

现代统计信息系统 理论与实践

杜 栋 主编



清华大学出版社

现代统计信息系统理论与实践

杜 栋 主编

清华大学出版社

北 京

内 容 简 介

本书从不同层次统计人员的实际需要出发,系统地介绍了统计信息系统的基础理论知识、统计信息系统的核心信息技术、统计信息系统的开发和应用,并提供了统计信息系统的大量背景资料和典型案例。

本书可作为参与统计信息与决策研究的科技人员和管理工作者的参考书,也可作为统计学专业研究生和高年级本科生“统计信息系统”课程的教材。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

现代统计信息系统理论与实践/杜栋主编. —北京:清华大学出版社, 2004

ISBN 7-302-08436-X

I. 现… II. 杜… III. 统计-管理信息系统 IV. C8

中国版本图书馆 CIP 数据核字(2004)第 028803 号

出 版 者: 清华大学出版社 地 址: 北京清华大学学研大厦

<http://www.tup.com.cn> 邮 编: 100084

社 总 机: 010-62770175 客户服务: 010-62776969

组稿编辑: 欧振旭

文稿编辑: 鲁秀敏

版式设计: 张红英

印 装 者: 北京鑫海金澳胶印有限公司

发 行 者: 新华书店总店北京发行所

开 本: 185 × 230 印张: 13.25 字数: 170 千字

版 次: 2004 年 5 月第 1 版 2004 年 5 月第 1 次印刷

书 号: ISBN 7-302-08436-X/F · 778

印 数: 1 ~ 4000

定 价: 18.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话: (010)62770175-3103 或(010)62795704。

前 言

现代社会，人们对统计数据和信息的需求越来越大，大量的管理决策活动已无法离开统计信息的支持。传统的计算机数据处理系统虽然在统计信息的准确性、时效性和丰富统计内容方面起到一定的作用，但只是把统计原始数据转换成预信息，还没有成为决策用信息。因此，它只是信息系统的低级阶段。

统计工作要利用计算机增强和扩大为社会服务的能力，不仅仅只是提供统计的汇总信息，而是要建立统计信息系统，采用现代科学的统计方法进行深度开发，及时准确地提供有效的统计信息，满足各种层次的管理决策需要。统计信息系统以计算机为工具，运用多种方法，对统计信息收集、存储、分析、加工、传递和使用，把统计人员从繁重的手工作业中解放出来，提高了统计工作的质量、效率和水平，使统计工作现代化和科学化。

为了反映最新科技尤其是现代信息技术在统计工作中的应用，使统计人员熟悉当今统计信息系统的相关知识，特别是探讨基于数据仓库以及在网络环境下的统计信息系统开发和应用，我们编写了这本书。本书从统计人员的实际需要出发，系统地介绍了统计信息系统的基础理论知识、统计信息系统的核心信息技术、统计信息系统的开发和应用。最后，提供了大量统计信息系统的背景资料和典型案例。

本书注重突出以下四个特色。

- 广泛性。从各类统计人员的实际需要出发，尤其是从国家、省市、企业三个层面来展现统计信息系统。
- 先进性。强调最新的理论和技术，特别是网络环境下基于数据仓库的

统计信息系统，具有一定的学术价值。

- 实用性。重视案例，这是开展理论与实践结合的比较有效的方式，也是本套丛书的主要特色之一。
- 通俗性。本书在编写过程中，注意深入浅出，用简洁、流畅的语言描述，可读性比较强。

本书由杜栋主编。徐绪堪参编了第 3.1 节，陈建鑫提供了一些材料。本书的编写参阅和引用了许多作者的研究成果（在参考文献中仅列出了主要的书目和文章），在此向他们表示由衷的谢意。莫潇、向子梦、沈晶、张海燕、杭卫、苏伟文、韦发明、胡婧婧、杨波等同学搜集和提供了大量背景资料和典型案例，在此也对这些资料和案例的原创作者表示敬意。

当代信息技术高度发展并广泛应用，统计信息化的步伐不断加速，给统计信息系统的理论和实践提出了许多新课题。本书虽然提出了一些新的思想和观点，比如基于数据仓库技术和网络环境下的统计信息系统，但仍需要进一步探索和研究。由于水平有限，书中的缺点和不足之处在所难免，恳请广大读者批评指正。

编 者

2003 年 12 月于龙城

目 录

第1章 总论	1
1.1 统计信息的基本概念	1
1.2 统计信息系统的基本概念	5
1.3 统计信息系统的统计理论基础	11
1.4 统计信息系统的主要技术要求	13
第2章 统计信息系统的基础知识	16
2.1 计算机统计数据处理概述	16
2.2 计算机统计数据处理的全过程	17
2.3 计算机统计数据处理的实现方式	22
2.4 计算机统计数据处理与统计信息系统	23
第3章 统计信息系统的核心技术	25
3.1 数据库与数据仓库技术	25
3.1.1 数据库基础知识	25
3.1.2 数据仓库及其相关技术	29
3.2 基于数据仓库的统计信息系统	42
3.3 通信与计算机网络技术	44
3.3.1 通信基础知识	44
3.3.2 计算机网络及其相关技术	48
3.4 网络环境下的统计信息系统	60
第4章 统计信息系统开发	63
4.1 信息系统开发概述	63
4.1.1 信息系统开发的全过程	63

4.1.2 信息系统开发人员的组织	67
4.1.3 信息系统开发的项目管理	70
4.1.4 信息系统的开发方法	72
4.2 统计信息系统的分析	74
4.3 统计信息系统的设计	85
4.4 统计信息系统的实施	100
第5章 统计信息系统的应用	105
5.1 统计信息系统的转换	105
5.2 统计信息系统的运行	107
5.3 统计信息系统的维护	109
5.4 统计信息系统的管理	111
附件 统计信息系统的背景资料和典型案例	115
附件 1 2000 年统计信息化发展概况	115
附件 2 国家统计局信息工程实施方案简介	122
附件 3 国家统计局信息网络系统	125
附件 4 国家统计局信息自动化系统	133
附件 5 国家统计局 5000 家企业直报系统	142
附件 6 劳动和社会保险统计管理信息系统 (SMIS) 简介	144
附件 7 江苏省统计信息工程建设规划 (2002—2005 年)	154
附件 8 北京市统计信息系统	157
附件 9 长沙市统计信息系统建设	163
附件 10 某市统计信息系统的开发	172
附件 11 通用统计信息处理平台 ADT- EasyStat	180
附件 12 煤炭工业统计信息管理系统开发研究	184
附件 13 优尼泰通用统计管理系统	188

附件 14 汇创石化企业综合统计信息系统 (COM-STSYS)	193
附件 15 西北某厂计划统计信息系统的开发与应用	195
参考文献	201

第 1 章 总 论

本章首先介绍了统计信息和统计信息系统的基本概念、特点以及在国民经济管理中的作用。然后探讨了统计信息系统的统计理论基础和主要技术要求。

1.1 统计信息的基本概念

1. 统计信息的概念和作用

信息是事物运动状态和过程以及关于这种状态和过程的知识，它的作用是消除观察者在相应认识上的不确定性。

统计信息的定义有广义和狭义之分。广义的统计信息是指运用统计方法处理的对人类活动产生影响的以统计数据或资料形式表现的信息。狭义的统计信息是指统计工作的统计信息，即通过统计部门或专（兼）职统计人员在统计设计、调查、整理、分析等过程中形成的统计信息。需要强调的是，讨论统计信息的定义有着重要的现实意义。目前至少有两点是明确的：广义统计信息定义的确立，有利于纠正长期以来人们对统计的偏见；也有利于扩大统计的研究和应用领域。

在社会经济管理活动中，一般把统计信息看成是对统计数据进行统计加工处理的结果，把统计数据看成是生成统计信息的原始材料。也就是说，它们之间的关系是原材料和结果的关系（如图 1-1 所示）。

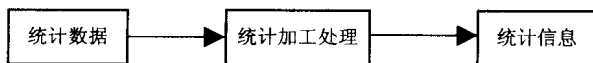


图 1-1 统计数据与统计信息

统计数据通常是指记录下来的统计事实。它具体是指统计指标的数值和统计指标的名称。它是一些孤立的原始事实，一般不能使用。

统计信息是指经过统计加工处理的对人类活动产生影响的统计数据。它是有价值、有意义的统计数据，也是社会经济信息的主体。

统计信息反映客观情况，往往表达或体现了人们对一事物的认识。统计信息往往与决策密切相关，并能指导人们的行动。也就是说，统计信息可以反映和监督社会经济活动，并帮助人们做出决策。

统计信息从形式上表现为数量特征的特点，这是由统计学研究社会经济活动数量方面的一般方法论和将其运用于认识社会经济活动数量方面及其规律性决定的。由于统计信息内容的数量特性，所以统计信息经常表现为统计数据的形式，但是统计信息决不局限于统计数据的形式，还有图表、文字、口头表达等形式。比如，统计图是依据统计资料所绘制的图形，用它表示统计资料的显著特点是形象生动，看后给人留下深刻的印象；统计表是传统的表现统计资料的最基本的形式之一，其特点是简明扼要，便于进一步的比较对照和分析研究。另外，在计算机硬件和软件飞速发展的今天，除了以上传统的文字、表格、图形之外，其他媒体和形式的统计信息也不断涌现，比如计算机存储介质、统计数据库、声音、图像、动画以及录像等。

统计信息的职能是指根据科学的统计指标体系和统计调查方法，系统地采集、加工、存储、传递和提供大量以数据描述为基本特征的社会经济信息。它从数量上反映社会经济现象之间的联系，不仅能够描述事物过去的发展过程，而且还可以反映事物现在的状况，依据这些统计信息，还可以运用科学的方法预测事物未来的发展趋势，从而有效地指导市场经济有序地运行。统计信息是

国家、地方、部门、企业和个人了解整个社会经济运行状况，参与社会经济活动，对复杂的社会经济运行进行全面、系统的定量监测和综合分析，准确描述、评价、预测社会经济发展，在全社会范围内合理开发、配置各种资源，有效调节和控制社会经济运行的根本保证。

目前，我国的国民经济正沿着社会主义市场经济的道路健康发展，市场经济越发达，对统计信息的要求越高，各级领导要加强和提高宏观调控能力，对统计工作的要求也越高。统计信息是经济现状和社会发展的基础材料，必须做好其组织开发和管理利用工作。

2. 统计信息的特征

统计信息主要有如下几个特征。

- (1) 事实性。统计信息是排除了不确定性反映客观事物的实际情况，并对接受者有价值的消息。事实性是统计信息最重要的特征。
- (2) 数量性。统计信息是从数量上反映社会、经济、科技的运行状况的，也是通过数量来认识和揭示它们的运行过程和运行规律。
- (3) 社会性。统计信息是一种人工信息，是经过人类加工处理过的信息，是反映人类活动状况的信息。因此，它具有被全社会所理解和认识的特性。
- (4) 有效性。统计信息是人们为了管理和决策的需要，有目的地收集、加工、处理、使用的。它的价值和效用也与适合管理和决策的需要成正比。
- (5) 动态性。统计信息产生于人类活动之中，人类活动（尤其是经济活动）是持续不断变化的。这就要求人们要注意信息的变化，不断进行信息的更新。

- (6) 时效性。统计信息在时间上往往滞后于事物发生的时间。依据及时的信息进行决策, 往往成功; 依据过时的信息进行决策, 则往往失败。
- (7) 不完全性。统计信息不可能反映客观事物的全部内容。因为统计方法不同, 处理方式不一, 以及时间上的偏差, 总会造成信息丢失。
- (8) 再生性。反映统计对象在一定时空条件下所处状态的原始统计信息, 经过不同科学方法的分析加工, 会产生不同的再生信息。

3. 统计信息的分类

统计信息的分类方法很多, 以下仅从实用的角度介绍几种。

- (1) 按照统计信息的来源, 可分成原始信息和经过加工的信息。原始信息是指基层单位的原始记录, 是统计信息的基础。原始信息经过加工处理, 以及分析提炼所生成的信息为经过加工的信息。
- (2) 按照统计信息反映的内容, 可分成宏观信息、中观信息和微观信息。反映国民经济运行的信息为宏观信息, 反映企业经济活动的信息为微观信息, 介于两者之间的反映地区(区域)、行业(产业)的信息为中观信息。
- (3) 按照统计信息的使用者和提供者, 可将统计信息划分为政府统计信息、企业统计信息和社会个人统计信息。这是因为在社会主义市场经济体制下, 由于社会经济利益主体多元化, 政府、企业和居民个人在不同层次上各自按自身的运作方式参与经济与社会活动, 因此它们各自有着不同的统计信息需求。
- (4) 按照统计信息反映经济活动的时态, 可分为有关过去的信息和有关未来的信息, 经常变动的信息和相对稳定的信息、即反映历史经济活动的信息、反映预测经济活动的信息、反映经济活动每时间变化的信

息、反映经济活动内在联系的具有规律性的信息。

- (5) 按照统计信息取得的形式,可分为常规性信息和偶然性信息、正式渠道获取的信息和非正式渠道获取的信息。常规性信息是指反映经济活动正常情况,按照一定程序以经常不断的形式进行收集和处理的的信息。偶然性信息是指反映经济活动中特殊的、突发的事件,而要进行特殊的、紧急处理的信息。正式渠道获取的信息是指按照制度规定的渠道获得的信息。非正式渠道获得的信息是指从正式渠道以外其他多种途径获取的信息。
- (6) 按照所属国民经济部门,可以分为工业统计信息、农业统计信息、劳动和工资统计信息、基本建设统计信息、住宅公用事业统计信息、运输和邮电统计信息、物资设备供应统计信息、商业统计信息、财政统计信息和新技术统计信息等。

1.2 统计信息系统的基本概念

1. 统计信息系统的概念和分类

系统是指若干相互依存、相互制约的要素为了实现确定的目标而组成的具有特定功能的有机整体。即系统存在需要3个基本条件:结构、功能和目标。

统计信息系统是在传统的人工统计工作基础上,应用统计理论、信息技术、系统工程方法建立的人/机复合系统。具体地说,统计信息系统是一个由人、计算机等组成的,根据统计指标和指标体系进行统计信息的搜集、传递、存储、加工、分析、维护和使用的系统。

从统计信息传播的角度看,统计信息系统由统计信息、统计信息需求方、

统计信息提供方和统计信息传播渠道 4 部分组成。这里的统计信息是指统计部门经过调查、整理、分析所获得的统计数据；统计信息提供方是指专门负责数据收集、加工、处理的单位或部门（甚至包括咨询公司和民间统计）；统计信息需求方是指作为社会经济主体的政府、企业以及广大社会公众；统计信息传播渠道是指包括 Internet 网络、光盘、图书、报纸、杂志等在内的各种信息媒体。

统计信息系统能对大量数据做进一步加工，它反映国民经济各地区、各部门、各单位的经济运行情况和发 展规模，并能对经济活动中的重大趋势性问题进行分析、预测和预警，及时向上级有关部门宏观决策提供可靠的统计数据，在更大程度上满足了社会各界日益增长的对经济信息的需要。

通常把国民经济按层次和范围分为宏观经济和微观经济，实际上介于宏观经济和微观经济之间的还有一个中观经济。因此，国民经济系统通常分为宏观经济系统、中观经济系统和微观经济系统，国民经济统计信息系统相应可分为国家统计信息系统、地区统计信息系统、行业统计信息系统、城市统计信息系统和企业统计信息系统等（如图 1-2 所示）。

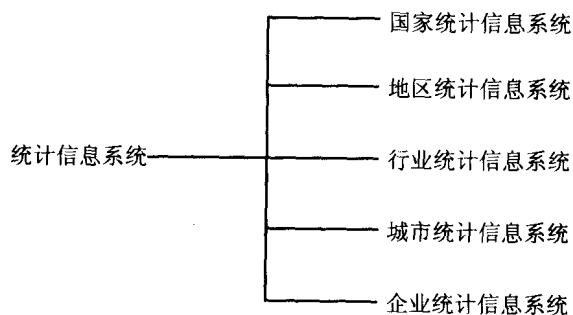


图 1-2 国民经济统计信息系统的分类

国家统计信息系统是一个大系统，它由地区和部门的统计信息子系统构成。这些子系统又由更小的子系统（例如一个工厂的统计信息系统）构成。国家统

计信息系统和基层统计信息系统虽然都是统计信息系统，且系统的目标都是提供信息，但它们在整個国民经济管理中所处的地位是不一样的，因而要求提供的数据和信息在内容、形式、时间等方面都有不同。在设计各级统计信息系统时，不仅要考虑国家统计信息系统对信息的总要求，还要考虑各级统计信息系统自身的要求。

也有人认为，统计信息系统是由国家统计信息系统、省级统计信息系统、地（市）级统计信息系统和县（区）统计信息系统构成。国家级与30个省、重点城市、国务院各部委统计专业司相连；省级除与国家级相连外，还要与地（市）级相连；地（市）级除与所属省级相连外，还要与县（区）级相连；县（区）级要与所属地（市）级相连。可见，这一系统是一个多层次、多地域、多部门的开放式互连系统。

通常把信息系统分为计算机数据处理系统、管理信息系统、决策支持系统和专家系统4种。所以，从现代统计信息系统开发来看，统计信息系统可分为：统计数据处理系统、统计管理信息系统、统计决策支持系统和统计专家系统4大类（如图1-3所示）。

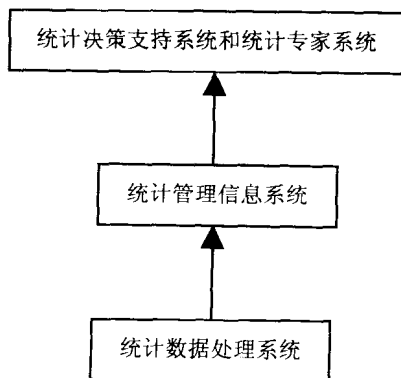


图 1-3 现代统计信息系统的分类

前两类系统属于结构化问题的系统，这两类系统建立和使用解决的是效率问题；后两类系统解决的是如何提高统计决策效果问题，这两类系统解决的是半结构化或非结构化问题。而且，一般认为统计信息系统是管理信息系统(MIS)的一个子系统，是决策支持系统(DSS)的一个重要组成部分。当然，它也可以以一个单独的统计信息系统形式出现。

以企业统计信息系统为例，企业的详细业务数据主要来自统计数据处理系统，它也为统计管理信息系统和统计决策支持系统提供基础数据。统计管理信息系统的主要目的是为企业各个业务领域的管理者提供信息支援，以便使他们能够更好地把握日常业务、对日常业务进行有效的管理。具体地说，它是负责向中层管理者提供事先定义好的、具有固定格式的业务报表的信息系统。它将相关的各种统计数据处理系统连接在一起，起着将事务处理数据库中的详细业务数据转换成管理者所需的综合管理信息的作用。统计决策支持系统的主要目的是帮助企业的高级管理者解决在经营决策时面临的特殊问题。与统计管理信息系统不同，统计决策支持系统不是基于面向业务操作的数据库，而是基于面向决策分析的数据仓库。它的作用是对数据和信息进行深层次的开发利用。而统计专家系统是人工智能应用的一个分支，是一种模拟人类专家解决领域问题的计算机程序系统。如果说决策支持系统是在大范围内的通用型支持，那么专家系统则是在某一专门领域的专用型支持。它不仅利用综合数据库，而且基于知识库。基于知识的系统可以扩展管理者解决问题的能力，促使企业的信息资源向知识资源转化。

总之，利用现代统计信息系统，不仅可以提高统计工作效率，而且还可以大大提高统计的决策效果。

2. 统计信息系统的特点

统计信息系统主要有如下几个特点。

- (1) 统计信息量大。统计信息系统是专门从事信息处理的系统，所汇集和提供的信息内容广泛，涉及社会经济活动的各个方面，具有综合性，因此信息量大。
- (2) 数据模型复杂。由于统计信息受多种因素影响，社会经济现象之间相互联系、相互制约，因此，要求所描述的数据模型必须结构化，避免不相容。
- (3) 及时性、准确性要求高。及时和准确是统计工作的生命，也是对统计信息系统的主要要求。即要求统计信息系统能及时提供准确可靠的信息。
- (4) 多路径检索。统计对信息的检索往往涉及信息的多个侧面，这就要求统计信息系统能够提供多路径检索功能。
- (5) 安全保护性好。统计信息系统必须具有对数据的安全性保护功能，这样就可以防止不合法的使用和修改。

3. 统计信息系统的作用

统计信息系统是统计工作现代化的保证，它代替传统的统计工作方式，在国民经济中产生了深远的影响。统计信息系统的主要功能包括以下几个方面。

- (1) 资料的收集与整理。系统按统计理论收集和积累资料，并合理组织。
- (2) 资料的统计与汇总。系统定期地、不定期地对历史资料进行加工汇总。
- (3) 生成统计报告。系统具有生成统计表格、统计图形等统计报告的功能。
- (4) 分析、预测与决策。系统具有由统计工作者直接使用的，与数据库自动联系的模型库、方法库及知识库，提供各种各样的手段进行分析、预测与决策。
- (5) 提供现代化的统计服务。系统提供统计服务主要包括查询服务、统计