

# SAS

## 统计分析系统

使用手册

SAS/PC系统入门  
SAS 语言

北京大学概率统计系

高惠璇 张平 栾世武 编

北京大学新技术公司

一九九〇年六月

## 前 言

SAS 系统是国际上享有盛名,非常流行的统计分析软件。我们系因承接国家经济信息中心的“七五”项目中的子项目“国外经济预测软件的比较研究”的课题,自 1988 年 3 月开始,对 SAS 软件系统进行开发研究,组织了部分教员,研究生参加这项工作,还举办了两届高年级部分学生参加的 SAS 系统开发讨论班。我们认为 SAS 系统是一个功能齐全,使用灵活方便的统计分析系统。值得把它介绍给国内的用户。

为了加快 SAS 系统的推广使用,让它为我们国内更多用户服务,我们根据《SAS 入门指南》和《SAS 语言指南》(6.03 版)编写了这本《SAS 统计分析系统使用手册》。本书适用于 SAS 系统的初学者及 Base SAS 软件的一般用户上机时使用。附录部分简单介绍了 Base SAS 软件的基本过程和 SAS/STAT 软件的功能。详细的使用手册在条件成熟时,我们也准备继续编写出版。

本手册的第十六章,第十七章及第二十章由张平同志编写;第十章及第十三章 § 8 由栾世武同志编写;其余部分由高惠璇同志编写。限于我们的水平,加上对 SAS 系统使用时间较短、对该系统的了解不是很深透,书中难免出现错误或不足之处,敬请用户批评指正。

本手册由北京大学新技术公司培训部用北大华光轻印刷系统排版,我们对他们的辛勤劳动表示衷心感谢。

编者

一九九〇年五月

# 目 录

第一部分 SAS/PC 系统入门 .....	(1)
第一章 SAS 系统概况 .....	(1)
§ 1 SAS 系统的特点 .....	(1)
§ 2 Base SAS 软件 .....	(2)
§ 3 SAS 系统的其它产品 .....	(3)
§ 4 SAS 系统的资料 .....	(4)
§ 5 安装 SAS/PC 系统对环境的要求 .....	(5)
第二章 SAS/PC 系统会话实例 .....	(6)
§ 1 启动 SAS 系统 .....	(6)
§ 2 例子 .....	(7)
§ 3 输入 SAS 语句 .....	(8)
§ 4 提交 SAS 程序 .....	(9)
§ 5 分析数据 .....	(10)
§ 6 结束 SAS 会话 .....	(11)
第三章 SAS 数据集 .....	(13)
§ 1 把外部数据文件转换为 SAS 数据集 .....	(13)
§ 2 创建永久的 SAS 数据集 .....	(15)
§ 3 数据整理 .....	(16)
§ 4 整理 SAS 数据集的数据 .....	(18)
第四章 SAS 过程 .....	(19)
§ 1 引言 .....	(19)
§ 2 使用 SAS 过程 .....	(20)
§ 3 使用 PROC SORT 来重新排列数据 .....	(22)
§ 4 使用 PROC PRINT 来显示数据 .....	(23)
§ 5 使用 PROC FREQ 来产生频率表 .....	(27)
§ 6 使用 PROC MEANS 来概括数据 .....	(30)
第五章 SAS/PC 系统的运行环境 .....	(33)
§ 1 SAS 显示管理系统 .....	(33)
§ 2 使用 SAS 过程的 MENU 系统 .....	(37)
§ 3 SAS 程序的其它运行环境 .....	(39)
第二部分 SAS 语言 .....	(41)
第六章 使用 SAS 语言 .....	(41)
§ 1 SAS 语句 .....	(41)
§ 2 SAS 程序 .....	(42)

§ 3	SAS 数据集 .....	(43)
§ 4	编写 SAS 程序 .....	(44)
§ 5	SAS 程序的输出 .....	(45)
第七章	SAS 表达式 .....	(48)
§ 1	引言 .....	(48)
§ 2	SAS 常数 .....	(48)
§ 3	SAS 算符 .....	(49)
§ 4	特殊常数和算符 .....	(53)
§ 5	两个特殊的专题 .....	(54)
第八章	SAS 函数 .....	(56)
§ 1	引言 .....	(56)
§ 2	SAS 函数的自变量和结果 .....	(57)
§ 3	SAS 函数的分类 .....	(58)
§ 4	概率函数 .....	(62)
§ 5	分位数函数 .....	(65)
§ 6	样本统计函数 .....	(67)
§ 7	随机数函数 .....	(70)
第九章	DATA 步入门 .....	(76)
§ 1	引言 .....	(76)
§ 2	什么是 DATA 步 .....	(76)
§ 3	DATA 步语句及其类型 .....	(78)
§ 4	DATA 步流程 .....	(80)
§ 5	例子 .....	(82)
第十章	用在 DATA 步的文件操作语句 .....	(84)
§ 1	DATA 语句 .....	(84)
§ 2	INPUT 语句 .....	(86)
§ 3	CARDS 与 CARDS4 语句 .....	(94)
§ 4	PUT 语句 .....	(95)
§ 5	BY 语句 .....	(97)
§ 6	SET 语句 .....	(99)
§ 7	MERGE 语句 .....	(102)
§ 8	UPDATE 语句 .....	(104)
§ 9	FILE 语句 .....	(105)
§ 10	INFILE 语句 .....	(108)
第十一章	用在 DATA 步的运行语句 .....	(110)
§ 1	赋值语句和累加语句 .....	(110)
§ 2	DELETE 语句和 LOSTCARD 语句 .....	(112)
§ 3	STOP 语句和 ABORT 语句 .....	(115)
§ 4	WHERE 语句 .....	(116)

§ 5	OUTPUT 语句	(120)
§ 6	其它运行语句	(122)
第十二章	用在 DATA 步的控制语句	(127)
§ 1	DO 语句	(127)
§ 2	END 语句	(131)
§ 3	SELECT 语句	(131)
§ 4	IF 语句	(133)
§ 5	GOTO 语句	(135)
§ 6	LINK 语句	(136)
§ 7	RETURN 语句	(138)
第十三章	用在 DATA 步的信息语句	(140)
§ 1	ARRAY 语句	(140)
§ 2	INFORMAT 语句和 FORMAT 语句	(147)
§ 3	LENGTH 语句和 LABEL 语句	(150)
§ 4	ATTRIB 语句	(153)
§ 5	语句标号	(154)
§ 6	DROP 语句和 KEEP 语句	(155)
§ 7	RENAME 语句和 RETAIN 语句	(157)
§ 8	WINDOW 语句	(160)
第十四章	PROC 步及用在 PROC 步的语句	(167)
§ 1	PROC 步入门	(167)
§ 2	PROC 语句	(168)
§ 3	VAR 语句和 MODEL 语句	(169)
§ 4	WEIGHT 语句和 FREQ 语句	(169)
§ 5	ID 语句和 WHERE 语句	(170)
§ 6	CLASS 语句和 BY 语句	(171)
§ 7	OUTPUT 语句和 QUIT 语句	(173)
§ 8	FORMAT 语句和 LABEL 语句	(174)
§ 9	ATTRIB 语句	(175)
第十五章	全程语句(用在任何地方的 SAS 语句)	(177)
§ 1	引言	(177)
§ 2	注释语句	(177)
§ 3	DM 语句和 X 语句	(178)
§ 4	TITLE 语句和 FOOTNOTE 语句	(180)
§ 5	RUN 语句和 ENDSAS 语句	(181)
§ 6	LIBNAME 语句	(182)
§ 7	FILENAME 语句	(184)
§ 8	%INCLUDE 语句	(186)
§ 9	OPTIONS 语句	(188)

第十六章 SAS 显示管理系统 .....	(191)
§ 1 引言 .....	(191)
§ 2 显示管理全程命令 .....	(191)
§ 3 SAS 文本编辑 .....	(196)
§ 4 三个基本窗口 .....	(201)
§ 5 HELP 窗口和 KEYS 窗口 .....	(203)
§ 6 MENU 窗口 .....	(204)
§ 7 其它窗口 .....	(206)
§ 8 特殊专题 .....	(212)
第十七章 SAS 输出和错误的处理 .....	(214)
§ 1 SAS 记录 .....	(214)
§ 2 输出 SAS 过程产生的结果 .....	(216)
§ 3 在 LOG 窗口显示的几类错误 .....	(218)
§ 4 错误的处理 .....	(222)
第十八章 SAS 输入格式和格式及缺失值 .....	(224)
§ 1 引言 .....	(224)
§ 2 使用 SAS 输入格式 .....	(225)
§ 3 使用 SAS 格式 .....	(228)
§ 4 使用 SAS 日期、时间和日期时间的输入格式和格式 .....	(232)
§ 5 缺失值 .....	(236)
第十九章 SAS 文件 .....	(241)
§ 1 SAS 文件的类型 .....	(241)
§ 2 规定 SAS 文件的名字 .....	(241)
§ 3 SAS 数据库 .....	(244)
§ 4 SAS 数据集 .....	(245)
§ 5 SAS 目录 .....	(250)
第二十章 SAS 全局选择项 .....	(254)
§ 1 引言 .....	(254)
§ 2 配置选择项 .....	(254)
§ 3 SAS 系统选择项 .....	(257)
第三部分 附录 .....	(261)
附录一 Base SAS 软件的基本 SAS 过程 .....	(261)
附录二 SAS/STAT 软件简介 .....	(264)

# 第一部分 SAS/PC 系统入门

## 第一章 SAS 系统概况

SAS (statistical Analysis System)系统是数据分析的软件系统。它于 1966 年开始研制, 1976 年由美国 SAS 公司实现商业化。1985 年推出了 SAS/PC 版本。自 SAS 系统推出以来, 它的版本更新很快,功能也不断增加。如自 1987 年推出 6.03 版本后,目前又推出最新版本 6.06。SAS 是国际上最著名的数据分析软件系统之一。

### § 1 SAS 系统的特点

#### (一)什么是 SAS 系统

SAS 系统是数据分析和生成报表的一个软件系统。所谓软件系统就是一组在一起作业的计算机程序。使用 SAS 系统用户可以根据自己的需要,随时作出明智的选择。不论你从事简单还是复杂的工作,不论你是 SAS 系统的初学者还是老用户,SAS 系统都能满足你的需要。

SAS 系统是一种组合软件系统。基本部分是 Base SAS 软件(基本 SAS 软件),此外还有用于统计分析计算的 SAS/STAT 软件;用于绘图的 SAS/GRAPH 软件;用于矩阵运算的 SAS/LML 软件;用于运筹学和线性规划的 SAS/OR 软件;用于经济预测和时间序列分析的 SAS/ETS 软件等等。SAS 系统的核心是包括了超过 125 个灵活应用的过程。可以现成地使用,或稍加修改,以便满足用户的特殊需要。用户只需要学会 SAS 系统,就可以替代你学习其它程序设计语言或其它统计程序包乃至程序。SAS 系统可成为你唯一需要的软件系统。

#### (二)SAS 系统的特点

SAS 系统的最大特点是把数据管理和统计分析融为一体。具体地说,它具有如下几方面的特点:

##### (1)操作简单

SAS 的宗旨是为所有需要进行数据处理、统计分析的非计算机专业人员提供一种易学易用,完整可靠的软件系统。由于 SAS 系统采用两个基本步骤(DATA 步用来整理数据, PROC 步用来分析数据和编写报告)作为任何 SAS 程序的基本组件,因此 SAS 系统不仅效能极高,且操作非常简单。

SAS 系统提供一种交互地显示管理方式来运行 SAS 程序。SAS 显示管理系统是用户和 SAS 系统的一个很友好的界面,它不仅让用户看到自己输入的内容,而且可以看到输出的信息——包括错误信息和输出结果。

##### (2)SAS 软件系统的基础是 SAS 语言

SAS 语言是一种近乎自然英语的非过程语言(或称第四代语言)。用户把要解决的问题,

用 SAS 语言表达出来,组成 SAS 程序,提交给 SAS 系统就可以方便地解决你提出的问题。SAS 语言的特点是用户不必告诉 SAS“怎样做”,只需告诉它你要“做什么”就行了。

用户只要熟悉这些用英文单词作基本指令(关键词),如 DATA, INPUT, PUT, KEEP, SELECT, RUN 等组成的 SAS 语句,就能容易地掌握 SAS 系统的操作:用 DATA 步整理数据,调用 SAS 过程分析处理数据等。

SAS 语言是 SAS 系统的基础。在这个基础上,用户可以独运匠心,按自己的需要将 SAS 系统改制成为适合你的专门工具。

### (3) 数据处理与统计分析融为一体

SAS 系统克服了通常软件或偏重于数据管理功能或偏重于统计分析的弊病,把数据管理功能与统计分析功能有机地结合在一起,它不仅具有一整套从数据输入、加工处理直至文件操作,打印输出等完备灵活的数据管理功能,而且还能对所存贮的数据进行各种统计分析,它包括 125 个过程(模块),提供丰富完备的统计分析算法。

### (4) 扩展性能强——没有上限的软件系统

SAS 系统的模块式结构可以让用户灵活地按自己的需要去选择所需功能的模块。随着你需要的增加和改变,你可以把各种功能的软件产品添加到 Base SAS 软件中。SAS 系统还设置一个过程扩充接口,用 SAS 的特定语句把新过程添加到 SAS 过程库中,成为整个系统的一个功能。不论你的工作变得多么复杂,SAS 系统都可以伴你一起工作。

### (5) 适用性强

SAS 系统可以在各种不同的操作系统下运行。(如 PC-DOS/MS-DOS, UNIX, VMS 等);可以在大型计算机,小型计算机及微机上运行。它可以在几十种不同的计算机系列上运行。就 SAS/PC 系统而言,适用的机器类型也很多。如 IBM-PC/XT/AT, 显示器用 CGA/EGA/单显; IBM-PC/XT/AT, 3270; IBM-PS/2, 显示器用 VGA; WANG, 单显/彩显; Compaq AT, 彩显; HP Vectra, 彩显; 及采用 ANSL SYS 的非 IBM 兼容机。

### (6) 应用面广

SAS 系统可广泛地应用于自然科学和社会科学的各个领域。如统计学、心理学、经济学、生物学、考古学、医疗卫生、商业金融、社会调查等。适当地组合 SAS 系统的模块,可用于:图形显示、图形分析、数据输入、数据检索、数据管理、报表生成、统计计算、工程计算、市场研究、销售管理、供应管理、订货追踪、调查分析、建立预测模型、部门信息管理等方面。

## § 2 Base SAS 软件

SAS 系统的基本部分——Base SAS 软件提供以下几方面功能:

### (一) 信息存贮和检索

SAS 系统可用任何格式读入数据值,然后把数据组成 SAS 数据集。它具有很强的与外部文件交换信息的功能,可以用文件操作管理方法把不同数据集的数据组合在一起,供 SAS 过程分析处理,生成报表。SAS 数据集用临时或永久数据集两种形式存贮,它具有自动生成文档的能力,即既包含数据值,又含有它们的描述信息。利用 SAS 数据库的特殊结构可使维



护工作降到最低的限度。

## (二)数据修改与程序设计

SAS 系统为用户提供完备的 SAS 语句和函数用于数据加工处理;有些语句用于执行标准操作,如建立新变量,累加求和及修改错误;另外还提供了 DO/END,IF-THEN/ELSE 等控制语句,构成一套完整的语言系统。由于 Base SAS 软件具有非常强的数据管理功能,所以许多人亦把基本 SAS 软件作为数据库管理系统。

## (三)报表生成

基本 SAS 软件输出数据与读入数据时一样可采用任何格式。除了 SAS 过程产生的固定格式的报表外,用户根据自己的需要设计报表的输出格式,包括输出文件到磁盘上。

## (四)描述统计计算

Base SAS 软件中有一些基本过程能够完成基本的统计计算。它包括:

- (1)计算简单的描述统计量,如均值、标准差、极差、相关系数等。
- (2)绘制水平的或垂直的直方图,饼图和散布图等。
- (3)生成并分析列联表;
- (4)对数据求秩和标准化。

## (五)文件操作

数据分析常常需要从几个数据集中抽取一些数据和观测进行组合。Base SAS 软件中有用于对数据进行编辑、调整、连结、合并及更新的功能,还能同时处理多个输入文件,而且对一次输入的数据能够生成几种报表等。

# § 3 SAS 系统的其它产品

在基本 SAS 软件的基础上,你可以附加上统计分析、绘图、预测、运筹学、质量控制、交互式矩阵运算等软件来构成一个总系统。SAS 的最大优点是用基本 SAS 软件和提供各种功能的 SAS 软件产品组合成总系统来满足用户的种种需要。

对于 SAS 系统 6.03 版,Base SAS 软件之外的其它产品有以下几种:

## (一)SAS/STAT 软件

它是一个高度可靠完整的统计分析软件。包括回归分析、方差分析、属性数据分析,多变量分析、判别分析、聚类分析、得分方法和残存分析八类方法共 26 个过程。每个过程还提供多种不同的算法及选择,从而组成一个庞大而完整的统计分析方法集。

Base SAS 和 SAS/STAT 是 SAS 系统的核心和精华,也是 SAS 系统用来解决实际问题的主要部分。

## (二) SAS/GRAPH 软件

绘图软件能够完成多种绘图功能,如生成等值线图、二维和三维曲线图、条形图、圆形图、区块图、星形图及各种映象图。非常形象、直观地表现各变量之间的关系及数据的分布状态,对解决各种实际问题起着重要的辅助作用。

SAS/GRAPH 软件的适用性强,对外部设备的要求少,可把图形输出到显示器、打印机、绘图仪、磁盘等外部设备上。

## (三) 其它软件

### (1) SAS/IML 软件

该软件是一个交互式的程序设计语言(交互式矩阵语言),它处理的基本数据元素是一个矩阵;它允许用户直接用矩阵代数的记号作为 IML 的程序语句。

### (2) SAS/FSP 软件

主要用来进行文字处理的一个交互式地菜单系统,可用来输入数据,编辑 SAS 文件和书写其它文本文件。

### (3) SAS/AF 软件

该软件是一个交互式地全屏幕软件应用系统。

### (4) SAS/ETS 软件

用于经济分析、预测和商务计划的专用软件。

### (5) SAS/OR 软件

用于运筹学和工程管理的专用软件。

### (6) SAS/QC 软件

用于质量控制的软件。

SAS 系统的 Base SAS 软件是 SAS 系统的基本软件,其它的软件产品均以 Base SAS 软件为基础,但它们相互之间是独立的,对不同需要的用户可以有选择地装入某种软件,当需要其它软件产品时,再装入所需的软件,这一点对用户是十分方便的。

## § 4 SAS 系统的资料

SAS 系统提供非常完整的文本资料。Base SAS 软件和其它软件产品(6.03 版本)的资料有:

SAS 入门指南 6.03 版

SAS 语言指南 6.03 版

SAS 过程指南 6.03 版

SAS/STAT 用户指南 6.03 版

SAS/GRAPH 用户指南 6.03 版

SAS/IML 用户指南 6.03 版

SAS/AF 用户指南 6.03 版

SAS/FSP 用户指南 6.03 版  
关于宏过程的 SAS 指南 第 6 版  
关于表过程的 SAS 指南 1987 年版  
等等。

这本使用手册是根据 SAS 入门指南和 SAS 语言指南两本资料编写的。附录部分参阅了其它有关的资料。

## § 5 安装 SAS/PC 系统对环境的要求

安装和使用 SAS/PC 系统有以下要求：

### (一) 内存

- (1) 安装 SAS/PC 系统, 内存空间不能少于 640K, 其中自由空间不少于 525K。
- (2) 使用显示管理系统运行基本 SAS 软件, 要求有 332K 的自由空间; 使用行方式运行基本 SAS 系统时, 你必须有 220K 的自由内存空间。

### (二) 磁盘空间

为了安装 SAS 系统, 你的计算机硬盘上必须有足够没有使用的空间。各个软件产品要求的存贮空间大约为

基本 SAS 软件—5.28MB  
SAS/STAT 软件—4.09MB  
SAS/GRAPH 软件—4.31MB  
SAS/IML 软件—0.69MB  
SAS/FSP 软件—0.80MB  
SAS/AF 软件—0.68MB

### (三) 操作系统

你必须有 PC-DOS 或 MS-DOS 的 2.0 版或 2.0 以后的版本。

### (四) 系统时钟

你的计算机时钟必须置为对安装和使用 SAS 软件两者都是正确的日期和时间。

### (五) 警告

- (1) 你安装其它产品之前必须安装基本 SAS 软件。
- (2) 不能使用同基本 SAS 软件不同版本的其它软件产品。例如 6.03 版的 SAS/STAT 软件不能同 6.02 版的基本 SAS 软件一起使用。

## 第二章 SAS/PC 系统会话实例

为了让用户对 SAS 系统的工作方式和环境有一个具体的认识,本章通过一个实例的介绍,让用户尽快熟悉 SAS 系统。

下面我们约定带有下划线的部分表示用户输入的命令,其它则为系统的响应信息。符号 '↵' 表示回车键(ENTER)。

### § 1 启动 SAS 系统

假定 SAS 系统安装在子目录 SAS 下。启动 SAS/PC 系统时键入如下命令:

C>CD SAS↵

C>SAS↵(或 C:\SAS\ \SAS)↵)

执行这一命令后,屏幕将显示 SAS 系统的三个基本窗口,即 OUTPUT 窗口,LOG 窗口和 PGM 窗口(PROGRAM EDITOR)。见屏幕 2.1

```
----- OUTPUT -----
command ==>

----- LOG -----
command ==>

NOTE:Copyright(c) 1985,86,87 SAS Institute Inc.,Cary,NC 27512-8000,U.S.A.
NOTE:SAS (r) Proprietary Software Release 6.03
      Licensed to SAS Institute 6.03,Site 000000000.

----- PROGRAM EDITOR -----
command ==>

00001
00002
00003
```

屏幕 2.1 SAS 显示管理系统屏幕

上面的 OUTPUT 窗口存放 SAS 过程处理数据后的输出结果;中间的 LOG 窗口的内容为 SAS 系统的运行记录;下面的 PGM 窗口是用户编辑 SAS 源程序,存贮或调入文件,提交 SAS 程序给 SAS 系统执行的场所,它是用户使用 SAS 系统的主要窗口。

屏幕出现这三个彩色的窗口后,光标位于 PGM 窗口的命令行,此时光标可以自由移动。在 LOG 窗口还显示关于 SAS 系统的版本等信息。至此 SAS 系统已启动完毕,等待用户输入程序和命令。

SAS 系统用来分析处理数据。它首先利用 DATA 步整理数据,把你要分析的数据组织成 SAS 系统能够读的灵活、容易操控的文件—SAS 数据集。然后利用 PROC 步根据用户的需要选择 SAS 系统提供的 SAS 过程来分析处理数据并编写报告。

SAS 系统要求用户把你想做什么通过一些近乎自然英语的指令—SAS 程序在 PGM 窗口采用会话方式提交给 SAS 系统执行。执行过程中的记录信息(包括出错信息)在 LOG 窗口显示;SAS 过程产生的结果在 OUTPUT 窗口输出。这就是 SAS 会话系统。

## § 2 例子

假定某公司生产两类机器:小型的机器(如打印机、复印机)和计算机。这些机器由公司的职员负责在四个地区销售。表 2.1 给出的数据包括销售人员的名字(name),销售的地区(region),机器的类型(type),和年销售量(sales)。利用这些资料,你想知道每个地区的年销售量是多少,每种类型的机器卖了多少?及每类机器的平均销售量等问题。

表 2.1 销售数据

序号	NAME	SALES	REGION	TYPE
1	stafer	9664	east	sm
2	young	22969	east	sm
3	stride	27253	esat	sm
4	topin	86432	east	c
5	spark	99210	east	c
6	vetter	38928	west	c
7	curci	21531	wect	sm
8	marco	79345	west	c
9	greco	18523	west	sm
10	ryan	32915	west	sm
11	tomas	42109	west	sm
12	thalman	94320	south	c
13	moore	25718	south	sm
14	allen	64700	south	c
15	steiam	27634	south	sm
16	farlow	32719	north	sm
17	smith	38712	north	sm
18	wilson	97214	north	c

表 2.1 中的每一项—如名字 'stafer', 地区 'east', 销售总量 22969, 类型 'c' (表示计算机)或 'sm' (表示小型机器)都是数据值。

表 2.1 的每一行给出职员的名字,年销售量,销售的地区和卖出的机器类型。每一行构成一个观测(observation)。

表 2.1 中每一项的信息形成一列;每列的值构成一个变量(variable)。例如这些职员的名字构成名字变量;所有年销售量构成销售变量等等。为了描述这些数据给 SAS 系统,对每个变量规定名字。比如称这些职员的名字为 name,年销售量为 sales,销售地区为 region,卖出

的机器类型为 type。每个变量的名字必须以字母开始,由 1 到 8 个字符组成。

表 2.1 的所有销售信息组成一个数据集。它有十八个观测和四个变量。

为了分析处理这些资料,用户要把这些数据读到 SAS 系统中形成 SAS 数据集,然后调用有关的 SAS 过程分析处理它们。

### § 3 输入 SAS 语句

把表 2.1 的销售数据变为 SAS 数据集的最常用的方法是在键盘上直接输入数据。

启动 SAS 系统后光标所在的 PGM 窗口是输入 SAS 语句和数据的窗口。首先在窗口的命令行上使用 'ZOOM' (放大) 命令让 PGM 窗口充满整个终端的屏幕:

```
command ==> ZOOM
```

再按 '↵' 把光标从命令行移到 PGM 窗口的第一行。然后开始插入 SAS 语句。SAS 语句是要求 SAS 系统执行一种操作或给 SAS 系统以信息的指令。在这个例子中,为了读入表 2.1 的数据并形成 SAS 数据集,在 PGM 窗口输入的 SAS 语句如屏幕 2.2 所示:

```
PROGRAM EDITOR
Command ==>
00001  data sale;
00002  input name $ sales region $ type $;
00003  cards;
00004  stafer      9664  east  sm
00005  young      22969 east  sm
00006  stride     27253 east  sm
00007  topin      86432 east  c
00008  spark      99210 east  c
00009  vetter     38928 west  c
00010  curoi      21531 west  sm
00011  maroo      79345 west  c
00012  greco      18523 west  sm
00013  ryan       32915 west  sm
00014  tomas      42109 west  sm
00015  thalman    94320 south c
00016  moore      25718 south sm
00017  allen      64700 south c
00018  steiam     27634 south sm
00019  farlow     32719 north sm
00020  smith      38712 north sm
00021  wilson     97214 north c
00022  run;
```

ZOOM

屏幕 2.2 在 PGM 窗口输入的 SAS 语句和数据行

下面我们摘要介绍每个 SAS 语句的作用:

(1) data sale; (或 DATA SALE;)

DATA 语句告诉 SAS 系统开始创建一个名字为 sale 的 SAS 数据集。DATA 语句是数据步的开始。

(2)input name \$ sales region \$ type \$ ;

INPUT 语句告诉 SAS 系统数据值是如何排列在数据行上及变量的名字是什么。名字 name,region 和 type 后面的 dollar 号 (\$)告诉 SAS 这三个变量的值包含字符。

(3)cards;

CARDS 语句告诉 SAS 下面就是数据行。

(4)data lines(数据行)

从 CARDS 语句后的第 4 行到第 21 行是数据行,每一行是一个观测,数值间用空格隔开。

(5)run;

RUN 语句告诉 SAS 系统执行上述这些 SAS 语句。它标志在这个 SAS 会话中 DATA 步的结束。

屏幕 2.2 的这些 SAS 语句就构成一个 SAS 程序。

## § 4 提交 SAS 程序

当你认为这些 SAS 语句和数据都已正确地输入时,你准备提交这个 SAS 程序给 SAS 系统以便创建 SAS 数据集。这时在 PGM 窗口的命令行上打入提交命令:

command===>SUBMIT✓

这时 PGM 窗口上的所有语句都消失了,在窗口右下角边上出现字母'R',表示正在运行这个 DATA 步。

当这个作业运行完成后,光标又出现在 PGM 窗口的命令行上。为了浏览运行的记录,打入命令:

command===>LOG✓

把 LOG 作为当前窗口(或活动窗口)。屏幕 2.3 列出这个 SAS 程序的运行记录。

SAS 记录告诉用户你的数据集有十八个观测和四个变量。这十八个观测对应十八个职员的信息;四个变量是由 INPUT 语句读入的这些变量。运行记录还告诉用户 SAS 系统处理这个 DATA 步用了多少时间。

如果 LOG 窗口显示你输入的 SAS 语句有错误时,你必须返回到 PGM 窗口,重新调进刚才输入的 SAS 程序对错误的语句作修改。即在命令行上打入命令:

command===>PGM;RECALL✓

修改完毕后再重新提交 SAS 程序。SAS 系统执行这个 SAS 程序后,产生一个名为 sale 的 SAS 数据集。

```
LOG
Command == =>
      Dr. Reeves for use on total of three machines.
NOTE: AUTOEXEC processing completed.
      1  data sale;
      2  input name $ sales region $ type $ ;
      3  cards;
      22 run;

NOTE: The data set WORK.SALE has 18 observations and 4 variables.
NOTE: The DATA statement used 8.00 seconds.

ZOOM
```

屏幕2.3 LOG窗口显示的信息

## § 5 分析数据

表 2.1 的销售数据转换为 SAS 数据集后,接着我们准备对这些数据作分析计算。利用 SAS 系统提供的 SAS 过程对 SAS 数据集的数据作分析处理。屏幕 2.4 列出的这些 SAS 语句可完成 4 种分析计算。

(1)PROC PRINT;(或 proc print;)

这个语句要求 SAS 显示刚才创建的 SAS 数据集中的数据值。关键词 PROC(PROCEDURE 的缩写)指示一个 PROC 步开始了,PRINT 是你使用的 SAS 过程的名字。

这个 PROC 步由 PROC PRINT 语句和 RUN 语句组成。RUN 语句告诉 SAS 执行这个 PROC 步。

(2)PROC PREQ;

该语句要求 SAS 计算频率。

(3)TABLES TYPE \* REGION;

TABLES 语句描述包含频率的表格。该语句要求做二维表,显示每种类型的机器在每个地区的销售量。

(4)PROC SORT;

该语句要求 SAS 对这个数据集排序。

(5)BY TYPE;

BY 语句描述用户想如何分类这些数据。这个语句告诉 SAS 按卖出机器的类型重新排列这些数据,类型为 'c' 的观测放在一起,'sm' 的观测放在后面。



```

PROGRAM EDITOR
Command ==>

00001  proc print;
00002  run;
00003  proc freq;
00004  tables type*region;
00005  run;
00006  proc sort;
00007  by type;
00008  run;
00009  proc means;
00010  by type;
00011  run
00012
00013
00014
00015
00016
00017
00018
00019
00020
00021

```

屏幕 2.4 分析数据集的 SAS 语句

(6) PROC MEANS;

该语句要求对数值变量计算简单统计量。

(7) BY TYPE;

该语句告诉 MEANS 过程,按 TYPE 分或二组来分别计算简单统计量。

当提交屏幕 2.4 的这些语句后(用 SUBMIT 命令),PGM 窗口的这些语句消失,并在右下边上出现字母'R'表示正在运行这些语句;在 LOG 窗口可以看到运行的记录;在 OUTPUT 窗口输出几个过程产生的结果。例如屏幕 2.5 是 PROC MEANS 的输出。

## § 6 结束 SAS 会话

当对数据的分析计算完成后,你想结束这个 SAS 会话,退出 SAS 系统。方法有两种:

(一)用命令行命令

当光标处在三个基本窗口之一时,使用以下任一条命令都可以退出 SAS 系统:

command ==> BYE ✓

或 command ==> ENDSAS ✓

键入退出命令后,系统或者立即退出 SAS 会话系统,或者提问是否决定退出 SAS 系统,根据用户的回答,系统或退出或继续等待命令。

(二)用 ENDSAS 语句

在 PGM 窗口插入语句