

肥料登记数据分析

目录

1 数据的预处理.....	2
1.1 规范化产品通用名称.....	2
1.2 计算总无机养分百分比.....	2
2.肥料产品的数据分析.....	2
2.1 筛选出复混肥料的产品并分组.....	2
2.1.1 绘制登记数量直方图并分析分布特点.....	2
2.2.筛选出有机肥料的产品并分组.....	3
2.2.1 汇总分组形成坐标分组标签.....	3
2.2.2 绘制有机肥料产品的分布热力图并分析分布特点.....	3
2.3 聚类分析.....	3
3 肥料产品的多维度对比分析 1.....	4
3.1 分析比较不同年份产品登记数量的变化趋势.....	4
3.2 筛选有效产品并区分出广西和湖北.....	4
3.2.1 分析结果.....	4
3.3 提取产品登记数量大于 10 的肥料企业.....	4
3.3.1 企业的原料集合.....	4
3.3.2 企业之间的杰卡德相似系数矩阵.....	4

任务 1

1 数据的预处理

1.1 规范化产品通用名称

1.2 计算总无机养分百分比

1	序号	正式登记证号	总无机养分百分比
2	1	皖农肥(2016)准字4255号	0.500
3	2	皖农肥(2016)准字4256号	0.500
4	3	皖农肥(2016)准字4257号	0.510
5	4	皖农肥(2016)准字4258号	0.510
6	5	皖农肥(2016)准字4259号	0.510
7	6	皖农肥(2016)准字4260号	0.510
8	7	皖农肥(2016)准字4261号	0.520
9	8	皖农肥(2016)准字4262号	0.520
10	9	皖农肥(2016)准字4263号	0.540
11	10	皖农肥(2016)准字4264号	0.540

图 1 部分结果示意图

任务 2

2.肥料产品的数据分析

2.1 筛选出复混肥料的产品并分组

产品通用名称	复混肥料	
行标签	计数项:总无机养分百分比	分组标签
0-0.055	3	1
0.22-0.275	332	2
0.275-0.33	704	3
0.33-0.385	624	4
0.385-0.44	1337	5
0.44-0.495	1927	6
0.495-0.55	928	7
0.55-0.605	95	8
0.605-0.66	3	9
0.715-0.77	1	10
总计	5954	

图 2 数据透视表结果

2.1.1 绘制登记数量直方图并分析分布特点

（2）登记数量最大的前 3 个分组的数据统计

在原数据透视表对计数列按照降序排序，选取登记数量最大的前三名，如图所示：

排名	一	二	三
分组标签	6	5	7
产品登记数量	1927	1337	928

图 7 登记数量最大的前三个分组

（3）分析复混肥料产品的分布特点

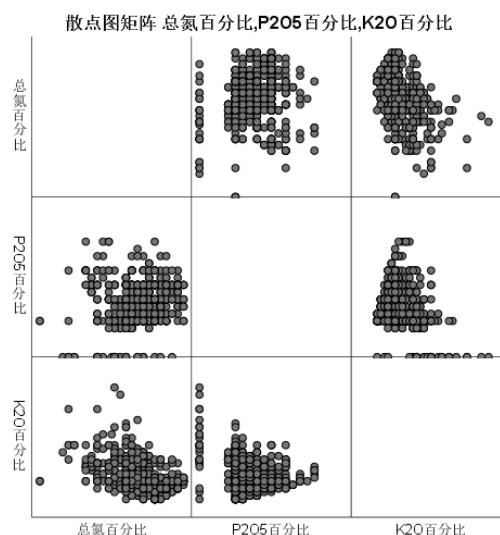
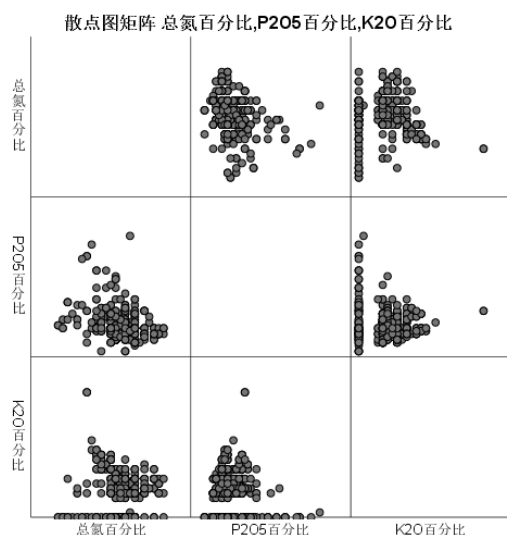
根据产品登记数量的直方图以及最大的前 3 个分组可知，登记产品的总无机养分百分比大部分处于（0.385，0.55），占比 70%左右。登记产品的总无机养分百分比大于 0.55%的很少，仅 4 个登记产品。

2.2.筛选出有机肥料的产品并分组

2.2.1 汇总分组形成坐标分组标签

2.2.2 绘制有机肥料产品的分布热力图并分析分布特点

2.3 聚类分析



任务 3

3 肥料产品的多维度对比分析 1

3.1 分析比较不同年份产品登记数量的变化趋势

3.2 筛选有效产品并区分出广西和湖北

3.2.1 分析结果

3.3 提取产品登记数量大于 10 的肥料企业

3.3.1 企业的原料集合

3.3.2 企业之间的杰卡德相似系数矩阵

	A	B	C	D	E	F	G	H	I	J	K
1		ID1	ID2	ID3	ID4	ID5	ID6	ID7	ID9	ID10	ID12
2	ID1	1	0.269231	0.3	0.333333	0.214286	0.3	0.333333	0.153846	0.227273	0.238095
3	ID2	0.269231	1	0.35	0.318182	0.206897	0.35	0.388889	0.192308	0.272727	0.285714
4	ID3	0.3	0.35	1	0.466667	0.333333	0.538462	0.636364	0.333333	0.3125	0.428571
5	ID4	0.333333	0.318182	0.466667	1	0.304348	0.375	0.428571	0.3	0.277778	0.294118
6	ID5	0.214286	0.206897	0.333333	0.304348	1	0.272727	0.3	0.28	0.208333	0.217391
7	ID6	0.3	0.35	0.538462	0.375	0.272727	1	0.8	0.263158	0.4	0.428571
8	ID7	0.333333	0.388889	0.636364	0.428571	0.3	0.8	1	0.294118	0.357143	0.5
9	ID9	0.153846	0.192308	0.333333	0.3	0.28	0.263158	0.294118	1	0.190476	0.263158
10	ID10	0.227273	0.272727	0.3125	0.277778	0.208333	0.4	0.357143	0.190476	1	0.3125
11	ID12	0.238095	0.285714	0.428571	0.294118	0.217391	0.428571	0.5	0.263158	0.3125	1

图 29 相似系数矩阵图