

肥料登记数据分析

一、背景介绍

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。

二、分析目标

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。

三、数据预处理与统计

1. 数据预处理

(1)任务1.1

将“附件1.xlsx”的“产品通用名称”进行处理，按要求将“掺混肥料”改为“复混肥料”，将“稻苗床土调酸剂”改为“床土调酸剂”，将其他不规范的名称“有机无机 复混肥料”、“有机一无机复混肥料”和“有机肥料”分别改为“有机一无机复混肥料”和“有机肥料”。将完整的结果保存到文件“result1_1.xlsx”中。

(2)任务1.2

在“附件1.xlsx”中新增一列并命名为“总无机养分百分比”，利用sum函数计算各肥料产品的氮、磷、钾养分百分比之和得到总无机养分百分比，计算结果保留3位小数（如 0.010），最后将结果保存到文件“result1_2.xlsx”中。

(3)任务4.1数据预处理

对“附件4.xlsx”中的数据针对“技术指标”进行筛选，观察到序号174的复混肥料“技术指标”为无效数据，序号121的复混肥料“技术指标”只给出了氮磷钾的配比无总养分含量，序号23的有机肥料“技术指标”缺失明确数据，这三行均无法进行下一步计算，故做删除处理。对序号61,99,100,135,136,172,189这七行数据，观察到“产品通用名称”均为“掺混肥料”，根据产品通用名称的命名规则，掺混肥料属于复混肥料的一种，故将其“产品通用名称”更改为“复混肥料”。序号26的有机一无机复混肥料的“技术指标”为“1、养分(N+P₂O₅+K₂O)≥25%2、有机质≥15%”还有多余数值，为方便进一步计算，将其更改为“养分(N+P₂O₅+K₂O)≥25% 有机质≥15%”。序号59的“技术指标”中“P₂O₅”为拼写错误，将其更改为“P₂O₅”，将序号38“技术指标”的内容“总养分(N+P₂O₅+K₂O)，% ≥10有机质 % ≥60”更改为“总养分(N+P₂O₅+K₂O)≥10%有机质≥60%”。

(4)任务4.2数据预处理

对“附件4.xlsx”的数据针对“原料与占比”进行异常值处理，序号23和51的“原料与占比”存在数据缺失，做删除处理。序号56、57、66和198存在空字符串情况，进行格式修改处理。

2. 肥料产品的数据分析

(1)任务2.1

从图1可以看出所有复混肥料中产品登记数量集中分在第5，6，7和8组。其中，第7组的数量最多，为2098个，第2,3组的数量最少，为0个。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如表1所示:

表1 产品登记数量最大的前3个分组表

排名	一	二	三
分组标签	7	6	5
产品登记数量	2098	1684	1040

(2)任务2.2

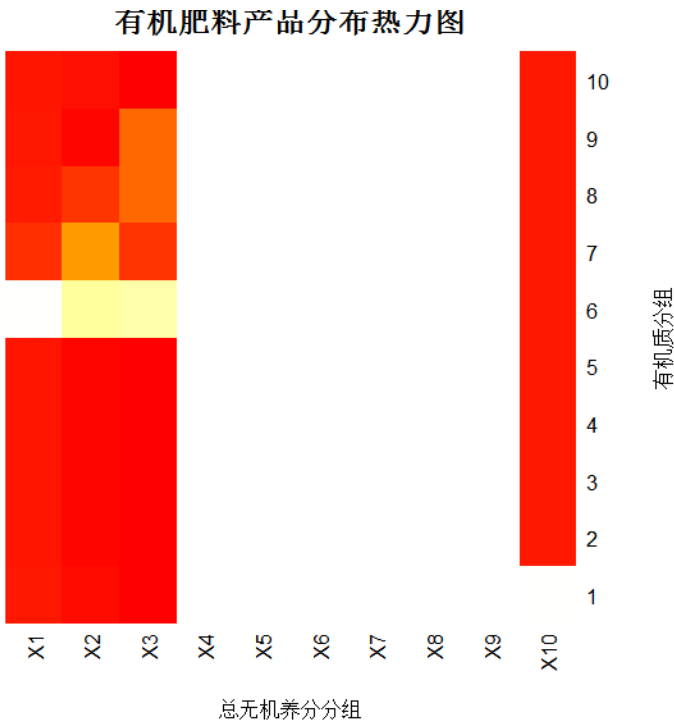


图2 有机肥料产品的分布热力图

3. 肥料产品的多维度对比分析

(1)任务3.1

4. 统计分析

- (1) 任务4.1
- (2) 任务4.2