KNIME

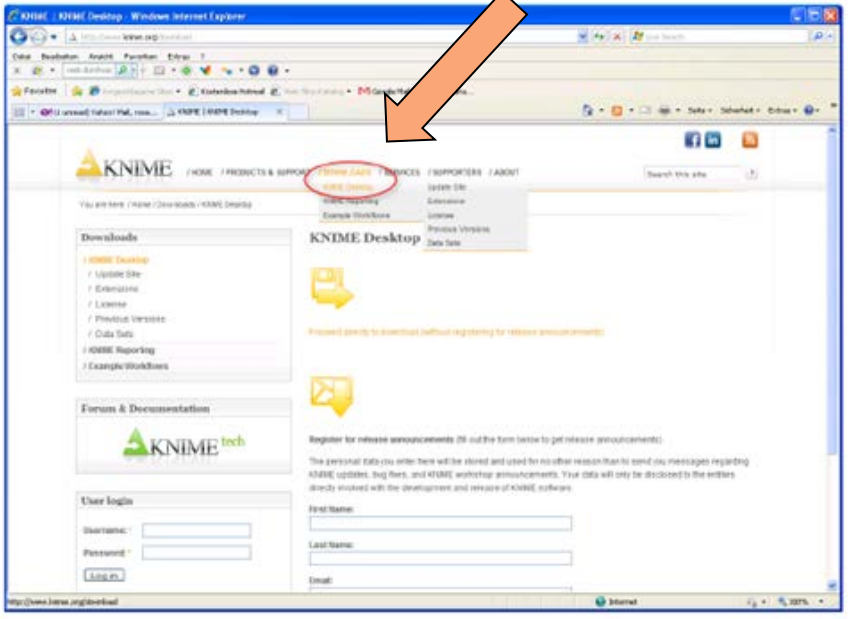
- Modern data analytics platform that enables sophisticated statistical and data mining analysis on small, medium or large data
- Visual workbench chains data access, data transformation, initial investigation, powerful predictive analytics and visualization
- Provides the ability to develop reports or automate the deployment of new insights back into production systems
- KNIME Analytics Platform is open source and available under GPL license
- It can be extended with Commercial Software to include professional support, productivity and collaboration functionality

KNIME Website Links

Web Links	
http://www.knime.org	This is the first place to look for information about KNIME products. The open source Desktop version can be downloaded here.
http://www.knime.org/learning-hub	This is the landing page to learn more specific KNIME functionalities. Learning material - as web sites, videos, webinars, courses, and more - is organized by topic, like text mining or chemistry, or basic KNIME functionalities, etc...
http://tech.knime.org/forum	In the www.knime.org site you can find a number of resources. What I find particularly useful is the KNIME Forum. Here you can ask questions about how to use KNIME or about how to extend KNIME with new nodes. Someone from the KNIME community answers always and quickly.
http://tech.knime.org/knime-labs	This site contains nodes still under development; i.e. the beta version of new nodes. You can already download them and use them, but they are not of product/release quality yet.
http://knime.org/supporters-0	This is the site where all contributing supporters (partners, providers, and sponsors) are listed.

DOWNLOAD KNIME

- Go to www.knime.org
- Select the "Download" tab
- Select "KNIME Desktop" in the submenu
- Select one of the two options (with registration or without registration, as you prefer)
- Choose the version that suits your environment (Windows/Linux, 32 bit/64 bit)
- Accept the terms and conditions
- Start downloading
- You will end up with a zipped or a self-extracting archive file (*.zip or *.exe)
- Unpack your file in the destination folder
- This creates a new subfolder in the selected destination folder named "knime_ 2.X.x where x represents the minor release



KNIME Download and Installation

Installation

License

Starting with Version 2.1, KNIME is released under the GNU General Public License, Version 3 (including certain additional permissions according to Sec. 7 of the GPL Ver. 3). It is also available - through the dual licensing model - under customized licenses. If you wish to receive KNIME under a different license than the GPL, please contact us at contact@knime.org to discuss licensing arrangements.

Versions

The table describes the KNIME versi nd on which platform they are available or not.

	Linux	Windows	Mac OS X
KNIME (32bit)	yes	yes	no
KNIME (64bit)	yes	yes	yes
KNIME Developer Version (32bit)	yes	yes	no
KNIME Developer Version (64bit)	yes	yes	yes

Installation

Download one of the above versions, unzip it to any directory. For Windows click the *knime.exe* file, and for Linux click *knime* in order to start KNIME.

Welcome Screen / Additional features

When KNIME is started for the first time a welcome screen appears.

From here you can

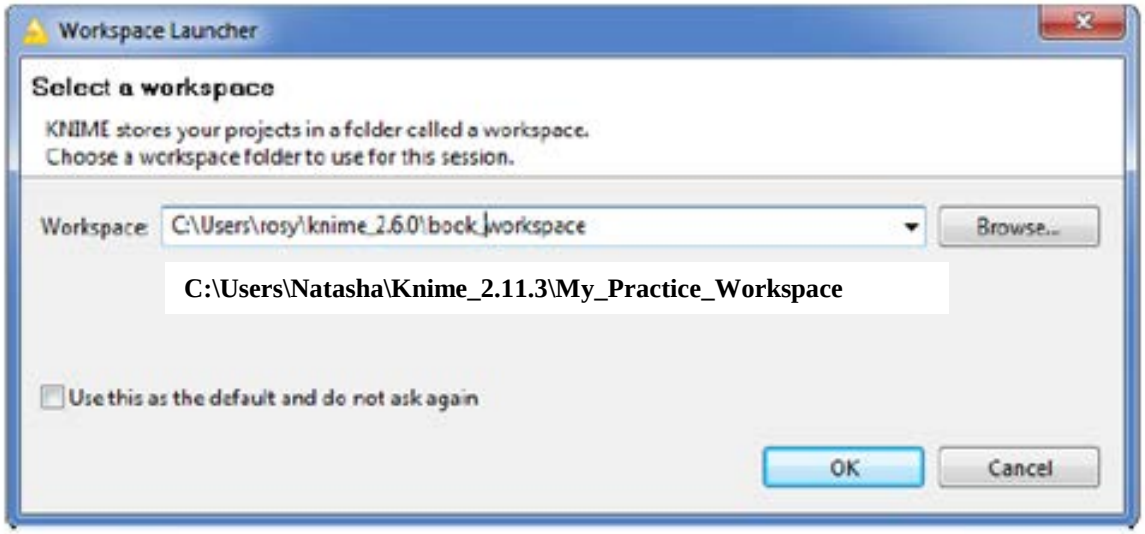
1. Open KNIME workbench: opens the KNIME workbench to immediately start exploring KNIME, build own workflows and explore your data.
2. Get additional nodes: in addition to the ready-to-start basic KNIME installation there are additional plug-ins for KNIME e.g. an R and Weka integration, or the integration of the Chemistry Development Kit with additional nodes for the processing of chemical structures, compounds, etc. You can download these features later from within KNIME (File, Update KNIME...) as well

KNIME Workspace

The Workspace is the directory where all current workflows and preferences are saved

It can be located anywhere(you specify) on the hard drive

By Default it is in the KNIME folder – but you can change that!



To install a new KNIME extension:

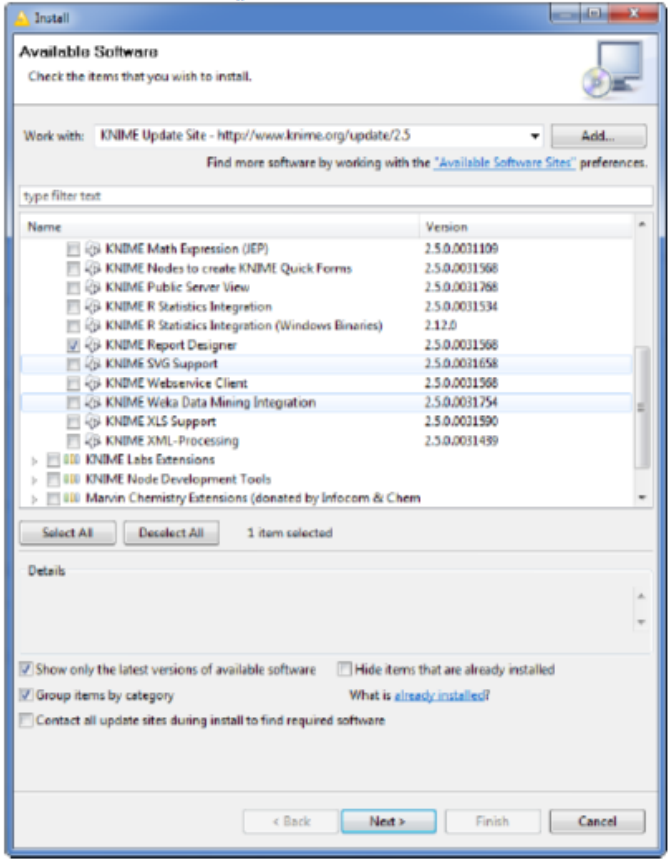
- From the Top Menu, select **"Help" -> "Install New Software"**
- In the "Available Software" window, in the "Work with" textbox, select the URL with the KNIME update site (usually named KNIME Update Site - http://www.knime.org/update/2.x)
- Select the extensions that you need. In this example, we chose to install the "KNIME Report Designer" in the category "KNIME & Extensions.
- Click the "**Next**" button on the bottom right and follow the wizard instructions

Once the KNIME extension has been installed and KNIME has been restarted, you should see a new node/category in the "Node Repository" in the KNIME workbench (see a few sections ahead).

For example, after installing the KNIME Report Designer, you should see a category "Reporting" in the "Node Repository" panel on the left.

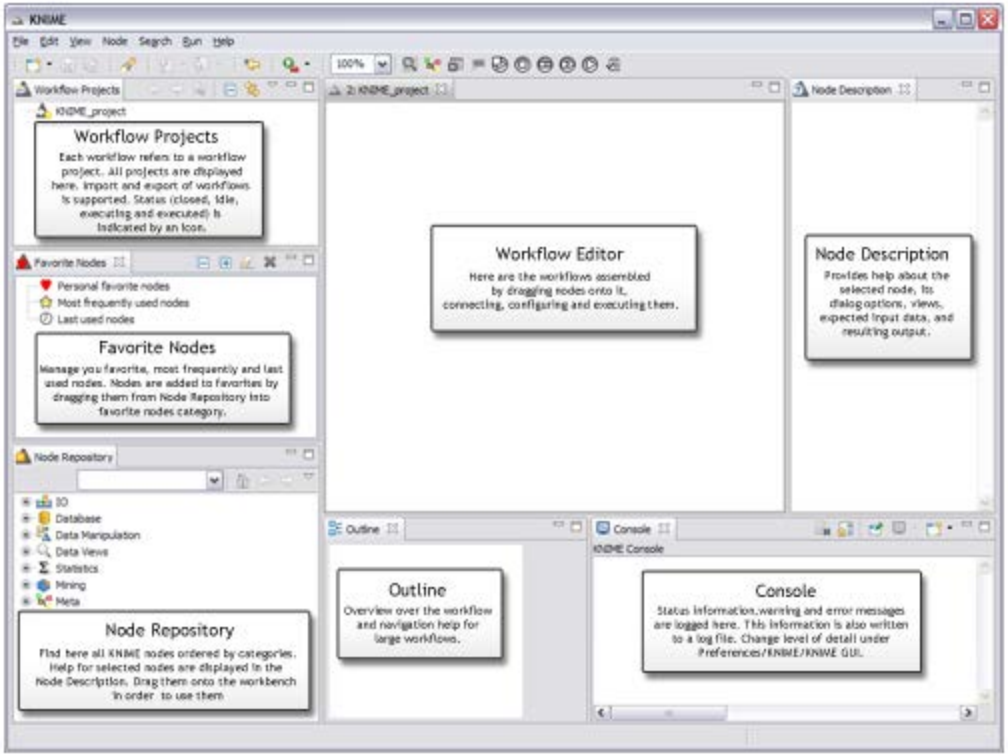
Another way to reach the "Available Software" window and then install a new KNIME extension is the following: from the Top Menu, select **"File" -> "Install KNIME Extensions".**

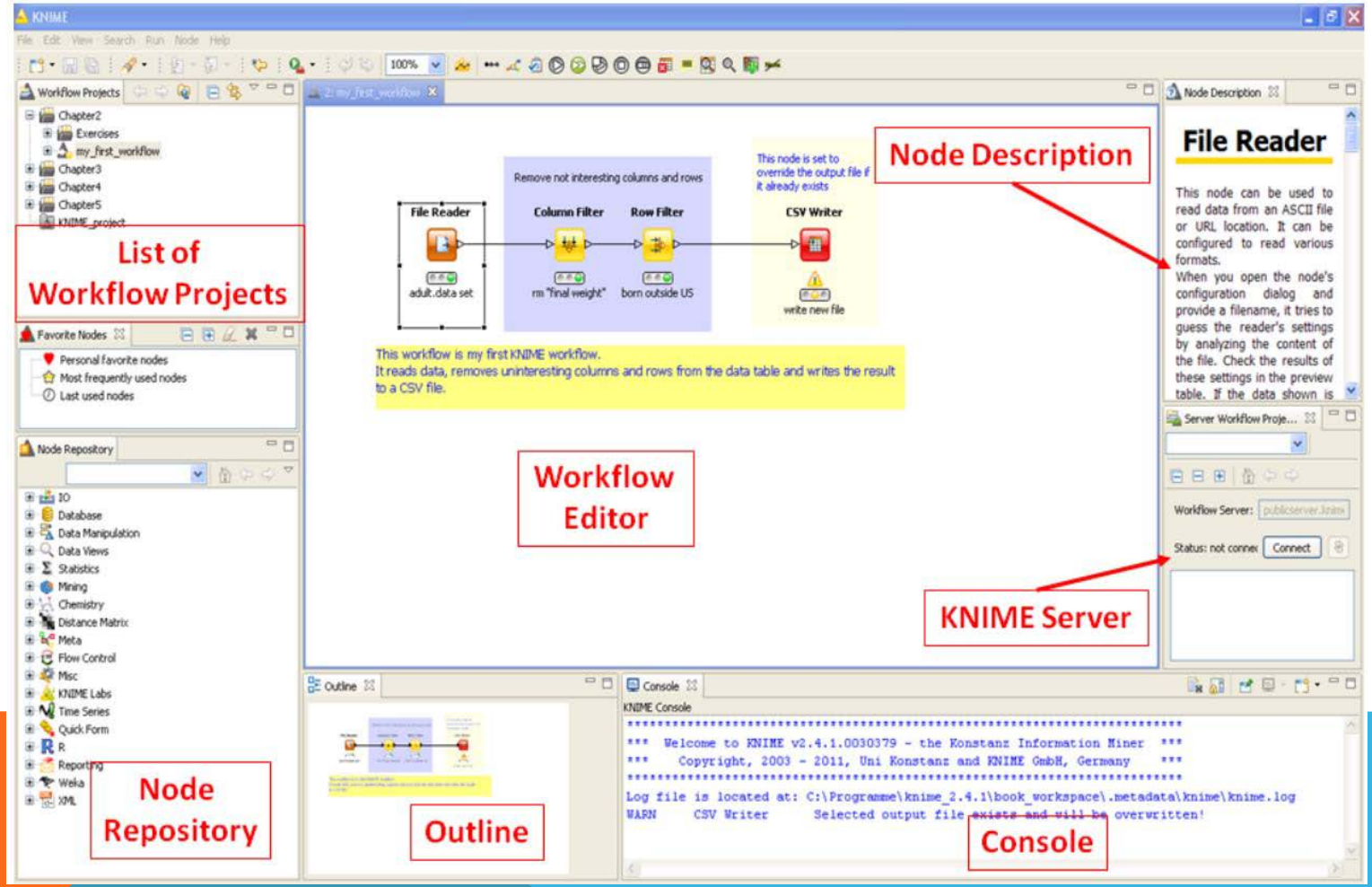
1.3. The „Available Software" window

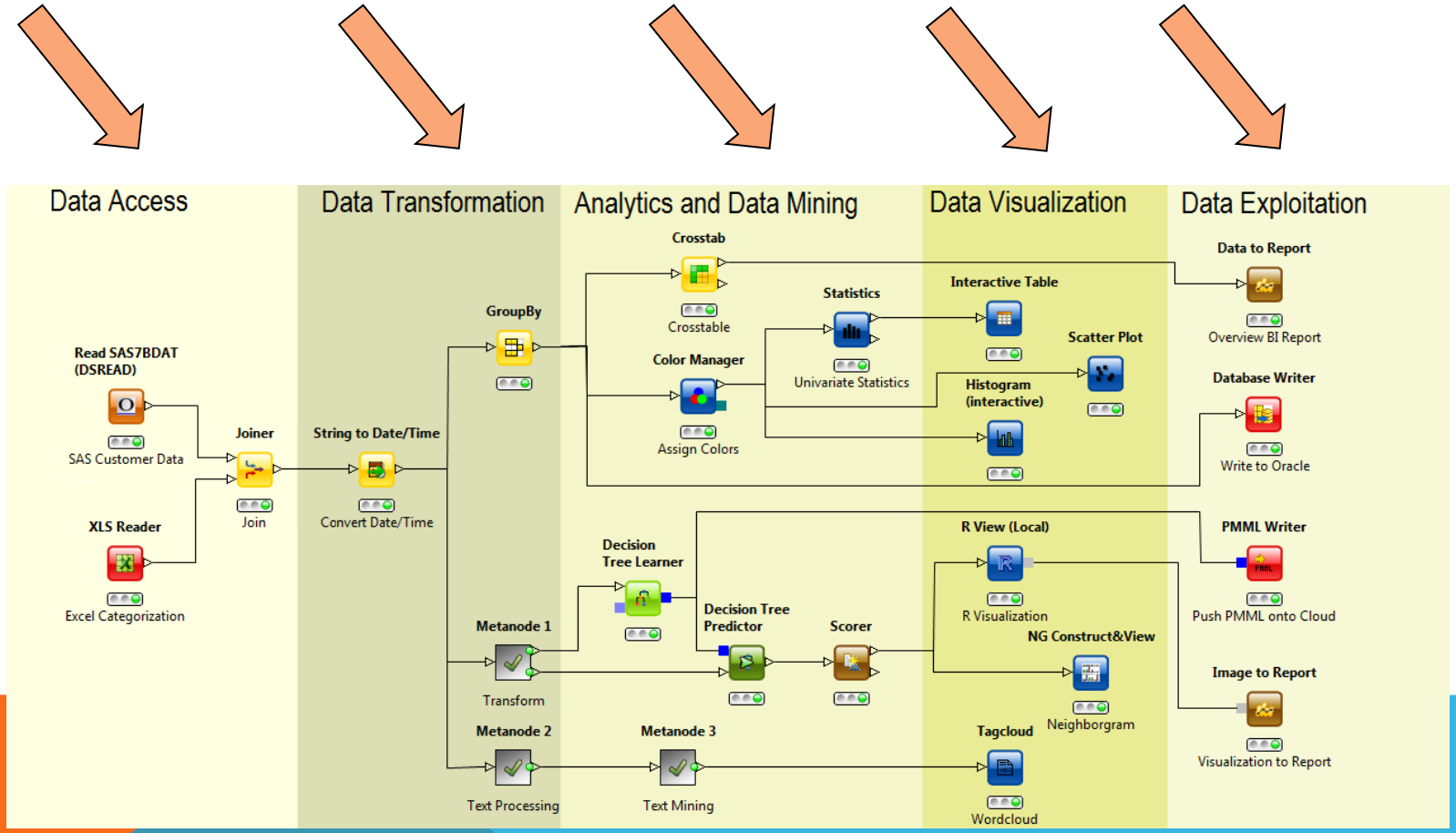


Workbench overview

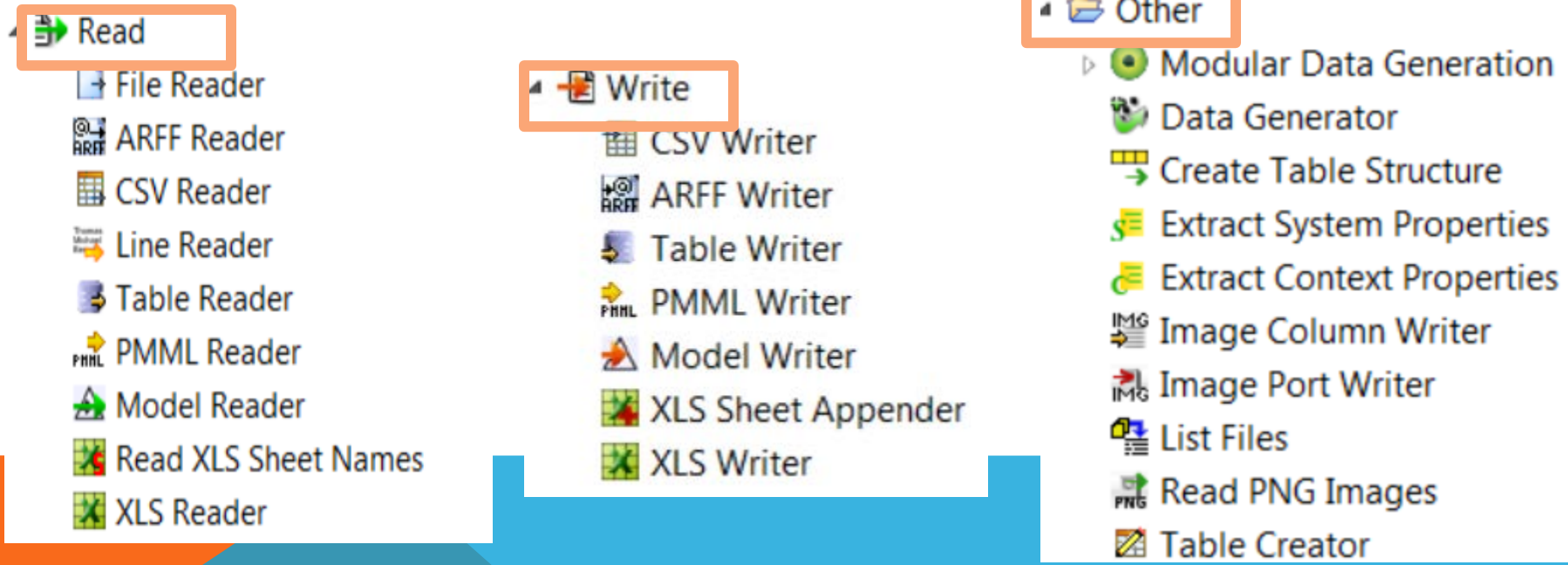
The KNIME Workbench is organized as follows:



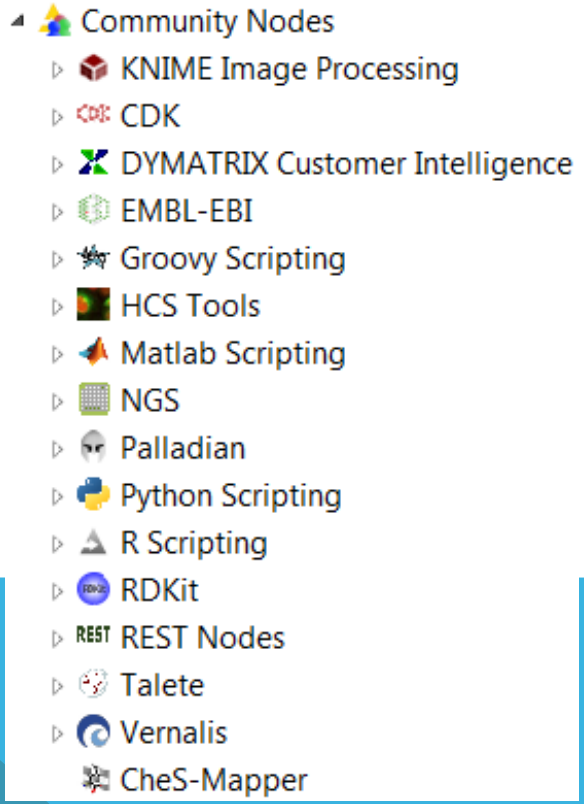

Node Description
List of Workflow Projects
Workflow Editor
KNIME Server
Node Repository
Outline
Console
File Reader
Remove not interesting columns and rows
Column Filter
Row Filter
CSV Writer
This node is set to override the output file if it already exists
adult.data set
rm "final weight"
born outside US
write new file
This workflow is my first KNIME workflow.
It reads data, removes uninteresting columns and rows from the data table and writes the result to a CSV file.
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
Workflow Projects
Chapter2
Exercises
my_first_workflow
Chapter3
Chapter4
Chapter5
KNIME_project
Favorite Nodes
Personal favorite nodes
Most frequently used nodes
Last used nodes
Node Repository
IO
Database
Data Manipulation
Data Views
Statistics
Mining
Chemistry
Distance Matrix
Meta
Flow Control
Misc
KNIME Labs
Time Series
Quick Form
R
Reporting
Weka
XML
Outline
Console
KNIME Console
*** Welcome to KNIME v2.4.1.0030379 - the Konstanz Information Miner ***
*** Copyright, 2003 - 2011, Uni Konstanz and KNIME GmbH, Germany ***
Log file is located at: C:\Programme\knime_2.4.1\book_workspace\.metadata\knime\knime.log
WARN CSV Writer Selected output file exists and will be overwritten!
Server Workflow Proje...
Workflow Server:
Status: not connec
Connect


Data Access
Data Transformation
Analytics and Data Mining
Data Visualization
Data Exploitation
Read SAS7BDAT (DSREAD)
SAS Customer Data
XLS Reader
Excel Categorization
Joiner
Join
String to Date/Time
Convert Date/Time
GroupBy
Crosstab
Crosstable
Color Manager
Assign Colors
Statistics
Univariate Statistics
Interactive Table
Scatter Plot
Histogram (interactive)
Data to Report
Overview BI Report
Database Writer
Write to Oracle
Decision Tree Learner
Decision Tree Predictor
Scorer
Metanode 1
Transform
Metanode 2
Text Processing
Metanode 3
Text Mining
R View (Local)
R Visualization
NG Construct&View
Neighborgram
Tagcloud
Wordcloud
PMML Writer
Push PMML onto Cloud
Image to Report
Visualization to Report

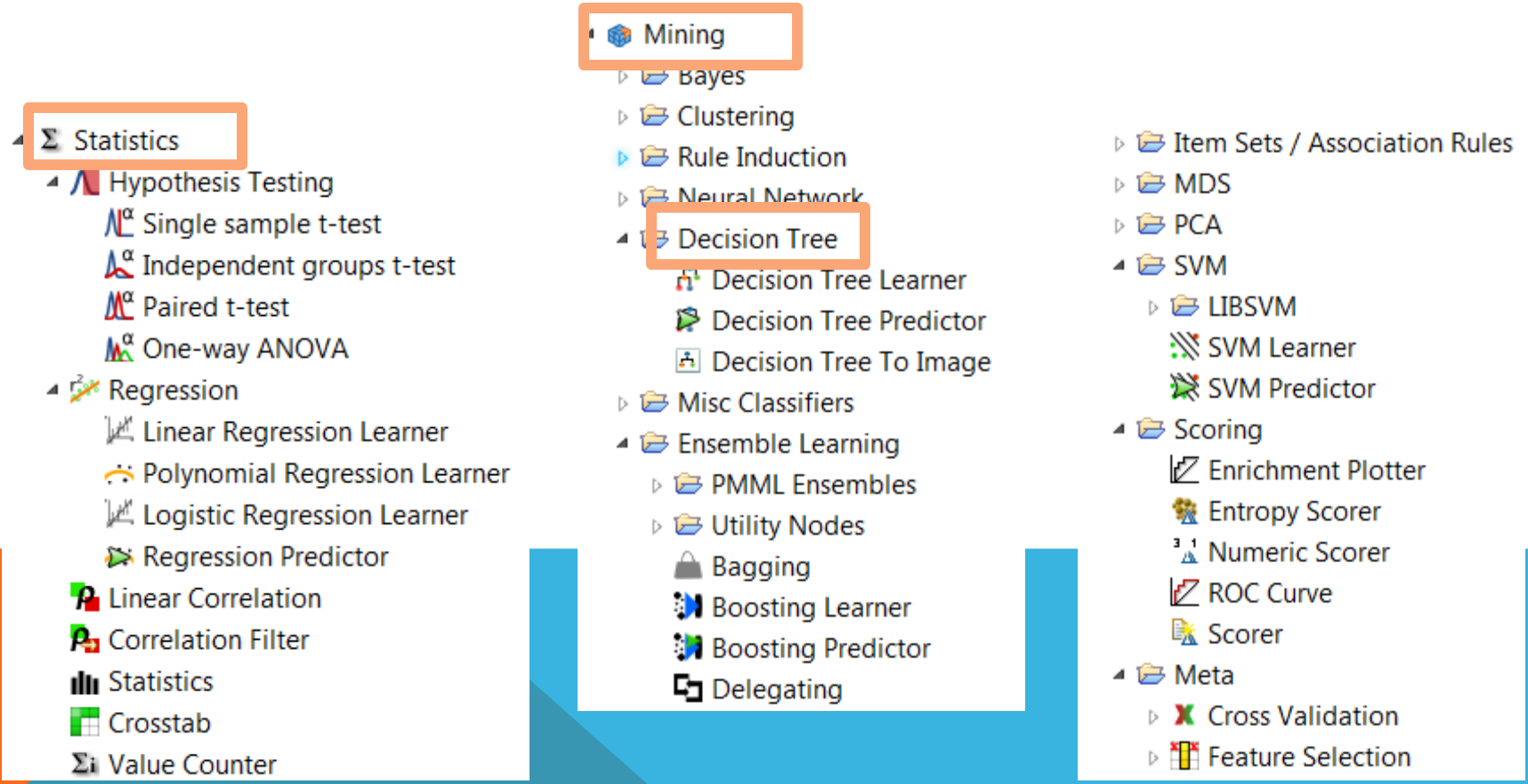
Many Available Data Nodes



Many Available Community Nodes



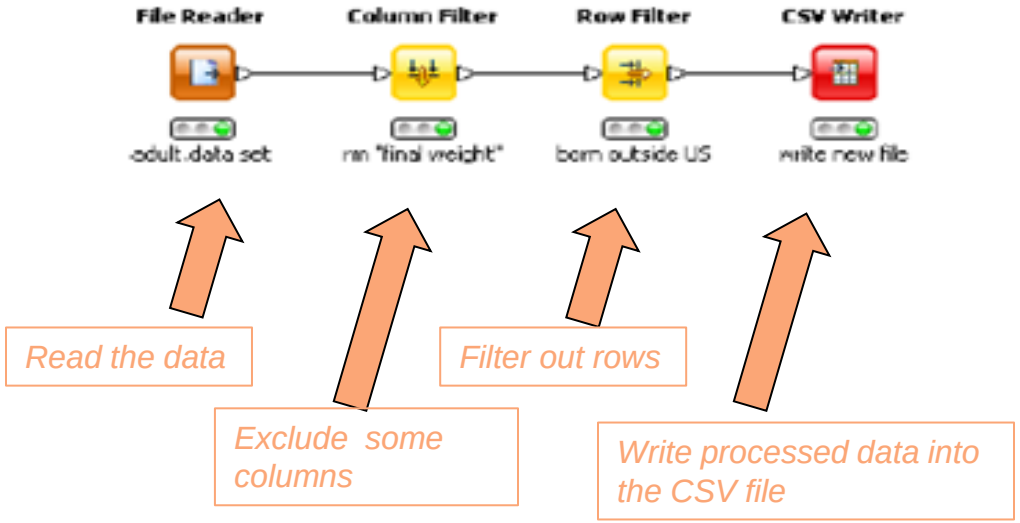
Many Statistics and Data Mining Nodes



WORKFLOWS IN KNIME

Sequence (pipeline) of the analysis steps

- Example
  - Read Data
  - Clean Data
  - Train a model
  - Etc.

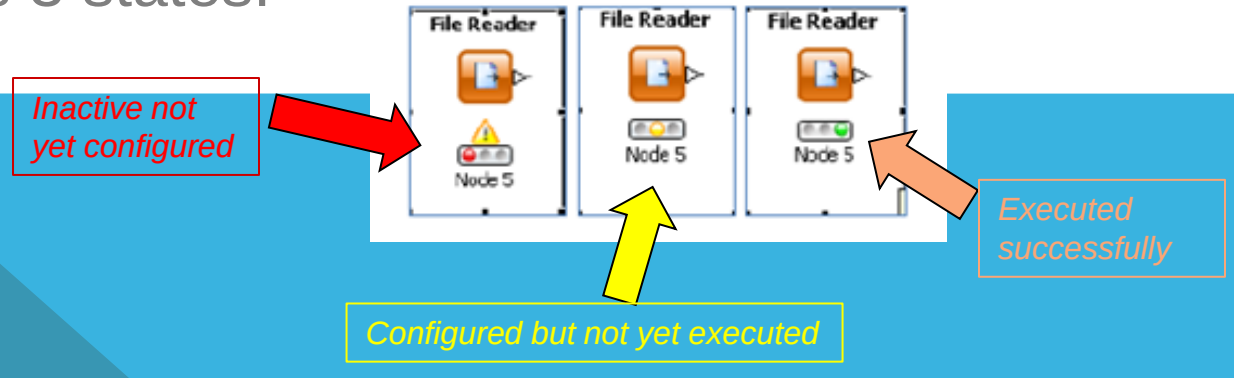


Building a Workflow

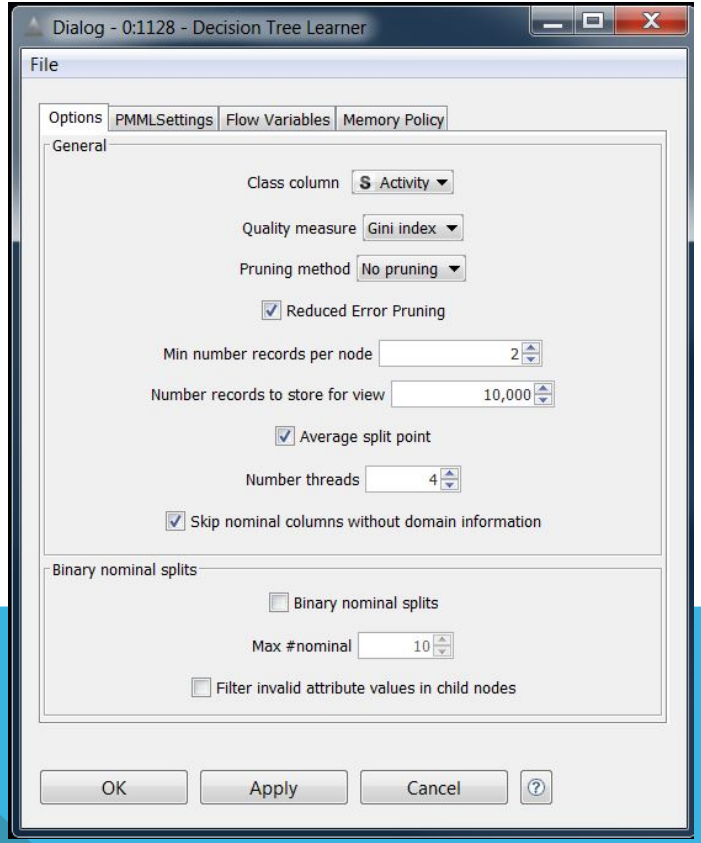
- A workflow is built by dragging nodes from the Node Repository to the Workflow Editor and connecting them
- Nodes are the basic processing units of a workflow
- Each node has a number of input- and/or output ports
- Data (or a model) is transferred via a connection from an out-port to the in-port of another node

Nodes in a Workflow

- Node is a single processing unit in the workflow
- The node takes data from the input, processes it and delivers it in the output port
- The processing can be anything from data manipulation, plotting to training Machine learning algorithms
- Every node has 3 states:



Decision Tree Node



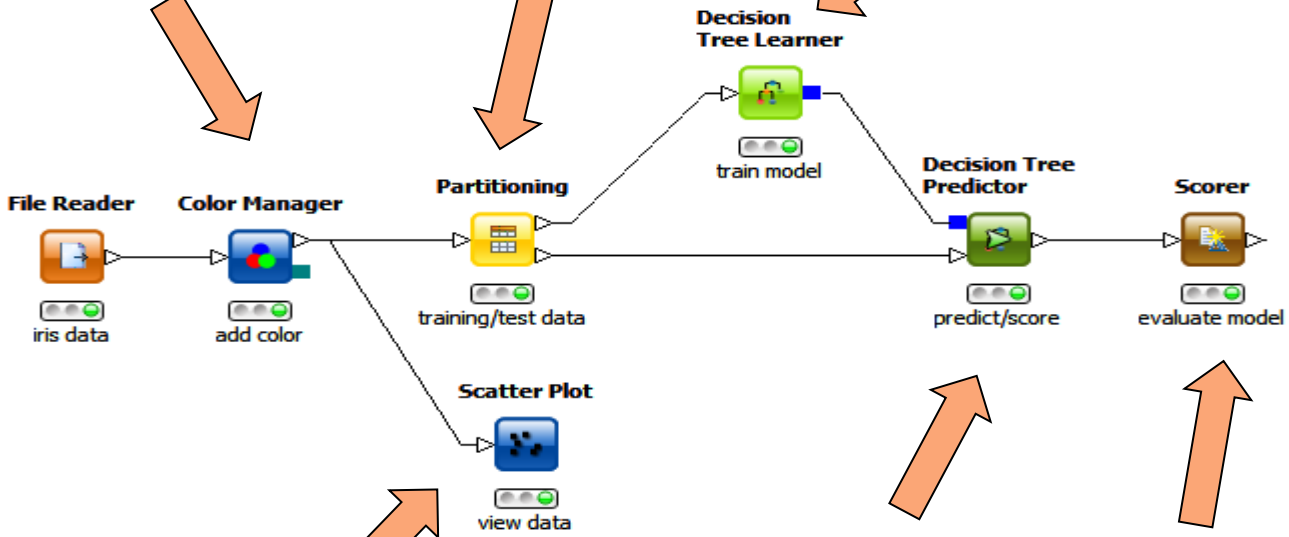
DECISION TREE WORKFLOW EXAMPLE

Read in Training data set

Assign colors to the data for easier visualization

Partition data into training and testing data sets

Train the decision tree model



Visualize data in a scatter plot

Make predictions on the model - scoring

Evaluate the model with the confusion matrix. etc

Additional Information

https://www.knime.org/learning-hub?src=knimeapp