基于 ARIMA、LSTM 残差拟合网络和价格弹性的 商超自动定价与补货决策模型

摘要

本文认为蔬菜销售量受三个因素影响:销量的自回归性、和其他蔬菜销量的互相关性、受当天定价影响的弹性。本文首先分析了蔬菜销量随时间的分布关系(自回归性)和不同蔬菜销量之间的影响(互相关性);其次先后引入差分自回归移动平均模型(Autoregressive Integrated Moving Average Model, ARIMA)、基于长短期记忆网络(Long Short-Term Memory, LSTM) 搭建的残差拟合网络 RFAN(Residual Fitting Autoregression Network)和价格弹性通过二次残差不断逼近拟合蔬菜的真实销售量,将其转化成本利润率的函数,以求解收益函数的最值;接着结合最大收益情况通过整数规划进行蔬菜选购决策;最后给出商超补充调研数据的建议,分析其对本模型求解问题的帮助,并通过增加疫情相关指标的实验证明我们建议的合理性。

针对问题一,首先对于每个品类进行时间序列分解,得出其总体趋势(先减后增)和季节性规律。然后对于不同品类销售量的时间序列进行 Spearman 相关性分析并绘制热力图,发现品类间销售量的相关性呈现整体正相关,少数负相关的特点。其次对具有复杂单品种类的品类内部进行数据清洗和整合,过滤长时间或近期无销售量的数据,同时合并不同销售方式的单品,随后进行 R 型层次聚类分析,减小数据维度进一步划分子类,并在同类单品中选择高销量的典型分析其销售量随时间变化特征。

针对问题二,首先采用 ARIMA 模型拟合销量曲线,其次采用 RFAN 网络结合历史定价、销售情况、损耗情况等对 ARIMA 拟合残差进行修正,得到第二天的销售量预测基准线,然后利用基于价格弹性的二次残差修正确定销售量波动情况和销售价格波动情况的函数关系,最优化总收益。最后循环使用该策略给出七天的补货量和定价策略。

针对问题三,首先根据各品类单品数量确定品类分布约束。然后结合问题一的分析 在多单品品类下细化种类分布约束,由此确定销售单品选择池,再基于问题二的销售量 预测模型预测池内各单品的7月1日销售量,结合上述定价策略算出各单品最大利润及 最终定价。最后增加单品订购种类数和订购量等条件约束建立整数规划模型确定最终单 品的选择,以最大化总利润。

针对问题四,我们从**宏观角度**和**微观角度**分析了影响蔬菜商品需求的可能因素,最后对数据采集给出以下重点建议: 1. 增加采集气候数据等会影响蔬菜产量和品质的分析指标 2. 增加采集社会经济数据等会影响销售量的经济学指标 3. 增加采集其他同类型商超定价和促销的数据。

关键字: ARIMA LSTM 残差拟合 RFAN 价格弹性 函数极值求解 整数 01 规划

一、问题背景与重述

1.1 问题背景

生鲜商超在运营时往往面临着蔬菜类商品定价与补货决策困难的问题。[1] 由于蔬菜类商品的新鲜度和品相随时间衰减明显,影响售价,且大部分品种无法保存到隔日继续销售,因此,商超通常需要根据各商品的历史销售和需求情况每天补货。[2] 然而由于商超销售的蔬菜品类繁多,产地也不尽相同,且进货时间通常在凌晨,因此商家每日做出各蔬菜品类的补货决策时,并不能确切知道具体单品和进货价格。蔬菜的定价一般采用"成本加成定价"[3]方法,即按成本利润率来确定所加利润的大小,具体公式为

价格 = 单位成本 + 单位成本 × 成本利润率 = 单位成本 × (1 + 成本利润率)

运损和品相变差的商品通常需要打折销售。要做出科学合理的补货决策和定价决策,可靠的市场需求分析极为重要。由于蔬菜类商品的供应和需求往往与时间有关,且商超的销售空间有限,因此,商超还需要安排合理的销售组合。

1.2 问题提出

现已知某商超经销的蔬菜商品信息和近期的损耗率数据,以及一段时间内的销售与成本数据,试解决以下问题:

- (1) **问题一**:分析并给出蔬菜类商品不同品类、不同单品销售量的分布规律,以及相互之间的关系。
- (2) **问题**二:若商超以品类为单位来规划补货策略,分析各蔬菜品类的销售总量与成本加成定价的关系,并设计一种各蔬菜品类未来一周(2023年7月1-7日)的日补货总量和定价策略,使得商超收益最大。
- (3) **问题**三:考虑到蔬菜类商品的销售空间有限,将可售单品总数控制在 27-33 个,且各单品订购量需满足最小陈列量 2.5 千克的要求。根据 2023 年 6 月 24-30 日的可售品种,在尽量满足市场对各品类蔬菜商品需求的前提下,给出 7 月 1 日的单品补货量和定价策略,使得商超收益最大。
- (4) **问题四**:说明为了更好地制定蔬菜商品的补货和定价决策,商超还需要采集哪些相 关数据,这些数据对解决上述问题有何帮助,并给出理由。

二、模型假设

1. 假设一:认为商品当日销售价格缺失是当日未进货导致的,不影响商品价格变化的整体趋势。

解释:某些蔬菜可能并不是每天都有进货,因此对这些蔬菜的分析在处理价格缺失情况的时候采用向前补齐的方式。

2. 假设二: 在本模型对历史数据的分析过程中不考虑当天某蔬菜未销售完的情况。

解释: 商超历史销售数据中并未给出每天某蔬菜的进货量, 且大部分蔬菜保质期断, 很难留到第二天继续销售。因此认为销售量即是前一天的进货量。

3. 假设三:认为商超过年附近几天的停产情况不影响各时间序列的变化情况。

解释:过年时商超会关门导致缺乏几天的销售数据,我们认为年后会正常恢复销售,因此直接删去过年时的几天。

三、符号和变量说明

| 符号或变量 | 意义 | 符号或变量 | 意义 |
|------------|------------------------|--------------|--------------------|
| E_t | 第t天的预估销售量基准线 | A_t | 第 t 天的预估销售量 |
| B_t | 第 t 天的批发价格 | P_t | 第 t 天的定价 |
| O_t | 第 t 天的理论率 | ϵ_t | 第 t 天的价格弹性 |
| ϵ | 价格弹性平均值 | W_{ij} | i 品类下 j 单品的最大收益 |
| a_{ij} | i品类下j单品处于最大收益下的预估销售量 | p_{ij} | i品类下j单品处于最大收益下的出售价 |
| Max_i | i品类的最大选取单品数 | Min_i | i品类的最小选取单品数 |
| MIN | 最小总选取单品数 | MAX | 最大总选取单品数 |
| x_{ij} | 是否选择订购 i 品类的 j 单品(0/1) | E_{ij} | 订购 i 品类的 j 单品的损耗率 |
| perc | 损耗单品的打折率 | Max_{ik} | i品类下K子品类的最大选取单品数 |
| Min_{ik} | i 品类下 K 子品类的最小选取单品数 | | |

四、问题分析

4.1 问题一分析

对于问题一,先需要对品类进行分析,通过绘制每个品类的折线图观察该品类蔬菜销量的季节性和趋势,然后对品类间求解相关系数,绘制热力图分析品类间的相互关系。然后需要对单品进行分析,先分析单品的数据以过滤掉不需要进行分析的单品,然后将不同销售方式或供应来源的相同单品合并分析。然后对单品数量较多的品类进行 R型层次聚类,以便于进一步进行单品的相关性分析和季节性分析。

4.2 问题二分析

对于问题二,我们认为新一天的蔬菜销售量由三类因素影响: 1. 历史销售量数据(自回归性) 2. 历史定价、其他蔬菜销售量数据(互相关性) 3. 新一天的成本定价,即利润率(价格弹性)。因此我们采用差分模型进行第一次拟合,再利用基于多元 LSTM 构建的时序神经网络进行残差拟合,最后通过成本定价的迎合进行二次残差拟合。最终建立定价和销售量的关系,得到总收益的函数表达式,进行极值求解。

4.3 问题三分析

为了进行蔬菜单品采购的选择,首先需要确定单品、品类分布约束。然后基于问题 二的销售量预测模型确定池内各单品的7月1日销售量和定价的关系,结合上述定价策 略算出各单品最大利润及最终定价。最后结合约束条件建立整数规划模型确定最终单品 的选择,以最大化总利润。

4.4 问题四分析

为了给出建议,需要查阅相关文献,总结得出影响蔬菜供应和需求的因素。除此之外,还应采集这方面的数据进行实验,检测拟合和预测结果是否会更好。同时还可以结合同类商超的销售数据,据此制定相应的补货和定价决策。

五、模型概览

结合上述分析,给出解决本问题的总模型框架图,如图1所示。

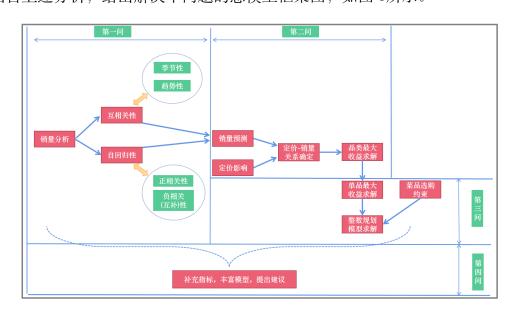


图 1 模型整体框架图

六、问题一的模型建立与求解

6.1 各品类自回归性分析

由于题中销售数据粒度太高,不便于分析,首先需要对数据进行汇总简化,以天为单位统计不同品类蔬菜的销售量数据总和。为了使绘制的折线图更加清晰,再以10天为一个单位绘图,如图2所示。



图 2 以 10 天为单位的各品类蔬菜销量变化情况

可以直观地看出,各种品类的蔬菜销量都存在明显的季节性变化。因此,我们使用 **滑动平均**的方法对其进行**时间序列分解**,得到六种品类蔬菜的趋势序列和季节性序列, 分别进行分析。

6.1.1 趋势分析

从总体变化趋势的角度来看,大致可分为两种情况,图 3展示了两种典型代表,剩余品类的分析图详见附录。

第一种情况:除了茄类蔬菜以外,其他品类蔬菜的总体趋势都类似图 3a的花叶类蔬菜,随着时间先下降后上升。猜测可能是因为该商超所在地 2021 年遭受了某种较为严重的自然灾害,导致蔬菜大量减产。

第二种情况: 茄类的销量总体呈下降趋势。猜测可能是由市场竞争等因素导致的。

6.1.2 季节性序列

从季节变化的角度来看,每种品类的蔬菜均存在明显的季节性变化:

• 花叶类: 销售量高峰在每年的冬季;

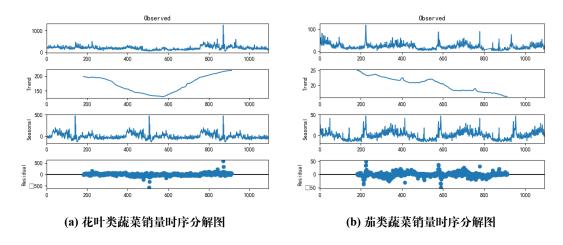


图 3 两类蔬菜品类销量时间序列分解的典型代表

• 茄类: 销售量高峰在每年的 2 月和 9 月左右;

• 花菜类: 销售量高峰在每年的2月和9月左右;

• 辣椒类: 销售量高峰在每年的 2 月和 9 至 11 月;

• 食用菌;销售量高峰在每年的2月和11月左右;

• 水生根茎类: 销售量高峰在每年的2月和12月左右。

6.1.3 品类间互相关性分析

接下来分析各品类蔬菜销量之间的相互关系。绘制热力图如图 4所示。



图 4 各品类蔬菜销量相互关系

可以得出,食用菌类的销售量变化与水生根茎类、花叶类、辣椒类分别存在较为显著的正相关性,而茄类则是唯一与其他品类存在负相关性的蔬菜。这与上文分析的茄类蔬菜销量逐渐下降的趋势相符合。

此外,我们还进行了 **Kendall 系数一致性检验**。结果显示,显著性水平为 1% 时,总体数据的显著性 P 值为 0.000,小于 0.05,拒绝原假设。同时模型的 **Kendall** 协调系数 W 为 0.732,因此数据呈现高度的一致性,即各品类的销量分布趋势整体相近。

6.2 各蔬菜单品销量的分布规律及相互关系分析

6.2.1 数据预处理

首先**过滤**在时间序列中非常稀疏或低频的数据。观察数据,可以发现存在某些单品的销量长时间为 0、总销量极低,或者近期销量极低。对于这类单品,我们不必讨论它们的分布规律。根据数据的特征,我们规定若超过 80% 的天数销售量为 0 或近一年内销售量不为 0 的天数少于 30 天,则删去,不予考虑。过滤情况见附录表 5。

其次,进一步**整合数据**。在观察数据的过程中,我们发现同一种蔬菜单品存在不同的销售方式(如"云南油麦菜"和"云南油麦菜(份)")和供应来源。为了探究不同销售方式和供应来源对销售量的影响,我们对所有这些蔬菜的不同销售种类做折线图,并进行了相关性分析。以下以云南油麦菜为例进行说明。

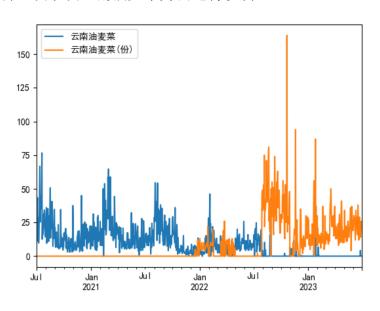


图 5 不同销售方式下油麦菜销售情况

由图 5可以看出,在"云南油麦菜(份)"销售方式出现后,云南油麦菜原本的销售方式被渐渐取代。结合其他多组数据可以发现,份装的销售方式均呈现后来居上的态势,和原来的销售方式结合所得的该蔬菜单品总销量变化较为稳定。我们对其进行时间序列的皮尔森相关性分析,计算得到相关性系数为-0.43,也很好的体现了互补性。

因此,在后面的分析中,我们**把不同销售方式、不同供应来源的同一蔬菜单品的销量相加作为该蔬菜单品的总销量进行分析**。我们对筛选处理后包含单品数量较多的花叶类和辣椒类进行编号,见附录表 6、表 7。

6.2.2 R 型层次聚类

由于不同品类所包含的单品数量差距较大,部分品类的蔬菜包含的单品数量过多,为了便于分析,根据单品数量的多少,我们采用不同的方法进行销量相关性分析:

- 对于所含蔬菜单品数量较少的品类,包括花菜类、茄类、食用菌和水生根茎类,直接进行相关性分析。
- 对于所含蔬菜单品较多的品类,包括花叶类、辣椒类,首先对其进行 R 型层次聚类 分析,随后再对每个子类进行相关性分析。

R型层次聚类分析的结果如图 6所示,其中横轴数字即为筛选后给出的的编号。结合聚类图,我们可以将花叶类蔬菜分为五类,将辣椒类蔬菜分为两类。

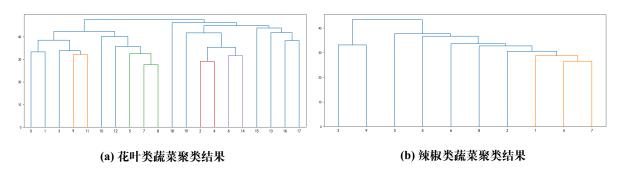


图 6 聚类结果

6.2.3 相关性分析

对每种品类的蔬菜绘制 **Spearman 相关系数热力图**,分析各单品间的相关关系。食用菌的情况如图 7所示,其余品类见附录。



图 7 食用菌相关系数热力图

可以发现,金针菇和虫草花、金针菇和杏鲍菇、金针菇和双孢菇之间存在较为显著的正相关性,西峡花菇和西峡香菇之间存在较为显著的负相关性。其余单品之间均不存在显著的相关性。

6.2.4 季节性分析

结合数据分析结果和生活常识,可以发现不少蔬菜的销量存在明显的季节性特征。 我们将 4-10 月作为暖季,11-次年 3 月作为冷季分析蔬菜的销量情况,绘制分析图。花 叶类情况如图 8所示,其余品类见附录。

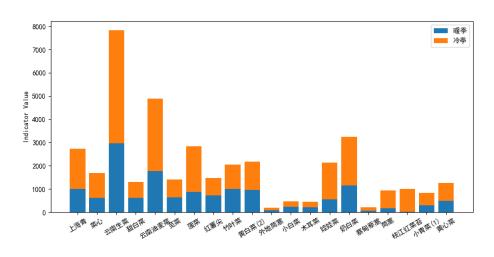


图 8 花叶类蔬菜各单品总销量及季节组成

可以发现,不同单品蔬菜的销量存在明显的差距。大部分单品冷季的销量均大于暖季。且存在有像枝江红菜苔这样基本仅在冷季有销量的蔬菜。

七、问题二的模型建立与求解

7.1 数据处理

针对商超以品类为单位的补货计划,需要计算每个品类的销售量、批发价格、定价 (销售价格)等数据。考虑到不同蔬菜的购买量不同,我们采用求总购买量和总花费的 方式计算品类的批发价格和销售价格,体现了销售量权重的影响。

7.2 整体思路

在本问题的求解中,整体的思路流程框架如图9、10所示。

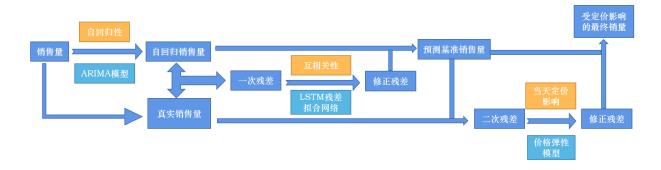


图 9 问题二训练拟合流程图

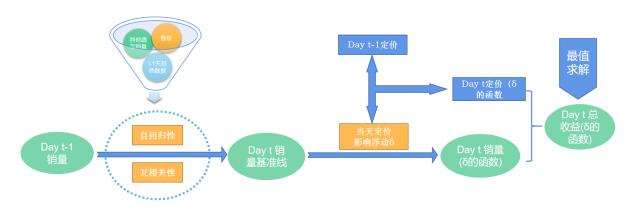


图 10 问题二预测求解流程图

模型步骤如下:

- 1. 利用 ARIMA 模型对品类的销售量进行初步时序预测。
- 2. 搭建以多变量 LSTM 网络为核心的 RFAN 网络,利用其对 ARIMA 模型的预测 结果和实际销售量的一次残差进行拟合,预测 ARIMA 模型的残差修正序列。以此来确定销售量的浮动基值。
 - 3. 利用价格弹性系数建立 RFAN 网络预测二次残差和售价变动比之间的关系。
- 4. 利用销量、定价等变量结合其相关关系确定第二天的总收益函数(化简为单变量),利用函数极值求解进行计算。

注:各蔬菜品类的销售总量和成本定价关系的分析在需求价格弹性模型介绍中进行了体现,由于模型逻辑并没有先进行分析。

7.3 模型介绍

7.3.1 ARIMA 模型

ARIMA模型,全称自回归差分移动平均模型 (Autoregressive Integrated Moving Average Model),其基本思想是利用数据本身的历史信息来预测未来,即试图通过数据的自相关性和差分的方式,提取出隐藏在数据背后的时间序列模式,然后用这些模式来预

测未来的数据 [4]。模型的一般形式如下式所示:

$$X_t = c + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \ldots + \beta_q \varepsilon_{t-q}$$

模型预测步骤如图 11所示。

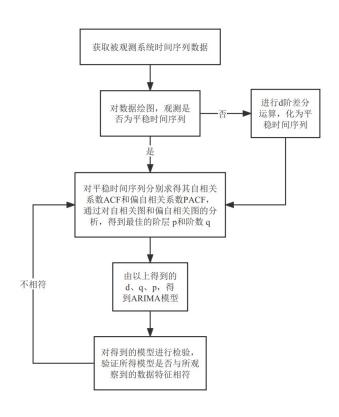


图 11 ARIMA 模型预测步骤图

7.3.2 RFAN 网络

RFAN 网络结构如图 12所示。

RFAN 网络的核心结构为一个多元 LSTM 网络,我们利用其去拟合 LSTM 模型预测值和真实值的残差。我们认为影响残差的因素有:

- 1. 历史数据之间的相关性,体现为历史数据矩阵和相似度矩阵相乘得到的结果。
- 2. 时间(即季节性),体现为365天进行embedding后的时间表征数据。
- 3. 其余特征量,如定价,损耗率等随时间变化的数据。
- 4. 历史残差序列。

因此构成了一个输入数据为前 n 天(n 为设置的步长,本模型为 30)各数据,预测值为 n+1 天残差的多元时序网络。

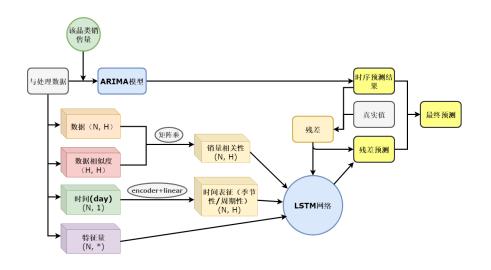


图 12 RFAN 网络结构图

7.3.3 需求价格弹性模型

价格变动会影响需求的变动。在本题中我们认为销售量反应的是需求量,因此定义 定价变动的百分比与销售量变动的百分比之比为需求对价格的弹性,简称需求弹性[6]。

在通过 AEIMA 和 RFAN 网络对销售量序列的时间特征进行拟合预测之后,我们认为当天的实际销售量还会受到定价变动百分比的影响,该影响体现在 RFAN 网络结果和真实值的残差中(二次残差),我们对二次残差和定价数据进行了相关性分析,检验得到相关系数为,证明确实存在较大影响。

因此,我们假设 t 天的预估销售量基准线 (由 RFAN 模型得到的)、预估销售量、批发价格、定价、利润率(定价/批发价)、价格弹性分别为 E_t , A_t , B_t , P_t , O_t , ϵ_t , 依照需求价格弹性模型,定义如下关系:

$$\frac{A_t - E_t}{E_t} = \epsilon_t \frac{P_t - P_{t-1}}{P_{t-1}}$$

作出定价(售价)弹性比和销量浮动比随时间变化的曲线如图13所示。

可以发现其存在较为明显的负相关性,因此我们依据预测日期前 100 天的历史数据 计算 ϵ_t ,作出其大小的直方图如图14所示。

可以看出弹性系数大部分的值在 $\epsilon_t = -0.5$ 附近波动,因为我们过滤去绝对值过大的峰值(这可能是由当天的其他因素如销售量基准线骤变导致的)后对其取平均得到一个定值,以便于后面的分析。

7.3.4 总收益函数设置

通过价格弹性,我们可以列出计算第t天某品类蔬菜总收益 M_t 的方程组

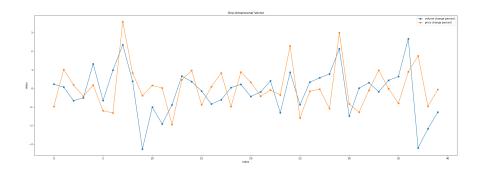


图 13 弹性关系

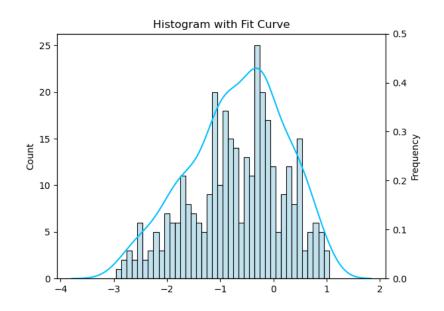


图 14 弹性系数直方图

$$\begin{cases} M_t = (P_t - B_t) \cdot A_t \\ \frac{A_t - E_t}{E_t} = \epsilon \frac{P_t - P_{t-1}}{P_{t-1}} \\ O_t = \frac{P_t}{B_t} \\ \epsilon = \frac{1}{100} \sum_{\tau=1}^{t-1} \epsilon_{\tau} \end{cases}$$

注: ε 计算式求和是在筛选了过大过小值之后进行的。

其中预估销售量和批发价格可以通过 RFAN 模型预测得到,前一天的定价也可以根据 RFAN 模型得到,存在实际销售量和定价(利润率)两个变量,将 M_t 化简为 O_t 的单变量方程如下:

$$M_t = B_t E_t (O_t - 1) (\frac{PB_t}{P_{t-1}} O_t - P + 1)$$

这是一个一元二次方程,因为P < 0,所以存在极大值,在 $O_t = \frac{1}{2} + \frac{(\epsilon-1)P_{t+1}}{2PB_t}$ 时取得。可以看出,该极大值点与 B_t ,即与当天批发价格无关,这符合我们在确认定价策略时不知道当天的批发价格的实际情况。

7.4 模型求解

7.4.1 ARIMA 模型和 RFAN 网络

近 100 天 ARIMA 模型预测结果、经过 RFAN 网络残差修正后的预测结果、销售量 真实值对比如图 15所示。

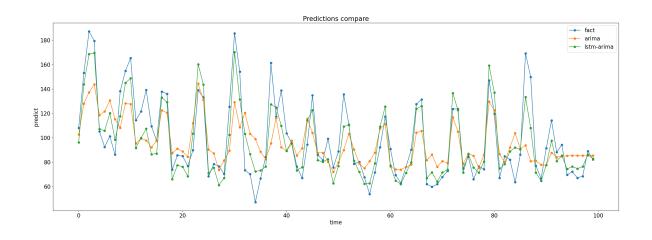


图 15 预测效果对比图

RFAN 网络在训练时的 loss 变化曲线,以及训练完成后的残差预测结果与真实结果的对比图如图 16、图 17所示。

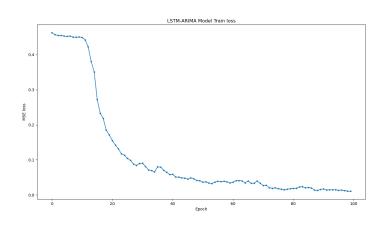


图 16 loss 曲线 (平稳 loss = 0.054)

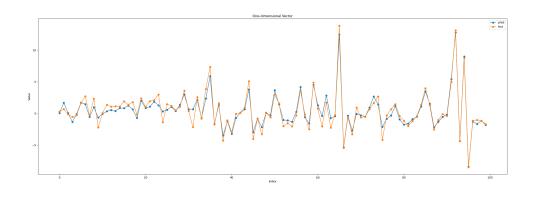


图 17 训练销量残差

可以发现, RFAN 网络拟合的效果较好, 因此经过 RFAN 修正后的预测结果也更接近真实值。

7.4.2 总收益函数求解

通过 RFAN 的训练结果得到残差加上 ARIMA 模型的拟合值得到 7月1日-7月7日 销量预测基准线后,就可以代入求解总收益函数的极值。对于 7月1日-7月7日的预测,我们将弹性定价结合每天预测到的销量基准循环 7次,得到 7天的预测结果,我们在下表 1中展示**辣椒类**的求解结果。其余品类的求解结果见附录。

表 1 辣椒类最大收益求解结果

| 日期 | 补货量 | 定价 | 定价/成本(利润率) | 收益 |
|------|-------------|-------------|-------------|-------------|
| | .,,,, | | | |
| 7月1日 | 87.69112983 | 5.9777356 | 1.940754339 | 254.096119 |
| 7月2日 | 95.73848934 | 5.847804198 | 1.773288621 | 244.141487 |
| 7月3日 | 82.72264912 | 5.710090986 | 1.81262709 | 211.763103 |
| 7月4日 | 90.95036716 | 5.582515942 | 1.75489197 | 218.4081536 |
| 7月5日 | 79.16701328 | 5.447697671 | 1.599425067 | 161.6323401 |
| 7月6日 | 79.41408848 | 5.316372202 | 1.57409329 | 153.9802205 |
| 7月7日 | 89.51359531 | 5.196558637 | 1.447547281 | 143.817256 |

7.5 模型合理性分析

- 1. 本模型采用以 LSTM 为核心的网络拟合 ARIMA 模型预测值和真实值的残差,充分考虑到了两种模型的优势和缺陷。ARIMA 模型适合拟合较为平稳的序列,可以识别时间序列的线性特征,但容易忽视模型的非线性特征,可以理解为 ARIMA 模型预测结果和真实结果的残差反映了该序列的非线性特征,我们用 LSTM 去拟合该残差来对ARIMA 进行修正,得到了更准确的预测结果。
- 2. 本模型结合经济学原理,通过价格弹性来判断定价不同对销售量的变化,以此进行决策,具有合理性。

八、问题三的模型建立与求解

8.1 单品分布约束确立

首先根据题意筛选出 6 月 24-30 号的可售品种。我们定义在 6 月 24-30 号有销量, 且平均销量大于 1.5 千克的即为可售品种,最终筛选出花叶类 12 种,水生根茎类 4 种, 花菜类 2 种, 茄类 3 种,辣椒类 8 种,食用菌 5 种。

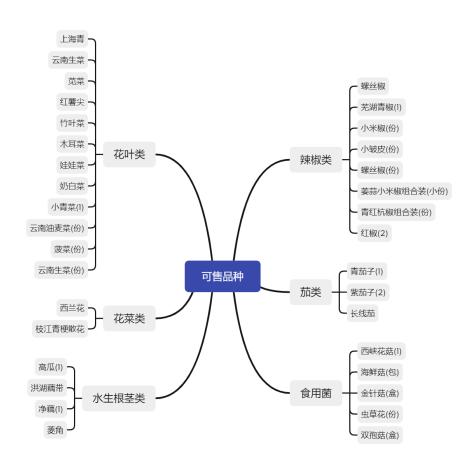


图 18 可售品种

然后结合第一问,我们发现存在互补型的单品销售数据(注意如果两种同类蔬菜都在近7天有销售,如高瓜(1)和高瓜(2),则仍然视为两种),将这些单品时间序列的数据进行加和用于后续模型。

最后我们对剩余种类数较多的水生根茎类、花叶类、辣椒类、食用菌类参考第一问的层次分析聚类,进一步细分为子类,聚类数量的标准为单品数量最多的子类不超过4个,根据分类结果做如下处理:

- 1. 对于单一单品构成的子类,证明该单品具有独特性,直接对其进行选择。
- 2. 对于多单品构成的子类,参考已经确定选择的单品数量为该子类确定单品选择数量。

根据该策略和聚类结果,得到花叶类细分结果,如表2所示。

上海青 云南生菜 苋菜 红薯尖 竹叶菜 木耳菜 花叶四类 花叶四类 花菜五类 花菜五类 花菜五类 花菜五类 娃娃菜 奶白菜 小青菜 云南油麦菜 菠菜 云南生菜(份) 花叶一类 花叶三类 花叶三类 花叶二类 花叶二类 花叶二类

表 2 花叶类次聚类结果

对该分类结果,直接选择娃娃菜(单品一类),并要求花叶三类、四类至少选择一个,花叶二类至少选择两个、以此类推。

8.2 各单品销售量预测和最大利润计算

对于上述单品,我们采用第二问的模型和算法计算每个单品的7月1日销售量和最大利润,表3给出花菜类的预测结果,其余品类见附录。

 単品
 成本
 定价
 补货量
 收益

 西兰花
 7.622778
 10.36467
 20.22377
 55.45136

 枝江青梗散花
 8.638414
 9.26241
 4.910001
 3.063821

表 3 花菜类单品补货和定价策略

8.3 整数规划模型

假设 i = 1, 2, 3, 4, 5, 6 表示六种品类, $j = 1, 2, 3, 4, ..., Num_i$ 表示每种品类的不同单品, Num_i 表示每种品类中的单品总数, $k = 1, 2, 3, 4, ..., Cluster_i$ 。具体变量和常量的定义见第三节说明。

订购蔬菜的整数规划模型为:

$$\max \sum_{ij} (x_{ij}[W_{ij} + \max(a_{ij} - 2.5, 0) \cdot P_{ij}] - E_{ij} \cdot a_{ij} \cdot perc)$$

$$s.t. \begin{cases} \sum_{ij} x_{ij} \leq MAX \\ \sum_{ij} x_{ij} \geq MIN \\ \sum_{j} x_{ij} \leq Max_{i}, \quad \forall i \\ \sum_{j} x_{ij} \geq Min_{i}, \quad \forall i \\ \sum_{j \in cluster_{k}} x_{ij} \leq Max_{ik}, \quad \forall k, i \\ \sum_{j \in cluster_{k}} x_{ij} \geq Min_{ik}, \quad \forall k, i \end{cases}$$

带入 W_{ij} , a_{ij} , p_{ij} 参数和实际约束条件求解,最终在所有可售品种中删去云南生菜、青茄子 (1)、双孢菇 (盒) 和高瓜 (1) 四类。以水生根茎类为例列出规划后的补货量和定价策略,以及规划前后的收益对比,如表 4所示,完整数据见附录。可以发现我们删去的四种蔬菜的修正收益均小于零,这体现了策略的合理性。

单品 原预测收益 修正收益 是否选择 补货量 定价 定价/成本(利润率) 10.56 高瓜(1) -3.99496 -5.10757 否 3.6 0.905 洪湖藕带 2.786025 0.521147 是 5.077 18.55 1.03 3.413812 是 5.67 10.98 净藕(1) 3.758942 1.064 9.421875 是 1.95 菱角 1.413328 14 1.528

表 4 最终补货计划和定价策略

九、问题四的模型建立与求解

9.1 数据采集建议

上文主要根据商超的历史销售数据分析了如何补货和定价能够使得收益最大化。除 了总结历史规律,商超还可以通过调研近期的蔬菜需求情况来更加全面地预估销售量, 进而制定决策计划。图 19给出了影响蔬菜商品需求的可能因素 [5]。

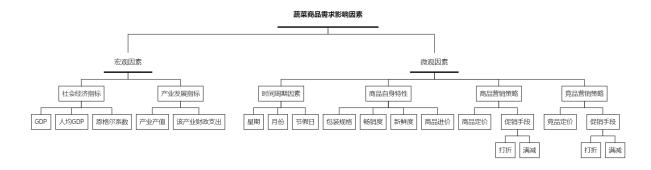


图 19 蔬菜商品需求影响因素

据此,从**宏观角度**,我们建议商家额外采集 GDP、恩格尔系数、蔬菜产业产值和产业财政投入等社会经济数据,了解近期经济状况和居民消费情况,及时更新产业的最新发展动态,在经济不景气或者产业衰退时可适当减少补货量,避免出现浪费。商家还可以多多关注时下流行趋势,收集各品类蔬菜的热门程度数据,保持对市场的敏感性。从微观角度,商家可以多多收集竞争对手的有关信息,包括对方的进货情况、竞品的定价和营销手段等,并及时采取相应的对策,例如避免补充大量相同商品、适当降低利润率等以提高自己的竞争力。

此外, 商家还可以采集有关数据以预测供货的数量和质量, 例如气候、环境、自然灾害等影响蔬菜产量和品质的数据。在供应量较多时, 批发价也相应较低, 可以适量增加补货量, 供应量较少时则相反。

9.2 模型验证

为了验证上节建议的有效性,本节在问题二模型的基础上添加疫情因素 (疫情反映了经济和环境的变化),我们增设代表疫情影响的 01 指标 (0 表示疫情加剧, 1 表示疫情恢复),重新拟合 RFAN 神经网络,训练 loss 变化如图 20所示:

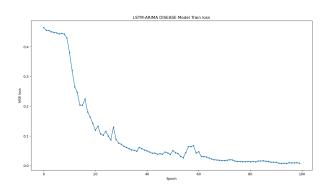


图 20 添加疫情指标后的 loss 下降趋势

可以发现, loss 变化的波动变大, 但最终的平稳 loss 为 0.047, 对比原来的 loss 有少量下降,证明模型拟合的结果更好。因此,指标的添加对模型预测效果提升有意义。

9.3 总结

最后,结合原理分析和实验验证,我们给出如下建议:

- 1. 增加采集气候数据等会影响蔬菜产量和品质的分析指标
- 2. 增加采集社会经济数据等会影响销售量的经济学指标
- 3. 增加采集其他同类型商超定价和促销的数据

十、模型评价与改进

10.1 优缺点分析

10.1.1 优点

- 1. 本模型充分考虑了影响蔬菜销售量的因素,并采用二次残差拟合适合预测销售量的结果更接近真实值。
- 2. 本模型结合神经网络、差分序列、整数规划等多种子模型,在逻辑上进行了串联和整合。
- 3. 本模型根据弹性模型拟合定价对实际销售量波动的影响, 充分结合了经济学知识。

10.1.2 缺点

- 1. 本模型缺乏对品类施加给单品蔬菜的影响的考虑。
- 2. 本模型对菜品选择的模型设置的较为简单。
- 3. 本模型将想同品类的商品的价格弹性系数的设置为常数,较为简单,没有考虑时间对其可能产生的影响。

10.2 模型改进

- 1. 可以建立不同蔬菜之间的销量影响矩阵, 充分考虑蔬菜之间同天销量的影响关系, 以此实现第三问更复杂的整数规划求解模型。
- 2. 可以为销售量最终预测值设定一个区间,在批发的时候适当多订购一些。
- 3. 可以将弹性系数进行更复杂的处理,如拟合多项式等,然后取一个相对更优的函数关系。

参考文献

- [1] Andala Barusman. Analysis of implementation cost plus pricing method in the decision on the determination of product sales prices. 03 2020.
- [2] Tijun Fan, Chang Xu, and Feng Tao. Dynamic pricing and replenishment policy for fresh produce. Computers Industrial Engineering, 139:106127, 2020.
- [3] Chris Guilding, Colin Drury, and Mike Tayles. An empirical investigation of the importance of cost-plus pricing. Managerial Auditing Journal, 20(2):125–137, 2005.
- [4] Domenico Piccolo. A distance measure for classifying arima models. <u>Journal of time series</u> analysis, 11(2):153–164, 1990.
- [5] 徐小峰, 余乐安, 林姿汝, and 孙玉萍. 基于特征融合的生鲜商品短期销量组合预测. 管理科学学报, 25(12):102–123, 2022.
- [6] 李敏. 基于价格弹性和行为选择的共享停车动态定价方法研究. Master's thesis, 宁波大学, 2020.

附录 A 支撑材料文件列表

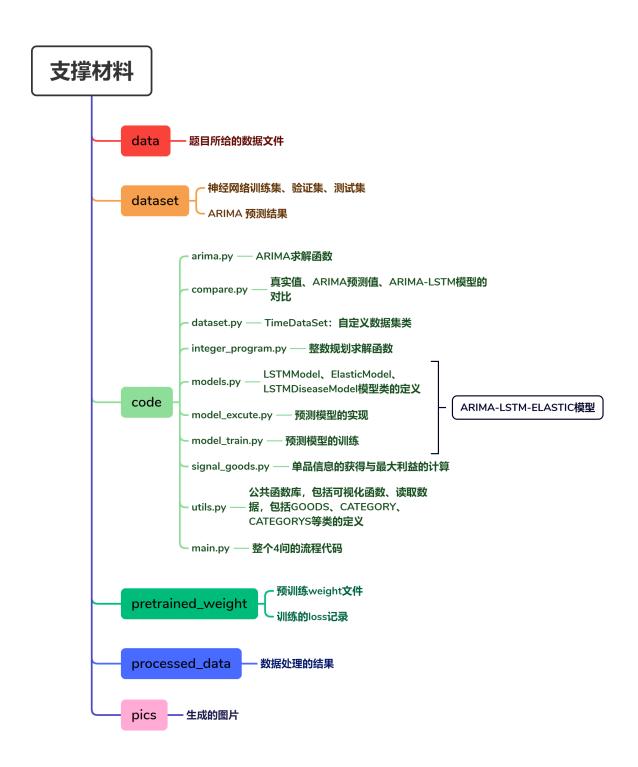


图 21 支撑材料结构图

附录 B 代码

这里只列举出部分关键代码、完整代码见支撑材料。

2.1 LSTMModel

Listing 1: LSTMModel

```
class LSTMModel(nn.Module):
      1stm Model
   def __init__(self, input_channel=18, hidden_channel=18,
              output_channel=18, lstm_hidden=50, lstm_layer=2):
      super(LSTMModel, self).__init__()
      self.input_channel = input_channel
      self.hidden_channel = hidden_channel
      self.output_channel = output_channel
      self.lstm_hidden = lstm_hidden
      self.lstm_layer = lstm_layer
      self.lin = nn.Linear(input_channel, hidden_channel).to(device)
      self.lstm = nn.LSTM(3*hidden_channel, lstm_hidden, lstm_layer,
           batch_first=True).to(device)
      self.output_layer = nn.Linear(lstm_hidden, output_channel).to(device)
      self.time_embed = nn.Embedding(366, 48).to(device)
      self.time_lin = nn.Linear(48, hidden_channel).to(device)
      self.diff_lin = nn.Linear(input_channel, hidden_channel).to(device)
   def forward(self, x:torch.Tensor, corr_x:torch.Tensor,
             t:torch.Tensor, diff_x: torch.Tensor):
      x: (N, 18)
      corr_x: (18, 18)
      t: (N, 1)
      diff_x: (N, 18)
      11 11 11
      x = F.relu(self.lin(torch.matmul(x, corr_x))) # (N, H)
      t = self.time_lin(self.time_embed(t.long())) # (N, H)
      diff_x = self.diff_lin(diff_x)
      y = torch.cat([x, t, diff_x], dim=1)
                                                # (N, 3H)
      # 1stm
      h0 = torch.zeros(self.lstm_layer, 1, self.lstm_hidden).to(x.device)
      c0 = torch.zeros(self.lstm_layer, 1, self.lstm_hidden).to(x.device)
      out, _ = self.lstm(y.unsqueeze(dim=0), (h0, c0))
```

```
out = self.output_layer(out[:, -1, :]).squeeze(dim=0)
return out
```

2.2 ARIMA

Listing 2: ARIMA

```
import pandas as pd
import numpy as np
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.tsa.stattools import adfuller
import warnings
warnings.filterwarnings("ignore")
def get_arima_d(data: pd.DataFrame):
   d = 0
    while True:
       adf_result = adfuller(data["diff_{}".format(d)].dropna())
       p = adf_result[1]
       if p < 0.05:
          return d
       if d >= 2:
          raise ValueError("too many diffs")
       d += 1
       data["diff_{}".format(d)] = data["diff_{}".format(d-1)].diff(1)
def get_arima_p_q(data: pd.DataFrame, arima_d):
    pmax = 2
    qmax = 2
    bic_matrix = []
    np_data = np.nan_to_num(data["diff_0"], nan=0).astype(float)
    for p in range(pmax+1):
       tmp = []
       for q in range(qmax+1):
           try:
              tmp.append(ARIMA(np_data, order=(p, arima_d, q)).fit().bic)
           except:
              tmp.append(None)
       bic_matrix.append(tmp)
    bic_matrix = pd.DataFrame(bic_matrix)
    p, q = bic_matrix.stack().idxmin()
    return p, q
def _arima(data: np.ndarray, predict_size=0):
   0.00
```

```
Predict the future data using ARIMA Modle
      # data process
      df = pd.DataFrame(data=data, columns=["diff_0"])
      # define the d
      d = get_arima_d(df)
      # define the p & q
      p, q = get_arima_p_q(df, d)
      # arima model
      model = ARIMA(data.astype(float), order=(p, d, q)).fit()
      # predict
      fittedvalues = model.fittedvalues
      predict = model.forecast(predict_size+1)
      result = np.concatenate([np.expand_dims(fittedvalues, axis=0),
                   np.expand_dims(predict, axis=0)], axis=1).squeeze(axis=0)
      result = result[1:]
      return result
   def arima(data: np.ndarray, predict_size=0):
      if data.ndim == 1:
          data = np.expand_dims(data, axis=0)
60
      arima_pred = list()
      for i in range(data.shape[1]):
          arima_pred.append(_arima(data[:, i], predict_size))
      arima_pred = np.array(arima_pred)
      return arima_pred
```

2.3 integer program

Listing 3: 整数规划

```
import pulp
import numpy as np
import pandas as pd

def lip():
    """

Linear Integer Programming
    """

# load data

df = pd.read_excel("processed_data/预测收益.xlsx")

data = np.array(df)

damage = data[:, -4]

benifits = data[:, -1]

volumes = data[:, -2]

prices = data[:, -3]
```

```
min_zero = np.empty(shape=(34,))
for i in range(34):
   min_zero[i] = min(volumes[i]-2.5, 0)
fix_benifits = benifits + min_zero * prices
fix_benifits -= (damage/100) * volumes * prices * 0.1
# create problem
problem = pulp.LpProblem("Goods Selected", pulp.LpMaximize)
# target
x = [pulp.LpVariable(f"x_{i}", lowBound=0, upBound=1, cat=pulp.LpInteger)]
    for i in range(34)]
# object
object = pulp.lpSum(x[i] * fix_benifits[i] for i in range(34))
problem += object
# constraint
problem += pulp.lpSum(x[i] for i in range(34)) <= 33</pre>
problem += pulp.lpSum(x[i] for i in range(34)) >= 27
# get result
problem.solve()
result = np.zeros(shape=(34, ))
for i in range(34):
   result[i] = pulp.value(x[i])
df = pd.read_excel("processed_data/预测收益.xlsx")
df["修正收益"] = fix_benifits
df["是否选择"] = result
df.to_excel("processed_data/选择结果.xlsx")
return result
```

附录 C 过滤单品表

表 5 各品类蔬菜不予考虑的单品表

| 蔬菜品类 | 删去的单品 |
|-------|---------------------|
| 花菜类 | 紫白菜(1),紫白菜(2) |
| 花叶类 | 东门口小白菜,南瓜尖,快菜等71种 |
| 辣椒类 | 组合椒系列,红线椒,小皱皮等 29 种 |
| 茄类 | 大龙茄子, 花茄子等6种 |
| 食用菌 | 金针菇(1),海鲜菇(1)等63种 |
| 水生根茎类 | 莲蓬(个),红莲藕带等14种 |

附录 D 花叶类、辣椒类过滤后单品编号

| 上海青 | 菜心 | 云南生菜 | 甜白菜 | 云南油麦菜 | 苋菜 | 菠菜 |
|-----|------|------|-------|-------|-----|-----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 红薯尖 | 竹叶菜 | 黄白菜 | 外地茼蒿 | 小白菜 | 木耳菜 | 娃娃菜 |
| 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 奶白菜 | 蔡甸藜蒿 | 茼蒿 | 枝江红菜苔 | 小青菜 | 黄心菜 | |
| 14 | 15 | 16 | 17 | 18 | 19 | |

表 6 花叶类预处理后编号结果

| 青线椒 | 小米椒 | 螺丝椒 | 芜湖青椒 | 小皱皮 |
|-----|-----|---------------|---------|-----|
| 0 | 1 | 2 | 3 | 4 |
| 红尖椒 | 七彩椒 | 姜蒜小米椒组合装 (小份) | 青红杭椒组合装 | 红椒 |
| 5 | 6 | 7 | 8 | 9 |

表 7 辣椒类预处理后编号结果

附录 E 蔬菜品类销量的时间序列分解图

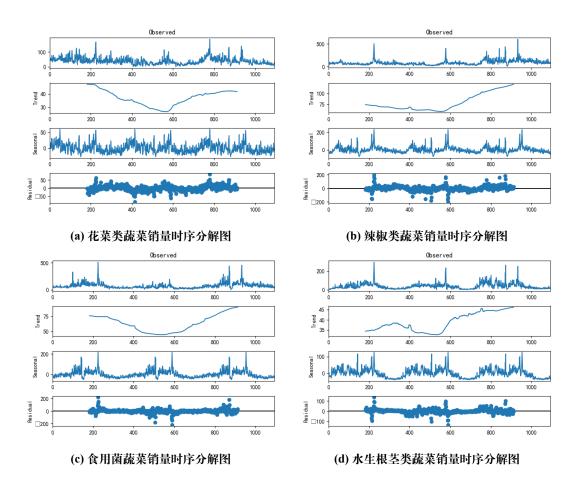
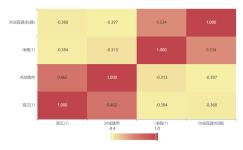
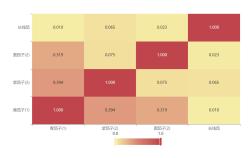


图 22 剩余四类蔬菜品类销量的时间序列分解图

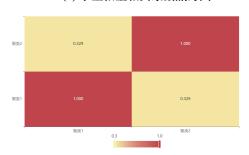
附录 F 剩余品类蔬菜各单品的相关系数热力图



(a) 水生根茎相关系数热力图



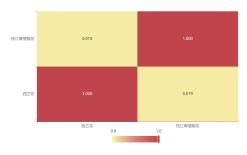
(b) 茄类相关系数热力图



(c) 辣椒相关系数热力图



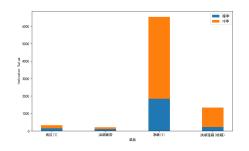
(d) 水生根茎类蔬菜销量时序分解图



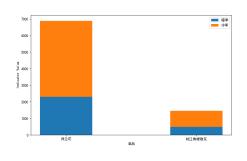
(e) 花菜相关系数热力图

图 23 剩余五类蔬菜品类销量的相关系数热力图

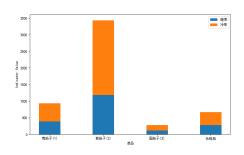
附录 G 剩余品类蔬菜各单品的相关系数热力图



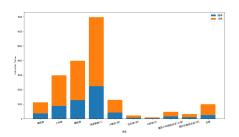
(a) 水生根茎类蔬菜各单品总销量及季节组成



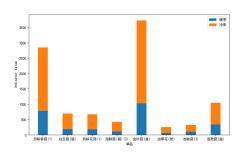
(b) 花菜类蔬菜各单品总销量及季节组成



(c) 茄类蔬菜各单品总销量及季节组成



(d) 辣椒类蔬菜各单品总销量及季节组成



(e) 食用菌各单品总销量及季节组成

图 24 剩余五类各单品总销量及季节组成

附录 H 问题二 LSTM-ARIMA 补货和定价策略

表 8 花叶类补货和定价策略

| 日期 | 补货量 | 定价 | 定价/成本 | 收益 |
|------|----------|----------|----------|----------|
| 7月1日 | 171.7394 | 5.021126 | 1.616885 | 329.0002 |
| 7月2日 | 168.6939 | 4.842233 | 1.473506 | 262.4937 |
| 7月3日 | 150.0436 | 4.657954 | 1.378587 | 191.9305 |
| 7月4日 | 135.7983 | 4.469925 | 1.30338 | 141.2898 |
| 7月5日 | 146.2728 | 4.297279 | 1.250907 | 126.0797 |
| 7月6日 | 145.8826 | 4.131042 | 1.187254 | 95.04978 |
| 7月7日 | 146.6226 | 3.971708 | 1.168966 | 84.17374 |

表9 水生根茎类补货和定价策略

| 日期 | 补货量 | 定价 | 定价/成本 | 收益 |
|------|----------|----------|----------|----------|
| 7月1日 | 21.15289 | 14.18478 | 1.154487 | 40.15098 |
| 7月2日 | 19.24945 | 13.76138 | 1.134113 | 31.32528 |
| 7月3日 | 27.8539 | 13.76138 | 1.112789 | 38.85096 |
| 7月4日 | 16.42205 | 13.76138 | 1.097481 | 20.07305 |
| 7月5日 | 19.61174 | 13.76138 | 1.076247 | 19.12 |
| 7月6日 | 16.90154 | 13.76138 | 1.087184 | 18.6518 |
| 7月7日 | 21.23252 | 13.76138 | 1.09367 | 25.02516 |

表 10 食用菌补货和定价策略

| 日期 | 补货量 | 定价 | 定价/成本 | 收益 |
|------|----------|----------|----------|----------|
| 7月1日 | 60.15094 | 6.370807 | 1.770528 | 166.7718 |
| 7月2日 | 59.91093 | 6.09068 | 1.648912 | 143.6019 |
| 7月3日 | 51.44075 | 5.769335 | 1.549797 | 105.2836 |
| 7月4日 | 43.36055 | 5.42629 | 1.476222 | 75.90242 |
| 7月5日 | 59.91107 | 5.167403 | 1.423832 | 92.15403 |
| 7月6日 | 55.69183 | 4.908879 | 1.421351 | 81.04312 |
| 7月7日 | 54.36715 | 4.659341 | 1.359297 | 66.95769 |

表 11 花菜类补货和定价策略

| 日期 | 补货量 | 定价 | 定价/成本 | 收益 |
|------|----------|----------|----------|----------|
| 7月1日 | 20.50974 | 11.33863 | 1.238742 | 44.81974 |
| 7月2日 | 10.24897 | 10.91221 | 1.260186 | 23.09099 |
| 7月3日 | 16.72362 | 10.59396 | 1.301836 | 41.07742 |
| 7月4日 | 23.64251 | 10.35407 | 1.309955 | 57.92248 |
| 7月5日 | 15.1288 | 10.02121 | 1.291447 | 34.21427 |
| 7月6日 | 20.21115 | 9.765837 | 1.251961 | 39.72304 |
| 7月7日 | 23.43295 | 9.543241 | 1.192618 | 36.1176 |

表 12 茄类补货和定价策略

| 日期 | 补货量 | 定价 | 定价/成本 | 收益 |
|------|----------|----------|----------|----------|
| 7月1日 | 28.07636 | 8.008852 | 1.655971 | 89.07235 |
| 7月2日 | 23.27489 | 7.903068 | 1.681061 | 74.52225 |
| 7月3日 | 24.57484 | 7.802006 | 1.776556 | 83.80905 |
| 7月4日 | 23.29864 | 7.699018 | 1.858737 | 82.87204 |
| 7月5日 | 17.23922 | 7.575764 | 1.650012 | 51.44917 |
| 7月6日 | 22.74519 | 7.474296 | 1.547625 | 60.15582 |
| 7月7日 | 13.5431 | 7.332377 | 1.571345 | 36.10689 |

附录 I 问题三补货和定价策略

表 13 辣椒类单品补货和定价策略

| 单品 | 成本 | 定价 | 补货量 | 收益 |
|--------------|----------|----------|----------|----------|
| 螺丝椒 | 8.97 | 12 | 6.205036 | 18.80126 |
| 芜湖青椒(1) | 3.63 | 4.312712 | 19.32815 | 13.19556 |
| 小米椒(份) | 2.082974 | 4.84498 | 24.69234 | 68.20039 |
| 小皱皮(份) | 2.1 | 2.10629 | 13.14729 | 0.082693 |
| 螺丝椒(份) | 4.29 | 4.452714 | 13.48246 | 2.193786 |
| 姜蒜小米椒组合装(小份) | 2.370378 | 3.420717 | 9.035011 | 9.48983 |
| 青红杭椒组合装(份) | 3.32 | 5.8 | 2.640583 | 6.548646 |
| 红椒 (2) | 12.72 | 18.91667 | 2.054464 | 12.73083 |

表 14 花叶类单品补货和定价策略

| 单品 | 成本 | 定价 | 补货量 | 收益 |
|---------|----------|----------|----------|----------|
| 上海青 | 4.239579 | 5.744762 | 5.658378 | 8.516895 |
| 云南生菜 | 5.350596 | 9.2 | 0.847355 | 3.261812 |
| 苋菜 | 2.21 | 2.798344 | 9.545281 | 5.615905 |
| 红薯尖 | 2.6 | 3.687641 | 6.952566 | 7.561895 |
| 竹叶菜 | 2.15 | 2.979612 | 15.83244 | 13.13478 |
| 木耳菜 | 3.12 | 4.399652 | 7.016334 | 8.978465 |
| 娃娃菜 | 4.4 | 5.784732 | 11.95356 | 16.55247 |
| 奶白菜 | 2.56 | 3.323194 | 10.10688 | 7.713513 |
| 小青菜 (1) | 2.7 | 3.933175 | 6.585132 | 8.120621 |
| 云南油麦菜 | 2.66 | 6.485372 | 23.35763 | 89.35164 |
| 菠菜 | 9.667564 | 12.16804 | 14.49005 | 36.23202 |
| 云南生菜(份) | 3.49 | 5.272247 | 37.75645 | 67.29133 |

表 15 茄类单品补货和定价策略

| 单品 | 成本 | 定价 | 补货量 | 收益 |
|--------|----------|----------|----------|----------|
| 青茄子(1) | 3.98 | 6 | 1.339973 | 2.706746 |
| 紫茄子(2) | 3.341439 | 5.452602 | 14.62051 | 30.86628 |
| 长线茄 | 7 | 8.118829 | 3.621417 | 4.051746 |

表 16 食用菌单品补货和定价策略

| 单品 | 成本 | 定价 | 补货量 | 收益 |
|---------|----------|----------|----------|----------|
| 西峡花菇(1) | 15.6 | 24 | 5.294879 | 44.47698 |
| 海鲜菇 (包) | 1.95 | 2.4 | 8.576086 | 3.859239 |
| 金针菇(盒) | 1.434275 | 1.938462 | 19.44152 | 9.802163 |
| 虫草花(份) | 2.701668 | 3.8 | 3.573939 | 3.925373 |
| 双孢菇 (盒) | 3.504003 | 4.888889 | 14.99507 | 20.76647 |

表 17 水生根茎类单品补货和定价策略

| 单品 | 成本 | 定价 | 补货量 | 收益 |
|--------|----------|----------|----------|----------|
| 高瓜 (1) | 11.67 | 10.56085 | 3.601807 | -3.99496 |
| 洪湖藕带 | 18 | 18.54874 | 5.077094 | 2.786025 |
| 净藕(1) | 10.31671 | 10.84323 | 5.674167 | 2.987552 |
| 菱角 | 9.16 | 14 | 1.946668 | 9.421875 |

附录 J 问题三规划前后收益对比表

表 18 花菜类规划前后收益对比

| 单品 | 原预测收益 | 修正收益 | 是否选择 |
|--------|----------|----------|------|
| 西兰花 | 54.45631 | 52.52452 | 是 |
| 枝江青梗散花 | 2.276796 | 1.855356 | 是 |

表 19 花叶类规划前后收益对比

| 单品 | 原预测收益 | 修正收益 | 是否选择 |
|---------|----------|----------|------|
| 上海青 | 9.150579 | 8.672373 | 是 |
| 云南生菜 | 3.261812 | -12.0614 | 否 |
| 苋菜 | 5.901089 | 5.40112 | 是 |
| 红薯尖 | 7.945938 | 7.726827 | 是 |
| 竹叶菜 | 13.46529 | 12.81828 | 是 |
| 木耳菜 | 9.435903 | 9.197505 | 是 |
| 娃娃菜 | 17.15512 | 16.98214 | 是 |
| 奶白菜 | 8.053193 | 7.521221 | 是 |
| 小青菜 (1) | 8.534878 | 8.263047 | 是 |
| 云南油麦菜 | 90.18053 | 88.22941 | 是 |
| 菠菜 | 37.5502 | 34.2622 | 是 |
| 云南生菜(份) | 68.15583 | 66.27052 | 是 |

表 20 辣椒类规划前后收益对比

| 単品 | 原预测收益 | 修正收益 | 是否选择 |
|--------------|----------|----------|------|
| 螺丝椒 | 18.80126 | 18.04325 | 是 |
| 芜湖青椒(1) | 13.86992 | 13.39094 | 是 |
| 小米椒 (份) | 69.03058 | 67.8946 | 是 |
| 小皱皮(份) | 0.38446 | 0.120479 | 是 |
| 螺丝椒 (份) | 2.833843 | 2.261691 | 是 |
| 姜蒜小米椒组合装(小份) | 9.979858 | 9.683791 | 是 |
| 青红杭椒组合装(份) | 6.548646 | 6.404222 | 是 |
| 红椒 (2) | 12.73083 | 3.936282 | 是 |

表 21 茄类规划前后收益对比

| 单品 | 原预测收益 | 修正收益 | 是否选择 |
|---------|----------|----------|------|
| 青茄子(1) | 2.706746 | -4.29369 | 否 |
| 紫茄子 (2) | 29.88611 | 29.40816 | 是 |
| 长线茄 | 2.33037 | 2.139376 | 是 |

表 22 食用菌规划前后收益对比

| 单品 | 原预测收益 | 修正收益 | 是否选择 |
|---------|----------|----------|------|
| 西峡花菇(1) | 44.47698 | 43.10455 | 是 |
| 海鲜菇 (包) | 3.859239 | 3.859239 | 是 |
| 金针菇(盒) | 0.172987 | 0.159088 | 是 |
| 虫草花(份) | 3.925373 | 3.797304 | 是 |
| 双孢菇 (盒) | -0.9119 | -0.92345 | 否 |