# KNIME SOP

This SOP is intended as a general guide to help beginners get started with the software. It details: the installation process of KNIME and basic extensions; setting up the workspace in logical manner; how to import and use existing workflows (e.g. a bespoke PAINS filter that can be modified to suit the needs of a project); and finally general and more advanced knowledge on how to use KNIME.

Downloading the software

1. Visit the KNIME website (https://www.knime.com/)
2. Click the download icon in the top right corner:
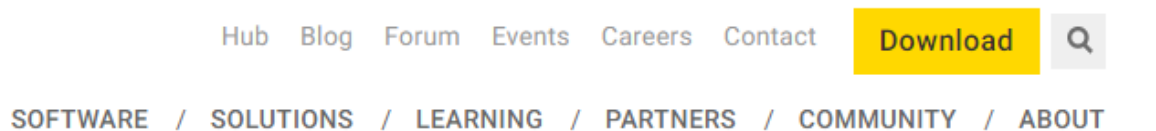
    a. Fill out registration details (only email is compulsory) and then click submit

    b. Download appropriate version. For windows select the 64 bit installer:

    c. Accept and download (takes a few minutes):

Home > Downloads

You decided to download the installer for KNIME Analytics Platform for Windows (installer) 64 Bit (441.03 MB).

The installer adds an icon to the desktop and suggests suitable memory settings

If you want to run the KNIME installer or self-extracting archive for Windows you might experience some difficulty because of the Microsoft SmartScreen filter which was introduced with Internet Explorer 9 and Windows 8. Find out how to solve the problem.

☐ I have read and accept the privacy policy and the terms and conditions *
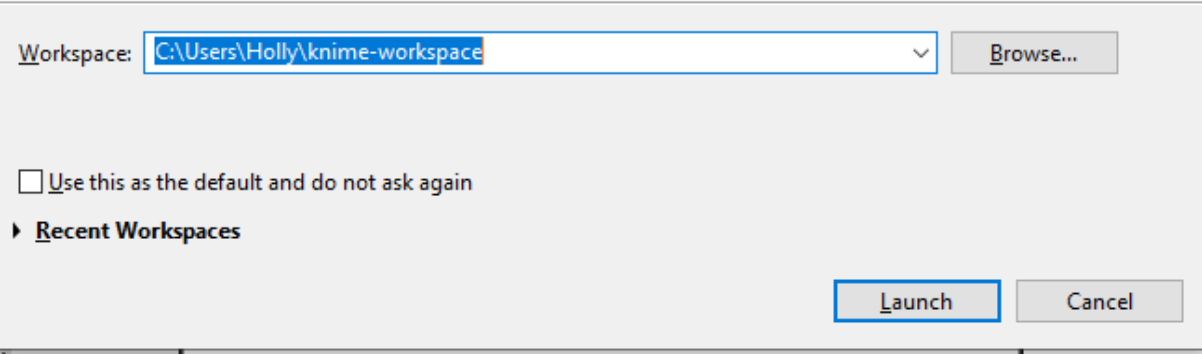
⊕ Download

       d.   Allow app to make changes to your device
       e.   Accept the agreement
       f.   Choose location to install KNIME (click next)
       g.   Start menu folder – either click next or choose not to have this
       h.   Additional tasks – make sure file extensions and URLs are all ticked
       i.   Proceed with suggested memory space
       j.   Install
       k.   Finish

3. Launch KNIME (either from desktop icon if you selected that or type 'KNIME' in the computer search bar)
   a.   The platform launcher should pop up, and click launch:

Tidying the workspace for efficient working

Once set up these modifications should be retained every time KNIME is launched. This is not the only way to configure the panel set up, but I find it most useful as gives the most space for the workflow.
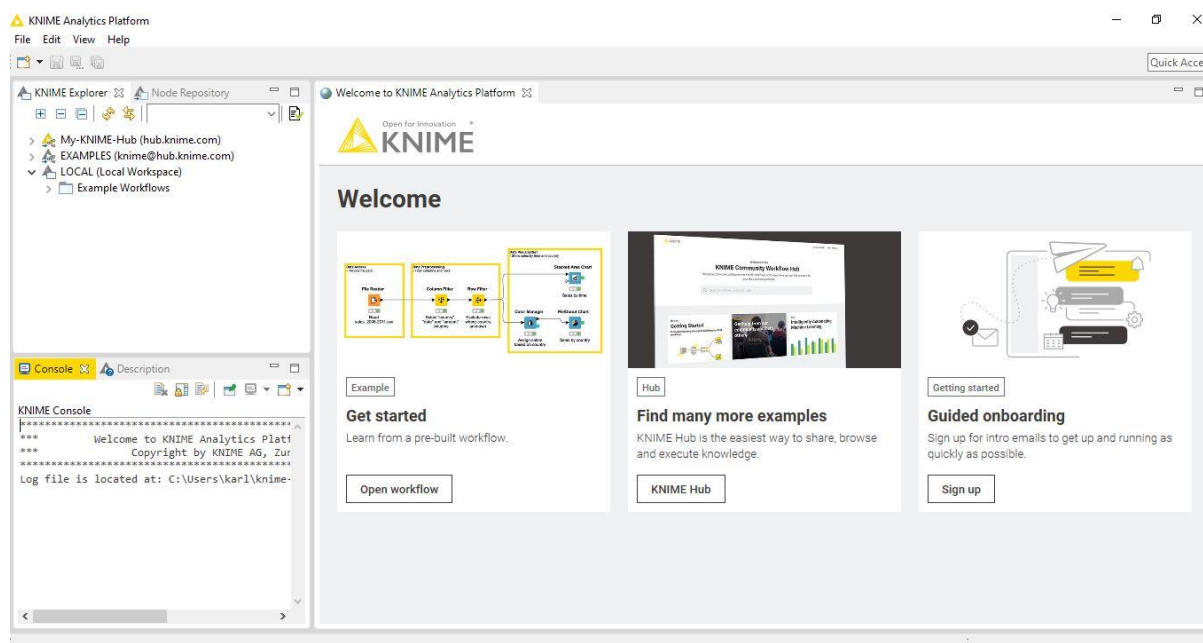
1. Close workflow coach (use the cross):
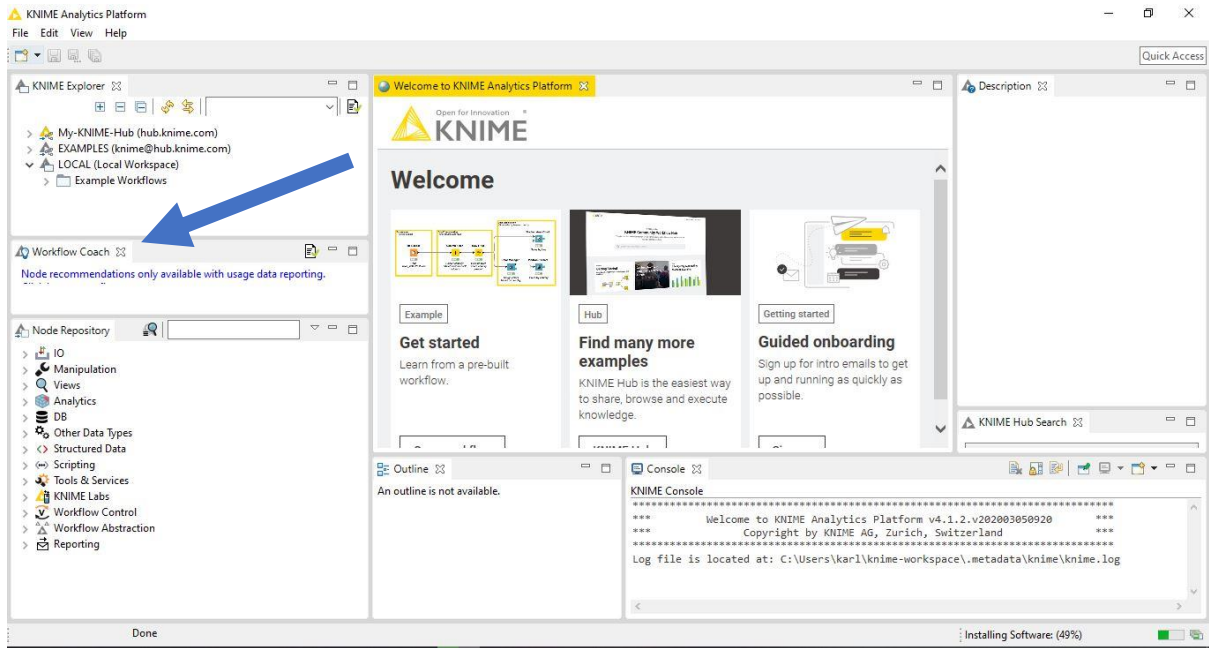
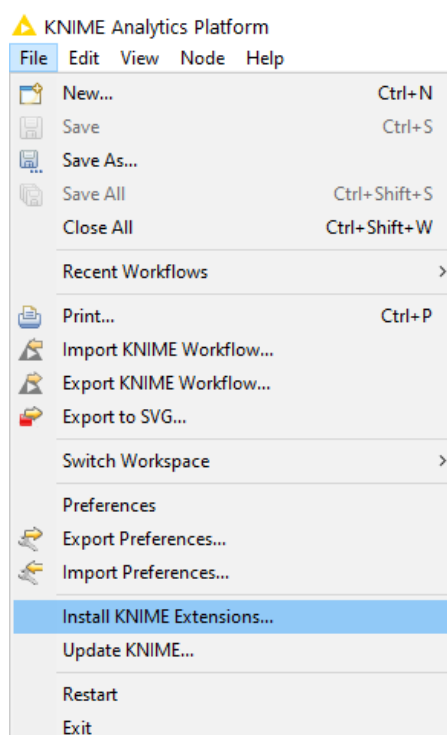2. Drag Node Repository next to KNIME explorer:
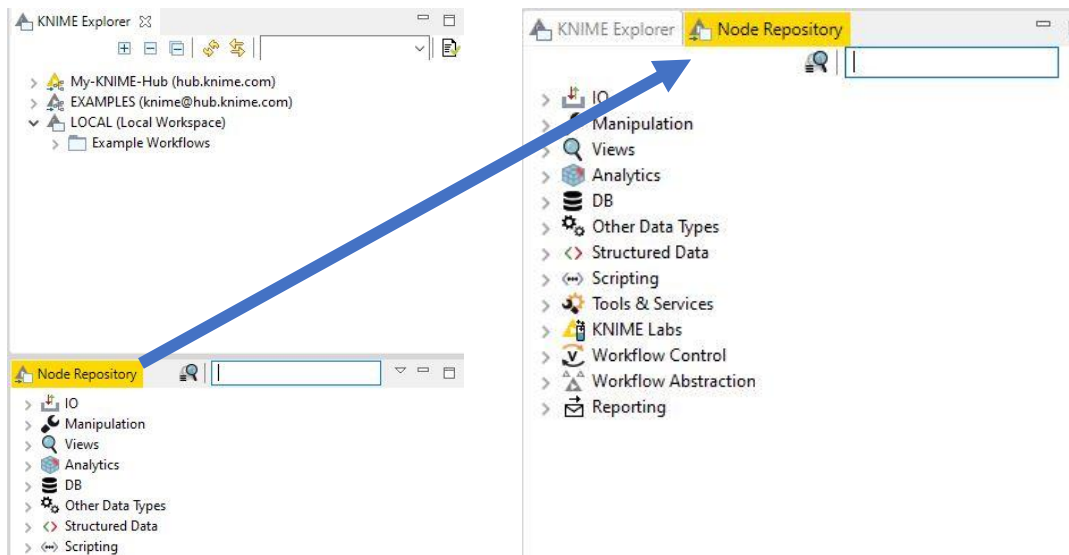
3. Close KNIME hub search
4. Close Outline
5. Drag 'Console' and 'Description' to bottom left of the screen

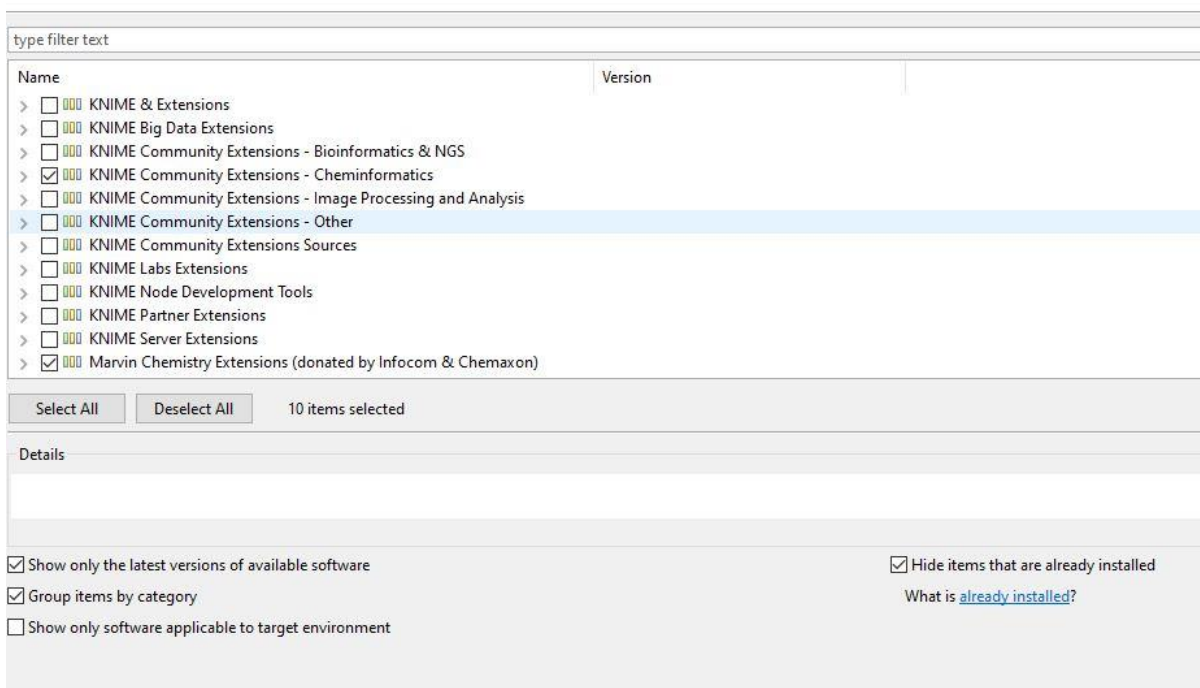Window should look like this upon completion:

Installing chemistry extensions

1. File > Install KNIME Extensions

2. Select extensions of interest using the tick boxes. The cheminformatics and Marvin chemistry extensions are recommended as a minimum:

3. Click next and then next, accept agreements, finish
4. Close KNIME and restart for changes to take effect
5. Other extensions can be added by clicking Help > Install New Software:
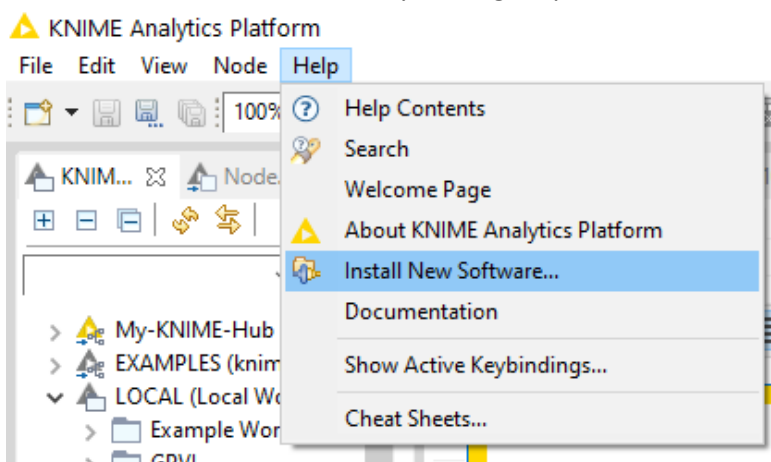
6. Type in web address for new extension in the 'work with' section then click 'add'

⚠ Install

Available Software

Select a site or enter the location of a site.

Work with: | type or select a site | ▼ | Add... | Manage... |

type filter text

Name	Version
☐ ⓘ There is no site selected.	

Select All   Deselect All

Details

☑ Show only the latest versions of available software   ☑ Hide items that are already installed

☑ Group items by category   What is _already installed_?

☑ Show only software applicable to target environment

☑ Contact all update sites during install to find required software

< Back   Next >   Finish   Cancel

Basic KNIME knowledge

- Headed arrows show the data inputs/outputs from nodes:
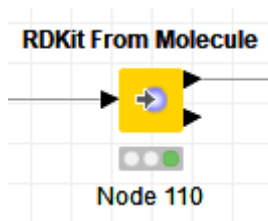
- Running nodes: right click > execute or execute and open views to see the output of this node. Alternatively select then hold fn and f7.
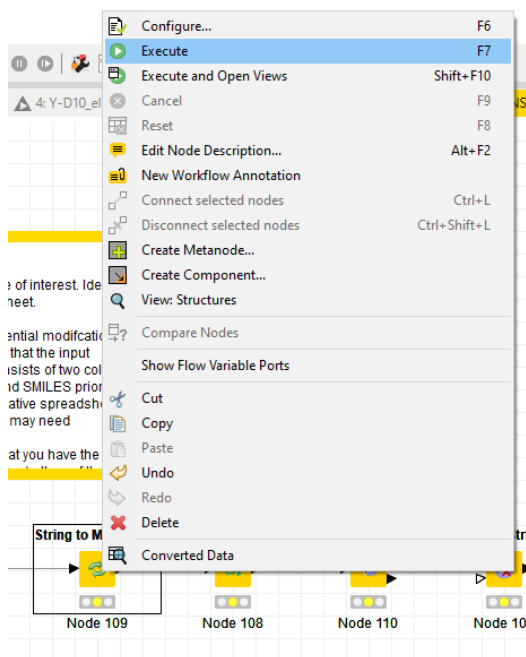
- Viewing node outputs: right click and then select the bottom option that has a spreadsheet and magnifying glass
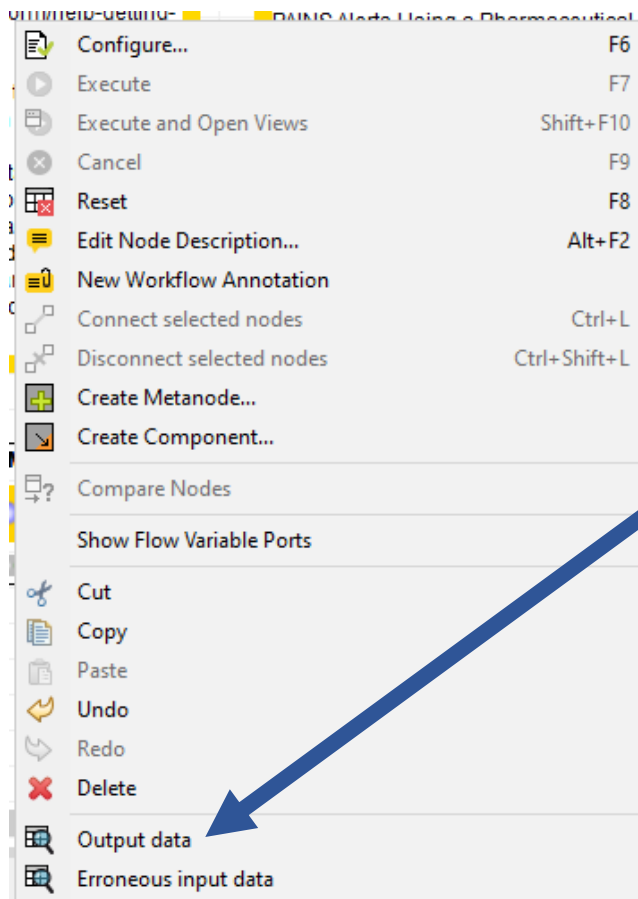
- Searching for nodes: Navigate to the node repository and type a key word in the search bar (e.g. filter), or navigate manually (e.g. to find column filters: manipulation > column > filter)
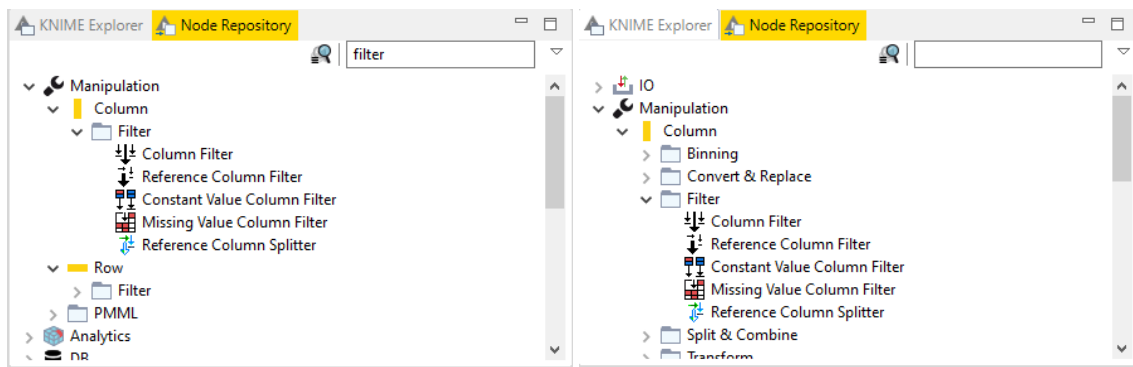
- Connecting nodes: Drag inputs/outputs from one node to another
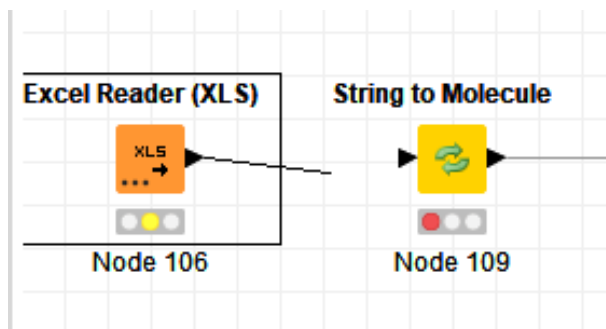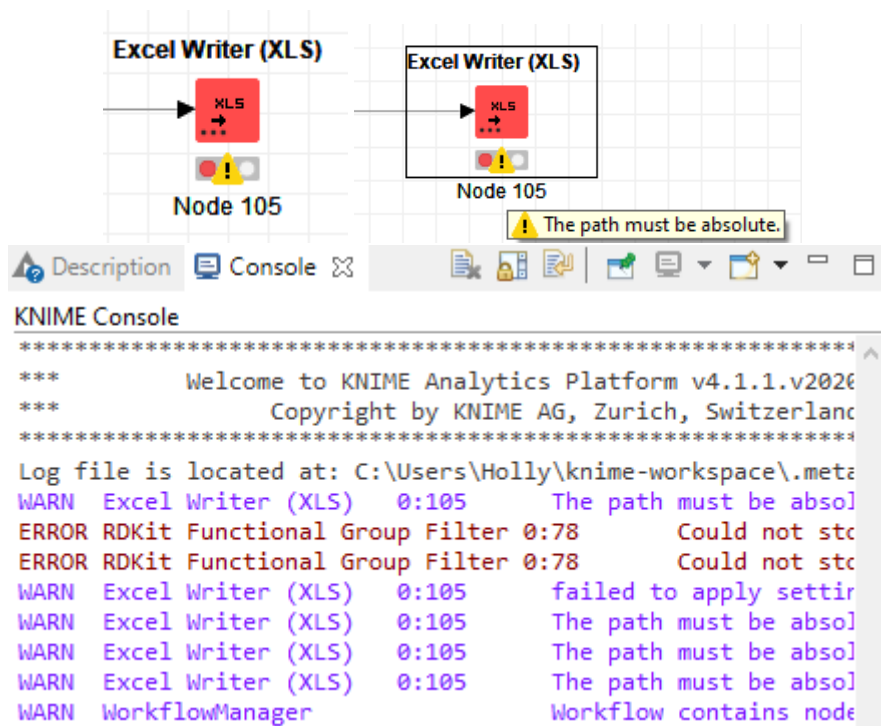
- Node traffic lights:

Understanding the traffic light system:

🔴⚪⚪ **Not configured:** Node is not yet configured and cannot be executed with its current settings

⚪🟡⚪ **Configured:** Node has been correctly configured and may be executed at any time

⚪⚪🟢 **Executed:** Node has been successfully executed and results can be viewed and used in downstream nodes
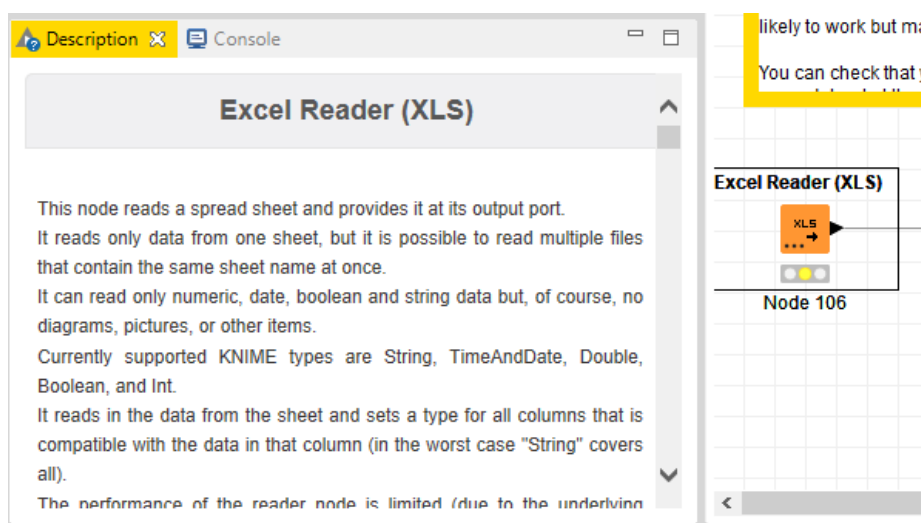
- Node errors are denoted with a warning triangle. If you hover over this it will detail the issue. The console may offer further description of the problem.

Excel Writer (XLS)

Excel Writer (XLS)

Node 105

Node 105

⚠ The path must be absolute.

🔎 Description 🖥 Console ⊠

KNIME Console

```
**********************************************************
***         Welcome to KNIME Analytics Platform v4.1.1.v2020
***              Copyright by KNIME AG, Zurich, Switzerland
**********************************************************
Log file is located at: C:\Users\Holly\knime-workspace\.meta
WARN   Excel Writer (XLS)    0:105       The path must be absol
ERROR  RDKit Functional Group Filter 0:78       Could not sto
ERROR  RDKit Functional Group Filter 0:78       Could not sto
WARN   Excel Writer (XLS)    0:105       failed to apply settin
WARN   Excel Writer (XLS)    0:105       The path must be absol
WARN   Excel Writer (XLS)    0:105       The path must be absol
WARN   Excel Writer (XLS)    0:105       The path must be absol
WARN   WorkflowManager                  Workflow contains node
```

- Node descriptions: Click on a node and select the 'Description' tab to understand the function of the node:

🔎 Description ⊠ 🖥 Console

Excel Reader (XLS)

This node reads a spread sheet and provides it at its output port. It reads only data from one sheet, but it is possible to read multiple files that contain the same sheet name at once.

It can read only numeric, date, boolean and string data but, of course, no diagrams, pictures, or other items.

Currently supported KNIME types are String, TimeAndDate, Double, Boolean, and Int.

It reads in the data from the sheet and sets a type for all columns that is compatible with the data in that column (in the worst case "String" covers all).

The performance of the reader node is limited (due to the underlying

likely to work but ma

You can check that y

Excel Reader (XLS)
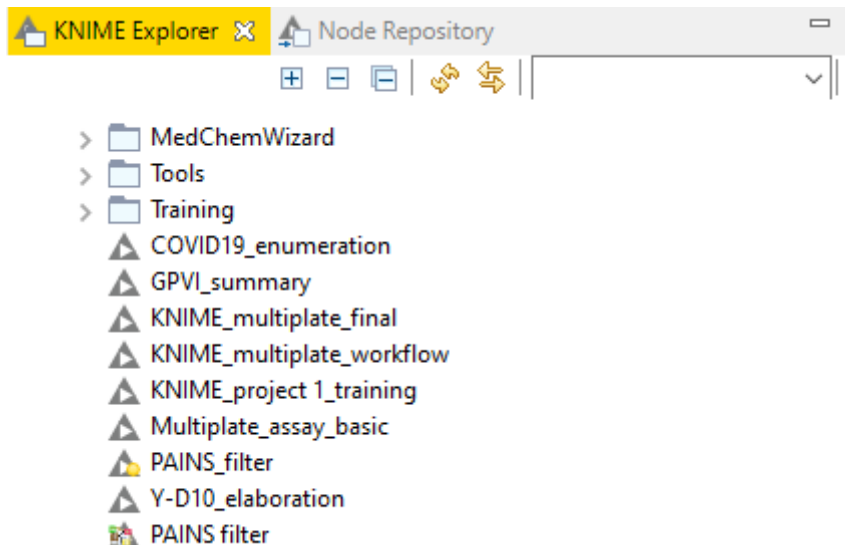
Node 106

Importing existing workflows

1. File > Import KNIME workflow
2. Browse to select workflow of interest > open > finish
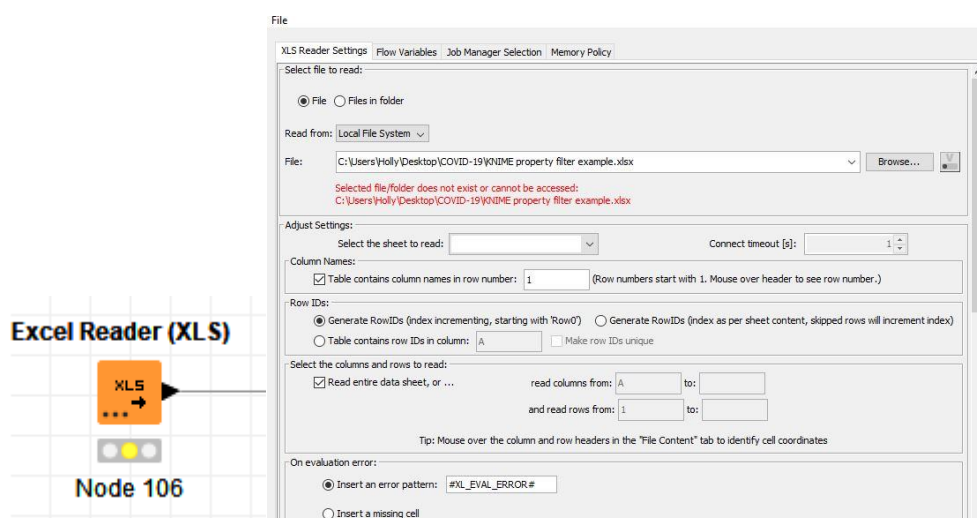3. Double click workflow on KNIME explorer

Using the PAINS filter

1. Import workflow:
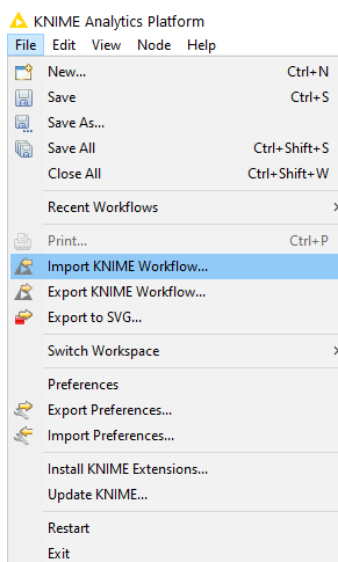

2. Double click workflow link in the 'KNIME explorer':

3. Double click on Excel reader node:

a. 'Browse' to the spreadsheet of interest
b. If you scroll down you can see a preview of what the spreadsheet will look like. Make sure this looks reasonable. Ideally the input spreadsheet should have one sheet consisting of the compound ID and SMILES.
4. Double click on the 'RDkit functional group filter' and select the functional groups to REMOVE then click okay. The pre-selected functional groups are detailed below the node.
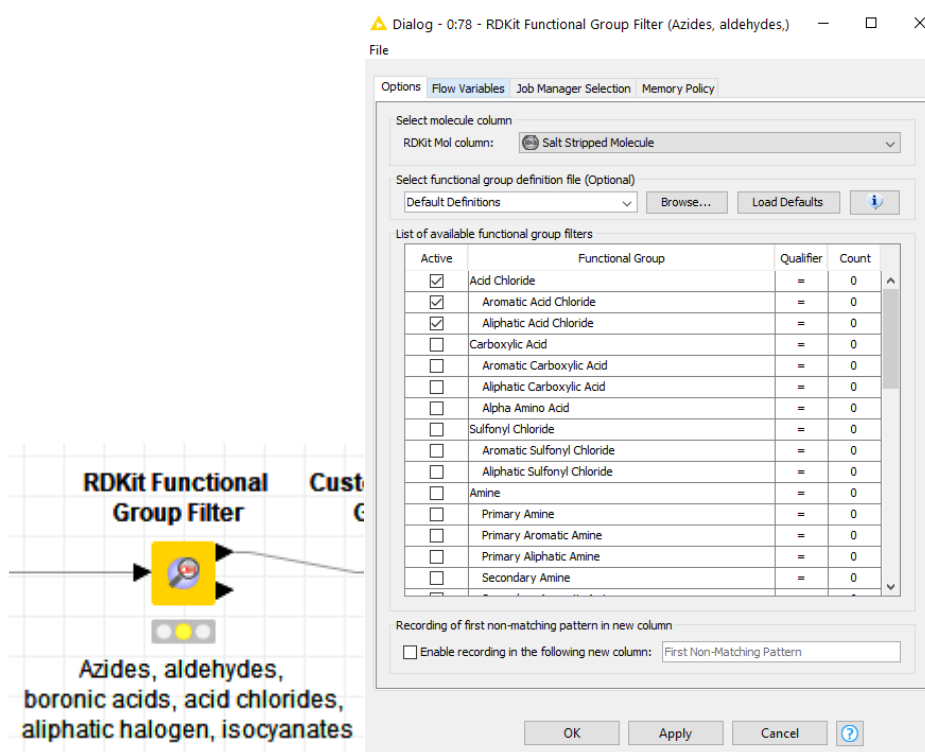
5. Double click the 'custom functional group filter' and select other functional groups that you wish to remove then click okay:
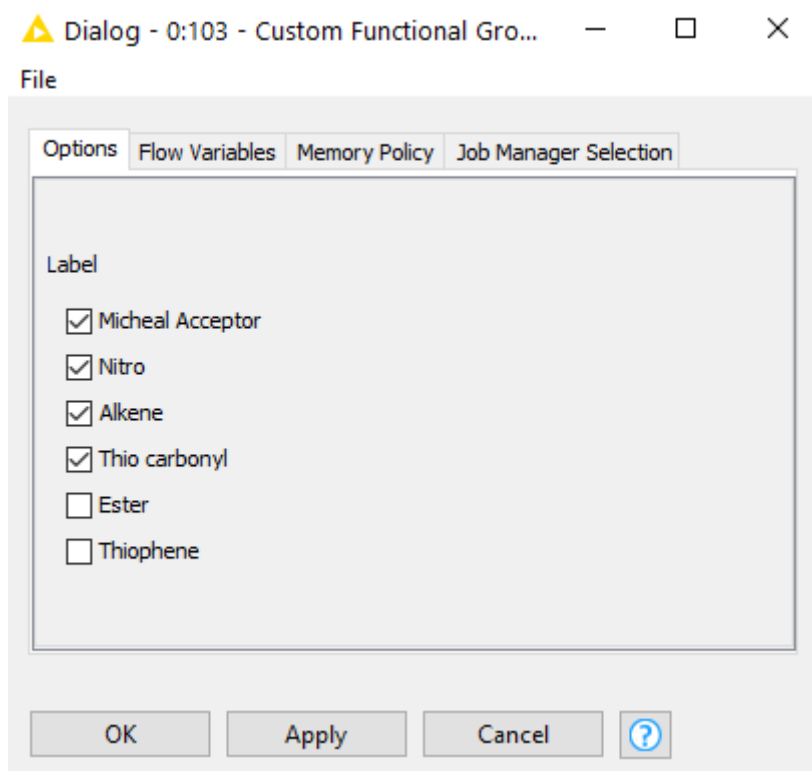
6. Double click on the property filter and set the parameters then click okay:

7. Right click on the property filter and click 'execute'
   a. Alternatively can select the node and then press fn > f7
   b. The traffic light underneath the node should now be green
8. Double click on the Excel writer, and specify the save location, click okay then run this node
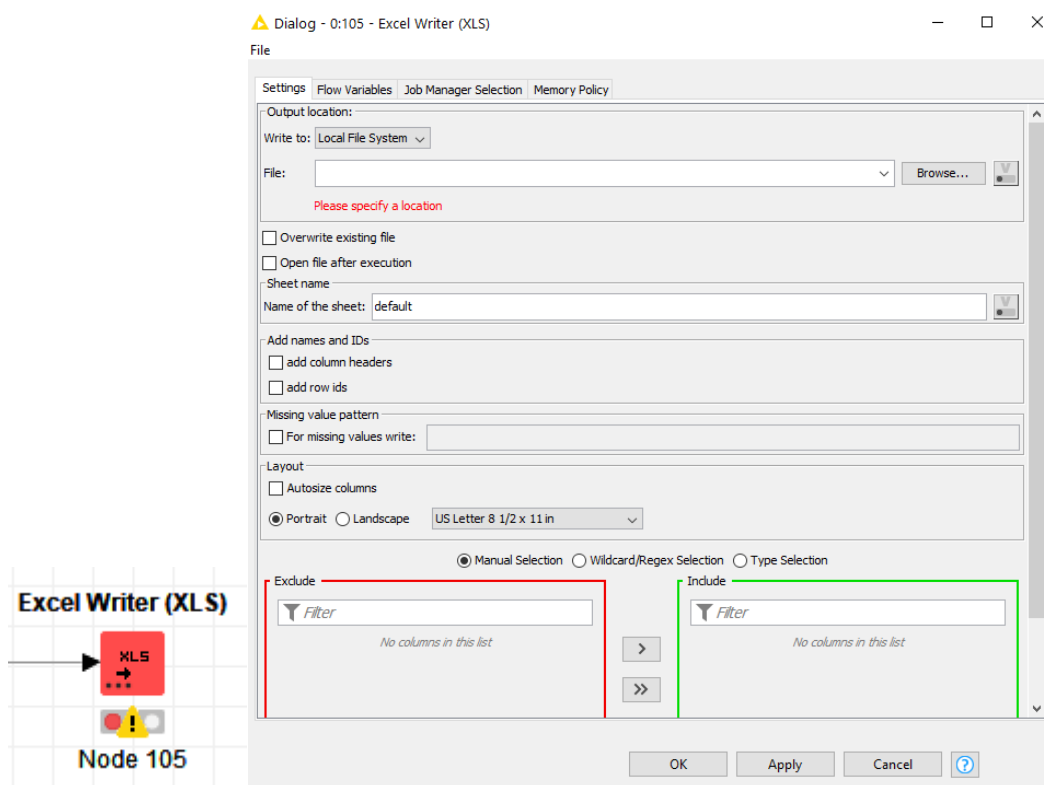
Advanced KNIME Knowledge

Metanodes

Components (GUIs)

Flow Variables

Useful Chemistry KNIME Knowledge

Glossary

Node	A node is the smallest programming unit in KNIME. Each node serves a dedicated task, from very simple ones - like changing the name of a data column - to very complex one - like training a random forest.

Training Resources

KNIME offers a number of **FREE** online training courses: https://www.knime.com/knime-introductory-course

KNIME also has three short one page 'cheat sheets' which are VERY helpful for new users: https://www.knime.com/learning/cheatsheets . See below for a sneak preview:

Version History