摘要

肥料是农业生产中一种重要的生产资料，是粮食的“粮食”。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，各省政府根据《中华人民共和国农业法》制定了相关的肥料登记管理方法。化肥促进了粮食和农业的生产发展，但目前仍存在化肥盲目使用及过度使用肥料等问题，带来了成本的增加和环境的污染。对肥料做好登记管理工作，通过分析各省的肥料使用情况，严把肥料产品有效、安全和适用准入关，促进了肥料科学、经济和生态施用，提高了农产品质量安全水平，维护了农民群众利益，保障生产、节本增效；在减少化肥不合理的投入之外，鼓励研制、生产和使用安全、高效、经济的肥料产品，确保粮食在化肥的作用下稳定增产、使得农民持续增效、确保农业的可持续发展。因此，通过对各省的肥料使用情况进行处理分析，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，并对非结构化数据进行结构化处理，对中国农业的可持续发展具有重大意义。

Abstract

Fertilizer is an important means of production in agricultural production and the "grain" of grain. In order to strengthen fertilizer management, protect the ecological environment, ensure human and livestock safety and promote agricultural production, provincial governments have formulated relevant fertilizer registration management methods in accordance with the agricultural law of the people's Republic of China. Chemical fertilizer has promoted the production and development of grain and agriculture, but there are still problems such as blind use and excessive use of chemical fertilizer, which has brought increased costs and environmental pollution. We have done a good job in the registration and management of fertilizers. By analyzing the use of fertilizers in various provinces, we have strictly controlled the effective, safe and applicable access of fertilizer products, promoted the scientific, economic and ecological application of fertilizers, improved the quality and safety level of agricultural products, safeguarded the interests of farmers, and ensured production, cost saving and efficiency increase; In addition to reducing the unreasonable input of chemical fertilizer, encourage the development, production and use of safe, efficient and economic fertilizer products to ensure the stable increase of grain production under the action of chemical fertilizer, enable farmers to increase efficiency and ensure the sustainable development of agriculture. Therefore, it is of great significance to the

sustainable development of China's agriculture to process and analyze the fertilizer use in each province, compare and analyze the fertilizer registration data from the dimensions of province, date, manufacturer and fertilizer composition, and structurally process the unstructured data.

For task one, preprocess the data provided in Annex 1. Data preprocessing is mainly divided into normalizing the generic names of non-standard products in the annex and calculating the percentage of total inorganic nutrients. First, use drop in pandas_ The duplicates () function finds the generic name of the non-standard product. Secondly, it normalizes the name according to the requirements of the standard name. For the calculation of the percentage of total inorganic nutrients, the percentage of nitrogen, phosphorus and potassium nutrients can be added by using sum () function; Then use the format() function to keep three decimal places.

For task two, firstly, the compound fertilizer products are sorted according to the total inorganic percentage and divided into 10 groups. Focus on the distribution characteristics of compound fertilizer products, draw the histogram of product registration quantity by using the distribution characteristics of products by tableau, and further analyze the distribution of compound fertilizer products. Secondly, organic fertilizer products were equally divided into 10 groups according to the percentage of total inorganic nutrients and the percentage of organic matter. Then use tableau to draw the thermal diagram of organic fertilizer products according to the grouping, and further analyze the organic fertilizer products. Finally, the birch clustering algorithm in machine learning is mainly used to draw the three-dimensional scatter diagram, scatter diagram matrix and radar diagram of clustering results of each compound fertilizer product according to the different inorganic components, and analyze the characteristics of each cluster.

For task three, first, analyze the product data of compound fertilizer in each year directly through the data in Annex 3, obtain the change trend of product registration quantity of compound fertilizer in each year from 2013 to 2020, and visually present the change trend by drawing a broken line diagram. Secondly, the validity period is timed, and the organic fertilizer products still valid on September 30, 2021 are obtained by comparing the time, and then grouped and sorted. Finally, create the matrix through dataframe () in pandas, and then use the for loop to calculate the similarity coefficient of each enterprise according to the use of raw materials, and make further analysis.

For task four, the replace() function is mainly used to split parentheses, replace parentheses with commas, and then replace multiple commas with one comma; Then use the replace () function to divide the comma into an array; Then string the data; Finally, the regular function is used to judge the data, load the new table, and finally analyze the content

of various raw materials.

目录

1、问题分析

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。

2、任务一 数据的预处理

2.1 产品通用名称规范化

2.2 计算各类养分的数据

3、任务二 肥料产品的数据分析

3.1 复混肥料的产品数据分析

3.1.1 复混肥料产品养分分析

3.1.2 复混肥料产品分布情况分析——基于纵向直方图

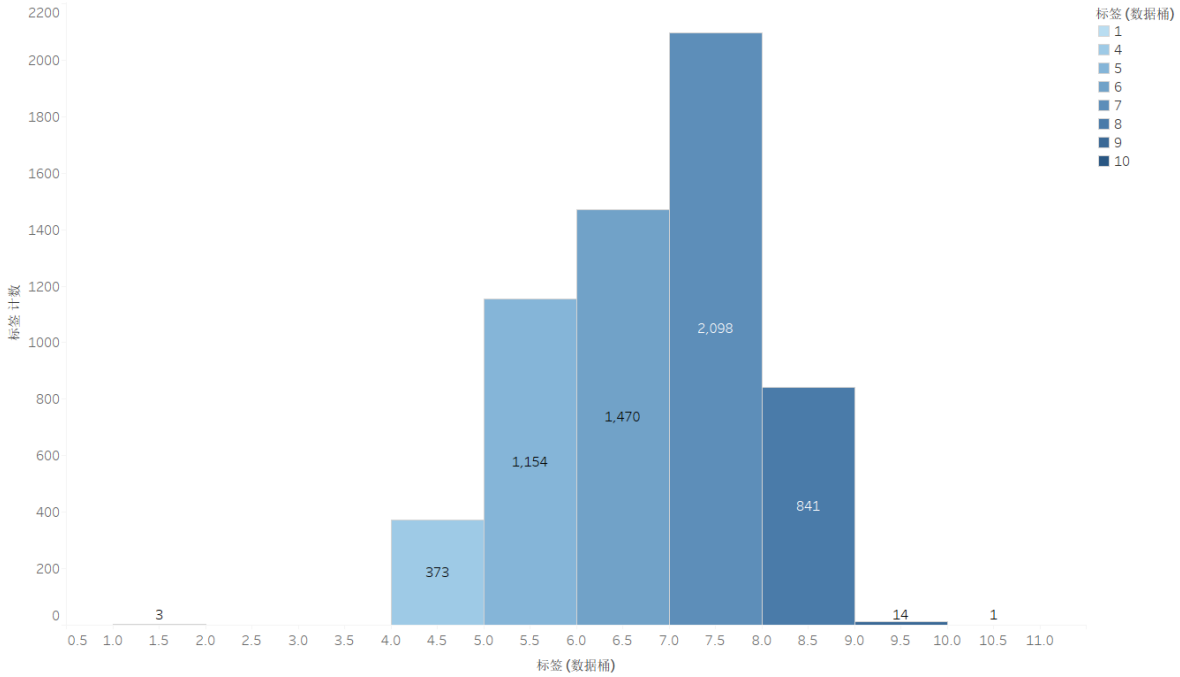

图 5 产品登记数量直方图

3.1.3 重点分组登记数量分布情况分析——基于横向直方图

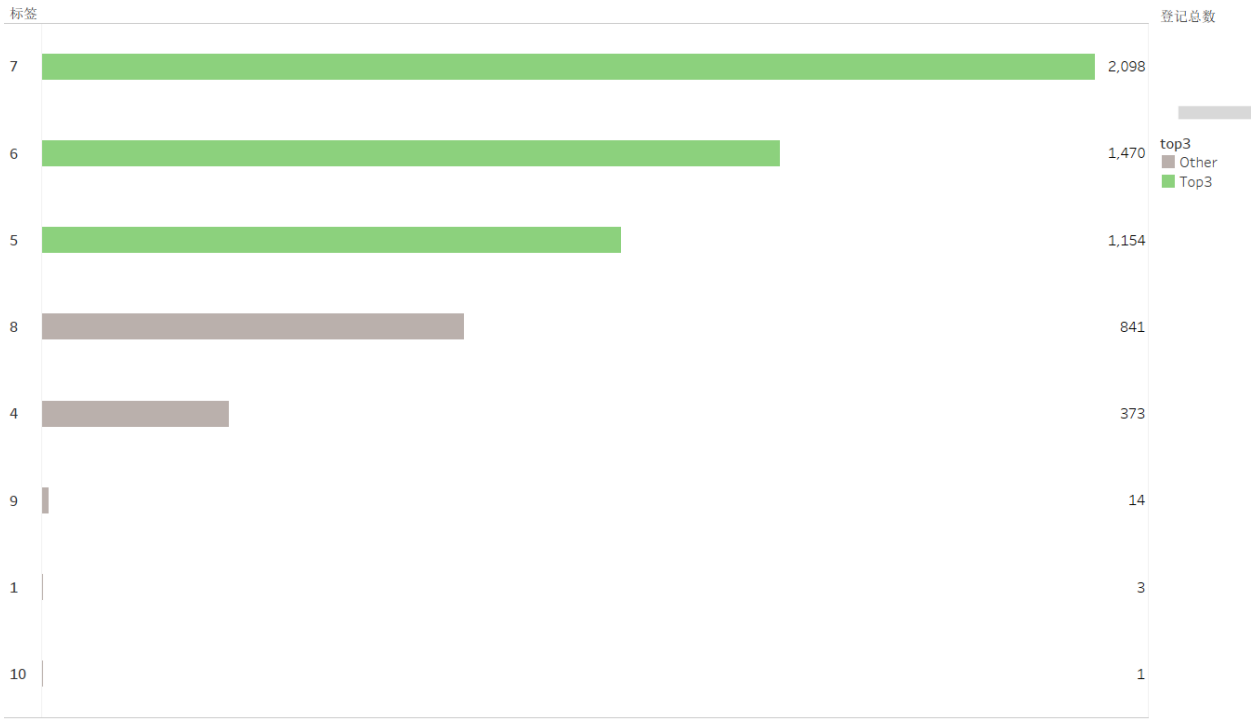

图 6 登记数量排行直方图

表 4 登记数量排行前三的具体数据

排名	一	二	三
分组标签	7	6	5
产品登记数量	2098	1470	1154

3.2 有机肥料的产品数据分析

有机肥料亦称“农家肥料”。凡以有机物质(含有碳元素的化合物) 作为肥料的均称为有机肥料。包括人粪尿、厩肥、堆肥、绿肥、饼肥、沼气肥等。具有种类多、来源广、肥效较长等特点。有机肥料所含的营养元素多呈有机状态，作物难以直接利用，经微生物作用，缓慢释放出多种营养元素，源源不断地将养分供给作物。施用有机肥料能改善土壤结构，协调土壤中的水、肥、气、热，提高土壤肥力和土地生产力。由于其来源广，种类多的特点，因此，研究有机肥料产品的数据并对其进行分析，对中国这个农业大国的农业持续发展具有重大意义。

3.2.1 有机肥料产品养分分析

3.2.2 有机肥料产品分布情况分析——基于热力图

3.2.3 重点分组登记数量分布情况分析——基于横向直方图

图 10 部分结果示意图

3.3.2 基于肥料产品散点图的分析

图 11 肥料产品的散点图矩阵

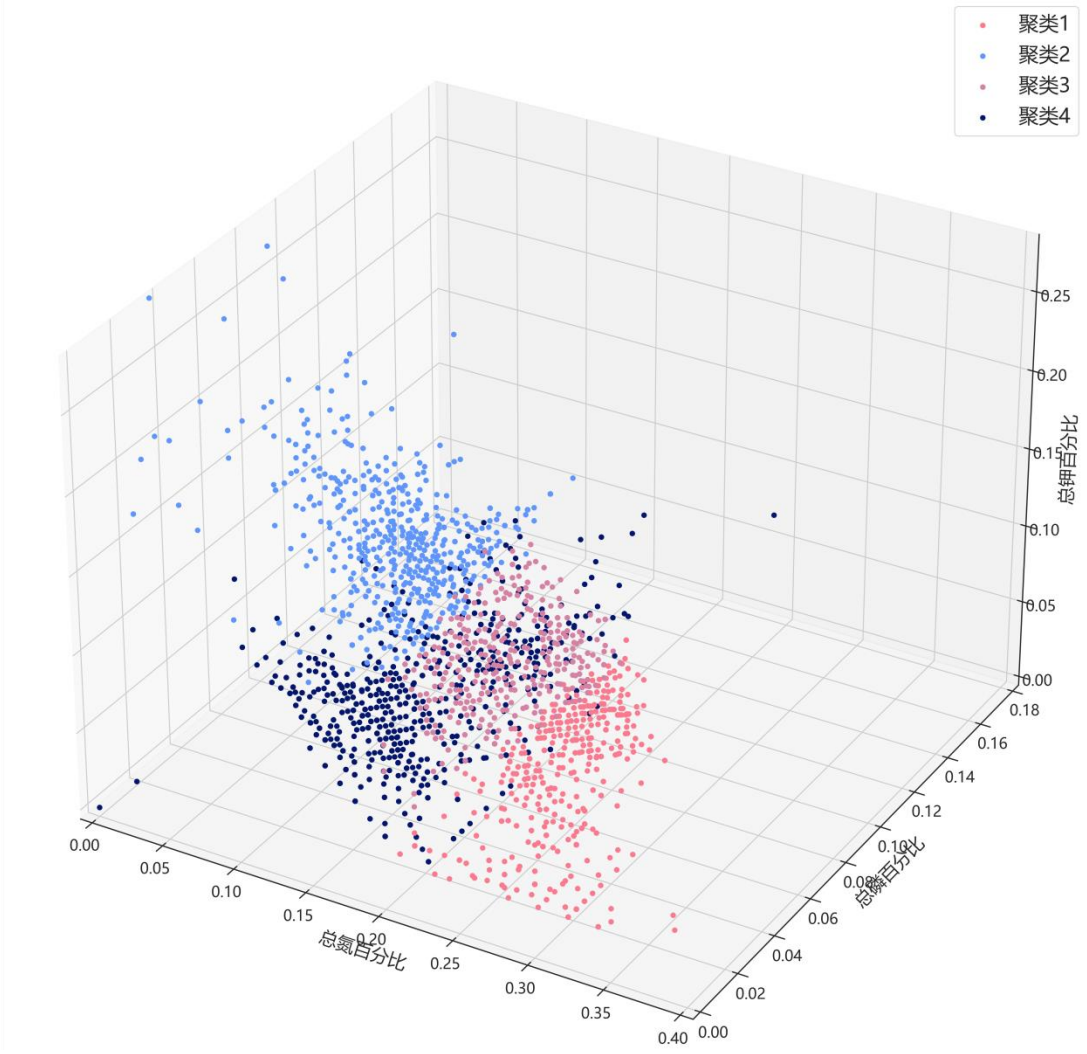

图 12 肥料产品的三维散点图

3.3.3 聚类特征

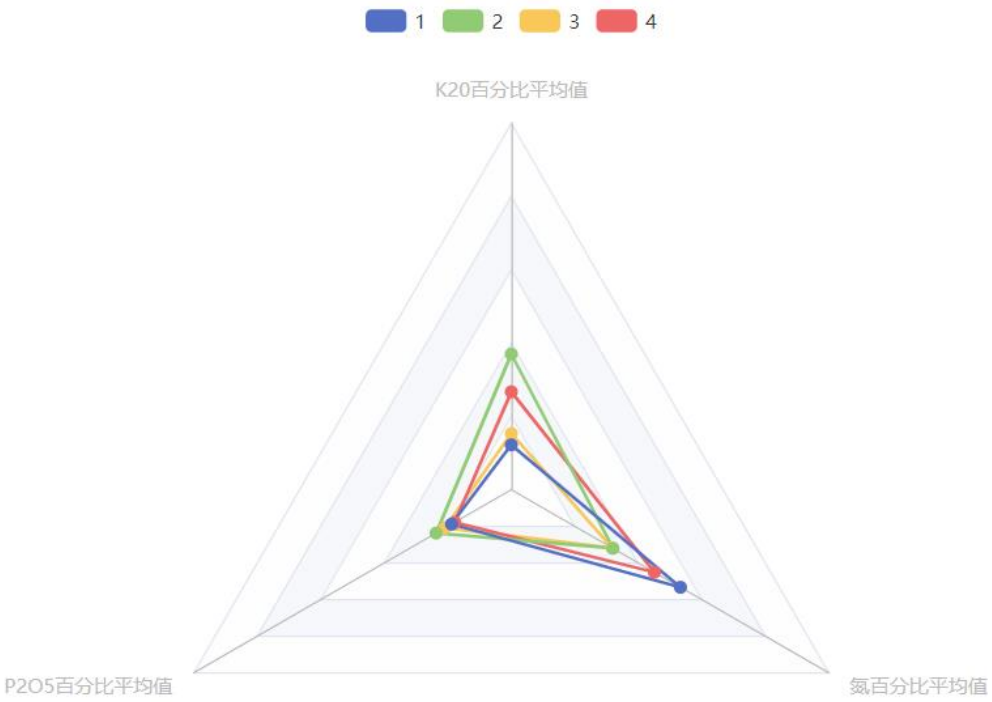

图 17 聚类结果的总雷达图

7、参考文献

[1]中华人民共和国农业部.农业部关于切实做好肥料登记管理工作的通知:中华人民共和国农业农村部,2009-05-20 第五期

[2]广东省人民政府.肥料登记管理方法：中华人民共和国司法部网站,2019-06-24

[3]中国国家标准《肥料术语及其定义》,GB/T6274-1997

[4]中华人民共和国农业部.行业分类农、林、牧、渔业：中华人民共和国农业农村部,2012-06-01

[5]产业洞察研究. 中国硫酸铵产业运行态势：产业洞察研究，2013