



from
Excel[®]
to
KNIME[®]

Author: Kathrin Melcher

版权所有 © 2022 KNIME出版社

版权所有。本出版物受版权保护，未经出版商事先获得许可，不得进行任何禁止的复制、存储于检索系统中或以任何形式或任何方式传输，包括电子、机械、复印、录音或类似方式。

本书已更新至KNIME 4.6版本。

有关许可和销售信息，请写信至：

KNIME出版社
Talacker 50
8001 Zurich
瑞士

knimepress@knime.com

目录

常规用法	5
电子表格 工作流程和节点.....	6
文件夹 工作区.....	7
KNIME工作台.....	8
构建KNIME工作流程.....	9
显示数据表.....	10
输入/输出	11
打开Excel文件 Excel读取器节点.....	12
打开CSV或txt文件 CSV读取器节点.....	13
将多个相同类型的文件内容导入到单个表中.....	14
将多个工作表的内容导入到单个表中.....	15
保存Excel文件 Excel写入节点.....	16
向Excel文件添加工作表 Excel写入节点.....	17
更新现有Excel工作表中的单元格 Excel单元格更新节点.....	18
Excel中的数据类型 KNIME中的数据类型.....	19
连接到数据库 数据库连接器节点.....	20
使用KNIME Analytics Platform读取数据的提示.....	21
追加/连接数据	22
追加数据 连接节点.....	23
VLOOKUP 过滤和连接节点.....	24
过滤和转换	25
过滤/删除具有特定值的行 行过滤器.....	26

过滤/删除具有不同值的行	基于规则的行过滤器.....	27
删除重复项	重复行过滤器.....	28
按多个关键列对行进行排序	排序节点.....	29
删除列	列过滤器节点.....	30
重新排序和重命名列	列重新排序和列重命名节点.....	31
更改数据类型	字符串转数字和数字转字符串.....	32
使用KNIME Analytics Platform进行数据操作的提示	33
数据聚合	34
数据透视表	数据透视节点.....	35
没有列的数据透视表	分组节点	36
解开	解开节点	37
数学函数和文本函数	38
数学函数	数学公式节点.....	39
数学函数	数学公式节点.....	40
连接和查找替换	字符串操作节点.....	41
格式化Excel表格	42
使用标签值创建XLS控制表.....	45
根据标签值添加格式化操作.....	49

常规用法

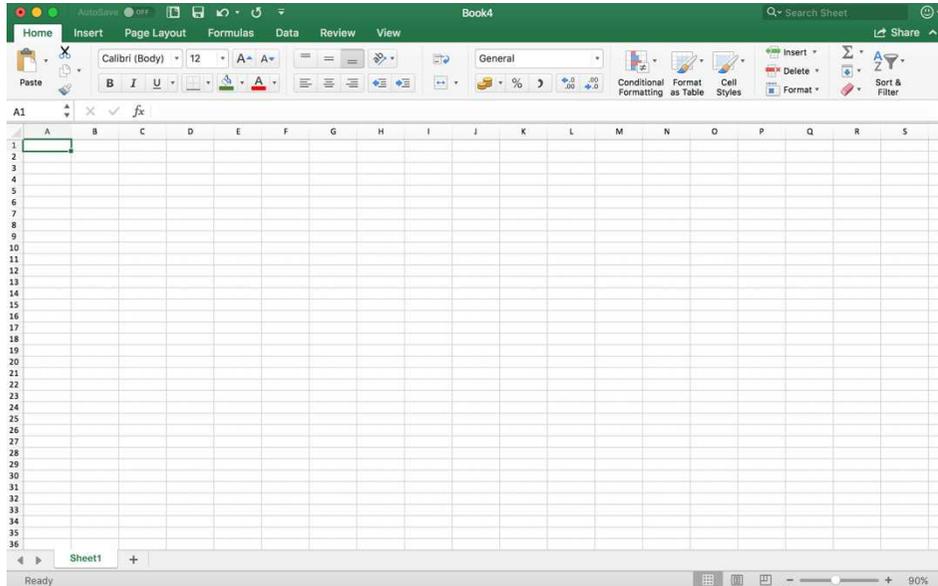
电子表格

工作流程和节点

Excel

Microsoft Excel是一款电子表格程序，具有计算、绘图工具、数据透视表和宏编程语言（Visual Basic for Applications, 简称VBA）等功能。

通过使用单元格数学、宏和VBA，您可以编辑一个工作表。这可以是简单的单元格数学，比如汇总单元格A1和B1的值（=SUM(A1,B1)），但也可以是非常复杂的嵌入式逻辑。

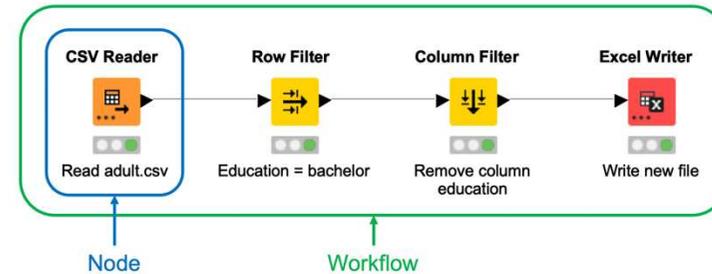


KNIME分析平台

KNIME分析平台实现了可视化编程。这意味着每个数据分析步骤都通过一个图标块（称为节点）在图形编辑器中表示。每个节点可以执行一个特定的任务。例如，Excel读取器节点可以读取Excel文件的一个工作表，而行过滤器节点允许根据过滤条件过滤行。

一系列连接的节点称为工作流，这是一个具有许多功能和/或VBA的Excel表格的对应概念。

数据通过数据表进行组织，每个数据单元格由列标题和行ID标识。要可视化数据表的内容，请参见第11页。



注意。节点有四种可能的状态，通过节点本身下方的一个小红绿灯来显示：

- 未配置 -> **红灯**
- 已配置 -> **黄灯**
- 成功执行 -> **绿灯**
- 执行出错 -> **带有交叉的红灯**

有关KNIME Analytics Platform的更多详细信息，请查看：

- S.Hayasaka, R. Silipo, [“KNIME初学者的幸运”](#), KNIME Press, 2021
- R.Silipo, V.Palacios, [“KNIME高级幸运”](#), KNIME Press, 2021

文件夹

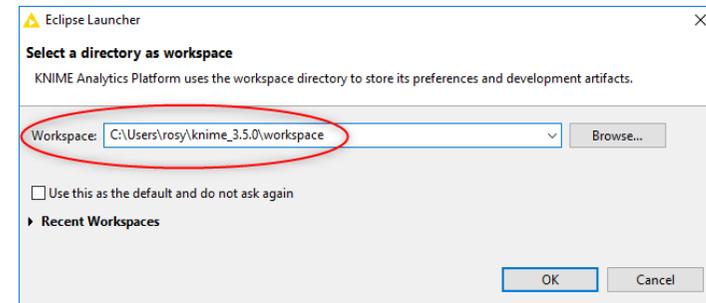
Excel

Excel文件通常保存在不同的文件夹中。单个Excel文件可以包含多个工作表。

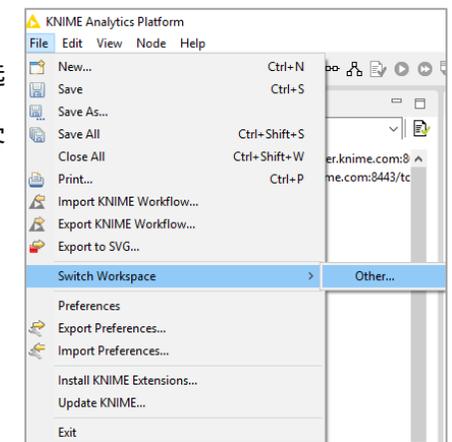
工作区

KNIME分析平台

工作区定义了文件夹，其中保存了所有的工作流程、数据和中间数据。一个工作流程对应一个带有所有公式、可视化和VBA的Excel表格。在工作区中保存的所有项目和数据集都可以在KNIME资源管理器中找到，位于KNIME工作台的左上角。工作区的路径在开始*KNIME Analytics Platform*之后选择。



在启动KNIME Analytics Platform之后，仍然可以更改工作区，方法是在顶部菜单中选择“文件”，然后选择“切换工作区”。您可以拥有多个工作区，例如用于不同的项目或客户。



KNIME工作台

下载并安装KNIME Analytics Platform后，您可以从桌面或安装文件夹启动它。您可以在下面看到的KNIME工作台包括以下面板：

“**KNIME资源管理器**”显示当前可用的工作流程和选定工作区的KNIME服务器以及My-KNIME-Hub挂载点。

“ **workflow 教练**”根据KNIME用户统计数据 and 当前在“ **workflow 编辑器**”中选择的节点推荐下一个节点。

“**节点仓库**”包含所有当前安装的节点。在此面板顶部有一个“搜索”框，可用于搜索节点。

“ **workflow 编辑器**”位于中心，允许创建和编辑 workflow。

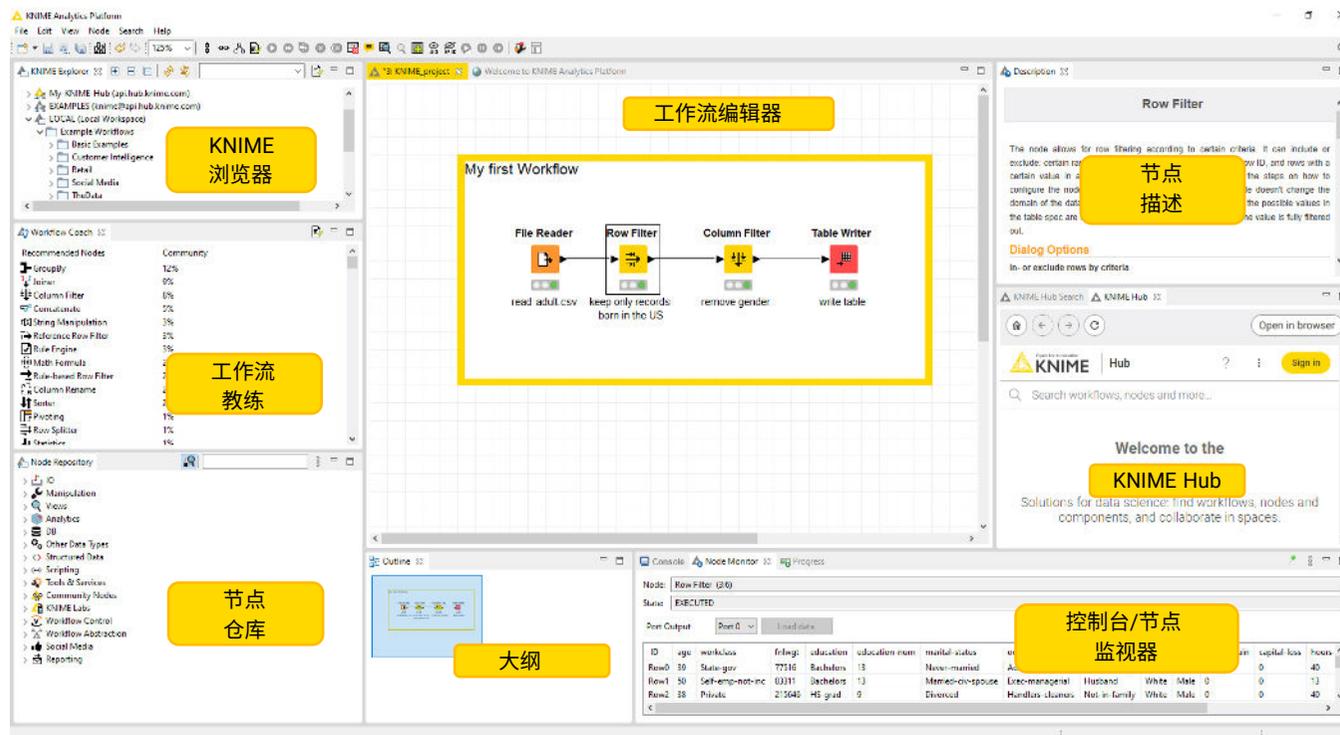
“**节点描述**”显示描述节点任务和配置设置的文本，对于“ **workflow 编辑器**”或“**节点仓库**”中的选定节点。

“**节点监视器**”显示在“ **workflow 编辑器**”中选择的节点的输出表格预览。

“**KNIME Hub**”允许使用KNIME Hub搜索节点、 workflow、组件和扩展。

“**大纲**”提供 workflow 的概述

“**控制台**”显示执行消息，例如错误和警告消息。



构建KNIME工作流程

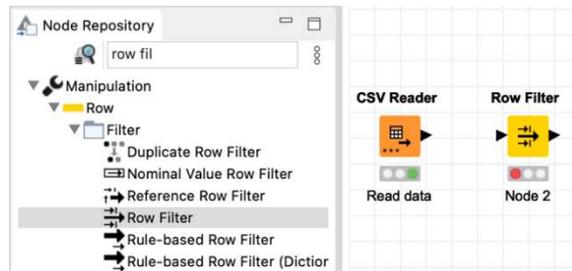
通过从“节点库”或“ workflow助手”面板将节点拖放到“ workflow编辑器”中，可以创建KNIME workflow。使用节点库顶部的搜索框或按不同类别浏览节点，以找到下一步的正确节点。

节点通过其输入和输出端口相互连接。只需单击第一个节点的输出端口，然后释放到第二个节点的输入端口。

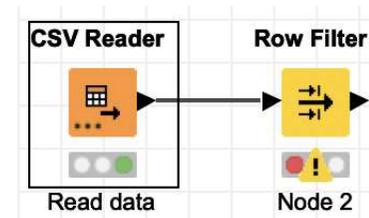
刚创建的节点显示为红色状态：尚未配置。要配置节点，请右键单击节点并选择“配置”选项，或者双击节点。节点“配置”窗口将打开。配置节点并关闭配置窗口。如果配置成功，节点状态将变为黄色交通灯。

节点现在已配置，但尚未执行。要执行节点，请右键单击节点并选择“执行”选项。如果执行成功，节点状态将变为绿色灯。

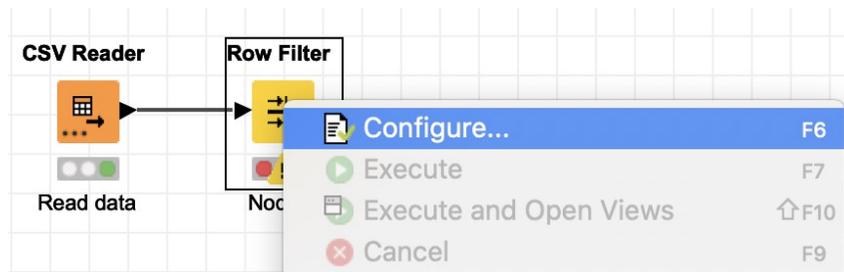
步骤 1：通过拖放搜索并创建一个节点



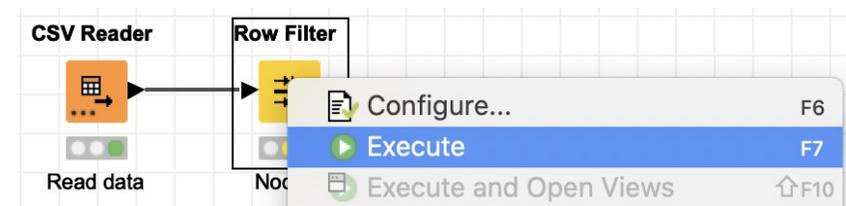
步骤 2：连接节点



步骤 3：配置节点



步骤 4：执行节点



注释 1：要创建一个新的空 workflow，请右键单击 KNIME Explorer 面板，在新窗口中选择“创建新的 KNIME workflow”，并定义新 workflow 的名称和目标。

注释 2：单击节点库中搜索框旁边的放大镜，将搜索框的模式更改为模糊搜索。这样在开始阶段更容易找到正确的节点。

注释 3：《入门指南》逐步指导您构建第一个示例 workflow。

显示数据表

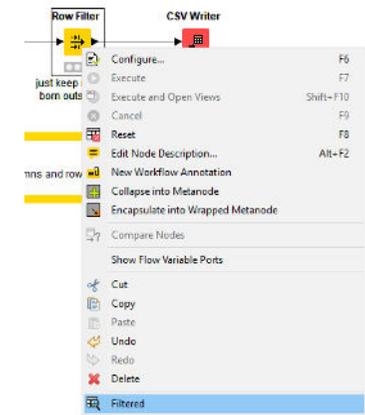
Excel

在 Excel 中所见即所得。这意味着您看到的数据表是最终的数据表。

KNIME分析平台

节点执行后产生的输出数据表始终可用。要查看它们：

- 右键单击 workflows 中的节点
- 选择上下文菜单中的最后一个选项



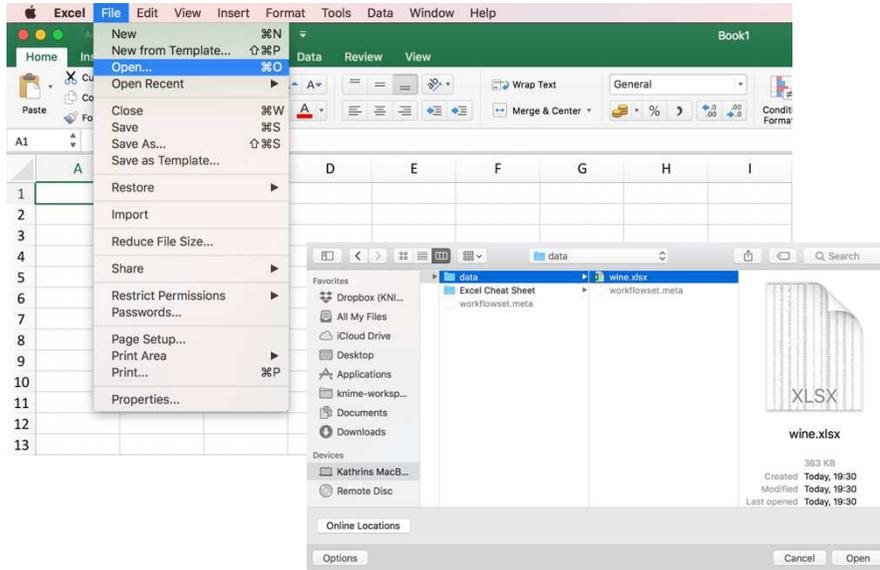
注意：某些节点，如绘图和建模节点，还具有更复杂的“查看”功能。导致此“查看”的选项通常显示在上下文菜单的中间。

输入/输出

打开Excel文件

Excel

要打开Excel文件，您可以双击文件或打开Excel，转到顶部菜单中的“文件”，选择“打开...”，然后浏览要打开的文件。

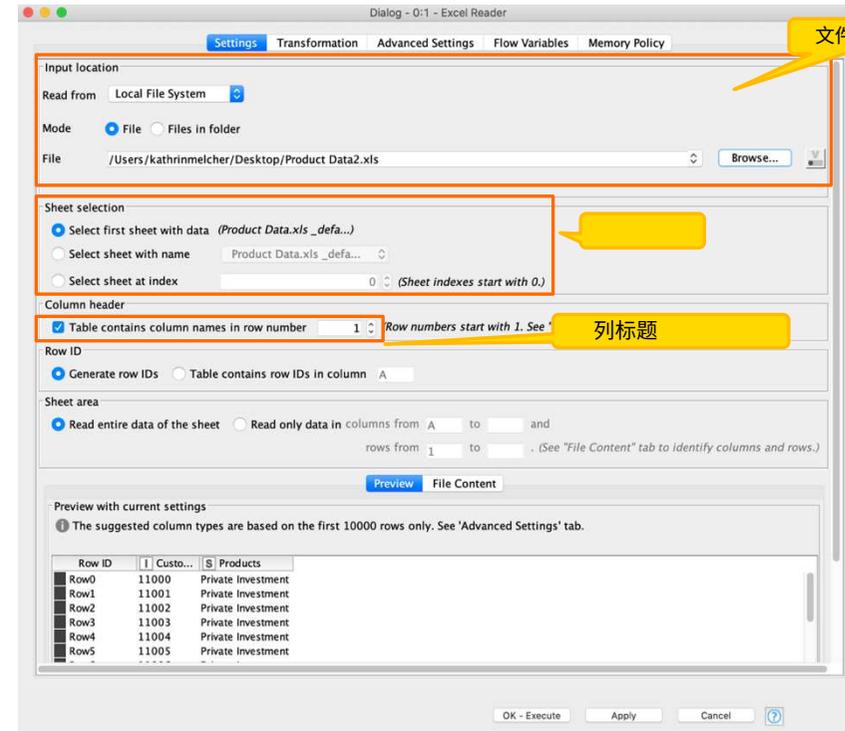


Excel读取器节点

KNIME分析平台

[Excel Reader](#)节点读取Excel文件的单个工作表。与Excel类似，您首先必须定义文件路径。接下来，您可以选择工作表。附加的设置选项使您能够定义表是否具有列标题和/或行ID，并指定要读取的工作表的哪个部分。

Excel Reader

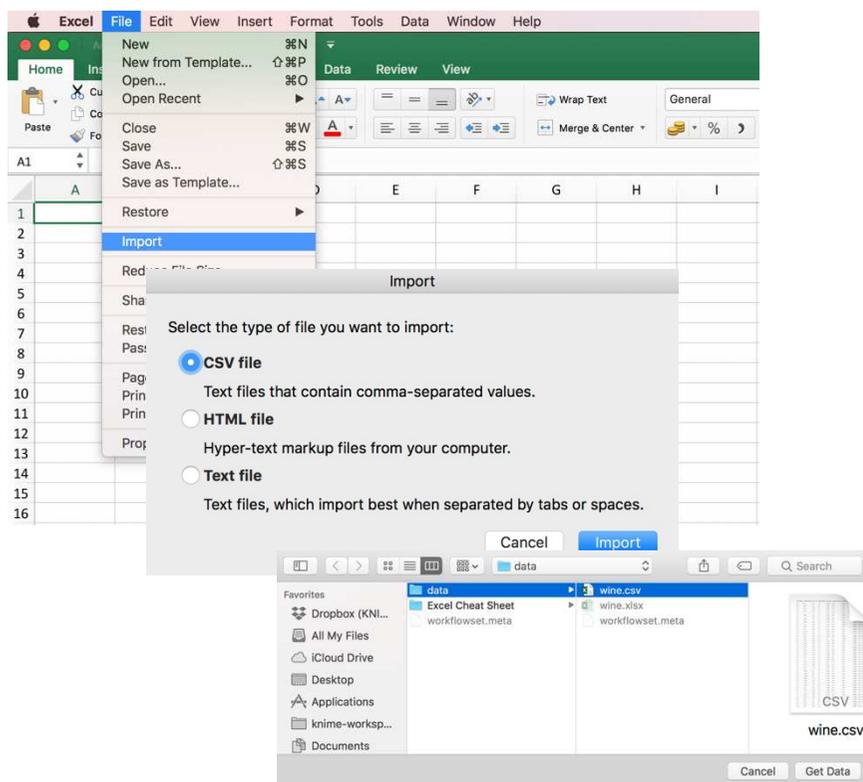


注意：第15页显示了如何读取和连接具有相同列标题的多个Excel文件。

打开CSV或txt文件

Excel

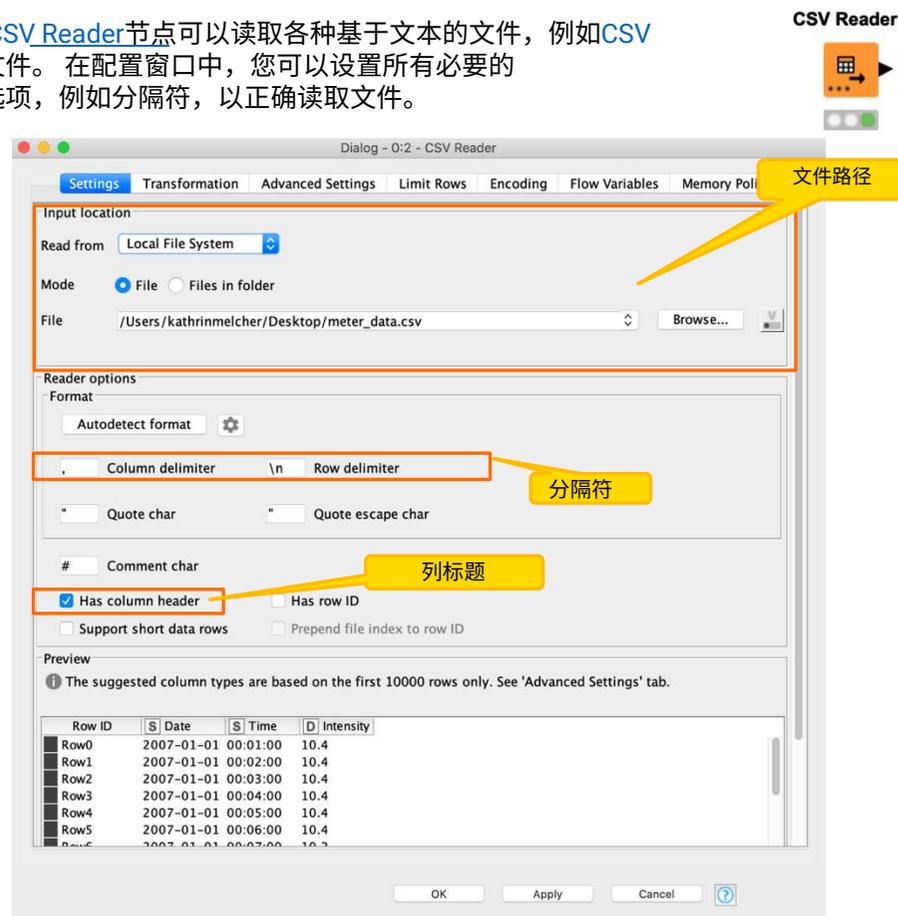
要打开CSV文件，请点击顶部菜单中的“文件”，然后选择“导入”。选择CSV文件然后选择要打开的文件。
填写所有必要的选项，例如分隔符，以正确读取文件。



CSV读取器节点

KNIME分析平台

CSV Reader节点可以读取各种基于文本的文件，例如CSV文件。在配置窗口中，您可以设置所有必要的选项，例如分隔符，以正确读取文件。



注意1：如果节点没有创建预览，请点击“自动检测格式”按钮。
注意2：查看其他选项卡以限制行数或更改编码方式。

将相同类型的多个文件的内容导入到单个表中

Excel

如果您有一个包含多个相同类型文件（例如CSV）的文件夹，您可以同时在不同的Excel实例中打开它们。按照上一页的说明，选择要打开的所有文件。

要将数据表一起移动到一个单独的工作表中，您必须手动进行复制和粘贴。

注意:在复制和粘贴之前，请确保所有文件具有相同的列顺序。

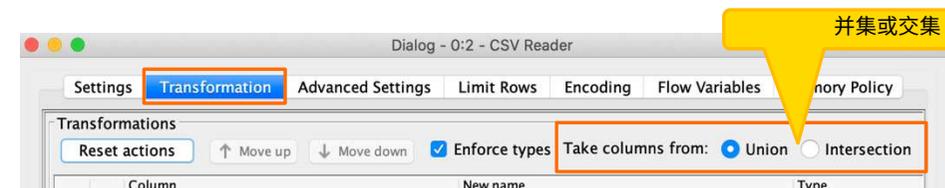
KNIME分析平台

KNIME Analytics Platform中的许多读取节点支持从文件夹中读取多个文件并将它们合并到一个表中。例如，CSV Reader可以读取多个基于文本的文件。

要读取多个文件，请选择“模式”“文件夹中的文件”并使用“浏览”按钮指定文件夹。一个小提示告诉您当前选择了多少个文件。如果您不想读取文件夹中的所有文件，请单击按钮“筛选选项”以根据文件扩展名和/或名称指定应包含哪些文件。最后，您还可以通过激活“包括子文件夹”复选框来包括子文件夹中的文件。



注意:在转换选项卡中，您可以定义是否要使用不同表格的列的并集或交集。



将内容从多个工作表导入到单个表中

Excel

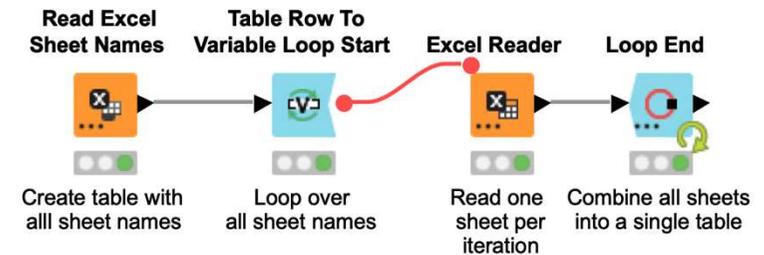
要将数据从多个Excel工作表移动到一个Excel工作表中，您需要手动进行，使用复制和粘贴。

注意:在复制和粘贴之前，请确保所有文件具有相同的列顺序。

KNIME分析平台

通过简单的循环，您可以自动读取Excel文件的所有工作表。

读取[Excel工作表名称节点](#)创建一个包含所有工作表名称的列表。循环（蓝色节点之间的部分）在每次迭代中读取[Excel文件的一个工作表](#)。因此，每次迭代时，[Table Row to Variable Loop Start](#)节点都会创建一个流变量，并将工作表名称作为其值。该流变量用于控制[Excel Reader](#)节点的工作表选择。[Loop End](#)节点将来自不同表格的内容连接起来。



注意1: 免费的[KNIME](#)自学课程[L2-DW KNIME Analytics Platform for Data Wranglers](#)课程的第3课介绍了流变量。

注意2: 免费的[KNIME](#)自学课程[L2-DW KNIME Analytics Platform for Data Wranglers](#)课程的第4课介绍了KNIME中的循环。

保存Excel文件

Excel写入节点

Excel

要保存Excel文件的工作表，您有不同的选项：

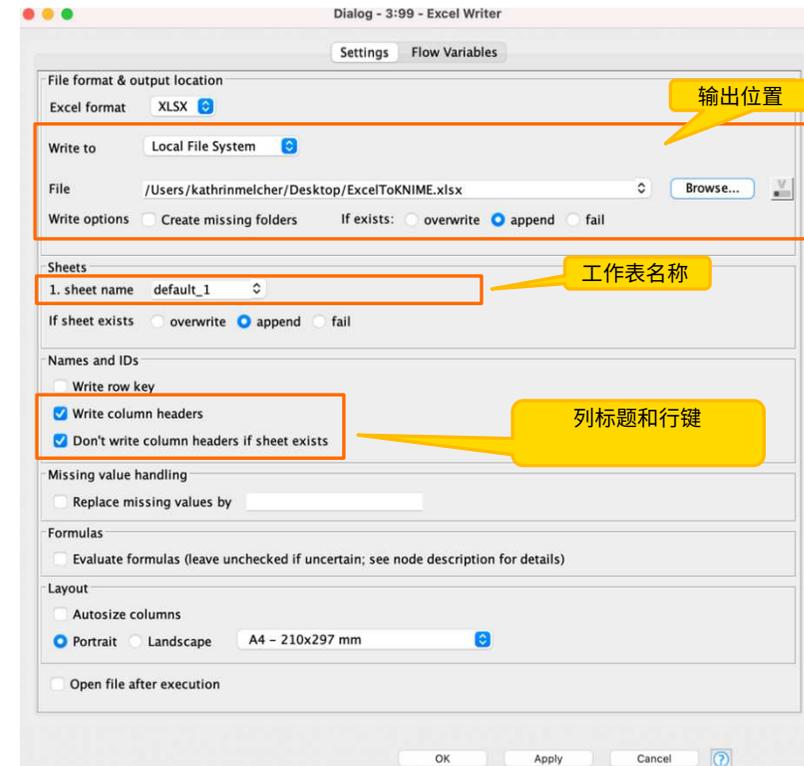
- 点击文件->另存为...并定义输出位置
- 或按Shift+Ctrl+s并定义输出位置

KNIME分析平台

Excel写入节点将输入数据表写入或追加到Excel文件的工作表中，可以是xls或xlsx格式。

在配置窗口中，您可以设置输出位置和工作表名称。其他设置选项使您可以覆盖现有文件，并定义是否要将列标题/行ID写入Excel工作表的每一行/列。

Excel Writer



注意1：要将多个表写入不同的工作表中，您可以添加动态输入端口并为每个输入表定义一个工作表名称。

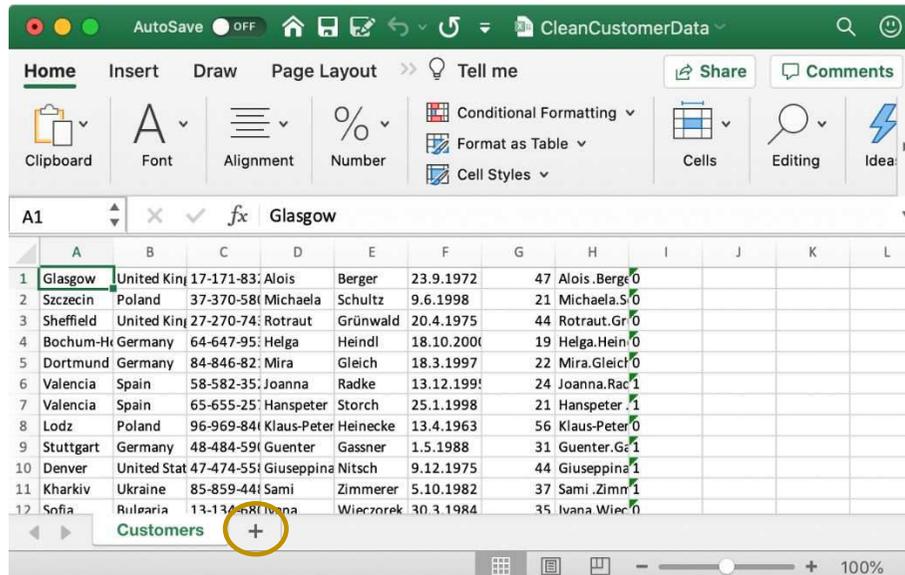
注意2：选择“追加”作为“写入选项”和“如果工作表存在”的设置，以将输入数据追加到现有工作表的最后一行之后。

向Excel文件添加工作表

Excel

要向现有的Excel文件添加新工作表，您必须点击表格下方的加号，位于已有工作表旁边。

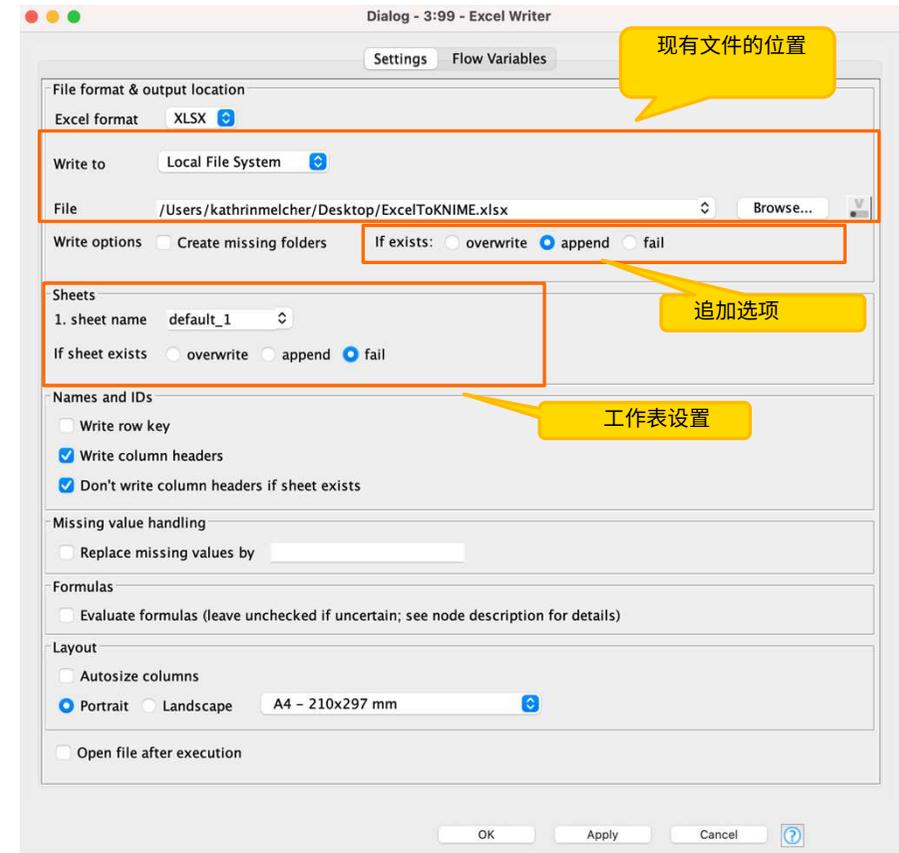
要从另一个Excel文件中添加表格，您可以手动复制并粘贴表格。



ExcelWriter节点

KNIME分析平台

Excel Writer节点还可以向现有的Excel文件添加工作表。在配置窗口中，您可以设置现有Excel文件的位置，并选择“追加”作为“如果存在”设置选项。在“工作表”部分，您可以定义新工作表的名称，并指定节点在已存在具有指定工作表名称的情况下是覆盖还是失败。



更新现有Excel工作表中的单元格

Excel单元格更新节点

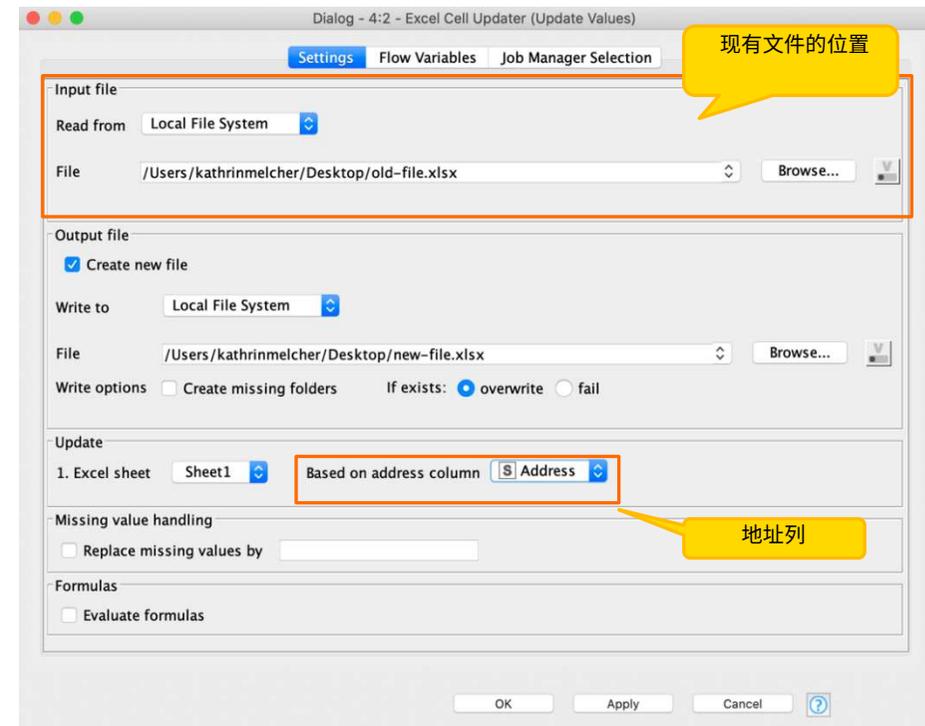
Excel

要编辑现有Excel工作表的内容，您可以打开它并手动编辑不同的单元格。

KNIME分析平台

Excel Cell Updater节点根据输入的数据表在现有Excel工作表中更新单元格。

输入表格需要一列包含应该更新的单元格地址，例如B2或2:2。此外，表格需要每种数据类型的一列，其中包含新的单元格内容。每行只允许有一个值。



注意1：勾选“评估公式”复选框以使用新单元格内容评估所有公式。

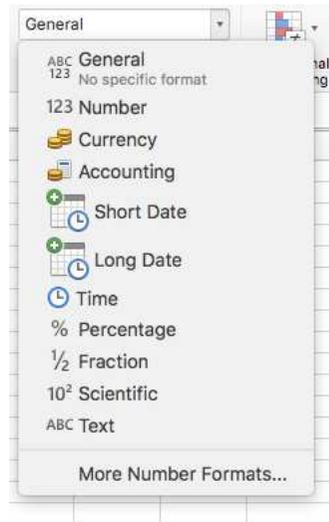
注意2：勾选“创建新文件”复选框以保留原始文件。

注意3：[KNIME Hub](#)上的示例工作流。

Excel中的数据类型

Excel

右侧的截图显示了Excel中可用的不同数据类型。



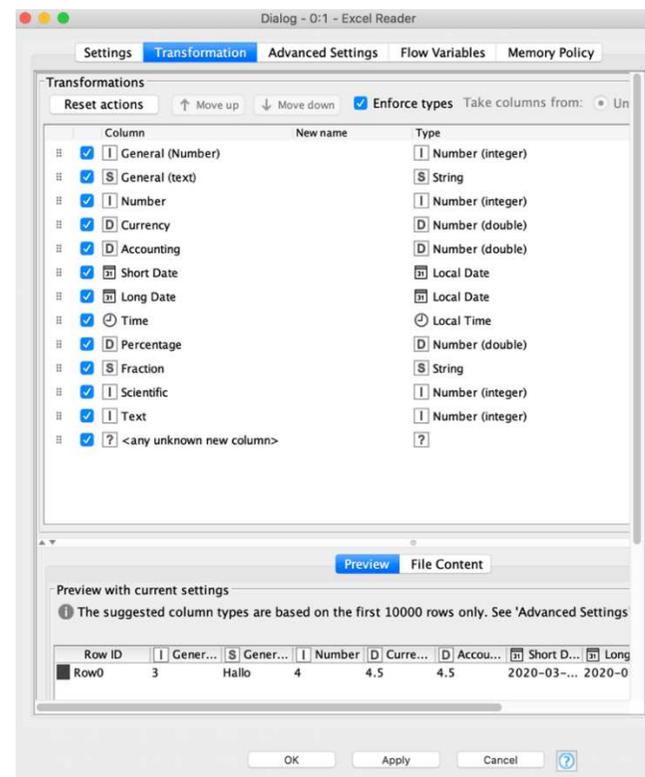
它们可以映射到KNIME Analytics Platform中的以下数据类型

Excel	KNIME分析平台
常规	数字或字符串
数字	数字
货币 (例如 50,25 €)	数字 (整数或浮点数) (例如 50,25)
会计 (例如 50,25 €)	数字 (整数或浮点数) (例如 50,25)
短日期	本地日期
长日期	本地日期
时间	本地时间
百分比 (例如 50%)	数字 (浮点数) (例如 0,5)
分数	数字 (浮点数)
科学计数法 (例如 5,00E+02)	整数 (例如 500)

KNIME中的数据类型

KNIME分析平台

KNIME Analytics Platform支持从Excel读取所有数据类型，甚至更多。在读取文件时，KNIME Analytics Platform会尝试自动检测正确的数据类型。如果您想在读取过程中更改此设置，可以使用Excel Reader节点的“转换”选项卡。



注意1： 转换选项卡也适用于其他读取节点。

注意2： 在转换选项卡中，您还可以重命名、删除和更改列的顺序。

注意3： 要在工作流程中稍后更改数据类型，可以使用以下节点之一：字符串转数字、数字转字符串或表操作器。

连接到数据库

数据库连接器节点

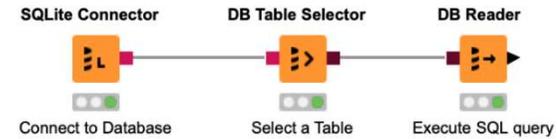
Excel

您可以在Excel中连接到数据库。有关如何执行此操作的说明，请通过以下链接获取：

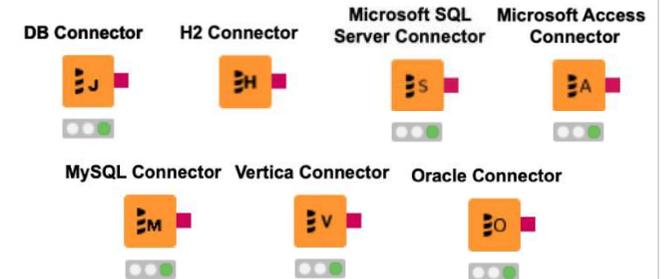
<http://www.erpsoftwareblog.com/2017/01/microsoft-excel-connections-sql-databases/>

KNIME分析平台

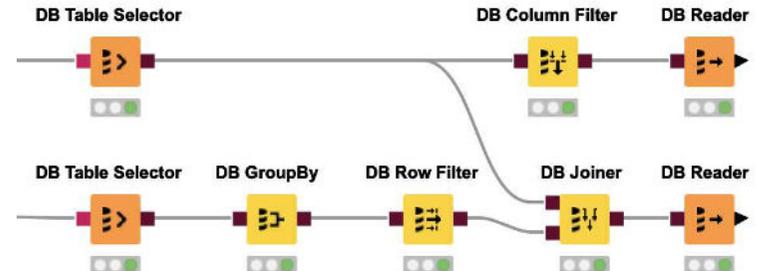
从数据库读取数据需要三个步骤：连接、选择和提取。工作流程是逐步构建的，使用连接器、数据库表选择器和数据库读取器节点。数据库节点仅构建SQL查询，它们不执行查询。只有最终节点，例如数据库读取器节点，才执行SQL查询并提取数据。



有多个数据库连接器节点可用于连接到最常用的数据库。然而，DB连接器节点允许您连接到所有符合JDBC标准的数据库。



还有更多的数据库节点可用于帮助构建用于数据库内的SQL查询。您可以在DB表选择器和DB读取器节点之间使用它们。



使用KNIME Analytics Platform读取数据的提示

所有读取器节点都需要输入文件位置的路径。让我们收集一些提示和技巧：

提示和技巧1：从KNIME资源管理器中使用拖放：

保存在工作区文件夹中的数据文件可在KNIME资源管理器面板（左上面板）中使用。要读取其中一个文件，只需将文件从KNIME资源管理器面板拖放到 workflow 编辑器中。KNIME 会自动创建正确的读取器节点并设置输入位置的路径。

提示和技巧2：定义文件路径的不同选项：

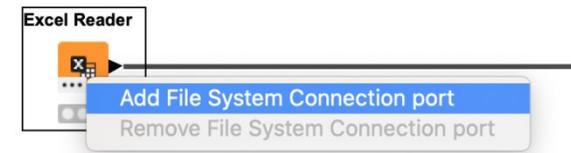
在KNIME中，我们有不同的选项来提供文件路径。当您开始共享 workflow 或将其导出到其他 KNIME Analytics Platform 安装或KNIME服务器时，这变得很重要。KNIME Analytics Platform 中有四个默认文件系统可用。

- 本地文件系统：允许您从本地系统中选择文件/文件夹。
- 挂载点：您可以通过KNIME资源管理器中的其他挂载点连接到KNIME服务器或KNIME Hub。要从LOCAL或其他挂载点读取数据，请选择“挂载点”。选择后，将出现一个新的下拉菜单，您可以选择挂载点。未连接的挂载点会变灰，但仍然可以选择（请注意，在这种情况下，浏览被禁用）。转到KNIME资源管理器并连接到挂载点以启用浏览。
- 相对于：允许您选择是相对于当前挂载点、当前 workflow 还是当前 workflow 的数据区来解析路径。选择后，将出现一个新的下拉菜单，以选择三个选项中的哪一个使用。
- 自定义/KNIME URL：允许指定URL（例如file://、http://或knime://协议）。对于此选项，浏览被禁用。

技巧3：从另一个文件系统中读取：

KNIME分析平台允许您连接和读取许多不同的来源/文件系统，例如Amazon S3，Microsoft SharePoint Online，Databricks等等。需要三个步骤（文件处理指南提供更多信息）。

步骤1：点击“...”在读取器节点图标左下角添加一个文件系统连接端口



步骤2：通过专用连接器节点连接到所需的文件系统，并将其与读取器节点连接起来



步骤3：在连接的文件系统中选择文件/文件夹

追加/连接数据

追加数据

连接节点

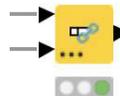
Excel

要将数据追加到表中，请手动选择要追加的区域，并将内容复制并粘贴到第一个表格下方。

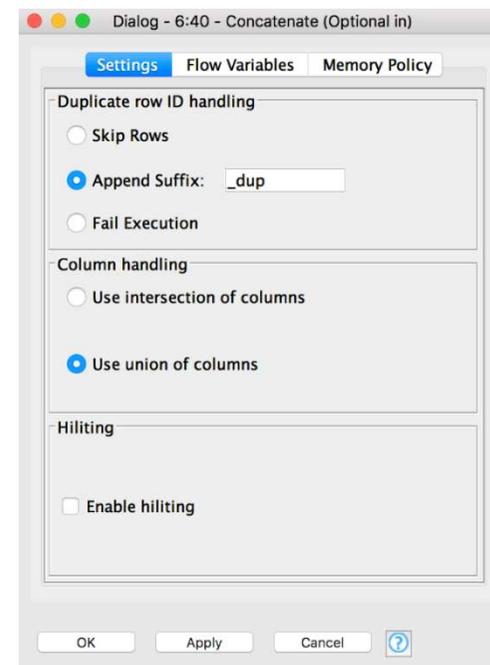
注意1：在复制和粘贴之前，请确保所有表格具有相同的列结构。

KNIME分析平台

Concatenate



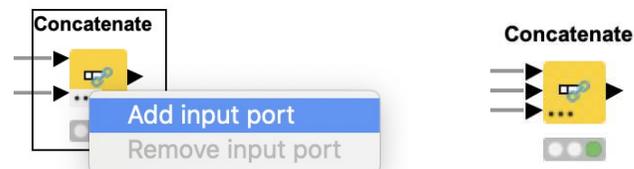
连接节点将两个或更多表格写在彼此下方。



注意1：列的顺序不必相同。

注意2：您可以决定您想要的列的交集还是并集。

注意3：连接节点有选项可以添加更多的输入端口，以连接超过两个表。要添加额外的输入端口，请单击节点左下角的三个点，并选择“添加输入端口”。



VLOOKUP

过滤和连接节点

Excel

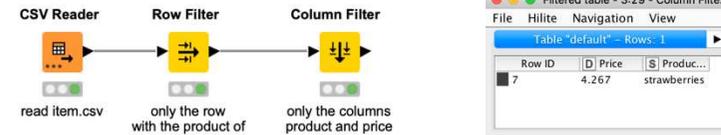
VLOOKUP函数用于不同的任务。最常见的用途有：

1. 查找某个值，例如某个产品的价格。
2. 基于主键（查找值）连接列，例如基于产品ID连接产品信息。

第二个任务的替代函数是INDEX MATCH。

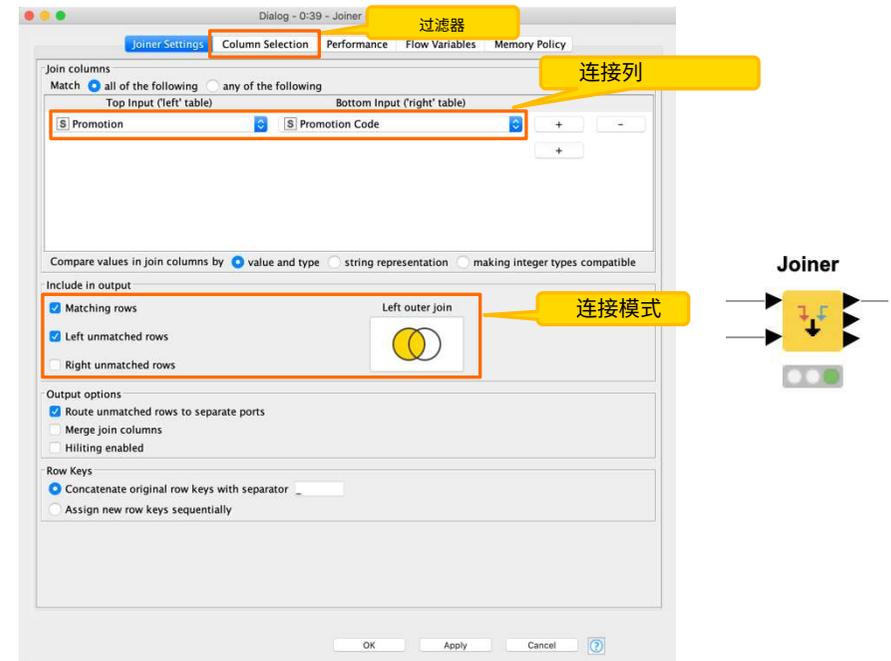
KNIME分析平台

1. 查找某个值，例如某个产品的价格。



注意1：您完整的原始表仍然可在表读取器节点的输出端口中找到。有关行过滤器和列过滤器节点的更多信息，请参阅第26页和第30页。

2. 根据连接列合并列，例如根据产品ID合并产品信息。



注意1：“连接列”是查找值列。

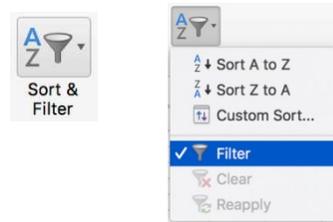
注意2：对于任务2，您还可以使用单元格替换节点

过滤和转换

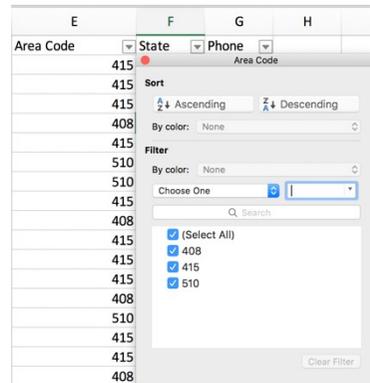
过滤/删除具有特定值的行

Excel

要筛选数据，请在数据表中选择一个随机单元格，转到主页选项卡，单击“排序和筛选”按钮，然后选择“筛选”选项。



从下拉菜单中选择您感兴趣的值。

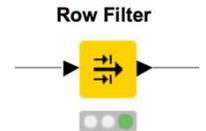


要删除行，请选择要删除的行，右键单击并选择“删除行”。

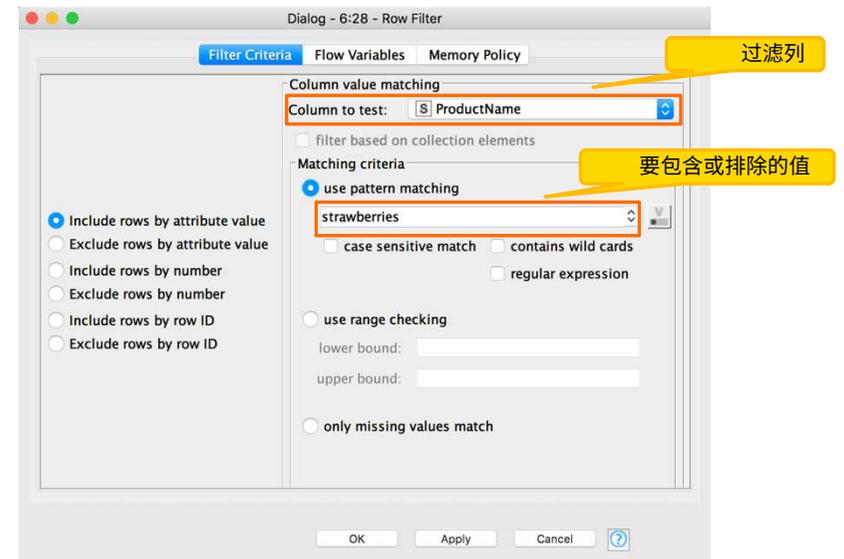
行过滤器

KNIME分析平台

在KNIME Analytics Platform中，过滤和删除行之间没有区别，因为原始表未被删除，并且仍可在前一个节点的输出端口中使用。



行过滤器节点根据过滤条件（例如，在过滤列中包含/排除具有特定值的所有行）对表进行过滤。



注意1： 在左侧，您可以选择是否包括或排除具有匹配值的行

注意2： 如果您只对具有特定值的行感兴趣，可以使用行过滤器节点。

注意3： 如果您想根据不同的值包括行，可以使用基于规则的行过滤器。（请参见下一页）。

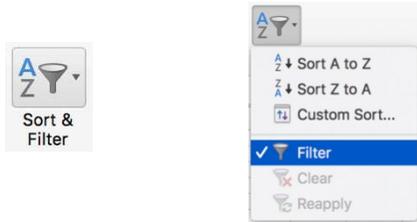
注意4： 还提供其他过滤选项，例如按数值范围过滤行，按行号或行ID过滤行，或仅过滤缺失值。

过滤/删除具有不同值的行

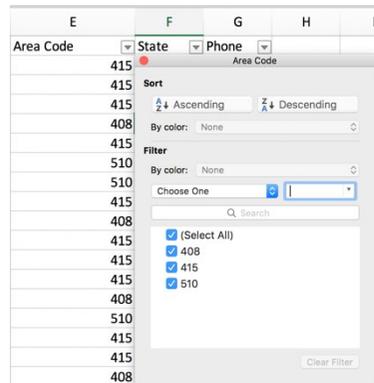
基于规则的行过滤器

Excel

要过滤数据，请在数据表中选择一个随机单元格，点击“排序和过滤”，然后选择“过滤”。



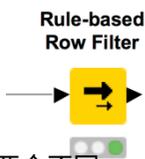
从下拉菜单中选择您感兴趣的值。



要删除行，您必须选择要删除的行，右键单击并选择“删除行”。

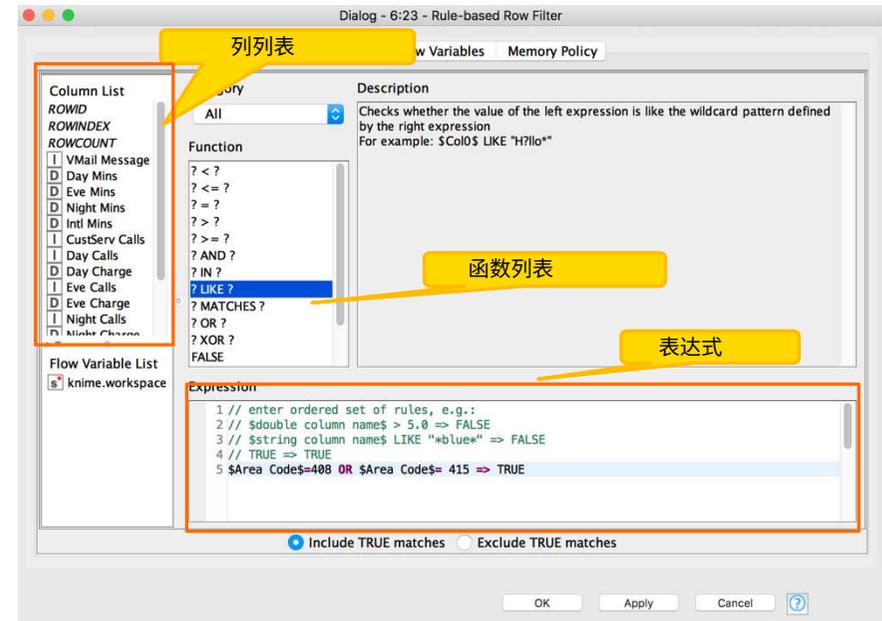
KNIME分析平台

在KNIME Analytics Platform中，过滤和删除行之间没有区别，因为原始表未被删除，并且仍可在前一个节点的输出端口中使用。



基于规则的行过滤器节点根据一组规则过滤行。要包括/排除具有两个不同值的行，可以使用以下表达式

\$列名\$ = “值 1” 或 \$列名\$ = “值 2”=> 真



注释 1: 在配置窗口底部，您可以选择包括或排除真实匹配。

注释 2: 列由其用\$括起来的名称给出。通过双击列表中的列名将它们添加到表达式框架中。注释 3: 基于规则的行过滤器节点具有许多高级过滤选项的不同函数。

注释 4: 表达式框架中的不同行起到OR连接的作用。

删除重复项

重复行过滤器

Excel

要删除重复项

- 选择可能有重复项的单元格范围，然后删除。
- 转到数据选项卡，选择“删除重复项”。

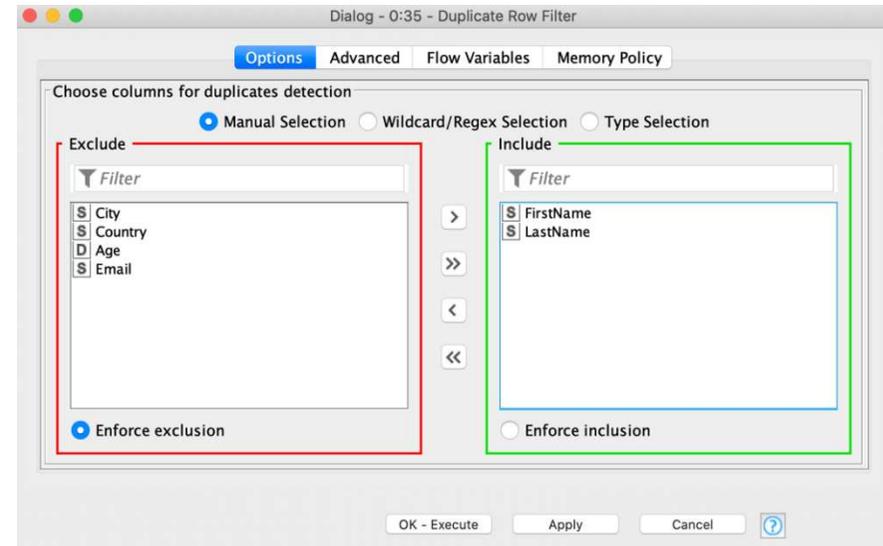
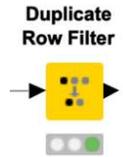


- 选择用于检测重复项的列，然后点击“确定”。



KNIME分析平台

重复行过滤节点检测和处理重复项。默认处理方式类似于Excel中删除重复行。在“包括”框架中的列对应于Excel中选择的用于检测重复项的列。这意味着具有相同值的行将被检测为重复项。



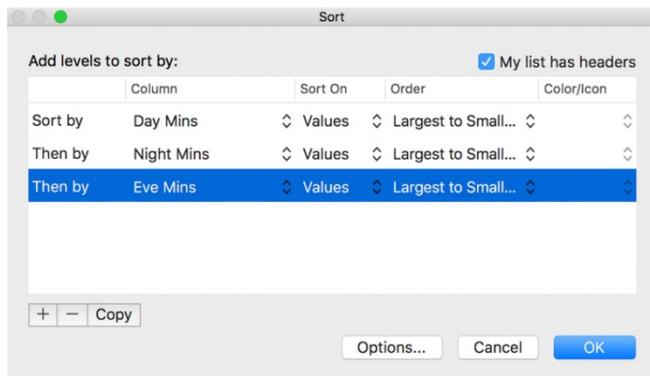
注意1：在“高级”选项卡中，您可以更改重复项的处理方式，例如保留重复行并添加一列显示哪些行是唯一的、选择的或重复的。

按多个关键列对行进行排序

排序节点

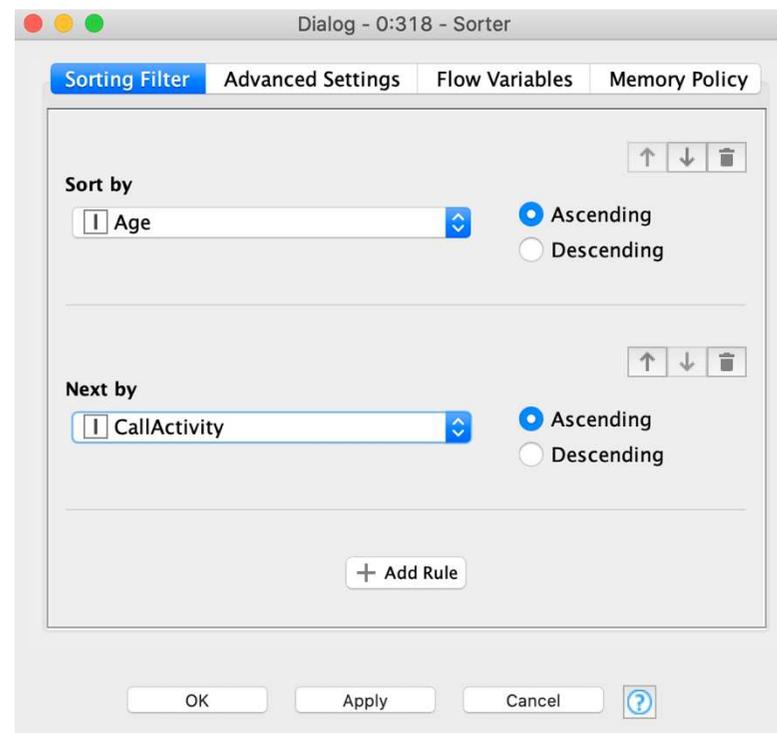
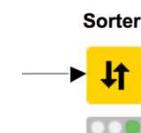
Excel

要按多个列对行进行排序，请选择要排序的列。然后单击“排序和筛选”，选择“自定义排序”。



KNIME分析平台

使用排序器节点，您可以按升序或降序对一个或多个列进行排序。



注意1：您可以通过点击“添加规则”按钮添加任意数量的关键列。

注意2：您可以临时对节点的输出表进行排序。单击要排序的列标题，并选择是升序还是降序排序。

删除列

列过滤器节点

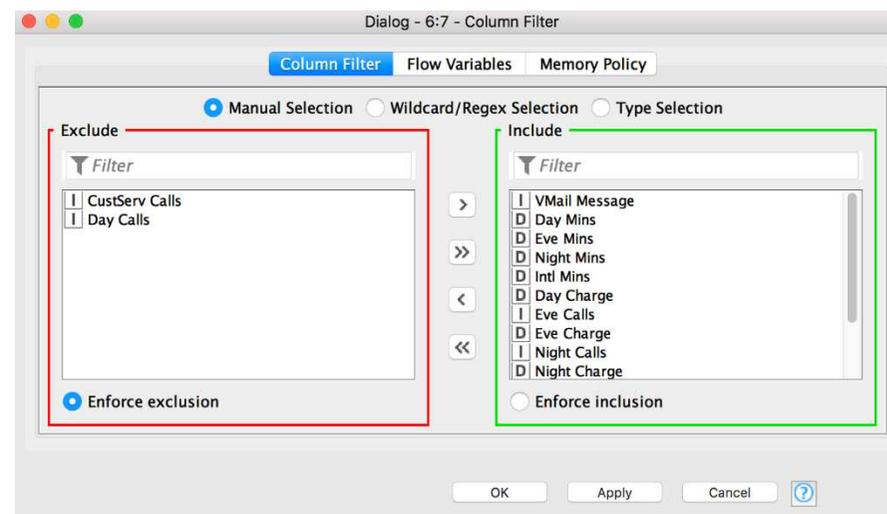
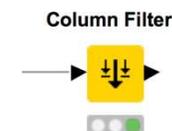
Excel

要删除列，只需选择并删除多余的列。

记住：如果您删除了用于计算字段的列，这将破坏您的公式。

KNIME分析平台

使用列过滤器节点，您可以删除多余的列。请记住，完整的数据表仍将在前一个节点的输出端口和上游计算中可用。



注意1：您可以使用中间的箭头按钮将列从包含框架移动到排除框架，反之亦然。

注意2：您可以使用通配符/正则表达式选择自动删除名称模式的列。

注意3：您可以使用类型选择自动删除数据类型的列。

重新排序和重命名列

列重新排序器和列重命名节点

Excel

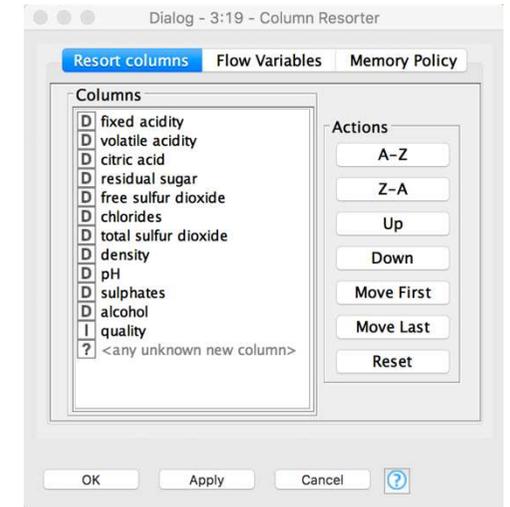
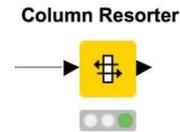
要重新排序列：

- 选择要移动的列的列标题
- 按下ctrl + shift + 下箭头
- 单击并按住要移动的列的绿色轮廓
- 将列拖动到所需位置

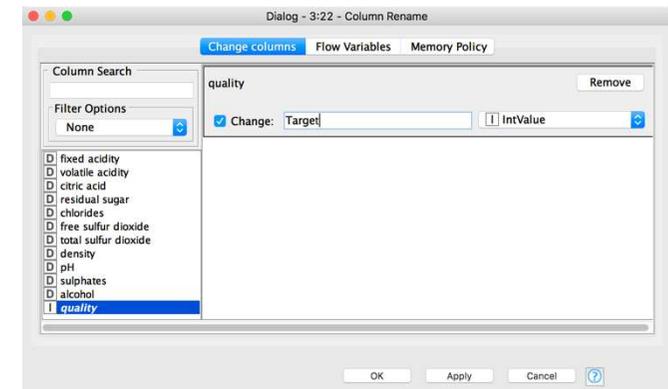
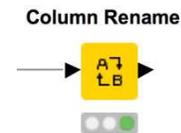
要重命名列，只需单击列单元格并更改单元格值。

KNIME分析平台

要重新排序列，您可以使用列重新排序节点。通过选择要移动的列并使用右侧的“操作”按钮来定义所需的列顺序。



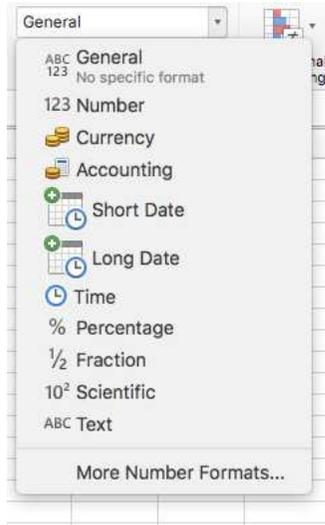
要重命名列，您可以使用列重命名节点。双击要重命名的列，激活复选框“更改”并在文本框中定义列标题。



更改数据类型

Excel

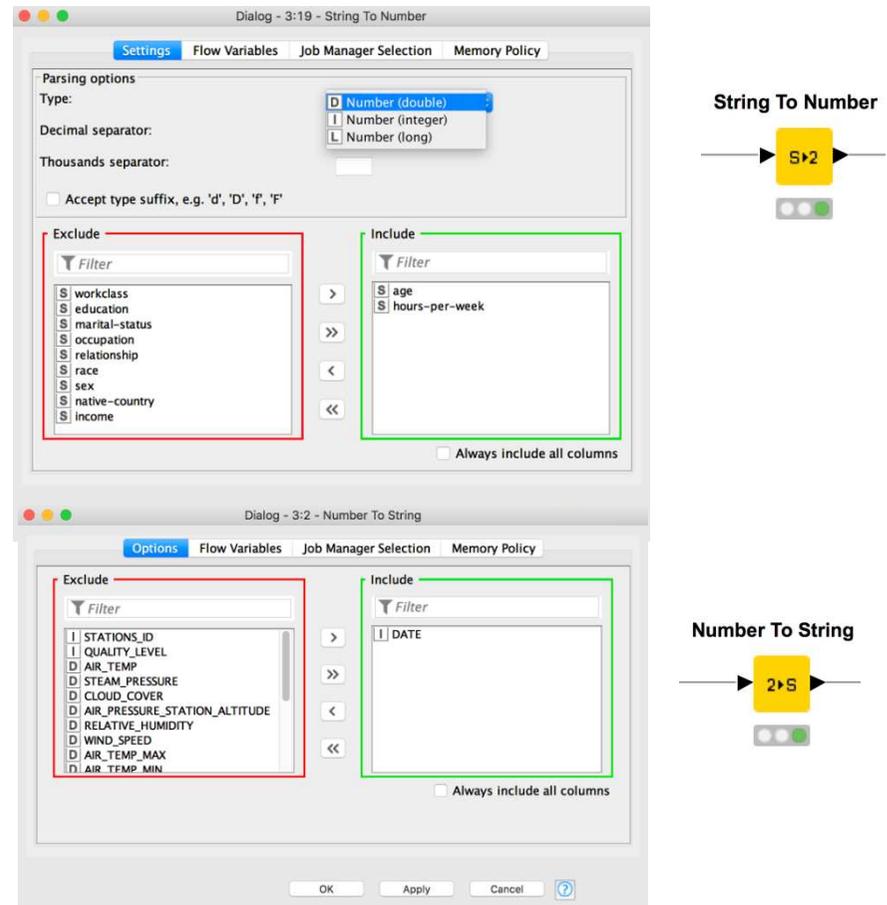
要更改列或单元格的数据类型，请选择它并转到主页选项卡。然后使用中间的下拉菜单。



字符串转数字和数字转字符串

KNIME分析平台

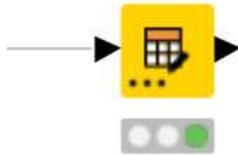
要更改列的数据类型，您可以使用字符串转数字或数字转字符串节点。



注意1：在字符串转数字节点中，您可以选择不同的数值类型，例如Double、Integer和Long。

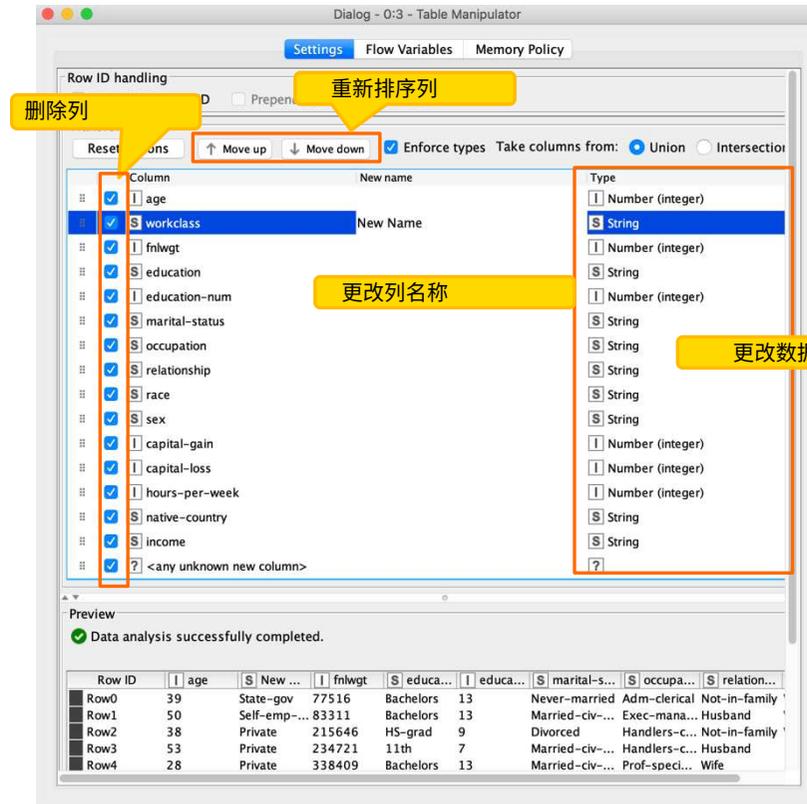
使用KNIME Analytics Platform进行数据操作的提示

Table Manipulator



表操作节点允许您在一个节点中执行许多描述的转换。您可以使用它来：

- 连接多个表格（在节点左下角点击“...”添加动态输入端口）
- 过滤、排序和重命名列
- 更改列的数据类型



数据聚合

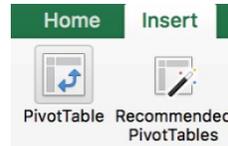
数据透视表

数据透视节点

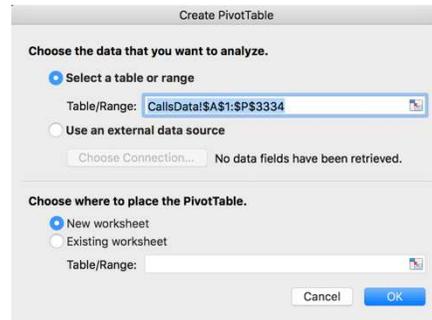
Excel

在Excel中创建数据透视表

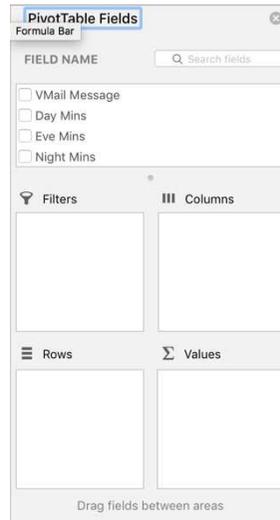
- 点击数据集中的任意单元格
- 转到“插入”选项卡，点击“数据透视表”



- 选择表格/范围和输出位置



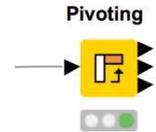
- 通过将其拖动到“列”、“行”和“值”中选择数据透视表字段



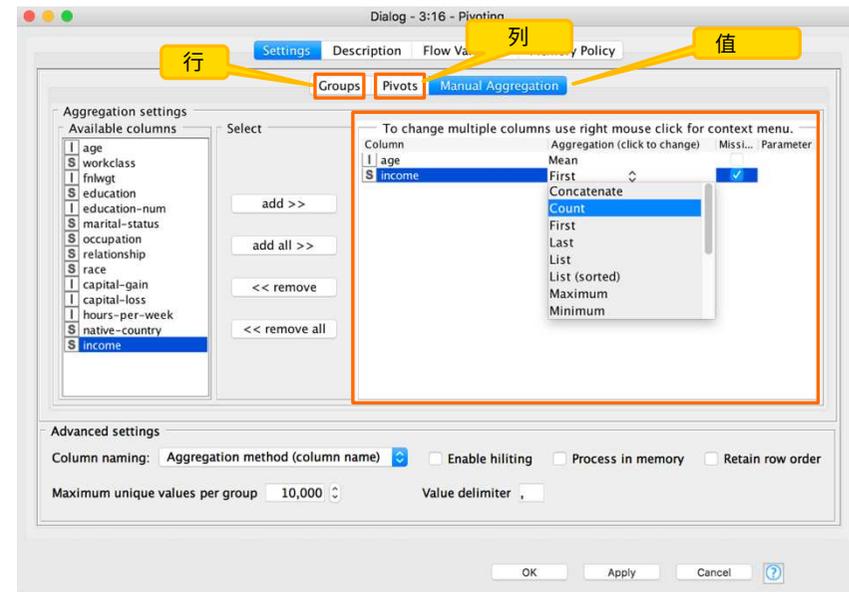
KNIME分析平台

通过三个选项卡配置数据透视节点：

- “分组”定义分组列，即Excel中的“行”框架（最终行ID）
- “透视”定义透视列，即Excel中的“列”框架（最终列标题）
- “手动聚合”对应于“值”设置选项。从可用列列表选择一个或多个列进行聚合，并为每个选定的列选择聚合方法。



枢轴节点生成三个输出表：枢轴表和列和行的总值。



注1：枢轴节点没有“筛选”选项，但是您可以在之前简单地使用行筛选节点。

注2：在KNIME中，您必须至少选择一列作为组和枢轴。如果您只想选择“行”，可以使用GroupBy节点。

Excel

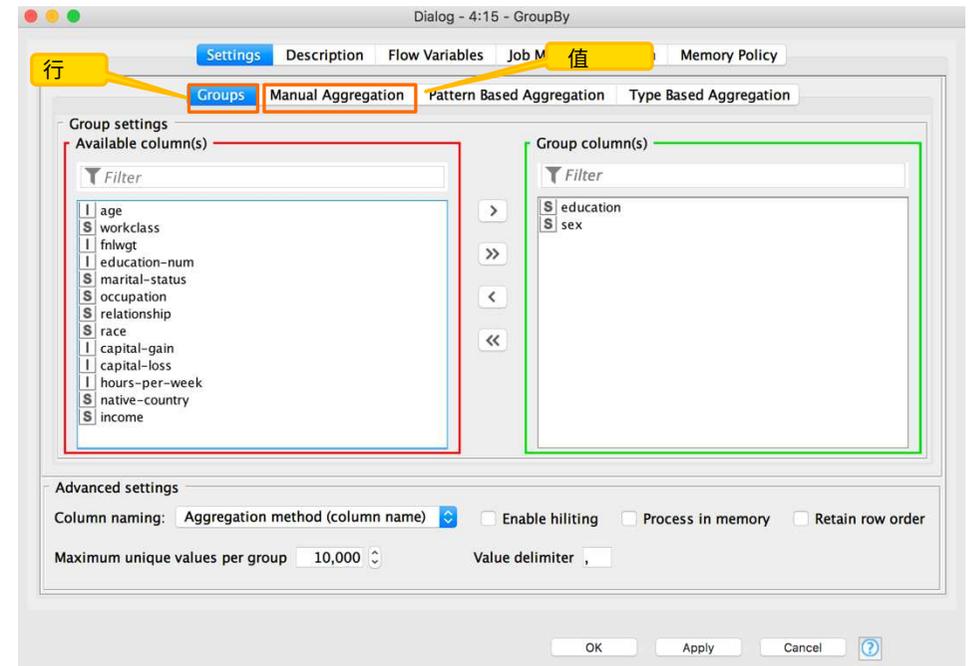
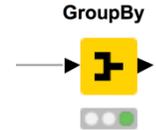
按照前一页所述创建一个枢轴表，并将列仅拖放到“行”和“值”框中。

KNIME分析平台

GroupBy节点通过两个选项卡进行配置：

- “组”定义组列，也就是Excel中的“行”框中的列（最终行ID）
- “手动聚合”对应于“值”设置选项。选择一个或多个列用于从可用列列表中聚合并为每个选定的列选择一个聚合方法。

“值”设置选项对应于“手动聚合”选项卡。



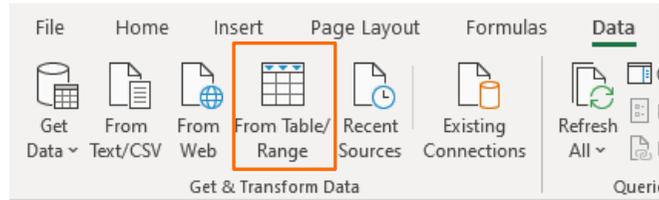
解开

解开枢轴节点

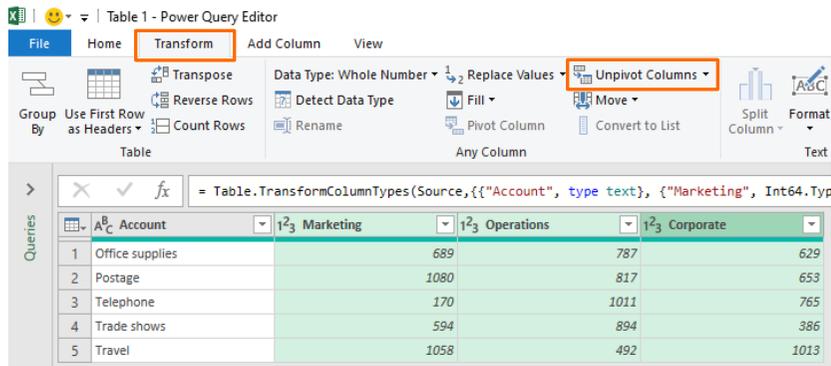
Excel

在Excel中创建一个解开枢轴表

- 将数据集存储在一个表中。
- 选择表中的任意单元格。
- 点击“数据”选项卡，然后选择“从表/范围”选项。



- 这将打开“Power Query Editor”。通过按住Shift键选择要解开枢轴的列。
- 点击Power Query Editor的“转换”选项卡，然后选择“解开枢轴列”。



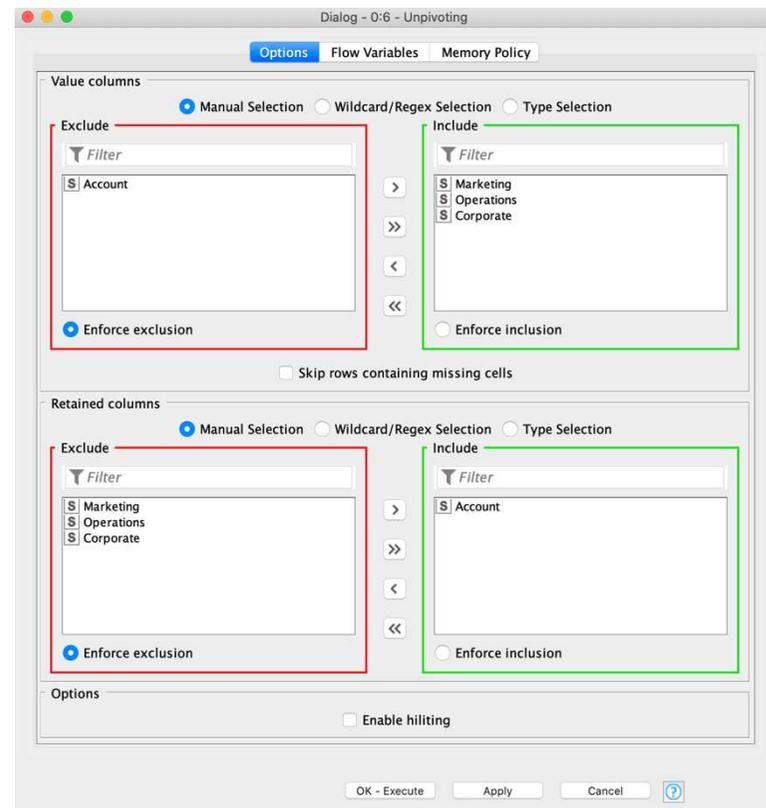
- 点击“Power Query Editor”的“主页”选项卡，然后选择“关闭并加载”以将解开枢轴的数据保存回Excel工作簿中。

注意1: 在Excel 2016 for Windows中，无需任何额外下载即可使用解开枢轴命令。如果您使用的是其他版本，则可能需要先从Microsoft网站下载免费的Power Query插件。目前不支持在Mac上进行Power Query Editor的编写。

KNIME分析平台

通过一个选项卡配置Unpivoting节点：

- 在上部分，“Value columns”定义了解开枢轴的列，也就是Excel中的列选择。
- 在下部分，“Retained columns”定义了保持不变的列，也就是Excel中的未选择列。



注意1: Unpivoting节点需要在“Value columns”部分选择至少一列。

数学函数和文本函数

Excel

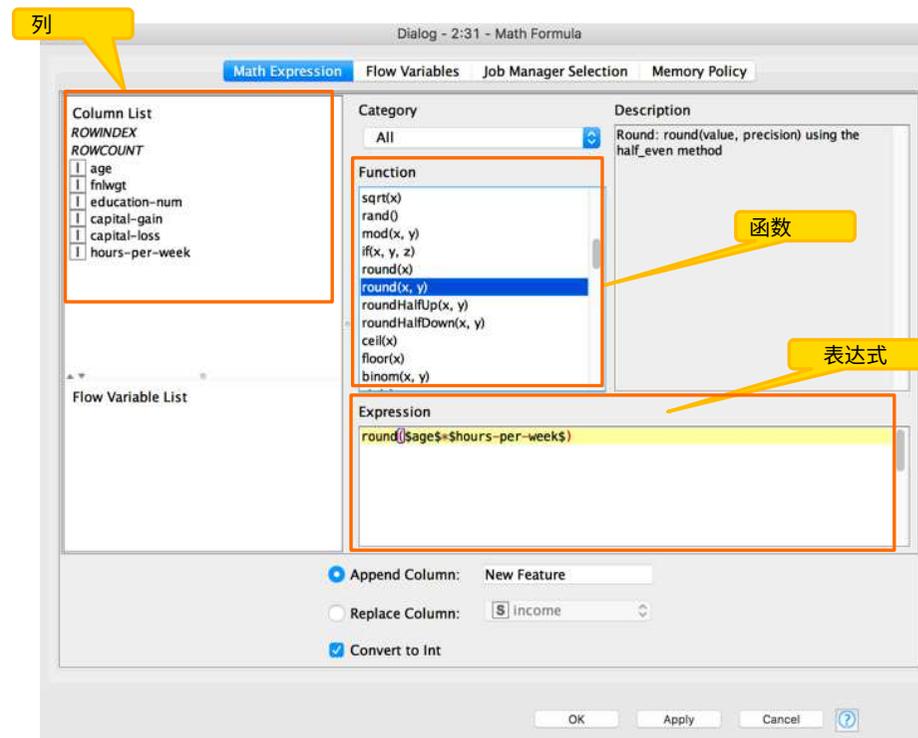
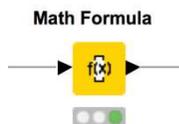
这是Excel中经常使用的数学公式列表。你将在下一页找到KNIME公式的翻译：

- 求和
- 平均值
- 中位数
- 乘积求和
- 绝对值
- 条件求和
- 四舍五入函数
 - ROUND
 - ROUNDUP
 - ROUNDDOWN

如果你经常使用的数学公式缺失，请发送邮件至kathrin.melcher@knime.com。

KNIME分析平台

Math Formula节点根据行中的值计算数学表达式。它有一个很长的函数列表。下一页的表格提供了最常用的Excel函数的翻译。



注意1：您可以通过在表达式框下方使用复选框来决定是追加新列还是替换列之一。

注意2：通过激活“转换为整数”复选框，您可以确保输出的追加/替换列为整数类型。

注意3：要对多个列执行相同的数学表达式，可以使用数学公式（多列）节点。

数学函数

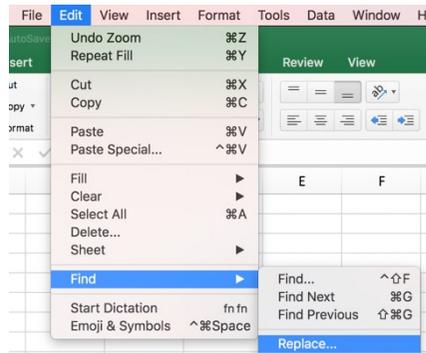
数学公式节点

Excel	KNIME分析平台
SUM(数字1;数字2;...)	COL_SUM(列) = 选定列中的值的总和。 Col1 + Col2。 = Col1中的值和Col2中的值的总和，对于每一行。
AVERAGE(数字1;数字2;...)	COL_MEAN(列名) = 选定列中的值的平均值。 average(Col1, Col2,...) = 选定列的值的平均值，对于每一行。
MEDIAN(数字1;数字2;...)	COL_MEDIAN(列名) = 选择列中值的中位数。 中位数(列1, 列2, ...) = 每行所选列的值的的中位数。
SUMPRODUCT(数字1;数字2;...)	两个“数学公式”节点的序列： 第一个： 使用表达式Col1*Col2将两列相乘，并附加一个新列。 第二个： 在新列上使用COL_SUM。
ABS(数字1)	abs(列) = 所选列中所有值的绝对值
SUMIF(范围; 条件)	两个“数学公式”节点的序列： 第一个： 附加一个新列，如果满足条件，则为列求和，否则为0。 第二个： 在新列上使用COL_SUM。
舍入函数： ROUND(数字1, 小数位数) ROUNDUP(数字1, 小数位数) ROUNDDOWN(数字1, 小数位数)	round(列, 小数位数) = 小数位数是可选的。 ceil(列*10^(小数位数))/10^小数位数 floor(列*10^(小数位数))/10^小数位数

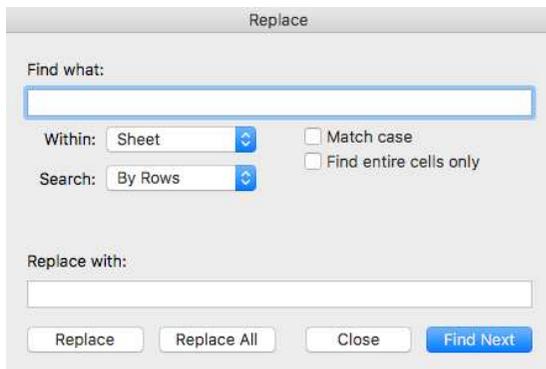
连接和查找替换

Excel

1. 函数CONCATENATE (text1, text2, ..) 将不同的文本组合成一个单元格。
2. 要查找和替换特定值，您必须转到“编辑->查找->替换...”以打开下面的对话框。



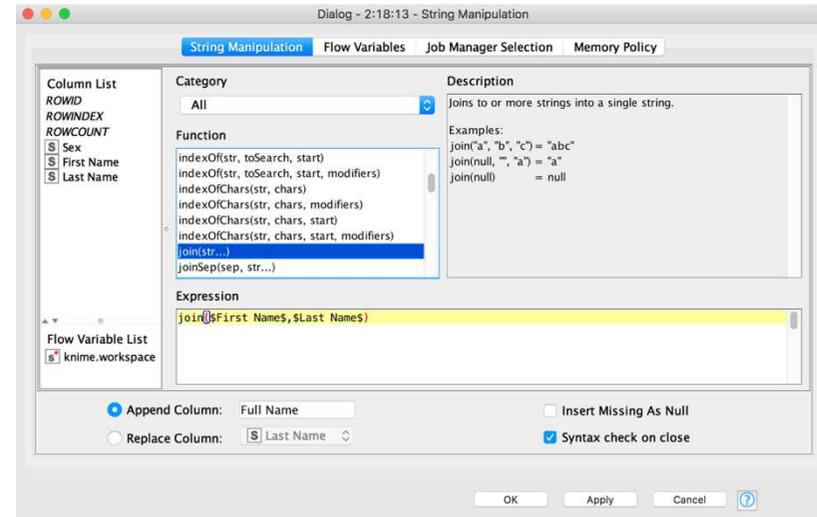
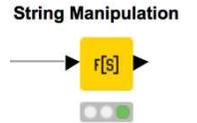
在打开的对话框中，您可以定义要替换的值和要替换为的值。



字符串操作节点

KNIME分析平台

字符串操作节点根据定义的表达式对类型为String的列进行操作。有很多不同的函数可用。



1. 来自Excel的CONCATENATE (text1, text2, ..)

join (col1, col2, ...) 或 join (col1, "your own string", ...)

注意1：您可以将不同列的值按行连接，或者还可以将另一个String值添加到列中。

2. 查找和替换

Replace(string, search, replace)

注意1：“string”对应于您想要替换某些内容的列。

注意2：要对多个字符串列执行相同的操作，可以使用字符串操作（多列）节点

格式化Excel表格

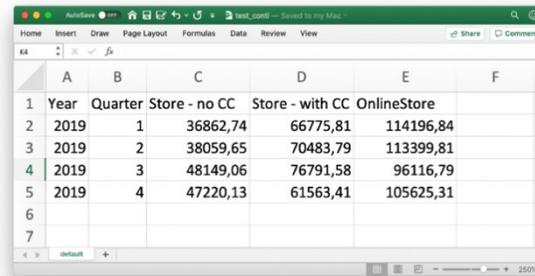
在第1章中，我们介绍了[Excel Writer](#)，您可以使用它将结果表写入Excel表格。默认情况下，这是一个简单的表格，没有任何格式，如颜色、边框单元格等。在本章中，我们想向您展示如何使用KNIME的社区扩展[Continental Nodes](#)的XLS Formatting节点。这些节点使您能够向已有的XLS文件添加格式化指令和高级设置，以便创建具有您所需外观和感觉的Excel报告。

您格式化的Excel表格的关键是与原始文件大小相同的附加XLS控制表，其中包含一个或多个逗号分隔的标签值，例如标题、边框等。不同的XLS Formatter节点根据这些标签为单元格分配不同的格式化指令，例如您可以更改所有标记为“header”的单元格的背景颜色。

然后，您可以使用[XLS格式化程序（应用）](#)节点将格式应用于已存在的Excel表格。

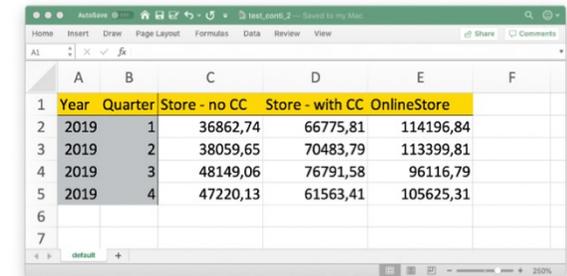
提示：使用流变量连接确保Excel文件已经被写入。

Standard written table



Year	Quarter	Store - no CC	Store - with CC	OnlineStore
2019	1	36862,74	66775,81	114196,84
2019	2	38059,65	70483,79	113399,81
2019	3	48149,06	76791,58	96116,79
2019	4	47220,13	61563,41	105625,31

Styled table



Year	Quarter	Store - no CC	Store - with CC	OnlineStore
2019	1	36862,74	66775,81	114196,84
2019	2	38059,65	70483,79	113399,81
2019	3	48149,06	76791,58	96116,79
2019	4	47220,13	61563,41	105625,31

Continental Nodes for KNIME
XLS Formatter Nodes

图1：左侧是由Excel Writer节点创建的没有格式的Excel表格，右侧是在添加了XLS Formatter节点的格式化信息后的样式化表格，例如标题的黄色背景。

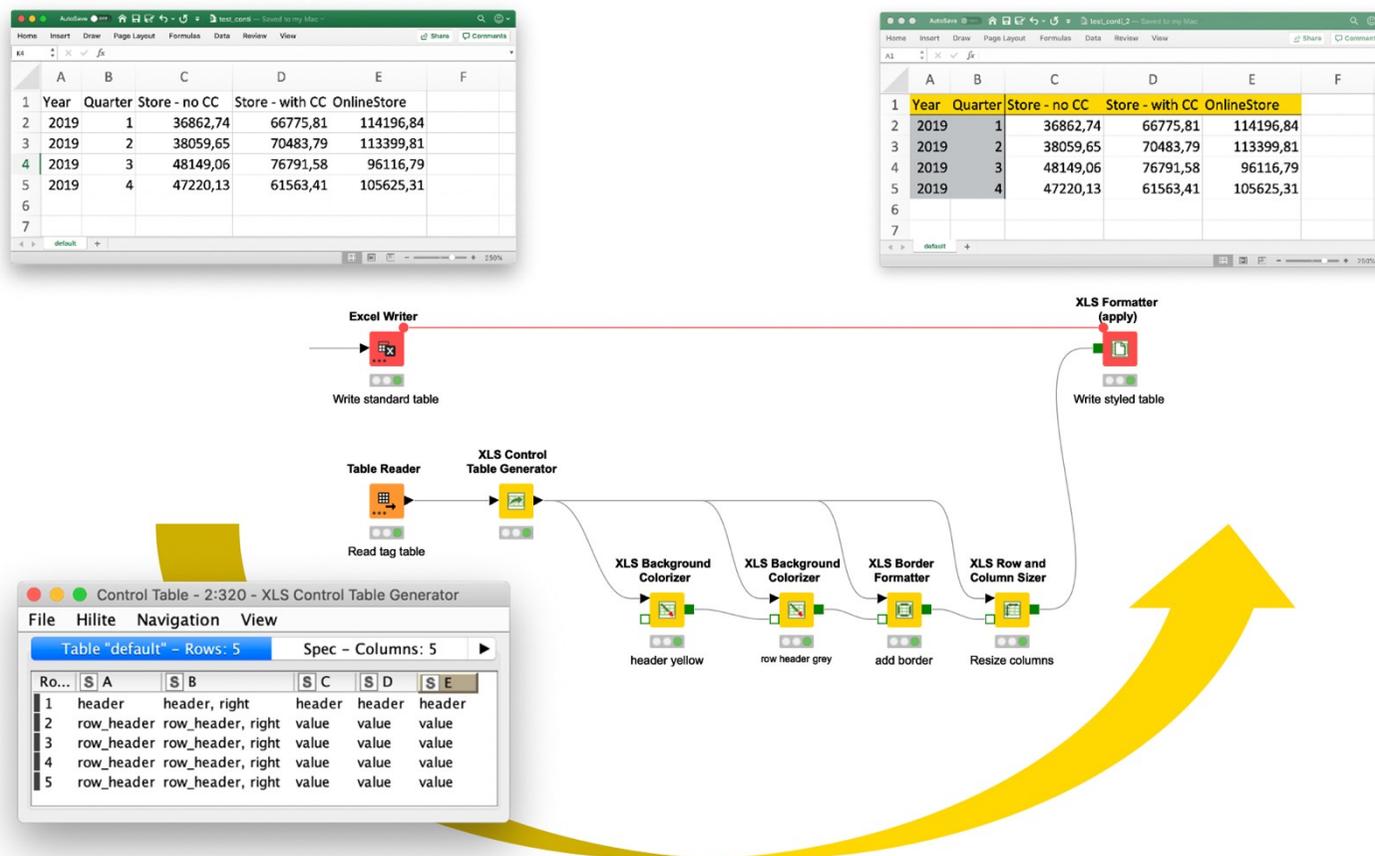


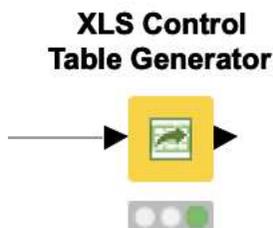
图2左下角显示了一个带有标签值的控制表，这是您样式化表格的关键。基于标签值，黄色的XLS格式化节点收集格式化指令，然后由XLS格式化程序（应用）节点应用，生成样式化的表格（右上角）。

总结一下，这意味着我们需要执行以下步骤：

1. 将表格写入Excel表格
2. 创建一个带有标签值的XLS控制表
3. 根据标签值添加格式化指令
4. 将格式化指令应用于现有的Excel表格

本章分为两个部分。本章的第一部分展示了两种创建带有标签的XLS控制表的方法。第二部分介绍了一些可用于添加格式化指令的节点。

提示：你无法在节点库中找到节点吗？[KNIME的大陆节点是一个社区扩展](#)，您可以通过将扩展从[KNIME Hub](#)拖到[KNIME Analytics Platform](#)或按照此视频中的说明安装扩展。[_____](#)



使用标签值创建XLS控制表

俗话说，条条大路通罗马。本节介绍了创建XLS控制表的两种不同方法或途径。（第二种方法恰好是我个人最喜欢的！）两个示例中的“关键节点”都是XLS控制表生成器节点。
[_____](#)

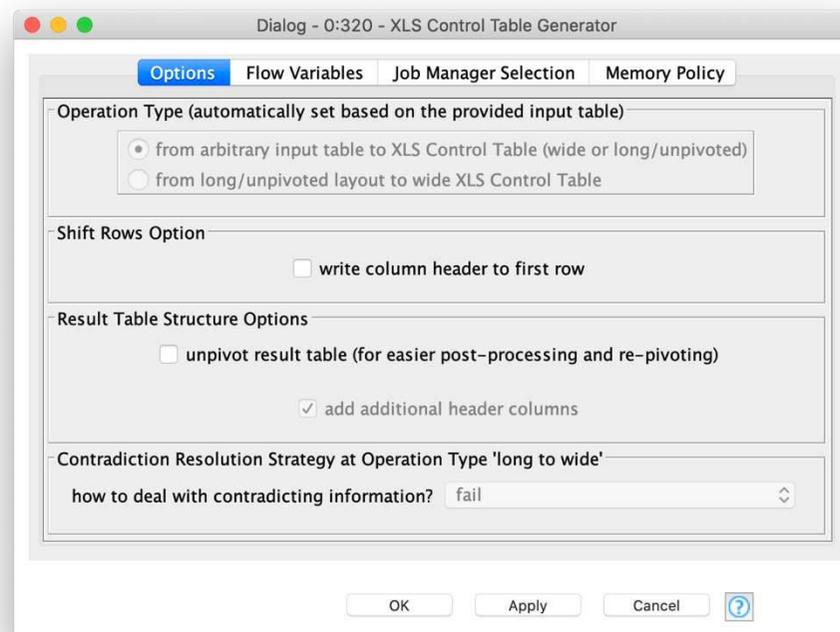


图3: XLS控制表生成器节点的配置对话框

方法1：表格创建器 + XLS控制表生成器

创建带有标签的XLS控制表的第一种方法涉及表格创建器和XLS控制表生成器节点的组合。

这是一种简单的方法，但缺点是它需要大量手动工作来创建标签表，并且标签表是静态的。因此，只建议在行数和列数不会改变的小表格中使用这种方法。

打开Table Creator节点的配置窗口，为每个单元格添加一个或多个标签值。如果您想输入多个标签，请记得用逗号分隔它们。XLS Control Table Generator节点将表格转换为XLS控制表，并用字母替换列名和数字替换行ID。

复选框“将列标题写入第一行”可以选择保留列标题，类似于Excel Writer节点中的“添加列标题”选项。

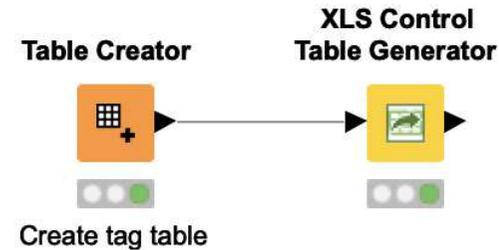


图4：这里您可以看到使用Table Creator节点和XLS Control Table Generator节点创建控制表的一个选项。

方法2：XLS Control Table Generator + Rule Engine

我们要展示的第二种方法是使用两个XLS控制表生成器节点的组合，激活unpivot选项和一个规则引擎节点。

在第一个XLS控制表生成器节点的配置窗口中，勾选“unpivot结果表”复选框，以输出一个包含输入表中每个单元格的行数、列标题等信息的表格。

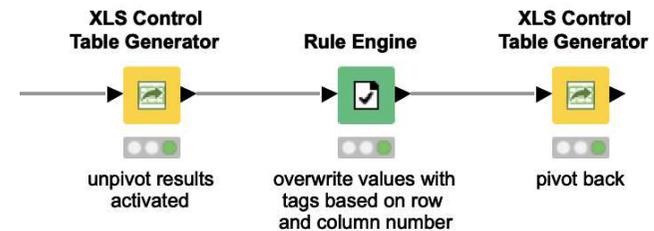


图5：在这里，您可以看到使用两个XLS控制表生成器节点和一个规则引擎节点组合创建控制表的另一种选项。规则引擎节点用于根据行号和列号定义标签值。

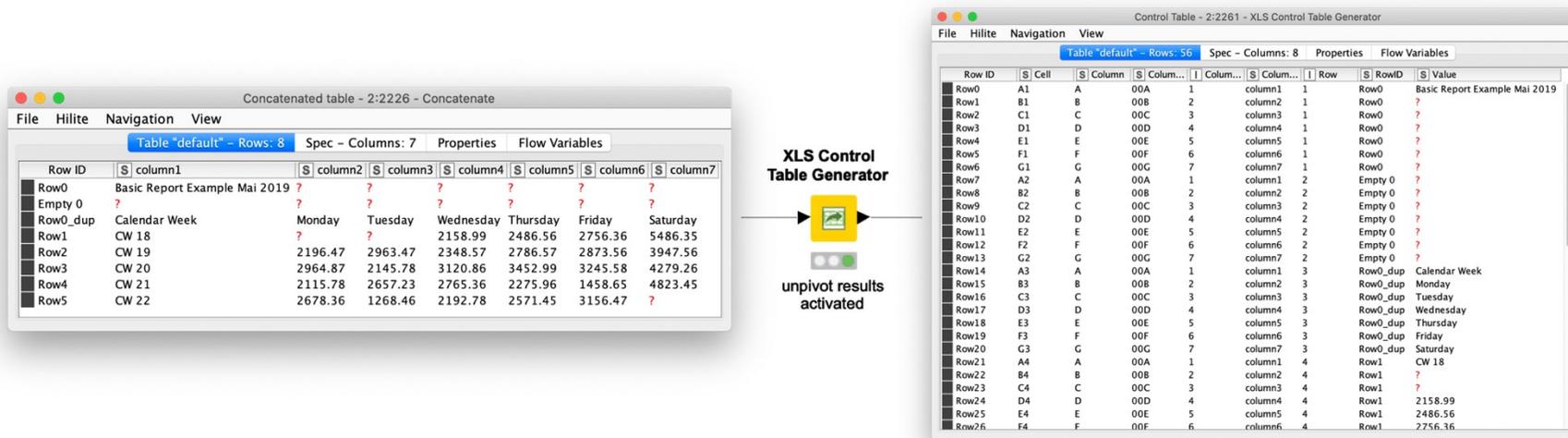
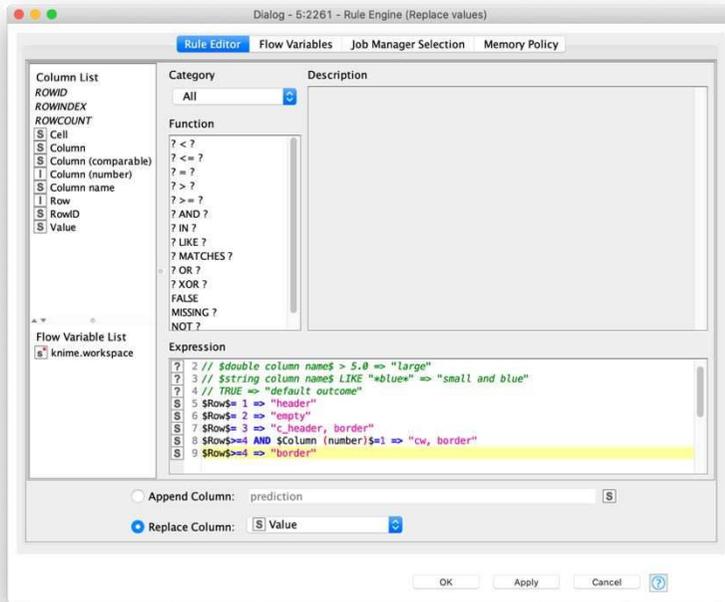


图6在XLS控制表生成器节点中激活“unpivot结果表”复选框后，您可以看到生成的表格。该节点为每个单元格创建一行，包括值、行号、列标题等。

这个表格是将值转换为标签的理想基础，可以使用规则引擎节点进行转换。
 例如，我们可以将第一行中的所有值替换为“header”标签，或者将第一列中行号大于3的所有值替换为“cw”标签。

提示：勾选“替换列”复选框，并选择“值”列。



Row ID	Cell	Column	Column...	Column...	Column...	Row	RowID	Value
Row0	A1	A	00A	1	column1	1	Row0	header
Row1	B1	B	00B	2	column2	1	Row0	header
Row2	C1	C	00C	3	column3	1	Row0	header
Row3	D1	D	00D	4	column4	1	Row0	header
Row4	E1	E	00E	5	column5	1	Row0	header
Row5	F1	F	00F	6	column6	1	Row0	header
Row6	G1	G	00G	7	column7	1	Row0	header
Row7	A2	A	00A	1	column1	2	Empty 0	empty
Row8	B2	B	00B	2	column2	2	Empty 0	empty
Row9	C2	C	00C	3	column3	2	Empty 0	empty
Row10	D2	D	00D	4	column4	2	Empty 0	empty
Row11	E2	E	00E	5	column5	2	Empty 0	empty
Row12	F2	F	00F	6	column6	2	Empty 0	empty
Row13	G2	G	00G	7	column7	2	Empty 0	empty
Row14	A3	A	00A	1	column1	3	Row0_dup	c_header,...
Row15	B3	B	00B	2	column2	3	Row0_dup	c_header,...
Row16	C3	C	00C	3	column3	3	Row0_dup	c_header,...
Row17	D3	D	00D	4	column4	3	Row0_dup	c_header,...
Row18	E3	E	00E	5	column5	3	Row0_dup	c_header,...
Row19	F3	F	00F	6	column6	3	Row0_dup	c_header,...
Row20	G3	G	00G	7	column7	3	Row0_dup	c_header,...
Row21	A4	A	00A	1	column1	4	Row1	cw, border
Row22	B4	B	00B	2	column2	4	Row1	border
Row23	C4	C	00C	3	column3	4	Row1	border
Row24	D4	D	00D	4	column4	4	Row1	border
Row25	E4	E	00E	5	column5	4	Row1	border
Row26	F4	F	00F	6	column6	4	Row1	border

图7左侧是Rule Engine节点的配置对话框。在表达式部分，您可以看到根据行和列号定义的规则，用标签替换原始值。在右侧，您可以看到应用规则并用不同标签替换值的输出表格。

第二个XLS控制表生成器节点可以将此表格转换回原始形式，其中值被不同的标签替换。当节点检测到由XLS控制表生成器节点以非枢轴模式创建的输入表格时，此功能会自动激活。

与第一种方法相比，这种方法需要更少的手动工作，并且可以优雅地处理表格尺寸的变化。

提示：创建静态XLS控制表的另一个有用节点是XLS控制表从单元格范围节点。

根据标签值添加格式化操作

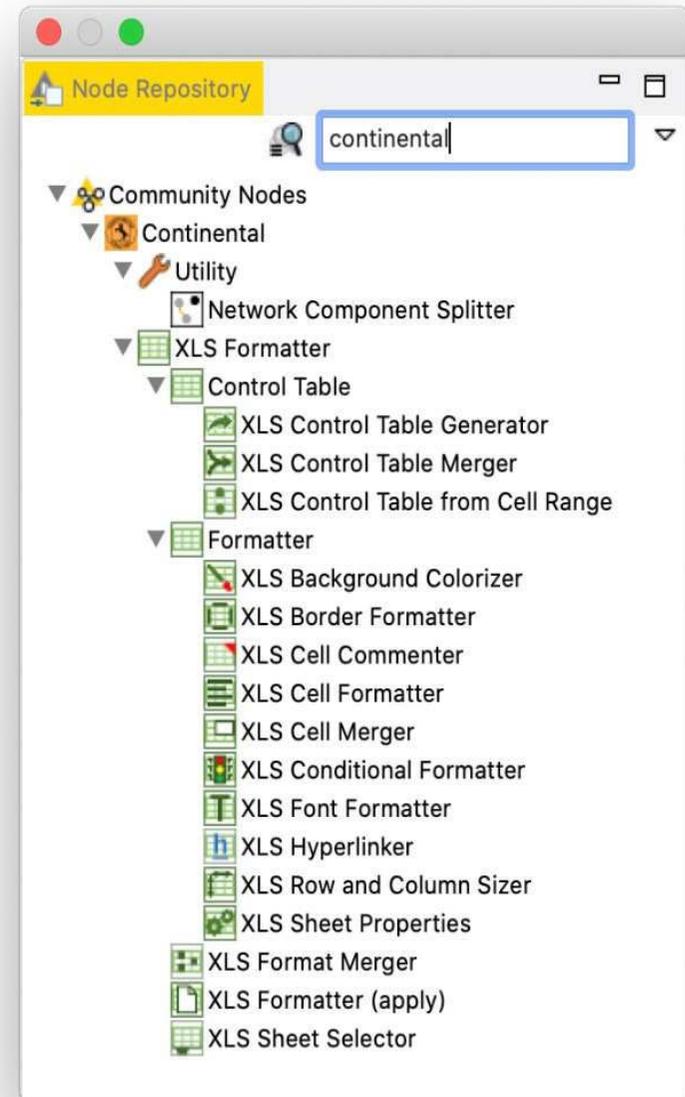
为您的表格添加背景颜色或边框等的下一步是一系列的XLS格式化节点，类似于图2中的工作流程。

正如您所看到的，示例工作流程中的所有节点都有两个输入端口和一个输出端口：

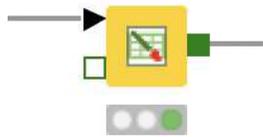
- 一个数据输入端口
- 一个可选的XLS格式化器输入端口（带有绿色边框的正方形）
- 一个输出端口，即XLS格式化器端口

绿色正方形是扩展的特殊端口类型，用于收集不同的格式化指令。数据输入端口期望带有标签值的表格。可选的输入端口可用于提供具有先前的XLS格式化指令的XLS格式化表格，节点的指令将添加到其中。

右侧的图表向您展示了大陆扩展中的所有节点的概述。我将介绍我最喜欢的几个，并留给您去探索其他的。

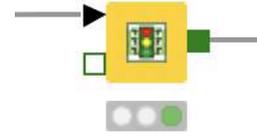


XLS Background Colorizer



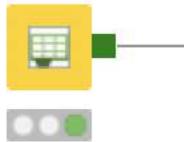
XLS背景颜色节点可以更改单元格的背景颜色。您可以分配一个静态颜色和/或图案填充。一种选项是为具有特定标签值的所有单元格分配相同的颜色，例如所有具有“header”标签的单元格应具有黄色背景。另一种选项是使用十六进制语法#FFD800或十进制语法R/G/B作为标签，并将它们用作背景颜色。

XLS Conditional Formatter



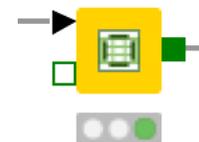
XLS条件格式化节点根据其数值更改具有特定标签值的单元格的背景。在配置窗口中，您可以通过设置最小和最大值并为每个值分配颜色来定义颜色比例。您还可以设置中间值并为其分配颜色。值高于或低于阈值的单元格将具有最小/最大值的背景颜色。

XLS Sheet Selector



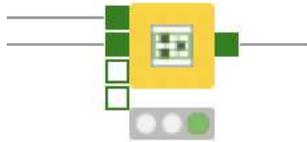
如果您的EXCEL文件有多个工作表，XLS工作表选择器和XLS合并器节点非常有用。默认情况下，格式化始终应用于第一个工作表。因此，如果您的Excel文件只有一个工作表，您不必担心这两个节点。但是，如果您有多个工作表，XLS工作表选择器允许您定义XLS控制表所属的工作表。

XLS Border Formatter



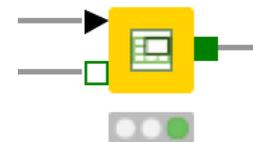
XLS边框格式化节点可以为由特定标签或所有标签指定的给定范围添加边框。通过激活相应的复选框，您可以为顶部、右侧、底部和左侧添加边框。除了在标签指定的范围周围添加边框外，该节点还可以选择在每个单元格中使用内部垂直和水平边框线。

XLS Format Merger

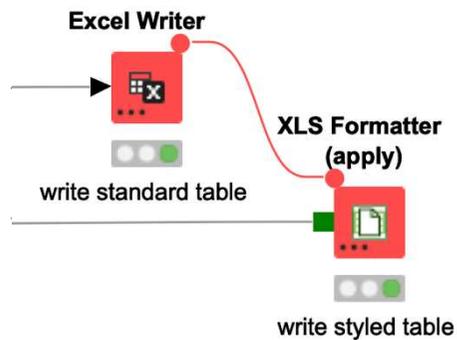


XLS格式合并节点允许您在使用XLS格式化（应用）节点之前，将不同工作表的格式指令合并，或者在应用于同一工作表时，以最低的详细级别合并属性（例如，控制表一中单元格A1的格式指令为粗体，控制表二中为斜体。单元格A1的后续格式指令为斜体和粗体）。因此，在冲突信息的情况下（例如，同一单元格的两种不同字体颜色），上方输入端口将覆盖下方输入端口。

XLS Cell Merger



XLS单元格合并节点将给定的输入标签的矩形范围内的单元格合并为一个单元格。例如，我们可以合并第一行的所有单元格，并使用XLS字体格式化节点将标题居中。此节点仅适用于严格的矩形范围。合并单元格的值是合并范围最左上角单元格的值。



XLS格式化程序（应用）节点读取未格式化的Excel文件，应用所有收集到的格式化指令，并将漂亮的Excel文件保存在定义的输出位置。

这是一个简短的介绍。您可以在文档<https://www.knime.com/community/continental-nodes-for-knime-xls-formatter>中找到有关不同XLS格式化程序节点的更多信息，或者在KNIME Analytics Platform内查看每个单独节点的节点描述。

The KNIME Booklet for Excel Users

您是一位有经验的Excel用户，想要开始使用KNIME Analytics Platform吗？

从一个软件工具切换到另一个软件工具有时候很困难。但是这本小册子是一个完美的起点，因为它将最常用的Excel函数和技术映射到它们在KNIME中的等效物。例如，了解在KNIME中如何进行数据读取、过滤、排序和vlookup。

要了解KNIME的完整介绍，请参考我的书“KNIME初学者的幸运”
在KNIME Press的<https://www.knime.com/knimepress>下可获得。

关于作者

Kathrin Melcher目前是KNIME的数据科学家。她拥有德国康斯坦茨大学的数学硕士学位。她于2017年5月加入了KNIME传教团队，并对数据科学、机器学习和算法有浓厚的兴趣。她喜欢教学并分享她在这些领域的知识。