



杨池然 仲文明 周志勇 编著

SAS 9.2

从入门到精通



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

SAS 9.2 从入门到精通

杨池然 仲文明 周志勇 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书以 SAS 9.2 为开发平台, 通过 SAS 与关系数据库 ORACLE 相结合的应用, 与实际开发应用案例相结合, 详细讲解 SAS 编程基础语法、数据步处理机制、数据步对各类外部文件处理、SAS 数据集数据装载到 ORACLE 数据库、ORACLE 数据库数据生成 SAS 数据集、过程步处理机制、常用统计过程、SQL 过程、ODS 输出控制、数据挖掘、SAS 编程效率优化和综合案例分析。每个章节和知识点都提供了实际应用中的经典案例, 添加详细的注释和程序解读, 旨在帮助初学者和有经验的读者更深入地掌握 SAS 编程语言。SAS 数据文件处理和数据集处理及 SAS 编程效率优化是所有模块都需要的知识点, 其中 SAS 编程效率优化、SAS 与关系数据库 ORACLE 结合的实际应用案例、数据挖掘和经典程序综合讲解与分析是本书的一个亮点。

本书适合金融、统计学、经济、数学和计算机等专业的读者阅读, 也可作为各类 SAS 工作者的参考用书和学校教材。

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

SAS 9.2 从入门到精通/杨池然, 仲文明, 周志勇编著. —北京: 电子工业出版社, 2011.8

ISBN 978-7-121-14284-0

I. ①S… II. ①杨… ②仲… ③周… III. ①统计分析—应用软件, SAS 9.2 IV. ①C812

中国版本图书馆 CIP 数据核字 (2011) 第 158537 号

责任编辑: 戴 新

文字编辑: 徐 萍

印 刷: 三河市鑫马印装有限公司

装 订:

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

北京市海淀区翠微东里甲 2 号 邮编: 100036

开 本: 787×1092 1/16 印张: 25.5 字数: 650 千字

印 次: 2011 年 8 月第 1 次印刷

定 价: 52.00 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。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前 言

本书结合作者多年的实践和处理大数据量的数据经验而写。让初学者知道如何入门，如何提高兴趣，语言基础是关键。本书结合开发应用案例讲述，使读者更能深刻理解 SAS 语言，而不是一开始就把大家带入统计领域。

本书通过实例讲解让初学者快速入门，循序渐进地学习并有所收获；让有基础的读者深入提高；让有经验的读者学会优化程序，提高效率。本书理论与实践互相结合，相得益彰。

本书重点对 SAS 的核心模块 BASE SAS (SAS 所有产品的基础模块)、SAS/ACCESS TO PC FILES (用于读取和生成多种格式的文件模块)、SAS/ACCESS (该模块提供 SAS 软件访问各类数据库的接口) 和 SAS/EM (企业数据挖掘模块) 进行详细讲解，并结合 ORACLE 关系数据库进行介绍。书中对用 SAS 读取各类数据文件进行了详细的讲解并给出了每一个类型文件的实例，以帮助读者理解 SAS 读取各类数据文件的方法。本书通过实例来讲解 SAS 与数据库互相结合的奥妙，使读者能够深入理解 SAS 在实际中的应用价值。

第 1~第 2 章主要讲解 SAS 9.2 的新特性、发展历史和 SAS 编程语言，从此版本开始支持自定义函数是 SAS 系统的一个大跨越，编程语言结合丰富的实例和注释进行了详细讲解。

第 3~第 5 章主要讲解 SAS 9.2 各类函数，其中 SAS 9.2 自定义函数单独一节通过实例进行详细的讲解；数据步处理重点讲解数据步对各类文件的处理，以及根据业务需求对函数的处理；过程步部分讲解过程的语法、基本语句和实际调用。

第 6~第 9 章主要讲解 SQL 过程、ODS 输出控制将分析结果存储到各类型格式的文件中和宏过程，数据文件的装载是结合关系数据库 ORACLE 进行的讲解，是 SAS 与数据库结合应用的讲解。

第 10~第 13 章主要讲解 SAS 系统的常用过程、数据挖掘、SAS 编程优化与调用、经典程序综合应用与分析。

本书具有以下特点：

- ◆ SAS 编程语言语法详细，实例丰富，并添加有程序解读和详细注释。
- ◆ 通过案例讲解 SAS 9.2 自定义函数，并与实际应用相结合。
- ◆ 专门讲解 SAS 对各类文件处理，以及 SAS 与 ORACLE 数据库结合应用。
- ◆ 数据挖掘案例 SAS/EM 的实际应用。
- ◆ 通过经典案例将理论与实际应用相结合对 SAS 语言进行讲解。

本书每一部分的讲解都引入实际开发应用案例进行说明和注释，使理论和实践相吻合。看每一章都有收获，都有反思，让读者在探索中思考。

参加本书编写的人员还有孙更新、宾晟、满在龙、宫生文、李海涛、李晓娜、王洪臣，在此表示感谢，同时也感谢同事、朋友和家人的支持。由于作者水平有限，加上时间仓促，书中难免存在不妥与疏漏之处，欢迎广大读者批评指正，以便于以后修订。

为方便读者阅读，若需要本书配套资料，请登录“北京美迪亚电子信息有限公司”(<http://www.medias.com.cn>)，在“资料下载”页面进行下载。

目 录

第 1 章 SAS 概述	1	2.2.8 格式修饰符	28
1.1 SAS 发展历史与应用领域	1	2.2.9 指针控制读取	30
1.1.1 SAS 发展历史	1	2.2.10 INPUT 语句	32
1.1.2 SAS 应用领域	1	2.2.11 INFILE 读取文件	36
1.2 SAS 模块和处理流程	2	2.2.12 DELIMITER 应用	40
1.2.1 SAS 模块	2	2.2.13 日期与时间	41
1.2.2 SAS 处理流程	3	2.2.14 INFORMAT 和 FORMAT 读取数据定义格式	43
1.3 SAS 9.2 新特性	3	2.2.15 LENGTH 应用	45
1.3.1 SAS 9.2 新增 Base SAS 过程	3	2.2.16 LABEL 定义变量标签	45
1.3.2 SAS 9.2 ACCESS 新增功能	4	2.2.17 FILE 写数据到外部文件	46
1.3.3 SAS 9.2 其他模块新增功能	4	2.2.18 RETAIN 语句	49
1.4 SAS 9.2 安装步骤与启动	4	2.2.19 RENAME 语句	50
1.4.1 SAS 9.2 安装步骤	4	2.2.20 SAS 编程注释语句	51
1.4.2 SAS 9.2 启动与运行	7	第 3 章 SAS 9.2 函数	52
1.5 SAS 开发工具与操作界面	8	3.1 函数概述	52
1.5.1 SAS 9.2 操作界面	8	3.1.1 函数定义	52
1.5.2 编辑器	8	3.1.2 函数参数书写形式	53
1.5.3 输出	8	3.1.3 字符处理函数	54
1.5.4 日志	9	3.1.4 数学运算函数	68
1.6 SAS 程序书写要求	9	3.1.5 日期和时间函数	80
第 2 章 SAS 编程语言	10	3.1.6 统计函数	90
2.1 SAS 逻辑库	10	3.1.7 SAS 9.2 自定义函数	97
2.1.1 逻辑库定义	10	3.2 函数综合应用案例	99
2.1.2 语法定义逻辑库	11	第 4 章 SAS 数据步的建立	100
2.1.3 图形操作定义逻辑库	12	4.1 数据步概述	100
2.1.4 逻辑库引用	13	4.1.1 数据步的作用	100
2.1.5 删除逻辑库	14	4.1.2 数据步基本语句应用	102
2.2 SAS 编程语法	14	4.1.3 数据集整理	111
2.2.1 变量	15	4.2 数据步读 DAT 格式文件建立 数据集	122
2.2.2 观测值	18	4.2.1 DAT 格式文件建立数据集	122
2.2.3 常量	18	4.2.2 DAT 格式文件特殊分隔符 的处理	124
2.2.4 SAS 操作符	20		
2.2.5 宏变量	22		
2.2.6 数组	22		
2.2.7 选择语句与循环语句	23		

4.2.3	DAT 格式文件特殊需求的 字段处理	125	6.2.11	SAS 与关系数据库的连接 和断开语句	174
4.3	数据步读 Excel 格式文件建立 数据集	126	6.2.12	子查询语句	175
4.3.1	Excel 格式文件建立数据集	126	6.3	SQL 过程与 ORACLE 数据库连接	176
4.3.2	Excel 格式文件数据集处理	127	6.3.1	SAS 建立与 ORACLE 数据库 连接方式	177
4.4	数据步读主机 COBOL 语言生成 文件建立数据集	128	6.3.2	SELECT 语句综合应用	177
4.5	数据步读 TXT 格式文件建立 数据集	131	6.3.3	数据库表中的字段值赋值给 所定义的宏变量	178
4.6	数据步读入 CARDS 或 DATALINES 语句建立数据集	132	6.3.4	数据库查询出来的值赋值给 数据集集中的变量	179
4.6.1	CARDS 语句中的数据建立 数据集	132	6.4	SQL 过程索引应用	181
4.6.2	DATALINES 语句中的数据 建立数据集	133	6.4.1	索引的优点与缺点	181
4.7	ORACLE 数据库表中的数据建立 数据集	133	6.4.2	数据步创建索引与 SQL 过程 创建索引比较	181
4.8	批量文件处理案例	134	6.4.3	删除索引	183
第 5 章	SAS 过程步应用	136	6.5	SQL 过程综合应用案例	184
5.1	过程步概述	136	第 7 章	SAS 输出控制	187
5.1.1	过程步处理机制与作用	136	7.1	数据步方式输出到指定文件	187
5.1.2	过程步基本语句	138	7.1.1	数据步方式输出语法	187
5.1.3	过程步变量属性语句	147	7.1.2	数据集解数到指定目录 DAT 格式文件	187
5.1.4	过程步选项参数	150	7.1.3	ORACLE 数据库表数据根据 需求解数到 DAT 格式文件	189
5.2	信用卡风险分析案例	151	7.2	PRINTTO 过程输出到指定文件	190
第 6 章	SQL 过程	154	7.2.1	PRINTTO 用途与定义	190
6.1	SQL 概述	154	7.2.2	PRINTTO 过程应用实例	190
6.2	数据操纵语句	156	7.3	ODS 输出控制	191
6.2.1	数据操纵语句概述	156	7.3.1	ODS 概述与处理流程	192
6.2.2	创建表语句与应用实例	157	7.3.2	ODS 常用语法	192
6.2.3	创建索引语句与应用实例	159	7.3.3	ODS 输出到 HTML 格式 文件	198
6.2.4	创建视图语句与应用实例	160	7.3.4	ODS 输出到 PDF 格式文件	199
6.2.5	查询语句与应用实例	162	7.3.5	ODS 输出到 RTF 格式文件	201
6.2.6	删除语句与应用实例	170	7.3.6	ODS 输出到 DAT 格式文件	202
6.2.7	修改表语句与应用实例	171	7.3.7	ODS 输出到 TXT 格式文件	202
6.2.8	插入语句与应用实例	172	7.3.8	ODS 输出到 CSV 格式文件	204
6.2.9	更新语句与应用实例	173	7.3.9	ODS SELECT 选择输出对象	205
6.2.10	显示表定义信息语句与 应用实例	174	7.3.10	ODS EXCLUDE 排除输出 对象	206

7.3.11 ODS PRINTER 输出到 打印机文件.....	207	9.1.3 宏编程语句.....	249
7.3.12 ODS 与 ORACLE 数据库 交互.....	209	9.1.4 宏变量创建方法与分类.....	252
7.4 ODS 输出综合应用案例.....	210	9.1.5 宏函数.....	256
第 8 章 数据装载与数据集管理	212	9.1.6 宏嵌套.....	266
8.1 APPEND 过程.....	212	9.1.7 宏过程调用数据步.....	267
8.1.1 APPEND 过程语法.....	212	9.1.8 宏过程调用过程步.....	268
8.1.2 APPEND 过程应用实例.....	213	9.1.9 SQL 过程引用宏变量.....	269
8.1.3 SAS 与 ORACLE 数据库 连接.....	216	9.1.10 宏过程内部编译顺序.....	270
8.1.4 APPEND 与 ORACLE 数据库结合应用.....	217	9.2 宏过程综合应用案例.....	271
8.1.5 宏过程调用 APPEND 过程应用.....	222	第 10 章 常用过程	276
8.1.6 APPEND 过程综合应用 案例.....	223	10.1 统计过程概述.....	276
8.2 DBLOAD 过程.....	225	10.1.1 统计过程语法定义.....	276
8.2.1 DBLOAD 过程语法.....	225	10.1.2 常用统计过程与应用实例 分析.....	278
8.2.2 DBLOAD 过程应用实例.....	226	10.2 GLM 过程与 MIXED 过程.....	289
8.3 数据步数据装载.....	227	10.2.1 GLM 过程与 MIXED 过程 对比.....	289
8.4 SQL 过程数据装载.....	228	10.2.2 GLM 过程语法定义.....	289
8.5 SORT 过程数据集整理.....	229	10.2.3 GLM 过程应用实例.....	291
8.5.1 SORT 过程语法.....	229	10.2.4 MIXED 过程语法定义.....	293
8.5.2 SORT 过程应用实例.....	230	10.2.5 MIXED 过程应用实例.....	293
8.6 DATASETS 过程数据集整理.....	233	10.3 CHART 制图过程.....	294
8.6.1 DATASETS 过程语法.....	233	10.3.1 CHART 过程语法定义.....	294
8.6.2 DATASETS 过程基本语句 与应用实例.....	234	10.3.2 CHART 过程应用实例.....	295
8.7 DAT 格式文件数据装载综合应用 案例.....	237	10.4 TABULATE 制表过程.....	296
8.8 Excel 格式文件数据装载综合应用 案例.....	241	10.4.1 TABULATE 过程语法 定义.....	296
8.8.1 Excel 格式文件装载到 数据库表.....	241	10.4.2 TABULATE 过程应用 实例.....	297
8.8.2 Excel 格式文件装载到 数据库表举例.....	241	10.5 PRINT 打印过程.....	299
第 9 章 宏过程	243	10.5.1 PRINT 过程语法定义.....	299
9.1 宏过程概述.....	243	10.5.2 PRINT 过程应用实例.....	300
9.1.1 宏过程执行机制与定义.....	243	10.6 CORR 相关过程.....	301
9.1.2 宏过程调用.....	248	10.6.1 CORR 过程语法定义.....	301
		10.6.2 CORR 过程应用实例.....	302
		10.7 REG 回归过程.....	303
		10.7.1 REG 过程语法定义.....	303
		10.7.2 REG 过程应用实例.....	305
		10.8 LOGISTIC 回归.....	306
		10.8.1 LOGISTIC 过程语法定义.....	306
		10.8.2 LOGISTIC 过程应用实例.....	309

第 11 章 数据挖掘	314
11.1 数据挖掘简介	314
11.1.1 数据挖掘的概念	314
11.1.2 数据挖掘产生的条件	314
11.1.3 数据挖掘的功能	315
11.2 SAS 数据挖掘 EM	315
11.2.1 SAS 数据挖掘 EM 模块概述	315
11.2.2 SAS/EM 数据挖掘流程	317
11.3 SAS/EM 控件	319
11.3.1 Sample 数据取样控件	319
11.3.2 Explore 数据探索控件	319
11.3.3 Modify 数据调整控件	320
11.3.4 Model 模型化控件	320
11.3.5 Assess 评价控件	321
11.3.6 Scoring 得分值控件	321
11.3.7 Utility 类控件	321
11.4 SAS 数据挖掘案例	322
11.4.1 业务分析	322
11.4.2 目标定位	322
11.4.3 模型开发与应用流程	323
11.4.4 结果评价与分析报告	343
第 12 章 SAS 编程优化与调用	344
12.1 SAS 程序效率概述	344
12.1.1 提高效率的解决机制	344
12.1.2 数据步执行机制	345
12.1.3 数据步语句优化	345
12.1.4 逻辑库 LIBNAME 在 UNIX 环境下的优化	354
12.1.5 IF 与 WHERE 语句的差别	355
12.1.6 KEEP 语句位置与程序效率	356
12.1.7 IF 语句中函数判断使用	358
12.1.8 IF 语句中条件的放置位置	359
12.1.9 减少 I/O 输出	361
12.1.10 IF 与字段处理	362
12.1.11 DATASETS 过程优化数据集	363
12.1.12 索引应用	364
12.1.13 MODIFY 语句与 SET 语句的转换	364

12.1.14 LENGTH 语句应用节省存储空间	365
12.1.15 DROP 语句删除不需要的变量	366
12.2 shell 程序调用 SAS	366
12.2.1 shell 语言概述	366
12.2.2 shell 程序调用 SAS	367
第 13 章 经典程序综合讲解与分析	368
13.1 数据处理技巧案例	368
13.1.1 开发调试技巧	368
13.1.2 数据处理技巧	369
13.2 综合程序应用讲解	376
13.2.1 学校学生成绩加分处理	376
13.2.2 根据信用度级别处理发卡级别	376
13.2.3 CATS 函数的应用	377
13.2.4 日期显示格式	378
13.2.5 宏函数判断目录存在与否	379
13.2.6 取出数据库中数据集的前 16 条记录	379
13.2.7 删除数据集中的最后一条记录	380
13.2.8 SAS 与数据库连接接口安全性加密	380
13.2.9 SAS 调用关系数据库 ORACLE 的存储过程	382
13.2.10 插入数据到 ORACLE 数据库表	382
13.2.11 根据条件查询 ORACLE 数据库表	383
13.2.12 表格数据处理方法	384
13.2.13 员工工资计算	385
13.2.14 写文件头信息解数到文件	386
13.2.15 HASH 应用	388
13.2.16 数据库信息解数到指定文件	389
13.2.17 连接数据库的两种方式	389
13.2.18 宏过程综合应用案例	393
13.2.19 等级显示处理案例	395
参考文献	397

第 1 章 SAS 概述

SAS 最初是“Statistical Analysis Software”三个英文单词的简写，随着 SAS 的发展，SAS 已经不再单是一个统计软件，原来的三个英文单词已经不能更好地描述 SAS，现在普遍称为“Statistics Analysis System”三个英文单词的简写，发展成为一个组合软件系统。本章主要讲解 SAS 语言和 SAS 工具的发展历史及 SAS 软件各个模块在各个应用领域的应用概况。通过本章的学习让初学者了解如何学习和运用 SAS 语言和 SAS 工具，可以整体把握 SAS 语言和 SAS 工具的架构，知道哪些是学习的重点，哪些是必须掌握的，并根据应用需求学习 SAS 对应的模块。

1.1 SAS 发展历史与应用领域

SAS 软件从 20 世纪 70 年代诞生到成为世界公认的国际标准软件，足以说明了 SAS 的强大功能和应用领域的广泛性。随着时间的发展，SAS 语言和 SAS 工具将越来越强大，只会在原来的基础上提升功能，改善应用，而不会过时。因为 SAS 系统是把数学理论和实际应用结合起来的一个组合软件系统。

1.1.1 SAS 发展历史

SAS (Statistics Analysis System) 软件是在 20 世纪 70 年代由北卡罗莱纳州立大学编写出来的，最初是北卡罗莱纳州立大学两个学生编制了程序，当时用来处理数据，做生物分析用。随着 SAS 的发展，SAS 研究院于 1975 年成立。SAS 公司逐渐发展壮大，现在已经成为世界最大的独立软件公司之一。SAS 从版本 5 系列发展到现在的最新版本 9 系列，随着版本的发展，SAS 内部自带函数和过程一直在增加，功能逐渐强大和完善，图形化的操作也变得更灵活。

最新版本 9 系列中，SAS 9.2 开始支持自定义函数，这是 SAS 软件的一个大跳跃，也是 SAS 9 系列的一个突出点，显示了 SAS 的强大功能。

1.1.2 SAS 应用领域

SAS 主要应用在金融、数学、通信、经济、生物医药、教育机构和科研院所等领域。现在全世界很多国家都在运用 SAS。SAS 在金融领域可以用来做风险分析，信用卡开卡的评分系统，底层数据抽取、转换与装载的工作，简称为 ETL (Extract, Transform and Load)，处理大数据量的数据。在 BI (Business Intelligence, 商业智能) 方面，SAS 通过前端展现工具 SAS/Enterprise Guide (简称 EG) 用于报表的开发，SAS /EG 组件就是用来开发报表的组件。

在统计分析领域，SAS 是世界上公认的最具权威的统计分析工具之一，其统计分析指标得到国际公认。

SAS 在通信领域可以对客户短信消费类型进行分析,挖掘潜在客户。

SAS 在生物医药领域可以用来对生物医药研发的指标进行统计分析。

通过以上所列举的领域,可以看到 SAS 的应用领域非常广泛,发展前景有目共睹。SAS 工具集成了很多数学领域的公式,通过封装,用户可以直接调用,运用封装好的数学公式来进行指标的分析。只需要知道这些分析出来的指标代表什么意义,就可以根据这些指标写出分析报告,做出对未来的预测和决策。

随着信息时代的发展,SAS 语言和 SAS 工具将会越来越强大。SAS 语言不会随时代的变迁而过时,只会在原来的基础上越来越完善,功能更强大,适合领域更广泛。

1.2 SAS 模块和处理流程

本节对 SAS 软件主要模块和处理流程进行介绍。由于 SAS 工具包含的模块众多,因此重点掌握 SAS 工具的核心模块 Base SAS 模块,其他模块根据需求可以有选择地重点学习。

1.2.1 SAS 模块

SAS 软件由很多不同功能的模块组成,其中 Base SAS 模块是 SAS 系统的核心,是 SAS 所有产品的基础模块,提供数据管理功能,支持关系数据库标准 SQL 语言对数据的处理,进行统计分析,是 SAS 中不可缺少的模块。其他模块有 SAS/STAT (统计模块)、SAS/ACCESS TO PC FILES (用于读取和生成多种格式的文件模块)、SAS/GRAPH (图形模块)、SAS/EM (数据挖掘模块)、SAS/ACCESS (该模块提供 SAS 软件访问各类数据库的接口)、SAS/QC (质量控制模块)、SAS/ETS (经济计量和时间序列模块),等等。表 1-1 所示为 SAS 主要模块与功能详解。

表 1-1 SAS 主要模块与功能详解

SAS 主要模块	模块功能与应用领域
BASE SAS	SAS 系统核心模块,是所有产品模块的基础,实现模块的调度、数据管理、数据访问,支持关系数据库 SQL 语言的处理,是统计分析的基础和报表生成的基础模块
SAS/ACCESS	连接其他数据库的接口,如连接 ORACLE 数据库、DB2、Sybase 和 Informix 等主流关系数据库,就用到此模块,实现 SAS 与数据库的交互操作
SAS/ACCESS TO PC FILES	SAS 对外部文件的读取和生成各类型的文件需要调用此模块,如读取.dat 格式文件、.txt 格式文件、.csv 格式文件和 COBOL 语言生成的 IBM 主机文件等都需要调用此模块。通过此模块把文件转换成 SAS 能够识别的数据集,再运用其他模块进行分析、生成报表、做图等。该模块实现对文件的处理
SAS/EM	实现数据挖掘的功能。在这个模块里实现了 SAS 的 SEMMA 数据挖掘模式,同时还有可以直接应用的数据挖掘模型
SAS/STAT	统计分析应用模块,是统计学学习者应用的模块,主要运用在统计领域,如回归分析、相关分析、因子分析、聚类分析和判别分析等分析功能
SAS/EG	报表开发模块,实现 OLAP 分析,多维模型的报表开发就需要调用此模块

调用模块的方法是运行 SAS 9.2 工具菜单栏上的“解决方案”,在弹出的子菜单项中选择“分析(S)”,即可选择调用上面叙述的某些模块,如图 1-1 所示。

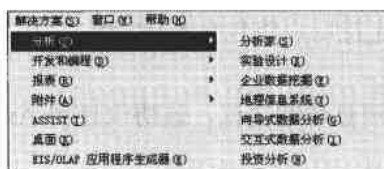


图 1-1 “分析”菜单

1.2.2 SAS 处理流程

为了便于对 SAS 处理流程有个清晰的认识,可参考如图 1-2 所示的 SAS 处理流程图中从左边到右边的顺序进行处理数据,首先 SAS 对外围数据文件或 SAS 表(数据集)通过数据步(DATA STEP)进行数据处理,生成 SAS 数据集,然后过程步(PROC STEP)根据业务需求调用 SAS 不同的模块来对数据集进行分析处理,实现满足业务需求的分析。

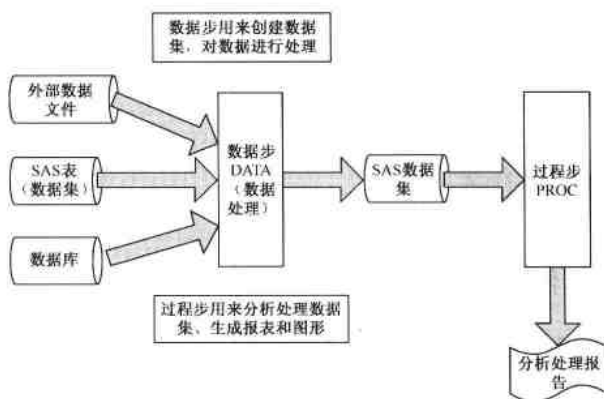


图 1-2 SAS 处理流程

1.3 SAS 9.2 新特性

SAS 9.2 新特性是 SAS 的新增功能,和 SAS 以前的版本比较, SAS 9.2 最大的改进是增加了自定义函数,这是 SAS 的一个大跳跃。SAS 9.2 以前的版本只能调用 SAS 内部已经编译好的函数,用户不能自定义函数。

1.3.1 SAS 9.2 新增 Base SAS 过程

SAS 9.2 在 Base SAS 里新增了 FCMP (Function Compiler Procedure) 过程,即 SAS 函数编译器过程,通过这个过程可以根据需求自定义函数,来实现特定功能。可以用 C、C++或 SAS 语言自行编写函数,然后在 SAS 程序中调用它们。新增的 FUNCTIONS DICTIONARY (函数类库)表存储了 SAS 所有函数的信息。在第 2 章函数精讲中会单独介绍这个自定义函数。

① 新增 JAVAINFO 过程, JAVAINFO 过程将有关 SAS 正在使用的 Java 环境的诊断信息报告给用户,通过这个可以确认是否配置了正确的 SAS Java 环境。

② 新增 PROTO 过程,支持以批处理模式注册使用 C 或 C++程序设计语言编写的外部函数。通过在 PROC PROTO 中注册 C 语言函数后,可以从 FCMP 过程中声明的任何 SAS 函数或子例程中调用它们,也可以从 COMPILE 过程中声明的任何 SAS 函数、子例程或方法块中调用它们。

③ 新增 SCAPROC 过程, 支持用户将包含 SAS 代码分析器的输出的文件或文件引用名, 将输出写入到该文件。

④ 新增 SOAP 过程, 支持消息传递协议, 该协议使用 Axis2 Java 客户端通过 Java 本机接口调用 Web 服务。

1.3.2 SAS 9.2 ACCESS 新增功能

SAS 9.2 ACCESS 模块与 PC 文件模块的接口提供对导入和导出向导、PROC IMPORT 过程及 PROC EXPORT 过程提供全新的和增强的 PC 文件特定支持。针对“PC 文件服务器”新增高级视图和安装选项。支持以下新的文件格式: 带有 memo 的 dBase、JMP、Paradox DB、SPSS SAV 和 Stata DTA。

SAS 在不同版本文件格式读入 PROC IMPORT 和 LIBNAME 引擎使用不同格式的 Microsoft Access 或 Excel 文件中的日期/时间值, 现在可以通过 USEDATE=指定为 LIBNAME 选项或 PROC IMPORT 中的可选语句改为所需的格式。

SAS 9.2 ACCESS 与关系型数据库的接口新增内容如下:

- AUTHDOMAIN=LIBNAME 选项。
- DBIDIRECTEXEC=系统选项。
- SASTRACE=系统选项的简要跟踪功能。

1.3.3 SAS 9.2 其他模块新增功能

SAS 9.2 connect 新增针对 OLE DB 的 SAS 提供程序, 提供对本地 Base SAS 系统的访问。

SAS 9.2 IntrNet 在编程和配置方面新增应用程序调度器, 提供后台处理、针对 64 位主机和基于 IPv6 的主机支持。

SAS 9.2 QC 在 CAPABILITY 过程新增的 CLASS 语句支持用户将数据分组为不同的分类水平, 可以指定一个或两个分类变量。

SAS 9.2 QC 在 FACTEX 过程新增 BLOCK 和 UNITEFFECT 语句, 支持为多个阶段的实验构造设计。SAS 9.2 SHARE 在 PROC SERVER 语句中新增或改进了多个选项。

1.4 SAS 9.2 安装步骤与启动

SAS 9.2 安装步骤帮助并指导安装 SAS 9.2, 在安装过程的每一步都配有详细的安装说明和注意事项, 要严格按照安装 SAS 9.2 的步骤执行。

1.4.1 SAS 9.2 安装步骤

SAS 9.2 在 Windows 下的安装步骤如下:

(1) 要安装虚拟光区 Daemon Tools 或 UltraISO_v9.32, 并把所有的杀毒软件和防火墙软件停掉。把系统时间修改到 SAS 9.2 ISO 所授权的有效期内, 如果不改系统时间, SAS 9.2 安装就会失败。

(2) 准备 SAS 9.2 安装文件, 这个笔者用 UltraISO_v9.32 解压了, 可以看到 setup.exe 安装文件, 如图 1-3 所示。

(3) 双击 setup.exe, SAS 9.2 安装开始启动, 如图 1-4 所示。

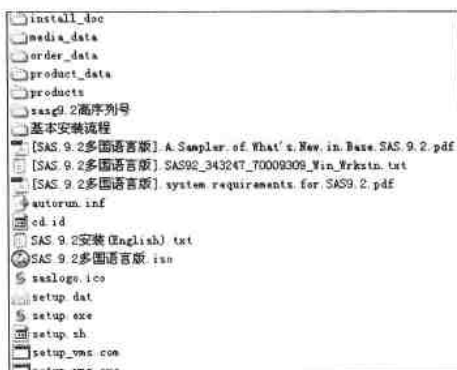


图 1-3 SAS 安装文件截图



图 1-4 SAS 安装启动窗口

(4) 在出现的“选择语言”窗口，要选择所使用的语言，从下拉列表框中选择“简体中文”，如图 1-5 所示。

(5) 出现安装向导界面，如图 1-6 所示。

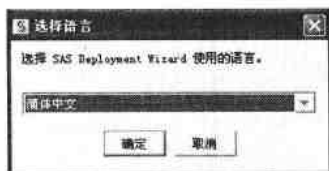


图 1-5 选择语言窗口

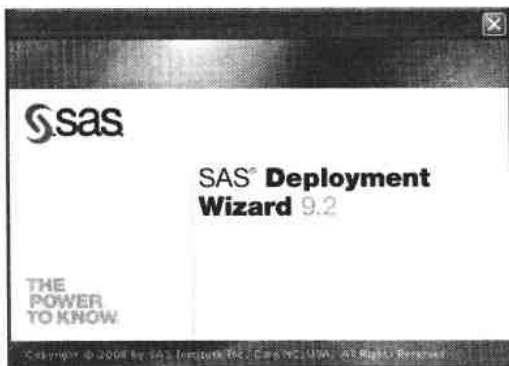


图 1-6 SAS 安装向导界面

(6) 选择部署任务，此处个人单机用户选择“安装 SAS 软件”，如图 1-7 所示。

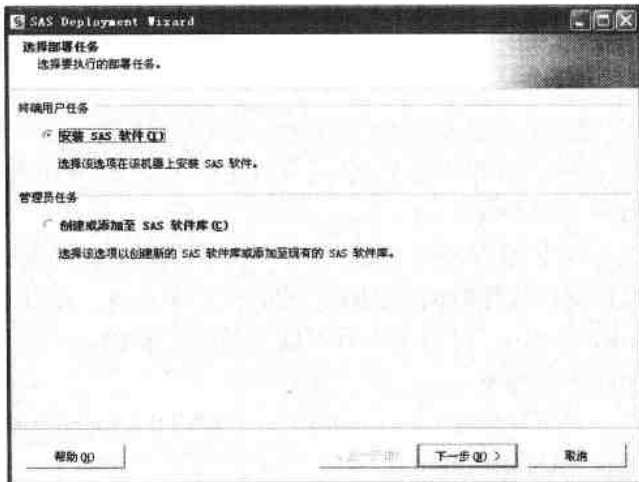


图 1-7 选择要执行的部署任务窗口

(7) 选择本机器要安装 SAS 组件产品，这里用默认选项，如图 1-8 所示。

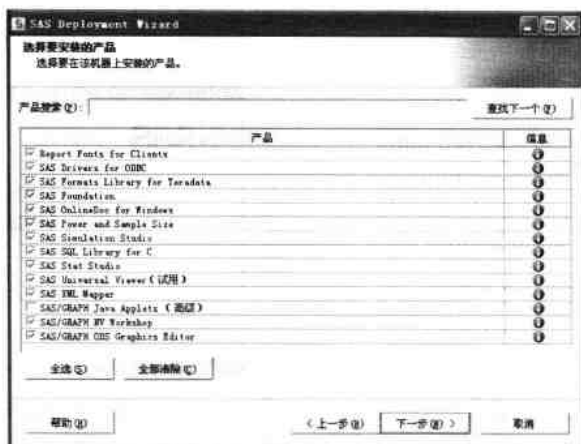


图 1-8 选择要安装的产品窗口

(8) 单击“下一步”按钮，出现选择 SAS 9.2 工具对应的安装数据 (SID) 文件窗口，从浏览界面中选择本机 SAS 9.2 目录下的文件 “[SAS.9.2 多国语言版].SAS92_343247_70009309_Win_Wrkstn.txt”，这一步很关键，如图 1-9 所示。



图 1-9 选择指定 SAS 安装数据 (SID) 文件窗口

(9) 单击“下一步”按钮，出现部署汇总窗口，如图 1-10 所示。

(10) 单击“开始”按钮，开始安装 SAS 9.2 的组件，SAS 9.2 即按所示的顺序安装组件，直到出现完成为止，如图 1-11 所示。

【提示】 SAS 9.2 安装后修改时间问题。SAS 9.2 安装好后，如果有最新的 SAS 9.2 sid 文件，按下面的步骤操作就可以把时间改到现在的时间。调用时一定要看 SAS 9.2 sid 文件授权给您的时间是到什么时候为止，过了这个时间就启动不了 SAS。

安装好后调整时间的步骤如下：

(1) 单击“开始”→“所有程序”→“SAS”→“SAS 9.2 License Renewal & Utilities”。

(2) 单击右键，更新 SAS 软件许可。

(3) 选择运行方式，然后再导入上述 SAS 9.2 sid 文件（下载的最新的 SAS 9.2 sid 文件）即可。

(4) 再把时间改回正常就可以了。



图 1-10 SAS 部署汇总窗口



图 1-11 SAS 正在安装组件窗口

1.4.2 SAS 9.2 启动与运行

SAS 9.2 的启动为单击“开始”→“所有程序”→“SAS”→“SAS 9.2 (简体中文)”，即可启动和运行 SAS 9.2，如图 1-12 所示。也可以把 SAS 9.2 (简体中文) 发送到桌面快捷方式启动。



图 1-12 SAS 9.2 的启动

运行 SAS 9.2 后，即启动了 SAS 工具，如图 1-13 所示。

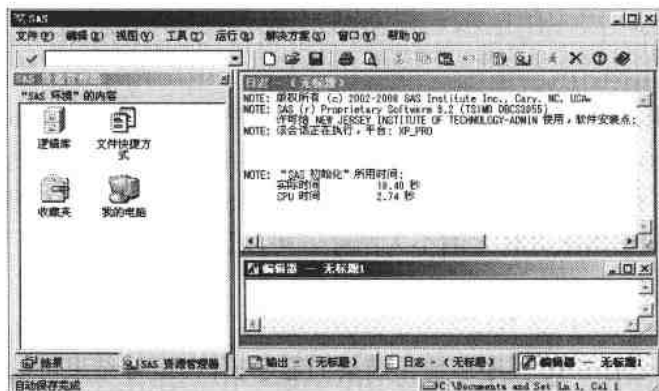


图 1-13 SAS 操作界面窗口

1.5 SAS 开发工具与操作界面

SAS 9.2 开发工具操作界面包括菜单栏、工具栏、编辑器、日志窗口、输出窗口和资源管理器。

1.5.1 SAS 9.2 操作界面

SAS 9.2 开发工具操作界面显示了 SAS 的图形界面, 左边显示的是 SAS 资源管理器窗口, 此窗口显示了逻辑库、文件快捷方式、收藏夹和我的电脑等资源; 上面是菜单栏, 通过单击菜单栏上的按钮可以弹出菜单项对应的子菜单; 命令窗口可以输入命令, 运行 SAS 相关的模块; 编辑窗口是编辑程序的编辑器, 可以在编辑窗口书写 SAS 程序; 日志窗口是对编辑窗口运行 SAS 程序后显示程序运行后的信息, 如果程序有错误可以在日志窗口看到相关的错误信息, 根据日志信息就可以修改和调试编辑窗口编写的 SAS 程序; 输出窗口显示程序运行后的输出结果, 如果运行 SAS 编辑器后有输出, 输出窗口就会显示输出结果。SAS 操作界面各功能窗口如图 1-14 所示。

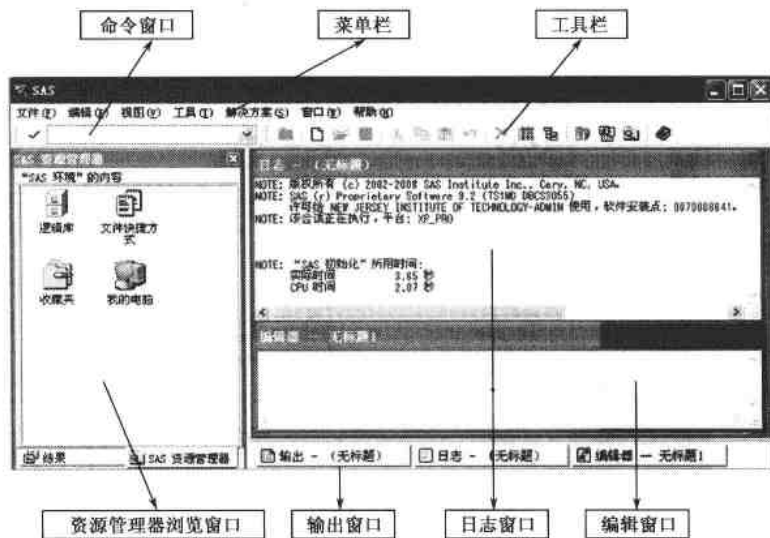


图 1-14 SAS 操作界面各功能窗口

1.5.2 编辑器

SAS 编辑器是用来书写 SAS 程序和调试 SAS 程序的编辑窗口, 通过运行和编译来调试程序, 如图 1-15 所示。编译好的 SAS 程序可以选择菜单栏上的“文件”菜单, 在弹出的下拉菜单中选择“保存”子菜单, 起好名字, 扩展名为“.sas”, 选择好存储位置就可以保存 SAS 程序。按快捷键 F8 或单击工具栏上的小人图标就可以运行 SAS 程序。

1.5.3 输出

SAS 输出窗口是运行 SAS 程序后输出结果的显示窗口。运行结果可以在输出窗口中看到, 如图 1-16 所示。