

目录

- 1. 背景.....1
- 2. 数据的预处理.....2
- 3. 肥料产品的数据分析.....2
 - 3.1. 复混肥料分组.....2
 - 3.2. 有机肥料分组.....3
 - 3.3. 聚类算法.....4
- 4. 多维度对比分析.....6
 - 4.1. 年份对比.....6
 - 4.2. 区域分布差异对比.....6
 - 4.3. 杰卡德系数.....7
- 5. 肥料成分对比分析.....7

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

3. 肥料产品的数据分析

3.1. 复混肥料分组

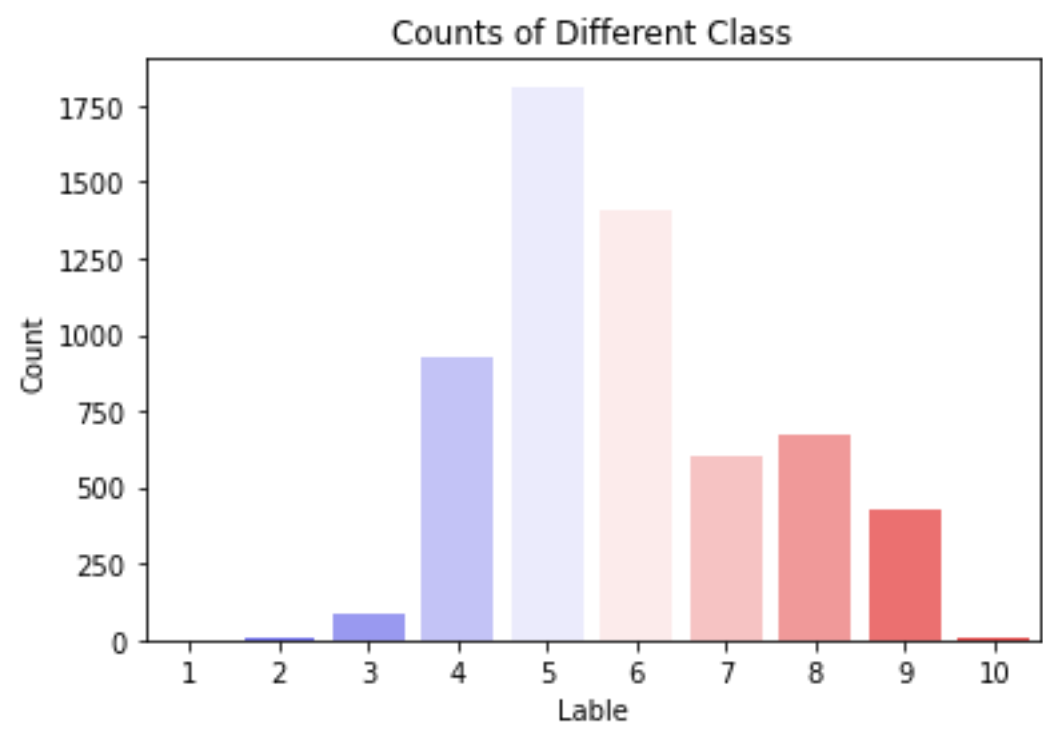


图 5.总无机养分百分比分布图

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	一	二	三
分组标签	5	6	4
产品登记数量	1920	1495	930

3.2. 有机肥料分组

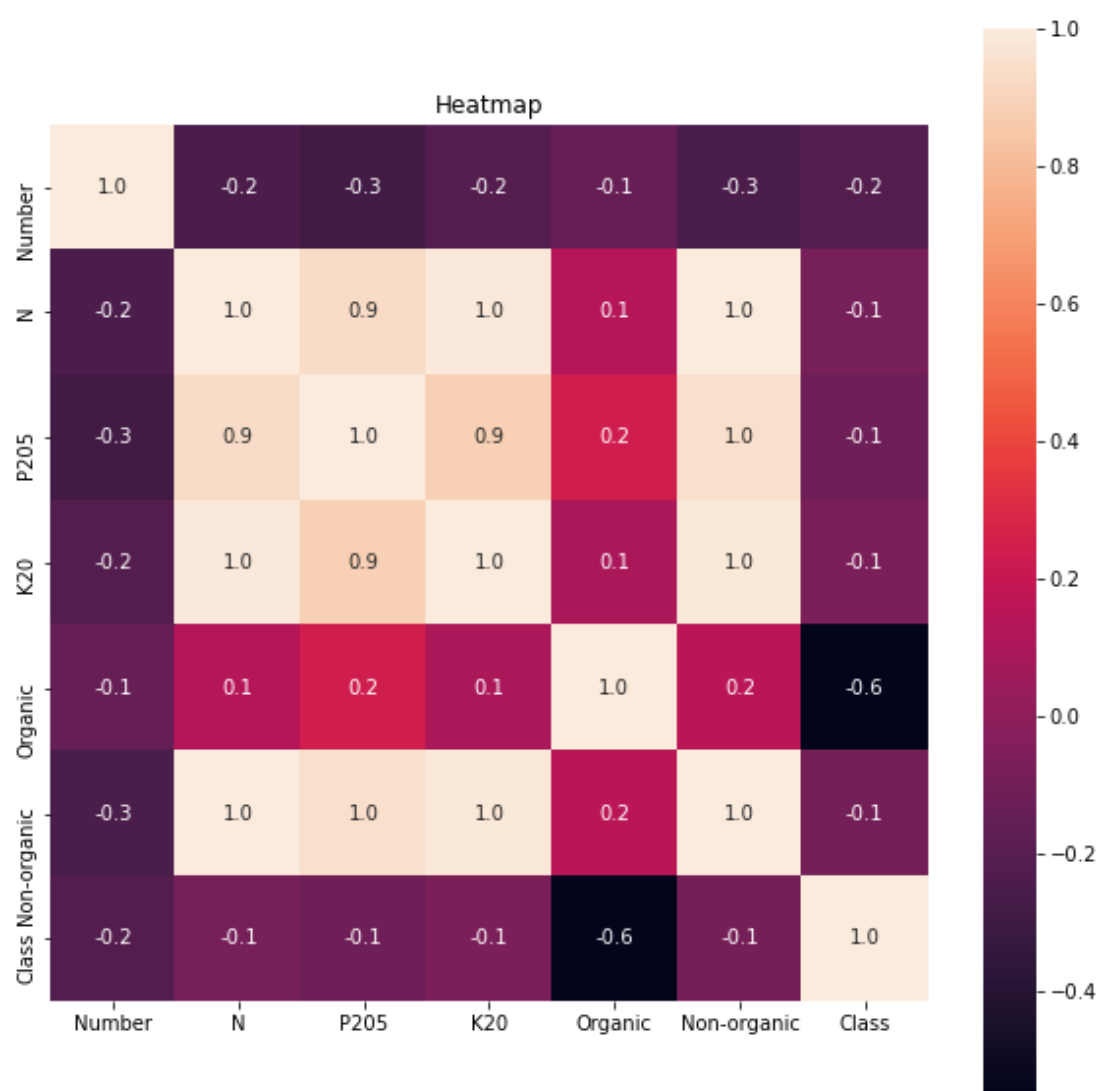


图 8.所有属性的热力分布图

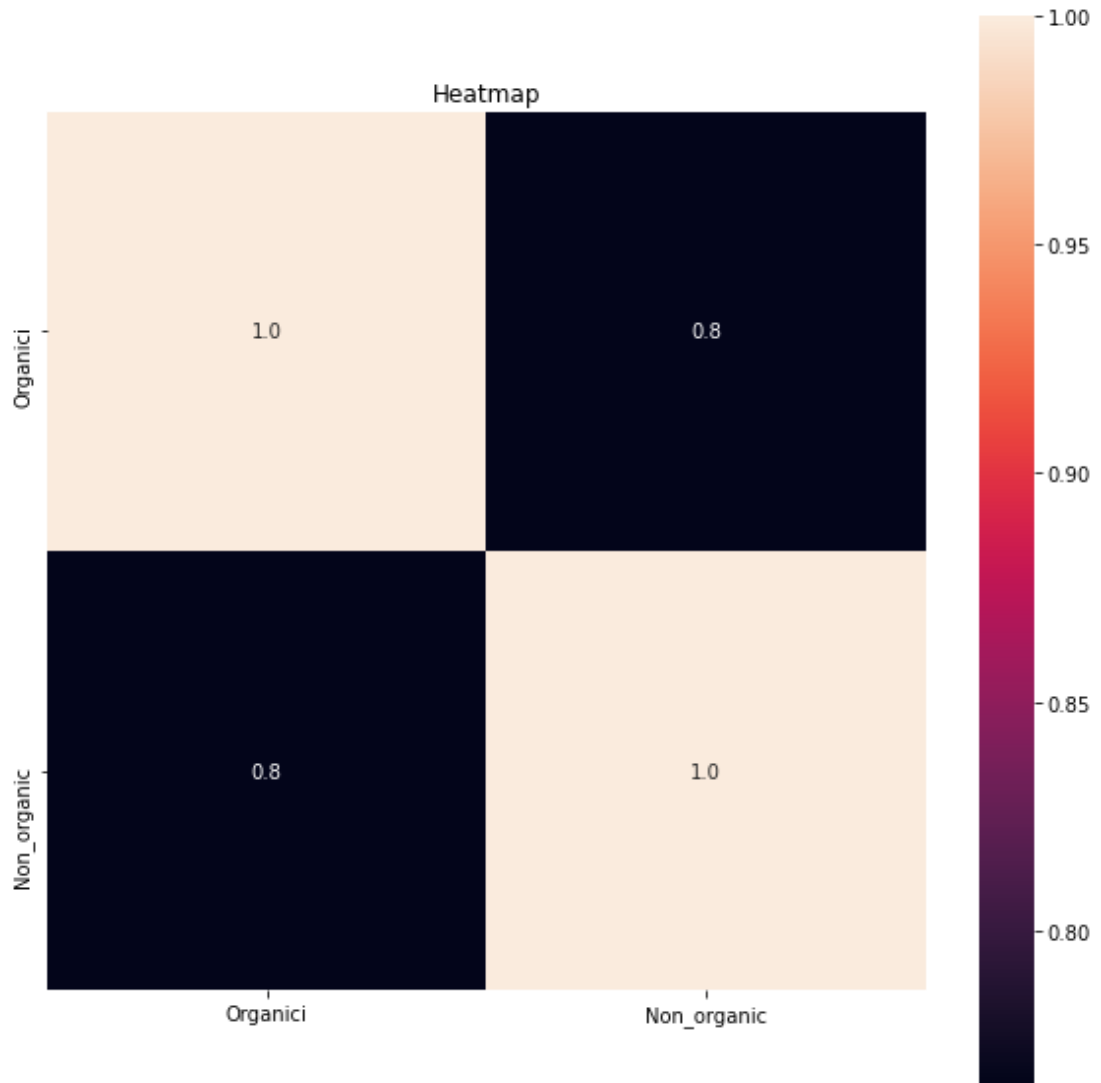


图 9.仅有有机肥料、无机养分两个属性的热力分布图

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

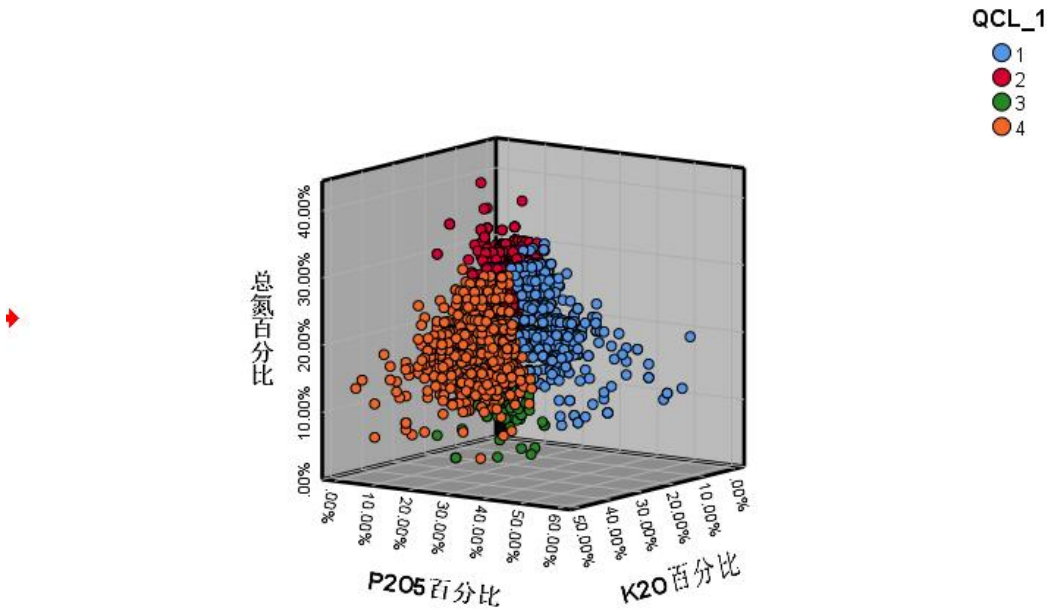


图 10.三位散点分布图

4组聚类结果氮磷钾百分含量散点图矩阵

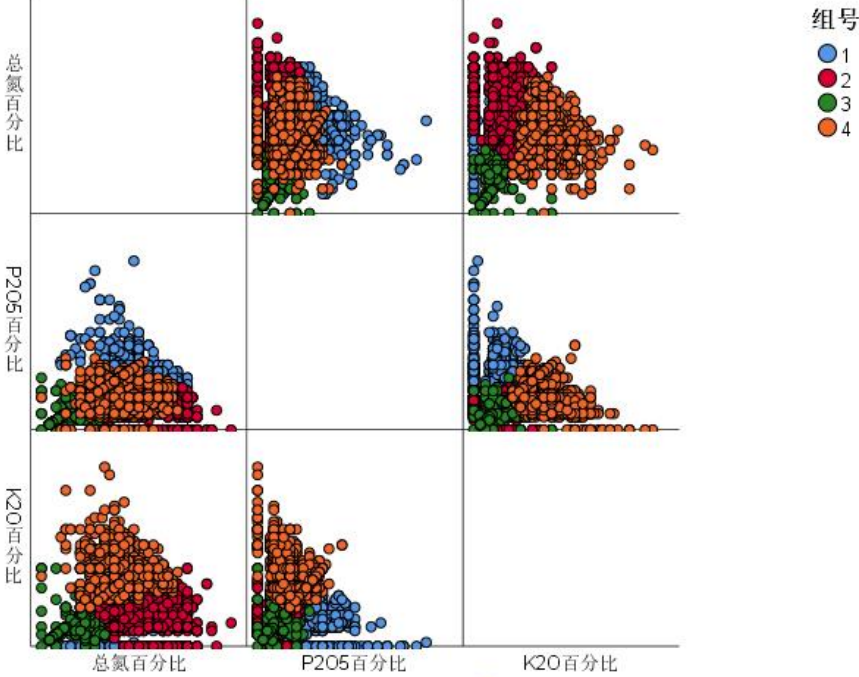


图 11.散点矩阵图

4. 多维度对比分析

4.1. 年份对比

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。

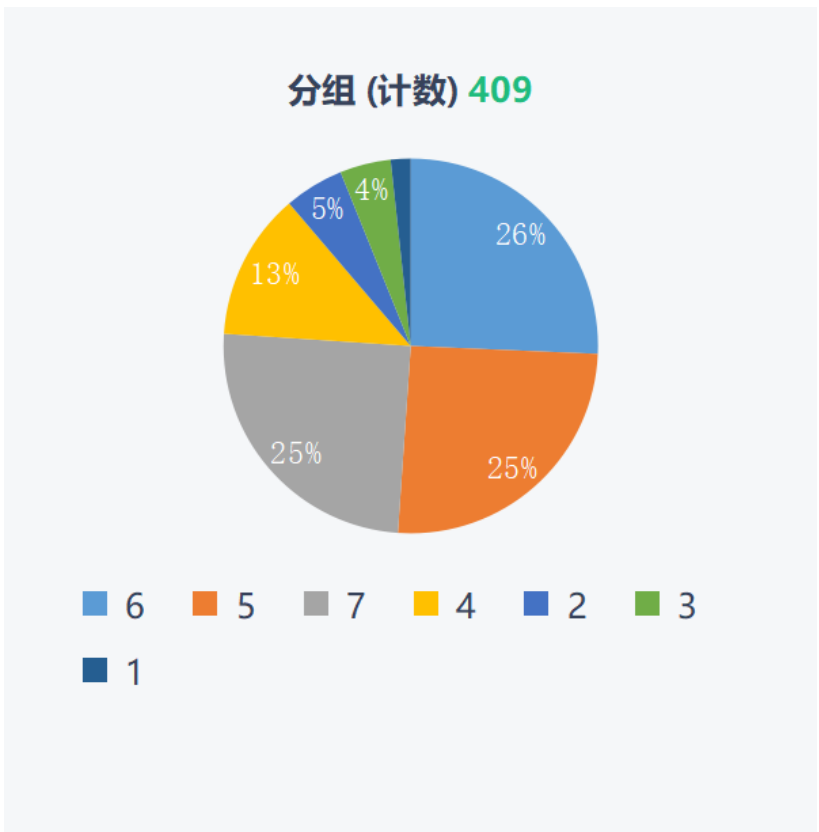


图 13.湖北分布图

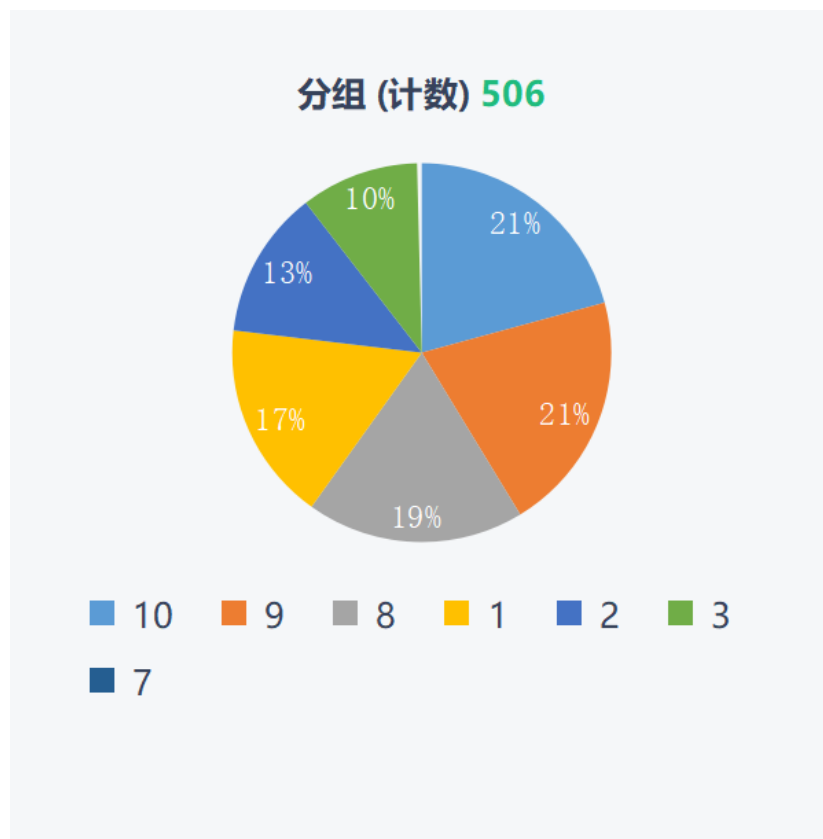


图 14.广西分布图

4.3. 杰卡德系数

5. 肥料成分对比分析