

IBM SPSS Statistics – Essentials for Python: 安装说明 Linux

以下说明用于在 Linux® 操作系统上安装 IBM® SPSS® Statistics – Essentials for Python。

概述

IBM® SPSS® Statistics – Essentials for Python 为您提供了开发与 IBM® SPSS® Statistics 一起使用的自定义 Python 应用程序所需的一切工具。它包括以下内容：

- ▶ 用于 SPSS Statistics 20 的 IBM® SPSS® Statistics – Integration Plug-In for Python
- ▶ 一组以 Python 实现的自定义过程，以供在 SPSS Statistics 中使用

安装 IBM SPSS Statistics 应用程序

IBM® SPSS® Statistics – Essentials for Python 针对以下应用程序设计：

- 用于 Linux 的 IBM SPSS Statistics

没有其他的操作系统和硬件要求。随 Essentials for Python 安装的组件可通过任何有效的 IBM® SPSS® Statistics 许可证运行。

如果您尚未完成这一准备工作，请按照随软件提供的说明在将要安装 Essentials for Python 的计算机上安装一个 SPSS Statistics 应用程序。

下载和安装 Python 2.7

IBM® SPSS® Statistics – Essentials for Python 的版本 20 需要 Python 版本 2.7（建议安装版本 2.7.1）。在将要安装 Essentials for Python 的计算机上安装 Python。它可从以下站点获得：<http://www.python.org>。您可以在命令提示符处输入 `python -V`，以检查您的计算机上的默认 Python 版本。

- ▶ 你的分布存储库可能包括 Python 2.7。如果这样，您可以使用分布的标准程序包管理器（如 RPM Package Manager 或 Synaptic Package Manager）安装 Python。不过，请注意，该 Python 可能未使用 IBM® SPSS® Statistics – Integration Plug-In for Python 所需的配置选项进行构建。因此，如果您从分布存储库中安装 Python，插件可能无法工作。这同样适用于在计算机上预先安装的 Python 版本。安装 Python 的替代方法是下载 Python 2.7 的源代码，并自行构建。该建议方法可以确保采用正确的配置选项来构建 Python。说明如下。

从源代码构建和安装 Python

如果您选择自行构建 Python，则基本步骤如下。不过，建议您阅读随 Python 源代码附带的 README 文件。

注意：Python 可以在命令语法或脚本编写中使用。脚本编写功能需要 Tkinter 模块。如果您从源代码构建 Python，默认情况下不会在 Python 中构建 Tkinter 支持。要包含 Tkinter 支持，应在构建 Python 之前安装 Tk/Tcl 与 BLT 开发程序包和头文件。这些程序包的名称以 `-dev` 或 `-devel` 结尾。例如，在 Debian 上，您可以输入 `apt-get install tk-dev` 来安装 `tk-dev` 程序包及其所有相关性。在安装 Tk/Tcl 与 BLT 开发程序包之后，您可以继续以下步骤。

- ▶ 创建临时目录用于解压缩和拆开 Python 源。例如，在命令提示符处键入：

```
mkdir ~/pysource
```

- ▶ 从站点 <http://www.python.org> 下载源代码，并将其保存到临时目录。
- ▶ 切换到临时目录。例如，在命令提示符处键入：

```
cd ~/pysource
```

- ▶ 解压缩并拆开 Python 源到临时目录。例如，在命令提示符处键入：

```
tar xzf Python-2.7.1.tgz
```

- ▶ 切换到源目录。例如，在命令提示符处键入：

```
cd Python-2.7.1
```

- ▶ 在继续之前，请阅读此目录中的 README 文件。

注意：要将 Python 安装到默认目录，您必须通过以 `root` 身份登录或使用 `sudo` 命令以 `root` 身份运行以下步骤。

- ▶ 配置、构建和安装 Python。确保使用 `--enable-shared` 和 `--enable-unicode=ucs2` 参数来配置 Python。例如，在命令提示符处键入：

```
./configure --enable-shared --enable-unicode=ucs2 && make && make install
```

下载和安装 IBM SPSS Statistics – Essentials for Python

确保使用与您计算机上 IBM® SPSS® Statistics 版本兼容的 IBM® SPSS® Statistics – Essentials for Python 版本。在 SPSS Statistics 的主版本中（如 20），您必须使用具有相同主版本的 Essentials for Python 版本。在安装较新版本前，不需要卸载 Essentials for Python 的先前版本。

对于在分布式模式中工作的用户（使用 SPSS Statistics Server），请同时在客户端和服务端计算机上安装 Essentials for Python。

- ▶ 下载 Essentials for Python 的版本 20，它可从以下站点获得：SPSS 社区 <http://www.ibm.com/developerworks/spsdevcentral>。
- ▶ 启动终端应用程序。
- ▶ 更改为您下载 Essentials for Python 的目录。
- ▶ 在命令提示符处键入：

```
./<<filename>>
```

其中 <<filename>> 为下载文件的名称。

注意：您必须通过以 root 身份登录或使用 `sudo` 命令以 root 身份运行先前的命令。

- ▶ 按照屏幕上显示的说明进行操作。当提示 Python 的位置时，请注意，这是 Python 2.7 的安装位置，例如 `/opt/Python2.7`，而不是 Python 2.7 可执行文件的位置。

无提示安装

作为上述手动安装的替代方法，您还可以运行无提示安装。这对于需要安装到多个最终用户的网络管理员最为有用。要运行无提示安装，请执行下列操作：

- ▶ 启动终端应用程序。
- ▶ 更改为您下载 Essentials for Python 的目录。
- ▶ 使用文本编辑器，创建名为 `installer.properties` 的响应文件。
- ▶ 将下列属性和相关值添加到响应文件中：

```
INSTALLER_UI=silent
USER_SPSS_HOME=<IBM SPSS Statistics location>
USER_PYTHON_HOME=<Python 2.7 home directory>
LICENSE_ACCEPTED=true
```

其中 <IBM SPSS Statistics location> 为 IBM SPSS Statistics 的安装位置，<Python 2.7 home directory> 为 Python 2.7 的安装位置。例如：

```
USER_SPSS_HOME=/opt/IBM/SPSS/Statistics/20
USER_PYTHON_HOME=/opt/Python2.7
```

注意 `LICENSE_ACCEPTED=true` 表示接受许可协议。

- ▶ 将 `installer.properties` 保存到包含 Essentials for Python 的 `.bin` 文件的目录中，并切换到此目录。
- ▶ 通过如下命令来运行安装程序：

```
./<installer_name>
```

其中 <installer_name> 为 Essentials for Python 的 `.bin` 文件的名称。注意：您必须通过以 root 身份登录或使用 `sudo` 命令以 root 身份运行先前的命令。

注意：要使用其他响应文件（不同于 `installer.properties`），通过如下命令运行安装程序：

```
./<installer_name> -f <response file name>
```

在开始使用 IBM SPSS Statistics – Integration Plug-In for Python 之前

安装 IBM® SPSS® Statistics – Essentials for Python 后，您可以开始使用 IBM® SPSS® Statistics – Integration Plug-In for Python 开发 Python 应用程序。“IBM® SPSS® Statistics 帮助”系统中的“Integration Plug-In for Python”主题提供了有关插件的完整文档。它还以两个 PDF 文件形式提供（分别针对插件安装的两个接口），并可从 SPSS Statistics 帮助菜单上的“可编程性 > Python 插件和可编程性 > 脚本编写”获得。

从外部 Python 过程中运行 IBM SPSS Statistics

您可以从任何外部 Python 过程中运行采用 IBM® SPSS® Statistics 处理程序的 Python 程序，例如 Python IDE 或 Python 翻译器。在此模式下，Python 程序无需 SPSS Statistics 客户端的相关实例即可启动 SPSS Statistics 处理器的新实例。可以通过所选的 Python IDE，使用此模式调试您的 Python 程序。

要从外部 Python 过程中驱动 SPSS Statistics 处理程序，应将以下位置添加到 LD_LIBRARY_PATH 环境变量中：

```
LD_LIBRARY_PATH=<PYTHON_HOME>/lib:<SPSS_HOME>/lib:$LD_LIBRARY_PATH
```

其中 <PYTHON_HOME> 为 Python 的安装位置 — 通常为 /usr/local—，<SPSS_HOME> 为 SPSS Statistics 的安装位置 — 例如 /opt/IBM/SPSS/Statistics/20。

下载和安装 IBM SPSS Statistics 的辅助 Python 模块

IBM SPSS 创建了大量 Python 模块，这些模块在 IBM® SPSS® Statistics – Integration Plug-In for Python 提供的功能上进行构建或有所扩展。这些可选模块需要 Integration Plug-In for Python 才能运行。模块 spssdata、spssaux、namedtuple 和 extension 随 IBM® SPSS® Statistics – Essentials for Python 附带。其他模块可从以下站点下载：SPSS 社区 <http://www.ibm.com/developerworks/spssdevcentral>。

您可能需要阅读文章如何使用下载的 Python 模块，该文章也可从以下站点获得：SPSS 社区。此外，您可能需要下载辅助模块的更新版本，以便利用 IBM® SPSS® Statistics 20 提供的新功能。请检查 SPSS 社区 以获得更新。

运行多个版本的 IBM SPSS Statistics 和 IBM SPSS Statistics – Integration Plug-In for Python

您可以在一台计算机上拥有多个版本的 IBM® SPSS® Statistics，并分别与不同 IBM® SPSS® Statistics – Integration Plug-In for Python 版本配合使用。有关使用多版本的详细信息，请参阅插件相关文档。

卸载 IBM SPSS Statistics – Essentials for Python 组件

- ▶ 启动终端程序。
- ▶ 将目录更改为 IBM® SPSS® Statistics 安装目录中的 `Uninstall_IBM_SPSS_Statistics_Essentials_for_Python_20`。
- ▶ 在命令提示符处键入：

```
./Uninstall_IBM_SPSS_Statistics_Essentials_for_Python_20
```

重要事项：您必须有权删除安装目录，否则卸载过程将失败。

通过 Essentials for Python 包装的自定义 Python 过程。

IBM® SPSS® Statistics – Essentials for Python 包含一组以 Python 实现的自定义过程，提供在内置 SPSS Statistics 过程中的可用功能以外的功能。所有自定义过程，除了 FUZZY 以外，均可在安装 Essentials for Python 后从 SPSS Statistics 菜单中使用。所有自定义过程，包括 FUZZY，均可从 SPSS Statistics 命令语法中使用。下表提供了菜单位置、相关命令名称，以及每个自定义过程的简短说明。这些命令的运行方式与任何内置命令（例如 FREQUENCIES）相同。

菜单位置	命令名称	描述
无	FUZZY	执行精确或模糊大小写控制匹配。
文件 > 收集变量信息	GATHERMD	从多个数据集构建变量信息数据集。
数据 > 比较数据集	SPSSINC COMPARE DATASETS	比较两个打开的数据集。
转换 > 创建虚拟变量	SPSSINC CREATE DUMMIES	创建一组代表变量值的虚拟变量。
实用程序 > 合并查看器表	SPSSINC MERGE TABLES	将查看器某个枢轴表的内容与其他枢轴表的内容进行合并。
实用程序 > 修改输出标题	SPSSINC MODIFY OUTPUT	修改输出标题。
实用程序 > 修改表外观	SPSSINC MODIFY TABLES	修改枢轴表的外观。
转换 > 可编程性转换	SPSSINC TRANS	对个案数据应用 Python 函数。
分析 > 描述统计 > TURF 分析	SPSSINC TURF	执行 TURF（总到达率及频率）分析。

附注

- 从菜单上可访问的每个过程的帮助信息，可以通过相关对话框上的帮助按钮进行访问。不过，这些帮助并未与 SPSS Statistics 帮助系统集成在一起。
- 要获得上面列出的每个命令的完整语法帮助，可在执行命令时包含 `/HELP` 子命令，例如：

```
SPSSINC COMPARE DATASETS /HELP.
```

不过，命令语法帮助并未与 SPSS Statistics 帮助系统集成在一起，也没有包含在命令语法参考中。

- 如果在您的 IBM® SPSS® Statistics 产品中不存在“分析 > 描述统计”菜单，请查看“自定义”菜单以获得“TURF 分析”分析对话框。

- 在 SPSS Statistics 中，对话框通过“自定义对话框生成器”进行创建。您可以从“自定义对话框生成器”（从“实用程序 > 自定义对话框 > 自定义对话框生成器”访问）中查看和设计任何对话框和/或进行自定义。要查看对话框设计，从“自定义对话框生成器”中选择“文件 > 打开已安装的”。
- 语法命令实现为扩展命令。这些扩展命令的实现代码（Python 模块）和 XML 规范文件位于 SPSS Statistics Server 安装目录/Library/Application Support/IBM/SPSS/Statistics/20/extensions 目录下的extensions目录中。
如果已使用 SPSS_EXTENSIONS_PATH 环境变量为扩展命令指定了备用位置，则文件将位于该变量中的第一个可写位置而不是 extensions 目录中。
- 在 Essentials for Python 未包含的其他扩展命令可从以下站点下载： SPSS 社区。
还可以在此处下载通过 Essentials for Python 包装的 Python 过程的新版本。