



Central European Institute of Technology
BRNO | CZECH REPUBLIC

2019-11-20

Processing of proteomics data in **KNIME**



Workshop outline

- morning session – theoretical part
 - 10:00 – 10:15 Opening and introduction
 - 10:15 – 10:45 Software container running KNIME
 - 10:45 – 11:00 Coffee break
 - **11:10 – 11:30 Introduction to KNIME**
 - 11:30 – 11:45 Coffee break
 - 11:45 – 13:00 Practical applications, our KNIME metanodes
 - 13:00 – 14:00 Lunch break, visit of our laboratories for interested people

Introduction to KNIME

Applications KNIME Analytics Platfo... [Terminal]



15:41 knimeuser



Trash



File System



Home



Knime



KNIME application

Workflows
templates r...Metanodes
templates r...

File Edit View Help



Quick Access

KNIME Explorer



- EXAMPLES (knime-guest@http://p...)
- LOCAL (Local Workspace)
 - 3.5.3c
 - _Metanodes templates

Workflow Coach



Recommended Nodes

File Reader

Node Repository



IO

Manipulation

Column

Binning

Auto-Binner

Auto-Binner (Apply)

Numeric Binner

Binner (Dictionary)

CAIM Binner

CAIM Applier

Convert & Replace

Filter

KNIME Analytics Platform

Welcome to KNIME Analytics Platform



Welcome to KNIME Analytics Platform!



New to KNIME? Looking for resources to get started?

- Register for emails with introductory tips [here](#).
- Explore our [Quickstart Guide](#).
- Check out [7 things to do after installing KNIME Analytics Platform](#)
- Find more hints and how-tos in the [Learning Hub](#).
- And register for our release and event emails right [here](#).

This page will be displayed upon startup but you can customize the content using the checkboxes at the bottom.

Updates for the following components are available:

- KNIME Public Server Access
- KNIME ServerSpace
- OpenMS

Outline

Console

An outline is not available.

KNIME Console

```
*****
***      Welcome to KNIME Analytics Platform v3.7.2.v201904170949 ***
***      Copyright by KNIME AG, Zurich, Switzerland
*****
Log file is located at: /home/knimeuser/knime-workspace/.metadata/knime/knime.log
```

This node allows to group numeric data in intervals - called bins. There are two naming options for the bins and two methods which define the number and the range of values that fall in a bin. Please use the "Numeric Binner" node if you want to define custom bins.

Dialog Options

Column Selection:

Columns in the include list are processed separately. The columns in the exclude list are omitted by the node.

Binning Method:

Use **Fixed number of bins** for bins with equal **width** over the domain range or bins that have an equal **frequency** of element

Applications

KNIME Analytics Platfor...

09:51 knimeuse

File Edit View Node Help



75%



Quick Access

KNIME Explorer

-

- Workflow_templates
 - MQ_PGs_LFQ_general_0.6a
 - graph1_files
 - graph_files
 - hostman_files

Workflow Coach

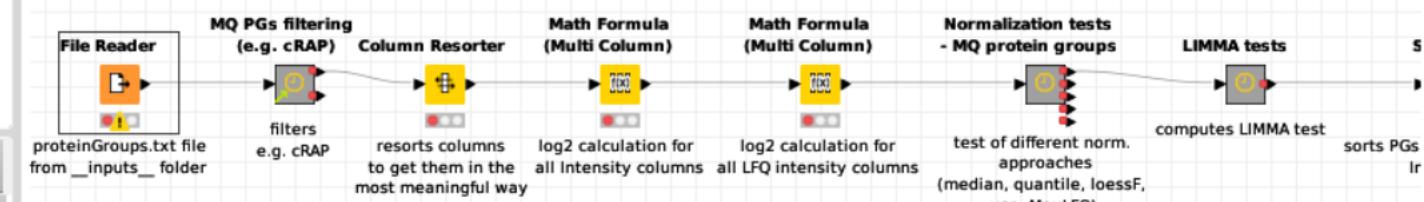
Recommended Nodes

Node Repository

ungroup

- Ungroup - /Manipulation/Row/
- GroupBy - /Manipulation/Row/
- Group Loop Start - /Workflow Control
- Database GroupBy - /Database
- Round Double - /Manipulation/
- MRMTransitionGroupPicker - /
- Grouped ScatterPlot - /Scriptlet
- Independent groups t-test - /AI
- Unpivoting - /Manipulation/Row/

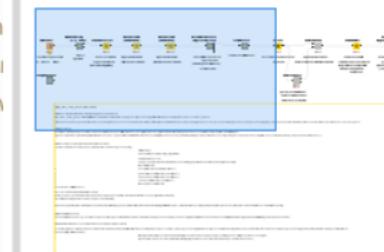
*0: MQ_PGs_LFQ_general_0.6a



Example dataset

```
# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
```

Outline



KNIME Analytics Platform

Node Description

File Reader

This node can be used to read data from an ASCII file or URL location. It can be configured to read various formats.

When you open the node's configuration dialog and provide a filename, it tries to guess the reader's settings by analyzing the content of the file. Check the results of these settings in the preview table. If the data shown is not correct or an error is reported, you can adjust the settings manually (see below).

The file analysis runs in the background and can be cut short by clicking the "Quick scan", which shows if the analysis takes longer. In this case the file is not analyzed completely, but only the first fifty lines are taken into account. It could happen then, that the preview appears looking fine, but the execution of the File Reader fails, when it reads the lines it didn't analyze. Thus it is recommended you check the settings, when you cut an analysis short.

Console

KNIME Console

```
WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unknown
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc
```



330M of 408M



KNIME Analytics Platform

File Edit View Node Help

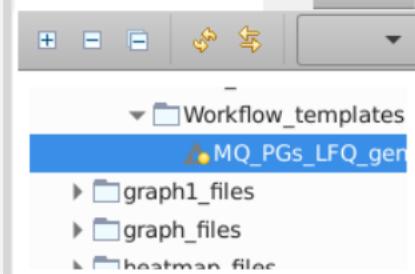


75%



Quick Access

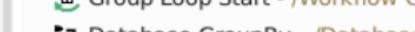
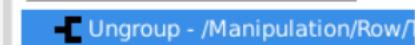
KNIME Explorer



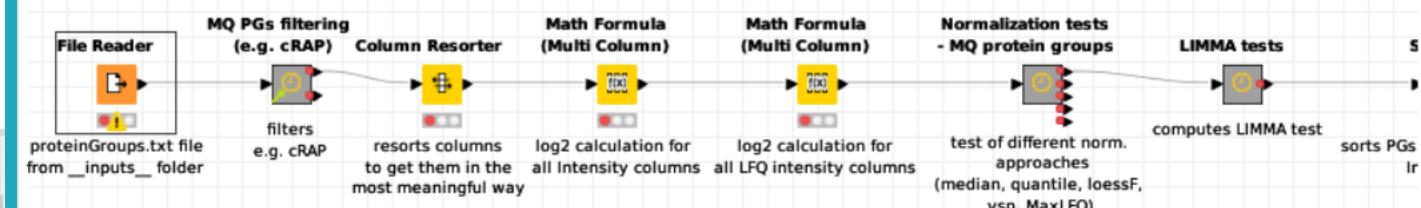
Workflow Coach



Recommended Nodes



*0: MQ_PGs_LFQ_general_0.6a

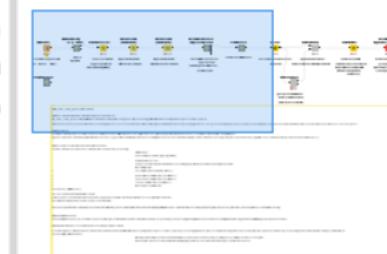


Example dataset



```
# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
```

Outline



Console

KNIME Console

```
WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unkn
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc
```

Workflow editor

File Reader

This node can be used to read data from an ASCII file or URL location. It can be configured to read various formats.

When you open the node's configuration dialog and provide a filename, it tries to guess the reader's settings by analyzing the content of the file. Check the results of these settings in the preview table. If the data shown is not correct or an error is reported, you can adjust the settings manually (see below).

The file analysis runs in the background and can be cut short by clicking the "Quick scan", which shows if the analysis takes longer. In this case the file is not analyzed completely, but only the first fifty lines are taken into account. It could happen then, that the preview appears looking fine, but the execution of the File Reader fails, when it reads the lines it didn't analyze. Thus it is recommended you check the settings, when you cut an analysis short.



75%

.6a x

KNIME Explorer

Workflow_templates

MQ_PGs_LFQ_gen

graph1_files

graph_files

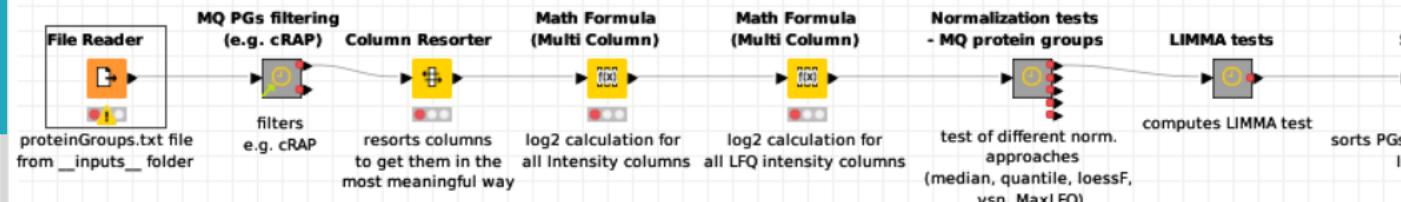
hostman_files

Workflow Coach

Recommended Nodes

Node Repository

ungroup

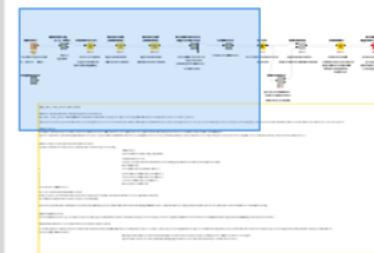


Example dataset



```
# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
```

Outline



Console

KNIME Console

```
WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unkn
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc
```



330M of 408M



Applications

KNIME Analytics Platfor...

09:51 knimeuse

File Edit View Node Help



Quick Access

KNIME Explorer



- Workflow_templates
- MQ_PGs_LFQ_general_0.6a
- graph1_files
- graph_files
- hostman_files

Workflow Coach



Recommended Nodes



Ungroup - /Manipulation/Row/

GroupBy - /Manipulation/Row/

Group Loop Start - /Workflow C

Database GroupBy - /Database

Round Double - /Manipulation/

MRMTransitionGroupPicker - /

Grouped ScatterPlot - /Scriptin

Independent groups t-test - /Al

Unpivoting - /Manipulation/Ro



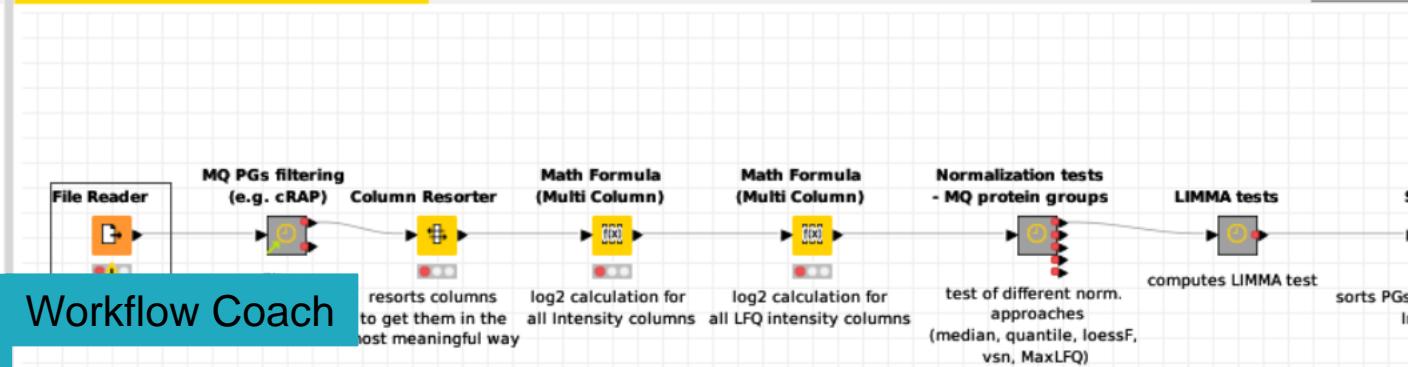
Outline



Console

```
KNIME Console
WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unkn
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc
```

*0: MQ_PGs_LFQ_general_0.6a



Workflow Coach

Example dataset



```
# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
```

This node can be used to read data from an ASCII file or URL location. It can be configured to read various formats.

When you open the node's configuration dialog and provide a filename, it tries to guess the reader's settings by analyzing the content of the file. Check the results of these settings in the preview table. If the data shown is not correct or an error is reported, you can adjust the settings manually (see below).

The file analysis runs in the background and can be cut short by clicking the "Quick scan", which shows if the analysis takes longer. In this case the file is not analyzed completely, but only the first fifty lines are taken into account. It could happen then, that the preview appears looking fine, but the execution of the File Reader fails, when it reads the lines it didn't analyze. Thus it is recommended you check the settings, when you cut an analysis short.

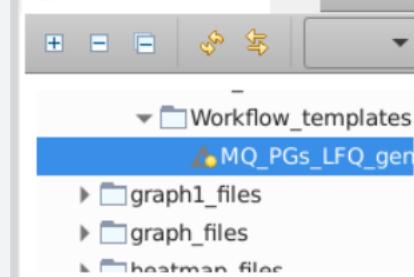
File Edit View Node Help



75%



KNIME Explorer

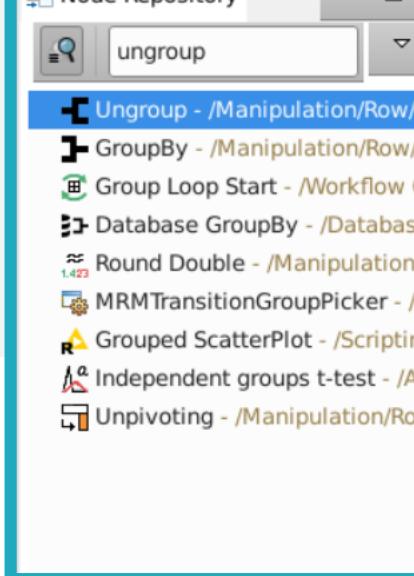


Workflow Coach

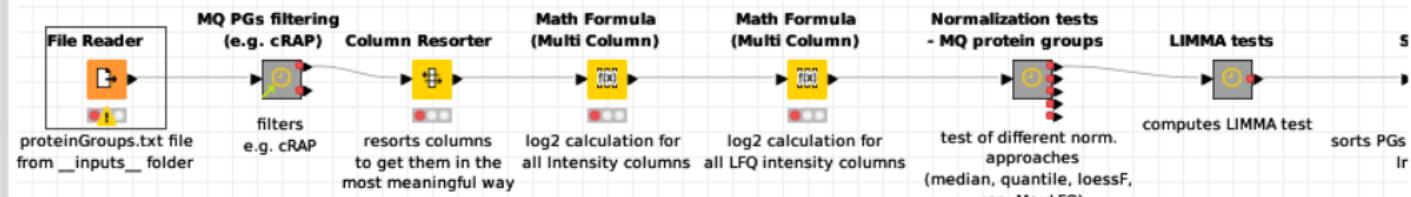


Recommended Nodes

Node Repository



*0: MQ_PGs_LFQ_general_0.6a



Example dataset

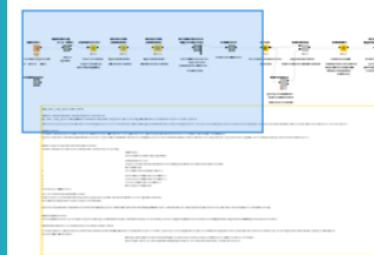
Node repository

```

# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data

```

Outline



Console

KNIME Console

```

WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unknown
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc

```

Applications

KNIME Analytics Platfor...

09:51 knimeuse

File Edit View Node Help



75%



KNIME Explorer



- Workflow_templates
 - MQ_PGs_LFQ_general_0.6a
 - graph1_files
 - graph_files
 - hostman_files

Workflow Coach



Recommended Nodes

Node Repository



Ungroup - /Manipulation/Row/

GroupBy - /Manipulation/Row/

Group Loop Start - /Workflow C

Database GroupBy - /Database

Round Double - /Manipulation/

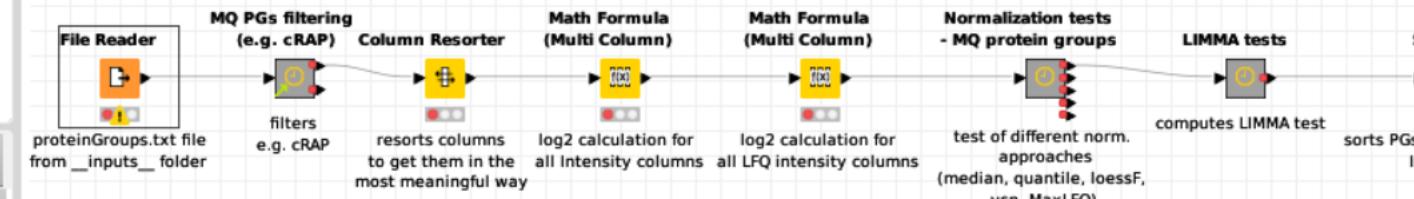
MRMTransitionGroupPicker - /

Grouped ScatterPlot - /Scriptin

Independent groups t-test - /Al

Unpivoting - /Manipulation/Ro

*0: MQ_PGs_LFQ_general_0.6a



Example dataset

```
# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
```

I understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr

ojects in the workflow are not necessarily the ones I used in my study, so please do not be confused if they do not make sense.

The workflow consists of several steps: reading the proteinGroups.txt file, filtering protein groups using cRAP, resorting columns, calculating log2 values for intensity and LFQ intensity, normalizing using various methods, and finally performing LIMMA tests to identify differentially expressed proteins.

The workflow is designed to handle large datasets efficiently, with a preview table showing the first 50 lines of the input file. It also includes a configuration dialog for each node, allowing users to fine-tune parameters according to their specific needs.

The workflow is intended to be a starting point for mass spectrometry data analysis and can be modified or extended as required. I hope it will be useful for other researchers in the field.

If you have any questions or need further assistance, please feel free to contact me. I am happy to help!

Best regards,

John Doe

Mass Spectrometry Research Group

Institute for Biochemistry and Molecular Biology

University of Regensburg

Germany

John.Doe@uni-regensburg.de

Phone: +49 941 944-1234

Fax: +49 941 944-1235

Mobile: +49 151 1234-5678

E-mail: John.Doe@uni-regensburg.de

Office: Room 101, Building A, University of Regensburg

Address: Universitätsstraße 31, 93040 Regensburg, Germany

Postal Address: D-93040 Regensburg, Germany

ZIP Code: 93040

City: Regensburg

Country: Germany

Latitude: 48.2081

Longitude: 11.3415

Altitude: 390 m

Timezone: Europe/Berlin

Daylight Saving Time: Yes

Summer Time: Yes

Winter Time: Yes

Spring Time: Yes

Autumn Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Spring Daylight Saving Time: Yes

Autumn Daylight Saving Time: Yes

Summer Daylight Saving Time: Yes

Winter Daylight Saving Time: Yes

Applications

KNIME Analytics Platfor...

09:51 knimeuse

File Edit View Node Help



75%



Quick Access

KNIME Explorer

-
- Workflow_templates
 - MQ_PGs_LFQ_general_0.6a
 - graph1_files
 - graph_files
 - hostman_files

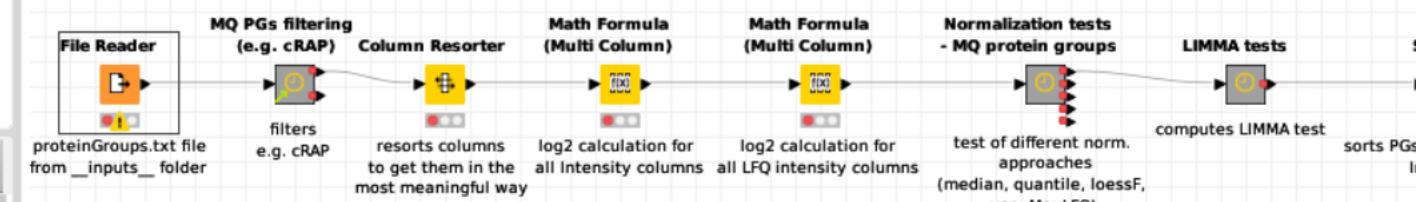
Workflow Coach

Recommended Nodes

Node Repository

- ungroup
- Ungroup - /Manipulation/Row/
- GroupBy - /Manipulation/Row/
- Group Loop Start - /Workflow C
- Database GroupBy - /Database
- Round Double - /Manipulation/
- MRMTransitionGroupPicker - /
- Grouped ScatterPlot - /Scriptin
- Independent groups t-test - /A
- Unpivoting - /Manipulation/Ro

*0: MQ_PGs_LFQ_general_0.6a

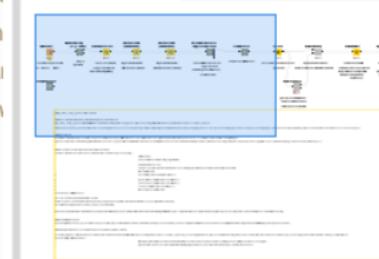


Example dataset



```
# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
```

Outline



Console

Console

KNIME Console

```
WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unknown
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc
```



330M of 408M



Applications

KNIME Analytics Platfor...

09:51 knimeuse

File Edit View Node Help



75%

KNIME Analytics Platform

Quick Access

KNIME Explorer

-
- Workflow_templates
 - MQ_PGs_LFQ_general_0.6a
 - graph1_files
 - graph_files
 - hostman_files

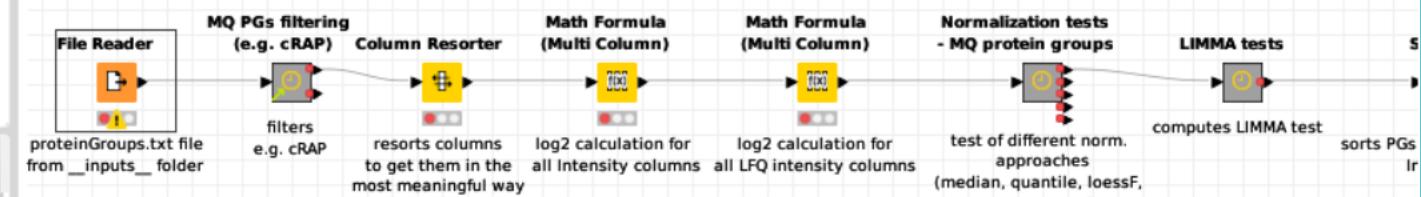
Workflow Coach

Recommended Nodes

Node Repository

- Ungroup - /Manipulation/Row/**
- GroupBy - /Manipulation/Row/**
- Group Loop Start - /Workflow Control**
- Database GroupBy - /Database**
- Round Double - /Manipulation/**
- MRMTransitionGroupPicker - /**
- Grouped ScatterPlot - /Scriptlet**
- Independent groups t-test - /AI**
- Unpivoting - /Manipulation/**

*0: MQ_PGs_LFQ_general_0.6a



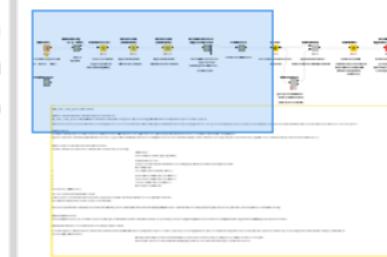
Example dataset



```

# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
  
```

Outline



Console

KNIME Console

```

WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unkn
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unkn
WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc
  
```

Node description

Node Description

File Reader

This node can be used to read data from an ASCII file or URL location. It can be configured to read various formats.

When you open the node's configuration dialog and provide a filename, it tries to guess the reader's settings by analyzing the content of the file. Check the results of these settings in the preview table. If the data shown is not correct or an error is reported, you can adjust the settings manually (see below).

The file analysis runs in the background and can be cut short by clicking the "Quick scan", which shows if the analysis takes longer. In this case the file is not analyzed completely, but only the first fifty lines are taken into account. It could happen then, that the preview appears looking fine, but the execution of the File Reader fails, when it reads the lines it didn't analyze. Thus it is recommended you check the settings, when you cut an analysis short.



330M of 408M



Applications

KNIME Analytics Platfor...

09:51 knimeuse

File Edit View Node Help



75%



Quick Access

KNIME Explorer

Workflow_templates
 MQ_PGs_LFQ_general
 graph1_files
 graph_files
 hostman_files

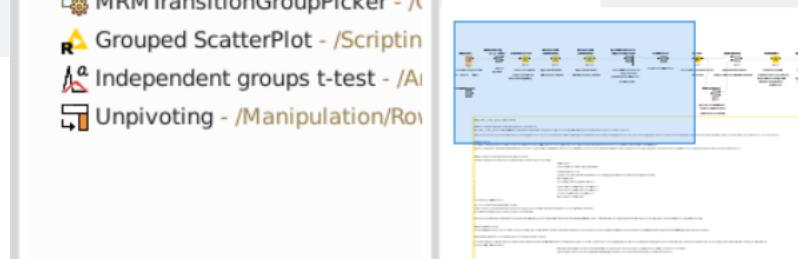
Workflow Coach

Recommended Nodes

Node Repository

ungroup
 Ungroup - /Manipulation/Row/
 GroupBy - /Manipulation/Row/
 Group Loop Start - /Workflow C
 Database GroupBy - /Database/
 Round Double - /Manipulation/
 MRMTransitionGroupPicker - /
 Grouped ScatterPlot - /Scriptin
 Independent groups t-test - /A
 Unpivoting - /Manipulation/Ro

Outline



Example dataset



```

# MQ_PGs_LFQ_general workflow
## Description of the workflow and general remarks
MQ_PGs_LFQ_general is a KNIME workflow developed for the general processing of label-free bottom-up mass spectrometry data.
Please note, that you should understand e.g. data structure and experimental design used within the study to apply the correct processing approach! The pr
## Input data
  
```

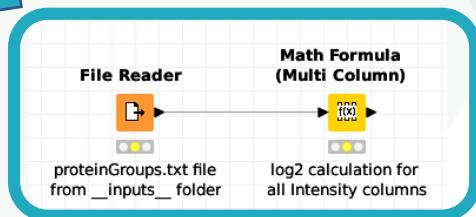
Console

KNIME Console

WARN Column Rename (Regex) 0:844:0:0:850 Pattern replace resulted in duplicate column names; resolved c
 WARN CASE Switch Data (Start) 0:844:0:0:73 Errors overwriting node settings with flow variables: Unknown
 WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
 WARN Column Rename (Regex) 0:844:239:0:850 Pattern replace resulted in duplicate column names; resolved c
 WARN CASE Switch Data (Start) 0:844:239:0:73 Errors overwriting node settings with flow variables: Unknown
 WARN Column Rename (Regex) 0:844:300:277:0:428 Pattern did not match any column name, leaving input unc
 WARN Column Rename (Regex) 0:844:300:290:0:428 Pattern did not match any column name, leaving input unc

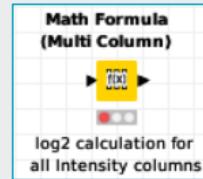
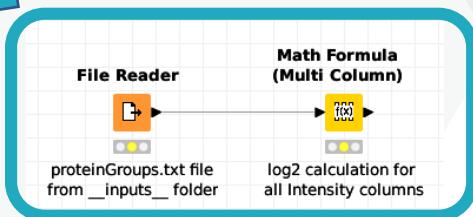
KNIME environment and metanodes

Nodes

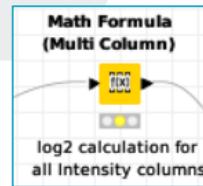


KNIME environment and metanodes

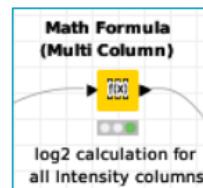
Nodes



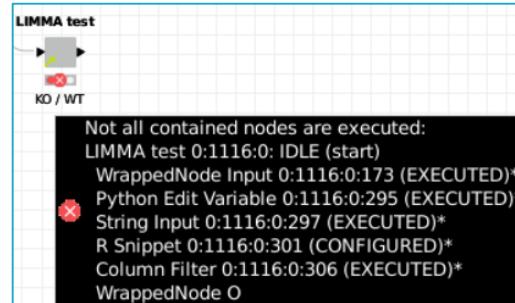
Not connected to other nodes
Not executed



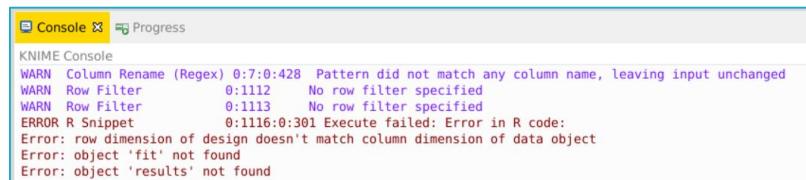
Connected to other nodes
Not executed



Connected to other nodes
Successfully executed

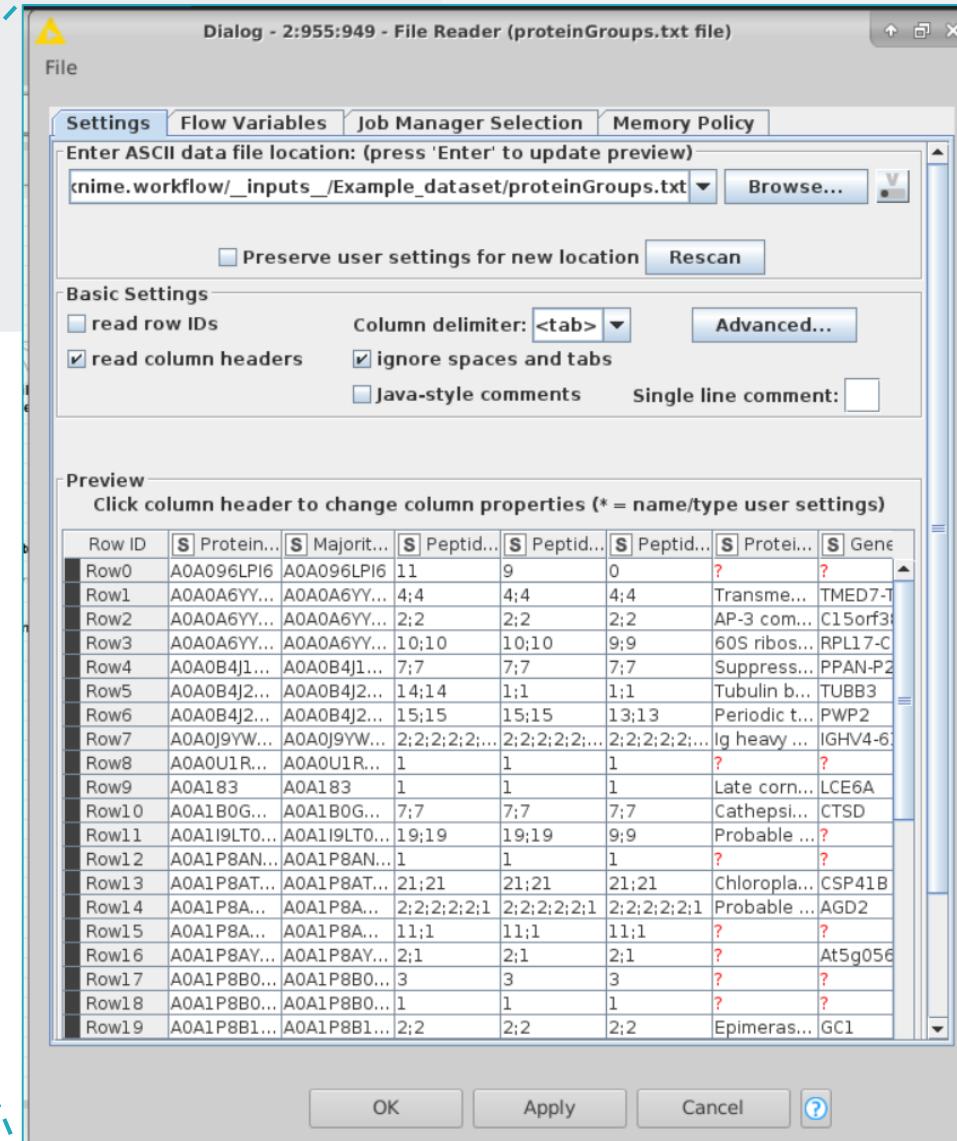
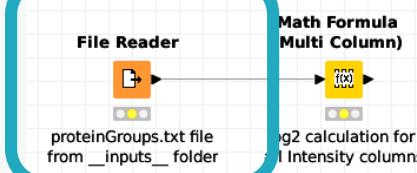


Connected to other nodes
Error while node execution

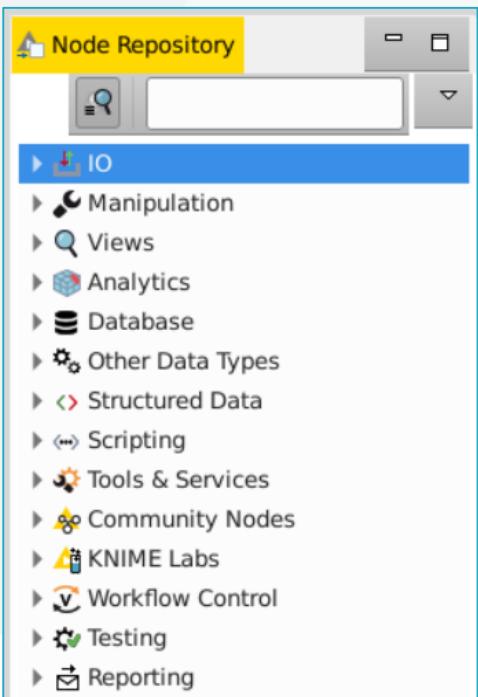


KNIME environment and metanodes

Nodes



In-built KNIME metanodes



IO

- File Reader
- Excel Reader
- CSV Reader
- Table Reader

- CSV Writer
- Table Writer
- Excel Sheet Appender

- Table Creator

Manipulation

- Column/Row Filter
- Column Rename
- Math Formula
- String Manipulation
- Column Splitter
- Column Appender
- Joiner
- Concatenate
- Column Resorter
- Missing Value
- Row Filter/Splitter
- GroupBy
- Pivoting
- Rule Engine
- Transpose
- Sorter
- Shuffle

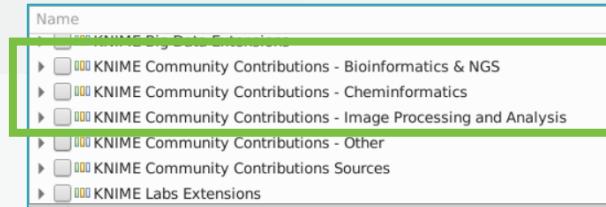
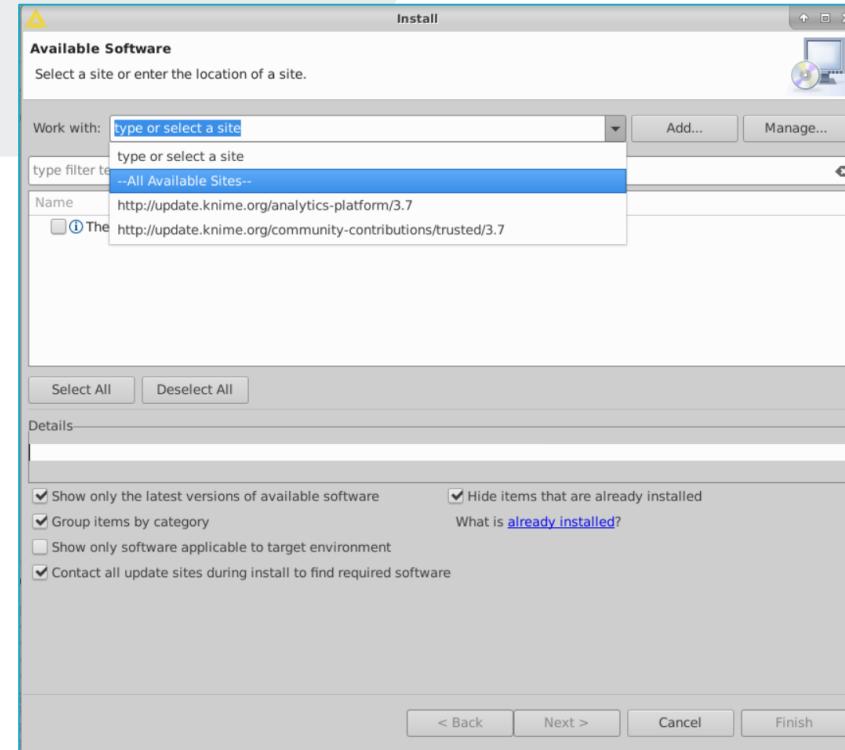
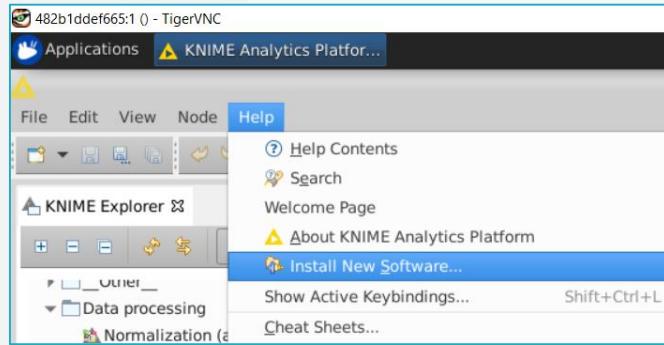
Views

- Bar Chart
- Box Plot
- Scatter Plot
- ROC Curve
- Histogram
- Line Plot

Scripting

- Java
- Python
- R

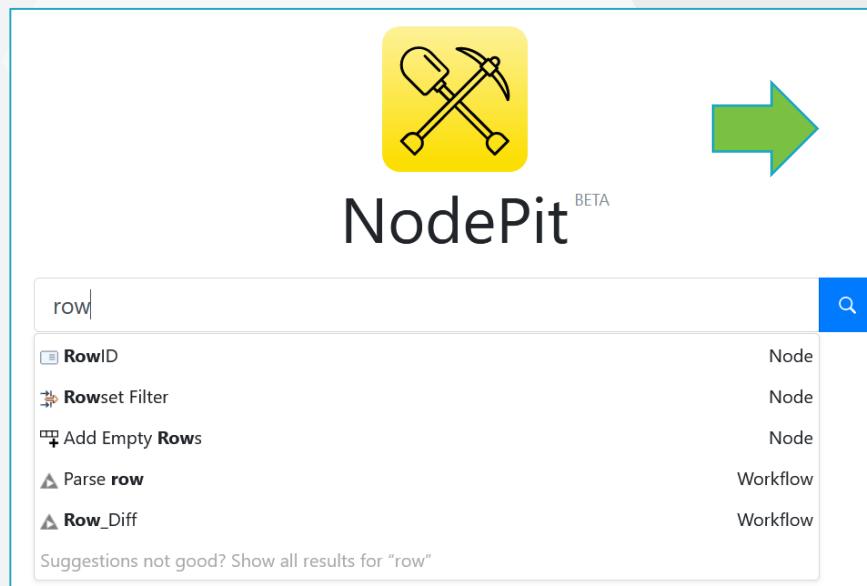
Installation of additional plugins / software



OpenMS
PIA

NodePit

Where to seek help or inspiration



The screenshot shows the NodePit documentation page for the "Row Filter" node. At the top, there are navigation links: Nodes (636), Workflows (763), Products (0), Installable Units (2), and Users (2). The main content area shows the "Row Filter" node with the following details:

Row Filter Streamable
Nodes > Manipulation > Row > Filter
Allows filtering of data rows by certain criteria, such as row ID, attribute value, and row number range.

Row Splitter Streamable
Nodes > Manipulation > Row > Filter
Allows splitting of the input table by certain criteria, such as row ID, attribute value, and row number range.

Key-Collection HiLite Translator
Nodes > Manipulation > Row
Translates hilite events from a row containing a collection cell with row keys to the original rows.

Duplicate Row Filter
Nodes > Manipulation > Row > Filter
Removes or marks duplicate rows.

RowID Streamable
Nodes > Manipulation > Row > Other
Node to replace the RowID and/or to create a column with the values of the current RowID.

Spark Row Filter
Nodes > Tools & Services > Apache Spark > Row
The Spark Row Filter allows rows to be excluded from the input Spark DataFrame/RDD.

NodePit

RowID

Streamable KNIME Base Nodes version 4.0.2.v201909251340 by KNIME AG, Zurich, Switzerland

This node can be used to replace the RowID of the input data with the values of another column (by converting the values to a string) or a generated row id of the format: Row0, Row1, Row2, ... The user has additional options to ensure uniqueness and to handle missing values.

It can also be used to create a new column, which contains the RowID as a value.

If both options are selected the node appends a new column with the values of the current RowID to the table and replaces the current RowID with the values of the selected column or the generated row id.

Note: Highlighting does not work across this node if the "Enable hiliting" option is disabled.

Options

Replace RowID with selected column values: If selected the node replaces each RowID with the corresponding string value of the selected column or with a generated key.

New RowID column: The column to replace the current RowID. If none is selected a row id is generated with the format: Row0, Row1, Row2, ...

Ensure uniqueness: If checked the node ensures uniqueness by appending a counter (x) to duplicates where x is incremented for each appearance.

Handle missing values: All missing values are replaced with '?' if this check box is ticked. We recommend activating the 'Ensure uniqueness' check box to handle any duplicate missing values.

Enable hiliting: If enabled, a map is maintained joining the old with the new RowID. Depending of the number of rows, enabling this feature might consume a lot of memory.

Create new column with the RowID values: If this is selected the node creates a new column with the value of the current RowID.

New column name: The name of the new column that is being added to the table.

Input Ports

► The data table to be processed.

Output Ports

► Data with the replaced RowID and/or a new column with the current RowID values.

Best Friends (Incoming)

- Transpose (7 %)
- Row Filter (6 %) Streamable
- Column Filter (6 %) Streamable
- Joiner (5 %)
- GroupBy (4 %)

[Show all 444 recommendations](#)

Best Friends (Outgoing)

- Joiner (14 %)
- Column Filter (8 %) Streamable
- Transpose (6 %)
- Row Filter (4 %) Streamable
- String Manipulation (3 %) Streamable

[Show all 590 recommendations](#)

Workflows

- GASIC
- Factiva_parser
- PhD_Articles
- wrong-column-name-13973
- Amazon Reviews Preprocessing

[Show all 348 workflows](#)

Information about the node, options, following nodes provided

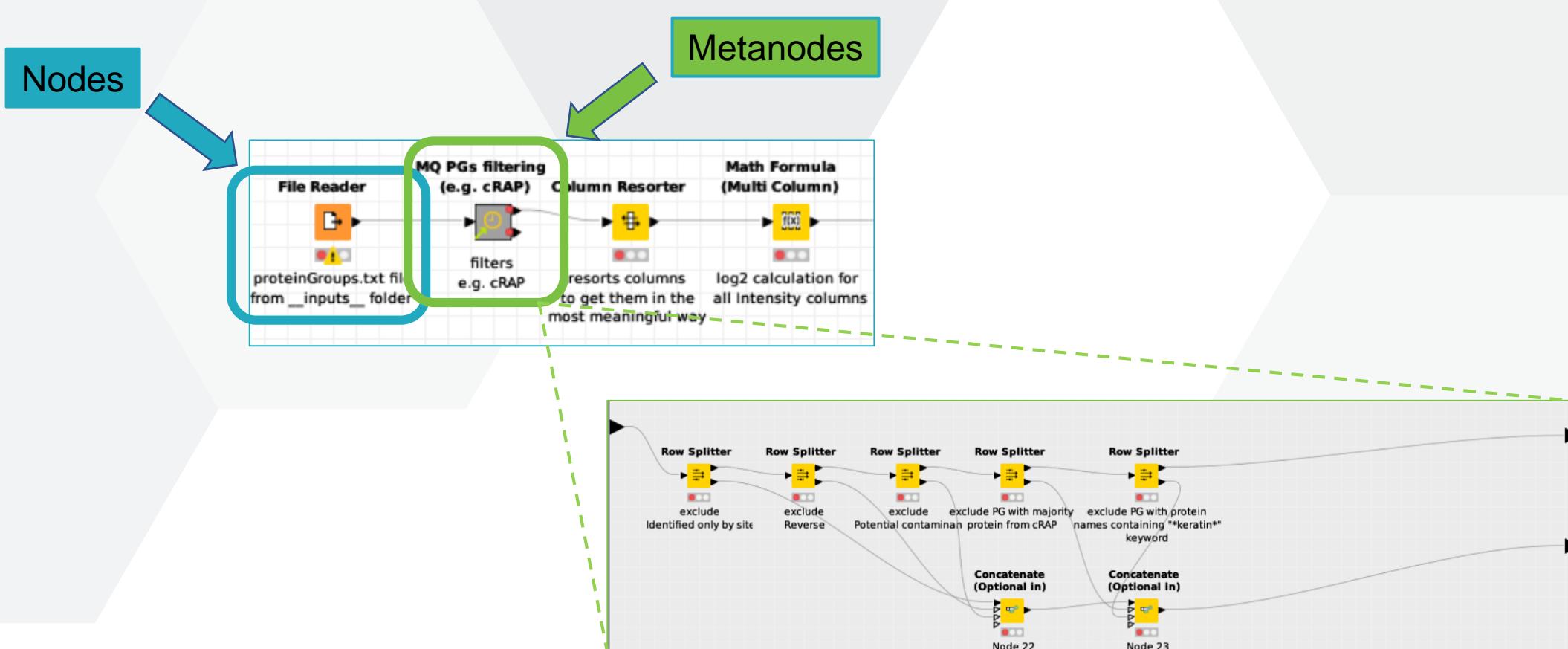
KNIME community and help

The screenshot shows the KNIME forum homepage. At the top, there is a navigation bar with links to Hub, Blog, Forum, Events, Career, Contact, Download, Log In, and a search icon. Below the navigation bar, there is a breadcrumb menu: SOFTWARE / SOLUTIONS / LEARNING / PARTNERS / COMMUNITY / ABOUT. The main content area has tabs for all categories, Categories (which is selected), Latest, and Top. On the left, there is a sidebar with categories: KNIME Analytics Platform, KNIME Extensions, Community Extensions, Partner Extensions, KNIME Server, Special Interest Groups, and KNIME Development. Each category has a brief description and a list of sub-topics. On the right, there is a "Latest" section displaying ten recent posts, each with a user icon, title, and timestamp.

Category	Topics
KNIME Analytics Platform	61 / week
For discussions related to KNIME Analytics Platform	
KNIME Extensions	8 / week
For discussions related to KNIME Extensions and Integrations	
Text Processing Scripting Reporting Image Processing REST Big Data Deep Learning	
Community Extensions	1 / week
For discussions related to extensions developed by the KNIME community	
RDKit Scripting Extensions Indigo Erlwood HCS Tools Palladian & Selenium CDK OpenMS Vernalis Seqan	
Partner Extensions	87
For discussions related to extensions developed by our KNIME partners	
JChem Extensions Schrödinger Extensions	
KNIME Server	4 / week
For discussions related to KNIME Server	
Special Interest Groups	1 / week
For discussions related to various special interest groups	
Cheminformatics BioInformatics	
KNIME Development	2 / week
For discussions related to KNIME development	

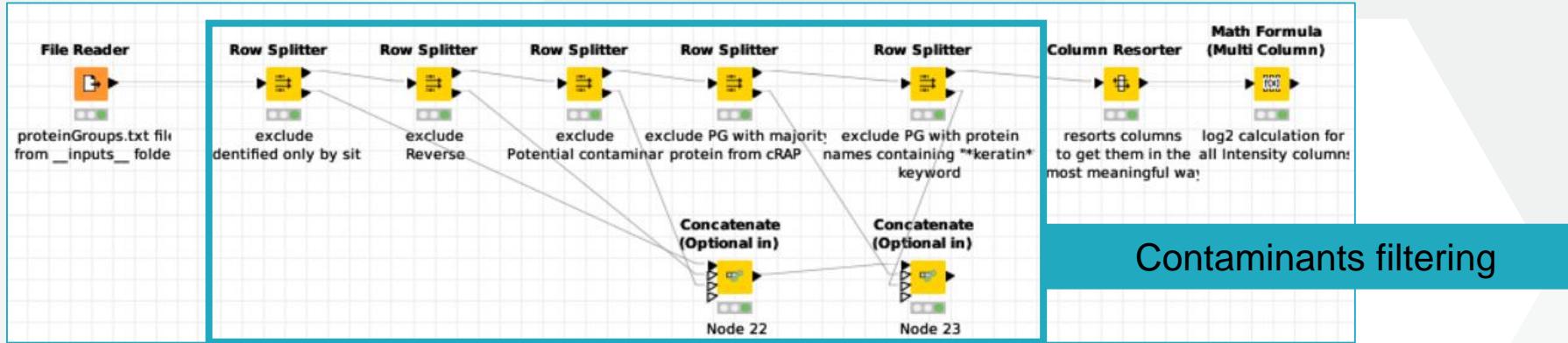
Topic	Responses	Age
Don't save start / end nodes for variable connection	1	26m
KNIME Analytics Platform		
DB Writer writing only 25 rows per second to MySQL database	5	36m
KNIME Analytics Platform		
workflow not executed on scheduled time..skiped few workflow and jumped to next..	37	1h
KNIME Server		
How to read all files in a folder with a loop	6	3h
KNIME Analytics Platform		
Obtain a distance vector from a Data Frame	2	5h
KNIME Analytics Platform		
Parsing and Pivoting JSON	4	5h
KNIME Development		
How to handle merged cells importing from Excel	2	6h
KNIME Analytics Platform		
Get Request Node	10	11h
KNIME Analytics Platform		

KNIME environment and metanodes



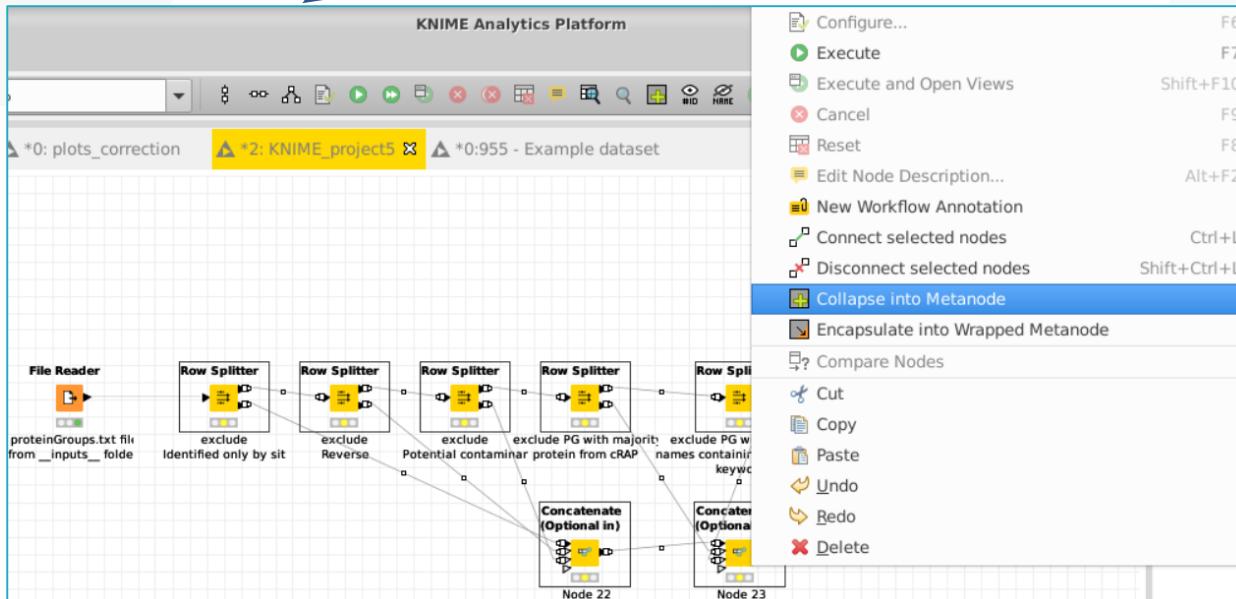
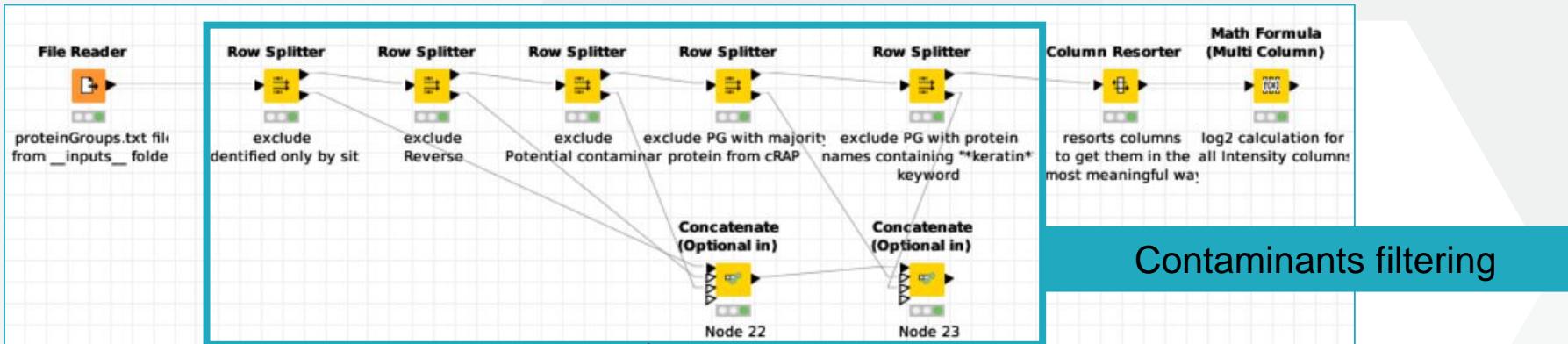
Another layer of data processing
Making workflow more structured and organized

KNIME environment and metanodes



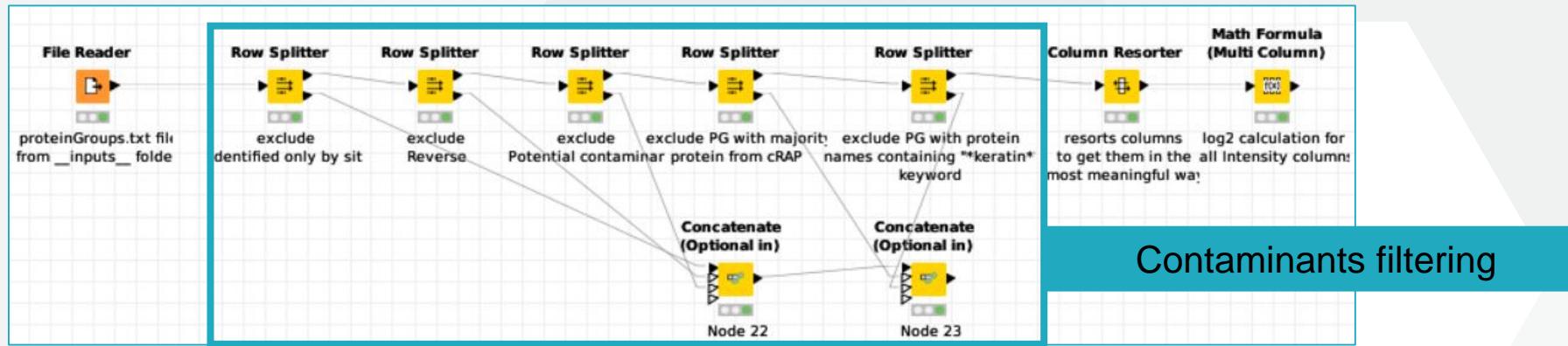
Collapsing nodes into Metanode
Structuring the workflow, making it easily readable

KNIME environment and metanodes

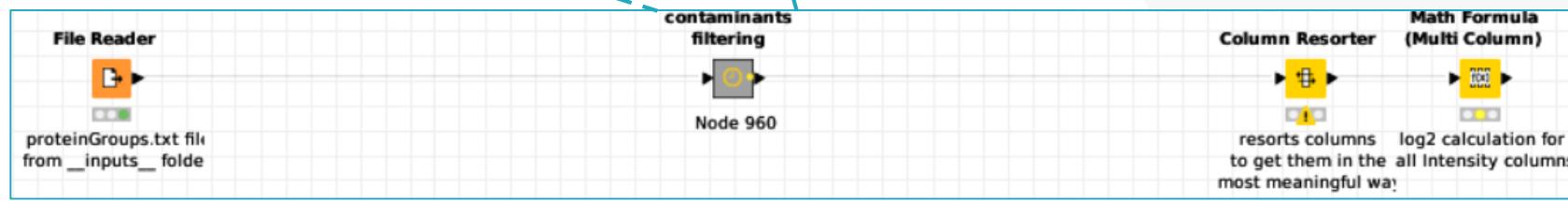


Collapsing nodes into Metanode
Structuring the workflow, making it
easily readable

KNIME environment and metanodes



Collapsing nodes into Metanode
Structuring the workflow, making it
easily readable



Different port types

tables holding data
variables used when scripting

Data
Flowvariable
PMML
Database Connection
Database Query
Distance Measure
DocumentVectorPortObject
FileStorePrefixURIPort
FileStoreReferenceURIPort
FileStoreURIPort
FilterDefinition
Gradient Boosting Model
Image

images

KnimeConnection
Moz API Connection
OpenNlpNerTaggerModelPortObject
Outlier
Python

whole R workspace

R Workspace
Regression Tree
Regression Tree
Remote Connection
StanfordNERModelPortObject
Tree Ensembles
URI Object

file located on the hard drive

URIPortObject managed by FileStore
VectorHashingPortObject
de.mpicbg.knime.scripting.r.port.RPortType2

Table To Image

Node 02

URI Port to Variable

Node 03

Table to R

Node 03

File Table - Wide table format

File Hilite Navigation View

Table "proteinGroups.txt" - Rows: 3855 Spec - Columns: 81 Properties Flow Variables

Columns: characteristics of particular e.g. protein groups

Rows: individual protein groups

Row ID	Protein IDs	Majority protein IDs	Peptide counts (all)	Peptide counts (Razor...)	Peptide counts (unique)	Protein names	Gene names	Fasta headers	Number of proteins	Peptides	Razor + unique peptides	Unique peptides	Peptides WT_1	Peptides WT_2	
Row0	A0A096LPI6	A0A096LPI6	11		0	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=2	1	11	9	0	7	6	
Row1	A0A0A6YY...	A0A0A6YY...	4;4		4;4	Transme...	TMED7-TI...	Protein TMED7-TICAM2 OS=Homo sapiens GN=TMED7-TICAM2 PE=...	2	4	4	4	3	1	
Row2	A0A0A6YY...	A0A0A6YY...	2;2		2;2	AP-3 com...	C15orf38...	Protein C15orf38-AP3S2 OS=Homo sapiens GN=C15orf38-AP3S2 ...	2	2	2	2	1	0	
Row3	A0A...														
Row4	A0A...														
Row5	A0A0B4J2...	A0A0B4J2...	14;14		1;1	Tubulin b...	TUBB3	Uncharacterized protein OS=Homo sapiens PE=1 SV=1;Tubulin b...	2	14	1	1	11	11	
Row6	A0A0B4J2...	A0A0B4J2...	15;15		13;13	Periodic t...	PWF...					15	13	6	4
Row7	A0A0J9YW...	A0A0J9YW...	2;2;2;2;2;...		2;2;...;2;2;2;...	Ig heavy ...	IGHV					2	2	1	1
Row8	A0A0U1R...	A0A0U1R...	1		1	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=1	1	1	1	1	0	0	
Row9	A0A1B0G...	A0A1B0G...	7;7		7;7	Cathepsi...	CTSD	Uncharacterized protein OS=Homo sapiens PE=1 SV=1;Cathepsi...	2	7	7	7	1	2	
Row10	A0A1W2P...	A0A1W2P...	14		0	NA	NA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1	1	14	2	0	11	11	
Row11	A0A1W2P...	A0A1W2P...	9;2		2;2	NA	NA	RPS10-NUDT3 readthrough OS=Homo sapiens GN=RPS10-NUDT3 ...	2	9	2	2	7	8	
Row12	A0A1W2P...	A0A1W2P...	3;2;1		3;2;1	Ester hyd...	C11orf54	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Ester hyd...	3	3	3	3	1	1	
Row13	A0AVT1	A0AVT1	29		29	Ubiquitin-...	UBA6	Ubiquitin-like modifier-activating enzyme 6 OS=Homo sapiens GN=...	1	29	29	29	9	8	
Row14	A0FGR8	A0FGR8	1		1	Extended...	ESYT2	Extended synaptotagmin-2 OS=Homo sapiens GN=ESYT2 PE=1 S...	1	1	1	1	0	0	
Row15	A1L0T0	A1L0T0	6		6	Acetolact...	ILVBL	Acetolactate synthase-like protein OS=Homo sapiens GN=ILVBL P...	1	6	6	6	1	0	
Row16	A3KN83	A3KN83	2		2	Protein s...	SBN01	Protein strawberry notch homolog 1 OS=Homo sapiens GN=SBNO...	1	2	2	2	0	0	
Row17	A4D1E9	A4D1E9	4		4	Protein A	GTPBP10	Protein A OS=Homo sapiens GN=GTPBP10 PE=1 SV=1	1	4	4	4	0	1	
Row18	A5YKK6	A5YKK6	20		20	CCR4-NO...	CNOT1	CCR4-NOT transcription complex subunit 1 OS=Homo sapiens GN...	1	20	20	20	6	6	
Row19	A6NDG6	A6NDG6	13		13	Phospho...	PGP	Glycerol-3-phosphate phosphatase OS=Homo sapiens GN=PGP P...	1	13	13	13	7	7	
Row20	A6NFQ2;A...	A6NFQ2;A...	2;1		2;1	TRPM8 ch...	TCAF2	TRPM8 channel-associated factor 2 OS=Homo sapiens GN=TCAF2...	2	2	2	2	1	1	
Row21	A6NHQ2	A6NHQ2	4		1	rRNA/tRN...	FBLL1	rRNA/tRNA 2'-methyltransferase fibrillarin-like protein 1 OS=Hom...	1	4	1	1	3	4	
Row22	A6NHR9	A6NHR9	14		14	Structura...	SMCHD1	Structural maintenance of chromosomes flexible hinge domain-co...	1	14	14	14	2	3	
Row23	A6NJ78	A6NJ78	1		1	Probable ...	METTL15	Probable methyltransferase-like protein 15 OS=Homo sapiens GN...	1	1	1	1	0	0	
Row24	A6NKT7;Q...	A6NKT7;Q...	18;16		1;1	RanBP2-li...	RGPD3;R...	RanBP2-like and GRIP domain-containing protein 3 OS=Homo sapi...	2	18	1	1	5	8	
Row25	A8CG34;Q...	A8CG34;Q...	4;3		4;3	Nuclear e...	POM121C...	Nuclear envelope pore membrane protein POM 121C OS=Homo s...	2	4	4	4	0	0	
Row26	A8MXV4;R...	A8MXV4	3;1;1		3;1;1	Nucleosi...	NUDT19	Nucleoside diphosphate-linked moiety X motif 19 OS=Homo sapie...	3	3	3	3	1	0	
Row27	A9UHW6	A9UHW6	2		2	MIF4G do...	MIF4GD	MIF4G domain-containing protein OS=Homo sapiens GN=MIF4GD ...	1	2	2	2	0	0	
Row28	B01T2	B01T2	39		38	Unconven...	MYO1G	Unconventional myosin-Ig OS=Homo sapiens GN=MYO1G PE=1 SV...	1	39	39	38	25	30	
Row29	B4DLN1;Q...	B4DLN1;Q...	14;7		7;7	Mitochon...	SLC25A10	cDNA FLJ60124, highly similar to Mitochondrial dicarboxylate carrie...	2	14	14	7	10	10	
Row30	P0CG08;B...	P0CG08;B...	1;1		1;1	Golgi pH ...	GPR89B;...	Golgi pH regulator B OS=Homo sapiens GN=GPR89B PE=1 SV=1;...	2	1	1	1	0	0	
Row31	C9JAW5;Q...	C9JAW5;Q...	1;1		1;1	HIG1 dom...	HIGD1A	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;HIG1 dom...	2	1	1	1	0	0	
Row32	E7ENX8;P...	E7ENX8	11;2;1;1		1;1	2;2;1;1	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=3 SV=1	4	11	11	2	7	6	
Row33	E7EVH7;Q...	E7EVH7;Q...	11;11;3		1;1			Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Kinesin lig...	3	11	11	11	4	4	
Row34	E9PAV3;Q...	E9PAV3	5;1		4;3;1	NEDD8	NEDD8-MDP1	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Type peptide-associ...	2	5	5	5	3	5	
Row35	E9PL57;Q...	E9PL57;Q...	4;3;1		1;1	Uncharac...	C11orf98	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Uncharac...	2	4	4	4	2	3	
Row36	E9PLD3;E...	E9PLD3;E...	1;1		1;1	Vitamin K...	VKORC1	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=...	2	1	1	1	1	1	
Row37	E9PLN8;Q...	E9PLN8;Q...	1;1		1;1	2 oxipo...	PCKDHA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;2 oxipo...	2	1	1	1	1	1	
Row38	EEUEP2;P...	EEUEP2;P...	2;2		2;2	2 oxipo...	PCKDHA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;2 oxipo...	2	1	1	1	1	1	

File Table - Wide table format

File Hilite Navigation View

Table "proteinGroups.txt" - Rows: 3855 Spec - Columns: 81 Properties Flow Variables

Columns: characteristics of particular e.g. protein groups

Data processing -> “column-wise”
Adding new columns, not new rows (in comparison to Excel)

Rows: individual protein groups

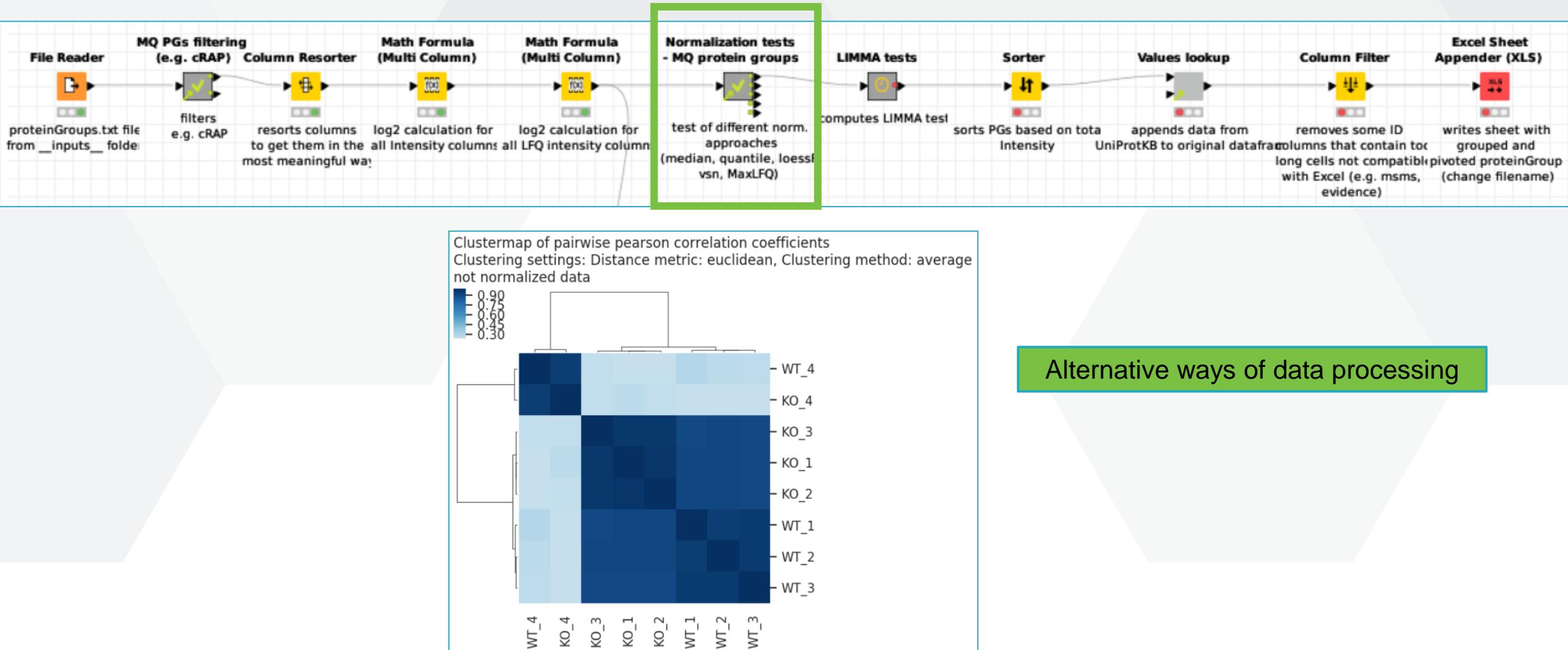
Row ID	Protein IDs	Majority protein IDs	Peptide counts (all)	Peptide counts (Razor...)	Peptide counts (unique)	Protein names	Gene names	Fasta headers	Number of proteins	Peptides	Razor + unique peptides	Unique peptides	Peptides WT_1	Peptides WT_2
Row0	A0A096LPI6	A0A096LPI6	11		0	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=2	1	11	9	0	7	6
Row1	A0A0A6YY...	A0A0A6YY...	4;4		4;4	Transme...	TMED7-TI...	Protein TMED7-TICAM2 OS=Homo sapiens GN=TMED7-TICAM2 PE=...	2	4	4	4	3	1
Row2	A0A0A6YY...	A0A0A6YY...	2;2		2;2	AP-3 com...	C15orf38...	Protein C15orf38-AP3S2 OS=Homo sapiens GN=C15orf38-AP3S2 ...	2	2	2	2	1	0
Row3	A0A...													
Row4	A0A...													
Row5	A0A0B4J2...	A0A0B4J2...	14;14		1;1	Tubulin b...	TUBB3	Uncharacterized protein OS=Homo sapiens PE=1 SV=1;Tubulin b...	2	14	1	1	11	11
Row6	A0A0B4J2...	A0A0B4J2...	15;15		13;13	Periodic t...	PWF...							
Row7	A0A0J9YW...	A0A0J9YW...	2;2;2;2;2;...		2;2;... 2;2;2;2;2;...	Ig heavy ...	IGHV							
Row8	A0A0U1R...	A0A0U1R...	1		1	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=1	1	1	1	1	0	0
Row9	A0A1B0G...	A0A1B0G...	7;7		7;7	Cathepsi...	CTSD	Uncharacterized protein OS=Homo sapiens PE=1 SV=1;Cathepsi...	2	7	7	7	1	2
Row10	A0A1W2P...	A0A1W2P...	14		0	NA	NA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1	1	14	2	0	11	11
Row11	A0A1W2P...	A0A1W2P...	9;2		2;2	NA	NA	RPS10-NUDT3 readthrough OS=Homo sapiens GN=RPS10-NUDT3 ...	2	9	2	2	7	8
Row12	A0A1W2P...	A0A1W2P...	3;2;1		3;2;1	Ester hyd...	C11orf54	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Ester hyd...	3	3	3	3	1	1
Row13	A0AVT1	A0AVT1	29		29	Ubiquitin-...	UBA6	Ubiquitin-like modifier-activating enzyme 6 OS=Homo sapiens GN=...	1	29	29	29	9	8
Row14	A0FGR8	A0FGR8	1		1	Extended...	ESYT2	Extended synaptotagmin-2 OS=Homo sapiens GN=ESYT2 PE=1 S...	1	1	1	1	0	0
Row15	A1L0T0	A1L0T0	6		6	Acetolact...	ILVBL	Acetolactate synthase-like protein OS=Homo sapiens GN=ILVBL ...	1	6	6	6	1	0
Row16	A3KN83	A3KN83	2		2	Protein s...	SBN01	Protein strawberry notch homolog						
Row17	A4D1E9	A4D1E9	4		4	Protein A	GTPBP10	Protein A OS=Homo sapiens GN=G						
Row18	A5YKK6	A5YKK6	20		20	CCR4-NO...	CNOT1	CCR4-NOT transcription complex su						
Row19	A6NDG6	A6NDG6	13		13	Phospho...	PGP	Glycerol-3-phosphate phosphatase OS=Homo sapiens GN=PGP P...	1	13	13	13	7	7
Row20	A6NFQ2;A...	A6NFQ2;A...	2;1		2;1	TRPM8 ch...	TCAF2	TRPM8 channel-associated factor 2 OS=Homo sapiens GN=TCAF2...	2	2	2	2	1	1
Row21	A6NHQ2	A6NHQ2	4		1	rRNA/tRN...	FBLL1	rRNA/tRNA 2'-methyltransferase fibrillarin-like protein 1 OS=Hom...	1	4	1	1	3	4
Row22	A6NHR9	A6NHR9	14		14	Structura...	SMCHD1	Structural maintenance of chromosomes flexible hinge domain-co...	1	14	14	14	2	3
Row23	A6NJ78	A6NJ78	1		1	Probable ...	METTL15	Probable methyltransferase-like protein 15 OS=Homo sapiens GN...	1	1	1	1	1	0
Row24	A6NKT7;Q...	A6NKT7;Q...	18;16		1;1	RanBP2-li...	RGPD3;R...	RanBP2-like and GRIP domain-containing protein 3 OS=Homo sapi...	2	18	1	1	5	8
Row25	A8CG34;Q...	A8CG34;Q...	4;3		4;3	Nuclear e...	POM121C...	Nuclear envelope pore membrane protein POM 121C OS=Homo s...	2	4	4	4	0	0
Row26	A8MXV4;R...	A8MXV4	3;1;1		3;1;1	Nucleosi...	NUDT19	Nucleoside diphosphate-linked moiety X motif 19 OS=Homo sapie...	3	3	3	3	1	0
Row27	A9UHW6	A9UHW6	2		2	MIF4G do...	MIF4GD	MIF4G domain-containing protein OS=Homo sapiens GN=MIF4GD ...	1	2	2	2	0	0
Row28	B01T2	B01T2	39		38	Unconven...	MYO1G	Unconventional myosin-Ig OS=Homo sapiens GN=MYO1G PE=1 SV...	1	39	39	38	25	30
Row29	B4DLN1;Q...	B4DLN1;Q...	14;7		7;7	Mitochon...	SLC25A10	cDNA FLJ60124, highly similar to Mitochondrial dicarboxylate carrie...	2	14	14	7	10	10
Row30	P0CG08;B...	P0CG08;B...	1;1		1;1	Golgi pH ...	GPR89B;...	Golgi pH regulator B OS=Homo sapiens GN=GPR89B PE=1 SV=1;...	2	1	1	1	0	0
Row31	C9JAW5;Q...	C9JAW5;Q...	1;1		1;1	HIG1 dom...	HIGD1A	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;HIG1 dom...	2	1	1	1	0	0
Row32	E7ENX8;P...	E7ENX8	11;2;1;1		1;1	2;2;1;1	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=3 SV=1	4	11	11	2	7	6
Row33	E7EVH7;Q...	E7EVH7;Q...	11;11;3		1;1	2;2;1;1	NA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Kinesin lig...	3	11	11	11	4	4
Row34	E9PAV3;Q...	E9PAV3	5;1		1;1	2;2;1;1	NA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Type peptide-associ...	2	5	5	5	3	5
Row35	E9PL57;Q...	E9PL57;Q...	4;3;1		1;1	4;3;1	NEDD8	NEDD8-MDP1 readthrough (Fragment) OS=Homo sapiens GN=NE...	3	4	4	4	2	3
Row36	E9PLD3;E...	E9PLD3;E...	1;1		1;1	Uncharac...	C11orf98	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Uncharac...	2	1	1	1	1	1
Row37	E9PLN8;Q...	E9PLN8;Q...	1;1		1;1	Vitamin K...	VKORC1	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=...	2	1	1	1	1	1
Row38	EEUEP2;P...	EEUEP2;P...	2;2		2;2	2;2;1;1	NA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;2;2;1;1	2	2	2	2	1	1

Row ID	S Protein IDs	S Majority protein IDs	S Peptide counts (all)	S Peptide counts (razor...)	S Peptide counts (uniq...)	S Protein names	S Gene names	S Fasta headers	I Number of proteins	I Peptides	I Razor + unique peptides	I Unique peptides	I Peptides WT_1	I
Row0	A0A096LPI6	A0A096LPI6	11	9	0	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=2	1	11	9	0	7	6
Row1	A0A0A6YY...	A0A0A6YY...	4;4	4;4	4;4	Transmembrane domain-containing protein TMED7 TL	TMED7 TL	Protein TMED7_TICAM2 OS=Homo sapiens GN=TMED7_TICAM2 PE=...	2	4	4	4	3	1
Row2	A0A0A6YY...	A0A0A6YY...	2;2	2;2	2;2	AP-2-associated protein kinase	AP-2-APK	Protein AP-2-APK OS=Homo sapiens PE=1 SV=1	2	2	2	2	1	0
Row3	A0A0A6YY...	A0A0A6YY...	10;10;1	10;10;1	10;10;1	60S ribosomal protein L10	L10	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=2	3	10	10	10	9	9
Row4	A0A0B4J1...	A0A0B4J1...	7;7	7;7	7;7	Superoxide dismutase	Superoxide dismutase	Superoxide dismutase OS=Homo sapiens PE=3 SV=1;S...	2	7	7	7	2	3
Row5	A0A0B4J2...	A0A0B4J2...	14;14	1;1	1;1	Tubulin	Tubulin	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1;Tubulin b...	2	14	1	1	11	11
Row6	A0A0B4J2...	A0A0B4J2...	15;15	15;15	13;13	Peroxiredoxin 2	PEROXIREDOXIN 2	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1;Peroxiredoxin 2...	2	15	15	13	6	4
Row7	A0A0J9YW...	A0A0J9YW...	2;2;2;2;2;...	2;2;2;2;2;...	2;2;2;2;2;...	Ig heavy chain	Ig heavy chain	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=1	14	2	2	2	1	1
Row8	A0A0U1R...	A0A0U1R...	1	1	1	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=1	1	1	1	1	0	0
Row9	A0A1B0G...	A0A1B0G...	7;7	7;7	7;7	Cathepsin D	Cathepsin D	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1;Cathepsin D...	2	7	7	7	1	2
Row10	A0A1W2P...	A0A1W2P...	14	2	0	NA	NA	Uncharacterized protein OS=Homo sapiens PE=4 SV=1	1	14	2	0	11	11
Row11	A0A1W2P...	A0A1W2P...	9;2	2;2	2;2	NA	NA	RPS10-NUDT3 readthrough OS=Homo sapiens GN=RPS10-NUDT3 ...	2	9	2	2	7	8
Row12	A0A1W2P...	A0A1W2P...	3;2;1	3;2;1	3;2;1	Ester hydrolase	C11orf54	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Ester hydrolase...	3	3	3	3	1	1
Row13	A0AVT1	A0AVT1	29	29	29	Ubiquitin...	UBA6	Ubiquitin-like modifier-activating enzyme 6 OS=Homo sapiens GN=...	1	29	29	29	9	8
Row14	A0FGR8	A0FGR8	1	1	1	Extended...	ESYT2	Extended synaptotagmin-2 OS=Homo sapiens GN=ESYT2 PE=1 S...	1	1	1	1	0	0
Row15	A1L0T0	A1L0T0	6	6	6	Acetolactate synthase	ILVBL	Acetolactate synthase-like protein OS=Homo sapiens GN=ILVBL P...	1	6	6	6	1	0
Row16	A3KN83	A3KN83	2	2	2	Protein s...	SBNO1	Protein strawberry notch homolog 1 OS=Homo sapiens GN=SBNO...	1	2	2	2	0	0
Row17	A4D1E9	A4D1E9	4	4	4	Protein A	GTPBP10	Protein A OS=Homo sapiens GN=GTPBP10 PE=1 SV=1	1	4	4	4	0	1
Row18	A5YKK6	A5YKK6	20	20	20	CCR4-NOT complex subunit 1	CNOT1	CCR4-NOT transcription complex subunit 1 OS=Homo sapiens GN...	1	20	20	20	6	6
Row19	A6NDG6	A6NDG6	13	13	13	Phosphoglycerate kinase	PGP	Glycerol-3-phosphate phosphatase OS=Homo sapiens GN=PGP P...	1	13	13	13	7	7
Row20	A6NFQ2;A...	A6NFQ2;A...	2;1	2;1	2;1	TRPM8 channel-associated factor 2	TCAF2	TRPM8 channel-associated factor 2 OS=Homo sapiens GN=TCAF2...	2	2	2	2	1	1
Row21	A6NHQ2	A6NHQ2	4	1	1	rRNA/tRNA...	FBLL1	rrRNA/tRNA 2'-O-methyltransferase fibrillarin-like protein 1 OS=Hom...	1	4	1	1	3	4
Row22	A6NHR9	A6NHR9	14	14	14	Structural protein of chromatin	SMCHD1	Structural maintenance of chromosomes flexible hinge domain-complex...	1	14	14	14	2	3
Row23	A6NJ78	A6NJ78	1	1	1	Probable membrane protein	METTL15	Probable methyltransferase-like protein 15 OS=Homo sapiens GN...	1	1	1	1	1	0
Row24	A6NKT7;Q...	A6NKT7;Q...	18;16	1;1	1;1	RanBP2-like protein	RGPD3;R...	RanBP2-like and GRIP domain-containing protein 3 OS=Homo sapi...	2	18	1	1	5	8
Row25	A8CG34;Q...	A8CG34;Q...	4;3	4;3	4;3	Nuclear envelope protein	POM121C	Nuclear envelope pore membrane protein POM121C OS=Homo s...	2	4	4	4	0	0
Row26	A8MXV4;R...	A8MXV4	3;1;1	3;1;1	3;1;1	Nucleoside diphosphate-linked moiety X motif 19	NUDT19	Nucleoside diphosphate-linked moiety X motif 19 OS=Homo sapi...	3	3	3	3	1	0
Row27	A9UHW6	A9UHW6	2	2	2	MIF4G domain	MIF4GD	MIF4G domain-containing protein OS=Homo sapiens GN=MIF4GD ...	1	2	2	2	0	0
Row28	B0I1T2	B0I1T2	39	39	38	Unconventional myosin-Ig	MYO1G	Unconventional myosin-Ig OS=Homo sapiens GN=MYO1G PE=1 SV...	1	39	39	38	25	30
Row29	B4DLN1;Q...	B4DLN1;Q...	14;7	14;7	7;7	Mitochondrial membrane protein	SLC25A10	cDNA FLJ60124, highly similar to Mitochondrial dicarboxylate carrier	2	14	14	7	10	10
Row30	P0CG08;B...	P0CG08;B...	1;1	1;1	1;1	Golgi pH regulator B	GPR89B;...	Golgi pH regulator B OS=Homo sapiens GN=GPR89B PE=1 SV=1;...	2	1	1	1	0	0
Row31	C9JAW5;Q...	C9JAW5;Q...	1;1	1;1	1;1	HIG1 domain	HIGD1A	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;HIG1 domain...	2	1	1	1	0	0
Row32	E7ENX8;P...	E7ENX8	11;2;1;1	11;2;1;1	2;2;1;1	NA	NA	Uncharacterized protein (Fragment) OS=Homo sapiens PE=3 SV=1	4	11	11	2	7	6
Row33	E7EVH7;Q...	E7EVH7;Q...	11;11;3	11;11;3	11;11;3	Kinesin light chain	KLC1	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Kinesin lig...	3	11	11	11	4	4
Row34	E9PAV3;Q...	E9PAV3	5;1	5;1	5;1	Nascent polypeptide-associated complex subunit alpha	NACA	Nascent polypeptide-associated complex subunit alpha, muscle-specific	2	5	5	5	3	5
Row35	E9PL57;Q...	E9PL57;Q...	4;3;1	4;3;1	4;3;1	NEDD8	NEDD8-M...	NEDD8-MDP1 readthrough (Fragment) OS=Homo sapiens GN=NE...	3	4	4	4	2	3
Row36	E9PLD3;E...	E9PLD3;E...	1;1	1;1	1;1	Uncharacterized protein	C11orf98	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;Uncharact...	2	1	1	1	1	1
Row37	E9PLN8;Q...	E9PLN8;Q...	1;1	1;1	1;1	Vitamin K-dependent protein	VKORC1	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=...	2	1	1	1	1	1
Row38	EEHEP2;P...	EEHEP2;P...	2;2	2;2	2;2	2-oxoisocoumarin	PKCDH4	Uncharacterized protein OS=Homo sapiens PE=4 SV=1;2-oxoisocou...	2	2	2	2	1	1

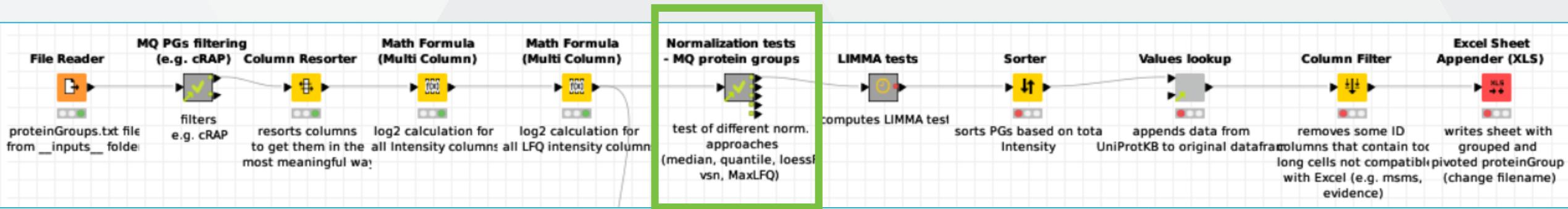
Columns can be of different **data types**:

- **S (string):** any text
 - **I (integer):** “whole” numbers
 - **D (double):** numbers with decimal places
 -  **(date and time)**

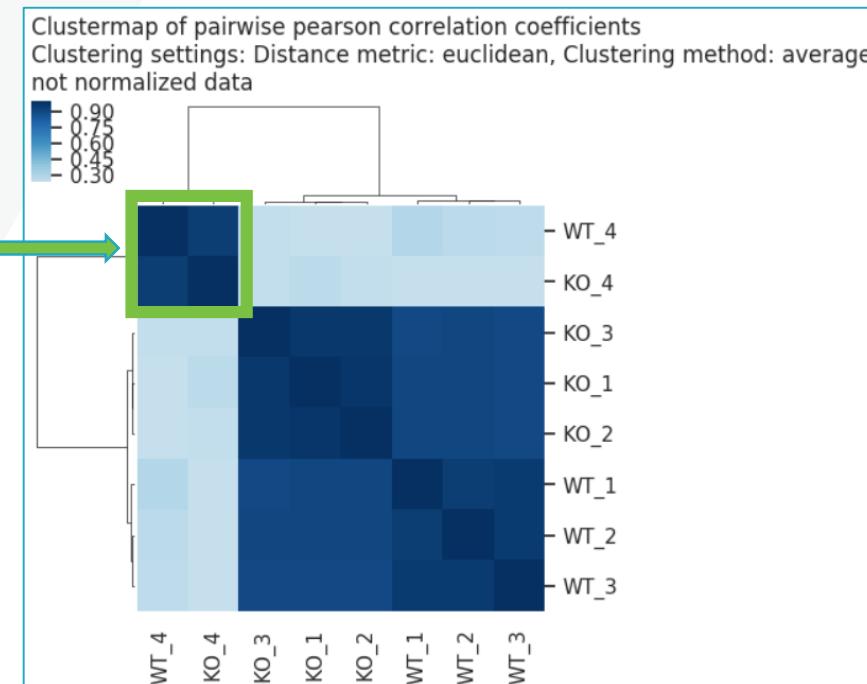
KNIME environment and metanodes



KNIME environment and metanodes

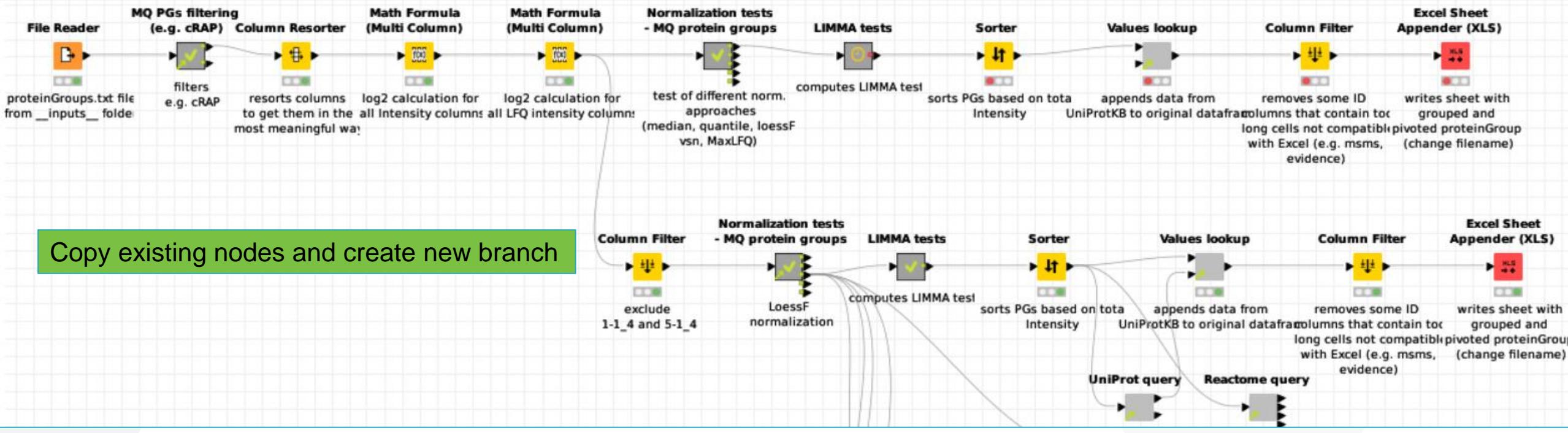


How will the clustering change
when outliers are excluded?

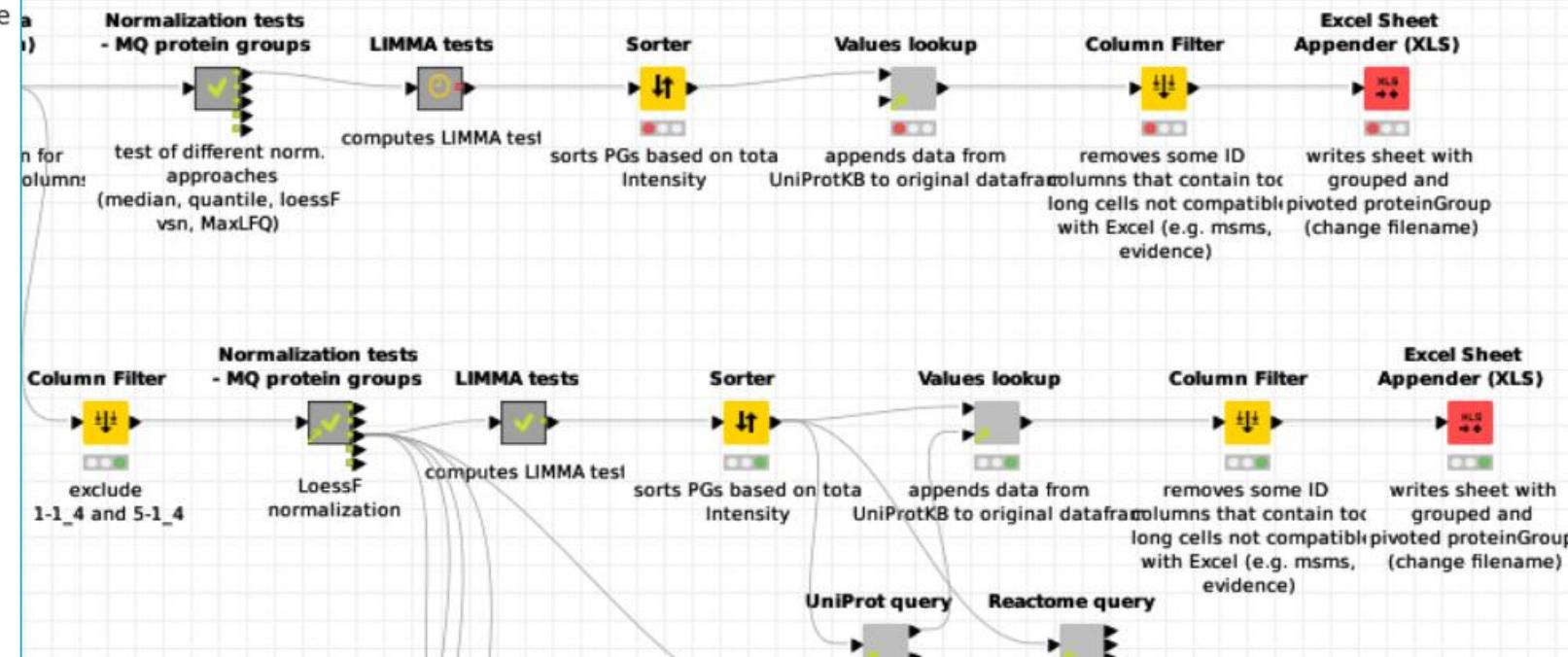
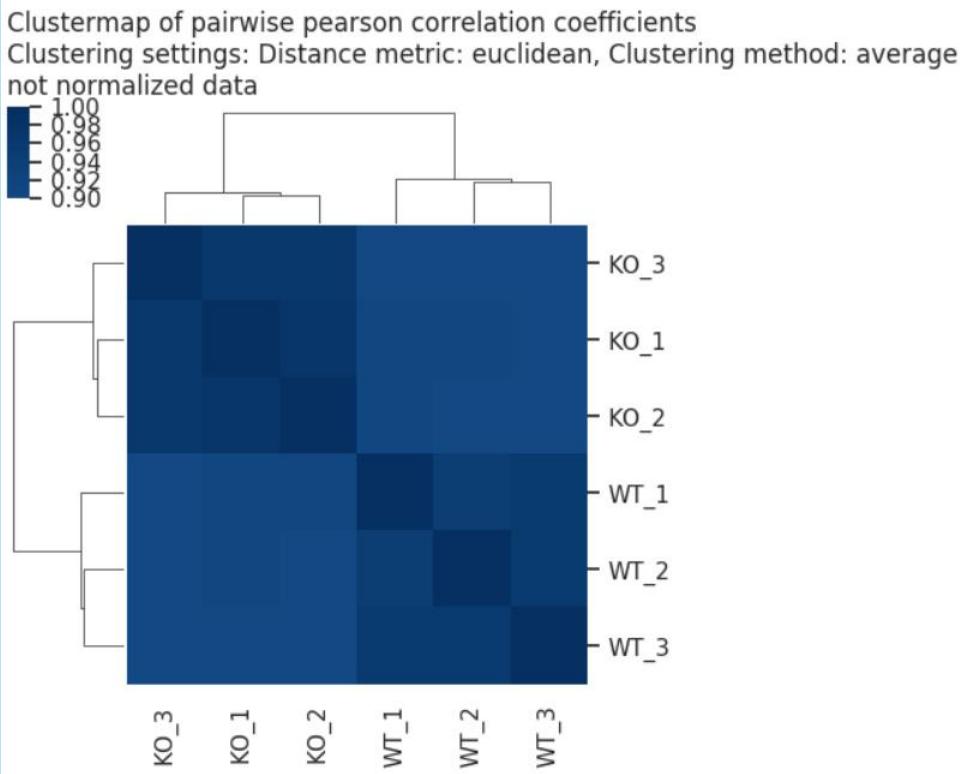


Alternative ways of data processing

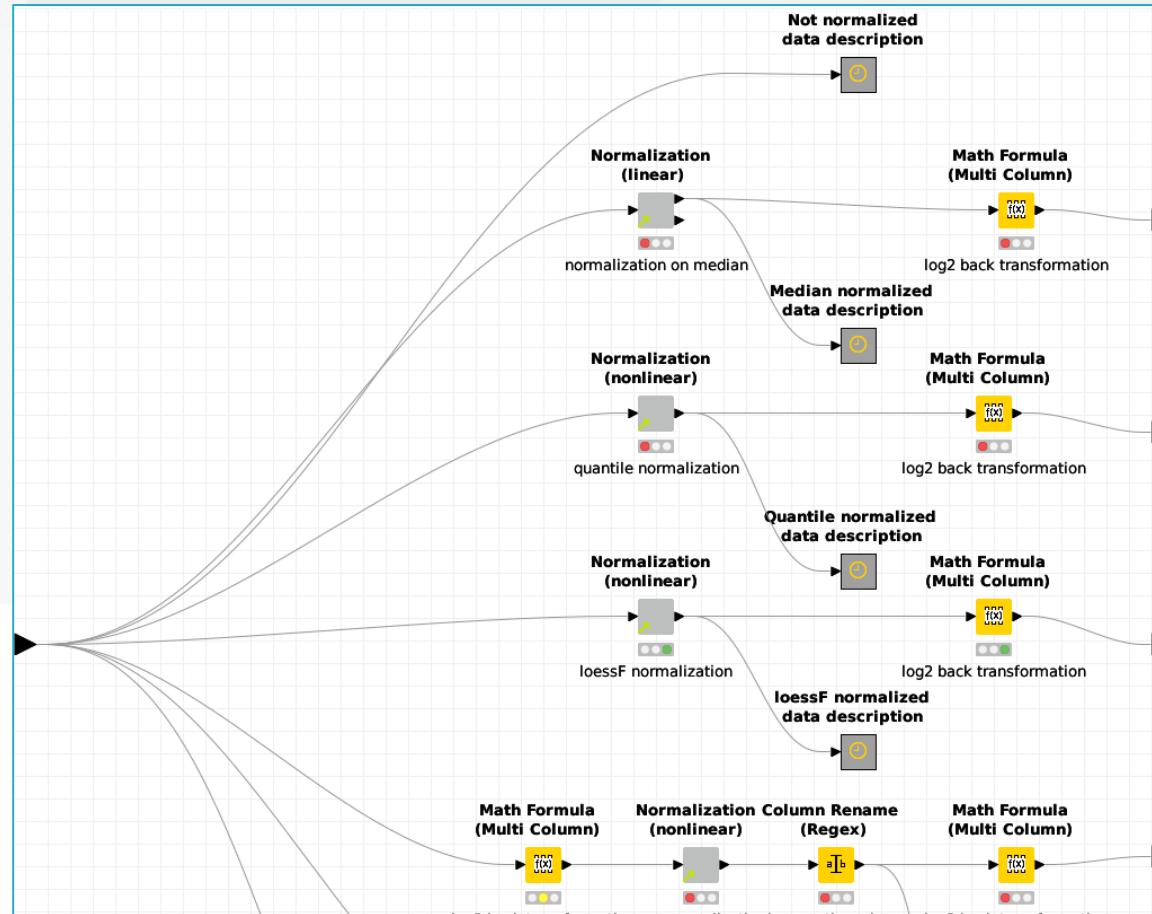
KNIME environment and metanodes



KNIME environment and metanodes

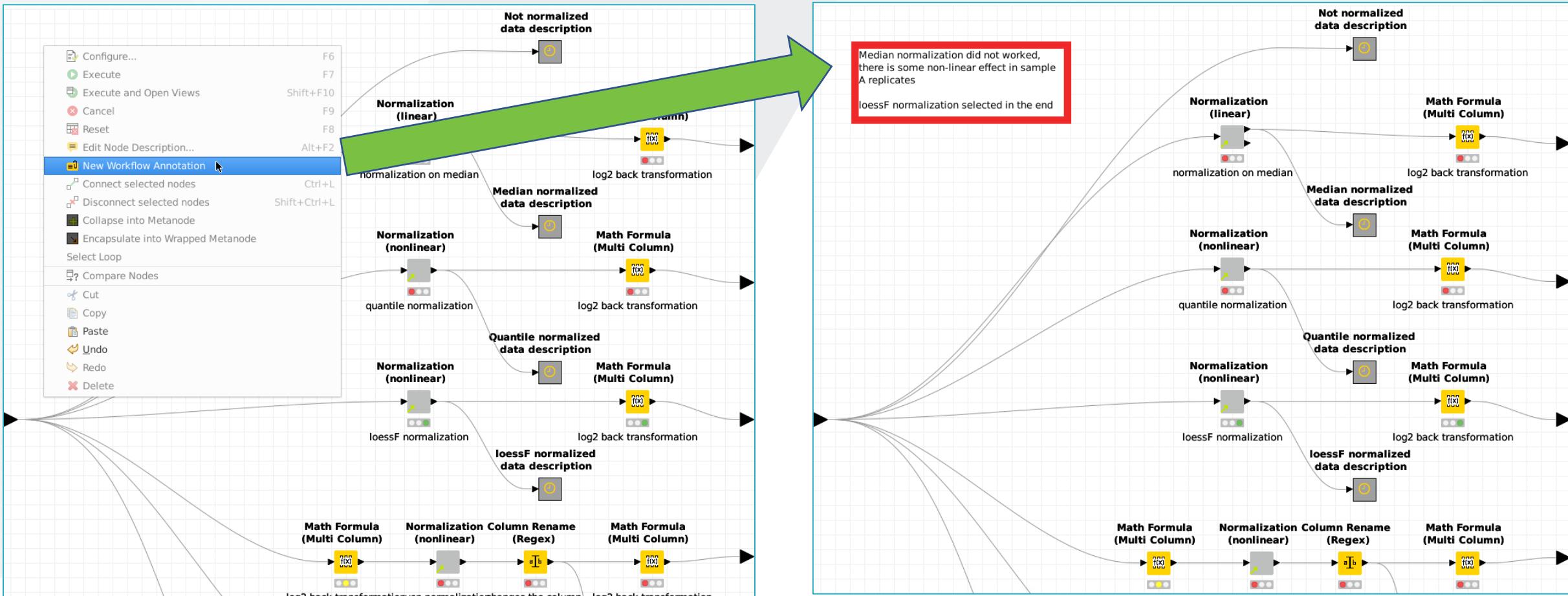


KNIME environment and metanodes



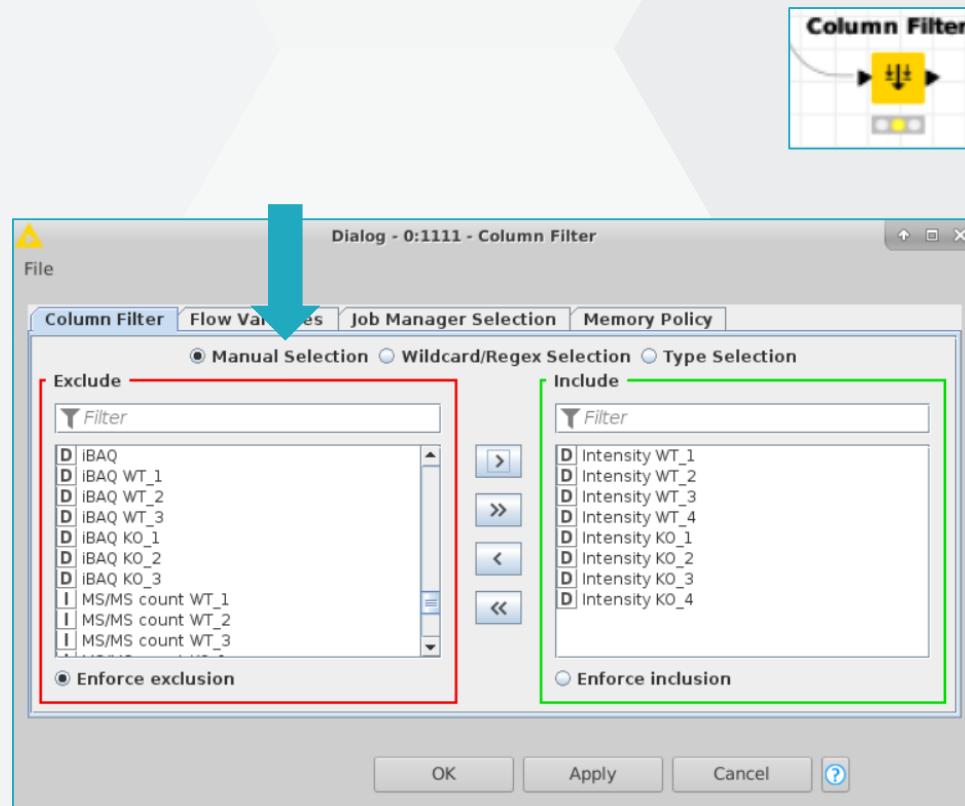
Alternative ways of data processing
Normalization metanode

KNIME environment and metanodes



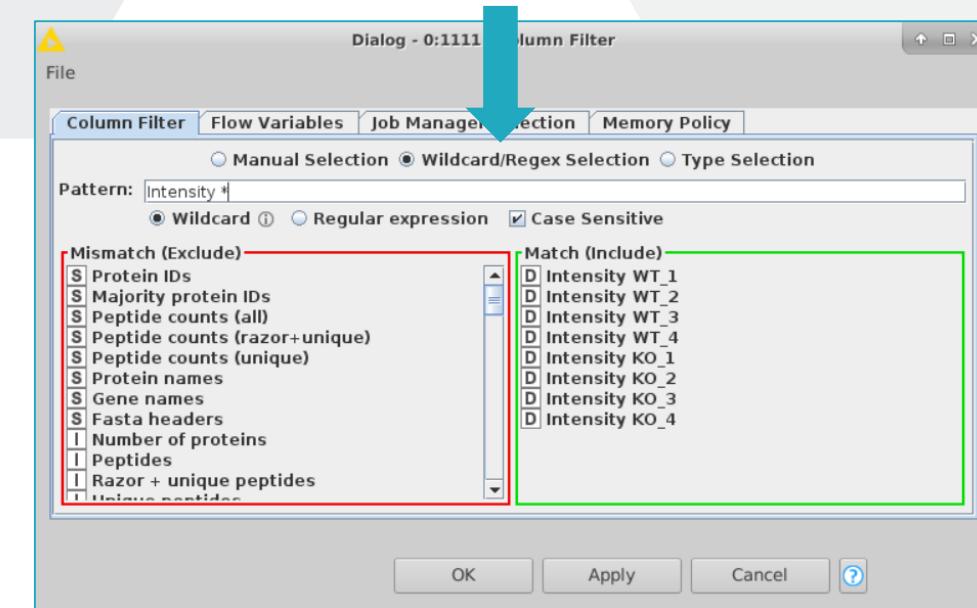
Add workflow annotation to describe your partial results or your reasons for the next processing steps

KNIME environment and metanodes



Manual Selection

Selecting particular columns



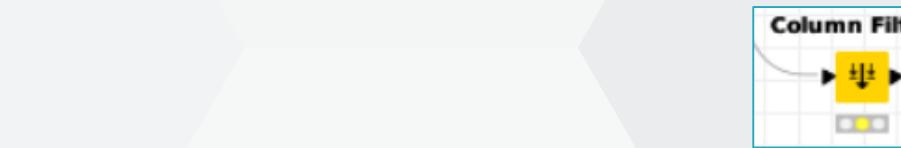
Wildcard/Regex Selection

* any number of characters

? single character

○ Manual Selection ○ Wildcard/Regex Selection ○ Type Selection
Pattern: Intensity\w[A-Za-z]{2}\d{1}
○ Wildcard ○ Regular expression Case Sensitive

KNIME environment and metanodes



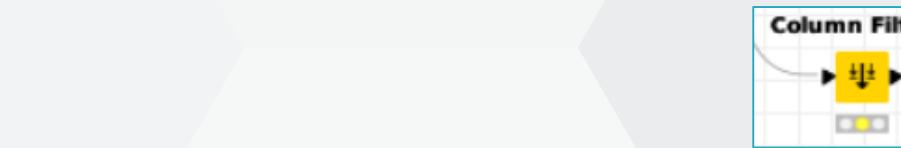
Selecting particular columns

Table "proteinGroups.txt"								- Rows: 3855	- Spec - Columns: 81	Properties	Flow Variables
Row ID	Protein IDs	Majority protein IDs	Peptide counts (all)	Peptide counts (razor...)	Peptide counts (uniq...)	Protein names	Gene names				
Row0	A0A096LP16	A0A096LP16	11	9	0	NA	NA				
Row1	A0A0A6YY...	A0A0A6YY...	4;4	4;4	4;4	Transme...	TMED7-TI...				
Row2	A0A0A6YY...	A0A0A6YY...	2;2	2;2	2;2	AP-3 com...	C15orf38...				
Row3	A0A0A6YY...	A0A0A6YY...	10;10;1	10;10;1	10;10;1	60S ribos...	RPL17-C1...				
Row4	A0A0B4J1...	A0A0B4J1...	7;7	7;7	7;7	Suppress...	PPAN-P2...				
Row5	A0A0B4J2...	A0A0B4J2...	14;14	1;1	1;1	Tubulin b...	TUBB3				
Row6	A0A0B4J2...	A0A0B4J2...	15;15	15;15	13;13	Periodic t...	PWP2				
Row7	A0A0J9YW...	A0A0J9YW...	2;2;2;2;2...	2;2;2;2;2...	Ig heavy ...	IGHV4-61...					
Row8	A0A0U1R...	A0A0U1R...	1	1	1	NA	NA				
Row9	A0A1B0G...	A0A1B0G...	7;7	7;7	7;7	Cathepsi...	CTSD				
Row10	A0A1W2P...	A0A1W2P...	14	2	0	NA	NA				
Row11	A0A1W2P...	A0A1W2P...	9;2	2;2	2;2	NA	NA				
Row12	A0A1W2P...	A0A1W2P...	3;2;1	3;2;1	3;2;1	Ester hyd...	C11orf54				
Row13	A0AVT1	A0AVT1	29	29	29	Ubiquitin...	UBA6				
Row14	A0FGR8	A0FGR8	1	1	1	Extended...	ESYT2				
Row15	A1L0T0	A1L0T0	6	6	6	Acetolact...	ILVBL				
Row16	A3KN83	A3KN83	2	2	2	Protein s...	SBNO1				
Row17	A4D1E9	A4D1E9	4	4	4	Protein A	GTPBP10				
Row18	A5YKK6	A5YKK6	20	20	20	CCR4-NO...	CNOT1				
Row19	A6NDG6	A6NDG6	13	13	13	Phospho...	PGP				
Row20	A6NFQ2;A...	A6NFQ2;A...	2;1	2;1	2;1	TRPM8 ch...	TCAF2				
Row21	A6NHQ2	A6NHQ2	4	1	1	rRNA/tRN...	FBL1				
Row22	A6NRH9	A6NRH9	14	14	14	Structura...	SMCHD1				
Row23	A6NJ78	A6NJ78	1	1	1	Probable...	METTL15				
Row24	A6NKT7;Q...	A6NKT7;Q...	18;16	1;1	1;1	RanBP2-li...	RGPD3;R...				
Row25	A8CG34;Q...	A8CG34;Q...	4;3	4;3	4;3	Nuclear e...	POM121C...				
Row26	A8MXV4;R...	A8MXV4	3;1;1	3;1;1	3;1;1	Nucleosi...	NUDT19				
Row27	A9UHW6	A9UHW6	2	2	2	MIF4G do...	MIF4GD				
Row28	B01T2	B01T2	39	39	38	Unconven...	MYO1G				
Row29	B4DLN1;Q...	B4DLN1;Q...	14;7	14;7	7;7	Mitochon...	SLC25A10				
Row30	P0CG08;B...	P0CG08;B...	1;1	1;1	1;1	Golgi pH ...	GPR89B;...				
Row31	C9JAW5;Q...	C9JAW5;Q...	1;1	1;1	1;1	HIG1 dom...	HIGD1A				
Row32	E7ENX8;P...	E7ENX8	11;2;1;1	11;2;1;1	2;2;1;1	NA	NA				
Row33	E7EVH7;Q...	E7EVH7;Q...	11;11;3	11;11;3	11;11;3	Kinesin li...	KLC1				
Row34	E9PAV3;Q...	E9PAV3	5;1	5;1	5;1	Nascent ...	NACA				
Row35	E9PL57;Q...	E9PL57;Q...	4;3;1	4;3;1	4;3;1	NEDD8	NEDD8-M...				
Row36	E9PLD3;E...	E9PLD3;E...	1;1	1;1	1;1	Uncharac...	C11orf98				
Row37	E9PLN8;Q...	E9PLN8;Q...	1;1	1;1	1;1	Vitamin K...	VKORC1				
Row38	EEUEP2;P...	EEUEP2;P...	2;2	2;2	2;2	2;2;2;2;2...	EEUEP2;P...				
Row39	EEUEP2;P...	EEUEP2;P...	2;2	2;2	2;2	2;2;2;2;2...	EEUEP2;P...				

Filtering
Intensity * columns

Table "default"												- Rows: 3855	- Spec - Columns: 8	Properties	Flow Variables
Row ID	Intensity WT_1	Intensity WT_2	Intensity WT_3	Intensity WT_4	Intensity KO_1	Intensity KO_2	Intensity KO_3	Intensity KO_4							
Row0	13,805,142.7...	12,334,974.4...	71,774,050.5...	15,572,388.0...	46,310,787.8...	65,411,765.7...	27,839,936.0...	28,964,138...							
Row1	?	?	?	?	13,352,870.9...	50,866,275.16...	39,509,763.1...	20,286,929.2...	48,905,794...						
Row2	4,160,334,638...	?	?	?	3,590,791.711	3,540,430.878	1,125,920.393	?							
Row3	1,028,354,513...	314,222,598...	1,280,566,98...	15,563,821.8...	894,136,536...	1,134,082,69...	1,635,044,01...	17,004,683...							
Row4	6,383,802,317	4,694,509,839	34,252,056.6...	?	15,392,710.3...	29,521,666.2...	20,631,702.7...	26,533,305...							
Row5	?	?	?	?	?	22,386,282.3...	5,698,109,921	?							
Row6	45,039,556,163	6,157,646,112	83,816,383.7...	9,455,654,763	55,917,428.8...	44,384,862.8...	29,949,435.3...	6,222,054.0...							
Row7	46,969,859.13	14,405,226.6...	28,810,180.6...	?	45,626,826.2...	66,037,583.4...	25,067,767.0...	?							
Row8	?	?	?	?	?	?	?	?							
Row9	12,895,069,479	3,612,318,917	32,890,061.6...	391,393,647...	15,571,478.06	13,145,711.7...	4,972,123,133	430,975,71...							
Row10	?	?	?	17,645,879.3...	?	4,431,229,258	1,242,353,534	20,573,911...							
Row11	?	1,454,052,925	?	3,559,576,992	?	?	?	?							
Row12	7,278,034,617	2,113,210,345	6,789,013,352	213,736,654...	9,675,544,988	4,720,476,968	2,074,351,321	203,337,43...							
Row13	54,171,556,037	14,595,970.3...	68,329,068.9...	54,171,556.0...	58,593,352.1...	84,210,793.1...	38,816,331.2...	58,593,352...							
Row14	?	?	?	137,976,535...	?	3,716,122,177	?	?							
Row15	5,320,155,427	?	?	?	?	17,536,605.8...	13,779,263,62	10,612,251.5...	4,291,075.3...						
Row16	?	?	?	?	?	?	?	?							
Row17	?	2,046,101,616	10,655,091.6...	?	?	?	?	?							
Row18	55,347,664,163	13,088,101.8...	41,248,787.4...	359,827,171...	39,173,530.8...	32,866,689.33	15,099,435.2...	296,568,00...							
Row19	71,992,970,049	19,154,319.73	88,855,174.1	200,024,787...	85,563,561.2...	94,532,283.7	44,716,978.1...	246,184,02...							
Row20	890,579,777	?	?	1,366,094,69...	972,197,046	1,597,896,889	268,900,452	1,290,778.4...							
Row21	?	13,058,093.19	66,441,285.2...	5,275,453.284	50,187,124.4...	?	26,998,669.1...	11,795,860...							
Row22	59,882,689,882	6,212,733,741	56,689,054.7...	?	6,227,628,293	23,367,589.1...	5,757,035,496	7,194,820.0...							
Row23	98,544,334,268	20,479,323,18	84,373,772.8...	6,439,973,522	2,158,196.13	?	1,016,594,482	11,247,742...							
Row24	6,754,534,287	?	?	377,031,701...	8,765,007,713	8,480,396,246	2,499,980,107	404,221,96...							
Row25	?	?	?	2,766,398,83...	4,019,159,526	7,448,896,21	2,245,272,389	2,648,083,2...							
Row26	8,186,897,83	?	?	54,040,722,7...	5,461,522,912	5,859,424,259	4,621,569,045	43,169,566...							
Row27	101,043,892.8...	20,432,110.4...	89,429,608	1,206,242,33...	2,484,254,773	2,595,930,984	?	1,256,513.7...							
Row28	494,047,248.7...	11,268,416...	480,779,680...	494,047,248...	491,505,055...	522,048,431...	495,165,972...	491,505,05...							
Row29	204,876,470.3...	46,750,630.2...	172,037,542...	67,321,686.8...	130,567,297...	171,180,153...	109,584,291...	68,693,558...							
Row30	?	?	?	83,022,355.4...	?	?	?	?							
Row31	?	?	?	10,928,040.4...	5,932,887,497	?	?	?							
Row32	59,777,456.03	13,411,808.9...	73,261,192.3...	?	59,858,974.6...	66,362,341.8...	38,006,238.8...	50,187,124...							
Row33	14,671,669,374	4,404,999,811	34,091,238.5...	39,142,061.8...	33,545,136.1...	34,370,554,78	16,886,550.2...	34,878,282...							
Row34	138,538,281.43	40,677,775.3...	165,100,708.5...	1,198,523,51...	117,652,941...	134,983,192...	88,565,189.3...	1,244,078,1...							
Row35	77,916,388,388	21,069,492.6...	80,497,529.8...	?	72,284,195.78...	68,722,025.2...	38,381,032.2...	?							
Row36	31,876,367,938	5,416,680,068	27,488,776.4...	143,689,405...	23,858,141.79	20,890,186.3...	8,719,334,296	131,808,08...							
Row37	18,562,401,605	3,906,078,279	12,935,077.1...	?	18,680,678.3...	16,908,523.8...	7,826,320,112	4,950,995.7...							
Row38	2,568,251,553	645,107,373	1,699,465,973	?	2,626,481,715	2,596,831,245	617,488,994	?							
Row39	21,242,586,846	3,108,789,822	23,171,114.1...	?	20,206,887.1...	24,492,980.7...	7,895,203,846	2,732,346.1...							
Row40	?	?	?	?	?	?	?	?							

KNIME environment and metanodes



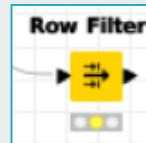
Selecting particular columns

Table "proteinGroups.txt"									
	- Rows: 3855 Spec - Columns: 81							Properties	Flow Variables
Row ID	Protein IDs	Majority protein IDs	Peptide counts (all)	Peptide counts (razor...)	Peptide counts (uniq...)	Protein names	Gene names		
Row0	A0A096LP16	A0A096LP16	11	9	0	NA	NA		
Row1	A0A0A6YY...	A0A0A6YY...	4;4	4;4	4;4	Transme...	TMED7-TI...		
Row2	A0A0A6YY...	A0A0A6YY...	2;2	2;2	2;2	AP-3 com...	C15orf38...		
Row3	A0A0A6YY...	A0A0A6YY...	10;10;1	10;10;1	10;10;1	60S ribos...	RPL17-C1...		
Row4	A0A0B4J1...	A0A0B4J1...	7;7	7;7	7;7	Suppress...	PPAN-P2...		
Row5	A0A0B4J2...	A0A0B4J2...	14;14	1;1	1;1	Tubulin b...	TUBB3		
Row6	A0A0B4J2...	A0A0B4J2...	15;15	15;15	13;13	Periodic t...	PWP2		
Row7	A0A0J9YW...	A0A0J9YW...	2;2;2;2;2...	2;2;2;2;2...	Ig heavy ...	IGHV4-61...			
Row8	A0A0U1R...	A0A0U1R...	1	1	1	NA	NA		
Row9	A0A1B0G...	A0A1B0G...	7;7	7;7	7;7	Cathepsi...	CTSD		
Row10	A0A1W2P...	A0A1W2P...	14	2	0	NA	NA		
Row11	A0A1W2P...	A0A1W2P...	9;2	2;2	2;2	NA	NA		
Row12	A0A1W2P...	A0A1W2P...	3;2;1	3;2;1	3;2;1	Ester hyd...	C11orf54		
Row13	A0AVT1	A0AVT1	29	29	29	Ubiquitin...	UBA6		
Row14	A0FGR8	A0FGR8	1	1	1	Extended...	ESYT2		
Row15	A1L0TO	A1L0TO	6	6	6	Acetolact...	ILVBL		
Row16	A3KN83	A3KN83	2	2	2	Protein s...	SBNO1		
Row17	A4D1E9	A4D1E9	4	4	4	Protein A	GTPBP10		
Row18	A5YKK6	A5YKK6	20	20	20	CCR4-NO...	CNOT1		
Row19	A6NDG6	A6NDG6	13	13	13	Phospho...	PGP		
Row20	A6NFQ2;A...	A6NFQ2;A...	2;1	2;1	2;1	TRPM8 ch...	TCAF2		
Row21	A6NHQ2	A6NHQ2	4	1	1	rRNA/tRN...	FBL1		
Row22	A6NRH9	A6NRH9	14	14	14	Structura...	SMCHD1		
Row23	A6NJ78	A6NJ78	1	1	1	Probable...	METTL15		
Row24	A6NKT7;Q...	A6NKT7;Q...	18;16	1;1	1;1	RanBP2-li...	RGPD3;R...		
Row25	A8CG34;Q...	A8CG34;Q...	4;3	4;3	4;3	Nuclear e...	POM121C...		
Row26	A8MXV4;R...	A8MXV4	3;1;1	3;1;1	3;1;1	Nucleosi...	NUDT19		
Row27	A9UHW6	A9UHW6	2	2	2	MIF4G do...	MIF4GD		
Row28	B01T2	B01T2	39	39	38	Unconven...	MYO1G		
Row29	B4DLN1;Q...	B4DLN1;Q...	14;7	14;7	7;7	Mitochon...	SLC25A10		
Row30	P0CG08;B...	P0CG08;B...	1;1	1;1	1;1	Golgi pH ...	GPR89B;...		
Row31	C9JAW5;Q...	C9JAW5;Q...	1;1	1;1	1;1	HIG1 dom...	HIGD1A		
Row32	E7ENX8;P...	E7ENX8	11;2;1;1	11;2;1;1	2;2;1;1	NA	NA		
Row33	E7EVH7;Q...	E7EVH7;Q...	11;11;3	11;11;3	11;11;3	Kinesin li...	KLC1		
Row34	E9PAV3;Q...	E9PAV3	5;1	5;1	5;1	Nascent ...	NACA		
Row35	E9PL57;Q...	E9PL57;Q...	4;3;1	4;3;1	4;3;1	NEDD8	NEDD8-M...		
Row36	E9PLD3;E...	E9PLD3;E...	1;1	1;1	1;1	Uncharac...	C11orf98		
Row37	E9PLN8;Q...	E9PLN8;Q...	1;1	1;1	1;1	Vitamin K...	VKORC1		
Row38	EEUEP2;P...	EEUEP2;P...	2;2	2;2	2;2	2;2;2;2;2...	EEUEP2;P...		
Row39	EEUEP2;P...	EEUEP2;P...	2;2	2;2	2;2	2;2;2;2;2...	EEUEP2;P...		

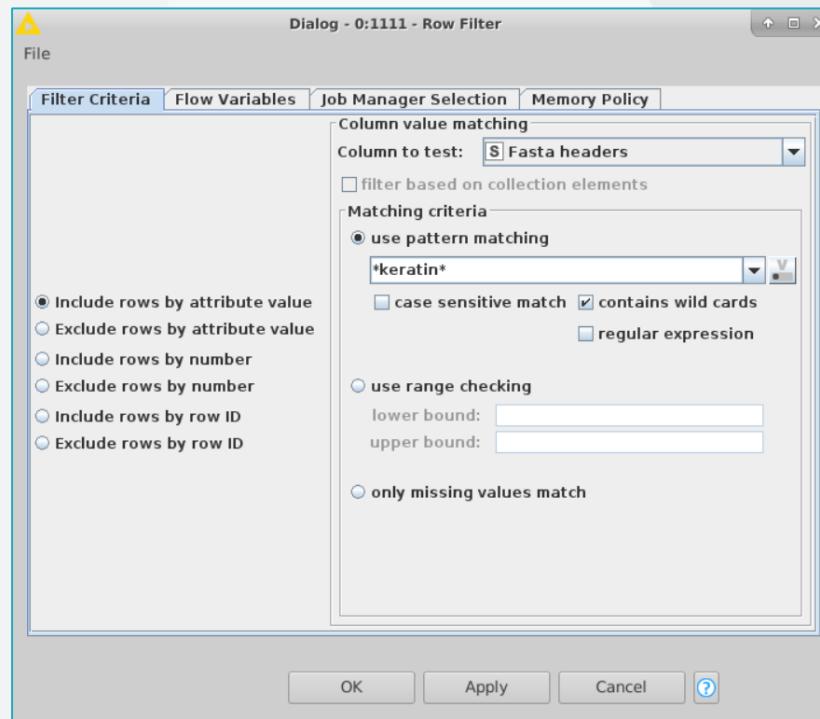
Filtering
Intensity * columns

Table "default"									
	- Rows: 3855	Spec - Columns: 8	Properties	Flow Variables					
Row ID	Intensity WT_1	Intensity WT_2	Intensity WT_3	Intensity WT_4	Intensity KO_1	Intensity KO_2	Intensity KO_3	Intensity KO_4	
Row0	138,005,142.7...	12,334,974.4...	71,774,050.5...	15,572,388.0...	46,310,787.8...	65,411,765.7...	27,839,936.0...	28,964,138...	
Row1	?	?	?	?	13,352,870.9...	50,866,275.16	39,509,763.1...	20,286,929.2...	48,905,794...
Row2	4,160,334,638	?	?	?	3,590,791.711	3,540,430.878	1,125,920.393	?	
Row3	1,028,354,513...	314,222,598...	1,280,566,98...	15,563,821.8...	894,136,536...	1,134,082,69...	1,635,044,01...	17,004,683...	
Row4	6,383,802,317	4,694,509,839	34,252,056.6...	?	15,392,710.3...	29,521,666.2...	20,631,702.7...	26,533,305...	
Row5	?	?	?	?	?	22,386,282.3...	5,698,109,921	?	
Row6	45,039,556,163	6,157,646,112	83,816,383.7...	9,455,654,763	55,917,428.8...	44,384,862.8...	29,949,435.3...	6,222,054.0...	
Row7	46,969,859.13	14,405,226.6...	28,810,180.6...	?	45,626,826.2...	66,037,583.4...	25,067,767.0...	?	
Row8	?	?	?	?	?	?	?	?	
Row9	12,895,069,479	3,612,318,917	32,890,061.6...	391,393,647...	15,571,478.06	13,145,711.7...	4,972,123,133	430,975,71...	
Row10	?	?	?	?	17,645,879.3...	4,431,229,258	1,242,353,534	20,573,911...	
Row11	?	1,454,052,925	?	?	3,559,576,992	?	?	?	31,459,265...
Row12	7,278,034,617	2,113,210,345	6,789,013,352	213,736,654...	9,675,544,988	4,720,476,968	2,074,351,321	203,337,43...	
Row13	54,171,556,03	4,475,5,970.3...	68,1...	793.1...	38,816,331.2...	58,593,352...			
Row14	?	?	?	?	22,177	?	?	?	
Row15	5,320,155,427	?	?	?	26,632	10,612,251.5...	4,291,075.3...		
Row16	?	?	?	?	?	?	?	?	
Row17	?	101.616	10,655,091.6...	?	?	?	?	?	
Row18	55,347,664,163	13,088,101.8...	41,248,787.4...	359,827,171...	39,173,530.8...	32,866,689.33	15,099,435.2...	296,568,00...	
Row19	71,992,970,049	19,154,319.73	88,855,174.1	200,024,787...	85,563,561.2...	94,532,283.7	44,716,978.1...	246,184,02...	
Row20	890,579,777	?	?	1,366,094,69...	972,197,046	1,597,896,889	268,900,452	1,290,778.4...	
Row21	?	13,058,093.19	66,441,285.2...	5,275,453.284	50,187,124.4...	?	26,998,669.1...	11,795,860...	
Row22	59,882,689,882	6,212,733.741	56,689,054.7...	?	6,227,628,293	23,367,589.1...	5,757,035,496	7,194,820.0...	
Row23	98,544,334,268	20,479,323.18	84,373,772.8...	6,439,973,522	2,158,196.13	?	1,016,594,482	11,247,742...	
Row24	6,754,534,287	?	?	377,031,701...	8,765,007,713	8,480,396,246	2,499,980,107	404,221,96...	
Row25	?	?	?	2,766,398,83...	4,019,159,526	7,448,896,21	2,245,272,389	2,648,083,2...	
Row26	8,186,897,83	?	?	54,040,722.7...	5,461,522,912	5,859,424,259	4,621,569,045	43,169,566...	
Row27	101,043,892.8...	20,432,110.4...	89,429,608	1,206,242,33...	2,484,254,773	2,595,930,984	?	1,256,513,7...	
Row28	494,047,248.7...	11,268,416...	480,779,680...	494,047,248...	491,505,055...	522,048,431...	495,165,972...	491,505,05...	
Row29	204,876,470.3...	46,750,630.2...	172,037,542...	67,321,686.8...	130,567,297...	171,180,153...	109,584,291...	68,693,558...	
Row30	?	?	?	83,022,355.4...	?	?	?	?	39,242,273...
Row31	?	?	?	10,928,040.4...	5,932,887,497	?	?	?	25,248,767...
Row32	59,777,456.03	13,411,808.9...	73,261,192.3...	?	59,858,974.6...	66,362,341.8...	38,006,238.8...	50,187,124...	
Row33	14,671,669,374	4,404,999,811	34,091,238.5...	39,142,061.8...	33,545,136.1...	34,370,554,78	16,886,550.2...	34,878,282...	
Row34	138,916,388.281.43	40,677,775.3...	165,100,708.5...	1,198,523,51...	117,652,941...	134,983,192...	88,565,189.3...	1,244,078,1...	
Row35	77,916,388.388	21,069,492.6...	80,497,529.8...	?	72,284,195.8...	68,722,025.2...	38,381,032.2...	?	
Row36	31,876,367,938	5,416,680,068	27,488,776.4...	143,689,405...	23,858,141.79	20,890,186.3...	8,719,334,296	131,808,08...	
Row37	18,562,401,605	3,906,078,279	12,935,077.1...	?	18,680,678.3...	16,908,523.8...	7,826,320,112	4,950,995.7...	
Row38	2,568,251,553	645,107,373	1,699,465,973	?	2,626,481,715	2,596,831,245	617,488,994	?	
Row39	21,242,586,846	3,108,789,822	23,171,114.1...	?	20,206,887.1...	24,492,980.7...	7,895,203,846	2,732,346.1...	
Row40	?	?	?	?	?	?	?	?	50,988,242...

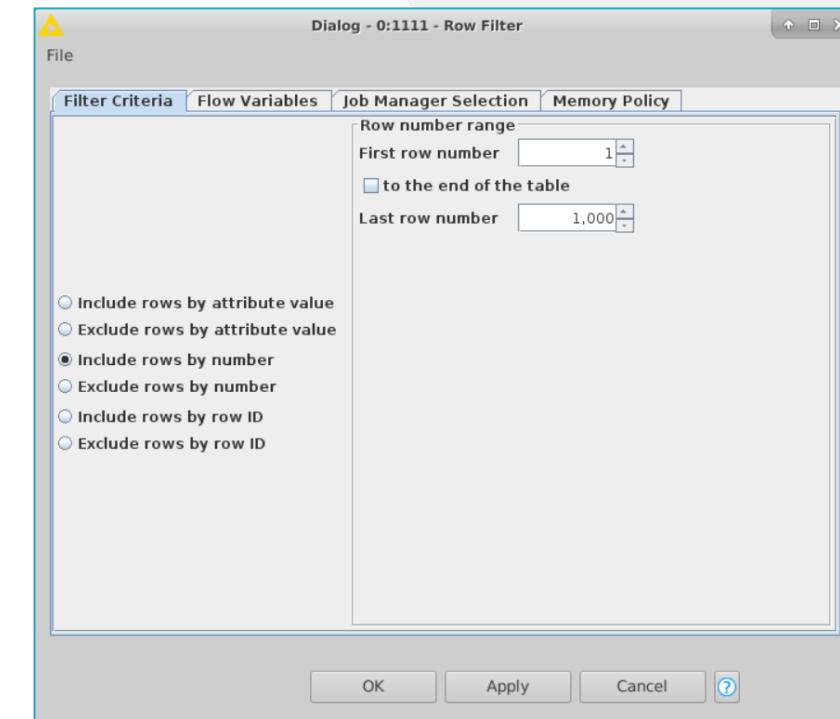
KNIME environment and metanodes



Selecting particular rows

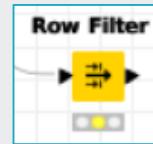
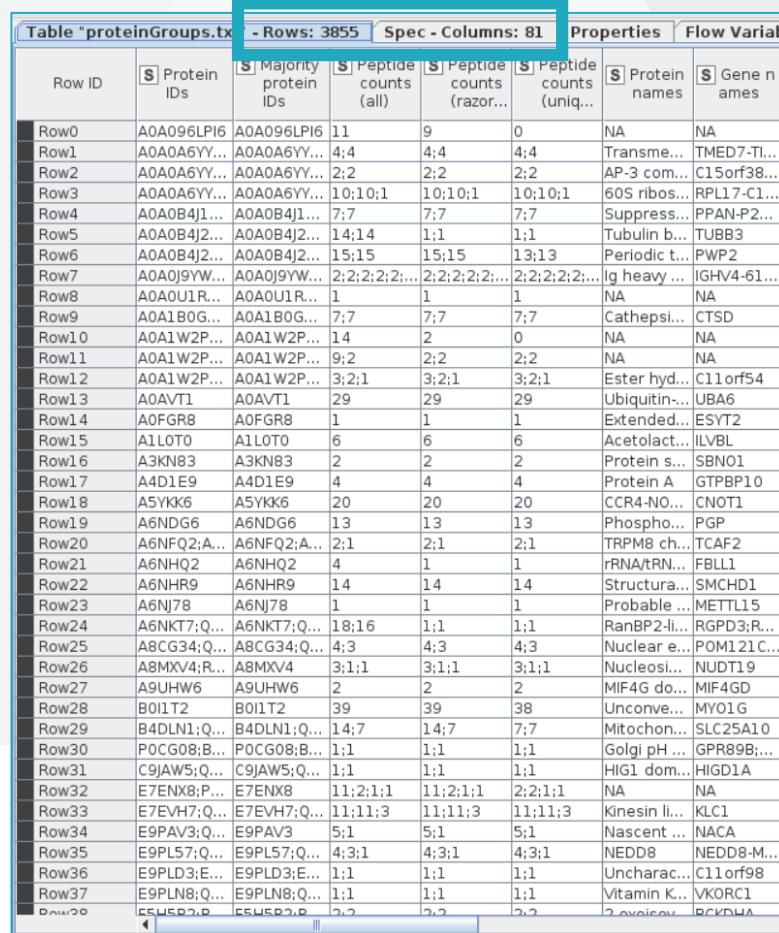


Filtering using matching criteria

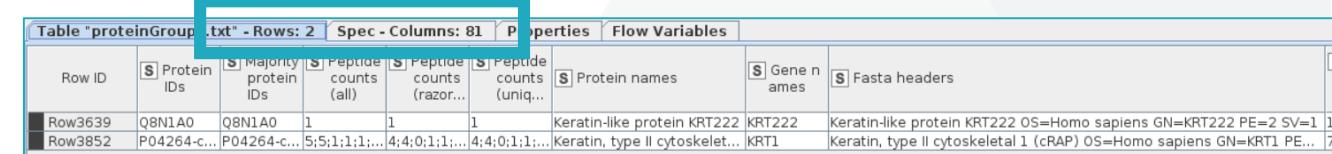


Filtering using rows range

KNIME environment and metanodes



Selecting particular rows



Filtering rows containing *keratin* keyword in Fasta headers column

When you just need or want to script...

Visual programming
(in-built KNIME nodes)

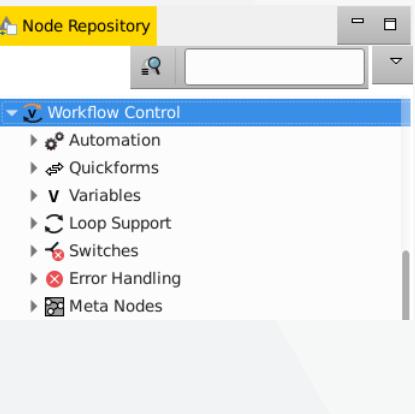


Scripting using snippets

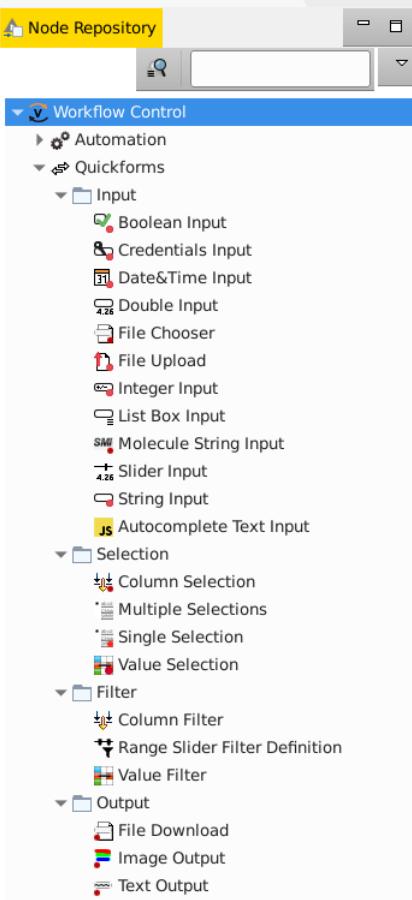


Visual programming concepts in KNIME

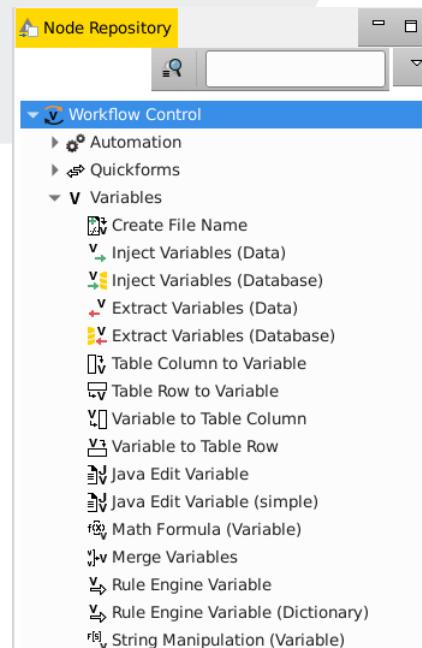
Workflow control nodes



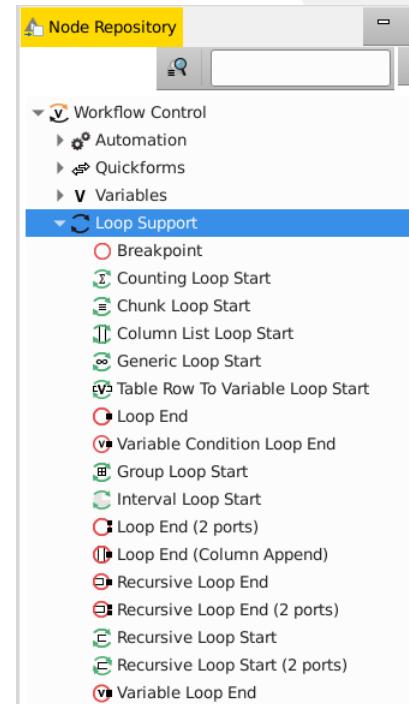
inputs support



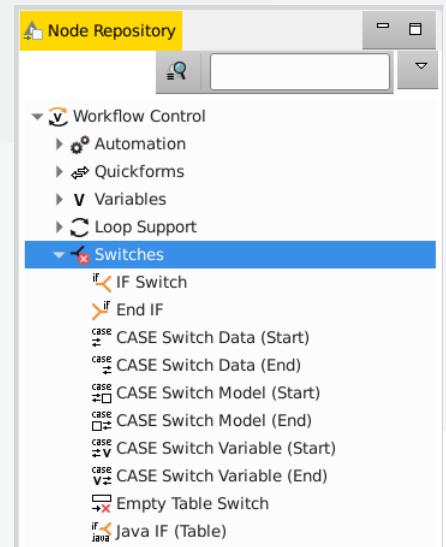
variables support



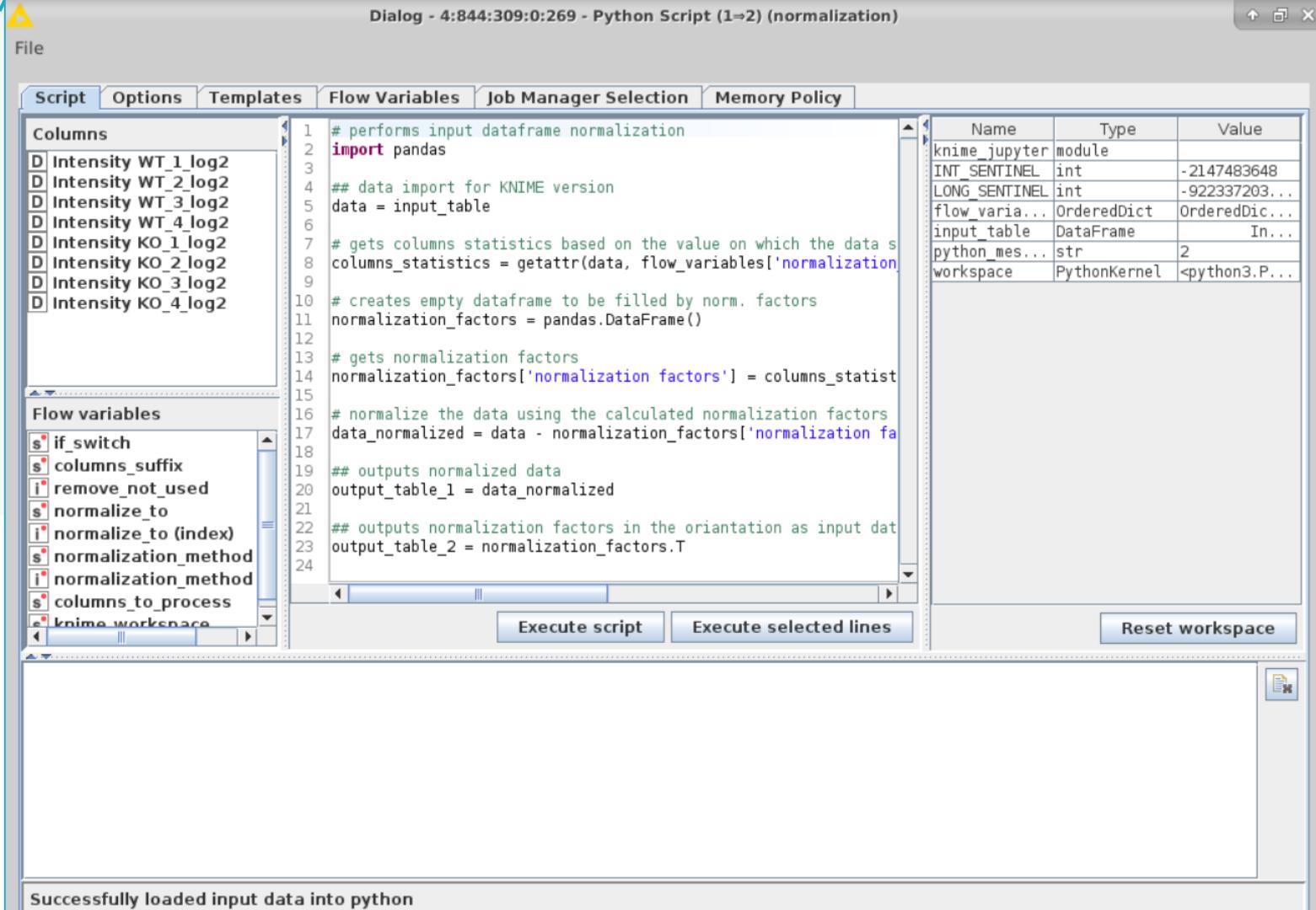
looping support



switches support



Scripting using snippets



The screenshot shows the KNIME Python Script dialog (version 4.8.4) for a node named "normalization". The dialog has tabs for Script, Options, Templates, Flow Variables, Job Manager Selection, and Memory Policy. The Script tab is active, displaying the following Python code:

```
# performs input dataframe normalization
import pandas

## data import for KNIME version
data = input_table

# gets columns statistics based on the value on which the data is
columns_statistics = getattr(data, flow_variables['normalization'])

# creates empty dataframe to be filled by norm. factors
normalization_factors = pandas.DataFrame()

# gets normalization factors
normalization_factors['normalization factors'] = columns_statist

# normalize the data using the calculated normalization factors
data_normalized = data - normalization_factors['normalization fa

## outputs normalized data
output_table_1 = data_normalized

## outputs normalization factors in the orientation as input dat
output_table_2 = normalization_factors.T
```

The Flow Variables panel on the left lists variables used in the script, such as if_switch, columns_suffix, remove_not_used, normalize_to, normalize_to_index, normalization_method, normalization_method, columns_to_process, and knime_workspace.

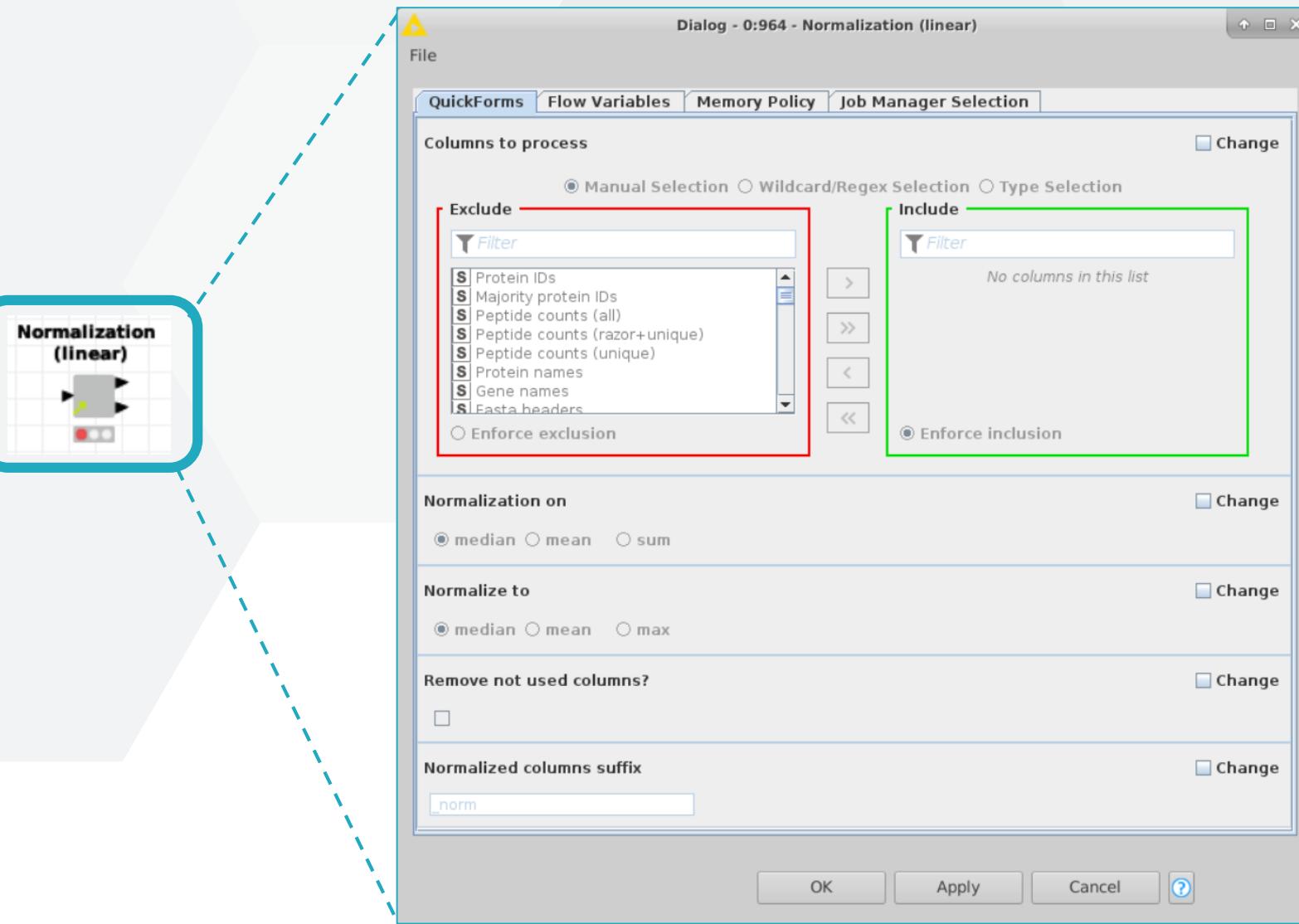
The Workspace panel on the right shows the current state of variables:

Name	Type	Value
knime_jupyter	module	
INT_SENTINEL	int	-2147483648
LONG_SENTINEL	int	-922337203...
flow_varia...	OrderedDict	OrderedDic...
input_table	DataFrame	In...
python_mes...	str	2
workspace	PythonKernel	<python3.P...

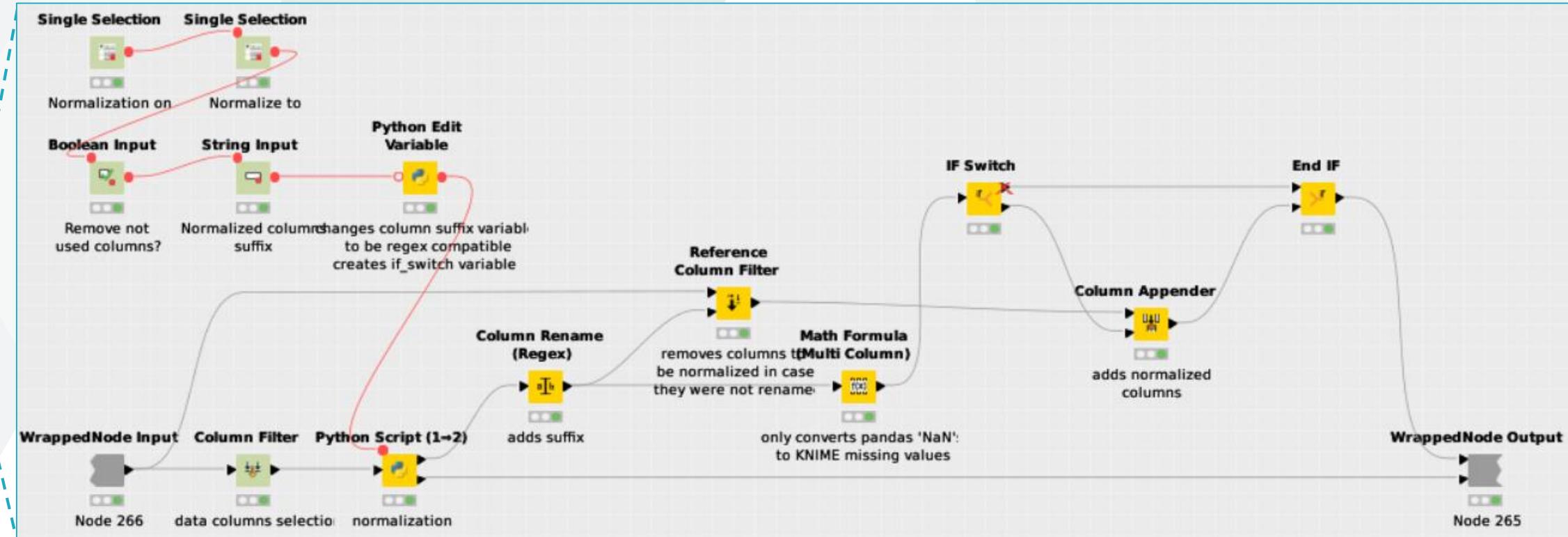
At the bottom of the dialog, there are buttons for Execute script, Execute selected lines, and Reset workspace. A message at the bottom says "Successfully loaded input data into python".

- R
- Python
- Java

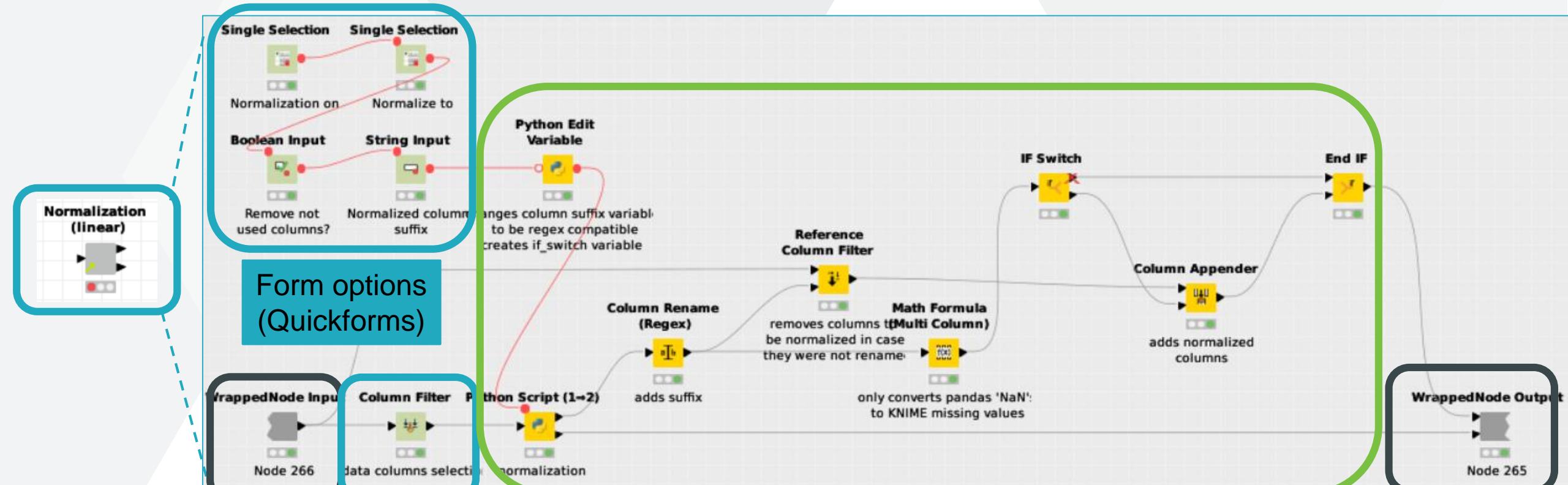
Visual programming concepts in KNIME



Visual programming concepts in KNIME



Visual programming concepts in KNIME



Scripting:

- Native KNIME nodes
- R
- Python
- Java

Visual programming concepts in KNIME

The image shows a KNIME workflow and its corresponding configuration dialog.

Workflow Diagram:

- Normalization (linear) Node (Node 266):** A node highlighted with a blue rounded rectangle. It has a "WrappedNode Input" port and a "Normalized column suffix" port.
- Single Selection Nodes:** Two nodes labeled "Single Selection". The first has a "Normalization on" output port connected to the "Normalized column suffix" port of the "Normalization (linear)" node. The second has a "Normalize to" output port connected to the "Normalized column suffix" port of the "Normalization (linear)" node.
- Boolean Input and String Input Nodes:** These nodes receive inputs from the "Single Selection" nodes and feed into a "Python Edit Variable" node.
- Python Script (1→2) Node:** This node receives inputs from the "Boolean Input" and "String Input" nodes. It has a "normalization" output port connected to the "Normalized column suffix" port of the "Normalization (linear)" node.
- Column Rename (Regex) Node:** This node receives input from the "Python Script (1→2)" node and adds a suffix to the columns.
- Column Filter Node:** This node is part of a section labeled "data columns selection". It has an "adds suffix" port connected to the "Normalized column suffix" port of the "Normalization (linear)" node.

Dialog - 0:964:0:258 - Single Selection (Normalization on):

Job Manager Selection Tab:

- Label:** Normalization on
- Description:** on what value within each data column the normalization should be done
- Variable Name:** normalization_method
- Parameter Name:** single-selection
- Selection Type:** Radio buttons (horizontal)
- Possible Choices:** median, mean, sum
- Default Value:** median
- Limit number of visible options:**
- Number of visible options:** 5

Control Tab: This tab is currently selected.

Flow Variables Tab: This tab is also present in the dialog.

Specification of settings for particular form options: A green bar at the bottom of the dialog.

Visual programming concepts in KNIME

The image shows a KNIME workflow and a configuration dialog box.

Workflow Diagram:

- Normalization (linear) Node:** A node highlighted with a blue border. It has an output port labeled "Normalized column suffix".
- Single Selection Nodes:** Two nodes highlighted with a green border. The first is labeled "Normalization on" and the second is labeled "Normalize to". They receive input from the "Normalized column suffix" port of the "Normalization (linear)" node.
- Boolean Input and String Input Nodes:** These nodes receive input from the "Single Selection" nodes.
- Python Edit Variable Node:** This node receives input from the "String Input" node and outputs to a "Column Rename (Regex)" node.
- Column Rename (Regex) Node:** This node receives input from the "Python Edit Variable" node and outputs to a "Python Script (1→2)" node.
- Python Script (1→2) Node:** This node receives input from the "Column Filter" node and outputs to the "Column Rename (Regex)" node.
- Column Filter Node:** This node receives input from the "WrappedNode Input" node and outputs to the "Python Script (1→2)" node.
- WrappedNode Input Node:** This node receives input from the "Node 266" node and outputs to the "Column Filter" node.

Dialog - 0:964:0:258 - Single Selection (Normalization on):

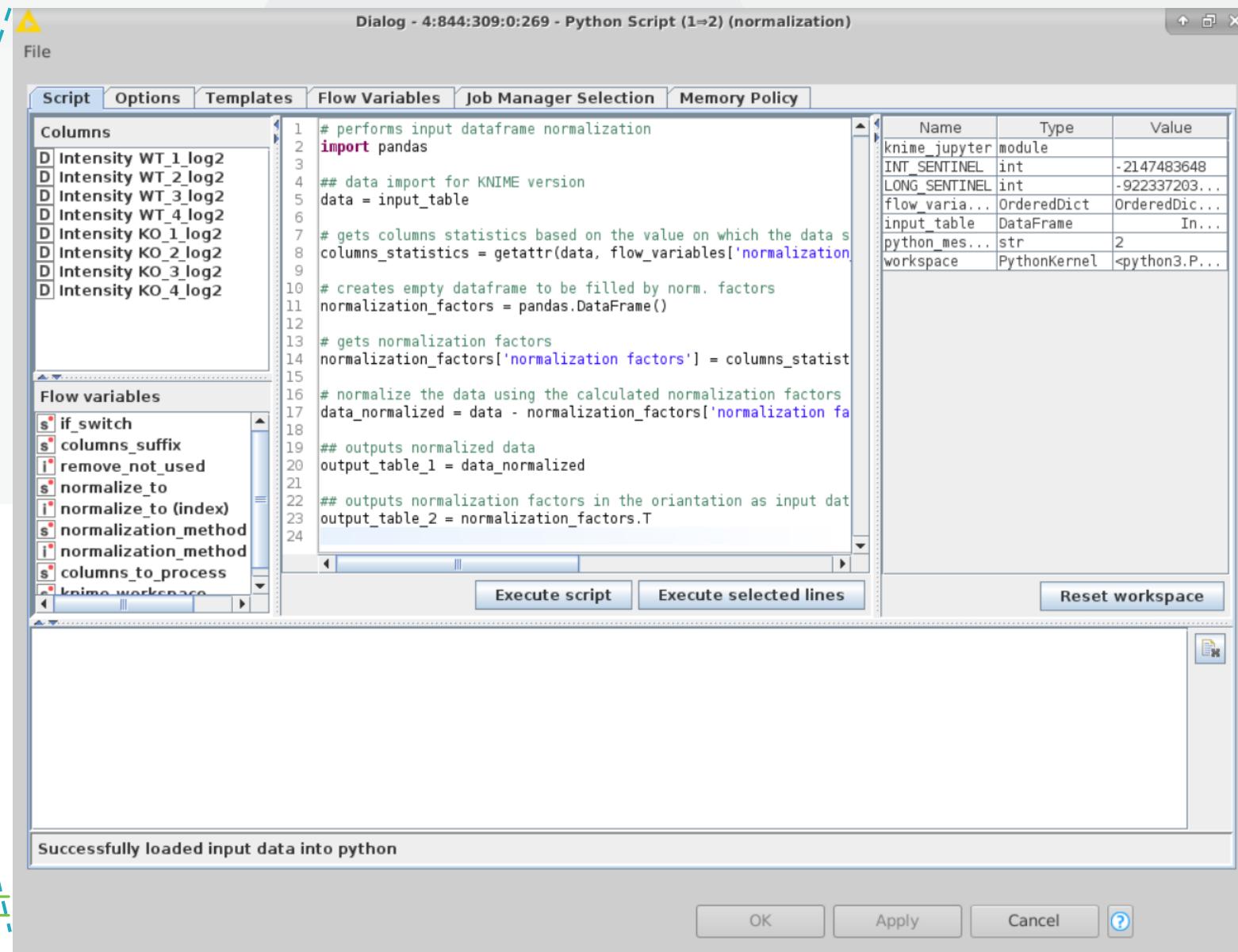
Job Manager Selection Tab:

- Label:** Normalization on
- Description:** on what value within each data column the normalization should be done
- Variable Name:** normalization_method
- Parameter Name:** single-selection
- Selection Type:** Radio buttons (horizontal)
- Possible Choices:** median, mean, sum
- Default Value:** median
- Limit number of visible options:**
- Number of visible options:** 5

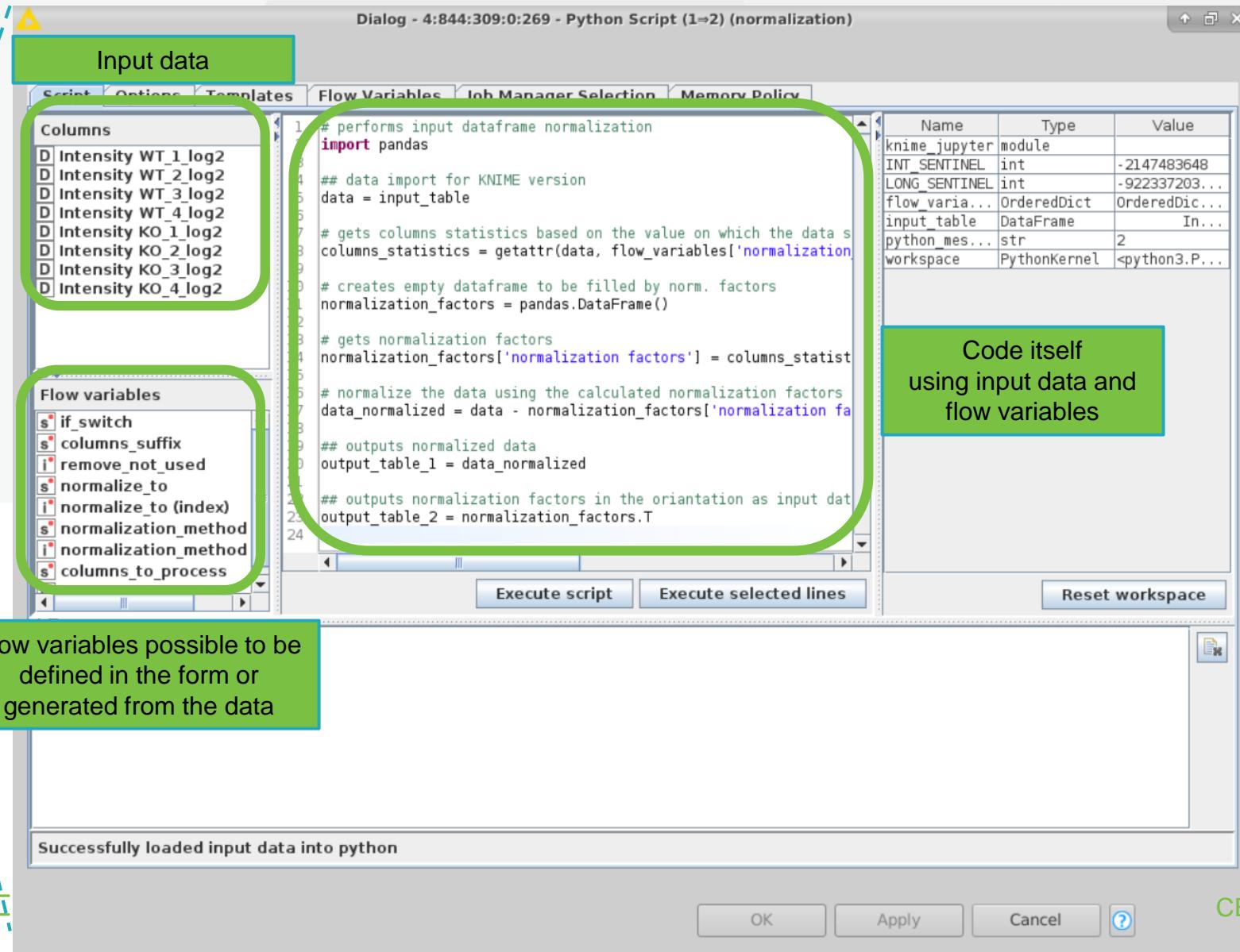
OK **Apply**

**Specification of settings for particular form options
Variables can be further used for scripting**

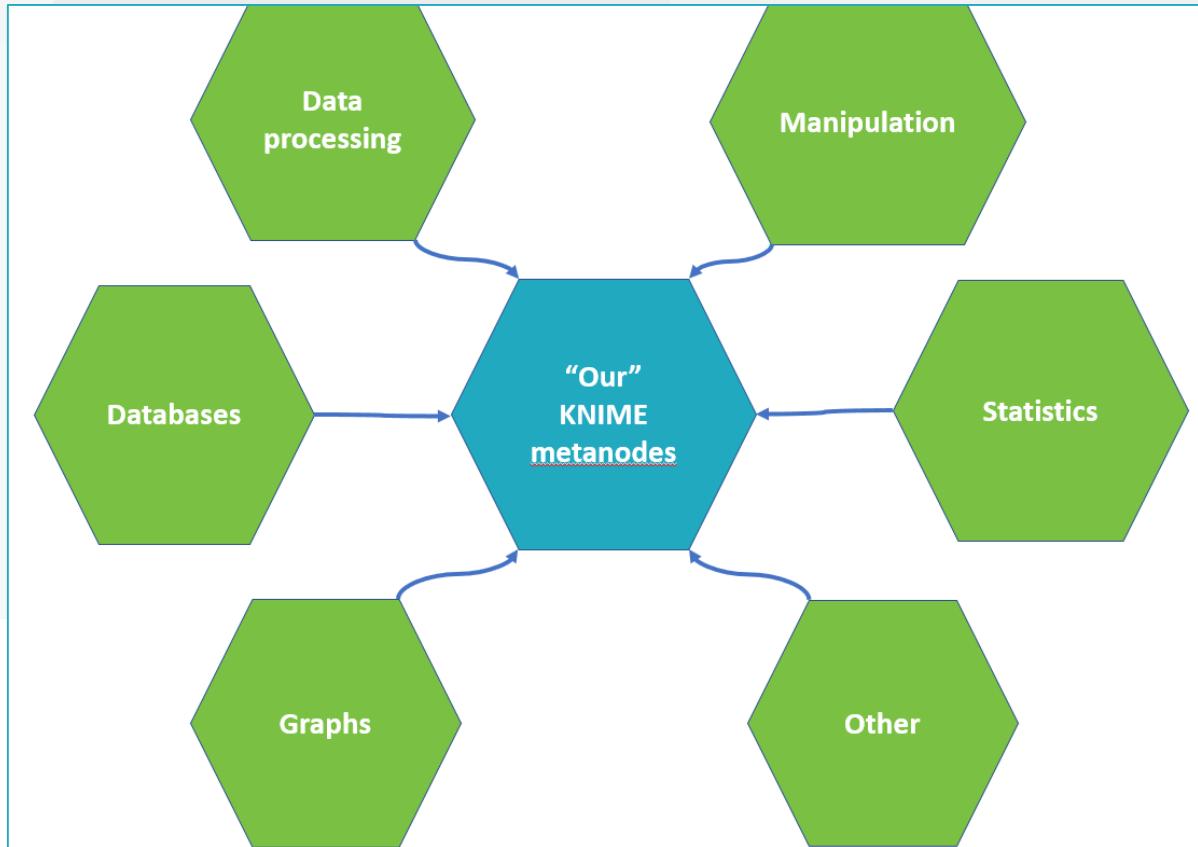
When you just need or want to script...



When you just need or want to script...

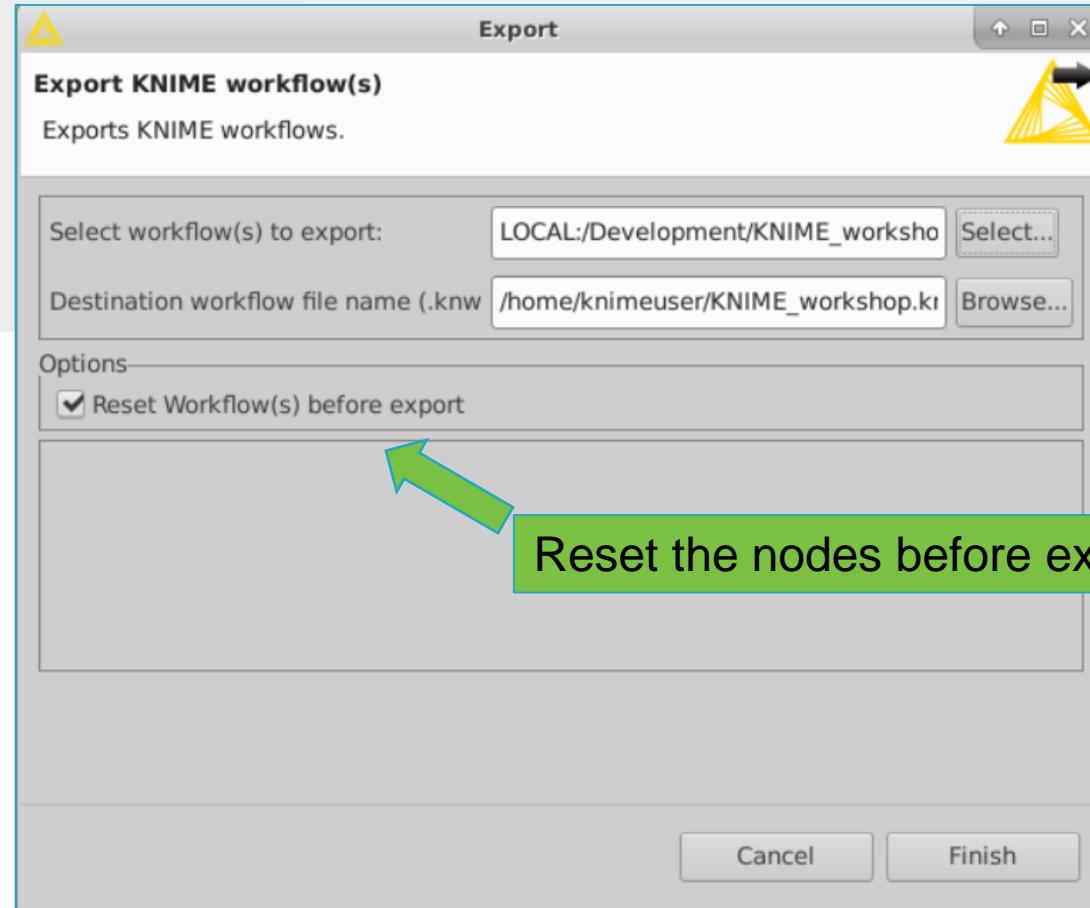
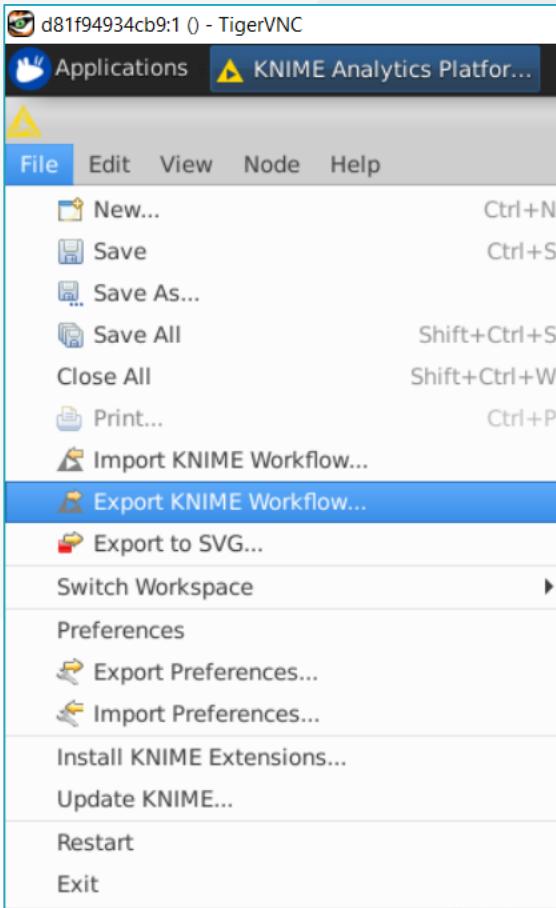


Metanodes sharing



Saving metanode as a template for sharing it with your colleagues or re-using in your workflows

Workflows sharing



Exporting workflows for easy sharing with your colleagues



Thank you
for your
attention

Workshop outline

- morning session – theoretical part

- 10:00 – 10:15 Opening and introduction
- 10:15 – 10:45 Software container running KNIME
- 10:45 – 11:00 Coffee break
- 11:10 – 11:30 Introduction to KNIME
- **11:30 – 11:45** **Coffee break**
- 11:45 – 13:00 Practical applications, our KNIME metanodes
- 13:00 – 14:00 Lunch break, visit of our laboratories for interested people