

---

## 摘要

我们借用 `pandas` 库来进行大部分的数据处理，例如对数据进行读取和规范化，或者同过这个库下的一些函数进行对表格、对数据的处理，例如依据表格某一列的值筛选出对应的数据，亦或者对表格进行增删改查等操作。

除去使用 `pandas` 库来进行数据处理外，我们还使用了 `numpy` 库作为辅助处理数据使用，尽管对于 `numpy` 库我们的使用不如 `pandas` 库那么频繁，但对于矩阵和一些表格数据还是较为有效的。

如果说数据处理占 60% 的话，那作图最少占据了 30% 的内容，其中我们使用最多的就是 `matlab` 库，无论是折线还是矩阵图，都能较好的实现绘制，但有时这个库并不能满足我们的所有需求，我们便补充使用了 `pyecharts` 库，此时便能较好的绘制出我们想要的图形。

通过使用不同的库来满足我们对数据分析的要求，并且绘制出相应的热力图、散点图矩阵等。其中还有一部分使用了聚类算法，通过聚类算法的结果来绘制了雷达图等。

---

# 目录

摘要 .....	1
一、 数据预处理.....	3
1.1 任务：规范化处理附件 1 .....	3
1.2 任务：计算总无机养分百分比，结果保留三位小数.....	3
二、肥料产品的数据分析 .....	3
2.1 任务：将复混肥料按总无机养分百分比分组 .....	3
2.2 任务：筛选有机肥料按照总无机养分百分比和有机质百分比等距分组并打上分组标签 .....	4
2.2.1 读取文件观察数据基本情况.....	4
1.2.4 绘制分布热力图 .....	4
2.3 任务 .....	4
2.3.4 散点图矩阵.....	5
2.3.5 雷达图的绘制 .....	5
三、肥料产品的多维度对比分析（日期、省份、生产商等） ....	6
3.1 任务：提取年份，分析比较复混肥料中各组别不同年份产品登记数量的变化趋势 .....	6
3.1.1 读取数据并提取相应数据.....	6

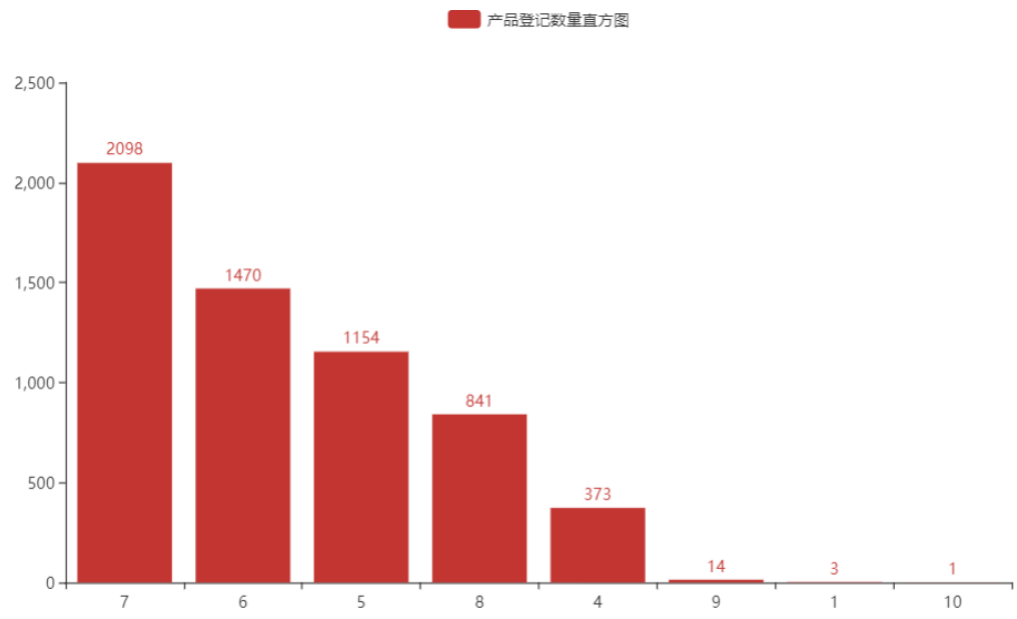
# 一、数据预处理

## 1.1 任务：规范化处理附件 1

## 1.2 任务：计算总无机养分百分比，结果保留三位小数

# 二、肥料产品的数据分析

## 2.1 任务：将复混肥料按总无机养分百分比分组



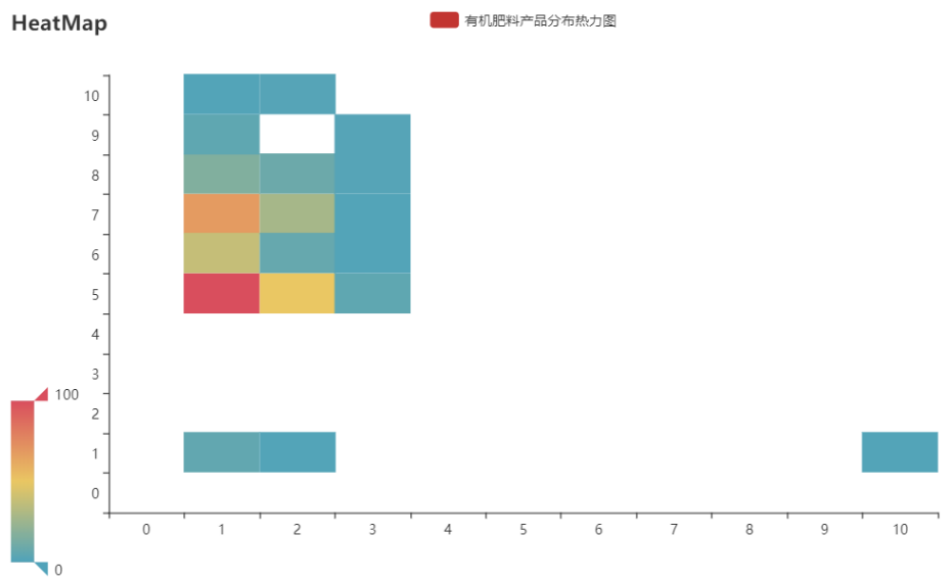
此时通过已经完成排序的数据查看登记数量最大的前 3 个分组：

排名	一	二	三
分组标签	7	6	5
产品登记数量	2098	1470	1154

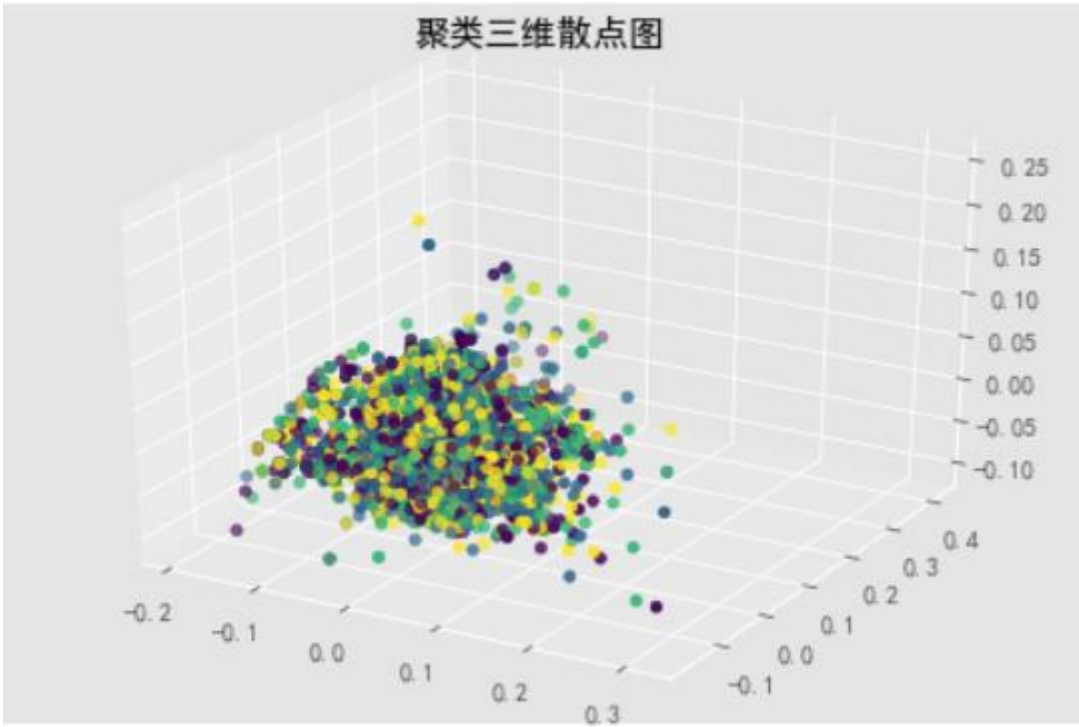
## 2.2 任务：筛选有机肥料按照总无机养分百分比和有机质百分比等距分组并打上分组标签

### 2.2.1 读取文件观察数据基本情况

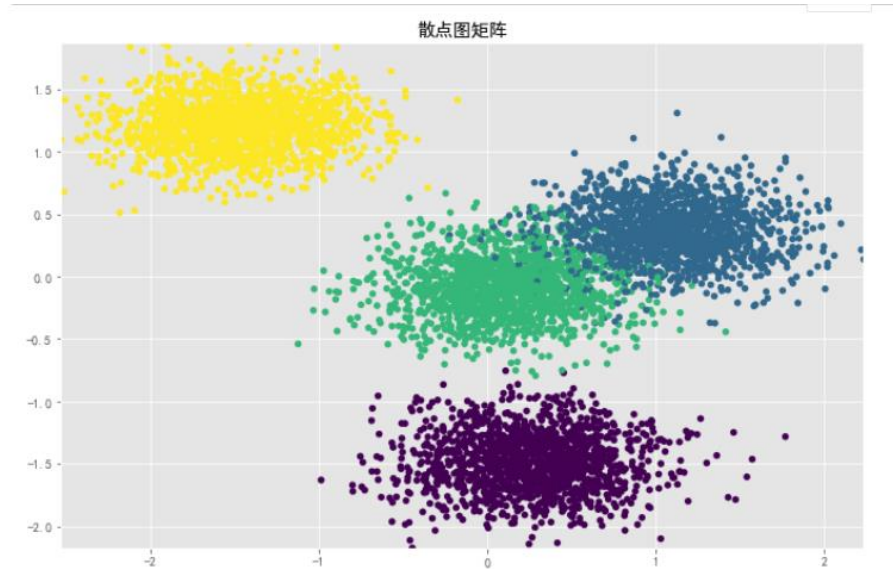
### 1.2.4 绘制分布热力图



## 2.3 任务

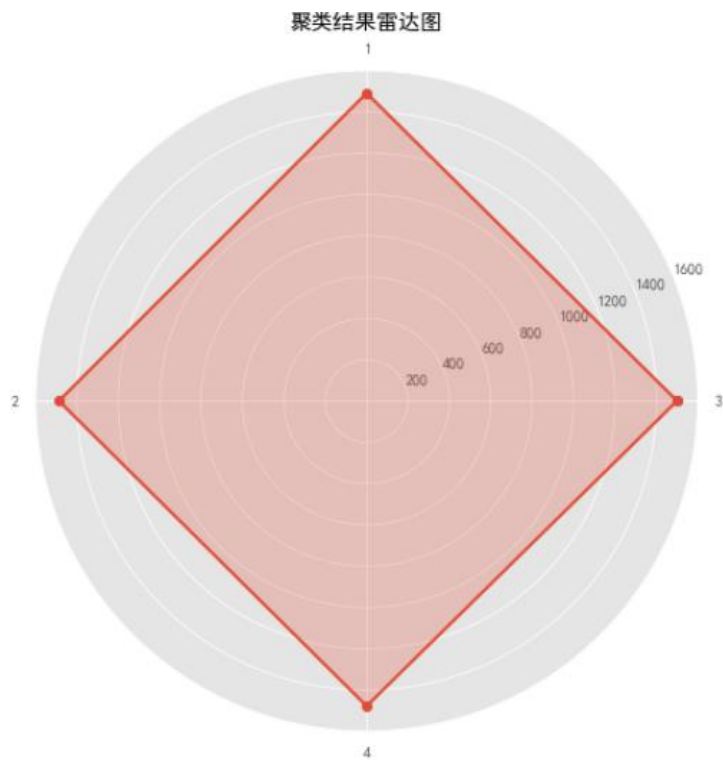


### 2.3.4 散点图矩阵



内绘制。

### 2.3.5 雷达图的绘制



### 三、肥料产品的多维度对比分析（日期、省份、生产商等）

#### 3.1 任务：提取年份，分析比较复混肥料中各组别不同年份产品登记数量的变化趋势

##### 3.1.1 读取数据并提取相应数据

