1.	背景	. 1
	数据预处理	
	7.1 夕称 抑	. – 2

1. 背景

肥料在农业生产中具有十分重要的意义,是农业中的重要生产资料,同时根据《肥料登记管理办法》规定,肥料需要按要求在农业行政管理部门进行登记。各省、市、自治区、直

辖市人民政府的农业主管部门进行登记。

本次竞赛使用的数据均来自农业部门官方网站,并记载了不同省份、企业生产的肥料登记信息及其含量组成、营养成分百分比等,以下为本组的数据分析报告。

2. 数据的预处理

2.1. 名称规范

1	序▼	企业名称 🔻	产品通用名称了	产品形▼	总氮百分 🕶	P205百分 ▼	K20百分▼	含氯情▼	有机质百分 🔻	正式登记证号	•	发证日↓▼	有效其▼
24	23	合肥佳明环保有限公司	有机肥料	粉状	1.67%	1.67%	1.67%	无氯	45.00%	皖农肥(2016)准字4277号		2016-01-17	2021-01
25	24	合肥佳明环保有限公司	有机肥料	颗粒	1.67%	1.67%	1.67%	无氯	45.00%	皖农肥(2016)准字4278号		2016-01-17	2021-01
171	170	怀远县新城区友好有机肥加工	有机肥料	颗粒	2.33%	2.33%	2.33%	无氯	45.00%	皖农肥(2016)准字4423号		2016-04-29	2021-04
174	173	淮南市益康达有机肥料有限公	有机肥料	粉状	1.67%	1.67%	1.67%	无氯	45.00%	皖农肥(2016)准字4426号		2016-05-11	2021-05
175	174	安徽省龙晨肥业有限公司	有机肥料	颗粒	2.00%	2.00%	2.00%	无氯	45.00%	皖农肥(2016)准字4427号		2016-05-11	2021-05
262	261	安徽昶春生物科技有限公司	有机肥料	颗粒	1.67%	1.67%	1.67%	无氯	45.00%	皖农肥(2016)准字4514号		2016-06-29	2021-06
263	262	滁州力大王农业科技有限公司	有机肥料	颗粒	2.00%	2.00%	2.00%	无氯	50.00%	皖农肥(2016)准字4515号		2016-06-29	2021-06
264	263	滁州力大王农业科技有限公司	有机肥料	粉状	2,00%	2,00%	2.00%	无氯	50,00%	皖农肥(2016)准字4516号		2016-06-29	2021-06
378	377	宿州市东方红肥业有限公司	有机肥料	颗粒	1,67%	1.67%	1.67%		45,00%	皖农肥(2016)准字4630号		2016-08-04	2021-08
521	520	定远县远景生物有机肥业有限:	有机肥料	颗粒	1.67%	1.67%	1.67%	无氯	45,00%	皖农肥(2016)准字4773号		2016-12-02	2021-12
624	623	安徽莱姆佳生物科技股份有限。	有机肥料	颗粒	2,67%	2, 67%	2.67%	无氯	45, 00%	皖农肥(2017)准字4876号		2017-01-25	2022-01
636	635	青阳县依本肥料有限公司	有机肥料	粉状	1.67%	1.67%	1.67%		60,00%	皖农肥(2017)准字4888号		2017-02-11	2022-02
669	668	安徽富民生物有机肥有限公司	有机肥料	颗粒	2, 67%	2, 67%			60,00%	皖农肥(2017)准字4921号		2017-03-06	2022-03
670			有机肥料	颗粒	1.67%	1.67%	1.67%			皖农肥(2017)准字4922号		2017-03-06	
671	670	灵璧县绿飨园现代农业有限公司	有机肥料	颗粒	1.67%	1.67%	1.67%	无氯	45, 00%	皖农肥(2017)准字4923号		2017-03-06	2022-03
738		安徽省绿福农业科技股份有限。		粉状	1.67%		1.67%			皖农肥(2017)准字4990号		2017-03-17	
766	765	安徽乐林农林科技有限公司	有机肥料	颗粒	1.67%	1.67%	1.67%		45, 00%	皖农肥(2017)准字5018号		2017-03-23	2022-03
767		阜阳市义丰肥业有限公司	有机肥料	颗粒	2.33%	2, 33%	2, 33%			皖农肥(2017)准字5019号		2017-03-23	2022-03
768		宿州騰飞有机肥料有限公司	有机肥料	颗粒	2,00%	2,00%	2.00%			皖农肥(2017)准字5020号		2017-03-23	
769		宿州天雨养殖(集团)有限公		颗粒	2,00%	2,00%	2,00%			皖农肥(2017)准字5021号		2017-03-23	
770		安徽友源有机肥有限公司	有机肥料	颗粒	1,67%	1.67%	1.67%			皖农肥(2017)准字5022号		2017-03-23	2022-03
876		安徽省文胜生物工程股份有限		粉状	1.67%	1.67%	1.67%			皖农肥(2017)准字5128号		2017-03-30	
977		· · · · · · · · · · · · · · · · · · ·		447.54	2.07%	2.07%	2.07%			台中間(2017)沿中日20日		2017-09-90	

图 1.筛选后使用自动填充工具完成名字的修改

2.2. 计算总无机养分百分比之和

	F2	• @ fx =SI	JM (B2:D2)				
4	Α	В	С	D	E	F	G
1	序号	总氮百分比。	P205百分比。	K20百分比。	正式登记证号	总无机养分百分比	
2	1	13.00%	17.00%	20.00%	皖农肥(2016)准字4255号 0	0.500	
3	2	13.00%	17.00%	20.00%	皖农肥(2016)准字4256号	0.500	
4	3	20.00%	15.00%	16.00%	皖农肥(2016)准字4257号	0.510	
5	4	26.00%	13.00%	12.00%	皖农肥(2016)准字4258号	0.510	
6	5	26.00%	13.00%	12.00%	皖农肥(2016)准字4259号	0.510	
7	6	20.00%	15.00%	16.00%	皖农肥(2016)准字4260号	0.510	
8	7	14.00%	26.00%	12.00%	皖农肥(2016)准字4261号	0. 520	
9	8	14.00%	26.00%	12.00%	皖农肥(2016)准字4262号	0. 520	
10	9	14.00%	20.00%	20.00%	皖农肥(2016)准字4263号	0.540	
11	10	27.00%	15.00%	12.00%	皖农肥(2016)准字4264号	0.540	
12	11	27.00%	15.00%	12.00%	皖农肥(2016)准字4265号	0.540	
13	12	28.00%	15.00%	12.00%	皖农肥(2016)准字4266号	0.550	
14	13	28.00%	15.00%	12.00%	皖农肥(2016)准字4267号	0.550	
15	14	20.00%	10.00%	15.00%	皖农肥(2016)准字4268号	0.450	
16	15	30.00%	0.00%	8.00%	皖农肥(2016)准字4269号	0.380	
					10.11.1m / > >80.13.1 = ==		

图 2 自动求和总无机养分百分比及其结果

3. 肥料产品的数据分析

3.1.复混肥料分组

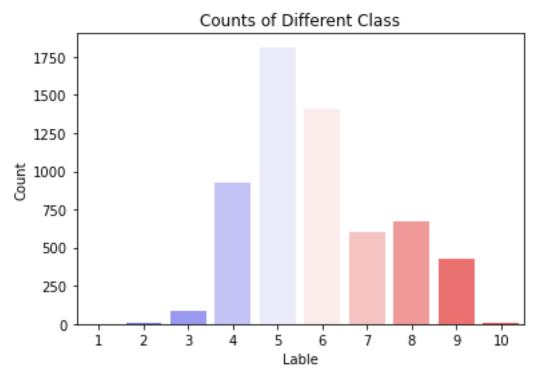


图 5.总无机养分百分比分布图

由图可知,大部分肥料集中于 5 与 6 两类中,而含量较高的类别里肥料数量有较大差值,相反在较低含量的分类中仍有相当一部分的肥料登记信息,可以发现肥料生产商对于控制成本的选择往往集中于降低养分原料的占比,从而达到更高的盈利能力。在图中可知,最高数量的组别分别是 5、6、4 三组,可以得知,虽然部分厂商为保证盈利而偷工减料,但是大部分的厂商可以将肥料中的养分保持在中等较高水平,这一现象可能是相关农业政策或农业部门的相关规定起到了积极作用。

排名	_	1_	111
分组标签	5	6	4
产品登记数量	1920	1495	930

3.2. 有机肥料分组

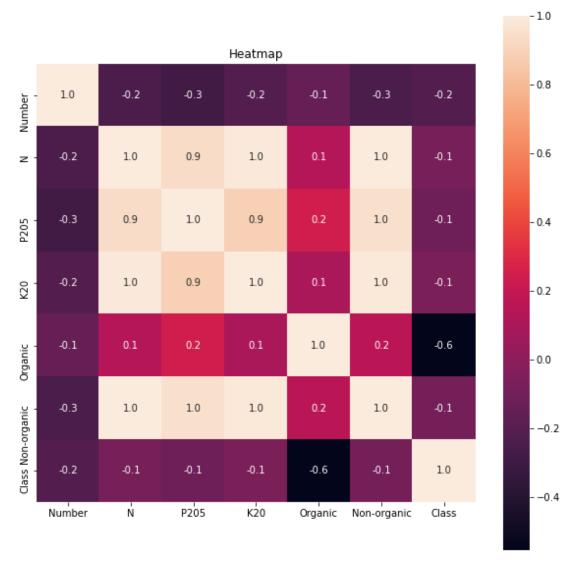


图 8.所有属性的热力分布图

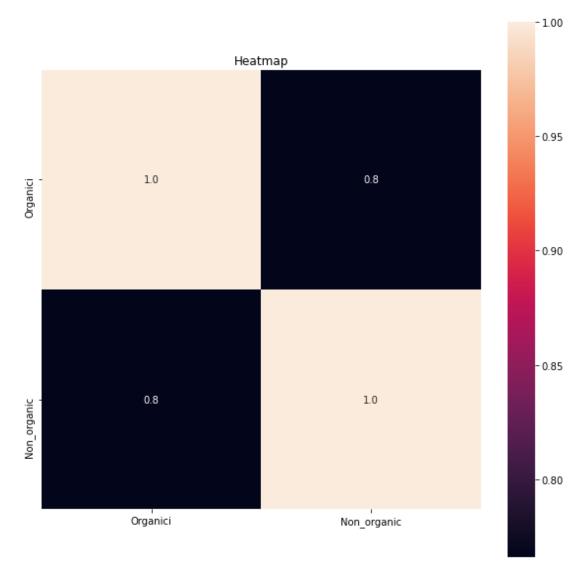


图 9.仅有有机肥料、无机养分两个属性的热力分布图

3.3. 聚类算法

在该部分任务中,我们使用 SPSS 作为分析工具,通过聚类选项设定 4 个分类后,得到分类完毕的数据集,并且在此基础上我们绘制了肥料产品的三维散点图与散点图矩阵如下所示。

复混肥料氮磷钾百分比3D散点图



组号

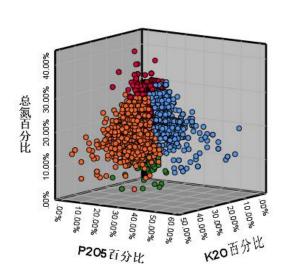


图 10.三位散点分布图

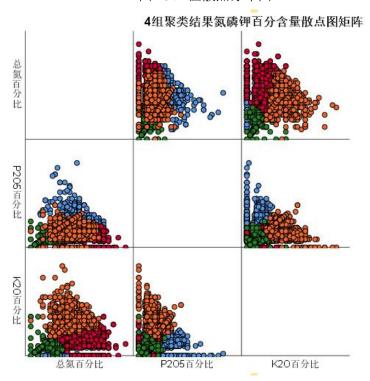


图 11.散点矩阵图

4. 多维度对比分析

4.1. 年份对比

年份	组1	组2	组3	组4	组5	组6	组7	组8	组9	组10
2012	3	4	3	0	22	3	5	0	5	2
2013	88	77	60	34	127	119	26	35	38	39
2014	116	89	88	71	126	103	57	70	50	53
2015	100	160	163	206	72	55	179	111	79	60
2016	106	134	149	161	117	67	192	146	128	129
2017	133	89	96	62	106	144	53	114	59	118
2018	29	17	11	24	15	68	26	49	79	69
2019	11	8	17	30	3	24	38	45	84	64
2020	8	17	9	6	7	10	18	32	73	62

图 12.年份分类结果

4.2. 区域分布差异对比

在该部分任务中,我们首先使用 FIND()函数提取出各公司所处省份,然后依据分组结果作出如下所示的地区分布图。



图 13.湖北分布图

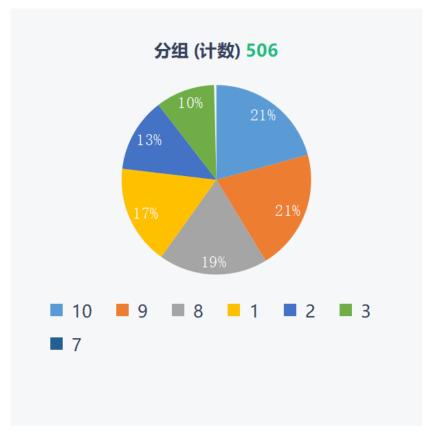


图 14.广西分布图

4.3. 杰卡德系数

D181	ID182	ID185	ID	186	ID187	ID188	ID189	ID192	ID193	ID194		ID195	ID196	ID198	ID200	ID201	ID202	ID204	ID205
0.047619048	0.052632		0 0.	. 210526316	0	0. 235294	(0.047619	0.05		0	0.157894737		0 (0.0	5 0	0.043478	0	
0	0		0	0.15	0	0. 235294	(0	0.05		0	0.047619048		0 () (0	0	0	
0	0		0 0.	. 230769231	0	0.4	(0	0		0	0.071428571		0 () (0	0	0	
0	0		0	0.2	0	0.230769	(0	0		0	0.0625		0 () (0	0	0	
0.045454545	0		0 0.	. 142857143	0	0.157895	(0.095238	0		0	0		0 () (0	0	0.047619	
0	0		0 0.	. 230769231	0	0.4	(0	0		0	0.071428571		0 () (0	0	0	
0	0		0 0.	. 272727273	0	0.5	(0	0		0	0.083333333		0 () (0	0	0	
0	0		0	0.25	0	0.3	(0	0		0	0.076923077		0 () (0	0	0	
0	0		0 0.	. 1111111111	0	0.2	(0	0		0	0.05555556		0 () (0	0	0	
0	0		0 0.	. 307692308	0	0.25	(0	0		0	0		0 () (0	0	0	
0. 1111111111	0.142857		0	0	0	0	(0.111111	0.125		0	0.25		0 (0. 12	5 0	0.090909	0	
0	0		0 0.	. 230769231	0	0.4	(0	0		0	0.071428571		0 () (0	0	0	
0	0		0 0.	. 176470588	0	0.285714	(0	0		0	0.05555556		0 () (0	0	0	
0	0		0	0.2	0	0. 333333	(0	0.066667		0	0.214285714		0 (0.06666	7 0	0	0	
0.052631579	0		0 0.	. 166666667	0	0.266667	0.0625	0	0.055556		0	0.111111111		0 0.058824	4 0. 055556	6 0	0	0	
0	0		0	0	0.090909	0. 090909	(0	0		0	0.083333333		0 () (0	0.071429	0	
0	0		0 0.	. 272727273	0	0.5	(0	0		0	0.083333333		0 () (0	0	0	
0.166666667	0		0	0.25	0	0.3	(0.076923	0		0	0.076923077		0 () (0	0	0	
0	0		0	0.3	0	0.571429	(0	0		0	0. 2		0 () (0	0	0	
0	0		0 0.	. 214285714	0	0. 25	(0	0		0	0		0 () (0	0	0	
0	0		0 0.	. 307692308	0	0.25	(0	0		0	0		0 () (0	0	0	
0	0		0 0.	. 142857143	0	0.272727	(0	0		0	0.071428571		0 () (0	0	0	
0	0		0	0. 25	0	0.2	(0.055556	0		0	0		0 () (0.055556	0	0.058824	
0	0		0 0.	. 230769231	0	0.4	(0	0.076923		0	0.153846154		0 (0.07692	3 0	0	0	
0	0		0	0.5	0	0.428571	(0.1	0		0	0		0 () (0	0	0.111111	
0	0		0 0.	. 285714286	0.066667	0.333333	(0	0		0	0.0625		0 () (0	0	0	
0	0		0	0	0	0	(0	0		0	0		0 () (0	0	0	
0.05555556	0		0 0.	. 1111111111	0.058824	0.125	(0.055556	0		0	0		0 () (0	0	0	0.062
0	0.076923		0 0.	. 214285714	0.071429	0.25	(0	0		0	0		0 () (0	0	0	0.07692
0	0		0	0.5	0	0.428571	(0	0		0	0		0 () (0	0	0	
0. 25	0		0	0	0	0.125	(0	0.125		0	1		0 (0. 12	5 0	0	0	
0	0		0	0. 25	0	0.3	(0	0.083333		0	0.076923077		0 (0.08333	3 0	0	0	
0.083333333	0.1	0.11111	1 0.	. 272727273	0	0. 333333	(0.083333	0		0	0		0 0.	1 (0	0	0.090909	0.
0	0		0	0.3	0	0.571429	(0	0		0	0.090909091		0 () (0	0	0	
0	0		0 0.	. 181818182	0	0.375	(0	0		0	0.090909091		0 () (0	0	0	
0	0		0 0.	. 272727273	0	0.5	(0	0		0	0.083333333		0 () (0	0	0	
0	0		0 0	. 333333333	0	0.666667	(0	0		0	0. 1		0 () () (0	0	

图 15.部分杰卡德系数求解结果

5. 肥料成分对比分析