

清华
电脑学堂

DVD

超值多媒体光盘

大容量、高品质多媒体教程
语音视频演示讲解
实例素材、效果和模板

- ✓ 总结了作者多年SPSS应用经验和教学心得
- ✓ 系统讲解了SPSS统计分析的要点和难点
- ✓ 实例丰富、实用性强
- ✓ 附大容量、高品质多媒体语音视频教程光盘

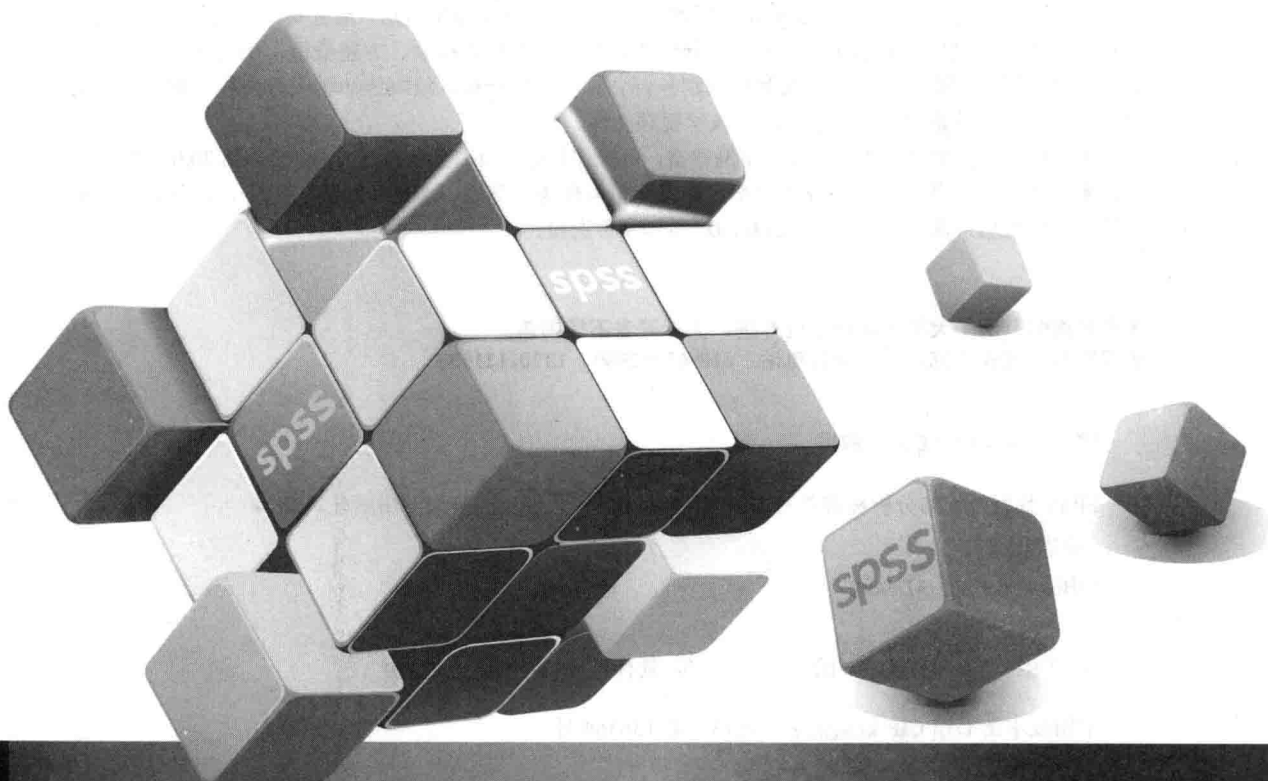


SPSS 数据统计与分析

标准教程

■ 夏丽华 谢金玲 等编著

清华大学出版社



SPSS 数据统计与分析

标准教程

■ 夏丽华 谢金玲 等编著

清华大学出版社
北 京

内 容 简 介

SPSS 20.0 是 IBM 公司最新推出的数据统计分析软件,是各行业办公人员必备的统计分析工具之一。本书汇集了用户在使用 SPSS 20.0 分析数据时最常用的基础分析方法,通过实用案例,详细向用户解释了管理数据文件、基本统计分析、差异性均值分析、非参数检验、方差分析、相关分析、回归分析、信度分析、对数线性分析、聚类和判别分析、时间序列分析、绘制统计图表等内容。配书光盘提供了本书实例完整素材文件和全程配音教学视频文件。

本书结构编排合理,图文并茂,实例丰富,可操作性强,可有效帮助用户提升 SPSS 20.0 的操作水平。本书适合作为高职高专院校学生学习使用,也可作为计算机、金融、教育、医疗,以及具备一定统计学基础的用户深入学习 SPSS 20.0 的培训和参考资料。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

SPSS 数据统计与分析标准教程/夏丽华等编著. —北京:清华大学出版社, 2014

(清华电脑学堂)

ISBN 978-7-302-32890-2

I. ①S… II. ①夏… III. ①统计分析-软件包-教材 IV. ①C819

中国版本图书馆 CIP 数据核字(2013)第 136364 号

责任编辑:冯志强

封面设计:吕单单

责任校对:徐俊伟

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京鑫丰华彩印有限公司

装 订 者:三河市溧源装订厂

经 销:全国新华书店

开 本:185mm×260mm 印 张:23 插 页:1 字 数:549 千字

(附光盘 1 张)

版 次:2014 年 5 月第 1 版

印 次:2014 年 5 月第 1 次印刷

印 数:1~3000

定 价:49.00 元

前 言

SPSS 20.0 是 IBM 公司最新推出的数据分析软件,也是当前世界上最为流行的三大统计分析软件之一。它以强大的统计分析功能、方便易用的用户操作方式、灵活的表格式分析报告和精美的图形展示形式,以及可以直接读取 Excel 及 DBF 数据文档优势,充分赢得了各个领域广大数据分析人员的喜爱,并得到了广泛的应用。

本书以 SPSS 20.0 中的基础分析方法出发,配以大量实例,采用知识点讲解与动手练习相结合的方式,详细介绍了 SPSS 20.0 中的数据分析基础应用知识与具体使用方法。每一章都配合了丰富的插图说明,生动具体、浅显易懂,使用户能够迅速上手,轻松掌握功能强大的 SPSS 20.0 在日常生活与数据分析中的应用,为工作和学习带来事半功倍的效果。

1. 本书内容介绍

全书系统全面地介绍了 SPSS 20.0 的基础分析知识,每个知识点下面都提供了具体实例,用来巩固所学技巧。本书共分为 13 章,内容概括如下:

第 1 章:全面讲解了初始 SPSS 20.0 的基础知识,介绍了 SPSS 发展历史、SPSS 的功能性、SPSS 软件安装、SPSS 软件介绍等内容。

第 2 章:全面讲解了管理数据文件的应用知识,介绍了定义变量、编辑变量、设置 SPSS 数据、拆分和合并数据文件等内容。

第 3 章:全面讲解了描述统计分析的应用知识,介绍了频数分析、描述分析、交叉表分析、探索性统计分析、比率分析等内容。

第 4 章:全面讲解了差异性均值分析的应用知识,介绍了假设检验概述、均值比较、单样本 T 检验、独立样本 T 检验等内容。

第 5 章:全面讲解了非参数检验的应用知识,介绍了二项式检验、卡方检验、K-S 检验、游程检验、Wilcoxon 符号秩检验、独立样本非参数检验、相关样本非参数检验等内容。

第 6 章:全面讲解了方差分析的应用知识,介绍了单因素方差分析、双因素方差分析、多元方差分析、重复测量的双因素方差分析、协方差分析等内容。

第 7 章:全面讲解了相关分析的应用知识,介绍了双变量相关分析、偏相关分析、距离相关分析等内容。

第 8 章:全面讲解了回归分析的应用知识,介绍了一元线性回归分析、多元线性回归分析、非线性回归分析、曲线估计回归分析等内容。

第 9 章:全面讲解了信度分析的应用知识,介绍了重测和复本信度、内部一致性信度、分半信度、评分信度、多维尺度分析等内容。

第 10 章:全面讲解了对数线性分析的应用知识,介绍了常规模型、Logit 模型、模型选择等内容。

第 11 章：全面讲解了聚类 and 判别分析的应用知识，介绍了二阶聚类分析、K-均值聚类分析、层次聚类分析、普通判别分析、决策树分析等内容。

第 12 章：全面讲解了时间序列分析的应用知识，介绍了时间序列数据的预处理、指数平滑模型、ARIMA 模型、季节分析模型等内容。

第 13 章：全面讲解了绘制统计图的应用知识，介绍了条形图、线图、面积图和饼图、散点图、高低图、箱图、编辑统计图等内容。

2. 本书主要特色

- **系统全面** 本书提供了 50 多个实用案例，通过实例分析、设计过程讲解 SPSS 20.0 的应用知识，涵盖了 SPSS 20.0 中的各个分析工具。
- **课堂练习** 本书各章都安排了课堂练习，全部围绕实例讲解相关内容，灵活生动地展示了 SPSS 20.0 各分析方法的具体功能。课堂练习体现本书实例的丰富性，方便读者组织学习。每章后面还提供了思考与练习，用来测试读者对本章内容的掌握程度。
- **全程图解** 各章内容全部采用图解方式，图像均做了大量的裁切、拼合、加工，信息丰富，效果精美，阅读体验轻松，上手容易。
- **随书光盘** 本书使用 Director 技术制作了多媒体光盘，提供了本书实例完整素材文件和全程配音教学视频文件，便于读者自学和跟踪练习图书内容。

3. 本书使用对象

本书从 SPSS 20.0 基础分析知识入手，全面介绍了 SPSS 20.0 面向数据分析应用的知识体系。本书制作了多媒体光盘，图文并茂，能有效吸引读者学习。本书适合作为高职高专院校学生学习使用，也可作为数据分析爱好者及办公人员等用户深入学习 SPSS 20.0 的培训和参考资料。

参与本书编写的人员除了封面署名人员之外，还有王海峰、马玉仲、席宏伟、祁凯、徐恺、王泽波、王磊、张仕禹、夏小军、赵振江、李振山、李文才、李海庆、王树兴、何永国、李海峰、王蕾、王曙光、牛小平、贾栓稳、王立新、苏静、赵元庆、郭磊、何方、徐铭、李大庆等人。由于时间仓促，水平有限，疏漏之处在所难免，敬请读者朋友批评指正。

编 者

目 录

第 1 章 初识 SPSS	1	2.5 课堂练习: 读取 Excel 数据文件	38
1.1 SPSS 概述	2	2.6 课堂练习: 管理员工信息数据文件	40
1.1.1 SPSS 发展历史	2	2.7 思考与练习	42
1.1.2 SPSS 的功能性	2	第 3 章 描述统计分析	44
1.1.3 SPSS 的产品特点	3	3.1 频数分析	45
1.2 SPSS 软件安装	4	3.1.1 频数分析概述	45
1.2.1 SPSS 软件的环境要求	4	3.1.2 设置频率统计量	46
1.2.2 安装 SPSS 软件	4	3.1.3 设置频率格式	47
1.3 SPSS 软件介绍	6	3.1.4 设置频率图表	48
1.3.1 启用 SPSS 软件	6	3.1.5 Bootstrap 分析方法	48
1.3.2 SPSS 数据编辑窗口	7	3.1.6 显示分析结果	50
1.3.3 SPSS 结果输出窗口	10	3.2 描述分析	50
1.3.4 SPSS 语法编辑窗口	10	3.2.1 描述分析概述	50
1.3.5 SPSS 脚本编辑窗口	11	3.2.2 SPSS 描述统计分析方法	52
第 2 章 管理数据文件	12	3.3 交叉表分析	54
2.1 变量与常量	13	3.3.1 交叉表分析概述	54
2.1.1 理解常量	13	3.3.2 添加交叉表变量	55
2.1.2 理解操作符与表达式	13	3.3.3 交叉表中的精确检验	55
2.1.3 定义变量	14	3.3.4 交叉表中的统计量	56
2.1.4 编辑变量	19	3.3.5 交叉表中的单元显示	57
2.1.5 指定加权变量	20	3.3.6 显示交叉表分析结果	58
2.1.6 创建变量	20	3.4 探索性统计分析	59
2.2 设置 SPSS 数据	22	3.4.1 探索性统计分析概述	59
2.2.1 输入数据	22	3.4.2 设置统计变量	60
2.2.2 读取数据文件	24	3.4.3 设置探索统计量	60
2.2.3 保存数据文件	29	3.4.4 设置探索图	61
2.3 编辑 SPSS 数据	31	3.4.5 显示探索分析结果	62
2.3.1 复制与移动数据	31	3.5 比率分析	63
2.3.2 设置字体格式	32	3.5.1 设置比率变量	63
2.3.3 查找和替换数据	33	3.5.2 设置比率统计量	63
2.3.4 编辑个案	34	3.5.3 保存/显示分析结果	64
2.4 拆分与合并数据文件	36	3.6 P-P 图和 Q-Q 图	65
2.4.1 拆分数据文件	36	3.6.1 P-P 图	65
2.4.2 合并数据文件	36	3.6.2 Q-Q 图	66

3.7 课堂练习: 交叉表分析工作年限	67	5.3.4 显示分析结果	100
3.8 课堂练习: 频数分析考试成绩	69	5.4 卡方检验	101
3.9 课堂练习: 销售比率分析	71	5.4.1 卡方检验概述	101
3.10 思考与练习	72	5.4.2 添加分析变量	102
第4章 差异性均值分析	74	5.4.3 设置检验选项	102
4.1 假设检验概述	75	5.4.4 显示分析结果	103
4.1.1 假设检验基本原理	75	5.5 K-S 检验	103
4.1.2 两种类型的错误	75	5.5.1 K-S 检验概述	103
4.1.3 单侧与双侧检验	76	5.5.2 添加分析变量	104
4.2 均值比较	77	5.5.3 设置检验选项	104
4.2.1 添加均值比较变量	77	5.5.4 显示分析结果	105
4.2.2 设置均值比较选项	77	5.6 Wilcoxon 符号秩检验	106
4.2.3 设置 Bootstrap 选项	78	5.6.1 添加分析变量	106
4.2.4 显示均值比较结果	79	5.6.2 设置分析选项	106
4.3 单样本 T 检验	79	5.6.3 显示分析结果	106
4.3.1 单样本 T 检验概述	79	5.7 游程检验	107
4.3.2 使用单样本 T 检验	80	5.7.1 游程检验概述	107
4.4 独立样本 T 检验	81	5.7.2 添加分析变量	108
4.4.1 独立样本 T 检验概述	81	5.7.3 设置分析选项	108
4.4.2 使用独立样本 T 检验	82	5.7.4 显示分析结果	109
4.5 配对样本 T 检验	84	5.8 独立样本非参数检验	110
4.5.1 配对样本 T 检验概述	84	5.8.1 独立样本非参数检验概述	110
4.5.2 使用配对样本 T 检验	84	5.8.2 设置检验目标	111
4.6 课堂练习: 均值比较考试成绩	85	5.8.3 设置检验字段	112
4.7 课堂练习: 独立样本 T 检验身高数据	88	5.8.4 设置检验类型	112
4.8 课堂练习: 配对样本 T 检验体重数据	90	5.9 相关样本非参数检验	114
4.9 思考与练习	91	5.9.1 相关样本非参数检验概述	114
第5章 非参数检验	94	5.9.2 设置检验目标	116
5.1 非参数检验概述	95	5.9.3 设置检验字段	117
5.2 单样本非参数检验概述	95	5.9.4 设置检验类型	117
5.2.1 设置检验目标	96	5.10 课堂练习: 分析考试成绩的 差异性	119
5.2.2 设置检验字段	96	5.11 课堂练习: 评价招聘方案	122
5.2.3 设置检验类型	97	5.12 课堂练习: 分析血糖值的 合格度	124
5.3 二项式检验	98	5.13 思考与练习	126
5.3.1 二项式检验概述	98	第6章 方差分析	129
5.3.2 添加分析变量	99	6.1 方差分析概述	130
5.3.3 设置检验选项	99	6.1.1 方差分析的因素	130
		6.1.2 方差分析的假设条件	130

6.1.3	方差分析的基本流程	131
6.2	单因素方差分析	133
6.2.1	单因素方差概述	133
6.2.2	正态分布检验	133
6.2.3	设置对比方式	134
6.2.4	设置方差齐性	135
6.2.5	设置分析选项	136
6.2.6	设置附加选项	137
6.2.7	显示分析结果	137
6.3	双因素方差分析	139
6.3.1	设置单变量模型	139
6.3.2	设置对比方式	141
6.3.3	设置轮廓图	142
6.3.4	设置均值的比较方式	142
6.3.5	设置保存选项	143
6.3.6	设置分析选项	145
6.3.7	显示分析结果	146
6.4	多元方差分析	147
6.4.1	设置变量模型	147
6.4.2	设置对比方式	149
6.4.3	设置轮廓图	149
6.4.4	设置两两比较方式	149
6.4.5	设置保存方式	150
6.4.6	设置分析选项	152
6.4.7	显示分析结果	153
6.5	重复测量的双因素方差分析	155
6.5.1	添加测量因子	155
6.5.2	设置轮廓图	155
6.5.3	设置选项	156
6.5.4	显示分析结果	156
6.6	协方差分析	158
6.6.1	协方差概述	158
6.6.2	检验交互性	159
6.6.3	协方差分析数据	160
6.6.4	显示分析结果	161
6.6	课堂练习：重复测量法分析考试成绩	162
6.7	课堂练习：协方差分析空腹血糖	165
6.8	思考与练习	168

第7章	相关分析	170
7.1	相关分析概述	171
7.1.1	函数关系和相关关系	171
7.1.2	相关关系的分类	171
7.1.3	相关系数	172
7.2	双变量相关分析	173
7.2.1	双变量相关分析方法	173
7.2.2	设置分析变量	174
7.2.3	设置分析选项	174
7.2.4	显示分析结果	175
7.3	偏相关分析	176
7.3.1	偏相关分析概述	176
7.3.2	设置分析变量	176
7.3.3	设置分析选项	177
7.3.4	显示分析结果	177
7.4	距离相关分析	178
7.4.1	距离相关分析概述	178
7.4.2	非相似性测量	179
7.4.3	相似性测量	182
7.5	课堂练习：分析小麦的产量因素	184
7.6	课堂练习：分析公司间的执行能力	186
7.7	思考与练习	189
第8章	回归分析	191
8.1	线性回归分析	192
8.1.1	一元线性回归分析概述	192
8.1.2	一元线性回归分析	192
8.1.3	多元线性回归概述	198
8.1.4	多元线性回归分析	198
8.2	非线性回归分析	202
8.2.1	非线性回归概述	202
8.2.2	初步分析数据	204
8.2.3	设置参数	205
8.2.4	设置损失函数	206
8.2.5	设置参数约束	206
8.2.6	设置保存和分析选项	207
8.2.7	显示分析结果	208
8.3	曲线估计回归分析	209

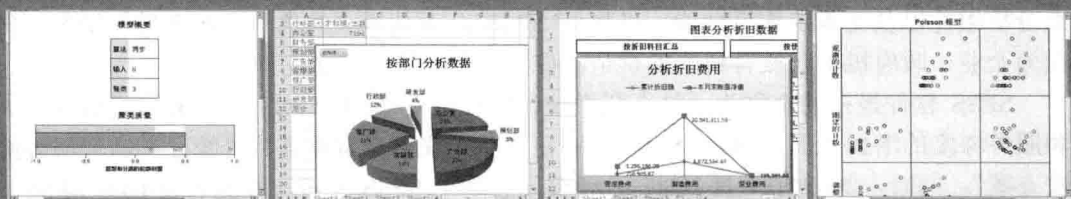
8.3.1 曲线估计回归模型	209	9.9 思考与练习	238
8.3.2 设置分析变量和模型	210	第 10 章 对数线性分析	240
8.3.3 设置保存选项	211	10.1 对数线性分析概述	241
8.3.4 显示分析结果	211	10.1.1 列联表分析概述	241
8.4 课堂练习:非线性回归分析		10.1.2 对数线性模型的表现形式	241
推广费用	213	10.2 常规模型	242
8.5 课堂练习:曲线估计分析		10.2.1 常规模型概述	242
企业利润	216	10.2.2 添加分析变量	242
8.6 思考与练习	218	10.2.3 设置模型选项	243
第 9 章 信度分析	220	10.2.4 设置保存选项	244
9.1 信度分析概述	221	10.2.5 设置分析选项	244
9.1.1 信度分析的基本理论	221	10.2.6 显示分析结果	245
9.1.2 信度系数	222	10.3 Logit 模型	246
9.2 重测和复本信度	222	10.3.1 Logit 模型概述	246
9.2.1 重测信度分析	222	10.3.2 添加分析变量	247
9.2.2 复本信度分析	223	10.3.3 设置分析选项	247
9.3 内部一致性信度	224	10.3.4 设置保存选项	248
9.3.1 内部一致性信度概述	224	10.3.5 设置模型选项	248
9.3.2 设置分析变量	225	10.3.6 显示分析结果	248
9.3.3 设置统计量	226	10.4 模型选择	249
9.3.4 显示分析结果	227	10.4.1 设置分析变量	249
9.4 分半信度	228	10.4.2 设置分析模型	250
9.4.1 分半信度概述	228	10.4.3 设置分析选项	250
9.4.2 双变量执行信度分析	228	10.4.4 显示分析结果	250
9.4.3 分半信度分析法	229	10.5 课堂练习:分析早餐的影响因素	251
9.5 评分信度	230	10.6 课堂练习:分析日常娱乐项目	254
9.5.1 评分信度概述	230	10.7 思考与练习	256
9.5.2 设置分析变量	230	第 11 章 聚类和判别分析	258
9.5.3 设置检验的精确程度	231	11.1 聚类和判别分析概述	259
9.5.4 设置分析选项	232	11.1.1 聚类分析概述	259
9.5.5 显示分析结果	232	11.1.2 判别分析概述	261
9.6 多维尺度分析	232	11.2 二阶聚类分析	262
9.6.1 设置分析变量	232	11.2.1 添加分析变量	262
9.6.2 设置分析模型	233	11.2.2 设置分析选项	263
9.6.3 设置分析选项	234	11.2.3 设置输出选项	264
9.6.4 显示分析结果	234		
9.7 课堂练习:分析成绩的内部一致性信度	235		
9.8 课堂练习:尺度分析城市距离	236		

11.2.4 显示分析结果·····	264	12.2 时间序列数据的预处理·····	302
11.3 K-均值聚类分析·····	267	12.2.1 缺失值替换·····	302
11.3.1 添加分析变量·····	268	12.2.2 定义日期变量·····	303
11.3.2 设置迭代选项·····	268	12.2.3 创建时间序列·····	304
11.3.3 设置保存选项·····	269	12.3 指数平滑模型·····	306
11.3.4 设置分析选项·····	269	12.3.1 添加分析变量·····	306
11.3.5 显示分析结果·····	270	12.3.2 设置分析条件·····	306
11.4 层次聚类分析·····	271	12.3.3 设置统计量选项·····	307
11.4.1 设置分析变量·····	271	12.3.4 设置图表选项·····	308
11.4.2 设置统计量选项·····	272	12.3.5 设置输出过滤选项·····	309
11.4.3 设置分析图·····	272	12.3.6 设置保存选项·····	309
11.4.4 设置聚类方法·····	273	12.3.7 设置分析选项·····	310
11.4.5 设置保存选项·····	275	12.3.8 显示分析结果·····	311
11.4.6 显示分析结果·····	276	12.4 ARIMA 模型·····	311
11.5 普通判别分析·····	277	12.4.1 添加分析变量·····	311
11.5.1 设置分析变量·····	277	12.4.2 设置模型类别·····	312
11.5.2 设置统计量·····	278	12.4.3 设置离群值·····	313
11.5.3 设置步进法分析方法·····	279	12.4.4 设置统计量和图表 选项·····	314
11.5.4 设置分类选项·····	280	12.4.5 显示分析结果·····	314
11.5.5 设置保存选项·····	281	12.4.6 更改分析设置·····	315
11.5.6 显示分析结果·····	281	12.5 季节分析模型·····	316
11.6 决策树分析·····	282	12.5.1 添加分析变量·····	316
11.6.1 设置分析变量·····	283	12.5.2 设置保存选项·····	316
11.6.2 设置输出选项·····	284	12.5.3 显示分析结果·····	317
11.6.3 设置验证条件·····	287	12.5.4 制作分析图表·····	317
11.6.4 设置标准条件选项·····	288	12.6 课堂练习: 分析景区游客量·····	318
11.6.5 设置保存选项·····	290	12.7 课堂练习: 分析国民经济指标·····	320
11.6.6 设置分析选项·····	290	12.8 思考与练习·····	323
11.6.7 显示分析结果·····	291	第 13 章 绘制统计图·····	325
11.7 课堂练习: 二阶聚类分析 学历水平·····	292	13.1 条形图·····	326
11.8 课堂练习: 判别分析学校水平·····	295	13.1.1 简单条形图·····	326
11.9 思考与练习·····	297	13.1.2 复式条形图·····	328
第 12 章 时间序列分析·····	300	13.1.3 堆积条形图·····	329
12.1 时间序列分析概述·····	301	13.1.4 3-D 条形图·····	330
12.1.1 时间序列的分类·····	301	13.1.5 简单误差条形图·····	331
12.1.2 时间序列的特点·····	301	13.1.6 复式误差条形图·····	333
12.1.3 时间序列的构成因素·····	302	13.2 线图·····	334
		13.2.1 简单线图·····	334
		13.2.2 多线线图·····	335
		13.2.3 垂直线图·····	337
		13.3 面积图和饼图·····	338

13.3.1 简单面积图	338	13.6.1 简单箱图	346
13.3.2 堆积面积图	338	13.6.2 复式箱图	347
13.3.3 饼图	339	13.6.3 对比箱图	348
13.4 散点图	340	13.7 编辑统计图	348
13.4.1 简单分布散点图	340	13.7.1 添加数据标签	348
13.4.2 矩阵散点图	340	13.7.2 美化统计图	349
13.4.3 重叠散点图	341	13.7.3 设置排序方式	350
13.4.4 3-D 散点图	342	13.7.4 设置坐标轴	351
13.5 高低图	342	13.8 制作交互式统计图表	352
13.5.1 简单高低关闭图	343	13.8.1 使用图表构建程序	352
13.5.2 简单范围栏图	344	13.8.2 使用图形画板模板	353
13.5.3 聚类高低关闭图	344	13.9 课堂练习：绘制 3-D 散点图	354
13.5.4 差别面积图	345	13.10 课堂练习：绘制复式箱图	355
13.6 箱图	346	13.11 思考与练习	357

第1章

初识 SPSS



SPSS (Statistical Product and Service Solutions, SPSS), 即统计产品与服务解决方案软件, 是当前世界上最为流行的三大统计分析软件之一。它以强大的统计分析功能、方便易用的用户操作方式、灵活的表格式分析报告和精美的图形展示形式, 以及可以直接读取 Excel 及 DBF 数据文档优势, 充分赢得了各个领域广大数据分析人员的喜爱, 并得到了广泛的应用。

在本章中, 将从 SPSS 软件的产生、发展、功能与特点, 以及 SPSS 软件的安装、界面介绍等方面, 详细介绍 SPSS 软件的基础知识, 帮助用户认识并熟悉 SPSS 软件。

本章学习目标:

- SPSS 概述
- SPSS 软件安装
- SPSS 软件介绍
- SPSS 数据统计分析基础步骤

1.1 SPSS 概述

SPSS 是世界上最早的统计分析软件，它是一个组合式的软件包，集数据整理、分析功能于一身，不仅可以为用户提供数据管理、统计分析、趋势分析、制表、绘图服务，还可用于计划、经济、教育、心理、医学、生物、气象及其他社会科学领域。

1.1.1 SPSS 发展历史

SPSS 是由美国斯坦福大学中的 Norman H 等三位研究生于 1968 年研发成功，同时于 1975 年在芝加哥成立了 SPSS 公司。1984 年 SPSS 总部首先推出了世界上第一个统计分析软件的微机版本 SPSS/PC+，并得到了广泛的应用与高度的评价。

SPSS 公司于 1992 年推出了 Windows 版本统计软件，同时开始了全球化的发展。1994—1998 年，SPSS 公司先后并购了 SYSTAT、BMDP、Quantime、ISL 等公司，通过收购引进了数据挖掘技术、为 IBM 的 eServer iSeries 开发的商业智能套件、Web 分析、复杂分析组件以及文本挖掘等功能，进而使 SPSS 公司从原来的单一产品开发与销售，转向为企业、政府机构和教育科研提供全面信息统计决策的支出与服务。

SPSS 软件最初的全称为“社会科学统计软件包”，随着 SPSS 产品服务领域的扩大和服务深度的增加，于 2000 年 SPSS 公司正式将英文名全称更改为“统计产品与服务解决方案”。同时，2009 年 7 月 28 日 IBM 公司收购了 SPSS 公司，并更名为 IBM SPSS，对于 SPSS 而言，SPSS 预测分析技术与 IBM 的数据库软件、商业智能软件和咨询顾问服务更好地结合，能够更完善地实现预测分析技术及市场应用解决方案。而适用所有行业的软件功能进一步进行完善，包括了需求预测、员工聘用与保留、客户盈利率、信用评分与欺诈检测等服务。如今 SPSS 已研发出版本 20.0，而目前中最新版本 21.0 只有 Linux 版本。迄今，SPSS 公司已有 40 余年的历史，其具体发展历史概述如表 1-1 所述。

表 1-1 SPSS 发展历史概述

年 份	发展概述
1968 年	斯坦福大学三位研究生创建了 SPSS
1968 年	诞生了第一个用于大型机的统计分析软件
1975 年	在芝加哥成立了 SPSS 总部
1984 年	推出了用于个人电脑的 SPSS/PC+
1992 年	推出了 Windows 版本，同时自 SPSS 11.0 版本开始，SPSS 全名更改为“Statistical Product and Service Solutions”，即“统计产品和解决服务方案”
2009 年	SPSS 公司重新包装 SPSS 产品线，并定位为预测统计分析软件（Predictive Analytics Software）PASW，包括统计分析、数据挖掘、数据收集和企业应用服务 4 部分
2010 年	SPSS 软件的名称由 PASW SPSS 统一更改为 IBM SPSS

1.1.2 SPSS 的功能性

SPSS 是一个组合式软件包，集数据整理、数据分析功能于一身，也是世界上最早采

用图形菜单驱动界面的一款统计软件，它与 SAS 和 BMDP 并称为最有影响力的三大统计软件。SPSS 为方便用户使用，将大部分经常使用的功能以统一、规范的界面进行展现，使用 Windows 窗口方式展示各种统计与分析数据的功能，并使用对话框展示各种功能的设置选项。

SPSS 基本功能包括数据管理、统计分析、图表分析、输出管理等。而 SPSS 软件的统计分析过程，则包括描述性统计、聚类统计、数据简化、时间序列分析、回归分析、相关分析等分析类型。其中，SPSS 中比较突出的一些功能如下所述。

1. 数据管理

从 SPSS 10.0 版本之后，每个新增版本都会对数据管理功能方法进行一些有效的改进，以帮助用户可以更方便地使用 SPSS 软件。例如，将变量名由 64 个字符的规定进行放大，以可以兼容各种复杂的数据库。另外，改进 Autorecorda 过程，可以使用自动编码模块，方便用户按照自定义的顺序进行变量值的重编码。

2. 结果报告

在 SPSS 16.0 版本中，SPSS 在推出了全新的常规图功能之后，完善了 SPSS 的报表功能。在常规中的统计图中，除了引用了图组、带误差线的分类图形、线形、误差线条形图、堆积和分段饼图等交互图功能之外，还引用了人口金字塔、点密度图等新图形。而在常规中的统计表中，所有过程的输出都更改为美观的枢轴表。利用枢轴表不仅可以对统计数据进行排序、合并或省略表格中若干小类的输入等功能，而且还可以将枢轴表直接导出到 PowerPoint 中。

3. 统计建模

在 SPSS 12.0 版本中引用了 Complex Samples 模板，以满足用户进行复杂抽样设计的需求，而该建模的操作方式相似于完全随机抽样数据的分析操作。另外，在 13.0 版本中又为该模块引进了一般线性模型，以满足用户对复杂抽样研究中各种连续性变量的建模预测。同时，还为该模块引入了用于分类数据的 Logistic 回归模型。这样一来，用户在进行多阶段分层整群抽样，或者更复杂的 PPS 抽样时，不仅可以在该模块中轻松地实现从抽样设计、统计描述到复杂统计建模的整个分析过程，还可以使用相对复杂的方差分析模型、线性回归模型等统计模型。

1.1.3 SPSS 的产品特点

SPSS 分析结果清晰、直观、易学易用，并采用了通用的类似于 Excel 表格方式的输入与管理数据接口，可以方便、快速地直接读取 Excel 与 DBF 数据文件；但是 SPSS 软件却很难与 Office 等一般办公软件相兼容，用户只能使用复制的方式对其进行交互。其中，SPSS 软件主要包括下面一些特点：

- **操作简便** 具有便于操作与友好的界面，方便用户使用鼠标进行各种操作。
- **编程方便** SPSS 具有第 4 代语言的特点，用户只需了解统计学原理，即可对各

类数据进行专业分析。在该软件中，用户无需学习统计专业中的各种计算方法，只需花费些时间记忆软件中的命令、过程与统计信息的选项即可。

- ❑ **功能强大** SPSS 自带了 11 种类型的 130 多个函数，并提供了多种统计分析方法，以帮助用户完成数据输入、编辑、统计分析、报表与图形制作等数据统计与分析操作。
- ❑ **数据接口** SPSS 软件还具有读取与输出 Dbase、FoxBASE、FoxPRO 产生的*.dbf 文件，以及文本编辑器软件产生的 SCSII 数据文件、Excel 产生的*.xls 文件等多种格式的文件。另外，SPSS 软件还可以将结果保存为*.txt 及 html 格式的文件。
- ❑ **模块组合** SPSS 软件为用户提供了多种功能模块，便于用于根据分析需求选择不同的模板类型，从而体现了自身软件的灵活性与互动性。

1.2 SPSS 软件安装

通过对前面 SPSS 概述的了解，用户已经掌握了 SPSS 软件的功能与特点，在本小节中重点介绍 SPSS 软件的安装过程。

1.2.1 SPSS 软件的环境要求

在安装 SPSS 软件之前，为了保证该软件的正常运行，还需要查看本地计算机的硬件与软件配置。

SPSS 软件对计算机配置的要求比较低，一般比较主流的计算机都适合安装 SPSS 软件。随着 SPSS 软件版本不断升级，为了保证 SPSS 软件运行的流畅性，还需要保持计算机最低硬件为：

- ❑ **CPU** 一般要求 CPU 位于 1GHz 或更高。
- ❑ **内存** 内存要求位于 516MB 以上。
- ❑ **运行环境** 运行环境没有特殊要求，SPSS 软件可以在 Windows XP 或 Windows Vista，以及 Windows 7 以上系统中运行。
- ❑ **硬盘空间** 硬盘要求大于 650MB 即可。

1.2.2 安装 SPSS 软件

SPSS 软件的安装过程与其他软件安装大体一致，用户直接运行光盘文件，或打开包含安装文件的文件夹，双击“setup”文件，弹出 SPSS 安装向导对话框，并单击【下一步】按钮，如图 1-1 所示。

在弹出的要求信息界面中，直接单击【下一步】按钮。选中【单个用户许可证】选项，单击【下一步】按钮，如图 1-2 所示。

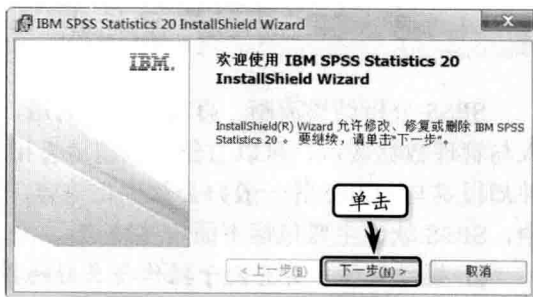


图 1-1 安装向导

提示

【单个用户许可证】表示在下面输入认证码时，需要提供个人用户的认证码。否则需要提供站点或网络要求的认证码。

在弹出的【软件许可协议】向导界面中，选中【我接受许可协议中的全部条款】选项，并单击【下一步】按钮，如图 1-3 所示。

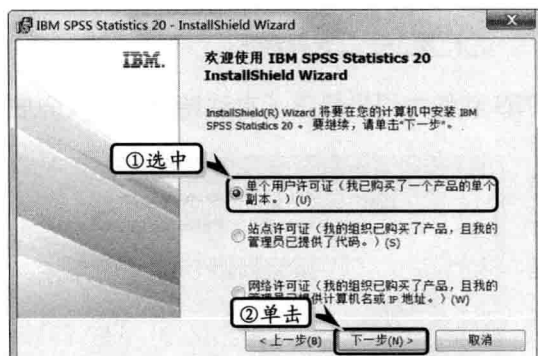


图 1-2 选中用户类别

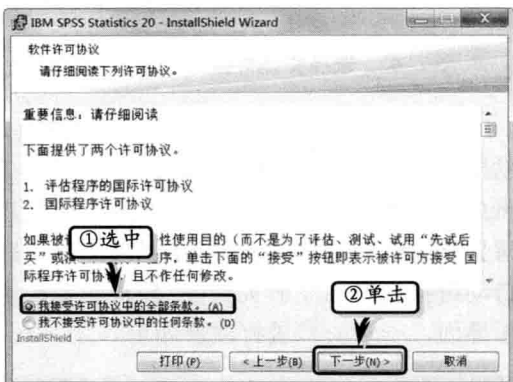


图 1-3 接受许可协议

在弹出的【客户信息】向导界面中，在【用户姓名】与【单位】文本框中，分别输入用户的姓名与单位名称，并单击【下一步】按钮，如图 1-4 所示。

此时，系统自动弹出【帮助语言】向导界面，保持选择【简体中文】选项，禁用其他语言选项，并单击【下一步】按钮，如图 1-5 所示。

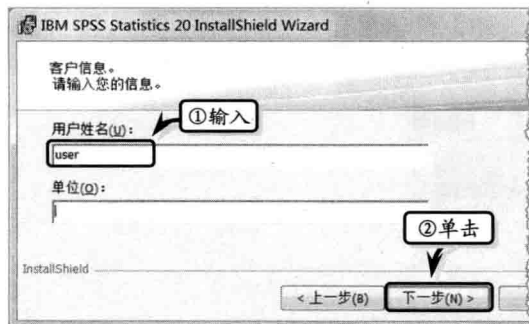


图 1-4 输入客户信息

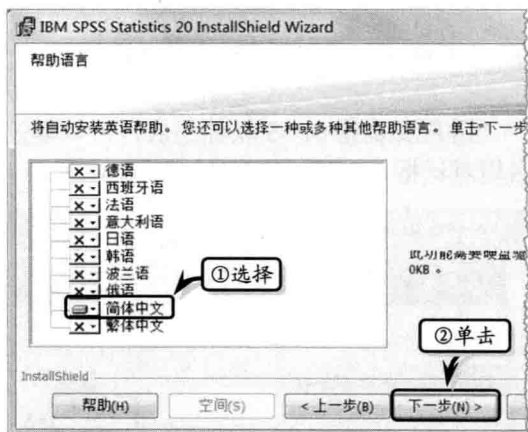


图 1-5 选择语言类别

在弹出的【目的地文件夹】向导界面中，系统会将软件默认并保存在 C 盘内，用户可选择默认的保存位置，也可以通过单击【更改】按钮，重新选择安装路径。在此，保存默认文件安装路径，并单击【下一步】按钮，如图 1-6 所示。

在弹出的【已做好安装程序的准备】向导界面中，单击【安装】按钮，系统便会自动安装软件，如图 1-7 所示。



图 1-6 选择安装路径

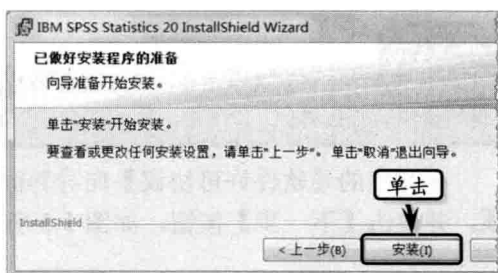


图 1-7 安装软件

最后,系统会自动弹出安装完成界面,SPSS 软件为用户提供了自动链接注册页面的功能,如图 1-8 所示。在该界面中,启用【单击此处进行注册,以获取产品更新与优惠提醒】复选框,在单击【确定】按钮之后,即可自动链接到产品注册页面中,用户只需按照提示界面,一步一步进行注册即可。

1.3 SPSS 软件介绍

安装完 SPSS 软件后,用户便可以使用该软件进行数据分析了。但是,对于新手来讲,还需要先熟悉一下 SPSS 软件的基本功能,为日后进行各种复杂的数据分析打下坚实的基础。

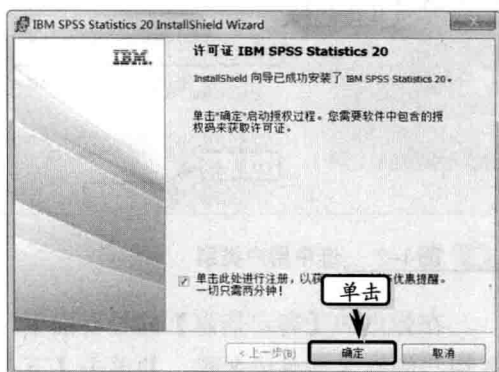


图 1-8 完成安装

1.3.1 启用 SPSS 软件

用户安装完 SPSS 软件之后,可以通过【开始】菜单运行 SPSS 软件,进入到 SPSS 启用对话框,如图 1-9 所示。

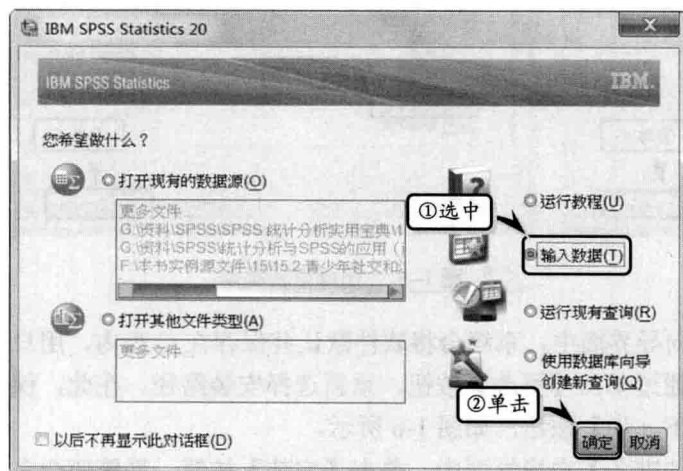


图 1-9 启动对话框