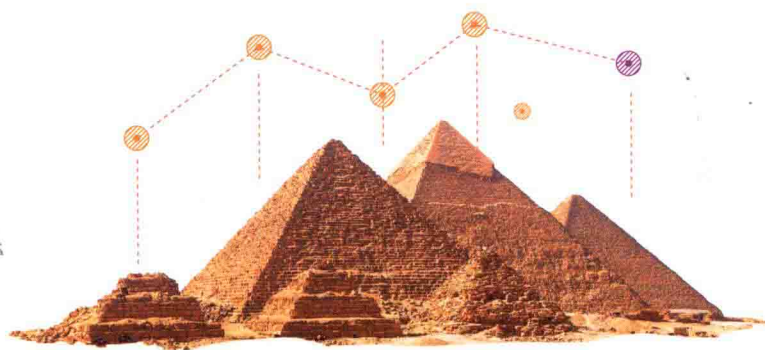


涵盖电商、咨询、教育、电信、保险、餐饮、零售、旅游、
物流、制造、房地产、农业等行业应用



SPSS statistical analysis in practice

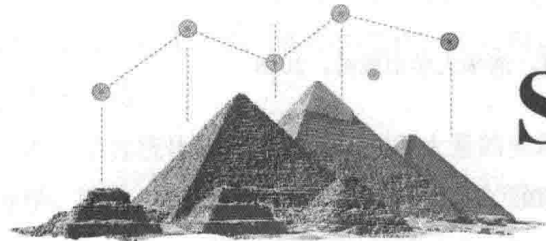
SPSS 统计分析 与行业应用实战

从数据建模到数据处理，从SPSS操作到结果分析，
掌握数据分析的整体思路并使用SPSS解决实际问题

• 王国平 编著 •

清华大学出版社





SPSS统计分析 与行业应用实战

• 王国平 编著



清华大学出版社
北京

内 容 简 介

本书以全实例为主线,阐述了 SPSS 在社会学、管理学、经济学、教育学、医学、制造业、农业、采矿业等方面的应用。本书结合编者多年工作经验,由浅入深、循序渐进,理论与实际相结合,对 SPSS 24 进行比较全面的介绍。配书资源包含每个案例采用的数据源文件及数据分析结果,方便读者在阅读本书时进行操作练习和参考。

本书既适于咨询、零售、农业、互联网等行业的数据分析用户使用,也可供高等院校相关专业学生以及从事数据分析的研究者参考,还可作为 SPSS 软件培训和自学的教材。

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 统计分析与行业应用实战 / 王国平编著. —北京:清华大学出版社, 2018

ISBN 978-7-302-50327-9

I. ①S... II. ①王... III. ①统计分析—软件包 IV. ①C819

中国版本图书馆 CIP 数据核字 (2018) 第 114922 号

责任编辑:王金柱

封面设计:王翔

责任校对:闫秀华

责任印制:刘海龙

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

印装者:三河市君旺印务有限公司

经 销:全国新华书店

开 本:190mm×260mm

印 张:26

字 数:666 千字

版 次:2018 年 8 月第 1 版

印 次:2018 年 8 月第 1 次印刷

定 价:79.00 元

产品编号:076193-01

前言

目前,无论是个人、政府还是企业都需要在数据中获取有价值的信息,并据此做出科学的评估和决策。因此,对信息的采集、处理、分析并给出评估和预测报告就变得十分重要,SPSS 正是为此而设计的一整套集数据处理、评估和预测于一身的数据统计分析工具。

SPSS 的全名为社会学统计软件包,是世界上应用最广泛的专业统计软件,以其推理严谨、结论可靠、操作方便的独特优势在科学研究中发挥着重要作用,集数据录入、数据管理、统计分析、报表制作、图形绘制为一体,是非专业统计人员运用最多的一款统计分析软件。

使用 SPSS 软件对数据进行统计分析,只需要通过选择菜单、填写对话框和选择按钮等简单操作即可完成,免去了编写程序的工作,不仅可以得到分析后的数字结果,还可以得到直观、清晰、漂亮的统计图。

本书结合统计分析的特点,应用大量的实例,分 20 章阐述 SPSS 在社会学、管理学、经济学、教育学、医学、制造业、农业、采矿业等方面的应用。本书内容翔实、语言简练、思路清晰、图文并茂、深入浅出、理论与实际应用相结合,对 SPSS 24 进行了比较全面的介绍。

本书的内容

第 1 章介绍 SPSS 在电商行业中的应用,分析消费者网络购买行为的影响因素及影响程度,其中商家信誉、商品质量、商品价格、售后服务等是影响消费者购买行为的主要因素。

第 2 章介绍 SPSS 在电信行业中的应用,使用神经网络模型分析电信客户流失的问题,并通过聚类分析对客户进行细分,针对不同的客户提出不同的营销方式。

第 3 章介绍 SPSS 在咨询行业中的应用,探究各因素对居民消费率的影响。影响居民消费率的因素主要有居民总收入、人均 GDP、人口结构状况、居民消费价格指数增长率等。

第 4 章介绍 SPSS 在教育行业中的应用,综合评定学生的体质健康水平,进行因子分析,其中体重、机能和耐力等是体质检测的重要因素,是衡量身体素质的重要指标。

第 5 章介绍 SPSS 在医疗行业中的应用,结合肺表面活性物质对胎粪吸入综合征患儿血管内血小板源性生长因子的影响,对胎粪吸入综合征患儿性别、分娩方式等进行统计检验。

第 6 章介绍 SPSS 在烟草行业中的应用,通过分析城市空气污染物及其影响因素,运用相关性分析与多元线性回归分析的方法,找出影响空气质量的关键因素。

第 7 章介绍 SPSS 在保险行业中的应用,通过对财产保险公司的基本情况的相关分析、聚类

分析等，找出构成财产保险公司基本特征的各变量之间的内在联系。

第 8 章介绍 SPSS 在电力行业中的应用，通过对变量进行相关分析，找出某市的短期电力负荷影响因素，并应用时间序列模型对某地的短期电力负荷进行预测。

第 9 章介绍 SPSS 在农业生产中的应用，首先提取影响冬小麦产量的气候因子，构建冬小麦气候产量模型，然后利用该模型估算了气候要素对冬小麦的贡献率。

第 10 章介绍 SPSS 在物流业中的应用，构建物流业竞争力评价指标体系，对某市物流业竞争力进行因子分析，并根据因子综合得分来评价该市物流业的竞争力。

第 11 章介绍 SPSS 在餐饮业中的应用，通过联合分析对某地餐饮市场消费行为展开定量研究，测量不同影响因素对消费者消费决策的相对重要性等。

第 12 章介绍 SPSS 在零售业中的应用，通过实地问卷调查收集某地普通消费者对洗发水的使用情况，通过交叉分析等方法研究产品的目标消费群体。

第 13 章介绍 SPSS 在航空业中的应用，针对我国民航客运量的影响因素，分阶段对客运量的变化趋势及成因进行研究，并运用多元回归分析的方法对民航客运量进行预测。

第 14 章介绍 SPSS 在旅游业中的应用，对影响旅游业发展的评价指标进行主成分分析，并利用聚类分析方法得出各省市旅游业竞争力水平排名。

第 15 章介绍 SPSS 在制造业中的应用，从某企业的产品质量数据入手，找出产品生产的瓶颈缺陷问题，然后设计正交试验找出最佳的工艺组合，从而提高产品的质量等级。

第 16 章介绍 SPSS 在采矿业中的应用，挖掘选矿厂浮选生产现有生产条件下潜藏的精矿品位和回收率可能达到的更优值，大大提高了选矿的效益与效率。

第 17 章介绍 SPSS 在银行业中的应用，通过建立影响银行用户贷款风险的指标体系，应用相关分析和神经网络模型研究影响贷款风险的因素。

第 18 章介绍 SPSS 在政府税收中的应用，提出税收分析建模的一般步骤，并针对入库税收历史数据进行时间序列 ARIMA 模型预测的实证分析和研究。

第 19 章介绍 SPSS 在 P2P 行业中的应用，基于 P2P 网络借贷平台数据，探讨问题平台的特征，并对 P2P 网络借贷平台出现问题的原因进行实证分析。

第 20 章介绍 SPSS 在房地产业中的应用，结合城市地铁交通的特点、房地产价格的影响因素等，采用定性和定量的综合研究方法研究城市地铁交通对周边房地产价格的影响。

本书的特色

(1) 内容全面，讲解详细

本书对于初次学习 SPSS 软件的读者来说帮助较大，书中详细列出 SPSS 在实际分析过程中的每一步操作以及注意事项，便于读者练习。

（2）案例丰富，高效学习

为了使读者快速提高数据分析的实际应用能力，本书几乎列出了 SPSS 在所有行业中的应用，每章都是结合案例进行建模理论和 SPSS 操作的介绍。

本书的下载资源

本书相关资源下载地址（注意数字与字母大小写）如下：

<https://pan.baidu.com/s/1XpeK3yFhOiCkoYmvnKyAlg>

也可扫描右边二维码获取下载地址。如果下载有问题，请发送邮件至电子邮箱 booksaga@126.com，邮件主题为“SPSS 统计分析与行业应用实战”。



本书的读者对象

本书既适于咨询、零售、农业、互联网等行业的数据分析用户使用，也可供高等院校相关专业学生以及从事数据分析的研究者参考，还可作为 SPSS 软件培训和自学的教材。

由于编者水平有限，书中难免存在不妥之处，敬请广大读者批评指正。

编者
2018 年 5 月

目 录

第 1 章 SPSS 在电商行业中的应用 1

- 1.1 研究背景及意义 1
- 1.2 数据建模思路 2
- 1.3 调查内容与方法 3
- 1.4 问卷调查数据处理 4
 - 1.4.1 问卷编码录入 4
 - 1.4.2 问卷一般性分析 5
- 1.5 SPSS 操作步骤——
多重响应分析 6
 - 1.5.1 为变量定义数据集 6
 - 1.5.2 多重响应下的频率分析 8
 - 1.5.3 多重响应下的交叉分析 9
- 1.6 SPSS 结果介绍 11
 - 1.6.1 了解电商网站的渠道 11
 - 1.6.2 上电商网站购物的原因 12
 - 1.6.3 浏览网站没有购物的原因 13
 - 1.6.4 购物时主要考虑的因素 14
- 1.7 模型效果评估 16
- 1.8 电商政策建议 16

第 2 章 SPSS 在电信行业中的应用 18

- 2.1 研究背景及意义 18
- 2.2 数据建模思路 19
- 2.3 神经网络模型 20
 - 2.3.1 神经元 21
 - 2.3.2 多层感知器 22
 - 2.3.3 径向基函数 23
- 2.4 模型数据处理 25
 - 2.4.1 导入建模数据 25

2.4.2 添加目标变量 25

2.4.3 设置变量角色 27

2.5 SPSS 操作步骤——

神经网络模型 27

2.5.1 “变量”选项 28

2.5.2 “分区”选项 29

2.5.3 “体系结构”选项 30

2.5.4 “输出”选项 31

2.5.5 “保存”选项 33

2.5.6 “导出”选项 33

2.5.7 “选项”选项 34

2.6 SPSS 结果介绍 34

2.7 模型效果评估 37

2.8 模型变量预测 38

2.8.1 导入预测数据 39

2.8.2 打开评分向导 39

2.8.3 匹配模型字段 40

2.8.4 选择评分函数 41

第 3 章 SPSS 在咨询行业中的应用 43

3.1 研究背景及意义 43

3.2 数据建模思路 44

3.3 非参数检验 44

3.3.1 单样本 45

3.3.2 两独立样本 46

3.3.3 多独立样本 47

3.3.4 两配对样本 48

3.3.5 多配对样本 49

3.4 模型数据处理 50

3.5 SPSS 操作步骤—— 非参数检验	51	5.5.2 Kaplan-Meier	86
3.5.1 “精确”选项	52	5.5.3 Cox 回归	89
3.5.2 “统计”选项	53	5.6 SPSS 结果介绍	94
3.6 SPSS 结果介绍	53	5.6.1 寿命表	94
3.7 居民消费建议	54	5.6.2 Kaplan-Meier	98
5.6.3 Cox 回归	100	5.7 随访服务建议	101
第 4 章 SPSS 在教育行业中的应用	56	第 6 章 SPSS 在烟草行业中的应用	103
4.1 研究背景及意义	56	6.1 研究背景及意义	103
4.2 数据建模思路	57	6.2 数据建模思路	104
4.3 判别分析模型	58	6.3 广义线性模型	105
4.3.1 模型定义	58	6.3.1 模型定义	105
4.3.2 模型应用	59	6.3.2 模型应用	106
4.4 模型数据处理	60	6.4 模型数据处理	107
4.4.1 指标选取	60	6.5 SPSS 操作步骤—— 广义线性模型	108
4.4.2 创建体质指数	62	6.5.1 “重复”选项	109
4.5 SPSS 操作步骤——判别分析	67	6.5.2 “模型类型”选项	110
4.5.1 “统计”选项	68	6.5.3 “响应”选项	113
4.5.2 “方法”选项	68	6.5.4 “预测变量”选项	114
4.5.3 “分类”选项	69	6.5.5 “模型”选项	115
4.5.4 “保存”选项	71	6.5.6 “估算”选项	116
4.5.5 “自助抽样”选项	71	6.5.7 “统计”选项	118
4.6 SPSS 结果介绍	72	6.5.8 “EM 平均值”选项	119
4.7 体质锻炼建议	74	6.5.9 “保存”选项	121
第 5 章 SPSS 在医疗行业中的应用	76	6.5.10 “导出”选项	122
5.1 研究背景及意义	76	6.6 SPSS 结果介绍	124
5.2 数据建模思路	77	6.7 公共场所禁烟	125
5.3 生存分析模型	78	第 7 章 SPSS 在保险行业中的应用	127
5.3.1 模型定义	78	7.1 研究背景及意义	127
5.3.2 模型应用	80	7.2 数据建模思路	128
5.4 模型数据处理	80	7.3 聚类分析模型	129
5.4.1 指标选取	80	7.3.1 聚类定义	129
5.4.2 导入数据	82	7.3.2 聚类应用	131
5.5 SPSS 操作步骤——生存分析	83		
5.5.1 寿命表	83		

7.4 模型数据处理	132	9.5.3 “保存”选项	163
7.5 SPSS 操作步骤——聚类分析	133	9.6 SPSS 结果介绍	164
7.5.1 “迭代”选项	134	9.7 农业病虫害防治	166
7.5.2 “保存”选项	135		
7.5.3 “选项”选项	135	第 10 章 SPSS 在物流业中的应用 ..	168
7.6 SPSS 结果介绍	136	10.1 研究背景及意义	168
7.7 保险业风险管理	138	10.2 数据建模思路	169
第 8 章 SPSS 在电力行业中的应用 ..	140	10.3 因子分析介绍	170
8.1 研究背景及意义	140	10.3.1 模型定义	170
8.2 数据建模思路	141	10.3.2 模型应用	172
8.3 指数平滑模型	142	10.4 模型数据处理	172
8.3.1 模型定义	142	10.4.1 评价指标体系	172
8.3.2 模型应用	143	10.4.2 导入建模数据	173
8.4 模型数据处理	144	10.5 SPSS 操作步骤——因子分析	174
8.5 SPSS 操作步骤——		10.5.1 “描述”选项	176
指数平滑模型	145	10.5.2 “提取”选项	177
8.5.1 “变量”选项	146	10.5.3 “旋转”选项	178
8.5.2 “统计”选项	148	10.5.4 “得分”选项	179
8.5.3 “图”选项	149	10.5.5 “选项”选项	180
8.5.4 “输出过滤”选项	151	10.6 SPSS 结果介绍	180
8.5.5 “保存”选项	152	10.7 物流业发展建议	182
8.5.6 “选项”选项	152	第 11 章 SPSS 在餐饮业中的应用 ..	184
8.6 SPSS 结果介绍	154	11.1 研究背景及意义	184
8.7 电力负荷预测	156	11.2 数据建模思路	185
第 9 章 SPSS 在农业生产中的应用 ..	157	11.3 联合分析介绍	186
9.1 研究背景及意义	157	11.3.1 模型定义	186
9.2 数据建模思路	158	11.3.2 模型应用	187
9.3 曲线回归模型	158	11.4 模型数据处理	188
9.3.1 模型定义	158	11.4.1 建立评价指标	188
9.3.2 模型应用	159	11.4.2 数据处理方式	189
9.4 模型数据处理	160	11.5 SPSS 操作步骤——联合分析	189
9.5 SPSS 操作步骤——曲线回归	161	11.5.1 生成正交文件	192
9.5.1 设置分析变量	162	11.5.2 联合分析程序	193
9.5.2 选择曲线模型	162	11.6 SPSS 结果介绍	194
		11.7 模型结论与建议	195

第 12 章 SPSS 在零售业中的应用 ··· 197

- 12.1 研究背景及意义 ··· 197
- 12.2 数据建模思路 ··· 198
- 12.3 客户细分理论 ··· 198
 - 12.3.1 顾客细分的必要性 ··· 198
 - 12.3.2 顾客细分理论原理 ··· 199
- 12.4 模型数据处理 ··· 200
- 12.5 SPSS 操作步骤——交叉分析 ··· 201
 - 12.5.1 “精确”选项 ··· 203
 - 12.5.2 “统计”选项 ··· 204
 - 12.5.3 “单元格”选项 ··· 206
 - 12.5.4 “格式”选项 ··· 208
 - 12.5.5 “自助抽样”选项 ··· 208
- 12.6 SPSS 结果介绍 ··· 209
 - 12.6.1 功能和性别交叉分析 ··· 209
 - 12.6.2 生活费 and 价格交叉分析 ··· 211
- 12.7 零售业发展建议 ··· 213

第 13 章 SPSS 在航空业中的应用 ··· 214

- 13.1 研究背景及意义 ··· 214
- 13.2 数据建模思路 ··· 215
- 13.3 RFM 模型 ··· 216
 - 13.3.1 模型定义 ··· 216
 - 13.3.2 模型应用 ··· 217
- 13.4 模型数据处理 ··· 218
- 13.5 SPSS 操作步骤——RFM 模型 ··· 218
 - 13.5.1 “变量”选项 ··· 220
 - 13.5.2 “分箱化”选项 ··· 220
 - 13.5.3 “保存”选项 ··· 222
 - 13.5.4 “输出”选项 ··· 222
- 13.6 SPSS 结果介绍 ··· 223
- 13.7 模型效果评估 ··· 226

第 14 章 SPSS 在旅游业中的应用 ··· 228

- 14.1 研究背景及意义 ··· 228
- 14.2 数据建模思路 ··· 229

14.3 主成分分析 ··· 230

- 14.3.1 模型定义 ··· 230
- 14.3.2 模型应用 ··· 231
- 14.4 模型数据处理 ··· 233
 - 14.4.1 指标选取 ··· 233
 - 14.4.2 导入数据 ··· 234

14.5 SPSS 操作步骤——

- 主成分分析 ··· 235
 - 14.5.1 “描述”选项 ··· 236
 - 14.5.2 “提取”选项 ··· 237
 - 14.5.3 “旋转”选项 ··· 239
 - 14.5.4 “得分”选项 ··· 239
 - 14.5.5 “选项”选项 ··· 240

14.6 SPSS 结果介绍 ··· 241**14.7 模型效果评估 ··· 244****第 15 章 SPSS 在制造业中的应用 ··· 246**

- 15.1 研究背景及意义 ··· 246
- 15.2 数据建模思路 ··· 247
- 15.3 质量控制理论 ··· 248
 - 15.3.1 质量控制 ··· 248
 - 15.3.2 控制图 ··· 249
- 15.4 模型数据处理 ··· 250
 - 15.4.1 二极管不良品 ··· 251
 - 15.4.2 反向击穿电压 ··· 251
- 15.5 SPSS 操作步骤——质量控制 ··· 252
 - 15.5.1 二极管不良品分析 ··· 253
 - 15.5.2 反向击穿电压分析 ··· 255
- 15.6 SPSS 结果介绍 ··· 257
 - 15.6.1 二极管不良品控制图 ··· 257
 - 15.6.2 反向击穿电压控制图 ··· 258
- 15.7 模型效果评估 ··· 259

第 16 章 SPSS 在采矿业中的应用 ··· 260

- 16.1 研究背景及意义 ··· 260
- 16.2 数据建模思路 ··· 261

16.3 线性回归分析	262	18.4 模型数据处理	299
16.3.1 模型定义	262	18.5 SPSS 操作步骤——ARIMA	
16.3.2 模型应用	264	模型	299
16.4 模型数据处理	265	18.5.1 模型定义日期	300
16.5 SPSS 操作步骤——线性回归	266	18.5.2 生成序列图	301
16.5.1 “统计”选项卡	268	18.5.3 生成自相关图	302
16.5.2 “图”选项	269	18.5.4 创建传统模型	304
16.5.3 “保存”选项	270	18.6 SPSS 结果介绍	314
16.5.4 “选项”选项	272	18.7 模型效果评估	318
16.5.5 “自助抽样”选项	273	18.8 模型变量预测	320
16.6 SPSS 结果介绍	274		
16.7 模型变量预测	276	第 19 章 SPSS 在 P2P 行业	
第 17 章 SPSS 在银行业中的应用	278	中的应用	321
17.1 研究背景及意义	278	19.1 研究背景及意义	321
17.2 数据建模思路	279	19.2 数据建模思路	322
17.3 Logistic 回归	280	19.3 方差分析模型	323
17.3.1 模型定义	280	19.3.1 模型定义	323
17.3.2 模型应用	282	19.3.2 模型应用	324
17.4 模型数据处理	283	19.4 模型数据处理	325
17.5 SPSS 操作步骤——		19.4.1 建立评价指标体系	325
Logistic 回归	284	19.4.2 平台经营经验分析	326
17.5.1 “分类”选项	286	19.4.3 资本充足程度分析	327
17.5.2 “保存”选项	287	19.4.4 行业竞争程度分析	327
17.5.3 “选项”选项	288	19.4.5 平台生存月数分析	328
17.6 SPSS 结果介绍	289	19.5 SPSS 操作步骤——	
17.7 模型效果评估	292	方差分析模型	328
17.8 模型变量预测	293	19.5.1 “对比”选项	329
第 18 章 SPSS 在政府税收		19.5.2 “事后比较”选项	330
中的应用	296	19.5.3 “选项”选项	332
18.1 研究背景及意义	296	19.5.4 “自助抽样”选项	333
18.2 数据建模思路	297	19.6 SPSS 结果介绍	333
18.3 ARIMA 模型	298	19.6.1 平台经营经验方差分析	333
18.3.1 模型定义	298	19.6.2 资本充足程度方差分析	334
18.3.2 模型应用	298	19.6.3 行业竞争程度方差分析	335
		19.7 风险控制建议	336

第 20 章 SPSS 在房地产业

中的应用 338

20.1 研究背景及意义 338

20.2 数据建模思路 339

20.3 决策树模型 340

20.3.1 模型定义 340

20.3.2 模型应用 341

20.4 模型数据处理 342

20.5 SPSS 操作步骤——
决策树模型 343

20.5.1 “输出”选项 344

20.5.2 “验证”选项 349

20.5.3 “条件”选项 350

20.5.4 “保存”选项 353

20.5.5 “选项”选项 354

20.6 SPSS 结果介绍 355

20.6.1 模型摘要 355

20.6.2 分类树 355

20.6.3 节点的增益 356

20.7 模型变量预测 357

20.7.1 导入预测数据 357

20.7.2 打开评分向导 357

20.7.3 匹配模型字段 358

20.7.4 选择评分函数 358

附录 A SPSS 软件新增功能简介 360

A.1 SPSS 24.0 新增 360

A.2 SPSS 23.0 新增 361

A.3 SPSS 22.0 新增 361

附录 B SPSS 24 的安装与卸载 363

B.1 SPSS 软件的安装 363

B.2 SPSS 软件的卸载 372

B.3 SPSS 中英文切换 373

附录 C SPSS 如何读取数据 377

C.1 从 Excel 中读取数据 377

C.2 从数据库中读取数据 380

C.3 从文本文件读取数据 384

附录 D SPSS 调用 Python 程序 390

D.1 安装 Python Essentials 390

D.2 如何运行 Python 代码 392

附录 E SPSS 调用 R 语言程序 395

E.1 安装 R Essentials 395

E.2 调用 R 语言的方法 400

参考文献 403

第 1 章

SPSS 在电商行业中的应用

1.1 研究背景及意义

近年来，随着互联网的发展，网络购物的方式越来越流行，尤其是对于广大年轻人，预计未来网购群体在整体网民中的比例会更高，逐渐成为人们的生活消费习惯。

根据艾瑞咨询数据显示，2016 年中国网络购物市场交易规模为 5.0 万亿元，较 2015 年增长 30.7%，预计 2017 年市场交易规模为 6.2 万亿元、2018 年将达到 7.5 万亿元，如图 1-1 所示。



图 1-1 2011—2018 年中国网络购物市场规模

根据艾瑞咨询数据，2011—2016 年我国网络购物人均消费额持续增长，2016 年网络购物人均消费额为 10830.2 元，同比 2015 年增长 17.6%，预计 2017 年为 12539.9 元、2018 年将达到 14240.7 元，如图 1-2 所示。



图 1-2 2011—2018 年网络购物人均消费额

以前由于受到利益驱使、信用障碍、意识淡薄和缺乏管理等方面因素的影响，导致电子商务市场缺失诚信。《网络购买商品七日无理由退货暂行办法》自 2017 年 3 月 15 日正式施行后，根据办法网络购买商品可以在七日内无理由退货，只有消费者定做的商品、鲜活易腐的商品、在线下载或者消费者拆封的音像制品以及计算机软件等数字化商品、交付的报纸期刊等 4 类商品不适用。

虽然各交易平台推出卖家信用度来显示商家的一些交易特性，在一定程度上消除卖方信用不确定、商家商品质量等问题，但是卖家信用度的计算简单，不能全面反映卖家的真实信息，而且消费者的交易行为不仅仅由卖家信用度决定，因此了解消费者网络交易行为的影响因素、影响方式，以及影响程度对于规范电子商务活动和促进电子商务的健康发展具有重要意义。

1.2 数据建模思路

本实例通过设计问卷，调查消费者网络购买行为的主要影响因素及其影响程度，问卷涉及被调查者的性别与学历、了解电子商务网站的渠道、上电子商务网站购物的原因、浏览网站但是没有购物的原因和购物时主要考虑的因素等，其中除了被调查者的性别是单选外，其他均为多选。

对于多选问题最有效的分析方法就是利用 SPSS 进行多重响应分析，因此本文首先采用多重响应的方法对问卷数据进行分析，然后对 SPSS 的分析结果进行详细介绍，同时还对模型的效果进行评估，具体数据建模过程的流程图如图 1-3 所示。

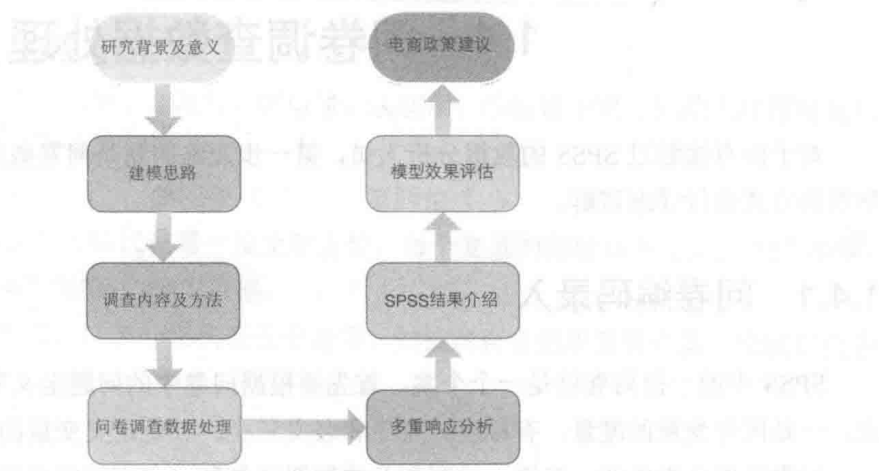


图 1-3 建模过程流程图

1.3 调查内容与方法

为了分析电子商务诚信情况以及消费者购买行为影响因素,我们采用问卷调查的形式收集相关数据。经过两周的实地调查,共计发放 230 份调查问卷,其中回收 200 份,回收率为 86.96%。根据英国著名社会学家莫泽(C.A.Moser)标准,如果问卷的回收率达不到 50%,那么问卷调查就应该终止,问卷的回收率达不到 70%,问卷调查的结果就不足以可信,由于本次问卷调查的有效率大于该指标,因此视为有效调查。

《电子商务诚信情况以及消费者购买行为影响因素的调查问卷》

- 第 1 题 您的性别? [单选题]
A.男 B.女
- 第 2 题 您是通过什么渠道了解电子商务网站的? [多选题]
A.网上广告 B.网上论坛 C.电视广告 D.户外广告 E.亲朋推荐
- 第 3 题 吸引您去上电子商务网站购物的原因? [多选题]
A.商品齐全 B.价格便宜 C.比较方便 D.买到二手货 E.交朋友
- 第 4 题 您浏览电子商务网站没有购物的原因? [多选题]
A.随便看看 B.没有尝试过 C.网络安全隐患 D.卖家信誉 E.其他
- 第 5 题 您在电子商务网站购物时主要考虑的因素有哪些? [多选题]
A.网站安全性 B.网站知名度 C.卖家信用度 D.产品质量与服务 E.配送是否及时

1.4 问卷调查数据处理

对于没有接触过 SPSS 的数据分析人员，第一步面临的的就是问卷数据的编码问题。下面将通过举例的方式进行详细讲解。

1.4.1 问卷编码录入

SPSS 中的一份问卷就是一个个案，首先要根据问卷中的问题定义变量。定义变量需要注意两点：一是区分变量的度量，有标度、有序和名义三类；二是定义变量的数据类型。

问卷可以分为单选、多选、排序和开放题四种类型，它们的定义和处理方法不同。

(1) 单选题：答案只能有一个选项。

例一：您的性别？

A.男 B.女

编码：只定义一个变量，值 1、0 分别代表男、女两个选项。

录入：录入选项对应值，若选 1 则录入男，若选 0 则录入女。

(2) 多选题：答案可以有多个选项，其中又有项数不定多选和项数定多选。

① 方法一（二分法）。

例二：您是通过什么渠道了解 C2C 网站的？

A.网上广告 B.网上论坛 C.电视广告 D.户外广告 E.亲朋推荐

编码：把每一个选项定义为一个变量，每一个变量的值用 0 和 1 表示（“0”未选，“1”选）。

录入：被调查者选了的选项录入 1，没选录入 0。例如，选择被调查者选 A、C，那么五个变量就分别录入为 1、0、1、0、0。

② 方法二。

例三：您认为开展保持党员先进性教育活动的最重要的目标是哪三项？

1 () 2 () 3 ()

A. 提高党员素质 B. 加强基层组织 C. 坚持发扬民主 D. 激发创业热情 E. 服务人民群众
F. 促进各项工作

编码：定义三个变量，分别代表题目中的 1、2、3 三个括号。三个变量的值均以对应的选项定义，即“1” A、“2” B、“3” C，“4” D，“5” E，“6” F。

录入：录入的数值 1、2、3、4、5、6 分别代表选项 A、B、C、D、E、F，相应录入到每个括号对应的变量下。若被调查者三个括号分别选 A、C、F，则在三个变量下分别录入 1、3、6。

注：能用方法二编码的多选题也能用方法一编码，但是项数不定的多选只能用二分法，即方法一是多选题一般的处理方法。

(3) 排序题：对选项重要性进行排序。

例四：您购买商品时，在①品牌、②流行、③质量、④实用、⑤价格中对它们的关注程度先后顺序是（请填写代号重新排列）什么？

第一位（ ） 第二位（ ） 第三位（ ） 第四位（ ） 第五位（ ）

编码：定义五个变量，分别可以代表第一位至第五位，每个变量的都做如下定义：“1” 品牌，“2” 流行，“3” 质量，“4” 实用，“5” 价格。

录入：录入的数字 1、2、3、4、5 分别代表五个选项，如被调查者把质量排在第一位就在代表第一位的变量下输入“3”。

(4) 选择排序题。

例五：把例三中的问题改为“您认为开展保持党员先进性教育活动的最重要的目标是哪三项，并按重要性从高到低排序”，选项不变。

编码：以“ABCDEF”六个选项分别对应定义六个变量，每个变量都做同样的定义，即“0” 未选、“1” 排第一、“2” 排第二、“3” 排第三。

录入：以变量的值录入。比如三个括号里分别选的是 E、C、F，则该题的六个变量的值应该分别录入 0（代表 A 选项未选）、0（代表 B 选项未选）、2（代表 C 选项排在第二位）、0（代表 D 选项未选）、1（代表 E 选项排在第一位）、3（代表 F 选项排在第三位）。

(5) 开放性数值题和量表题：这类题目要求被调查者自己填入数值，或者打分。

例六：您的年龄是_____。

编码：一个变量，不定义变量的值。

录入：录入被调查者实际填入的数值。

(6) 开放性文字题。

如果可能的话可以按照含义相似的答案进行编码，转换成为封闭式选项进行分析。如果答案内容较为丰富、不容易归类，应对这类问题直接做定性分析。

1.4.2 问卷一般性分析

下面具体介绍 SPSS 中间卷的一般处理方法，操作以版本 SPSS 24.0 为例，以下提到的菜单项均在“分析”主菜单下。

(1) 频数分析：可以做单变量的频数分布表，显示数据文件中由用户指定的变量的特定值发生的频数，获得某些描述统计量和描述数值范围的统计量。

适用范围：单选题（例一），排序题（例四），多选题的方法二（例三）。

频数分析也是问卷分析中最常用的方法。

实现：分析——描述统计——频率

(2) 描述分析：可以计算单变量的描述统计量。这些统计量有平均值、算术和、标准差、最大值、最小值、方差、范围和平均数标准误等。