

# 肥料登记数据分析报告

## 摘要

由于肥料在农业生产中扮演着越来越重要的角色，研究肥料登记数据有利于农业行政部门提高管理水平。在此背景之下，对肥料数据进行分析具有重要意义；本次报告的研究主要针对企业名称、产品通用名称、正式登记编号、有效期、营养成分表、含氯情况等指标进行分析，希望通过各指标之间的关系提取重要信息以便于行政部门提高管理水平以及发现各种肥料的分布情况。

**关键词：**热力图 K-均值聚类 杰卡德相似矩阵

## 一、背景及意义

目前肥料在农业生产中占有重要地位，其生产销售在农业行政管理部门必须进行登记处理。分析肥料登记数据中存在的问题，然后提出相应的解决对策，对提高肥料管理水平具有一定的帮助。

## 二、研究目标

1. 对肥料的登记数据进行预处理
2. 依据无机和有机养分百分比对肥料进行分组
3. 从省份、日期、生产商、肥料构成及各种维度对肥料数据进行相关性分析
4. 对非结构化数据进行相应的结构化处理

## 三、任务一

### 1. 数据预处理

#### 1.1 规范化处理

依据题目要求，产品通用名称分为四个类别；附件一中产品通用名称存在不属于四个类别中的名称，因此需要将名称、格式不符合的改为对应的四种类别中的名

3.2 总无机养分百分比的计算

由题目要求知总无机养分百分比=含氮百分比+含磷百分比+含钾百分比，依据附件一中的氮、磷、钾养分百分比，通过简单的数学运算并利用 Excel 软件即可得出总无机养分百分比，

任务 2

2.1 复混肥料的等间距分组

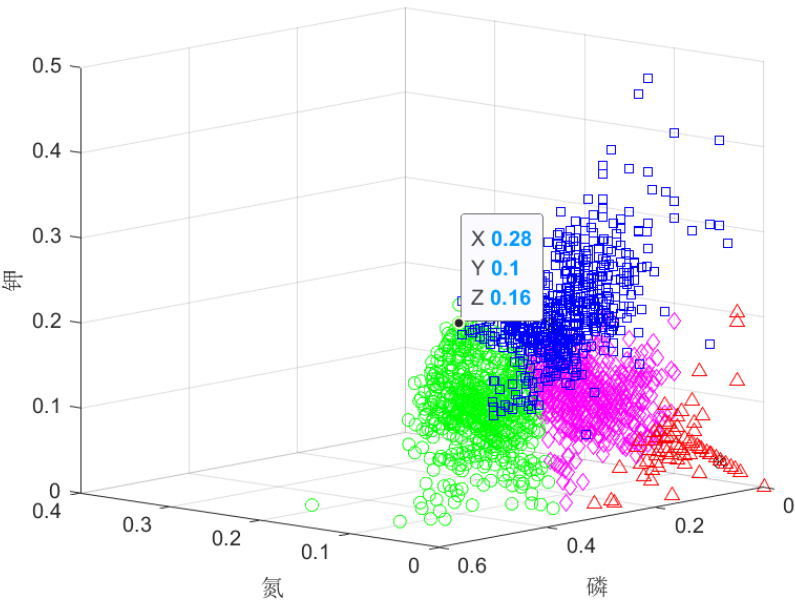
根据组别登记数量列表可知总复混肥料的无机养分百分比基本处于 24%至 56%之间，为了更加直观的看出复混肥料等级数量的差异，绘制直方图如下  
最后按登记数量列出登记数量三名的组别及登记数量如下表所示

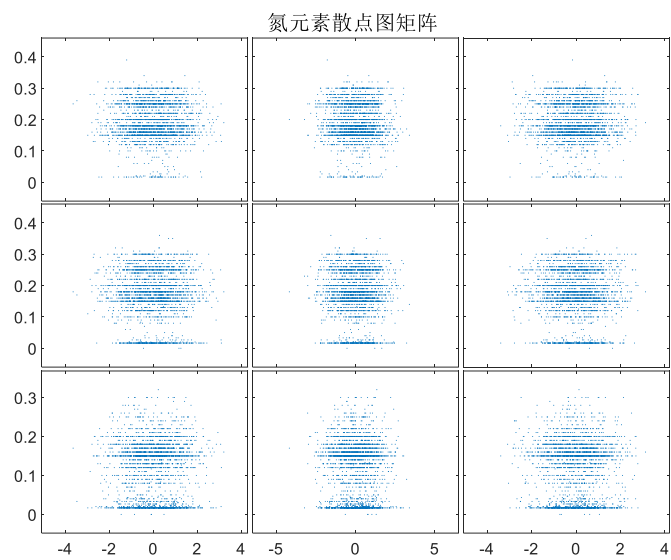
表格 1 登记数量前三名组别

排名	一	二	三
分组标签	6	7	4
产品登记数量	2760	1417	952

2.2 有机肥料等间距分析

2.21 依据总无机养分和有机质百分比分组





通过三维散点图、氮元素散点图矩阵可知登记数量含钾元素量较多。复混肥料登记数量与氮磷钾含量具有密切的联系。