

基于 ARIMA 模型的蔬菜类商品的自动定价与补货决策研究

夏孟瑶

(云南大学, 云南 昆明 650091)

摘要: 随着人工智能时代的到来, 餐饮行业面临成本挑战、菜品浪费和供应短缺等问题。为了应对这些挑战, 传统的销售管理方式已不够高效。因此, 企业需要采用信息化的管理模式来提高管理效率。本文基于 ARIMA 模型, 针对蔬菜品类分别建立工作日与周末的预测模型, 以有效预测蔬菜短期销量。同时, 通过皮尔逊相关性分析, 推荐替补菜品, 增加销售蔬菜的多样性。最后, 对商超提出相关数据收集意见, 以帮助其提升市场竞争力。

关键词: ARIMA 时间序列预测模型; 蔬菜销售量预测; 推荐策略

一、研究意义

蔬菜行业在中国发展迅速, 商户面临着机遇和挑战。商户必须不断变革自身的产业链, 打破传统的蔬菜营销理念, 紧跟时代的发展步伐。随着人们对蔬菜需求品质、安全、优质的转变, 商家需要控制蔬菜的进货量以保证蔬菜新鲜度。因此, 蔬菜销售预测变得越来越重要。有效的销售预测可以避免蔬菜滞销和折价处理, 确保消费者能够购买到所需的蔬菜品种。因此对于市场需求的严谨分析于蔬菜类商品的补货决策及定价决策尤为重要, 从需求侧及供给侧来看合理的销售组合与时间对于商品的销售量均会产生影响。文章基于 2023 年全国大学生数学建模竞赛 C 题所提供的数据进行研究, 建立了较为可靠的预测模型, 通过对历史蔬菜销售量的分析来对未来短期的蔬菜销售量进行预测, 并进行合理性和鲁棒性检验。

二、数据预处理

因蔬菜销售数据在销售单价上存在量纲差异, 故本文使用 Min-Max 标准化将不同特征的数值统一到 $[0, 1]$ 区间内, 能在保留原有数据之间关系的同时, 消除数据取值的影响。其次对 6 种蔬菜品类、252 种蔬菜单品进行重新编码, 简化分析处理过程。

三、聚类分析

对蔬菜商品的自动定价与补货决策需要考虑到多个因素, 如蔬菜的品种、产地、季节性、销售情况、供需关系等等。而聚类分析可以帮助我们将相似的蔬菜商品进行分类, 进而更好地理解它们之间的关系和差异。然后, 对每个群组中的商品进行进一步分析, 以确定最佳的定价和补货决策。对于每个群组, 我们可以考虑其销售历史数据、当前市场供需情况、成本和利润等因素, 来制定最佳的定价和补货策略。依据季度, 分析六类蔬菜品类的均值、最大值、最小值、标准差、变异系数、偏度系数、峰度系数描述统计量, 进行聚类分析。通过数据分析可知, 在第二季度, 蔬菜销售量极值相差较大, 方差较大, 峰度为四个季度中的最大值, 说明蔬菜销售量的波动更大, 商家在该季度更需要注重销售量的预测, 其中第一季度与第二季度的平均销售量最大, 说明在以上两个季度, 蔬菜的需求量较大。本文按单品总销售量将数据分为三类, 如水果辣椒(橙色)、活体银耳、芥兰等 213 种总销售量低的蔬菜分为第一类, 甜白菜、牛首油菜、娃娃菜等 30 种总销售量中等的蔬菜分为第二类, 大白菜、云南生菜、芜湖青椒(1)等 8 种总销售量高的分为第三类。对于不同蔬菜单品, 依据日总销售量进行层次聚类, 可分为家常日用菜, 炒菜类, 配菜类; 对于蔬菜单品和品类销售量, 利用数据可视化, 发现其随季节有明显周期性变化, 通过傅里叶函数拟合, 家常日用菜的销量变换周期为 87.7193 天, 炒菜类的销量变换周期为 232.5581 天, 配菜类

的销量变换周期为 172.4138 天, 商家可依靠此周期来推测销售者的需求迭代。

四、回归分析

了解到成本加成定价是指在产品成本基础上设定的价格加上一定的利润空间, 根据相关数据, 处理批发价格及销售价格数据, 得出各单品的销售总量及成本加成率, 再通过回归分析, 生成各蔬菜品类的销售总量与成本加成率的线性回归方程, 其后根据方程进行关系分析。其中横坐标表示销售总量, 纵坐标表示平均定价, 根据图形发现销售总量为 0 时仍存在平均定价, 为了防止此类情况影响回归函数, 便选择排除此类情况后进行分析。R 方表示样本数据中因变量的变异程度可以由自变量解释的比例, 且 R 方的取值范围 0 到 1, 值越接近 1 表示模型拟合程度越好。RMSE 是残差平方和的平方根, 表示模型预测值与实际观测值之间的平均差异, 且 RMSE 越小表示模型的预测效果越好。由此可得该模型的拟合程度及预测效果良好, 可根据函数分析其相关关系。可知, 不同品类拟合曲线呈单调递减趋势, 结合函数可得其斜率近似为 -0.01 , 则各蔬菜品类的销售总量与成本加成率呈负相关, 但并不显著。

五、ARIMA 时间预测

筛选不同蔬菜品类的日定价可知, 不同蔬菜定价变化较小, 短时间内未有大幅度变化, 影响商超收入的仅有销售量, 故蔬菜定价采用过往销售单价的众数。对蔬菜数据进行差分, 发现其为平稳数据, 适用于 ARIMA 模型。ARIMA(自回归移动平均模型)是一种经典的时间序列预测模型。它结合了自回归(AR)和移动平均(MA)的特性, 可以用于预测具有一定规律性的时间序列数据。发现不同品类蔬菜, 工作日与周末的销售情况, 差别较大, 故分为两类情况进行拟合。得到日补货总量策略后, 则可结合过往数据对蔬菜进行定价。以不同蔬菜品类中的代表叶类以莴菜为例, 花菜类以西兰花为例, 水生根茎类以荸荠为例, 茄类以青茄子(1)为例, 辣椒类以红尖椒为例, 食用菌类以白玉菇(1)为例。

六、寻找替补菜品

通过对日销售量做皮尔逊相关性分析, 筛选出与可出售商品具有强相关性的蔬菜单品, 作为 33 种蔬菜单品的替补商品, 以增加市场蔬菜种类, 满足市场需求。

七、意见的提出与理由

(一) 天气数据

1. 意见

天气数据为未来时间段天气数据, 可通过爬虫获取, 主要包括气温、风力大小、天气等。可大致分为三类, 如多云/晴天、阴天/小雨、大雨。

2. 理由

针对补货决策：(1) 预测客流量：天气数据可从侧面反映出客流量情况，例如雨天时前往商超购买蔬菜的人数减少，从而导致蔬菜销售量降低。因此商超可根据天气数据预估当天的客流量从而做出对应的蔬菜补货决策。(2) 判断蔬菜的质量和保鲜情况：如高温天气容易导致蔬菜变质腐烂，而潮湿的天气则会增加蔬菜霉菌和细菌的滋生，影响蔬菜的保鲜期和口感。因此，商超需要密切关注天气情况，及时调整蔬菜的供货和贮存方式，以确保蔬菜的品质和供应能够满足消费者的需求。

针对定价决策：了解蔬菜的供需关系：例如，在气温较高的夏季，市场上的蔬菜供应通常较大，价格可能较为稳定或下跌；而在气温较低的冬季，蔬菜供应相对较少，价格可能会上涨。商超可以利用天气数据了解到不同季节的市场供需情况，合理制定蔬菜的定价策略，以提高销售效益。

(二) 供应链数据

1. 意见

收集各个蔬菜商品供应链环节的数据，包括供应商数据、采购信息、物流运输数据等。

2. 理由

针对补货决策：(1) 可帮助商超了解蔬菜的销售情况和库存水平：如果某种蔬菜的销售量较大，并且库存水平较低，商超可以根据数据判断是否需要补货，以保证货架上的供应充足，同时避免过多的库存。此外，供应链数据还可以分析蔬菜的销售趋势和季节性需求，帮助商超预测未来的蔬菜需求，进而调整补货计划。(2) 确保蔬菜商品的及时补货：通过分析供应链数据，商超可以了解每个蔬菜商品的供货情况，判断供应商的供货能力和稳定性，进而优化供应链管理，确保蔬菜商品的及时补货。

针对定价决策：可提供有关蔬菜的市场需求和价格趋势等信息：商超可以通过分析数据了解市场对蔬菜的需求程度和价格敏感度，进而制定合理的定价策略。当供应链数据显示某种蔬菜的需求量较大，但市场价格较高时，商超可以适当提高售价，以提高利润；反之，如果市场竞争激烈且价格压力较大，商超可以降低售价以吸引消费者增加销量。

(三) 消费者偏好数据

1. 意见

商超可以了解消费者对不同种类蔬菜的偏好及消费者对有机、无农药、本地特产等特殊属性的偏好。

2. 理由

针对补货决策：优化补货决策：消费者喜好数据可以帮助商超了解消费者对不同种类蔬菜的需求量和偏好程度。商超可以收集并分析这些数据，以便更好地调整库存，选择合适的蔬菜品种和数量进行补货。如果消费者对某个特定种类的蔬菜有较高的偏好，商超可以相应地增加该种类的补货量，以满足消费者的需求，提高销售额。

针对定价决策：区别定价：如果消费者对某些高质量或有机蔬菜有较高的偏好，商超可以考虑提高这些产品的价格，因为消费者可能愿意为更好的品质支付更高的价格。另一方面，如果消费者对一些经济型蔬菜有较高的偏好，商超可以保持这些产品的较低价格，以吸引更多的消费者购买。

(四) 补充说明

除上诉原因补货决策及定价决策还受到节假日（法定节假日/非节假日）、星期（周末/非周末）、季节（春/夏/秋/冬）、消费者购买习惯等因素的影响，但其均可通过附件数据得出结论，例与日期相关的分析可对附件二销售日期进行分类，其后结合销

售数据分析；与购买习惯相关的分析可对不同蔬菜间进行关联分析，提供组合促销的思路，故不对此类因素进行相关说明。

(五) 结论与建议

随着人工智能时代的到来，餐饮行业面临着成本挑战以及菜品浪费和供应短缺等问题。为了应对这些挑战，企业需要采用信息化、自动化管理模式，以提高管理效率和准确性。

1. 预测销售量

一种有效的管理模式是建立 ARIMA 模型进行销售预测。通过分别建立不同蔬菜品类在工作日和周末的销量预测模型，可以有效地预测短期销量。此外，还可以利用皮尔逊相关性分析，找出与可出售商品具有强相关性的蔬菜单品，作为替补商品，以增加市场蔬菜种类，满足市场需求。

2. 收集多维数据

同时，商超应该收集相关数据来支持销售管理决策。通过收集和分析销售数据，可以了解蔬菜销售量的变化趋势和周期性，例如家常日用菜、炒菜类和配菜类的销量变化周期。商家可以根据这些周期性变化来推测消费者的需求迭代，从而合理安排进货和销售策略。

3. 合理定价

此外，根据不同蔬菜品类的日定价情况，可以发现其定价变化较小。因此，在短时间内可以采用过往销售单价的众数来确定蔬菜的定价。重要的是要意识到销售量对商超收入的影响更大，因此定价策略应重点考虑销售量的变化。

在数据分析中，还发现不同品类蔬菜在工作日和周末的销售情况存在差异。因此，可以将销售数据分为两类进行拟合，以更好地理解不同情况下的销售量变化。通过皮尔逊相关性分析，可以筛选出与可出售商品具有强相关性的蔬菜单品，作为替补商品，以增加市场蔬菜种类，满足市场需求。

4. 总结

综上所述，基于信息化、自动化管理模式，并结合 ARIMA 模型、皮尔逊相关性分析以及数据可视化等方法，可以有效提高餐饮行业的管理效率和准确性。商家应该利用数据分析来预测销售量，合理安排进货和销售策略，并根据销售数据对蔬菜定价进行调整。此外，商家还可以通过筛选替补商品，增加市场蔬菜种类，满足消费者需求。这些措施将有助于提高顾客满意度和商超的市场竞争力。

参考文献：

- [1] 王君霞. 基于易腐品分组策略的联合补货研究综述 [J]. 中国储运, 2023 (08): 97-98.
- [2] 张鹤. 基于集成学习的蔬菜销售预测 [D]. 云南师范大学, 2021.
- [3] 杨盼盼, 赵慧峰. 基于 ARMA 模型河北省羊肉产量分析与预测 [J]. 北方牧业, 2023 (08): 8-9.
- [4] 康莎. 不同需求特征下生鲜农产品双渠道销售定价与库存补货联合决策研究 [D]. 重庆交通大学, 2023.
- [5] 欧燕锋. 具有时变变质率的易腐品联合补货和成本分摊策略研究 [D]. 东南大学, 2022.D
- [6] 李亚丽. 考虑保鲜技术投资的联合补货与定价协同决策 [D]. 重庆交通大学, 2023.

作者简介：夏孟瑶，云南大学数学与统计学院本科生，研究领域为统计学。