

肥料登记数据分析

摘要

我国社会经济高度增长，促进了农业领域的发展。土壤肥料在农业生产过程中有着重要的作用，他能给土地、农作物带去各种营养物质，大幅增加区域内农作物的产量^[1]，但因大量使用土壤肥料，导致的土壤问题，也不利于农业的可持续性发展，为增强对可持续化理念，国家设定了《肥料登记管理办法》，让土壤肥料不断科学化、规范化。^[2] 因此，对各肥料登记资料进行数据处理、数据分析有重要意义。

目录

1、 问题分析	4
2、 数据说明	4
3、 数据预处理	4
3.1 缺失值	4
3.1.1 缺失值说明	4
3.1.2 缺失值处理	4
3.2 异常值	4
3.2.1 异常值说明	4
3.2.2 异常值处理	4
3.3 重复值	4
3.3.1 重复值说明	4
3.3.2 重复值处理	4
4、 任务一	5
任务 1.1	5
任务 1.2	5
5、 任务二	5
任务 2.1	5
任务 2.2	6
任务 2.3	7
6、 任务三	8
任务 3.1	8
任务 3.2	8

任务 3.3	8
7、任务四.....	8
任务 4.1	8
任务 4.2	8
8、参考文献.....	8

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、数据说明

本次数据分析采用四个数据表，分别为安徽肥料登记数据、广西湖北肥料登记数据、某省登记肥料产品配方、某省肥料登记数据营养成分及原料构成，后文分别称作数据一、数据二、数据三、数据四。

3、数据预处理

3.1 缺失值

3.1.1 缺失值说明

3.1.2 缺失值处理

3.2 异常值

3.2.1 异常值说明

3.2.2 异常值处理

3.3 重复值

3.3.1 重复值说明

3.3.2 重复值处理

4、任务一

任务 1.1

任务 1.2

5、任务二

任务 2.1

从数据 2 筛选出复混肥料的产品，并查看结构，结果如图 17。将所有复混肥料按照总无机养分百分比的取值等距分为 10 组。根据每个产品所在的分组，为其打上分组标签（标签用 1~10 表示），结果如图 18。绘制产品登记数量的直方图，如图 19。

	G	H	I	J	K	L	M	N	U	F
1	K20百分比	含氯情况	有机质百分比	正式登记证号	发证日期	有效期	品商品名	适用作物	总无机养分百分比	分组标签
2	0.01	无氯	0	鄂农肥（2015）准字1379号	2015/11/30	2020-11			0.72	1
3	0.21	无氯	0	鄂农肥（2015）准字1359号	2015/11/27	2020-11			0.63	1
4	0.26	中氯	0	鄂农肥（2016）准字1991号	2016-08-22	2021-08			0.62	1
5	0.26	低氯	0	桂农肥（2017）准字2633号	2017-02-20	2022-02			0.62	1
6	0.2	无氯	0	鄂农肥（2013）准字0229号	2013-05-20	2018-05			0.6	1
7	0.2	无氯	0	鄂农肥（2013）准字0406号	2013-06-30	2018-06			0.6	1
8	0	无氯	0	鄂农肥（2014）准字0518号	2014-10-15	2019-10			0.6	1
9	0.2	无氯	0	鄂农肥（2015）准字0916号	2015-05-10	2020-05			0.6	1
10	0.06	无氯	0	鄂农肥（2015）准字0956号	2015-05-25	2020-05			0.6	1
11	0.2	无氯	0	鄂农肥（2016）准字1990号	2016-08-22	2021-08			0.6	1
12	0.2	无氯	0	鄂农肥（2017）准字2484号	2017-01	2022-01			0.6	1
13	0.2	无氯	0	桂农肥（2019）准字3639号	2019-11-04	2024-11	冻美		0.6	1
14	0.46	无氯	0	鄂农肥（2014）准字0586号	2014-12-15	2019-12			0.59	1
15	0.2	低氯	0	鄂农肥（2014）准字0167号	2014-05-25	2019-05			0.58	2
16	0.44	无氯	0	鄂农肥（2015）准字0730号	2015-02-25	2020-02			0.58	2
17	0.19	低氯	0	鄂农肥（2010）准字0437号	2015-02-25	2020-02			0.57	2
18	0.16	低氯	0	鄂农肥（2010）准字0883号	2015/08/20	2020-08			0.57	2
19	0.16	低氯	0	鄂农肥（2010）准字0887号	2015/08/20	2020-08			0.57	2
20	0.19	低氯	0	鄂农肥（2010）准字1034号	2015/10/23	2020-10			0.57	2
21	0.19	低氯	0	鄂农肥（2012）准字0500号	2015/07/30	2020-07			0.57	2
22	0.19	中氯	0	鄂农肥（2013）准字0173号	2013-04-19	2018-04			0.57	2
23	0.19	低氯	0	鄂农肥（2013）准字0389号	2013-07-30	2018-07			0.57	2
24	0.19	低氯	0	鄂农肥（2013）准字0661号	2013-11-15	2018-11			0.57	2
25	0.19	低氯	0	鄂农肥（2013）准字0803号	2013-12-25	2018-12			0.57	2
26	0.19	低氯	0	鄂农肥（2013）准字0805号	2013-12-25	2018-12			0.57	2
27	0.19	低氯	0	鄂农肥（2014）准字0005号	2014-01-20	2019-01			0.57	2
28	0.27	中氯	0	鄂农肥（2014）准字0029号	2014-01-15	2019-01			0.57	2
29	0.19	低氯	0	鄂农肥（2014）准字0184号	2014-05-05	2019-05			0.57	2
30	0.19	低氯	0	鄂农肥（2015）准字0759号	2015-02-25	2020-02			0.57	2
31	0.19	低氯	0	鄂农肥（2015）准字0760号	2015-02-25	2020-02			0.57	2

图 1 复混肥料产品分组

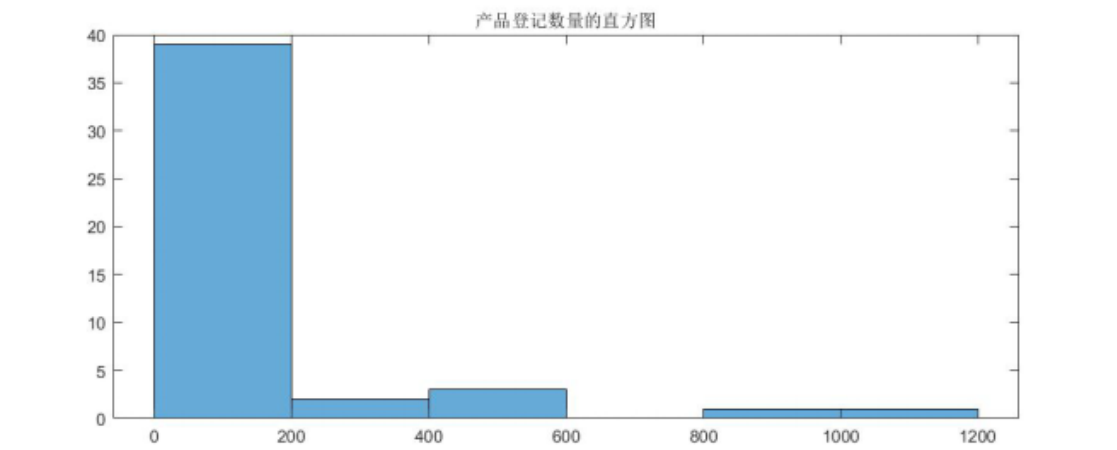


图 2 绘制产品登记数量直方图

排名	一	二	三
分组标签	7	8	6
产品登记数量	2012	1906	1020

任务 2.2

从数据二中筛选出有机肥料的产品，如图 20，将产品按照总无机养分百分比和有机质百分比分别等距分为 10 组，并为每个产品打上分组标签，部分结果如图 21。根据分组情况绘制有机肥料产品分布热力图，如图 22。为了利于分析，绘制了各公司登记的有机肥料

序号	企业名称	产品通用名	产品形态	总氮百分比	总磷百分比	总钾百分比	有机质百分比	含氯情况	登记证号	发证日期	有效期	商品名	适用作物	总无机养分百分比	分组标签
3137	湖北炎帝	有机肥料	粉状	0.16	0.07	0.2	无氯		0	鄂农肥2012016/01/2021-1				0.43	(1, 10)
25	广西宁明	有机肥料	粉剂	0.02	0.02	0.02	无氯		0.8	桂农肥20202020-12-12025-12		花生麸有机肥		0.06	(10, 1)
4855	湖北双赢	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.8	鄂农肥2012018/01/2023-01-01				0.0501	(10, 1)
1196	南宁市绿	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.8	桂农肥2012018-01-2023-01		绿佳益有机肥料		0.0501	(10, 1)
2235	广西群星	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.45	桂农肥2012020-12-12024-02		精制有机肥		0.0501	(10, 10)
3276	鄂州市广	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0	鄂农肥2012016/05/2021-05				0.0501	(10, 10)
3277	湖北杨林	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0	鄂农肥2012016/05/2021-05				0.0501	(10, 10)
3278	湖北政鑫	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0	鄂农肥2012016/05/2021-05				0.0501	(10, 10)
3279	黄石市华	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0	鄂农肥2012016/05/2021-05				0.0501	(10, 10)
3280	宜昌牧童	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0	鄂农肥2012016/05/2021-05				0.0501	(10, 10)
3290	武汉润土	有机肥料	粒状	0.0167	0.0167	0.0167	无氯		0	鄂农肥2012016/05/2021-05				0.0501	(10, 10)
2177	广西田东	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.75	桂农肥20202020-09-2025-09		有机肥料		0.0501	(10, 2)
1827	武汉东壮	有机肥料	粉状、粒剂	0.0167	0.0167	0.0167	无氯		0.7	鄂农肥2012014-04-2019-04				0.0501	(10, 2)
2526	湖北中瑞	有机肥料	粒状	0.0167	0.0167	0.0167	无氯		0.7	桂农肥2012015-04-12020-04				0.0501	(10, 2)
5105	湖北慈缘	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.7	鄂农肥2012018-12-12023-09-11				0.0501	(10, 2)
895	广西东土	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.7	桂农肥2012017-07-2022-07				0.0501	(10, 2)
1106	南宁正丰	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.7	桂农肥2012017-11-2022-11		有机肥料		0.0501	(10, 2)
1276	桂林正丰	有机肥料	粉、粒剂	0.0167	0.0167	0.0167	无氯		0.7	桂农肥2012019-03-2024-03		黑色金子		0.0501	(10, 2)
168	广西美耕	有机肥料	粉状	0.02	0.02	0.02	无氯		0.6	桂农肥2012017-06-2022-06		美耕有机肥料		0.06	(10, 5)
959	广西快生	有机肥料	粉状	0.02	0.02	0.02	无氯		0.6	桂农肥2012017-08-2022-08				0.06	(10, 5)
1352	广西邦禾	有机肥料	粉状	0.02	0.02	0.02	无氯		0.6	桂农肥2012018-04-2023-04				0.06	(10, 5)
1401	广西丰硕	有机肥料	粉剂	0.02	0.02	0.02	无氯		0.6	桂农肥2012018-06-2023-06		丰硕		0.06	(10, 5)
1404	广西桂林	有机肥料	粉剂	0.02	0.02	0.02	无氯		0.6	桂农肥2012018-06-2023-06		桔丰精制有机肥		0.06	(10, 5)
1702	广西大发	有机肥料	粉状	0.02	0.02	0.02	无氯		0.6	桂农肥2012019-06-2024-06		有机肥料		0.06	(10, 5)
2060	广西森荷	有机肥料	粉状	0.02	0.02	0.02	无氯		0.6	桂农肥20202020-06-2025-06		森荷二有机肥		0.06	(10, 5)
3241	湖北中瑞	有机肥料	粒状	0.0167	0.0167	0.0167	无氯		0.6	鄂农肥2012016/03/2021-03				0.0501	(10, 5)
4768	湖北万安	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.6	鄂农肥2012017-11-2022-11-12				0.0501	(10, 5)
4769	湖北万安	有机肥料	粒状	0.0167	0.0167	0.0167	无氯		0.6	鄂农肥2012017-11-2022-11-12				0.0501	(10, 5)
4987	湖北源丰	有机肥料	粒状	0.0167	0.0167	0.0167	无氯		0.6	鄂农肥2012018-05-12023-05-13				0.0501	(10, 5)
5073	湖北庄康	有机肥料	粉状	0.0167	0.0167	0.0167	无氯		0.6	鄂农肥2012018-08-2023-08-06				0.0501	(10, 5)
5074	湖北庄康	有机肥料	粒状	0.0167	0.0167	0.0167	无氯		0.6	鄂农肥2012018-08-2023-08-06				0.0501	(10, 5)

图 3 分组结果（部分）

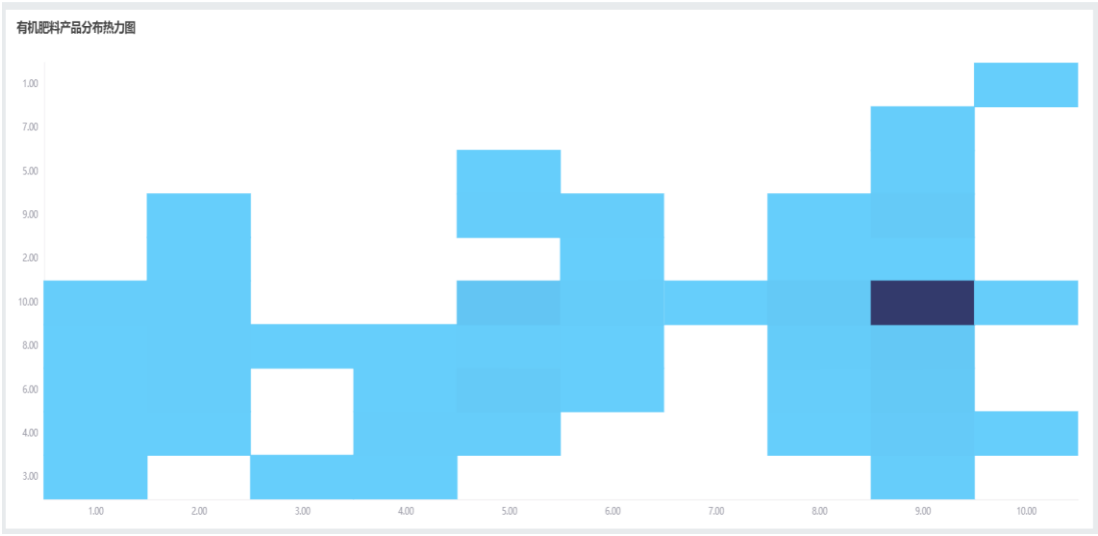


图 4 有机肥料产品分布热力图

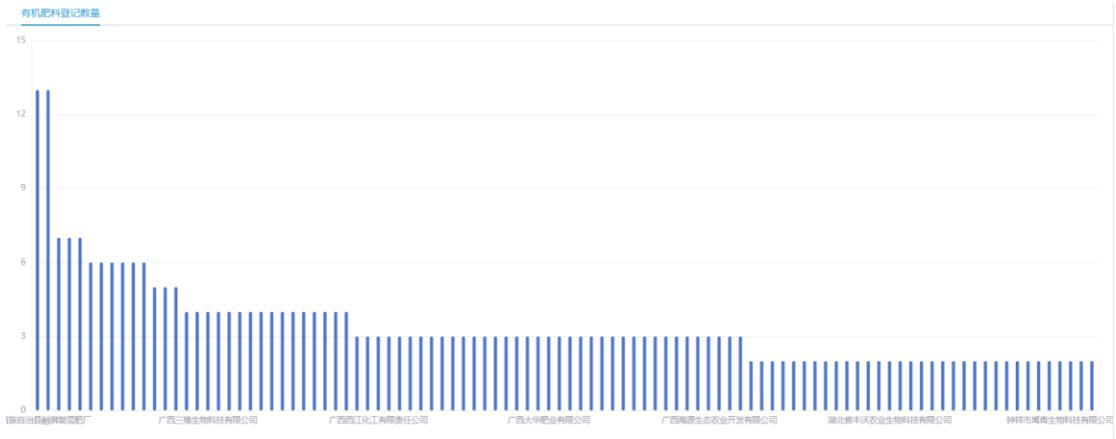


图 5 有机肥料登记图

任务 2.3

从附件 2 中筛选出复混肥料的产品，按照氮、磷、钾养分的百分比，使用聚类算法将这些产品分为 4 类, 部分结果如图 24，根据聚类标签绘制肥料产品的三维散点图，如图 25 并通过绘制聚类结果的雷达图如图 26。

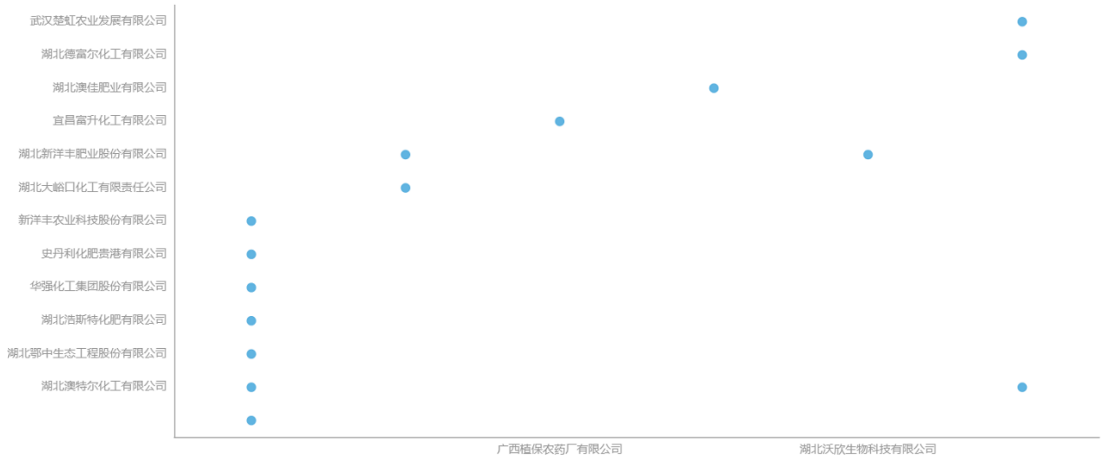


图 6 三维散点图

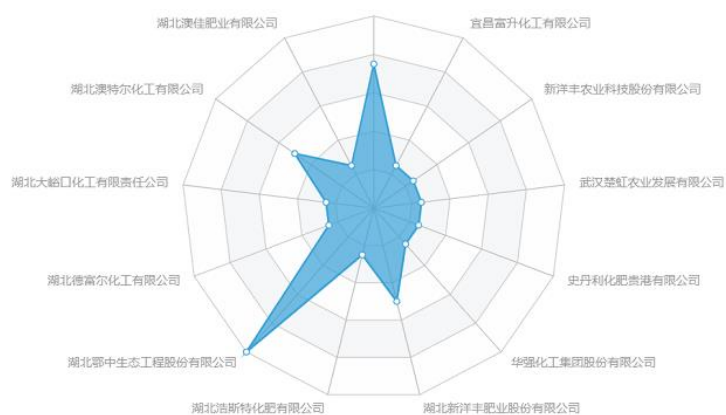


图 7 聚类结果的雷达图

6、任务三

任务 3.1

任务 3.2

任务 3.3

7、任务四

任务 4.1

任务 4.2

8、参考文献

- [1] 关菁,田伟,金萌.浅谈加强肥料登记管理促进农业可持续发展[J].科技风,2018(23):214.
- [2] 孙蓊锋,王旭,刘红芳,保万魁,侯晓娜.我国土壤调理剂中重金属元素及其相关原料农业资源化利用现状[J].中国土壤与肥料,2017(06):149-154.
- [3] 杨伟婷,李保育,左文斌.基于机器视觉的图像处理方法[J].信息技术与信息化,202107:143-145.
- [4] 张晓琳,付英姿,褚培肖.杰卡德相似系数在推荐系统中的应用
- [5] 季炼.Excel 在企业不同场景中的应用[J].中国市场,202127:101-102.