

# 怎样审核工业生产者价格调查数据

■ 万钰淇 丁肖婷

工业生产者价格指数是国民经济核算、宏观经济分析和调控等的重要依据,其数据质量对制定工业经济政策有直接影响。笔者结合自身工作实践,就如何做好工业生产者价格调查全流程数据审核谈几点思考。

在工业生产者价格调查工作中,数据审核的难点主要有三个。一是国内外市场环境多变,价格传导差异大。近年来,国际国内大宗商品价格波动大,价格传导机制复杂多变,上游产品价格变化传导到下游产品价格的过程中,存在滞后性和差异性,数据审核难度较大。二是企业生产经营活动复杂,价格波动有时差。部分企业为降低生产成本会一次性采购大量原材料囤积备用,而后再次采购时由于新旧采购时间跨度大,导致当月上报价格环比波动较大。三是工业产品规格多、迭代快,价格波动难掌握。工业生产者价格调查遵循“同质可比”原则,但工业产品精细化、定制化发展趋

势明显,企业统计员不一定能完全掌握其企业主流规格品的价格波动趋势和价格影响因素。

针对上述难题,笔者认为可以这样去做。

紧跟形势,关注重要时间节点,做好事前预判预审。一是建立大宗商品价格旬度台账。根据国家统计局每月发布的流通领域重要生产资料市场价格变动情况,整理汇总大宗商品价格旬度台账,重点关注本地区行业权重高、与调查规格品相关度高的大宗商品价格走势。二是关注其他经济统计指标对工业生产产品价格波动的影响并进行关联预判。例如,如果近期住宅销售价格指数持续上涨态势,表明房地产市场需求呈上升趋势,则预判本月“水泥”“制作砖”等相关规格品价格有可能上涨。三是及时积累汇总重要时间节点对特定规格品价格的影响。

纵横结合,内外兼顾,做好事中即报即审。一是横向对比外部市场、相关行业价格走势,审核价格变动方向一致性。对于变动方向相反的,需

进一步核实原因。关注出厂价格与购进价格的匹配情况与衔接情况、上下游产品价格关联性。例如,如果规格品“纯碱”价格上涨,则关注其下游产品“玻璃”类产品价格是否同向变动。二是纵向对比规格品价格变动情况,审核价格变动幅度合理性。重点比较同一规格品5日和20日价格的差异是否异常,比较平均价格的环比波动是否合理,对价格变动幅度超过阈值的,在核实企业规格品符合“同质可比”原则且数据填写无误的前提下,进一步了解具体影响因素,填写登记价格变动说明。对于长期不变价格的规格品,要逐一核实具体原因。如一些机械化批量生产的陶瓷工艺产品,企业会通过一次性大批量购进原材料等方式控制产品成本,出厂的陶瓷工艺产品价格一旦确定,基本上整个生命周期的价格不再变动。三是报表内部筛选排序比较,审核变动幅度差异性。通过相似规格品筛选对比,检查该波动情况是否在相同分类中普遍存在,与同类规格品

变动幅度是否相近,对于变动幅度明显偏离其他同类规格品的,则需要进一步核实数据误差和变动原因。

勤走访、重监督,做好事后检查互审。一是带着问题去走访。把价格波动异常、价格波动情况说明描述不清晰、不详细的企业作为重点走访对象,规范企业报送,做到数出有据,并建立企业走访档案,以便复查。二是探索创新数据审核机制。探索开展企业台账集中互审工作。每年随机抽取2-3个月的统计台账和原始凭证,调市、区、镇街三级人员力量,集中互审。汇总整理存在问题并分别反馈到区、镇街以及企业,达到既培训又审核的目的。三是充分发挥数据质量核查和执法检查监督作用。每年开展至少一次集中专项检查,由队领导带队,纪检组全程监督,对于数据检查过程中发现的问题,现场反馈并详细讲解正确的数据采集方法,并限时整改,全面防范和严肃惩治统计造假、弄虚作假。

(作者单位:国家统计局佛山调查队、国家统计局顺德调查队)

## 探讨交流

# 提高居民消费价格数据质量之我见

■ 陈积恒

开展居民消费价格调查,搞准居民消费价格指数能及时反映居民消费市场价格走势,为国家进行宏观调控、国民经济核算提供重要参考依据。笔者结合自身工作实际,就如何提高居民消费价格调查源头数据质量浅谈几点看法。

当前提升居民消费价格调查数据质量主要受以下几方面因素制约:

一是部分采价员业务水平有待提高。由于有的采价员工作能力和专业水平不高,对采价流程不够熟练,导致漏拍规格品变价照片、录错价格或沿用错误价格等问题,增加后台审核人员工作负担。

二是数据审核经验不足。部分后台工作人员对于不同类别规格品价格趋势变化把握不准,不能准确了解市场变化趋势,难以发现价格变动异常数据,容易出现漏审,影响调查数据质量。

三是扫描数据质量不够稳定。扫描数据一般由调查网点专人负责每月数据报送,但因部分网点负责人更换频繁,会出现报送当月数据不及时、漏报规格品价格等情况,对数据质量造成影响。

针对上述问题,笔者认为可通过以下措施解决。

提升基层人员数据采集能力。一是多形式开展培训。在做好现场集中培训的基础上,对采价员开展闭卷考试,将成绩列入采价员年度考核项目,根据考试结果有针对性地进行讲解,加深采价员印象。采用“以老带新”的方式,让新采价员熟悉采价路线、采价方法、收集价格凭证等要求,同时跟随新采价员开展工作,及时纠正错误步骤,有效提升采价实操经验。二是丰富培训内容。除统计法律法规、典型案例作为必讲内容,强调采价原则和要求,牢固树立采价员法律意识和责任意识外,重点讲解居民消费价格基础知识、采价技巧,积极总结探讨解决采价工作中的突出问题,增加采价员业务知识储备和提高其专业素养。

提高数据审核评估质量。一是严格落实“密切督导,即报即审”数据质量管理机制。按照“采价员自审、工作人员复审、科长再审、分管领导终审”的审核流程,加大数据审核力度。在每个采价日,工作人员对数据采集管理系统中上报的价格数据做好实时审核。检查数据上报的及时性、准确性和完整性,并将检查情况记录到后台审核台账中,保证数据审核有迹可循。二是采用“横向+纵向”数据审核方式。横向对比全国、全区的指数,从八大类的环比、同比、累计比指数比较中发现趋势差异。此外,从大类找基本分类,对不同规格品价格进行细致纵向审核,结合服务和商品价格的季节淡旺特点、节假日价格变动规律和周期性规律审核同类规格品,快速排查异常数据,反馈给采价员核实,同步修正异常数据,及时做好数据质量评估。

加强电子扫描数据应用。一是强化线上沟通,线下摸底。利用走访网点、普法宣传等契机向相关企业公司积极宣传应用扫描数据的重要意义和优势,对意向企业、计划应用扫描数据的采价大类进行推广。二是寻找合适调查网点替代不符合推广条件的企业,同时根据工作量给予网点辅助调查员适当的补贴和慰问品,提高其工作配合意愿,建立稳定的扫描数据来源,稳步扩大应用扫描数据的普及面。(作者单位:国家统计局贺州调查队)

# 如何做好投资领域的四项调研

■ 万国强

固定资产投资对拉动地方经济增长提供了强大动力,做好固定资产投资统计具有重要意义。在投资统计工作中,调查研究作为基本功,是进一步提高数据质量的有力抓手。笔者结合多年投资统计工作经验,就如何做好投资领域的四项调研,提升调研工作效能谈几点心得体会。

梳理潜在对象,做好新增项目调研。结合区域重点支撑项目台账和主管部门、属地政府提供的新增项目线索,梳理新增项目台账,制定调研计划安排,确保新增项目调研覆盖率。对新增项目做好入库准备工作,指导填报对象梳理好入库材料,及时关注项目推进进展,满足开工要求后及时入库。开展培训指导工作,及时对新增项目的调查对象进行制度培训,明确填报凭证的获取途径,将问题消除在萌芽阶段。同时,加强项目的全面了解,及时与建设单位、施工单位相关负责人进行沟通,对新增项目投资发展情况、潜在问题等进行全方位分析研判。

强化数据核查,做好重点项目调研。及时梳理本地区内全部填报项目,对数据增量高、变动大的项目进行数据核查。通过现场查阅凭证,梳理调查对象的填报工作流程,查找报表环节内存在的问题,提升填报人员对制度的掌握程度。进一步深入了解调查对象存在的问题,完善年报培训及现场指导的关注重点。借助数据核查积极了解项目建设进展,对填报人员提出的问题进行及时解答,进一步夯实项目填报工作基础。

紧盯调查对象,做好需求项目调研。通过各种形式邀请调查对象,就工作存在的问题、难点及时开展需求调研。通过本地区内各种会议及培训,鼓励



原永红 绘图

调查对象对自身工作中存在的问题及时与统计部门联系,不可“带病”填报。梳理点题项目台账,及时到项目建设地核实情况,协调各方面存在的问题,提供最优解决方案。事后加大项目数据审核力度,及时开展总结工作,对共性问题梳理成册下发到相关单位。

深入重点领域,做好特色专题调研。结合区域

发展定位加强对重点投资领域的分析研究,针对重点投资领域的发展现状、存在问题、未来趋势开展深层分析,提出有利于区域投资发展的意见建议,切实做好调研工作的成果转化。加强与政府部门之间的沟通,定期撰写监测分析、预警研判等,为各级政府部门决策做好统计服务工作。

(作者单位:北京市怀柔区统计局)

## 统计科普

# 什么是网络数据抓取?

■ 赵军利 董倩

大数据是政府统计数据的重要补充来源,在政府统计中的应用越来越广泛。大数据的特点是数据来源丰富且数据类型多样,传统的数据采集方法难以应对,需要通过新技术来采集数据。网络数据抓取是获取大数据的重要技术之一。

### 一、什么是网络数据抓取

网络数据抓取(Web Scraping)是指采用技术手段从大量网页中提取结构化或非结构化信息,按照一定规则和筛选标准进行数据处理,并保存到结构化数据库中的过程。目前网络数据抓取采用的技术主要是对垂直搜索引擎(针对某一个行业的专业搜索引擎)的网络爬虫(或数据采集机器人)、分词系统、任务与索引系统等技术的综合运用。

### 二、网络数据抓取有什么作用

科学研究离不开详实可靠的数据,互联网的发展提供了新的获取数据的手段。面对海量的互联网数据,网络数据抓取技术被视为一种行之有

效的技术手段。相比于传统的数据采集方法,网络抓取数据无论时效性,还是灵活性均有一定的优势。利用网络数据抓取技术,可以在短时间内快速地抓取目标信息,构建大数据集以满足分析研究需要。

### 三、网络数据抓取流程

网络数据抓取的一般步骤包括:(一)确定数据抓取的目标网站。根据研究需求确定所需信息的来源网站。(二)网站的源代码分析。逐个分析各来源网站的数据信息组织形式,包括信息的展示方式以及返回方式,比如在线校验格式化的工具(JSON),在线格式美化工具(XML)等,根据研究需求确定抓取字段。

### (三)编写代码。

分析时尽量找出各来源网站信息组织的共性,这样更便于编写服务器端和数据抓取端的代码。

### (四)抓取环境测试。

对抓取端进行代码测试,根据测试情况对代码进行修改和调整。

### (五)数据抓取。

将测试好的代码在目标网站进行正式数据抓取。

### (六)数据存储。

将抓取的数据以一定格式存储,比如将文本数据内容进行过滤和整理后,以excel、csv等格式存储,如果数据量较大也可以存储在关系型数据库(如MySQL、Oracle等),或非关系型数据库(如Hadoop、Spark等)提供的分布式文件系统中。数据存储完成后,基于整理好格式的数据,可以根据分析目标执行各类数据挖掘和机器学习算法,如分类、建模、预测等等。

### 四、应用案例

目前世界上将网络抓取数据用于研究分析编制指数的成功案例有麻省理工学院的“10亿价格项目”(The Billion Prices Project, BPP),在经济学家阿尔伯特·卡瓦洛(Alberto Cavallo)和罗伯特·瑞格本(Roberto

Rigobon)带领下,项目研究团队利用网上商品价格数据计算“每日网上价格指数”(Daily Online Price Index),以反映日常通货膨胀程度。通过网络抓取技术,研究人员每天在网上抓取多于50万条商品价格信息,其数据量是美国政府价格统计收集数据的5倍,抓取的数据包括世界70个国家、300个零售商、共500万种在线商品的价格。由于价格信息不是通过访问实体商店的传统方法获得,网络抓取数据的成本相对较低。相比传统CPI的月发布机制,“10亿价格项目”仅有3天的滞后期,几乎实现了通货膨胀的实时预测,网上价格指数与传统CPI在趋势上高度吻合,在官方统计数据发布之前就能够大致了解通货膨胀的走势。

此外,通过网络数据抓取技术,所获得的搜索数据被广泛应用于舆情调查分析。基于网络搜索的关键词在一定程度上反映了社情民意,可以发挥网络搜索数据在制定公共政策方面的参考价值。

(摘自《领导干部统计知识问答》(第二版),中国统计出版社)

# 浅谈畜牧业全面统计调查

■ 周雨晗

畜牧业全面统计调查数据反映畜牧业生产成果,是“菜篮子工程”建设的重要数据支撑,是地方分析畜牧业发展变化规律,做出正确决策的重要依据。笔者结合基层统计调查工作经验,总结当前调查存在的难点,并就难点提出几点解决建议。

当前,畜牧业全面统计调查工作存在的困难主要体现在以下几方面:

报表周期短、频率高。畜牧业生产情况报表分季报和年报,均采用全面调查、逐级汇总的调查方法。调查对象是养殖场(户)、生产单位,每个报告期由村级统计员入户采集数据,汇总至镇级,再经县(区)统计局、市级调查队逐级上报至省级、国家。数据来源涉及千家万户,部分乡镇畜禽养殖场(户)、生产单位达几百家甚至更多。每季度需村统计员逐户调查,工作量较大。

统计人员业务能力有待提高。目前,涉农统计工作含农牧业、人口普查等工作,村、镇及县(区)统计人员从事统计业务时间短,参加统计调查培训少,统计业务经验不足。

联网直报平台审核功能有待完善。目前的畜牧业生产情况联网直报审核功能,仅限于统计指标数值的简单加减逻辑关系审核,包含与被包含指标数值的大小关系审核,其他需人工审核。

统计人员法律意识有待提升。部分统计人员统计法律意识较为淡薄,主要表现在村、镇级统计人员对统计工作重要性认识不足、责任心不强,对存在明显逻辑关系问题的数据不核实、不纠正等。

针对上述难点,笔者建议可从健全报表制度、加强队伍建设等方面着手加大工作力度。

进一步完善畜牧业统计报表制度。一是以多种调查方式替代畜牧业全面统计调查。以每年畜禽监测样本轮换为基础,采用抽样调查、重点调查、典型调查等非全面调查相结合的统计调查方法。二是建立调查队、统计局与农村局三个部门互动的工作模式,互通工作信息,共享调查成果。三是扩大主要畜禽监测调查样本量。在主要畜禽监测调查样本量的基础上,实行对每个县(区)开展主要畜禽监测调查。四是建议加快推进畜牧业全面统计调查信息化,改变农村调查纸质报表多级上报汇总的现状,减少数据传递的中间环节。

稳定乡镇统计人员队伍。可采取奖励制度,通过表扬先进和业务考核,提升乡镇统计人员工作积极性。积极争取各级政府对畜牧业调查工作的重视和支持,以地方配套资金提高乡镇统计人员的调查补贴。

完善统计联网直报平台功能。在原来的统计联网直报平台上,完善报表的审核功能。除具有对报表的简单逻辑关系审核功能外,还应增加相关指标间的匹配性审核,报表间的关联性审核等功能,如肉产量与出栏量之间的胴体重审核,出栏与存栏的比例关系审核,奶牛存栏与牛奶产量的比例关系审核,蛋鸡存栏与鸡蛋产量的比例关系审核等。

强力推进统计法治建设。以检查数据质量为契机,深入基层加强督导检查,坚决贯彻执行统计法,不断增强法治意识。加强对县(区)、乡镇的执法检查,重点核查调查数据是否准确,账机是否一致,逻辑是否错误等。加强统计法律法规宣传,增强县(区)、乡镇统计法律意识。

(作者单位:国家统计局北海调查队)