

# Analysis of Economic Level Cluster Measurement of 179 Countries and Regions

Zihao Zhang

School of Geography and Environmental Science, Zhejiang Normal University, Jinhua, Zhejiang, 321004, China

## Abstract

Based on the data of GDP, population, land area and the proportion of the three industries of 179 countries and regions in 2012, this paper uses SPSS software for cluster analysis, conducts K-mean clustering and systematic clustering for different factors in each country and region, and finally draws the following conclusions according to the analysis results: 1) In terms of the overall economic volume in 2012, China, the United States, India, Germany and Japan are the five largest countries with GDP far exceeding other countries. 2) From the perspective of the six influencing factors of GDP, population and land area in 2012, China, the United States and India are three relatively unique countries, which are closely related to the economic strength, population and the proportion of the three industries of the three countries. China can be attributed to the population and economic level of the parallel type, the United States can be attributed to the economic level and industrial proportion of the dominant type, and India can be attributed to the population and economic level of the mismatch type.

## Keywords

cluster analysis; system clustering; China; United States; India

## 179 个国家和地区经济水平聚类测度分析

张子皓

浙江师范大学地理与环境科学学院, 中国·浙江·金华 321004

## 摘要

论文基于2012年世界179个国家和地区GDP、人口数量、国土面积、三个产业占比情况的数据基础上,运用SPSS软件进行聚类分析,分别对各个国家和地区不同因素进行K-均值聚类和系统聚类后进行分析,最终根据分析结果得出以下结论:①在2012年整体经济体量上来看,中国、美国、印度、德国、日本为远超其他国家的五个GDP大国。②在2012年综合了GDP、人口、国土面积三个产业占比情况这六个影响因素来看,中国、美国和印度为三个较为独特的国家,这与三个国家的经济实力、人口数量和三个产业的占比情况密不可分,中国可归结为人口与经济水平并驾齐驱型、美国可归结为经济水平与产业占比优势型,而印度则可归结为人口与经济水平不匹配型。

## 关键词

聚类分析; 系统聚类; 中国; 美国; 印度

## 1 引言

2006年美国爆发次贷危机,2008年危机达到顶峰,世界经济受到威胁,2008—2012年是世界经济发展的分水岭年份<sup>[1]</sup>,在2006年的前30年世界经济持续高速发展,而在2012年以后,世界各国经济逐渐出现各种各样的问题。因此,论文以分水岭年份的2012年作为研究年份,分析2012年不同类型国家的GDP、人口数量、国土面积等基础数据并进行聚类分析,从而对各个国家在经济危机下的基本情况和经济发展状况进行分析,并得出不同国家经济类型和发展模式,这对于经济危机下各国的应对措施和防灾水平深入研

究具有重大的理论和现实意义<sup>[2]</sup>。

## 2 研究数据与方法

### 2.1 数据来源

论文研究对象为世界179个主要国家和地区,主要采用国外数据库Kaggle中的数据并查询资料加以补充。

### 2.2 研究方法

聚类分析法,聚类分析法是一种将研究对象分为相对同质的群组的统计分析技术,它能够将一批样本数据按它们在性质上的亲疏程度在没有先验知识的情况下自动进行分类<sup>[3]</sup>。

## 3 实证分析

### 3.1 数据选取

论文选取2012年179个国家和地区GDP<sup>[4]</sup>、人口数量、

【作者简介】张子皓(2000—),男,中国山西太原人,硕士,从事人文地理研究。

国土面积、第一产业占比、第二产业占比和第三产业占比作为研究对象。GDP (Gross Domestic Product) 是指国内生产总值,是衡量一个国家或地区经济实力的重要指标。人口数量是一个国家的人口总数,是一个国家经济水平高低的重要表征。国土面积是指国家陆地面积与内陆水域的总和,是反映国家经济发展的核心指标之一。第一产业占比是指农业占国家产业总量的比重,第二产业占比是指工业占国家产业总量的比重,第三产业占比是指服务业占国家产业重量的比重,三者的比例极大地反映了一个国家的发达程度。综合来看,选取的指标能够均衡、综合、有效地反映一个国家或地区的综合经济水平。

### 3.2 GDP 的系统聚类分析和 K- 均值聚类分析

#### 3.2.1 系统聚类分析

单独对各个国家 GDP 进行系统聚类分析,则可分为美国,中国,印度、日本、法国、英国、巴西、德国,其他国家四种类型。由此可以看出,美国、中国、印度、日本、法国、英国、巴西、德国在世界上的综合经济地位较高。GDP 系统聚类分析如图 1 所示。

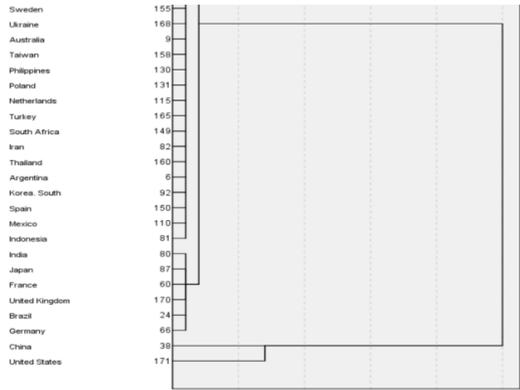


图 1 GDP 系统聚类分析

#### 3.2.2 K- 均值聚类分析

对各个国家 GDP 进行 K- 均值聚类分析,将其分为十种类型,则中国、美国、印度、德国、日本会被分为单独一类。这五个国家也是 2012 年 GDP 位居世界前列的国家。综上所述,在各个指标的表征能力上,GDP 在 2012 年的国家综合经济水平测度上依然是最重要的因素之一。每个聚类中的个案数如表 1 所示。

### 3.3 综合数据的系统聚类分析和 K- 均值聚类分析

#### 3.3.1 系统聚类分析

通过对各个国家 GDP、人口、国土面积、第一产业占

比、第二产业占比和第三产业占比标准化处理后进行系统聚类分析,若按 6 为标度进行分类,则可分为美国、印度、中国、澳大利亚和巴西、其他国家四类,若按照更小标度进行分类,则法国、英国、德国、墨西哥、日本被分为一类,阿根廷、伊朗、乌克兰、泰国等被分为一类,卡塔尔和赤道几内亚被分为一类。其他如巴拿马、斐济、冰岛等其他国家被分为一类。法国、英国、德国、日本、墨西哥被分为一类。综合数据系统聚类分析特殊区域如图 2 所示。

表 1 每个聚类中的个案数

聚类	1	27.000
	2	1.000
	3	1.000
	4	129.000
	5	1.000
	6	4.000
	7	3.000
	8	1.000
	9	11.000
	10	1.000
有效		179.000
缺失		.000

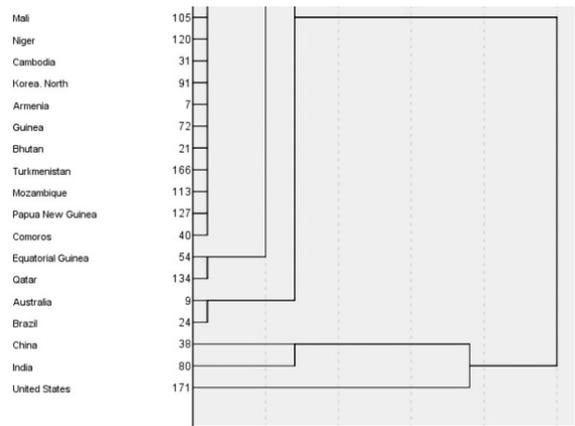


图 2 综合数据系统聚类分析特殊区域

#### 3.3.2 K- 均值聚类分析

对各个国家 GDP、人口、国土面积、第一产业占比、第二产业占比、第三产业占比标准化处理后进行 K- 均值聚类分析,将其分为十种类型,则中国、美国、印度会被分为单独一类。综合数据的 K- 均值聚类分析最终聚类中心如表 2 所示,每个聚类中的个案数如表 3 所示。

表 2 综合数据的 K- 均值聚类分析最终聚类中心

	聚类									
	1	2	3	4	5	6	7	8	9	10
Zscore(GDP)	2.74139	-.24703	-.14248	-.25290	-.25153	10.38492	3.13556	5.94166	.29982	-.25302
Zscore(Population)	8.05333	-1.6090	.74106	-.25932	-.23658	2.00533	.70770	9.71252	-1.0587	-2.5556
Zscore: Area (sq. mi.)	1.95134	-.31935	.25765	-.40405	-.32444	6.49675	-.13352	6.47206	5.10345	-.38414
Zscore(Agriculture)	.19225	1.22302	.74067	-.94425	4.04443	-.97068	-.92442	-.21081	-.78567	-.83853
Zscore(Industry)	-.08572	-.92672	1.41809	-1.82474	-1.66794	-.59888	-.21401	1.31831	-1.8550	4.40435
Zscore(Service)	-.09965	-.33505	-1.87426	2.41137	-2.27868	1.40334	1.02911	-.91452	.87820	-2.97283

表3 每个聚类中的个案数

聚类	1	1.000
	2	43.000
	3	40.000
	4	61.000
	5	22.000
	6	1.000
	7	5.000
	8	1.000
	9	2.000
	10	3.000
有效		179.000
缺失		.000

## 4 结果分析

根据 SPSS 软件中的聚类分析方法,对 2012 年世界上 179 个主要国家和地区的主要经济数据进行分析,发现了几个特殊的地区和国家,现对其分类依据进行分析。

### 4.1 以 GDP 将 179 个国家和地区分为美国、中国、印度、日本、法国、英国、巴西、德国,其他国家四个类型的分类依据

GDP 作为表征各国经济水平最重要的指标,将其单独作为聚类依据进行聚类分析。根据原始数据及基本情况不难看出,美国在 2012 年 GDP 独占鳌头,是中国 GDP 的将近一倍之多,而根据分类结果分析,中国居于世界第二位,且与日本、印度等国不在同一类型,因此可以判断,美国、中国两个国家的 GDP 即使在经济危机背景下,也已经与日、印、法、英等国家形成的第三梯队和其他国家形成的第四梯队拉开差距。另外,美国在世界经济中的领先地位依然存在并可能长期保持,经过经济危机的冲击与洗礼后,其经济实力虽然有所下滑,但以 2012 年作为衡量年份,其 GDP 水平依然独树一帜。

### 4.2 法国、英国、德国、日本、墨西哥在综合指标分类中被分为一类的依据

法国、英国、德国、日本、墨西哥被分为一类,前四者比较好理解,都是较为发达的资本主义国家,而墨西哥被分为此类中,有这么几方面可能的原因。其一就是其产业结构较为合理,趋向于向第三产业发展,而从数据可见,墨西哥产业结构和法国等发达国家明显不同<sup>[5]</sup>,第一种可能性可以基本排除;其二是其国土面积、人口和这几个发达国家高度相似,从而导致其被分为一类,从数据可得,这个原因的可能性较大;其三是其 GDP 与几个发达国家接近,从而致

使其被分为一类,从数据可知,墨西哥经济与其他几个发达国家经济还有一定差距,因此,第三种可能性也不大。依照选取指标可以得出,未来墨西哥的发展势头甚至可以和英法德日等发达国家相比。

### 4.3 综合指标的 K-均值聚类分析中国、美国、印度被分为三个独立类型的依据

中国被分为单独一类,因为其 GDP 总量、人口规模、国土面积三者均居于世界前列,且人口增长和 GDP 增长基本保持同步,可以归结为人口与经济水平并驾齐驱型。另外,中国的整体综合水平较高,GDP 总量位居世界第二,人口规模位居世界第一,国土面积也位居世界第三,因此中国是世界上综合国力最强的国家之一。

美国被分为单独一类,是因为其 GDP 总量在全世界独占鳌头,且同样拥有着较高的人口规模和土地面积,最后是由于较为合理的产业结构,其三个产业的占比分别为 1%、20.4%、78.7%,其第三产业占比经标准化后的数据为 1.4,位居全球第十。

印度被分为单独一类,是因为其人口虽然位居世界第二位,且与排在第三位的美国相差将近 9 亿人,但其 GDP 却排到了全球的第 11 位,不均衡的人口与经济模式造成了其也被分为了单独一类。

## 5 结论与讨论

在 2012 年整体经济体量上来看,中国、美国、印度、德国、日本为远超其他国家的五个 GDP 大国。

在 2012 年综合了 GDP、人口、国土面积三个产业占比情况这六个影响因素来看,中国、美国和印度为三个较为独特的国家,这与三个国家的经济水平、人口数量和三个产业的占比情况密不可分,中国可归结为人口与经济水平并驾齐驱型、美国可归结为经济水平与产业占比优势型,而印度则可归结为人口与经济水平不匹配型。

### 参考文献

- [1] 张军,吴桂英,张吉鹏.中国省际物质资本存量估算:1952—2000[J].经济研究,2004(10):35-44.
- [2] 王小鲁,樊纲,刘鹏.中国经济增长方式转换和增长可持续性[J].经济研究,2009,44(1):4-16.
- [3] 孙吉贵,刘杰,赵连宇.聚类算法研究[J].软件学报,2008(1):48-61.
- [4] 谈儒勇.中国金融发展和经济增长关系的实证研究[J].经济研究,1999(10):53-61.
- [5] 吴国平.中国与墨西哥产业和贸易结构比较[J].拉丁美洲研究,2009,31(5):12-22+79.