

# from Excel® to KNINE®

Author: Kathrin Melcher

knimepress@knime.com

KNIME出版社 Talacker 50 8001 Zurich 瑞士

版权所有 © 2022 KNIME出版社

本书已更新至KNIME 4.6版本。

有关许可和销售信息,请写信至:

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。

2

常规用法		5
电子表格	工作流程和节点	6
文件夹     工作团	٢	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	容导入到单个表中	14
将多个工作表的内容导入到	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 Plat	tform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	日行 基于规则的行过滤器	 '
删除重复项	重复行过滤器	 ;
按多个关键列对行进行	排序       排序节点	 )
删除列	列过滤器节点	 )
重新排序和重命名列	列重新排序和列重命名节点	
更改数据类型	字符串转数字和数字转字符串	 <u>.</u>
使用KNIME Analytics I	Platform进行数据操作的提示	 ;
数据聚合		 ŀ
数据透视表	数据透视节点	 j
没有列的数据透视表	分组节点	 )
解开 解开 解开	开节点	 '
数学函数和文本函数		 ;
数学函数	数学公式节点	 J
数学函数	数学公式节点	 J
连接和查找替换	字符串操作节点	
格式化Excel表格		 !
使用标签值创建XLS控制	制表	 j
根据标签值添加格式化	操作	 )



```
电子表格
```

#### 工作流程和节点

Excel	KNIME分析平台
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。这可以是非常简单的单 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,E xcel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。
Autore     Callbri (Body)     I     I     A     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I     I	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。
34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34       34 <td< td=""><td>Node     Workflow       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     -&gt;红灯       - 已配置     -&gt;黄灯       - 成功执行     -&gt;绿灯       - 执行出错     -&gt;绿灯       - 执行出错     -&gt;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</td></td<>	Node     Workflow       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     ->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	KNIME分析平台
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。
	本 Eclipse Launcher Seicer a directory as workspace (villsers/vosy/knime_3.5.0/workspace) () Use this as the default and do not ask again () Necent Workspaces () Use this as the default and do not ask again () Recent Workspaces () CK Cancel         在启动KNIME Analytics Platform之后, for Karou U更改工作区,方法是在顶部菜单中选 Yarohr, 然后选择"切换工作区"。您可以以 拥有多个工作区,例如用于不同的项目或apa po         Swee Launcher () Swee Launcher (
	Update KNIME Exit

#### 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。 "工作流教练"根据KNIME用户统计数据和当前在"工作流编辑器"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。 "工作流编辑器"位于中心,允许创建和编辑工作流程。 "节点描述"显示描述节点任务和配置设置的文本,对于"工作流编辑器"或"节点仓库"中的选定节点。

"节点监视器"显示在"工作流编辑器"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、工作流程、组件和扩展。 "大纲"提供工作流程的概述 "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"工作流助手"面板将节点拖放到"工作流编辑器"中,可以创建KNIME工作流。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 3:配置节点

步骤 1: 通过拖放搜索并创建一个节点







步骤 4:执行节点



注释 1: 要创建一个新的空工作流,请右键单击 KNIME Explorer 面板,在新窗口中选择"创建新的 KNIME 工作流",并定义新工作流的名称和目标。

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工作流程。

### 显示数据表

Excel	KNIME分析平台
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。	<complex-block><complex-block></complex-block></complex-block>



## 打开Excel文件

### Excel读取器节点

Excel	KNIME分析平台
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 Reader节点读取Excel文件的单个工作表。       Excel Reader         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。       Excel Reader         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。         Dialog - 0:1 - Excel Reader       Image: Comparison of Comparison
A1 \$ Save As 0385 Save as Template 0385 Save as Template 0 E F G H I 1 Restore 2 Import 3 Reduce File Size 5 Share 6 Restrict Permissions 7 Passwords 9 Print Area 9 Print Area 9 Print Area 10 Properties 11 Properties 12 Obumloads 13 Center International Control Print Area 14 Control Drive 15 Documents 16 Documents 16 Documents 17 Properties 18 Documents 19 Desktop 10 Print Area 10 Print Are	Input location         Read from local File System          Mode       File         File       File in folder         File       /Users/kathrinmelcher/Desktop/Product Data2.xis         Sheet selection       Image: Select first sheet with data (Product Data.xis_defa)         Select first sheet with name       Product Data.xis_defa Image: Select sheet with name         Select sheet at index       0 0 0 (Sheet indexes start with 0.)         Column header       Image: Table contains row IDs in row numbers start with 1. See       列标题         Row ID       Generate row IDs       Table contains row IDs in column A       Sheet area         Read entire data of the sheet       Read only data in columns from A to and       and
Options Cancel Open	Image: Work and The Content and Detriving Content and De

## 打开CSV或txt文件

### CSV读取器节点

Excel	KNIME分析平台
<complex-block></complex-block>	CSV Reader 节点可以读取各种基于文本的文件,例如CSV     文字。 在配置窗口中,您可以设置所有必要的     达项,例如分隔符,以正确读取文件。
Desktop     Applications     CSV     knime-worksp     Documents	OK Apply Cancel
Cancel Get Data	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	KNIME分析平台
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。	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	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。
顺序。	Dialog - 0:2 - CSV Reader           Settings         Transformation         Advanced Settings         Limit Rows         Filter options
	Input location       File filter options         Read from       Local File System       資         Mode       File       Filter options       Include subfolders         File       Files in folder       Filter options       Include subfolders
	Folder       /Users/kathrinmelcher/Desktop/         ① Selected 26 of 26 files         Reader options         Format    Folder files Folder f
	Autodetect format     Folder name       , Column delimiter     \n Row delimiter       Quote char     Quote escape char
	Comment char     OK Cancel     J Has column header Has row ID
	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。
	● ● Dialog - 0:2 - CSV Reader 并集或交集 Settings Transformation Advanced Settings Limit Rows Encoding Flow Variables nory Policy
	Transformations Reset actions Move up Move down Enforce types Take columns from: Union Intersection
	тем паше туре

## 将内容从多个工作表导入到单个表中

Excel	KNIME分析平台
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
	注意2: <u>免费的KN</u> IME自学课程L2-DW KNIME Analytics Platform for Data Wran gles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	KNIME分析平台
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	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 Excel Writer ,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。其他设置选项使念 可以覆盖现有文件,并定义是否要将列标题/行ID写入Excel工作表的第 一行/列。
	Sations - Elow Variables
	File format & output location
	Excel format XLSX 3 输出位置
	Write to Local File System 😝
	File /Users/kathrinmelcher/Desktop/ExcelToKNIME.xlsx    Browse
	Write options     Create missing folders     If exists:     overwrite     append     fail
	Sheets $T h = 2 \pi$
	1. sheet name default_1
	If sheet exists Overwrite O append Tail
	Names and IDs
	Write row key
	☑ Don't write column headers if sheet exists
	Missing value handling
	Replace missing values by
	Formulas
	Layout
	Autosize columns
	O Portrait Landscape A4 - 210x297 mm
	Open file after execution
	OK Apply Cancel
	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

#### ExcelWriter节点

Excel	KNIME分析平台
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。	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。
Home Insert Draw Page Layout $\gg$ $\bigcirc$ Tell me $\bigcirc$ Share $\bigcirc$ Comments	
Clipboard       A       Tont       Alignment       %       Tont       Conditional Formatting ~       Tont       Tont       Idea         Clipboard       Font       Alignment       Number       Tont       Cells       Cells       Editing       Idea	Dialog - 3:99 - Excel Writer       Settings       File format & output location
A1 $f_x$ Glasgow A B C D E F G H I J K L 1 Glasgow United Kin 17-171-83: Alois Berger 23.9.1972 47 Alois Berge 0 2 Szczecin Poland 37-370-58 (Michaela Schultz 9.6.1998 21 Michaela.5 0 3 Sheffield United Kin 27-270-74 (Kotraut Grünwald 20.4.1975 44 Rotraut.Gr 0 4 Bochum-Ht Germany 64-647-95 Helga Heindi 18.10.200(19 Helga.Hein 0 5 Dortmund Germany 84-846-82 Mira Gleich 18.3.1997 22 Mira.Gleich 0 6 Valencia Spain 58-582-35 Joanna Radke 13.12.1991 24 Joanna.Rac 1 7 Valencia Spain 58-582-55 Hanspeter Storch 25.1.1998 21 Hanspeter 1 8 Lodz Poland 96-969-84 (Klaus-Peter Heinecke 13.4.1963 56 Klaus-Peter 0 9 Stuttgart Germany 48-484-59 Guenter Gasner 1.5.1988 31 Guenter.Ge 1 10 Denver United Stat 47-474-55 Giuseppina Nitsch 9.12.1975 44 Gliseppina 1 11 Kharkiv Ukraine 88-889-441 Sami Zimmerre 5.10.1982 37 Sami Zimm 1 12 Sofia Rulearia 13.13 (h8/thoma Wieczorek 30.3.1984 35 Ivana Wiec 0 Customers + 1 12 Sofia Rulearia 13.13 (h8/thoma Wieczorek 30.3.1984 35 Ivana Wiec 0 4 100	Excel format       XLSX ●         Write to       Local File System ●         File       /Users/kathrinmelcher/Desktop/ExcelToKNIME.xlsx       ● Browse         Write options       Create missing folders       If exists: overwrite ● append fail         Sheets       .       .         1. sheet name       default_1 ◆       ibin选项         If sheet exists       overwrite ● append ● fail          Names and IDs       工作表设置         Write column headers       ●       On't write column headers         ●       Don't write column headers if sheet exists          Missing value handling
	Replace missing values by   Formulas   Evaluate formulas (leave unchecked if uncertain; see node description for details)   Layout   Autosize columns   Portrait   Landscape   A4 - 210x297 mm     Open file after execution   OK   Apply   Cancel

#### 更新现有Excel工作表中的单元格

#### Excel单元格更新节点

Excel	KNIME分析平台
EXCEI     KNINE J J // 1 // 1 // 1       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。     Excel Cell Updater节点根据输入 输入表格需要一列包含应该更新 种数据类型的一列,其中包含新       『	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。
	Read from       Local File System         File       /Users/kathrinmelcher/Desktop/old-file.xlsx         Output file         Image: Create new file         Write to       Local File System         File       /Users/kathrinmelcher/Desktop/new-file.xlsx         File       /Users/kathrinmelcher/Desktop/new-file.xlsx         Write options       Create missing folders         If exists:       If exists:         Voldate       Image: Create missing folders         I. Excel sheet       Sheet1         Based on address column       SAddress         Missing value handling          Wateplace
	Formulas

#### Excel中的数据类型

#### KNIME中的数据类型

;1		KNIME分析平台		
]截图显示了Excel中可用的	9不同数据类 General 123 No specific format 123 Number General 123 Number General 123 Number Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Currency Cure	KNIME Analytics Platform 件时, KNIME Analytics I 在读取过程中更改此设置 Settings Transformations Reset actions Comm Comm Comm Comm Comm Comm Comm Com	m支持从Excel读取 Platform会尝试自动 了,可以使用Excel F Dialog - 0:1 - Excel F Transformation Advanced Settings 个 Move up	所有数据类型,甚至更多。在读 加检测正确的数据类型。如果您 Reader节点的"转换"选项卡。 Reader Flow Variables Memory Policy Flow Variables Memory Policy Flow Variables Memory Policy Inforce types Take columns from: • Un Type String Number (integer) String Number (double) Number (double) Number (double) Number (double) D Number (double) D Number (double) D Number (double) D Local Date C Local Time
「以映射到KNIME Analytics	s Platform	II V D Percent	n -	S String
「以映射到KNIME Analytics 【下数据类型 「 <b>Fxcel</b>	s Platform KNIME分析平台	ii v D Percent ii v B Fractior ii v S Fractior ii v I Scientifi ii v I Text ii v ? <any td="" u<=""><td>tage n fic unknown new column&gt;</td><td>Number (double) S String Number (integer) Number (integer)</td></any>	tage n fic unknown new column>	Number (double) S String Number (integer) Number (integer)
「以映射到KNIME Analytics 「下数据类型 <b>Excel</b> 営切	s Platform KNIME分析平台 数字或字符串	ii v D Percent ii v S Fractior ii v I Scientifi ii v I Scientifi ii v I Text ii v ? <any td="" u<=""><td>tage n fic unknown new column&gt;</td><td>S String Number (integer) Number (integer) Number (integer)</td></any>	tage n fic unknown new column>	S String Number (integer) Number (integer) Number (integer)
「以映射到KNIME Analytics 「下数据类型 <mark>Excel</mark> 常规 数字	s Platform KNIME分析平台 数字或字符串 数字	<ul> <li>I U Percent</li> <li>I S Fraction</li> <li>I Scientifi</li> <li>I Scientifi</li> <li>I Text</li> <li>I Text</li> <li>I ? &lt; any u</li> </ul>	tage n fic unknown new column>	Number (double)  S String  Number (integer)  Number (integer)
「以映射到KNIME Analytics 「下数据类型 <b>Excel</b> 常规 数字 货币	s Platform KNIME分析平台 数字或字符串 数字 数字 数字(整数或浮点数)	<ul> <li>I of D Percent</li> <li>I of D Percent</li></ul>	tage n fic unknown new column> Previe	Number (double) S String Number (integer) Number (integer)
「以映射到KNIME Analytics 「下数据类型 <b>Excel</b> 常规 数字 货币 (例如 50,25 €)	s Platform KNIME分析平台 数字或字符串 数字 数字 数字 数字 (例如 50,25)		tage n fic unknown new column> Previe rent settings d column types are based on the first	Number (double)     String     Number (integer)     Number (integer)     Number (integer)     P     File Content t 10000 rows only. See 'Advanced Settings
「以映射到KNIME Analytics 「下数据类型 <b>Excel</b> 常规 数字 货币 (例如 50,25 €) 会计	s Platform  KNIME分析平台  数字或字符串  数字  数字 (整数或浮点数)  (例如 50,25)  数字 (整数或浮点数)	<ul> <li>■ ✓ D Percent</li> <li>■ ✓ S Fraction</li> <li>■ ✓ I Scientifi</li> <li>■ ✓ I Text</li> <li>■ ✓ Text</li> <li>■ ✓ ? <any li="" u<=""> <li>■ ✓ The suggested</li> <li>■ ✓ D = U</li> </any></li></ul>	tage n n fic unknown new column> Previe rent settings d column types are based on the first	Winder (double) S String Number (integer) Number (integer) ? W File Content t 10000 rows only. See 'Advanced Settings
「以映射到KNIME Analytics 「大数据类型 <b>Excel</b> 常规 数字 货币 (例如 50,25 €) 会计 (例如 50,25 €)	s Platform  KNIME分析平台  数字或字符串 数字  数字 (整数或浮点数) (例如 50,25)  数字(整数或浮点数) (例如 50,25)		tage n n fic unknown new column> Previe Previe d column types are based on the first G Gener [] Number [D Halio 4 4.5]	Winder (double) S String Number (integer) Number (integer) P File Content t 10000 rows only. See 'Advanced Settings Curre D Accou [5] Short D [5] Long 5 4.5 2020-03 2020-0
(以映射到KNIME Analytics) 下数据类型          Excel         常规         数字         货币         (例如 50,25 €)         会计         (例如 50,25 €)         短日期	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期	C D Percent     C S Fraction     C S Fraction     C S Fraction     C S C S T S C S S T     C S S T S S S S S S S S S S S S S S	tage n n fic unknown new column> Previe rent settings d column types are based on the first 1 Gener   S  Gener    1 Number   D i Hallo 4 4.:	Winder (double) S String Number (integer) Number (integer) P V File Content t 10000 rows only. See 'Advanced Settings' Curre D Accou D Short D D Long 5 4.5 2020-03 2020-0
以映射到KNIME Analytics 下数据类型 Excel 常规 数字 货币 (例如 50,25 €) 会计 (例如 50,25 €) 短日期 长日期	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期 本地日期	Image: Constraint of the state of the s	tage n n fic unknown new column> Previe ent settings d column types are based on the first I Gener S Gener I Number D Hallo 4 4.;	Number (double) S String Number (integer) Number (integer) R Hele Content File Content t 10000 rows only. See 'Advanced Settings Curre D Accou ID Short D ID Long S 4.5 2020-03 2020-0
<ul> <li>以映射到KNIME Analytics</li> <li>下数据类型</li> <li>Excel</li> <li>常规</li> <li>数字</li> <li>货币</li> <li>(例如 50,25 €)</li> <li>会计</li> <li>(例如 50,25 €)</li> <li>短日期</li> <li>长日期</li> <li>时间</li> </ul>	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期 本地日期 本地日期 本地时间	■ C D Percent ■ C S Fraction ■ C Scientifi ■ C Text ■ C Text ■ C C <any u<br="">• • • • • • • • • • • • • • •</any>	tage n fic unknown new column> Previe rent settings d column types are based on the first GenerSGenerINumber b Hallo 4 4.3 OK	Number (double)         S String         Number (integer)         Number (integer)         Number (integer)         P         File Content         t 10000 rows only. See 'Advanced Settings         Curre       D Accou         D Short D       D Long         S       4.5       2020-03         Apply       Cancel
<ul> <li>以映射到KNIME Analytics</li> <li>下数据类型</li> <li><b>Excel</b></li> <li>常规</li> <li>数字</li> <li>货币</li> <li>(例如 50,25 €)</li> <li>会计</li> <li>(例如 50,25 €)</li> <li>短日期</li> <li>长日期</li> <li>时间</li> <li>百分比</li> </ul>	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期 本地日期 本地日期 本地时间 数字(浮点数)	Image: Constraint of the second state of the second sta	tage n n fic unknown new column>  Previe rent settings d column types are based on the first Gener Gener Hallo 4 . OK	Number (double)         S String         Number (integer)         Number (integer)         Number (integer)         Pile Content         t 10000 rows only. See 'Advanced Settings         Curre D Accou In Short D In Long         S         4.5         2020-03 2020-0
[以映射到KNIME Analytics] 下数据类型          Excel         常规         数字         货币         (例如 50,25 €)         会计         (例如 50,25 €)         短日期         长日期         时间         百分比         (例如 50%)	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期 本地日期 本地日期 本地日期 (例如 0,5)	・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	tage n fic unknown new column> Previe rent settings d column types are based on the first i Gener	Number (double)         S String         Number (integer)         Number (integer)         Number (integer)         P         *         File Content         t 10000 rows only. See 'Advanced Settings'         Curre       D Accou         S 4.5       2020-03         2020-03         Apply       Cancel
<ul> <li>以映射到KNIME Analytics</li> <li>下数据类型</li> <li>Excel</li> <li>常规</li> <li>数字</li> <li>货币</li> <li>(例如 50,25 €)</li> <li>会计</li> <li>(例如 50,25 €)</li> <li>短日期</li> <li>长日期</li> <li>时间</li> <li>百分比</li> <li>(例如 50%)</li> <li>分数</li> </ul>	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期 本地日期 本地日期 本地时间 数字(浮点数) (例如 0,5) 数字(浮点数)	注意1:转换选项卡也适, 注意2: 女林梅浩西上内	tage n fic unknown new column> rent settings d column types are based on the first d column types are based on the first I Gener S Gener I Number D Hallo 4 4.1 のK 用于其他读取节点。 ので、コンチークク	Number (integer) String Number (integer) Number (integer) P W File Content t 10000 rows only. See 'Advanced Settings Curre D Accou 页 Short D 页 Long 5 4.5 2020-03 2020-0 Apply Cancel ⑦
<ul> <li>以映射到KNIME Analytics</li> <li>下数据类型</li> <li>Excel</li> <li>常规</li> <li>数字</li> <li>货币</li> <li>(例如 50,25 €)</li> <li>会计</li> <li>(例如 50,25 €)</li> <li>短日期</li> <li>长日期</li> <li>时间</li> <li>百分比</li> <li>(例如 50%)</li> <li>分数</li> <li>科学计数法</li> </ul>	s Platform KNIME分析平台 数字或字符串 数字 数字(整数或浮点数) (例如 50,25) 数字(整数或浮点数) (例如 50,25) 本地日期 本地日期 本地日期 本地日期 数字(浮点数) (例如 0,5) 数字(浮点数)	注意1:转换选项卡也适, 注意2:在转换选项卡也适,	n fic unknown new column> rent settings d column types are based on the first I GenerSi GenerI Number D Hallo 4 4.1 OK 用于其他读取节点。 ,您还可以重命名、	Number (integer) Number (integer) Number (integer) Number (integer) File Content 10000 rows only. See 'Advanced Settings Curre D Accou I Short D ID Long 4.5 2020-03 2020-0 Apply Cancel ? ml除和更改列的顺序。

#### 连接到数据库

#### 数据库连接器节点



#### 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工作流编辑器中。 K NIME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。 当您开始共享工作流或将其导出到其他 KNIME Analytics Platform安装或KNIME服务器时,这变得很重要。KNIME Analytics P latform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、当前工作流还是当前工作流的数据
 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Am azon 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"…"在读取器节点图标的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

-Input locati	ion		
Read from	Amazon S3	$\sim$	
Mode	File Files in f	folder	
File			

## 追加/连接数据

追加数据

#### 连接节点



#### VLOOKUP

### 过滤和连接节点

Excel	KNIME分析平台
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。	1. 查找某个值,例如某个产品的价格。
第二个任务的替代函数是INDEX MATCH。	interest 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 这题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。 If 使 a log a for the following and of the following and the following



#### 过滤/删除具有特定值的行

#### 行过滤器



#### 过滤/删除具有不同值的行

#### 基于规则的行过滤器

Excel	KNIME分析平台		
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"。	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 <u>规则的行过滤器节点</u> 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 S列名\$ = "值 1" 或 S列名\$ = "值 2"=> 真		
从下拉荧单由洗择你感兴趣的值	Dialog - 6:23 - Rule-based Row Filter		
E       F       G       H       I         415       F**** Code       15       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I       I	Column List ROWDDX ROWCOUNT     Ory     Description       All     Checks whether the value of the left expression is like the wildcard pattern defined by the right expression For example: SColOS LIKE "H7ilor"     For example: SColOS LIKE "H7ilor"       Wight Mins Day Mins Ever Calls Day Charge Ever Calls Day Charge Eve		
	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



#### 按多个关键列对行进行排序

#### 排序节点

Excel			KNIME分析平台	
要按多个列对行进行排/ 然后点击"排序和筛选",	序,请选择要排序的 选择"自定义排序"。	列。	使用排序 <u>器节点</u> ,您可以按升序或降序对一个 序。	<sup>▶</sup> 或多个列进行排 Sorter
	Sort			
Add levels to	o sort by:	My list has headers	<b>Dialog</b> - 0:318	- Sorter
Orthu	Column Sort On	Order Color/Icon		
Sort by	Day Mins 🗘 Values	C Largest to Small	Sorting Filter Advanced Settings	How Variables Memory Policy
Then by	Eve Mins   C Values	Chargest to Small      C	Sort by	
			1 Age	Ascending
+ - Co	ру			Descending
		Options Cancel OK		
			Next by	
				Ascending
				Descending
			+ Add Ru	le
			OK Apply	Cancel
			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 <b>2:</b> 您可以临时对节点的输出表进行排序 序还是降序排序。	。  点击要排序的列标题,并选择是升

### 删除列

### 列过滤器节点

Excel	KNIME分析平台
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。	使用列过 <u>滤器节点,您</u> 可以删除多余的列。请记住,完 Column Filter 整的数据表仍将在前一个节点的输出端口和上游计算中 ————————————————————————————————————
	Oialog - 6:7 - Column Filter
	Column Filter Flow Variables Memory Policy
	Manual Selection Wildcard/Regex Selection Type Selection     Fxclude
	T Filter
	I CustServ Calls       I VMail Message         Day Calls       D ay Mins         Night Mins       Intl Mins         Intl Mins       Intl Mins         Int Verage       Intl Mins         Int Verage       Intl Mins
	Enforce exclusion     Enforce inclusion
	OK Apply Cancel
	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	KNIME分析平台
要重新排序列: • 选择要移动的列的列标题 • 按下ctrl + shift +下箭头	要重新排序列,您可以使用列重新排 <u>序节点。 通过选</u> 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。
<ul> <li>单击并按住要移动的列的绿色轮廓</li> <li>将列拖动到所需位置</li> </ul>	Dialog - 3:19 - Column Resorter
要重命名列,只需单击列单元格并更改单元格值。	要重命名列, 您可以使用列重命名节点 <u>。双击要重命名</u> 的列, 激活复选框"更 改"并在文本框中定义列标题。 Column Rename <u> </u>
	OK Apply Cancel

#### 更改数据类型

#### 字符串转数字和数字转字符串

#### KNIME分析平台 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。 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。 然后使用中间的下拉菜单。 ... Dialog - 3:19 - String To Number Flow Variables Job Manager Selection Memory Policy General \* Parsing options ABC General 123 No specific format Type: String To Number tal DN ng I Number (integer) Decimal separator: L Number (long) 123 Number S+2 Thousands separator: Currency Accept type suffix, e.g. 'd', 'D', 'f', 'F' J Accounting Include 0-Short Date Exclude · T Filter T Filter 0-Long Date S age S hours-per-week S workclass > S education S marital-status ( Time >> S occupation S relationship % Percentage < S race S sex 1/2 Fraction S native-country ~ S income 10<sup>2</sup> Scientific Always include all columns ABC Text More Number Formats... Dialog - 3:2 - Number To String . . Flow Variables Job Manager Selection Memory Policy Exclude Include T Filter **T** Filter Number To String I STATIONS\_ID I DATE > 1 QUALITY\_LEVEL OUALIT\_LEVEL AIR\_TEMP STEAM\_PRESSURE CLOUD\_COVER AIR\_PRESSURE\_STATION\_ALTITUDE RELATIVE\_HUMIDITY UNADIM COMPARING >> 2+5 < D WIND\_SPEED << D AIR\_TEMP\_MAX Always include all columns Apply Cancel OK 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、Integ er和Long。

#### 使用KNIME Analytics Platform进行数据操作的提示

## Table Manipulator



#### 表操<u>作节点允许您在一</u>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

		Column New name	Type
7	~	New Name	S String
	-	1 fnlwgt	I Number (integer)
	<b>Z</b>	S education	S String
= (	<ul> <li>Image: A start of the start of</li></ul>	I] education-num    更改列名称	I Number (integer)
=	<b>v</b>	S marital-status	S String
		S occupation	S String 百次数据米刑
		S relationship	S String
		S race	S String
		S sex	S String
		L capital-gain	Number (integer)
	2	I hours-per-week	I Number (integer)
•	~	S native-country	S String
8		S income	S String
	$\checkmark$	? <any column="" new="" unknown=""></any>	?
		0	
evie	w		
0.0-	ta a	analysis successfully completed.	

ſ



#### 数据透视表

#### 数据透视节点

Excel			KNIME分析平台
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"	Home Insert	通过三个选项卡配置数据透视节点: <ul> <li>"分组"定义分组列,即Excel中的"行"框架(最终行ID)</li> <li>"透视"定义透视列,即Excel中的"列"框架(最终列标题)</li> </ul> <li>"手动聚合"对应于"值"设置选项。从可用列列表中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。</li>
<text><text><text></text></text></text>		Create PivotTable  vou want to analyze. r range  alsobatiSAS1:SP\$3334  t data source  action No data fields have been retrieved.  action Cancel  Coc  PivotTable Fields  FielD NAME  FielD NAME  FielD NAME  Fields  Fields  VMail Message  Age Mins  VMail Message  Field NAME  VMail Message  VMa	<complex-block></complex-block>
			注2. 在KNIME中,忽必须主少选择一列作为组和枢轴。如果恣只想选择"行",可以使用GroupBy节点。

#### 没有列的数据透视表

## GroupBy节点

Excel	KNIME分析平台
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。	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。
	Dialog - 4:15 - GroupBy         Image: Settings Description Flow Variables Job M Image: Memory Policy         Image: Coups Manual Aggregation Pattern Based Aggregation Type Based Aggregation         Image: Variable column(s)         Image: Variable column(s) <t< td=""></t<>

### 解开

## 解开枢轴节点

Excel	KNIME分析平台
在Excel中创建一个解开枢轴表 <ul> <li>将数据集存储在一个表中。</li> <li>选择表中的任意单元格。</li> <li>点击"数据"选项卡,然后选择"从表/范围"选项。</li> </ul>	<ul> <li>通过一个选项卡配置Unpivoting节点:</li> <li>在上部分, "Value columns"定义了要unpivot的列, 也就是Excel中的列选择。</li> <li>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。</li> </ul>
Get From Table/ Recent Data ~ Text/CSV Web Get & Transform Data       From Table/ Recent Existing Connections Get & Transform Data       Refresh All ~ 日         • 这将打开"Power Query Editor"。通过按住Shift键选择要解开枢轴的列。	Value columns Manual Selection Wildcard/Regex Selection Type Selection Exclude Filter S Account S Account S Corporate
<ul> <li>点击Power Query Editor的"转换"选项卡,然后选择"解开枢轴 列"。</li> <li>■ ● マ マ Table 1 - Power Query Editor</li> <li>File Home Transform Add Column View</li> <li>□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □</li></ul>	Enforce exclusion      Skip rows containing missing cells
Group Use First Row By as Headers + Count Rows Table Any Column Text	Retained columns O Manual Selection O Wildcard/Regex Selection Type Selection Exclude Include T Filter
$\frac{1}{5}$ $\frac{1}{7}$ $\frac{1}$	S Marketing       S         S Operations       S         S Corporate       S         C Enforce exclusion       Enforce inclusion
<ul> <li>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。</li> </ul>	Options Enable hillting
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。	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

数学函数

#### 数学公式节点

Excel	KNIME分析平台
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 • <sup>绝对值</sup> • 条件求和 • 四舍五入函数 • ROUND • ROUNDUP • ROUNDDUP • ROUNDDOWN 如果你经常使用的数学公式缺失,请发送邮件至kathrin.melche r@knime.com。	Math Formula 甘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

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

#### 字符串操作节点

Excel	KNIME分析平台						
1. 函数CONCATENCATE(text1,text2,)将不同的文本组合成一个单元格。	字符 <u>串操作节点根据定义</u> 的表达式对类型为String的列进行操   String Manipulation 作。有很多不同的函数可用。						
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。							
	String Manipulation Flow Variables Job Manager Selection Memory Policy						
File Edit View Insert Format Tools Data Window H	Column List Category Description						
artoSave Undo Zoom #22 sert Repeat Fill #2 Review View	ROWID ROWINDEX         All         Image: Construction of the string of t						
ut Cut %X = = = %	ROWCOUNT     Function     Examples: join("a", "b", "c") = "abc"						
xmat Paste 3£V ≡ ≡ ≡ •≡	S     First Name     indexOf(str, tosearch, start)     joinfnull, ", "a") = "a"       S     Last Name     indexOf(str, tosearch, start, modifiers)       indexOf(bars(str, chars))     joinfnull, "," a") = "a"						
× √ Paste Special ^≋∨	indexOfChars(str, chars, modifiers) indexOfChars(str, chars, start)						
Fill E F Clear	indexOfChars(str, chars, start, modifiers) join(str)						
Select All 36A Delete	joinSep(sep, str)						
Sheet	joir()\$First Name\$, \$Last Name\$)						
Find ► Find ^①F Find Next %G	Flow Variable List						
Emoji & Symbols ^%Space							
在打开的对迁框中,你可以完议更基地的值和策略地为的值	Append Column: Full Name     Insert Missing As Null						
	Replace Column: Di Last Name 🗘 🧹 Syntax check on close						
	OK Apply Cancel						
Replace							
Find what:	<b>1</b> .来自Excel的CONCATENATE(text <b>1</b> ,text2,)						
Within: Sheet	join (col1, col2,) 或join (col1, "your own string",)						
Search: By Rows	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						
	0						
Replace with:	2.查找和替换						
Replace Replace All Close Find Next	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。

#### Standard written table

к4	• • ×	√ fx				
	А	В	С	D	E	F
1	Year	Quarter	Store - no CC	Store - with CC	OnlineStore	
2	2019	1	36862,74	66775,81	114196,84	
3	2019	2	38059,65	70483,79	113399,81	
4	2019	3	48149,06	76791,58	96116,79	
5	2019	4	47220,13	61563,41	105625,31	
6						
7						

#### Styled table

A1	¢ ×	√ fx	Capour Formulas Data	1011011 11011		
	A	В	С	D	E	F
1	Year	Quarter	Store - no CC	Store - with CC	OnlineStore	
2	2019	1	36862,74	66775,81	114196,84	
3	2019	2	38059,65	70483,79	113399,81	
4	2019	3	48149,06	76791,58	96116,79	
5	2019	4	47220,13	61563,41	105625,31	
6						
7						

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,<u>右侧是在添</u>加了XLS Fo rmatter节点的格式化信息后的样式化表格,例如标题的黄色背景。

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。



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。

atically set based on th ry input table to XLS C inpivoted layout to win	ne provided input ta	able) or long/unpivoted)
ry input table to XLS C inpivoted layout to wi	Control Table (wide	or long/unpivoted)
inpivoted layout to wi	1 1/1 0 0 1 1 1 1	
	de XLS Control Tabl	le
write column	header to first row	
n Strategy at Operatio	Il header columns	e'
radicting information?	fail	4
	♥ write column ptions result table (for easier ✓ add additiona n Strategy at Operatio radicting information?	<ul> <li>write column header to first row</li> <li>ptions</li> <li>result table (for easier post-processing a</li> <li>✓ add additional header columns</li> <li>In Strategy at Operation Type 'long to wid</li> <li>radicting information? fail</li> </ul>

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<u>每个单元格添</u>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 Ta ble Generator节点将表格转换为XLS控制表,并用<u>字母替换列名和数</u>字 <u>替换行ID。</u>

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 T able 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 <u>控制表生成器节点的组合</u>,激活unpivot选项和一个规则 <u>引擎节点</u>。

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。\_\_\_\_\_

												Control Ta	ble - 2:2261	1 - XLS Cont	rol Table Ger	nerator	
								F	ile Hilite	Navigation	View						
											Table "defau	lt" - Rows: !	6 Spec -	- Columns:	8 Propert	ies Flow \	/ariables
								[	Row ID	S Cell	S Column	S Colum	I Colum	. S Colum.	Row	S RowID	S Value
	0		0.0000						Row0	A1	A	00A	1	column1	1	Row0	Basic Report Example Mai 201
	Concate	nated table -	2.2226 - 0	oncatenate					Row1	B1	В	00B	2	column2	1	Row0	?
le Hilite	Navigation View								Row2	C1	С	00C	3	column3	1	Row0	?
	Table Block and Block	6 C.	I	Description	<b>F</b> 1	a la la casa			Row3	51	D	000	4	column4	1	Row0	3
	Table default - Rows: 8	spec – Co	iumns: 7	Properties	Flow var	lables		VIS Control	Row5	F1	F	OOF	6	column6	1	Row0	2
Row ID	S column1	S column2	S column	3 S column4	S column	5 S column	6 S column7	Table Concreter	Row6	G1	G	00G	7	column7	1	Row0	?
Dawe	Basis Basart Framela Mai 2010	2	2		2		2	lable Generator	Row7	A2	A	00A	1	column1	2	Empty 0	?
KOWO	Basic Report Example Mai 2019	-	-	-	-	-	-		Row8	B2	в	00B	2	column2	2	Empty 0	?
Empty 0	<i>r</i>	f	<u></u>	· · · ·	<u></u>	1	1		Row9	C2	С	00C	3	column3	2	Empty 0	7
Row0_dup	Calendar Week	Monday	luesday	Wednesday	Thursday	Friday	Saturday		Row10	52	D	000	4	column4	2	Empty 0	3
Row1	CW 18	7	?	2158.99	2486.56	2756.36	5486.35		Row12	E2	F	005	6	column6	2	Empty 0	2
Row2	CW 19	2196.47	2963.47	2348.57	2786.57	2873.56	3947.56		Row13	G2	G	00G	7	column7	2	Empty 0	?
Row3	CW 20	2964.87	2145.78	3120.86	3452.99	3245.58	4279.26		Row14	A3	A	00A	1	column1	3	Row0_dup	Calendar Week
Row4	CW 21	2115.78	2657.23	2765.36	2275.96	1458.65	4823.45	unpivot results	Row15	B3	В	00B	2	column2	3	Row0_dup	Monday
Row5	CW 22	2678.36	1268.46	2192.78	2571.45	3156.47	?	activated	Row16	C3	С	00C	3	column3	3	Row0_dup	Tuesday
-								adavated	Row17	D3	D	00D	4	column4	3	Row0_dup	Wednesday
									Row18	E3	E	OOE	5	column5	3	Row0_dup	Thursday
									Row20	63	c c	000	7	column7	3	Row0_dup	Saturday
									Row21	A4	A	00A	1	column1	4	Row1	CW 18
									Row22	B4	В	00B	2	column2	4	Row1	?
									Row23	C4	С	00C	3	column3	4	Row1	?
									Row24	D4	D	00D	4	column4	4	Row1	2158.99
									Row25	E4	E	00E	5	column5	4	Row1	2486.56
									Row26	F4	F	OOF	6	column6	4	Row1	2756.36

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 <u>节点为每个单元格创建一行</u>,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。 <u>例如,</u>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。

	Rule Editor Flow Variables Job Manager Selection Memory Policy				Classified	values - 2:226	i0 - Rule En	gine		
Column List	Category Description	File Hilite	Navigation	View						
ROWID			Table "d	lefault" – Row	vs: 56 S	pec - Columr	ns: 8 Pro	perties	Flow Variable	s
ROWINDEX ROWCOUNT S Cell		Row ID	S Cell	S Column	S Colum	Colum	S Colum.	. I Row	S RowID	S Value
	Function	Row0	A1	A	00A	1	column1	1	Row0	header
Column	7<7	Row1	B1	В	008	2	column2	1	Row0	header
Column (comparable)	1 <= 1 2 = 2	Row2	C1	c	000	3	column3	1	Row0	header
Column (number)	7 > 7	Row3	D1	D	00D	4	column4	1	Row0	header
Row	? >= ?	Row4	E1	E	OOE	5	column5	1	Row0	header
RowID	? AND ?	Row5	F1	F	OOF	6	column6	1	Row0	header
JValue	7107	Row6	G1	G	00G	7	column7	1	Row0	header
	/ UKE / JANE / J	Row7	A2	A	00A	1	column1	2	Empty 0	empty
1	7 OR 7	Row8	B2	в	008	2	column2	2	Empty 0	empty
	7 XOR 7	Row9	C2	с	00C	3	column3	2	Empty 0	empty
	FALSE	Row10	D2	D	00D	4	column4	2	Empty 0	empty
0	MISSING ?	Row11	E2	E	OOE	5	column5	2	Empty 0	empty
w Variable List		Row12	F2	F	OOF	6	column6	2	Empty 0	empty
nime.workspace	Expression	Row13	G2	G	00G	7	column7	2	Empty 0	empty
	? 2// \$double column name\$ > 5.0 => "large"	Row14	A3	A	00A	1	column1	3	Row0_dup	c_header
	? 3// \$string column name\$ Like "#blue#" => "small and blue"	Row15	B3	В	00B	2	column2	3	Row0_dup	c_header
	\$ 5 \$Row\$= 1 => "header"	Row16	C3	С	00C	3	column3	3	Row0_dup	c_header
	S 6 \$Row\$= 2 => "enpty"	Row17	D3	D	00D	4	column4	3	Row0_dup	c_header
	S / SKowS=3 ⇒ **C_header, border* S / SKowS=4 AMD SColumn (number)S=1 ⇒ "cw. border"	Row18	E3	E	00E	5	column5	3	Row0_dup	c_heade
	S 9 \$Row\$>=4 => "border"	Row19	F3	F	00F	6	column6	3	Row0_dup	c_heade
		Row20	G3	G	00G	7	column7	3	Row0_dup	c_heade
~		Row21	A4	A	00A	1	column1	4	Row1	cw, bord
A	peraction S	Row22	B4	В	00B	2	column2	4	Row1	border
<b>A D</b>	nlace Column: Si Value	Row23	C4	c	00C	3	column3	4	Row1	border
U Ki		Row24	D4	D	00D	4	column4	4	Row1	border
		Row25	E4	E	00E	5	column5	4	Row1	border
		Row26	F4	F	OOF	6	column6	4	Row1	border

#### 图7左侧是*Rule Engine*节点的配置对话框。在表达式<u>部分,您可以</u>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XL<u>S控制表生成器节点可以将此</u>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<u>围节点。</u>

0

#### 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 <u>作流程。</u>

正如您所看到的,示例工作流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<u>背景颜色节点可以更改单元</u>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<u>您的EXCEL文件有多个工作表,XLS工作</u>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<u>边框格式化节点可以为</u>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 <u>,以最低的详细级别合</u>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<u>格式化程序(应用)节</u>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/communit y/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<u>,或者在KNIME Analytics Platform内查看每个单独节点的节</u> 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KN IME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.knim<u>e.com/knimepress下可获得。</u>

#### 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。