

Challenges and Opportunities of Government Statistics Services in the Era of Big Data

Chunxia Zhang

People's Government of Lingzi Town, Zichuan District, Zibo City, Shandong Province, Zibo, Shandong, 255000, China

Abstract

In the era of big data, people's dependence on data is increasing, and their way of thinking is constantly changing. At present, the statistical department using big data in the national economic accounting, economic census, census, agricultural statistics, residents and real estate price statistics, and other fields for the beneficial exploration, data of traditional statistics provides an opportunity, big data of traditional statistics is supplement rather than alternative, traditional statistics is still playing a leading role in the social statistics and economic analysis. Statistics will usher in a new development in the era of big data, and big data is becoming an important driving force for the development of statistics in China.

Keywords

big data; government statistics; challenges; opportunity

大数据时代对政府统计服务的挑战和机遇

张春霞

山东省淄博市淄川区岭子镇人民政府, 中国·山东 淄博 255000

摘要

大数据时代人们对数据的依赖不断加强,思维方式也随之在不断地发生变化。当前,统计部门运用大数据在国民经济核算、经济普查、人口普查、农业统计、居民和房地产价格统计等多个领域进行了有益的探索,数据对传统统计学提出提供了契机,大数据对传统统计学是补充而非替代,传统统计学仍在社会统计和经济分析中发挥着主导作用。统计学将在大数据时代迎来新的发展,大数据正在成为中国统计发展的重要驱动力。

关键词

大数据; 政府统计; 挑战; 机遇

1 引言

随着新产品、新业态的应运而生,电子数据日新月异,纷繁多样,其规模呈指数级增长,催生了大量新产业、新业态、新技术、新模式,大数据战略正式上升为国家战略,实时海量的交互交易大数据既为统计工作提供了全新而丰富的基础数据来源,也给统计数据的权威性构成了极大的挑战^[1]。大数据已成为推动社会发展的重要因素,通过大数据获取信息的方式越来越多,速度越来越快,人们对大数据的依赖不断加强,思维方式也随之在不断地发生变化。当前,统计部门运用大数据在国民经济核算、经济普查、人口普查、农业统计、居民和房地产价格统计等多个领域进行了有益的探索,大数据正在成为中国统计发展的重要驱动力。

2 大数据的概念

大数据,又称巨量资料,指的是所涉及的数据资料量规模巨大到无法通过人脑甚至主流软件工具,在合理时间内达到撷取、管理、处理、并整理成为帮助企业经营决策。大数据应用,是指对特定的大数据集合,集成应用大数据技术,获得有价值信息的行为。

3 统计的概念

统计是人类对事物数量的认识形成的定义,指对某一现象有关的数据的搜集、整理、计算、分析、解释、表述等的活动。统计一词起源于国情调查,最早意为国情学。一般来说,统计包括统计工作、统计资料和统计科学。统计工作的成果是统计资料,统计资料和统计科学的基础是统计工作,统计科学既是统计工作经验的理论概括,又是指导统计工作的原理、原则和方法。

4 大数据”时代对于政府统计的冲击力影响

大数据与经济社会活动的方方面面紧密相连,数据时

【作者简介】张春霞(1972-),女,中国山东淄博人,本科,高级会计师,从事统计研究。

代的到来将对政府统计工作产生重要影响。其影响主要体现在4个方面：

①可提供丰富的统计数据来源，拓展了统计调查的范围。

大数据所搜集的数据既包含结构化数据，也包括非结构化数据、半结构化数据，数量之大、类型之多、变化之快，不仅为政府统计提供了新的数据源和更加丰富的海量原始资料，而且极大地拓展统计调查的空间与范围。

②可辅助验证政府统计数据质量，有助于补充完善数据质量控制体系。

大数据具有量大、源多和速度快的特点，可以为传统的统计指标数据提供验证或补充，能够在一定程度上加强政府统计对数据质量的控制，有效提高统计数据的真实性、准确性、完整性和及时性。从世界范围看，大数据的利用已经在价格统计、就业统计、人口统计、交通运输统计、经济活动预测等多个专业领域显示出广泛前景和应用潜力^[2]。

③可促进统计设计标准化进程，有效提高政府统计工作效能。

面对大数据条件下日益丰富的电子化行政记录，以及以指数式倍增的互联网信息、数以亿计的电子化企业生产经营记录、各类信息公司积累的大量数据，这就要求统计部门必须加快统计设计标准化进程，及时制定和规范数据标准，对大数据进行有效整合利用。通过大数据搜集和分析挖掘技术等的应用，极大地缩短数据生产时间，减少报表填报程序，减轻基层负担，从而为及时发布数据和进行数据分析等赢得更多时间。

④大数据为统计调查提供了更多的数据收集途径。

大数据时代数据的收集不再局限于传统的模式，而是增加了更多的数据收集途径，比如对于一些传统模式中无法提取的数据，运用大数据技术能更好地对数据进行提取。在大数据时代，数据采集可以依靠互联网、手机和数字化行政商业记录等多种途径更好地获取想要的信息，现场数据采集的调查模式也可能被淘宝、抖音、拼多多、网络直播、银行转账支付记录、支付宝、快递物流等多种途径所代替。大数据为政府数据收集提供了极大的便利，节省了大量的人力、物力和财力，同时也极大程度地节省了时间。

山东省淄博市“四上企业”通过互联网开展活动比例见图1。



图1 山东省淄博市“四上企业”通过互联网开展活动比例 (%)

5 大数据存在缺陷

①“大数据的分析模式淡化了数据之间的因果关系。在“大数据”时代之前对于数据的收集一般都采用普查、问卷调查等的形式。在现在这个数据大爆发的社会，在“大数据”时代之后社会中存在着海量的数据，那么想要在这些海量数据之间寻找一定的因果关系就不是那么容易。

②数据安全性难以保证。数据泄露，将对人们的生活甚至财产安全产生许多不可弥补的伤害。

③大数据产生的数据质量问题难以保障。与传统数据相比，大数据时代下的数据信息系统更容易产生数据质量问题。

6 统计可以有效弥补大数据的部分缺陷

6.1 统计调查因果推断具有大数据不可替代性

大数据更重视相关关系而忽视了因果关系。《大数据时代》这本书的观点指出，数据关系是一种相关关系而不是因果关系，它知道是什么，不需要知道为什么，得出的结论往往存在一定的问题。此外，大数据所拥有的大量数据会使聚类过程变得困难，而统计调查方法可以弥补大数据的不足，是进行数据间因果关系分析的有效方法，因此抽样调查方法的因果推断能力是其他方法无法替代的。

6.2 统计调查可以作为大数据分析比较和验证的基础

大数据的结果是从大量且混乱的信息中获得的，统计调查可以用作比较和验证大数据分析结果是否可靠的工具。统计调查在研究数据质量、判断数据结果可靠性、探索数据关系等方面效果显著。调查可以有效验证大数据推理结论的可靠性，从而及时发现问题、分析问题和解决问题，以避免大数据中出现错误推断结论造成的不利影响。

6.3 统计法可以提高数据来源真实提高统计数据质量，有效防止数据信息泄密

大数据应用相关立法工作有待加强和完善，大数据从数据生成、信息收集到数据的分析和应用，在法律法规上还存在一定的空白和欠缺，如关于用户隐私、政府信息收集和管控、敏感数据管理、数据质量方面都需要进一步通过法律来进行规范和保障。可通过修订《统计法实施细则》，统计数据质量控制也要贯穿于统计工作的全过程。全体统计工作者都要参与其中，牢牢树立依法统计意识，各个主要的工作环节都要落实专人负责。只有人人依法，大家才会树立对数据质量的责任心，才能从源头上保障数据质量，才能产生优质的统计数据。规范行政记录和商业记录的使用目的、内容、方法和工作流程，建立配套的保密措施和发布机制。

7 传统统计存在的缺陷

传统统计观念与方法手段落后，这会造成统计数据方面的准确性、及时性、权威性下降，间接影响到其他职能的弱化。统计报表往往是“数字多、问卷少”“硬性指标多、可发挥指标少”，导致统计信息缺少前瞻性和可预见性。数

据统计队伍综合素质低。在基层统计人员中存在业务素质偏低,新手多,外行多,复合型人才少,尖子人才少的问题。由于基层统计人员缺乏必要统计基础知识,不能运用科学方法搜集、整理、论证统计数据,仍然沿用过去的统计方法,导致即使拥有海量正确原始数据,也不会合理运用。过度追求成绩也会影响到数据质量的真实性。企业统计人员基础薄弱。企业统计人员不稳定,也会严重影响到数据质量。

8 如何利用大数据技术提高政府统计工作的效率

8.1 提升统计大数据人员的职业素质,为统计大数据提供人才保障

大数据时代对各行各业的人员都提出了更高要求的职业素质,特别是统计人员。很多政府统计人员不是统计相关专业毕业的,甚至有些人不懂统计知识,也不懂计算机软件,因此必须加强对统计人员的培训,提高其统计软件应用能力,使其掌握对大数据进行搜集、筛选、整理和分析的方法,经常进行统计大数据专项培训。

8.2 拓展数据收集途径,为抽样调查提供数据方面的支持与便捷

在大数据时代,很多的数据可以利用现代计算机技术,通过支付记录、通信记录、行程记录、行政记录等多种渠道获取,并可以利用统计学的相关知识,对数据进行分析 and 处理。这样大大地增加了数据信息获取和收集的渠道,改变了政府统计部门的工作形式,从而降低数据调查过程中的成本,并提高调查工作的效率和质量,还可以避免人为的误差和影响,同时提高调查数据的可信度。

8.3 提高数据挖掘分析能力

“大数据”的概念来源于互联网,其要求的数据分析已经远非目前的统计数据处理技术能够实现。我们要能从这些海量数据中提取隐含的关系、模式和趋势等信息和知识供统计使用。目前,结构化的海量数据挖掘已有比较完善的方法论和挖掘工具及算法,在一些商业调查领域已经开始发挥作用。为了在未来的数据竞争中站稳脚跟,我们必须在数据挖掘等现代分析能力上下功夫。大数据潮流势不可挡,谁能率先实现大数据,谁对大数据的挖掘更为深刻,谁就将抢占未来先机。

8.4 政府统计借势改革,大数据让政府统计服务更“聪明”

打造大数据综合服务平台。大数据的发展是大势所趋,

是发展社会经济不可或缺的力量,亦是提升政务服务质量的必备要素。“大数据”是一把“双刃剑”,对于统计业务的再造、数据的采集以及数据质量的提高都有帮助,其带来的革命性冲击已经波及现行的统计制度。大数据希望官方统计进行技术创新、管理创新和应用创新。统计应当借助于“大数据”所带来的有利条件,主动顺应数据社会化的趋势,加快推进政府统计的改革。推动整体智治、整体智服,数字治理力,聚焦高质量发展,进一步打开经济发展空间。让统计服务更“聪明”。如,杭州市统计局开发的“杭数统”整体智治平台,打造数据底座、城市比较、专业分屏等子场景,基本实现主要指标“一屏览”、风险监控“一图管”、企业直报“一键直达市级部门160余万条数据,统计专业320”,实现统计全流程精准化管理。应用平台于2022年6月上线,已归余万条数据;根据统计业务需求,开发工业、投资、服务业、贸易等各专业经济指标大屏10余个;及时回流国家统计局联网直报单位数据,实现2.6万家规模以上企业的实时直报监测。为杭州建设数字经济第一城、数字治理第一城提供强劲动力。可以说,杭州深厚的数字底蕴与联合国数据论坛的主题“可持续发展数据”高度契合,为中国统计事业现代化发展过程中提供了“杭州样本”。

9 结语

统计学在大数据时代新的思路和方法已经带来了巨大的收获,但是如果继续有所突破就必须在统计学的教学方面进行改进,如果继续用传统的教学方法、沿用原来的统计学科目,忽视了对互联网技术、计算机、数字挖掘技术和管理智能的培养,会让统计学在新形势下的发展遇到瓶颈。数据对传统统计学提出提供了契机,统计学将在大数据时代迎来新的发展,应该清醒认识到大数据对传统统计学是补充而非替代。传统统计学仍在社会统计和经济分析中发挥着主导作用。统计学家不应该固守传统数据的环境,必须积极学习新生事物,适应新的大数据环境,扩展统计学的应用领域,创造出迎合大数据的新统计方法^[1]，“机遇”与“挑战”并存。在大数字时代为统计学带领的新机遇下,创造统计学发展史上的新的里程碑。

参考文献

- [1] 邱东.大数据时代对统计学的挑战[J].统计研究,2014,31(1):16-22.
- [2] 姜奇平.2013全球大数据——大数据的时代变革力量[J].互联网周刊,2013(5).
- [3] 李宜昶,康子威,杨泽华.“大数据”时代背景下财经新闻的转变[J].商业经济,2014(2):19-22.