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 121-139, 298-300.
- [5] 王燕. 应用时间序列分析[M]. 北京: 中国人民大学出版社, 2022.