

## 摘要

肥料是农业生产中一种重要的生产资料，其生产销售必须遵循国家颁布的相关条例。依法有序在相关部门完成登记工作。肥料与当今社会的食品发展密切相关，与人们的健康有着间接的联系。因此，利用数据分析这些肥料的成分与组成，对某地区甚至我国乃至全世界有着重要作用。

## Abstract

Fertilizer is an important means of production in agricultural production, its production and sales must follow the relevant regulations promulgated by the state. To complete registration in relevant departments in accordance with the law and in an orderly manner. Fertilizer is closely related to food development in today's society and indirectly related to people's health. Therefore, using data to analyze the composition and composition of these fertilizers has an important role in some regions and even in China and the world.

---

摘要.....	1
Abstract.....	2
1. 问题分析.....	4
2. 任务一 数据的预处理.....	4
2.1 任务 1.1.....	4
2.2 任务 1.2.....	4
3 任务二 肥料产品的数据分析.....	5
3.1 任务 2.1.....	5
3.2 任务 2.2.....	6
3.3 任务 2.3.....	8
4.任务三 肥料产品的多维度对比分析.....	10
4.1 任务 3.1.....	10
4.2 任务 3.2.....	10
4.3 任务 3.3.....	10
5.任务四 肥料产品的多维度对比分析.....	10
5.1 任务 4.1.....	10
参考文献.....	10

# 1.问题分析

对不同省份、不同日期、不同生产商、不同肥料作为维度，对肥料登记数据进行多维度分析，同时结合分析结果，给出合理意见。

## 2.任务一 数据的预处理

### 2.1 任务 1.1

2.1.1 缺失值、重复值、异常值处理

2.1.2 通用名称的规范化处理

### 2.2 任务 1.2

2.2.1 计算总无机养分百分比

## 3 任务二 肥料产品的数据分析

### 3.1 任务 2.1

#### 3.1.1 缺失值、重复值、异常值处理

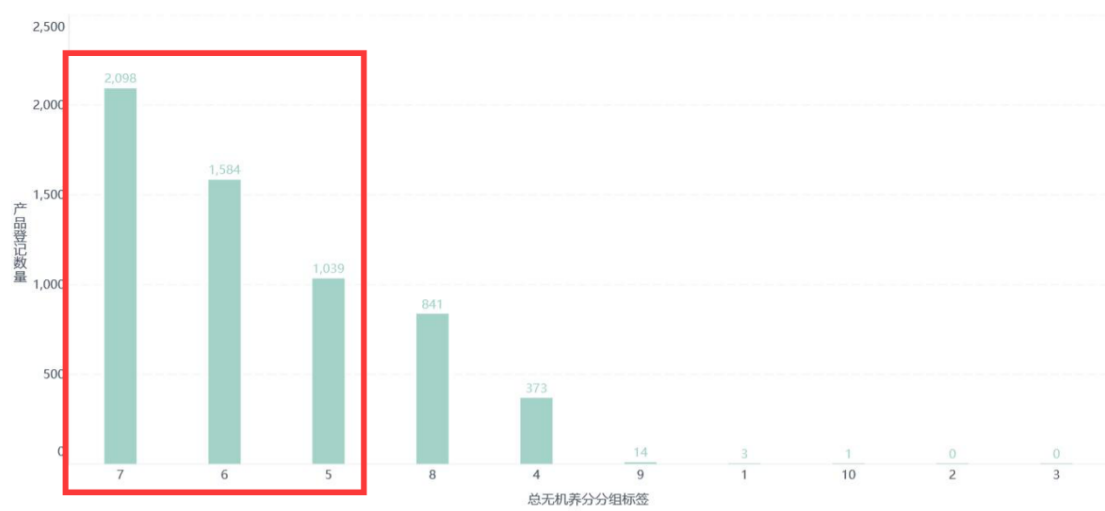
#### 3.1.2 等距分组

#### 3.1.3 绘制直方图，分析产品分布特点

**分析：**根据图标可以看出，标签为 5、6、7 组的登记产品数量的总和为前三，对应的总无机养分分组对应的为 0.36、0.432、0.504，总无机养分中包含着氮、磷、钾这三种元素总和。说明并不是无机养分含量越高，化肥质量越好。因此没有必要去生产一些养分含量过高的化肥，这也许会生产作物产生营养过度的作用甚至起到副作用，同时还会增加成本。但是对于某些特定的生产作物也许会需要养分低一点或者养分更高一点的化肥，但是受众范围比较小。

但是判断一种肥料的好坏，并不能单一的使用总无机养分去判断，过于武断。

不同标签的产品登记数量的直方图：

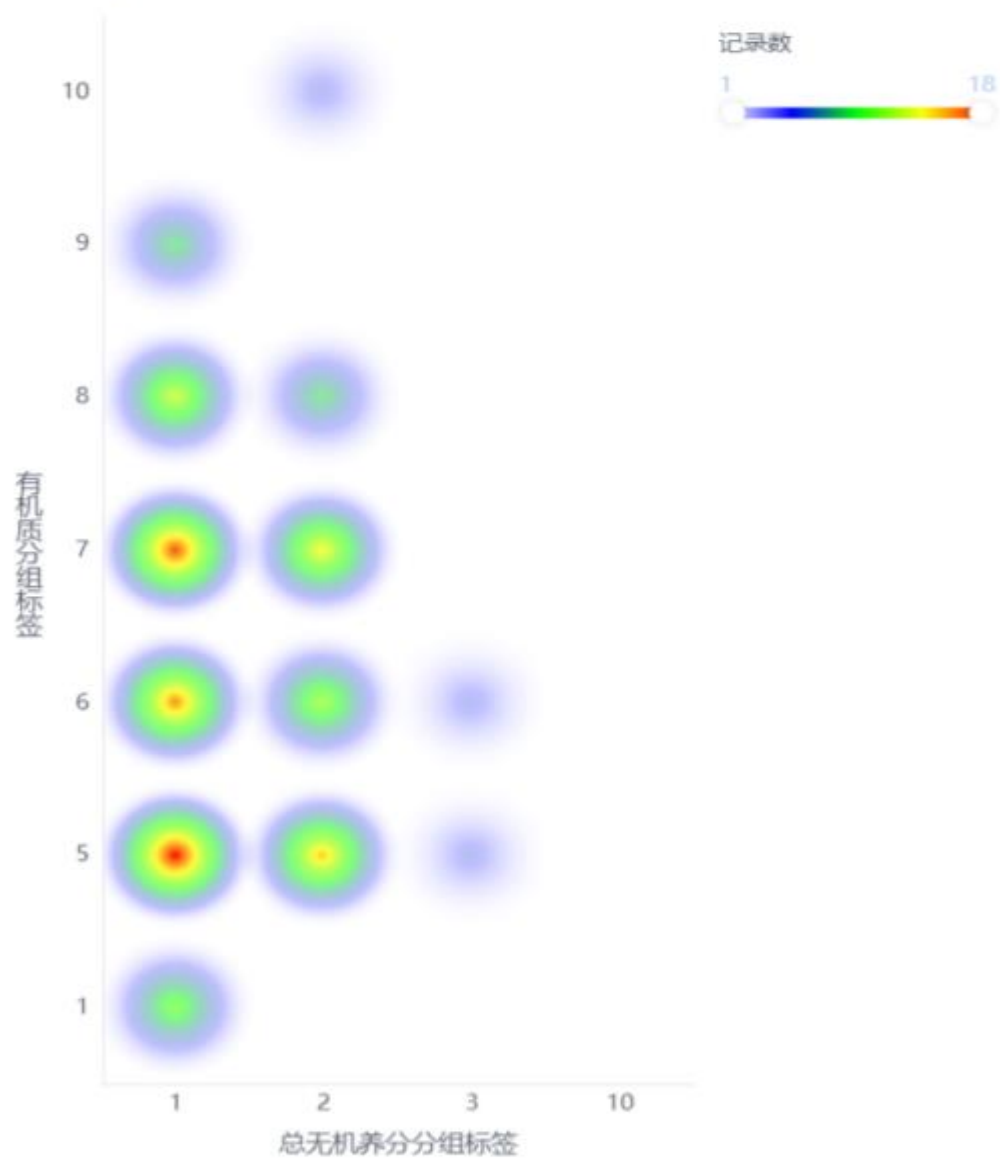


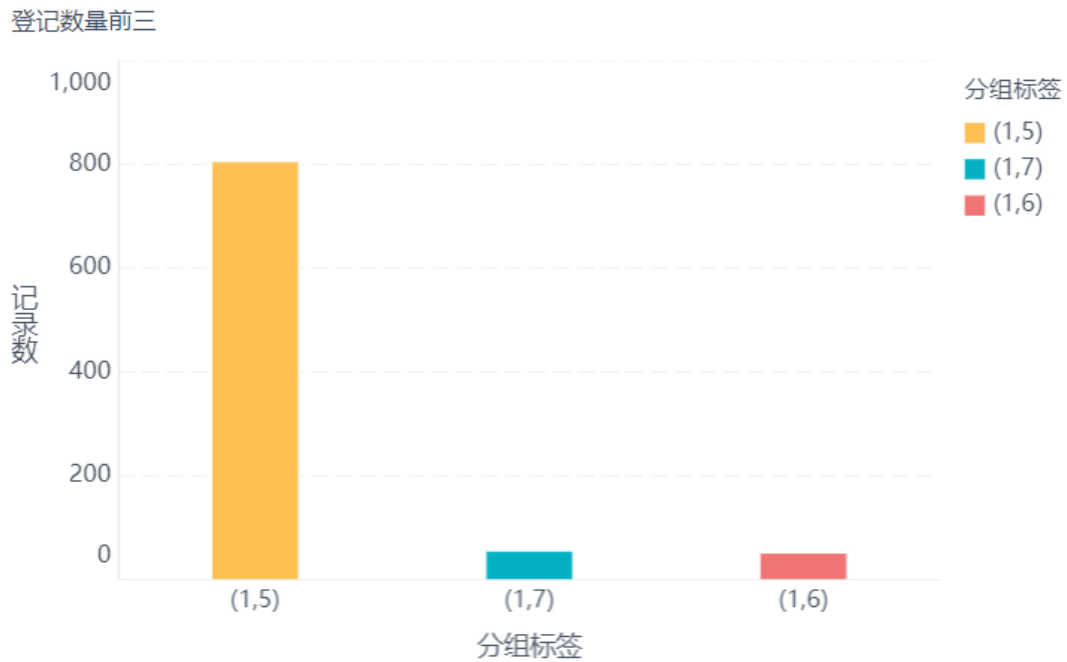
排名	一	二	三
分组标签	7	6	5
产品登记数量	2098	1584	1039

3.2任务 2.2

3.2.1 有机肥料等距分组

分布热力图





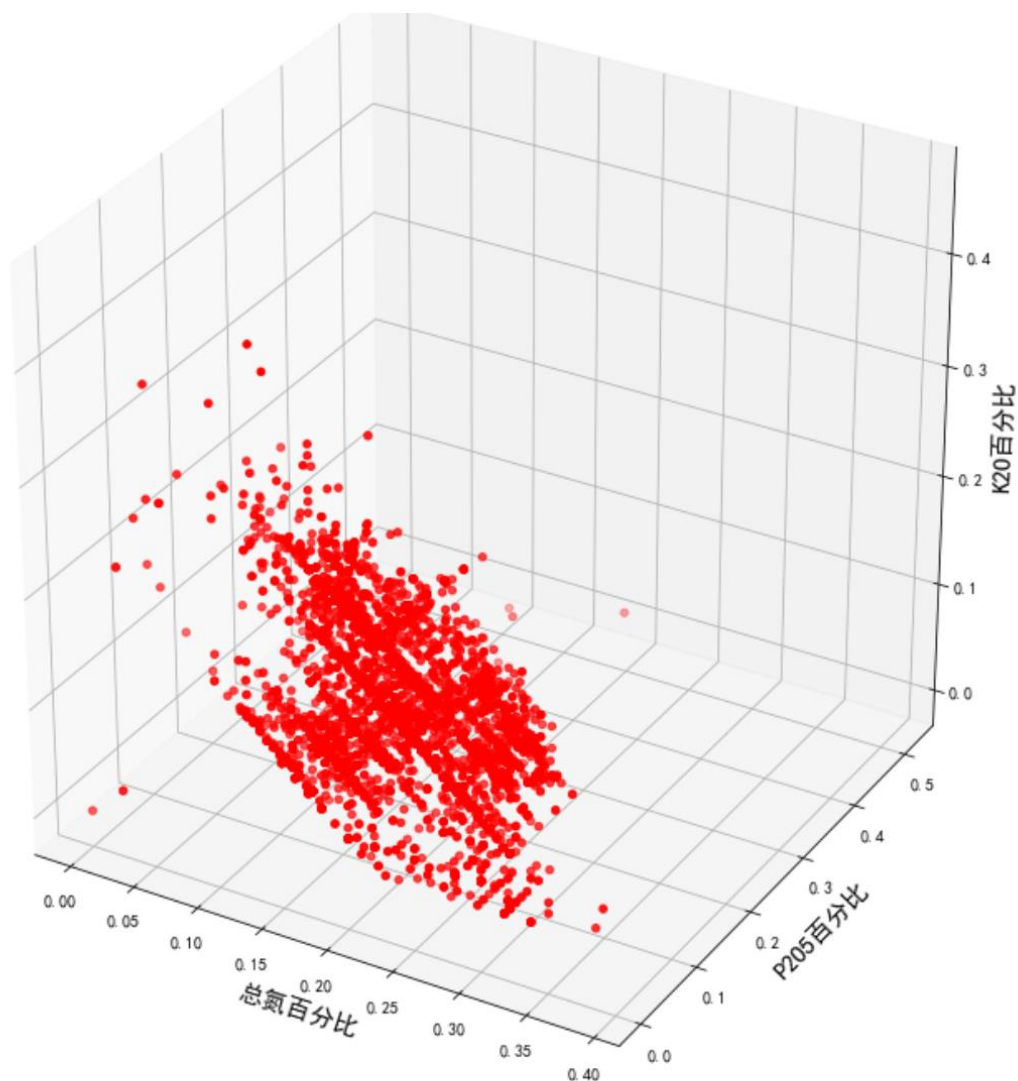
分析:通过分布热力图和筛选出排名前三的记录数可以看出,(1, 5)出现的频率最高,805次,其次是(1, 7),(1, 6),说明总无机养分在0-0.092和有机质养分在0-0.1之间的有机化肥最多。说明通过大数据分析来看,大多数的有机肥料的无机养分不能存在过多,同时有机质养分只能保持适中,这样的肥料才是最受欢迎的,适合商家去生产,也说明这样的肥料对生产作物比较有用。

### 3.3任务 2.3

3.3.1 复混肥料的聚类算法分组, 绘制雷达图

3.3.2 绘制三维散点图、三点矩阵图, 分析特征





分析：从三维矩阵散点图可以看出，复混肥料中的总氮占比区间大部分为 0.05-0.30 之间，P2O5 占比区间在 0.1-0.3 之间，K2O5 占比区间大部分在 0.0-0.1 之间，说明大部分的复混肥料中氮和磷占比多于钾，并且大多数的复混肥料的个元素占比都较集中，说明复混肥料的配料较统一，不存在较大差异，但是也存在个别肥料的配方，差异较大，但是由于土地、食品等种类不同的原因，差异可以理解。

## 4.任务三 肥料产品的多维度对比分析

### 4.1任务 3.1

4.1.1 不同年份不同组别的产品的等级数量

4.1.2 处理思路及分析

### 4.2任务 3.2

### 4.3任务 3.3

## 5.任务四 肥料产品的多维度对比分析

### 5.1任务 4.1

## 参考文献

[1]廖彬,张陶,于炯,国冰磊,刘继.基于二维划分的杰卡德相似系数批量计算效率优化[J].计算机科学,2017,44(01):219-225.

[2]刘春晓,苗庆选,怀婷婷,任爱芝,郭满,马海林,李海云,司东霞.氮磷钾肥料施用对核桃蛋白质和粗脂肪含量的影响[J/OL].中国土壤与肥料:1-8[2021-11-14].