

基于统计评价的肥料登记数据分析

摘要

我国为农业大国，肥料使用量较大，肥料的安全使用需引起重视，对农药的生产销售也有严格的规定。分析肥料登记数据有重要意义。

关键词：规范化处理；**k-means++**聚类；布朗运动预测；

一、任务重述

1.1 任务背景

人口的日益增长，增加了对粮食的需求，我国身为农业大国，对肥料的需求与使用量都较大，但肥料的安全使用一直存在问题，所以对农药的生产销售也有严格的规定。并制定了相关方案对此进行约束。合理规范肥料使用也是推动农业高质量发展的必要举措。根据产品登记信息，初步了解肥料产品品种、研发创新等情况有重要意义。

1.2 任务重述

对各地区农业行政管理部门的肥料登记数据进行分析处理，得到有效信息。

任务一：

由于产品的名称会不规范，需要规范化处理附件 1 中的肥料数据，便于后续分析。在此基础上计算出总无机养分百分比，完成数据预处理。

任务二：

对附件 2 中的肥料产品以总无机养分百分比的取值进行细分，将其进行分组处理。并分析其分布特点，绘制直方图进行展示。

对附件 2 中的肥料产品以总无机养分百分比和有机质百分比的取值进行细分，将其进行分组处理。绘制热力图进行展示并分析其分布特点。

对附件 2 中的肥料产品以氮、磷、钾养分的百分比的取值进行聚类，并绘制三维散点图和散点图矩阵，再通过雷达图分析聚类特征。

任务三：

分析不同组别的复混肥料在不同年份产品登记数量的变化趋势。

提取到 2021 年 9 月 30 日仍有效的有机肥料产品，并分析在广西和湖北两省的分布差异。

分析提取产品登记数量大于 10 的企业的原料集合。并计算这些企业之间的杰卡德相似系数矩阵。

任务四：

设计算法得出氮、磷、钾养分，有机质的百分比和肥料含氯的程度。

设计算法得出各原料的名称及其百分比。

二、模型假设

联系实际，为保证模型求解的准确性与合理性，排除一些干扰，本文提出以下假设：

- 1.假设有效期的天数均位于月末。
- 2.假设各省的肥料均只为本省提供。
- 3.假设磷与钾的含量只存在于 P_2O_5 与 K_2O 中。

三、任务求解

5.1 任务一求解

5.1.1 任务一的分析

任务要求首先对肥料数据进行数据预处理，由于产品的名称存在不规范的情况，对附件 1 中的肥料数据进行规范化处理，对各肥料名称进行分类并对不规范的名称进行替换便于后续分析。在此基础上计算出各肥料产品的氮、磷、钾养分百分比之和，得到总无机养分比，完成数据预处理。



图 1 任务一分析图

5.1.2 数据预处理

(1) 规范化处理

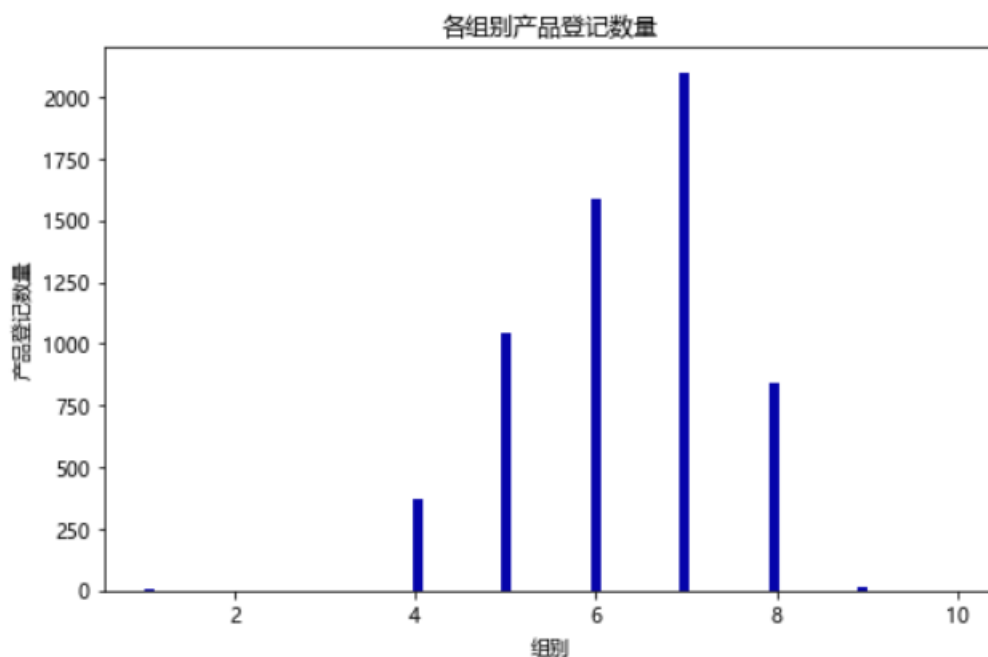


图 5 直方图

◆ 分布特点

根据直方图展示数据可知，都没有分布在 2, 3 两组的数据，即 0.072-0.216 范围内，集中分布在 5, 6, 7, 三组的范围内，即 0.360-0.504 范围内。其中第 7 组集中分布最多，占比为 35.2%。第 1 组，第 9 组和第十组中分布数据较少。

第七组的占比最大，有 2098 组数据，占比为 35.2%；

其次为第六组，有 1584 组数据，占比为 26.6%；

然后为第五组，有 1040 组数据，占比 17.4%。

排名	一	二	三
分组标签	7	6	5
产品登记数量	2098	1584	1040

(2) 有机肥料分析

在附件 2 中筛选出有机肥料，以总无机养分百分比和有机质百分比作为关键词，描述性统计出最大值与最小值，依照公式（1）将最大值与最小值的和等值划分为十组。

◆ 处理结果

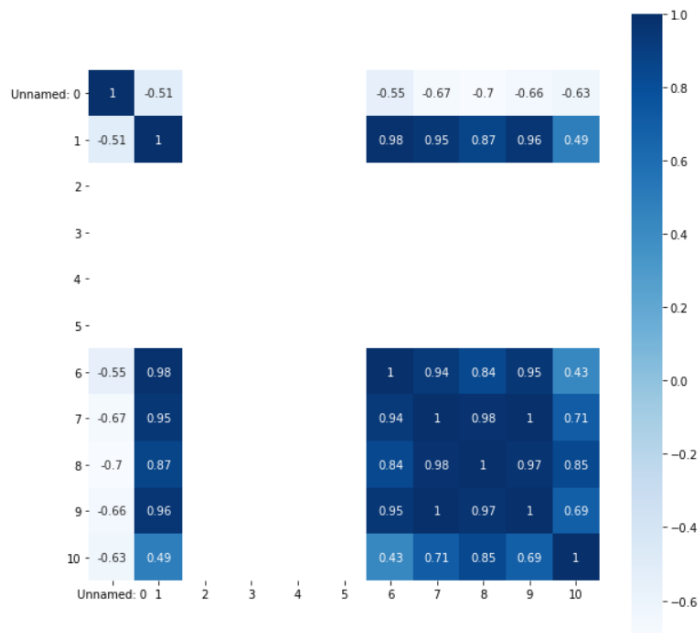


图 6 热力图

◆ 分布特点

将有机肥料产品的产品数量设为温度，颜色越深，温度越高，数量越大，由图得到产品集中分布在（1,6），（1,7），（2,6）。其中（1,6）中的产品分布数量为 843；（1,7）中的产品分布数量为 68；（2,6）中的产品分布数量为 60. 为登记数量最大的前 3 个分组。

（2）复混肥料分析

◆ k-means++聚类

问题要求使用聚类算法按照氮、磷、钾养分的百分比，将复混肥料分为 4 类。k-means++均值聚类能便捷、快速的将数据进行分群，计算量小且可以通过最终的聚类中心与指标间的距离进行判断。

重心公式：

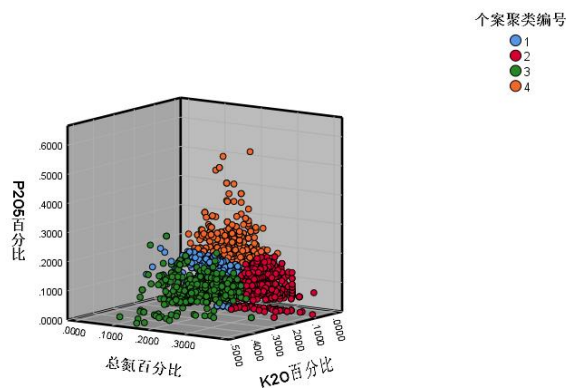


图 8 三维散点图

四、参考文献

[1] 杨真真,孙艳彬. 基于模糊聚类的城市空气环境质量分析——以黑龙江省为例

[J]. 中国科技信息. 2021(05)

[2] 佟汉文,朱展望,刘易科等. 三种肥料运筹模式对湖北西部坡耕地机条播小麦生产的影响[J]. 应用生态学报.2021(01)

[6] 梁运献,余焘,阮坤良等. 广西常用无机肥料中 5 种有害元素含量调查分析[J]..中国土壤与肥料 . 2021(02)