

菜鸟学

每一个案例都有详细的阐述
每一个结果都有完整的解读

SPSS 数据分析

60 个案例带你轻松学

但婉欣 蒋胡英 郭佳桦 著



中国工信出版集团



电子工业出版社
ELECTRONIC INDUSTRY PRESS
<http://www.phei.com.cn>

菜鸟学 SPSS 数据分析

但婉欣 蒋胡英 郭佳桦 著



電子工業出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书深入浅出地介绍了 SPSS 24.0 的基本操作和数据分析方法。全书共有 12 章，主要分为两大部分：第一部分为 SPSS 与数据分析基础知识（第 1~5 章），主要介绍 SPSS 界面功能、SPSS 与数据分析、数据的管理、报表的生成、图形的生成；第二部分为数据分析方法（第 6~12 章），主要介绍描述性分析、假设检验、参数检验、非参数检验、卡方检验、方差分析、相关分析。

本书内容不仅涵盖统计分析理论知识，而且在每一种统计分析方法中都结合实际案例、完整的操作步骤、结果分析等，并且每一步操作都有文字和操作界面图详细阐述，产生的结果都配有界面图和文字的完整解读，简单易懂、清晰全面，十分适合初学者阅读。

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

菜鸟学 SPSS 数据分析 / 但婉欣, 蒋胡英, 郭佳桦著. —北京: 电子工业出版社, 2020.1
ISBN 978-7-121-38169-0

I. ①菜… II. ①但… ②蒋… ③郭… III. ①统计分析—软件包 IV. ①C819

中国版本图书馆 CIP 数据核字（2019）第 270588 号

责任编辑：王 静

印 刷：北京盛通商印快线网络科技有限公司

装 订：北京盛通商印快线网络科技有限公司

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：787×980 1/16 印张：16.75 字数：424 千字

版 次：2020 年 1 月第 1 版

印 次：2020 年 6 月第 2 次印刷

定 价：69.90 元

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）88254888，88258888。

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。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。



SPSS学堂 (ID: spss2333)

菜鸟也能学会数据分析!

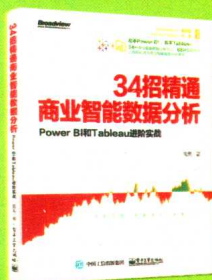
这里有详细的SPSS教程,专业的交流大社群,
可以随时帮你答疑解惑。

“菜鸟系列”“信效度问卷系列”“神经网络
系列”……关注我们就可以免费获取相关教
学资源!

SPSS学堂是你学习SPSS、熟写论文、掌握
数据分析技巧的好帮手。



好书分享，助你成为数据分析高手！



《34招精通商业智能数据分析：Power BI和Tableau进阶实战》
34种模板化分析方法，让你轻松成为商业数据分析高手！



《快乐机器学习》
一本有趣、轻松、学得会的机器学习书，让你快乐学习！



《企业经营数据分析：思路、方法、应用与工具》
用数据治理企业、改变企业！



《从数据到Excel自动化报表：Power Query和Power Pivot实战》
快速入门，打好坚实基础，轻松驾驭数据！



《人人都会数据分析：从生活实例学统计》
帮助读者建立统计理论知识框架，形成数据分析思维逻辑，学会数据分析方法！

编委会

但婉欣 蒋胡英 郭佳桦

陈 联 李玉星 李培江 X

SPSS 数据分析的利器

哈佛大学社会学教授加里·金说：“这是一场革命，庞大的数据资源使得各个领域开始了量化进程，无论是学术界、商界还是政府，所有领域都将开始这种进程。”

大数据时代来临，对数据分析技能的要求越来越高。作为一款功能强大、方便易用的数据分析软件，SPSS 广受数据分析爱好者和从业者的欢迎，所有的功能都以对话框选项的形式呈现，界面统一、规范，用户只要掌握一定的 Windows 操作技能，通晓一定的统计分析原理，就可以使用该软件。随着版本的升级，SPSS 的功能愈发强大，被广泛应用于诸多行业及学术研究中。

本书的作者均来自 SPSS 学堂（微信公众号），最初通过微信公众号发布文章得到了诸多粉丝的关注，恰逢出版社的邀约，学堂将推文重新汇编整理成书。SPSS 学堂一直致力于为数据分析爱好者带来高质量的文章，希望通过对本书的学习，各位读者都能对 SPSS 的基本操作有更加深刻的认识，能熟练掌握并运用于实际的研究或工作中。

本书内容及服务

本书适合于想系统学习 SPSS 软件却不知如何着手的读者朋友们。

本书的内容循序渐进，重在对基础性的操作进行讲解，从理论到思路，每一节都配有案例，每一部分都有细致的讲解，力求使读者（尤其是初学者）能更好地理解、运用 SPSS。SPSS 学堂提供售后服务，若有不理解之处，可以关注微信公众号并留言，我们会对大家的疑问进行解答。当然，除本书的内容外，我们的公众号还有其他精品文章，欢迎关注。另外，本书还有 SPSS 社群供各位读者讨论和交流，希望大家加入。

对于本书的阅读有如下建议：

- 若时间充裕，请从头到尾精读一遍，尤其是案例，需要实际动手去操作练习；
- 若时间不充裕，可就自己急切想要知道的内容有针对性地查阅目录，根据需求到相应的章节重点研读。

联系本书作者

本书由 SPSS 学堂的多位作者共同编写，其中，蒋胡英参与编写了 1.3、2.1、3.5、3.10、3.15、3.16、5.3、6.2、7.2、9.3、10.1、11.2、12.1、12.3 节；郭佳桦参与编写了 1.2、3.1、3.3、3.8、3.13、5.1、5.7、6.3、9.4、9.5、9.8、10.4、11.6 节；但婉欣参与编写了 1.4、3.4、3.9、3.14、5.2、5.4、6.1、7.3、9.2、9.9、11.1、11.3、12.2 节；陈联参与编写了 1.5、2.2、3.6、3.12、4.1、4.2、5.5、6.4、9.6、10.2、10.3、11.4、12.4 节；李玉星参与编写了 1.1、3.2、3.7、3.11、4.3、5.6、8.1、8.2、8.3、9.1、9.7、11.5 节；7.1 节由李玉星、蒋胡英、陈联共同编写。

需要特别说明的是，一本书的编写需要投入大量的时间与精力，各位作者在编写时都尽可能地保持正确与严谨，但也难免会存在疏漏之处，还请各位专家和广大读者不吝赐教。

联系我们：请关注公众号 SPSS 学堂（微信号 spss2333）并在后台留言。

相关案例下载：请关注公众号 SPSS 学堂（微信号 spss2333）并回复“书籍案例”即可获得。



目 录

第 1 章 SPSS 与数据分析	1
1.1 数据分析思路概述	1
1.1.1 数据分析的一般思路	1
1.1.2 使用 SPSS 进行数据分析的优势	2
1.2 明确数据分析目的	2
1.2.1 为什么要明确数据分析目的	3
1.2.2 如何明确数据分析目的	3
1.2.3 明确数据分析目的需要注意什么	3
1.3 数据采集	4
1.3.1 采集方法	4
1.3.2 采集流程	4
1.4 数据分析	5
1.4.1 数据的预处理	5
1.4.2 数据的描述统计	6
1.4.3 数据的推论统计	6
1.5 撰写 SPSS 报告	8
1.5.1 撰写报告的逻辑思路	8
1.5.2 撰写报告的注意事项	9
第 2 章 SPSS 界面功能介绍	10
2.1 SPSS 基本界面介绍	10
2.1.1 数据编辑窗口	10
2.1.2 数据视图	11
2.1.3 变量视图	12
2.1.4 结果输出窗口	12
2.2 常用参数设置	13
2.2.1 “常规”选项卡	13
2.2.2 “语言”选项卡	14

2.2.3	“查看器”选项卡	14
2.2.4	“数据”选项卡	15
2.2.5	“货币”选项卡	16
2.2.6	“输出”选项卡	17
2.2.7	“图表”选项卡	17
2.2.8	“透视表”选项卡	18
2.2.9	“文件位置”选项卡	18
2.2.10	“脚本”选项卡	19
2.2.11	“多重插补”选项卡	20
2.2.12	“语法编辑器”选项卡	20
第3章	数据的管理	22
3.1	数据的录入与保存	22
3.1.1	定义变量	22
3.1.2	录入数据	26
3.1.3	保存数据	27
3.2	其他格式数据的导入	27
3.2.1	Excel 数据的导入	28
3.2.2	文本数据的导入	30
3.3	数据的重构	34
3.3.1	数据结构	34
3.3.2	数据重构(横向结构→纵向结构)	35
3.3.3	数据重构(纵向结构→横向结构)	39
3.4	数据的合并	41
3.4.1	添加个案(纵向合并)	42
3.4.2	添加变量(横向合并)	44
3.5	标识重复个案	47
3.5.1	为何要标识重复个案	47
3.5.2	“双十一”交易明细分析:删除重复数据	48
3.6	数据的拆分	50
3.6.1	拆分文件(比较组)	51
3.6.2	拆分文件(按组来组织输出)	53
3.7	数据的筛选	53
3.7.1	数据筛选概述	53
3.7.2	收入与生活满意度分析:数据的筛选	54
3.8	数据的加权	56

3.8.1	什么是加权	56
3.8.2	“春节回家吗？”分析：数据的加权	56
3.8.3	为什么要加权	57
3.8.4	总结	59
3.9	数据的汇总	59
3.9.1	化学成绩分析：简单数据汇总	59
3.9.2	化学成绩分析：多重数据汇总	61
3.10	新变量的生成	63
3.10.1	课程成绩分析：生成新变量	64
3.10.2	“如果”按钮	65
3.11	个案中值的计数	67
3.11.1	满意度分析：统计个案中值	67
3.11.2	拓展知识	70
3.12	对变量重新赋值	71
3.12.1	信用评分分析：重新编码为不同变量	71
3.12.2	职工奖金分析：重新编码为相同的变量	73
3.13	可视化分段	76
3.13.1	什么是可视分箱	76
3.13.2	北京月度价格分析：可视分箱分组	76
3.14	个案排秩	80
3.14.1	化学成绩分析：利用个案排秩	81
3.14.2	化学成绩分析：利用自动重新编码	83
3.15	缺失值的处理	84
3.15.1	人均消费性支出分析：替换缺失值	85
3.15.2	删除缺失值	86
3.16	数据的排序	88
3.16.1	多变量的单向排序	89
3.16.2	多变量的混合排序	90
第 4 章	报表的生成	93
4.1	交叉表	93
4.1.1	交叉表概述	93
4.1.2	制作交叉表	93
4.2	定制表	97
4.2.1	定制表概述	97
4.2.2	定制表的各项功能	97

4.3 定制表的实例操作	101
4.3.1 问卷调查分析: 制作定制表	101
4.3.2 购物频次分析: 制作定制表	103

第5章 图形的生成..... 106

5.1 图表构建器介绍.....	106
5.1.1 如何使用“图表构建器”	106
5.1.2 使用图表构建器总结.....	108
5.2 直方图.....	109
5.2.1 利用“图表构建器”绘制直方图.....	109
5.2.2 利用“直方图”菜单绘制直方图.....	110
5.3 折线图.....	113
5.3.1 利用“旧对话框”绘制折线图.....	113
5.3.2 利用“图表构建器”绘制折线图.....	115
5.4 饼图.....	117
5.4.1 利用“图表构建器”绘制饼图.....	117
5.4.2 利用“饼图”菜单绘制饼图.....	119
5.5 条形图.....	120
5.5.1 利用“旧对话框”绘制条形图.....	120
5.5.2 利用“图表构建器”绘制条形图.....	123
5.6 箱图.....	125
5.6.1 利用“图表构建器”绘制箱图.....	125
5.6.2 利用“旧对话框”绘制箱图.....	127
5.6.3 利用“描述统计”制作箱图.....	128
5.7 散点图.....	130
5.7.1 简单散点图.....	130
5.7.2 矩阵散点图.....	132
5.7.3 重叠散点图.....	133
5.7.4 三维散点图.....	135

第6章 描述性分析..... 137

6.1 描述统计与变量分布形态.....	137
6.1.1 集中趋势.....	137
6.1.2 离散趋势.....	138
6.1.3 分布形态.....	140
6.2 频率分析.....	141

6.2.1	分类变量的频率分析.....	142
6.2.2	连续变量的频率分析.....	144
6.3	描述统计.....	145
6.3.1	描述统计概述.....	145
6.3.2	考生信息分析: 描述统计.....	145
6.4	探索性分析.....	148
6.4.1	身高、体重分析: 探索性分析.....	148
6.4.2	结果解读.....	151
第7章	假设检验	155
7.1	假设检验概述.....	155
7.1.1	假设检验的基本思想.....	155
7.1.2	假设检验的四大步骤.....	155
7.1.3	假设检验的意义.....	156
7.1.4	单尾检验和双尾检验.....	156
7.1.5	假设检验需要注意的问题.....	157
7.2	Z分布与t分布.....	158
7.2.1	Z分布.....	158
7.2.2	t分布.....	159
7.2.3	Z分布与t分布的关系.....	160
7.3	第一类错误和第二类错误.....	160
7.3.1	判断假设检验的四种情况.....	161
7.3.2	婴儿奶粉市场投放分析: 第一类错误和第二类错误.....	161
7.3.3	如何控制两类错误.....	162
第8章	参数检验——t检验	163
8.1	单样本t检验.....	163
8.1.1	概述.....	163
8.1.2	初中生体重分析: 单样本t检验.....	163
8.2	独立样本t检验.....	165
8.2.1	适用条件及假设检验.....	165
8.2.2	健株与病株差异性分析: 独立样本t检验.....	166
8.3	成对样本t检验.....	170
8.3.1	概述.....	171
8.3.2	检测血磷值分析: 成对样本t检验.....	171

第 9 章 非参数检验.....	174
9.1 非参数检验综述.....	174
9.1.1 基本概念.....	174
9.1.2 单样本的非参数检验.....	174
9.1.3 两个独立样本的非参数检验.....	175
9.1.4 多个独立样本的非参数检验.....	176
9.1.5 两配对样本的非参数检验.....	177
9.1.6 多配对样本的非参数检验.....	177
9.2 游程检验.....	178
9.2.1 游程检验概述.....	178
9.2.2 消费额数据分析：游程检验.....	179
9.3 单样本 K-S 检验.....	181
9.3.1 单样本 K-S 检验概述.....	181
9.3.2 耗油量数据分析：单样本 K-S 检验.....	182
9.4 二项分布检验.....	184
9.4.1 二项分布检验概述.....	184
9.4.2 检验答题正确率分析：二项分布检验.....	184
9.5 两个独立样本的非参数检验.....	185
9.5.1 概述.....	185
9.5.2 不同工艺类型产品的使用寿命数据分析：曼-惠特尼 U 检验、 K-S 检验、游程检验及极端反应检验.....	186
9.6 多个独立样本的非参数检验.....	189
9.6.1 检验方法.....	189
9.6.2 股票收益率对比分析：Kruskal-Wallis H 检验、中位数检验及 Jonckheere-Terpstra 检验.....	189
9.7 两个相关样本的非参数检验.....	191
9.7.1 概述.....	191
9.7.2 患者血磷值差异分析：Wilcoxon 符号秩检验及符号检验.....	191
9.8 多个相关样本的非参数检验.....	193
9.8.1 概述.....	193
9.8.2 产品销售额差异分析：Friedman 检验及 Kendall W 检验.....	193
9.9 非参数检验与参数检验的比较.....	195
9.9.1 非参数检验与参数检验的区别.....	195
9.9.2 非参数检验与参数检验的优缺点.....	195
9.9.3 非参数检验与参数检验的分类.....	195

第 10 章	卡方检验	196
10.1	卡方检验综述	196
10.1.1	卡方检验原理	196
10.1.2	数字偏好分析: 卡方检验	196
10.2	拟合优度检验	199
10.2.1	概述	199
10.2.2	专业学科分布比例分析: 拟合优度检验	199
10.3	独立性检验	201
10.3.1	独立性检验的适用条件	201
10.3.2	学历与消费水平分析: 独立性检验	202
10.4	一致性检验	204
10.4.1	概述	204
10.4.2	评价结果分析: 一致性检验	204
第 11 章	方差分析	208
11.1	单因素完全随机方差分析	208
11.1.1	F 检验	208
11.1.2	前提条件	209
11.1.3	媒体推广效果分析: 单因素完全随机方差分析	210
11.2	单因素重复测量方差分析	214
11.2.1	F 检验	214
11.2.2	重复测量设计的优缺点	215
11.2.3	前提条件	215
11.2.4	满意度指标分析: 单因素重复测量方差分析	215
11.3	两因素完全随机方差分析	221
11.3.1	分析思路	221
11.3.2	假设检验	222
11.3.3	阅读速度影响因素分析: 两因素完全随机方差分析	222
11.4	两因素重复测量方差分析	228
11.4.1	两因素重复测量方差分析概述	228
11.4.2	命题反应分析: 两因素重复测量方差分析	228
11.5	单因素多元方差分析	231
11.5.1	分析思路	232
11.5.2	身高和体重数据分析: 单因素多元方差分析	232
11.6	两因素多元方差分析	238

11.6.1	两因素多元方差分析概述	238
11.6.2	大学生支出情况分析：两因素多元方差分析	238
第 12 章	相关分析	243
12.1	皮尔逊相关分析	243
12.1.1	理论概述	243
12.1.2	影响超市销售的因素分析：皮尔逊相关分析	244
12.2	肯德尔等级相关分析	247
12.2.1	肯德尔相关系数	247
12.2.2	大学排名数据分析：肯德尔等级相关分析	247
12.3	斯皮尔曼等级相关分析	249
12.3.1	斯皮尔曼等级相关系数	249
12.3.2	知名度和服务质量相关性分析：斯皮尔曼等级相关分析	250
12.4	偏相关分析	251
12.4.1	理论概述	251
12.4.2	投资额与收益率的相关性分析：偏相关分析	251
参考文献	254

第 1 章

SPSS 与数据分析

1.1 数据分析思路概述

我们知道，在信息化时代背景下，大量纷繁复杂的数据和信息充斥在社会的各个领域，为了更加准确、快捷地从中筛选出有价值的部分并根据需要进行一系列数据处理，数据分析过程显得尤为重要。

1.1.1 数据分析的一般思路

1. 明确数据分析原因

在开始数据分析之前，首先应弄清楚为什么要做数据分析，明确数据分析的目的。这是数据分析的基础，也是接下来一系列分析过程的核心。

2. 明确数据分析目标和预期结果

应充分理解数据分析的目标和想要达到的最终效果，牢记正确的分析因子，通过全方面分析，找到真正的问题。

3. 明确所需数据

与研究问题相关的数据数量庞大、种类繁多，在研究问题时，应该弄清楚需要哪些数据，先整理好思路，最好列一个表格，尽可能一次性明确所需数据。

4. 数据采集与整理

采集数据有多种方法，可以直接从数据库中调取，或者找相关技术人员导出等。接下来就是对采集的数据进行处理，如清洗数据、转换数据、数据分组等，这些均可通过一些软件如 SPSS、Excel 等来完成。