

IBM SPSS Modeler 扩展

IBM

注释

在使用本资料及其支持的产品之前，请阅读第 51 页的『声明』中的信息。

产品信息

此版本适用于 IBM SPSS Modeler V18.0.0，及所有后续发行版和修订版，除非在新版本中另有说明。

目录

第 1 章 支持的语言	1
R	1
Python for Spark	1
使用 Python for Spark 编写脚本	1
"扩展"节点	11
"扩展导出"节点	11
"扩展输出"节点	13
"扩展模型"节点	15
扩展模型块	16
"扩展变换"节点	18
"扩展导入"节点	19
第 2 章 扩展	21
扩展中心	21
"探索"选项卡	22
"已安装"选项卡	22
设置	23
扩展详细信息	23
安装本地扩展束	23
扩展的安装位置	24

必需的 R 软件包	24
创建和管理定制节点	24
定制对话框构建器布局	25
构建定制节点对话框	26
对话框属性	26
在对话框画布上布置控件	27
构建脚本模板	27
预览定制节点对话框	28
控件类型	28
扩展属性	44
管理定制节点对话框	46
创建定制节点对话框的本地化版本	48
使用 Python for Spark 导入和导出数据	49
使用 R 导入和导出数据	49
声明	51
商标	52
产品文档的条款和条件	52
索引	55

第 1 章 支持的语言

IBM® SPSS® Modeler 支持 R 和 Apache Spark (通过 Python) 。请参阅以下部分以获取更多信息。

R

IBM SPSS Modeler 支持 R。

允许的语法

- 在各种"扩展"节点的语法选项卡上的语法字段中，仅允许使用 R 能够识别的语句和函数。
- 对于"扩展变换"节点和扩展模型块，通过 R 脚本（批处理）传递数据。因此，用于模型评分节点和流程节点的 R 脚本不应包括跨数据行或组合数据行的操作，例如排序或汇总。实施此限制的目的是确保能够在 Hadoop 环境中以及数据库内挖掘期间分割数据。"扩展输出"和"扩展模型构建"节点没有此限制。
- 在"扩展变换"节点和扩展模型块中增加非批量数据传输方式，这意味着您可以在 SPSS Modeler Server 中跨越或组合数据行。
- 所有 R 节点均视为独立的全局 R 环境。因此，在两个单独的 R 节点中使用 library 函数需要在两个 R 脚本中装入 R 库。
- 要显示 R 脚本中定义的 R 对象的值，必须调用 printing 函数。例如，要显示称为 data 的 R 对象的值，请在 R 脚本中包含以下行：

```
print(data)
```
- 由于 IBM SPSS Modeler 将 R setwd 函数用于控制 R 脚本输出文件的路径，因此您无法在 R 脚本中包含对该函数的调用。
- 如果在 R 脚本中使用定义用于 CLEM 表达式和脚本编制的流参数，那么无法识别这些流参数。

Python for Spark

IBM SPSS Modeler 支持 Apache Spark 的 Python 脚本。

注：

- Python 节点取决于 Spark 环境。
- Python 脚本必须使用 Spark API，因为将以 Spark DataFrame 形式表示数据。
- V17.1 中创建的旧节点仍将仅针对 IBM SPSS Analytic Server 运行（数据源自 IBM SPSS Analytic Server 源节点并且尚未抽取到 IBM SPSS Modeler 服务器）。V18.0 或更高版本中创建的新 Python 和"定制对话框构建器"可针对 IBM SPSS Modeler 服务器运行。
- 安装 Python 时，请确保所有用户都有权访问 Python 安装。
- 如果要使用"机器学习库 (MLlib)"，必须安装包含 NumPy 的 Python 版本。然后，必须配置 IBM SPSS Modeler 服务器（或 IBM SPSS Modeler 客户机中的本地服务器）以使用 Python 安装。有关详细信息，请参阅『使用 Python for Spark 编写脚本』。

使用 Python for Spark 编写脚本

在使用 Apache Spark 框架来处理数据时，IBM SPSS Modeler 可执行 Python 脚本。此文档针对提供的接口提供 Python API 描述。

IBM SPSS Modeler 安装包含 Spark 分发版（例如，IBM SPSS Modeler 18.1 包含 Spark 2.1.0）。如果您更喜欢使用不同版本的 Spark，那么可以通过将 `eas_spark_path` 和 `eas_spark_version` 参数添加到 `options.cfg` 来对它进行配置。`eas_spark_version` 参数可能值的值为 1.x 或 2.x。例如：

```
eas_spark_path, "C:/spark_1.5"  
eas_spark_version, "1.x"
```

先决条件

- 如果您计划针对 IBM SPSS Analytic Server 执行 Python/Spark 脚本，必须具有到 Analytic Server 的连接，并且 Analytic Server 必须有权访问 Apache Spark 的兼容安装。请参阅 IBM SPSS Analytic Server 文档以获取有关使用 Apache Spark 作为执行引擎的详细信息。
- 如果您计划针对 IBM SPSS Modeler 服务器（或 IBM SPSS Modeler 客户机包含的本地服务器，需要 Windows 64、Linux 64 或 Mac64）执行 Python/Spark 脚本，那么不再需要安装 Python 并编辑 `options.cfg` 以使用您的 Python 安装。启动 V18.1，IBM SPSS Modeler 现在将包含 Python 分发版。但是，如果您需要未随附在缺省 IBM SPSS Modeler Python 分发版中的特定模块，那么可以转至 `<Modeler_installation_directory>/python` 并安装附加软件包。

即使 Python 分发版现在包含在 IBM SPSS Modeler 中，如果需要，您仍然可以通过向 `options.cfg` 添加以下选项，像前发行版一样指向自己的 Python 安装：

```
# Set to the full path to the python executable  
(including the executable name) to enable use of PySpark.  
eas_pyspark_python_path, ""
```

例如：

```
eas_pyspark_python_path, "C:/Your_Python_Install/python.exe"
```

- 如果在 Linux 上使用 IBM SPSS Modeler 客户机，请为存储定制对话框构建器扩展创建文件夹。然后，将以下命令添加到 `modelerclient.sh` 文件，其中 `<your_CDB_folder>` 是您创建的文件夹：

```
export IBM_SPSS_MODELER_EXTENSION_PATH=<your_CDB_folder>
```

IBM SPSS Analytic Server 上下文对象

Analytic Server 上下文对象定义 Python/Spark 脚本的执行上下文。针对 IBM SPSS Modeler 服务器运行时，上下文对象用于 IBM SPSS Modeler 服务器安装中包含的 Analytic Server 的嵌入版本。要获取上下文对象，脚本必须包含以下内容：

```
import spss.pyspark.runtime  
asContext = spss.pyspark.runtime.getContext()
```

从 Analytic Server 上下文，您可以获取 Spark 上下文和 SQL 上下文：

```
sparkContext = asc.getSparkContext()  
sqlContext = asc.getSparkSQLContext()
```

请参阅 Apache Spark 文档以获取有关 Spark 上下文和 SQL 上下文的信息。

访问数据

在 Python/Spark 脚本和 Spark SQL DataFrame 格式的执行上下文之间传输数据。使用数据的脚本（也就是，除源节点之外的任何节点）必须从上下文检索数据帧：

```
inputData = asContext.getSparkInputData()
```

生成数据的脚本（也就是，终端节点外的任何节点）必须将数据帧返回到上下文：

```
asContext.setSparkOutputData(outputData)
```

您可以使用 SQL 上下文以从需要的 RDD 创建输出数据：

```
outputData = sqlContext.createDataFrame(rdd)
```

定义数据模型

生成数据的节点还必须定义数据模型，此数据模型描述节点的字段可视下游。在 Spark SQL 术语中，数据模型即模式。

Python/Spark 脚本定义 `pyspark.sql.types.StructType` 对象格式的数据数据模型。`StructType` 描述输出数据帧中的行，并且是从 `StructField` 对象构造的。每个 `StructField` 都描述输出数据模型中的一个字段。

您可以使用输入数据帧的 `:schema` 属性获取输入数据的数据模型：

```
inputSchema = inputData.schema
```

可将未经更改传递的字段从输入数据模型复制到输出数据模型。可使用 `StructField` 构造函数创建输出数据模型中新建或修改的字段：

```
field = StructField(name, dataType, nullable=True, metadata=None)
```

请参阅 Spark 文档以获取有关构造函数的信息。

您必须至少提供字段名称及其数据类型。您可以指定元数据以提供资源的测量、角色和描述（请参阅第 6 页的『数据元数据』）。

DataModelOnly 方式

在执行节点前，IBM SPSS Modeler 需要指定节点的输出数据模型以支持下游编辑。要获取 Python/Spark 节点的输出数据模型，IBM SPSS Modeler 以无可用数据的“仅数据模型”方式执行脚本。脚本可在 Analytic Server 上下文对象上使用 `isComputeDataModelOnly` 方法来标识此方式。

变换节点的脚本可采用以下一般模式：

```
if asContext.isComputeDataModelOnly():
    inputSchema = asContext.getSparkInputSchema()
    outputSchema = ... # construct the output data model
    asContext.setSparkOutputSchema(outputSchema)
else:
    inputData = asContext.getSparkInputData()
    outputData = ... # construct the output data frame
    asContext.setSparkOutputData(outputData)
```

构建模型

构建模型的节点必须将某些内容返回给执行上下文，此内容充分描述模型，从而使应用模型的节点可在以后精确重新创建它。

按照键/值对定义模型内容，其中仅构建和评分节点了解键和值的含义，而 Modeler 不会以任何方式进行解释。（可选）节点可将 MIME 类型分配给值，旨在 Modeler 可显示具有模型块中用户已知类型的这些值。

此上下文中的值可以是 PMML、HTML 或图像等。要向模型内容添加值（在构建脚本中）：

```
asContext.setModelContentFromString(key, value, mimeType=None)
```

要从模型内容检索值（在评分脚本中）：

```
value = asContext.getModelContentToString(key)
```

作为快捷方式，其中模型或模型的一部分存储到文件系统中的文件或文件夹，您可以通过一个调用绑定存储到此位置的所有内容（在构建脚本中）：

```
asContext.setModelContentFromPath(key, path)
```

请注意，在此情况下，无选项指定 MIME 类型，因为捆绑软件可能包含各种内容类型。

如果在构建模型时需要临时位置来存储内容，那么可以从上下文获取相应的位置：

```
path = asContext.createTemporaryFolder()
```

要将现有内容检索到文件系统中的临时位置（在评分脚本中）：

```
path = asContext.getModelContentToPath(key)
```

错误处理

要引发错误，请从脚本抛出异常并向 IBM SPSS Modeler 用户显示。模块 `spss.pyspark.exceptions` 中预定义了一些异常。例如：

```
from spss.pyspark.exceptions import ASContextException
if ... some error condition ...:
    raise ASContextException("message to display to user")
```

Analytic Server 上下文

"上下文"针对 Analytic Server 上下文界面提供支持以与 IBM SPSS Analytic Server 交互。

AnalyticServerContext 对象

`AnalyticServerContext` 对象设置上下文环境，此环境提供多个接口以供与 IBM SPSS Analytic Server 交互。想要构造此上下文实例的应用程序必须使用 `spss.pyspark.runtime.getContext()` 接口执行此操作，而不是直接实施接口。

返回 Pyspark Python `SparkContext` 实例：

```
cxt.getSparkContext() : SparkContext
```

返回 Pyspark Python `SQLContext` 实例：

```
cxt.getSparkSQLContext() : SQLContext
```

返回 `True` 以描述是否只执行来计算输出数据模型。否则，返回 `False`：

```
cxt.isComputeDataModelOnly() : Boolean
```

如果脚本在 `Spark` 环境中运行，那么返回 `True`。当前，总是返回 `True`：

```
cxt.isSparkExecution() : Boolean
```

从上游临时文件装入输入数据并生成 `pyspark.sql.DataFrame` 实例：

```
cxt.getSparkInputData() : DataFrame
```

返回从输入数据模型生成的 `pyspark.sql.StructType` 实例。如果输入数据模型不存在，那么返回 `None`：

```
cxt.getSparkInputSchema() : StructType
```

将输出数据帧序列化为 Analytic Server 上下文并返回上下文：

```
cxt.setSparkOutputData( outDF ) : AnalyticServerContext
```

参数：

- `outDF (DataFrame)` : 输出数据帧值

异常:

- `DataOutputNotSupported` : 如果在函数 `pyspark:buildmodel` 中调用此接口
- `ASContextException` : 如果输出数据帧为 `None`
- `InconsistentOutputDataModel` : 对于两个对象公共的字段名称和存储类型信息不一致

将 `outSchema StructType` 实例转换为数据模型, 将其序列化为 `Analytic Server` 上下文, 并返回上下文:

```
cxt.setSparkOutputSchema(outSchema) : AnalyticServerContext
```

参数:

- `outSchema(StructType)` : `StructType` 对象的输出

异常:

- `ASContextException` : 如果输出模式实例为 `None`
- `InconsistentOutputDataModel` : 对于两个对象公共的字段名称和存储类型信息不一致

将模型构建输出的位置存储到 `Analytic Server` 上下文并返回上下文:

```
cxt.setModelContentFromPath(key, path, mimetype=None) : AnalyticServerContext
```

路径可以是在将目录下的所有项打包为模型内容时应使用 `cxt.createTemporaryFolder()` API 生成的目录路径。

参数:

- `key (string)` : 键字符串值
- `path (string)` : 模型构建输出字符串路径的位置
- `mimetype (string, optional)` : 内容的 MIME 类型

异常:

- `ModelOutputNotSupported` : 当不是从 `pyspark:buildmodel` 函数调用此 API 时
- `KeyError` : 如果密钥属性为 `None` 或者字符串为空

将模型构建内容、元数据或其他属性存储到 `Analytic Server` 上下文并返回上下文:

```
cxt.setModelContentFromString(key, value, mimetype=None) : AnalyticServerContext
```

参数:

- `key (string)` : 键字符串值
- `value (string)` : 模型元数据字符串值
- `mimetype (string, optional)` : 内容的 MIME 类型

异常:

- `ModelOutputNotSupported` : 当不是从 `pyspark:buildmodel` 函数调用此 API 时
- `KeyError` : 如果密钥属性为 `None` 或者字符串为空

返回 `Analytic Server` 管理的临时文件夹位置; 这可用于存储模型内容:

```
cxt.createTemporaryFolder() : string
```

异常:

- `ModelOutputNotSupported` : 当不是从 `pyspark:buildmodel` 函数调用此 API 时

返回匹配输入键的模型的位置：

```
cxt.getModelContentToPath(key) : string
```

参数：

- key (string) : 键字符串值

异常：

- ModelInputNotSupported : 当不是从 pyspark:applymodel 函数调用此 API 时
- KeyError : 如果密钥属性为 None 或者字符串为空
- IncompatibleModelContentType : 如果模型内容类型不是容器

返回匹配输入键的模型内容、模型的元数据或其他模型属性：

```
cxt.getModelContentToString(key) : string
```

参数：

- key (string) : 键字符串值

异常：

- ModelInputNotSupported : 当不是从 pyspark:applymodel 函数调用此 API 时
- KeyError : 如果键属性为 None、字符串为空或者键不存在
- IncompatibleModelContentType : 如果模型内容类型不一致

返回分配给输入键的 MIME 类型。如果指定的内容无 MIME 类型，那么返回 None：

```
cxt.getModelContentMimeType(key) : string
```

参数：

- key (string) : 键字符串值

异常：

- ModelInputNotSupported : 当不是从 pyspark:applymodel 函数调用此 API 时
- KeyError : 如果键属性为 None、字符串为空或者键不存在

数据元数据

此部分描述如何基于 pyspark.sql.StructField 设置数据模型属性。

spss.datamodel.Role 对象

该类枚举数据模型中每个字段的有效角色。

BOTH: 指示该字段可以是前提条件或结果。

FREQWEIGHT: 指示该字段用作频率权重；不向用户显示。

INPUT: 指示该字段是预测变量还是前提条件。

NONE: 指示在建模期间不之间使用该字段。

TARGET: 指示该字段是预测还是结果。

PARTITION: 指示该字段用于标识数据分区。

RECORDID: 指示该字段用于标识记录标识。

SPLIT: 指示该字段用于拆分数据。

spss.datamodel.Measure 对象

该类枚举数据模型中字段的测量级别。

UNKNOWN: 指示测量类型未知。

CONTINUOUS: 指示测量类型为连续。

NOMINAL: 指示测量类型为名义。

FLAG: 指示字段值为两个值之一。

DISCRETE: 指示字段值应解释为值集合。

ORDINAL: 指示测量类型有序。

TYPELESS: 指示字段可以具有与其存储相兼容的任何值。

pyspark.sql.StructField 对象

表示 StructType 中的字段。StructField 对象有四个字段组成:

- name (string): StructField 名称
- dataType (pyspark.sql.DataType): 特定数据类型
- nullable (bool): 如果 StructField 的值可包含 None 值
- metadata (dictionary): 用于存储选项属性的 Python 字典

您可以使用元数据字典实例来存储指定字段的测量、角色、标签属性。这些属性的关键字为:

- measure: measure 属性的关键字
- role: role 属性的关键字
- displayLabel: label 属性的关键字

示例:

```
from spss.datamodel.Role import Role
from spss.datamodel.Measure import Measure
_metadata = {}
_metadata['measure'] = Measure.TYPELESS
_metadata['role'] = Role.NONE
_metadata['displayLabel'] = "field label description"
StructField("userName", StringType(), nullable=False,
metadata=_metadata)
```

日期、时间和时间戳记

对于使用日期、时间或时间戳记类型数据的操作，将基于值 1970-01-01:00:00:00（使用全球标准时间）将值转换为实数值。

对于日期，该值基于值 1970-01-01（使用全球标准时间）表示天数。

对于时间，该值表示 24 小时制的秒数。

对于时间戳记，该值基于值 1970-01-01:00:00:00（使用全球标准时间）来表示秒数。

异常

此部分描述可能的异常实例。

MetadataException 对象

Python 异常的子类。

如果操作元数据对象期间发生错误，那么将抛出此异常。

UnsupportedOperationException 对象

Python 异常的子类。

如果特定操作不允许执行，那么将抛出此异常。

InconsistentOutputDataModel 对象

Python 异常的子类。

如果调用 `setSparkOutputSchema` 和 `setSparkOutputData` 但是两个对象的公共字段名称和存储类型不一致性，那么将抛出此异常。

IncompatibleModelContentType 对象

Python 异常的子类。

在以下场景中，将抛出此异常：

- 使用 `setModelContentFormString` 来设置模型但是使用 `getModelContentToPath` 以获取值
- 使用 `setModelContentFormPath` 来设置模型但是使用 `getModelContentToString` 来获取值

DataOutputNotSupported 对象

Python 异常的子类。

在函数 `pyspark:buildmodel` 处理的执行的 `setSparkOutputData` 中发出此异常。

ModelInputNotSupported 对象

Python 异常的子类。

仅当脚本未在 `pyspark:applymodel` 函数中调用 `getModelContentPathByKey` 和 `getModelContentToString` API 时发生此异常。

ModelOutputNotSupported 对象

Python 异常的子类。

仅当脚本未在 `pyspark:buildmodel` 函数中调用 `setModelContentFromPath` 和 `setModelContentFromString` API 时发生此异常。

ASContextException 对象

Python 异常的子类。

如果发生意外的运行时异常，那么抛出此异常。

示例

该部分包含 Python for Spark 脚本编制示例。

用于处理数据的基本脚本编制示例

```
import spss.pyspark.runtime
from pyspark.sql.types import *

cxt = spss.pyspark.runtime.getContext()

if cxt.isComputeDataModelOnly():
    _schema = cxt.getSparkInputSchema()
    cxt.setSparkOutputSchema(_schema)
else:
    _structType = cxt.getSparkInputSchema()
    df = cxt.getSparkInputData()
    _newDF = df.sample(False, 0.01, 1)
    cxt.setSparkOutputData(_newDF)
```

使用 LinearRegressionWithSGD 算法的示例模型构建脚本

```
from pyspark.context import SparkContext
from pyspark.sql.context import SQLContext
from pyspark.sql import Row
from pyspark.mllib.regression import
LabeledPoint, LinearRegressionWithSGD, LinearRegressionModel
from pyspark.mllib.linalg import DenseVector
import numpy
import json

import spss.pyspark.runtime
from spss.pyspark.exceptions import ASContextException

ascontext = spss.pyspark.runtime.getContext()
sc = ascontext.getSparkContext()
df = ascontext.getSparkInputData()

# field settings amd algorithm parameters

target = '%target_field%'
predictors = [%predictor_fields%]
num_iterations=%num_iterations%
prediction_field = "$LR-" + target

# save linear regression model to a filesystem path

def save(model, sc, path):
    data =
sc.parallelize([json.dumps({"intercept":model.intercept,"weights":model.weights.tolist()})])
    data.saveAsTextFile(path)

# print model details to stdout

def dump(model,predictors):
    print(prediction_field+" = " + str(model.intercept))
    weights = model.weights.tolist()
    for i in range(0,len(predictors)):
        print("\t+ "+predictors[i]+"*"+ str(weights[i]))
```

```

# check that required fields exist in the input data

input_field_names = [ty[0] for ty in df.dtypes[:]]
if target not in input_field_names:
    raise ASContextException("target field "+target+" not found")
for predictor in predictors:
    if predictor not in input_field_names:
        raise ASContextException("predictor field "+predictor+" not found")

# define map function to convert from dataframe Row objects to mllib LabeledPoint

def row2LabeledPoint(target,predictors,row):
    pvals = []
    for predictor in predictors:
        pval = getattr(row,predictor)
        pvals.append(float(pval))
    tval = getattr(row,target)
    return LabeledPoint(float(tval),DenseVector(pvals))

# convert dataframe to an RDD containing LabeledPoint

training_points = df.rdd.map(lambda row:
row2LabeledPoint(target,predictors,row))

# build the model

model = LinearRegressionWithSGD.train(training_points,num_iterations,intercept=True)

# write a text description of the model to stdout

dump(model,predictors)

# save the model to the filesystem and store into the output model content

modelpath = ascontext.createTemporaryFolder()
save(model,sc,modelpath)
ascontext.setModelContentFromPath("model",modelpath)

```

使用 **LinearRegressionWithSGD** 算法的示例模型评分脚本

```

import json
import spss.pyspark.runtime
from pyspark.sql import Row
from pyspark.mllib.regression import
LabeledPoint,LinearRegressionWithSGD, LinearRegressionModel
from pyspark.mllib.linalg import DenseVector
from pyspark.sql.context import SQLContext
import numpy
from pyspark.sql.types import DoubleType, StructField

ascontext = spss.pyspark.runtime.getContext()
sc = ascontext.getSparkContext()

prediction_field = "$LR-" + '%target_field%'
predictors = ['%predictor_fields%']

# compute the output schema by adding the prediction field
outputSchema = ascontext.getSparkInputSchema()
outputSchema.fields.append(StructField(prediction_field,
DoubleType(), nullable=True))

# make a prediction based on a regression model and Dataframe Row object
# return a list containing the input row values and the predicted value
def predict(row,model,predictors,infields,prediction_field_name):
    pvals = []
    rdict = row.asDict()
    for predictor in predictors:
        pvals.append(float(rdict[predictor]))

```

```

estimate = float(model.predict(pvals))
result = []
for field in infields:
    result.append(rdict[field])
result.append(estimate)
return result

# load a serialized model from the filesystem

def load(sc, path):
    js = sc.textFile(path).take(1)[0]
    obj = json.loads(js)
    weights = numpy.array(obj["weights"])
    intercept = obj["intercept"]
    return LinearRegressionModel(weights,intercept)

ascontext.setSparkOutputSchema(outputSchema)

if not ascontext.isComputeDataModelOnly():
    # score the data in the input data frame
    indf = ascontext.getSparkInputData()

    model_path = ascontext.getModelContentToPath("model")
    model = load(sc,model_path)

    # compute the scores
    infield_names = [ty[0] for ty in indf.dtypes[:]]
    scores_rdd = indf.rdd.map(lambda row:predict(row,model,predictors,infield_names,prediction_field))

    # create an output DataFrame containing the scores
    sqlCtx = SQLContext(sc)
    outdf = sqlCtx.createDataFrame(scores_rdd,schema=outputSchema)

    # return the output DataFrame as the result
    ascontext.setSparkOutputData(outdf)

```

"扩展"节点

作为对 IBM SPSS Modeler 及其数据挖掘功能的补充, "扩展"节点使专业用户能够输入自己的 R 脚本或 Python for Spark 脚本以执行数据处理、模型构建和模型评分。

"扩展导出"节点

使用"扩展导出"节点, 您可以运行 R 或 Python for Spark 脚本以导出数据。

"扩展导出"节点 -"语法"选项卡

选择语法类型 - **R** 或 **Python for Spark**。请参阅以下部分以获取更多信息。语法就绪时, 您可以单击运行来执行"扩展导出"节点。

R 语法

R 语法。您可以在此字段中输入或粘贴用于数据分析的定制 R 脚本语法。

转换标志字段。指定标志字段的处理方式。共有两个选项: 将字符串转换为因子, 将整数和实数转换为双精度数和逻辑值 (**True** 和 **False**)。如果选择逻辑值 (**True** 和 **False**), 那么标志字段的原始值将丢失。例如, 如果某个字段的值为 Male 和 Female, 那么这些值将更改为 True 和 False。

将缺失值转换为 R 的"不可用"值 (NA)。如果选择了此选项，那么所有缺失值都将转换为 R 的 NA 值。R 使用值 NA 来标识缺失值。您使用的某些 R 函数可能有一个自变量来控制数据包含 NA 时函数的行为。例如，该函数可能允许您选择自动排除包含 NA 的记录。如果未选择此选项，那么所有缺失值都将按原样传递到 R，并可能导致执行 R 脚本时发生错误。

将日期/时间字段转换为具有特殊时区控件的 R 类。如果选择此选项，那么会将带有日期或日期时间格式的变量转换为 R 日期/时间对象。必须选择下列选项之一：

- **R POSIXct**。会将带有日期或日期时间格式的变量转换为 R POSIXct 对象。
- **R POSIXlt** (列表)。会将带有日期或日期时间格式的变量转换为 R POSIXlt 对象。

注：POSIX 格式是高级选项。仅当您的 R 脚本指定以需要这些格式的方式处理日期时间字段时才使用这些选项。POSIX 格式不适用于带有时间格式的变量。

Python 语法

Python 语法。您可以针对数据分析向此字段中输入或粘贴定制的 Python 脚本语法。有关 Python for Spark 的更多信息，请参阅第 1 页的『Python for Spark』和第 1 页的『使用 Python for Spark 编写脚本』。

"扩展导出"节点 -"控制台输出"选项卡

控制台输出选项卡包含当"语法"选项卡上的 R 脚本或 Python for Spark 脚本运行时接收到的任何输出（例如，如果使用 R 脚本，当执行语法选项卡上的 R 语法字段中的 R 脚本时，它显示从 R 控制台接收到的输出）。此输出可能包括执行 R 或 Python 脚本时生成的 R 或 Python 错误消息或警告。输出可主要用于调试脚本。控制台输出选项卡还包含 R 语法或 Python 语法字段中的脚本。

每次执行"扩展导出"脚本时，都会使用从 R 控制台或 Python for Spark 接收到的输出来覆盖控制台输出选项卡的内容。输出不能编辑。

发布流

使用以下任何导出节点从 IBM SPSS Modeler 直接完成发布流："数据库"、"平面文件"、"Statistics 导出"、"扩展导出"、"数据收集导出"、"SAS 导出"、Excel 和"XML 导出"节点。导出节点的类型确定每次使用 IBM SPSS Modeler Solution Publisher Runtime 或外部应用程序执行已发布的流时要写入的结果的格式。例如，如果您希望每次运行已发布的流时将结果写入数据库，请使用"数据库"导出节点。

要发布流

1. 按常规方式打开或构建流并在最后附加导出节点。
2. 在导出节点中的"发布"选项卡上，为已发布的文件（即 .pim、.par 和 .xml 扩展名将附加到的文件名）指定根目录名称。
3. 单击**发布**以发布流，或选择**发布流**以在每次执行节点时自动发布流。

已发布的名称。为已发布的图像和参数文件指定根目录名称。

- 与导出时完全一样，**图像文件** (*.pim) 提供 Runtime 执行已发布的流所需的全部信息。如果您认为将不需要更改流（例如，输入数据源或输出数据文件）的任何设置，那么您可以只部署图像文件。
- **参数文件** (*.par) 包含有关数据源、输出文件和执行选项的可配置信息。如果您希望能够在不重新发布流的情况下控制流的输入或输出，您将需要参数文件以及图像文件。
- **元数据文件** (*.xml) 描述图像及其数据模型的输入和输出。此文件专供嵌入运行时库并且需要知道输入和输出数据结构的应用程序使用。

注：仅当选择**发布元数据**选项时，才生成此文件。

发布参数。如果需要，您可以在 *.par 文件中包含流参数。编辑 *.par 文件或通过运行时 API 执行图像时，您可以更改这些流参数值。

此选项启用**参数按钮**。单击此按钮时，显示"发布参数"对话框。

通过在**发布列**中选择相关选项来选择要包含在已发布的图像中的参数。

在**流执行时**。指定在执行节点时是否自动发布流。

- **导出数据。**按标准方式执行导出节点，但不发布流。（节点在 IBM SPSS Modeler 中的执行方式基本与 IBM SPSS Modeler Solution Publisher 不可用时相同。）如果您选择此选项，那么将不发布流，除非您通过单击导出节点对话框中的**发布**来明确执行该操作。或者，您也可以使用工具栏上的"发布"工具或使用脚本来发布当前流。
- **发布流。**使用 IBM SPSS Modeler Solution Publisher 为部署发布流。如果您希望在每次执行流时自动发布流，请选择此选项。

注：

- 如果您计划使用新数据或更新的数据来运行已发布的流，务必注意输入文件中的字段顺序必须与已发布流中指定的源节点输入文件中的字段顺序相同。
- 发布到外部应用程序时，请考虑过滤外部字段或其余字段以符合输入需求。使用导出节点之前的"过滤器"节点可以完成这两个操作。

"扩展输出"节点

如果在"扩展输出"节点对话框的**输出选项卡**上选择了**输出到屏幕**，那么将在输出浏览器窗口中显示屏幕上的输出。此输出还将添加到输出管理器中。输出浏览器窗口具有它自己的菜单集，使用这些菜单，可以打印或保存输出，也可以将输出导出为其他格式。**编辑菜单**仅包含**复制**选项。"扩展输出"节点的输出浏览器有两个选项卡：用于显示文本输出的**文本输出**选项卡，以及用于显示图形和图表的**图形输出**选项卡。

如果在"扩展输出"节点对话框的**输出选项卡**上选择了**输出到文件**，那么在成功执行"扩展输出"节点后，不会显示输出浏览器窗口。

"扩展输出"节点 -"语法"选项卡

选择语法类型 - **R** 或 **Python for Spark**。请参阅以下部分以获取更多信息。语法就绪时，您可以单击**运行**来执行"扩展输出"节点。输出对象将添加到输出管理器，或者选择性地添加到**输出选项卡**上的**文件名字段**中指定的文件中。

R 语法

R 语法。您可以在此字段中输入或粘贴用于数据分析的定制 R 脚本语法。

转换标志字段。指定标志字段的处理方式。共有两个选项：**将字符串转换为因子**，**将整数和实数转换为双精度数和逻辑值 (True 和 False)**。如果选择**逻辑值 (True 和 False)**，那么标志字段的原始值将丢失。例如，如果某个字段的值为 Male 和 Female，那么这些值将更改为 True 和 False。

将缺失值转换为 R 的"不可用"值 (NA)。如果选择了此选项，那么所有缺失值都将转换为 R 的 NA 值。R 使用值 NA 来标识缺失值。您使用的某些 R 函数可能有一个自变量来控制数据包含 NA 时函数的行为。例如，该函数可能允许您选择自动排除包含 NA 的记录。如果未选择此选项，那么所有缺失值都将按原样传递到 R，并可能导致执行 R 脚本时发生错误。

将日期/时间字段转换为具有特殊时区控件的 R 类。如果选择此选项，那么会将带有日期或日期时间格式的变量转换为 R 日期/时间对象。必须选择下列选项之一：

- **R POSIXct**. 会将带有日期或日期时间格式的变量转换为 R POSIXct 对象。
- **R POSIXlt** (列表)。会将带有日期或日期时间格式的变量转换为 R POSIXlt 对象。

注：POSIX 格式是高级选项。仅当您的 R 脚本指定以需要这些格式的方式处理日期时间字段时才使用这些选项。POSIX 格式不适用于带有时间格式的变量。

Python 语法

Python 语法。您可以针对数据分析向此字段中输入或粘贴定制的 Python 脚本语法。有关 Python for Spark 的更多信息，请参阅第 1 页的『Python for Spark』和第 1 页的『使用 Python for Spark 编写脚本』。

"扩展输出"节点 -"控制台输出"选项卡

控制台输出选项卡包含当"语法"选项卡上的 R 脚本或 Python for Spark 脚本运行时接收到的任何输出（例如，如果使用 R 脚本，当执行语法选项卡上的 **R 语法** 字段中的 R 脚本时，它显示从 R 控制台接收到的输出）。此输出可能包括执行 R 或 Python 脚本时生成的 R 或 Python 错误消息或警告。输出可主要用于调试脚本。控制台输出选项卡还包含 **R 语法** 或 **Python 语法** 字段中的脚本。

每次执行"扩展输出"脚本时，都会使用从 R 控制台或 Python for Spark 接收到的输出来覆盖控制台输出选项卡的内容。输出不能编辑。

"扩展输出"节点 -"输出"选项卡

输出名称。指定执行此节点时生成的输出的名称。选择**自动**时，输出名称将自动设置为"R 输出"或"Python 输出"，具体取决于脚本类型。（可选）可以选择**自定义**以指定其他名称。

输出到屏幕。选择此选项可以在新窗口中生成并显示输出。此输出还将添加到输出管理器中。

输出到文件。选择此选项可将输出保存到文件中。执行此操作将启用**输出图形**和**输出文件**单选按钮。

输出图形。只有在选择了**输出到文件**的情况下才会启用此按钮。选择此选项可以将执行"扩展输出"节点所产生的任何图形保存到文件中。请在**文件名**字段中指定要用于生成的输出的文件名。单击省略号按钮 (...) 以选择特定文件和位置。请在**文件类型**下拉列表中指定文件类型。可用的文件类型如下所示：

- 输出对象 (.cou)
- HTML (.html)

输出文本。只有在选择了**输出到文件**的情况下才会启用此按钮。选择此选项可以将执行"扩展输出"节点所产生的任何文本输出保存到文件中。请在**文件名**字段中指定要用于生成的输出的文件名。单击省略号按钮 (...) 以指定特定文件和位置。请在**文件类型**下拉列表中指定文件类型。可用的文件类型如下所示：

- HTML (.html)
- 输出对象 (.cou)
- 文本文档 (.txt)

扩展输出浏览器

如果在"扩展输出"节点对话框的**输出**选项卡上选择了**输出到屏幕**，那么将在输出浏览器窗口中显示屏幕上的输出。此输出还将添加到输出管理器中。输出浏览器窗口具有它自己的菜单集，使用这些菜单，可以打印或保存输出，也可以将输出导出为其他格式。**编辑**菜单仅包含**复制**选项。"扩展输出"节点的输出浏览器有两个选项卡：

- **文本输出**选项卡显示文本输出
- **图形输出**选项卡显示图形和图表

如果在"扩展输出"节点对话框的输出选项卡上选择了输出到文件，而不是输出到屏幕，那么在成功执行"扩展输出"节点后，不会显示输出浏览器窗口。

扩展输出浏览器 -"文本输出"选项卡： 文本输出选项卡显示执行"扩展输出"节点的语法选项卡上的 R 脚本或 Python for Spark 脚本时生成的任何文本输出。

注：由于执行扩展输出脚本而产生的 R 或 Python for Spark 错误消息或警告始终显示在"扩展输出"节点的控制台输出选项卡上。

扩展输出浏览器 -"图形输出"选项卡： 图形输出选项卡显示执行"扩展输出"节点的语法选项卡上的 R 脚本或 Python for Spark 脚本时生成的任何图形或图表。例如，如果您的 R 脚本包含对 R plot 函数的调用，那么产生的图形会显示在此选项卡上。

"扩展模型"节点

使用"扩展模型"节点，您可以运行 R 或 Python for Spark 脚本以构建模型并对其进行评分。

"扩展模型"节点 -"语法"选项卡

选择语法类型 - **R** 或 **Python for Spark**。然后，向以下某个字段输入或粘贴定制的脚本语法。语法就绪时，您可以单击运行来执行"扩展模型"节点。

R 语法

R 模型构建语法。 您可以针对模型构建向此字段输入或粘贴定制的 R 脚本语法。

R 模型评分语法。 您可以针对模型评分向此字段输入或粘贴定制的 R 脚本语法。

Python for Spark 语法

Python 模型构建语法。 您可以针对模型构建向此字段输入或粘贴定制的 Python 脚本语法。

Python 模型评分语法。 您可以针对模型评分向此字段输入或粘贴定制的 Python 脚本语法。

有关 Python for Spark 的更多信息，请参阅第 1 页的『Python for Spark』和第 1 页的『使用 Python for Spark 编写脚本』。

"扩展模型"节点 -"模型选项"选项卡

模型名称。 您可以根据目标或标识字段（或未指定此类字段时的模型类型）自动生成模型名称，或者指定定制名称。

"扩展模型"节点 -"控制台输出"选项卡

控制台输出选项卡包含当"语法"选项卡上的 R 脚本或 Python for Spark 脚本运行时接收到的任何输出（例如，如果使用 R 脚本，当执行语法选项卡上的 **R** 语法字段中的 R 脚本时，它显示从 R 控制台接收到的输出）。此输出可能包括执行 R 或 Python 脚本时生成的 R 或 Python 错误消息或警告。输出可主要用于调试脚本。控制台输出选项卡还包含 **R** 语法或 **Python** 语法字段中的脚本。

每次执行"扩展模型"脚本时，都会使用从 R 控制台或 Python for Spark 接收到的输出来覆盖控制台输出选项卡的内容。输出不能编辑。

"扩展模型"节点 -"文本输出"选项卡

如果通过在"扩展模型"节点对话框的模型选项选项卡上选中显示 R 文本输出复选框进行请求，那么文本输出选项卡存在于"扩展模型"节点中。此选项卡仅能显示文本输出。此选项卡显示由执行 R 模型构建脚本生成的所有文本输出。如果您再次执行模型构建脚本，而未先对模型指定其他名称，那么来自先前执行的文本输出选项卡的内容将被覆盖。文本输出不能编辑。

如果您在脚本中包含对 R sink 函数的调用，那么在此函数之后生成的所有输出会保存到指定文件并且不会显示在文本输出选项卡。

注：由于执行模型构建脚本而产生的 R 或 Python for Spark 错误消息或警告始终显示在"扩展模型"节点的控制台输出选项卡上。

扩展模型块

执行"扩展模型"节点后会生成扩展模型块并将其置于"模型"选用板上，该模型块包含定义模型构建和模型评分的 R 或 Python for Spark 脚本。缺省情况下，扩展模型块包含用于模型评分的脚本、用于读取数据的选项以及 R 控制台或 Python for Spark 的任何输出。（可选）扩展模型块还可以包含各种其他格式的模型输出，例如图形和文本输出。在生成扩展模型块且将其添加到流画布后，可将输出接点与其进行连接。此时，该输出接点通常会在 IBM SPSS Modeler 流中用于获取有关数据和模型的信息以及导出各种格式的数据。

要将此节点与 R 配合使用，您必须安装 IBM SPSS Modeler - Essentials for R。请参阅 *IBM SPSS Modeler - Essentials for R: Installation Instructions* 以获取安装指示信息和兼容性信息。另外，还必须在计算机上安装 R 的兼容版本。

扩展模型块 -"语法"选项卡

语法选项卡始终存在于扩展模型块中。

R 模型评分语法。如果使用 R，那么用于模型评分的 R 脚本显示在此字段中。缺省情况下，该字段为启用状态，但不可编辑。要编辑 R 模型评分脚本，请单击编辑。

Python 模型评分语法。如果使用 Python for Spark，那么用于模型评分的 Python 脚本显示在此字段中。缺省情况下，该字段为启用状态，但不可编辑。要编辑 Python 模型评分脚本，请单击编辑。

编辑。单击编辑以使评分语法字段可编辑。然后，您可以通过在评分语法字段中进行输入来编辑模型评分脚本。例如，如果您在执行扩展模型块后在模型评分脚本中识别了错误，那么您可能希望编辑您的模型评分脚本。如果您通过执行"扩展模型"节点重新生成模型，那么您在扩展模型块中对模型评分脚本所做的任何更改都将丢失。

扩展模型块 -"模型选项"选项卡

模型选项选项卡始终存在于扩展模型块中。

阅读数据选项。这些选项仅适用于 R，不适用于 Python for Spark。借助这些选项，您可以指定格式为日期或日期时间的缺失值、标志字段和变量的处理方式。

- **以批处理方式读取数据。**如果要处理大量数据（例如，数据太大无法适合 R 引擎的内存），请使用此选项将数据划分为可单独发送和处理的批次。指定要包含在每个批次中的数据记录的最大数量。

对于"扩展变换"节点和扩展评分块，通过 R 脚本（批处理）传递数据。因此，用于在 Hadoop 或"数据库"环境中运行的模型评分节点和流程节点的脚本不应包括跨数据行或组合数据行的操作，例如排序或汇总。实施此限制的目的是确保能够在 Hadoop 环境中以及数据库内挖掘期间分割数据。如果在 SPSS Modeler Server 中运行模型评分脚本，那么此限制不适用。"扩展输出"和"扩展模型"节点没有此限制。

- **转换标志字段。**指定标志字段的处理方式。共有两个选项：**将字符串转换为因子，将整数和实数转换为双精度数和逻辑值 (True 和 False)**。如果选择**逻辑值 (True 和 False)**，那么标志字段的原始值将丢失。例如，如果某个字段的值为 Male 和 Female，那么这些值将更改为 True 和 False。
- **将缺失值转换为 R 的"不可用"值 (NA)。**如果选择了此选项，那么所有缺失值都将转换为 R 的 NA 值。R 使用值 NA 来标识缺失值。您使用的某些 R 函数可能有一个自变量来控制数据包含 NA 时函数的行为。例如，该函数可能允许您选择自动排除包含 NA 的记录。如果未选择此选项，那么所有缺失值都将按原样传递到 R，并可能导致执行 R 脚本时发生错误。
- **将日期/时间字段转换为具有特殊时区控件的 R 类** 如果选择此选项，那么会将带有日期或日期时间格式的变量转换为 R 日期/时间对象。必须选择下列选项之一：
 - **R POSIXct.** 会将带有日期或日期时间格式的变量转换为 R POSIXct 对象。
 - **R POSIXlt (列表)。** 会将带有日期或日期时间格式的变量转换为 R POSIXlt 对象。

注：POSIX 格式是高级选项。仅当您的 R 脚本指定以需要这些格式的方式处理日期时间字段时才使用这些选项。POSIX 格式不适用于格式为时间的变量。

针对数据库运行扩展模型块时，未识别您为**转换标志字段、将缺失值转换为 R"不可用"值 (NA)** 以及将**日期/时间字段转换为具有时区特殊控件的 R 类**控件选择的选项。针对数据库运行节点时，将改为使用这些控件的缺省值：

- **转换标志字段**设置为将字符串转换为因子，将整数和实数转换为双精度数。
- 选择了**将缺失值转换为 R 的"不可用"值 (NA)**。
- 未选择将日期/时间字段转换为具有时区特殊控件的 R 类。

扩展模型块 -"图形输出"选项卡

如果通过在"扩展模型"节点对话框的**模型选项**选项卡上选中以 **HTML 格式显示 R 图形**复选框进行请求，那么**图形输出**选项卡存在于扩展模型块中。由执行模型构建 R 脚本产生的图形可显示在此选项卡上。例如，如果您的 R 脚本包含对 R plot 函数的调用，那么产生的图形会显示在此选项卡上。如果您再次执行模型构建脚本，而未先对模型指定其他名称，那么来自先前执行的**图形输出**选项卡的内容将被覆盖。

扩展模型块 -"文本输出"选项卡

如果通过在"扩展模型"节点对话框的**模型选项**选项卡上选中**显示 R 文本输出**复选框进行请求，那么**文本输出**选项卡存在于扩展模型块中。此选项卡仅能显示文本输出。此选项卡显示由执行"扩展模型"脚本生成的任何文本输出。如果您再次执行"扩展模型"脚本，而未先对模型指定其他名称，那么来自先前执行的**文本输出**选项卡的内容将被覆盖。文本输出不能编辑。

注：

- 如果您在脚本中包含对 R sink 函数的调用，那么在此函数之后生成的所有输出会保存到指定文件并且不会显示在**文本输出**选项卡。
- 由于执行"扩展模型"脚本而产生的错误消息或警告始终显示在"扩展模型"节点的**控制台输出**选项卡上。

扩展模型块 -"控制台输出"选项卡

控制台输出选项卡始终存在于扩展模型块中。它包含当"语法"选项卡上的 R 脚本或 Python for Spark 脚本运行时接收到的任何输出（例如，如果使用 R 脚本，当执行扩展模型块语法选项卡上的 **R 模型评分**语法字段中的 R 脚本时，它显示从 R 控制台接收到的输出）。此输出包括在执行 R 或 Python 脚本时生成的任何 R 或 Python 错误消息或警告，以及 R 控制台的任何文本输出。输出可主要用于调试脚本。

每次执行模型评分脚本时，都会使用从 R 控制台或 Python for Spark 接收到的输出来覆盖**控制台输出**选项卡的内容。控制台输出不能编辑。

"扩展变换"节点

借助"扩展变换"节点，您可以使用 R 脚本编制或 Python for Spark 脚本编制从 IBM SPSS Modeler 流中获取数据并对该数据应用变换。完成修改后，数据将返回到流中以进行进一步处理、模型构建和模型评分。"扩展变换"节点支持使用以 R 或 Python for Spark 编写的算法来变换数据，并让用户能够开发针对特定问题进行定制的数据变换方法。

要将此节点与 R 配合使用，您必须安装 IBM SPSS Modeler - Essentials for R。请参阅 *IBM SPSS Modeler - Essentials for R: Installation Instructions* 以获取安装指示信息和兼容性信息。另外，还必须在计算机上安装 R 的兼容版本。

"扩展变换"节点 -"语法"选项卡

选择语法类型 - **R** 或 **Python for Spark**。请参阅以下部分以获取更多信息。语法就绪时，您可以单击运行来执行"扩展变换"节点。

R 语法

R 语法。您可以在此字段中输入或粘贴用于数据分析的定制 R 脚本语法。

转换标志字段。指定标志字段的处理方式。共有两个选项：**将字符串转换为因子**，**将整数和实数转换为双精度数和逻辑值 (True 和 False)**。如果选择**逻辑值 (True 和 False)**，那么标志字的原始值将丢失。例如，如果某个字段的值为 Male 和 Female，那么这些值将更改为 True 和 False。

将缺失值转换为 R 的"不可用"值 (NA)。如果选择了此选项，那么所有缺失值都将转换为 R 的 NA 值。R 使用值 NA 来标识缺失值。您使用的某些 R 函数可能有一个自变量来控制数据包含 NA 时函数的行为。例如，该函数可能允许您选择自动排除包含 NA 的记录。如果未选择此选项，那么所有缺失值都将按原样传递到 R，并可能导致执行 R 脚本时发生错误。

将日期/时间字段转换为具有特殊时区控件的 R 类。如果选择此选项，那么会将带有日期或日期时间格式的变量转换为 R 日期/时间对象。必须选择下列选项之一：

- **R POSIXct.** 会将带有日期或日期时间格式的变量转换为 R POSIXct 对象。
- **R POSIXlt (列表)。**会将带有日期或日期时间格式的变量转换为 R POSIXlt 对象。

注：POSIX 格式是高级选项。仅当您的 R 脚本指定以需要这些格式的方式处理日期时间字段时才使用这些选项。POSIX 格式不适用于带有时间格式的变量。

Python 语法

Python 语法。您可以针对数据分析向此字段中输入或粘贴定制的 Python 脚本语法。有关 Python for Spark 的更多信息，请参阅第 1 页的『Python for Spark』和第 1 页的『使用 Python for Spark 编写脚本』。

"扩展变换"节点 -"扩展输出"选项卡

控制台输出选项卡包含当"语法"选项卡上的 R 脚本或 Python for Spark 脚本运行时接收到的任何输出（例如，如果使用 R 脚本，当执行语法选项卡上的 **R 语法** 字段中的 R 脚本时，它显示从 R 控制台接收到的输出）。此输出可能包括执行 R 或 Python 脚本时生成的 R 或 Python 错误消息或警告。输出可主要用于调试脚本。控制台输出选项卡还包含 **R 语法** 或 **Python 语法** 字段中的脚本。

每次执行"扩展变换"脚本时，都会使用从 R 控制台或 Python for Spark 接收到的输出来覆盖控制台输出选项卡的内容。输出不能编辑。

"扩展导入"节点

使用"扩展导入"节点，您可以运行 R 或 Python for Spark 脚本以导入数据。

"扩展导入"节点 -"语法"选项卡

选择语法类型 - R 或 **Python for Spark**。然后，为导入数据输入或粘贴定制脚本。语法就绪时，您可以单击运行来执行"扩展导入"节点。

"扩展导入"节点 -"控制台输出"选项卡

控制台输出选项卡包含当"语法"选项卡上的 R 脚本或 Python for Spark 脚本运行时接收到的任何输出（例如，如果使用 R 脚本，当执行语法选项卡上的 **R 语法** 字段中的 R 脚本时，它显示从 R 控制台接收到的输出）。此输出可能包括执行 R 或 Python 脚本时生成的 R 或 Python 错误消息或警告。输出可主要用于调试脚本。控制台输出选项卡还包含 **R 语法** 或 **Python 语法** 字段中的脚本。

每次执行"扩展导入"脚本时，都会使用从 R 控制台或 Python for Spark 接收到的输出来覆盖控制台输出选项卡的内容。输出不能编辑。

过滤或重命名字段

您可以重命名或排除流中任何点处的字段。例如，作为医学研究人员，您可能不关心病人（记录级数据）的钾水平（字段级数据）；因此，您可以过滤出 K（钾）字段。使用单独的"过滤器"节点或者使用源节点或输出节点上的"过滤器"选项卡，可以完成此操作。不论从哪个节点进行访问，功能都是相同的。

- 从源节点（例如，"变量文件"、"固定文件"、"Statistics 文件"、XML 或"扩展导入"）中，您可以在将数据读入 IBM SPSS Modeler 时重命名或过滤字段。
- 使用"过滤器"节点，您可以重命名或过滤流中任何点处的字段。
- 从"Statistics 导出"、"Statistics 变换"、"Statistics 模型"和"Statistics 输出"节点，您可以过滤或重命名字段以符合 IBM SPSS Statistics 命名标准。有关更多信息，请参阅为 IBM SPSS Statistics 重命名或过滤字段主题。
- 您可以使用以上任何节点中的"过滤器"选项卡来定义或编辑多响应集。有关更多信息，请参阅编辑多响应集主题。
- 最后，您可以使用"过滤器"节点将字段从一个源节点映射到另一个节点。

第 2 章 扩展

扩展是扩展 IBM SPSS Modeler 的功能的定制组件。扩展打包在扩展束（.mpe 文件）中并且安装到 IBM SPSS Modeler。任何用户都可创建扩展并通过共享关联的扩展束来与其他用户共享。

提供实用程序以用于处理扩展：

- 『扩展中心』（从**扩展 > 扩展中心**访问）是从 GitHub 上的 IBM SPSS Predictive Analytics 集合搜索、下载和安装扩展的界面。从“扩展中心”对话框，您还可以查看计算机上安装的扩展的详细信息，获取已安装的扩展的更新，以及除去扩展。
- 您可以从**扩展 > 安装本地扩展束**安装存储在本地计算机上的扩展束。
- 您可以使用针对扩展的定制对话框构建器来创建包含用户界面的扩展，此用户界面称为定制节点对话框。定制节点对话框会生成 R 脚本或 Python for Spark 脚本，用于执行与此扩展关联的任务。指定生成的脚本作为设计定制对话框的一部分。

扩展中心

从“扩展中心”对话框，您可以执行以下任务：

- 探索 GitHub 上的 IBM SPSS Predictive Analytics 集合中提供的扩展。您可以选择扩展以立即安装或者下载选中的扩展并稍后安装。
- 获取计算机上已安装的扩展的更新版本。
- 查看有关计算机上已安装的扩展的详细信息。
- 除去计算机上安装的扩展。

要下载或除去扩展：

1. 从菜单，选择：**扩展 > 扩展中心**
2. 选择想要下载或除去的扩展，然后单击**确定**。在单击**确定**时，将处理在“探索”和“已安装”选项卡上生成的所有选择。

缺省情况下，将在计算机上下载和安装针对下载选中的扩展。从“设置”选项卡，您可以将选中的扩展下载到指定的位置，而无需安装。然后，您可以通过选择**扩展 > 安装本地扩展束**以稍后进行安装。对于 Windows，您可以通过双击扩展束文件来安装扩展。

要点：

- 对于 Windows 7 和更新版本，安装更新版本的现有扩展束可能需要使用管理员特权运行 IBM SPSS Modeler。您可以通过右键单击 IBM SPSS Modeler 的图标并选择**以管理员身份运行**，使用管理员特权来启动 IBM SPSS Modeler。尤其是，如果收到指示无法安装一个或多个扩展束的错误消息时，那么尝试使用管理员特权运行。
- 如果要通过代理连接到因特网，当您尝试通过**扩展 > 扩展中心**菜单选项打开“扩展中心”时，您可能会收到类似“因为未检测到因特网连接，一些功能不可用”的错误。要解决此问题，请向 jvm.cfg 文件（位于 SPSS Modeler 安装的 config 目录中）中的 # JVM 选项添加以下参数。保存文件并重新启动 SPSS Modeler。

```
options, "-DproxyHost=proxyIP"  
options, "-DproxyPort=proxyPort"
```

注：以后随时可通过单击“已安装”选项卡上扩展的**更多信息...**来查看安装扩展时同意的许可证。

"探索"选项卡

"探索"选项卡显示 GitHub (<https://ibmpredictiveanalytics.github.io/>) 上的 IBM SPSS Predictive Analytics 集合的所有扩展。从"探索"选项卡,您可以选择新扩展以下载和安装,并且您可以选择计算机上已安装的扩展的更新。"探索"选项卡需要因特网连接。

- 对于每个扩展,将显示最新版本号以及此版本的关联日期。还提供扩展的简述。对于计算机上已安装的扩展,还将显示已安装的版本号。
- 您可以单击**更多信息**来查看有关扩展的详细信息。当更新可用时,**更多信息**显示有关更新的信息。
- 您可以通过单击**必备软件**来查看运行扩展的必备软件,例如,是否需要 IBM SPSS Modeler - Integration Plug-in for R。在更新可用时,**必备软件**显示有关更新的信息。

优化依据

您可以优化显示的扩展集。您可以按常规扩展类别、实施扩展的语言、提供扩展的组织类型或扩展的状态进行优化。对于每个组,例如,"类别",您可以选择用来优化显示的扩展列表的多个项。您还可以按搜索项进行优化。搜索不区分大小写,星号(*)将作为任何其他字符,并且不指示通配符搜索。

- 要优化显示的扩展列表,请单击**应用**。在光标置于搜索框中时按 Enter 键与单击**应用**的效果相同。
- 要重置列表以显示所有可用扩展,请删除搜索框中的任何文本,取消选中所有项,然后单击**应用**。

如何获取集成插件

要获取 **IBM SPSS Modeler - Integration Plug-in for R**,请执行以下操作:

安装从 https://github.com/IBMPredictiveAnalytics/R_Essentials_Modeler/releases/ 或 IBM SPSS Predictive Analytics 社区 (<https://developer.ibm.com/predictiveanalytics/predictive-extensions/>) 获取的 IBM SPSS Modeler - Essentials for R。IBM SPSS Modeler - Essentials for R 包含 IBM SPSS Modeler - Integration Plug-in for R。Essentials for R 不包含 R 编程语言。在安装 IBM SPSS Modeler - Essentials for R 前,如果尚未安装 R,您将需要进行安装。可从 <http://www.r-project.org/> 获取。我们建议安装 R 3.3.3。

有关 安装和配置 的详细信息,请参阅 IBM SPSS Statistics - Integration Plug-in for R 概述 Wiki。

注:如果要在没有因特网访问权的计算机上安装 Essentials for R 并且计划使用 Essentials for R 中包含的 R 脚本,那么您必须获取这些脚本所需的任何 R 软件包并在 R 中进行手动安装。要确定特定 R 脚本需要哪些 R 软件包,请打开"扩展中心"对话框(扩展 > 扩展中心),单击**已安装**选项卡,然后为所需扩展单击**更多信息**。"扩展详细信息"对话框上列出了所需的 R 软件包。R 软件包可以从通过 <http://www.r-project.org/> 访问的任何 R CRAN 镜像站点中获取。请确保获取与您的 R 版本相匹配的软件包版本。特定于版本的软件包可以通过 CRAN 镜像站点的"已添加的软件包"页面上的链接获取。

"已安装"选项卡

"已安装"选项卡会显示计算机上已安装的所有扩展。从"已安装"选项卡中可以选择 GitHub 上的 IBM SPSS Predictive Analytics 集合提供的已安装扩展的更新和移除扩展。要获取已安装的扩展的更新,必须具有因特网连接。

- 对于每个扩展,将显示已安装的版本号。当因特网连接可用时,会显示最新版本号和该版本的关联日期。还提供扩展的简述。
- 您可以单击**更多信息**来查看有关扩展的详细信息。当更新可用时,**更多信息**显示有关更新的信息。
- 您可以通过单击**必备软件**来查看运行扩展的必备软件,例如,是否需要 IBM SPSS Modeler - Integration Plug-in for R。在更新可用时,**必备软件**显示有关更新的信息。

优化依据

您可以优化显示的扩展集。您可以按常规扩展类别、实施扩展的语言、提供扩展的组织类型或扩展的状态进行优化。对于每个组，例如，“类别”，您可以选择用来优化显示的扩展列表的多个项。您还可以按搜索项进行优化。搜索不区分大小写，星号 (*) 将作为任何其他字符，并且不指示通配符搜索。

- 要优化显示的扩展列表，请单击**应用**。在光标置于搜索框中时按 Enter 键与单击**应用**的效果相同。
- 要重置列表以显示所有可用扩展，请删除搜索框中的任何文本，取消选中所有项，然后单击**应用**。

专用扩展

专用扩展是计算机上已经安装但是 GitHub 上的 IBM SPSS Predictive Analytics 集合尚未提供的扩展。优化显示的扩展集合和查看运行扩展的先决条件的功能对专用扩展不可用。

注：在没有因特网连接的情况下使用“扩展中心”时，“已安装”选项卡的部分功能可能不可用。

设置

“设置”选项卡指定是下载并安装针对下载选中的安装，还是下载但不安装。此设置应用于新扩展并更新为现有扩展。如果下载扩展以分发给组织中的其他用户，那么您可以选择下载扩展而不必安装。如果您没有运行扩展的必备软件但计划获取必备软件，那么您还可以选择下载但不安装扩展。

如果选择下载扩展而不安装，那么稍后可通过选择**扩展 > 安装本地扩展束**来安装。对于 Windows，您可以通过双击扩展束文件来安装扩展。

扩展详细信息

“扩展详细信息”对话框显示另一个扩展提供的信息。除了必需的信息，例如，“摘要”和“版本”，作者可能包含相关性位置的 URL，例如，作者的主页。如果扩展是从“扩展中心”下载的，那么它包含许可证，可以通过单击**查看许可证**来查看。

定制节点。“定制节点”表列出了扩展包含的定制节点对话框。

注：安装包包含定制节点对话框的扩展可能需要重新启动 IBM SPSS Modeler，以便在定制节点表中查看节点对话框的输入。

依赖性。“依赖性”组列出运行扩展中包含的组件所需的附加组件。

- **R 的集成插件。**扩展的组件可能需要 Integration Plug-in for R。
- **R 软件包。**列出扩展所需的任何 R 软件包。有关更多信息，请参阅第 24 页的『必需的 R 软件包』主题。

安装本地扩展束

要安装存储在本地计算机上的扩展束：

1. 从菜单中选择：

扩展 > 安装本地扩展束...

2. 选择扩展束。扩展束的文件类型为 mpe。

要点：对于 Windows 7 和更新版本的 Windows 用户，安装更新版本的现有扩展束可能需要使用管理员特权运行 IBM SPSS Modeler。您可以通过右键单击 IBM SPSS Modeler 的图标并选择**以管理员身份运行**，使用管理员特权来启动 IBM SPSS Modeler。尤其是，如果收到指示无法安装一个或多个扩展束的错误消息时，那么尝试使用管理员特权运行。

扩展的安装位置

缺省情况下，扩展会安装到操作系统的常规用户可写位置。

通过使用 `IBM_SPSS_MODELER_EXTENSIONS_PATH` 环境变量来定义路径，您可以覆盖缺省位置。指定的位置在目标计算机上必须存在。在设置 `IBM_SPSS_MODELER_EXTENSIONS_PATH` 后，必须重新启动 IBM SPSS Modeler 以使更改生效。

要在 Windows 上创建环境变量，请从控制面板：

Windows 7

1. 选择"用户帐户"。
2. 选择**更改我的环境变量**。
3. 单击**新建**，在**变量名称**字段中输入环境变量的名称（例如，`IBM_SPSS_MODELER_EXTENSIONS_PATH`），然后在"变量值"字段中输入路径。

Windows 8 或更高版本

1. 选择"系统"。
2. 选择"高级"选项卡并单击**环境变量**。从"高级系统设置"访问"高级"选项卡。
3. 在"用户变量"部分中，单击**新建**，在**变量名称**字段中输入环境变量的名称（例如，`IBM_SPSS_MODELER_EXTENSIONS_PATH`），然后在"变量值"字段中输入路径。

要点：对于 Windows 7 和更新版本的 Windows 用户，安装更新版本的现有扩展束可能需要使用管理员特权运行 IBM SPSS Modeler。您可以通过右键单击 IBM SPSS Modeler 的图标并选择**以管理员身份运行**，使用管理员特权来启动 IBM SPSS Modeler。尤其是，如果收到指示无法安装一个或多个扩展束的错误消息时，那么尝试使用管理员特权运行。

必需的 R 软件包

如果您不具有因特网访问权，那么将需要从具有该访问权的人员处获取特定扩展的任何在计算机上找不到的所需 R 软件包。安装扩展后，可以从"扩展详细信息"对话框查看所需 R 软件包的列表。请参阅主题 第 23 页的『扩展详细信息』以获取更多信息。可从 下载软件包，然后从 R 中安装。有关详细信息，请参阅随 R 一起分发的《R 安装和管理指南》。

注：对于 UNIX（包括 Linux）用户，以源格式下载软件包，然后进行编译。这需要您在机器上安装相应的工具。请参阅《R 安装和管理指南》以获取详细信息。尤其是，Debian 用户应从 `apt-get install r-base-dev` 安装 `r-base-dev` 软件包。

创建和管理定制节点

针对扩展的定制对话框构建器会创建节点，以便在 SPSS Modeler 流中使用。

使用针对扩展的定制对话框构建器，您可以：

- 创建定制节点对话框，以便执行以 R 或 Apache Spark（通过 Python）实现的节点。请参阅第 27 页的『构建脚本模板』，以了解更多信息。

- 打开包含定制节点对话框（此对话框可能由其他用户创建）规范的文件，并将此对话框添加到您的 IBM SPSS Modeler 安装中，从而选择性地对您自己的修改。
- 保存定制节点对话框的规范，以便其他用户可以将其添加到他们的 IBM SPSS Modeler 安装中。
- 创建定制节点和编写 Python for Spark 脚本，以从数据源所在位置读取数据，将数据写出到 Apache Spark 支持的任何数据格式。请参阅第 49 页的『使用 Python for Spark 导入和导出数据』，以了解更多信息。
- 创建定制节点和编写 R 脚本，以从数据源所在位置读取数据，将数据写出到 R 支持的任何数据格式。请参阅第 49 页的『使用 R 导入和导出数据』，以了解更多信息。

在"针对扩展的定制对话框构建器"中，可以在扩展中创建或修改定制节点对话框。打开"针对扩展的定制对话框构建器"时，会创建包含空的定制节点对话框的新扩展。从"针对扩展的定制对话框构建器"保存或安装定制节点对话框时，会将其作为扩展的一部分进行保存或安装。

注：

- 无法针对标准 IBM SPSS Modeler 节点创建您自己的节点对话框版本。
- 针对使用"定制对话框构建器"创建的节点（包括定制对话框构建器 R 节点和定制对话框构建器 Python 节点）不支持脚本编制。

如何启动"针对扩展的定制对话框构建器"

从菜单中选择 **扩展 > 定制节点对话框构建器**

注：

- Python 节点取决于 Spark 环境。
- Python 脚本必须使用 Spark API，因为将以 Spark DataFrame 形式表示数据。
- V17.1 中创建的旧节点仍将仅针对 IBM SPSS Analytic Server 运行（数据源自 IBM SPSS Analytic Server 源节点并且尚未抽取到 IBM SPSS Modeler 服务器）。V18.0 或更高版本中创建的新 Python 和"定制对话框构建器"可针对 IBM SPSS Modeler 服务器运行。
- 安装 Python 时，请确保所有用户都有权访问 Python 安装。
- 如果要使用"机器学习库 (MLlib)"，必须安装包含 NumPy 的 Python 版本。然后，必须配置 IBM SPSS Modeler 服务器（或 IBM SPSS Modeler 客户机中的本地服务器）以使用 Python 安装。有关详细信息，请参阅第 1 页的『使用 Python for Spark 编写脚本』。

定制对话框构建器布局

对话框画布

对话框画布是定制对话框构建器的一个区域，您可以在其中设计节点对话框的布局。

属性窗格

属性窗格是定制对话框构建器的一个区域，您可以在其中指定节点对话框中包含的控件的属性以及对话框自身的属性，例如节点类型。

工具选用板

工具选用板提供了一组可以包括在定制节点对话框中的控件。通过从查看菜单中选择"工具选用板"，可以显示或隐藏工具选用板。

脚本模板

脚本模板指定定制节点对话框生成的 R 脚本或 Python for Spark 脚本。您可以通过单击**移至新窗口**将"脚本模板"窗格移至单独的窗口。要将单独的"脚本模板"窗口移动回"定制对话框构建器"，请单击**复原到主窗口**。

构建定制节点对话框

构建定制节点对话框时涉及的基本步骤如下所示：

1. 指定节点对话框自身的属性，例如启动此节点对话框时显示的标题，以及新节点在 IBM SPSS Modeler 选用板中的位置。有关更多信息，请参阅『对话框属性』主题。
2. 指定构成此节点对话框以及任何子对话框的控件，例如字段选择器和复选框。有关更多信息，请参阅第 28 页的『控件类型』主题。
3. 创建脚本模板，用于指定节点对话框所生成的 R 代码或 Python for Spark 代码。有关更多信息，请参阅第 27 页的『构建脚本模板』主题。
4. 指定包含节点对话框的扩展的属性。有关更多信息，请参阅第 44 页的『扩展属性』主题。
5. 将包含节点对话框的扩展安装到 IBM SPSS Modeler 和/或将扩展保存到扩展束 (.mpe) 文件。请参阅主题第 46 页的『管理定制节点对话框』以获取更多信息。

在构建节点对话框时，可以进行预览。有关更多信息，请参阅第 28 页的『预览定制节点对话框』主题。

对话框属性

"定制对话框构建器"窗口显示了节点对话框的属性以及所选用户界面控件的属性。要查看和设置对话框属性，请单击画布上任何控件外的区域。如果画布上没有任何控件，那么对话框属性始终可见。

对话框名称。"对话框名称"属性是必填项，它指定要与节点对话框相关联的唯一名称。为了尽量降低名称冲突可能性，您可以使用贵公司的标识（例如 URL）作为名称的前缀。

标题。"标题"属性指定要显示在节点对话框的标题栏中的文本。

帮助文件。"帮助文件"属性是可选项，用于指定节点对话框的帮助文件的路径。这是用户单击对话框中的**帮助**按钮时将要启动的文件。帮助文件必须采用 HTML 格式。安装或保存节点对话框后，节点对话框的规范随附指定帮助文件的副本。如果没有相关联的帮助文件，那么运行时对话框上的"帮助"按钮将处于隐藏状态。

- 添加帮助文件时，与帮助文件存在于同一目录中的帮助文件的本地化版本会自动添加到节点对话框。帮助文件的本地化版本命名为 <Help File>_<language identifier>.htm。有关更多信息，请参阅主题第 48 页的『创建定制节点对话框的本地化版本』。
- 通过首先保存节点对话框，可以将支持文件（如图像文件和样式表）添加到节点对话框。然后，将支持文件手动添加到节点对话框文件 (.cfe)。有关访问和手动修改定制节点对话框文件的信息，请参阅主题第 48 页的『创建定制节点对话框的本地化版本』中标题为"本地化对话框字符串"的部分。

脚本类型。指定可用于构建脚本模板的脚本的类型。在 IBM SPSS Modeler 中，可使用 R 脚本编制或 Python for Spark 脚本编制。

使用模型进行评分。指定使用模型构建脚本构建的模型是否用于评分。

节点类型。指定安装节点对话框时将创建的节点的类型。

选用板。指定安装节点对话框时要将新创建的节点添加到的选用板。

节点图标。单击省略符 (...) 按钮可以选择要用作新创建的节点的节点图标的图像。选择的图像必须是 .gif 文件。

在对话框画布上布置控件

通过将控件从工具选用板拖放到对话框画布上，可以将这些控件添加到定制节点对话框。为确保与内置节点对话框的一致性，对话框画布划分为四个可在其中放置控件的功能列。

- 第一个（最左侧）列主要用于字段选择器控件。
- 子对话框列必须位于最右列（例如，如果仅使用三列那么为第三列），并且任何其他控件都不能位于与“子对话框”按钮相同的列中。因此，第四列只能包含“子对话框”按钮。

虽然未在对话框画布上显示，但将节点对话框安装到 IBM SPSS Modeler 时，相应的按钮将添加到对话框中（例如，**确定**、**取消**、**应用**、**重置**以及适当时添加的**帮助**和**运行**）。这些按钮的存在和位置都是自动的。但是，如果没有与节点对话框相关联的帮助文件（由对话框属性中的“帮助文件”属性指定），那么**帮助**按钮将处于隐藏状态。

通过向上或向下拖动控件，可以更改这些控件在一列中的垂直顺序，但是各个控件的确切位置将自动确定。在运行时，当对话框自身调整大小时，各个控件将相应地调整大小。字段选择器之类的控件将自动展开，以填充其下方的可用空间。

构建脚本模板

脚本模板指定定制节点对话框将生成的 R 脚本或 Python for Spark 脚本。可以使用单个定制节点对话框来指定一项或多项操作，这些操作将按顺序运行。

脚本模板可以包含静态文本。静态文本与静态文本控件不同；这是节点运行时始终生成的 R 或 Python for Spark 代码。例如，与用户输入无关的命令名和子命令规范是静态文本。另外，脚本模板还可以包含控件标识，在运行时，这些标识将替换为相关联定制节点对话框控件的值。例如，在字段选择器中指定的字段集将通过该字段选择器控件的控件标识来表示。

要构建脚本模板

1. 对于与用户指定的值无关的静态文本，请像在“R 构建”节点的 **R 模型构建语法** 字段之类的位置输入 R 脚本或 Python for Spark 脚本一样输入该脚本。
2. 在要插入由控件生成的 R 脚本或 Python for Spark 脚本的位置添加格式为 `%Identifier%` 的控件标识，其中 `Identifier` 是控件的“标识”属性值。
 - 您可以通过选中标识表中的某个行、右键单击并选择**添加到脚本模板**来插入控件标识。您还可以通过右键单击画布上的控件并选择**添加到脚本模板**来插入控件标识。
 - 您还可以按 `Ctrl + 空格` 键来从可用控件标识列表中进行选择。此列表依次包含控件标识以及通过脚本自动补全功能提供的项。

如果您手动输入标识，请保留所有空格，这是因为标识中的所有空格都有意义。

在运行时，对于除复选框、复选框组和静态文本控件以外的所有控件，每个标识都将替换为相关联控件的脚本属性的当前值。如果控件在运行时为空，那么不会生成任何脚本。对于复选框和复选框组，该标识将替换为相关联控件的“选中 R 脚本”或“未选中 R 脚本”属性的当前值，具体取决于该控件的当前状态 - 选中或未选中。有关更多信息，请参阅第 28 页的『控件类型』主题。

示例：在 R 脚本模板中包含运行时值

在此示例中，定制节点对话框将生成并运行 R 脚本，并借助此处显示的特征符对 `R lm` 函数进行调用，从而构建线性回归模型并对该模型进行评分。

```
lm(formula,data)
```

- *formula* 用于指定 $Na \sim Age$ 之类的表达式，其中 *Na* 是模型的目标字段，*Age* 是模型的输入字段。
- *data* 是一个数据帧，其中包含公式中指定的字段的值。

假定有一个定制节点对话框，其中包含单个字段选择器控件，该控件允许用户选择线性模型的输入字段。您在脚本选项卡上输入脚本模板，此脚本模板生成并运行用于构建模型的 R 脚本，并且可能如下所示：

```
modelerModel <- lm(Na~%input%,data=modelerData)
```

- `%input%` 是字段选择器控件的"标识"属性的值。在运行时，它将替换为该控件的脚本属性的当前值。
- 将字段选择器控件的脚本属性定义为 `%ThisValue%` 表示，在运行时，此属性的当前值将是该控件的值，也就是从字段选择器中选择的字段。

假设此定制节点对话框的用户选择 *Age* 字段作为模型的输入字段。节点对话框将生成以下 R 脚本：

```
modelerModel <- lm(Na~Age,data=modelerData)
```

您在评分脚本选项卡上输入脚本模板，此脚本模板生成并运行用于对模型进行评分的 R 脚本，并且可能如下所示：

```
result <- predict(modelerModel,newdata=modelerData)
var1 <-c(fieldName="predicted", fieldLabel="",fieldStorage="real",fieldMeasure="",fieldFormat="",
fieldRole="")
modelerDataModel<-data.frame(modelerDataModel,var1)
```

这个 R 脚本不依赖于任何由用户指定的值，而仅依赖于使用模型构建 R 脚本构建的模型。因此，就像在"R 构建"节点的 **R 模型评分语法** 字段中输入模型评分 R 脚本一样输入该脚本。

预览定制节点对话框

您可以预览当前在定制对话框构建器中打开的节点对话框。此对话框的显示和运行方式就像是从 IBM SPSS Modeler 中的节点运行一样。

- 字段选择器以虚字段进行填充。
- **确定**按钮用于关闭预览。
- 如果指定了帮助文件，那么**帮助**按钮将处于启用状态，并可以打开指定的文件。如果未指定帮助文件，那么帮助按钮在预览时处于禁用状态，而在实际对话框运行时处于隐藏状态。

要预览定制节点对话框，请从定制对话框构建器的菜单中选择**文件 > 预览对话框**。

控件类型

工具选用板提供了定制节点对话框中可能需要的所有标准控件。

- **字段选择器**：活动数据集中所有字段的列表。有关更多信息，请参阅第 29 页的『字段选择器』主题。
- **复选框**：单个复选框。有关更多信息，请参阅第 30 页的『复选框』主题。
- **组合框**：这是用于创建下拉列表的组合框。有关更多信息，请参阅第 31 页的『组合框』主题。
- **列表框**：这是用于创建单选列表或多选列表的列表框。有关更多信息，请参阅第 31 页的『组合框』主题。
- **文本控件**：这是接受任意文本作为输入的文本框。有关更多信息，请参阅第 33 页的『"文本"控件』主题。
- **数字控件**：这是只能接受数字值作为输入的文本框。有关更多信息，请参阅第 34 页的『"数字"控件』主题。

- **数据控件**：用于指定日期/时间值（包括日期、时间和日期时间）的微调控件。有关更多信息，请参阅第 35 页的『"日期"控件』主题。
- **受保护的文本**：用星号屏蔽用户输入的文本框。有关更多信息，请参阅第 35 页的『受保护的文本』主题。
- **静态文本控件**：这是用于显示静态文本的控件。有关更多信息，请参阅第 36 页的『"静态文本"控件』主题。
- **颜色选择器**：用于指定颜色和生成关联的 RGB 值的控件。有关更多信息，请参阅第 36 页的『颜色选择器』主题。
- **表控件**：具有固定列数和可在运行时添加的可变行数的表。有关更多信息，请参阅第 37 页的『表控件』主题。
- **项目组**：这是一个容器，用于对一组控件（例如一组复选框）进行分组。有关更多信息，请参阅第 39 页的『项目组』主题。
- **单选按钮组**：一组单选按钮。有关更多信息，请参阅第 39 页的『单选按钮组』主题。
- **复选框组**：这是包含一组控件的容器，这些控件由单个复选框成组启用或禁用。有关更多信息，请参阅第 40 页的『复选框组』主题。
- **文件浏览器**：这是用于浏览文件系统以打开或保存文件的控件。有关更多信息，请参阅第 41 页的『文件浏览器』主题。
- **选项卡**：单个选项卡。有关更多信息，请参阅第 42 页的『选项卡』主题。
- **子对话框按钮**：这是用于启动子对话框的按钮。有关更多信息，请参阅第 42 页的『子对话框按钮』主题。

字段选择器

"字段选择器"控件显示适用于最终用户的节点对话框的字段列表。您可以显示活动数据集中的所有字段（缺省情况），也可以根据类型和测量级别对列表进行过滤 - 例如，测量级别为比例的数字字段。您还可以将任何其他"字段选择器"指定为当前"字段选择器"的字段源。"字段选择器"控件具有下列属性：

标识。 控件的唯一标识。

标题。 控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

标题位置。 指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。仅在设置选择器类型以选择单个字段时此属性才适用。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。仅当鼠标指针悬停在此控件的标题区域上时，才会显示指定的文本。如果将鼠标指针悬停在列出的某个字段上，那么将显示字段名称和标签。

助记键。 这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 Alt+[助记键] 激活。

选择器类型。 指定定制节点对话框中的字段选择器可用于从字段列表中选择单个字段还是多个字段。

分隔符类型。 指定生成的脚本中选中的字段之间的定界符。允许的允许为空格、逗号和加号 (+)。您还可以输入任意单个字符以用作分隔符。

最小字段数。 必须为控件（如果有）指定的最小字段数。

最大字段数。 可以为控件（如果有）指定的最小字段数。

执行所必需。指定此控件中是否需要值才能继续执行。如果指定了 **True**，那么节点对话框的用户必须为此控件指定值，否则单击**确定**按钮将出错。如果指定了 **False**，那么此控件中缺少值并不会对**确定**按钮的状态造成任何影响。

变量过滤器。允许您对此控件中显示的字段集进行过滤。可以按字段类型和测量级别进行过滤，并可以指定将多个响应集包括在字段列表中。单击省略符 (...) 按钮将打开"过滤器"对话框。您也可以通过双击画布上的"字段选择器"控件来打开"过滤器"对话框。有关更多信息，请参阅『对字段列表进行过滤』主题。

资源源。指定另一个"字段选择器"是当前"字段选择器"的字段源。在未设置"字段源"属性时，字段源为活动数据集。单击省略符 (...) 按钮以打开对话框并指定字段源。

脚本。指定由此控件生成并运行脚本，并且可插入到脚本模板中。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 `%%ThisValue%%` 指定此控件的运行时值，即字段列表。这是缺省值。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

为"字段选择器"指定"字段源"："字段源"对话框可指定"字段选择器"中显示的资源的源。源可以是任何其他"字段选择器"。您可以从活动数据集选择显示包含在选定控件中的字段或不包含在选定控件中的的字段。

对字段列表进行过滤

通过使用与字段选择器控件相关联的"过滤器"对话框，可以对活动数据集中能够在列表中显示的 字段类型进行过滤。您还可以指定是否包括多个与此活动数据集相关联的响应集。数字字段包括除日期格式和时间格式以外的所有数字格式。

复选框

"复选框"控件是一个简单的复选框，它可以针对选中和未选中状态生成并运行不同的 R 脚本或 Python for Spark 脚本。"复选框"控件具有下列属性：

标识。控件的唯一标识。

标题。控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 `Alt+[助记键]` 激活。

缺省值。这是复选框的缺省状态 - 选中或未选中。

选中/未选中脚本。指定在选中以及未选中此控件的情况下，生成并运行的 R 脚本或 Python for Spark 脚本。要将脚本包括在脚本模板中，请使用"标识"属性的值。生成的脚本将插入到该标识的指定位置，而与它是根据"选中脚本"属性还是"未选中脚本"属性生成无关。例如，如果标识为 `checkbox1`，那么在运行时，脚本模板中 `%%checkbox1%%` 的实例将替换为"选中的脚本"属性的值（在选中框时）和"未选中的脚本"属性的值（在未选中框时）。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

组合框

使用"组合框"控件可以创建下拉列表，此下拉列表可以生成并运行特定于所选列表项的 R 脚本和 Python for Spark 脚本。此控件只允许进行单选。"组合框"控件具有下列属性：

标识。控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

标题位置。指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

列表项。单击省略符 (...) 按钮以打开列表项属性对话框，这允许您指定控件的列表项。您还可以通过双击画布上的"组合框"控件来打开"列表项属性"对话框。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 Alt+[助记键] 激活。

可编辑。指定"组合框"控件是否可编辑。在控件可编辑时，可在运行时输入定制值。

脚本。指定运行时此控件生成且可插入脚本模板的 R 脚本或 Python for Spark 脚本。

- 值 `%%ThisValue%%` 指定控件的运行时值，也是缺省值。如果手动定义列表项，那么运行时值是所选列表项的"脚本"属性的值。如果列表项以目标列表控件为基础，那么运行时值是所选列表项的值。对于多选列表框控件，运行时值是以空格分隔的所选项列表。有关更多信息，请参阅『指定组合框和列表框的列表项』主题。
- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。

引号处理。指定在脚本属性包含 `%%ThisValue%%` 作为加引号的字符串的一部分时，如何处理运行时值 `%%ThisValue%%` 中的引号。在此上下文中，加引号的字符串是用单引号或双引号引起来的字符串。引号处理仅应用于用 `%%ThisValue%%` 引起来的类型为引号的引号。以下类型的引号处理可用。

Python

匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，如果脚本属性为"`%%ThisValue%%`"，并且组合框的运行时值为 Combo box's value，那么生成的脚本为 `'Combo box\'s value'`。请注意，在三重引号括起 `%%ThisValue%%` 时，不执行引号处理。

R 匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，如果脚本属性为"`%%ThisValue%%`"，并且组合框的运行时值为 Combo box's value，那么生成的脚本为 `'Combo box\'s value'`。

无 匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将保留不做任何修改。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

指定组合框和列表框的列表项： 使用"列表项属性"对话框可以指定组合框或列表框控件的列表项。

手动定义的值。允许您显式指定各个列表项。

- **标识。**列表项的唯一标识。
- **名称。**在列表中针对这个项显示的名称。名称是必填字段。
- **缺省。**对于组合框，指定此列表项是否为组合框中显示的缺省项。对于列表框，指定缺省情况下是否选中此列表项。
- **脚本。**指定选择此列表项时生成的 R 脚本或 Python for Spark 脚本。
- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。

注：可以在现有列表底部的空白行中添加新的列表项。输入除标识以外的任何属性都将生成一个唯一标识，您可以保留该唯一标识，也可以对其进行修改。通过单击列表项的标识单元格并按 Delete 键，可以删除该项。

列表框

使用"列表框"控件可以显示项目列表，这些项目支持单选或多选，并将生成特定于所选项目的 R 脚本和 Python for Spark 脚本。"列表框"控件具有以下属性：

标识。控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

列表项。单击省略符 (...) 按钮以打开列表项属性对话框，这允许您指定控件的列表项。您还可以通过双击画布上的"列表框"控件来打开"列表项属性"对话框。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带下划线。快捷键通过按下 Alt+[助记键] 激活。

列表框类型。指定列表框是仅支持单选还是支持多选。您还可以指定将这些项目显示为复选框列表。

分隔符类型。指定生成的脚本中所选列表项之间的的定界符。允许的允许为空格、逗号和加号 (+)。您还可以输入任意单个字符以用作分隔符。

所选最小值。在控件（如果有）中必须选择的最小项数。

所选最大值。在控件（如果有）中可以选择的最大项数。

脚本。指定运行时此控件生成且可插入脚本模板的 R 脚本或 Python for Spark 脚本。

- 值 `%ThisValue%` 指定控件的运行时值，也是缺省值。如果手动定义列表项，那么运行时值是所选列表项的"脚本"属性的值。如果列表项以目标列表控件为基础，那么运行时值是所选列表项的值。对于多选列表框控件，运行时值是以指定的分隔符类型分隔（缺省以空格分隔）的所选项的列表。有关更多信息，请参阅第 31 页的『指定组合框和列表框的列表项』主题。
- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。

引号处理。指定在脚本属性包含 `%ThisValue%` 作为加引号的字符串的一部分时，如何处理运行时值 `%ThisValue%` 中的引号。在此上下文中，加引号的字符串是用单引号或双引号引起来的字符串。引号处理仅应用于用 `%ThisValue%` 引起来的类型为引号的引号。以下类型的引号处理可用。

Python

匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，如果脚本属性为"`%%ThisValue%%`"，并且所选列表项为 `List item's value`，那么生成的脚本为 `'Listitem\'s value'`。请注意，在三重引号括起 `%%ThisValue%%` 时，不执行引号处理。

R 匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，如果脚本属性为"`%%ThisValue%%`"，并且所选列表项为 `List item's value`，那么生成的脚本为 `'Listitem\'s value'`。

无 匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将保留不做任何修改。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

"文本"控件

"文本"控件是一个简单的文本框，它可以接受任意输入并具有下列属性：

标识。控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

标题位置。指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 `Alt+[助记键]` 激活。

文本内容。指定内容是否为任意内容，或者指定文本框是否必须包含符合 IBM SPSS Modeler 字段名称规则的字符串。

缺省值。文本框的缺省内容。

宽度。指定控件文本区域的宽度（字符）。允许的值为正整数。空值表示自动确定宽度。

执行所必需。指定此控件中是否需要值才能继续执行。如果指定了 **True**，那么节点对话框的用户必须为此控件指定值，否则单击**确定**按钮将出错。如果指定了 **False**，那么此控件中缺少值并不会对**确定**按钮的状态造成任何影响。缺省值为 **False**。

脚本。指定由此控件在运行时生成并运行且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 `%%ThisValue%%` 指定控件的运行时值，即文本框的内容。这是缺省值。
- 如果"脚本"属性包含 `%%ThisValue%%`，并且文本框的运行时值为空，那么文本框控件不会生成任何脚本。

引号处理。指定在脚本属性包含 `%%ThisValue%%` 作为加引号的字符串的一部分时，如何处理运行时值 `%%ThisValue%%` 中的引号。在此上下文中，加引号的字符串是用单引号或双引号引起来的字符串。引号处理仅应用于用 `%%ThisValue%%` 引起来的类型为引号的引号。以下类型的引号处理可用。

Python

匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，

如果脚本属性为 '%ThisValue%', 并且文本控件的运行时值为 Text box's value, 那么生成的脚本为 'Text box\'s value'。在三重引号括起 %ThisValue% 时, 不执行引号处理。

R 匹配嵌套引号的 %ThisValue% 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如, 如果脚本属性为 '%ThisValue%', 并且文本控件的运行时值为 Text box's value, 那么生成的脚本为 'Text box\'s value'。

无 匹配嵌套引号的 %ThisValue% 运行时值中的引号将保留不做任何修改。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时, 才会显示"启用规则"属性。

"数字"控件

"数字"控件是用于输入数字值的文本框, 它具有下列属性:

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 控件上方显示的可选标题。对于多行标题, 请使用 \n 来指定换行符。

标题位置。 指定标题相对于控件的位置。值为"顶部"和"左侧", 其中"顶部"为缺省值。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。 这是标题中的一个可选字符, 用作控件的键盘快捷键。在标题中, 此字符显示为带下划线。快捷键通过按下 Alt+[助记键] 激活。

数字类型。 指定对可以输入的内容的任何限制。值"实数"指定输入的值必须是数字, 此外没有任何限制。值"整数"指定值必须是整数。

下拉输入。 指定控件是否显示为下拉列表。缺省值为 False。

增量。 在控件显示为下拉列表时的增量。

缺省值。 缺省值 (如果有)。

最小值。 允许的最小值 (如果有)。

最大值。 允许的最大值 (如果有)。

宽度。 指定控件文本区域的宽度 (字符)。允许的正整数。空值表示自动确定宽度。

执行所必需。 指定此控件中是否需要值才能继续执行。如果指定了 **True**, 那么节点对话框的用户必须为此控件指定值, 否则单击**确定**按钮将出错。如果指定了 **False**, 那么此控件中缺少值并不会对**确定**按钮的状态造成任何影响。缺省值为 **False**。

脚本。 指定由此控件在运行时生成并运行且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本, 单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 %ThisValue% 指定控件的运行时值, 这是数字值。这是缺省值。
- 如果"脚本"属性包含 %ThisValue%, 并且"数字"控件的运行时值为空, 那么"数字"控件不会生成任何脚本。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时, 才会显示"启用规则"属性。

"日期"控件

"日期"控件是一种用于指定日期/时间值的微调控件，其中包含日期、时间和日期时间。"日期"控件具有下列属性：

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

标题位置。 指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。 这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下列划线。快捷键通过按下 Alt+[助记键] 激活。

类型。 指定控件是用于日期、时间还是日期时间值。

日期 控件指定 yyyy-mm-dd 格式的日历日期。"缺省值"属性指定缺省运行时值。

时间 控件指定 hh:mm:ss 格式的一天的时间。缺省运行时值是一天的当前时间。

日期时间

控件指定 yyyy-mm-dd hh:mm:ss 格式的日期和时间。缺省运行时值是当前日期和一天的时间。

缺省值。 在类型为"日期"时控件的缺省运行时值。您可以指定是显示当前日期还是特定日期。

脚本。 指定由此控件在运行时生成并运行且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 %ThisValue% 指定控件的运行时值。这是缺省值。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

注："日期"控件在 IBM SPSS Modeler R18 之前的发行版中不受支持。

受保护的文本

"受保护的文本"控件是用星号屏蔽用户数输入的文本框。

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

标题位置。 指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。 这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下列划线。快捷键通过按下 Alt+[助记键] 激活。

宽度。 指定控件文本区域的宽度（字符）。允许的正整数。空值表示自动确定宽度。

执行所必需。 指定此控件中是否需要值才能继续执行。如果指定了 **True**，那么节点对话框的用户必须为此控件指定值，否则单击**确定**按钮将出错。如果指定了 **False**，那么此控件中缺少值并不会对**确定**按钮的状态造成任何影响。缺省值为 **False**。

脚本。 指定由此控件在运行时生成并运行且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 `%%ThisValue%%` 指定控件的运行时值，即文本框的内容。这是缺省值。
- 如果脚本属性包含 `%%ThisValue%%` 并且受保护的文本控件的运行时值为空，那么受保护的文本控件不会生成任何 R 脚本或 Python for Spark 脚本。

引号处理。 指定在脚本属性包含 `%%ThisValue%%` 作为加引号的字符串的一部分时，如何处理运行时值 `%%ThisValue%%` 中的引号。在此上下文中，加引号的字符串是用单引号或双引号引起来的字符串。引号处理仅应用于用 `%%ThisValue%%` 引起来的类型为引号的引号，并且仅在 `Encrypt passed value=False` 时适用。以下类型的引号处理可用。

Python

匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，如果脚本属性为 `'%%ThisValue%%'`，并且控件的运行时值为 `Secured Text's value`，那么生成的脚本为 `'Secured Text\'s value'`。在三重引号括起 `%%ThisValue%%` 时，不执行引号处理。

R 匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将使用反斜杠字符 (\) 进行转义。例如，如果脚本属性为 `'%%ThisValue%%'`，并且控件的运行时值为 `Secured Text's value`，那么生成的脚本为 `'Secured Text\'s value'`。

无 匹配嵌套引号的 `%%ThisValue%%` 运行时值中的引号将保留不做任何修改。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

注："受保护文本"控件在 IBM SPSS Modeler R18 之前的发行版中不受支持。

"静态文本"控件

使用"静态文本"控件可以向节点对话框添加文本块，此控件具有下列属性：

标识。 控件的唯一标识。

标题。 文本块的内容。对于多行内容，使用 \n 来指定换行符。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

颜色选择器

"颜色选择器"控件是用于指定颜色和生成关联的 RGB 值的用户界面。"颜色选择器"控件具有下列属性：

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

标题位置。 指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 Alt+[助记键] 激活。

脚本。指定由此控件在运行时生成并运行且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 `%ThisValue%` 指定控件的运行时值，即选中的颜色的 RGB 值。RGB 值表示为以下顺序的空格分隔的整数列表：R 值、G 值、B 值。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

注："颜色选择器"控件在 IBM SPSS Modeler R18 之前的发行版中不受支持。

表控件

"表"控件创建具有固定列数和可在运行时添加的可变行数的表。"表"控件具有以下属性：

标识。控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。控件上方显示的可选标题。对于多行标题，请使用 \n 来指定换行符。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 Alt+[助记键] 激活。

重新排序按钮。指定是否向表添加上移和下移按钮。在运行时可使用这些按钮来重新排序表行。

表列。单击省略符 (...) 按钮以打开表列对话框，您可以在其中指定表列。

最小行数。必须位于表中的最小行数。

最大行数。可位于表中的最大行数。

执行所必需。指定此控件中是否需要值才能继续执行。如果指定了 **True**，那么节点对话框的用户必须为此控件指定值，否则单击**确定**按钮将出错。如果指定了 **False**，那么此控件中缺少值并不会对**确定**按钮的状态造成任何影响。

脚本。指定运行时此控件生成且可插入脚本模板的 R 脚本或 Python for Spark 脚本。

- 值 `%ThisValue%` 指定控件的运行时值，也是缺省值。运行时值是表中每列生成的脚本的空格分隔的列表，从最左列开始。如果脚本属性包含 `%ThisValue%` 并且没有任何列生成脚本，那么整个表不会生成任何脚本。
- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

注："表"控件在 IBM SPSS Modeler R18 之前的发行版中不受支持。

指定表控件的列： "表列"对话框指定表控件的列的属性。

标识。 列的唯一标识。

列名。 表中显示的列名。

内容。 指定列的数据类型。值**实数**指定输入的值必须是数字，此外没有任何限制。值**整数**指定值必须是整数。值**任何**指定对输入值没有任何限制。值**变量名称**指定该值必须满足 IBM SPSS Statistics 中的有效变量名称的需求。

缺省值。 在运行时将新行添加到表时此列的缺省值（如果有）。

分隔符类型。 在生成的脚本中列的值之间指定定界符。允许的允许为空格、逗号和加号 (+)。您还可以输入任意单个字符以用作分隔符。

加引号。 指定在生成的脚本中是否使用双引号将列中每个值引起。

引号处理。 指定"加引号"属性为 true 时处理列的单元格输入的引号。引号处理仅适用于单元格值中双引号。以下类型的引号处理可用。

Python

单元格值中的双引号使用反斜杠字符 (\) 进行转义。例如，如果单元格值为 This "quoted" value，那么生成的脚本为 "This \"quoted\" value"。

R

单元格值中的双引号使用反斜杠字符 (\) 进行转义。例如，如果单元格值为 This "quoted" value，那么生成的脚本为 "This \"quoted\" value"。

无

单元格值中的双引号将保留不做修改。

宽度（字符）。 指定列的宽度（字符数）。允许的值为正整数。

脚本。 指定在运行时此列生成的 R 脚本或 Python for Spark 脚本。针对表整体生成的脚本是一个表中每个列（从最左侧的列开始）生成的脚本的以空格分隔的列表。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 %%ThisValue%% 指定列的运行时值，即列中的值的列表，以指定的分隔符来定界。
- 如果列的脚本属性包含 %%ThisValue%% 并且列的运行时值为空，那么此列不生成任何脚本。

注： 您可以在"表列"对话框中的现有列表底部的空白行中为新的"表列"添加行。输入标识以外的任何属性会生成唯一的标识，您可以保留或修改此标识。您可以通过单击"表"列的标识单元并按 Delete 键来删除"表"列。

链接至控件

您可以将"表"控件链接至"字段选择器"控件。将"表"控件链接至"字段选择器"时，在表中针对"字段选择器"中每个字段存在一个行。通过向"字段选择器"添加字段将行添加到表。通过从"字段选择器"中除去字段从表中删除行。例如，已链接的"表"控件可用于指定"字段选择器"中选中的字段的属性。

要启用链接，表必须包含针对"内容"属性的"变量名称"列，并且画布上必须存在至少一个"字段选择器"控件。

要将"表"控件链接至"字段选择器"，请从"表列"对话框上的"链接至控件"组中的可用控件列指定"字段选择器"。然后，选择定义链接的表列（称为**已链接的列**）。呈现表时，已链接的列会显示"字段选择器"中的当前字段。仅可以链接多字段字段选择器。

项目组

"项目组"控件是包含其他控件的容器，使用此控件可以对多个控件生成的脚本进行分组和控制。例如，您有一组用于为子命令指定可选设置的复选框，但希望只有在至少选中了一个复选框的情况下，才生成该子命令的脚本。这通过使用"项目组"控件作为复选框控件的容器来实现。项目组中可以包含以下类型的控件：字段选择器、复选框、组合框、列表框、文本控件、数字控件、静态文本、单选按钮和文件浏览器。"项目组"控件具有下列属性：

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 组的可选标题。对于多行标题，请使用 \n 来指定换行符。

脚本。 指定由此控件在运行时生成并运行且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 可以包括任何包含在项目组中的控件的标识。在运行时，这些标识将替换为控件所生成的 R 脚本或 Python 脚本。
- 值 `%%ThisValue%%` 将生成项目组中各个控件所生成的 R 脚本或 Python 脚本的列表（以空格分隔），并按其相应控件在此组中的显示顺序排列（从上到下）。这是缺省值。如果脚本属性包含 `%%ThisValue%%` 并且项目组中的任何控件未生成脚本，那么整个项目组不会生成任何脚本。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

单选按钮组

"单选按钮组"控件是包含一组单选按钮的容器，其中每个单选按钮都可以包含一组嵌套控件。"单选按钮组"控件具有下列属性：

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 组的可选标题。对于多行标题，请使用 \n 来指定换行符。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

单选按钮。 单击省略符 (...) 按钮可以打开单选按钮组属性对话框，使用此对话框可以指定单选按钮的属性以及在组中添加或删除按钮。在给定单选按钮下方嵌套控件的能力是单选按钮的一个属性，此属性在"单选按钮组属性"对话框中设置。请注意，您还可以双击画布上的"单选组按钮"控件来打开"单选组属性"对话框。

脚本。 指定运行时此控件生成且可插入脚本模板的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 `%%ThisValue%%` 指定单选按钮组的运行时值，即选中的单选按钮的脚本属性的值。这是缺省值。如果脚本属性包含 `%%ThisValue%%` 并且选中的单选按钮未生成任何脚本，那么单选按钮不会生成任何脚本。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

定义单选按钮： 使用"单选按钮组属性"对话框可以指定一组单选按钮。

标识。 单选按钮的唯一标识。

列名。在单选按钮旁边显示的名称。名称是必填字段。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。这是名称中的一个可选字符，用作助记符。指定的字符在名称中必须存在。

嵌套组。指定是否可以将其他控件嵌套在此单选按钮之下。缺省值为 `false`。如果将嵌套组属性设置为 `true`，那么将在相关单选按钮下方显示嵌套并缩进的矩形放置区。单选按钮下可嵌套以下控件：字段选择器、复选框、文本控件、静态文本、数字控件、组合框、列表框和文件浏览器。

缺省。指定此单选按钮是否为缺省选择。

启用规则。指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

脚本。指定选择此单选按钮时生成的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 对于包含嵌套控件的单选按钮，值 `%ThisValue%` 将生成各个嵌套控件所生成的 R 脚本或 Python for Spark 脚本的列表（以空格分隔），并按其相应控件在此单选按钮下的显示顺序排列（从上到下）。

可以在现有列表底部的空白行中添加新的单选按钮。输入除标识以外的任何属性都将生成一个唯一标识，您可以保留该唯一标识，也可以对其进行修改。通过单击单选按钮的标识单元格并按 `Delete` 键，可以删除该按钮。

复选框组

"复选框组"控件是包含一组控件的容器，这些控件由单个复选框成组启用或禁用。"复选框组"中可包含以下类型的控件：字段选择器、复选框、组合框、列表框、文本控件、数字控件、静态文本、单选按钮和文件浏览器。"复选框组"控件具有下列属性：

标识。控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。组的可选标题。对于多行标题，请使用 `\n` 来指定换行符。

复选框标题。随控制复选框一起显示的可选标签。支持使用 `\n` 来指定换行符。

工具提示。这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 `Alt+[助记键]` 激活。

缺省值。这是控制复选框的缺省状态 - 选中或未选中。

选中/未选中 R 脚本。指定在选中以及未选中此控件的情况下，生成的 R 脚本。要将 R 脚本包括在脚本模板中，请使用"标识"属性的值。生成的 R 脚本将插入到该标识的指定位置，而与它是根据"选中 R 脚本"属性还是"未选中 R 脚本"属性生成无关。例如，如果标识为 `checkboxgroup1`，那么在运行时，脚本模板中 `%checkboxgroup1%` 的实例将替换为"选中的 R 脚本"属性的值（在选中框时）和"未选中的 R 脚本"属性的值（在未选中框时）。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 可以包括任何包含在复选框组中的控件的标识。在运行时，这些标识将替换为控件所生成的 R 脚本。

- 值 `%ThisValue%` 可以用于"选中 R 脚本"属性，也可以用于"未选中 R 脚本"属性。它将生成复选框组中各个控件所生成的 R 脚本的列表（以空格分隔），并按其相应控件在此组中的显示顺序排列（从上到下）。
- 缺省情况下，"选中 R 脚本"属性的值为 `%ThisValue%`，而"未选中 R 脚本"属性的值为空。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

文件浏览器

"文件浏览器"控件包含用于输入文件路径的文本框，以及用于打开标准 IBM SPSS Modeler 对话框以打开或保存文件的浏览按钮。"文件浏览器"控件具有下列属性：

标识。 控件的唯一标识。这是在脚本模板中引用控件时使用的标识。

标题。 控件上方显示的可选标题。对于多行标题，请使用 `\n` 来指定换行符。

标题位置。 指定标题相对于控件的位置。值为"顶部"和"左侧"，其中"顶部"为缺省值。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

助记键。 这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 `Alt+[助记键]` 激活。

文件系统操作。 指定通过浏览按钮启动的对话框是适合于打开文件还是保存文件。值"打开"表示浏览对话框将验证指定的文件是否存在。值"保存"表示浏览对话框不验证指定的文件是否存在。

浏览器类型。 指定浏览对话框是用于选择文件（查找文件）还是选择文件夹（查找文件夹）。

文件过滤器。 单击省略符 (...) 按钮可以打开文件过滤器对话框，使用此对话框可以指定适用于打开或保存对话框的文件类型。缺省情况下，允许使用所有文件类型。请注意，您还可以双击画布上的"文件浏览器"控件来打开"文件过滤器"对话框。

文件系统类型。 在分布式分析方式下，此属性指定打开或保存对话框是浏览运行 IBM SPSS Modeler Server 的文件系统还是浏览本地计算机的文件系统。选择**服务器**表示浏览服务器的文件系统，选择**客户机**表示浏览本地计算机的文件系统。在本地分析方式下，此属性无效。

执行所必需。 指定此控件中是否需要值才能继续执行。如果指定了 **True**，那么节点对话框的用户必须为此控件指定值，否则单击**确定**按钮将出错。如果指定了 **False**，那么此控件中缺少值并不会对**确定**按钮的状态造成任何影响。缺省值为 **False**。

缺省。 控件的缺省值。

脚本。 指定运行时此控件生成且可插入脚本模板的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本。对于多行脚本或长脚本，单击省略符 (...) 按钮并在脚本属性对话框中输入您的脚本。
- 值 `%ThisValue%` 指定文本框的运行时值，这是嵌套双引号的文件路径，通过手动指定或由浏览器对话框填充。这是缺省值。
- 如果"脚本"属性包含 `%ThisValue%`，并且文本框的运行时值为空，那么文件浏览器控件不会生成任何脚本。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

文件类型过滤器： 使用"文件过滤器"对话框可以指定一些文件类型，这些文件类型将显示在通过"文件系统浏览器"控件访问的打开及保存对话框的"文件类型"和"另存为类型"下拉列表中。缺省情况下，允许使用所有文件类型。

要指定未在对话框中显式列出的文件类型，请完成下列步骤：

1. 选择"其他"。
2. 输入文件类型的名称。
3. 使用 *.suffix 格式输入文件类型 - 例如 *.xls。您可以指定多种文件类型，各种文件类型之间以分号分隔。

选项卡

"选项卡"控件用于向节点对话框添加选项卡。可以向新选项卡添加任何其他控件。"选项卡"控件具有下列属性：

标识。 控件的唯一标识。

标题。 选项卡的标题。

位置。 在节点对话框上指定选项卡相对于节点对话框上的其他选项卡的位置。

脚本。 指定在运行时由此控件生成和运行，并且可以插入到脚本模板中的 R 脚本或 Python for Spark 脚本。

- 您可以指定任何有效的 R 脚本或 Python for Spark 脚本，并可以使用 \n 作为换行符。
- 值 %ThisValue% 生成选项卡中每个控件生成的 R 脚本或 Python for Spark 脚本的空格分隔的列表，并按照选项卡中出现的顺序进行排序（从上到下，从左到右）。这是缺省值。
- 如果"脚本"属性包含 %ThisValue%，并且此选项卡中的所有控件均未生成 R 脚本或 Python for Spark 脚本，那么选项卡整体不生成任何脚本。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

子对话框按钮

"子对话框按钮"控件指定用于启动子对话框的按钮，并允许针对这个子对话框来访问对话框构建器。"子对话框按钮"具有下列属性：

标识。 控件的唯一标识。

标题。 在按钮中显示的文本。

工具提示。 这是用户将鼠标指针悬停在控件上时显示的可选工具提示文本。

子对话框。 单击省略符 (...) 按钮可以针对子对话框打开定制对话框构建器。另外，您也可以通过双击子对话框按钮来打开此构建器。

助记键。 这是标题中的一个可选字符，用作控件的键盘快捷键。在标题中，此字符显示为带有下划线。快捷键通过按下 Alt+[助记键] 激活。

启用规则。 指定确定何时启用当前控件的规则。单击省略符 (...) 按钮以打开对话框并指定规则。仅当画布上存在可用于指定启用规则的其他控件时，才会显示"启用规则"属性。

注： 不能向子对话框添加"子对话框按钮"控件。

子对话框的对话框属性： 要查看和设置子对话框的属性，请完成下列步骤：

1. 通过在主对话框中双击子对话框的按钮来打开子对话框，或者单击子对话框按钮，然后单击"子对话框"属性的省略符 (...) 按钮。
2. 在子对话框中，单击画布上任何控件外的区域。如果画布上没有任何控件，那么子对话框的属性始终可见。

子对话框名称。子对话框的唯一标识。"子对话框名称"属性是必填项。

注：如果在 %My Sub-dialog Name% 中，指定"子对话框名称"作为脚本模板中的标识，那么在运行时将替换为子对话框中每个控件生成的脚本的空格分隔的列表，采用显示顺序排序（从上到下，从左到右）。

标题。指定要在子对话框的标题栏中显示的文本。"标题"属性是可选项，但建议您指定此属性。

帮助文件。指定子对话框的可选帮助文件的路径。这是用户单击子对话框中的帮助按钮时启动的文件，并可以就是对主对话框指定的帮助文件。帮助文件必须采用 HTML 格式。有关更多信息，请参阅对话框属性的"帮助文件"属性的相关描述。

为控件指定"启用规则"

您可以指定用于确定何时启用控件的规则。例如，您可以指定填充字段选择器时启用单选按钮组。用于指定启用规则的可用选项取决于用于定义规则的控件的类型。

字段选择器

您可以指定在使用至少一个字段（非空字段）填充字段选择器时启用当前控件。还可以指定在未填充字段选择器时启用当前控件。

复选框或复选框组

您可以指定在选中"复选框"或"复选框组"时启用当前控件。或者，您可以指定在未选中"复选框"或"复选框组"时启用当前控件。

组合框或单选列表框

您可以指定选中"组合框"或"单选列表框"中的某个特定值时启用当前控件。或者，您可以指定未选中"组合框"或"单选列表框"中的某个特定值时启用当前控件。

多选列表框

您可以指定"多选列表框"中的选定值中包含某个特定值时启用当前控件。或者，您可以指定"多选列表框"中的选定值中不包含某个特定值时启用当前控件。

单选按钮组

您可以指定选中某个特定单选按钮时启用当前控件。或者，您可以指定未选中某个特定单选按钮时启用当前控件。

可以为其指定启用规则的控件具有关联"启用规则"属性。

注：

- 无论是否启用定义规则的控件，启用规则都适用。例如，请考虑指定填充字段选择器时启用单选按钮组的规则。只要填充字段选择器，就会启用单选按钮组，无论是否启用字段选择器都是如此。
- 禁用选项卡控件时，会禁用选项卡上的所有控件，不管这些控件中任何控件是否具有达到需求的启用规则。
- 禁用"复选框组"时，会禁用组中所有控件，无论是否选中控制复选框都是如此。

扩展属性

"扩展属性"对话框指定有关"扩展的定制对话框构建器"内的当前扩展的信息，如扩展的名称和扩展中的文件。

- 在"扩展的定制对话框构建器"中创建的所有定制节点对话框都是扩展的一部分。
- 必须先指定"扩展属性"对话框的"必需"选项卡上的字段，然后才能安装扩展及其包含的定制节点对话框。

要指定扩展的属性，请从"扩展的定制对话框构建器"中的菜单选择：

扩展 > 属性

扩展的必需属性

名称 与扩展相关联的唯一名称。它最多可包含三个字，并且不区分大小写。字符限制为 7 位 ASCII。为尽量降低名称冲突的可能性，您可能想要使用多字名称，其中第一个字是组织的标识，例如 URL。

摘要 扩展的简短描述，旨在以单行形式显示。

版本 x.x.x 格式的版本标识，其中，标识的每个组件必须是一个整数，例如，1.0.0。如果未提供，那么将隐含零。例如，版本标识 3.1 暗示 3.1.0。版本标识与 IBM SPSS Modeler 版本无关。

最低 SPSS Modeler 版本

运行扩展所需的 SPSS Modeler 的最低版本。

文件 "文件"列表显示扩展中当前包含的文件。单击添加以向扩展中添加文件。您也可以从扩展中移除文件，并且可以将文件抽取到指定的文件夹。

- 定制节点对话框具有的文件类型为 .cfe。
- 从"可选"选项卡上的"本地化设置"添加扩展的组件的转换文件。
- 可以向扩展中添加自述文件。请将文件名指定为 ReadMe.txt。用户可以从显示扩展的详细信息对话框访问自述文件。可以包含自述文件的本地化版本（指定为 ReadMe_<language identifier>.txt，例如 ReadMe_fr.txt 表示法语版本）。

扩展的可选属性

常规属性

描述 比为摘要字段提供的更详细的扩展描述。例如，您可能会列出扩展随附的主要功能部件。

日期 扩展的当前版本的可选日期。未提供格式化。

作者 扩展的作者。您可能希望包含电子邮件地址。

链接 要与扩展关联的一组 URL；例如，作者的主页。此字段的格式为任意，因此请务必以空格、逗号或某个其他合理定界符分隔多个 URL。

关键字

要与扩展关联的一组关键字。

平台 有关任何适用于在特定操作系统平台上使用扩展的限制的信息。

依赖性

最高 SPSS Modeler 版本

可以在其之上运行扩展的 IBM SPSS Modeler 的最高版本。

需要 Integration Plug-in for R

指定是否需要 Integration Plug-in for R。

如果扩展需要 CRAN 软件包库中的任何 R 软件包，那么在“所需 R 软件包”控件中输入这些软件包的名称。名称区分大小写。要添加第一个软件包，请单击“所需 R 软件包”控件中的任意位置以突出显示输入字段。按 **Enter** 键（光标位于给定行中）将创建新行。通过选择一行并按删除可删除该行。

本地化

定制节点

可以为扩展内的定制节点对话框添加属性文件（指定节点对话框中显示的所有字符串）的已转换版本。要为特定节点对话框添加转换，请单击**添加转换**，然后选择包含已转换版本的文件夹。特定节点对话框的所有已转换文件必须在同一文件夹中。有关创建转换文件的指示信息，请参阅主题第 48 页的『创建定制节点对话框的本地化版本』。

转换目录文件夹

您可以为扩展提供在最终用户查看扩展主数据中心内的扩展详细信息时显示的**摘要**和**描述**字段的本地化版本。扩展的所有本地化文件集都必须在名为 lang 的文件夹中。请浏览至包含本地化文件的 lang 文件夹，然后选择该文件夹。

要提供**摘要**和**描述**字段的本地化版本，请为将其提供转换的每种语言创建名为 `<extension name>_<language-identifier>.properties` 的文件。在运行时，如果找不到当前用户界面语言的 `.properties` 文件，那么会使用“必需”和“可选”选项卡上指定的**摘要**与**描述**字段的值。

- `<extension name>` 是扩展的**名称**字段的值，其中任何空格都替换为下划线字符。
- `<language-identifier>` 是特定语言的标识。以下显示了 IBM SPSS Modeler 支持的语言的标识。

例如，名为 MYORG MYSTAT 的扩展的法语转换存储在文件 MYORG_MYSTAT_fr.properties 中。

`.properties` 文件必须包含以下两行，其中指定两个字段的本地化文本：

```
Summary=<localized text for Summary field>  
Description=<localized text for Description field>
```

- 关键字 Summary 和 Description 必须为英语，并且本地化文本必须与关键字在同一行上且没有换行符。
- 文件必须采用 ISO 8859-1 编码形式。必须使用 Unicode 转义 (“\u”) 编写此编码中无法直接表示的字符。

包含本地化文件的 lang 文件夹必须具有名为 `<language-identifier>` 的子文件夹，其中包含特定语言的本地化 `.properties` 文件。例如，法语 `.properties` 文件必须在 lang/fr 文件夹中。

语言标识

de。德语

en。英语

es。西班牙语

fr。法语

it。意大利语

ja。日语

ko。韩国语

pl。波兰语

pt_BR。巴西葡萄牙语

ru。俄语

zh_CN。简体中文

zh_TW。繁体中文

管理定制节点对话框

通过"扩展的定制对话框构建器"，您可以管理由您或其他用户创建的扩展内的定制节点对话框。在可以使用前，必须将定制节点对话框安装到 SPSS Modeler 客户机的所有实例或者需要的 SPSS Modeler Batch。请注意，无需将任何项安装到 SPSS Modeler Server 即可以服务器方式使用定制对话框节点。

注：您只能对 IBM SPSS Modeler 中创建的定制节点对话框进行修改。

打开包含定制节点对话框的扩展

您可以打开包含一个或多个定制节点对话框的规范的扩展束文件 (.mpe)，也可以打开已安装的扩展。您可以打开扩展中的任何节点对话框并保存或安装扩展。安装扩展将会安装扩展中包含的节点对话框。保存扩展将会保存已对扩展中的任何节点对话框进行的更改。

要打开扩展束文件，请从"扩展的定制对话框构建器"中的菜单选择：

文件 > 打开

要打开已安装的扩展，请从"扩展的定制对话框构建器"中的菜单选择：

文件 > 打开已安装的项

注：如果您是打开已安装的扩展来对其进行修改，那么选择文件 > 安装会重新安装该扩展，从而替换现有版本。对于通过定制对话框构建器创建的节点，使用它的上下文菜单中的编辑并不会在定制对话框构建器中打开该节点对话框。

保存到扩展束文件

保存"扩展的定制对话框构建器"中打开的扩展还会保存扩展中包含的定制节点对话框。扩展会保存到扩展束文件 (.mpe)。

从"扩展的定制对话框构建器"中的菜单选择：

文件 > 保存

安装扩展

安装"扩展的定制对话框构建器"中打开的扩展还会安装扩展中包含的定制节点对话框。安装现有扩展将会替换现有版本，包括替换已安装的扩展中的所有定制节点对话框。

要安装当前打开的扩展，请从"扩展的定制对话框构建器"中的菜单选择：

文件 > 安装

缺省情况下，扩展会安装到操作系统的常规用户可写位置。有关更多信息，请参阅主题第 24 页的『扩展的安装位置』。

注：在打开的流中，将不会替换扩展中包含的节点对话框的现有版本。您打开包含已重新安装的"定制对话框构建器"节点的流时，将接收到一条警告消息。

卸载扩展

从"扩展的定制对话框构建器"中的菜单选择：

文件 > 卸载

卸载扩展将会卸载扩展中包含的所有定制节点对话框。您也可以从"扩展主数据中心"卸载扩展。

导入定制对话框软件包文件

您可以将定制对话框软件包 (.cfd) 文件导入到"扩展的定制对话框构建器"中。.cfd 文件会转换为将添加到新扩展中的 .cfe 文件。

从"扩展的定制对话框构建器"中的菜单选择：

文件 > 导入

您也可以从"扩展属性"对话框将 .cfe 文件添加到扩展，该对话框通过"扩展的定制对话框构建器"内的扩展 > 属性进行访问。

向扩展中添加定制节点对话框

您可以向扩展中添加新的定制节点对话框。

从"扩展的定制对话框构建器"中的菜单选择：

扩展 > 新建对话框

在扩展中的多个定制节点对话框之间切换

如果当前扩展包含多个定制节点对话框，那么可以在其之间切换。

从"扩展的定制对话框构建器"中的菜单选择：

扩展 > 编辑对话框，然后选择要处理的定制节点对话框。

创建新扩展

在"扩展的定制对话框构建器"中创建新扩展时，会向该扩展中添加新的空定制节点对话框。

要创建新扩展，请从"扩展的定制对话框构建器"中的菜单选择：

文件 > 新建

SPSS Modeler Batch 或 IBM SPSS Collaboration and Deployment Services 中的扩展

要在 SPSS Modeler Batch 或 IBM SPSS Collaboration and Deployment Services 安装中使用扩展，请确保在目标环境中定义环境变量 `IBM_SPSS_MODELER_EXTENSION_PATH`，并且该环境变量指向包含扩展的位置。如果已在定义 `IBM_SPSS_MODELER_EXTENSION_PATH` 环境变量前将包含定制节点的流保存到 IBM SPSS Collaboration and Deployment Services Repository，那么必须先将流复原到存储库，然后才能成功运行。

注：请确保 SPSS Modeler 的 SPSS Modeler Batch 或 IBM SPSS Collaboration and Deployment Services 适配器版本与其中已创建扩展的 SPSS Modeler Client 的版本匹配。

创建定制节点对话框的本地化版本

您可以针对 IBM SPSS Modeler 所支持的任何语言创建定制节点对话框的本地化版本。可以将定制节点对话框中显示的任何字符串本地化，并可以将可选的帮助文件本地化。

要本地化对话框字符串

您必须为计划部署的每种语言创建与定制节点对话框关联的属性文件的副本。此属性文件包含所有与节点对话框相关联的可本地化字符串。

通过在"扩展属性"对话框中（在"扩展的定制对话框构建器"内）选择定制节点对话框文件（.cfe）并单击抽取来从扩展中抽取该文件。然后，抽取 .cfe 文件的内容。.cfe 文件即是 .zip 文件。.cfe 文件的所抽取内容包含每种受支持语言的属性文件，其中特定语言的文件的名称由 <Dialog Name>_<language identifier>.properties 给定（请参阅下面的表中的语言标识）。

1. 使用支持 UTF-8 的文本编辑器（如 Windows 上的记事本）打开计划转换的每个属性文件。请修改与任何需要本地化的属性相关联的值，但不要修改属性的名称。与特定控件相关联的属性以该控件的标识作为前缀。例如，标识为 *options_button* 的控件的"工具提示"属性为 *options_button_tooltip_LABEL*。标题属性简单地命名为 <identifier>_LABEL，例如 *options_button_LABEL*。
2. 通过"扩展属性"对话框的"可选"选项卡上的"本地化设置"将属性文件的本地化版本重新添加到定制节点对话框文件（.cfe）。有关更多信息，请参阅主题第 44 页的『扩展的可选属性』。

启动节点对话框后，IBM SPSS Modeler 会按照"选项"对话框中"常规"选项卡上的"语言"下拉列表所指定，搜索其语言标识与当前语言匹配的属性文件。如果找不到此类属性文件，那么会使用缺省文件 <Dialog Name>.properties。

要查找帮助文件

1. 创建与定制节点对话框关联的帮助文件的副本，然后针对所需语言将文本本地化。
2. 使用下表中的语言标识将副本重命名为 <Help File>_<language identifier>.htm。例如，如果帮助文件为 myhelp.htm，并且您希望创建此文件的德语版，那么本地化的帮助文件应该名为 myhelp_de.htm。

将帮助文件的所有本地化版本与非本地化版本存储在同一目录中。从"对话框属性"的"帮助文件"属性中添加非本地化帮助文件时，会将本地化版本自动添加到节点对话框。

如果存在也需要本地化的补充性文件（如图像文件），那么必须将主帮助文件中的相应路径手动修改为指向本地化版本。必须将补充性文件（包括本地化版本）手动添加到定制节点对话框（.cfe）文件。请参阅标题为"本地化对话框字符串"的先前部分，以获取有关访问和手动修改定制节点对话框文件的信息。

启动节点对话框后，IBM SPSS Modeler 会按照"选项"对话框中"常规"选项卡上的"语言"下拉列表中所指定，搜索其语言标识与当前语言匹配的帮助文件。如果找不到这样的帮助文件，那么将使用对节点对话框指定的帮助文件（对话框属性的"帮助文件"属性中指定的文件）。

语言标识

de. 德语

en. 英语

es. 西班牙语

fr。法语

it。意大利语

ja。日语

ko。韩国语

pl。波兰语

pt_BR。巴西葡萄牙语

ru。俄语

zh_CN。简体中文

zh_TW。繁体中文

注意：定制节点对话框及相关帮助文件中的文本不限于 IBM SPSS Modeler 所支持的语言。您可以采用任何语言任意编写节点对话框和帮助文本，而不必创建特定于语言的属性文件和帮助文件。随后，所有节点对话框用户都将以该语言查看文本。

使用 Python for Spark 导入和导出数据

通过使用针对扩展的定制对话框构建器，您可以创建定制节点和编写 Python for Spark 脚本，以从数据源所在位置读取数据，将数据写出到 Apache Spark 支持的任何数据格式。

例如，用户希望将他的数据写入数据库。他使用针对扩展的定制对话框构建器和 Python for Spark 来创建定制导出 JDBC 节点，然后运行模型以将数据写入数据库。要从数据库读取数据，还可以创建定制导入 JDBC 节点。例如，还可以使用此方法将数据从 JSON 文件读入 SPSS Modeler。然后，在将其数据读入 SPSS Modeler 后，他可以使用所有可用的 SPSS Modeler 节点来处理业务问题。

注：如果您希望将 JDBC 与 Python for Spark 导入和导出功能配合使用，必须将 JDBC 驱动程序文件复制到 IBM SPSS Modeler 安装目录中的 `as/lib` 目录。

使用 Python for Spark 导入/导出数据

1. 转至扩展 > 定制节点对话框构建器。
2. 在对话框属性下，为脚本类型选择 **Python for Spark**，为节点类型选择**导入或导出**。
3. 根据需要输入其他属性，如对话框名称。
4. 在"脚本"部分中，为导入或导出数据输入或粘贴 Python for Spark 脚本。
5. 单击**安装**以安装 Python for Spark 脚本。新的定制导入节点将添加到"源"选用板，新的定制导出节点将添加到"导出"选用板。

使用 R 导入和导出数据

通过使用针对扩展的定制对话框构建器，您可以创建定制节点和编写 R 脚本，以从数据源所在位置读取数据，将数据写出到 R 支持的任何数据格式。

例如，用户希望将她的数据写入数据库。她使用针对扩展的定制对话框构建器和 R 脚本编制来创建定制导出 JDBC 节点，然后运行模型以将数据写入数据库。要从数据库读取数据，还可以创建定制导入 JDBC 节点。例如，还可以使用此方法将数据从 JSON 文件读入 SPSS Modeler。然后，在将其数据读入 SPSS Modeler 后，用户可以使用所有可用的 SPSS Modeler 节点来处理业务问题。

使用 R 导入/导出数据

1. 转至扩展 > 定制节点对话框构建器。
2. 在对话框属性下，为脚本类型选择 **R**，为节点类型选择导入或导出。
3. 根据需要输入其他属性，如对话框名称。
4. 在"脚本"部分中，为导入或导出数据输入或粘贴 R 脚本。
5. 单击**安装**以安装 R 脚本。新的定制导入节点将添加到"源"选用板，新的定制导出节点将添加到"导出"选用板。

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档中讨论的产品、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

[D]

- 定制对话框构建器
 - 帮助文件 26
 - 表控件 37
 - 表控件列 38
 - 布局规则 27
 - 单选按钮组 39
 - 对话框属性 26
 - 对字段列表进行过滤 30
 - 复选框 30
 - 复选框组 40
 - 将对话框和帮助文件本地化 48
 - 脚本模板 27
 - 列表框 32
 - 列表框列表项 31
 - 启用规则 43
 - 受保护的文本 35
 - 文本控件 33
 - 文件类型过滤器 42
 - 文件浏览器 41
 - 选项卡 42
 - 颜色选择器 36
 - 预览 28
 - 子对话框按钮 42
 - 子对话框属性 43
 - 字段选择器 29
 - 字段源 30
 - 组合框 31
 - 组合框列表项 31
 - Python for Spark 数据导出 49
 - Python for Spark 数据导入 49
 - R 数据导出 49
 - R 数据导入 49
 - "静态文本"控件 36
 - "日期"控件 35
 - "数字"控件 34
 - "项目组"控件 39

[F]

- 发布流
 - IBM SPSS Modeler Solution Publisher 12

[K]

- 扩展 21
 - 查看已安装的扩展 22
 - 查找和安装新扩展 22
 - 除去扩展 22

- 扩展 (续)
 - 将更新安装到扩展 22
 - 扩展详细信息 23
- 扩展模型块 16
 - 关于 16
 - "控制台输出"选项卡 17
 - "模型选项"选项卡 16
 - "图形输出"选项卡 17
 - "文本输出"选项卡 17
 - "语法"选项卡 16
- 扩展输出浏览器 14, 15
- 扩展束
 - 安装扩展束 23

[Y]

- 用于扩展的定制对话框构建器 24
 - 安装包含节点对话框的扩展 46
 - 保存包含节点对话框的扩展 46
 - 打开包含节点对话框的扩展 46
 - 扩展束文件 46
 - 修改已安装的扩展中的节点对话框 46
- 允许的 R 语法 1

A

- Apache Spark 1

I

- IBM SPSS Modeler 扩展 1
- IBM SPSS Modeler Solution Publisher 12

P

- Python for Spark 1
 - API 1, 4, 6, 7, 8, 9

R

- R
 - 允许的语法 1

[特别字符]

- .par 文件 12
- .pim 文件 12

- "控制台输出"选项卡
 - 扩展模型块 17
- "扩展变换"节点 18
 - "控制台输出"选项卡 18
- "扩展导出"节点 11
 - "控制台输出"选项卡 12
- "扩展导入"节点 19
 - "控制台输出"选项卡 19
- "扩展模型"节点 15
 - "控制台输出"选项卡 15
 - "模型选项"选项卡 15
 - "文本输出"选项卡 16
- "扩展输出"节点 13
 - "控制台输出"选项卡 14
 - "输出"选项卡 14
 - "语法"选项卡 13
- "扩展"节点 11
- "图形输出"选项卡 15
 - 扩展模型块 17
- "文本输出"选项卡 15
 - 扩展模型块 17
- "语法"选项卡
 - 扩展模型块 16
- "IBM SPSS Modeler 扩展"节点 11



Printed in China