|  |
| --- |
|  |
| **基于因子分析法的我国农村居民幸福指数研究** |
|  |
|  |
|  |
|  |
| **10175501112 陈诺**  **华东师范大学 上海浦东 200241** |

|  |
| --- |
|  |

**基于因子分析法的我国农村居民幸福指数研究**

**摘要**：本文从我国30个省市、自治区[[1]](#footnote-1)2015年的农村经济数据出发，选取14个指标，运用因子分析方法，利用python实现构建农村居民幸福指数的评价体系，分析影响农村居民主观幸福感的主要因素，总结出各地区间幸福感的差异及其原因，并提出政策建议。

**关键词**：幸福指数，因子分析，农村居民，python

**分类号**：C913.7 , TP3-05

**Abstract:** This paper incorporates the rural economic data for 31 provinces and autonomous regions in China, based on which 14 indexes are selected and positive analyses are conducted by the factor analysis and realized by python. In the context, this work aims to construct an evaluation system on the happiness index of rural residents and identify the factors affecting their subjective well-being. Furthermore, the paper exposes the difference of such well-being in diverse regions and sum up the reasons as well, thus putting forward policy suggestions.

**Keywords:** Happiness index, Factor analysis, Rural residents, Python

**Classification:** C913.7 , TP3-05

目录

[1 引言 3](#_Toc534480479)

[1.1 研究背景 3](#_Toc534480480)

[1.2 研究现状 3](#_Toc534480481)

[1.3 研究方法 3](#_Toc534480482)

[2 居民幸福指数体系构建 4](#_Toc534480483)

[2.1 生活环境 4](#_Toc534480484)

[2.2 社会环境 4](#_Toc534480485)

[2.3 福利水平 4](#_Toc534480486)

[3 建模方法 4](#_Toc534480487)

[3.1 基本思想 5](#_Toc534480488)

[3.2 数学原理 6](#_Toc534480489)

[4 数据来源与数据处理 7](#_Toc534480490)

[4.1 数据来源 7](#_Toc534480491)

[4.2 数据清洗及预处理 9](#_Toc534480492)

[4.2.1 数据清洗 9](#_Toc534480493)

[4.2.2 缺失值处理 10](#_Toc534480494)

[4.2.3 数据标准化处理及相关检验 10](#_Toc534480495)

[5 因子分析过程 11](#_Toc534480496)

[5.1 确定因子个数及命名 12](#_Toc534480497)

[5.1.1 保留因子个数 12](#_Toc534480498)

[5.1.2 命名 13](#_Toc534480499)

[5.2 因子得分 15](#_Toc534480500)

[5.3 总分排名 17](#_Toc534480501)

[6 结果分析及政策建议 18](#_Toc534480502)

[7 附录 20](#_Toc534480503)

[探索：python与SPSS在结果上的不同 20](#_Toc534480504)

[参考文献 22](#_Toc534480505)

# 1 引言

## 1.1 研究背景

“国民幸福指数”这个概念最早由不丹国王于20世纪70年代提出，相关可考据文献最远可至1972年，N.T Feather教授发表的关于学生幸福指数的调查研究；国内对幸福指数的研究起步较晚，2004年才开始全国范围的幸福指数抽样调查，2006年我国学术界对幸福学的探讨才达到热议阶段。一些学者认为，政府过去为了推动国家发展，一味地采取提高GDP的方法是不妥的，相反的，幸福指数是比GDP更需要国家政府重视的指标，对国家发展也更有意义。

## 1.2 研究现状

居民幸福指数是反映居民主观幸福感的直接表现，也是现今最有表现力的一种测度水平。了解居民幸福指数对改善居民生活状况、推动国家基础设施建设、制定社会发展相关决策、促进经济体制改革具有重要意义。目前，学术界对于幸福感指数的调查研究对象主要是广大城镇居民，而对农村居民的幸福指数关注较少据统计，在中国知网的经济与管理科学学科下，对非农村居民幸福指数的相关研究有11392条结果，而农村仅有697条。此外，现有的农村幸福指数调查研究较为局限，多为区域性调查；现有文献多关注于幸福感的影响因素，缺少对地区间幸福感差异的原因的探索，而这正是本文寻求创新的突破口。

## 1.3 研究方法

基于既有文献，目前幸福指数的研究方法主要有两种，一是专家赋权法，二是因子分析法，本文采用后者。专家赋权法也称为主观赋权法，它根据决策者（专家）对每个变量的重视程度来确定权重，之后采用Delphi方法建立指标体系，其原始数据大部分取决于决策者（专家）经验判断，随意性高，客观性差；因子分析法作为一种更为客观的分析方法，从原始数据出发，研究变量内部之间的依赖程度，对变量进行降维并得出公共因子，其运算过程科学严谨，结论更加客观并具有说服力。

# 2 居民幸福指数体系构建

本文基于国内外相关假说、文献，并且考虑到农村较城市设施、经济落后的现实情况，以及现有数据的局限，将从三大模块建立居民幸福指数体系，分别为：生活环境、社会环境、福利水平。

## 2.1 生活环境

生活环境指收入、消费水平，以及房屋建设。它是测量居民生活水平的重要指标，真实反映了居民的生活条件，尤其是对农村居民而言，收入和消费水平在更大程度上影响他们对生活的满意度。从中国的基本国情来看，区别于国外的历史文化，国内对住房问题更加看重，因此房屋建设也是衡量标准之一。

## 2.2 社会环境

社会环境指卫生条件和文化方面。对于部分偏僻山区，卫生条件远远低于我国平均水平，使得农村居民的疾病得不到及时医治，影响国民健康水平，因此，完善的医疗保障体系对提高农村居民的幸福指数尤为重要。文化可以影响一个人的价值观、期望目标、生活态度等等，这是影响居民幸福感的主要内在因素，所以文化普及对农村居民生活也有相应改善。

## 2.3 福利水平

福利水平包括社会救助和养老院建设等。由于目前农村留守问题频生，居民多为老人儿童，他们中的很大一部分人都要靠社会救济金来保障最低生活水平，因此这方面的影响也需要考虑到幸福指数指标中。

# 3 建模方法

在科学研究中，往往希望尽可能多地收集反映研究对象的多个变量，以期能对问题有比较全面、完整的把握与认识。多变量的大样本虽然能为科学研究提供大量的信息，但是在一定程度上增加了数据采集的工作量，更重要的是在大多数情况下，许多变量之间可能存在相关性，这意味着表面上看来彼此不同的变量并不能从各个侧面反映事物的不同属性，而恰恰是事物同一种属性的不同表现。  
   如何从众多相关的指标中找出少数几个综合性指标来反映原来指标所包含的主要信息，这就需要进行因子分析（Factor Analysis），它是用少数几个因子来描述许多指标或因素之间的联系，即：用较少几个因子反映原始数据的大部分信息的统计方法。

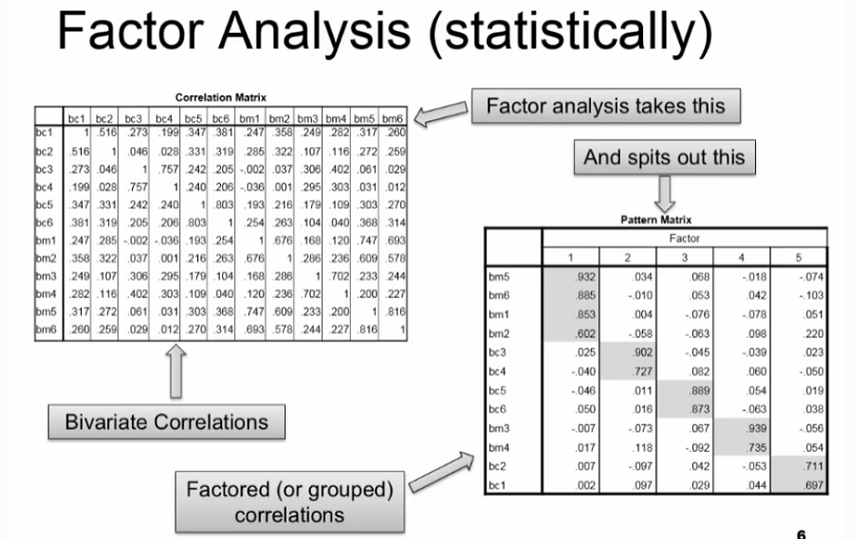


图3.1 因子分析

## 3.1 基本思想

因子分析法的基本思想是将观测变量进行分类，将相关性较高，即联系比较紧密的分在同一类中，而不同类变量之间的相关性则较低，那么每一类变量实际上就代表了一个基本结构，即公共因子。对于所研究的问题就是试图用最少个数的不可测的所谓公共因子的线性函数与特殊因子之和来描述原来观测的每一分量。这样，就能相对容易地以较少的几个因子反映原资料的大部分信息，从而达到浓缩数据，以小见大，抓住问题本质和核心的目的。

## 3.2 数学原理

因子分析法的核心是对若干综合指标进行因子分析并提取公共因子，再以每个因子的方差贡献率作为权数与该因子的得分乘数之和构造得分函数。因子分析法的数学表示为矩阵：，即:

 (k≤p)………………(1式)

模型中，向量X是可观测随机向量，即原始观测变量。F是X 的公共因子，即各个原观测变量的表达式中共同出现的因子，是相互独立的不可观测的理论变量。公共因子的具体含义必须结合实际研究问题来界定。A是公共因子F的系数，称为因子载荷矩阵，（i=1,2,.....,p;j=1,2,....,k)称为因子载荷，是第i个原有变量在第j个因子上的负荷，或可将看作第i个变量在第j公共因子上的权重。是xi与fj的协方差，也是xi与fj的相关系数，表示xi对fj的依赖程度或相关程度。的绝对值越大，表明公共因子fj 对于xi的载荷量越大。B是X的特殊因子，是不能被前k个公共因子包含的部分，这种因子也是不可观测的。各特殊因子之间以及特殊因子与所有公共因子之间都是相互独立的。

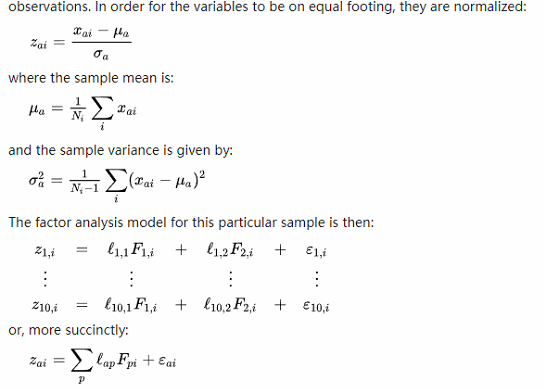


图3.2 数学原理

# 4 数据来源与数据处理

## 4.1 数据来源

为了对我国农村居民的幸福感进行综合评价，本文数据完全来源于EPS数据平台中的中国三农数据库（分地区），选取2015年我国31个省市、自治区[[2]](#footnote-2)有关衡量幸福感的14个指标，如表3.1所示：

|  |  |
| --- | --- |
| 符号表示 | 指标 |
| x1 | 农村居民家庭人均可支配收入（元/人） |
| x2 | 卫生厕所普及率（%） |
| x3 | 建筑工程（亿元） |
| x4 | 住宅（亿元） |
| x5 | 食品支出（元/人） |
| x6 | 衣着支出（元/人） |
| x7 | 居住支出（元/人） |
| x8 | 文教娱乐支出（元/人） |
| x9 | 医疗保健支出（元/人） |
| x10 | 卫生院（个） |
| x11 | 乡镇文化站（个） |
| x12 | 养老院单位数（个） |
| x13 | 农村最低生活保障支出（万元） |
| x14 | 农村居民最低生活保障人数（人） |

表 3.1

由于调查方向指向不同城市间的差异，所以需要全国范围所有省市的数据。

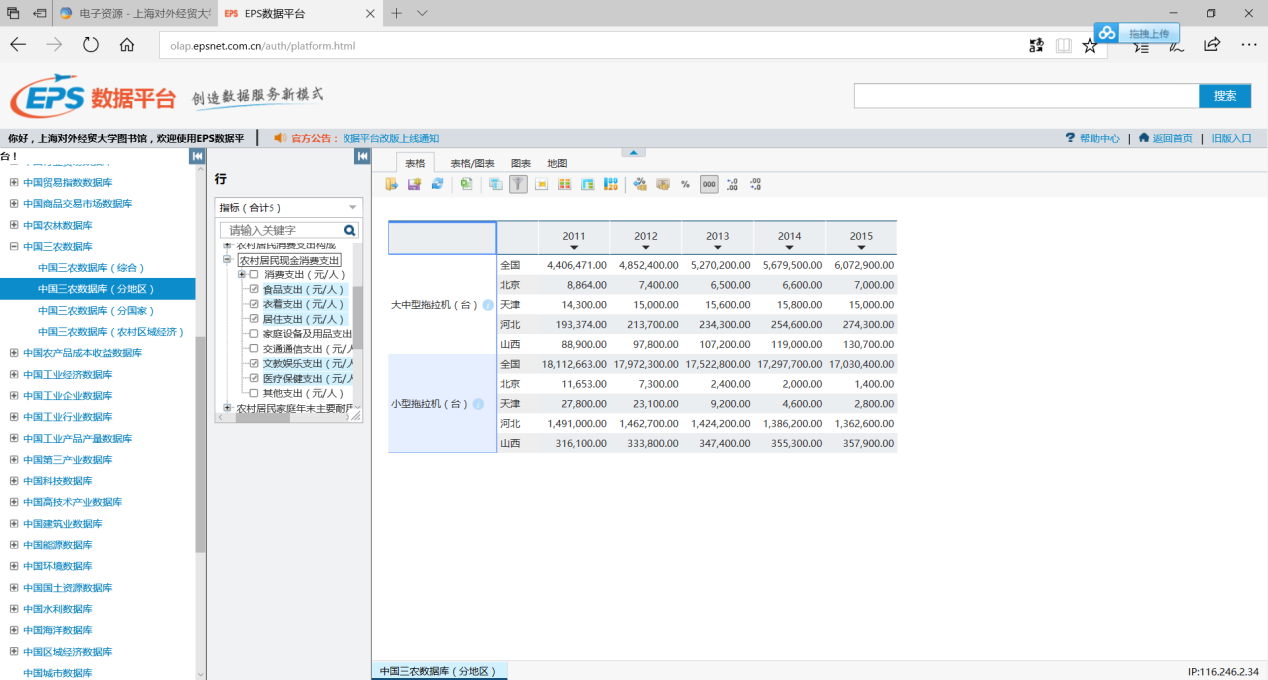


图4.1 筛选数据

为让数据直观显示并便于分析，调整数据位置，转置行列并生成表格。

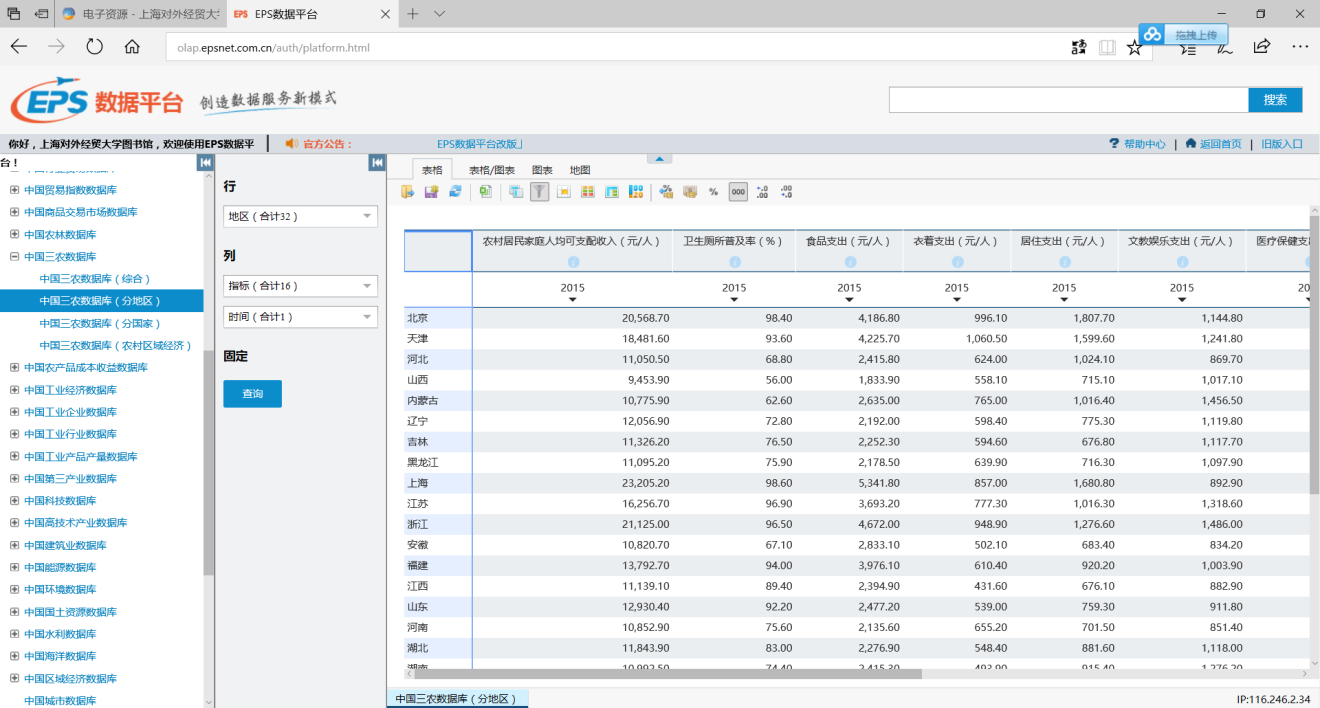


图4.2 整理数据

本文利用python的factor\_analyzer库对该14个指标进行因子分析，并使用其数学原理脱离库也做了一次分析。在使用因子分析前，先对指标进行相关预处理。

## 4.2 数据清洗及预处理

通过利用EPS数据平台对所需数据筛选并调整后导入，故不存在错误数据、多余数据与重复数据，故只简化表格与缺失值处理。

### 4.2.1 数据清洗

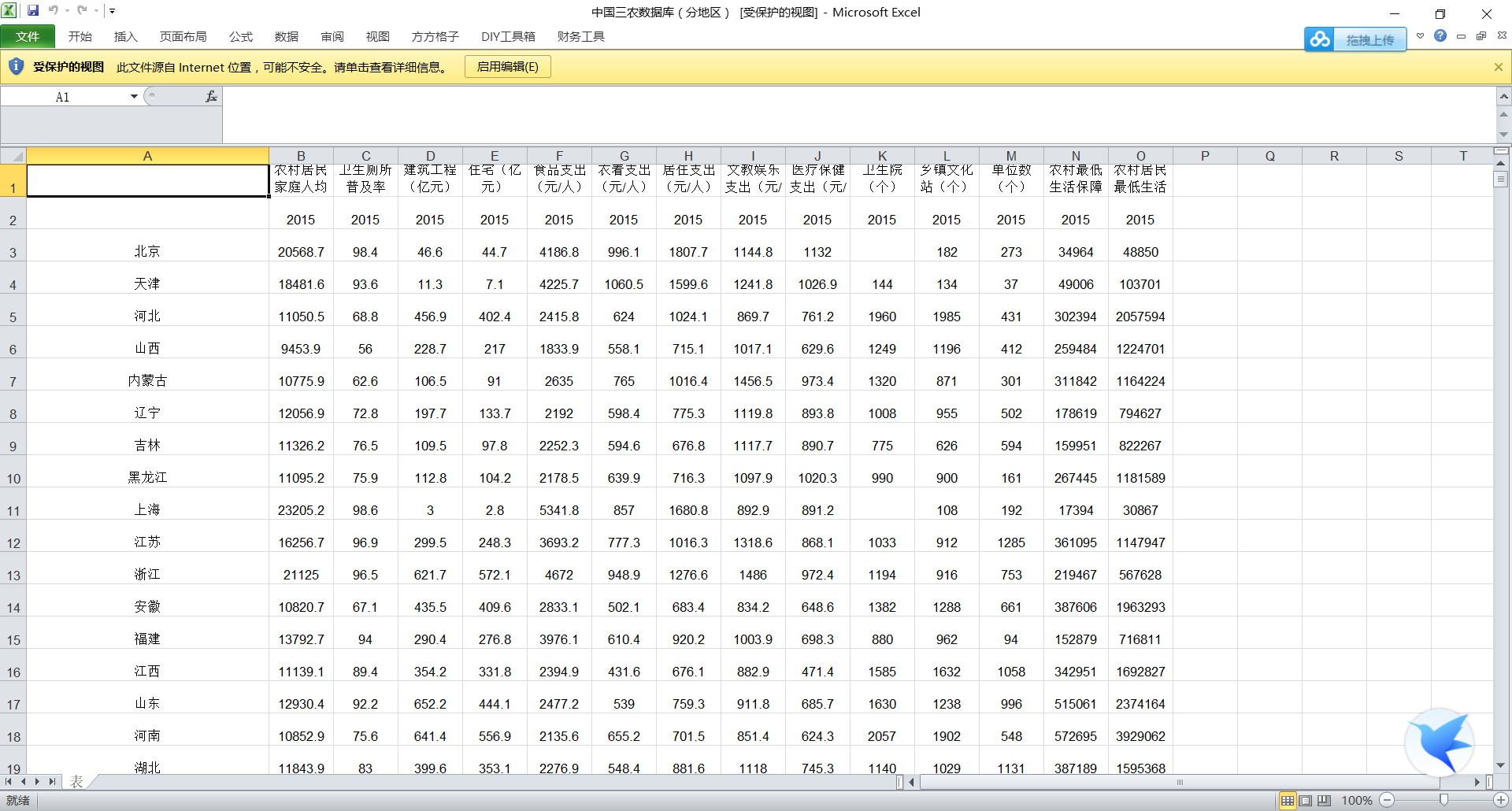


图4.3

搜集到的数据如图4.3所示。将该csv文件导入后，先将所有空行及年份所在行去除。可以看到图中存在缺失值，因为所分析问题不针对具体地区，所以考虑剔除所有缺失值大于1的地区。剩下的缺失值考虑用缺失值插补的办法补齐。清洗代码如图4.4所示。

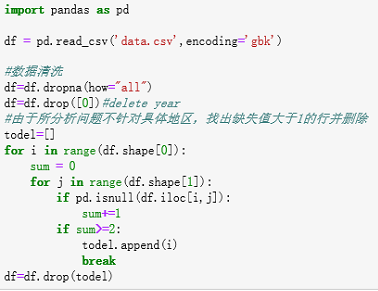


图4.4 数据清洗

### 4.2.2 缺失值处理

在现代的统计研究方法中，处理缺失值的方法主要有均值插补法、近邻插补法、拉格朗日插值法等方法。拉格朗日插值是根据若干个不同的地方得到相应的观测值，找到一个[多项式](https://baike.baidu.com/item/%E5%A4%9A%E9%A1%B9%E5%BC%8F/10660961)，其恰好在各个观测点取到观测到的值的方法。利用该多项式可较为精确地估计缺失值。故采用该方法作缺失值处理，处理过程如图4.5所示。

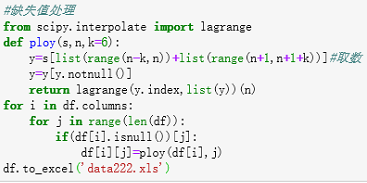


图4.5

从导出的文件中可以看到相关缺失的估计值为：北京x10为288.51，上海x10为140.98。

### 4.2.3 数据标准化处理及相关检验

将补充好缺失值的数据通过Pearson积矩相关系数标准化。标准化处理后计算其相关系数矩阵，并进行KMO检验和Bartlett检验，以判断样本数据是否适用于因子分析。相关代码实现如图4.6、图4.7所示。



图4.6 Kaiser-Meyer-Olkin Test

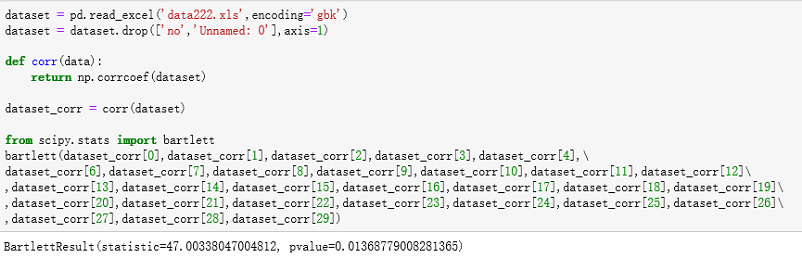


图4.7 Bartlett Test

|  |  |
| --- | --- |
| KMO和Bartlett检验 表4.1 | |
| KMO取样适切性量数 | 0.742 |
| Bartlett的球形度检验 | 0.014 |

KMO和Bartlett检验结果如表4.1所示。检验结果显示该相关系数矩阵不是一个单位阵，KMO值为0.742，大于0.7，表现为适合分析，显著性水平为0.014，小于0.05，通过了Bartlett球体检验，因此选取的变量适合做因子分析。

# 5 因子分析过程

本文在分析过程中使用了factor\_analyzer库，并根据其数学原理重复了过程。

## 5.1 确定因子个数及命名

### 5.1.1 保留因子个数

本文选择14个农村幸福感衡量指标以主成分方法作为因子提取方法。首先采取特征值大于1 的方法提取主成分，但是发现方差累积率较低，所以采取固定因子数量的方法提取四个主成分。由总方差解释表5.3可以看出到第五个因子时，方差累积率为86.4725%，已经超过经验所要求的85%的水平，因此选择保留四个因子。



图5.1

|  |  |
| --- | --- |
| 公因子方差 表5.1 | |
| 农村居民家庭人均可支配收入（元/人） | 0.979015 |
| 卫生厕所普及率（%） | 0.714303 |
| 建筑工程（亿元） | 0.977686 |
| 住宅（亿元） | 0.934984 |
| 食品支出（元/人） | 0.908247 |
| 衣着支出（元/人） | 0.788777 |
| 居住支出（元/人） | 0.775803 |
| 文教娱乐支出（元/人） | 0.476636 |
| 医疗保健支出（元/人） | 0.920935 |
| 卫生院（个） | 0.999084 |
| 乡镇文化站（个） | 0.995501 |
| 单位数（个） | 0.610154 |
| 农村最低生活保障支出（万元） | 0.729777 |
| 农村居民最低生活保障人数（人） | 0.773618 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 总方差解释 表5.2 | | | | | |
|  | Factor1 | Factor2 | Factor3 | Factor4 | Factor5 |
| SS Loadings  特征值 | 6.834592 | 3.194403 | 0.961788 | 0.593738 | 0.457663 |
| Proportion Var  总方差（即贡献率） | 0.488185 | 0.228172 | 0.068699 | 0.042410 | 0.032690 |
| CumulativeVar  累积率 | 0.488185 | 0.716357 | 0.785056 | 0.827466 | 0.864725 |

不借助factor\_analyzer库求因子个数，即当特征值的比重大于85%时，输出所求的因子个数。可以看到因子个数亦为4。

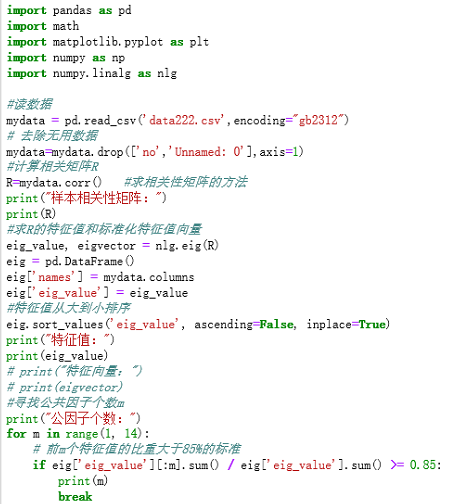


图5.2

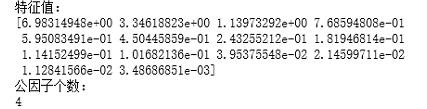


图5.3

### 5.1.2 命名

从旋转后的成分矩阵表4.2可以得出，x10、x11、x13、x14与第一因子的相关系数较大，故将第一因子命名为社会服务因子；x1、x2、x5、x6、x12与第二因子的相关系数较大，故将第二因子命名为生活水平因子；x7、x8、x9与第三因子的相关系数较大，故将第三因子命名为经济实力因子；x3、x4与第四因子的相关系数较大，故将第四因子命名为农村建设因子。



图5.4

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 成分矩阵 表5.3 | | | | |
|  | Factor1 | Factor2 | Factor3 | Factor4 |
| 农村居民家庭人均可支配收入（元/人） | -0.741840 | 0.619400 | -0.181992 | -0.109138 |
| 卫生厕所普及率（%） | -0.480037 | 0.554495 | -0.389450 | -0.157262 |
| 建筑工程（亿元） | 0.619959 | 0.656901 | -0.187991 | 0.855637 |
| 住宅（亿元） | 0.626392 | 0.634952 | -0.191613 | 0.820526 |
| 食品支出（元/人） | -0.712575 | 0.533618 | -0.277939 | -0.196181 |
| 衣着支出（元/人） | -0.723884 | 0.391067 | 0.334416 | 0.001439 |
| 居住支出（元/人） | -0.774069 | 0.412782 | 0.663750 | -0.046562 |
| 文教娱乐支出（元/人） | -0.464631 | 0.330300 | 0.439945 | 0.306728 |
| 医疗保健支出（元/人） | -0.737966 | 0.285507 | 0.511962 | 0.180892 |
| 卫生院（个） | 0.785132 | 0.472499 | 0.320779 | -0.237691 |
| 乡镇文化站（个） | 0.794532 | 0.445168 | 0.277130 | -0.298739 |
| 单位数（个） | 0.434615 | 0.630926 | 0.109449 | -0.105913 |
| 农村最低生活保障支出（万元） | 0.835680 | 0.171910 | -0.019683 | 0.038422 |
| 农村居民最低生活保障人数（人） | 0.866442 | 0.144452 | -0.006787 | -0.044546 |

不借助factor\_analyzer库求成分矩阵，可以看出除了数据有微小差别外原始变量与共同因素的对应完全一致。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因子载荷矩阵（成分矩阵） 表5.4 | | | | |
|  | factor1 | factor2 | factor3 | factor4 |
| 0 | -0.738795 | 0.599196 | -0.202794 | -0.063935 |
| 1 | -0.496873 | 0.584711 | -0.517871 | -0.057619 |
| 2 | 0.611422 | 0.651266 | -0.119407 | 0.895104 |
| 3 | 0.621681 | 0.637594 | -0.129750 | 0.893243 |
| 4 | -0.716630 | 0.525524 | -0.320552 | -0.105393 |
| 5 | -0.741144 | 0.398609 | 0.332963 | -0.230594 |
| 6 | -0.793158 | 0.421061 | 0.630402 | -0.169792 |
| 7 | -0.498850 | 0.379829 | 0.473675 | 0.630747 |
| 8 | -0.741195 | 0.279273 | 0.498156 | -0.024852 |
| 9 | 0.774395 | 0.471547 | 0.243862 | -0.352917 |
| 10 | 0.783951 | 0.444199 | 0.191720 | -0.411541 |
| 11 | 0.450954 | 0.704585 | 0.126008 | -0.152386 |
| 12 | 0.859246 | 0.194538 | 0.007139 | 0.090829 |
| 13 | 0.885451 | 0.161296 | -0.010026 | -0.067972 |



图5.5

## 5.2 因子得分

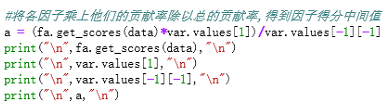
将四个因子确定后, 还需要根据因子载荷矩阵计算各个因子的得分。因此,用因子旋转后的成分得分系数矩阵对30个省市、自治区进行分析。根据因子得分表5.5，可以进一步计算每个省市的各因子的得分，并可以据此对各省市农村幸福感水平进行进一步比较分析。



图5.6

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因子得分 表5.5 | | | | |
|  | Factor1 | Factor2 | Factor3 | Factor4 |
| 0 | -2.137689 | 0.651531 | 0.668025 | -0.404874 |
| 1 | -1.897958 | 0.235468 | 0.258827 | -0.679857 |
| 2 | 0.677230 | 0.792481 | 0.527338 | -0.194066 |
| 3 | 0.349539 | -0.558523 | 0.458334 | -0.247075 |
| 4 | -0.330139 | -0.361002 | 2.143390 | 0.374028 |
| 5 | -0.275593 | 0.025416 | 0.580240 | -0.216648 |
| 6 | -0.652702 | -0.803876 | 0.628110 | 0.597715 |
| 7 | -0.467175 | -0.319421 | 1.478733 | -0.302540 |
| 8 | -2.269976 | 0.766835 | -1.204037 | -1.567623 |
| 9 | -0.728556 | 0.889314 | -0.205469 | 0.134404 |
| 10 | -0.971272 | 2.465195 | -0.364523 | 1.266881 |
| 11 | 0.255340 | -0.005137 | -0.581653 | 1.331314 |
| 12 | -0.269420 | 0.381454 | -1.077252 | -0.963790 |
| 13 | 0.745918 | 0.147666 | -0.732965 | -1.320629 |
| 14 | 0.768625 | 1.292123 | -0.900911 | 1.515853 |
| 15 | 1.070255 | 0.778949 | -0.373868 | 1.297140 |
| 16 | 0.076411 | 0.175205 | -0.056886 | 1.143417 |
| 17 | 1.071789 | 1.104903 | -0.087108 | 1.145331 |
| 18 | 0.115090 | 0.436508 | -1.369246 | -0.358769 |
| 19 | 0.888438 | -0.334590 | -0.745636 | 0.292708 |
| 20 | -0.298438 | -1.338634 | -1.774191 | -0.463067 |
| 21 | -0.225594 | -0.915992 | -0.248929 | -0.272123 |
| 22 | 1.981596 | 1.809696 | 2.098686 | -3.010878 |
| 23 | 1.144308 | -1.012603 | -0.511748 | -0.915535 |
| 24 | 1.203013 | -0.613506 | -1.067749 | 0.181934 |
| 25 | 0.442723 | -0.241226 | 1.413078 | 0.747978 |
| 26 | 0.777608 | -1.352772 | 0.229954 | -0.967042 |
| 27 | -0.747881 | -1.703114 | 0.941792 | 1.464923 |
| 28 | -0.476471 | -1.634306 | -0.032314 | 0.597008 |
| 29 | 0.180981 | -0.758044 | -0.092021 | -0.206118 |

根据该表得到以上四个因子的表达式后，可以计算出30个省市对于四个因子的得分情况。



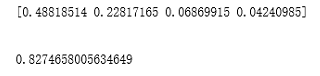


图5.7

再根据四个因子的方差贡献率组成表达式：

而不利用factor\_analyzer库求得的方差贡献率求得的表达式为：

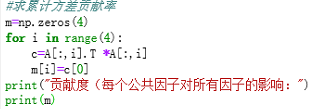




图5.8

以求得30个省市在幸福感方面的综合得分F。因此，可以建立对幸福指数测算的评价模型为：

农村居民幸福指数=(\*社会服务因子+\*生活水平因子+\*经济实力因子+\*农村建设因子)/

## 5.3 总分排名

最后，对各省市因子得分和总分进行汇总及排名，如表5.6所示。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **因子得分及排名 表5.6** | | | | | | | |
|  |  | **Factor1** | **Factor2** | **Factor3** | **Factor4** | **score** | **rank** |
| **22** | 四川 | 1.169095 | 0.499019 | 0.17424 | -0.15432 | 1.688039 | 1 |
| **17** | 湖南 | 0.63233 | 0.304674 | -0.00723 | 0.058701 | 0.988474 | 2 |
| **15** | 河南 | 0.631425 | 0.214793 | -0.03104 | 0.066482 | 0.88166 | 3 |
| **14** | 山东 | 0.45347 | 0.3563 | -0.0748 | 0.077692 | 0.812665 | 4 |
| **2** | 河北 | 0.399549 | 0.218525 | 0.043781 | -0.00995 | 0.651909 | 5 |
| **24** | 云南 | 0.709749 | -0.16917 | -0.08865 | 0.009325 | 0.461253 | 6 |
| **19** | 广西 | 0.524157 | -0.09226 | -0.06191 | 0.015002 | 0.384992 | 7 |
| **13** | 江西 | 0.440074 | 0.040718 | -0.06085 | -0.06769 | 0.352253 | 8 |
| **25** | 陕西 | 0.261196 | -0.06652 | 0.117319 | 0.038336 | 0.350333 | 9 |
| **23** | 贵州 | 0.675115 | -0.27922 | -0.04249 | -0.04692 | 0.306481 | 10 |
| **11** | 安徽 | 0.150645 | -0.00142 | -0.04829 | 0.068233 | 0.169171 | 11 |
| **16** | 湖北 | 0.045081 | 0.048312 | -0.00472 | 0.058603 | 0.147273 | 12 |
| **10** | 浙江 | -0.57303 | 0.679771 | -0.03026 | 0.064931 | 0.141411 | 13 |
| **3** | 山西 | 0.20622 | -0.15401 | 0.038053 | -0.01266 | 0.077598 | 14 |
| **18** | 广东 | 0.0679 | 0.120366 | -0.11368 | -0.01839 | 0.056199 | 15 |
| **26** | 甘肃 | 0.45877 | -0.37302 | 0.019092 | -0.04956 | 0.055275 | 16 |
| **4** | 内蒙古 | -0.19477 | -0.09955 | 0.177952 | 0.01917 | -0.0972 | 17 |
| **5** | 辽宁 | -0.16259 | 0.007009 | 0.048174 | -0.0111 | -0.11852 | 18 |
| **29** | 新疆 | 0.106775 | -0.20903 | -0.00764 | -0.01056 | -0.12046 | 19 |
| **12** | 福建 | -0.15895 | 0.105185 | -0.08944 | -0.0494 | -0.1926 | 20 |
| **9** | 江苏 | -0.42983 | 0.245226 | -0.01706 | 0.006889 | -0.19478 | 21 |
| **7** | 黑龙江 | -0.27562 | -0.08808 | 0.12277 | -0.01551 | -0.25644 | 22 |
| **21** | 重庆 | -0.1331 | -0.25258 | -0.02067 | -0.01395 | -0.42029 | 23 |
| **6** | 吉林 | -0.38508 | -0.22167 | 0.052148 | 0.030635 | -0.52396 | 24 |
| **28** | 宁夏 | -0.28111 | -0.45066 | -0.00268 | 0.030598 | -0.70385 | 25 |
| **20** | 海南 | -0.17607 | -0.36912 | -0.1473 | -0.02373 | -0.71623 | 26 |
| **27** | 青海 | -0.44123 | -0.46963 | 0.078191 | 0.075081 | -0.75759 | 27 |
| **0** | 北京 | -1.26119 | 0.179658 | 0.055462 | -0.02075 | -1.04682 | 28 |
| **1** | 天津 | -1.11975 | 0.06493 | 0.021489 | -0.03484 | -1.06818 | 29 |
| **8** | 上海 | -1.33923 | 0.211453 | -0.09996 | -0.08034 | -1.30809 | 30 |

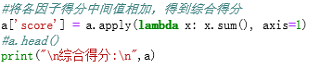


图5.9

利用地图慧工具，将分数可视化，效果如图5.10所示。



图5.10

# 6 结果分析及政策建议

由表5.6可得，在社会服务和生活水平的影响下，农村居民的幸福感水平普遍偏低。在社会服务方面，除四川外，其他省市在该方面发展严重不足。在生活水平方面，浙江，四川，山东发展更好，能给农村居民带来较大的幸福感，但其他地区相对欠缺。

社会服务水平对农村居民幸福指数的影响最大。社会服务水平包括社会保障、居住环境、医疗服务、文化体制、教育制度等，这些与居民的生活密切相关，所以会直接影响居民的幸福感及对政府的满意度。从30个省市第一因子得分情况来看，大多数省市在该方面的建设存在很大程度的缺失，不能在一定程度上给居民带来幸福感，甚至会削弱居民的整体幸福感水平。在这一方面，四川省的幸福指数远远超过其他的省市，这也正是四川省农村居民幸福感排在全国第二名的重要原因。在徐仲安、靳共元关于“四川省震后农村居民幸福感影响因素的实证分析”中提到，四川省震后灾区居民的幸福感程度远高于非灾区的居民。这说明近年来政府对灾区的经济援建工作十分重视，社会保障制度十分完善，本文建议四川省政府可以将该种工作体制转变为常态化机制，并同时加大对非灾区农村地区的基础设施建设投入，改善地区间经济社会发展不平衡的状况，从而提高四川省农村居民的整体幸福感水平。

浙江农村居民在生活水平方面的幸福感指数最高。党的十六大以来，浙江大力实施统筹城乡发展方略，把解决“三农”问题作为全面建成小康社会的重中之重，出台一系列统筹城乡兴“三农”的政策，建立健全以工促农、以城带乡的发展机制，充分发挥工业化、城市化、市场化对“三农”的带动作用，对居民生活水平有显著提高，城乡一体化进程明显加速。

在经济实力方面内蒙古位居首位。内蒙古具有丰富的矿产资源，工业水平发展迅速，使得该地区人均GDP水平屡次位于我国前列，这也正是其经济水平高的重要原因。然而内蒙古东部区缺少投资，资源匮乏，经济发展缓慢，以农牧业为主，拉低了内蒙古整体幸福感水平。所以，政府要加大对内蒙古东部地区的建设，一视同仁，共同发展。

在农村建设方面各省市较幸福指数相对平均，且因子载荷也较小。可以说就提升幸福感而言，社会服务、生活水平、经济实力占大头，做好这三方面的工作比加强农村建设更为重要。

而上海这样的大城市农村幸福指数最低，似乎有些出乎意料。笔者认为上海作为我国经济发展中心，其经济实力很强，这种较高的生活水平惠及上海的各个地区，但是由于过分重视中心城区的发展，而相对忽视对农村的建设。在繁华的大上海背后还隐藏着众多的老城区，有些城区居住条件简陋、生活不便，因为种种原因使得这些问题得不到有效的解决，故而降低了农村居民整体幸福感水平。上海物价高，房价高，居民的经济压力不小，这也在一定程度上影响幸福感水平。

综上所述，四川、湖南、云南等一些南方地区的农村幸福感水平高于北方地区，这些地区的幸福感水平是正值，说明人们是感到幸福的。其中，四川农村的幸福感水平是最高的。西部地区人们总体幸福感差，主要是因为其经济实力不强，各方面建设不完善，国家应加大对其人力与物质投入，使人们的幸福感增强,从而带动经济、社会服务和农村建设。

# 7 附录

### 探索：python与SPSS在结果上的不同

笔者利用SPSS对该数据集用同样的实验过程分析了一遍，发现结果不尽相同。

|  |  |  |  |
| --- | --- | --- | --- |
| 总方差解释 表 7.1 | | | |
| 组件 | 初始特征值 | | |
| 总计 | 方差百分比 | 累积% |
| 1 | 6.983 | 49.880 | 49.880 |
| 2 | 3.346 | 23.901 | 73.781 |
| 3 | 1.140 | 8.141 | 81.922 |
| 4 | .769 | 5.490 | 87.412 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 旋转后的成分矩阵 表7.2 | | | | |
|  | 组件 | | | |
| 1 | 2 | 3 | 4 |
| 农村居民家庭人均可支配收入（元/人） | -.099 | **.896** | .368 | -.028 |
| 卫生厕所普及率（%） | -.045 | **.914** | .016 | .148 |
| 建筑工程（亿元） | .632 | .051 | -.048 | **.716** |
| 住宅（亿元） | .628 | .043 | -.066 | **.717** |
| 食品支出（元/人） | -.150 | **.906** | .233 | -.036 |
| 衣着支出（元/人） | -.037 | .526 | **.662** | -.372 |
| 居住支出（元/人） | -.152 | **.715** | .481 | -.247 |
| 文教娱乐支出（元/人） | -.139 | .157 | **.884** | .247 |
| 医疗保健支出（元/人） | -.136 | .332 | **.803** | -.321 |
| 卫生院（个） | **.950** | -.208 | -.111 | .087 |
| 乡镇文化站（个） | **.943** | -.193 | -.180 | .057 |
| 养老院单位数（个） | **.807** | .142 | .080 | .231 |
| 农村最低生活保障支出（万元） | **.621** | -.380 | -.301 | .402 |
| 农村居民最低生活保障人数（人） | **.662** | -.378 | -.380 | .296 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 因子得分及排名 表4.4 | | | | | | | | | | |
| 省份 | 第一因子 | | 第二因子 | | 第三因子 | | 第四因子 | | 综合得分 | |
| F1 | 排名 | F2 | 排名 | F3 | 排名 | F4 | 排名 | F | 排名 |
| 浙江 | 0.16125 | 10 | 1.55634 | 3 | 1.98369 | 2 | 2.0582 | 1 | 1.24 | 1 |
| 四川 | 4.04414 | 1 | 0.30402 | 10 | -0.77185 | 22 | -2.0414 | 30 | 1.02 | 2 |
| 湖南 | 1.3195 | 2 | -0.36105 | 18 | 1.09741 | 4 | 1.48632 | 2 | 0.78 | 3 |
| 江苏 | 0.32613 | 7 | 0.92378 | 7 | 1.00944 | 5 | 0.54059 | 9 | 0.68 | 4 |
| 山东 | 0.54077 | 5 | 0.58708 | 8 | -0.41004 | 18 | 1.38611 | 3 | 0.47 | 5 |
| 北京 | -0.40163 | 21 | 1.75853 | 2 | 0.97002 | 6 | -1.33327 | 28 | 0.42 | 6 |
| 河南 | 1.05283 | 3 | -0.12471 | 14 | -0.28991 | 16 | 0.96517 | 6 | 0.39 | 7 |
| 湖北 | 0.16423 | 9 | -0.039 | 13 | 0.48081 | 11 | 0.92297 | 8 | 0.28 | 8 |
| 广东 | -0.03075 | 15 | 1.07187 | 5 | -1.02516 | 25 | 0.97587 | 5 | 0.24 | 9 |
| 上海 | -0.66457 | 25 | 2.65232 | 1 | -0.70068 | 21 | -1.51747 | 29 | 0.23 | 10 |
| 天津 | -0.72235 | 27 | 1.28767 | 4 | 1.12235 | 3 | -1.15216 | 26 | 0.23 | 10 |
| 河北 | 0.85264 | 4 | -0.13751 | 15 | 0.08966 | 13 | -0.29662 | 17 | 0.22 | 12 |
| 江西 | 0.39364 | 6 | 0.45935 | 9 | -1.1168 | 26 | 0.37004 | 13 | 0.09 | 13 |
| 福建 | -0.61179 | 24 | 1.06023 | 6 | -0.5691 | 20 | 0.40488 | 11 | 0.06 | 14 |
| 安徽 | 0.20074 | 8 | -0.15241 | 17 | -0.43195 | 19 | 0.53091 | 10 | 0.01 | 15 |
| 内蒙古 | -0.13056 | 16 | -1.2326 | 28 | 2.25354 | 1 | -0.42234 | 20 | -0.01 | 16 |
| 辽宁 | -0.28619 | 19 | -0.58524 | 21 | 0.81665 | 8 | -0.35087 | 18 | -0.16 | 17 |
| 陕西 | 0.09331 | 12 | -1.32056 | 29 | 0.78017 | 9 | 0.02788 | 14 | -0.21 | 18 |
| 黑龙江 | -0.33248 | 20 | -0.73636 | 23 | 0.87952 | 7 | -0.64848 | 23 | -0.25 | 19 |
| 吉林 | -0.52106 | 23 | -0.56337 | 19 | 0.6694 | 10 | -0.387 | 19 | -0.26 | 20 |
| 广西 | -0.46828 | 22 | -0.14553 | 16 | -1.26141 | 28 | 1.34594 | 4 | -0.28 | 21 |
| 山西 | -0.19182 | 17 | -1.15113 | 27 | 0.36054 | 12 | -0.24857 | 16 | -0.38 | 22 |
| 新疆 | -0.23668 | 18 | -0.00083 | 12 | -1.11922 | 27 | -0.7826 | 24 | -0.43 | 23 |
| 云南 | 0.00135 | 14 | -0.96955 | 26 | -1.31572 | 29 | 0.96349 | 7 | -0.44 | 24 |
| 贵州 | 0.10829 | 11 | -1.34931 | 30 | -0.89055 | 23 | 0.01116 | 15 | -0.57 | 25 |
| 重庆 | -0.70725 | 26 | -0.5647 | 20 | -0.3379 | 17 | -0.5668 | 22 | -0.57 | 25 |
| 甘肃 | 0.00895 | 13 | -0.92259 | 25 | -0.95938 | 24 | -0.78471 | 25 | -0.59 | 27 |
| 宁夏 | -1.23507 | 29 | -0.60693 | 22 | 0.07322 | 15 | -0.56644 | 21 | -0.67 | 28 |
| 青海 | -0.93758 | 28 | -0.8359 | 24 | 0.07448 | 14 | -1.28744 | 27 | -0.74 | 29 |
| 海南 | -1.78973 | 30 | 0.13808 | 11 | -1.46124 | 30 | 0.39666 | 12 | -0.82 | 30 |

笔者认为原因有二。其一SPSS默认用的提取方法是主成分分析，即PCA中的principal 函数，而模块factor\_analyzer中用的提取方法是fa函数，其二SPSS选择了最大平衡法旋转，而python的代码中， rotation=None，没有旋转。

# 参考文献

1. [国民幸福指数测算方法研究](http://kns.cnki.net/kcms/detail/detail.aspx?filename=DBDS201504008&dbcode=CJFQ&dbname=CJFDTEMP&v=)[J]. 李刚,王斌,刘筱慧.东北大学学报(社会科学版).2015(04)

1. [基于因子分析、聚类分析对我国居民幸福指数的研究——论国民幸福指数](http://kns.cnki.net/kcms/detail/detail.aspx?filename=ZGSM201627078&dbcode=CJFQ&dbname=CJFDTEMP&v=" \t "kcmstarget)[J].赵一纯.中国商论.2016(27)
2. [国内外幸福指数研究现状评述及展望](http://kns.cnki.net/kcms/detail/detail.aspx?filename=JJGL201404006&dbcode=CJFQ&dbname=CJFD2014&v=)[J].姜文华,朱孔来,李俊杰.经济与管理.2014(04)
3. 基于因子分析的苏州居民幸福指数构建及实证分析[J].宣春霞.湖北师范学院学报（哲学社会科学版）.2015(01)
4. 基于因子分析法的城镇居民幸福感研究[J].武娇艳,杨苹.公共管理.2016(04)
5. [基于因子分析法的我国农村居民主观幸福感研究](http://kns.cnki.net/kcms/detail/detail.aspx?filename=AHNY201224136&dbcode=CJFQ&dbname=CJFD2012&v=)[J].郭玉彬.安徽农业科学.2012(24)
6. [我国农民幸福指数的实证研究](http://kns.cnki.net/kcms/detail/detail.aspx?filename=NJWT201412004&dbcode=CJFQ&dbname=CJFD2014&v=)[J]. 熊彩云,孟荣钊,史亚峰.农业经济问题.2014(12)
7. 我国商业银行经营绩效评价实证研究——基于因子分析和缺失值插补[J].谢英欣,张世玲.统计与管理.2015(09)
8. 影响居民幸福感因素的调查与分析——以江苏省常州市武进区为例[J].李素君.才智. 2010(29)

[10]农村居民幸福感影响因素的实证分析[J].徐仲安,靳共元.中国农村观察.2013(04)

[11] 因子分析—SPSS操作及其原理[J].陶鑫.2008

1. 因数据有限，本文所选数据不包含西藏、香港、台湾地区 [↑](#footnote-ref-1)
2. 因数据有限，不包含西藏、香港、台湾地区 [↑](#footnote-ref-2)