

● 统计数据分析与应用丛书

基于EXCEL

的统计应用

薛 薇 陈欢歌 编著

● 统计数据分析与应用丛书

C819

49

2006

基于 EXCEL

的统计应用

薛 薇 陈欢歌 编著

 中国人民大学出版社

图书在版编目 (CIP) 数据

基于 SPSS 的数据分析/薛薇编著.

北京: 中国人民大学出版社, 2006

(统计数据分析与应用丛书)

ISBN 7-300-07659-9

I. 基…

II. 薛…

III. 统计分析-软件包, SPSS

IV. C819

中国版本图书馆 CIP 数据核字 (2006) 第 115529 号

统计数据分析与应用丛书

基于 SPSS 的数据分析

薛薇 编著

出版发行	中国人民大学出版社		
社 址	北京中关村大街 31 号	邮政编码	100080
电 话	010-62511242 (总编室)		010-62511398 (质管部)
	010-82501766 (邮购部)		010-62514148 (门市部)
	010-62515195 (发行公司)		010-62515275 (盗版举报)
网 址	http://www.crup.com.cn http://www.ttrnet.com (人大教研网)		
经 销	新华书店		
印 刷	北京鑫丰华彩印有限公司		
规 格	170 mm×228 mm 16 开本	版 次	2006 年 10 月第 1 版
印 张	25.25 插页 1	印 次	2006 年 10 月第 1 次印刷
字 数	407 000	定 价	35.00 元

版权所有 侵权必究 印装差错 负责调换

总序

Prologue

《管子》中有一句话：不明于计数，而欲举大事，犹无舟楫而欲经于水险也。意思是说在不清楚具体数据的情况下想做大事，就如同没有桨的船只航行在激流险滩之中。

对于国家政府来说，所谓大事就是引导社会经济沿着正确的方向健康和谐地发展；对于企事业单位来说，所谓大事就是在经营管理中做出科学有效的决策从而在激烈的市场竞争中生存发展；对于科学研究来说，所谓大事就是透过事物外在的表象深入探索其内在的规律性；对于个人生活来说，所谓大事就是在人生成长和个人理财的重要关口做出明智合理的抉择。

这些都离不开数据，离不开数据分析，离不开统计应用。通过数据处理进行科学的定量分析是成大事者的基本方法和首要条件。

当前，中国面临着国际化大背景下转轨的经济体制和转型的社会环境的综合考验，瞬息万变的信息时代对国家、企事业单位和个人都提出了严峻的挑战，而作为“信息密集”的统计应用也将迎来快速发展的机遇期。

这是我们撰写这套“统计数据分析与应用丛书”的初衷，希望它能够从事统计应用的实际工作者提供及时有效的帮助，也能够为有志于进行数据分析的在校本科生和研究生打开一扇统计应用的大门。



《论语》中有一句话：工欲善其事，必先利其器。意思是说做事情要取得较好的成就，应当首先利用先进的工具或手段。

突飞猛进的信息技术已经将统计应用引领到一个崭新的水平，并渗透到统计数据处理的各个环节，同时深刻影响着统计工作的全貌。计算机化和网络化是统计应用的必然趋势，它使得统计应用摆脱了复杂公式和计算的羁绊，可以为更多人所理解和使用，所以毫不夸张地说，现代信息技术是解决统计实际问题、掌握统计分析算法、建立统计应用系统的必要条件。

这是我们撰写这套丛书的另外一个初衷，就是让更多的统计应用读者能够使用计算机等信息技术实现统计数据分析与应用的目标，同时也可以让更多的计算机应用读者能够了解一些统计应用的特征和方法。同时，近年来在西方发达国家，信息技术人才和统计应用人才一直排名在就业需求榜的前列，我们认为具备统计知识和计算机知识的复合型应用人才在未来将具有巨大的发展前景和明显的从业优势。

在这套丛书的编写过程中，我们注重引进当前统计应用和计算机应用的前沿技术和理论方法，结合在统计应用、科研与教学中的丰富案例和实际经验，着力突出以下特点：

1. 针对性。针对不同的读者群，由浅入深地展开统计应用的论述，读者也可以按照《基于 EXCEL 的统计应用》、《基于 SPSS 的数据分析》和《基于信息技术的统计信息系统》所构成的应用体系，不断提高统计应用能力和统计理论水平。

2. 可操作性。以计算机和网络等信息技术为基础，实现统计应用的各种需求，读者可以按照丛书指引的策略和具体的方法解决自己工作生活中的数据处理问题。

3. 通俗性。以案例说明原理，以应用解释算法，以发展脉络说明理论形成，努力体现深入浅出的结构安排和文字风格。

不断改革前进是我们这个时代的主旋律，从目前国内相关论著和教材情况看，丛书的上述特点也是许多统计应用学者和同仁所追求的共同方向，在此特别感谢中国人民大学统计学院领导的鼎力支持，感谢中国人民大学出版社各位编辑所付出的辛勤劳动。

统计作为数据处理的方法论，具有广泛的应用领域，而它根本的生命力也在于应用。我们很想为相关读者奉献一套具有一定理论高度，且具有一定



指导性和实战性的统计应用书籍，它应该以统计应用案例为主线，以计算机技术为实现工具，可以使普通经营管理人员、基层科研人员、高层决策者、一般数据处理工作者和高年级的高校学生们从中获益。

薛薇

2006 年 10 月

于中国人民大学统计学院

前言

Preface

建立统计应用整体解决方案，掌握和普及统计分析方法是满足用户统计应用与统计分析需求的重要途径，而选择功能强大且简单易用的实施工具更是极为关键的。Excel 作为微软公司推出的 Office 套装软件中的主要成员，因其具有电子表格管理、数据清单管理、商业统计图表处理以及数据分析与决策功能，目前已受到广大用户的青睐，并逐步成为政府机构和企事业单位进行数据处理分析的主要工具。然而，应当指出的是，目前人们对 Excel 的应用大多还停留在一般的电子表格处理上。事实上，进一步开发 Excel 的统计应用功能，利用 Excel 建立面向统计应用的数据处理和分析整体解决方案，已成为一种必然趋势和迫切要求。

本书以统计应用的实际需求为线索。从建立统计应用的数据环境到统计数据的加工汇总，从统计数据的描述性分析到数据间的相关性研究，从数据推论到数据之间差异性剖析以及数据的预测，以案例的形式，简明而系统地介绍了 Excel 的主要统计应用功能。同时，以统计应用中极为普遍的问卷调查系统为例，介绍了利用 Excel 构建统计应用系统的框架、步骤和技术。

同时，由于统计本身具有一定的专业性，而且不同分析方法有不同的适用性，不了解方法的基本原理，就无法正确运用方法解决实际问题，就无法解读统计分析



结果所包含的实际意义。因此，本书立足于普通管理人员，以由浅入深的分析需求为线索，用极其通俗的语言和案例讲解了统计分析方法的基本思想和原理。

本书适合于初步掌握 Excel，并有采集数据、加工整理数据、分析数据等统计应用需求的读者。本书还可作为非统计学专业本科生和研究生的教材或选读书籍。

在本书第八章、第九章的编写过程中，我院的硕士研究生王敏、黄雅琴同学提供了一些案例和文字素材，本科生吴潇同学对全书进行了文字审阅，这里一并对她们表示衷心感谢。

请到人大经管图书在线 <http://www.rdjg.com.cn> 中的“课件下载”栏下载本套丛书的案例数据，也可到教育部应用统计科学研究统计数据库研究室 www.t1soft.com 中的“服务直通车”栏下载。本书疏漏之处，恳请读者斧正。

编者

2006 年 7 月

CONTENTS 目录

第 1 章 建立统计应用的数据环境	(1)
1.1 统计应用的数据环境	(2)
1.1.1 统计台账	(3)
1.1.2 统计指标	(4)
1.1.3 统计台账的基本组织形式：二维表	(4)
1.1.4 统计指标的基本组织形式：统计表	(6)
1.2 Excel 的数据组织和相关概念	(8)
1.2.1 工作簿、工作表、单元格	(8)
1.2.2 工作表的使用管理	(10)
1.3 利用 Excel 建立统计台账	(14)
1.3.1 不同类型数据的输入	(15)
1.3.2 数据的基本编辑	(20)
1.3.3 数据清单	(26)
1.4 利用 Excel 建立统计表	(28)
1.4.1 搭建统计表格框架	(28)
1.4.2 添加表格的说明信息	(31)
1.4.3 统计表格的保护	(33)
1.5 利用 Excel 进行统计分析的附加说明	(38)
第 2 章 加工统计数据	(39)
2.1 统计台账数据的派生	(40)
2.1.1 公式的输入	(40)
2.1.2 公式的复制	(45)
2.1.3 公式中名称的使用	(47)



2.1.4	公式中的函数	(51)
2.1.5	公式中的数组计算方式	(58)
2.2	统计台账数据的排序	(59)
2.3	统计台账数据的筛选查询	(61)
2.3.1	自动筛选	(62)
2.3.2	高级筛选	(63)
2.4	统计台账数据的统计	(65)
2.4.1	计数统计	(65)
2.4.2	汇总统计	(82)
2.5	统计台账数据的个性化输出	(86)
2.5.1	按工资条方式输出统计台账	(87)
2.5.2	按原始凭证方式输出统计台账	(88)
2.6	外部统计台账数据的导入	(91)
2.7	统计报表接收的自动化管理	(94)
2.7.1	建立宏	(95)
2.7.2	编辑宏	(97)
2.7.3	指定宏	(99)
2.7.4	运行宏	(102)
2.8	统计报表的汇总	(102)
2.9	利用 Excel 发布统计数据	(105)
2.9.1	利用 HTML 文档发布统计数据	(105)
2.9.2	统计数据的导航式发布	(107)
第 3 章	描述数据特征	(111)
3.1	刻画数据的集中趋势	(112)
3.1.1	平均数	(112)
3.1.2	中位数	(115)
3.1.3	众数	(116)
3.2	刻画数据的分布形状	(117)
3.2.1	频数分析	(117)
3.2.2	偏态系数	(128)
3.2.3	峰度系数	(130)



3.3	刻画数据的离散趋势	(131)
3.3.1	极差	(132)
3.3.2	四分位差	(133)
3.3.3	箱线图	(134)
3.3.4	标准差和方差	(138)
3.4	利用“数据分析”命令进行数据描述性分析	(139)
第4章	分析数据间的相关性	(143)
4.1	数据间的相关性	(143)
4.1.1	函数关系和统计关系	(143)
4.1.2	线性关系和非线性关系	(144)
4.1.3	正线性相关和负线性相关	(145)
4.2	分析两数值型数据间的相关性	(145)
4.2.1	图形分析方法：散点图	(145)
4.2.2	数值分析方法：简单相关系数	(149)
4.3	分析两品质型数据间的相关性	(151)
4.3.1	图形分析法：复式柱形图	(151)
4.3.2	数值分析法：列联分析	(152)
4.4	相关的可靠性检验	(155)
4.4.1	数值型数据间相关的可靠性检验	(156)
4.4.2	品质型数据间相关的可靠性检验	(158)
4.5	数据的线性回归分析	(162)
4.5.1	线性回归分析基础	(162)
4.5.2	确定回归分析中的自变量和因变量	(163)
4.5.3	根据收集到的数据确定自变量和因变量之间的 数学关系式	(163)
4.5.4	线性回归方程的统计检验	(165)
4.5.5	利用 Excel 的工具进行线性回归分析	(173)
4.5.6	利用回归方程预测	(175)
第5章	由数据到推论	(177)
5.1	数据与推论	(177)
5.1.1	用于推论的数据	(178)



5.1.2	推论什么	(179)
5.2	推论基础	(180)
5.2.1	总体分布和样本分布	(180)
5.2.2	抽样分布	(182)
5.2.3	如何推论	(184)
5.3	一个总体均值的推论	(186)
5.3.1	一个总体均值的推论：应用场合一	(187)
5.3.2	一个总体均值的推论：应用场合二	(188)
5.4	一个总体比例的推论	(191)
5.5	两个总体均值的比较	(193)
5.5.1	独立样本的均值检验	(194)
5.5.2	配对样本的均值检验	(197)
5.5.3	利用数据分析工具实现两总体均值的对比推断	(198)
第 6 章	分析数据间的差异性	(203)
6.1	数据间的差异性	(203)
6.1.1	研究数据间差异性的意义	(203)
6.1.2	方差分析基础	(204)
6.2	单因素方差分析	(205)
6.2.1	单因素方差分析的思路	(205)
6.2.2	利用 Excel 进行单因素方差分析	(208)
6.3	双因素方差分析	(210)
6.3.1	无交互作用的方差分析	(211)
6.3.2	利用 Excel 实现无交互作用的方差分析	(213)
6.3.3	有交互作用的方差分析	(215)
6.3.4	利用 Excel 实现有交互作用的方差分析	(219)
第 7 章	从数据看未来	(223)
7.1	如何从数据看未来	(223)
7.1.1	什么是时间序列	(223)
7.1.2	时间序列的类型	(225)
7.1.3	时间序列分析的主要任务	(227)
7.2	时间序列的预处理	(228)



7.2.1	数据准备	(228)
7.2.2	图形化观察数据	(229)
7.3	时间序列的趋势外推分析	(231)
7.3.1	线性趋势的外推	(232)
7.3.2	非线性趋势的外推	(233)
7.4	时间序列的移动平均分析	(240)
7.4.1	移动平均分析的基本思想	(240)
7.4.2	利用 Excel 实现简单移动平均分析	(242)
7.5	时间序列的指数平滑分析	(244)
7.5.1	指数平滑分析的基本思路	(245)
7.5.2	利用 Excel 实现指数平滑分析	(246)
7.6	时间序列的季节调整分析	(248)
7.6.1	季节调整分析的基本思想	(248)
7.6.2	用 Excel 实现季节调整分析	(249)
第 8 章	统计应用集成系统的开发工具	(255)
8.1	初步认识 Excel VBA	(255)
8.1.1	从“宏”认识 Excel VBA	(256)
8.1.2	认识 Excel VBA 的集成开发环境	(256)
8.2	对象、属性和事件	(261)
8.2.1	什么是对象、属性和事件	(261)
8.2.2	如何设置对象的属性值	(262)
8.2.3	如何编写对象的事件	(264)
8.3	Excel VBA 程序设计基础	(266)
8.3.1	Excel VBA 构成元素和基本词汇	(266)
8.3.2	数据类型	(267)
8.3.3	常量和变量	(267)
8.3.4	数组	(269)
8.4	Excel VBA 程序结构	(271)
8.4.1	顺序结构的程序	(272)
8.4.2	分支结构的程序	(276)
8.4.3	循环结构的程序	(280)



8.4.4	Excel VBA 的 Function 函数和 Sub 过程	(282)
8.5	常用的 Excel VBA 对象	(288)
8.5.1	Application 对象	(288)
8.5.2	Workbooks 对象	(290)
8.5.3	Worksheets 对象	(291)
8.5.4	Range 对象	(293)
8.5.5	Cells 对象	(294)
8.6	Excel VBA 宏的安全性	(295)
第 9 章	基于 Excel 的统计应用集成系统	(297)
9.1	利用 Excel VBA 实现用户窗口设计	(297)
9.1.1	插入用户窗口	(298)
9.1.2	Excel VBA 的控件	(298)
9.2	一个简单的系统示例	(301)
9.2.1	示例系统的功能	(301)
9.2.2	示例系统的数据录入	(301)
9.2.3	示例系统的数据浏览	(310)
9.2.4	示例系统的数据分析	(313)
9.3	利用 Excel VBA 开发统计应用集成系统的问题	(315)



建立统计应用的数据环境

中国企业正面临着市场化的挑战。在过去改革开放的 20 多年里，我国的经济体制由计划经济步入了社会主义市场经济的运行轨道，这一客观现实要求中国企业必须能够独立地面对市场经济的大风大浪。只有利用一切可能的条件和手段，及时获取各种千变万化的市场信息，准确理解信息中蕴涵的深层含义，并迅速制定出符合企业自身特点的经营战略，在动态的市场环境中不断快速调整竞争策略和方向，企业才有可能在残酷的市场竞争中处于不败之地，并不断发展。所以，吸收先进的管理理论，利用先进的信息技术和数量分析方法，“多、快、好、省”地获取市场信息，了解市场变化，对市场信息进行科学分析和科学决策，已经成为中国企业能否成为竞争“赢家”的关键。加入 WTO 后，中国企业在面临更广泛的客户选择的同时，也面临着更强大对手的竞争。在这种大趋势下，如果说工业化时代企业信奉“时间是金钱，效率是生命”的企业精神，那么，进入 21 世纪的信息化时代，企业则应树立“信息是金钱，决策是生命”的科学理念。

中国企业面临着客户需求个性化的挑战。一方面，随着企业生产技术和服务水平的普遍提高，产品的“同质化”特征越来越明显，市场竞争也已呈白热化趋势。此时，企业拓展生存空间、提升竞争力的有效途径是为客户提



供满足客户需要的、竞争对手无法“效仿”的个性化服务；同时，随着现代社会人们自我意识的不断加强和对个性化需求的不懈追求，对企业提供个性化服务的呼声也愈来愈高。所有这些都强烈地冲击着传统的企业经营模式，企业的经营战略正逐步由“以产品为中心”的卖方市场向“以客户为中心”的买方市场转移，从关心重量、数量、质量的工业化思路向满足个性化服务的方向转移。

市场化和个性化对企业的挑战，要求企业承担起信息化和定量分析的重任，只有信息化才能够迅速全面地感知市场化和个性化的要求，只有定量分析才能对所感知的信息做出科学分析和正确决策。

以数据库技术、管理信息系统等为代表的信息化技术，为收集、存储和管理大量的市场和客户数据，建立企业级的数据库管理信息系统打下了坚实基础。而以统计方法为主要代表的数量分析方法则是基于数据库基础之上，实现深层次定量分析并辅助企业级决策的有效工具。

建立统计应用整体解决方案，掌握和普及统计分析方法是满足用户上述需求的重要途径，而选择功能强大且简单易用的实施工具则是极为关键的环节。Excel 作为微软公司推出的 Office 套装软件中的主要成员，因其具有电子表格管理、数据清单管理、商业统计图表处理以及数据分析与决策功能，目前已受到广大用户的青睐，并逐步成为政府机构和企事业单位进行数据处理分析的主要工具。然而，应当指出的是，目前人们对 Excel 的开发应用大多还停留在一般的电子表格处理上。事实上，进一步开发 Excel 的统计应用功能，利用 Excel 建立面向统计应用的数据处理和分析整体解决方案，已成为一种必然趋势和迫切要求。

我们认为，统计应用整体解决方案是面向管理服务决策，以统计方法为工具，以计算机技术为依托，涉及数据采集与存储、数据传输与共享、数据整理与加工、数据分析与发布等诸多环节的综合应用系统。

§ 1.1 统计应用的数据环境

统计应用整体解决方案的核心是数据，统计分析的对象是数据。数据环



境的规范和建设是数据采集和存储的重要内容。针对企事业单位的统计分析需求,如何从体系上建立全面、科学、合理的统计指标体系和基础数据体系,从统计制度上规范统计数据的采集过程,从技术上选择面向统计分析的数据组织和管理策略,建立综合统计信息管理系统等,是统计应用数据环境建设中的重要内容,它们是获得高质量的统计数据,揭示统计数据的规律性,进一步挖掘统计数据深层内涵的重要前提和保障。

内容完备、结构合理的统计应用数据环境是企事业单位实现全面科学管理决策的核心。

统计应用数据环境一般包括统计台账和统计指标。统计台账是统计原始数据,统计指标是在统计台账基础上加工生成后的汇总数据。统计台账是统计指标的依据,统计指标是对统计台账的概括和总结。

1.1.1 统计台账

统计台账,又称统计基础数据,是企事业单位各种业务和各项管理的基本数据。

单位的人事档案、工资数据,学校学生各门课程的成绩,公司每种产品的每笔销售记录,保险公司的每张保单,医院每个病人的病历等都可以看做统计台账。此外,诸如为评估某产品或服务的市场占有情况而进行的问卷调查和客户满意度调查数据等,也可纳入统计台账的范畴。

统计台账涉及范围广,数据类型复杂,内容详尽,通常与企事业单位的业务和管理直接相关,因此,统计台账的存储管理往往被纳入企事业单位的管理信息系统中。以往,使用和关注统计台账的人员往往是企事业单位基层(业务层)的业务人员,他们通过对统计台账的操作,实现日常业务的管理,而企业中层或高层管理者则更多关注综合性强的统计指标,并通过它们制定各种管理决策。然而,随着管理的不断深入和决策正确性要求的不断提高,越来越多的中层或高层管理者的注意力不再仅仅停留于统计指标,也开始关注那些具有个案特征的统计台账,并尝试通过它们挖掘出更丰富的利于决策的信息。