

# 摘要

肥料是农业生产中一种重要的生产资料，对农业生产起着关键作用，国内外农化专家普遍认为,在其他生产因素不变的情况下,农作物施用化肥可增加产量 40%~60%。可以说肥料的施用解决了中国人甚至是世界的挨饿问题。而不同类别的肥料对各省份地区、各品种农作物又起到不同的效果，正因为如此，对不同类别肥料的生产登记数量进行分析，对各省肥料生产企业起到重要决策作用。

# 目录

1 目标 .....	1
2 任务一 数据的预处理.....	1
2.1 规范产品通用名称.....	1
2.1.1 处理思路.....	1
2.1.2 规范过程及结果.....	1
2.2 总无机养分百分比.....	1
2.2.1 处理思路.....	1
2.2.2 过程及结果.....	1
3 任务二 肥料产品的数据分析.....	1
3.1 复混肥料产品分析.....	1
3.1.1 处理思路.....	2
3.1.2 过程及结果分析.....	2
3.2 有机肥料产品分析.....	2
3.2.1 处理思路及过程.....	2
3.2.2 分布特点分析.....	3
3.3 复混肥料聚类处理分析.....	3
3.3.1 处理思路及过程.....	3
3.3.2 分析聚类特征.....	4
4 任务三 肥料产品的多维度对比分析.....	5
4.1 复混肥料变化趋势分析.....	5
4.1.1 处理思路.....	5
4.1.2 分析过程.....	5
4.2 广西和湖北有效产品分布差异.....	6
4.2.1 处理过程.....	6
4.2.2 分析结果.....	6
4.3 企业间的杰卡德相似系数矩阵.....	6
5 任务四 肥料产品的多维度对比分析.....	7
5.1 从技术指标中提取信息.....	7
5.1.1 氮、磷、钾养分和有机质的百分比.....	7
5.1.2 肥料含氯的程度.....	7
5.2 提取各原料名称及其百分比.....	7

---

## 1 目标

1. 对肥料登记数据进行预处理。
2. 根据养分的百分比对肥料产品进行细分。
3. 从省份、日期、生产商、肥料构成等维度对肥料登记数据进行对比分析。
4. 对非结构化数据进行结构化处理。

## 2 任务一 数据的预处理

### 2.1 规范产品通用名称

#### 2.1.1 处理思路

#### 2.1.2 规范过程及结果

### 2.2 总无机养分百分比

#### 2.2.1 处理思路

#### 2.2.2 过程及结果

## 3 任务二 肥料产品的数据分析

### 3.1 复混肥料产品分析

3.1.1 处理思路

3.1.2 过程及结果分析

图 1 产品登记数量的直方图

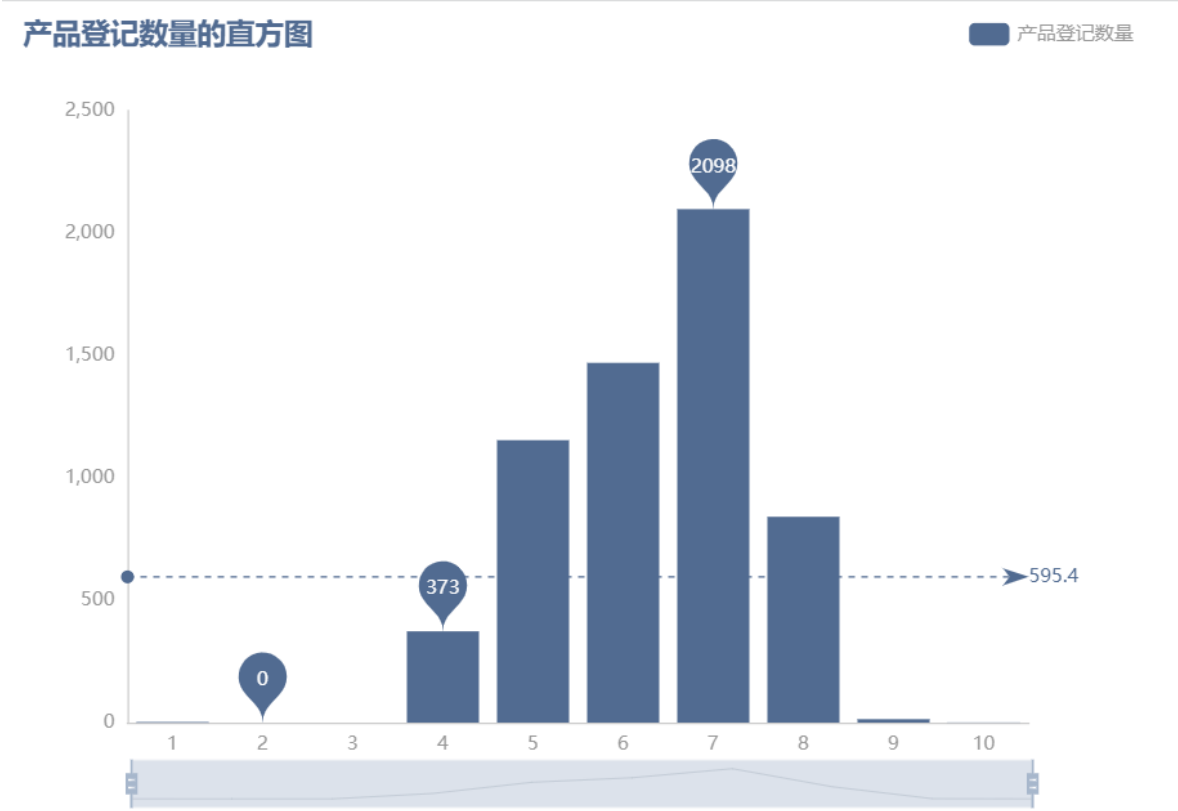


表 1 登记数量最大的前 3 个分组及相应的产品登记数量

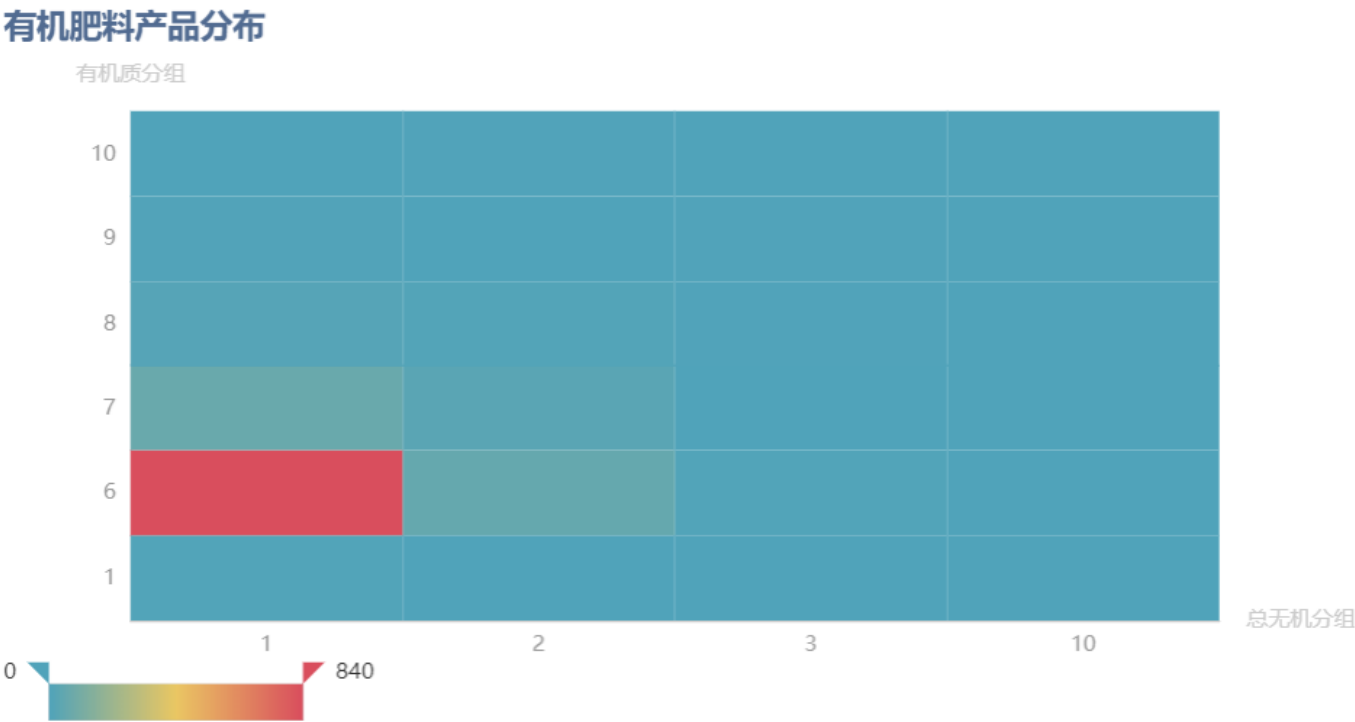
排名	一	二	三
分组标签	7	6	5
产品登记数量	2098	1470	1154

3.2 有机肥料产品分析

3.2.1 处理思路及过程

### 3.2.2 分布特点分析

图 2 有机肥料产品的分布热力图

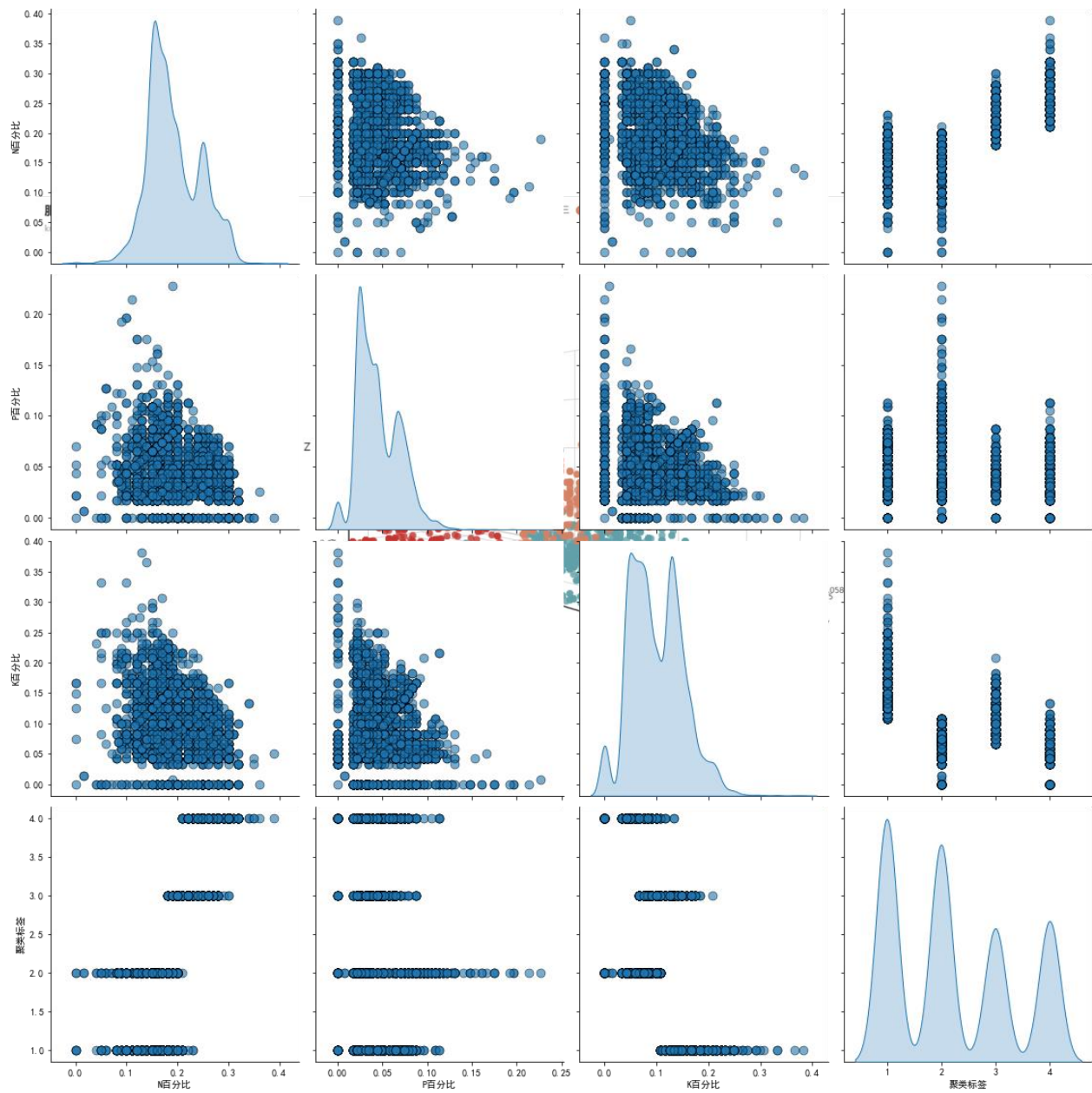


### 3.3 复混肥料聚类处理分析

#### 3.3.1 处理思路及过程

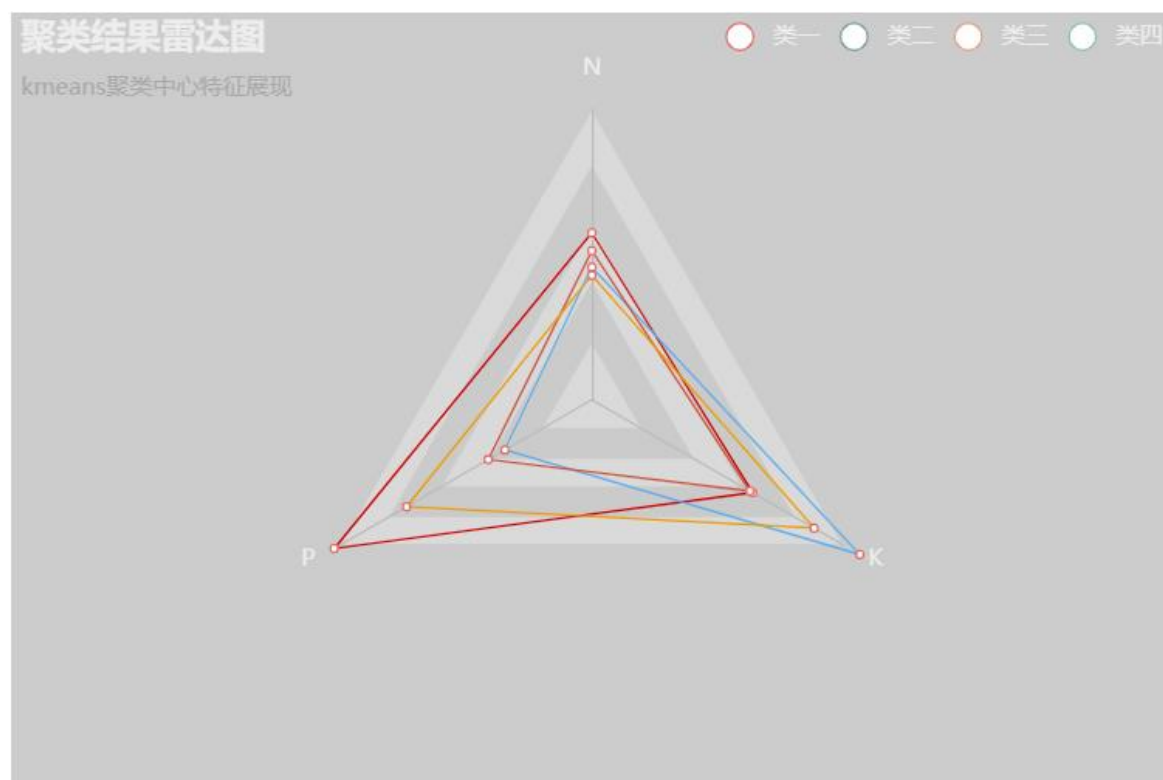
图 3 根据聚类标签绘制肥料产品的三维散点图

图 4 根据聚类标签绘制肥料产品的散点图矩阵



### 3.3.2 分析聚类特征

图 5 聚类结果的雷达图



## 4 任务三 肥料产品的多维度对比分析

### 4.1 复混肥料变化趋势分析

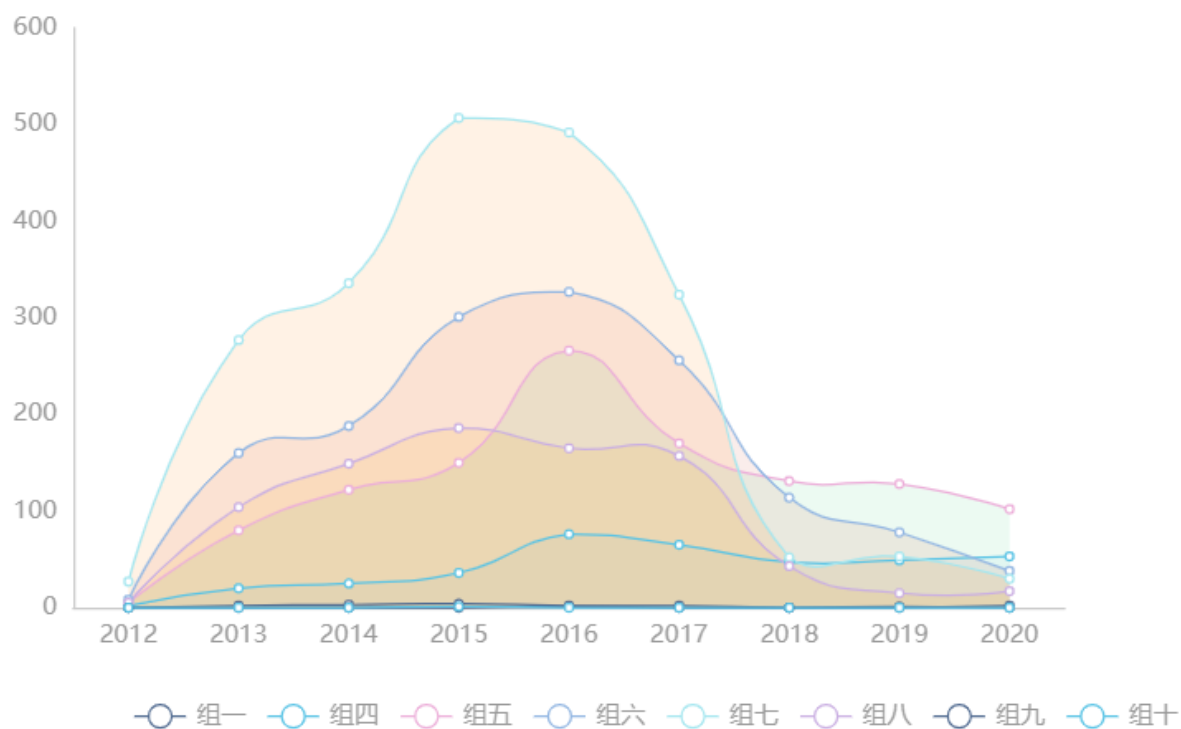
#### 4.1.1 处理思路

#### 4.1.2 分析过程

图 6 复混肥料各组别不同年份产品登记数量变化趋势

## 混肥料各组别变化趋势

不同年份产品登记数量



## 4.2 广西和湖北有效产品分布差异

### 4.2.1 处理过程

### 4.2.2 分析结果

图 7 广西和湖北产品登记数量在前 5 的组别

## 4.3 企业间的杰卡德相似系数矩阵

计算公式：集合  $A$  与  $B$  的杰卡德相似系数定义为  $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$ ，其中  $|S|$  表示集合  $S$  中元素的个数。



---

## 5 任务四 肥料产品的多维度对比分析

### 5.1 从技术指标中提取信息

#### 5.1.1 氮、磷、钾养分和有机质的百分比

#### 5.1.2 肥料含氯的程度

### 5.2 提取各原料名称及其百分比