

蔬菜类商品的自动定价与补货决策

摘要：本研究通过分析某商超 3 年内的销售、批发和损耗数据，通过搭建多种数学模型，分析各种蔬菜销量和单价等关键因素之间的关系，利用迭代优化算法求出特定时间条件下的最优销售单价与进货量，使得商超总利润达到最大。该研究对于各种商业模式下的数据分析和营销策略制定等都具有一定参考价值。

关键词：自动定价与补货决策，数据分析，数学建模

如需word论文和完整源代码
淘宝店铺：亦是远方

1. 问题分析

1.1. 研究背景

在生鲜超市领域，通常情况下，蔬菜类商品的保鲜期相对较短，其品质随着销售时间的延长逐渐下降。大多数蔬菜品种，如果当天未售出，第二天通常就无法再销售。因此，生鲜超市通常会根据历史销售数据和需求情况，每日进行商品的补货。由于涉及多种蔬菜品类和其他因素，蔬菜的进货通常在凌晨3:00-4:00进行。因此，商家需要在不确切知晓具体单品和进货价格的情况下，做出当天各种蔬菜品类的补货决策。

蔬菜的定价通常采用“成本加成定价”方法，对于运输损耗和品质下降的商品，生鲜超市通常会进行折扣销售。从需求角度来看，蔬菜类商品的销售量通常与时间存在一定的关联。从供应角度来看，4月至10月期间蔬菜供应较为充裕，但由于商超的销售空间有限，因此合理的销售组合变得尤为重要。

本研究基于上述内容和给定的某商超相关销售数据，通过数据处理、分析等方式得出结论，以解决提出的各个问题。

1.2. 研究内容

1. 蔬菜类商品不同品类或不同单品之间可能存在一定的关联关系，分析蔬菜各品类及单品销售量的分布规律及相互关系。

2. 考虑商超以品类为单位做补货计划，分析各蔬菜品类的销售总量与成本加成定价的关系，并给出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. 因蔬菜类商品的销售空间有限，商超希望进一步制定单品的补货计划，要求可售单品总数控制在2733个，且各单品订购量满足最小陈列量2.5千克的要求。根据2023年6月24-30日的可售品种，给出7月1日的单品补货量和定价策略，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。

4. 为了更好地制定蔬菜商品的补货和定价决策，给出商超还需要采集的相关数据，这些数据对解决上述问题的帮助，并给出意见和理由。

1.3. 研究方法

问题一聚焦于附件一，要求分析蔬菜各品类及单品销售量的分布规律及相互关系。要解决这一问题，即需要利用时序分布拟合方法对附件一表格数据进行归纳分类与分析，提炼出蔬菜各品类及单品销售量的分布规律等，以得出相关结论。

问题二至三聚焦于附件二、三和四，要求根据给出的一定时间段内的销售数据探究各蔬菜品类的销售总量与成本加成定价的关系，并预测未来一段时间内可获取最佳销售量的补货与定价策略。此类问题包括关系探究与数据预测两个部分。要探究各蔬菜品类的销售总量与成本加成定价的关系，分析方法与问题一近似，可采用时序分布拟合。要预测未来一段时间内可获取最佳销售量的补货与定价策略，首先要为给定的销售数据找到合适的数学模型进行拟合。直观看来，蔬菜销售情况通常与季节等时间因素有关，故我们搭建高斯过程模型(GPR)、决策树模型(DTR)和神经网络模型(NN)等多个模型进行预拟合，并计算各类参数和结果探究它们的拟合性能，最终选出拟合性能最佳的模型进行预测，以得出相关结论。此外，我们还训练了时间序列 ARIMA 神经网络模型^[1]学习给定的销售数据，探究其拟合性能并进行预测。

问题四为开放类问题，要求在已有的分析基础上，探究还可以收集哪些相关数据以优化补货和定价策略。通过查找有关资料并总结方法和规律，我们可以得到针对该问题的答案。

2. 数据处理

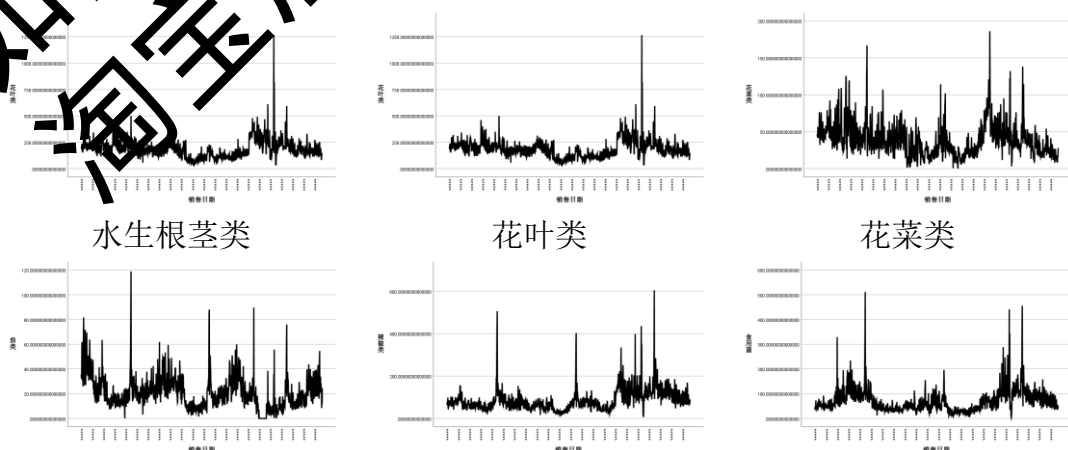
附件蔬菜品类、销售情况与损耗率等数据为 .xlsx 格式，我们编写 Python 和 R 脚本对数据进行预处理，包括单品编码对应等步骤。

3. 研究结果

3.1. 蔬菜各品类及单品销售量的分布规律及相互关系

3.1.1. 蔬菜各品类销售量的分布规律

2016 年内，各蔬菜品类的销售量时序分布图如下：



茄类

辣椒类

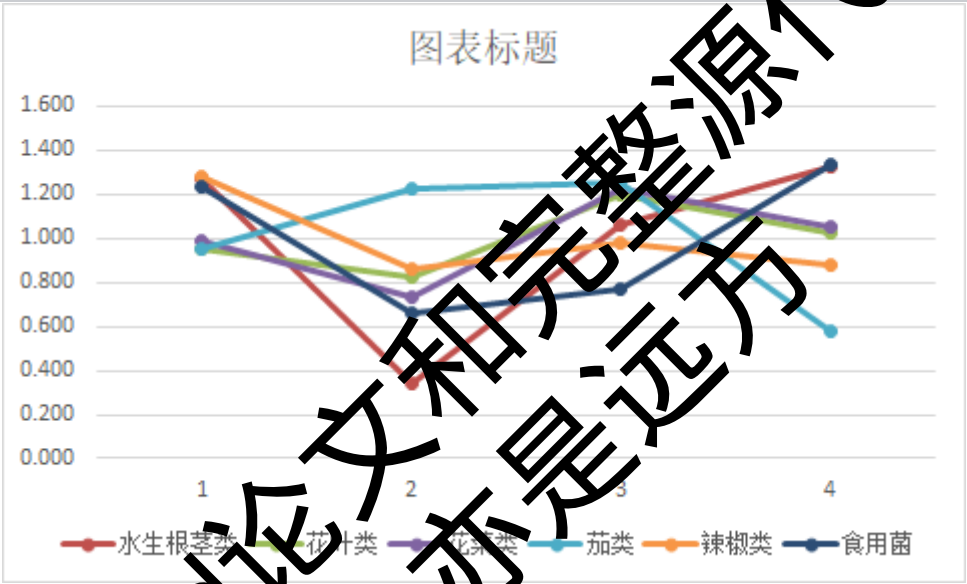
食用菌

为了解各品类蔬菜在时序上的分布规律，我们定义尺度系数的计算如下：

尺度系数 = 指定尺度的平均销售单价 / 平均销售单价

3.1.1.1. 以季度为尺度

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|-------|-------|-------|-------|-------|-------|
| 1 | 1.270 | 0.952 | 0.989 | 0.949 | 1.280 | 1.237 |
| 2 | 0.340 | 0.820 | 0.736 | 1.221 | 0.864 | 0.639 |
| 3 | 1.065 | 1.201 | 1.223 | 1.257 | 0.976 | 0.856 |
| 4 | 1.326 | 1.026 | 1.052 | 0.574 | 0.879 | 1.339 |

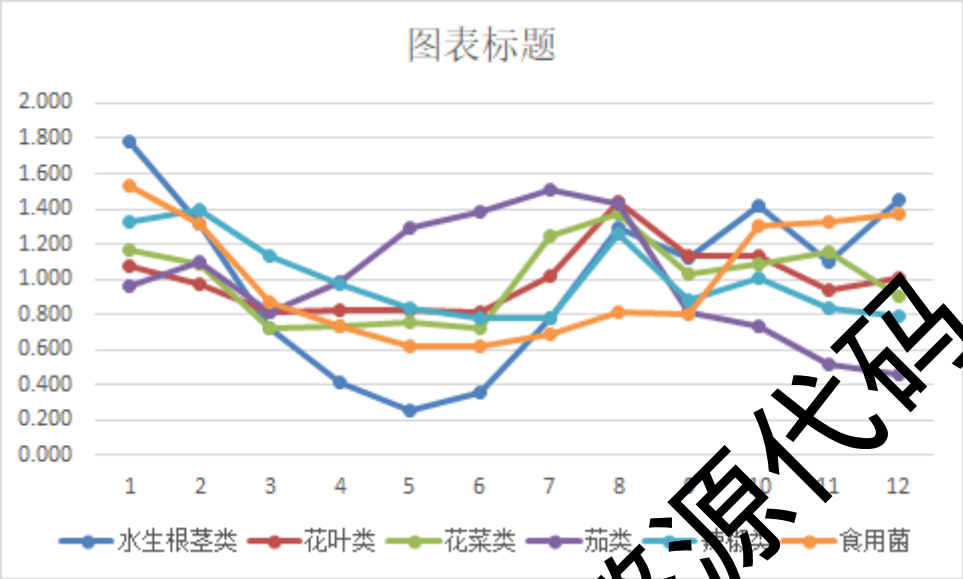


观察数据得知，除茄类以外的蔬菜的销售量均在第二季度达到最小值，其他季度略有浮动。

3.1.1.2. 以月份为尺度

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|-------|-------|-------|-------|-------|-------|
| 1 | 1.783 | 1.074 | 1.171 | 0.959 | 1.325 | 1.532 |
| 2 | 0.327 | 0.972 | 1.088 | 1.094 | 1.396 | 1.314 |
| 3 | 0.721 | 0.818 | 0.720 | 0.812 | 1.132 | 0.874 |
| 4 | 0.413 | 0.824 | 0.731 | 0.985 | 0.971 | 0.739 |
| 5 | 0.254 | 0.830 | 0.755 | 1.292 | 0.841 | 0.619 |
| 6 | 0.355 | 0.808 | 0.720 | 1.384 | 0.780 | 0.617 |
| 7 | 0.785 | 1.022 | 1.248 | 1.513 | 0.780 | 0.689 |
| 8 | 1.291 | 1.442 | 1.376 | 1.427 | 1.257 | 0.808 |
| 9 | 1.116 | 1.138 | 1.035 | 0.818 | 0.887 | 0.798 |
| 10 | 1.414 | 1.130 | 1.090 | 0.730 | 1.008 | 1.303 |
| 11 | 1.099 | 0.939 | 1.159 | 0.522 | 0.837 | 1.331 |

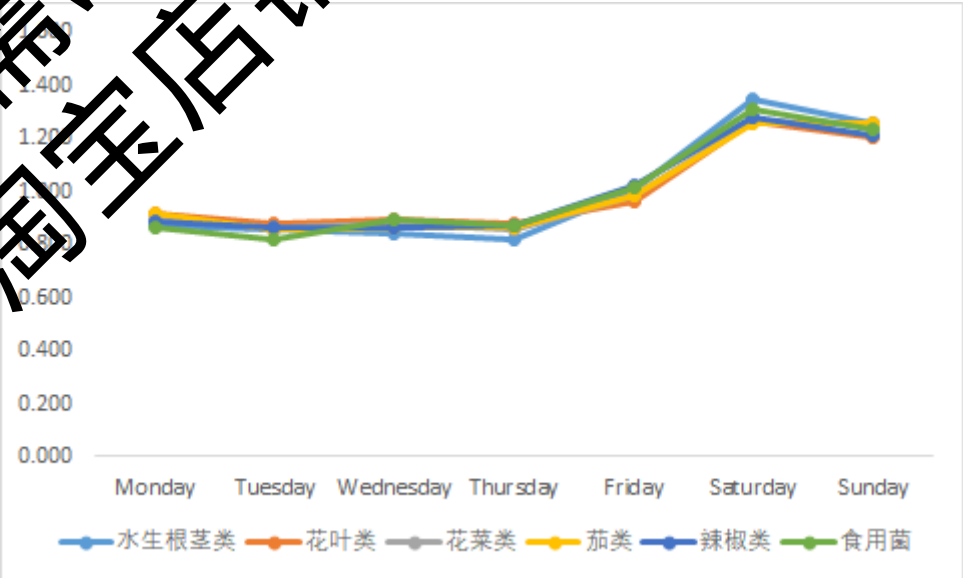
| | | | | | | |
|----|-------|-------|-------|-------|-------|-------|
| 12 | 1.450 | 1.004 | 0.907 | 0.464 | 0.786 | 1.377 |
|----|-------|-------|-------|-------|-------|-------|



观察数据得知,大部分蔬菜的销量在1到6月逐渐递减,7到8月显著增加,9到12月略有波动。而茄类的销量在1到4月略有波动,5到7月显著增加,8到12月迅速下降。

3.1.1.3. 以星期为尺度

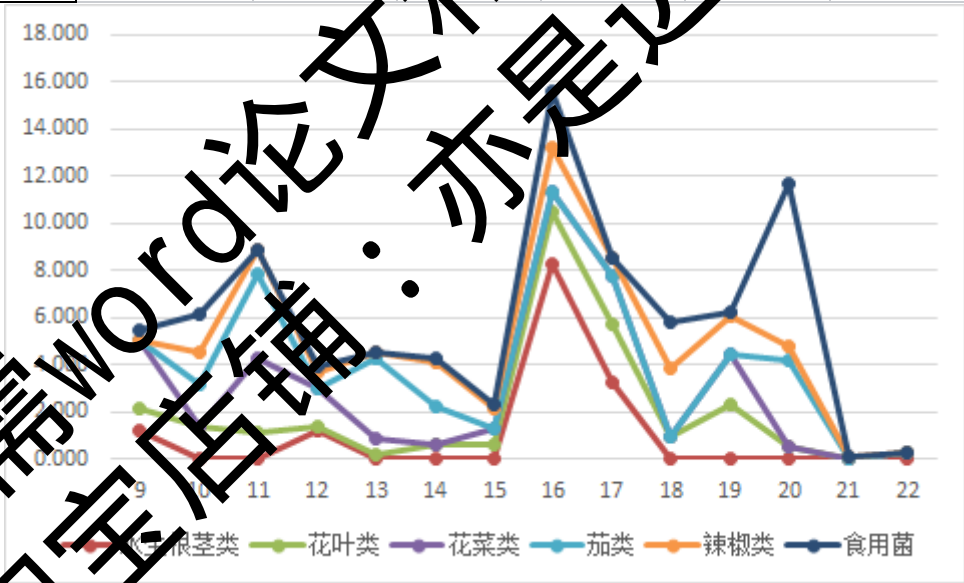
| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|-------|-------|-------|-------|-------|-------|
| 周一 | 0.872 | 0.919 | 0.897 | 0.918 | 0.888 | 0.867 |
| 周二 | 0.854 | 0.877 | 0.862 | 0.856 | 0.862 | 0.816 |
| 周三 | 0.845 | 0.896 | 0.885 | 0.866 | 0.866 | 0.892 |
| 周四 | 0.817 | 0.876 | 0.856 | 0.864 | 0.869 | 0.871 |
| 周五 | 1.007 | 0.962 | 0.983 | 0.985 | 1.023 | 1.016 |
| 周六 | 1.347 | 1.208 | 1.265 | 1.255 | 1.279 | 1.306 |
| 周日 | 1.258 | 1.200 | 1.256 | 1.254 | 1.214 | 1.232 |



观察数据得知，各品类蔬菜在周五到周日的销售量明显高于周一到周四，并在周六达到最大值。

3.1.1.4. 以小时为尺度

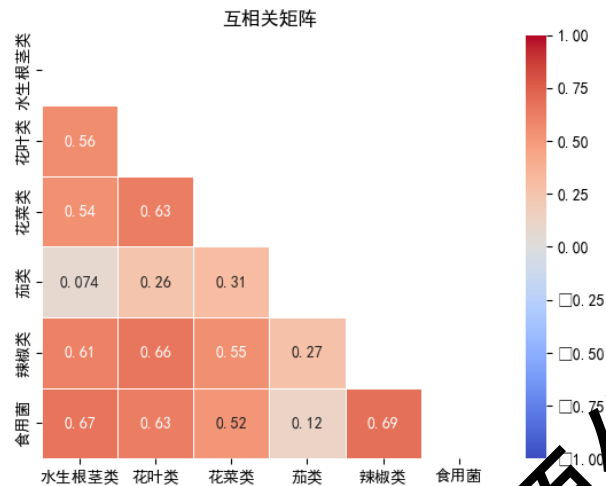
| 小时 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|----|-------|-------|-------|-------|-------|-------|
| 9 | 0.439 | 3.000 | 1.293 | 0.000 | 0.000 | 0.290 |
| 10 | 0.000 | 4.392 | 0.000 | 0.331 | 2.100 | 1.222 |
| 11 | 0.000 | 3.651 | 1.369 | 0.661 | 1.661 | 0.000 |
| 12 | 0.462 | 0.377 | 0.735 | 0.000 | 1.000 | 0.000 |
| 13 | 0.000 | 0.540 | 0.314 | 0.638 | 0.395 | 0.000 |
| 14 | 0.000 | 2.106 | 0.000 | 0.289 | 3.006 | 0.125 |
| 15 | 0.000 | 2.000 | 0.294 | 0.000 | 1.384 | 0.156 |
| 16 | 3.000 | 7.000 | 0.361 | 0.000 | 2.936 | 1.841 |
| 17 | 1.193 | 7.565 | 0.911 | 0.000 | 1.215 | 0.000 |
| 18 | 0.000 | 3.002 | 0.000 | 0.000 | 4.562 | 1.492 |
| 19 | 0.000 | 7.386 | 0.922 | 0.000 | 2.803 | 0.134 |
| 20 | 0.000 | 1.631 | 0.000 | 0.378 | 1.900 | 5.181 |
| 21 | 0.000 | 0.000 | 1.000 | 0.000 | 0.227 | 0.000 |
| 22 | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 |



观察数据得知，该地区居民习惯于在上午 11 时、下午 16 时、晚上 20 时购买蔬菜。晚上 20 时，部分蔬菜由于新鲜程度下降从而打折销售，通常导致更多人选择购买。

3.1.2. 蔬菜各品类销售量的相互关系

为了解蔬菜各品类销售量之间的相互关系，考虑到销售量作为连续变量，我们选择使用 Pearson 相关系数来衡量不同蔬菜品类之间的相关性，绘制相关系数矩阵如下图：



相关系数负值表示负相关，正值表示正相关。相关系数绝对值处在 0.8 到 1 为极强相关,0.6 到 0.8 为强相关，0.4 到 0.6 为中等相关，0.2 到 0.4 为弱相关，0 到 0.2 为极弱相关^[2]。

3.1.3. 蔬菜各单品销售量的分布规律

与蔬菜品类同理，但考虑到蔬菜品类一共有 152 种，限于篇幅问题，我们将结果以 .xlsx 文件的形式置于附件中。

3.1.4. 蔬菜各单品销售量的相互关系

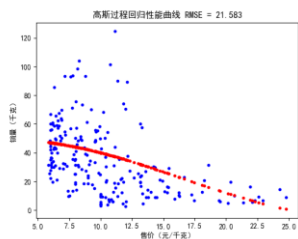
与蔬菜品类同理，但限于篇幅，我们将结果以 .xlsx 文件的形式置于附件中。

3.2. 各蔬菜品类的销售总量与成本加成定价的关系

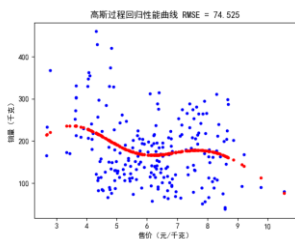
为了解各蔬菜品类的销售总量与成本加成定价的关系，考虑到销售量与成本加成定价均为连续变量，我们针对水生根茎类、花叶类、花菜类、茄类、辣椒类和食用菌这 6 种蔬菜品类，分别搭建高斯过程模型(GPR)、决策树模型(DT)和神经网络模型(NN)，并分别保存为.pkl 文件, .joblib 文件和.h5 文件，以备后续调用。

为确保数据集统一，我们规定采样种子为 42，80%为训练集，20%为验证集。为确保每种模型都能达到各自的最佳表现，我们对其进行多次调参优化。训练完成后，以均方误差(RMSE)为参考指标，将模型预测结果进行比较，其值越小拟合效果越好。经过比较后，从中选出效果最佳的模型作为最终结果。

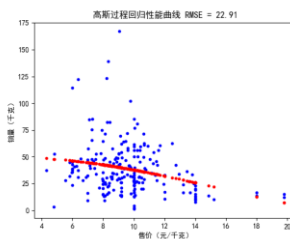
3.2.1. 高斯过程模型预测结果



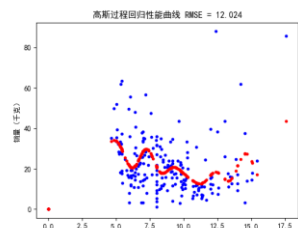
水生根茎类



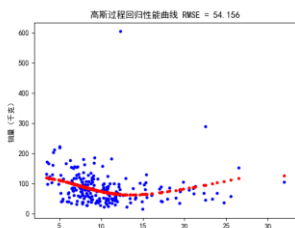
花叶类



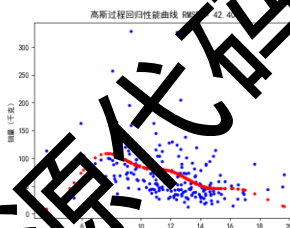
花菜类



茄类

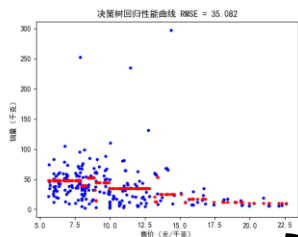


辣椒类

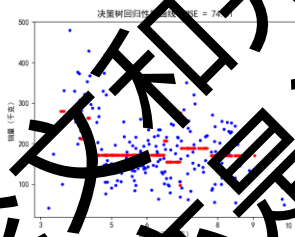


食用菌

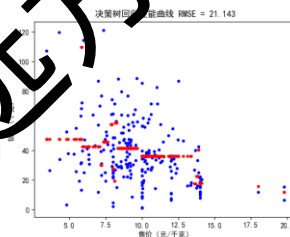
3.2.2. 决策树模型预测结果



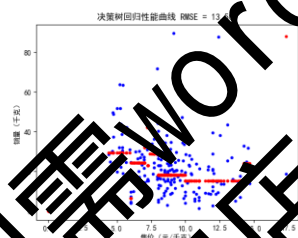
水生根茎类



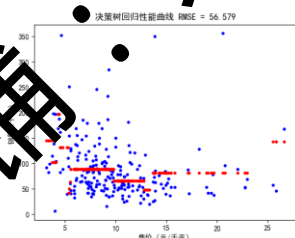
花叶类



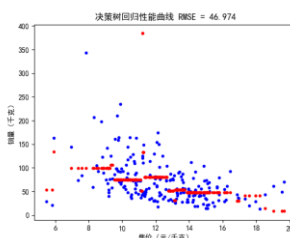
花菜类



茄类

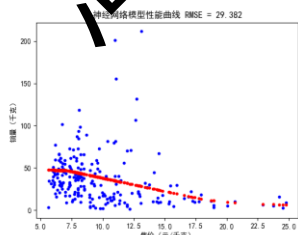


辣椒类

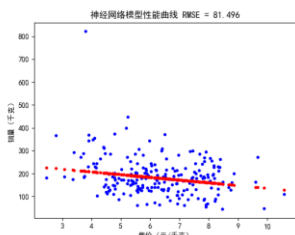


食用菌

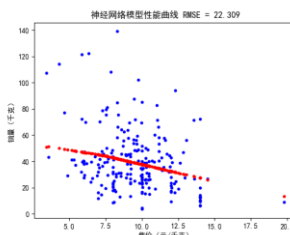
3.2.3. 神经网络模型预测结果



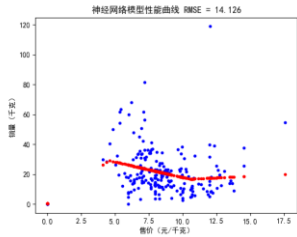
水生根茎类



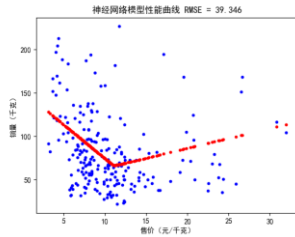
花叶类



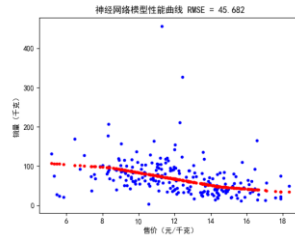
花菜类



茄类



辣椒类



食用菌

最终，3 种模型在 6 种蔬菜品类上拟合表现如下，其中在每种蔬菜品类上表现最佳的模型被加粗标红：

| 均方误差 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|--------|---------------|---------------|---------------|---------------|---------------|---------------|
| 高斯回归模型 | 21.583 | 74.525 | 22.910 | 12.024 | 14.556 | 42.404 |
| 决策树模型 | 35.082 | 74.010 | 21.143 | 14.500 | 56.579 | 46.974 |
| 神经网络模型 | 29.382 | 81.496 | 22.309 | 14.126 | 39.346 | 45.682 |

从表现上看，高斯回归模型效果最好，决策树模型次之，神经网络最差，因此我们最终选择高斯回归模型。

3.3. 各蔬菜品类未来一周的日补货总量和定价策略

为了使得商超未来一周的收益达到最大，我们构建每种蔬菜品类收益与销售单价之间的模型如下：

$$\text{总利润} = \text{销售单价} * \text{预测销售量} - \text{批发价格} * \text{补货总量}$$

其中，预测销售量与销售单价和星期有关，批发价格和星期有关，补货总量与销售单价有关。

1. 简化模型，我们做出以下假设：

1. 预测销售量与星期的关系如下：

$$\text{预测销售量} = \text{星期系数} * \text{平均销售量}$$

其中，平均销售量由高斯过程模型得到，星期系数在之前已经求出。

2. 批发价格与星期的关系如下：

$$\text{批发价格} = \text{指定星期的平均批发价格} / \text{平均批发价格}$$

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|-------|-------|-------|-------|-------|-------|
| 周一 | 1.003 | 1.002 | 1.000 | 0.997 | 0.990 | 1.000 |
| 周二 | 0.997 | 1.002 | 0.992 | 0.996 | 0.991 | 1.012 |
| 周三 | 1.021 | 1.015 | 0.996 | 1.008 | 1.006 | 0.990 |
| 周四 | 0.992 | 1.003 | 1.007 | 1.002 | 0.990 | 1.002 |

| | | | | | | |
|----|-------|-------|-------|-------|-------|-------|
| 周五 | 0.999 | 0.978 | 1.013 | 0.996 | 1.007 | 1.015 |
| 周六 | 0.979 | 1.005 | 0.998 | 0.991 | 1.008 | 0.995 |
| 周日 | 1.009 | 0.995 | 0.994 | 1.010 | 1.010 | 0.986 |

观察数据得知，批发价格不存在明显的周期性，这是合理的。

3. 补货总量与预测销售量的关系如下：

$$\text{补货总量} = \text{预测销售量} / (1 - \text{损耗率})$$

设销售单价为 x ，星期为 y ，预测销售量为 $b(x,y)$ ，批发价格为 $c(y)$ ，补货总量为 $d(x,y)$ ，则总利润 $F(x,y)$ 可以表示为：

$$F(x,y) = x \cdot b(x,y) - c(y) \cdot d(x,y)$$

设蔬菜品类为 z ，则总利润 $F(x,y,z)$ 可以表示为：

$$F(x,y,z) = x(z) \cdot b(x,y) - c(y,z) \cdot d(x,y)$$

当前模型共有 3 个自变量，为了简化计算，我们分别对每种星期进行计算。考虑到在之前我们得出过大部分蔬菜品类销售量之间都具有一定相关性，我们将所有蔬菜品类合计在一起进行计算，将总利润直接求最优优化，而不是对每种蔬菜品类分开计算各自利润。

我们采用迭代网格优化算法来求解最优情况。为了简化模型计算，我们对于各品类蔬菜的销售单价进行初始化，考虑到往年平均售价如下：

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|--------|--------|-------|-------|-------|--------|--------|
| 平均销售单价 | 10.372 | 6.241 | 9.466 | 8.613 | 10.247 | 12.147 |

因此，我们取各网格初始化参数如下，精确到 0.5：

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|-------|-----|------|-----|------|------|
| 取值 1 | 9.5 | 5.5 | 8.5 | 7.5 | 9.0 | 11.0 |
| 取值 2 | 10.0 | 6.0 | 9.0 | 8.0 | 9.5 | 11.5 |
| 取值 3 | 10.5 | 6.5 | 9.5 | 8.5 | 10.0 | 12.0 |
| 取值 4 | 11.0 | 7.0 | 10.0 | 9.0 | 10.5 | 12.5 |
| 取值 5 | 11.5 | 7.5 | 10.5 | 9.5 | 11.0 | 13.0 |

我们先关注 2023 年 7 月 1 日的训练过程。经过第一次优化后，得到结果如下，全局最优解被加粗标红：

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|----------|---------|--------|--------|--------|---------------|
| 销售单价 | 11.5 | 7.0 | 10.5 | 7.5 | 11.0 | 12.0 |
| 补货总量 | 40.218 | 194.962 | 40.093 | 30.408 | 74.574 | 83.809 |
| 总利润 | 4448.415 | | | | | |

从结果上看，食用菌的取值在网格范围内，因此无需再次迭代。而其他蔬菜

品类的取值均在网格边缘，表明可能尚未达到最优解，仍需继续迭代。将需要继续迭代的网格顺次向上或向下平移，得到第二次优化的参数如下表：

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|-------|-----|------|-----|------|------|
| 取值 1 | 11.5 | 7.0 | 10.5 | 5.5 | 11.0 | 12.0 |
| 取值 2 | 12.0 | 7.5 | 11.0 | 6.0 | 11.5 | |
| 取值 3 | 12.5 | 8.0 | 11.5 | 6.5 | 12.0 | |
| 取值 4 | 13.0 | 8.5 | 12.0 | 7.0 | 12.5 | |
| 取值 5 | 13.5 | 9.0 | 12.5 | 7.5 | 13.0 | |

第二次迭代优化结果如下：

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|----------|---------|--------|--------|-------|--------|
| 销售单价 | 13.0 | 8.5 | 12.0 | 7.5 | 11.0 | 12.0 |
| 补货总量 | 35.798 | 187.612 | 35.304 | 36.408 | 68.69 | 83.809 |
| 总利润 | 4814.217 | | | | | |

依此类推，经过多次迭代优化后，得到最终结果如下：

| 蔬菜品类 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 | 食用菌 |
|------|----------|---------|--------|--------|---------|--------|
| 销售单价 | 13.0 | 8.5 | 12.0 | 7.5 | 32.5 | 12.0 |
| 补货总量 | 35.798 | 187.612 | 35.304 | 30.408 | 134.698 | 83.809 |
| 总利润 | 8774.301 | | | | | |

注意到，除辣椒类以外，各品类销售单价均在合理范围内，后续可以通过减小网格宽度的方式继续迭代优化。根据一份市场调研报告显示，2023 年广东省辣椒类市场价格大约在 35 元/千克^[3]。虽然根据《中华人民共和国价格法》第十一条规定，经营者进行价格活动，享有自主制定属于市场调节的价格的权利^[4]。

7 月 2 日到 4 日的补货总量和定价策略以此类推，限于篇幅，结果放在附件中。

3.3 各蔬菜单品未来一天的补货总量和定价策略

基于问题二中的利润计算方法，针对问题三我们继续搭建数学模型。

首先做出如下假设：

1. 认为今年市场环境是温和的。即到 2023 年 7 月为止与农业方面如农业生产、海关进出口相关的国内国际环境是稳定的，因此 2023 年 7 月 1 日的销售情况只与蔬菜是否旺季有关。

2. 认为 6 月 24 到 30 日的批发价格, 销售倾向是稳健的。根据近一周的销售成功产品的批发价格波动大小可知是可行的。

现在进行一个初步的变量选择,对每个单品采用如下的评分细则:

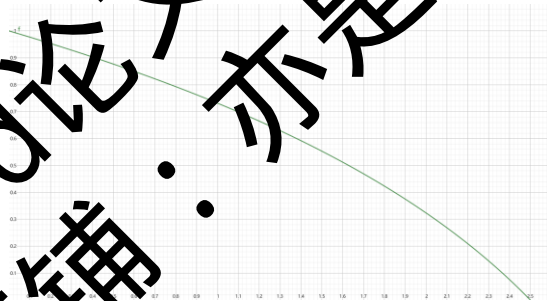
$$score = \sum_{i=1}^7 (\omega_i \times Y_i + I_{(x_i \leq 2.5)} (x_i - 2.5) (\alpha(x_i)))$$

其中, x_i 代表旺季以来即 6 月 24 日以来的第 i 天的销售量, $\alpha(x)$ 是一个凹函数, ω_i 是第 i 天的天数因子, Y_i 为该天总利润。

因为研究对象是 6 月 24 到 30 日的数据与 7 月 1 日的关系, 考虑间隔的天数我们提出天数因子, 以天数因子 2 为例计算如下从 4 月 1 号开始将 4 月 3 号与其各放到一组, 4 月 2 日与 4 月 4 号分别放入前面的第一组与第二组, 依此类推, 最后计算两个小组之间的协相关系数, 这样计算完 1 到 7 天间隔之后再求得间隔两天占总和的比例即为天数因子; I 是一个惩罚项, 对销售量未达到 2.5 千克的产品进行惩罚, α 代表惩罚力度, 满足如下要求:

1. $\alpha(0)$ 为批发价, $\alpha(2.5)$ 为 0
2. 越靠近 0 惩罚力度越大, 惩罚下降越慢, 越靠近 1 惩罚力度越小, 惩罚下降越快。

$$\alpha(x) = \frac{\text{批发价}}{\lg(3.5)} \times \lg(3.5 - x)$$



最终我们选取分数不为负数的单品, 接着基于问题二我们提出如下模型, 首先对与单品有关的以下符号做出规定与假设:

α : 日损失率

β : 批发价格

z : 销售量

y : 日利润

x : 定价

于是根据第二问提出的

$$\text{总利润} = \text{销售单价} * \text{预测销售量} - \text{批发价格} * \text{补货总量}$$

最终对每个单品有如下模型:

$$x = \frac{y + A\beta}{1 - \alpha} \times \frac{1}{A}$$

对于上述模型各变量处理,根据假设,我们首先剔除了最近几天销售为 0 的单品,然后对剩下的单品我们用 4 月 1 日以来的数据采用 ARIMA 模型做出预测,得到销售量 95% 的置信区间,同时,我们采用多项式样条方法对 4 月 1 日以来的数据做处理得到销售量与销售总利润的关系,在 A 的置信区间范围内对 y 进行预测并得到 y 的最大值;对于 β , 根据假设认为 β 是在前几天的平均情况下加上一个微小扰动。

最终根据总利润及上述定价公式,我们选择的单品及其对应的补货量和销售定价为:

| 产品名称 | 定价 | 补货量 |
|----------|--------|--------|
| 螺丝椒 | 9.354 | 43.613 |
| 小米椒 | 4.241 | 41.537 |
| 紫茄子 | 5.949 | 34.951 |
| 西兰花 | 10.758 | 26.762 |
| 芜湖青椒 | 3.536 | 37.480 |
| 西峡花菇 | 23.834 | 9.583 |
| 云南油麦菜 | 4.189 | 38.052 |
| 洪湖藕带 | 24.515 | 10.124 |
| 长线茄 | 11.894 | 9.272 |
| 金针菇 | 3.544 | 41.487 |
| 七彩椒 | 21.278 | 3.116 |
| 双孢菇 | 4.948 | 24.430 |
| 红薯尖 | 5.857 | 13.406 |
| 长耳莩 | 6.253 | 11.665 |
| 云南生菜 | 4.906 | 51.802 |
| 竹叶菜 | 3.427 | 29.805 |
| 娃娃菜 | 6.218 | 26.211 |
| 姜蒜小米椒组合装 | 4.135 | 16.091 |
| 枝江青梗散花 | 11.347 | 17.599 |
| 菠菜 | 7.091 | 37.316 |
| 高瓜 | 17.223 | 7.135 |

| | | |
|-----------|--------|--------|
| 上海青 | 5.992 | 14.634 |
| 净藕 | 12.902 | 15.007 |
| 奶白菜 | 3.577 | 22.516 |
| 小皱皮 | 2.436 | 20.848 |
| 苋菜 | 3.372 | 16.563 |
| 小青菜 | 3.889 | 14.271 |
| 红椒 | 17.653 | 4.239 |
| 蟹味菇与白玉菇双拼 | 4.178 | 5.019 |
| 青茄子 | 6.819 | 5.317 |
| 圆茄子 | 7.152 | 3.831 |
| 海鲜菇 | 2.273 | 24.969 |
| 菜心 | 3.629 | 6.628 |

计算得到的总利润为 1129.719 元。

3.4. 关于另需额外采集数据的研究

可以观察到，适当且充分的数据分析可以有效应对问题一至问题三。然而，这些数据仅反映了相关蔬菜商品在 2020 年 7 月至 2023 年 6 月这一特定时间段内的销售情况，具有品类和时间上的局限性。换言之，若要研究其他蔬菜品类的销售策略或是距离 2020 年 7 月至 2023 年 6 月较远时段的销售策略，基于提供的数据构建的数学模型或是神经网络均有可能引发较大误差甚至明显的错误，因而无法得出合理可行的结果。

为了更好地制定蔬菜商品的补货和定价决策，通过深入查询相关资料以及根据已有研究和实践经验，我们坚信商超还需要采集以下相关数据，这些数据对解决上述问题具有各自的帮助。

3.4.1. 顾客反馈数据

商超应收集顾客信息，包括购物偏好和历史，因为不同顾客群体对蔬菜需求各异。有些注重品质，有些注重价格。通过分析这些数据，商超能更好理解各群体需求，制订个性化的补货和定价策略，提升销售^[5]。同时，顾客反馈和建议的收集可改进品质和品种，提高满意度与忠诚度，更好满足市场需求。这些信息为制定精确策略、增加竞争力，最大化利润提供全面帮助。

3.4.2. 竞争对手数据

了解附近竞争对手的蔬菜商品种类、定价策略和销售数据等，对商超的竞争

力至关重要。相关研究表明，商超在价格竞争激烈的市场中需要根据竞争对手的策略来调整自己的定价和营销策略，以吸引更多顾客并保持盈利能力。

3.4.3. 供应链数据

供应链的稳定性和可靠性对商超非常重要。收集蔬菜的产地、供应商信息、运输时间和条件等数据可以帮助商超更好地管理供应链。研究表明，及时了解供应链中的变化和延误，可以帮助商超做出迅速的补货和定价决策，以满足市场需求^[6]。

3.4.4. 市场趋势数据

收集与蔬菜相关的市场趋势数据，如天气情况、季节性变化、政策变化等。这些数据可以帮助商超更好地预测需求波动和价格波动，从而及时调整补货策略。

3.4.5. 损耗率数据的持续监测

商超已经提供了近期的损耗率数据，但研究显示，持续监测这些数据对于补货决策非常重要。通过追踪每个蔬菜品类的损耗率，商超可以及时调整补货策略，减少损失，并提高库存管理的效率。

3.4.6. 库存数据

持续监测并统计每个蔬菜品类的库存水平，以确保库存处于合理水平。研究发现，过高或过低的库存都会对商超的盈利能力产生负面影响。因此，保持适当的库存水平是关键。

4. 总结

本文研究了蔬菜类商品的自动定价与补货决策问题。通过建立多种数学模型，分析某超市一年内的销售、批发和损耗等数据，研究了蔬菜销量、单价等各个关键因素之间的关系。研究过程中利用迭代优化算法，求解了在特定时间条件下的最优销售单价和进货量，以达到商超总利润最大化的目标。此研究对不同商业模式下的数据分析和营销策略制定等都有研究价值。

研究方法包括数据处理和数学模型建立，以及通过分析销售、定价和补货数据来解决问题。研究结果包括销售量的分布规律、各蔬菜品类的销售总量与成本加成定价的关系、未来一周的补货总量和定价策略等。最后，提出了需要采集额外数据的建议，以优化蔬菜商品的补货和定价策略。

总体而言，这篇论文通过深入的数据分析和数学建模，为商超提供了有关蔬菜类商品的自动定价与补货决策的实用建议。这些建议有助于商超提高销售效益并最大化利润，同时也为相关领域的研究提供了有价值的参考。

参考文献

- [1] 娄新新,冯爱芬,荣博.基于 ARIMA 模型的农产品销售问题研究[J].Modeling and Simulation, 2023, 12: 3918.
- [2] 许正厚. 城市商住混合街区格局对微气候影响机理研究[D].北方工业大学,2021.DOI:10.26926/d.cnki.gbfgu.2021.000633.
- [3] "2023 年广东辣椒市场价格行情," Vipveg, <http://www.vipveg.com/lajiao/price/guangdong/>
- [4] "中华人民共和国价格法," Gov, https://www.gov.cn/banshi/200509/12/content_69757.htm
- [5] 王梓润,王仟怡,袁文佳等.基于混合智能算法对销售量的预测研究[J].电脑与电信,2023(03):33-37.DOI:10.15966/j.cnki.dnydx.2023.03.020.
- [6] 王志刚,冯云超,江勇.基于深度学习的连锁超市销售预测研究[J].现代计算机,2020(34):31-35.

如需word论文和完整源代码
淘宝店铺：亦是远方