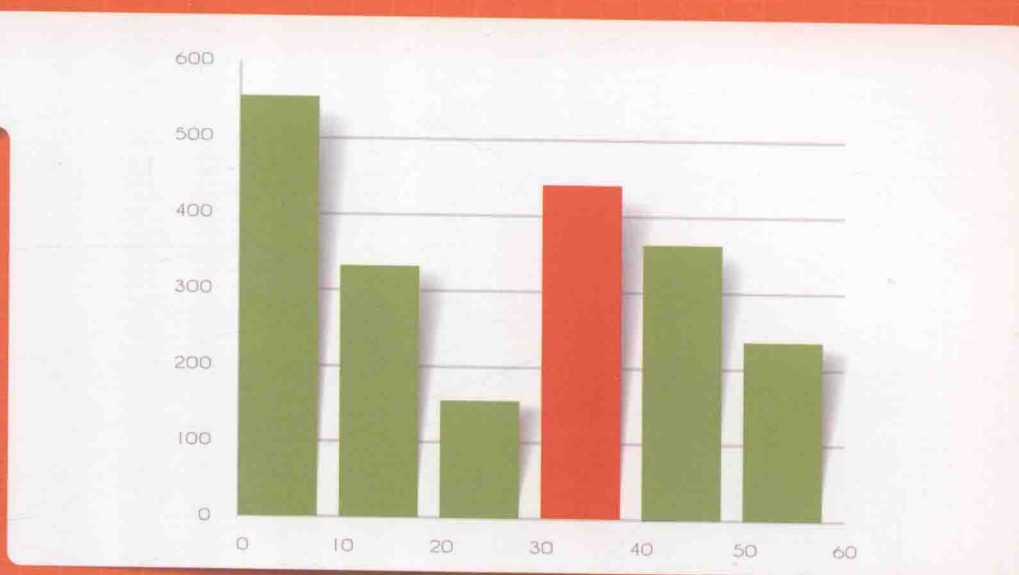




“十三五”普通高等教育规划教材



SPSS RUANJIAN
JIANMING CAOZUO JIAOCHENG

SPSS软件 简明操作教程

主编 薛伟

68%
DESCRIPTION



中国财经出版传媒集团
中国财政经济出版社

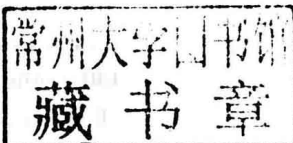


“十三五”普通高等教育规划教材

SPSS 软件简明操作教程

主 编 薛 伟

副主编 韩 存 杨秀艳



中国财经出版传媒集团
中国财政经济出版社

图书在版编目 (CIP) 数据

SPSS 软件简明操作教程/薛伟主编. —北京: 中国财政经济出版社, 2018. 7

“十三五”普通高等教育规划教材

ISBN 978-7-5095-8297-8

I. ①S… II. ①薛… III. ①统计分析—软件包—高等学校—教材 IV. ①C819

中国版本图书馆 CIP 数据核字 (2018) 第 121381 号

责任编辑: 蔡 宾 田明晖

责任校对: 李 静

封面设计: 肖玉坤

版式设计: 王志强

中国财经出版传媒集团 出版
中国财政经济出版社

URL: <http://www.cfeph.cn>

E-mail: jiaoyu@cfeph.cn

(版权所有 翻印必究)

社址: 北京市海淀区阜成路甲 28 号 邮政编码: 100142

营销中心电话: 010-82333010 编辑部门电话: 010-88190670

北京时捷印刷有限公司印刷 各地新华书店经销

787×1092 毫米 16 开 17 印张 424 000 字

2018 年 7 月第 1 版 2018 年 7 月北京第 1 次印刷

定价: 42.00 元

ISBN 978-7-5095-8297-8

(图书出现印装问题, 本社负责调换)

本社质量投诉电话: 010-88190744

打击盗版举报电话: 010-88190414、QQ: 447268889

前 言

SPSS 是“社会科学统计软件包”(Statistical Package for the Social Sciences)的简称,是一种集成化的计算机数据处理应用软件。于 2000 年正式更名为“统计产品与服务解决方案”(Statistical Product and Service Solutions),简称依然是 SPSS。SPSS 是世界上公认的三大数据分析软件之一(SAS、SPSS 和 SYSTAT)。SPSS 软件以其友好的界面、简便的操作、强大的功能和易学易掌握等特点,备受广大企业和个人用户的青睐,尤其是非专业的用户更加易学易懂。

本教材软件选用 IBM SPSS Statistics 21 中文版本,操作系统选用 Windows 7 进行编写。全书共 11 章,第 1 章到第 4 章为基础操作部分,第 5 章到第 11 章主要介绍各种统计分析方法。

本教材的特点突出简明性,即弱化了对方法理论的详细介绍,以实例直接切入操作和应用主题,方法理论可通过其他专业教材来深入学习。读者使用本教材可以在短时间内掌握 SPSS 软件的应用。

本教材由薛伟主编,韩存、杨秀艳为副主编,陈芳老师、李有才老师、硕士研究生李军参编。本教材力求深入浅出介绍 SPSS 软件的操作与方法分析,是一本上手较快的软件学习教程。

由于水平有限,加之时间仓促,有待改进的地方仍然很多,不妥之处在所难免。恳请广大读者批评指正,反馈意见请发电子邮件至:bestxuewei@sina.com。

编 者

2018 年 5 月

目 录

第 1 章 SPSS 软件概述	1
1.1 SPSS 的发展历程	1
1.2 SPSS 软件特点	2
1.3 SPSS 的运行环境	4
1.4 SPSS 安装和卸载	4
1.5 SPSS 启动与退出	8
1.6 SPSS 的三种运行管理方式	9
1.7 SPSS 软件常用窗口介绍	10
1.8 SPSS 帮助	13
1.9 常用系统设置对话框及其设置	14
第 2 章 SPSS 数据文件的建立与编辑	17
2.1 SPSS 变量属性定义	17
2.2 SPSS 变量属性浏览	22
2.3 SPSS 数据的录入	23
2.4 创建 SPSS 数据文件实例	25
2.5 SPSS 数据的编辑	27
2.6 读取数据文件	31
2.7 读取数据库文件	45
第 3 章 SPSS 数据整理	50
3.1 定义变量属性	50
3.2 复制数据属性	53
3.3 定义日期	55
3.4 标识重复个案	56
3.5 排序个案	59
3.6 转置	61
3.7 数据文件合并	62

3.8	分类汇总	71
3.9	拆分文件	73
3.10	选择个案	75
3.11	加权个案	81
第 4 章	SPSS 数据转换	86
4.1	SPSS 基本运算及表达式	86
4.2	SPSS 函数	88
4.3	计算变量	99
4.4	对个案内的值计数	104
4.5	重新编码	107
4.6	可视离散化	113
4.7	创建时间序列	117
4.8	替换缺失值	120
第 5 章	描述统计	123
5.1	基本统计量的定义和计算	123
5.2	频率分析过程	124
5.3	描述统计过程	129
5.4	探索分析过程	132
第 6 章	比较均值过程	142
6.1	比较均值基本原理	142
6.2	均值过程	143
6.3	单样本 T 检验过程	145
6.4	独立样本 T 检验过程	147
6.5	配对样本 T 检验过程	150
第 7 章	方差分析过程	153
7.1	方差分析基本原理	153
7.2	单因素方差分析过程	154
7.3	单变量多因素方差分析	159
第 8 章	非参数检验	168
8.1	卡方检验过程	168
8.2	二项检验过程	173
8.3	游程检验过程	176
8.4	单个样本的柯尔莫哥洛夫-斯米尔诺夫检验	179



8.5	两个独立样本的检验过程	181
8.6	多个独立样本检验过程	184
8.7	两个相关样本检验	186
8.8	多个相关样本检验	188
第 9 章	相关与回归分析	191
9.1	相关与回归分析基本原理	191
9.2	双变量相关分析	192
9.3	偏相关分析	195
9.4	线性回归分析	199
9.5	曲线回归过程	210
第 10 章	时间序列分析与预测	214
10.1	时间序列分析与预测基本原理	214
10.2	时间序列预分析	215
10.3	时间序列创建模型过程	216
10.4	季节性分解过程	236
第 11 章	多元统计分析	239
11.1	多元统计分析基本原理	239
11.2	系统聚类分析	239
11.3	判别分析	247
11.4	因子分析	255
参考文献		264

第1章

SPSS 软件概述

1.1 SPSS 的发展历程

SPSS 是“社会科学统计软件包”(Statistical Package for the Social Sciences)的简称,是一种集成化的计算机数据处理应用软件。SPSS 是世界上公认的三大数据分析软件之一(SAS、SPSS 和 SYSTAT)。SPSS 软件以其友好的界面、简便的操作、强大的功能和易学易掌握等特点,备受广大企业和个人用户的青睐,尤其使非专业的用户更加易学易懂。

1968 年,美国斯坦福大学三位大学生开发了最初的 SPSS 统计软件,并于 1975 年在芝加哥成立了 SPSS 公司。SPSS 软件广泛应用于通讯、医疗、银行、证券、保险、制造、商业、市场研究、科研、教育等多个领域和行业。

1994 至 1998 年间,SPSS 公司陆续购并了 SYSTAT 公司、BMDP 公司等,由原来单一统计产品开发转向为企业、教育科研及政府机构提供全面信息统计决策支持服务。伴随 SPSS 服务领域的扩大和深度的增加,2000 年正式更名为“统计产品与服务解决方案”(Statistical Product and Service Solutions),简称依然是 SPSS。

2009 年,SPSS 公司宣布重新包装旗下的 SPSS 产品,定位为预测统计分析软件(Predictive Analytics Software)PASW,包括四部分: PASW Statistics (formerly SPSS Statistics): 统计分析; PASW Modeler (formerly Clementine): 数据挖掘; Data Collection family (formerly Dimensions): 数据收集; PASW Collaboration and Deployment Services (formerly Predictive Enterprise Services): 企业应用服务。

2010 年,随着 SPSS 公司被 IBM 公司并购,各子产品名称前面不再以 PASW 为名,统一替换为 IBM SPSS 字样。

本教材选用 IBM SPSS Statistics 21 中文版本(以下简称 SPSS 21.0),简介方法理论,重点以实例为基础详细介绍 SPSS 软件的操作过程和输出结果。为用户尤其是非专业用户快速掌握 SPSS 软件提供便捷易懂的学习教材。本教材操作系统选用 Windows 7。

1.2 SPSS 软件特点

1.2.1 Windows 风格的友好界面

SPSS 最突出的特点就是操作界面极为友好,输出结果美观漂亮。SPSS 是第一个采用人机交互界面的统计软件,非常容易学习和使用。自从 1995 年 SPSS 公司与微软公司合作开发 SPSS 界面后,SPSS 界面变得越来越友好,操作也越来越简单,这就使熟悉微软公司产品的用户学习 SPSS 操作时,很容易上手。

SPSS 界面完全是菜单式,使用下拉菜单来选择所需要执行的复杂的统计命令,使用 Windows 的窗口方式展示各种管理和分析数据方法的功能,使用对话框展示出各种功能选择项,只要掌握一定的 Windows 操作技能,粗通统计分析原理,就可以使用该软件为特定的研究工作服务。

开放式的命令语句窗口,可以通过拷贝和粘贴的方法来学习和使用,尤其是每种方法对话框中的“粘贴”按钮,可将菜单操作过程直接转化成命令语句,用户只需在此基础上进行简单修改即可达到分析要求,大大节省了编写程序的时间,也不要牢记众多的命令代码。

1.2.2 易学易用的统计分析软件

SPSS 易于操作、易于入门,结果易于阅读,对统计软件的学习不会冲淡统计的主题,这样研究人员就可以将精力集中在社会研究方法、市场研究方法、营销的业务问题上,而不是忙于编程和统计。一般稍有统计基础的人经过几天的培训即可用 SPSS 做简单的数据分析,包括绘制图表、简单回归、相关分析等等。当然,真正应用好 SPSS 软件的关键在于如何进行科学地研究方案设计、严谨地收集数据、严密深入地对数据进行统计分析及解释,适度保守的下研究结论和进行决策。这一方面要求研究者掌握数理统计的基本知识,另一方面也要求研究者多进行实践,在实践中了解各种统计结果的实际意义。

从某种意义上讲,SPSS 软件还可以帮助数学功底不够的使用者学习运用现代统计技术。使用者仅需要关心某个问题应该采用何种统计方法,并初步掌握对计算结果的解释,而不需要了解其具体运算过程,就可以在使用手册的帮助下完成对数据定量分析。因此 SPSS 特别受非统计专业数据分析人员的青睐。

SPSS 采用类似 EXCEL 表格的方式输入与管理数据,数据接口较为通用,能方便地从其他数据库中读入数据。其统计方法包括了常用的、较为成熟的统计方法,完全可以满足非统计专业人士的工作需要,是非专业统计人员的首选统计软件。

1.2.3 功能全面的统计分析软件

SPSS 针对初学者、熟练者及精通者都比较适用。提供了数据获取、数据管理与准备、数据分析、结果报告这样一个数据分析的完整过程,因此,非常全面地涵盖了数据分析的整个流程,特别适合设计调查方案、对数据进行统计分析,以及制作研究报告中的相关图表。

此外,SPSS 具有完整的数据输入、编辑、统计分析、报表、图形制作等功能。仅 SPSS Base 模块就提供了从简单的统计描述到复杂的多因素统计分析方法,如数据的探索性分析、统计描述、列联表分析、二维相关、秩相关、偏相关、一元方差分析、非参数检验、多元回归、生存分析、协方差分析、判别分析、因子分析、聚类分析等常见的分析方法。

1.2.4 强大的编程能力,支持二次开发

对于常见的统计方法,SPSS 的命令语句、子命令及选择项的选择绝大部分由“对话框”的操作完成。因此,用户无需花大量时间记忆大量的命令、过程、选择项。

由于 SPSS 具备强大的 Syntax 编程功能,以及包括了 SPSS Programmability Extension 模块的编程扩展功能,那些熟练或精通者也较喜欢 SPSS,因为他们可以通过编程,在 SPSS 命令语法语言的基础上提供与其他编程语言的结合功能,来实现更强大的功能。例如,用其他语言编写的程序代码,如 Python 和 R,可以管理使用 SPSS 语法所编写的任务流。使用 SPSS 提供的扩展编程功能和特性,让 SPSS 成为了最强大的统计开发平台之一。

1.2.5 支持丰富的数据源,具备强大数据访问和管理能力

SPSS 可以同时打开多个数据集,方便研究时对不同数据库进行比较分析和进行数据库转换处理。软件提供了更强大的数据管理功能帮助用户通过 SPSS 使用其他的应用程序和数据库。能够读取及输出多种格式的文件。比如由 dBASE、FoxBASE、FoxPRO 产生的 *.dbf 文件,文本编辑器软件生成的 ASC II 数据文件,Excel 的 *.xls 文件等均可转换成可供分析的 SPSS 数据文件。支持 Excel、文本、Dbase、Access、SAS 等格式的数据文件。能够把 SPSS 的图形转换为 7 种图形文件。结果可保存为 *.txt,word,PPT 及 html 格式的文件。

此外,通过使用 ODBC(Open Database Capture)的数据接口,可以直接访问以结构化查询语言(SQL)为数据访问标准的数据库管理系统,通过数据库导出向导功能可以方便地将数据写入到数据库中等等。

从 SPSS 19.0 开始,增加了 Salesforce.com 数据库驱动程序,允许分析人员访问 Salesforce.com 中的数据,就像访问 SQL 数据库中的数据一样。

1.2.6 灵活的配置方案

对于商业用户而言,SPSS 是一种按照模块进行配置的软件产品,主要包括 SPSS Statistics Base 模块和其他一系列扩充功能模块。SPSS Statistics Base 是基础的软件平台,具备强大的数据管理能力和输入输出界面管理能力,并具备常见的统计分析功能,而其他每个独立扩充功能模块均在 SPSS Statistics Base 的基础上,为其增加某方面的分析功能家族。

1.2.7 支持多种操作系统

客户端支持 Windows XP(32 位)、Windows Vista and 7(32 位和 64 位)及以上系统、Linux 和 Mac OS。服务器端支持 Windows Server 2003(32 位和 64 位)、Windows Server 2008(32 位和 64 位)、AIX、HP-UX、Solaris。

1.3 SPSS 的运行环境

SPSS 21.0 对计算机的运行环境要求并不太高,但如果对于大数据量的分析,则需要较大的内存空间。

对硬件的要求:内存在 1GB 以上,硬盘应保留 800MB 左右的剩余空间。这对目前市面上的绝大部分计算机来说,完全能够保证。

对软件的要求:支持 Windows XP(32 位)、Windows Vista and 7(32 位和 64 位)及以上系统、Linux 和 Mac OS。

1.4 SPSS 安装和卸载

SPSS 21.0 的安装和卸载和 Windows 操作系统下的其他应用软件基本步骤相同。

1.4.1 软件安装

以下为主要安装步骤过程:

第一步,双击安装程序,弹出安装界面如图 1-1 所示。

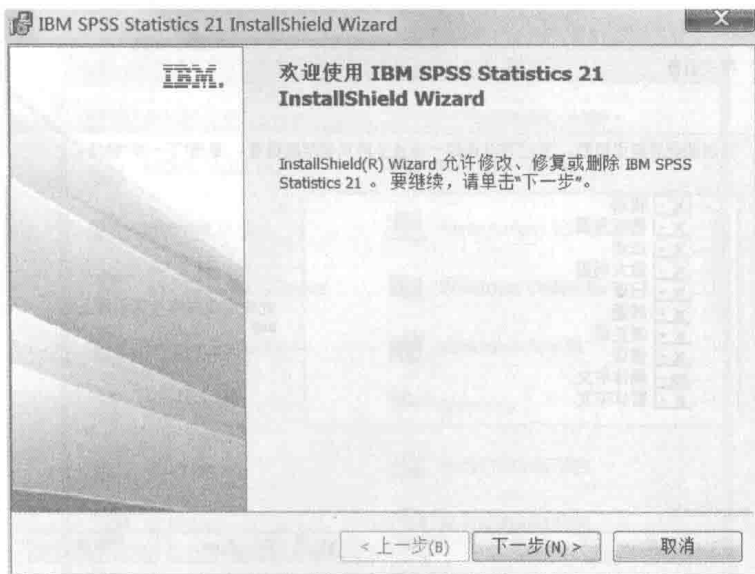


图 1-1 初始安装界面

第二步,按照安装向导提示,单击“下一步”按钮,弹出安装界面如图 1-2 所示,如果是初次安装,则单击“修改”单选按钮。

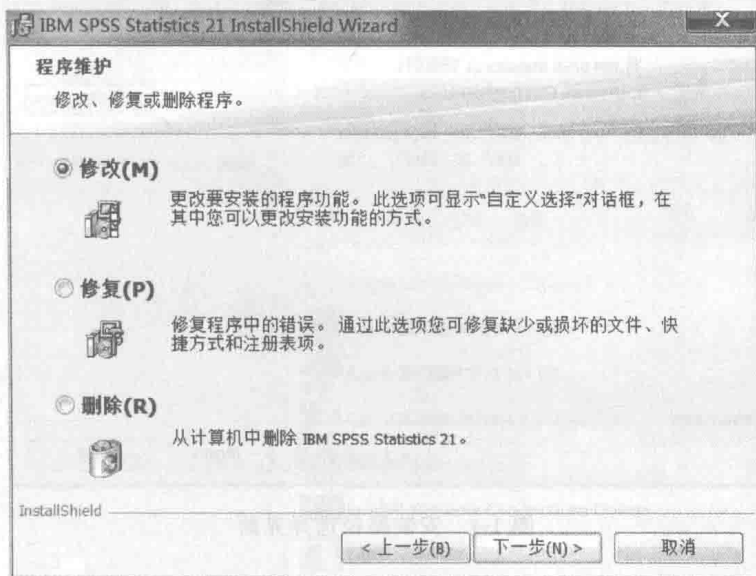


图 1-2 安装过程界面

第三步,按照安装向导提示,单击“下一步”按钮,弹出安装界面如图 1-3 所示,选择安装语言。

第四步,按照安装向导提示,单击“下一步”按钮,弹出安装界面如图 1-4 所示,选择安装目标文件夹路径。然后单击“下一步”按钮,完成整个安装即可。

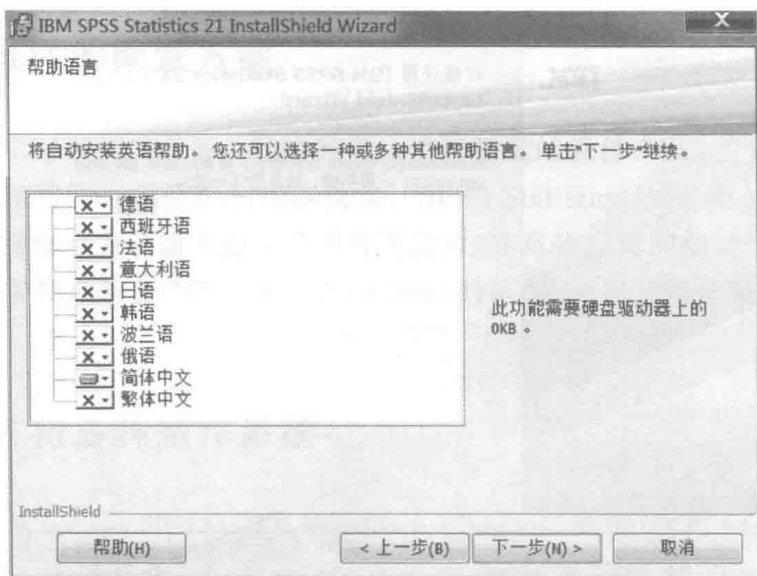


图 1-3 选择语言界面

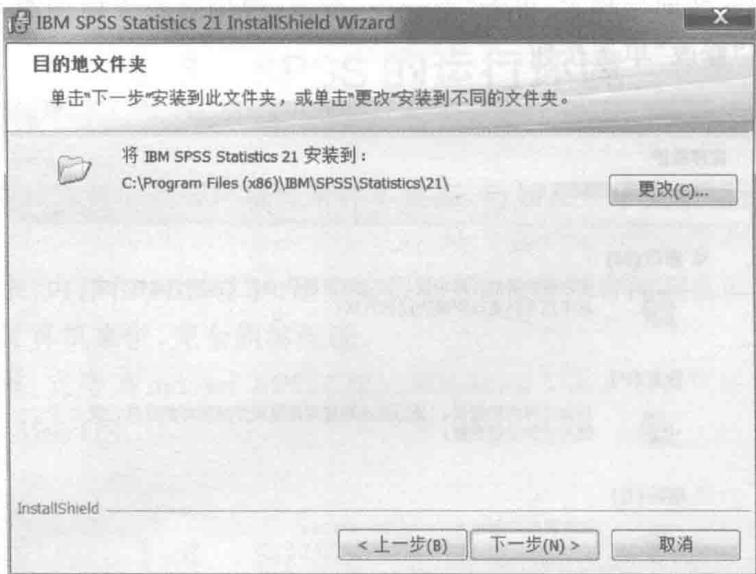


图 1-4 安装路径选择界面

1.4.2 软件卸载

第一步,在 Windows 系统中单击“开始”菜单,选取“控制面板”子菜单,如图 1-5 所示。
第二步,点开“程序和功能”对话框。如图 1-6 所示。



图 1-5 控制面板



图 1-6 程序和功能对话框

第三步,选中“IBM SPSS Statistics 21”并双击完成卸载即可。

1.5 SPSS 启动与退出

1.5.1 SPSS 启动

SPSS 21.0 安装完毕后,系统会自动在“开始”菜单中创建 IBM SPSS Statistics 21 快捷菜单。选择【开始】→【所有程序】→【IBM SPSS Statistics】→【IBM SPSS Statistics 21】启动 SPSS。如图 1-7 所示。

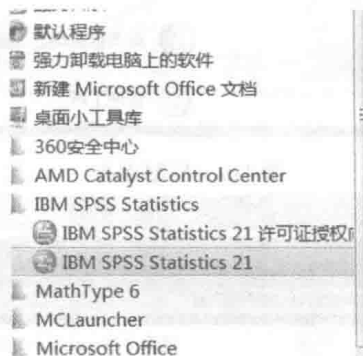


图 1-7 启动 SPSS 21.0

启动后,弹出如图 1-8 所示的初始界面。

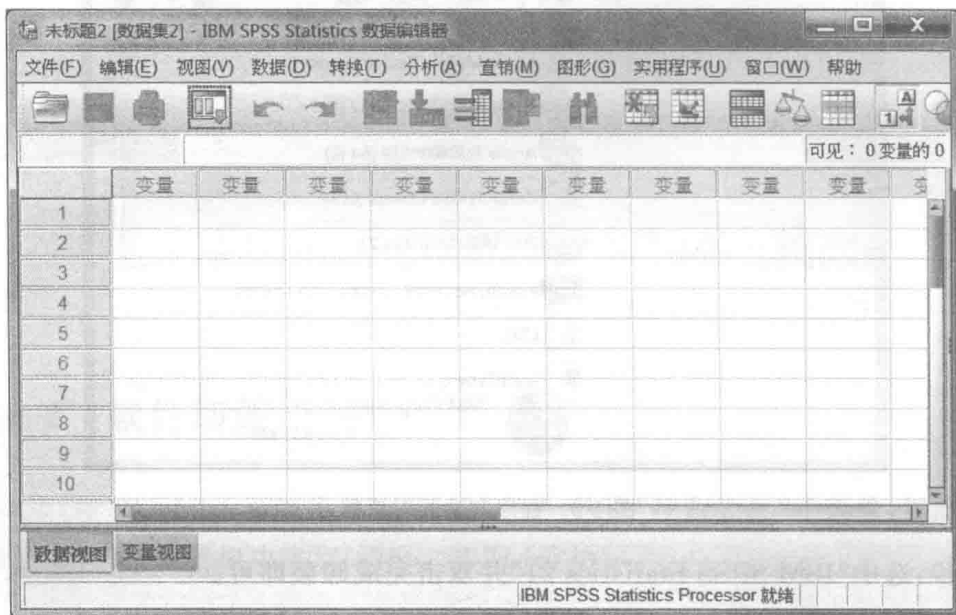


图 1-8 SPSS 初始界面

1.5.2 SPSS 退出

如图 1-8 所示,直接单击数据编辑窗口右上角“×”按钮,或者选择【文件】→【退出】路径菜单,即可退出 SPSS。如果本次打开了数据窗口或输出窗口,且修改后没有保存,则系统会自动提示是否继续关闭,待用户选择保存或放弃后,才能完全退出软件。

1.6 SPSS 的三种运行管理方式

SPSS 为用户提供了三种运行管理方式,即完全窗口菜单运行方式、程序运行方式和混合运行方式。用户可根据不同的分析需求选择使用。

1.6.1 完全窗口菜单运行方式

完全窗口菜单运行方式是用户在使用 SPSS 软件过程中,只需要通过菜单、按钮和输入对话框的方式实现数据录入和分析的整个操作过程。其特点是操作简单、直观,SPSS 中几乎所有的统计功能都可以通过菜单操作来完成。用户不需要记忆大量的统计功能命令和参数,就可以完成相应的统计分析工作。对于初学者和非专业用户来说,这是最容易掌握的一种运行管理方式。本教材也是主要在这种管理运行方式下介绍 SPSS 软件的应用。

1.6.2 程序运行方式

这种运行方式是在语句(Syntax)窗口中直接运行编写好的程序或者在脚本(Script)窗口中运行脚本程序的一种运行方式。这种模式要求掌握 SPSS 的语句或脚本语言。

程序运行方式适合专业用户的大规模和多步骤地统计分析工作,编好的程序还可以应用于其他数据文件。但是这种方式需要用户记忆大量的命令和参数,对于初学者或非专业用户较为困难。

1.6.3 混合运行方式

SPSS 软件为用户提供了结合完全窗口菜单运行方式和程序运行方式的混合运行方式。要实现混合运行方式,用户首先要按照菜单运行方式选择好统计分析的菜单和选项,并在统计分析对话框中单击“粘贴”按钮,SPSS 软件系统会自动打开一个语句窗口,将菜单选项转化成命令程序,供用户修改使用。这种方式大大提高了编程效率,使用户从记忆大量命令和参数中解脱出来。

1.7 SPSS 软件常用窗口介绍

SPSS 软件的常用窗口主要有四个,即数据编辑窗口、输出窗口、语句窗口和图形编辑窗口。对于初学者或非专业用户主要会用到数据编辑窗口、输出窗口和图形编辑窗口三个主要窗口。

1.7.1 数据编辑窗口

数据编辑窗口也是 SPSS 默认的启动用户界面,它是 SPSS 的基本工作台面,用户可以在这里建立、读取、编辑数据文件,选取设置各种所需的统计分析方法并运行。数据编辑窗口有两视图窗口,分别为数据视图和变量视图。如图 1-9 所示。



图 1-9 SPSS 数据编辑窗口

- (1)标题栏:显示数据编辑的数据文件名。
- (2)菜单栏:通过对这些菜单的选择,用户可以进行几乎所有的 SPSS 操作。关于菜单的详细的操作步骤将在后续内容中分别介绍。
- (3)工具栏:为了方便用户操作,SPSS 软件把菜单项中常用的命令放到了工具栏里。当鼠标停留在某个工具栏按钮上时,会自动跳出一个文本框,提示当前按钮的功能。另外,如果用户对系统预设的工具栏设置不满意,也可以用【视图】→【工具栏】→【设定】命令对工具栏按钮进行定义。
- (4)编辑栏:可以输入数据,以使它显示在内容区指定的方格里。
- (5)数据编辑区:此区域可以录入编辑数据。数据编辑区列向为变量,顶端显示变量