

# 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《肥料登记管理办法》，依法在农业行政管理部门进行登记。各省、自治区、直辖市人民政府农业行政主管部门主要负责本行政区域内销售的肥料登记工作，相关数据可从政府网站上自由下载。 目标任务：

- 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。

# 目录

- 1、任务 1 数据的预处理 ..... 4
  - 1.1 任务 1.1 ..... 4
    - 1.1.1 任务描述 ..... 4
    - 1.1.2 处理思路 ..... 4

1.1.3 关键过程 .....	4
1.1.4 结果分析 .....	5
1.2 任务 1.2 .....	6
1.2.1 任务描述 .....	6
1.2.2 处理思路 .....	6
1.2.3 关键过程 .....	6
1.2.4 结果分析 .....	7
<b>2、任务 2 肥料产品的数据分析 .....</b>	<b>8</b>
2.1 任务 2.1 .....	8
2.1.1 任务描述 .....	8
2.1.2 思路步骤 .....	8
2.1.3 结果分析 .....	9
2.2 任务 2.2 .....	10
2.2.1 任务描述 .....	10
2.2.2 思路步骤 .....	10
2.2.3 结果分析 .....	12
2.3 任务 2.3 .....	15
2.3.1 任务描述 .....	15
2.3.2 处理思路 .....	15
2.3.3 K-means 聚合算法 .....	15
2.3.3.1 原理 .....	15
2.3.3.2 实现 .....	17
2.3.4 结果分析 .....	18
<b>3、任务 3 肥料产品的多维度对比分析 .....</b>	<b>20</b>
3.1 任务 3.1 .....	20
3.1.1 任务描述 .....	20
3.1.2 处理思路 .....	20
3.1.3 结果分析 .....	20
3.2 任务 3.2 .....	22
3.1.1 任务描述 .....	22
3.1.2 处理思路 .....	22
3.1.3 结果分析 .....	22
3.3 任务 3.3 .....	
3.1.1 任务描述 .....	23
3.1.2 处理思路 .....	23
3.1.3 结果分析 .....	23
<b>4、任务 4 肥料产品的多维度对比分析 .....</b>	<b>25</b>
4.1 任务 4.1 .....	25
3.1.1 任务描述 .....	25
3.1.2 处理思路 .....	26
3.1.3 结果分析 .....	26

4.2 任务 4.2 .....	26
3.1.1 任务描述 .....	26
3.1.2 处理思路 .....	26
3.1.3 结果分析 .....	27

## 1、任务 1 数据的预处理

### 1.1 任务 1.1

**1.1.1 任务描述：**附件 1 的产品通用名称存在不规范的情况。请按照复混肥料（掺混肥料归入这一类）、有机-无机复混肥料、有机肥料和床土调酸剂这 4 种

类别对附件 1 进行规范化处理。请在报告中给出处理思路、过程及必要的结果，同时将完整的结果保存到文件“result1\_1.xlsx”中。

## 1.2 任务 1.2

**1.2.1 任务描述：**计算附件 1 中各肥料产品的氮、磷、钾养分百分比之和，称为总无机养分百分比。请在报告中给出处理思路、过程及必要的结果，同时将完整的结果保存到文件“result1\_2.xlsx”中，结果保留 3 位小数（例如 1.0%，即 0.010）。

## 2、任务 2 肥料产品的数据分析

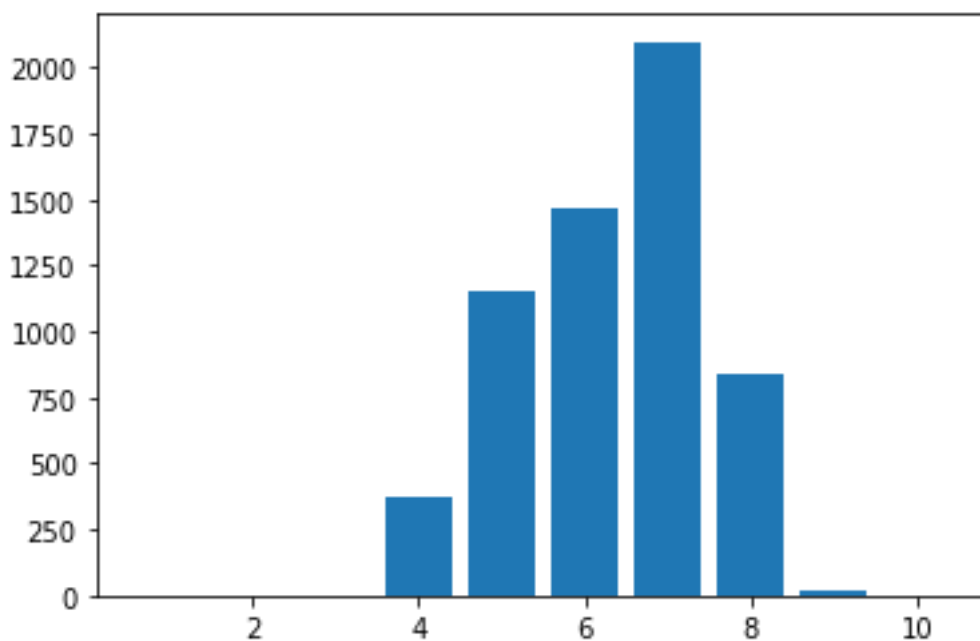
### 2.1 任务 2.1

#### 2.1.1 任务描述

从附件 2 中筛选出复混肥料的产品，将所有复混肥料按照总无机养分百分比等距分为 10 组。根据每个产品所在的分组，为其打上分组标签（标签用 1~10 表示），将完整的结果保存到文件“result2\_1.xlsx”中。分析复混肥料产品的分布特点，在报告中绘制产品登记数量的直方图，给出处理思路及过程，并按登记数量从大到小列出登记数量最大的前 3 个分组及相应的产品登记数量。

#### 2.1.2 思路步骤：

2) 绘制直方图



由分组各自的登记数量和直方图可见，复混肥料产品的分布特点：

①分组 7（复混肥料产品总无机养分百分比在 0.432 - 0.504 间）的产品数量最多

②99%的复混肥料产品总无机养分百分比在 0.21 - 0.576 区间，只有极少部分产品总无机养分百分比低于 0.21 或者高于 0.576，整体分布呈正态分布

3) 登记数量从大到小列出登记数量最大的前 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

## 2.2 任务 2.2

**2.2.1 任务描述：**从附件 2 中筛选出有机肥料的产品，将产品按照总无机养分百分

比和有机质百分比分别等距分为 10 组，并为每个产品打上分组标签 (1, 1), (1, 2), ..., (10, 10)，将完整的结果保存到文件“**result2\_2.xlsx**”中。

请在报告中给出处理思路及过程，并根据分组情况绘制有机肥料产品的分布热力图，其中横轴代表总无机养分分组，纵轴代表有机质分组。在此基础上，分析有机肥料产品的分布特点，并按登记数量从大到小列出登记数量最大的前 3 个分组及相应的产品登记数量。

### **2.2.3 结果分析**

## **2.3 任务 2.3**

### **2.3.1 任务描述**

从附件 2 中筛选出复混肥料的产品，按照氮、磷、钾养分的百分比，使用聚类算法将这些产品分为 4 类。根据聚类结果为每个产品打上聚类标签（标签用 1~4 表示），并将完整的结果保存到文件“**result2\_3.xlsx**”中。请在报告中给出处理思路及过程，根据聚类标签绘制肥料产品的三维散点图和散点图矩阵，并通过绘制聚类结果的雷达图分析每个聚类的特征。