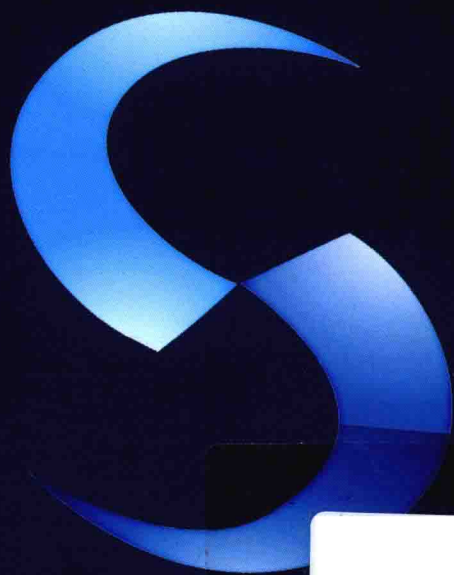


继《深入解析SAS：数据处理、分析优化与商业应用》之后，资深SAS专家的又一倾力之作
首次系统全面阐述SAS最新分析平台SAS Viya的整体架构和技术细节，深入细致展现灵活运用
SAS Viya强大分析能力的多种方式和详细步骤

SAS Viya

潘红莲 杨光辉 张涛 夏坤庄 著



OPEN ARCHITECTURE
POWERFUL
ADAPTIVE
UNIFIED
MACHINE
LEARNING
FACTIMAC
NNET
GRADBOOST
FOREST
ACTION
VISUAL
STATISTICS
PROGRAMMING CASL
LUA PYTHON
RESTful JAVA
GENSELECT
LOGSELECT
TREESPLIT
KELVIS
PCA
CAS
DATA
MANAGEMENT
CASUTIL
DATA STEP
DATA
SGPLOT
SGPANEL
SGSCATTER
EXPLORATION
CODE EDITOR
SAS TASK and UTILITY
STUDIO SNIPPET

SAS VIYA



机械工业出版社
China Machine Press

SAS Viya

潘红莲 杨光辉 张涛 夏坤庄 著

OPEN ARCHITECTURE
POWERFUL
ADAPTIVE
UNIFIED
MACHINE
LEARNING
FACTMAC
NNET
GRADBOOST
FOREST
ACTION
VISUAL
STATISTICS
CAS
DATA
MANAGEMENT
CASUTIL
DATA STEP
SAS VIYA
DATA
SGPLOT
SGPANEL
SGSCATTER
EXPLORATION
CODE EDITOR
SAS TASK and UTILITY
STUDIO SNIPPET
PROGRAMMING CASL
LUA PYTHON
RESTful JAVA



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

SAS Viya/ 潘红莲等著. —北京: 机械工业出版社, 2017.8
(SAS 核心技术丛书)

ISBN 978-7-111-57889-5

I. S… II. 潘… III. 统计分析 - 应用软件 - 基本知识 IV. C819

中国版本图书馆 CIP 数据核字 (2017) 第 209490 号

SAS Viya

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 陈佳媛 杨绣国

责任校对: 李秋荣

印 刷: 北京诚信伟业印刷有限公司

版 次: 2017 年 9 月第 1 版第 1 次印刷

开 本: 186mm×240mm 1/16

印 张: 20

书 号: ISBN 978-7-111-57889-5

定 价: 79.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

为什么要写这本书

大数据、数据挖掘和机器学习、人工智能以及数据可视化发展到了一个全新的阶段，开始影响着企业、组织和个人理解数据并进行决策的各个方面。这些概念并不是代表崭新的理论或突然出现的实际需求，而是经历了长时间的理论发展和实际应用探索之后的必然结果。

关于海量数据及数据挖掘的概念与理论，乃至相应的产品和工具，早在至少二十年之前就已出现，即使在国内，也已经不算是新生事物了。至于人工智能，更是人类很早就梦想要突破的领域，无须检索任何文献去考查究竟是在何时开始探索这一领域的，仅凭笔者个人的经验，就可以追溯到 20 世纪 90 年代。彼时，我的研究生导师华克强老师，就已经开始带领实验室的老师和学生们探讨如何把诸如神经网络、遗传算法、模糊数学、支持向量机等算法运用于水下机器人的控制当中。数据可视化，更全面地说，数据分析过程和方法的可视化，以及分析结果的可视化，也是一直都存在并且为业界所努力发展的领域。就 SAS 而言，从最初主机上的命令行运行方式，到出现图形化界面的版本，到提供集成的开发环境，再到出现了分析专家和业务专家都可以使用的可视化分析和建模产品，一直到现在的全线产品与技术可视化，其一直在可视化的道路上前进和发展着。

时至今日，在运算方法上以迭代与分步为特征的数据分析，包括机器学习，从单机处理（SMP）模式到大规模并行处理（MPP）模式的突破、GPU 的出现及其在数据分析领域的广泛应用，使得更为庞大的数据处理、更为复杂的深度学习和更为广泛的可视化用户，都得到了很好的计算支撑。在这个过程中，SAS 作为领导者和推动者，一直在引领数据分析各个领域的潮流，包括大数据、可视化以及机器学习。其新一代分析平台 SAS Viya 正是 SAS 公司在这些领域所做贡献和突破的集中体现。

SAS Viya 是一个以大数据、人工智能以及数据可视化为导向的、全新的统一平台和环境。

既集中提供了大数据、人工智能以及数据可视化所需要的专业功能，也提供了支持这些专业功能所必需的高性能、可扩展性和开放性。从而可以利用人工智能的优势，并结合传统的分析方法来有效地处理大数据，并且使得这一切在可视化的环境和方法下得到实现。

继上一本书《深入解析 SAS：数据处理、分析优化与商业应用》全面介绍 SAS 9，帮助国内广大的用户学习和使用 SAS 之后，本书旨在全面快速地介绍 SAS Viya，让更多用户学习和使用 SAS 的最新平台和技术，能够快速利用 SAS Viya 开发自己的大数据分析和人工智能应用并实现可视化展现。本书在保证全面讲解各技术点的基础之上，包含了必要的细节，使得读者能够迅速系统地了解 SAS Viya，并为进一步深入学习提供良好的开端。

请注意，本书是基于 SAS Viya 3.1 编写的，在功能上，SAS Viya 的新版本会不断提高和增强，这些也会在本书的后续版本中不断得到体现。

读者对象

本书主要适合于以下读者：

- 使用和学习 SAS 的科技工作和工程技术人员。
- 需要使用第三方开发语言调用 SAS Viya 提供的分析功能的工程技术人员。
- 使用 SAS Viya 进行大数据平台规划和建设的管理人员和技术人员。
- 使用 SAS 进行项目规划、实施和管理的系统架构师、系统管理员和项目管理人员。
- 团队的工作涉及 SAS 产品与技术的管理人员。

如何阅读本书

本书共有 9 章，下面介绍各个章节的主要内容，以方便大家查阅。

第 1 章主要阐述了 SAS Viya 的架构和基本组成。第 2 章阐述了 SAS 应用开发的基本概念，旨在帮助读者从整体上理解 SAS Viya 以及基于 SAS Viya 的应用开发。建议每一位读者都仔细研读这两章。

第 3 章和第 4 章，介绍了 CAS 的数据读写和管理以及数据处理。需要利用 SAS Viya 处理数据的读者可仔细阅读这两章。

第 5 章的目的在于保证本书结构的完整性，介绍了 SAS Viya 中的数据探索和展现。熟悉 SAS 9 的读者可以略过本章。

第 6 章是关于用 SAS Viya 进行数据分析的章节。SAS Viya 在数据分析方面添加了很多机器学习 and 深度学习的方法，建议分析建模领域的读者仔细阅读。本书的后续版本也会继续加

强这方面的内容。

有很多的 SAS 用户有着长期使用 SAS 9 的宝贵经验，在 SAS 9.4 系统上也开发部署了很多成熟的应用。为此，第 7 章专门阐述了 SAS Viya 和 SAS 9 之间的集成和交互。

第 8 章是关于 CAS 编程语言的介绍。利用 CAS 编程语言可以最灵活且最大限度地充分利用 SAS Viya 的所有数据分析和处理功能。如果读者期望能够直接调用 CAS 中的 ACTION，进行灵活自由且功能齐全的数据分析和数据处理，建议仔细研读本章。

第 9 章阐述了如何用 Java、Lua、Python 开发基于 SAS Viya 的应用。介绍了 SAS Viya 作为一个开放系统，如何支持第三方开发语言，让更多的数据分析开发者借助于 SAS Viya 的分析和处理能力，使用他们自己熟悉的开发语言建立满足自己需求的应用。

致谢

忙碌繁杂的日常工作之余，其实是不适合写书的。为了让需要学习的读者能够得到一个正确的路线图，为了让系统性的知识可以得到完整的传播，本书的每一位作者都做出了不懈的努力。

特别感谢机械工业出版社华章公司的杨绣国编辑（Lisa Yang）和 SAS 出版社的主编 Julie Platt。他们的帮助和宝贵建议，促成了本书的完成。

夏坤庄（Kansun Xia）

中国，北京，2017 年 7 月

Preface

Purpose

Big data, data mining and machine learning, artificial intelligence and data visualization have entered a new era and began to affect all aspects of decision-making based on information by businesses, organizations and individuals. These concepts are not representation of the new theory nor the sudden emergence of the actual needs, they are results from a long period of theoretical development and exploration in practical application.

The concept and technology of mass data and data mining, from theory to practice and even products and tools, appeared at least 20 years ago. It is not a new thing even in China. As for artificial intelligence, human being has been dreaming to break through in this field for a long time. When was this field explored? Without studying any literature on the history of science, but just based on the author's personal experience, it can be traced back to the nineties of last century. At that time, my graduate tutor, Keqiang Hua, had begun to lead teachers and students to explore how to apply algorithms such as neural networks, genetic algorithms, fuzzy mathematics and support vector machines to underwater robots. Huge efforts have been exerted in area of Data Visualization as well. In terms of SAS, from the original command line on the mainframe to the graphical interface version, to integrated development environment, to visual analysis and modeling products that can be used by analysts and business experts, until now the full range of products and technology visualization, it has been constantly pursuing and developing in the field of visualization.

Today, data analysis, characterized by iteration and step-by-step calculating, including in machine learning, has successfully extended from standalone processing mode (SMP) to large-scale parallel processing mode (MPP). And GPU has emerged and been widely used in the field of data

analysis. These technologies have provided a very good calculation support to mass data processing, more complex deep learning, and more extensive visualization users. In this process, SAS as a leader and promoter, has been leading the trend of data analysis in all areas including big data, visualization, and machine learning. A new generation analytics platform, SAS Viya, is the core of SAS contribution and breakthrough in these areas.

SAS Viya is a brand new, unified platform and environment oriented to big data, artificial intelligence and data visualization. It provides specialized functions required by big data, artificial intelligence and data visualization, as well as high performance, scalability and openness required to support these functions. Thus, users can take the advantages of artificial intelligence combined with traditional analytical methods to effectively process big data and implement all of these in the visual environment and methods.

The last book "Understand SAS - Data Processing, Analytical Optimization and Business Applications" is a comprehensive introduction to SAS 9. It is designed to help the majority of SAS users learn and use SAS. This new book is a quick introduction to SAS Viya and aimed to help SAS users to learn and use the latest SAS platform and technology to quickly develop their own big data analysis and artificial intelligence applications. This book contains appropriate amount of details that enable the readers to systematically understand SAS Viya and provides a good starting point for further learning.

Please note that this book is based on SAS Viya 3.1. New version of SAS Viya will continue to improve and enhance. These improvements and enhancements will be reflected in the subsequent revisions of this book.

Audience

This book is primarily targeted to following readers:

- ☐ Technical and engineering personnel who want to use or study SAS.
- ☐ Engineers who need to use the third-party programming language to invoke SAS Viya to provide analytical functions.
- ☐ System architects, system administrators, and project managers who use SAS for project planning, implementation, and administration.
- ☐ Management and technical staffs who use SAS Viya for big data platform planning and

implementation.

- ❑ Management staffs whose teams involve SAS products and technology in daily work.

Organization

There are totally 9 chapters in this book.

First chapter mainly describes the infrastructure and basic components of SAS Viya.

Second chapter introduces basic concepts of SAS application development, which aims to help readers understand SAS Viya and Viya-based application development as a whole. It is necessary for every reader to study these two chapters first.

Chapter 3 and Chapter 4 describe how to read, write CAS data and administration of CAS and data processing, which is very helpful for readers who need to use SAS Viya to process data.

Chapter 5 makes up the structural integrity of this book. It introduces data exploration and presentation in SAS Viya. Users who are familiar with SAS 9 can skip this chapter.

Chapter 6 is related to using SAS Viya to perform data analysis. SAS Viya has added a lot of machine learning and deep learning methods to data analysis. Readers in analysis and modeling fields will find it useful for their work. Subsequent versions of this book will continue to enhance this part.

A lot of SAS users have accumulated valuable experience over a long term use of SAS 9 and have developed a number of proven applications on SAS 9.4 system, given that Chapter 7 explains the integration and interaction between SAS Viya and SAS 9.

Chapter 8 introduces CAS programming language, which provides great flexibility and makes full use of data analysis and processing capabilities of SAS Viya. Those users who need to call the actions directly to achieve flexible, free and full-featured data analysis and processing can read this chapter carefully.

Chapter 9 explains how to use Java, Lua, and Python to develop SAS Viya-based applications. It describes how SAS Viya being an open system supports third-party programming language. With the analysis and processing capabilities of SAS Viya, more data analysis developers can use their familiar programming language to build applications as needed.

Acknowledgments

It is not an easy job to write a book after a busy work day, we have tried our best with the

expectation that the systematic fundamental knowledge about SAS Viya can be fully propagated and the readers can get the right road map to effectively learn more.

Special thanks to Lisa Yang of Huazhang company of China Machine Press and Julie Platt, chief editor of SAS Press. Their help and valuable advices contributed a lot to this book.

Kansun Xia
Beijing, China
July 2017

目 录 *Contents*

前 言

第1章 SAS Viya 的架构和基本组成 ··· 1

- 1.1 SAS Viya 的架构 ··· 1
 - 1.1.1 SAS Viya 平台的组成部分 ··· 2
 - 1.1.2 SAS Viya 平台的部署架构 ··· 7
 - 1.1.3 基于 SAS Viya 的产品 ··· 8
- 1.2 CAS 的数据访问 ··· 8
 - 1.2.1 客户端访问数据 ··· 10
 - 1.2.2 服务器串行访问数据 ··· 10
 - 1.2.3 服务器并行访问分布式数据源 ··· 11
 - 1.2.4 服务器并行访问 HDFS ··· 11
- 1.3 CAS 的监控和管理 ··· 12
 - 1.3.1 系统状态 ··· 13
 - 1.3.2 配置 ··· 14
 - 1.3.3 资源监控 ··· 16
- 1.4 CAS 的安全性 ··· 17
 - 1.4.1 CAS 中的身份和角色 ··· 17
 - 1.4.2 认证 ··· 18
 - 1.4.3 授权 ··· 19
 - 1.4.4 加密 ··· 22

- 1.4.5 审计日志 ··· 22

- 1.5 CAS 数据容错 ··· 23

- 1.6 本章小结 ··· 23

第2章 SAS Viya应用开发概览 ··· 24

- 2.1 如何运行 SAS 代码 ··· 24
 - 2.1.1 SAS 程序的基本组成 ··· 25
 - 2.1.2 SAS Studio ··· 27
 - 2.1.3 批处理、交互式行模式及如何执行操作系统命令 ··· 39
- 2.2 SAS Viya 中开发应用的基本概念 ··· 40
 - 2.2.1 CAS 会话 ··· 41
 - 2.2.2 数据访问 ··· 45
 - 2.2.3 CAS 语言 ··· 55
 - 2.2.4 SAS 选项 ··· 56
- 2.3 SAS Viya 中的编码 ··· 62
 - 2.3.1 SAS Viya 的 UTF-8 编码 ··· 63
 - 2.3.2 语言 / 区域选项 LOCALE= ··· 64
- 2.4 SAS 代码在哪里执行 ··· 65
 - 2.4.1 DATA 步 ··· 66
 - 2.4.2 CAS 语言 ··· 66

2.4.3 SAS Viya 的基本过程	67	第4章 CAS中的数据处理	92
2.4.4 SAS 可视化数据建模和机器学习过程	67	4.1 DATA 步在 SAS Viya 中的执行 ..	92
2.4.5 DS2 语言	67	4.1.1 DATA 步的执行会话	93
2.4.6 SAS 函数和 CALL 例程	68	4.1.2 DATA 步在 CAS 中的单线程和多线程处理	96
2.4.7 格式和输入格式	68	4.2 DATA 步的 VARCHAR 字符类型	101
2.4.8 宏	68	4.2.1 LENGTH 语句创建 VARCHAR 字符变量	101
2.5 SAS Viya 与 SAS 9 协同开发	69	4.2.2 字符变量类型转换	102
2.6 第三方开发语言开发 SAS Viya 应用	70	4.2.3 字符串处理函数	104
2.7 本章小结	70	4.3 CAS 数据的分区和排序	104
第3章 CAS 的数据读写和管理	71	4.3.1 DATA 步的 BY 组处理和排序	105
3.1 CAS 会话、CASLIB 和 CAS 引擎逻辑库	71	4.3.2 DATA 步的 PARTITION= 和 ORDERBY= 数据集选项分区和排序	106
3.2 CAS 中的数据类型	71	4.3.3 使用 CASUTIL 过程的 LOAD 语句进行分区和排序	107
3.3 加载数据到 CAS 服务器	73	4.3.4 查看 CAS 表的分区信息	108
3.3.1 加载客户端文件到 CAS 服务器	73	4.4 ACTION DATASTEP.RUNCODE 执行 DATA 步代码	108
3.3.2 加载 CAS 端文件到 CAS 服务器	76	4.5 在 CAS 中应用格式进行数据处理	110
3.3.3 CAS 端文件的串行加载与并行加载	79	4.5.1 FORMAT 过程创建格式逻辑库	111
3.4 管理 CAS 数据	81	4.5.2 CAS 中自定义格式的应用 ..	112
3.4.1 查看 CAS 数据的属性	81	4.5.3 FORMAT 语句为变量指定格式	112
3.4.2 保存 CAS 服务器中的数据	84	4.5.4 CAS 中带格式变量的处理 ..	115
3.4.3 卸载 CAS 服务器中的数据	87	4.5.5 CAS 格式逻辑库的管理	117
3.4.4 删除文件	87		
3.4.5 复制文件	88		
3.4.6 全局化 CAS 表	89		
3.5 本章小结	91		

4.5.6 CAS 格式逻辑库搜索列表管理	122	6.4.3 数据加工	179
4.6 本章小结	124	6.5 数据建模	186
第5章 SAS Viya 中的数据探索和展现	125	6.5.1 模型的建立	186
5.1 SAS Viya 展现数据的方法：列表	125	6.5.2 模型的评估	195
5.2 SAS Viya 展现数据的方法：图形	127	6.5.3 模型的实施	210
5.2.1 SGPLOT	128	6.6 本章小结	211
5.2.2 SGPANEL	131	第7章 SAS Viya 与 SAS 9 执行环境的交互	212
5.2.3 SGSCATTER	134	7.1 SAS Viya 与 SAS 9 的差异	212
5.2.4 图形的类型	137	7.2 SAS/CONNECT	214
5.3 图形化展现 CAS 服务器的分析结果	152	7.2.1 SAS/CONNECT 介绍	214
5.3.1 示例 1：图形化展示聚类分析中误差平方和的下降过程	152	7.2.2 SAS 会话编码的兼容性	218
5.3.2 示例 2：图形化展示数据的统计结果	153	7.3 SAS Viya 与 SAS 9 的协作	218
5.4 本章小结	155	7.3.1 从已有的 SAS 9 平台到新的 SAS Viya 平台	219
第6章 用SAS Viya进行数据分析	156	7.3.2 从新的 SAS Viya 平台到已有的 SAS 9 平台	222
6.1 SAS Viya 中数据挖掘和机器学习过程步概述	156	7.4 本章小结	224
6.2 SAS Viya 中数据挖掘的一般流程	158	第8章 CAS编程语言介绍	226
6.3 确定业务问题和数据准备	159	8.1 CASL 简介	226
6.4 数据处理	160	8.2 开始使用 CASL	228
6.4.1 数据抽样	161	8.2.1 最简单的 CASL 代码	229
6.4.2 数据探索	161	8.2.2 调用 ACTION 进行多线程分析的 CASL 代码	229
		8.3 CASL 编程	230
		8.3.1 变量	230
		8.3.2 表达式和运算符	235
		8.3.3 数组	237
		8.3.4 字典	240

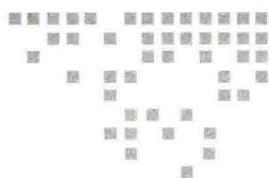
8.3.5	过程控制	241
8.3.6	ACTION	246
8.3.7	函数	255
8.3.8	输出	261
8.4	本章小结	261

第9章 用 Java、Lua、Python 和 REST API 开发基于 SAS Viya 的应用

9.1	SAS Viya 开放平台介绍	262
9.1.1	分析即服务和 SAS Viya 开放平台介绍	262
9.1.2	SWAT 架构介绍	263
9.1.3	基于 ACTION 及 ACTION 集合的数据挖掘和机器学习	264
9.2	用 Java 开发基于 SAS Viya 的应用	265
9.2.1	系统需求及开发环境配置	265
9.2.2	第一个 ACTION 调用：调用 ServerStatus 查看服务器状态	266
9.2.3	用 Java 将数据加载到 CAS 服务器	270
9.2.4	调用 ACTION 进行描述性统计分析	275
9.2.5	调用 ACTION 进行线性回归分析	277
9.2.6	对 CAS 服务器端事件的管理	279
9.3	用 Lua 开发基于 SAS Viya 的应用	284

9.3.1	系统需求及开发环境配置	284
9.3.2	第一个 ACTION 调用：调用 ServerStatus 查看服务器状态	285
9.3.3	用 Lua 将数据加载到 CAS 服务器	286
9.3.4	调用 ACTION 进行描述性统计分析	287
9.3.5	调用 ACTION 进行线性回归分析	288

9.4	用 Python 开发基于 SAS Viya 的应用	289
9.4.1	系统需求及开发环境配置	289
9.4.2	第一个 ACTION 调用：调用 ServerStatus 查看服务器状态	290
9.4.3	用 Python 将数据加载到 CAS 服务器	291
9.4.4	调用 ACTION 进行描述性统计分析	293
9.4.5	调用 ACTION 进行线性回归分析	294
9.5	用 CAS REST API 开发基于 SAS Viya 的应用	295
9.5.1	使用 Java 来调用通用类 REST Services	296
9.5.2	使用 Java 调用 REST Services 进行描述性统计分析	299
9.5.3	使用 Java 调用 REST Services 进行线性回归分析	303
9.6	本章小结	305



SAS Viya 的架构和基本组成

SAS Viya 平台是 SAS 公司于 2016 年推出的新一代的统一、开放及云就绪的高性能和可视化分析平台。SAS Viya 引入了 CAS (Cloud Analytics Services, 云分析服务) 服务器, 提供了基于内存处理的可扩展的高性能分析能力。

本章将从 SAS Viya 平台的架构开始, 介绍 SAS Viya 平台的基本特性、主要组成部分及各部分之间的交互关系, 并阐述在部署基于 SAS Viya 平台的产品和解决方案时可供选择的部署方式。在基本了解 SAS Viya 平台的架构之后, 本章将着重介绍在 SAS Viya 平台上的各种数据访问方式, 以帮助读者理解 SAS Viya 平台中的串行、并行数据加载及如何选择更有效的数据加载策略。本章最后还将介绍 CAS 服务器的监控和管理, 以及 SAS Viya 平台的安全性和数据容错性。

注意, SAS Viya 是一个平台, 基于这个平台, SAS 会推出各种用于数据处理和分析的产品和面向特定业务与行业的解决方案。基于 SAS Viya 所具备的开放性, 任何第三方都可以用其他开发工具 (例如 Java) 来开发自己的应用。但是 SAS Viya 本身并不是一个面向特定具体功能的应用或解决方案。

本书内容将基于 SAS Viya 平台的 SAS 可视化数据挖掘与机器学习产品来介绍, 在描述时, 如果不会涉及该产品的特有功能, 那么本书将不会提及该产品。

1.1 SAS Viya 的架构

SAS Viya 平台基于内存 (In-Memory) 处理, 可提供高性能分析和可视化分析功能, 同时还提供了便捷高效的环境管理功能和安全性, 并且可以部署在云环境中。SAS Viya 平台

具备如图 1.1 所示的一些特征。

（1）统一性

SAS Viya 以完整集成的架构体系、通用共享的低层代码实现和可视化的界面，使得采用 SAS Viya 平台的用户可以统一实现和管理整个分析生命周期中的所有活动——从数据处理和管理，到基本的报表和数据探索，再到分析模型的开发和管理等。SAS Viya 为用户提供了一个可管理可掌控的平台环境以创建灵活适用的分析资产。

（2）开放性

SAS Viya 的开放性是其重要特性。开发人员不但可以使用 SAS 进行应用的开发，而且还可以利用 SAS Viya 所提供的第三方开发语言接口使用 Python、Java、Lua 语言，以及 REST API 进行应用的开发。而这其中不变的是，无论采用哪种方式和语言，都可以自由访问和调用 SAS Viya 所提供的强大分析能力。此外，公共的 REST API 使得开发人员可以从他们的应用程序中直接调用 SAS Viya 的分析功能，使用 REST 服务的任何应用程序都可以利用 SAS Viya 的分析能力。SAS Viya 的这种开放性可以帮助那些虽不熟悉 SAS 编码但是了解 Python、Lua 或 Java 语言的数据分析从业者自由利用 SAS 全面丰富且业界领先的分析功能。

（3）强大的分析能力

SAS 强大、领先的分析能力在 SAS Viya 中得到了进一步加强。SAS Viya 中的内存分析引擎除了支持大数据的快速分析之外，还可将数据置换到磁盘以支持超过可用内存大小的数据。其内存分析引擎的分布式架构可以根据需要灵活地动态扩展其处理能力，并且 SAS Viya 还提供了新的更为强大的分析算法，例如在基于 SAS Viya 的可视化数据挖掘和机器学习产品提供的自动聚类中，不需要预先指定多少个唯一类等新功能。此外，SAS Viya 还将自己的分析能力扩展到了库内（In-Database）、Hadoop 内（In-Hadoop）、流内（In-Stream）和设备内（In-Device）。

（4）云就绪

SAS Viya 支持公有云和私有云的部署，以降低平台总体成本。它能识别并利用额外的计算、内存和磁盘资源以满足分析的需求，其提供的 API 服务可提供自助访问，来预定和快速运行 SAS，以及从其他应用中调用 SAS 的能力。SAS Viya 支持工业标准的云平台即服务、Cloud Foundry，还支持快速和自动部署、灵活扩展和管理式更新。



图 1.1 SAS Viya 特征

1.1.1 SAS Viya 平台的组成部分

图 1.2 给出了 SAS Viya 平台的架构示意图。该架构包括了客户端、SAS Viya 应用和 CAS 服务器、数据层和原 SAS 系统。其中 SAS Viya 应用和 CAS 服务器为 SAS Viya 平台的组成部分。

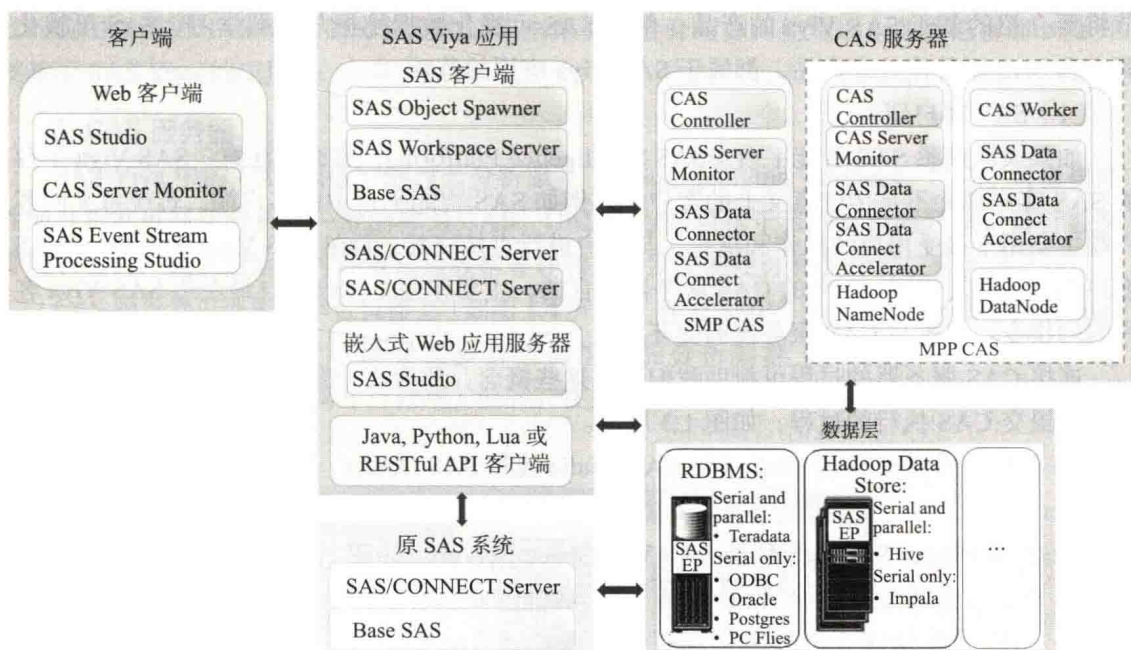


图 1.2 SAS Viya 平台架构

1. 客户端

SAS Viya 提供了基于 Web 的客户端开发和管理 SAS 应用、对 SAS Viya 环境进行监控和管理。同时，基于 SAS Viya 平台的 SAS 产品和解决方案也有自己对应的 Web 客户端。

(1) Web 客户端

SAS Studio 是 SAS 提供的 SAS Viya 平台的图形化开发界面。用户通过该界面进行 SAS 和 CAS 语言的开发、管理，提交运行并获取结果，同时还可利用 SAS Studio 提供的交互式方式指导完成分析过程。SAS 可视化数据挖掘与机器学习 (SAS Visual Data Mining and Machine Learning) 也使用 SAS Studio 进行开发。第 2 章会详细介绍 SAS Studio。

当 SAS Viya 中部署了其他 SAS 产品或解决方案时，例如 SAS 事件流处理 (SAS Event Stream Processing)、SAS 可视化调查 (SAS Visual Investigator) 等，其对应的客户端，如 SAS Event Stream Processing Studio 也包含在 Web 客户端之列。

(2) CAS 服务器监视器

CAS 服务器监视器 (CAS Server Monitor) 是 SAS 提供的监控和管理 CAS 服务器的图形化界面。本章后面将会专门介绍该功能。

2. SAS Viya 应用

SAS Viya 应用，又称为 SAS Viya 产品和解决方案，是 SAS Viya 的重要组成部分。其中包含了 Base SAS、SAS Workspace Server 和 SAS Object Spawner 组成的 SAS 客户端、SAS/CONNECT 服务器 (SAS/CONNECT Server)，以及嵌入式的 Web 应用服务器。1.1.3