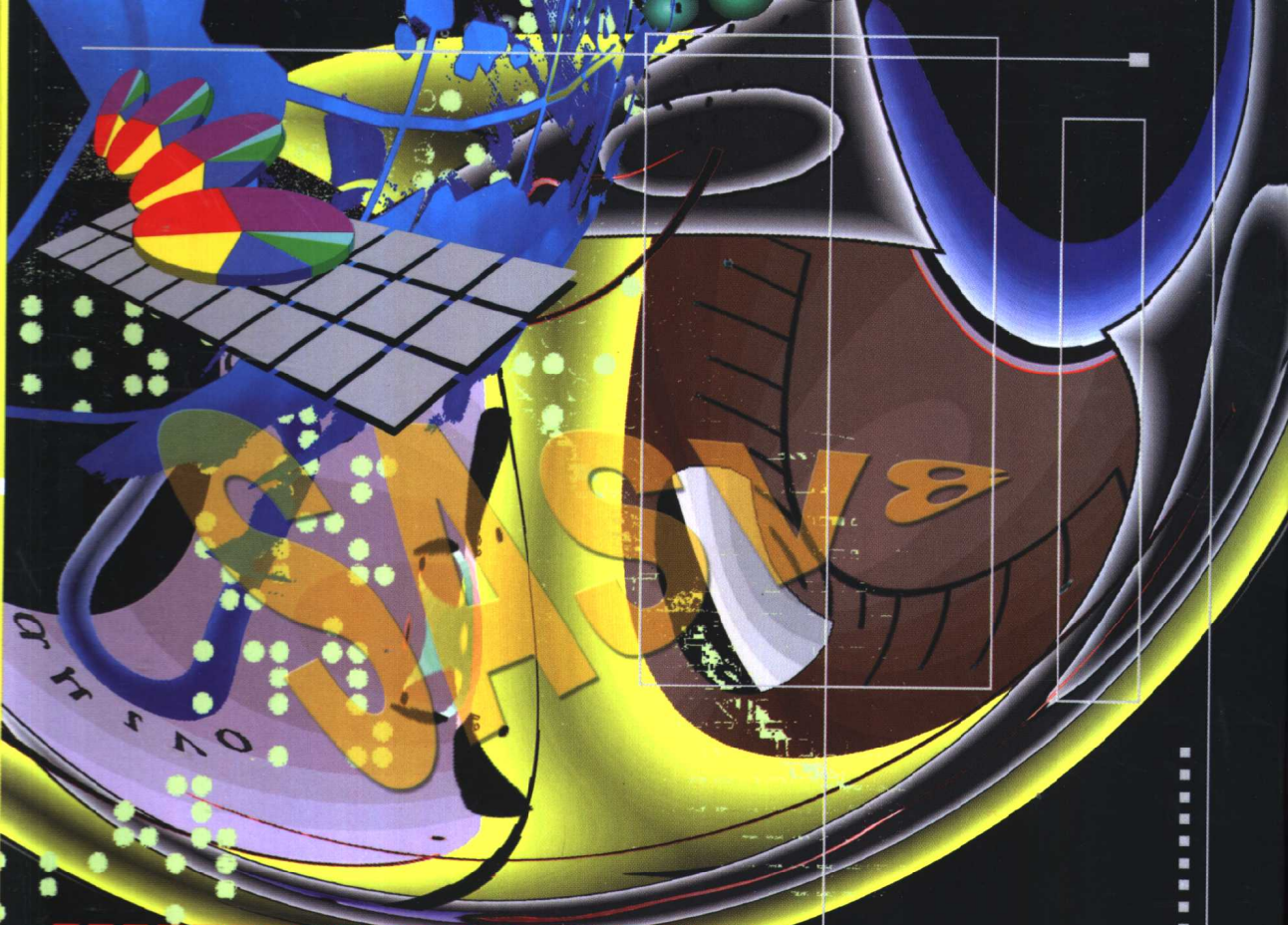




现代统计分析方法及应用系列丛书

汪嘉冈 编

SAS V8基础教程

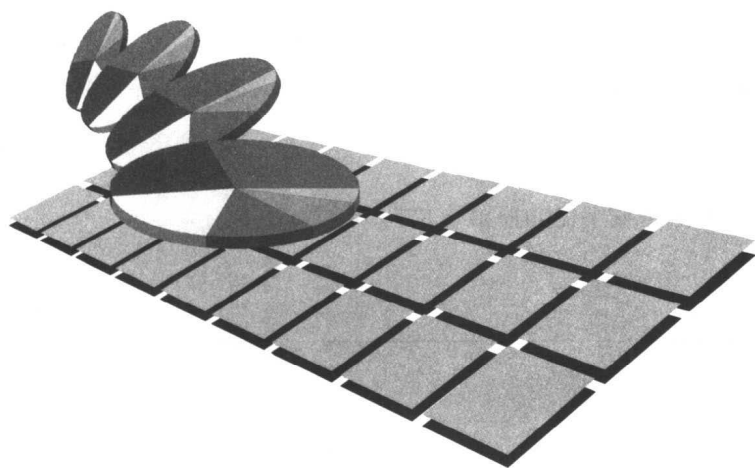
中国统计出版社
China Statistics Press



现代统计分析及应用系列丛书

SAS V8基础教程

汪嘉冈 编



中国统计出版社
China Statistics Press



NJS77/01

(京)新登字 041 号

图书在版编目(CIP)数据

SAS V8 基础教程/汪嘉冈编.
- 北京:中国统计出版社, 2001.9
ISBN 7-5037-3666-6

I . S…
II . 汪…
III . 统计分析-应用软件, SAS V8-教材
IV . C812

中国版本图书馆 CIP 数据核字(2001)第 068751 号

作 者/汪嘉冈
责任编辑/吕 军
封面设计/张建民
出版发行/中国统计出版社
通信地址/北京市西城区月坛南街 75 号 邮政编码/100826
办公地址/北京市丰台区西三环南路甲 6 号
电 话/(010)63459084, 63266600-22500(发行部)
印 刷/北京顺义兴华印刷厂
经 销/新华书店
开 本/787×1092mm 1/18
印 张/32
印 数/1-2500 册
字 数/650 千字
版 别/2003 年 12 月第 1 版
版 次/2003 年 12 月北京第 1 次印刷
书 号/ISBN 7-5037-3666-6/C·1982
定 价/50.00 元

中国统计版图书, 版权所有, 侵权必究。

中国统计版图书, 如有印装错误, 本社发行部负责调换。

序 言

赛仕研究所有限公司 (SAS Institute Inc.) 自 1976 年成立起至今, 始终致力于一个不变的主题: 应用当代最先进的信息技术与计算机技术, 做好业务数据与历史数据深加工。围绕这一主题, 伴随着信息技术与计算机技术的发展, 赛仕研究所有限公司走过了四分之一多个世纪不平凡的里程: 从初期的统计分析系统与报表工具开发商发展成为当今世界闻名的高度集成化的、严肃的、科学的数据仓库/数据挖掘/决策支持技术与行业应用系统的主要提供商。

今天, 拥有近万名员工的赛仕研究所有限公司一如既往地分布在全球 100 多个国家金融服务业、保险业、电信业、制造业、零售业、医药、能源、政府和教育等领域的 40000 多个用户点 3500000 多个用户提供其享有盛誉的技术支持与行业咨询服务。1999 年, SAS 公司年收入突破 10 亿美元。2002 年, SAS 公司年收入为 11 亿 8 千万美元。财富 (Fortune) 500 强中有 90% 的公司在使用 SAS 软件和解决方案, 2003 年福布斯 (Forbes) 500 强前 100 名公司中有 98 家公司在使用 SAS 软件和解决方案。SAS 公司和 SAS 客户保持着长期战略合作伙伴关系。98% 的

SAS 客户选择继续使用 SAS 软件和解决方案，60% 以上的 SAS 客户追加在 SAS 软件和解决方案上的投资。

享誉世界统计分析与决策支持领域的 SAS 软件系统，在中国拥有相当数量的使用者与商业用户。为了增强国内使用者与商业用户对 SAS 软件的了解，汪嘉冈教授在我公司的支持下，从学习与使用的角度系统地编写了这本《SAS V8 基础教程》。特别需要在此强调的是：本书不是 SAS 手册的中文编译，而是作者汪嘉冈教授与赛仕研究所有限公司一些资深员工在使用与帮助 SAS 用户的过程中的实践心得，是为了帮助那些想真正学习与使用 SAS 的人们减少学习弯路，是立足 SAS 相关技术资料的再创作，是学习与使用 SAS V8 的国内目前唯一一本中文 SAS 基础教材。边做边学是最有效的学习方法。本书在每章都附有习题，书中的例子和习题都是为了帮助读者在学习一般原理和做法的同时，进行实际操作。

赛仕研究所有限公司对汪嘉冈教授的严谨认真、谦和好学的治学态度深表钦佩。相信读过汪先生以前主笔的《SAS 基础教程》(1997)、《SAS 应用统计教程》(1999) 与《SAS 软件与应用统计》(2000) 等书的读者也会有同感。我们深信他的这本书也一定会对那些学习与使用 SAS V8.0 版本的人们有实质性帮助。

最后，感谢中国统计出版社对本书出版工作的大力支持！

赛仕研究所有限公司

2003 年 10 月

目 录

第一章 SAS 系统的工作环境	(1)
§ 1.1 进入 SAS 系统	(1)
§ 1.2 浏览 SAS 主窗口与发布命令	(3)
§ 1.3 SAS 的子窗口及其切换	(10)
§ 1.4 获取更多的帮助	(22)
第二章 SAS 系统对数据文件的管理	(25)
§ 2.1 SAS 文件和 SAS 逻辑库	(25)
§ 2.2 使用浏览器管理 SAS 库及其成员	(28)
§ 2.3 SAS 数据集	(33)
§ 2.4 输入和输出格式	(40)
第三章 利用窗口菜单操作和建立 SAS 数据集	(51)
§ 3.1 利用 Viewtable 浏览 SAS 数据集	(51)
§ 3.2 利用 Viewtable 修改与输出 SAS 数据集	(59)
§ 3.3 用 Viewtable 建立新的 SAS 数据集	(64)
§ 3.4 利用 Import 菜单转换外部文件为 SAS 数据集	(69)
第四章 用编程读入数据建立 SAS 数据集	(82)
§ 4.1 SAS 编程基本概念	(82)
§ 4.1.1 概述	(82)
§ 4.1.2 指定逻辑库名与逻辑文件名	(86)
§ 4.1.3 查看逻辑库和数据集的属性和内容	(91)
§ 4.2 用数据步读入数据生成 SAS 数据集	(94)
§ 4.2.1 数据文件格式与数据步的一般形式	(94)
§ 4.2.2 按列输入模式	(97)
§ 4.2.3 格式化输入模式	(102)

§ 4.2.4	列举模式输入·····	(106)
§ 4.2.5	命名模式输入及混合使用各种模式·····	(110)
§ 4.3	用数据步读入数据生成 SAS 数据集(二)·····	(112)
§ 4.3.1	数据步程序的编译和执行阶段·····	(112)
§ 4.3.2	几种常见程序错误的排除·····	(119)
§ 4.3.3	数据步读入数据的一些补充·····	(128)
§ 4.4	读入外部文件·····	(135)
§ 4.4.1	过程 IMPORT·····	(136)
§ 4.4.2	用 LIBNAME 语句连接 DBMS·····	(142)
§ 4.4.3	用过程 ACCESS 访问外部文件·····	(144)

第五章 数据加工····· (157)

§ 5.1	变量和观测的选择·····	(157)
§ 5.1.1	SET 语句的使用·····	(157)
§ 5.1.2	对变量的选择·····	(160)
§ 5.1.3	对观测的选择·····	(165)
§ 5.1.4	按变量值对数据集排序·····	(171)
§ 5.2	用 SAS 语句建立新的变量·····	(173)
§ 5.2.1	赋值语句·····	(173)
§ 5.2.2	条件语句·····	(182)
§ 5.3	SAS 函数·····	(194)
§ 5.3.1	函数的检索与调用·····	(194)
§ 5.3.2	与日期时间有关的函数·····	(197)
§ 5.3.3	字符运算函数·····	(200)
§ 5.3.4	数值与字符转换函数·····	(211)
§ 5.3.5	函数 LAG 和 DIF·····	(216)
§ 5.4	循环语句和数组·····	(218)
§ 5.4.1	循环语句·····	(218)
§ 5.4.2	数组·····	(226)
§ 5.5	写入 SAS 数据集和外部文件·····	(234)
§ 5.5.1	用 OUTPUT 语句写入 SAS 数据集·····	(234)
§ 5.5.2	写入文本格式文件·····	(239)
§ 5.5.3	小结·····	(245)
§ 5.6	使用 Analyst 菜单加工 SAS 数据集·····	(246)

第六章 面向多个数据集的加工	(257)
§ 6.1 数据集的拼接	(257)
§ 6.1.1 数据集的串接	(258)
§ 6.1.2 数据集的并接(Merge)	(265)
§ 6.1.3 用 Analyst 菜单拼接数据集	(269)
§ 6.1.4 数据集拼接的一些例子	(274)
§ 6.2 用过程 SQL 查询和加工数据	(284)
§ 6.2.1 数据的查询与加工	(285)
§ 6.2.2 用 SQL 拼接数据表	(291)
§ 6.3 使用 SQL 查询窗口	(295)
§ 6.3.1 单表的查询	(297)
§ 6.3.2 多表的查询	(306)
第七章 数据汇总与报表制作	(309)
§ 7.1 使用过程 PRINT 制作报表	(309)
§ 7.1.1 过程 PRINT 的基本语句	(309)
§ 7.1.2 改进列表报告	(317)
§ 7.1.3 使用菜单实现数据集列表输出	(327)
§ 7.1.4 过程 PRINTTO	(332)
§ 7.1.5 过程 FORMAT	(334)
§ 7.2 频数与描述统计量计算	(338)
§ 7.2.1 频数及百分数的计算	(338)
§ 7.2.2 使用菜单进行频数统计	(352)
§ 7.2.3 计算描述统计量	(359)
§ 7.2.4 用菜单计算描述统计量	(371)
§ 7.3 使用过程 TABULATE 制作汇总表	(375)
§ 7.3.1 过程 TABULATE 的基本用法	(375)
§ 7.3.2 对表格形式的进一步修饰	(382)
§ 7.3.3 用菜单制作汇总表	(389)
第八章 制作图形展现数据	(396)
§ 8.1 使用过程 GPLOT 制作散点图和连线图	(397)
§ 8.1.1 单一连线图形的制作	(398)
§ 8.1.2 多条连线的图形	(413)

§ 8.1.3	用菜单制作散点图和连线图·····	(423)
§ 8.2	使用过程 GCHART 制作柱状图和饼图 ·····	(428)
§ 8.2.1	制作基本的汇总图·····	(429)
§ 8.2.2	柱状图的修饰·····	(438)
§ 8.2.3	饼图的修饰·····	(446)
§ 8.2.4	用菜单制作柱状图和饼图·····	(453)
§ 8.3	图形的存储利用和输出 ·····	(459)
第九章	ODS 初步 ·····	(468)
§ 9.1	概述 ·····	(468)
§ 9.2	输送至目标 Listing, Output 与 Printer 等 ·····	(470)
§ 9.3	用 ODS 生成 HTML 输出 ·····	(488)
第十章	SAS 宏功能 ·····	(507)
§ 10.1	概述·····	(507)
§ 10.2	SAS 宏变量的使用与定义 ·····	(511)
§ 10.2.1	自动宏变量 ·····	(513)
§ 10.2.2	用户自定义宏变量 ·····	(515)
§ 10.2.3	宏变量定义和使用的进一步介绍 ·····	(518)
§ 10.3	宏程序·····	(534)
§ 10.3.1	简单的宏定义 ·····	(534)
§ 10.3.2	含参数的宏 ·····	(539)
§ 10.3.3	宏的 %IF-%THEN-%ELSE 语句 ·····	(545)
§ 10.3.4	宏循环语句 ·····	(549)
§ 10.3.5	宏的存储与调用 ·····	(557)
编写后记 ·····		(561)
参考书目 ·····		(563)

第一章 SAS 系统的工作环境

本章内容

进入 SAS 系统

SAS 主窗口与发布命令

浏览 SAS 常用子窗口的功能与提交程序

进入帮助菜单

§ 1.1 进入 SAS 系统

SAS 系统是一个可以在许多不同的操作系统下运行的跨平台的应用软件系统。但无论在什么操作环境下，SAS 系统都提供一个大体相同的运行环境和许多共同的功能。本书虽然只介绍 Windows 操作系统下 SAS 的使用方法，但所涉及 SAS 操作和编程的大部分内容(除了与操作系统的特有功能和文件系统相联系的部分)，对其它操作系统下也是适用的。

Windows 下的 SAS 系统是 Windows 操作系统下的一个应用系统，所以它和其它许多应用系统一样，保留了 Windows 为应用系统提供的一些功能，例如在 SAS 中保留了 Windows 的一些热键(Ctrl + C 等)和剪贴板的功能等。

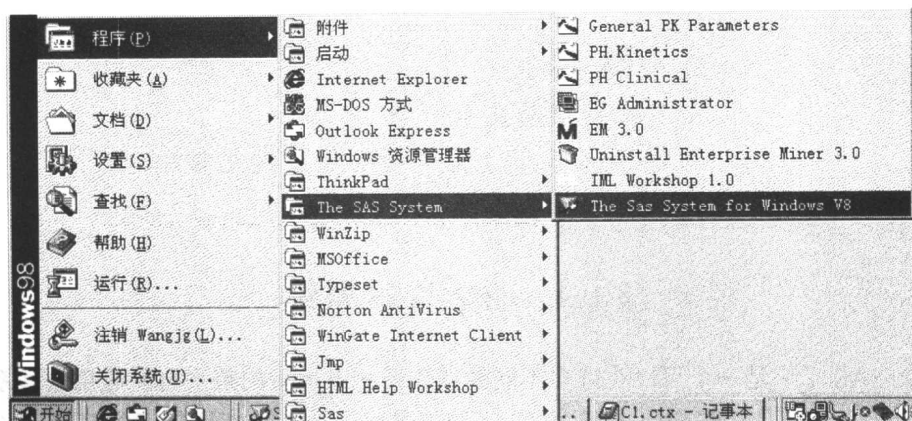
SAS 系统可以有多种方式在各种操作系统下运行，常用的有交互方式和批作业方式。在 Windows 系统下也是如此。在交互运行方式下，操作者在启动了 SAS 的操作环境后直接在 SAS 的操作环境下发布命令，提交程序或进行其它操作，SAS 系统随即执行这些任务并返回结果。操作者根据返回的结果继续在 SAS 的操作环境下进行操作，并不断地根据返回的结果进行新的操作。在批作业方式下，操作者向操作系统提交启动 SAS 作业的要求和全部要运行的 SAS 程序，再由操作系统安排这一作业运行。操作者可在事后获得 SAS 程序运行的全部结果。本书将只介绍 Windows 操作系统下如何在交互方式下运行 SAS 完成各项任务的基本方法。

启动 SAS

在 Windows 的桌面操作环境下，若已正确安装了 SAS 版本 8，则安装程序会在桌面左下方“开始”按钮弹出的程序菜单下加入一个 The SAS System 的文件夹。这时可按下列步骤启动 SAS 系统：

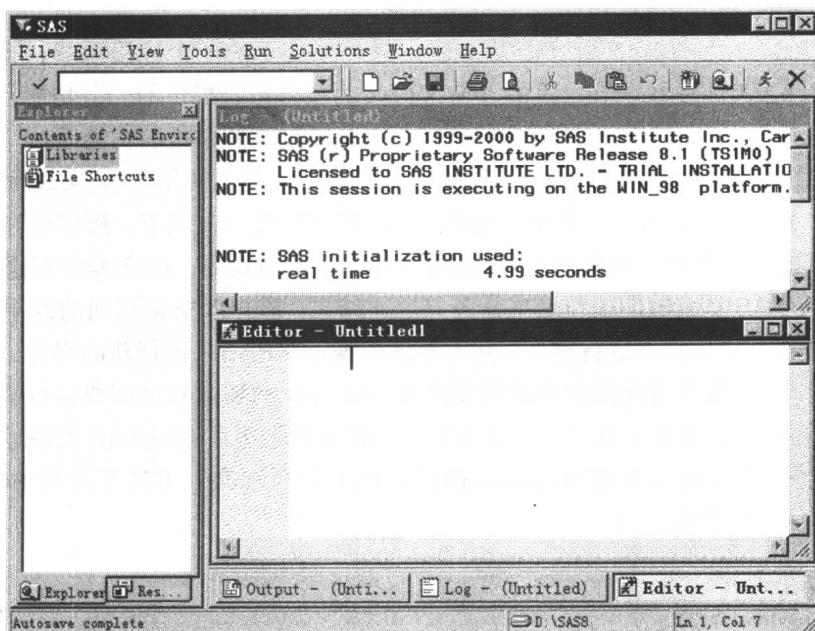
开始⇒程序⇒The SAS System⇒The SAS System for Windows V8

以上的操作表示点击桌面左下方“开始”按钮，并在弹出菜单中选“程序”。再在弹出的子菜单中选“The SAS System”等等。这里及以后使用鼠标点击时，若不指明是哪一个键时，都是指点击或双击鼠标器的‘选中键’即‘左键’。



在完成上述操作后经过几秒钟就会在屏幕上弹出如下的 SAS 系统的主窗口。它就是交互方式下运行 SAS 的工作界面。

2

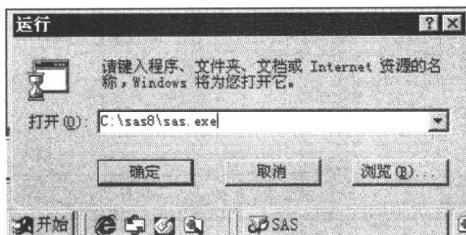


如果在 SAS 系统安装时，已经在 Windows 的桌面上建立了启动 SAS 系统的图标(一般为 )，则双击该图标也可启动 SAS。

若在操作系统中 SAS 是安装在 C:\SAS8 目录之下,那么双击图标相当于运行

C:\SAS8\sas.exe

所以也可以用直接运行上述程序来启动 SAS. 例如在桌面的左下方点击**开始**→**运行**,并在弹出的运行框中键入 C:\SAS8\sas.exe

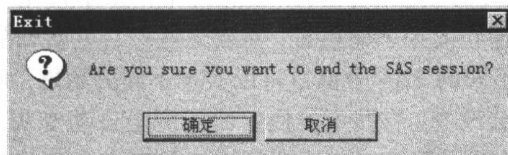


在键入完毕后,再按**确定**即可启动 SAS.

☞ 利用桌面的图标或利用安装程序提供的菜单启动 SAS 系统还包括了指定起始位置、指定设置文件(Sasv8.cfg)和自动执行文件(autoexec.sas),所以利用图标或菜单启动 SAS 系统比简单的运行 sas.exe 要包含更多的内容.

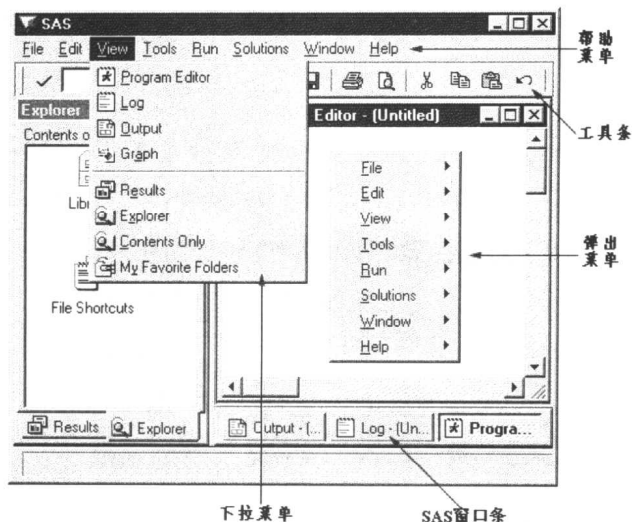
退出 SAS

在 SAS 的操作环境下,要结束 SAS 进程,只要点击 SAS 主窗口右上角的按钮  并在确认框中点击按钮**确定**,就可退出 SAS,返回 Windows 操作系统或其它应用.



§ 1.2 浏览 SAS 主窗口与发布命令

启动 SAS 后即展示如下的 SAS 主窗口,它是交互方式下运行 SAS 的工作界面,也称为应用工作空间(AWS: Application Work Space).它也是 SAS 的显示管理界面(DMS: Display Management System).

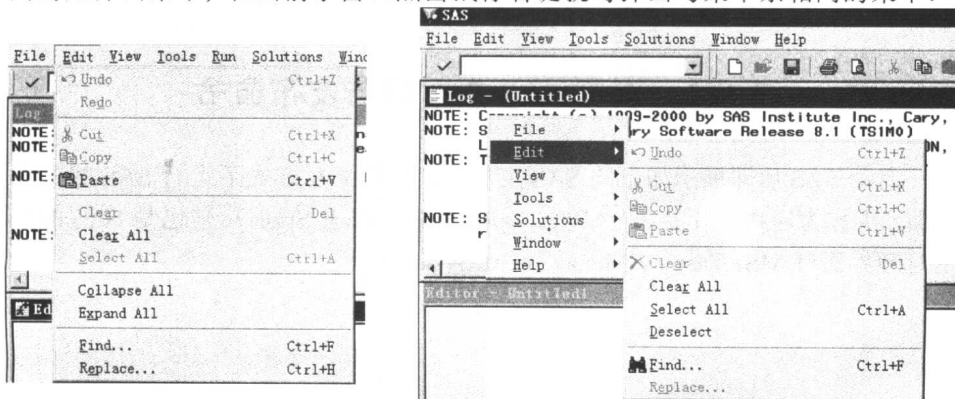


在 SAS 系统中要完成各种应用任务，可以通过提交 SAS 程序或者通过 SAS 的一些子菜单系统提交任务来实现。对交互操作中一些简单的任务和动作，可以在交互环境下发布命令来实现。所以在 SAS 的主窗口中：

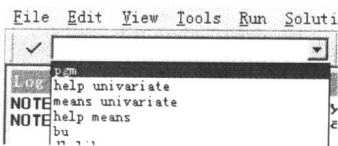
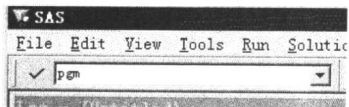
- 提供了菜单、命令框、工具条来发布命令和调用子菜单。
- 提供了多个子窗口和窗口条作为程序编写、提交和展示结果的环境。
- 提供了状态条展示系统信息。

菜单条(下拉菜单)

在 SAS 主窗口的上部就是菜单条(也称为下拉菜单)，它表示在当前(激活)的子窗口下可进行操作的菜单的栏目。点击这些栏目就会伸展出可操作的栏目或子菜单。由于在不同的 SAS 子窗口可以实现不同的功能，所以激活不同的子窗口菜单的内容也会改变，有些不能使用的功能会自动变灰。与菜单条功能相同的是弹出菜单，在当前子窗口点击鼠标右键就可弹出与菜单条相同的菜单。

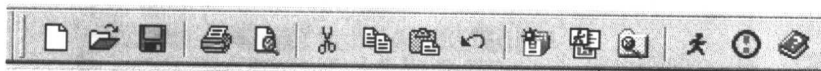


命令框



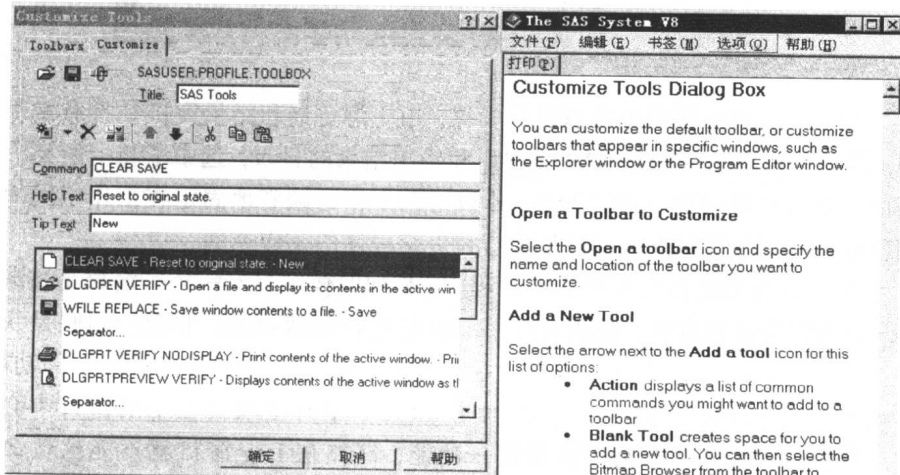
命令框提供键入和提交 SAS 命令的场合。例如键入 pgm 并回车就可使程序编辑器窗口激活。键入 bye 就可退出 SAS 返回 Windows 操作系统或其它应用。按下命令框右边的按钮  就可显示其保留的最近曾提交的若干条命令。若要重复使用只需选中所需的命令并点击左边的钩号按钮  就完成了提交命令。

工具条



在工具条中排列了若干个按钮，每个按钮相当于一个常用的命令。设置这一工具条可以让用户快速、直接发布命令。只需点击工具条中的相应按钮就可发出命令。当光标移至各个按钮时，按钮边上会弹出一个标签显示按钮的简单功能，同时在主窗口下方的信息条上显示该按钮功能的说明。

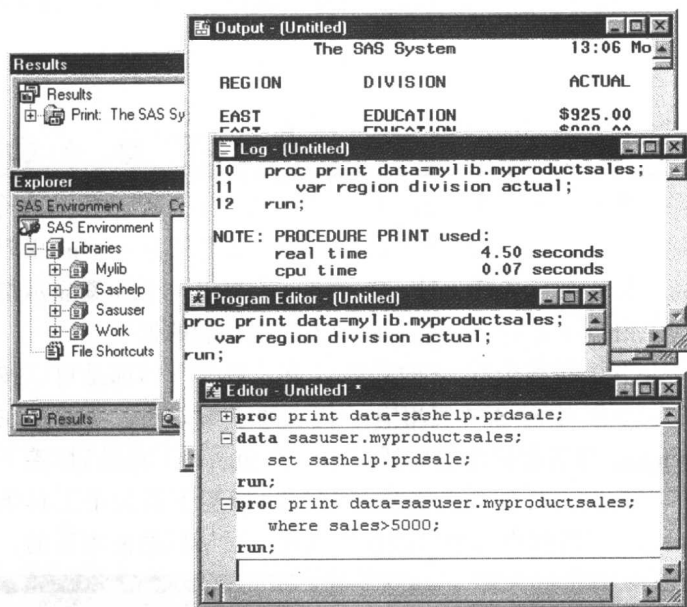
工具条所包含的按钮个数、它的功能以及采用的图标都是可以按用户的需要自行设定的。当您对 SAS 的使用有了初步的了解以后，就可以在菜单中选 **Tools→Customize**。然后在弹出的对话框中，Toolbars(工具条)标签下有关于工具条安排的许多设置，而在 Customize(客户化)标签下有关于工具条中增减或修改按钮的设置。对话框中的帮助按钮可以提供对话框的使用帮助。



子窗口工作空间

在 SAS 主菜单的中部是一个供打开 SAS 子窗口的工作空间。缺省地，它打开了以下窗口：

- 浏览器(Explorer),
- 结果(Results)窗口,
- 增强型编辑器(Enhanced Editor), 它只在 Windows 操作环境下才有。一般的有 程序编辑器(Program Editor)。
- 日志 Log 窗口,
- 输出(Output)窗口。

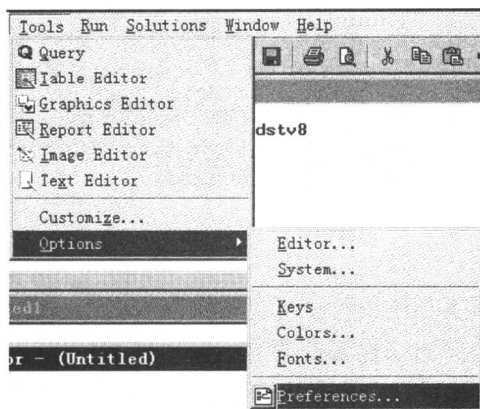


SAS 子窗口的排列也可以自行设置。由于浏览器和结果窗口是具有索引性的；即他们提供的分别是 SAS 文件库和输出结果的目录，选中了其中的某一项还由其它窗口显示进一步的细节。所以这两个窗口缺省地固定在工作空间的左侧(Docked)。而其它几个窗口是编写和提交程序的工作空间。

在 SAS 进程进行交互操作的过程中，任何时候都只有一个子窗口是被激活的，或称它为当前窗口。激活窗口上部的标题栏是蓝色的，非激活的则是灰色的。用鼠标点击某个子窗口就可激活该窗口。由于不同的窗口有不同的功能，系统为不同窗口提供了不同的命令。所以菜单中有些项目在个别窗口中会变灰，表示在当前窗口中无法使用这一功能。例如在 Log 窗口就不能执行命令

Open(打开), 也不能使用菜单中 File 下 Open 这个项目。

SAS 子窗口的许多特性是可由用户自行设置的。例如也可设置使浏览器和结果窗口不固定在左侧, 让激活的其它窗口有较大的显示空间。当您对 SAS 的使用有了初步了解以后, 就可以将窗口的功能特性安排得符合自己的使用习惯。这只需在菜单中选 **Tools**⇒**Options**⇒**Preferences**, 然后在弹出的对话框中选择不同的标签对感兴趣的项目进行设置即可。



窗口条



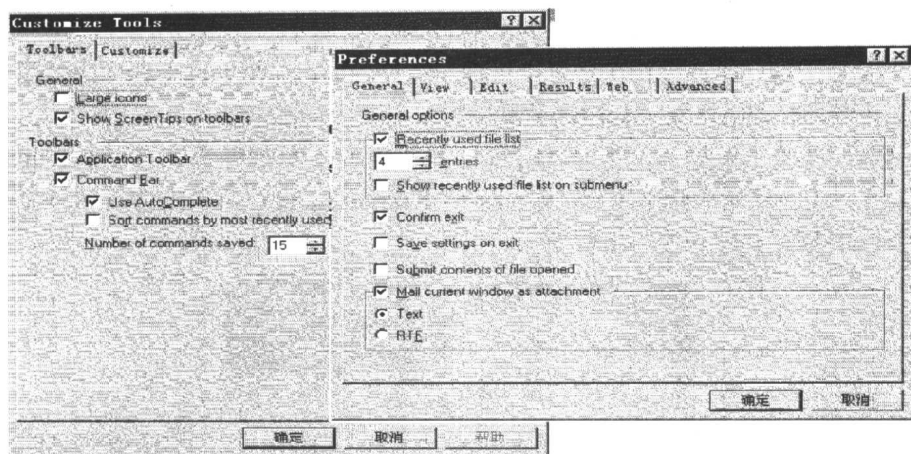
在 SAS 子窗口工作空间的下方是窗口条, 它包含了所有打开的子窗口的标签。对非激活的窗口, 点击这个窗口的标签就使它激活并置于显示的前台。对激活的窗口, 点击这个窗口的标签就使它最小化而让其它窗口激活。

状态条



在 SAS 主窗口的最下方是状态条。它的左部是信息显示区, 中部是运行 SAS 的工作文件夹名称, 右部是增强型编辑器光标所在的位置。信息显示区提供菜单和工具图标的帮助信息, 在子窗口操作时也提供一些提示信息。要改变 SAS 的工作文件夹位置只需用鼠标左键双击原有工作文件夹名称, 然后在弹出的 Change Folder 对话框中修改即可。增强型编辑器光标的位置提示光标在增强型编辑器中的行列位置。

SAS 主菜单中菜单、命令框、工具条、窗口条和状态条的出现与否是可以由用户自行设置的。这只要在菜单中分别选 **Tools**⇒**Customize** 和 **Tools**⇒**Options**⇒**Preferences**。然后可分别在 Toolbars 和 View 标签下进行设置。



提交命令

SAS 的不同窗口有不同的特征可以帮助您的工作。要在各个子窗口实现各种功能以及实现窗口间的切换都是靠在 SAS 的应用工作空间中发布命令来实现的。通常可有如下几种方式来发出命令：

- 在命令框中键入命令，并按回车。例如在命令框中键入命令 help 并按回车即可弹出帮助窗口。
- 使用菜单（或弹出菜单）。例如在菜单中选 **Help**→**SAS System**，就可弹出关于 SAS 系统的帮助菜单窗口。
- 使用工具条。例如点击工具条最右边的图标  就会弹出关于 SAS 系统的帮助菜单或其它与当前窗口相关的帮助菜单窗口。
- 使用功能键。例如在缺省的设置下，按 F1 键也会弹出关于 SAS 系统的帮助菜单或其它与当前窗口相关的帮助菜单窗口。

以上几种发布命令的方式是等价的。利用命令框是最基本的，但要求操作者记住常用的命令。下拉菜单的许多项目就是某一个命令，只是在下拉菜单中已经把常用的命令按其功能分为 File(文件)，Edit(编辑)，View(查看)，Tools(工具)，Run(运行)，Solutions(解决方案)，Window(窗口)和 Help(帮助)几个栏目，便于使用者按功能查找使用，而且只要点击鼠标可不必键入命令。工具条也是将调用最常用功能的命令用图标形式置于其中，使操作者单击鼠标就可发出命令。SAS 系统也可将键盘的功能键和组合键设置为一些常用的命令的热键，操作者只要按下这些热键就可发出命令。

在使用字符界面的 SAS 操作环境中，常在每个窗口的顶部有一个用于键入命令的命令行。