

# 全国数据资源调查报告(2024年)

NATIONAL DATA RESOURCE SURVEY REPORT



全国数据资源统计调查工作组

## 版权声明

本报告版权属于全国数据资源统计调查工作组，并受法律保护。转载、摘编或利用其他方式使用本报告文字或者观点的，应注明“来源：《全国数据资源调查报告（2024年）》”。违反上述声明者，将追究其相关法律责任。

# 专家委员会

(排名不分先后)

## 顾问

邬贺铨 江小涓 梅宏 刘韵洁 尹浩 张平文 李骏 高新民

## 委员

蒋艳 姜奇平 廖凯 汪玉凯 李晓东 吴志刚 冯海红 黄罡  
蔡跃洲 欧阳日辉 王晨 黄丽华 赵刚 孙毅 胡良霖 林伟  
郑爱军

## 工作组

国家工业信息安全发展研究中心

复旦大学上海数学中心

# 前言

## INTRODUCTION

习近平总书记指出，我国数据资源丰富，产业体系完备，应用场景广阔，市场空间巨大，要深化数据资源开发利用和开放共享。随着数字中国建设深入推进，数据领域新技术、新模式、新业态不断涌现，数据生产、存储、计算、流通和应用等各环节发生新变化，数据要素对经济社会发展的乘数效应进一步释放，我国海量数据资源和丰富应用场景优势逐渐转化为经济增长新动能、国家竞争新优势。

贯彻落实《数字中国建设整体布局规划》工作部署，2024年国家数据局联合相关部门首次开展全国数据资源调查，对数据资源进行摸底。2025年2月，国家数据局联合相关部门组织开展2024年度全国数据资源统计调查工作，有效样本数量1.44万个，覆盖除国际组织外的所有国民经济行业门类。本年度调查在上年度数据规模、开发利用和应用情况等指标基础上，新增人工智能（AI）、大模型、低空经济等热点领域调查指标，并对数据资源地区、行业分布情况做了分析测算，形成了《全国数据资源调查报告（2024年）》。

报告显示，我国数据资源规模优势持续扩大，数据资源开发利用活跃度稳步提升，各类主体加快人工智能布局投入，数据要素市场化、价值化进程进一步提速。

**一是数据资源规模加速增长。**2024年，全国年度数据生产总量<sup>1</sup>达41.06泽字节（ZB），同比增长25%，增速提高2.56个百分点。人均年度数据生产量约为31.31太字节（TB），同比增长25.17%，年度数据生产总量和人均产量实现同步跃升，智能家居、智能网联汽车、无人机等智能设备数据增速位居前列。

**二是数据存储结构持续优化。**2024年，全国数据存储总量<sup>2</sup>为2.09泽字节（ZB），同比增长20.81%。存储空间利用率<sup>3</sup>达61%，同比提高2个百分点。结构化数据同比增长36%，占存储总量比重为18.70%，大幅超过非结构化数据存储增速，表明数据资源开发利用的程度明显提高。

<sup>1</sup> 全国数据生产总量指2024年度全国数据创建、复制、捕获和消费产生的数据总量。

<sup>2</sup> 全国数据存储总量指全国累计存储的数据量，包括算力中心存储量，以及政府、企事业单位、个人终端的数据存储量。

<sup>3</sup> 存储空间利用率=累计数据存储总量/累计数据存储总空间。

**三是算力基础设施加快布局。**2024年，全国算力总规模<sup>4</sup>达280EFLOPS（每秒百亿亿次浮点运算），八大国家枢纽节点算力总规模达到175EFLOPS。智能算力规模达90EFLOPS，在算力总规模中占比提升至32%，为海量数据计算提供智能底座。

**四是公共数据资源开发利用步伐加快。**政务数据共享持续深化，共享枢纽平台累计支撑调用超5400亿次，有力支撑“高效办成一件事”改革。2024年，全国地市级以上地方公共数据开放平台数量增长7.5%，开放的数据量增长7.1%<sup>5</sup>。国家公共数据资源开发利用“1+3”政策文件发布后，各省（自治区、直辖市）、计划单列市中，超六成启动授权运营工作。

**五是企业数据资源开发利用提速。**2024年，66.1%的行业龙头企业购买过数据，企业数据与公共数据融合应用场景加速增长。活跃数据<sup>6</sup>总量同比提升22.73%，活跃数据总量占存储数据总量为62.04%。企业数据要素在企业内部关键环节，如技术研发、生产协同、产品营销、售后服务等环节中发挥了重要作用，产业链供应链韧性不断提升。

**六是数据促进人工智能加速发展。**2024年，新增238款人工智能服务在国家网信办完成备案。开发或应用人工智能的企业数量同比增长36%，高质量数据集数量同比增长27.4%，有力支撑人工智能训练和应用。利用大模型的数据技术企业同比增长57.21%，数据应用企业同比增长37.14%。

报告不足之处在所难免，敬请各位专家和广大读者指正。

---

<sup>4</sup> 数据来源于数字中国建设峰会官方网站。

<sup>5</sup> 数据来源于复旦大学数字与移动治理实验室。

<sup>6</sup> 活跃数据指一年中至少被使用1次的数据。

# 目录

## CONTENTS

<b>一、重点指标</b>	01
(一) 数据资源规模优势持续扩大	02
(二) 数据资源开发利用活跃度稳步提升	02
(三) 数据要素价值潜力持续释放	02
<b>二、调查分析</b>	03
(一) 数据生产总量增速加快	04
(二) 数据存储结构和空间持续优化	05
(三) 智能算力加速布局	05
(四) 公共数据资源开发利用步伐加快	06
(五) 企业用数活力逐步提升	07

(六) 人工智能与数据资源开发利用互相促进 .....	07
(七) 数据资源区域和行业分布不均衡 .....	07
<b>三、发展趋势</b> .....	<b>08</b>
(一) 数据生产总量将保持增长态势 .....	09
(二) 数据资源流通利用进入加速期 .....	09
(三) 数据融合应用加快价值化进程 .....	09
<b>附 录</b> .....	<b>10</b>
一、数据来源 .....	10
二、样本分布 .....	11
三、统计方法 .....	12

# 一、重点指标

## （一）数据资源规模优势持续扩大

我国数据生产、存储总量增长再提速，数据存储空间利用率提高，智能算力加快布局引领算力规模增长。

- 2024年，全国数据生产总量达41.06泽字节（ZB），同比增长25%。
- 2024年，全国数据存储总量为2.09泽字节（ZB），同比增长20.81%。存储空间利用率为61%，同比增长2个百分点。
- 2024年，全国算力总规模达280EFLOPS（每秒百亿亿次浮点运算），八大国家枢纽节点算力总规模达175EFLOPS。智能算力规模达90EFLOPS，在算力总规模中占比提升至32%。

## （二）数据资源开发利用活跃度稳步提升

公共数据资源供给和企业用数需求双向拉动，企业加快应用大模型，高质量数据集建设进入加速期。

- 2024年，全国地市级以上公共数据开放平台数量增长7.5%，开放的数据量增长7.1%。超六成省（自治区、直辖市）、计划单列市启动公共数据授权运营工作。
- 2024年，约66%的行业龙头企业和30%的数据技术企业购买过数据。
- 2024年，利用大模型的数据技术企业和数据应用企业数量分别增长57.21%、37.14%。
- 2024年，企业高质量数据集建设增速超27%。

## （三）数据要素价值潜力持续释放

数据“产-存-算-流-用”各环节转换过程中，数据要素价值持续释放，数据产业发展潜力巨大。

- 2024年，活跃数据总量为1.3泽字节（ZB），同比提升22.73%，活跃数据总量占存储数据总量为62.04%。
- 2024年，利用数据获得直接收益的企业实现小幅提升。

## 二、调查分析

## （一）数据生产总量增速加快

数据生产总量呈高速增长态势。2024年，全国数据生产总量达41.06泽字节（ZB），同比增长25%，增速较去年提高2.56个百分点（图2-1）。人均数据生产量约为31.31太字节（TB），同比增长25.17%，数据生产总量和人均产量实现同步跃升。

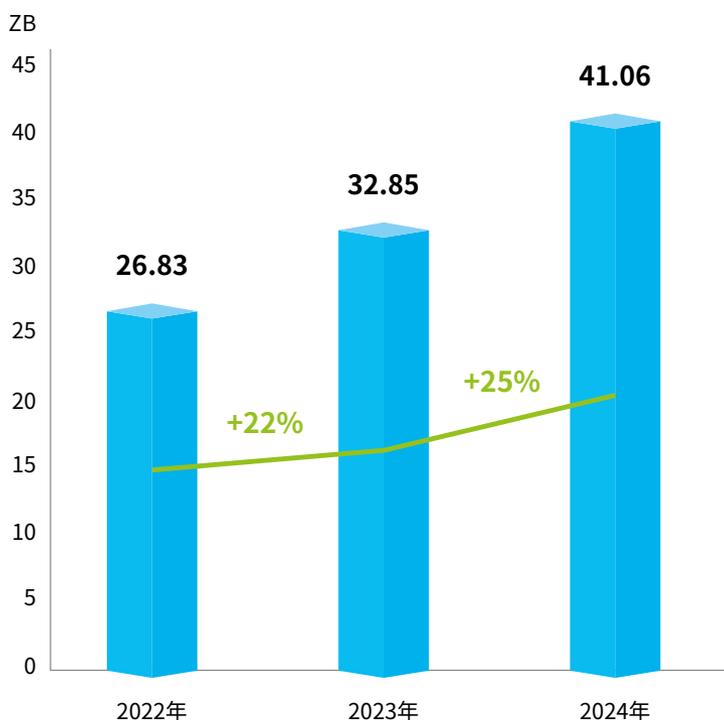


图2-1 数据生产总量（2022年—2024年）

数据来源：2024年度全国数据资源统计调查

**智能应用推动数据资源规模加速增长。**随着人工智能快速发展，用于开发、训练和推理的数据量同比增长40.95%。智能家居、智能网联汽车等智能设备数据增速位居前列，分别为51.43%、29.28%。

**战略性新兴产业激发数据新增长点。**无人机应用范围快速拓展，机器人产业化进程不断加快，带动低空经济和机器人数据生产量增速超过30%。

## （二）数据存储结构和空间持续优化

**数据存储结构持续优化。**2024年，全国数据存储总量为2.09泽字节（ZB），同比增长20.81%，其中新增数据存储总量1.15泽字节（ZB）。数据存储空间利用率为61%，同比提升2个百分点。结构化数据同比增长36%，占存储总量比重为18.70%，大幅超过非结构化数据存储增速，表明数据资源开发利用的程度明显提高。

**各行业数据上云差异明显。**2024年，全国数据云存储率为30.39%。政务、气象、金融数据云存储占比分别为47.39%、42%、26.38%，数据弹性灵活调用更敏捷。科学数据仍然以本地化存储为主，云存储占比约为5%。

## （三）智能算力加速布局

**智能算力规模大幅提升。**2024年，全国算力总规模达到280EFLOPS，其中智能算力规模达90EFLOPS，占比提升至32%（图2-2），为海量数据计算提供智能底座。

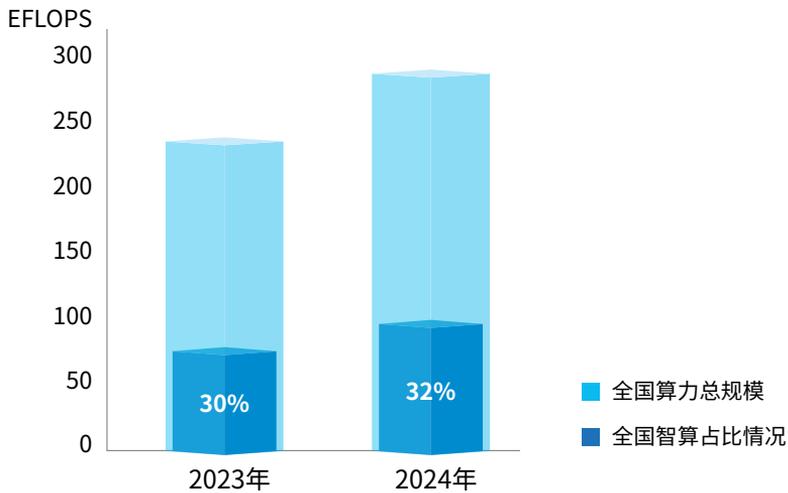


图2-2 2023—2024年算力中心算力规模及智能算力占比

数据来源：2024年度全国数据资源统计调查

**企业加快布局智能算力。**随着人工智能应用推广，市场对智能算力的需求呈爆发式增长，企业布局智能算力提速。其中，中央企业<sup>7</sup>算力规模增长近3倍，智能算力占比为40.22%；数据技术企业算力规模同比增长近1倍，智能算力占比为43.63%。

<sup>7</sup> 中央企业有效样本数量为101家。

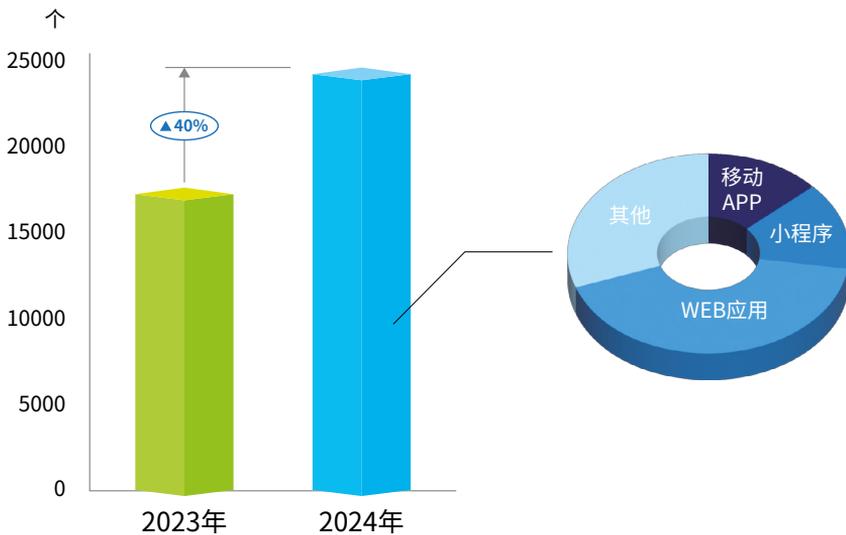
#### (四) 公共数据资源开发利用步伐加快

**政务数据共享持续深化。**建成全国一体化政务数据共享枢纽，累计支撑调用超5400亿次，推动国务院部门201个垂管系统向基层“回流”数据8.2亿条，为经济社会发展、数字政府建设、“高效办成一件事”改革提供了有力的数据支撑。

**公共数据集开放总量稳步增长。**公共数据开放程度逐渐提升，2024年全国地市级以上公共数据开放平台数量增长7.5%，开放的数据量增长7.1%。

**公共数据授权运营探索步伐加快。**国家公共数据资源开发利用“1+3”政策文件发布后，超六成省（自治区、直辖市）、计划单列市启动授权运营工作，公共数据授权运营逐步规范化、有序化。其中，市级政府部门公共数据授权运营数据量为省级政府部门的2.53倍，公共数据资源供给不断扩大。

**公共数据服务场景日益丰富。**公共数据服务产品数量同比增长超过40%（图2-3），涵盖交通、环保、医疗等经济社会各领域，应用在低空经济、多式联运、保险理赔、征信服务等场景。



**图2-3 公共数据服务产品数量变化及分布**

数据来源：2024年度全国数据资源统计调查

## （五）企业用数活力逐步提升

**企业数据应用需求逐步释放。**2024年，存储数据中活跃数据占比为62.04%，数据资源开发利用活跃度逐步提升。企业用数需求旺盛，约66%的行业龙头企业 and 30%的数据技术企业购买过数据，企业数据与公共数据融合应用场景更加丰富。

**平台企业成为数据流通的关键枢纽。**平台企业发挥生态优势，促进气象、交通、地理等公共数据加速与企业数据融合，在客户服务、精准营销、研发创新等场景得到广泛应用。

## （六）人工智能与数据资源开发利用互相促进

**大模型应用促进企业数据资源开发利用。**企业加快人工智能投入和大模型应用，利用大模型的企业数量同比增长超过37%，比例提升至约10%。大模型有望帮助企业降低非结构化数据加工成本，推动数据资源开发利用提档加速。

**人工智能发展带动高质量数据集需求激增。**数据技术企业积极推动数据集建设，高质量数据集数量增速达27.4%。人工智能加快向垂直领域发展，企业用于人工智能的数据量占数据存储量约为7%，企业数据集建设潜力空间巨大。

## （七）数据资源区域和行业分布不均衡

**区域数据资源水平和经济发展高度相关。**经济大省数据资源集聚效应明显，广东、江苏、山东、浙江、河南和四川6个省份数据生产量占全国比重达57.45%。东部地区数据资源开发利用活跃，北京、上海企业人工智能投入平均水平位居全国前二，浙江应用大模型的企业比例位居第一，江苏、广东企业平均数字化投入是全国平均水平的1.71倍。

**行业数据资源开发利用水平与数字化投入密切相关。**随着我国数字化、网络化、智能化进程加速，制造、金融、交通物流行业数据生产总量领先其他行业，金融、采矿、交通物流行业企业平均数据生产量位居前三。交通物流、电力行业数据资源开发利用活跃度远高于全国平均水平，分别为78.24%、75.83%。

## 三、发展趋势

### （一）数据生产总量将保持增长态势

预计2025年，全国数据生产总量有望突破50泽字节（ZB）。AI计算机的普及应用与AI服务器性能的快速提升，推动数据资源结构持续优化，计算数据、合成数据将超过影音视听、监测摄像等数据。无人驾驶、具身智能、低空经济等战略性新兴产业和未来产业的数据资源规模将保持高速增长。

### （二）数据资源流通利用进入加速期

可信数据空间将成为数据资源规模化流通利用的新模式新路径。数据治理、数据标注、数据智能等相关企业迎来发展机遇期，信息技术服务商加速向数据运营服务商转型，数据产品和服务供给能力将大幅提升，催生一批数智应用新产品新服务新业态。

### （三）数据融合应用加快价值化进程

人工智能向千行百业渗透，数据资源开发利用模式和数据要素价值机理正在发生深刻变革。公共数据资源“1+3”政策体系落地见效，市场化配置改革破冰提速，融合应用的“催化剂”作用将进一步发挥。企业数据价值挖掘重心从服务自身数字化转型向数据要素化、市场化拓展，数据跨主体流通、跨行业应用、跨场景赋能的乘数效应加速释放。



# 附录

## 一、数据来源

有效样本数量1.44万个，覆盖27个国家行业主管部门、31个省（自治区、直辖市）、新疆生产建设兵团、5个计划单列市、12个国家科学数据中心、5个国家实验室、32个行业协会商会、206个全国重点实验室、40个数据交易机构、101个中央企业、5911个数据技术企业和8067个数据应用企业（见表1）。

**表1 有效样本情况**

序号	调查对象	调查主体
1	政府部门	国家行业主管部门
2		省/自治区/直辖市及新疆生产建设兵团数据管理部门
3		计划单列市数据管理部门
4	科研机构	国家实验室及全国重点实验室
5		国家科学数据中心
6	算力中心	算力中心
7	电信运营商	电信运营商
8	行业协会商会	行业协会商会
9	数据交易机构	数据交易机构
10	企业	中央企业
11		数据技术企业
12		数据应用企业

## 二、样本分布

报告有效样本覆盖全国31个省（自治区、直辖市）的14079家企业，其中河北、江西、江苏、广东、浙江等省份样本企业较多（见图1）。

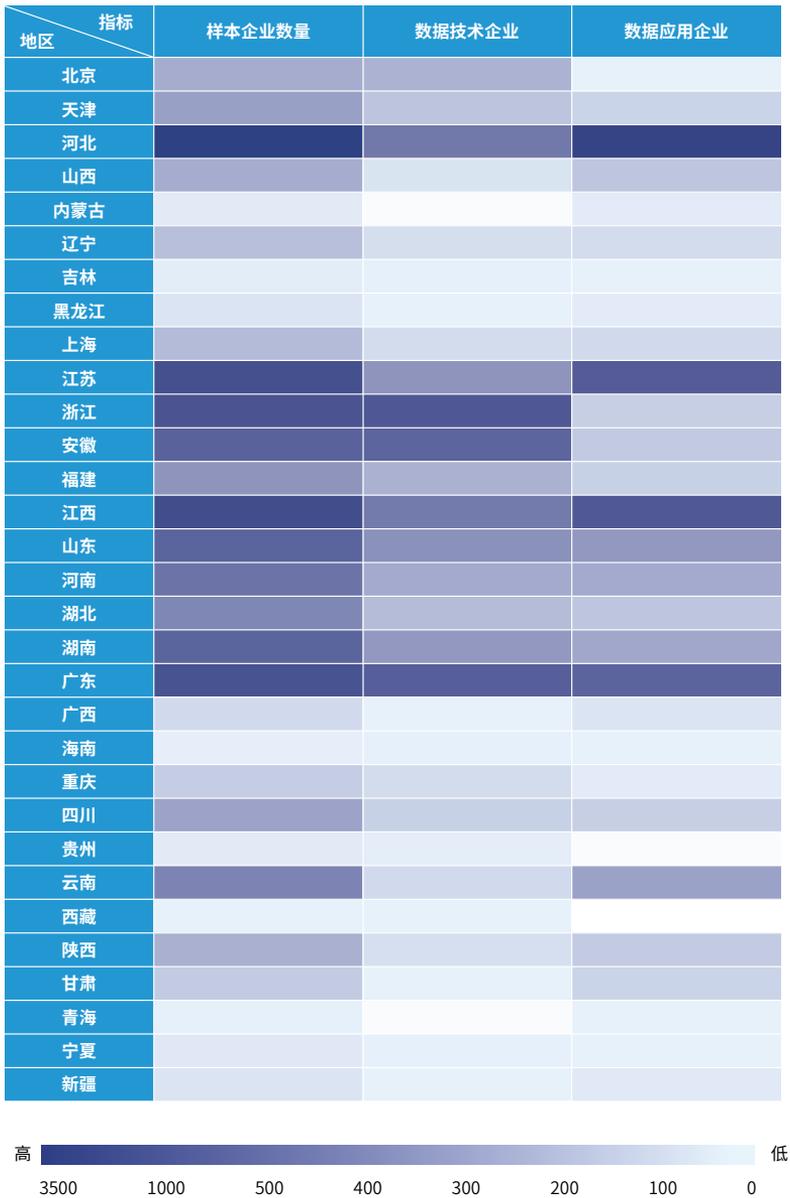


图1 企业区域分布图

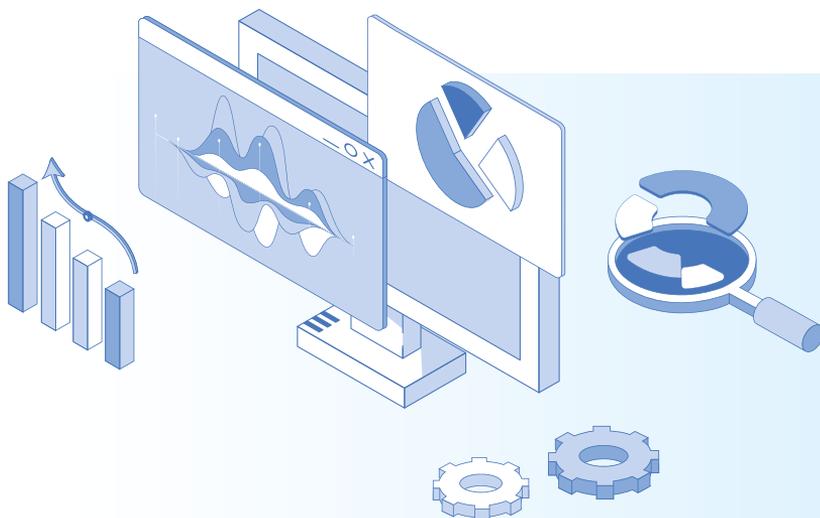
### 三、统计方法

本次调查采用分层抽样统计调查方法，将总体样本划分为不同层级，在每个层级中按比例进行独立、随机抽样，再将各层级的样本数据结合，实现对总体目标量的估计。考虑到全国企业数量多且规模差异明显，分层抽样方法相对其他随机抽样方法更能兼顾总体中各类差异，保证测算结果准确性。

本次调查将企业规模划分为大型企业、中型企业和微型企业，覆盖不同行业及规模特征。对各层级独立抽样，结合各层级的企业总数，加权推算总体企业的数据生产和存储规模。

区域测算旨在综合反映各省在数据发展方面的整体水平，基于调查的生产、存储、流通等数据，综合考虑经济基础、数字基础设施和企业数字化水平等关键变量进行加权计算。

汇总各区域数据资源相关数据，确保数据全面覆盖关键影响因素。利用机器学习模型，量化调查指标对各省数据资源发展水平的影响程度，明确决定性影响因素，结合模型输出指标重要度，为各项指标综合赋权，按照加权方式综合计算各省在各关键变量上的表现。



全国数据资源统计调查工作组

电话: 010-88680027、88686123

传真: 010-68632953