

IBM SPSS Modeler 18.0 批处理用户指南

The IBM logo is centered on the page, consisting of the letters 'IBM' in a bold, black, sans-serif font. Each letter is composed of eight horizontal bars of varying lengths, creating a distinctive striped pattern.

注释

在使用本信息及其支持的产品前，请阅读第 15 页的『声明』中的信息。

产品信息

本版本适用于 IBM SPSS Modeler V18.0.0 以及所有后续发行版和修订版，直到在新版本中另有声明为止。

目录

前言	v
关于 IBM Business Analytics	v
技术支持	v
第 1 章 关于 IBM SPSS Modeler	1
IBM SPSS Modeler 产品	1
IBM SPSS Modeler	1
IBM SPSS Modeler Server	1
IBM SPSS Modeler Administration Console	2
IBM SPSS Modeler Batch	2
IBM SPSS Modeler Solution Publisher	2
IBM SPSS Collaboration and Deployment Services 的 IBM SPSS Modeler Server 适配器	2
IBM SPSS Modeler 版本	2
IBM SPSS Modeler 文档	3
SPSS Modeler Professional 文档	3
SPSS Modeler Premium 文档	4
应用程序示例	4
Demos 文件夹	4
许可证跟踪	4
第 2 章 批处理模式执行	5
批处理模式简介	5
在批处理模式下工作	5

调用软件	5
命令行参数的使用	6
批处理模式日志文件	6
在批处理模式下编写脚本	6
在批处理模式下使用参数	7
在批处理模式下对输出进行操作	8
第 3 章 命令行自变量	9
调用软件	9
命令行自变量的使用	9
系统自变量	10
参数自变量	11
服务器连接自变量	11
IBM SPSS Collaboration and Deployment Services Repository 连接自变量	12
IBM SPSS Analytic Server 连接自变量	13
组合多个参数	13
声明	15
商标	16
产品文档的条款和条件	16
索引	19

前言

IBM® SPSS® Modeler 是 IBM Corp. 企业级数据挖掘工作平台。SPSS Modeler 通过深入了解数据来帮助组织改进与客户和市民的关系。组织通过借助源自 SPSS Modeler 的洞察力可以留住优质客户，识别交叉销售机遇，吸引新客户，检测欺诈，降低风险，促进政府服务交付。

SPSS Modeler 的可视化界面让用户可以应用自己的业务专长，从而生成更强大的预测模型并缩短问题解决时间。SPSS Modeler 提供了多种建模方法，例如预测、分类、分段和关联检测算法。模型创建成功后，通过 IBM SPSS Modeler Solution Publisher，在广泛的企业内交付给决策者，或通过数据库交付。

关于 IBM Business Analytics

IBM Business Analytics 软件提供完整、一致和正确的信息，决策者依据此信息来提高业务性能。企业智能、预测分析、财务业绩和战略管理的完整产品组合，和分析应用程序一起提供对当前业绩的清晰、直接和实用的洞察力，以及预测未来结果的能力。结合丰富的行业解决方案，久经证明的实践和专业服务，各种规模的组织都能够实现最高生产力、确信地自动作出决策以及获得更好的结果。

作为此产品服务组合的组成部分，IBM SPSS Predictive Analytics 软件可帮助组织预测未来事件，并在深入了解的基础上提前行动以实现更好的业务结果。减少欺诈和降低风险时，全球的商业、政府和学术客户都依赖 IBM SPSS 技术作为吸引、保留和增加客户的竞争优势。通过在日常运营中融入 IBM SPSS 软件，组织将成为预测型企业 - 即可以指引并实现决策的自动化，以满足企业目标并实现可衡量的竞争优势。有关详细信息或要联系一位代表，请访问 <http://www.ibm.com/spss>。

技术支持

技术支持可供维护客户使用。客户可就 IBM Corp. 产品使用问题或某一受支持硬件环境的安装帮助寻求技术支持。要寻求技术支持，请访问 IBM Corp. 网站 <http://www.ibm.com/support>。请求帮助时，请准备好您自己、您的组织和支持协议的相关信息。

第 1 章 关于 IBM SPSS Modeler

IBM SPSS Modeler 是一组数据挖掘工具，通过这些工具可以采用商业技术快速建立预测性模型，并将其应用于商业活动，从而改进决策过程。IBM SPSS Modeler 参照行业标准 CRISP-DM 模型设计而成，可支持从数据到更优商业成果的整个数据挖掘过程。

IBM SPSS Modeler 提供了各种借助机器学习、人工智能和统计学的建模方法。“建模”选用板中提供的方法使您可以根据数据派生新信息，并开发预测模型。每种方法各有所长，并且最适合于解决特定类型的问题。

SPSS Modeler 可以作为独立产品购买，也可以作为客户机与 SPSS Modeler Server 一起使用。另外，还提供了很多其他选项，以下各节概述了这些选项。有关更多信息，请参阅<http://www.ibm.com/software/analytics/spss/products/modeler/>。

IBM SPSS Modeler 产品

IBM SPSS Modeler 系列产品及关联的软件包括以下各项。

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Collaboration and Deployment Services 的 IBM SPSS Modeler Server 适配器

IBM SPSS Modeler

SPSS Modeler 是具有完整功能的产品，它安装并运行于个人计算机上。您可以在本地方式作为独立产品运行 SPSS Modeler，也可以在分布方式下将其与 IBM SPSS Modeler Server 一起使用来提高大型数据集的性能。

借助 SPSS Modeler，您可以快速直接地构建准确的预测模型，而不进行编程。通过使用唯一可视界面，您可以轻松地查看数据挖掘过程。借助该产品随附的高级分析支持，您可以发现数据中先前隐藏的模式和趋势。您可以构建结果模型并了解影响结果的因素，从而利用业务机会并降低风险。

SPSS Modeler 推出了两个版本：SPSS Modeler Professional 和 SPSS Modeler Premium。请参阅主题第 2 页的『IBM SPSS Modeler 版本』，了解更多信息。

IBM SPSS Modeler Server

SPSS Modeler 使用客户端/服务器体系结构将资源集约型操作的请求分发给功能强大的服务器软件，因而使数据集的传输速度大大加快。

SPSS Modeler Server 是一个单独授权的产品，在分布分析方式下，该产品在安装了一个或多个 IBM SPSS Modeler 的服务器主机上持续运行。这种运行方式大大提高了 SPSS Modeler Server 对大型数据集的处理速度，因为在服务器上可以运行耗用内存的操作，并且无需将数据下载到客户端计算机上。IBM SPSS Modeler Server 还提供对 SQL 优化和数据库内建模功能的支持，从而在性能和自动化方面提供更多优势。

IBM SPSS Modeler Administration Console

Modeler Administration Console 是一个图形应用程序，用于管理多个 SPSS Modeler Server 配置选项，这些选项还可以通过选项文件进行配置。此应用程序提供了一个用于监视和配置 SPSS Modeler Server 安装的控制台用户界面，并且可供当前 SPSS Modeler Server 用户免费使用。应用程序只能安装在 Windows 计算机上；但是它可以管理安装在任何受支持平台上的服务器。

IBM SPSS Modeler Batch

数据挖掘通常是交互过程，因此，还可以从命令行运行 SPSS Modeler 而不需要图形用户界面。例如，您可能具有长时间运行或重复任务，并且希望在用户不进行干预的情况下执行这些任务。SPSS Modeler Batch 是该产品的一个特殊版本，可提供对 SPSS Modeler 完整分析性能的支持，而无需访问常规的用户界面。要使用 SPSS Modeler Batch，需要 SPSS Modeler Server。

IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher 是一个支持创建 SPSS Modeler 流的打包版本的工具，该版本的流可以由外部运行时引擎运行或嵌入到外部应用程序中。通过这种方式，您可以发布和部署完整的 SPSS Modeler 流以用于未安装 SPSS Modeler 的环境。SPSS Modeler Solution Publisher 作为 IBM SPSS Collaboration and Deployment Services - Scoring 服务的组成部分分发，需要单独的许可证。通过此许可证，您可以接收 SPSS Modeler Solution Publisher Runtime，它使您能够执行已发布的流。

有关 SPSS Modeler Solution Publisher 的更多信息，请参阅 IBM SPSS Collaboration and Deployment Services 文档。IBM SPSS Collaboration and Deployment Services Knowledge Center 包含名为“IBM SPSS Modeler Solution Publisher”和“IBM SPSS Analytics Toolkit”的部分。

IBM SPSS Collaboration and Deployment Services 的 IBM SPSS Modeler Server 适配器

IBM SPSS Collaboration and Deployment Services 的一些适配器使 SPSS Modeler 和 SPSS Modeler Server 能够与 IBM SPSS Collaboration and Deployment Services 存储库进行交互。通过这种方式，部署到存储库的 SPSS Modeler 流可以由多个用户共享，或者从瘦客户机应用程序 IBM SPSS Modeler Advantage 进行访问。请将适配器安装在托管存储库的系统上。

IBM SPSS Modeler 版本

SPSS Modeler 推出了下列版本。

SPSS Modeler Professional

SPSS Modeler Professional 提供处理大多数类型的结构化数据所需要的所有工具，例如 CRM 系统中跟踪的行为和交互、人口统计信息、采购行为和销售数据。

SPSS Modeler Premium

SPSS Modeler Premium 是一个单独授权的产品，它对 SPSS Modeler Professional 进行了扩展，以便后者能够处理专门的数据（例如用于实体分析或社交网络的数据）和非结构化文本数据。SPSS Modeler Premium 由下列组件组成。

IBM SPSS Modeler Entity Analytics 在 IBM SPSS Modeler 预测分析的基础上添加了额外的维度。预测性分析尝试根据过去的行为预测未来的行为，而实体分析侧重于通过在记录本身中解决身份冲突改进当前数据

的连贯性和一致性。身份可以指个人、组织、对象或可能存在不确定性的任何其他实体的身份。解析身份在多个领域十分关键，包括客户关系管理、内部欺诈检测、反洗钱以及国家和国际安保。

IBM SPSS Modeler Social Network Analysis 将关于关系的信息转换为字段，这些字段可描述个人和组社交行为的特征。使用介绍社交网络之下关系的数据，IBM SPSS Modeler Social Network Analysis 可识别影响网络中他人行为的社交领导。另外，您可以确定哪些人员受其他网络参与者的影响最大。通过将这些结果与其他度量结合，您可以创建作为预测模型基础的复杂个人档案。包含此社交信息的模型比不包含此社交信息的模型执行地更好。

IBM SPSS Modeler Text Analytics 采用了先进语言技术和自然语言处理 (NLP)，以快速处理大量非结构化文本数据，提取和组织关键概念，以及将这些概念分为各种类别。提取的概念和类别可以和现有结构化数据中进行组合（例如人口统计学），并且可用于借助 IBM SPSS Modeler 的一整套数据挖掘工具来进行建模，以此实现更好更集中的决策。

IBM SPSS Modeler 文档

可以从 SPSS Modeler 的帮助菜单中获取在线帮助格式的文档。此文档包括 SPSS Modeler 和 SPSS Modeler Server 的文档，以及《应用程序指南》（也称为“教程”）和其他支持材料。

每个产品的 PDF 格式的完整文档（包括安装指示信息）都可以 PDF 格式在独立的压缩文件夹中作为部分产品下载提供。也可以从 Web 下载 PDF 文档：<http://www.ibm.com/support/docview.wss?uid=swg27046871>。

这两种格式的文档都可以从 SPSS Modeler Knowledge Center (http://www-01.ibm.com/support/knowledgecenter/SS3RA7_18.1.0) 获得。

SPSS Modeler Professional 文档

SPSS Modeler Professional 文档套件（安装指示信息除外）如下。

- **IBM SPSS Modeler 用户指南**。使用 SPSS Modeler 的一般使用介绍，包括如何构建数据流、处理缺失值、生成 CLEM 表达式、处理项目和报告以及将用于部署的流打包为 IBM SPSS Collaboration and Deployment Services 或 IBM SPSS Modeler Advantage。
- **IBM SPSS Modeler Source、Process 和 Output 节点**。描述用于以不同格式读取、处理和输出数据的所有节点。实际上这表示所有节点而非建模节点。
- **IBM SPSS Modeler Modeling 节点**。有关用于创建数据挖掘模型的所有节点的描述。IBM SPSS Modeler 提供了各种借助机器学习、人工智能和统计学的建模方法。
- **IBM SPSS Modeler 算法指南**。描述 IBM SPSS Modeler 中使用的建模方法的数学基础。本指南仅以 PDF 格式提供。
- **IBM SPSS Modeler 应用程序指南**。本指南中的示例旨在为具体的建模方法和技术提供具有针对性的简介。还可以在“帮助”菜单中查阅本指南的在线版本。有关更多信息，请参阅第 4 页的『应用程序示例』主题。
- **IBM SPSS Modeler Python 脚本编制和自动化**。通过编写 Python 脚本实现系统自动化的相关信息，其中包括可以用于处理节点和流的属性的信息。
- **IBM SPSS Modeler 部署指南**。有关在 IBM SPSS Collaboration and Deployment Services Deployment Manager 中以处理作业的步骤形式运行 IBM SPSS Modeler 流的信息。
- **IBM SPSS Modeler CLEF 开发者指南**。CLEF 提供了将第三方案序（例如，数据处理例程或建模算法）作为节点集成到 IBM SPSS Modeler 的功能。
- **IBM SPSS Modeler 数据库内挖掘指南**。有关如何利用数据库的功能通过第三方算法来改进性能并增强分析功能的信息。

- **IBM SPSS Modeler Server 管理与性能指南。** 提供有关如何配置和管理 IBM SPSS Modeler Server 的信息。
- **IBM SPSS Modeler Administration Console 用户指南。** 有关安装和使用控制台用户界面以监视和配置 IBM SPSS Modeler Server 的信息。控制台实现为 Deployment Manager 应用程序的插件。
- **IBM SPSS Modeler CRISP-DM 指南。** 借助 CRISP-DM 方法进行 SPSS Modeler 数据挖掘的分步指南。
- **IBM SPSS Modeler Batch 用户指南。** 提供在批处理方式下使用 IBM SPSS Modeler 的完整指导，包括批处理方式执行和命令行自变量的详细信息。本指南仅以 PDF 格式提供。

SPSS Modeler Premium 文档

SPSS Modeler Premium 文档套件（安装指示信息除外）如下。

- **IBM SPSS Modeler Entity Analytics 用户指南。** 提供有关将实体分析与 SPSS Modeler 配合使用的信息，包括存储库安装和配置、实体分析节点和管理任务。
- **IBM SPSS Modeler Social Network Analysis 用户指南。** 本指南指导使用 SPSS Modeler 执行社交网络分析，包括组分析和传播分析。
- **SPSS Modeler Text Analytics 用户指南。** 提供有关将文本分析与 SPSS Modeler 配合使用的信息，包括文本挖掘节点、交互式工作台、模板和其他资源。

应用程序示例

SPSS Modeler 中的数据挖掘工具可以帮助解决很多业务和组织问题，应用程序示例将提供有关特定建模方法和技术的简要的针对性说明。此处使用的数据集比某些数据挖掘器管理的大量数据存储要小得多，但涉及的概念和方法应可扩展到实际的应用程序。

要访问示例，请单击 SPSS Modeler 中“帮助”菜单上的[应用程序示例](#)。

数据文件和样本流安装在产品安装目录下的 Demos 文件夹中。有关更多信息，请参阅『[Demos 文件夹](#)』。

数据库建模示例。 请参阅 *IBM SPSS Modeler 数据库内挖掘指南* 中的示例。

脚本编制示例。 请参阅 *IBM SPSS Modeler 脚本编写和自动化指南* 中的示例。

Demos 文件夹

应用程序申请使用的数据文件和样本流安装在产品安装目录的 Demos 文件夹下（例如：C:\Program Files\IBM\SPSS\Modeler\\Demos）。可以在 Windows 的“开始”菜单中的 IBM SPSS Modeler 程序组访问此文件夹，也可以通过单击文件 > 打开流对话框中最近访问的目录列表中的 Demos 来进行访问。

许可证跟踪

当您使用 SPSS Modeler 时，会定期跟踪和记录许可证使用情况。记录的许可证度量为 *AUTHORIZED_USER* 和 *CONCURRENT_USER*，而记录的度量类型由您所拥有的 SPSS Modeler 许可证类型决定。

日志文件可以通过 IBM License Metric Tool 处理生产，您可以通过此工具生产许可证使用情况报告。

许可证日志文件在记录 SPSS Modeler Client 日志文件的相同目录中创建（缺省情况下，目录为 %ALLUSERSPROFILE%\IBM\SPSS\Modeler\\log）。

第 2 章 批处理模式执行

批处理模式简介

数据挖掘通常是交互式过程 - 用户与数据和模型进行交互，以提高对数据及其所代表域的了解程度。但是，IBM SPSS Modeler 流也可用于在没有可视用户界面的**批处理模式**中处理数据和执行数据挖掘任务。批处理模式在没有您的干预且屏幕不显示用户界面的情况下，即可执行长期任务或重复性任务。

对于希望以批处理模式独占运行 IBM SPSS Modeler 的客户，IBM SPSS Modeler Batch 是该产品的一个特殊版本，可提供对 IBM SPSS Modeler 完整分析性能的支持，而无需访问常规的用户界面。使用批处理模式需要获得 IBM SPSS Modeler Server 许可。

适用于批处理模式的任务示例包括：

- 在后台运行比较耗时的建模训练。
- 在计划的时间（例如在夜间，此时计算机上的合成加载不会有困难）运行流。
- 针对大量数据运行数据预处理流（例如，在后台和/或在夜间）。
- 运行定期的计划任务，例如每月报告。
- 作为另一个过程的嵌入部分运行流，如评分引擎工具。

注意：可以使用相应的操作系统命令或实用程序（例如 Windows NT 系统上的 `at` 命令）来计划以批处理模式执行 IBM SPSS Modeler 操作。

在批处理模式下工作

在批处理模式下工作通常包括：

1. 使用 `clemb` 命令以批处理模式调用 IBM SPSS Modeler。
2. 连接到服务器。
3. 加载现有的流或脚本文件。
4. 执行流或脚本。

注：SPSS Modeler Batch 需要 4GB 可用内存。

一旦执行完成，您随后即可以查询在批处理模式下默认生成的日志文件，并查看图形、输出节点以及模型结果。有关这些步骤的详细信息，请参阅下列主题。

调用软件

您可以使用操作系统的命令行来如下启动 IBM SPSS Modeler：

1. 在安装了 IBM SPSS Modeler 的计算机上，打开 DOS 或命令提示符窗口。
2. 要以交互方式启动 IBM SPSS Modeler 界面，请输入 `modelerclient` 命令，然后输入所需的参数；例如：
`modelerclient -stream report.str -execute`

可用参数（标记）允许您连接到服务器、装入流、运行脚本或根据需要指定其他参数。

命令行参数的使用

为了让 IBM SPSS Modeler 能够以批处理模式打开和执行文件（如流和脚本），您需要更改启动该软件的初始命令（`clemb`）。有一些命令行参数（也称为标志）可用于执行下列操作：

- 连接服务器。
- 加载流、脚本、模型、状态、项目和输出文件。（如果您获得了 IBM SPSS Collaboration and Deployment Services Repository 许可，则可以连接存储库，并从中加载对象。）
- 指定日志文件选项。
- 设置在 IBM SPSS Modeler 中使用的默认目录。

上述所有操作都需要使用加在 `clemb` 命令后面的标志。这些标志的格式为 `-flag`，其中参数前面为短划线。例如，将初始参数 `clemb` 与标志 `-server` 一并使用时，会连接使用其他标志选项指定的服务器。

为了以批处理模式加载和执行流，您可以将 `clemb` 命令与若干其他启动标志一起使用，如 `-server`、`-stream` 和 `-execute`。下列命令会在不调用用户界面的情况下加载和执行流 `report.str`：

```
clemb -server -hostname myserver -port 80
-username dminer -password 1234 -stream report.str -execute
```

有关命令行参数的完整列表，请参阅第 9 页的第 3 章，『命令行自变量』。

- IBM SPSS Modeler 状态和脚本也可以此方式执行，它们分别使用 `-state` 和 `-script` 标记。通过为每项指定相关标志，可以加载多个状态和流。
- 可以将多个参数组合在一个命令文件之中，这些参数在启动时通过使用 `@` 符号指定。请参阅主题 第 13 页的『组合多个参数』，以了解更多信息。

批处理模式日志文件

在批处理模式下运行会生成一个日志文件。缺省情况下，此日志文件的名称为 `clem_batch.log`，但您可以使用 `-log` 标记来指定其它名称。例如，下列命令以批处理模式执行 `report.str` 并发送日志信息到 `report.log` 文件：

```
clemb -server -hostname myserver -port 80
-username dminer -password 1234 -stream report.str
-execute -log report.log
```

通常情况下，该日志文件会覆盖具有相同名称的任何现有文件，但您可以使用 `-appendlog` 标志让 IBM SPSS Modeler 追加该日志文件。还可使用 `-nolog` 标记来将日志压缩到一起。

注意：日志参数仅在批处理模式下可用。

在批处理模式下编写脚本

在其最简单的形式中，通过使用本指南中讨论的命令行参数，IBM SPSS Modeler 流的批处理模式一次只执行一个脚本。执行某个给定的流时不会明显更改其节点参数。这种方式对于每月流失报告或预测的自动化生成来说可能非常好用，但它无法处理很多高级数据挖掘器都要进行自动化的复杂过程。

例如，某个金融研究机构可能要使用各种数据或建模参数来构建一些模型，然后针对另一个数据集检验这些模型，并针对这些结果生成报告。因为此过程需要对流以及节点的创建和删除进行反复的修改，所以对该过程进行自动化需要使用脚本。通过脚本，那些原来需要用户干预的复杂过程现在就可以以批处理模式进行自动化和执行了。

以批处理模式执行脚本

1. 将 `-script` 标志放在 `clemb` 命令的最后，并指定要执行的脚本的名称。

2. 对于上述参数也要使用 `-execute` 标志，以执行指定的脚本。这样将完整运行这个独立的脚本。

例如，要加载和执行一个运行下列任务的脚本：运行生成流失评分的模型，并且这些评分存储为数据仓库的输出；您要使用下列命令：

```
clemb -server -hostname myserver -port 80
-username dminer -password 1234
-script clemscript.txt -execute
```

在批处理模式下使用参数

通过对启动 IBM SPSS Modeler 的命令行提供参数，可以修改以批处理模式执行流的结果。这些参数可以是直接在 CLEM 表达式中使用的简单参数，也可以是节点属性（也称为通道参数，此类参数用于修改流中的节点设置）。

例如，下列流从文件中选择一个数据子集，将其传递到神经网络，然后将结果发送到一个文件：



图 1. 用户界面中的流操作

字段 `Month` 的值决定了选择的数据，选择节点中的表达式为：

```
Month == '$P-mth'
```

以批处理模式运行这同一个流时，则通过在命令行中设置参数 `mth` 的值来选择相应的月份：

```
clemb -server -hostname myserver -port 80
-username dminer -password 1234
-stream predict.str -Pmth=Jan -execute
```

注意：在命令行参数中，`-P` 标记用于表示参数。

有时，流所需的命令行控件包括使用通道参数修改流中的节点设置。以下面的流为例，该流会读取一个文件，处理它的内容，然后将报告发送给另一个文件：



图 2. 用户界面中的流操作

假设您要每月生成一次报告，这期间要读取相应月份的数据，并将报告发送到一个其名称表明相关月份的文件中。您可能希望为源数据和报告设置文件名。下列命令会设置相应的通道参数并执行该流：

```
clemb -stream report.str -Porder.full_filename=APR_orders.dat
-Preport.filename=APR_report.txt -execute
```

注意：此命令不包含特定于操作系统的、用来计划每月运行一次的代码。

在批处理模式下对输出进行操作

对可视输出（如表、图形和图表）进行操作通常需要用户界面。因为批处理模式不启动 IBM SPSS Modeler 用户界面，所以输出对象会传递到一个文件中，以便您可以日后在用户界面中或另一个软件包中查看这些对象。使用可用于节点的属性（通道参数），可以控制在批处理模式期间创建的输出对象的格式和文件名。

第 3 章 命令行自变量

调用软件

您可以使用操作系统的命令行来如下启动 IBM SPSS Modeler:

1. 在安装了 IBM SPSS Modeler 的计算机上, 打开 DOS 或命令提示符窗口。
2. 要以交互方式启动 IBM SPSS Modeler 界面, 请输入 `modelerclient` 命令, 然后输入所需的参数; 例如:
`modelerclient -stream report.str -execute`

可用参数 (标记) 允许您连接到服务器、装入流、运行脚本或根据需要指定其他参数。

命令行自变量的使用

您可以将命令行自变量 (也称为标记) 附加到最初的 `modelerclient` 命令以更改对 IBM SPSS Modeler 的调用。

存在多种可用的命令行自变量类型, 本节的随后内容将对其进行描述。

表 1. 命令行自变量类型。

自变量类型	描述位置
系统自变量	有关更多信息, 请参阅第 10 页的『系统自变量』主题。
参数自变量	有关更多信息, 请参阅第 11 页的『参数自变量』主题。
服务器连接自变量	有关更多信息, 请参阅第 11 页的『服务器连接自变量』主题。
IBM SPSS Collaboration and Deployment Services Repository 连接自变量	有关更多信息, 请参阅第 12 页的『IBM SPSS Collaboration and Deployment Services Repository 连接自变量』主题。
IBM SPSS Analytic Server 连接自变量	有关更多信息, 请参阅第 13 页的『IBM SPSS Analytic Server 连接自变量』主题。

例如, 可以使用 `-server`、`-stream` 和 `-execute` 标记来连接到服务器, 然后装入并运行流, 如下所示:

```
modelerclient -server -hostname myserver -port 80 -username dminer  
-password 1234 -stream mystream.str -execute
```

请注意, 针对本地客户机安装运行时, 不需要指定服务器连接自变量。

可以用双引号括起包含空格的参数值, 例如:

```
modelerclient -stream mystream.str -Pusername="Joe User" -execute
```

还可以用此种方式执行 IBM SPSS Modeler 状态和脚本, 但要分别使用 `-state` 和 `-script` 标记。

注: 如果您在命令中使用结构化参数, 那么必须在引号之前加上反斜杠。这可以避免在解释字符串的过程中除去引号。

调试命令行自变量

要调试命令行，请使用 `modelerclient` 命令在指定所需自变量的情况下启动 IBM SPSS Modeler。这样可以验证命令是否将按期望方式执行。另外，您还可以在“会话参数”对话框（“工具”菜单 ->“设置会话参数”）中对通过命令行传递的任何参数的值进行确认。

系统自变量

下表描述可用于用户界面命令行调用的系统自变量。

表 2. 系统自变量

自变量	行为/描述
@ <commandFile>	@ 符号后跟文件名，此文件用于指定命令列表。当 <code>modelerclient</code> 遇到以 @ 开头的参数时，它将在该文件中对命令进行操作，就如同在命令行中一样。有关更多信息，请参阅第 13 页的『组合多个参数』主题。
-directory <dir>	设置缺省工作目录。在本地模式下，该目录将同时用于数据操作和输出。示例: <code>-directory c:/</code> or <code>-directory c:\</code>
-server_directory <dir>	为数据设置缺省服务器目录。通过 <code>-directory</code> 标记指定的工作目录将用于输出。
-execute	在启动后执行启动时所装入的流、状态或脚本。如果在流或状态之外还装入了脚本，那么脚本将单独执行。
-stream <stream>	启动时装入指定的流。可以指定多个流，但是最后一个指定的流将被设置为当前流。
-script <script>	启动时装入指定的独立脚本。如下所述，除流或状态之外，此标记还可用于指定脚本，但在启动时仅可装入一个脚本。
-model <model>	在启动时装入指定的已生成模型（.gm 格式的文件）。
-state <state>	在启动时，装入指定的已保存状态。
-project <project>	装入指定工程。在启动时仅可装入一个工程。
-output <output>	在启动时装入已保存的输出项目（.cou 格式的文件）。
-help	显示命令行自变量列表。指定此选项后，将忽略所有其他参数并显示帮助屏幕。
-P <name>=<value>	用于设置启动参数。另外，还可用于设置节点属性（槽参数）。

注意：也可以在用户界面中设置缺省目录。要访问上述选项，请在“文件”菜单中选择**设置工作目录**或**设置服务器目录**。

装入多个文件

命令行模式下，您可以通过在启动时重复输入每个装入对象的相关参数来装入多个流、状态和输出。例如，要装入和运行两个称为 `report.str` 和 `train.str` 的流，您可以使用如下命令：

```
modelerclient -stream report.str -stream train.str -execute
```

从 IBM SPSS Collaboration and Deployment Services Repository 装入对象

因为可以从某个文件或 IBM SPSS Collaboration and Deployment Services Repository（如果已获许可）装入特定对象，可以使用文件名前缀 `spsscr:` 以及选择性地使用 `file:`（对于磁盘上的对象）来指示 IBM SPSS Modeler 在什么位置查找对象。前缀可与以下标记配合使用：

- `-stream`
- `-script`
- `-output`
- `-model`

- -project

您可以使用前缀创建 URI 以指定对象的位置，例如 `-stream "spsscr:///folder_1/scoring_stream.str"`。如果指定了 `spsscr:` 前缀，那么要求已在同一命令中指定了有效的 IBM SPSS Collaboration and Deployment Services Repository 连接。因此，完整的命令应形如以下的示例：

```
modelerclient -spsscr_hostname myhost -spsscr_port 8080
-spsscr_username myusername -spsscr_password mypassword
-stream "spsscr:///folder_1/scoring_stream.str" -execute
```

注意，在命令行中 **必须**使用 URI。不支持像 `REPOSITORY_PATH` 这样的简单路径。（此种路径仅适用于脚本。）

参数自变量

参数可用作在 IBM SPSS Modeler 的命令行执行期间的标记。在命令行自变量中，`-P` 标记用于表示形如 `-P <name>=<value>` 的参数。

形式参数可以是：

- **简单参数**（即，直接在 CLEM 表达式中使用的参数）。
- **槽参数**，也称为**节点属性**。此类参数用于修改流中各个节点的设置。
- **命令行自变量**，用于更改对 IBM SPSS Modeler 的调用。

例如，您可以提供数据源用户名和密码作为命令行标志，如下所示：

```
modelerclient -stream response.str -P:databasenode.datasource="{\"ORA 10gR2\", user1, mypsw, true}"
```

其格式与 `databasenode` 节点属性的 `datasource` 参数相同。

注：如果已命名节点，那么您必须将节点名括在双引号中并使用反斜杠对引号进行转义。例如，如果以上示例中数据源节点的名称为 `Source_ABC`，那么此条目将如下所示：

```
modelerclient -stream response.str -P:databasenode.\"Source_ABC\".datasource="{\"ORA 10gR2\", user1, mypsw, true}"
```

标识结构化参数的引号的前面也需要一个反斜杠，如以下 `TM1` 数据源示例中所示：

```
clemb -server -hostname 9.115.21.169 -port 28053 -username administrator
-execute -stream C:\Share\TM1_Script.str -P:tmlimport.pm_host="http://9.115.21.163:9510/pmhub/pm"
-P:tmlimport.tm1_connection="{\"SData\", \"\", \"admin\", \"apple\"}"
-P:tmlimport.selected_view="{\"SalesPriorCube\", \"salesmargin%\"}"
```

服务器连接自变量

`-server` 标记指示 IBM SPSS Modeler 应连接到公共服务器，标记 `-hostname`、`-use_ssl`、`-port`、`-username`、`-password` 和 `-domain` 用于指示 IBM SPSS Modeler 如何连接到公共服务器。如果未指定 `-server` 参数，那么使用缺省服务器。

示例

连接到公共服务器：

```
modelerclient -server -hostname myserver -port 80 -username dminer
-password 1234 -stream mystream.str -execute
```

连接到服务器集群：

```
modelerclient -server -cluster "QA Machines" \
-spsscr_hostname pes_host -spsscr_port 8080 \
-spsscr_username asmith -spsscr_epassword xyz
```

请注意，连接到服务器集群需要通过在整个 IBM SPSS Collaboration and Deployment Services 中使用过程协调器，因此 `-cluster` 参数必须与存储库连接选项 (`spsscr_*`) 结合使用。有关更多信息，请参阅『IBM SPSS Collaboration and Deployment Services Repository 连接自变量』主题。

表 3. 服务器连接自变量.

自变量	行为/描述
<code>-server</code>	以服务器模式运行 IBM SPSS Modeler，同时使用标志 <code>-hostname</code> 、 <code>-port</code> 、 <code>-username</code> 、 <code>-password</code> 和 <code>-domain</code> 连接到公共服务器。
<code>-hostname<name></code>	服务器的主机名。仅在服务器方式下可用。
<code>-use_ssl</code>	指定连接应使用 SSL（安全套接字层）。此标记为可选项，缺省设置为不使用 SSL。
<code>-port<number></code>	所指定服务器的端口号。仅在服务器方式下可用。
<code>-cluster<name></code>	指定指向服务器集群（而不是已命名的服务器）的连接；此参数可用于替代 <code>hostname</code> 、 <code>port</code> 和 <code>use_ssl</code> 参数。名称为聚类名，或标识 IBM SPSS Collaboration and Deployment Services Repository 中聚类的唯一 URI。服务器集群由 IBM SPSS Collaboration and Deployment Services 中的过程协调器管理。有关更多信息，请参阅『IBM SPSS Collaboration and Deployment Services Repository 连接自变量』主题。
<code>-username<name></code>	这是用于登录服务器的用户名。仅在服务器方式下可用。
<code>-password<password></code>	这是用于登录服务器的密码。仅在服务器方式下可用。注意：如果未使用 <code>-password</code> 自变量，那么系统将提示您输入密码。
<code>-epassword<encodedpasswordstring></code>	这是用于登录服务器的经过编码的密码。仅在服务器方式下可用。注意：可以通过 IBM SPSS Modeler 应用程序的“工具”菜单生成经过编码的密码。
<code>-domain<name></code>	这是用于登录服务器的域。仅在服务器方式下可用。
<code>-P <name>=<value></code>	用于设置启动参数。另外，还可用于设置节点属性（槽参数）。

IBM SPSS Collaboration and Deployment Services Repository 连接自变量

如果想通过命令行来存储或检索 IBM SPSS Collaboration and Deployment Services 中的对象，那么必须指定一个指向该 IBM SPSS Collaboration and Deployment Services Repository 的有效连接。例如：

```
modelerclient -spsscr_hostname myhost -spsscr_port 8080
-spsscr_username myusername -spsscr_password mypassword
-stream "spsscr:///folder_1/scoring_stream.str" -execute
```

下表列出了可用于建立连接的自变量。

表 4. IBM SPSS Collaboration and Deployment Services Repository 连接自变量

自变量	行为/描述
<code>-spsscr_hostname <hostname or IP address></code>	安装 IBM SPSS Collaboration and Deployment Services Repository 的服务器的主机名或 IP 地址。
<code>-spsscr_port <number></code>	IBM SPSS Collaboration and Deployment Services Repository 接受连接的端口号（通常，缺省值为 8080）。
<code>-spsscr_use_ssl</code>	指定连接应使用 SSL（安全套接字层）。此标记为可选项，缺省设置为不使用 SSL。

表 4. IBM SPSS Collaboration and Deployment Services Repository 连接自变量 (续)

自变量	行为/描述
-spsscr_username<name>	登录到 IBM SPSS Collaboration and Deployment Services Repository 的用户名。
-spsscr_password<password>	登录到 IBM SPSS Collaboration and Deployment Services Repository 的密码。
-spsscr_epassword<encoded password>	登录到 IBM SPSS Collaboration and Deployment Services Repository 的加密密码。
-spsscr_domain<name>	登录到 IBM SPSS Collaboration and Deployment Services Repository 所使用的域。此标志是可选项，除非您使用 LDAP 或 Active Directory 进行登录，否则请勿使用此标志。

IBM SPSS Analytic Server 连接自变量

如果想要通过命令行从 IBM SPSS Analytic Server 存储或检索对象，必须指定到 IBM SPSS Analytic Server 的有效连接。

注：Analytic Server 的位置是从 SPSS Modeler Server 获取的，并且无法在客户机上进行更改。

下表列出了可用于建立连接的自变量。

表 5. IBM SPSS Analytic Server 连接自变量

自变量	行为/描述
-analytic_server_username	用于登录到 IBM SPSS Analytic Server 的用户名。
-analytic_server_password	用于登录到 IBM SPSS Analytic Server 的密码。
-analytic_server_epassword	用于登录到 IBM SPSS Analytic Server 的经过编码的密码。
-analytic_server_credential	用于登录到 IBM SPSS Analytic Server 的凭证。

组合多个参数

通过使用后跟文件名的 @ 符号，可以在调用时指定的单个命令文件中组合使用多个自变量。这使您能够缩短命令行调用，并且可以克服操作系统在命令长度方面的限制。例如，以下启动命令使用了 <commandFileName > 的引用文件中的指定参数。

```
modelerclient @<commandFileName>
```

如果需用空格，那么请用引号将命令文件的文件名和路径括起来，如下所示：

```
modelerclient @ "C:\Program Files\IBM\SPSS\Modeler\mn\scripts\my_command_file.txt"
```

命令文件可以包含先前在启动时逐个指定的所有自变量，并且每个自变量各占一行。例如：

```
-stream report.str
-Porder.full_filename=APR_orders.dat
-Preport.filename=APR_report.txt
-execute
```

当写入和引用命令文件时，必须遵循以下限制：

- 每条命令各占一行。
- 不要在命令文件中嵌入 @CommandFile 参数。

声明

本信息是为在美国提供的产品和服务编写的。IBM 可能以其他语言提供本材料。然而，您可能需要拥有产品或产品版本的该语言副本才能进行访问。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。但是，评估和验证任何非 IBM 产品、程序或服务，则由用户自行负责。

IBM 公司可能已拥有或正在申请与本文档内容有关的各项专利。提供本文档并未授予用户使用这些专利的任何许可。您可以用书面方式将许可查询寄往：

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

有关双字节 (DBCS) 信息的许可查询，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

International Business Machines Corporation“按现状”提供本出版物，不附有任何种类的（无论是明示的还是暗含的）保证，包括但不限于暗含的有关非侵权、适销和适用于某种特定用途的保证。某些管辖区域在某些交易中不允许免除明示或暗含的保证。因此本条款可能不适用于您。

本信息中可能包含技术方面不够准确的地方或印刷错误。此处的信息将定期更改；这些更改将编入本资料的新版本中。IBM 可以随时对本出版物中描述的产品进行改进和/或更改，而不另行通知。

本信息中对非 IBM Web 站点的任何引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

本程序的被许可方如果要了解有关程序的信息以达到如下目的：(i) 使其能够在独立创建的程序和其它程序（包括本程序）之间进行信息交换，以及 (ii) 使其能够对已经交换的信息进行相互使用，请与下列地址联系：

IBM Director of Licensing
IBM Corporation

North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

只要遵守适当的条件和条款，包括某些情形下的一定数量的付费，都可获得这方面的信息。

本资料中描述的许可程序及其所有可用的许可资料均由 IBM 依据 IBM 客户协议、IBM 国际软件许可协议或任何同等协议中的条款提供。

所引用的性能数据和客户示例仅作参考用途。实际的性能结果可能会因特定的配置和运营条件而异。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

关于 IBM 未来方向或意向的声明都可随时更改或收回，而不另行通知，它们仅仅表示了目标和意愿而已。

本信息包含在日常业务操作中使用的数据和报告的示例。为了尽可能完整地说明这些示例，示例中可能会包括个人、公司、品牌和产品的名称。所有这些名字都是虚构的，若与实际个人或业务企业相似，纯属巧合。

商标

IBM、IBM 徽标和 ibm.com 是 International Business Machines Corp. 在全球许多行政管辖地区的商标或注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。IBM 商标的最新列表可从 Web 上的“Copyright and trademark information”处获得，网址为：www.ibm.com/legal/copytrade.shtml。

Adobe、Adobe 徽标、PostScript 和 PostScript 徽标是 Adobe Systems Incorporated 在美国和/或其他国家或地区的注册商标或商标。

Intel、Intel 徽标、Intel Inside、Intel Inside 徽标、Intel Centrino、Intel Centrino 徽标、Celeron、Intel Xeon、Intel SpeedStep、Itanium 和 Pentium 是 Intel Corporation 或其子公司在美国和其他国家或地区的商标或注册商标。

Linux 是 Linus Torvalds 在美国和/或其他国家或地区的注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国和/或其他国家或地区的商标。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

Java 和所有基于 Java 的商标和徽标是 Oracle 和/或其子公司的商标或注册商标。

产品文档的条款和条件

只有遵守以下条款和条件才会授予使用这些出版物的许可权。

适用性

除用于 IBM Web 站点的任何条款外，还需要遵守这些条款和条件。

个人使用

您可以为了个人使用而非商业性使用复制这些出版物，但前提是保留所有专有权声明。没有 IBM 的明确同意，您不能对这些出版物或者其中的任何部分进行分发、展示或从中派生内容。

商业使用

您仅可在贵公司内部复制、分发和显示这些出版物，但前提是保留所有专有权声明。未经 IBM 的明确许可，您不得制作这些出版物的演绎作品，也不得在贵公司外部复制、分发或显示这些出版物或其部分出版物。

权利

除非本许可权中明确授予，否则不得授予对这些出版物或其中包含的任何信息、数据、软件或其他知识产权的任何许可权、许可证或权利，无论明示的还是暗含的。

只要 IBM 认为这些出版物的使用会损害其利益或者 IBM 判定未正确遵守上述指示信息，IBM 将有权撤销此处授予的许可权。

您不得下载、出口或再出口这些信息，除非完全遵守所有适用的法律和法规（包括所有的美国出口法律和法规）。

IBM 对这些出版物的内容不作任何保证。本出版物以“按现状”的基础提供，不附有任何形式的（无论是明示的，还是暗含的）保证，包括但不限于暗含的有关适销性、非侵权以及适用于某特定用途的保证。

索引

[A]

安全性
 加密密码 11

[B]

标记
 命令行参数 6
标志
 命令行自变量 9
 组合多个标志 13

[F]

服务器
 命令行自变量 11

[J]

脚本编制
 批处理模式 6

[L]

流
 计划 5

[M]

密码
 已编码 11
命令行
 参数 11
 参数列表 10, 11, 12, 13
 多个参数 13
 运行 IBM SPSS Modeler 6, 9

[P]

批处理模式
 参数使用 6
 对数 6
 计划流 5
 脚本编制 6
 输出 8
 parameters 7

[R]

日志文件 6

[S]

示例
 概述 4
 应用程序指南 3

[W]

文档 3

[X]

系统
 命令行自变量 10

[Y]

应用程序示例 3

[Z]

自变量
 服务器连接 11
 命令文件 13
 系统 10
 IBM SPSS Analytic Server Repository
 连接 13
 IBM SPSS Collaboration and
 Deployment Services Repository 连接
 12

I
IBM SPSS Analytic Server Repository
 命令行参数 13
IBM SPSS Collaboration and Deployment
 Services Repository
 命令行自变量 12
IBM SPSS Modeler 1
 从命令行运行 6, 9
 文档 3
IBM SPSS Modeler Server 1



Printed in China