

# 问题2

## 前置数据处理

1. 提取数据+计算利润率+计算加成率
2. 3sigma异常值处理

## 因果检验

1. ADF检验确认各时间序列均是平稳序列(避免伪回归)
2. 建立滞后p阶的VAR向量自回归模型,从而找到最优滞后阶数(通过对FPE、AIC、SC、HQ的指标进行评价)
3. 格兰杰因果检验变量之间是否存在因果关系

## 各蔬菜品类的销售总量与成本加成定价的关系

### 方案1

1. 对于每个类进行建立单独的ADF检验
2. 寻找最优滞后阶数
3. 对于每个类进行建立单独的VAR向量自回归模型,得出各蔬菜品类的销售总量与成本加成定价的关系
4. VAR模型稳定性检验

### 方案2(备选)

1. 通过训练集数据来建立XGBoost回归模型。
2. 通过建立的XGBoost来计算特征重要性。
3. 将建立的XGBoost回归模型应用到训练、测试数据,得到模型评估结果。

## 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

### 日补货总量

1. 前提:忽略库存,理想情况下日补货总量将等于日销售量为最佳

2. 建立季节性ARIMA模型预测未来7天的日销量

## 定价策略\_方案甲(备选)

1. 建立XGBoost回归模型得到利润为因变量,日销量和售价为自变量的回归模型
2. 启发式搜索得到使得商超收益最大

## 定价策略\_方案乙

1. 建立线性回归(最小二乘)模型得到利润为因变量,日销量和售价为自变量的回归模型