

跟踪统计应用 聚焦产业发展

本版由国家统计局统计科学研究所指导

探索“智慧收支调查”新路径——

鸡西队研发新版电子编码表

■ 贾春哲

针对居民收支调查旧版电子编码表不能在Linux系统运行的实际问题,国家统计局鸡西调查队积极探索,运用Python语言研发新版住户收支电子编码表,在“智慧收支调查”道路上突破发力。新版电子编码表可以在多平台运行,很好地解决了不兼容问题,并且用户界面友好,查找速度快,可以有效提高居民收支调查的效率和效能,进一步推动基层基础规范化建设。

——实时更新,匹配最新编码数据。新版电子编码表数据库能保持编码要求的最新状态。每年国家统计局都要对一些编码进行修订和调整,2022年也有一些编码发生了变化。目前鸡西队研发的新版电子编码表数据库经逐条核对,在编码、记账项目和记账项目解释等方面全部与2022年新版调查员手册保持一致。并且,在今后编码发生修订时也可以进行实时更新,与国家最新的编码要求时刻保持同步,确保编码工作准确无误。

——检索快捷,实现三种查询功能。根据住户收支调查编码逐级多分类的特点,新版电子编码表设定了树形结构的多层分类,可以由收支大类到具体



工作人员在入户调查

小类逐级查看记账项目及具体项目解释。对于能确定大类但是不好确定具体小类的编码情况,可以选定大类,然后在具体的下级分类指标里查看各项子类的记账项目解释,进行比对判断,来确定最适合的项目进行编码。在新版电子编码表中,还可以搜索编码相关的关键字,编码表自动在编码、记账项目和

记账项目解释中进行匹配查找,将匹配到的编码项目以数据列表的形式展示出来,点击搜索结果中每一项目,可将选择项相关的编码、记账项目、记账项目解释、单位数量及录入控制码等全部编码信息显示出来,从而甄别选取最适合的编码。为进一步提高搜索精度,提高工作效率,新版电子编码表可以在搜索时使用联

合搜索功能。将多个关键字以空格分割后进行匹配搜索,编码表将搜索包含所有关键字的编码项目数据,并以数据列表形式展示出来,可查看每个选择项相关的编码、记账项目、记账项目解释、单位数量及录入控制码等信息,方便进行比对甄别,有效提高收支编码的准确性。

——质效双升,既快又准。在常规的编码工作中,常用收支编码大家都比较熟悉,可以直接进行编码。但是住户收支编码有200多个分类,6位编码指标有800多个,数量和分类繁多,对于一部分不常用的编码,则需要详细查看具体解释,对编码进行判断确认。如果通过手翻查阅纸质编码本或者调查员手册,效率较低。利用新版电子编码表,通过分类查找或关键字搜索实现编码的迅速查找定位,并根据编码表提供的详细记账项目解释进行判断甄别,可以有效提高编码效率,节省编码工作时间。电子版编码表设定多层逐级分类,在处理相近编码或者可能根据不同具体情况有不同编码的项目时,根据各级分类的用途对比,准确地甄别正确编码,防止混编错编,提高了编码准确率,进而提升了基础数据质量。通过新版电子编码表,目前鸡西队收支调查编码工作效率和准确性明显提升。

算法驱动 数据赋能
济南局为纳统提质
插上智慧翅膀

■ 吴春燕 王瑞云

今年以来,山东省济南市统计局全面提升统计智慧化水平,选取了天桥区先行先试,搭建“准四上”单位跟踪监测系统,在运算方法和数据处理上开拓创新,实现多个技术突破,用大数据为“四上”企业纳统工作提质增效。



天桥区统计网格员走访企业

济南市“准四上”单位跟踪监测系统和企业微信整合了市场监管、税务等部门的数据资源,形成“1+3+3+N”的智能信息化体系,即以建设监测系统1个平台为中心,利用统计地理信息、行业代码、产业代码智能编码3个数据库,形成基本单位名录库、准“四上”企业培育库、扶持政策库3个数据库,实现企业信息电子采集、名录库智能维护、数据可视化等多个应用功能。



针对在调查普查中行政区划代码不准确、基层走访核实任务无法有效分解等问题,济南局运用正则表达式和经纬度匹配相结合的方法,实现企业经营地址精准匹配,为山东省首创。首先收集各街道地址信息,根据关键字编写正则表达式模板,进而实现对地址字段的批量处理。同时,利用高德、腾讯、百度等第三方地图供应商提供的API开放接口,将地址转换为经纬度坐标,使用Geopandas、Polygon等Python三方扩展包计算经营地坐标所属的行政区域,与正则表达式匹配结果相结合,实现了普查调查任务的精细划分和企业地址的精准定位。

以跨行业、跨部门的数据共享机制为依托,济南局将“准四上”单位跟踪监测系统与统计基本单位名录库对接,构建了名录库智能更新机制,实现准“四上”企业库自动生成。综合使用多元Logistic回归统计方法和深度学习算法,在整合部门数据的基础上,以企业主要经营指标为变量,训练多层神经网络模型,实现横向、纵向间数据匹配性整合、验证和控制,判断企业经营规模和活跃度,智能筛选生成准“四上”企业储备库。将名录库维护等工作交给系统来完成,切实减轻基层工作负担,显著提高“上规入库”效率,为行业部门开展优质企业孵化培育提供了信息化技术支持。

广西总队用信息技术提升调查质效

本报讯 今年以来,国家统计局广西调查总队探索创新信息技术应用方法,不断提高调查业务信息化水平,提升统计调查业务工作质量和效率。

——积极推广自主开发平台。依托华南数据中心落户广西的独特优势,广西总队大力研发并积极推广使用广西国家调查综合信息平台,充分发挥平台集成优势,不断收集业务处室使用过程中的问题和建议,优化平台相关功能,让平台功能更加贴合调查业务需求。今年积极推进应用平台开展畜牧业调查电子化记账工作,目前,畜牧业调查信息化覆盖全部调查企业,实现广西畜禽调查台账电子化,切实减轻调查员工作负担。

——积极研发业务数据处理程序。根据相关业务处室的需求和实际工作流程量身订制业务程序,着力提高工作质效。一是编写消费价格调查数据处理程序。目前消费价格调查使用的CPI20程序生成月报需要人工进行大量的指标挑选,通过研究分析CPI20的数据存储规律和汇总数据结构,使用SQL语言有针对性地编写了自动筛选生成报表程序,替代后的程序报表处理耗时由天计缩短到1-2分钟。二是研发房地产价格调查数据调整程序。对房地产价格调查HPI程序进行了逆向分析和研究,针对HPI程序缺少的数据调整工作流程,进行补充完善,数据处理工作缩短到半天。三是对采购经理指数调查的程序进行完善,在程序上满足业务分类需求,达到无需人工分类汇总的目的。

——积极采购智能化服务。根据劳动力调查业务的需求,积极采购智能外呼机器人服务用于劳动力调查电话回访工作。在采购过程中,多次召开劳动力调查回访智能外呼机器人服务功能演示会议,多次测试查找智能外呼机器人在电话呼叫和访问过程中存在的问题,推动智能外呼机器人在劳动力调查回访工作中的使用达到预期效果,有效提高调查质量。

黄丽婷 唐俊



米易队开展无人机驾驶员业务培训

近日,国家统计局米易调查队前往米易县撒莲镇高王官村,开展无人机驾驶员业务培训暨玉米长势情况实地调查。

胡云波 文 冷丁 摄

利用信息技术 创新工作方式

江苏总队提升房价统计效能

本报讯 国家统计局江苏调查总队充分利用信息技术开展房地产价格统计,从夯实工作基础、做好市场调研、深化审核分析、加强部门合作四个方面,积极创新工作方式,提升房价统计效能,及时监测分析房价变动情况,为房地产市场调控提供统计数据支撑。

江苏总队陆续建立房价指数、网签数据、问卷调研结果、土拍情况、政策变化等方面的数据库,按月及时更新。全省房价数据库的建立既方便管理保存资料,又利于对数据的完整性、连续性进行检查分析。各市级调查队建立房地产开发企业和经纪机构数据信息库,及时掌握市场重要主体的生产经营情况,每月根据网签成交量、市场变化情况科学选取样本跟踪调研,充分了解各城市房地产市场活动情况,对房价指数统计生产具有重要意义。

今年以来,疫情多点散发,经济下行压力较大,各地住宅销售量下滑明显,直接影响房价调查的样本量和数据质量。江苏总队使用在线调查平台加强线上调研。一方面,邀请房地产开发企业和中介机构扫描二维码填写调研问卷,通过问卷了解新建商品住宅和二手住宅市场发展现状、未来发展信心、居民购房需求等方面情况;另一方面,开展典型调研,获取鲜活的个体案例,详细了解房地产市场面临的风险和问题,剖析企业、中介机构存在问题和意见建议,形成多篇调研报告。

房价统计调查每月需审核10多个文件,涉及的数据多而杂,且文件内部和不同文件之间存在着复杂的逻辑关系,人工审核工作量巨大且容易出错。江苏总队使用Python开发小程序、基于Excel使用小插件,解决人工审核不足问题,有效提高数据审核效率和数据质量。同时,增加问卷数据的汇总分析功能,实现各区域中位数、开盘、持销盘等价格计算功能,对各大房地产交易平台指数和当地有规模的中介平台网站的信息进行抓取分析,重点关注不同区域、不同面积楼盘的市场情况,提高分析的科学性,为决策提供有效数据支撑。

为做好房地产价格统计,各级调查队积极与住建、自然资源、税务等部门建立共享平台,及时交换住宅网签数据、土拍、房地产政策、市场走势变动情况。组建网签数据核验工作组,制订核验办法、核验内容,现场核验网签基础数据与商品房买卖合同、企业台账等资料,甄别新建商品住宅网签备案数据中项目名称、开发公司、房号、建筑面积、销售单价、销售总价、装修标准等信息是否准确、登记是否及时、签约时间及备案时间是否符合规定。按季度参加江苏省住房与房地产工作领导小组联络员会议,发挥多部门群策群力的优势,加强对房地产市场发展趋势的分析和研判,全力构建房地产市场健康发展长效机制。

孟映云 郝辉辉

黔南队开发劳动力调查审核程序

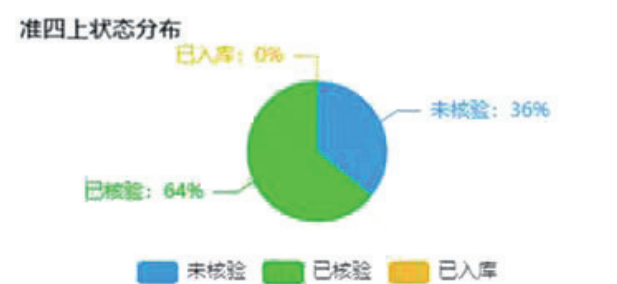
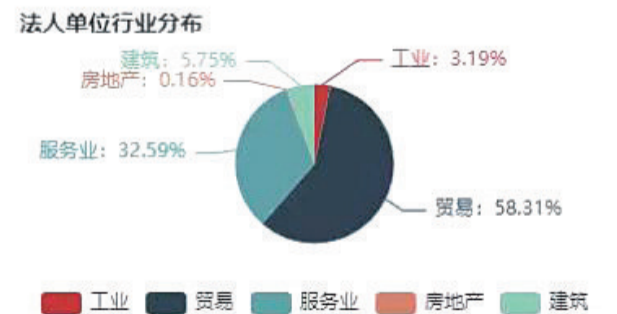
本报讯 为进一步强化大数据在统计工作中的运用,国家统计局黔南调查队创新方式方法,利用Python语言解释性、可扩充性、可移植性等特点,在严格遵守国家统计调查方案的前提下,于近期自行开发了月度劳动力调查审核程序。

该套审核程序充分利用Python语言在大数据处理方面的优势与pandas等模块在解决数据分析任务上的便捷性,使得审核效率大大提高。经测试,在黔南州劳动力调查数据量扩增后较以往增加一倍的情况下,审核效率可以提升十倍以上。除审核“规定动作”以外,还可扩充“自选动作”条件审核,同时针对黔南州不同县(市)电脑设备的不同操作系统进行了兼容性优化,可操作性更强。此外,利用Python中包含的Tkinter模块,针对程序UI界面进行了更改和开发,进一步拓展了数据审核程序的操作性与便捷性。目前,该系统已在黔南州辖区范围内县市区推广使用。

该审核程序有效地解决了劳动力调查数据审核速度慢、效率低的问题,及时发现漏审、误审等问题,实现了劳动力调查数据审核的标准化、规范化、清单化,较大程度减轻了基层业务人员的审核工作量。

据了解,黔南队将根据程序使用中出现的,进一步优化软件功能,简化操作流程和步骤,不断促进劳动力调查工作提质增效。

王祚轩



济南局利用地理信息系统等技术对系统内各行业统计数据进行分析挖掘,为主要经济指标赋予空间维度。利用莫兰指数计算评估统计指标的自相关程度,通过曲线图、散点图、冲积图等形式,形成行业热力分布、重点企业聚集度等多元可视化地理画像,有利于充分展示区域经济发展状况,实现监测平台数据多元化展示。

据济南局介绍,通过平台内嵌的综合数据分析模块,还可按照不同需求提供定制化服务,形成详细的发展宏观数据图像,帮助各级领导和统计人员准确分析研判趋势,更好地发挥数据监测和预测预警效能,为经济社会高质量发展提供更智能、更高效、更优质的统计服务。