

# 肥料登记数据分析

## 摘要

我国的肥料产品登记制度规定企业开发的新产品需要在农业农村部进行登记后方可生产、经营和使用。但是对于农田长期使用、有国家或行业标准的产品免于登记,具体包括硫酸铵(锌),尿素,硝酸铵,氰氨化钙,磷酸铵(磷酸一铵、二铵),硝酸磷肥,过磷酸钙,氯化钾,硫酸钾,硝酸钾,氯化铵,碳酸氢铵,钙镁磷肥,磷酸二氢钾,单一微量元素肥和高浓度复合肥。此外,“肥料和农药的混合物”“农民自制自用的有机肥料”“植物生长调节剂”也无需在农业农村部进行登记。

**关键词:** 数据分析 复混肥料 产品登记数量 pandas 数据分析 Matlab 可视化

# 目录

一、问题背景 .....	1
二、软件与技术介绍 .....	1
2.1 大数据 .....	1
2.2 大数据分析 .....	2
2.3 数据可视化 .....	3
三、任务一：数据预处理 .....	4
3.1 通用名称规范化处理 .....	4
3.2 氮、磷、钾养分百分比之和 .....	4
四、任务二：肥料产品数据分析 .....	4
4.1 基于复混肥料有机养分百分比分析 .....	4
4.2 筛选有机产品并分组 .....	4
4.2.1 基于总有机养分与无机养分百分比等距分组 .....	4
4.3 基于氮、磷、钾养分的百分比聚类算法 .....	4
五、任务三：肥料产品多维度对比分析 .....	4
5.1 不同年份产品登记数量变化趋势分析 .....	4
5.2 提取有效的有机产品 .....	4
5.3 基于产品登记数量大于 10 的肥料企业进行分析 .....	5
七、总结 .....	5
八、参考文献 .....	5

## 一、问题背景

肥料是指能供给作物生长发育所需养分，改善土壤性状，提高作物产量和品质的物质。是农业生产中的一种重要生产资料。一般分为有机肥料、无机肥料、生物性肥料。也可按来源分为农家肥料和化学肥料。按所含养料的多少分为完全肥料和不完全肥料；按供肥的特点，分为直接肥料和间接肥料；按所含成分，分为氮素肥料、钾素肥料、微量元素肥料和稀土元素肥料。

我国化肥产业自 1949 年开始迅速发展，发展顺序是先氮肥、后磷肥、再钾肥、复合（混）肥，常规化肥品种的发展趋势基本遵循着世界化肥产业发展的主流趋势，即向高浓度和复合化方向迈进。我国化肥的格局是：氮肥基本自给自足，磷肥少量靠进口，钾肥则大部分需要进口。

化肥历来是农作物增产的重要因素，对我国粮食增产的贡献率高达 40%。然而，长期以来化肥施用不合理，给耕地质量带来了一系列的问题，化肥实际利用率不足 30%，土壤污染问题严重，污染事故频发，全国耕种土地面积的 10%以上已受重金属污染，其中华南部分城市约有一半耕地遭受镉、砷、汞等有毒重金属和石油类有机物污染。伴随着这些问题的出现，新型肥料应运而生。新型肥料指具有一定肥料功能和一定增产作用，由肥料生产企业研制生产，符合行业或企业标准，含有机、无机、微生物菌剂或兼有上述成分，经物理、化学、生物方法加工处理的新品种肥料。新型肥料可有效改善传统肥料利用率低、过量施用等造成的一系列问题，正确施用可提高肥料利用率，达到增效增产、减量降污的目的，是发展高效、绿色、可持续农业的必然要求。农业部门对化肥用料进行登记处理，严格把关化肥的用料，以此来减少化肥使用过程中土壤的伤害。

## 二、软件与技术介绍

### 2.1 大数据

大数据（bigdata）或称巨量资料，是指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合，是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。大数据有大量（Volume）、高速（Velocity）、多样（Variety）、低价值密度（Value）、

真实性 (Veracity) 五大特点。它并没有统计学的抽样方法，只是观察和追踪发生的事情。大数据的用法倾向于预测分析、用户行为分析或某些其他高级数据分析方法的使用。

## 2.2 大数据分析

大数据分析是指对规模巨大的数据进行分析。IBM 提出了大数据 “5V” 特点：

1. Volume：数据量大，包括采集、存储和计算的量都非常大。大数据的起始计量单位至少是 P（1000 个 T）、E（100 万个 T）或 Z（10 亿个 T）。

2. Variety：种类和来源多样化。包括结构化、半结构化和非结构化数据，具体表现为网络日志、音频、视频、图片、地理位置信息等等，多类型的数据对数据的处理能力提出了更高的要求。

3. Value：数据价值密度相对较低，或者说是浪里淘沙却又弥足珍贵。随着互联网以及物联网的广泛应用，信息感知无处不在，信息海量，但价值密度较低，如何结合业务逻辑并通过强大的机器算法来挖掘数据价值，是大数据时代最需要解决的问题。

4. Velocity：数据增长速度快，处理速度也快，时效性要求高。比如搜索引擎要求几分钟前的新闻能够被用户查询到，个性化推荐算法尽可能要求实时完成推荐。这是大数据区别于传统数据挖掘的显著特征。

5. Veracity：数据的准确性和可信赖度，即数据的质量。

大数据作为时下最火热的 IT 行业的词汇，随之而来的数据仓库、数据安全、数据分析、数据挖掘等等，随着大数据时代的来临，大数据分析也应运而生。而 Python 为数据分析提供诸多工具，anaconda 是其中一个著名的科学计算发行版，包括近 200 多个工具包，常见的有 numpy, scipy, pandas, ipython, matplotlib, sklearn 等。Python 数据分析流程如图 1 所示。

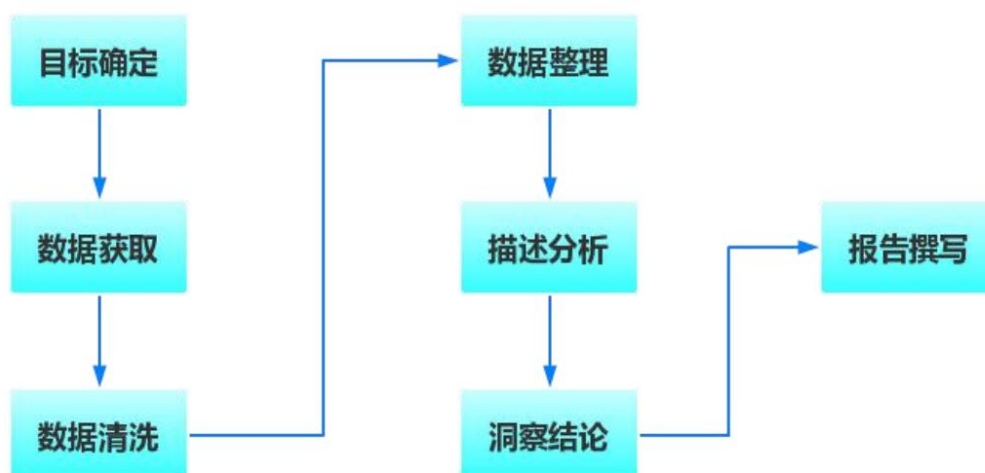


图 1 Python 数据分析流程

## 2.3 数据可视化

数据可视化主要旨在借助于图形化手段,清晰有效地传达与沟通信息。但是,这并不就意味着数据可视化就一定因为要实现其功能用途而令人感到枯燥乏味,或者是为了看上去绚丽多彩而显得极端复杂。为了有效地传达思想概念,美学形式与功能需要齐头并进,通过直观地传达关键的方面与特征,从而实现对于相当稀疏而又复杂的数据集的深入洞察。然而,设计人员往往并不能很好地把握设计与功能之间的平衡,从而创造出华而不实的数据可视化形式,无法达到其主要目的,也就是传达与沟通信息。其中数据用于聚焦于解决数据的采集,清理,预处理,分析,挖掘;图形用于聚焦于解决对光学图像进行接收、提取信息、加工变换、模式识别及存储显示;可视化用于聚焦于解决将数据转换成图形,并进行交互处理。数据可视化概念图如图 2 所示。

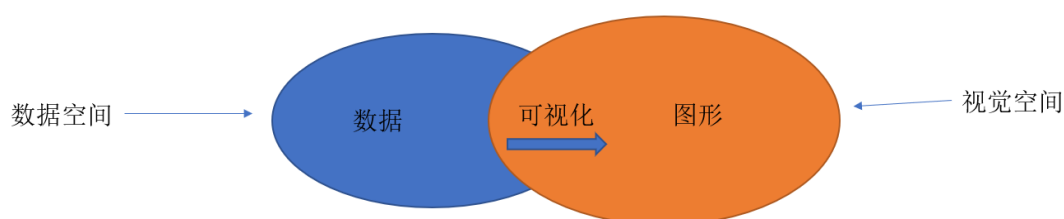


图 2 数据可视化概念图

三、任务一：数据预处理

3.1 通用名称规范化处理

3.2 氮、磷、钾养分百分比之和

四、任务二：肥料产品数据分析

4.1 基于复混肥料有机养分百分比分析

排名	二	一	三
分组标签	4	5	6
产品登记数量	1249	1586	849

复混肥料在第五组中分布最多，其次是第七组中所分布的较多，说明，复混肥料的无机养分总百分比是比较多的，产品登记数量的直方图如图 6 所示。

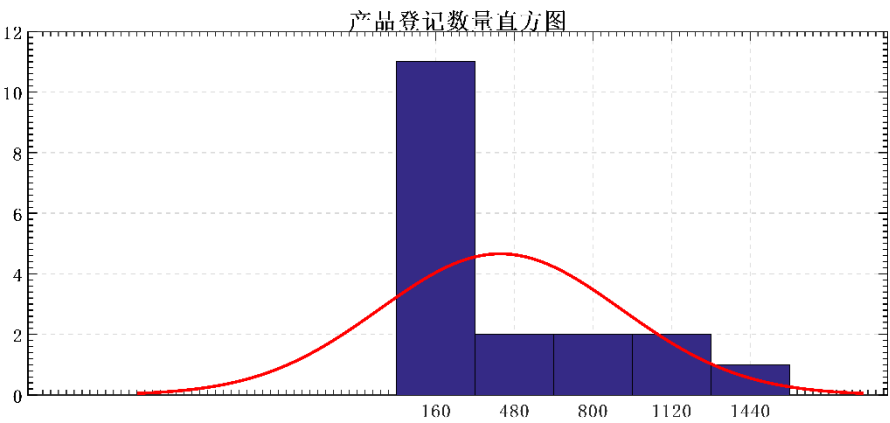


图 6 产品登记数量直方图

4.2 筛选有机产品并分组

4.2.1 基于总有机养分与无机养分百分比等距分组

4.3 基于氮、磷、钾养分的百分比聚类算法

五、任务三：肥料产品多维度对比分析

5.1 不同年份产品登记数量变化趋势分析

5.2 提取有效的有机产品

### 5.3 基于产品登记数量大于 10 的肥料企业进行分析

## 七、总结

## 八、参考文献

- [1]编辑. 国务院:取消七类肥料登记[J]. 中国农资, 2020(37).
- [2]丁焕新, 孙永泉, 陈吉, 等. “放管服”背景下地市级肥料登记管理工作探析[J]. 现代农业科技, 2019(22):2.
- [3]沈荣, 张保文. 大数据分析和大数据处理技术研究综述[J]. 电脑知识与技术, 2019, 15(11):19-22.
- [4]刘宗波. 一种解决农业霜冻的大数据分析处理系统及其工作方法:, CN110647194A[P]. 2020.
- [5]不公告发明人. 一种便于进行数据分析的统计用数据处理装置:, CN213186788U[P]. 2021.
- [6]谢雨呈, 谭长银, 张朝, 等. 典型肥料生产场地氨氮分布特征及风险控制目标确定[J]. 环境科学研究, 2019, 32(3):10.
- [7]王笑庸, 牛彦波, 殷博, 等. 生物肥料生产菌株对有机磷农药降解能力的比较研究[J]. 2021(2015-12):26-27.
- [8]郭利, 王学龙, 陈永德, 等. 有机肥料施用量对烟田土壤微生物的影响[J]. 湖北农业科学, 2009(11):3.
- [9] Oh T S , Kim S M , Kim C H . The Effect of Organic Fertilizer According to the Duration and Amount of Soil Chemical Changes on Yield Components of Rice[J]. Korean Journal of Cropence, 2014, 59(3):209-215.
- [10]西尾, 道德. 日本における化学肥料消費の動向と問題点[J]. 日本土壤肥料學雜誌, 2002, 73(2):219-225.