

SAS技术内幕



[SAS技术内幕 下载链接1](#)

著者:巫银良

出版者:清华大学出版社

出版时间:2018-12-1

装帧:平装

isbn:9787302502784

本书共27章，分为上下两卷：上卷介绍SAS编程基础与使用方法，是广大程序员快速掌握SAS编程技术的简明开发教程；下卷阐述数据分析的关键基础知识并提供相应SAS代码实现，目的是激发读者兴趣，跨越传统编程与数据分析的鸿沟，从程序员华丽转身为数据科学家。书中演示代码力图简洁清晰地解释相关概念，追求大道至简。本书兼顾编

程技术与数据分析，期许程序员、信息处理与统计分析人员以及对数据分析科学感兴趣的读者都能从本书中获益良多，循序渐进地掌握数据分析的要义和精髓，从数据中获取洞见与智慧。

作者介绍:

主要负责SAS中国研发中心大数据与可视化分析技术，可视化数据挖掘与机器学习，商业智能和移动应用等产品线的项目管理。

毕业于北京大学，获理学学士和硕士学位。拥有近20年计算机行业研发和管理经验，在企业应用和商业分析领域有深厚的技术功底。专注于研发管理和项目实践，涉及领域包括人工智能，大数据与可视化分析，金融量化分析与自动交易，商业智能与互联网，大型商业软件研发管理等。

目录: 上 卷

第1章 SAS语言入门 002

1.1 语言概述 005

1.2 编程环境 010

1.3 SAS Studio编程 016

第2章 数据集与DATA步 032

2.1 SAS 逻辑库 032

2.2 SAS 数据集 036

2.3 DATA 步 041

2.3.1 内嵌数据行或外部数据文件 043

2.3.2 通过已有 SAS 数据集生成 048

2.3.3 通过 PROC IMPORT 或 PROC SQL 生成 051

2.4 DATA 步的运行机制 054

2.4.1 编译阶段 055

2.4.2 运行阶段 056

2.5 DATA 步语句快速索引 057

第3章 变量与表达式 062

3.1 常量与变量 062

3.1.1 变量长度与缺失值 063

3.1.2 数值常量 065

3.1.3 日期/时间/日期时间常量 067

3.1.4 字符常量 067

3.2 表达式 068

3.2.1 运算符 068

3.2.2 运算符优先顺序 072

3.2.3 WHERE 语句特定运算符 072

3.2.4 赋值语句 073

3.2.5 累加赋值语句 074

3.2.6 RETAIN 语句 075

3.3 SAS数组 077

3.3.1 数组名称 079

3.3.2 数组元素变量列表 079

3.3.3 数组长度 080

3.3.4 隐式下标变量 080

3.3.5 多维数组 081

3.3.6 临时数组 082

3.3.7 数组排序 083

3.3.8 注意事项 084

- 第4章 流程控制 087
 - 4.1 DO-END语句块 087
 - 4.2 分支控制 088
 - 4.2.1 IF-THEN 分支控制 088
 - 4.2.2 ELSE-IF 多分支控制 090
 - 4.2.3 SELECT-WHEN 多分支控制 090
 - 4.3 循环控制 091
 - 4.3.1 指定次数的循环：DO-TO-BY 091
 - 4.3.2 指定条件的循环：DO-WHILE 与 DO-UNTIL 092
 - 4.3.3 指定集合的循环：DO-OVER 093
 - 4.4 特殊的流程控制语句 094
 - 4.4.1 跳出循环语句：LEAVE 094
 - 4.4.2 继续循环语句：CONTINUE 094
 - 4.4.3 返回语句：RETURN 095
 - 4.4.4 中止执行语句：STOP 与 ABORT 095
 - 4.4.5 跳转语句：GOTO 与 LINK 096
- 第5章 函数封装 099
 - 5.1 LINK-RETURN 技术 101
 - 5.2 SAS宏函数封装 103
 - 5.3 FCMP函数 105
 - 5.4 系统函数速查 110
- 第6章 SAS宏 113
 - 6.1 宏变量 114
 - 6.1.1 命名 114
 - 6.1.2 作用域 114
 - 6.1.3 系统宏 115
 - 6.1.4 宏代码调试 117
 - 6.1.5 宏表达式 117
 - 6.2 宏函数 119
 - 6.3 逻辑控制 121
 - 6.3.1 宏语句块 121
 - 6.3.2 条件分支 121
 - 6.3.3 循环控制 121
 - 6.4 系统宏函数 122
- 第7章 DS2 124
 - 7.1 程序结构 126
 - 7.1.1 变量声明与类型 126
 - 7.1.2 程序实体作用域 129
 - 7.1.3 变量数组与标准数组 130
 - 7.1.4 系统方法与用户自定义方法 131
 - 7.2 数据程序 136
 - 7.3 包程序 137
 - 7.4 线程程序 143
- 第8章 代码组织 149
 - 8.1 静态文件包含 149
 - 8.2 程序中动态扩展代码 151
 - 8.3 动态执行外部命令 153
- 第9章 文件读写 157
 - 9.1 二进制文件读写 157
 - 9.2 文本文件读写 159
 - 9.3 顺序读取多个文件 162
 - 9.4 并行读取多个文件 165
 - 9.5 共享缓冲区读写 166
- 第10章 按位运算 168

- 10.1 按位运算 168
- 10.2 实现方法 169
- 10.3 按位运算应用 173
- 第11章 扩展SAS功能 177
 - 11.1 通过 Module调用外部 DLL函数 177
 - 11.2 用 C 语言开发用户函数库 182
 - 11.2.1 准备64位 C 编译环境 182
 - 11.2.2 开发用户自定义动态库 183
 - 11.3 PROTO 编写 C 代码或注册外部 DLL 184
- 第12章 数据结构——数组 187
 - 12.1 数组 187
 - 12.1.1 DATA步数组 187
 - 12.1.2 FCMP 数组 189
 - 12.1.3 DS2 数组 192
 - 12.1.4 SAS宏数组 193
 - 12.2 数组应用：高精度数值计算 194
- 第13章 数据结构——队列与堆栈 196
 - 13.1 队列 196
 - 13.1.1 函数版实现与示例 196
 - 13.1.2 宏版实现与示例 199
 - 13.2 堆栈，函数版实现与示例 202
- 第14章 数据结构——链表 206
 - 14.1 基础知识 206
 - 14.2 如何在 SAS 代码中内嵌 C 语言代码 207
 - 14.3 单向链表和双向链表 209
 - 14.4 链表应用：约瑟夫斯问题 216
- 第15章 数据结构——二叉树 221
 - 15.1 PROTO 实现与封装 221
 - 15.2 FCMP二叉树实现 227
 - 15.3 二叉树应用：算术表达式求值 231
- 第16章 数据结构——矩阵运算 235
 - 16.1 FCMP 矩阵运算 236
 - 16.2 DS2 矩阵运算 243
 - 16.3 矩阵应用：线性方程组求解 246
 - 16.4 矩阵应用：非线性方程组求解 248
- 第17章 数据结构——图 255
 - 17.1 深度优先和广度优先遍历 256
 - 17.2 最短路径问题 260
 - 17.2.1 Dijkstra算法 261
 - 17.2.2 Bellman-Ford 算法 263
 - 17.2.3 Floyd-Warshall 算法 265
- 下 卷
- 第18章 统计学基础 270
 - 18.1 数据特征度量 270
 - 18.1.1 集中趋势度量 272
 - 18.1.2 离散趋势度量 274
 - 18.1.3 分布特征度量 277
 - 18.1.4 置信区间、置信水平与 p-值 279
 - 18.2 统计学上的变量类型 280
 - 18.3 基本数据处理 283
 - 18.3.1 排序与排名 284
 - 18.3.2 数据转置 285
 - 18.3.3 堆叠与拆分 286
 - 18.3.4 过滤数据 287

18.3.5 随机抽样	289
18.3.6 基本统计量	290
18.4 基本图形图表	292
18.5 SAS 产品与过程步概览	303
18.5.1 SAS核心产品功能简介	305
18.5.2 Base SAS 过程步速查	309
18.5.3 SAS/STAT过程步速查	318
第19章 大数定律与中心极限定理	327
19.1 大数定律	327
19.1.1 弱大数定律	327
19.1.2 三种大数定律	329
19.1.3 图形化证明	330
19.1.4 强大数定律	333
19.2 中心极限定理	334
19.2.1 大数定律与中心极限定理关系	335
19.2.2 图形化证明	336
19.2.3 实际用途	340
第20章 统计分布	342
20.1 均匀分布	342
20.2 离散型统计分布	345
20.2.1 伯努利分布	345
20.2.2 二项分布	347
20.2.3 几何分布	353
20.2.4 负二项分布	357
20.2.5 超几何分布	360
20.2.6 泊松分布	362
20.3 连续型统计分布	365
20.3.1 正态分布	365
20.3.2 对数正态分布	373
20.3.3 指数分布	376
20.3.4 卡方分布	379
20.3.5 学生t-分布	381
20.3.6 F分布	387
20.3.7 柯西分布	390
20.3.8 贝塔分布	392
20.3.9 伽马分布	395
20.3.10 爱尔朗分布	397
20.3.11 韦布尔分布	399
20.3.12 三角分布	400
20.3.13 Table 分布	401
附录：各统计分布之间的关系	403
第21章 方差分析	404
21.1 假设检验	404
21.2 方差分析	406
21.2.1 学生t-检验	406
21.2.2 单因子方差分析	408
21.2.3 双因子方差分析	418
第22章 数据标准化	421
22.1 常用标准化方法	421
22.2 SAS数据标准化	424
22.3 自定义数据标准化	429
第23章 主成分分析与因子分析	433
23.1 主成分分析	434
23.1.1 主成分分析原理	435

23.1.2 主成分分析的具体步骤	436
23.2 因子分析	443
23.2.1 因子分析原理	443
23.2.2 巴特利球度检验和 KMO 检验	443
23.2.3 因子分析的具体步骤	445
第24章 相关分析与回归分析	450
24.1 变量关系	450
24.2 相关分析	451
24.2.1 线性相关性度量	451
24.2.2 非参数关联度量	452
24.2.3 定量数据的相关分析	455
24.2.4 类别数据的相关分析	457
24.3 回归分析	460
第25章 聚类分析	467
25.1 聚类度量	469
25.1.1 距离系数	469
25.1.2 相似性/相关系数	471
25.1.3 SAS实践	473
25.2 聚类形成方法	475
25.2.1 一次形成分类系统	475
25.2.2 K-均值聚类	477
25.2.3 逐步形成分类系统	485
25.2.4 R 型聚类分析	491
25.3 自己实现聚类算法	494
25.3.1 K-均值方法	494
25.3.2 逐步形成分类系统	501
附录：聚类度量的自定义实现	509
第26章 神经网络	512
26.1 神经元模型	513
26.2 神经网络	517
26.2.1 训练神经网络	519
26.2.2 反向传播算法	519
26.3 SAS 代码实现与范例	524
第27章 π 高精度求解与探索分析	536
27.1 π 值计算	537
27.1.1 蒙特卡罗方法	543
27.1.2 蒲丰投针方法	544
27.1.3 微积分方法	545
27.1.4 幂级数方法	546
27.1.5 幂级数高精度方法	548
27.1.6 梅钦类公式高精度方法	550
27.1.7 迭代方法——贝拉公式	554
27.2 π 值分析	557
27.2.1 数字分布规律	558
27.2.2 可视化探索	561
附录	564

参考文献 598

• • • • • ([收起](#))

[SAS技术内幕_下载链接1](#)

标签

SAS技术内幕：从程序员到数据科学家

魔幻现实主义

霍乱时期的爱情

爱情

外国文学

加西亚·马尔克斯

评论

书的标题和文的内容不匹配，并且内容应用性差，深度不够，不能让想基础学习的人学到知识、也不能让想深度学习的人学到更多知识

[SAS技术内幕_下载链接1_](#)

书评

[SAS技术内幕_下载链接1_](#)