



配套光盘

视频讲解, PPT、SAS程序等素材文件

SAS

统计分析

刘荣 编著

与应用实例



NLIC2970943003



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

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

SAS 统计分析与应用实例

刘荣 编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

SAS (Statistical Analysis System, 统计分析系统) 作为当今国际最著名的数据分析软件系统, 功能强大、应用广泛。

本书基于 SAS 9.2 软件, 结合编程和菜单操作以实例为载体介绍 SAS 系统。主要内容有: SAS 系统简介、SAS 编程简介、SAS 菜单操作、定量数据描述性统计分析、参数估计与假设检验、方差分析、相关与回归分析、列联表分析、非参数检验、主成分分析与因子分析、典型相关分析、聚类分析、判别分析。

本书内容丰富、结构清晰, 采取从原理到实践的方式。随书附赠的光盘中收录了本书所有例题、实验、上机练习的数据与程序, 并含长达 8 小时的全程实录讲解视频。

本书的读者对象是高等院校各专业学习应用统计的大学生、研究生, 以及企事业单位数据分析工作者。

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 统计分析与应用实例/刘荣编著. —北京: 电子工业出版社, 2013.10
ISBN 978-7-121-21537-7

I. ①S… II. ①刘… III. ①统计分析—应用软件 IV. ①C819

中国版本图书馆 CIP 数据核字 (2013) 第 224599 号

策划编辑: 陈韦凯

责任编辑: 王凌燕

印 刷: 三河市鑫金马印装有限公司

装 订: 三河市鑫金马印装有限公司

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: 787×1092 1/16 印张: 17.75 字数: 454.4 千字

印 次: 2013 年 10 月第 1 次印刷

印 数: 4 000 册 定价: 49.00 元 (含 DVD 光盘 1 张)

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 zltz@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

SAS (Statistical Analysis System, 统计分析系统) 作为当今国际最著名的数据分析软件系统, 被广泛应用于金融、保险、医疗卫生、教育和科研等各行业与领域。

本书的编排遵循从理论到实践的原则, 首先介绍基本统计原理, 然后结合实例应用 SAS9.2 的编程和菜单操作的方式完成各项统计分析, 并结合例题实际背景对分析结果进行详细阐述。

以下介绍本书的基本结构: 本书共 13 章, 前 3 章为基础知识介绍, 后 10 章为统计分析实验。

第 1 章为 SAS 系统简介, 介绍了 SAS 系统的应用范围、主要模块、界面操作、文件管理及 SAS 逻辑库与数据集的部分操作。

第 2 章为 SAS 编程简介, 主要介绍了 SAS 数据步中的输入、赋值、控制语句, SAS 程序步中的 VAR、BY 等语句, 并引入 ODS 输出系统和 SAS 宏。

第 3 章为 SAS 菜单操作, 引入 SAS/ASSIST、SAS/INSIGHT 和 SAS/Analyst 三个菜单操作模块。

第 4 章为定量数据描述性统计分析, 介绍了连续型数据描述性统计分析方法, 以及统计图形的绘制。

第 5 章为参数估计与假设检验, 介绍了单样本均值和方差的区间估计、独立和配对样本 T 检验及正态分布拟合检验。

第 6 章为方差分析, 主要介绍单因素、多因素方差分析 (包括区组设计、析因设计和拉丁方设计) 和协方差分析。

第 7 章为相关与回归分析, 介绍了相关分析、线性回归和非线性回归。

第 8 章为列联表分析, 介绍了列联表的编制及其拟合优度检验、独立性检验、一致性检验、趋势检验和计算属性关联度。

第 9 章为非参数检验, 介绍了单样本、两独立 (配对) 样本和多个独立样本的位置检验。

第 10 章到第 13 章为多元回归分析部分, 主要包括主成分与因子分析、典型相关分析、判别与聚类分析。

本书内容丰富、层次清晰。在程序中添加了详细的注释; 菜单操作过程介绍清晰。随书附赠的光盘中分章收录了本书所有的例题、实验、上机练习的数据和程序; 并包含 8 小时随书全程实录, 力求让读者形象、直观地掌握 SAS 系统的一般统计分析方法。

本书的读者对象是高等院校各专业学习应用统计的大学生、研究生, 以及企事业单位数据分析工作者。

本书主要由刘荣编写, 同时, 参与本书编写工作的还有张玉兰、孙明、唐伟、王杨、顾辉、李成、陈杰、张霁芬、张计、陈军、张强、杨明、李建、李兵等人。

因作者水平有限, 书中错误、纰漏之处难免, 欢迎广大读者批评指正。

编 著 者

目 录

第 1 章 SAS 系统简介	(1)
1.1 SAS 系统概述.....	(1)
1.2 SAS 界面操作与文件管理.....	(2)
1.2.1 SAS 系统的启动与退出.....	(2)
1.2.2 SAS 系统界面简介.....	(2)
1.2.3 SAS 数据集和逻辑库.....	(4)
1.2.4 数据集文件操作.....	(6)
1.3 SAS 数据集整理.....	(6)
1.3.1 新建数据集.....	(6)
1.3.2 在数据集中增加、筛选变量和观测.....	(9)
1.3.3 对数据集排序.....	(10)
1.3.4 数据集纵向连接.....	(11)
1.3.5 数据集横向合并.....	(13)
1.3.6 数据集转置.....	(14)
1.3.7 数据集的导入与导出.....	(14)
1.4 SAS 帮助系统介绍.....	(17)
第 2 章 SAS 编程简介	(21)
2.1 SAS 程序简介.....	(21)
2.1.1 SAS 程序构成.....	(21)
2.1.2 SAS 程序基本规定.....	(21)
2.2 数据步中基本语言介绍.....	(21)
2.2.1 INPUT 语句.....	(22)
2.2.2 赋值语句.....	(23)
2.2.3 循环语句.....	(25)
2.2.4 分支结构.....	(27)
2.3 过程步中基本语句介绍.....	(28)
2.3.1 VAR、MODEL、BY、CLASS 语句.....	(29)
2.3.2 WHERE、FREQ、WEIGHT 语句.....	(30)
2.3.3 使用 OUTPUT、FORMAT、LABEL、ID 语句.....	(30)
2.4 SAS 函数.....	(31)
2.5 ODS 输出系统.....	(35)
2.6 SAS 宏简介.....	(36)
2.6.1 SAS 宏变量.....	(36)
2.6.2 创建和调用宏.....	(37)
练习题.....	(38)

第3章 SAS 菜单操作	(40)
3.1 SAS/ASSIST 视窗介绍.....	(40)
3.1.1 SAS/ASSIST 概述.....	(40)
3.1.2 SAS 实例——分性别描述某班学生英语成绩分布.....	(41)
3.2 SAS/INSIGHT 交互分析介绍.....	(43)
3.2.1 SAS/INSIGHT 概述.....	(43)
3.2.2 SAS 实例——绘制身高和体重的散点图.....	(44)
3.3 Analyst (分析家) 模块操作.....	(45)
3.3.1 Analyst 模块概述.....	(45)
3.3.2 应用 Analyst 整理数据.....	(46)
3.3.3 应用 Analyst 进行统计分析.....	(51)
3.3.4 SAS 实例——探索年龄和血压的相关关系.....	(52)
练习题.....	(53)
第4章 定量数据描述性统计分析	(55)
4.1 描述性统计分析指标.....	(55)
4.1.1 基本指标介绍.....	(55)
4.1.2 SAS 过程——MEANS 过程.....	(56)
4.1.3 SAS 过程——UNIVARIATE 过程.....	(58)
4.1.4 SAS 实例——描述小麦单穗粒数分布.....	(59)
4.2 描述性统计图形.....	(63)
4.2.1 常见统计图形介绍.....	(63)
4.2.2 SAS 过程——GPLOT 过程.....	(64)
4.2.3 SAS 过程——GCHART 过程.....	(66)
4.2.4 SAS 实例——绘制年龄和血压的散点图.....	(67)
4.2.5 SAS 实例——绘制某班学生成绩分布的直方图.....	(68)
4.2.6 SAS 实例——绘制国内生产总值的折线图.....	(69)
4.2.7 SAS 实例——绘制 2009 年 GDP 构成的饼图.....	(70)
4.2.8 SAS 实例——绘制某种玉米株高的条形图.....	(72)
练习题.....	(73)
第5章 参数估计与假设检验	(77)
5.1 TTEST 过程.....	(77)
5.2 基本的参数区间估计.....	(78)
5.2.1 总体均值的区间估计.....	(78)
5.2.2 总体方差的区间估计.....	(79)
5.2.3 SAS 实例——求均值和方差的 95%置信区间.....	(79)
5.3 基本假设检验.....	(82)
5.3.1 t 检验.....	(82)
5.3.2 两样本方差齐性检验.....	(84)
5.3.3 正态分布检验.....	(84)
5.3.4 SAS 实例——检验水稻单株产量是否为特定值.....	(85)

5.3.5	SAS 实例——比较不同方法的减肥效果	(87)
5.3.6	SAS 实例——检验某新药疗效是否显著	(89)
5.3.7	SAS 实例——检验射击环数是否服从正态分布	(91)
	练习题	(93)
第 6 章	方差分析	(96)
6.1	SAS 过程——ANOVA 过程	(96)
6.2	SAS 过程——GLM 过程	(98)
6.3	单因素方差分析	(100)
6.3.1	基本原理	(100)
6.3.2	SAS 实例——2009 年不同地区商品房销售差异分析	(104)
6.4	区组设计方差分析	(108)
6.4.1	基本原理	(108)
6.4.2	SAS 实例——检测不同化学试剂对布匹强度影响的差异性	(109)
6.5	拉丁方设计方差分析	(112)
6.5.1	基本原理	(112)
6.5.2	SAS 实例——研究不同电视组装方法的组装时间的差异性	(113)
6.6	析因设计方差分析	(116)
6.6.1	基本原理	(116)
6.6.2	SAS 实例——研究温度和压强对某化学物品产率的影响	(117)
6.7	协方差分析	(121)
6.7.1	基本原理	(121)
6.7.2	SAS 实例——比较不同化肥对桃子的产量的影响	(122)
	练习题	(126)
第 7 章	相关与回归分析	(129)
7.1	相关分析	(129)
7.1.1	基本原理	(129)
7.1.2	SAS 过程——CORR 过程	(130)
7.1.3	SAS 实例——考察航空公司航班正点率和顾客投诉次数的关系	(131)
7.2	直线回归	(133)
7.2.1	基本原理	(133)
7.2.2	SAS 过程——REG 过程	(137)
7.2.3	SAS 实例——考察沸点和气压的关系	(139)
7.2.4	SAS 实例——多元回归模型预测房屋售价	(146)
7.3	非线性回归	(153)
7.3.1	基本原理	(153)
7.3.2	SAS 过程——NLIN 过程	(153)
7.3.3	SAS 实例——拟合某微生物生长曲线	(154)
7.3.4	SAS 实例——非线性回归函数的参数估计	(162)
7.4	LOGISTIC 回归	(164)
7.4.1	基本原理	(164)

7.4.2	SAS 过程——LOGSTIC 过程	(165)
7.4.3	SAS 实例——结石病危险因素研究	(166)
	练习题	(171)
第 8 章	列联表分析	(174)
8.1	SAS 过程——FREQ 过程	(174)
8.2	拟合优度检验	(177)
8.2.1	基本原理	(177)
8.2.2	SAS 实例——检验各年龄阶层人口数是否满足特定分布	(178)
8.3	独立性检验	(179)
8.3.1	基本原理	(179)
8.3.2	SAS 实例——居住地与驾车类型关系探索	(179)
8.4	一致性检验	(182)
8.4.1	基本原理	(182)
8.4.2	SAS 实例——HR 对求职者评定等级的一致性研究	(182)
8.5	属性关联度	(185)
8.5.1	基本原理	(185)
8.5.2	SAS 实例——探索某原料的产地与质量等级的关系	(185)
	练习题	(187)
第 9 章	非参数检验	(189)
9.1	SAS 过程——NPAR1WAY 过程	(189)
9.2	单样本位置检验	(190)
9.2.1	基本原理	(190)
9.2.2	SAS 实例——检验某工地施工是否提高小区噪声水平	(191)
9.3	Wilcoxon 符号秩检验	(192)
9.3.1	基本原理	(192)
9.3.2	SAS 实例——情绪对血压值的影响	(193)
9.4	Wilcoxon 秩和检验	(194)
9.4.1	基本原理	(194)
9.4.2	SAS 实例——检验两地地表土壤的 pH 值的差异	(195)
9.5	Kruskal-Wallis 秩和检验	(197)
9.5.1	基本原理	(197)
9.5.2	SAS 实例——探索不同专业学生英语成绩差异	(198)
	练习题	(200)
第 10 章	主成分分析与因子分析	(202)
10.1	主成分分析	(202)
10.1.1	基本原理	(202)
10.1.2	SAS 过程——PRINCOMP 过程	(203)
10.1.3	SAS 实例——客户信誉的“5C”评级分析	(204)
10.2	因子分析	(210)
10.2.1	基本原理	(210)
10.2.2	SAS 过程——FACTOR 过程	(213)

10.2.3 SAS 实例——我国各省市发展情况分析	(215)
练习题	(224)
第 11 章 典型相关分析	(226)
11.1 基本原理	(226)
11.2 SAS 过程——CANCORR 过程	(228)
11.3 SAS 实例——生理指标和训练指标的相关分析	(230)
练习题	(237)
第 12 章 聚类分析	(238)
12.1 基本原理	(238)
12.1.1 样品(变量)间距离定义	(238)
12.1.2 类的性质	(239)
12.1.3 聚类方法	(240)
12.2 样品聚类	(242)
12.2.1 SAS 过程——CLUSTER 过程	(242)
12.2.2 SAS 过程——TREE 过程	(243)
12.2.3 SAS 实例——根据飞行距离对 10 所美国城市分类	(245)
12.3 变量聚类	(247)
12.3.1 SAS 过程——VARCLUSE 过程	(247)
12.3.2 SAS 实例——对 8 个身体素质指标进行聚类	(249)
练习题	(254)
第 13 章 判别分析	(256)
13.1 基本原理	(256)
13.1.1 距离判别分析法	(256)
13.1.2 Fisher 线性函数判别法	(259)
13.2 SAS 过程——DISCRIM 过程	(260)
13.3 SAS 实例——根据物质含量判断食物所属类别	(262)
练习题	(271)

第1章 SAS 系统简介

SAS (Statistical Analysis System) 系统于 1976 年由 SAS 软件研究所 (SAS Institute Inc.) 研制推出。历经多年发展, 最新版本 SAS9.3 于 2011 年 7 月问世, 它作为国际公认的著名数据统计分析软件系统之一, 受到许多国家和地区的机构青睐。本章将简介 SAS 系统主要模块、界面操作和文件管理, 并在引入 SAS 逻辑库和数据库后介绍新建逻辑库、数据集的方法, 及数据集的排序、连接、合并及与外部数据文件相互转换等操作的实现, 在章末将重点介绍 SAS 帮助系统的使用, “授人以鱼, 不如授人以渔”, 希望读者能举一反三, 将 SAS 软件灵活运用于学习和工作实践中。

1.1 SAS 系统概述

SAS 系统将数据管理和统计分析融为一体, 能够在不同的操作系统 (如 UNIX、MS-DOS、VMS 等) 和不同的机器类型下运行, 系统具备完备的数据存取、数据管理、数据分析和数据展示的功能。在 Windows 版的 SAS 运行环境下, 用户不仅能够以灵活的编程方式, 还可选择操作简单的菜单方式进行各种统计分析。目前, SAS 系统被广泛应用于金融、医疗卫生、生产、运输、通信、科研和教育等领域。它运用统计分析、时间序列分析、运筹决策等科学的方法进行质量控制、财务管理、生产优化、风险管理、市场调查和预测等业务, 并可将各种数据以灵活的报表、图形和三维透视的形式直观地表现出来。

SAS 系统包含众多模块以完成不同任务, 本书内容涉及的有:

- SAS/BASE (基础) ——完成数据整理和初步统计分析。
- SAS/STAT (统计) ——广泛的统计分析。
- SAS/ASSIST (面向任务的通用菜单驱动界面) ——交互式菜单操作。
- SAS/GRAPH (图形) ——提供了许多产生图形过程并支持众多图形设备。

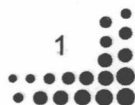
SAS/BASE 是 SAS 系统的核心平台, 提供了多个 SAS 过程, 可以实现简单报表计算、生成报表、计算得分 (标准化数据、数据排秩等) 及排序等功能。

SAS/STAT 提供 SAS 系统用来解决实际问题的具体统计分析过程, 主要包括方差分析、回归分析、属性数据分析、多变量分析、聚类分析、判别分析等。

SAS/GRAPH 具有强大的绘图功能, 能够用于绘制二维曲线图、条形图、饼图、区域图等。

本书将基于 Win7 系统下的 SAS9.2 多国语言版的 SAS (English (ODBS)) 版本 (该版本系统语言为英文, 但支持一般的中文字符输入), 首先概述 SAS 系统、简介 SAS 编程和菜单操作, 然后介绍具体统计方法基本原理及其在 SAS 系统中的实现。

注意: 建议读者先快速浏览通读本书, 对介绍的基本内容形成大致印象, 在遇到实际问题时再具体参考相应章节。若时间充裕, 也可以通过观看随书附赠的光盘中的教学视频练习本书



介绍的案例来熟练掌握 SAS 软件。

1.2 SAS 界面操作与文件管理

下面介绍 SAS 系统的启动和退出、界面特性及文件管理，完成以上操作前请用户将 SAS 软件安装在一台满足 SAS 系统配置的计算机中。

1.2.1 SAS 系统的启动与退出

SAS 系统的启动：

- 在系统的“开始”菜单列表中找到 SAS 系统文件夹，左键单击“SAS9.2 (Additional Languages)”文件夹下的项目“SAS 9.2 (English (DBCS))”即可启动系统，如图 1-1 所示。

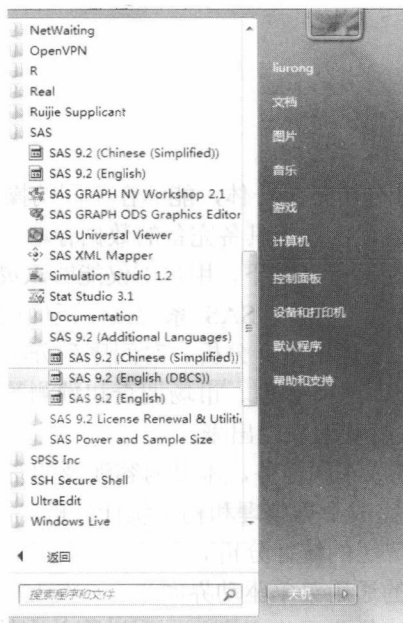



图 1-1 启动 SAS 系统

- 若右键单击项目“SAS 9.2 (English (DBCS))”，在下拉选项中选择“发送到”|“桌面快捷方式”，即可双击桌面上的快捷图标启动 SAS 系统。

SAS 系统的退出：

- 选择菜单 File|Exit，在弹出的确认对话框中左键单击“确定”按钮。
- 单击系统主界面右上角按钮，在弹出的确认对话框中单击“确定”按钮。

1.2.2 SAS 系统界面简介

启动 SAS 系统将出现如图 1-2 所示的工作界面。它在一个主窗口内包含有若干个子窗口，




并有菜单栏、工具栏、状态栏等。下面介绍 SAS 系统界面的主要窗口：Editor 程序编辑窗口、Log 运行记录窗口、Output 输出记录窗口、Explorer 窗口和 Results 窗口。

- Editor 程序编辑窗口——主要用于编辑 SAS 源程序文件，操作时光标可在整个窗口随意移动，且支持 Windows 系统常规编辑操作方式，如剪切、复制、粘贴等。SAS9.2 的智能编辑功能可根据用户输入的不同的 SAS 程序部分显示出不同的颜色。若用户输入有误，对应的颜色不对，以警告错误的发生。
- Log 运行记录窗口——用于显示和记录 SAS 程序的运行情况，说明其运行成功或提示错误信息。当程序运行不成功，Log 运行记录窗口将分别用绿色字符和红色字符显示警告和错误信息。
- Output 输出记录窗口——分页显示 SAS 程序运行的文本型输出结果，可使用主界面菜单的 File|Save As 将其保存在磁盘中，文件扩展名为.lst，该类型文件可用文字处理软件如 Word、写字板、记事本等将其打开和编辑。SAS 程序运行的图形输出结果将由 Graphics 窗口显示，可选择菜单 File|Export as Image 将图形导出并保存在磁盘内，并可在“保存类型”下拉列表中选择图形文件的保存格式。
- Explorer 窗口——用于显示 SAS 逻辑库（SAS 系统命名的库名和磁盘某文件的关联）及 SAS 数据集。
- Results 窗口——用于显示 SAS 成功运行时程序输出结果的目录。

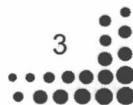


图 1-2 SAS 系统工作界面

可用左键单击窗口内的任意一处的方式切换到以上任一窗口。在使用时可根据需要直接单击窗口右上角的按钮  关闭窗口，也可使用主菜单 View 的下拉菜单打开或切换窗口。

SAS 主窗口标题栏下是主菜单，注意 SAS 菜单随着光标出现在不同的窗口是动态变化的。其主要命令及功能实现如下：

- File（文件）——实现 SAS 文件的调入、保存、转换及打印等功能。
- Edit（编辑）——实现窗口的编辑（清空、剪切、复制、粘贴等）功能。
- View（浏览）——实现打开或切换到 SAS 的各个工作窗口的功能。





- Tools (工具)——提供对各种输出结果进行编辑的工具,如表格、图形、报告等,并支持进行系统环境和状态的设置(如修改界面字体等)。
- Run (运行)——提交程序,仅在当前窗口为 Editor 程序编辑窗口时有效。
- Solutions (解决方案)——SAS 图形界面模块操作窗口,如 SAS/ASSIST。
- Help (帮助)——提供 SAS 软件自带的帮助系统。

主菜单下是一个命令输入栏和图表工具栏。在命令栏中可以输入 SAS 的显示管理命令,如在此输入“WSAVE”则表示永久保存对 SAS 系统的设置。工具栏图标提供了常见任务的快捷操作方式,其功能解释如下:

- □ New——建立新的编辑窗口。
- Open——打开文件到编辑窗口。注意:用户指定一个文件调入到编辑窗口内,以后的存盘操作将自动存入此文件。
- Save——保存编辑窗口内容。注意:若此窗口已经与一个文件相联系,存盘功能将覆盖文件原有内容。
- Print——打印当前窗口内容。
- Print preview——打印预览。
- Cut——剪切选定文本。
- Copy——复制选定文本。
- Paster——粘贴。

注意:这些操作是对 Windows 剪贴板进行的,所以它不仅支持 SAS 编辑窗口内的复制和剪切操作,还可用来与其他 Windows 应用程序交换文本、数据等。

- Undo——撤销编辑操作。
- New Library——建立新的 SAS 逻辑库。
- SAS Explorer——打开 SAS 管理窗口查看、管理 SAS 的各个逻辑库及其中的文件。
- Submit——提交 Editor 程序编辑窗口中的程序。
- Clear All——清空当前窗口内容。
- Break——中断正在运行的 SAS 程序。
- Help——进入 SAS 的帮助界面。

1.2.3 SAS 数据集和逻辑库

SAS 文件主要包括数据集(Database)文件、索引文件和 SAS 目录文件(Catalog)等。数据集是 SAS 使用和分析计算的原始数据来源,而正确合理地生成 SAS 数据集是数据分析的首要条件,因此数据集是 SAS 最重要的文件类型。SAS 目录文件主要用以保存各种不能表示成行列结构表格形式的数据,如系统设置、图像、声音等。

SAS 数据集可以看作由若干行和列组成的表格,数据集的每一行称为一个观测(Observation),每一列称为一个变量(Variable),变量可以取不同的类型值,如整数型、浮点值、时间值、字符串、货币值等。

如图 1-3 所示的数据集范例中包括了 3 条观测,代表了 3 个客户的情况;包含 5 个变量,分别为客户编号(ID)、姓名(name)、持卡类型(Type)、年龄(birth)和消费次数(N)。注意到该图中数据集名称为 Chap1.Example1, ID、name、Type、birth 和 N 为变量名。在 SAS 系统

中使用的数据集、变量名、逻辑库名等统称为“标识符”，SAS 系统对标识符有以下严格规定：

- SAS 标识符必须由英文字母、数字、下画线组成。
- 第一个字符必须是字母或下画线。
- 标识符中字母不区分大小写。
- 标识符的长度不宜过长。

ID	name	Type	birth	N
1	张三	金卡	10/06/64	20
2	李四	银卡	09/16/82	13
3	王五	普通卡	01/21/87	5

图 1-3 数据集范例

由此可知 name、area、ABC、X2、_Nall_ 等都是合法标识符，area 和 AREA 为同一标识符，但 number-3（不能有减号）、a bit（不能有空格）、team*（不能有特殊字符）等却不合法。

SAS 数据集存储在被称为 SAS 逻辑库（Library）的文件集中。通俗地说，SAS 逻辑库是一个连接，将磁盘中存储的文件和 SAS 系统联系起来。SAS 逻辑库命名遵循上述 SAS 命名规则，可用编程和菜单操作的方式建立逻辑库。

编程建立逻辑库：使用 Libname 命令可以指定逻辑库，命名格式为：

Libname 逻辑库标记"文件夹路径";

例如，要建立指向已存在的文件目录“E:\data\chap1”的逻辑库 chap1，可在 Editor 程序编辑窗口输入以下语句：

Libname chap1 'E:\data\chap1';

再选择 Run|Submit 菜单或左键单击工具栏上的 * 图标提交程序完成操作。

菜单方式建立逻辑库：此方式操作过程如下：进入 Explorer 窗口，双击 Library 图标，再单击右键，选择 new 命令，在弹出的窗口（如图 1-4 所示）的 Name 栏中输入逻辑库名，在 Path 栏中输入路径或单击右侧的 Browse（浏览）按钮选择磁盘中的文件夹，选中右侧的 Enable at startup（启动时可用）使其永久有效，最后单击 OK 按钮完成。

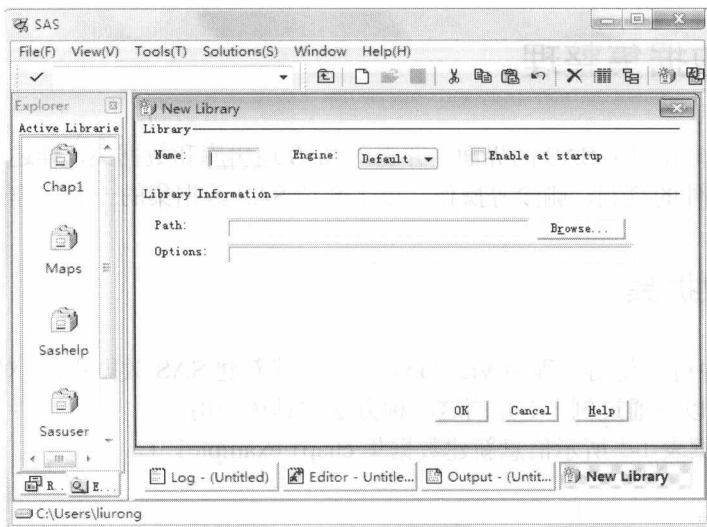


图 1-4 新建逻辑库

以上介绍了新建逻辑库的一般方法，接下来概述两种逻辑库类型：临时库和永久库，以及其对应的临时数据集和永久数据集的命名区别。

临时库和永久库：临时库仅有 WORK 逻辑库，它存储 SAS 临时文件，在每次启动 SAS 系统时自动生成，关闭 SAS 系统时库中的数据集被自动删除；永久库中存储 SAS 永久文件，SAS 预定义 Sasuser 和 Sashelp 两个永久逻辑库，其中 Sasuser 用以保存与个人设置有关的文件，Sashelp 逻辑库保存与 SAS 帮助系统、应用实例有关的文件。除此之外，用户可使用上述介绍的编程和菜单操作的方式建立 SAS 永久库。

临时数据集和永久数据集命名区别：临时数据集可以用“单水平名”，即只有数据集名，如 stu01，这种名字的数据集被保存在 WORK 逻辑库中，因此也可用 WORK.stu01 表示。永久数据集由两部分组成，前一部分为它的库名，后一部分为数据集名，中间用小数点连接，如放在 chap1 库（即“E:\data\chap1”子目录）中的数据集 ex1 需要用 chap1.ex1 表示。临时数据集在重启 SAS 系统时将会被自动删除，而永久数据集不会。

1.2.4 数据集文件操作

SAS 管理器（SAS Explorer）用来管理 SAS 逻辑库和文件，SAS 系统默认启动时在主界面左侧显示 SAS 管理器，若不慎将其关闭，可通过主菜单命令 View|Explorer 将其打开。以下介绍对逻辑库中数据文件的复制、删除、改名等操作。

- 文件复制：不同逻辑库之间的数据文件可以复制备份，操作为：左键双击逻辑库名，单击选择需要复制的数据文件，右键单击，在弹出的快捷菜单中选择 Copy（复制）命令实现复制，再左键双击打开目标逻辑库名，右键单击，在弹出的快捷菜单中选择 Paster（粘贴）命令完成备份。
- 文件删除：选择目标数据文件，右键单击，在弹出的快捷菜单中选择 Delete（删除）命令。
- 文件重命名：选择目标数据文件，右键单击，在弹出的快捷菜单中选择 Rename（更名）命令，再在弹出的对话框中输入改后的文件名，左键单击 OK 按钮保存设置并退出。

1.3 SAS 数据集整理

本章 1.2 节简单介绍了 SAS 的界面，并引入 SAS 逻辑库和数据集，详述了新建逻辑库的方法，以及数据集文件的复制、删除等操作。以下介绍 SAS 数据集的一般操作。

1.3.1 新建数据集

在 SAS 系统中可以使用编程和 Viewtable 表的方式新建 SAS 数据集，并对数据集进行修改、增删记录等操作。以下通过例 1-1 介绍这两种方法的具体使用。

例 1-1 请根据表 1-1 所示信息新建数据集 chap1.example1_1。

表 1-1 某公司客户信息表

编号 (ID)	姓名 (name)	持卡类型 (Type)	出生年月 (birth)	消费次数 (N)	登记时间 (Date)
0101	张三	金卡	1964/10/06	20	2011/01/02
0102	李四	银卡	1982/09/16	13	2011/04/03
0103	王五	消费卡	1987/01/27	5	2011/05/06


方法一：编写如下程序（其在光盘中的存储路径为“proc\chap1\example1_1.sas”）。

```
libname chap1 'E:\data\chap1'; /*新建指向路径为“E:\data\chap1”的逻辑库 chap1*/
data chap1.example1_1; /*新建永久 SAS 数据集 chap1.example1*/
input ID $1-4 name $5-11 Type$13-19 birth YYMMDD8. +1 N; /*定义变量输入格式*/
cards;
0101 张三 金卡 64/10/06 20
0102 李四 银卡 82/09/16 13
0103 王五 普通卡 87/01/27 5
; /*输入数据*/
proc print;
format birth YYMMDD8.; /*设置变量 birth 的输出格式*/
run;
```

选择 Run|Submit 命令提交程序，在 Output 输出记录窗口显示如图 1-5 所示，且在新建的逻辑库 chap1 中出现了数据集 example1。

Obs	ID	name	Type	birth	N
1	0101	张三	金卡	64-10-06	20
2	0102	李四	银卡	82-09-16	13
3	0103	王五	普通卡	87-01-27	5

图 1-5 结果输出

注意：日期格式的变量在 SAS 中默认保存为从 1960 年 1 月 1 日至某日期的天数，如 1964 年 10 月 6 号，默认输出为 1740。在打印输出日期格式的变量时，为了得到特定格式，务必使用 format 语句指定输出格式。而 SAS 数据集 chap1.example1，变量 birth 也将显示天数，此时右键单击变量名，在快捷菜单中选择 Column Attribute（变量属性），单击弹出的对话框中的 format 后的  按钮，在此可以选择设置变量的输出格式。

方法二：Viewtable 表新建数据集。

步骤一：打开新表

选择菜单 Tools|Table Editor 打开一个新表，如图 1-6 所示。

步骤二：定义变量

右键单击 A 列，选择 Column Attributes（变量属性）命令，弹出如图 1-7 所示对话框，在 Name 栏中输入变量名 ID，Type（类型）为默认 Character（字符型），Length（长度）为 8，Format（输出格式）和 Informat（输入格式）为 \$8。（默认设置）。左键单击 Apply（应用）按钮，再单击 Close（关闭）按钮；或者直接单击 Close 按钮完成第一列姓名变量的属性设置。

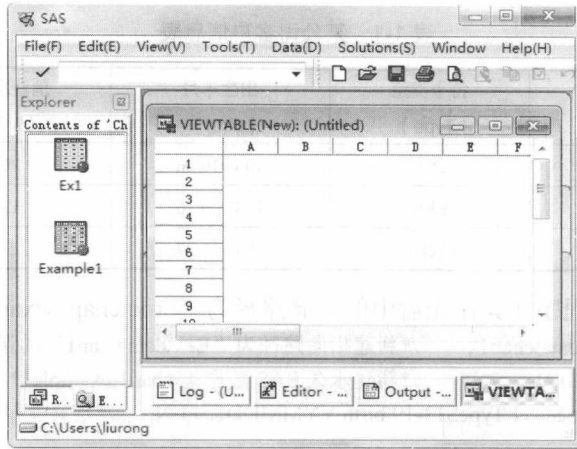


图 1-6 Viewtable 表视图

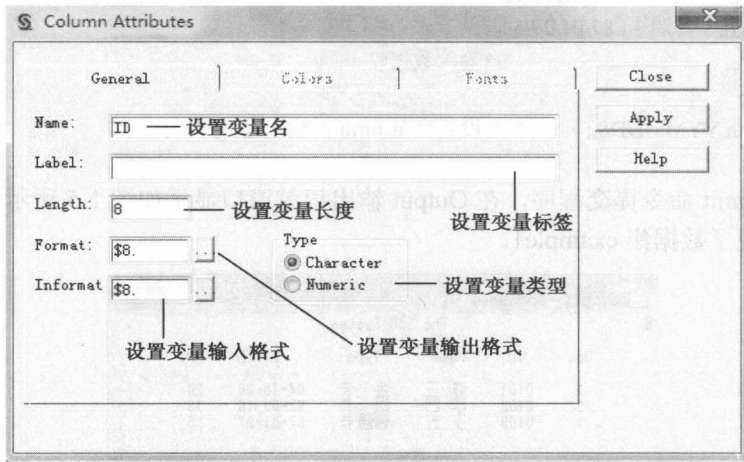



图 1-7 设置变量属性

重复以上操作，定义变量 Name、Type 时变量类型 (Type) 选择 Character，输出格式 (Format) 和输入格式 (Informat) 选择默认值 \$8.；定义变量 birth 时变量类型 (Type) 选择 Numeric，Format (输出格式) 和 Informat (输入格式) 选择 yy/mm/dd；定义变量 N 时变量类型 (Type) 选择 Numeric，Format (输出格式) 和 Informat (输入格式) 选择默认值 \$8.。

说明：用户根据变量性质选择设置变量类型 (Type) 为字符型 (Character) 和数值型 (Numeric)。Informat 和 Format 分别代表数据的输入格式和输出格式，可单击格式右侧的“...”按钮设置。在窗口的 Label 栏中可以输入变量标签。若不定义变量标签，则系统默认变量名为标签名。打开数据集时，Viewtable 表头为标签名，可通过主菜单 View|Columns Labels 和 View|Columns Names 进行切换。

步骤三：输入数据、保存数据集

直接在表中输入相应数据，然后直接单击工具栏上的保存图标，或者选择菜单 File|Save，在弹出的对话框中左键单击选择需要存入的库名 chap1，在 Member Name (对象名称) 框中输入数据集名 example1，单击 Save (保存) 按钮保存设置。

若要修改已有数据集的变量标签，可选择菜单 Editor|Edit Mode (编辑模式) 更改系统默认