面向提高全民科学素质服务的统计指标降维方法及应用研究

程　豪
（中国科协创新战略研究院，北京 100038）

摘　要：科学素质是国民素质的重要组成部分，是社会文明进步的基础。本文从《中国科学技术协会统计年鉴2022》中关于提高全民科学素质服务的指标数据出发，通过统计描述的方法刻画科普基础设施建设、科普宣讲活动、科普传播等方面的基本情况。在此基础上，聚焦基层服务全民科学素质工作开展情况，进一步反映基层科普的基本规律和特点。面对科普指标数量较多的问题，本文采用经典的主成分分析方法，分别对全指标体系和经业务判断删除部分指标后的指标体系进行主成分降维分析，以期为相关部门进行指标修订和删减提供参考与借鉴。

关键词：全民科学素质，基层组织，主成分分析

1. 引言

习近平总书记指出：“科技创新、科学普及是实现创新发展的两翼，要把科学普及放在与科技创新同等重要的位置。”这一重要指示精神是新发展阶段科普和科学素质建设高质量发展的根本遵循[1]。2021年06月25日，国务院发布《全民科学素质行动规划纲要（2021—2035年）》，指出科学素质是国民素质的重要组成部分，是社会文明进步的基础。开展全民科学素质服务的统计分析是全面了解掌握新时代我国公民科学素质的发展状况，推动科普和科学素质建设高质量发展的重要工作。公民具备科学素质是指崇尚科学精神，树立科学思想，掌握基本科学方法，了解必要科技知识，并具有应用其分析判断事物和解决实际问题的能力[2]。

习近平总书记指出：“没有全民科学素质普遍提高，就难以建立起宏大的高素质创新大军，难以实现科技成果快速转化。”自《全民科学素质行动计划纲要（2006—2010—2020年）》印发，全民科学素质行动取得显著成效，各项目标任务如期实现，2020年我国公民具备科学素质的比例达到10.56%。《全民科学素质行动规划纲要（2021—2035年）》提出2025年我国公民具备科学素质的比例超过15%，2035年达到25%[1]。

目前已有专家学者围绕《全民科学素质行动规划纲要（2021—2035年）》展开了研究工作[3]。研究成果可归纳为两个方面：一是国际国内形势急剧转变以及数字化时代的到来对全民科学素质提出更高要求[4–5]，二是从科学知识到科学精神和能力提升的科学素质内涵转变[6–8]。此外，整体层面存在不平衡不充分及制度安排不到位等问题[9]，

作者简介：程　豪，中国科协创新战略研究院，副研究员，研究方向为数据分析与统计建模，综合统计与调查研究。

项目来源：国家自然科学基金委青年科学基金项目（项目编号：72001197）；国家统计局全国统计科学研究项目（项目编号：2021LY052）；北京工商大学数字商科与首都发展创新中心项目（项目编号：SZSK202317）。

个体层面则表现出“对信息时代所需的核心能力表现不足”“对信息化社会的适应力不足”等显著问题，因而在公民科学素质提升上需要特别关注数字技术、互联网环境、信息素养、数字素养等方面[10-14]。

本文以《中国科学技术协会统计年鉴2022》中有关提高全民科学素质服务的指标数据为对象，完成两项统计分析工作：一是通过统计描述的方法，分别刻画整体和基层两个维度下服务全民科学素质工作开展情况和规律特点。二是采用经典的主成分分析方法，分别对全指标体系和经业务判断删除部分指标后的指标体系进行主成分降维分析，以期为相关部门进行指标修订和删减提供参考与借鉴。

2. 数据简介与变量说明

在《中国科学技术协会统计年鉴2022》中关于提高全民科学素质服务的指标基础上，根据指标间逻辑关系和具体内涵，对提高全民科学素质服务相关指标进行梳理，形成指标体系[15]，共包括43个指标。其中，1个指标是地区指标，42个指标是业务指标。在坚持指标体系表意不重复的基本原则基础上，主要表现在科普基础设施建设、科普宣讲活动和科普传播共三个方面。具体指标体系如表1所示。

表1　反映提高全民科学素质服务统计研究变量说明

一级标题	二级标题	三级标题
科普基础设施建设	实体科技馆建设及使用情况	实体科技馆个数
		实体科技馆建筑面积
		实体科技馆展厅面积
		实体科技馆参观人次
	数字科技馆建设及使用情况	数字科技馆/科技馆官方网站个数
		数字科技馆/科技馆官方网站日均页面浏览量
		数字科技馆/科技馆官方网站科普资源总量
科普基础设施建设	流动科技馆建设及使用情况	流动科技馆个数
		本年度流动科技馆巡展站点数
		本年度流动科技馆巡展受众人次
	科普活动站情况	科普（技）活动站（室、中心）个数
		全年参加活动（培训）人次
	科普大篷车数量及使用情况	科普大篷车数量
		科普大篷车下乡次数
		科普大篷车覆盖人数
		科普大篷车行驶里程
		科普大篷车展品数量
	中国科普e站个数	中国科普e站个数
	科普画廊建设及使用情况	科普画廊建筑面积
		科普画廊展示面积

表1　反映提高全民科学素质服务统计研究变量说明

（续表）

一级标题	二级标题	三级标题
科普宣讲活动	举办科普宣讲活动情况	举办科技志愿服务活动次数
		科普活动受众人数
		参加活动科技人员、专家人数
		参加科普宣讲活动的学会、协会、研究会个次
		科普宣讲活动覆盖（社区）个数
	举办实用技术培训情况	举办实用技术培训次数
		实用技术培训人数
	推广新技术、新品种数量	推广新技术、新品种个数
科普传播	纸质媒体情况	编著科技图书种数
		科技图书总印数
		主办科技报纸种数
		报纸总印数
		制作科普挂图种数
		科普挂图总印数
	非纸质媒体情况	主办科普APP或设置科普栏目的综合类APP个数
		科普APP下载安装数
		科普APP更新频次
		主办科普微信公众号个数
		关注数
		年度总阅读数
		主办科普微博个数
		粉丝数

3. 全民科学素质服务状况分析

3.1 基于科普基础设施建设的统计描述

由表2可知，2021年科普基础设施建设的科技馆中，流动科技馆数量最多，有242个，占比达79%，其次为数字科技馆及科技馆官方网站，数量为38个，占比12%，而实体科技馆仅有26个，占比9%。原因可能是由于实体科技馆建设难度大，需要耗费大量时间和人力物力，而流动科技馆是以巡回展出的方式，将展览资源送到尚未建设科技馆的县域地区，其实现难度要小于实体科技馆，因此其数据量占比较大。

由表2可知，实体科技馆全年参观人次最多，为12539291人次，全年流动科技馆巡展受众人数为10521891人次，在提高全民科学素质方面起到了至关重要的作用。实体科技馆建筑面积利用率达48%。科普画廊面积利用率达72%。

3.2 基于科普宣讲活动的统计描述

由图1可知，2021年科普宣讲活动共举办6269

表2　科普基础设施建设基本情况

指标名称	单位	数量	指标名称	单位	数量
实体科技馆	个	26	科普大篷车	辆	74
实体科技馆建筑面积	平方米	1036915	科普大篷车下乡次数	次	2553
实体科技馆展厅面积	平方米	493725	科普大篷车行驶里程	千米	563077
实体科技馆全年参观人次	人次	12539291	科普大篷车覆盖人次	人次	1931018
数字科技馆	个	38	科普大篷车展品数量	件	1880
数字科技馆日均页面浏览量	次	89605	科普中国e站	个	23490
数字科技馆科普资源总量	个	13445	科普画廊建筑面积	平方米	7141
流动科技馆	个	242	科普画廊展示面积	平方米	5141
全年流动科技馆巡展站点数	个	680			
全年流动科技馆巡展受众	人次	10521891			
科普（技）活动站（室、中心）	个	183			
全年参加活动（培训）人次	人次	191013			

次，科普活动的受众人数达154783254人，平均每次科普活动受众达24690人，本年举办的科普宣讲活动共覆盖20242个村，平均每次宣讲活动覆盖3个村。除此之外，参加宣讲活动的科技人员有20297人，参加活动的学会、协会、研究会有1110个。由图2可知，2021年实用技术培训共举办8191场，实用技术培训人数达1819939人，平均每场实用技术培训覆盖222人，共推广新技术、新品种1396项。

3.3 基于科普传播的统计描述

表3可知，纸质媒体中制作科普挂图种数最

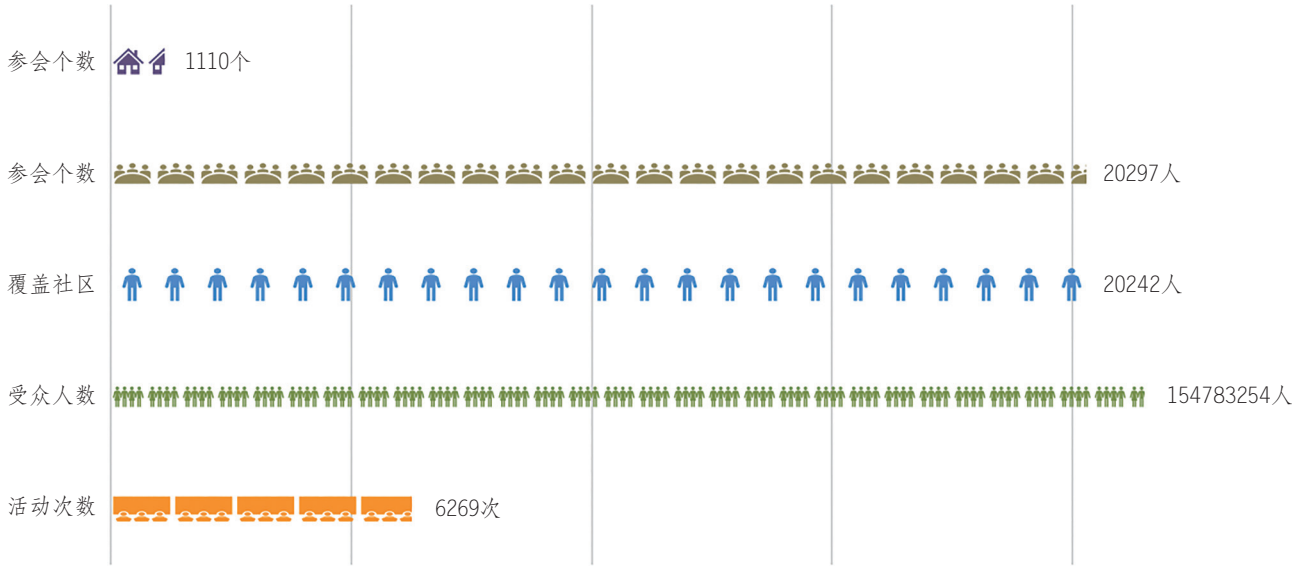


图1　2021年科普宣讲活动效果

注：数据来自《中国科学技术协会统计年鉴2022》中省级科协口径数据。活动次数：举办科普宣讲活动次数，受众人数：科普活动受众人数，覆盖社区：科普宣讲活动覆盖村（社区），专家人次：参加活动科技人员、专家人次，参会个数：参加活动的学会、协会、研究会。

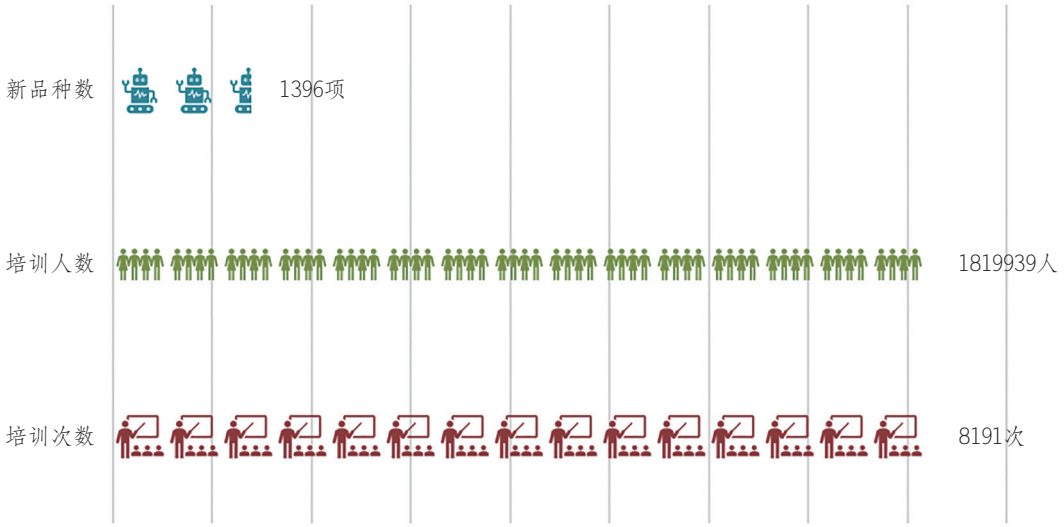


图2　2021年实用技术培训活动效果

注：数据来自《中国科学技术协会统计年鉴2022》中省级科协口径数据。培训次数：举办实用技术培训，培训人数：实用技术培训人数，新品种数：推广新技术、新品种。

多，为200种，但传播数量最少，为1006588张；主办科技报纸种数最少，为24种，但是传播数量最多，总印数达到了57261034份，编著科技图书数量及印刷量居中，分别为133种和1624097册，由此可以看出科技报纸传播效果最好。

非纸质媒体主要包含的指标为主办科普APP或设置科普栏目的综合类APP个数、科普APP下载安装数、科普APP更新频次、主办科普微信公众号个数、关注数、年度总阅读数、主办科普微博个数、粉丝数。非纸质媒体方面，其中主办科普微信公众号个数最多，为120个，关注数也最多，达到了16050689个，全年阅读量达到了59718133个。主办科普类APP有12个，科普APP的下载安装数和更新数分别为1746508次和9186次，下载安装数高说明用户对科普知识有较高的需求，但是更新次数少说明用户对APP的内容兴趣不高。主办科普微博共36个，其关注数有4493808个。通过对三种主要非纸质媒体科普传播情况的调查，发现科普微信公众号的科普传播效果最好。

表3　2021年纸质媒体科普传播情况和主办非纸质媒体传播情况

指标	编著科技图书（种）	主办科技报纸（种）	制作科普挂图（种）	科普APP（个）	微信公众号（个）	科普微博（个）
数量	133	24	200	12	120	36

注：数据来自《中国科学技术协会统计年鉴2022》中省级科协口径数据。

3.4 反映基层服务全民科学素质的统计描述

为进一步对细化分析全民科学素质服务情况在基层的表现，本文对2017—2021年来自县级科协体现全民科学素质的一些指标进行统计描述[16]。根据《中国科学技术协会统计年鉴》采用的“六大区”划分方式，先对不同地区内的省、自治区和直辖市数据进行均值计算，作为不同地区在相应反映基层服务提供全民科学素质方面的指标取值[17]。

经分析发现，2019—2021年科普人员数量整体呈现逐年递增的趋势，华东地区的“专职科普人员”和“兼职科普人员”均为最多，所有地区均在2021年达到峰值（如图5所示）。从指标“兼职科普人员”来看，华东地区的增长幅度最大、人数最多，2021年达到6.23万人。东北地区和西

北地区数据较低，其中西北地区人数在2021年达到最小值，为1.06万人。从指标“专职科普人员”来看，六大区整体发展较为平稳，其中华东地区人数最多，2021年数据为0.36万人，东北地区仅为0.08万人。

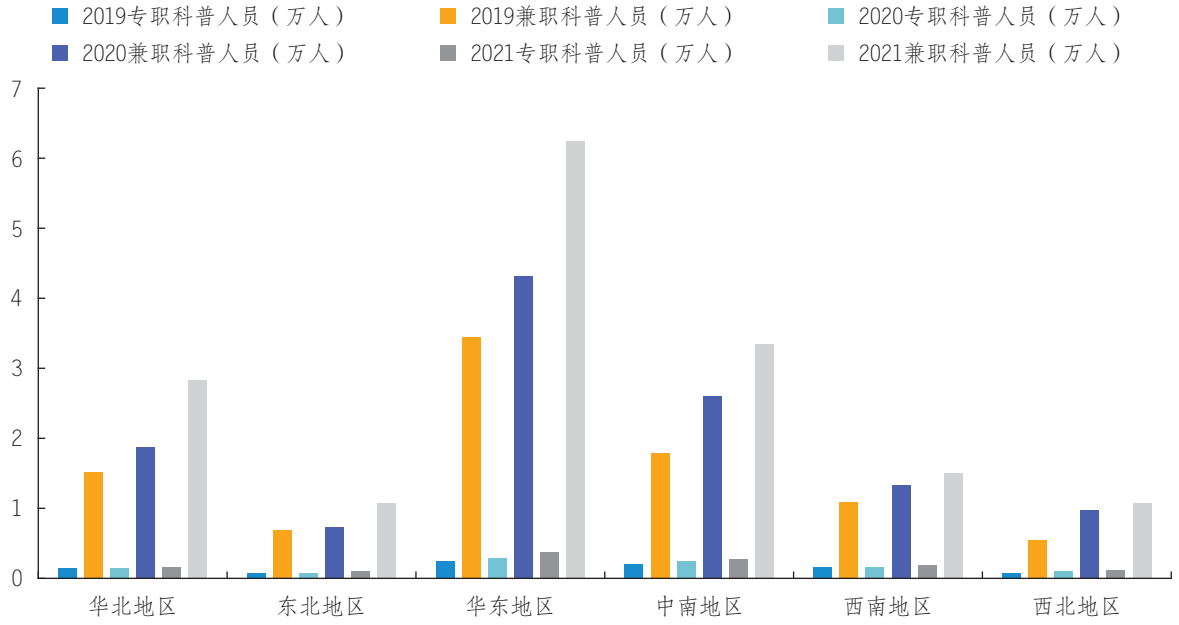


图3　2019—2021年科普人员基本情况

注：数据来自《中国科学技术协会统计年鉴2020—2022》中县级科协口径数据。

2019—2021年实体科技馆、科普（技）活动站（室、中心）数量在不同地区表现出不同规律（如图4所示）。从西北地区来看，活动站的个数在2019—2021年间呈现小幅度下降趋势，3年间活动站个数最多的是2019年和2020年，均为529个，2021年活动站的个数最少，为502.8个。与其他地区相比，西北地区2019到2021年的活动站个数变化情况最为平稳；实体科技馆的数量在2019—2021年间呈现小幅度增长趋势，3年间实体科技馆数量最多为2021年，科技馆数量为22座，2019年科技馆数量最少，仅为18座。从西南地区来看，活动站个数在2019—2021年间呈现下降趋势，3年间活动站个数最多的是2019年，为1476.6个，2021年活动站个数最少，为1084.8个；实体科技馆的数量在2019–2021年间呈波动发展趋势，3年间实体科技馆数量最多为2019年，科技馆数量为24.6座，2020年科技馆数量最少，仅为22.4座。从中南地区看，活动站个数在2019—2021年间呈波动发展趋势，3年间活动站个数最多的是2020年，为1827.67个，2021年活动站个数最少，为1434.17个；实体科技馆的数量在2019—2021年间呈小幅度增长趋势，3年间实体科技馆数量最多为2021年，科技馆数量为21座，2019年科技馆数量最少，仅为19.17座。从华东地区看，活动站个数在2019—2021年间呈波动发展趋势，3年间活动站的个数最多的是2020年，为3544个，2021年活动站个数最少，为2792.33个；实体科技馆的数量在2019—2021年间呈小幅度下降趋势，3年间实体科技馆数量最多为2019年，科技馆数量为33.67座，2021年科技馆数量最少，仅为29.5座。从东北地区看，活动站个数在2019–2021年间呈小幅度下降趋势，3年间活动站个数最多的是2019年，为1481个，2021年活动站个数最少，为1033.33个；实体科技馆的数量在2019—2021年间呈波动发展趋势，3年间实体科技馆数量最多为2019年，科技馆数量为17.67座，2020年科技馆数量最少，仅

为14.67座。从华北地区看，活动站个数在2019—2021年间呈波动发展趋势，3年间活动站个数最多的是2020年，为651.67个，2019年活动站个数最少，为500.33个；实体科技馆的数量在2019—2021年间呈小幅度增长趋势，3年间实体科技馆数量最多为2021年，科技馆数量为37.67座，2019年科技馆数量最少，仅为32座。

图5展示了2017—2021举办实用技术培训次数，显然，随着年份的推移，不同地区表现出不同规律。从西北地区看，举办实用技术培训次数在2017–2021年间呈波动发展趋势，5年间举办实用技术培训次数最多的是2021年，培训次数为2832.8次，举办实用技术培训次数最少的是2018年，培训次数为565.6次。从西南地区看，举办实用技术培训次数在2017—2021年间呈波动发展趋势，5年间举办实用技术培训次数最多的是2021年，培训次数为2780.4次，举办实用技术培训次数最少的是2018年，培训次数为792.6次。从中南地区看，举办实用技术培训次数在2017—2021年间呈波动趋势，其中2019—2020年间变化情况较为平稳，5年间举办实用技术培训次数最多的是2019年，培训次数为1180.33次，举办实用技术培训次数最少的是2018年，培训次数为533.67次。从华东地区看，举办实用技术培训

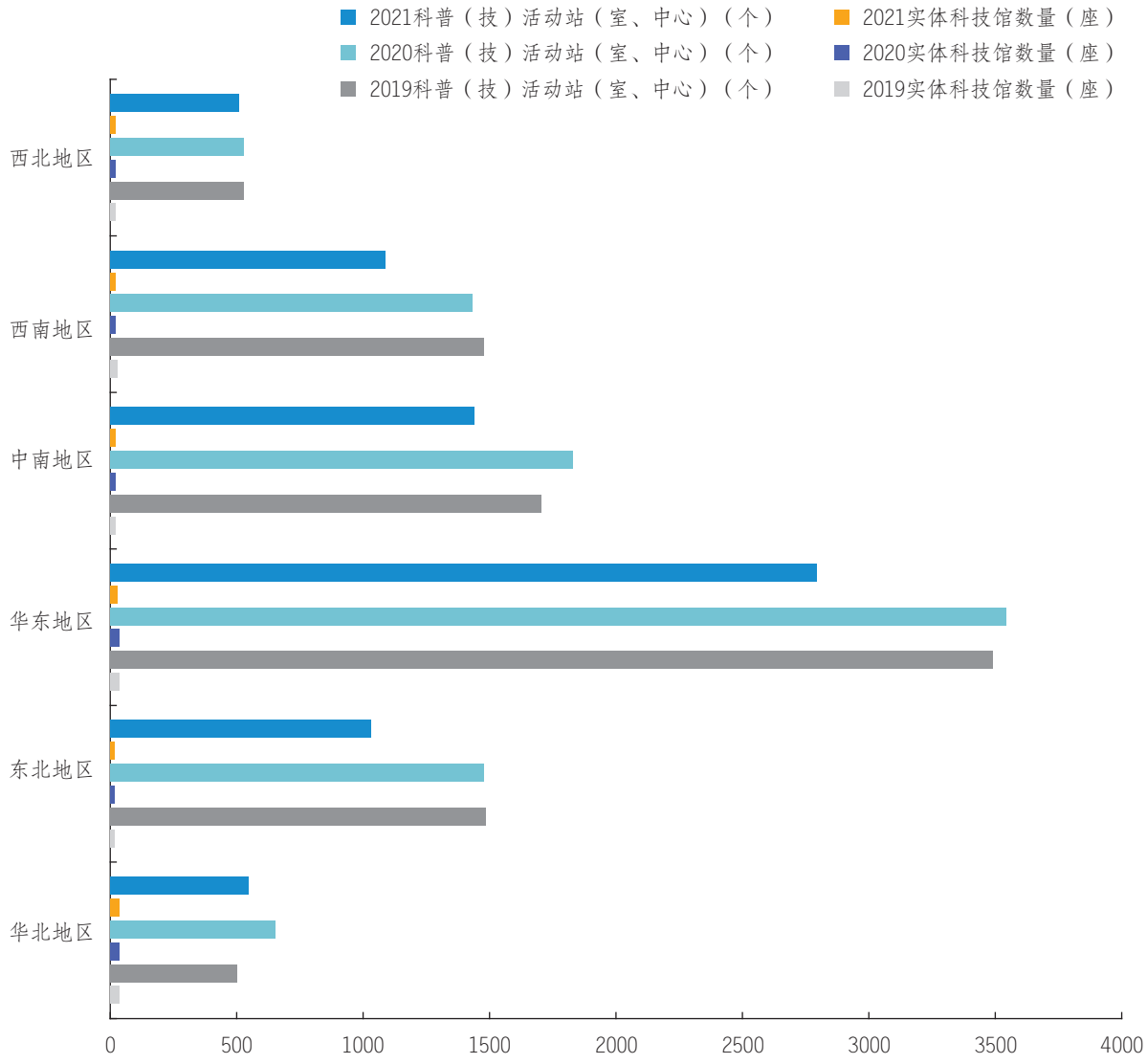


图4　2019—2021年实际科技馆、科普（技）活动站（室、中心）数量情况

注：数据来自《中国科学技术协会统计年鉴2020—2022》中县级科协口径数据。

次数在2017—2021年间呈波动发展趋势，5年间举办实用技术培训次数最多的是2021年，培训次数为1478.17次，举办实用技术培训次数最少的是2018年，培训次数为1004.67次。从东北地区看，举办实用技术培训次数在2017—2021年间呈波动发展趋势，5年间举办实用技术培训次数最多的是2017年，培训次数为1295.33次，举办实用技术培训次数最少的是2020年，培训次数为583.67次。从华北地区看，举办实用技术培训次数在2017—2021年间呈波动发展趋势，5年间举办实用技术培训次数最多的是2019年，培训次数为1109.67次，举办实用技术培训次数最少的是2018年，培训次数为502.67次。

图6分别展示出2017—2021推广新技术、新品种情况的项数。从西北地区看，推广新技术、新品种的项数在2017—2021年间呈波动发展趋势，5年间推广新技术、新品种的项数最多的是2019年，项数为495.8个，推广新技术、新品种的项数最少的是2018年，项数为140.8个。从西南地区看，推广新技术、新品种的项数在2017—2021年间呈先下降后迅速增长的趋势，5年间推广新技术、新品种的项数最多的是2021年，项数为997.6个，推广新技术、新品种的项数最少的是2018年，项数为201.4个。从中南地区看，推广新技

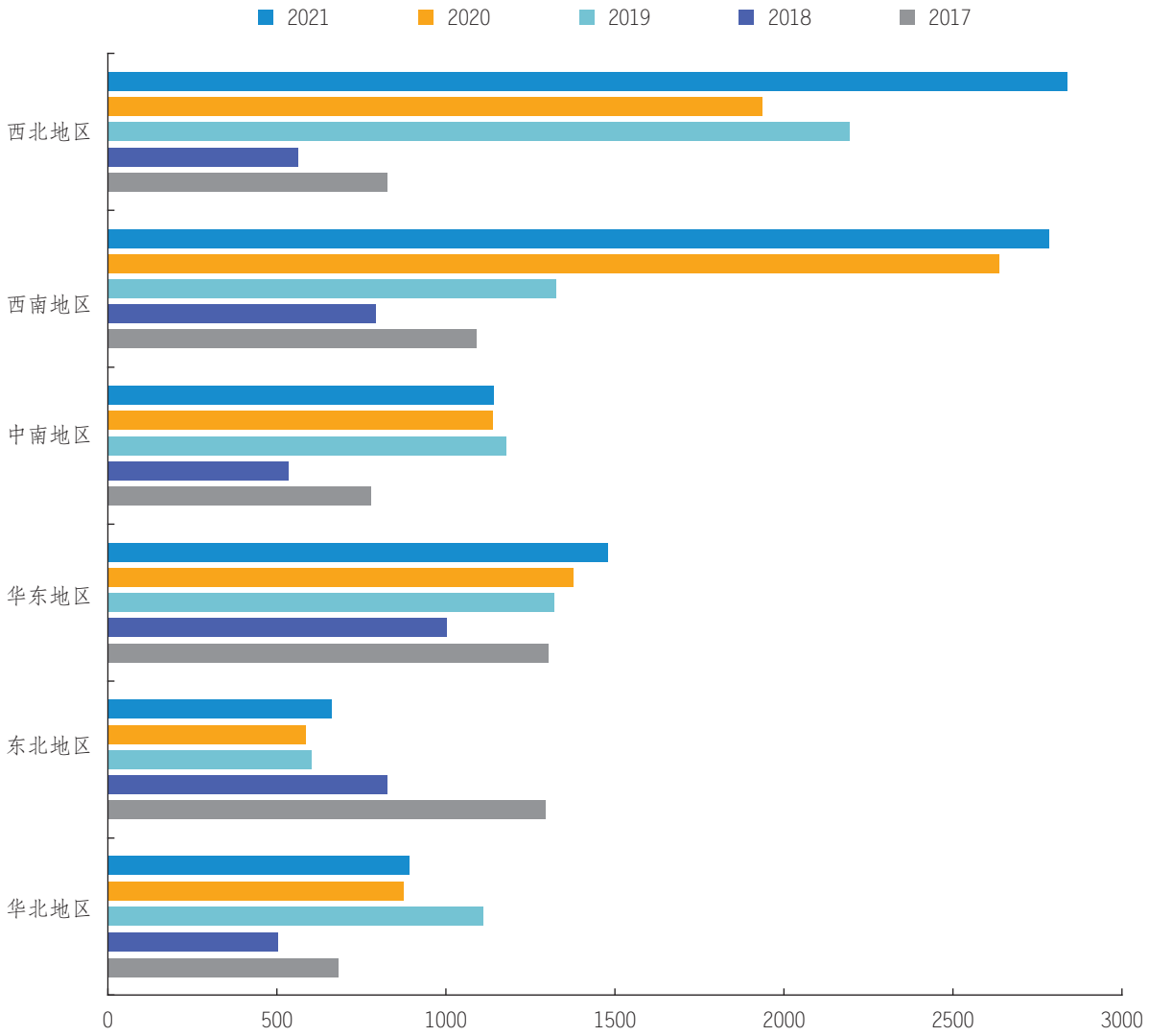


图5　2017—2021年举办实用技术培训情况（次）

注：数据来自《中国科学技术协会统计年鉴2018—2022》中县级科协口径数据，图6同。

术、新品种的项数在2017—2021年间呈先迅速下降后回升再波动的趋势，5年间推广新技术、新品种的项数最多的是2017年，项数为666.67个，推广新技术、新品种的项数最少的是2018年，项数为190.17个。从华东地区看，推广新技术、新品种的项数在2017—2021年间相对平稳，呈现出先下降后增长的发展趋势，5年间推广新技术、新品种的项数最多的是2017年，项数为359.33个，推广新技术、新品种的项数最少的是2018年，项数为267.5个。从东北地区看，推广新技术、新品种的项数在2017—2021年间呈先下降后增长的趋势，5年间推广新技术、新品种的项数最多的是2017年，项数为316个，推广新技术、新品种的项数最少的是2019年，项数为132.33个。从华北地区看，推广新技术、新品种的项数在2017—2021年间呈波动发展趋势，5年间推广新技术、新品种的项数最多的是2019年，项数为567个，推广新技术、新品种的项数最少的是2018年，项数为157.67个。

4. 全民科学素质服务状况指标降维分析

4.1 基于全民科学素质服务状况的主成分分析法

反映全民科学素质服务状况的指标存在数量众多的特点。对于这一问题，可以通过由Hotelling于1933年首先提出的主成分分析

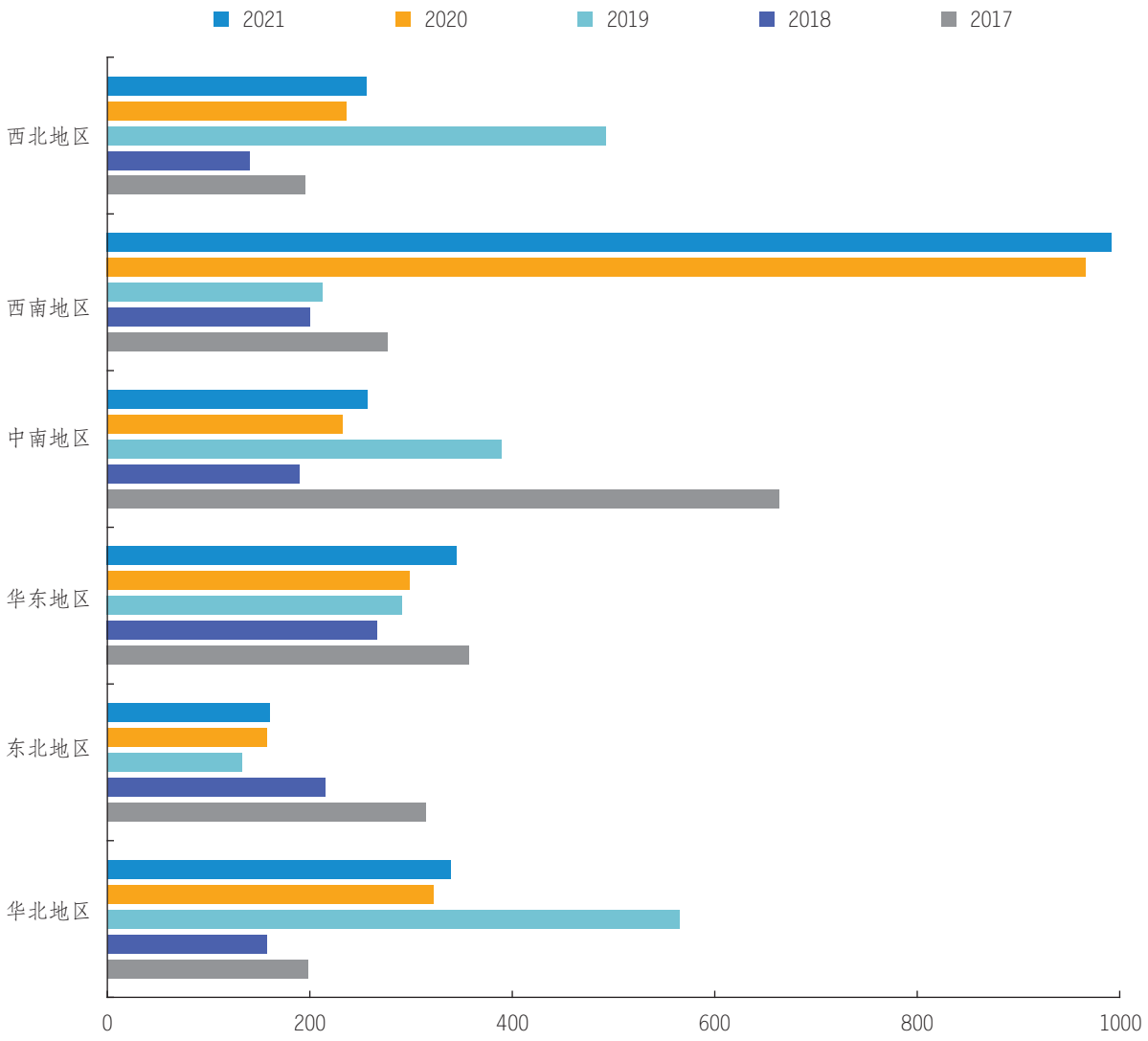


图6　2017—2021年推广新技术、新品种情况（个）

（Principal Component Analysis，PCA）方法，把多个指标压缩为少数几个综合指标（称为主成分），使得综合指标能够包含原来的多个变量的主要的信息[18-19]。另一方面，可以通过业务判断，先删除一部分表意重复的指标，再进行主成分分析。无论是否进行业务判断删除指标，主成分分析法都能够帮助实现将多个指标压缩为少数几个综合指标的降维目标。综上所述，主成分分析法的基本原理如下所示：

设多个随机向量集合为$X=(X_1,X_2,\cdots,X_p)$，其协方差矩阵为

$$\begin{bmatrix} \mathrm{cov}(X_1,X_1) & \mathrm{cov}(X_1,X_2) & \cdots & \mathrm{cov}(X_1,X_p) \\ \mathrm{cov}(X_2,X_1) & \mathrm{cov}(X_2,X_2) & \cdots & \mathrm{cov}(X_2,X_p) \\ \vdots & \vdots & & \vdots \\ \mathrm{cov}(X_p,X_1) & \mathrm{cov}(X_p,X_2) & \cdots & \mathrm{cov}(X_p,X_p) \end{bmatrix}$$

显然，协方差矩阵是一个$p \times p$的对称矩阵，主成分分析的目的是求解此矩阵的特征值与特征向量，设该协方差矩阵的特征值关系为$\lambda_1>\lambda_2>\cdots>\lambda_k$，相应的特征向量为：

$$\beta_1=\begin{pmatrix}\beta_{11}\\ \beta_{12}\\ \vdots\\ \beta_{1p}\end{pmatrix},\beta_2=\begin{pmatrix}\beta_{21}\\ \beta_{22}\\ \vdots\\ \beta_{2p}\end{pmatrix},\cdots,\beta_p=\begin{pmatrix}\beta_{p1}\\ \beta_{p2}\\ \vdots\\ \beta_{pp}\end{pmatrix}$$

那么该协方差矩阵为：

$$\begin{pmatrix} \mathrm{cov}(X_1,X_1) & \mathrm{cov}(X_1,X_2) & \cdots & \mathrm{cov}(X_1,X_p) \\ \mathrm{cov}(X_2,X_1) & \mathrm{cov}(X_2,X_2) & \cdots & \mathrm{cov}(X_2,X_p) \\ \vdots & \vdots & & \vdots \\ \mathrm{cov}(X_p,X_1) & \mathrm{cov}(X_p,X_2) & \cdots & \mathrm{cov}(X_p,X_p) \end{pmatrix}$$

$$\times(\beta_1,\beta_2,\cdots,\beta_p)=(\beta_1,\beta_2,\cdots,\beta_p)\times\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_1 & \cdots & 0\\ \vdots & \vdots & & 0\\ 0 & 0 & \cdots & \lambda_p\end{pmatrix}$$

经上述推导可得出主成分得分为：

$$\begin{cases} F_1=\beta_1\times X=\beta_{11}X_1+\beta_{12}X_2+\cdots+\beta_{1P}X_P \\ F_2=\beta_2\times X=\beta_{21}X_1+\beta_{22}X_2+\cdots+\beta_{2P}X_P \\ F_3=\beta_3\times X=\beta_{31}X_1+\beta_{32}X_2+\cdots+\beta_{3P}X_P \\ \vdots \\ F_P=\beta_P\times X=\beta_{P1}X_1+\beta_{P2}X_2+\cdots+\beta_{PP}X_P \end{cases}$$

由于特征向量$\beta_1,\beta_2,\cdots,\beta_p$两两正交，所以主成分得分之间无相关性。主成分分析之所以能够起到降维的效果，是因为在原则上单个主成分的解释变异(特征值)不应该小于1，并且选取少量主成分累积的解释变异(特征值)为80%～90%。主成分解释变异的能力可以以方差解释比例来计算（λ_k表示主成分k的特征值，分母表示所有特征值之和），即$\dfrac{\lambda_k}{\lambda_1+\lambda_2+\cdots+\lambda_p}$。

可以根据每一个主成分对应的主成分方程来解释主成分的含义，如对于下面的主成分方程：

$$F_1=\beta_{11}X_1+\beta_{12}X_2+\cdots+\beta_{1P}X_P \qquad \text{式（1）}$$

若规定$|\beta_{12}|$最大，则X_2所占权重最大，可以使用变量X_2的实际含义来解释该主成分[3]。

主成分个数的确定对于合理、正确运用主成分分析方法十分重要，主成分取得过少不能反映原始变量中的主要信息而取得过多有可能引入不必要的噪声和误差。一般情况下，确定主成分个数的方法主要有三种，分别为主成分累计方差贡献度超过85%；主成分特征值大于1.0；根据碎石图曲线选择主成分个数。本文采用方差累计贡献度大于85%的方法确定主成分个数。

4.2 全民科学素质服务状况指标降维应用

4.2.1 基于全指标的主成分降维分析结果 指标数据来自科协统计年鉴，经对部分存在问题的指标剔除后共保留42个业务指标变量和1个地区变量，32个省份样本，其中由于新疆建设兵团样本各指标数据均为0，因此将该样本剔除保留31个省份样本，43个指标。

首先对43个指标利用主成分分析法进行降维，43个指标共产生30项主成分，根据主成分分析的原理以累计贡献率大于85%的原则去提取主

成分，根据表4可知前13项主成分的方差贡献率累计达到了86.36%且大于85%，因此提取前13项主成分，可以利用所提取的13个新变量去替代原来的43个变量。

图7为主成分分析的特征矩阵热力图。由图可知，主成分1与大多数指标具有正向的关联关系，其中与科普大篷车数量（0.319）、科普大篷车展品数量（0.317）、科普大篷车覆盖人次（0.317）、流动科技馆数量（0.288）、全年流动科技馆巡展站点数（0.287）具有较强的正关联。

主成分2与主办科普App或设置科普栏目的综合类App（0.289）、科普App更新数（0.283）、全年流动科技馆巡展受众（0.243）、科普中国e站数（0.236）具有较强的正关联，与科普大篷车展品数量（–0.230）具有较强的负关联。

主成分3与数字科技馆及科技馆官方网站个数（0.385）、实体科技馆展厅面积（0.350）、实体科技馆建筑面积（0.336）、实体科技馆全年参观人次（0.288）、制作科普挂图数量（0.282）具有较强的正关联。

主成分4与科普活动受众（0.406）、科普微博关注数（0.382）、科普微信公众号关注数（0.316）、推广新技术、新品种（0.228）、科普微信公众号全年阅读量（0.219）具有较强的正关联。

主成分5与科普挂图总印数（0.389）、科普（技）活动站（室、中心）数量（0.341）、科普APP下载安装数（0.321）、科技图书总印数（0.241）具有较强的正关联，与流动科技馆数量（–0.247）具有较强的负关联。

主成分6与参加科普宣讲活动的学会、协会、研究会（0.274）、参加活动的科技人员、专家人次（0.260）具有较强的正关联，与科技图书总印数（–0.311）、地区（–0.288）、科普挂图总印数（–0.261）具有较强的负关联。

主成分7与科普画廊建筑面积（0.489）、举办科普宣讲活动次数（0.476）具有较强的正关联，与主办科技报纸总印数（–0.231）、主办科普微信公众号关注数（–0.225）、年度总阅读量（–0.262）具有较强的负关联。

表4　43个指标的相关系数矩阵的特征值及主成分的方差贡献率、累计方差贡献率

主成分	特征值	方差贡献率（%）	累计方差贡献率（%）
1	6.13	14.26	14.26
2	5.42	12.61	26.87
3	4.61	10.73	37.59
4	3.54	8.24	45.83
5	3.12	7.27	53.10
6	2.75	6.40	59.50
7	2.42	5.63	65.13
8	1.99	4.62	69.75
9	1.90	4.43	74.18
10	1.65	3.84	78.02
11	1.30	3.02	81.04
12	1.17	2.73	83.76
13	1.12	2.60	86.36

主成分8与实体科技馆数量（0.278）、科普资源总量（0.277）、主办科技报纸总印数（0.290）、主办科普微信公众号数量（0.244）具有较强的正关联，与科普宣讲活动覆盖村（社区）数量（–0.322）具有较强的负关联。

主成分9与地区（0.408）、科普资源总量（0.337）具有较强的正关联，全年参加活动（培训）人次（–0.277）、推广新技术、新品种（–0.259）、制作科普挂图种数（–0.270）具有较强的负关联。

主成分10与日均页面浏览量（0.396）具有较强的正关联，与科普（技）活动站（室、中心）数量（–0.238）、科普宣讲活动覆盖村（社区）数量（–0.367）、主办科技报纸总印数（–0.314）、主办科普微信公众号全年阅读量（–0.239）具有较强的负关联。

主成分11与科普大篷车行驶里程（0.253）、参加活动的科技人员、专家人次（0.257）、参加科普宣讲活动的学会、协会、研究会数量（0.408）、科普宣讲活动覆盖村（社区）数量（0.278）、主办科普微信公众号数量（0.258）具有较强的正关联。

主成分12与全年参加活动（培训）人次（0.298）、编著科技图书种数（0.267）具有较强的正关联，与实体科技馆建筑面积（–0.269）、实体科技馆展厅面积（–0.267）、主办科普微信公众号（–0.303）具有较强的负关联。

主成分13与参加科普宣讲活动的学会、协会、研究会个数（0.281）、举办实用技术培训次数（0.282）、主办科普微信公众号年度总阅读量（0.289）具有较强的正关联，与科普资源总量（–0.319）、主办科普微信公众号数量（–0.395）具有较强的负关联。

综上，将指标通过主成分分析法进行降维，前13个主成分作为最终的新变量代替原始43个变量。

4.2.2 基于业务删除后的主成分降维分析结果 对以上43个指标进行进一步分析，基于指标之间的相关性及业务判断综合考虑剔除部分指标，最终剩余33个指标，对以上指标再次利用主成分分析法进行降维处理，最终结果如下表5所示，前12项的方差贡献率累计达到了87.36%且大于85%，因

指标名称	主成分1	主成分2	主成分3	主成分4	主成分5	主成分6	主成分7	主成分8	主成分9	主成分10	主成分11	主成分12	主成分13
地区	-0.033	-0.159	-0.074	-0.142	0.020	-0.288	0.030	0.210	0.408	0.067	0.036	0.107	0.011
实体科技馆数量	0.160	0.088	0.144	0.076	0.088	-0.022	0.094	0.278	0.098	-0.202	-0.071	-0.143	0.187
实体科技馆建筑面积	0.121	0.082	0.336	-0.006	-0.035	0.093	0.026	-0.039	0.158	-0.190	-0.151	-0.269	0.019
实体科技馆展厅面积	0.134	0.097	0.350	-0.018	-0.084	0.073	0.020	0.030	0.180	-0.086	-0.120	-0.267	0.038
实体科技馆全年参观人次	0.059	0.141	0.288	0.081	-0.139	0.049	0.111	0.023	0.151	-0.139	0.085	-0.181	-0.117
数字科技馆及科技馆官方网站个数	0.117	0.024	0.385	-0.013	0.074	0.133	0.018	0.038	-0.166	-0.037	0.002	0.168	-0.009
日均页面浏览量	0.127	0.027	0.259	-0.078	0.041	-0.163	-0.154	0.109	-0.095	0.396	-0.109	-0.034	-0.020
科普资源总量	0.021	-0.118	0.009	-0.011	0.201	0.228	-0.010	0.277	0.337	-0.049	0.022	0.238	-0.319
流动科技馆数量	0.288	0.126	0.075	-0.008	-0.247	-0.065	-0.124	0.000	0.074	0.190	-0.008	0.094	-0.110
全年流动科技馆巡展站点数	0.287	0.205	-0.089	-0.174	-0.110	0.032	-0.043	0.007	-0.004	0.010	-0.019	-0.044	-0.037
全年流动科技馆巡展受众	0.277	0.243	-0.063	-0.118	-0.085	-0.056	-0.084	-0.029	0.107	-0.092	0.005	0.103	0.010
科普（技）活动站（室、中心）	0.057	0.088	0.000	-0.132	0.341	-0.225	-0.048	0.023	0.030	-0.238	0.174	-0.037	-0.258
全年参加活动（培训）人次	0.045	0.037	0.219	-0.151	0.098	-0.039	-0.008	0.243	-0.277	0.201	-0.159	0.298	-0.166
科普大篷车数量	0.319	-0.210	-0.077	0.034	-0.024	-0.037	0.001	-0.070	-0.168	-0.081	-0.016	0.007	0.038
科普大篷车下乡次数	0.265	-0.197	-0.097	-0.054	-0.134	-0.180	0.057	0.001	-0.033	0.122	0.139	-0.134	0.157
科普大篷车行驶里程	0.190	-0.133	-0.012	-0.094	-0.088	-0.235	-0.001	0.075	0.134	0.201	0.253	-0.184	0.067
科普大篷车覆盖人次	0.317	-0.194	-0.088	0.070	-0.041	0.007	-0.047	-0.082	-0.114	-0.096	-0.016	0.032	-0.011
科普大篷车展品数量	0.317	-0.230	-0.039	0.033	0.039	0.035	-0.064	-0.005	-0.028	-0.127	0.001	0.009	-0.089
科普中国e站	0.045	0.236	-0.197	-0.196	-0.049	0.128	-0.013	-0.126	0.086	0.077	-0.117	0.105	0.012
科普画廊建筑面积	-0.060	-0.071	0.005	-0.142	-0.167	-0.126	0.489	0.132	-0.091	-0.085	-0.083	0.122	0.013
科普画廊展示面积	0.115	0.140	-0.149	-0.128	0.038	0.242	0.039	0.236	0.072	0.072	0.228	0.006	0.246
举办科普宣讲活动次数	-0.010	0.009	0.022	-0.201	-0.171	-0.047	0.476	0.084	-0.026	-0.105	-0.080	0.249	0.161
科普活动受众	0.020	0.141	-0.010	0.406	0.039	-0.023	0.136	-0.097	0.091	0.218	0.062	0.108	-0.033
参加活动的科技人员、专家人次	0.019	0.210	-0.149	-0.105	-0.130	0.260	-0.078	0.041	-0.163	0.005	0.257	0.039	-0.238
参加科普宣讲活动的学会、协会、研究会	-0.065	-0.009	0.058	0.066	0.124	0.274	-0.063	-0.060	-0.017	0.098	0.408	-0.059	0.281
科普宣讲活动覆盖村（社区）	0.037	0.043	0.145	-0.078	0.094	0.013	0.117	-0.322	-0.001	-0.367	0.278	0.228	0.065
举办实用技术培训次数	0.166	-0.174	-0.096	0.072	0.207	0.182	0.020	0.209	0.071	0.063	0.021	-0.050	0.282
实用技术培训人次	0.186	-0.194	-0.068	0.093	0.194	0.243	-0.003	0.172	0.126	-0.072	-0.117	0.190	-0.210
推广新技术、新品种	0.246	-0.118	-0.090	0.228	0.045	0.105	0.052	-0.163	-0.259	-0.119	-0.202	0.014	0.011
编著科技图书种数	0.245	0.135	0.016	0.179	-0.112	-0.088	0.101	-0.194	0.133	0.066	0.187	0.267	-0.066
科技图书总印数	0.074	0.210	-0.096	-0.057	0.241	-0.311	-0.020	-0.129	0.143	-0.147	0.061	0.040	-0.031
主办科技报纸种数	0.031	0.071	-0.206	0.156	-0.152	-0.170	0.151	0.236	-0.198	-0.111	0.065	-0.105	-0.146
报纸总印数	-0.075	0.026	-0.079	0.103	-0.183	-0.110	-0.231	0.290	-0.082	-0.314	0.013	-0.009	0.076
制作科普挂图种数	0.015	0.109	0.282	0.092	0.069	-0.073	0.005	0.076	-0.270	0.073	0.244	0.211	0.131
科普挂图总印数	0.071	0.087	0.041	-0.106	0.389	-0.261	-0.049	0.115	-0.194	0.000	0.137	-0.014	0.089
主办科普App或设置科普栏目的综合类App个数	0.061	0.289	-0.091	-0.062	0.076	0.036	-0.158	0.129	-0.139	0.005	-0.235	-0.051	0.070
科普App下载安装数	-0.004	0.177	-0.154	0.146	0.321	-0.143	0.088	-0.091	-0.061	-0.046	-0.187	-0.165	0.007
科普App更新频次	0.079	0.283	-0.188	-0.191	-0.034	0.150	-0.016	-0.015	0.034	-0.013	-0.121	0.032	0.106
主办科普微信公众号个数	-0.006	0.102	-0.061	0.126	-0.081	0.062	0.181	0.244	-0.195	-0.014	0.258	-0.303	-0.395
关注数	-0.055	0.087	0.025	0.316	-0.210	-0.104	-0.225	0.121	0.126	-0.072	0.129	0.194	0.007
年度总阅读数	-0.058	0.071	-0.015	0.219	-0.135	-0.105	-0.262	0.223	-0.024	-0.239	-0.097	0.233	0.289
主办科普微博个数	0.086	0.194	-0.090	0.170	0.163	0.122	0.317	0.168	0.032	0.092	0.034	-0.026	0.185
粉丝数	0.041	0.163	-0.027	0.382	0.084	-0.096	0.170	-0.053	0.137	0.209	-0.172	0.043	-0.045

图7　基于43个指标的特征矩阵热力图

此选取前12项主成分代表原始33个指标，该12项主成分中已包含原始指标的大部分信息。

图8为主成分分析的特征矩阵热力图。由图可知，主成分1与大多数指标具有负向的关联关系，其中与科普大篷车数量（–0.372）、科普大篷车覆盖人次（–0.369）、科普大篷车展品数量（–0.369）、科普大篷车下乡次数（–0.314）、推广新技术、新品种（–0.288）具有较强的负关联。

主成分2与全年流动科技馆巡展受众人次（0.293）、实体科技馆全年参观人次（0.272）、流动科技馆数量（0.265）、全年流动科技馆巡展站点数（0.265）、主办科普App或设置科普栏目的综合类APP（0.251）具有较强的正关联。

主成分3与数字科技馆及科技馆官方网站个数（0.407）、实体科技馆建筑面积（0.283）具有较强的正关联，与科普中国e站（–0.315）、主办科技报纸种数（–0.286）、参加科普宣讲活动的科技人员、专家人次（–0.282）具有较强的负关联。

主成分4与科普画廊建筑面积（0.338）、科普宣讲活动次数（0.327）、地区（0.307）、科普大篷车行驶里程（0.286）具有较强的正关联，与科普宣讲活动受众人数（–0.268）具有较强的负关联。

主成分5与科普画廊建筑面积（0.395）、举办科普宣讲活动次数（0.321）、主办科技报纸种数（0.316）、主办科普微信公众号数量（0.291）、科普宣讲活动受众人数（0.278）具有较强的正关联。

主成分6与数字科技馆科普资源总量（0.469）、实用技术培训人次（0.314）、主办科普微博个数（0.294）、举办科普宣讲活动次数（0.293）、实体科技馆数量（0.263）具有较强的正关联。

主成分7与举办科普宣讲活动次数（0.272）、科普宣讲活动覆盖村（社区）数量（0.506）具有较强的正关联，与数字科技馆及科技馆官方网站日均页面浏览量（–0.369）、科普活动受众人数（–0.241）、科普大篷车行驶里程（–0.231）具有较强的负关联。

表5　33个指标的相关系数矩阵的特征值及主成分的方差贡献率、累计方差贡献率

主成分	特征值	方差贡献率（%）	累计方差贡献率（%）
1	5.92	17.94	17.94
2	4.42	13.40	31.34
3	3.55	10.76	42.09
4	2.82	8.55	50.65
5	2.34	7.11	57.76
6	1.97	5.98	63.74
7	1.72	5.22	68.96
8	1.60	4.85	73.82
9	1.25	3.80	77.61
10	1.13	3.42	81.03
11	1.12	3.40	84.43
12	0.97	2.93	87.36

主成分8与科普（技）活动站（室、中心）全年参加活动（培训）人次（0.416）、参加科普宣讲活动的科技人员、专家人次（0.241）具有较强的正关联，与地区（–0.340）、编著图书种数（–0.311）、科普宣讲活动受众人数（–0.267）具有较强的负关联。

主成分9与参加科普宣讲活动的学会、协会、研究会数量（0.334）、科普中国e站数量（0.266）、数字科技馆资源总量（0.229）具有较强的正关联，与科普（技）活动站（室、中心）数量（–0.598）、实体科技馆数量（–0.220）具有较强的负关联。

主成分10与参加科普宣讲活动的学会、协会、研究会数量（–0.422）、科普宣讲活动覆盖村（社区）数量（–0.395）、科普大篷车行驶里程（–0.365）、科普（技）活动站（室、中心）数量（–0.344）、制作科普挂图种数（–0.289）具有较强的负关联。

主成分11与主办科普微信公众号个数（0.450）、全年参观实体科技馆人次（0.344）、实体科技馆建筑面积（0.284）具有较强的正关联，与科普活动受众人数（–0.281）、主办科普微博个数（–0.252）具有较强的负关联。

主成分12与参加科普宣讲活动的学会、协会、研究会数量（0.406）、实体科技馆数量（0.363）、主办科普微博个数（0.258）具有较强的正关联，与科普资源总量（–0.340）、编著科技图书种数（–0.292）具有较强的负关联。

综上，将以上指标利用主成分分析法进行降维，前12个主成分作为最终的新变量代替原始33个变量。

5. 总结与讨论

全民科学素质是体现一个国家国民素质的重要方面，也是社会文明进步的主要基础。在“十四五”开局起步之际，国务院印发《全民科学素质行动规划纲要（2021—2035年）》。本文通过对《中国科学技术协会统计年鉴2022》中有关提高全民科学素质服务指标数据的统计分析与降维讨论，旨在为相关部门进行指标修订和删减提供参考与借鉴。

但是，在全民科学素质指标的深入讨论中发现，在切实反映相关部门业务需求的基础上，以数据为对象，以统计方法为工具，开展一些应用

指标名称	主成分1	主成分2	主成分3	主成分4	主成分5	主成分6	主成分7	主成分8	主成分9	主成分10	主成分11	主成分12
地区	-0.001	-0.224	-0.007	0.307	-0.085	0.229	-0.217	-0.340	-0.040	-0.008	-0.068	-0.053
实体科技馆数量	-0.127	0.165	0.148	-0.053	0.131	0.263	-0.142	-0.148	-0.220	0.148	0.109	0.363
实体科技馆建筑面积	-0.071	0.230	0.283	0.025	0.014	0.040	0.163	-0.152	0.096	0.229	0.284	0.151
实体科技馆全年参观人次	-0.008	0.272	0.187	0.013	0.190	0.004	0.038	-0.232	0.093	0.234	0.344	-0.004
数字科技馆及科技馆官方网站个数	-0.079	0.220	0.407	-0.034	0.029	0.057	0.103	0.189	0.077	-0.040	-0.004	-0.043
日均页面浏览量	-0.097	0.168	0.243	0.176	-0.178	-0.077	-0.369	0.155	0.040	-0.034	-0.164	-0.022
科普资源总量	-0.037	-0.140	0.141	-0.174	-0.140	0.469	-0.037	-0.134	0.229	-0.073	0.088	-0.340
流动科技馆数量	-0.257	0.265	-0.055	0.139	-0.086	-0.110	-0.133	-0.070	0.192	0.084	0.002	-0.182
全年流动科技馆巡展站点数	-0.233	0.265	-0.224	0.084	-0.156	0.108	0.066	0.028	0.033	0.011	0.037	0.088
全年流动科技馆巡展受众	-0.219	0.293	-0.194	0.078	-0.153	0.089	0.081	-0.144	-0.049	0.020	-0.038	0.033
科普（技）活动站（室、中心）	-0.023	0.062	0.023	0.030	-0.194	0.235	0.039	-0.120	-0.598	-0.344	0.094	-0.018
全年参加活动（培训）人次	-0.012	0.152	0.237	0.137	-0.087	0.172	-0.148	0.416	-0.083	0.067	-0.186	-0.217
科普大篷车数量	-0.372	-0.129	0.006	0.017	0.057	-0.092	0.099	0.140	-0.080	-0.002	-0.003	0.048
科普大篷车下乡次数	-0.314	-0.142	-0.059	0.238	0.070	-0.103	-0.078	0.026	0.119	-0.175	0.032	0.180
科普大篷车行驶里程	-0.216	-0.080	0.010	0.286	-0.032	-0.026	-0.231	-0.124	0.196	-0.365	0.133	0.138
科普大篷车覆盖人次	-0.369	-0.119	-0.011	-0.039	0.030	-0.109	0.104	0.085	-0.070	0.067	0.037	-0.015
科普大篷车展品数量	-0.369	-0.148	0.076	-0.047	-0.035	0.016	0.091	0.050	-0.033	0.013	0.120	-0.050
科普中国e站	0.008	0.169	-0.315	-0.012	-0.188	0.115	0.152	-0.027	0.266	0.120	-0.192	0.006
科普画廊建筑面积	0.051	-0.064	-0.028	0.338	0.395	0.244	0.144	0.161	0.058	0.007	-0.091	-0.022
举办科普宣讲活动次数	0.027	0.034	-0.042	0.327	0.321	0.293	0.272	0.102	0.121	0.000	-0.224	0.069
科普活动受众	-0.005	0.120	-0.052	-0.268	0.278	-0.164	-0.241	-0.267	-0.066	-0.008	-0.281	-0.183
参加活动的科技人员、专家人次	0.023	0.188	-0.282	-0.124	-0.112	0.073	0.122	0.241	0.212	-0.238	0.216	-0.161
参加科普宣讲活动的学会、协会、研究会	0.070	-0.013	0.120	-0.268	-0.032	-0.047	0.034	0.038	0.334	-0.422	0.026	0.406
科普宣讲活动覆盖村（社区）	-0.007	0.110	0.143	0.010	0.048	-0.022	0.506	-0.186	-0.167	-0.395	-0.087	-0.105
举办实用技术培训次数	-0.198	-0.195	0.052	-0.219	-0.015	0.196	-0.093	-0.008	0.071	0.031	-0.179	0.200
实用技术培训人次	-0.228	-0.187	0.092	-0.262	-0.060	0.314	0.016	0.019	0.086	0.041	-0.030	-0.248
推广新技术、新品种	-0.288	-0.073	-0.015	-0.239	0.165	-0.146	0.148	0.190	-0.149	0.133	-0.097	0.005
编著科技图书种数	-0.220	0.215	-0.068	-0.022	0.165	-0.139	0.011	-0.311	0.055	-0.149	-0.220	-0.292
主办科技报纸种数	-0.047	0.007	-0.286	0.022	0.316	0.044	-0.202	0.142	-0.194	-0.072	0.214	-0.090
制作科普挂图种数	0.022	0.233	0.271	-0.027	0.141	-0.085	-0.142	0.141	-0.009	-0.289	-0.163	-0.047
主办科普APP或设置科普栏目的综合类APP个数	0.001	0.251	-0.197	-0.092	-0.223	0.140	-0.083	0.200	-0.212	0.023	-0.100	0.216
主办科普微信公众号个数	0.018	0.097	-0.137	-0.121	0.291	0.113	-0.220	0.162	0.017	-0.157	0.450	-0.168
主办科普微博个数	-0.041	0.142	-0.116	-0.255	0.276	0.294	-0.146	-0.078	0.068	-0.044	-0.252	0.258

图8　基于33个指标的特征矩阵热力图

分析与讨论是对指标体系不断完善和优化的重要佐证。在业务工作持续高质量推进和统计数据分析深入加持的双向奔赴过程中，努力汇聚提升全民科学素质的强大合力，建立起宏大的高素质创新大军，推动实现科技成果快速转化。

责任编辑：李琦　校对：李琦 武琪

致谢

感谢中国科协战略发展部委托“综合统计调查数据分析方法及应用研究”项目，提供统计分析与研究平台。感谢北京久其软件股份有限公司团队（实施工程师陈毅、数据分析师张亚楠和刘倩）在数据提取与分析方面投入了大量的时间与精力。

参考文献

[1] 本报评论员. 提高全民科学素质服务高质量发展[N].人民日报, 2021–06–26(5).

[2] 新华社. 全民科学素质行动规划纲要(2021–2035年)[J].学会, 2021(07): 5–13.

[3] 李红林, 胡俊平, 曹金, 等. 数字化转型背景下公民科学素质提升的探讨与思考[J].科普研究, 2023, 18(04): 18–25+105–107.

[4] 王挺.夯实中华民族伟大复兴的科学根基——全面落实《科学素质纲要(2021–2035年)》的思考[J]. 科普研究,2021,16(4):5–13.

[5] 李正风,武晨箫,胡赛全. 关于新时代公民科学素质的再思考[J].科普研究,2021,16(2):18–23.

[6] 任定成. 新起点、新目标与新举措——《科学素质纲要(2021–2035年)》解读[J].科普研究,2021,16(4):18–24.

[7] 王挺,付文婷. 拓展科学普及新时代内涵 助力推进中国式现代化——科普中国智库专访武向平院士[J]. 科普研究,2023,18(2):5–8.

[8] 袁汝兵. 公民科学素质的现状、问题及其提升路径——以北京为例[J].科技智囊,2022(7):62–67.

[9] 高宏斌.《科学素质纲要(2021–2035年)》前言、指导思想和原则的解读[J].科普研究,2021,16(4):37–43.

[10] 任磊,郭凤林,黄乐乐. 我国公民网络科技信息辨识的行为模式与特征分析——基于2020年中国公民科学素质调查的分析[J].科普研究,2022,17(1):66–73.

[11] 王丽慧. 老年人科学素质提升行动的思考[J].科普研究,2021,16(4):69–73,86.

[12] 张增一."新"纲要新在何处?[J].科普研究,2021,16(4):25–30.

[13] 孟庆海. 新时代如何提升领导干部科学素质[J].中国党政干部论坛,2023(3):72–76.

[14] 全民科学素质行动规划纲要(2021–2035年)[M]. 北京:人民出版社,2021.

[15] 中国科学技术协会. 中国科学技术协会统计年鉴2022[M].北京:中国科学技术出版社, 2023.

[16] 中国科学技术协会组织人事部. 县级科协工作手册[M]. 北京:中国科学技术出版社, 2013.

[17] 程豪. 互联网统计学:方法与应用[M].北京:电子工业出版社, 2023.

[18] 贾俊平,何晓群,金勇进. 统计学[M]. 北京:中国人民大学出版社,2021.

[19] 何晓群. 多元统计分析[M]. 北京:中国人民大学出版社,2019.

Investigating statistical index dimension reduction methods and applications to enhance national scientific quality

Cheng Hao

(National Academy of Innovation Strategy CAST, Beijing 100038, China)

Abstract: Scientific literacy is an important component of national quality and the foundation of social civilization and progress. The starting point for this article is from the indicator data in the "Statistical Yearbook of the China Association for Science and Technology 2022" on improving the scientific literacy services forthe entire population, and uses statistical description methods to depict the basic situation of science popularization infrastructure construction, science popularization propaganda activities, and science popularization dissemination. On this basis, the article focus on the development of grassroots services for the scientific literacy of the whole people, and further reflect the basic laws and characteristics of grassroots science popularization. To address the issue of a large number of popular science indicators, this article utilizes the traditional principal component analysis method to conduct principal component dimensionality reduction analysis on the entire indicator system and the indicator system after deleting some indicators through business judgment, in order to provide reference and reference for relevant departments to revise and delete indicators.

Keywords: national scientific literacy; grassroots organizations; principal component analysis

（上接第43页）

Research and recommendations on the benefits and drawbacks of constructing a worldwide highly influential advanced photovoltaic industry cluster in Anhui

Zhu Huangwu[1], Xu Hongyao[2], Wu Bangyong[3]

(1. Anhui Earthquake Agency, Hefei 230031, China;
2. Administration for Culture, Radio, Television, News and Publication of Chuzhou City, Chuzhou 209000, China;
3. Economic and Information Bureau of Chuzhou City, Chuzhou 209000, China)

Abstract: Building a globally influential advanced photovoltaic industry cluster to promote high-quality development of the industrial economy is a major task of Anhui.Under the analysis of strengths, weaknesses, opportunities and threats for Anhui building a globally influential advanced photovoltaic industry cluster, the photovoltaic industry's development status at both home and abroad, as well as in Anhui province, is summarized in this paper. On this basis, it presents the goal of seizing the high ground of advanced photovoltaic industry development, and the policy measures of leading and providing motivation through science and technology innovation, cultivating backbone enterprises in the industrial chain through "talent and investment attraction" plan, and strengthening measures through optimizing industrial ecology.

Keywords: photovoltaic; impact; industry cluster; Anhui province