

肥料登记数据分析

摘要

近年来，肥料的管理逐渐被人们所重视。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》。《肥料登记管理办法》是为了加强肥料管理，保护生态环境，保障人畜安全，促进农业生产而制定的办法。现对安徽、广西、湖北和某省肥料登记数据进行处理和分析。

本文基于多个肥料登记数据，对肥料登记数据进行预处理；根据养分的百分比对肥料产品进行细分；从省份、日期等维度对肥料登记数据进行对比分析；对非结构化数据进行结构化处理。

本文主要工作包括以下三个方面：

1. 基于所给数据集进行数据预处理，包括不规范命名的规范化处理，计算肥料产品的养分百分比之并进行整合保存。
2. 针对任务二，对复混肥料的总无机养分百分比和有机肥料的总无机养分百分比和有机质百分比进行分组并统计产品登记数量，根据各复混肥料的氮，磷，钾含量对产品进行聚类分析，将复混肥料产品分成了四类，并绘制三维散点图，散点图矩阵和雷达图，对每个类别的产品的氮，磷，钾含量进行数据可视化分析。
3. 多维度对比分析不同肥料产品，利用 matplotlib 将不同组别各年份产品登记数量的变化趋势用折线图可视化。绘制条形图，直观的分析广西和湖北两个省份每个组别的产品登记数量情况以及差异。分析广西和湖北在不同组别的差异，以原料集合作为特征计算出杰卡德相似系数矩阵。
4. 利用 python 的 re 正则表达式模块从众多不同肥料产品的说明中提取出氮、磷、钾养分和有机质的百分比，以及肥料含氯的程度并进行综合分析。

目录

肥料登记数据分析1

1.数据预处理.....3

 1.1 分析方法与过程3

 1.2 分析方法与过程3

2. 肥料产品的数据分析.....3

 2.1 分析方法与过程3

 2.2 分析方法与过程4

 2.3 分析方法与过程4

3. 肥料产品的多维度对比分析5

 3.1 分析方法与过程5

 3.2 分析方法与过程5

 3.3 分析方法与过程5

4. 肥料产品的多维度对比分析.....5

 4.2 分析方法与过程5

1.数据预处理

1.1 分析方法与过程

本任务旨在处理产品通用名称不规范的数据，分 4 类进行规范化处理.

1.2 分析方法与过程

本任务旨在计算附件 1 中各肥料产品的氮、磷、钾养分百分比之和。

2. 肥料产品的数据分析

2.1 分析方法与过程

本任务旨在根据每个产品所在的分组，为其打上分组标签。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。

绘制产品登记数量的直方图，

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，
绘制排名图：

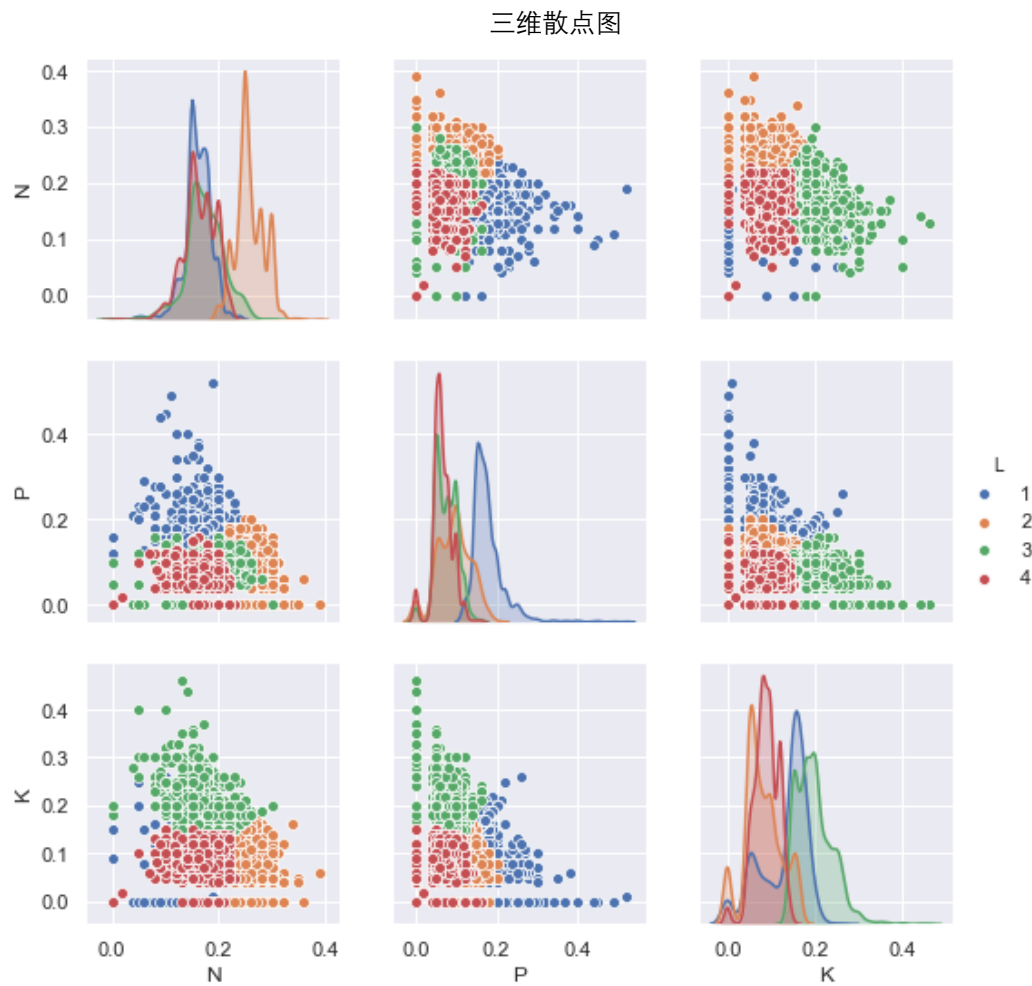
排名	一	二	三
分组标签	7	6	5
产品登记数量	2012	1501	1038

2.2 分析方法与过程

本任务旨在为每个产品打上分组标签。并根据分组情况绘制有机肥料产品的分布热力图，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。

2.3 分析方法与过程

本任务旨在筛选出复混肥料的产品，并通过氮，磷，钾对其进行聚类分析，并作图进行可视化。



3. 肥料产品的多维度对比分析

3.1 分析方法与过程

该任务旨在可视化不同组别的不同年份产品登记数量的变化趋势。

3.2 分析方法与过程

该任务旨在提取 result2_2.xlsx 中提取 2021 年 9 月 30 日仍有效的有机肥料产品，并筛选出广西和湖北产品登记数量在前 5 的组别，然后分析分析两个省份上述组别的分布差异。

3.3 分析方法与过程

该任务旨在提取产品登记数量大于 10 的肥料企业并给出这些企业所用到的原料集合（发酵菌剂除外）。所用到的原料集合（发酵菌剂除外）。以各企业用到的原料作为特征，计算企业之间的杰卡德相似系数矩阵。

4. 肥料产品的多维度对比分析

4.2 分析方法与过程

该任务旨在从技术指标中提取出氮、磷、钾养分和有机质的百分比，以及肥料含氯的程度。