

# 各国森林，草原聚类分析

## 数据处理：

显然数据的量纲不同，所以对数据进行描述性统计，并且保存标准化后的样本数据：

	N	最小值	最大值	均值	标准 偏差	方差
森林覆盖率（%）	21	8.6	84.0	32.100	19.3006	372.513
林木蓄积量（亿立方米）	21	.8	841.5	84.871	187.7405	35246.495
草原面积（万公顷）	21	58	45190	8287.71	13971.978	195216182.414
有效个案数（成列）	21					

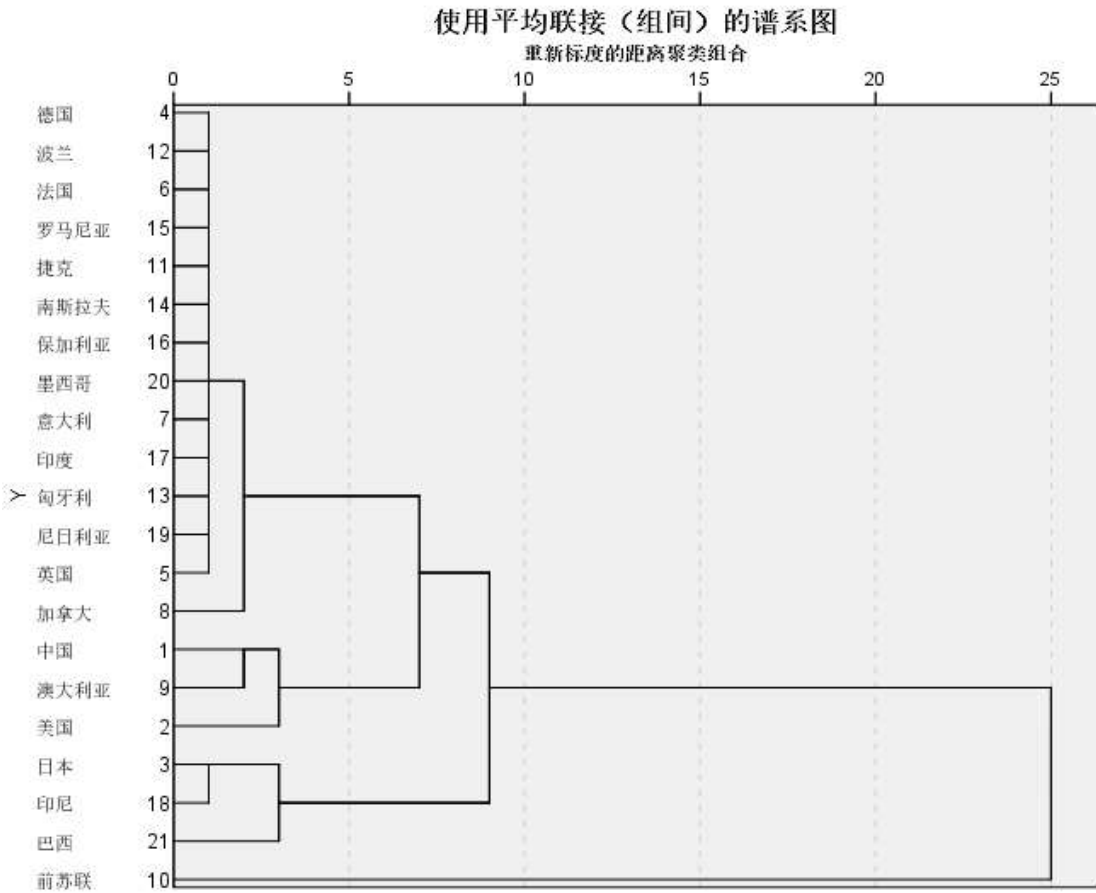
描述性统计

## 系统聚类分析

阶段	组合聚类		系数	首次出现聚类的阶段		下一个阶段
	聚类 1	聚类 2		聚类 1	聚类 2	
1	4	12	.001	0	0	5
2	11	14	.002	0	0	4
3	6	15	.004	0	0	5
4	11	16	.007	2	0	9
5	4	6	.007	1	3	9
6	7	17	.022	0	0	8
7	13	19	.024	0	0	8
8	7	13	.062	6	7	10
9	4	11	.186	5	4	11
10	5	7	.298	0	8	12
11	4	20	.412	9	0	12
12	4	5	.623	11	10	15
13	3	18	.767	0	0	16
14	1	9	1.104	0	0	17
15	4	8	1.289	12	0	18
16	3	21	2.794	13	0	19
17	1	2	2.830	14	0	18
18	1	4	6.626	17	15	19
19	1	3	8.306	18	16	20

20	1	10	24.832	19	0	0
----	---	----	--------	----	---	---

系统聚类分析结果谱系图：



## 分析聚类结果

这里我选择簇为 4 ( $K = 4$ ), 重新进行聚类分析:

国别	森林覆盖率 (%)	林木蓄积量 (亿立方米)	草原面积 (万公顷)	类别
中国	12.5	93.5	31908	1
美国	30.4	202	23754	1
澳大利亚	13.9	10.5	45190	1
日本	67.2	24.8	58	2
印尼	84	33.7	1200	2
巴西	67.6	238	15900	2
德国	28.4	14	599	3
英国	8.6	1.5	1147	3
法国	26.7	16	1288	3
意大利	21.1	3.6	514	3
加拿大	32.7	192.8	2385	3
捷克	35.8	8.9	168	3

波兰	27.8	11.4	405	3
匈牙利	17.4	2.5	129	3
南斯拉夫	36.3	11.4	640	3
罗马尼亚	26.7	11.3	447	3
保加利亚	34.7	2.5	200	3
印度	20.5	29	1200	3
尼日利亚	16.1	0.8	2090	3
墨西哥	24.6	32.6	7450	3
前苏联	41.1	841.5	37370	4

对结果进行说明：

- 类别 1：整体来说，该类别的国家草原面积较高，
- 类别 2：普遍来说：森林覆盖率较高
- 类别 3：整个类别中没有哪个国家有非常突出的
- 类别 4：三个指标综合来说最好，所以单独分为一类