



视频讲解

技能应用速成系列

SPSS技能速成，本书足矣！

升级版

SPSS 22.0

统计分析 | 从入门到精通

内容全面、案例丰富

基础操作→专题技能→实例应用→举一反三。

讲解细致、综合应用

通过案例详细讲解分析流程、方法、技巧、注意事项，再到综合应用。

视频教学、网络服务

案例视频教学、超值素材资源、网络支持。

网络服务: <http://www.catics.org/3117>

博客答疑: <http://blog.sina.com.cn/caxbook>

李 昕 张明明 编著



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

提供各种书籍的pd电子版代找服务，如果你找不到自己想要的书的pdf电子版，我们可以帮您找到，如有需要，请联系QQ1779903665.

PDF代找说明：

本人可以帮助你找到你要的PDF电子书，计算机类，文学，艺术，设计，医学，理学，经济，金融，等等。质量都很清晰，而且每本100%都带书签索引和目录，方便读者阅读观看，只要您提供给我书的相关信息，一般我都能找到，如果您有需求，请联系我QQ1779903665。

本人已经帮助了上万人找到了他们需要的PDF，其实网上有很多PDF,大家如果在网上不到的话，可以联系我QQ，大部分我都可以找到，而且每本100%带书签索引目录。因PDF电子书都有版权，请不要随意传播，如果您有经济购买能力，请尽量购买正版。

声明：本人只提供代找服务，每本100%索引书签和目录，因寻找pdf电子书有一定难度，仅收取代找费用。如因PDF产生的版权纠纷，与本人无关，我们仅仅只是帮助你寻找到你要的pdf而已。

技能应用速成系列

Moldflow 2015模流分析从入门到精通
ISBN 978-7-121-25537-3

UG NX 9.0模具设计从入门到精通
ISBN 978-7-121-25538-0

ANSYS Workbench 15.0有限元分析从入门到精通
ISBN 978-7-121-25539-7

MATLAB R2014a 从入门到精通
ISBN 978-7-121-25540-3

TArch 2014天正建筑设计从入门到精通
ISBN 978-7-121-25541-0

UG NX 9.0中文版从入门到精通
ISBN 978-7-121-25542-7

UG NX 9.0造型设计从入门到精通
ISBN 978-7-121-25543-4

UG NX 9.0数控加工从入门到精通
ISBN 978-7-121-25544-1

CATIA V5-6 R2014从入门到精通
ISBN 978-7-121-25545-8

ANSYS Fluent 15.0流体计算从入门到精通
ISBN 978-7-121-25546-5

SPSS 22.0统计分析从入门到精通
ISBN 978-7-121-25551-9

UG NX 9.0机械设计从入门到精通
ISBN 978-7-121-25552-6

AutoCAD 2014电气设计从入门到精通
ISBN 978-7-121-21284-0

AutoCAD 2014机械设计从入门到精通
ISBN 978-7-121-21261-1

AutoCAD 2014建筑水暖电设计从入门到精通
ISBN 978-7-121-21260-4

AutoCAD 2014建筑设计从入门到精通
ISBN 978-7-121-21259-8

AutoCAD 2014室内设计从入门到精通
ISBN 978-7-121-21222-2

AutoCAD 2014中文版从入门到精通
ISBN 978-7-121-21221-5



扫描二维码
直通图书网服小组

本书小组编号: 3117
www.catics.org/3117



请刮开获取网服VIP码

书友交流 编读互动 视频解析 勘误更新



策划编辑: 许存权
责任编辑: 许存权
封面设计: 朝天世纪

ISBN 978-7-121-25551-9



定价: 79.00元
(含DVD光盘1张)

技能应用速成系列

SPSS 22.0 统计分析从入门到精通

李 昕 张明明 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书以 SPSS 22.0 为平台,由浅入深地全面讲解 SPSS 软件的相关知识,通过图文并茂的方式讲解各项操作,深入浅出,实例引导,讲解翔实,清晰、直观、易学易用。

全书分为三部分共 19 章,详细介绍 SPSS 的界面、数据文件的编辑、数据文件的整理、基本统计分析、参数估计与假设检验、非参数检验、方差分析、相关分析、回归分析、聚类分析、判别分析、因子分析、对应分析、信度分析、生存分析、时间序列分析、统计图形的绘制、SPSS 在企业经济活动和房地产中的应用等内容。本书涉及面广,涵盖了一般用户需要使用的各种功能,全书按逻辑顺序编排,自始至终结合实例进行描述,内容完整且每章相对独立,是一本详细实用的 SPSS 参考书。

本书适合高等院校统计分析专业的学生、科研人员、SPSS 用户和爱好者,以及希望从事 SPSS 软件技术相关工作的人员学习。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

SPSS 22.0 统计分析从入门到精通 / 李昕, 张明明编著. —北京: 电子工业出版社, 2015.3
(技能应用速成系列)

ISBN 978-7-121-25551-9

I. ①S… II. ①李… ②张… III. ①统计分析—软件包 IV. ①C819

中国版本图书馆 CIP 数据核字(2015)第 030948 号

策划编辑: 许存权

责任编辑: 许存权 特约编辑: 马军令 冯彩茹

印 刷: 北京京师印务有限公司

装 订: 北京京师印务有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1 092 1/16 印张: 31 字数: 794 千字

版 次: 2015 年 3 月第 1 版

印 次: 2015 年 3 月第 1 次印刷

定 价: 79.00 元(含 DVD 光盘 1 张)

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。



前言

SPSS 是世界上最早的统计分析软件，以功能丰富、效率高、操作简便而著称，是非常适合进行数据分析的工具软件，在经济学、医学、教育学、管理学、心理学、广告学、统计学及商业、工业、林业、农业等各个领域有广泛应用。

本书基于 SPSS 最新产品的 SPSS 22.0 版本编写，该版本在界面设置、数据管理、报表和图标、编程能力等方面有很大改进和提高。本书采用“完全案例”的编写形式，与相关操作技巧结合紧密，与全书设计理念和创作构思相辅相成，专业性、层次性、技巧性等特点的组合搭配，使该书的实用价值达到了一个很高的层次。

1. 本书特点

本书由浅入深，适合各水平阶段的读者学习。书中结合作者的实际项目经验，讲解时穿插了大量应用技巧和实例。本书主要有如下特点：

循序渐进、通俗易懂。本书完全按照初学者的学习规律和习惯，由浅入深，由易到难安排每章节内容，可以让初学者在学习掌握 SPSS 的所有基础知识及其应用。

步骤详尽、内容新颖。本书结合作者多年 SPSS 的使用经验与实际工作应用案例，将 SPSS 软件的使用方法与技巧详细地讲解给读者。本书在讲解过程中步骤详尽、内容新颖，讲解过程辅以相应的图片，使读者一目了然，从而快速掌握书中所讲内容。

内容全面、结构合理。本书涉及基本统计分析、参数估计与假设检验、非参数检验、方差分析、相关分析、回归分析、聚类分析、判别分析各个统计分析方法的使用大全，思路清晰、内容丰富。

案例丰富、技术全面。本书的每一章都是 SPSS 的一个专题，每一个案例都包含多个知识点，读者对照本书进行学习，同时可以举一反三，达到入门并精通的目的。

视频教学、轻松易懂。本书配备了高清语音教学视频，作者手把手地精心讲解，并进行相关点拨，使读者领悟并掌握每个案例的操作难点，轻松掌握并且提高学习效率。

2. 本书内容

本书基于 SPSS 22.0 软件版本，讲解了 SPSS 的基础知识和综合应用，主要分为三个部分，即基础知识、数据分析和综合应用部分，其中基础知识部分包括第 1~3 章，数据分析部分包括第 4~17 章，综合应用部分包括第 18~19 章。

第一部分：基础知识。从 SPSS 的发展简史、数据类型、数据编辑、数据排序等方面入手，使读者掌握 SPSS 的基本知识及数据分析前的数据整理方法，为其后的数据分析打下基础。

第一部分 基础知识

第 1 章 SPSS 22.0 概述

第 2 章 数据文件操作

第 3 章 数据文件整理

第二部分：数据分析。全面讲解 SPSS 在统计分析中的应用，该部分以实用为目标，通过简明扼要的讲解，并实例引导，使读者全面掌握各种统计方法的操作。

第 4 章 SPSS 基本统计分析

第 5 章 参数检验

第 6 章 非参数检验

第 7 章 方差分析

第 8 章 相关分析

第 9 章 回归分析

第 10 章 聚类分析

第 11 章 判别分析

第 12 章 因子分析

第 13 章 对应分析

第 14 章 尺度分析

第 15 章 生存分析

第 16 章 时间序列分析

第 17 章 统计图形的绘制

第三部分：综合应用。该部分以实例作为导向，通过全面的数据分析，培养读者的统计分析思想，进一步提高对 SPSS 的理解。

第 18 章 SPSS 在企业经济活动中的应用

第 19 章 SPSS 在房地产市场中的应用

3. 光盘内容

本书附带 1 张 DVD 多媒体视频演示光盘，本书所有范例用到的数据文件等素材都收录在光盘中，光盘内容结构如下：

“/数据文件/”目录：书中所使用到的数据文件收录在光盘的该文件夹中。

“/视频文件/”目录：书中所有案例的多媒体教学文件，按章放置在光盘的该文件夹中，方便读者学习。

4. 读者对象

本书适合 SPSS 初学者和统计分析从业人员学习，具体包含如下：

★ 相关数据分析从业人员

★ 初学 SPSS 的技术人员

★ 大中专院校的师生

★ 相关培训机构的教师和学员

★ 参加工作实习的“菜鸟”

★ SPSS 爱好者

★ 广大科研工作人员

★ 初中级 SPSS 用户

5. 本书作者

本书主要由李昕、张明明编写，另外，参与编写的人员还有：吴光中、魏鑫、石良臣、刘冰、林晓阳、唐家鹏、丁金滨、王菁、吴永福、张小勇、刘成柱、乔建军、张迪妮、张岩、温光英、温正、郭海霞、王芳。虽然作者在编写过程中力求叙述准确、完善，但由于水平有限，书中欠妥之处，请读者及各位同行批评指正，在此表示诚挚的谢意。

6. 读者服务

为了方便解决本书疑难问题，读者在学习过程中如遇到与本书有关的技术问题，可以发邮件到邮箱 caxbook@126.com，或访问作者博客 <http://blog.sina.com.cn/caxbook>，作者会尽快给予解答，我们将竭诚为您服务。

编著者



目录

第一部分 基础知识

第 1 章 SPSS 22.0 概述	1	第 3 章 数据文件整理	29
1.1 SPSS 简介	2	3.1 数据排序	30
1.2 SPSS 基本操作	2	3.2 数据转置	33
1.3 SPSS 窗口及其功能	6	3.3 数据文件的合并	35
1.4 SPSS 帮助系统	9	3.4 数据文件的结构重组	41
1.5 本章小结	10	3.5 分类汇总	47
第 2 章 数据文件操作	11	3.6 文件拆分	49
2.1 数据类型及定义	12	3.7 选择个案	51
2.2 获取数据与查看数据的方法	16	3.8 个案加权	55
2.3 数据的编辑	26	3.9 计算新变量	57
2.4 数据的输出	27	3.10 对个案内的值计数	59
2.5 本章小结	28	3.11 变量的重新编码	62
		3.12 个案等级排秩	70
		3.13 本章小结	72

第二部分 数据分析

第 4 章 SPSS 基本统计分析	73	5.4 两独立样本的 t 检验	103
4.1 频数分析	74	5.5 配对样本的 t 检验	106
4.2 描述性统计分析	79	5.6 本章小结	108
4.3 探索性分析	81	第 6 章 非参数检验	109
4.4 列联表分析	89	6.1 卡方检验	110
4.5 本章小结	96	6.2 二项分布检验	113
第 5 章 参数检验	97	6.3 游程检验	115
5.1 假设检验	98	6.4 单样本 K-S 检验	117
5.2 平均值检验	99	6.5 两独立样本的非参数检验	120
5.3 单样本 t 检验	102	6.6 多个独立样本的非参数检验	124

6.7 两配对样本检验	128	第 11 章 判别分析	280
6.8 多个配对样本检验	132	11.1 判别分析的概述	281
6.9 本章小结	137	11.2 一般判别分析	283
第 7 章 方差分析	138	11.3 逐步判别分析	291
7.1 方差分析的概述	139	11.4 本章小结	298
7.2 单因素方差分析	140	第 12 章 因子分析	299
7.3 多因素方差分析	147	12.1 因子分析的概述	300
7.4 协方差分析	156	12.2 因子分析的实例分析	302
7.5 多元方差分析	159	12.3 本章小结	312
7.6 重复测量方差分析	163	第 13 章 对应分析	313
7.7 本章小结	169	13.1 对应分析的基本原理	314
第 8 章 相关分析	170	13.2 简单对应分析	315
8.1 相关系数	171	13.3 多元对应分析	322
8.2 两变量相关分析	172	13.4 本章小结	334
8.3 偏相关分析	176	第 14 章 尺度分析	335
8.4 距离分析	178	14.1 信度分析	336
8.5 本章小结	185	14.2 多维尺度分析	341
第 9 章 回归分析	186	14.3 本章小结	349
9.1 线性回归	187	第 15 章 生存分析	350
9.2 曲线回归	203	15.1 生存分析的概述	351
9.3 非线性回归	208	15.2 寿命表分析	353
9.4 二元 Logistic 回归	216	15.3 Kaplan-Meier 分析	358
9.5 多元 Logistic 回归	226	15.4 Cox 回归模型	364
9.6 有序回归	236	15.5 本章小结	375
9.7 概率单位回归	242	第 16 章 时间序列分析	376
9.8 加权回归分析	251	16.1 时间序列的预处理	377
9.9 本章小结	256	16.2 序列图	381
第 10 章 聚类分析	257	16.3 周期性分解	385
10.1 快速聚类	258	16.4 谱分析	389
10.2 系统聚类	264	16.5 自相关	391
10.3 两步聚类	273	16.6 创建时间模型	397
10.4 本章小结	279		

16.7	应用时间序列模型	409	17.7	箱图	440
16.8	交叉相关性	411	17.8	误差条形图	442
16.9	本章小结	414	17.9	人口金字塔图	445
第 17 章	统计图形的绘制	415	17.10	散点图	446
17.1	图形构建器的基本操作	416	17.11	直方图	451
17.2	条形图	418	17.12	P-P 概率图和 Q-Q 概率图	452
17.3	折线图	425	17.13	比较子组图	454
17.4	面积图	427	17.14	回归变量图	458
17.5	饼图	430	17.15	本章小结	461
17.6	高低图	433			
第三部分 综合应用					
第 18 章	SPSS 在企业经济活动中的应用	462	第 19 章	SPSS 在房地产市场中的应用	479
18.1	背景介绍	463	19.1	背景介绍	480
18.2	经济含义简介	464	19.2	相关分析的应用	480
18.3	主成分分析的应用	465	19.3	回归分析的应用	483
18.4	聚类分析的应用	470	19.4	本章小结	487
18.5	本章小结	478	参考文献		488

第一部分 基础知识

第1章

SPSS 22.0 概述

SPSS 是世界上应用最早的统计分析软件，目前已广泛应用于多个领域，如经济学、医学、教育学等。本章主要介绍 SPSS 的发展简史、安装与启动、退出与卸载、窗口及其功能及帮助系统。

学习目标

- (1) 了解 SPSS 的发展简史。
- (2) 掌握 SPSS 22.0 的安装与启动、退出与卸载。
- (3) 熟悉主界面的各项菜单。
- (4) 学会使用帮助系统。



Note

1.1 SPSS 简介

SPSS 是世界上应用最早的统计分析软件，由美国斯坦福大学的三位研究生 Norman H. Nie、C. Hadlai (Tex) Hull 和 Dale H. Bent 于 1968 年研究开发而成。当时世界上许多有影响的报刊杂志纷纷就 SPSS 的自动统计绘图、数据深入分析、使用方便、功能齐全等方面给予了高度的评价。

2009 年 7 月 28 日，IBM 公司宣布用 12 亿美元现金收购统计分析软件提供商 SPSS 公司，并更名为 IBM SPSS，如今 SPSS 已升级至版本 22.0。SPSS 是世界上最早采用图形菜单驱动界面的统计软件，它最突出的特点就是操作界面极为友好，输出结果美观漂亮。

Windows 的窗口方式展示各种管理和分析数据方法的功能，对话框展示出各种功能选择项。SPSS 采用类似 EXCEL 表格的方式输入管理数据，数据接口较为通用，能方便地从其他数据库中读入数据。

SPSS 是一个组合式软件包，它集数据录入、整理、分析功能于一身。用户可以根据实际需要和计算机的功能选择模块，以降低对系统硬盘容量的要求，有利于该软件的推广应用。SPSS 的基本功能包括数据管理、统计分析、图表分析、输出管理等。

SPSS 统计分析过程包括描述性统计、均值比较、一般线性模型、相关分析、回归分析、对数线性模型、聚类分析、数据简化、生存分析、时间序列分析、多重响应等几大类，每类又分多个统计过程，比如回归分析中又分线性回归分析、曲线估计、Logistic 回归、Probit 回归、加权估计、两阶段最小二乘法、非线性回归等多个统计过程，而且每个过程又允许用户选择不同的方法和参数。SPSS 也有专门的绘图系统，可以根据数据绘制各种图形。

SPSS 的分析结果清晰、直观、易用，而且可以直接读取 EXCEL 及 DBF 数据文件，现已推广到多种操作系统的计算机上，它和 SAS、BMDP 并称为国际上最有影响的三大统计软件。

SPSS 由于其操作简单，已经在我国的社会科学、自然科学的各个领域发挥着巨大作用。该软件还应用于经济学、数学、统计学、物流管理、生物学、心理学、地理学、医疗卫生、体育、农业、林业、商业等多个领域。

1.2 SPSS 基本操作

下面介绍 SPSS 22.0 的安装、启动、退出和卸载的操作方法。

1.2.1 安装 SPSS

在安装 SPSS 之前，首先要查看系统配置是否满足软件包的要求，比如硬盘剩余空

间、操作系统等方面。不同版本的 SPSS 有不同的安装过程,下面以 SPSS 22.0 为例介绍 SPSS 的安装过程。

(1) 启动计算机,将装有 SPSS 22.0 安装软件的光盘放入光驱。

(2) 打开资源管理器,双击光盘中的 setup.exe 文件,计算机便启动了 SPSS 22.0 的安装程序,自动运行后弹出如图 1-1 所示的界面。

(3) 单击 **下一步(N) >** 按钮,弹出如图 1-2 所示的界面,提示选择“单个用户许可证/网络许可证”,一般用户选择“单个用户许可证”即可。

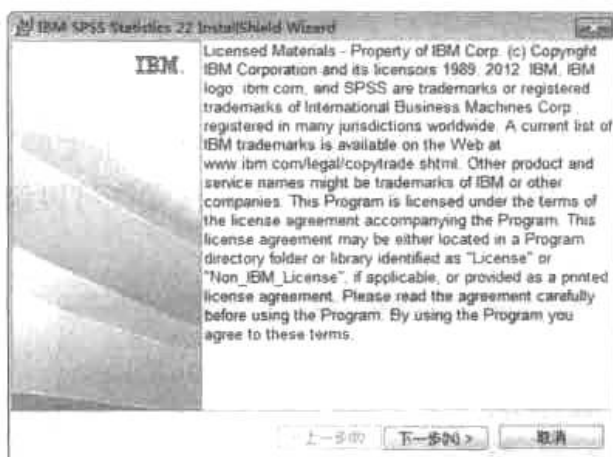


图 1-1 安装界面 1

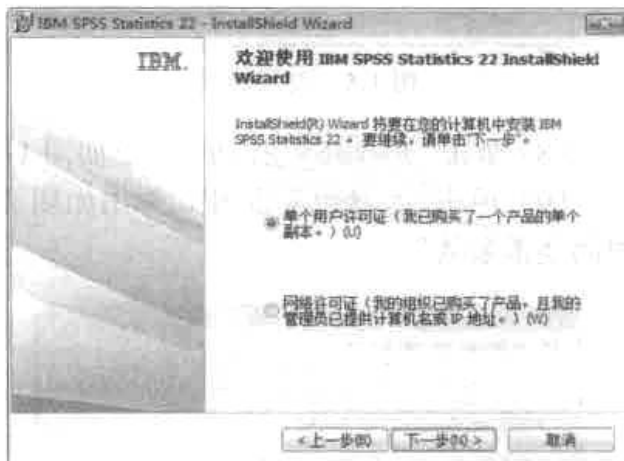


图 1-2 安装界面 2

(4) 单击 **下一步(N) >** 按钮,弹出如图 1-3 所示的界面,选择“我接受许可协议中的全部条款”。

(5) 单击 **下一步(N) >** 按钮,弹出如图 1-4 所示的界面,提示输入“客户信息”,采用默认信息即可。

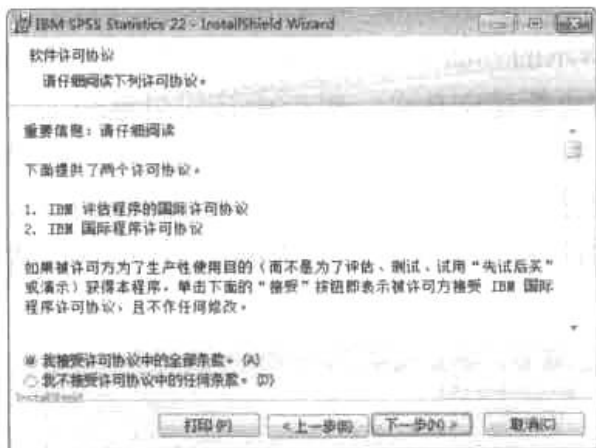


图 1-3 安装界面 3



图 1-4 安装界面 4

(6) 单击 **下一步(N) >** 按钮,弹出如图 1-5 所示的界面,提示选择“帮助语言”。用户根据需要进行选择,本例选中“简体中文”。

(7) 单击 **下一步(N) >** 按钮,弹出如图 1-6 所示的界面,该界面是为用户提供阅读软件,用户可根据需要进行选择。



Note



Note



图 1-5 安装界面 5

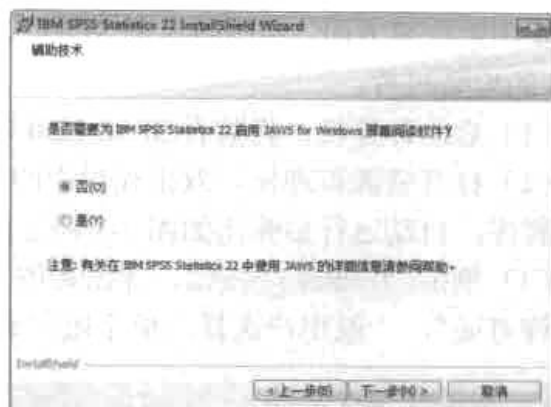


图 1-6 安装界面 6

(8) 单击 **下一步(N) >** 按钮，弹出如图 1-7 所示的界面，选择“是”进行 SPSS 的安装。

(9) 单击 **下一步(N) >** 按钮，弹出如图 1-8 所示的界面，单击选择“我接受许可协议中的全部条款”。

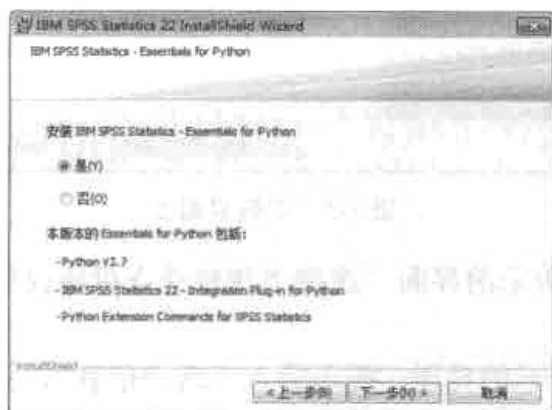


图 1-7 安装界面 7

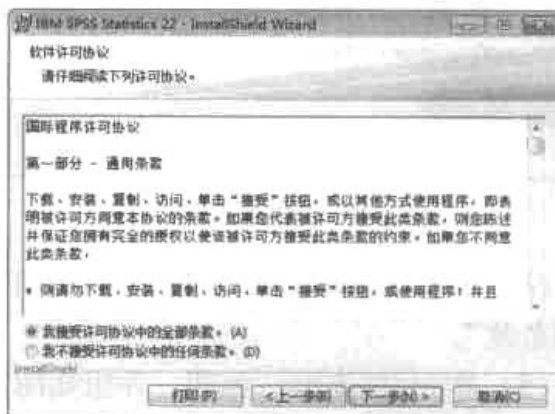


图 1-8 安装界面 8

(10) 单击 **下一步(N) >** 按钮，弹出如图 1-9 所示的界面。单击 **更改(C)...** 按钮，弹出更改安装路径的对话框（图 1-10），单击 **确定** 按钮可新建文件夹，单击 **返回** 按钮可返回上一级存储路径，单击 **确定** 按钮返回主界面；单击 **Available Space** 按钮，弹出如图 1-11 所示的对话框，在此可查看磁盘空间等信息，单击 **确定** 按钮返回主界面。

(11) 单击 **下一步(N) >** 按钮，弹出如图 1-12 所示的界面，单击 **安装(I)** 按钮执行软件安装。



图 1-9 安装界面 9



图 1-10 安装界面 10



Note



图 1-11 安装界面 11

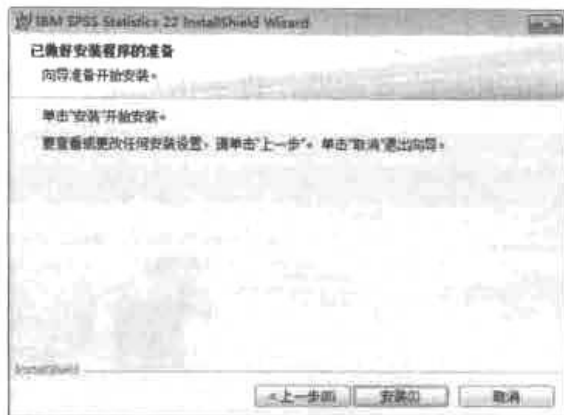


图 1-12 安装界面 12

(12) 安装完成后, 弹出如图 1-13 所示的界面, 若用户有授权代码, 则选择“立即授予产品许可证”; 反之, 选择“启用以用于临时使用”。单击 **下一步** 按钮, 弹出如图 1-14 所示界面, 单击 **确定** 按钮, 完成授权过程。

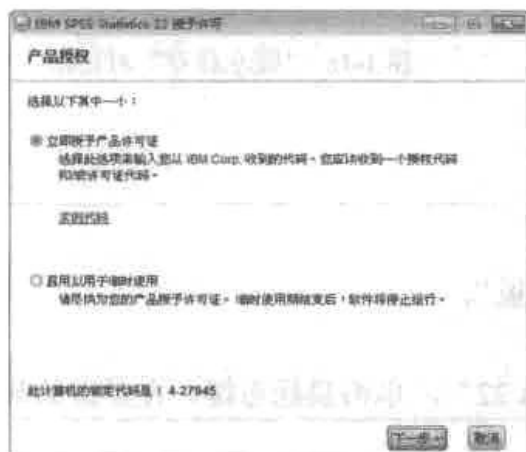


图 1-13 安装界面 13

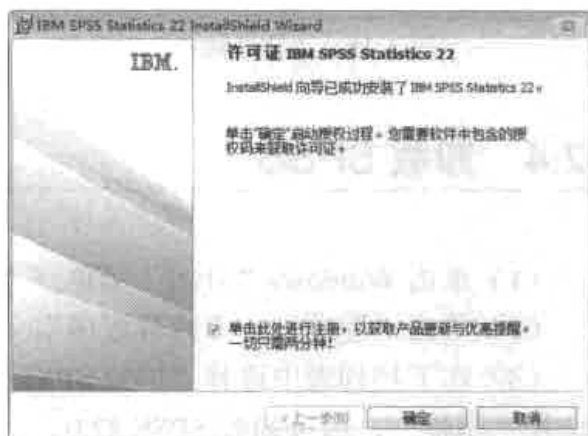


图 1-14 安装界面 14


1.2.2 启动 SPSS

用户可以在开始菜单中找到 SPSS 的执行程序: “开始” → “程序” → SPSS for Windows。为了方便, 用户可以设置 SPSS 的快捷方式: 右键单击“SPSS for Windows”, 在弹出的菜单中选择“发送到” → “桌面快捷方式”, 此时, 计算机桌面即会出现一个 (SPSS 22.0) 快捷键图标, 双击该图标即可启动 SPSS。

启动 SPSS 后, 首先会弹出一个文件选择对话框 (图 1-15), 该对话框中有新建文件、最近的文件、新增功能、模块和可编程性、教程五大部分, 用户可以根据自己的需求选择相应的选项, 单击“确定”按钮完成操作, 即进入 SPSS 的数据编辑窗口。若不想每次启动 SPSS 都看到这个对话框, 勾选最下方的复选框“以后不再显示此对话框”即可。

1.2.3 退出 SPSS

有两种方法退出 SPSS, 一是在数据窗口界面, 选择“文件” → “退出”; 二是单

击标题栏上的  按钮。当输出了新的结果，或者对文件进行了修改，退出时会弹出如图 1-16 所示的对话框，提示用户是否保存相关内容。



Note



图 1-15 “文件选择”对话框

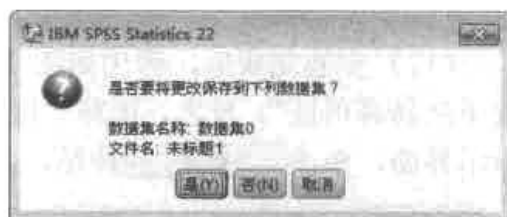


图 1-16 “提示保存”对话框

1.2.4 卸载 SPSS

- (1) 单击 Windows “开始”菜单→“控制面板”。
- (2) 单击“程序”→“卸载程序”。
- (3) 在下拉列表中选择“IBM SPSS Statistics 22”，单击鼠标右键，在弹出的列表中选择“删除”，即可卸载 SPSS 22.0。

1.3 SPSS 窗口及其功能

本节着重介绍 SPSS 22.0 常用的窗口，包括数据编辑窗口、结果输出窗口、命令语句窗口、脚本编辑窗口。

视频文件：视频文件\第一章\认识 SPSS.avi

1.3.1 数据编辑窗口

数据编辑窗口包括两种视图，分别为：数据视图、变量视图，如图 1-17 所示。前者是一个可以向下和向右扩展的二维表格，用于查看、录入和修改数据；后者用于输入和修改变量的定义。

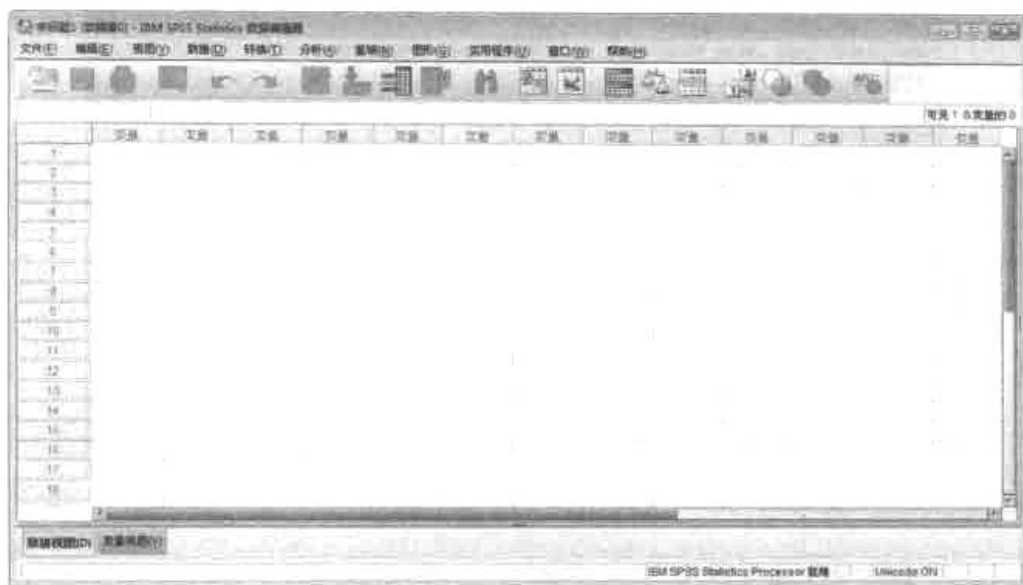


图 1-17 数据编辑窗口

数据编辑窗口包括标题栏、主菜单栏、快捷菜单栏、编辑栏、单元格信息栏、窗口切换标签页和状态栏。

- 标题栏：显示当前的文件名。
- 主菜单栏：包括文件、编辑、视图、数据、转换、分析、直销、图形、实用程序、窗口、帮助。
- 快捷菜单栏：有常用的快捷图标。
- 编辑栏：可输入和显示单元格中的数据。
- 单元格信息栏：在数据窗口中显示所有个案在各个变量中的取值，表中每一行表示一个个案，每一列表示一个变量；在变量窗口中显示所有变量的信息，表中每一行表示一个变量，每一列是关于变量一个方面的信息。
- 视图切换标签页：有“数据视图”、“变量视图”两个按钮，可以相互切换。
- 状态栏：用于查看当前程序运行的状态。

1.3.2 结果输出窗口

结果输出窗口是显示和管理 SPSS 统计分析结果（包括文本、表格及图形）的窗口，如图 1-18 所示。该窗口的内容可存为以.spv 为扩展名的 SPSS 文件。

在第一次产生分析结果的 SPSS 过程后，结果输出窗口自动打开，若要打开新的结果输出窗口，可单击“文件”→“新建”→“输出”。

结果输出窗口有标题窗、内容窗。前者用于显示已有的分析结果标题和内容索引；后者为统计分析的具体输出内容，包括文本、统计表和统计图。若要对内容窗中的结果进行编辑，可双击已选中内容，此时即可对结果进行修改。



Note



Note



图 1-18 结果输出窗口


1.3.3 语法编辑窗口

语法编辑窗口是编辑和运行命令文件的编辑器，如图 1-19 所示。该窗口不仅可以编辑对话框操作不能实现的特殊过程的命令语句，还可以将所有分析过程汇集在一个命令语句文件中，以避免处理较复杂资料时因数据的小小改动而大量重复分析过程。

该窗口的内容可存为以.sps 为扩展名的 SPSS 文件。用户根据自己的需求可以对命令文件进行修改、编辑，也可以编写针对当前数据文件的命令程序。



图 1-19 语法编辑窗口

在任何统计分析对话框上，都可以通过单击  按钮自动打开命令语句窗口，将执行 SPSS 过程的相应命令语句写在窗口中。

若要打开新的命令语句窗口，可单击“文件”→“新建”→“语法”。编写好命令文件后，可以单击菜单栏上的“运行”按钮，提交系统执行，显示输出窗口，得到分析结果。

1.3.4 脚本编辑窗口

脚本编辑窗口提供了 SPSS 内置语音 SaxBasic 的编程环境，其不仅可以开发 SPSS

的便捷功能或插件，还可以编写自动化数据处理程序，如图 1-20 所示。

在该编辑器中，可以利用程序或对话框编辑器编写出友好的 Windows 界面。基于 DDE 或 OLE 机制，实现与其他程序的接口。



Note

1.4 SPSS 帮助系统

SPSS 提供了相应的帮助系统，帮助读者在使用软件时如何获取帮助，下面着重介绍 SPSS 软件包提供的多种帮助工具。

1.4.1 “帮助”菜单

“帮助”菜单提供多种形式的帮助，有主题、教程、个案研究、统计指导、命令语法参考及算法等，如图 1-21 所示。

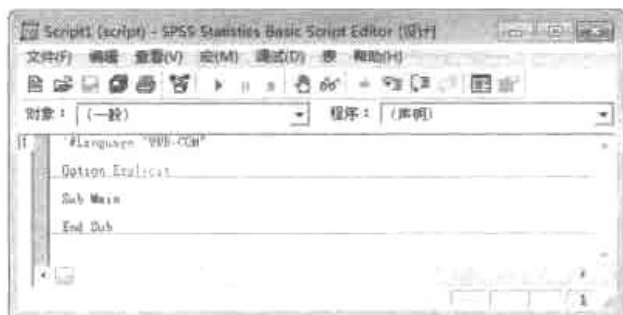


图 1-20 脚本编辑窗口



图 1-21 “帮助”菜单列表

在“帮助”菜单中，单击主题、教程、个案研究、统计指导，均可弹出浏览器形式的帮助窗口，且所弹出的窗口左侧的搜索、目录、索引、搜索结果、标签（后四项的切换可单击左下角的切换标签）都一样，其中的目录包括这几项所有帮助内容，而右侧部分则是具体的帮助内容。例如，单击“帮助”→“教程”，出现如图 1-22 所示界面。

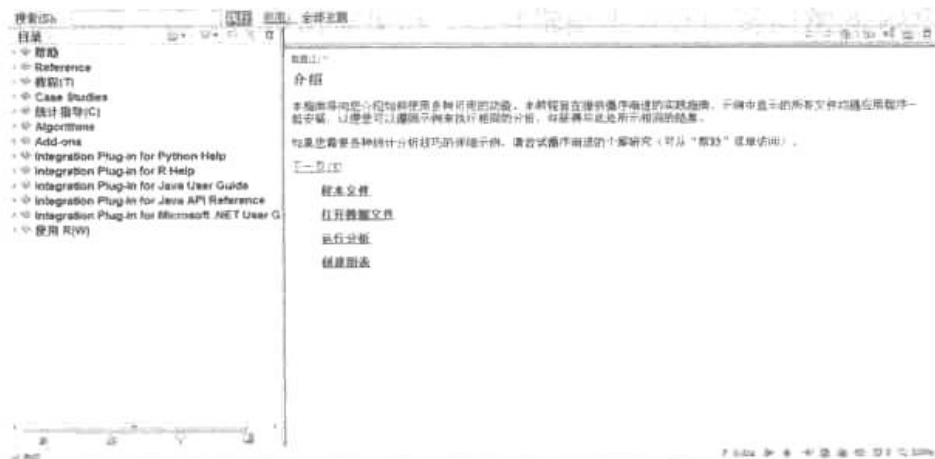


图 1-22 “帮助”对话框

1.4.2 其他帮助方法



Note

在用户界面中的许多地方都可以获得帮助。

(1) 对话框“帮助”按钮。单击对话框中的“帮助”按钮可直接进入该对话框的“帮助”主题。例如单击“编辑”→“选项”，弹出如图 1-23 所示的窗口，在右下角可看到“帮助”按钮。

(2) 命令语法。在命令语法窗口中，将光标放在命令语法块中的任意位置，然后按键盘上的 F1，此时将弹出浏览器形式的帮助窗口，显示该命令的完整命令语法图表。可从相关主题列表的链接中以及“帮助目录”选项卡中获得完整的命令语法文档。



图 1-23 “选项”对话框

1.5 本章小结

本章主要介绍了 SPSS 的发展简史、安装与启动、退出与卸载、窗口及其功能以及帮助系统。SPSS 的常用窗口有数据编辑窗口、结果输出窗口、命令语句窗口、脚本编辑窗口。

第2章

数据文件操作

本章主要介绍数据文件的建立、编辑和输出操作，这是对数据文件进行分析和管理的首要工作，熟练掌握本章内容有利于保证数据分析结果的正确性与科学性。本章分别介绍数据类型及定义、获取数据与查看数据的方法、数据的编辑和数据的输出。

学习目标

- (1) 熟练掌握数据类型及其定义。
- (2) 熟练掌握获得数据的方法，如读入 Excel 数据文件、纯文本数据文件和数据库文件。
- (3) 熟练掌握数据的编辑，如插入、撤销和删除。
- (4) 熟练掌握输出数据的方法。



Note

数据是统计分析的基础，用户在进行分析之前，需要区分不同的数据类型，同时，掌握定义变量的方法。

视频文件：视频文件\Chapter02\定义变量.avi

2.1.1 数据类型

数据包括常量和变量。常量指取值在一定阶段保存不变的量，如圆周率，SPSS 中的常量包括数值型、字符型和日期型；变量指在不同的记录行取不同的值，即取值可变的量。

2.1.2 定义变量

输入数据前首先要定义变量。定义变量即要定义变量名、变量类型、变量宽度、小数位数、变量标签、变量值标签和变量的格式等。



说明

变量的定义在“变量视图”窗口进行，每一行表示一个变量的定义信息。

1. 定义变量名

SPSS 默认的变量名为 VAR00001、VAR00002 等，用户可以根据自己的需要对变量进行命名。变量的命名有一定的规则，具体内容如下：

(1) 必须以字母、汉字或字符@开头，其他字符可以是任何字母、数字或_、#、\$ 等符号。

(2) 不能使用空白字符和其他特殊字符（如“!”、“?”等）。

(3) 变量命名不区分大小写且必须唯一。

(4) 用户定义的变量不能以“\$”为首写。以“\$”为首写的变量名特指 SPSS 的系统变量，它不可修改，而且在程序中不可用。

(5) 避免最后一个字符是“.”，因为英文句点有时会作为命令的结束标志，若这样定义变量，则容易引起歧义。

(6) 避免最后一个字符是“_”，因为下画线一般作为由程序或命令自动生成的变量名的结尾。

(7) SPSS 的保留字不能作为变量的名称，SPSS 的保留字有 ALL、AND、WITH、NOT、OR、BY、EQ、GE、GT、LE、LT、NE、TO。若使用了上述保留字作为变量名，系统会自动提示。



Note

2. 定义变量类型


单击“类型”相应的单元格中的按钮, 弹出“变量类型”对话框, 如图 2-1 所示, 在对话框中选择合适的变量类型并单击“确定”按钮, 即可定义变量类型。



图 2-1 “变量类型”对话框

SPSS 22.0 的变量类型包括数值型、字符串型、日期型, 具体介绍如下。

(1) 数值型变量。

数值型变量的长度是用字符个数度量的数字宽度, 小数点和其他分界符也计算在内, 常用的数值型变量如下。

- 数值: 可以定义数值的宽度和小数位数, SPSS 默认位数分别为 8、2。数值的宽度包括整数部分+小数点+小数部分的位数。
- 逗号: 指整数部分每 3 位数添加一个逗号, 如“1,234.00”。其余定义方式同数值型。
- 点: 以整数形式显示, 每 3 位数加一点 (这里不是小数点)。可以定义小数位数, 但所显示的小数部分均为 0, 且小数点用逗号表示。如自定义小数位数为 2, 则“9.8765”显示为“98.765,00”。
- 科学计数法: 指数值在数据窗口中以指数形式显示。如定义数值宽度为 8, 小数位数为 3, 则“123.456”显示为 1.23E+02。
- 美元: SPSS 提供了多种货币显示形式, 用户可根据需要选择, 并定义数据宽度和小数位数。其值在显示时, 有效数字前面带有“\$”, 输入时可以不输入“\$”, 显示时系统会自动加上“\$”和分隔符。
- 定制货币: 默认显示为整数部分每 3 位加一个逗号, 用户可以定义数据宽度和小数位数。如 9876543.21 显示为 9,876,543.21。
- 受限数值型: 指输入的数值位数限制为所设定的数据宽度。假定数据宽度设为 4, 则“112233”显示为“2233”, 而“11”显示为“0011”。

(2) 字符型变量。

用户自定义字符长度以便输入字符, 使用时需注意以下几个方面:

- SPSS 区分短字符串和长字符串, 短字符串最长 8 字节, 一个长字符串大于等于 8 字节, 长字符串变量不能定义用户缺失值。有些分析过程可以处理短字符串, 但不能处理长字符串。




Note

- 系统缺失值不能用于生成字符串变量。
- 当生成新变量、修改原变量时，可能产生缺失值或未定义的变量值，这时系统自动赋予值为空。变量值以空格表示时，若无特别定义，不能代表缺失值。
- 字符型变量不能参与算术运算。
- 字符串中的大小写字母是截然不同的两个字符，用户在使用时需注意。

(3) 日期型变量。

SPSS 中的日期型变量既可以表示日期，也可以表示时间。SPSS 提供了多种日期显示形式，用户可根据需要选择。

3. 定义变量宽度

系统默认宽度为 8。单击宽度列中的某一单元格，将出现  图标，单击上下箭头可调大调小变量的宽度。也可直接在单元格里输入数字。当变量类型为日期型时，变量长度的设定无效。

4. 定义变量小数



设置变量的小数位数，方法同变量宽度的设置。当变量类型为日期型时，变量小数位数的设定无效。

5. 定义变量标签

变量标签是对变量名的进一步描述，其可以输入 120 个字符，且可显示大小写。变量标签可以显示在输出窗口，便于查看结果时理解变量的实际意义。



6. 定义变量值标签

变量值标签是对变量的每一个可能取值的进一步描述，当变量是定性或定序变量时，变量值标签的描述是非常管用的。

例如，在统计分析中经常用 1 代表“男”，2 代表“女”等。具体操作为：单击“值”相应的单元格右侧的 ，弹出“值标签”对话框，如图 2-2 所示；在“值”中输入 1，“标签”中输入男；单击  按钮即完成变量值标签的添加。



注意

若要对已经设置好的值标签进行修改，只需选中该值标签，在值和标签列中进行修改后单击  按钮即可。若要删除值标签，则选择该值标签，单击  按钮即可。

7. 定义缺失值

SPSS 有两类缺失值：系统缺失值和用户缺失值，系统默认为无。在实际工作中常会因为某种原因出现记录数据失真、没有记录等缺失现象。

例如在统计过程中，一些被调查者没有回答的题目都将标为用户缺失值。定义缺失值的具体操作为：单击“缺失”下面的某一单元格，弹出“缺失值”对话框，如图 2-3 所示，其中有以下三种定义缺失值的方法：

- 没有缺失值，SPSS 默认方式。若当前变量的取值完整，则选择此项；

- 离散缺失值，可以指定 1~3 个离散的缺失值。例如指定 20、95 和 88 为缺失值；
- 范围加上一个可选离散缺失值，可以指定一个缺失值范围和一个离散的缺失值。

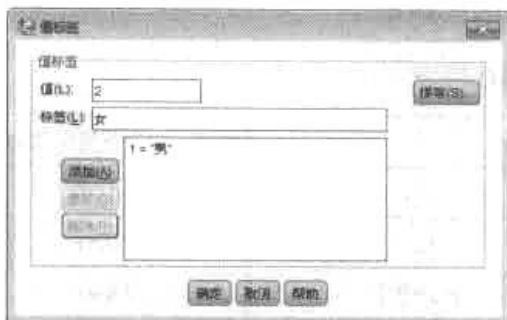


图 2-2 “值标签”对话框



图 2-3 “缺失值”对话框



Note

8. 定义变量的显示宽度

SPSS 默认为 8。用户可根据需要进行设置。

9. 定义变量显示的对齐方式

有左对齐、右对齐、居中对齐三种方式，SPSS 默认为右对齐。

10. 定义变量的测量尺度

可选测量方式有以下 3 种：

- 名义：一种分类变量，即它的取值只代表观测对象的不同类别，变量的取值之间没有内在的大小可比性。例如“性别”变量。
- 有序：一种分类变量，但是变量取值之间有内在的大小顺序或等级。例如“满意度”变量的取值为 1-很不满意、2-比较满意、3-非常满意，由小到大的取值代表满意度的提高。
- 度量：又称定距变量或刻度变量，一般为有刻度度量的连续变量，它的取值之间可以比较大小，且可以定义距离。例如“年龄”、“年份”等。

11. 定义变量的角色

用于定义变量在后续统计分析中的功能作用，SPSS 提供的角色选项有输入、目标、两者、无、分区和拆分。



技巧

如果有多个变量的类型相同，可以先定义一个变量，然后把该变量的定义信息复制给新变量。具体操作为：先定义好一个变量，在该变量的行号上单击右键，在弹出的快捷菜单中选择“复制”命令，然后选择其他同类型所在行，单击鼠标右键，在弹出的快捷菜单中选择“粘贴”即可。

2.1.3 操作符与表达式

SPSS 的基本运算有 3 种：数学运算、关系运算和逻辑运算。对应的操作符表示方法如表 2-1 所示。



Note

数学运算符也就是常用的算术运算符，可以连接数值型的常量、变量和函数，形成算术表达式，运算结果通常为数值。运算符的优先级为：括号>函数>乘方（幂）>乘或除>加或减，同一优先级的符号，位于左侧的优先级高。

表 2-1 操作符表

数学运算操作符		关系运算符		逻辑运算符	
+	加	<(LT)	小于	&(And)	与
-	减	>(GT)	大于	→(Or)	或
*	乘	<=(LE)	小于等于	~(Not)	非
/	除	>=(GE)	大于等于		
**	幂	!=(EQ)	等于		
()	括号	!=(NT)	不等于		

关系运算符用于建立两个量之间的比较关系，如果比较关系成立，则关系表达式的值为真（true），否则为假（false）。例如，假定表达式为“ $a < 0$ ”，那么如果 $a = 1$ ，则表达式“ $a < 0$ ”为假，表达式的值为 0（假）；如果 $a = -1$ ，那么表达式“ $a < 0$ ”为真，表达式的值为 1（真）。

在表 2-1 中，关系运算符和逻辑运算符均有两种表达方法，括号中的关系运算符与括号前的是等价的。



注意

相互比较的两个量的类型必须一致，无论进行比较的两个量是字符型还是数值型，比较的结果均是逻辑型。

逻辑运算符、逻辑型变量或值为逻辑型的表达式（如关系表达式）都称为逻辑表达式，逻辑表达式的值为逻辑型（true 或 false）。逻辑运算规则如表 2-2 所示。

表 2-2 逻辑运算规则

逻辑表达式	结 果	逻辑表达式	结 果
true AND true	=ture	true OR true	=true
true AND false	=false	true OR false	=true
false AND false	=false	false OR false	=false
true AND missing	=missing	true OR missing	=ture
missing AND missing	=missing	missing OR missing	=missing
false AND missing	=false	false OR missing	=missing

2.2 获取数据与查看数据的方法

可以通过直接录入的方法获取数据，同时，还可直接读取其他格式的数据文件，如 Excel、数据库和 SAS 等数据文件，本节着重介绍获取数据和查看数据的方法。

2.2.1 在 SPSS 数据窗口中直接录入数据

定义了所有变量后，单击左下方“数据视图”窗口切换按钮，即可在出现的数据视图窗口中输入数据。

用鼠标左键单击某单元格，激活该单元格后即可输入数据。输入完毕后，只需将光标移动到下一个单元格继续输入数据。数据录入时可以逐行或逐列录入。



Note



技巧

若录入的数据中有多个连续单元格的数值是相同的，则可以选中第一个单元格并右键单击选择复制；然后选中第二个单元格，并按住键盘上的 shift 键；再单击最后一个单元格，并单击鼠标右键选择粘贴。这样，就可以快速完成多个相同数据的录入工作。

2.2.2 直接读入 Excel 数据文件

当 Excel 文件的数据结构符合 SPSS 的数据结构时，SPSS 可以直接读取 Excel 文件数据。下面介绍读入一个 Excel 数据文件的实例。

实例一：“data02-01.xls”是一个 Excel 文件，现在需要转换成 SPSS 数据。

数据文件：数据文件\chapter02\data02-01.xls

视频文件：视频文件\chapter02\数据导入和查看.avi

(1) 确认 Excel 文件的数据结构是否符合 SPSS 的要求。我们发现，该文件的每一行数据是每个同学的信息，每一列是一个变量，所以，其数据结构符合 SPSS 要求。

(2) 选择“文件”→“打开”→“数据”，在弹出的“打开数据”对话框下方的“文件类型”项选择 Excel (*.xls,*.xlsx,*.xlsm)，如图 2-4 所示，当前目录下的 Excel 文件即可显示在对话框的文件列表中。

(3) 选定目标文件“语文成绩.xls”，如图 2-5 所示，单击 **打开(O)** 按钮即可进入下一级对话框。



图 2-4 “打开数据”对话框一



图 2-5 “打开数据”对话框二

(4) 弹出“打开 Excel 数据源”对话框，如图 2-6 所示，选中复选框“从第一行数

若 Excel 文件中有多个 sheet, 那么需要在“工作表”下拉框中选择要读入的 sheet; 若不选择, SPSS 则默认为 Sheet1。



提示

SPSS 一次只能读入一个 sheet 的数据。另外,若只需读取部分区域的数据,则可以在“范围”中输入要读取数据的范围。如要读入前 21 行的数据,则在该对话框中输入 A1:C21,表示读取的区域是以 A1 单元为左上角, C21 为右下角的矩形区域。

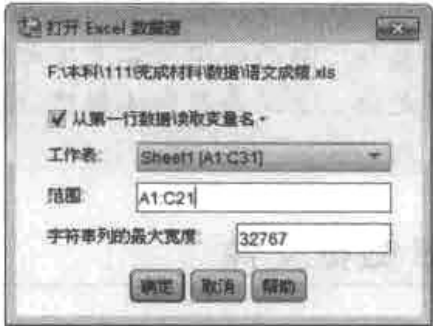



图 2-6 “打开 Excel 数据源”对话框

(5) 对话框中其他设置不变, 单击  按钮即可读入数据到 SPSS 中, 如图 2-7 所示。检查 SPSS 数据的变量与个案的数量是否丢失。

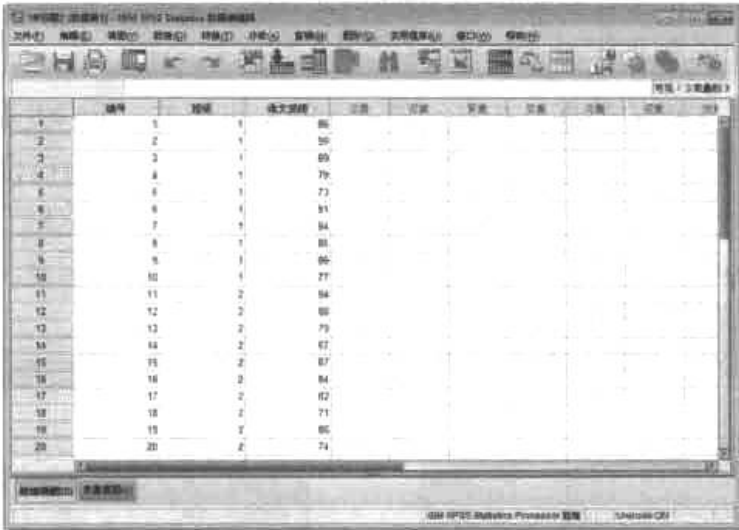


图 2-7 在 SPSS 数据窗口显示读入的数据

2.2.3 读入纯文本数据文件

纯文本数据文件是计算机各种软件中最通用的一种格式文件，其没有保存格式，因此其文件很小。根据纯文本数据文件中数据的排序方式，可以将其分为自由格式和固定格式。

前者的文本文件的每个个案的变量数目、排列顺序固定,但数据项的长度可以不同,且数据项之间必须有分隔符(逗号、空格、Tab 键等);后者的文本文件要求每个个案的

变量数目、排列顺序、变量取值长度都固定不变，且数据项之间不需要分隔符。现以实例说明两种不同的数据读入方法。

1. 以自由格式读入数据

实例二：“data02-02.txt”数据文件是纯文本数据，如图 2-8 所示。现要求以自由格式读入数据。

数据文件：数据文件\Chapter02\data02-02.txt

视频文件：视频文件\Chapter02\数据导入和查看.avi

(1) 打开“data02-02.txt”数据文件，选择“文件”→“打开文本数据”，弹出如图 2-9 所示的对话框，在其中选择目标文件“data02-02.txt”并单击 **打开(O)** 按钮打开数据文件（图 2-9），SPSS 随即启动文本数据导入的引导窗口。

1		86
2	1	90
3	1	89
4	1	79
5	1	73
6	1	91
7	1	94
8	1	85
9	1	86
10	1	77
11	2	94
12	2	88
13	2	79
14	2	67
15	2	87
16	2	84
17	2	82
18	2	71
19	2	86

图 2-8 “data02-02.txt”部分数据



图 2-9 “打开数据”对话框

(2) 弹出“文本导入向导-第 1 步，共 6 步”对话框，如图 2-10 所示。在该对话框中可以看出文本文件中的数据信息。单击 **下一步>(N)** 按钮进入下一步。

(3) 弹出“文本导入向导-第 2 步，共 6 步”对话框，如图 2-11 所示。在“变量是如何排列的？”和“变量名称是否包括在文件的顶部？”选框中分别选择“分隔”、“否”。单击 **下一步>(N)** 按钮进入下一步。



图 2-10 “文本导入向导-第 1 步，共 6 步”对话框



图 2-11 “文本导入向导-第 2 步，共 6 步”对话框



Note

(4) 弹出“文本导入向导-第3步,共6步”对话框,如图2-12所示。按照图示填写和选择相应内容。单击 **下一步>(N)** 按钮进入下一步。

(5) 弹出“文本导入向导-第4步,共6步”对话框,如图2-13所示。在“变量之间有哪些分隔符?”选框中选择文本文件的分隔符,有制表符(即Tab键分隔符)、空格、逗号、分号和其他,本次选择制表符。单击 **下一步>(N)** 按钮进入下一步。



图2-12 “文本导入向导-第3步,共6步”对话框



图2-13 “文本导入向导-第4步,共6步”对话框

(6) 弹出“文本导入向导-第5步,共6步”对话框,如图2-14所示。在此对话框中的下面是数据的文件的预览,这时用户需要注意查看各变量的长度是否正确。单击 **下一步>(N)** 按钮进入下一步。

(7) 弹出“文本导入向导-第6步,共6步”对话框,如图2-15所示。单击“完成”即成功读入自由格式的文本文件内容。



此时若想保存前几步操作中所定义的读取自由格式的文本文件信息,可以在该对话框中的“您要保存此文件格式以备以后使用吗?”中选择“是”选项,然后单击右边的“另存为”按钮,保存格式信息。单击 **完成** 按钮完成文本的导入工作。



图2-14 “文本导入向导-第5步,共6步”对话框



图2-15 “文本导入向导-第6步,共6步”对话框

(8) SPSS 顺利读入数据后,用户需要在变量视图中重新定义变量名,并进行宽度、小数位数等设置,如图 2-16 所示。



图 2-16 显示读入数据的变量视图



Note

2. 以固定格式读入数据

实例三:“data02-03.txt”数据文件是纯文本数据,如图 2-17 所示。现要求以固定格式读入数据。

数据文件: 数据文件\Chapter02\data02-03.txt

视频文件: 视频文件\Chapter02\数据导入和查看.avi

(1) 选择“文件”→“打开文本数据”,在弹出的对话框中选择目标文件“data02-03.txt”(图 2-18),并单击 **打开(O)** 按钮打开如图 2-17 所示的数据文件,SPSS 随即启动文本数据导入的引导窗口。

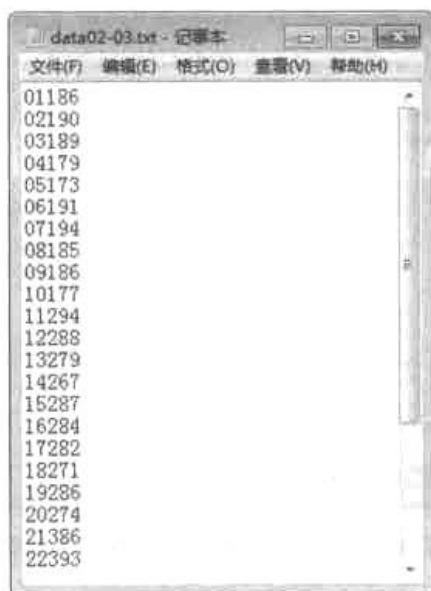


图 2-17 固定格式的纯文本数据



图 2-18 “打开数据”对话框

(2) 弹出“文本导入向导-第 1 步,共 6 步”对话框,如图 2-19 所示。在该对话框中可以看出文本文件中的数据信息。单击 **下一步>(N)** 按钮进入下一步。

(3) 弹出“文本导入向导-第 2 步,共 6 步”对话框,如图 2-20 所示。在“变量是如何排列的?”和“变量名称是否包括在文件的顶部?”选框中分别选择“固定宽度”、“否”。单击 **下一步>(N)** 按钮进入下一步。



Note



图 2-19 “文本导入向导-第 1 步, 共 6 步”对话框



图 2-20 “文本导入向导-第 2 步, 共 6 步”对话框

(4) 弹出“文本导入向导-第 3 步, 共 6 步”对话框, 如图 2-21 所示。按照图示填写和选择相应的内容。单击 **下一步(N) >** 按钮进入下一步。

(5) 弹出“文本导入向导-第 4 步, 共 6 步”对话框, 如图 2-22 所示。前两列数据属于第一个变量, 在第二列和第三列之间单击鼠标左键即可添加一条分割线; 利用同样的方法在第三列和第四列之间添加分割线将第二个、第三个变量分隔出来。单击 **下一步(N) >** 按钮进入下一步。



图 2-21 “文本导入向导-第 3 步, 共 6 步”对话框

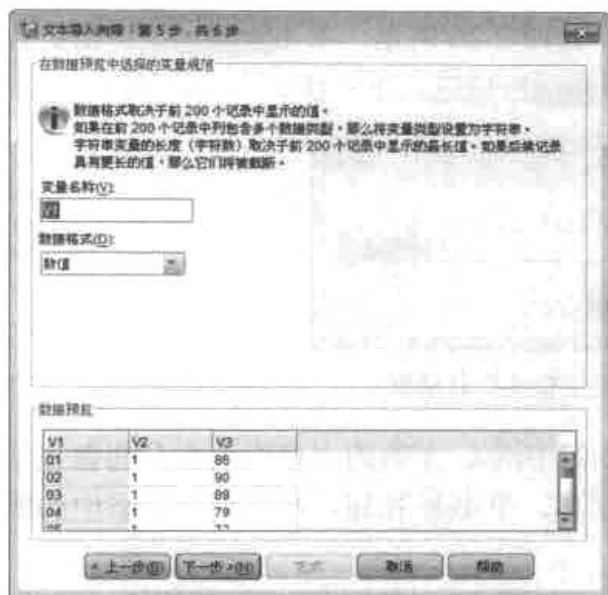


图 2-22 “文本导入向导-第 4 步, 共 6 步”对话框

(6) 弹出“文本导入向导-第 5 步, 共 6 步”对话框, 如图 2-23 所示。上一步已将前两列赋给第一个变量 V1, 中间一列赋给第二个变量 V2, 后面两列赋给第三个变量 V3。单击 **下一步(N) >** 按钮进入下一步。

(7) 弹出“文本导入向导-第 6 步, 共 6 步”对话框, 如图 2-24 所示。单击 **完成** 按钮即成功读入固定格式的文本文件内容。

(8) SPSS 顺利读入数据后, 用户需要在变量视图中重新定义变量名, 并进行宽度、小数位数等设置。



Note

图 2-23 “文本导入向导-第5步,共6步”对话框

图 2-24 “文本导入向导-第6步,共6步”对话框

2.2.4 读入数据库文件

数据库文件是一种常用的数据存储格式文件,在数据管理上起着很重要的作用,尤其是对于数据量非常庞大的数据文件。当需要对这些数据进行分析时,则可以采用 SPSS 对数据库文件进行导入。所以,掌握读入数据库文件的方法是非常有必要的。

实例四:“data02-04.mdb”数据文件是一个数据库文件,现在需要转换成 SPSS 数据。

数据文件: 数据文件\Chapter02\data02-04.mdb

视频文件: 视频文件\Chapter02\数据导入和查看.avi

(1) 选择“文件”→“打开数据库”→“新建查询”,弹出“数据库向导”对话框 1,如图 2-25 所示。右侧列表是 SPSS 提供的多种数据库的读入引擎。本例要导入的数据是存放在 Access 数据库中,所以选择 MS Access DataBase。单击“下一步>>”按钮进入下一步。



图 2-25 “数据库向导”对话框 1

(2) 弹出“ODBC 驱动程序登录”对话框,如图 2-26 所示,单击 **浏览(B)** 按钮指定数据库文件所在的路径和数据库文件名,并单击 **确定** 按钮进入下一步。



Note

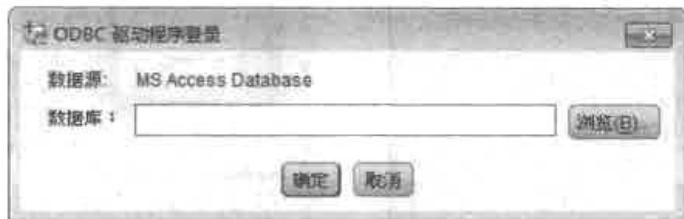


图 2-26 “ODBC 驱动程序登录”对话框

(3) 弹出“数据库向导”对话框 2,如图 2-27 所示。左侧的“可用表格”列表显示当前数据库中的所有数据表的名称。选中表格名称,单击 **+** 按钮,即可将该表格中的所有字段选入“以此顺序检索字段”列表中。



提示

本例将“员工表”选入右侧列表中,图 2-27 右侧所显示的即为员工表中所包含的字段。单击 **下一步>(N)** 按钮进入下一步。

(4) 弹出“数据库向导”对话框 3,如图 2-28 所示。该对话框用于指定导入条件,设定条件后,SPSS 只导入满足指定条件的记录。本例导入全部数据,则不需要设定条件。单击 **下一步>(N)** 按钮进入下一步。



图 2-27 “数据库向导”对话框 2



图 2-28 “数据库向导”对话框 3

(5) 弹出“数据库向导”对话框 4,如图 2-29 所示。该对话框显示被选中字段的名称、数据类型及是否重新编码数值的复选框。用鼠标左键双击各个字段的名称和数据类型可对其进行修改,勾选重新编码数值表示使用原始值作为值标签将字符串变量转换为数值。单击 **下一步>(N)** 按钮进入下一步。

(6) 弹出“数据库向导”对话框 5,如图 2-30 所示。该对话框显示所形成的数据库查询 SQL 语句,若用户想将其保存起来,可以单击 **浏览(W)** 按钮指定保存路径。若不保存,则直接单击 **完成** 按钮,读入的数据如图 2-31 所示。



图 2-29 “数据库向导”对话框 4



图 2-30 “数据库向导”对话框 5

ID	姓名	性别	年龄	职位	所属部门	入职时间	薪资
1	000001	李四	男	24	职员	04	3-Mar-1997
2	000002	张三	女	27	职员	04	6-Feb-1988
3	000003	赵五	男	28	职员	03	3-Jan-1989
4	000004	王六	男	25	主管	03	5-Jun-1990
5	000005	林七	女	30	经理	02	4-Aug-1986
6	000006	江八	女	30	主管	04	5-Jan-1987
7	000007	张九	男	19	职员	02	5-Jan-1990
8	000008	王十	男	19	职员	04	14-Feb-2000
9	000009	李十一	女	21	职员	03	14-Mar-2000
10	000010	王十二	女	22	职员	02	14-Apr-2000
11	000011	王十三	男	24	主管	04	18-May-1998
12	000012	李十四	男	28	经理	03	20-Jan-1998

图 2-31 读入的部分数据

2.2.5 查看文件信息

打开数据文件后，在数据编辑窗口选择“文件”→“显示数据文件信息”→“工作文件”，即可打开当前文件的相关信息，其在结果输出窗口显示，如图 2-32 所示。若要查看外部文件，选择“文件”→“显示数据文件信息”→“外部文件”即可。

变量	位置	标签	数据类型	角色	精度(位)	格式	初始格式	目标格式
srchid	1	Menu scenarios	字符串	输入	8	C	F3	F3
	2	Gender	字符串	输入	8	C	F3	F3
	3	Toast pop-up	字符串	输入	5	C	F3	F3
	4	Buttered toast	字符串	输入	5	C	F3	F3
EMM	5	English muffin and margarine	字符串	输入	5	C	F3	F3
	6	Jelly donut	字符串	输入	5	C	F3	F3
JO	7	Cinnamon toast	字符串	输入	5	C	F3	F3
	8	Blueberry muffin and margarine	字符串	输入	5	C	F3	F3
HRB	9	Hard rolls and butter	字符串	输入	5	C	F3	F3

工作文件中的变量

变量	位置	标签
srchid	1	Overall preference
	2	Breakfast, with juice, bacon and eggs, and beverage
	3	Breakfast, with juice, cold cereal, and beverage
	4	Breakfast, with juice, pancakes, sausage, and beverage
	5	Breakfast, with beverage only
	6	Snack, with beverage only
gender	1	Male
	2	Female

图 2-32 查看文件信息



Note



Note

另外，还可通过工具栏查看变量信息。选择“实用程序”→“变量”，弹出“变量”对话框，如图 2-33 所示。选中左侧变量列表中的某个变量，右侧就会显示该变量的所有信息，但这些信息不可编辑。

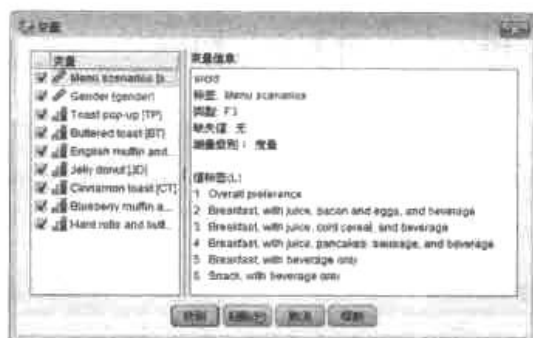


图 2-33 查看变量信息

2.3 数据的编辑

本节着重介绍数据的编辑方法，包括插入和删除变量、插入和删除个案、数据的复制和粘贴、撤销操作和数据的输出。

视频文件：视频文件\Chapter02\数据的编辑.avi

2.3.1 插入和删除变量

在数据视图窗口，右键单击变量栏中的某个变量名，在弹出的快捷菜单中选择“插入变量”，如图 2-34 所示，即可在选中变量列的左侧插入一个新的默认变量，SPSS 默认其名称为 VAR0000x（此处 x 指任何一个正整数，为 SPSS 定义的变量序号）。选择“编辑”→“插入变量”也可插入新变量。

在变量视图窗口，只需编辑指定的行，即可插入新的变量。或单击某一行的行号选中该行，然后单击鼠标右键并选中“插入变量”，如图 2-35 所示，此时即可在该行上面插入一个新的变量。

删除变量的方法同插入变量的方法类似。



图 2-34 在数据视图的列中插入变量



图 2-35 在变量视图的行中插入变量

2.3.2 插入和删除个案

在数据视图窗口，右键单击序列栏的某个行号，在弹出的快捷菜单中选择“插入个案”，如图 2-36 所示，即可在选中行的上面插入一个行的新记录行。选择“编辑”→“插入个案”也可插入新记录行。

删除个案的方法同插入个案的方法类似。



Note

2.3.3 数据的复制、剪切和粘贴

选择数据，右键单击“复制”或“剪切”即可复制或剪切相应内容，此时，在空白单元格右键单击“粘贴”即可完成内容的粘贴。

2.3.4 撤销操作

在用户对数据进行操作后，若想恢复至操作前的状态，可以单击按钮。同时，用户还可通过单击按钮恢复至操作后的状态。

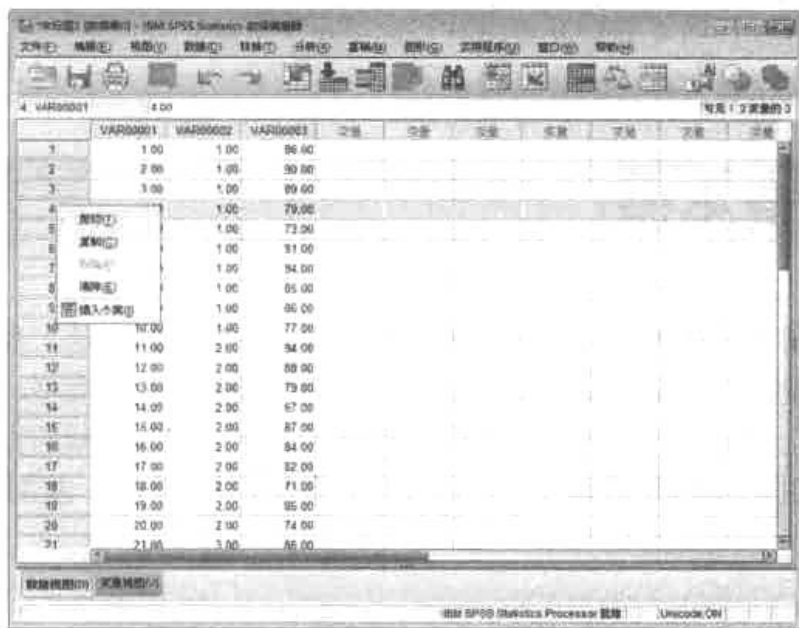


图 2-36 在变量视图的行中插入个案

2.4 数据的输出

SPSS 可以将数据保存为 SPSS (*.sav)、Excel (*.xls)、dBASE (*.dbf)、ASC II (*.dat)、Access (*.mdb) 及各类 SAS 数据文件。选择“文件”→“另存为”，即会弹出“将数据保存为”对话框，如图 2-37 所示。

选择保存路径，输入文件名称，选择数据类型，单击 **保存(S)** 按钮即可。若只需保存部分变量，可单击对话框中的 **变量(V)** 按钮，弹出“数据保存为：变量”对话框，如图 2-38 所示，选择所需保存的变量即可。



Note



图 2-37 “将数据保存为”对话框



图 2-38 “数据保存为：变量”对话框



技巧

SPSS 提供了数据文件的只读功能，对于因操作失误导致数据文件的不可逆转的改变起了很好的保护作用。方法为：选择“文件”→“将文件标记为只读”，此时数据处于锁定状态，虽然能进行编辑，但无法保存。若用户要将只读数据文件更改为可编辑数据文件，可选择“文件”→“将文件标记为只读”，这时就将数据文件切换为编辑状态。

2.5 本章小结

本章主要介绍了数据类型及定义、获取数据与查看数据的方法、数据的编辑和数据的输出。定义变量在“变量视图”窗口进行，包括变量名、变量类型、变量宽度、小数位数、变量标签、变量值标签和变量的格式等定义内容；获取数据的方法有直接录入数据、读取 Excel、纯文本和数据库等；数据编辑主要有插入、删除、复制、撤销及输出等。

第3章

数据文件整理

数据文件的整理是数据分析的基础。在统计分析时,最初的变量值可能不符合统计分析的要求,这就需要用户根据特定目的,对已有的数据进行整理,使数据格式符合相应的分析方法。本章分别介绍个案排序、数据变换、数据文件的合并、数据文件的结构重组、分类汇总、文件拆分、选择个案、个案加权、计算新变量、对个案内的值计数、变量的重新编码和个案等级排秩。

学习目标

- (1) 熟知本章所有数据文件整理的方法。
- (2) 熟记本章所涉及到的对话框的各项含义。
- (3) 熟练掌握各个数据文件整理方式的 SPSS 操作步骤。
- (4) 深刻理解本章中的各个实例。



Note

3.1 数据排序

数据排序在数据分析过程中非常重要，其便于了解数据的取值状况，SPSS 可以对数据基于一个或多个个案（变量）进行排序。数据排序包括个案排序和变量排序。

3.1.1 个案排序

在进行数据分析时，若要按照某个变量的取值重新排列个案在数据文件中出现的先后顺序，可通过“数据”→“排序个案”实现。当存在多个用于指定排序的变量时，数据首先按照主变量值的大小排序，然后对那些具有相同主排序的个案按照次变量值进行排序。

1. 参数设置

打开数据文件，选择“数据”→“排序个案”，弹出“排序个案”对话框，如图 3-1 所示。

- 排序依据：在左侧的变量列表中选中排序变量，单击 按钮，将其移动到“排序依据”列表中。
- 排列顺序：当将左侧的变量选入“排序依据”列表后，需在“排序顺序”选框中选择变量的排序方式，如果是针对多个变量排序，则需要依次指定每个排序变量及其排序方法。包括升序排列和降序排列。
 - 升序排列：指将所选变量从小到大的顺序对数据文件中的个案进行排序。
 - 降序排列：指将所选变量从大到小的顺序对数据文件中的个案进行排序。
- 保存已分类数据：即将分类的数据另存为一个文件或索引。当左侧的变量被选入右侧的列表时，激活该选框；勾选“保存带分类数据的文件”复选框，激活 按钮，单击它即可指定保存文件的路径。



图 3-1 “排序个案”对话框



Note

2. 个案排序的SPSS实现

实例一：针对“data03-01.sav”数据，如图 3-2 所示，其中包括“树种”、“树高”、“枝下高”和“冠幅”这四个变量。现要求根据树高进行升序排列，同时根据冠幅进行降序排列。


数据文件：数据文件\Chapter03\data03-01.sav


视频文件：视频文件\Chapter03\数据排序.avi

	树种	树高	枝下高	冠幅
1	1	13.2	2.7	5.4
2	1	14.8	2.6	5.3
3	1	17.3	4.2	5.9
4	1	16.7	2.3	5.1
5	1	15.4	3.2	4.7
6	1	17.6	6.6	4.8
7	1	15.4	4.5	6.3
8	1	14.1	4.2	5.0
9	1	16.0	2.5	5.8

图 3-2 “data03-01.sav”部分数据

(1) 打开“data03-01.sav”数据文件，选择“数据”→“排序个案”，弹出如图 3-1 所示的“排序个案”对话框。

(2) 在左侧的变量列表中选中“树高”变量，单击按钮，则将其指定为“排序依据”变量，然后在排列顺序中选择“升序”。

(3) 将“冠幅”变量添加为排序依据变量，方法同上一步。此时的排列顺序选“降序”。另外，选中某一个排序依据变量，并单击按钮，可将其从排序依据变量恢复至变量列表。

(4) 完成所有设置后，单击按钮执行命令。

3. 个案排序的结果分析

图 3-3 为排序完成后的结果，由图可以看出，所有个案先按树高升序排列，当树高相同时，再按冠幅降序排列。



个案排序是整行数据排序，而非只对某列变量排序；数据排序后，原有数据的排列次序被打乱，故用户需注意根据需要保留数据的原始排列，以免发生混乱。



Note

	树种	树高	枝下高	冠幅
1	2	7.8	2.5	4.7
2	2	8.4	2.5	4.7
3	2	8.7	2.7	6.0
4	3	9.5	2.5	5.0
5	3	10.5	4.6	4.4
6	3	11.0	4.0	8.5
7	3	12.0	2.0	8.2
8	2	12.0	6.0	5.6
9	3	12.5	4.0	7.4
10	3	12.5	2.5	7.3
11	3	12.5	3.0	5.9
12	3	13.0	2.8	8.3
13	3	13.0	5.0	7.2
14	3	13.0	5.0	5.3
15	1	13.2	2.7	5.4
16	2	14.1	8.5	6.5
17	1	14.1	4.2	5.0
18	2	14.5	5.2	8.0
19	2	14.8	6.4	7.7
20	1	14.8	2.6	5.3
21	2	15.1	8.3	3.2
22	1	15.4	4.5	6.3

图 3-3 排序后的结果

3.1.2 变量排序

变量排序是对“变量视图”窗口中各变量的信息进行排序。比如对“树木调查”数据文件中变量的“宽度”进行升序排列，那么选择“数据”→“排序变量”，弹出“排序变量”对话框，如图 3-4 所示。选中“变量视图列”中的“宽度”和“排列顺序”中的“升序”，单击“确定”按钮即可。



图 3-4 “排序变量”对话框



变量排序在实际应用中非常少，用户可以根据自身需求考虑是否学习。

3.2 数据转置

数据转置，也称数据变换，其可以将数据文件中原来的行变成列，原来的列变成行，即将观测量变成变量，将变量变成观测量。

由于可用于 SPSS 分析的数据都必须符合特定的数据结构，所以若用户在录入数据时错误地将行与列的位置颠倒了，则可利用“数据”→“变换”将数据调整至正确的结构。



Note

3.2.1 参数设置

打开数据文件，选择“数据”→“变换”，弹出“变换”对话框，如图 3-5 所示。其中各项含义如下：

- 变量。在左侧的变量列表中选中需要变换的变量，单击 按钮，将其移至“变量”列表。被选入“变量”列表的变量在变换后变成观测量，而未被选入“变量”列表的变量将不会出现在新数据文件中。
- 名称变量。在左侧的变量列表中选中某个变量，单击 按钮，将其移至“名称变量”列表，在新数据文件中该变量的值将作为变量名。一般选择标识观测量，如姓名等。若选择的变量是数值型，那么变换后新变量名为该变量各种冠以字母“K_”。若用户未选择“名称变量”，系统则自动赋予变换后的变量名为 VAR001、VAR002、…、VAR00n。



图 3-5 “变换”对话框

3.2.2 数据转置的 SPSS 实现

实例二：针对“data03-02.sav”数据，如图 3-6 所示，包括五位评委为七位球员的打分情况，现要求将数据文件中的行列变换。

数据文件：数据文件\Chapter03\data03-02.sav

视频文件：视频文件\Chapter03\数据转置.avi



Note

	球员	评委1	评委2	评委3	评委4	评委5
1	球员1	6	7	8	7	6
2	球员2	9	9	9	10	8
3	球员3	9	9	7	8	9
4	球员4	8	7	6	7	7
5	球员5	8	8	8	7	7
6	球员6	7	9	9	9	8
7	球员7	10	9	9	10	10

图 3-6 “data03-02.sav”数据

(1) 打开“data03-02.sav”数据文件，选择“数据”→“变换”，弹出如图 3-5 所示的“变换”对话框。

(2) 选中“球员”变量，单击下面的 按钮，将其移至“名称变量”；选中其余变量，单击上面的 按钮，将其移入变量列表中。此时可先选中“评委1”，然后按住键盘上的 Shift 键，再选中“评委5”，这样，所有评委变量都被快速选中，如图 3-7 所示。



图 3-7 “数据变换”对话框

(3) 完成所有设置后，单击 按钮执行命令。

3.2.3 数据转置的结果分析

图 3-8 为变换后的结果，由图可以看出，原来的行标识变量变成了新的列变量名称，列变量名称变成了新的行标识变量，且变量名称默认为“CASE_LBL”。

	CASE_LBL	球员1	球员2	球员3	球员4	球员5	球员6	球员7
1	评委1	6.00	9.00	9.00	8.00	8.00	7.00	10.00
2	评委2	7.00	9.00	9.00	7.00	8.00	9.00	9.00
3	评委3	8.00	9.00	7.00	6.00	8.00	9.00	9.00
4	评委4	7.00	10.00	8.00	7.00	7.00	9.00	10.00
5	评委5	6.00	8.00	9.00	7.00	7.00	8.00	10.00

图 3-8 变换后的数据



提示

①若用户只变换部分变量，那么在单击“确定”按钮时，SPSS 会弹出警告框，如图 3-9 所示，提示用户未被转置的变量将被丢失，此时单击 按钮则继续执行“变换”任务，单击 则放弃“变换”。②数据的变换是不能撤销的。



图 3-9 “数据变换”警告框



Note

3.3 数据文件的合并

在进行数据处理和统计分析时，常常需要将多个数据文件合并成一个数据文件。合并方式分为个案合并和变量合并。

- 个案合并指在两个数据文件拥有共同变量时，将其中一个数据文件的个案合并到另一个数据文件中。
- 变量合并指在两个数据文件拥有共同个案时，将其中一个数据文件的变量合并到另一个数据文件相应的个案中。

3.3.1 个案合并

1. 参数设置

打开第一个数据文件，选择“数据”→“合并文件”→“添加个案”，此时弹出“个案合并”对话框，如图 3-10 所示，其中各个选项含义如下。

- 打开的数据集：指从当前打开的数据集选择合并文件，下面显示的是当前打开的可用数据集名称，由对话框显示可知当前还打开了可用数据文件“data03-03 (2).sav”。
- 外部 SPSS Statistics：指读取外部的数据文件进行合并，该项需要用户指定文件路径和文件名。

用户根据需要选择其中一种方式打开第二个数据文件即可。

单击 **继续** 按钮，弹出“添加个案”对话框，如图 3-11 所示，其中各选项含义如下。



图 3-10 “个案合并”对话框

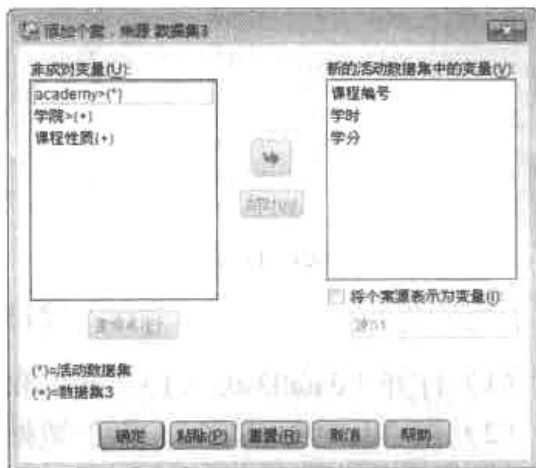


图 3-11 “添加个案”对话框



Note

- 非成对变量：该列表显示两个文件夹中不匹配的变量名，包括变量名不同的变量或变量名相同但变量定义不同的变量。显示在该列表的变量后面都有“*”或“+”，变量后带有“*”表示当前工作的数据文件中的变量，变量后带有“+”表示外部数据文件的变量。此时，可以对变量名进行更改后再选入“新的活动数据集中的变量”，操作方法为：单击 **重命名(N)** 按钮，在弹出如图 3-12 所示的对话框中输入新名称即可。



图 3-12 “重命名”对话框

- 新的活动数据集中的变量：该列表显示合并后的新数据集所包含的变量，默认显示的是两个文件中都有的变量名。
- 将个案源表示为变量：表示合并后的数据集生成一个新变量，用以表示每个个案的来源，SPSS 默认 0 代表个案来自源文件，1 代表来自被合并的文件。选中该复选框，则可激活下面的输入框并指定该新变量的名称，系统默认名称为“源 01”。
- 配对：该命令可以匹配来自两个数据文件中变量名不同、数据的含义和属性相同的变量。操作方法为：在“非成对变量”中选中这两个变量，激活并单击 **配对(P)** 按钮，即可把强行合并后的变量选入“新的活动数据集中的变量”。

将需要在新数据文件中显示的变量选入“新的活动数据集中的变量”，未选入该列表的变量将不进行个案合并。

2. 个案合并的SPSS实现

实例三：数据文件“data03-03 (1).sav”与“data03-03 (2).sav”分别为材料学院等 7 个学院、生物学院等 6 个学院的研究生课程信息，如图 3-13 所示，两个数据文件中包含相同的变量。现要求汇总所有学院的研究生课程信息，以便于后续数据分析。

数据文件：数据文件\Chapter03\data03-03 (1).sav;

数据文件\Chapter03\data03-03 (2).sav

视频文件：视频文件\Chapter03\数据文件的合并.avi

	academy	课程编号	学时	学分
1	1	1	32	2.0
2	1	2	32	2.0
3	1	3	32	2.0
4	1	4	16	1.0
5	1	5	16	1.0
6	1	6	16	1.0
7	1	7	16	1.0

(a) “data03-03 (1).sav” 数据

学院	课程编号	学时	学分	课程性质
8	93	32	2.0	2
8	94	32	2.0	1
8	95	48	3.0	1
8	96	32	2.0	1
8	97	48	3.0	1
8	98	32	2.0	1
8	99	32	2.0	1

(b) “data03-03 (2).sav”

图 3-13 原始数据

(1) 打开“data03-03 (1).sav”和“data03-03 (2).sav”两个数据文件。

(2) 在“data03-03 (1).sav”数据文件中选择“数据”→“合并文件”→“添加个案”，此时弹出如图 3-10 所示对话框。选择“打开的数据集”中的“data03-03 (2).sav”。

(3) 单击 **继续** 按钮进入下一步, 弹出如图 3-11 所示对话框, 同时选中左侧变量列表中的“academy”和“学院”两个变量, 激活并单击 **配对(A)** 按钮, 将两者强行合并后选入“新的活动数据集中的变量”, 显示为 academy & 学院>。



Note

(4) 勾选“将个案源表示为变量”复选框, 采用默认名称“源 01”。

(5) 为了说明当只有一个数据文件包括的变量在合并后所呈现的结果, 将只存在于一个数据文件的“课程性质”变量也选入“新的活动数据集中的变量”。

(6) 完成所有设置后, 单击 **确定** 按钮执行命令。

3. 个案合并的结果分析

图 3-14 为合并后的部分结果。

第一列是“academy”和“学院”两个变量强行合并后所呈现的结果, 因为两者仅变量名不一样, 所以能顺利合并。

最后一列是“源 01”变量, 即表示每个个案来源, 由此可以看出前部分是“data03-03 (1).sav”的数据, 后部分是“data03-03 (2).sav”的数据。

同时, 发现“课程性质”变量存在缺失值, 这是因为只有“data03-03 (2).sav”数据文件有“课程性质”变量, 而“data03-03 (1).sav”没有。

	academy	课程编号	学时	学分	课程性质	源01
90	7	90	32	2.0		0
91	7	91	32	2.0		0
92	7	92	18	1.0		0
93	8	93	32	2.0	2	1
94	8	94	32	2.0	1	1
95	8	95	48	3.0	1	1

图 3-14 个案合并后的部分结果

①若用户要强行合并的两个文件中的某两个变量(即两个变量来自不同的文件)仅具有相同的变量名, 而变量类型不同, 那么强行匹配将失败。②合并后的数据放在当前打开的源数据文件中, 保存后将直接更新此文件的内容, 不可撤销; 若用户希望保留源数据文件及保存合并后的数据文件, 则选择“文件”→“另存为”而非“保存”即可。



注意

3.3.2 变量合并

1. 参数设置

(1) 选择数据文件对话框的参数设置。

打开第一个数据文件, 选择“数据”→“合并文件”→“添加变量”, 弹出“变量合并”对话框, 如图 3-15 所示, 此界面各选项的含义与图 3-10 相似, 此处不再赘述。用户根据需要进行其中一种方式打开第二个数据文件即可。



Note



图 3-15 “变量合并”对话框

(2) 添加变量对话框的参数设置。

单击 **继续** 按钮，弹出“添加变量”对话框，如图 3-16 所示，其中各项设置的含义如下：



图 3-16 “添加变量”对话框

- 已排除的变量：显示的是出现在两个初始文件中，但不出现在合并后的文件里的变量。
- 新的活动数据集：显示合并后的数据集包含的变量。
- 关键变量：显示用以标识和匹配不同文件的个案的变量，当两个数据文件的排列顺序不一致时，可以指定关键变量，但需要先将数据文件按照关键变量值进行升序排列。若未进行排序操作，SPSS 将直接横向合并，这将导致合并错误。
- “匹配关键变量的个案”和“两个数据集中的个案都是按关键变量的顺序进行排序”复选框：选中这两个复选框，说明按关键变量选择个案，有以下三种可选项。
 - 非活动数据集为基于关键字的表：表示关键变量只出现在源文件中，以当前源文件为基准，外部文件匹配源文件的关键变量；若匹配成功，外部文件的新变量就加入到合并后的数据集的新变量中，匹配不成功则不加入。
 - 活动数据集为基于关键字的表：关键变量只出现在被合并的外部文件中，以外部文件为基准，源文件匹配外部文件的关键变量；如匹配成功，源文件的新变量就加入到合并后数据集的新变量中，匹配不成功则不加入。
 - 两个文件都提供个案：表示关键变量同时出现在两个文件中。
- 将个案源表示为变量：勾选“两个数据集中的个案都是按关键变量的顺序进行排序”复选框后激活该对话框，其与图 3-11 的“添加个案”对话框相似。

用户在合并变量时,若两个数据文件没有共同的变量,此时不需指定关键变量,直接将所需合并的变量选入“新的活动数据集”,即可按观测量出现的顺序一对一地合并。

若两个数据文件拥有相同的变量,也有不同的变量,此时可以将相同的变量指定为关键变量,并指定当关键变量值不同时的观测量处理方法。



Note

2. 变量合并的SPSS实现

实例四:“data03-04 (1) .sav”与“data03-04 (2) .sav”分别为课程信息和任课教师信息,如图 3-17 所示,其中“data03-04 (1) .sav”缺少“课程编号”为 2 和 3 的个案,“data03-04 (2) .sav”缺少“课程编号”为 8 和 9 的个案。现要求将后者中的部分变量合并到前者中,并使用三种方式处理关键变量值不同时的观测量。

数据文件: 数据文件\Chapter03\data03-04 (1) .sav;

数据文件\Chapter03\data03-04 (2) .sav

视频文件: 视频文件\Chapter03\数据文件的合并.avi

	学院	课程编号	学时	学分	课程性质
1	1	1	32	2.0	1
2	1	4	16	1.0	1
3	1	5	16	1.0	1
4	1	6	16	1.0	1
5	1	7	16	1.0	1
6	1	8	16	1.0	1
7	1	9	32	2.0	2
8	1	10	32	2.0	2
9	1	11	16	1.0	2
10	1	12	16	1.0	2

(a) “data03-04 (1) .sav” 数据

	学院	课程编号	任课教师姓名	任课教师所在学院
1	1	1	3	1
2	1	2	1	1
3	1	3	1	1
4	1	4	2	1
5	1	5	2	2
6	1	6	2	2
7	1	7	2	2
8	1	10	3	1
9	1	11	1	3
10	1	12	2	3

(b) “data03-04 (2) .sav” 数据

图 3-17 原始数据

- 第一种方式: 选择“非活动数据集为基于关键字的表”选项。

(1) 打开“data03-04 (1) .sav”, 选择“数据”→“合并文件”→“添加变量”, 弹出如图 3-15 所示对话框, 选择“打开的数据集”中的“data03-04 (2) .sav”。

(2) 单击 **继续** 按钮进入下一步, 弹出如图 3-16 所示“添加变量”对话框。勾选“匹配关键变量的个案”和“两个数据集中的个案都是按关键变量的顺序进行排序”复选框, 并选中“非活动数据集为基于关键字的表”, 在“已排除的变量”列表中选中“课程编号”, 然后单击下面的 **添加** 按钮, 将“课程编号”变量选入“关键变量”列表中, 同时“新的活动数据集”列表中的“课程编号”变量消失。

(3) 单击 **确定** 按钮, 此时, SPSS 会弹出一个警告框, 如图 3-18 所示, 提示用户两个文件的数据必须已按关键变量进行了升序排列, 否则关键字匹配将失败。因为在合并变量前, 两个数据文件的“课程编号”均进行升序排列, 所以单击 **确定** 按钮即可。结果如图 3-19 (a) 所示。



图 3-18 关键变量排序的警告框



Note

- 第二种方式：选择“活动数据集为基于关键字的表”选项。

在图 3-16 对话框中选择并勾选“活动数据集为基于关键字的表”，其余步骤与第一种方式相同，结果如图 3-19 (b) 所示。

- 第三种方式：选择“两个文件都提供个案”选项。

在图 3-16 对话框中选择“两个文件都提供个案”，勾选“将个案源表示为变量”并采用默认名称“源 01”。其余步骤与第一种方式相同，结果如图 3-19 (c) 所示。

3. 变量合并的结果分析

图 3-19 (a)、(b)、(c) 是分别选择“非活动数据集为基于关键字的表”、“活动数据集为基于关键字的表”和“两个文件都提供个案”选项处理关键变量值不同的观测量的合并结果。

我们知道，“data03-04(1).sav”缺少“课程编号”为 2 和 3 的个案，“data03-04(2).sav”缺少“课程编号”为 8 和 9 的个案，所以，图 3-19 (a) 图中个案 8 和 9 缺失“任课教师职称”和“任课教师所在学科”的观测量；图 3-19 (b) 图中个案 2 和 3 缺失“学时”和“学分”的观测量；图 3-19 (c) 图中个案 2 和 3 缺失“学时”和“学分”的观测量，个案 8 和 9 缺失“任课教师职称”和“任课教师所在学科”的观测量。

	学院	课程编号	学时	学分	课程性质	任课教师职称	任课教师所在学科
1	1	1	32	2.0	1	3	1
2	1	4	16	1.0	1	2	1
3	1	5	16	1.0	1	2	2
4	1	6	16	1.0	1	2	2
5	1	7	16	1.0	1	2	2
6	1	8	16	1.0	1		
7	1	9	32	2.0	2		
8	1	10	32	2.0	2	3	1

(a) 选择“非活动数据集为基于关键字的表”的部分结果

	课程编号	学院	学时	学分	课程性质	任课教师职称	任课教师所在学科
1	1 1		32	2.0	1	3	1
2	2					1	1
3	3					1	1
4	4 1		16	1.0	1	2	1
5	5 1		16	1.0	1	2	2
6	6 1		16	1.0	1	2	2
7	7 1		16	1.0	1	2	2
8	10 1		32	2.0	2	3	1

(b) 选择“活动数据集为基于关键字的表”的部分结果

	学院	课程编号	学时	学分	课程性质	任课教师职称	任课教师所在学科	源01
1	1	1	32	2.0	1	3	1	1
2		2				1	1	1
3		3				1	1	1
4	1	4	16	1.0	1	2	1	1
5	1	5	16	1.0	1	2	2	1
6	1	6	16	1.0	1	2	2	1
7	1	7	16	1.0	1	2	2	1
8	1	8	16	1.0	1			0
9	1	9	32	2.0	2			0
10	1	10	32	2.0	2	3	1	1

(c) 选择“两个文件都提供个案”的部分结果

图 3-19 变量合并的部分结果

提供各种书籍的pd电子版代找服务，如果你找不到自己想要的书的pdf电子版，我们可以帮您找到，如有需要，请联系QQ1779903665.

PDF代找说明：

本人可以帮助你找到你要的PDF电子书，计算机类，文学，艺术，设计，医学，理学，经济，金融，等等。质量都很清晰，而且每本100%都带书签索引和目录，方便读者阅读观看，只要您提供给我书的相关信息，一般我都能找到，如果您有需求，请联系我QQ1779903665。

本人已经帮助了上万人找到了他们需要的PDF，其实网上有很多PDF,大家如果在网上不到的话，可以联系我QQ，大部分我都可以找到，而且每本100%带书签索引目录。因PDF电子书都有版权，请不要随意传播，如果您有经济购买能力，请尽量购买正版。

声明：本人只提供代找服务，每本100%索引书签和目录，因寻找pdf电子书有一定难度，仅收取代找费用。如因PDF产生的版权纠纷，与本人无关，我们仅仅只是帮助你寻找到你要的pdf而已。