

高等学校计算机基础教育教材精选

SAS编程技术教程

(第2版)

朱世武 编著

清华大学出版社

高等学校计算机基础教育教材精选

SAS编程技术教程

(第2版)

朱世武 编著

清华大学出版社
北京

内 容 简 介

本书是 SAS 编程技术专著,是作者多年以来利用 SAS 软件进行教学与科研的工作结晶。本书内容全面、系统性强、层次明确、语言简练,可作为 SAS 编程技术词典使用。

本书可作为 SAS 编程技术、实证金融学、金融数据库、金融统计学、金融计算与建模、金融实验、数据分析工具及应用等课程的教材及参考书,亦可供研究生及各行业专业人员阅读。

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

SAS 编程技术教程/朱世武编著.--2 版.--北京:清华大学出版社,2013

高等学校计算机基础教育教材精选

ISBN 978-7-302-33309-8

I. ①S… II. ①朱… III. ①统计分析—应用软件—高等学校—教材 IV. ①C819

中国版本图书馆 CIP 数据核字(2013)第 173636 号

责任编辑:汪汉友 赵晓宁

封面设计:傅瑞学

责任校对:李建庄

责任印制:何 芊

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

课件下载: <http://www.tup.com.cn>,010-62795954

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×260mm 印 张:31.5

字 数:786 千字

版 次:2007 年 10 月第 1 版 2013 年 10 月第 2 版

印 次:2013 年 10 月第 1 次印刷

印 数:1~3000

定 价:56.00 元

产品编号:049846-01

出版说明

高等学校计算机基础教育教材精选

高等学校计算机基础教育教材精选在教育部关于高等学校计算机基础教育三层次方案的指导下,我国高等学校的计算机基础教育事业蓬勃发展。经过多年的教学改革与实践,全国很多学校在计算机基础教育这一领域中积累了大量宝贵的经验,取得了许多可喜的成果。

随着科教兴国战略的实施以及社会信息化进程的加快,目前我国的高等教育事业正面临着新的发展机遇,但同时也必须面对新的挑战。这些都对高等学校的计算机基础教育提出了更高的要求。为了适应教学改革的需要,进一步推动我国高等学校计算机基础教育事业的发展,我们在全中国各高等学校精心挖掘和遴选了一批经过教学实践检验的优秀教学成果,编辑出版了这套教材。教材的选题范围涵盖了计算机基础教育的三个层次,包括面向各高校开设的计算机必修课、选修课以及与各类专业相结合的计算机课程。

为了保证出版质量,同时更好地适应教学需求,本套教材将采取开放的体系和滚动出版的方式(即成熟一本、出版一本,并保持不断更新),坚持宁缺毋滥的原则,力求反映我国高等学校计算机基础教育的最新成果,使本套丛书无论在技术质量上还是文字质量上均成为真正的“精选”。

清华大学出版社一直致力于计算机教育用书的出版工作,在计算机基础教育领域出版了许多优秀的教材。本套教材的出版将进一步丰富和扩大我社在这一领域的选题范围、层次和深度,以适应高校计算机基础教育课程层次化、多样化的趋势,从而更好地满足各学校由于条件、师资和生源水平、专业领域等的差异而产生的不同需求。我们热切期望全国广大教师能够积极参与到本套丛书的编写工作中来,把自己的教学成果与全国的同行们分享;同时也欢迎广大读者对本套教材提出宝贵意见,以便我们改进工作,为读者提供更好的服务。

我们的电子邮件地址是 jiaoh@tup.tsinghua.edu.cn,联系人:焦虹。

第 2 版前言

SAS 编程技术教程(第 2 版)

本书是作者在《SAS 编程技术与金融数据处理》(清华大学出版社,2003 年版)和《SAS 编程技术教程》(清华大学出版社,2007 年版)的基础上,历经清华大学、北京大学、中国人民大学、外经贸大学等多所高校本科生与研究生的教学实践后,重新设计和完善的一本 SAS 编程专著,是作者多年来使用 SAS 软件进行教学与科研的工作结晶。

本书在体系安排上征询过多所院校专业老师的建议,循序渐近地展现了使用 SAS 软件编程解决实际问题的全过程。本书的内容架构对于学习同类软件的编程技术具有极大借鉴意义。全书共分 16 章,重点章节为第 1~第 5、第 10、第 13~第 16 章。第 1~第 5 章为核心内容,其余各章是对 SAS 编程技术的完善与提高。作为教材时,可将第 6~第 9、第 11、第 12 章设为自学内容。

本书特色:

- 倡导“为用而学,而不是为学而学”的理念。例题多且有实际意义,使读者更容易掌握 SAS 软件编程技术;避免了只是空洞地解释语句、创建没有实际意义例程的弊端。
- 练习题有应用价值。练习题选自作者多年使用 SAS 软件进行数据处理与金融建模的经典案例。完成这些练习题所掌握的算法及编程技巧会极大提高读者的编程能力。
- 语言简练、准确。一般情况下,一句话完成对 SAS 语句、过程相关选项的解释。

本书配套课件目录: ResDat2;SAS 编程技术教程(二)_ppt;SAS 编程技术教程(二)_Programs; SAS 编程技术教程(二)_Exercises。

本书学习方法: 阅读教材;利用样本数据 ResDat2,反复运行《SAS 编程技术教程(二)_Programs》中的各章例程;完成《SAS 编程技术教程(二)_Exercises》中的习题。

本书在修订过程中得到了数据提供商锐思数据公司研究人员的大力帮助,在此表示衷心的感谢,同时也感谢其他高校学生及金融机构专业人士提出的宝贵建议。

限于作者的水平,书中一定会存在不足之处,敬请读者提出宝贵建议并对有错误的地方进行指正,以便以后再版时加以改进。

作者电子邮件地址: zhushw@sem.tsinghua.edu.cn。 •

朱世武

2013 年 8 月于清华园

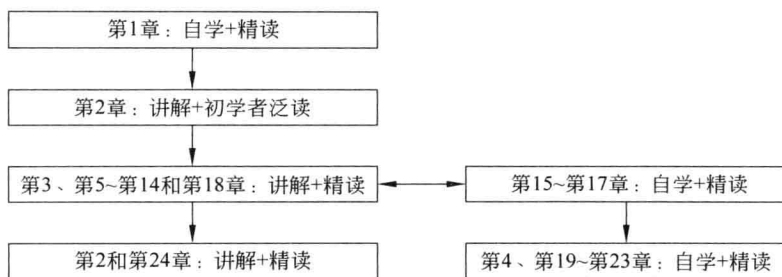
第 1 版前言

SAS 编程技术教程(第 2 版)

本书是作者在《SAS 编程技术与金融数据处理》(清华大学出版社,2003 年版)一书的基础上,历经清华大学本科生与研究生的教学实践、修改和完善的一本 SAS 编程专著,是作者多年来利用 SAS 软件进行教学与科研的工作结晶。

全书共分 24 章。重点章节为第 1、第 3、第 5~第 9、第 14 和第 18 章。最低要求是学好第 3、第 5~第 9 章内容,各章后面的习题可以帮助读者更好地掌握 SAS 编程的基本概念。

各章节内容的学习顺序与方法:



本书特色如下。

(1) 编程技术与实际问题相结合。配备了大量有实际意义的例子,加上多年来积累的练习题、水平测试题和综合练习题,帮助读者轻松地掌握 SAS 编程技术。克服许多编程专著只是空洞地解释语句、创建没有实际意义例程的弊端。

(2) 突出语句的重要应用功能,充分发挥 SAS 系统的优势。使用读者充分体会到 SAS 系统的强大功能,实现复杂的数据处理。

(3) 语言简练、准确。一般情况下,一句话完成对 SAS 语句、过程相关选项的解释。

(4) 内容全面、信息量大。可用作 SAS 编程技术词典使用。

(5) 专业金融数据网站的在线技术支持。本书得到了专业金融数据网站(www.resset.cn)的在线技术支持,提供配套数据库、程序下载与疑难问题解答等服务,方便读者学习。

本书适合多层次多专业的人士阅读,如数学、统计学、经济和金融等专业的本科生、研究生及相应部门的专业人员。

本书的写作过程中,得到了许多清华学生的帮助,他们是麦凌、陈健恒、许凯、陈明亮、何剑波、邢丽、刘海燕、李璇、王凯、邢艳丹、李文喆、张小红、赵宏旭、徐宇等。在此,特表示

衷心的感谢。同时也感谢其他高校学生及金融机构专业人士提出的宝贵建议。

限于作者的水平,书中一定会存在不少不足之处,敬请读者提出宝贵建议并对有错误的地方进行指正,以便以后再版时加以改进。

朱世武^①

2007年3月于清华园

^① 朱世武 数量经济专业博士、金融工程专业博士后。清华大学经济管理学院金融系副教授,金融量化分析与计算专业委员会副主任兼秘书长,中国现场统计学会理事,中国教育统计学会常务理事。研究领域为固定收益、风险管理、金融计算、金融建模与数据挖掘。主持完成了多项国家自然科学基金、社会科学基金、中国证监会、中国外汇交易中心、中国人寿资产管理有限公司、汇丰银行项目。国内外学术期刊发表论文50余篇,出版专著6部,译著2部。朱世武博士是我国著名的金融数据库专家,最早在国内开设金融数据库相关课程,并与国内外多家金融数据库公司进行过合作。讲授的课程有金融数据库、SAS编程技术、金融统计学、实证金融学,金融数据分析方法与应用,数据分析工具等,深受广大学生的欢迎。

目录

第 1 章 SAS 软件入门	1
1.1 SAS 软件功能模块	1
1.2 SAS 工作界面	3
1.2.1 工具栏	3
1.2.2 命令行窗口	4
1.2.3 功能窗口	4
1.3 SAS 窗口操作	8
1.3.1 资源管理器窗口操作	8
1.3.2 程序编辑窗口操作	9
1.3.3 功能键窗口操作	10
1.3.4 数据集导入和导出	10
习题 1	15
第 2 章 SAS 编程基础	16
2.1 SAS 语言组件	16
2.1.1 SAS 文件	16
2.1.2 SAS 外部文件	16
2.1.3 DBMS 文件	17
2.1.4 SAS 语言元素	17
2.1.5 SAS 宏工具	17
2.2 SAS 文件系统	17
2.2.1 逻辑库	17
2.2.2 数据集	19
2.2.3 数据文件	22
2.2.4 数据视图	23
2.3 SAS 词段和 SAS 名称	25
2.3.1 SAS 词段	25
2.3.2 SAS 名称命名规则	26
2.4 SAS 程序	26

2.4.1	书写规则	27
2.4.2	数据步	27
2.4.3	过程步	28
2.5	SAS 语言元素	28
2.5.1	语句	28
2.5.2	SAS 系统选项	29
2.5.3	数据集选项	29
2.6	SAS 变量	30
2.6.1	变量类型	30
2.6.2	变量属性	31
2.6.3	变量列表及其缩写规则	31
2.6.4	创建变量	32
2.6.5	输入输出格式	33
2.6.6	自动变量	34
2.7	表达式	35
2.7.1	引用 SAS 常数	36
2.7.2	SAS 算符	37
	习题 2	40

第 3 章	数据导入	41
3.1	菜单导入数据	41
3.2	数据步导入原始数据	41
3.2.1	原始数据呈现形式	41
3.2.2	使用 INPUT 语句读入原始数据	42
3.2.3	列方式输入	43
3.2.4	列表方式输入	44
3.2.5	格式化方式输入	47
3.2.6	命名方式输入	49
3.3	指针控制	50
3.3.1	列行指针控制	51
3.3.2	使用行固定说明符	52
3.3.3	读完数据后的指针位置	53
3.3.4	多个数据行构成一个观测	55
3.3.5	指针超过行结尾	56
3.3.6	指针移到第一列之前	57
3.4	过程步导入数据	57
3.4.1	句法与选项说明	57
3.4.2	导入程序模板	58

3.5	使用锐思 RESSET 数据库	59
3.5.1	下载 SAS 数据集	59
3.5.2	使用 RESSET 数据库系统自动产生 SAS 程序	62
习题 3	63
第 4 章	数据步数据集操作	64
4.1	创建数据集语句 DATA	64
4.1.1	语句格式	65
4.1.2	选项说明	65
4.1.3	特殊数据集名	66
4.1.4	一个 DATA 语句下多个数据集名	67
4.2	数据行引导语句 CARDS 与 CARDS4	67
4.2.1	CARDS 语句	67
4.2.2	CARDS4 语句	68
4.3	输出语句 PUT	68
4.3.1	语句格式	68
4.3.2	选项说明	69
4.3.3	应用举例	69
4.3.4	指针控制	70
4.3.5	列方式输出	72
4.3.6	列表方式输出	73
4.3.7	格式化输出	74
4.4	读入已存在 SAS 数据集语句 SET	74
4.4.1	语句格式	74
4.4.2	选项说明	75
4.4.3	应用举例	75
4.5	横向合并数据集语句 MERGE	80
4.5.1	语句格式	80
4.5.2	选项说明	80
4.5.3	应用举例	80
4.6	BY 语句	81
4.6.1	语句格式	81
4.6.2	选项说明	82
4.6.3	BY 语句概念	82
4.6.4	FIRST. 变量和 LAST. 变量	82
4.7	删除变量语句 DROP 与保留变量语句 KEEP	82
4.7.1	DROP 语句	82
4.7.2	KEEP 语句	83

4.7.3	DROP 和 KEEP 语句使用规则	84
4.7.4	数据集选项 DROP=和 KEEP=使用规则	84
4.8	更改变量名语句 RENAME 与保留数值语句 RETAIN	85
4.8.1	更改变量名语句 RENAME	85
4.8.2	保留数值语句 RETAIN	86
4.9	输出外部文件语句 FILE	89
4.9.1	语句格式	89
4.9.2	应用举例	92
4.10	定义外部数据文件语句 INFILE	92
4.10.1	语句格式	92
4.10.2	导入外部数据文件的标准程序	95
4.11	更新数据语句 UPDATE	96
4.11.1	语句格式	96
4.11.2	选项说明	97
4.11.3	MERGE 语句和 UPDATE 语句的比较	97
4.11.4	应用举例	97
4.12	修改数据语句 MODIFY	99
4.12.1	语句格式	99
4.12.2	选项说明	100
4.12.3	数据集访问方式	100
4.12.4	修改观测	101
4.12.5	与 UPDATE 等语句的比较	102
4.12.6	应用举例	102
4.13	数据集加密	105
4.13.1	数据集的三种密码	105
4.13.2	应用举例	105
4.14	数据导出	107
4.14.1	菜单导出数据	107
4.14.2	过程步导出数据	107
4.14.3	PUT 和 FILE 语句组合输出外部 TXT 文本	108
4.15	SAS 处理流程	110
4.15.1	SAS 处理流程	110
4.15.2	DATA 步处理流程	110
4.15.3	DATA 步的运行顺序	115
习题 4	117
第 5 章	DATA 步数据处理	124
5.1	基本语句	124

5.1.1	赋值语句	124
5.1.2	累加语句	126
5.2	观测的选择与输出	128
5.2.1	IF 语句	128
5.2.2	WHERE 语句	129
5.2.3	WHERE 和 IF 语句的比较	131
5.2.4	DELETE 语句	132
5.2.5	OUTPUT 语句	132
5.3	变量属性控制	135
5.3.1	INFORMAT 语句与 FORMAT 语句	135
5.3.2	LENGTH 语句与 LABEL 语句	137
5.3.3	变量类型转换	141
5.4	DO 语句及循环控制	144
5.4.1	简单 DO 语句	144
5.4.2	循环 DO 语句	145
5.4.3	DO OVER 语句	147
5.4.4	DO WHILE 语句	148
5.4.5	DO UNTIL 语句	148
5.5	选择控制语句 SELECT	149
5.5.1	语句格式	149
5.5.2	应用举例	149
5.6	数组	150
5.6.1	显式下标数组语句	151
5.6.2	引用显式下标数组元素	153
5.6.3	隐含下标数组语句	155
5.6.4	引用隐含下标数组元素	156
5.7	GO TO 语句与语句标号	158
5.7.1	GO TO 语句	158
5.7.2	语句标号	160
5.8	LINK 语句	160
5.8.1	语句格式	160
5.8.2	LINK 语句与 GOTO 语句的差别	162
5.9	STOP 语句与 ABORT 语句	162
5.9.1	STOP 语句	162
5.9.2	ABORT 语句	162
5.10	REMOVE 语句与 REPLACE 语句	166
5.10.1	REMOVE 语句	166
5.10.2	REPLACE 语句	167

5.11	MISSING 语句	168
5.11.1	语句格式	169
5.11.2	应用举例	169
5.12	其他语句	169
5.12.1	LIST 语句	169
5.12.2	PUT 语句与 LIST 语句比较	170
5.12.3	CALL 语句	170
5.12.4	CALL 语句调用子程序	170
5.12.5	NULL 语句	171
5.12.6	ERROR 语句	172
5.12.7	RETURN 语句	173
5.12.8	CONTINUE 语句与 LEAVE 语句	174
	习题 5	176
第 6 章	常用函数	181
6.1	SAS 函数定义	181
6.1.1	函数定义	181
6.1.2	函数用法	181
6.2	SAS 函数自变量与结果	182
6.2.1	函数自变量	182
6.2.2	函数结果	183
6.2.3	显示函数值的简单方法	184
6.3	SAS 函数分类	184
6.4	日期时间函数	185
6.4.1	日期时间函数	185
6.4.2	应用举例	186
6.5	概率分布函数	189
6.5.1	标准正态分布	189
6.5.2	卡方分布	190
6.5.3	伽马分布	190
6.5.4	贝塔分布	190
6.5.5	F 分布	191
6.5.6	t 分布	191
6.5.7	二项分布	191
6.5.8	泊松分布	191
6.5.9	负二项分布	192
6.5.10	超几何分布	192
6.6	分位数函数	193

6.6.1	卡方分布分位数	193
6.6.2	贝塔分布分位数	193
6.6.3	F 分布分位数	193
6.6.4	t 分布分位数	193
6.6.5	正态分布分位数	194
6.6.6	伽马分布分位数	194
6.7	样本统计函数	194
6.8	随机数函数	196
6.8.1	正态分布	196
6.8.2	均匀分布	197
6.8.3	二项分布	198
6.8.4	伽马分布	198
6.8.5	泊松分布	199
6.8.6	贝塔分布	199
6.8.7	指数分布	199
6.8.8	几何分布	200
6.8.9	极值分布	200
6.8.10	随机数函数自变量 SEED	201
6.9	SAS CALL 子程序	201
6.9.1	CALL 子程序类型	201
6.9.2	随机数子程序	202
	习题 6	203

第 7 章 通用语句 205

7.1	全局通用语句	205
7.1.1	注释语句	205
7.1.2	DM 语句与 X 语句	206
7.1.3	TITLE 语句与 FOOTNOTE 语句	208
7.1.4	RUN 语句与 ENDSAS 语句	209
7.1.5	LIBNAME 语句	210
7.1.6	FILENAME 语句	212
7.1.7	%INCLUDE 语句	215
7.1.8	%RUN 语句与 %LIST 语句	217
7.1.9	MISSING 语句	217
7.1.10	PAGE 语句与 SKIP 语句	218
7.1.11	OPTIONS 语句与 GOPTIONS 语句	218
7.2	过程步通用语句	220
7.2.1	PROC 语句	220

7.2.2	VAR 语句与 MODLE 语句	222
7.2.3	ID 语句与 WHERE 语句	222
7.2.4	CLASS 语句与 BY 语句	224
7.2.5	OUTPUT 语句与 QUIT 语句	225
7.2.6	FORMAT 语句与 ATTRIB 语句	226
7.2.7	LABEL 语句	227
习题 7	228
第 8 章	变量输入输出格式	230
8.1	输入格式	230
8.1.1	输入格式形式	230
8.1.2	输入格式使用方法	231
8.1.3	输入格式类型	232
8.1.4	数值变量输入格式	232
8.1.5	字符变量输入格式	235
8.2	输出格式	237
8.2.1	输出格式形式	237
8.2.2	输出格式使用方法	237
8.2.3	输出格式类型	238
8.2.4	数值变量输出格式	239
8.2.5	字符变量输出格式	241
8.3	日期时间存储方式	242
8.3.1	SAS 日期值存储方式	242
8.3.2	SAS 日期时间值存储方式	243
8.4	日期时间输入格式	243
8.4.1	日期时间输入格式	243
8.4.2	应用举例	244
8.5	日期时间输出格式	244
8.5.1	日期时间输出格式	244
8.5.2	应用举例	249
8.6	缺失值处理	251
8.6.1	读入含缺失值的数据	251
8.6.2	系统产生的缺失值	252
习题 8	252
第 9 章	输出控制	254
9.1	输出窗口与内容	254
9.2	日志输出控制	254

9.2.1	日志输出信息类型	254
9.2.2	日志输出信息控制	255
9.2.3	日志输出地点	255
9.2.4	定制日志格式	256
9.2.5	应用举例	256
9.3	运行结果输出控制	256
9.3.1	定制输出格式	257
9.3.2	运行结果输出地点	257
9.3.3	应用举例	258
9.4	图形存储利用与输出	258
9.4.1	图形存储	258
9.4.2	GREPLAY 过程	259
9.4.3	输出其他格式图形文件	260
9.5	输出传送系统	261
9.5.1	ODS 功能	261
9.5.2	ODS 对象与传送目标	261
9.5.3	ODS 语句	262
9.5.4	传送目标控制	262
9.5.5	选择输出对象	267
习题 9		268
第 10 章	数据管理	271
10.1	数据集排序	271
10.1.1	排序过程句法	271
10.1.2	PROC SORT 语句	271
10.1.3	BY 语句	272
10.1.4	应用举例	272
10.2	数据集转置	273
10.2.1	转置过程句法	273
10.2.2	PROC TRANSPOSE 语句	273
10.2.3	VAR 语句和 ID 语句	274
10.2.4	应用举例	274
10.3	添加观测	276
10.3.1	APPEND 过程句法	277
10.3.2	选项说明	277
10.3.3	应用举例	277
习题 10		277

第 11 章	统计量计算	279
11.1	相关性过程	279
11.1.1	相关过程句法	279
11.1.2	PROC CORR 语句	280
11.1.3	其他语句	280
11.1.4	应用举例	281
11.2	频数过程	283
11.2.1	频数过程句法	283
11.2.2	PROC FREQ 语句	283
11.2.3	TABLES 语句	284
11.2.4	WEIGHT 语句	286
11.2.5	BY 语句	286
11.2.6	OUTPUT 语句	286
11.2.7	应用举例	286
11.3	均值过程	288
11.3.1	均值过程句法	288
11.3.2	PROC MEANS 语句	289
11.3.3	其他语句	290
11.3.4	应用举例	292
11.4	单变量过程	295
11.4.1	单变量过程句法	295
11.4.2	PROC UNIVARIATE 语句	295
11.4.3	其他语句	296
11.4.4	应用举例	298
习题 11	301
第 12 章	制表与绘图	303
12.1	制表过程	303
12.1.1	制表过程句法	303
12.1.2	TABLE 语句	304
12.1.3	应用举例	304
12.2	作图过程	319
12.2.1	作图过程句法	319
12.2.2	PLOT 语句	319
12.2.3	SYMBOL 语句	320
12.2.4	AXIS 语句	320
12.2.5	应用举例	320
12.3	图表过程	329