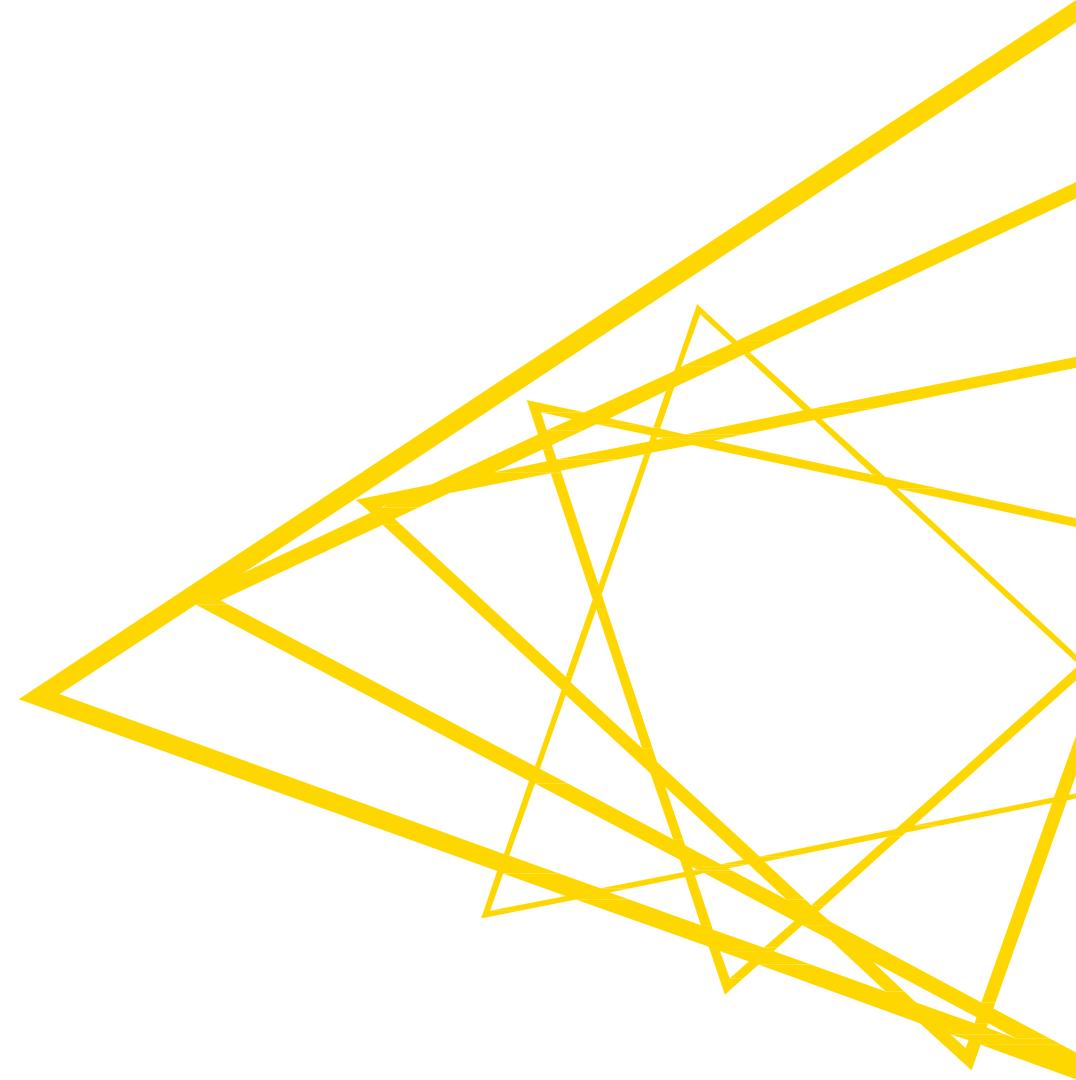




# Discover KNIME 101

Cevi Herdian  
Data-Driven  
Strategist

[itsmecivi.github.io](https://itsmecivi.github.io)



---

# Why KNIME?

# STRENGTHS

---

- No artificial limitations, it can run projects with hundreds of millions of rows (hard disk and memory)
- Build end to end data science workflows
- Open Source

# WEAKNESSES

---

- Intellectual property and patents issues. There is over 60 different licenses that comply with the open source definition
- Switching costs (more cheaper or more budget)
- Migration of data
- Partitioning ability is limited for dataset
- Less suitable option for large complex workflows

# OPPORTUNITIES

---

- Increasing users (personal, industry, institutions)
- Open platforms-> very committed users and advanced ecosystems-> breakthroughs in data-driven innovation
- Continue delivering extended functionality

# THREATS

---

- Intellectual property and patents issues
- Human Resources
- Migration of data
- Switching costs to open sources (more cheaper or more budget)

# Agenda

---

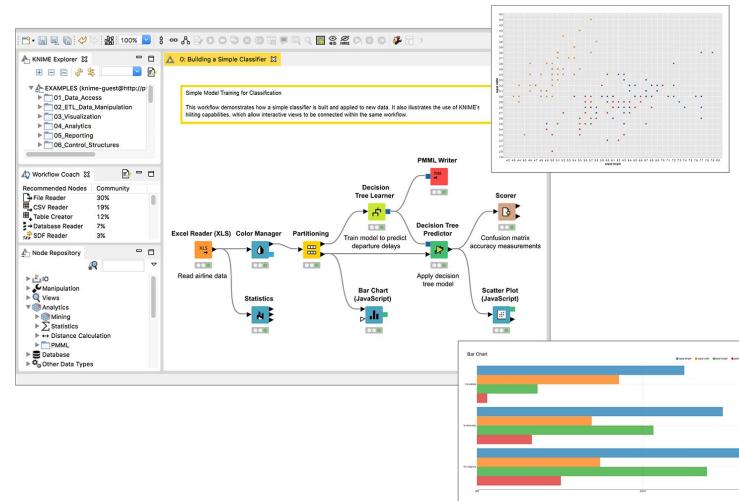
- 1. What is KNIME?
- 2. ETL (Extract, Transfrom, Load)
- 3. Integration
- 4. Big Data
- 5. Advanced Analytics
- 6. KNIME Server
- 7. KNIME GUI

---

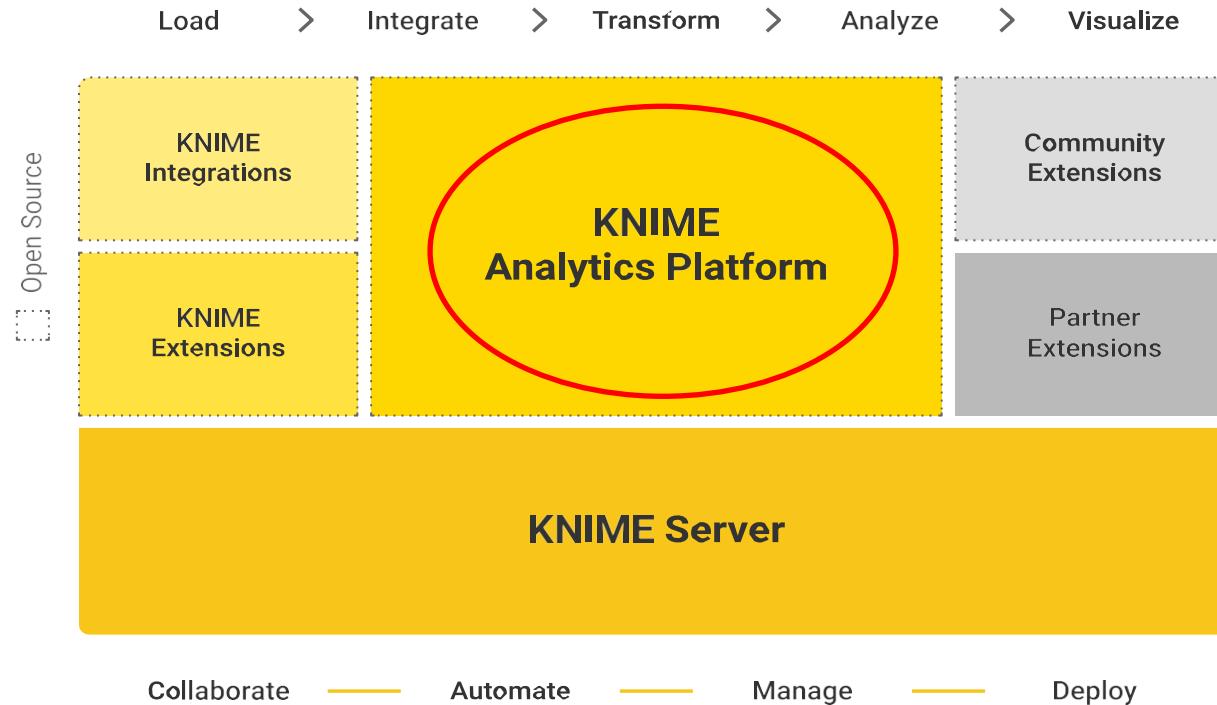
# 1. What/Who is KNIME?

# What is KNIME Analytics Platform?

- A tool for data analysis, manipulation, visualization, and reporting
- Based on the graphical programming paradigm (click & drag)
- Provides a diverse array of extensions:
  - Text Mining
  - Network Mining
  - Cheminformatics
  - Many integrations, such as Java, R, Python, Weka, Keras, Plotly, H2O, etc.



# KNIME Software Overview



## KNIME Fact Check:

software advanced data science  
individuals organizations drive innovation  
open sources software over a decade  
over 60 countries without the need for coding  
**R and Python**

## Commercial vs Open Sources:



- Code Free/Low Code



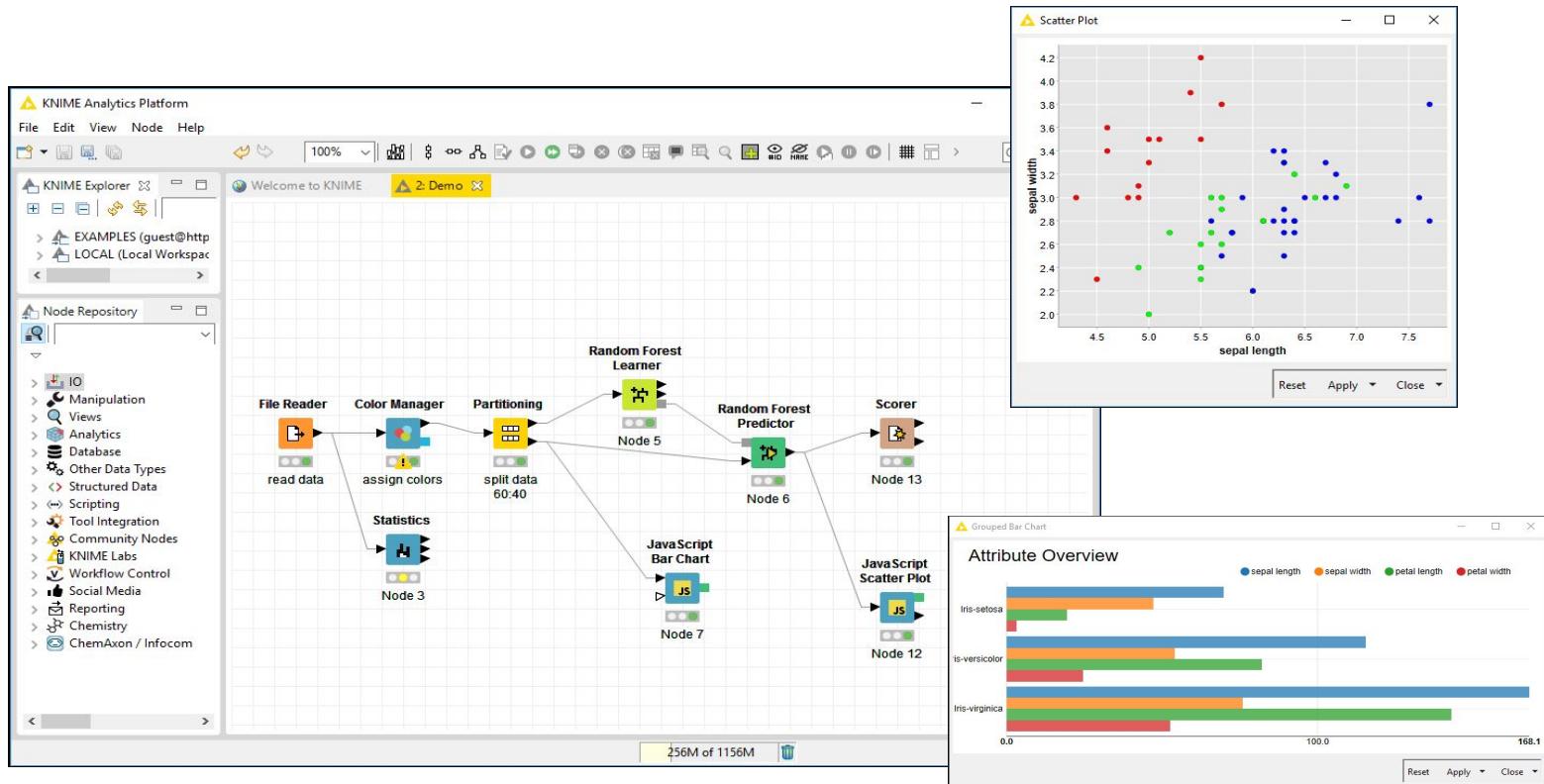
- Coding



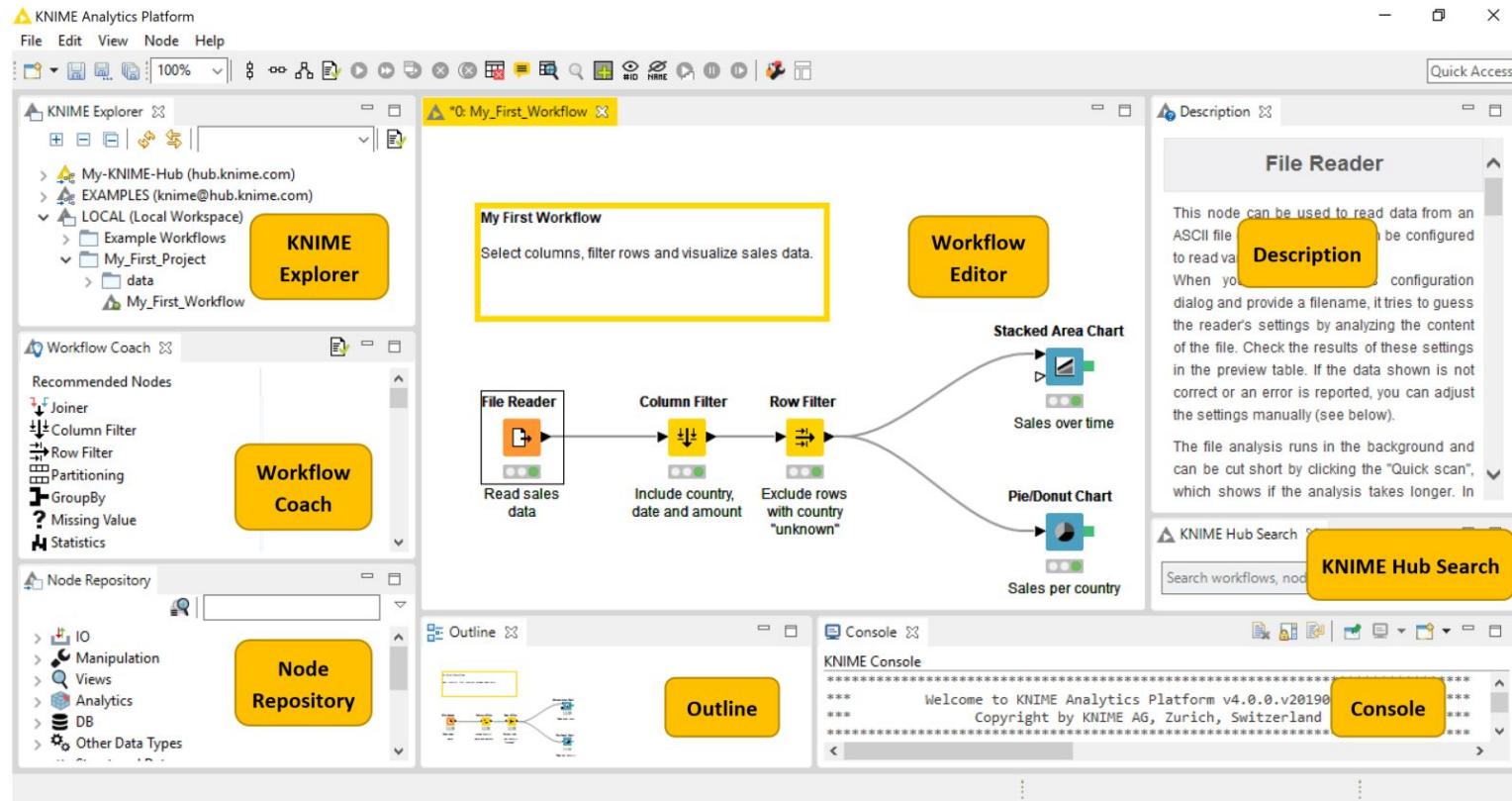
- Integrated



# KNIME GUI

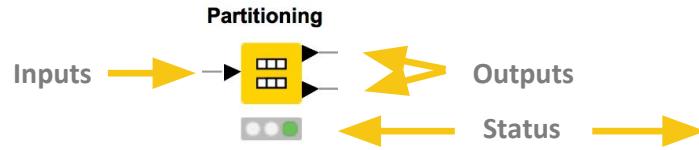


# The KNIME Workbench



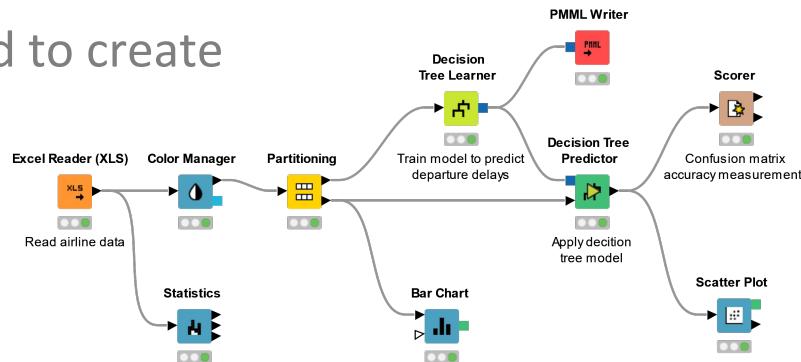
# Visual KNIME Workflows

NODES perform tasks on data

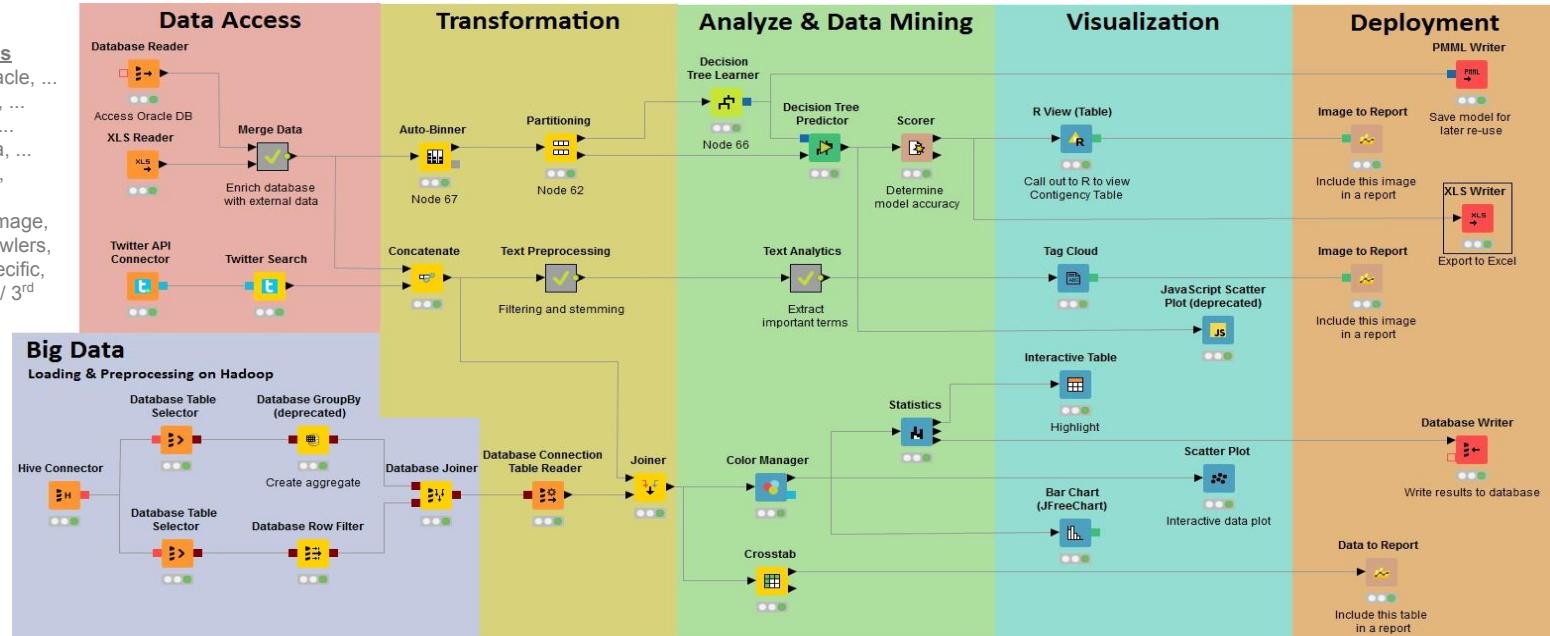


Not Configured  
Configured  
Executed  
Error

Nodes are combined to create  
**WORKFLOWS**



# Over 1500 native and embedded nodes included:



## Big Data

Hive, Impala, HDFS  
Vertica, Teradata/Aster,  
Spark, MLlib, Community / 3<sup>rd</sup> party, ...

## Transformation

Row, Column, Matrix  
Text, Image, Networks,  
Time Series, Java,  
Python, Community / 3<sup>rd</sup> party, ...

## Analysis & Mining

Statistics, Machine Learning,  
Data Mining, Web Analytics,  
Text Mining, Network Analysis,  
Social Media Analysis, R,  
Weka, Python, Community / 3<sup>rd</sup> party, ...

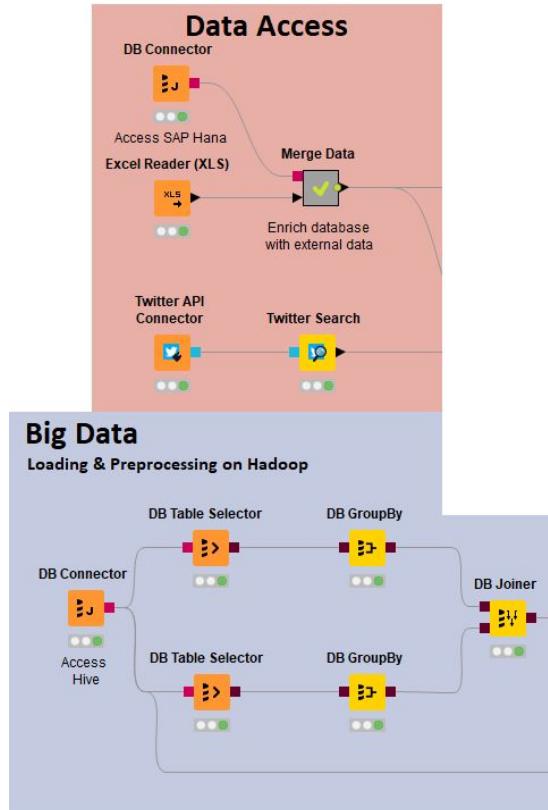
## Visualizati

on\_R,  
Python,  
JFreeChart,  
JavaScript,  
Community / 3<sup>rd</sup> party,  
...  
Node 66

## Deployment

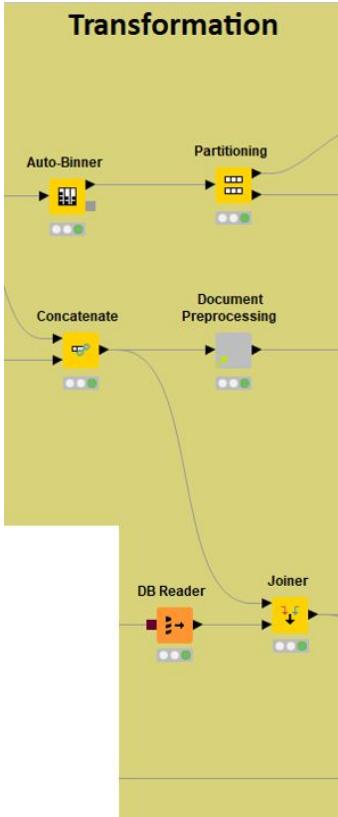
via BIRT  
PMML, XML, JSON  
Databases, Excel, Flat,  
etc. Text, Doc, Image  
Industry Specific  
Community / 3<sup>rd</sup> party,  
...  
Node 66

# Data Access/Data Sources/Extract



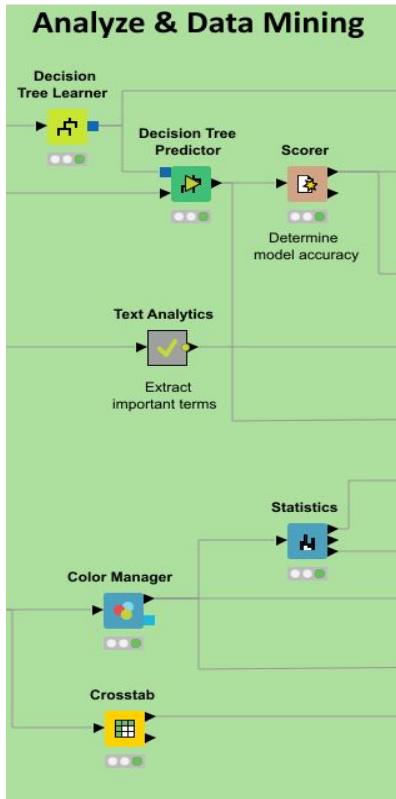
- Databases
  - MySQL, PostgreSQL
  - any JDBC (Oracle, DB2, MS SQL Server)
- Files
  - CSV, txt
  - Excel, Word, PDF
  - SAS, SPSS
  - XML
  - PMML
  - Images, texts
- Web, Cloud
  - REST, Web services
  - Twitter, Google

# Transformation



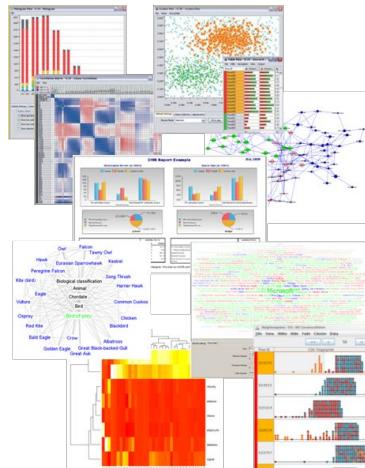
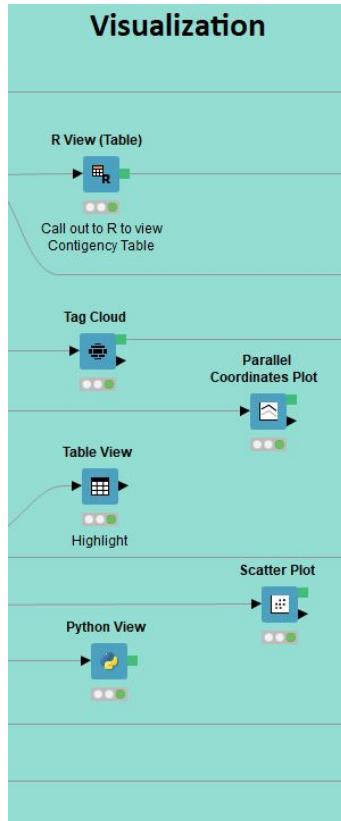
- Preprocessing
  - Row, column, matrix based
- Data blending
  - Join, concatenate, append
- Aggregation
  - Grouping, pivoting, binning
- Feature Creation and Selection

# Analyze & Data Mining



- Regression
  - Linear, logistic, time series analysis
- Classification
  - Decision tree, ensembles, SVM, MLP, Naïve Bayes, random forest, neural networks, model optimization
- Clustering
  - k-means, DBSCAN, hierarchical
- Validation
  - Cross-validation, scoring, ROC
- Misc
  - PCA, MDS, item set mining
- External
  - R, Python, H2O, Weka
- Deep Learning

# Visualization

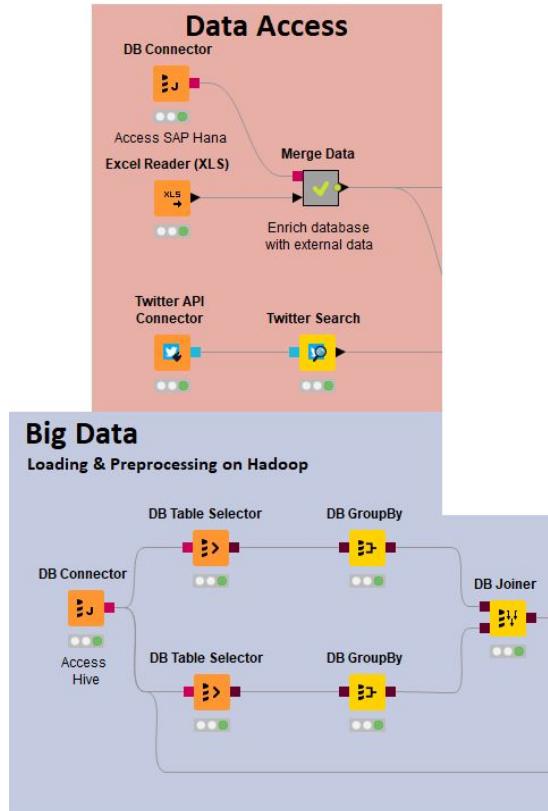


- Interactive Visualizations
- JavaScript-based nodes
  - Scatter Plot, Box Plot, Line Plot
  - Networks, ROC Curve, Decision Tree
  - Plotly Integration
  - Adding more with each release!
- Misc
  - Tag cloud, open street map, molecules
- Script-based visualizations
  - R, Python

---

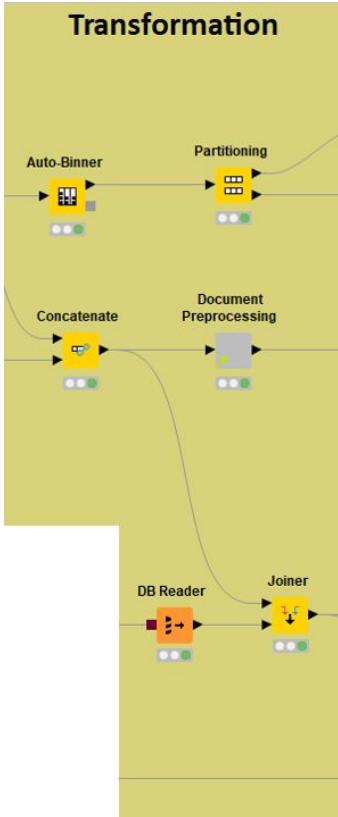
## 2. ETL

# Data Access/Data Sources/Extract



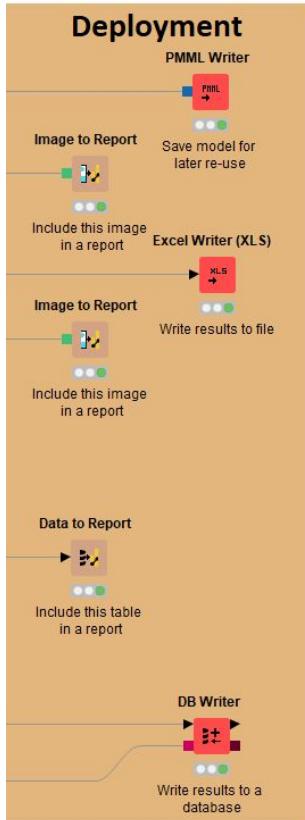
- Databases
  - MySQL, PostgreSQL
  - any JDBC (Oracle, DB2, MS SQL Server)
- Files
  - CSV, txt
  - Excel, Word, PDF
  - SAS, SPSS
  - XML
  - PMML
  - Images, texts
- Web, Cloud
  - REST, Web services
  - Twitter, Google

# Transformation



- Preprocessing
  - Row, column, matrix based
- Data blending
  - Join, concatenate, append
- Aggregation
  - Grouping, pivoting, binning
- Feature Creation and Selection

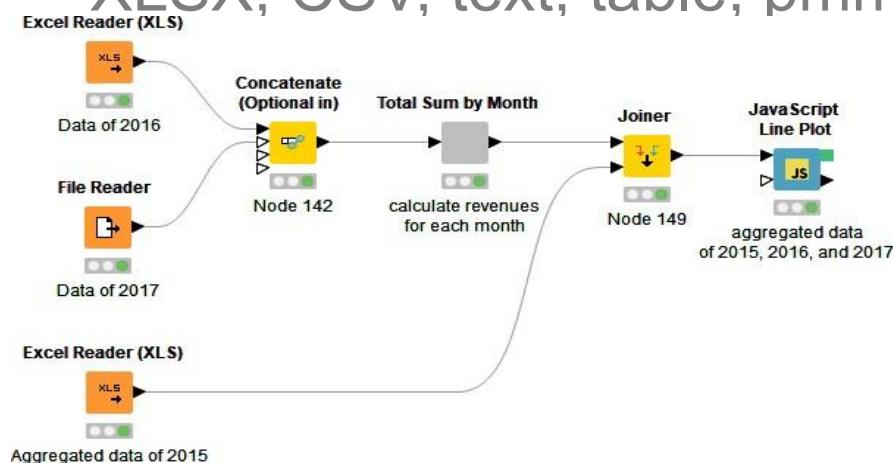
# Loaden/Deployment



- Database
- Files
  - Excel, CSV, txt
  - XML
  - PMML
  - to: local, KNIME Server, SSH-, FTP-Server
- BIRT Reporting

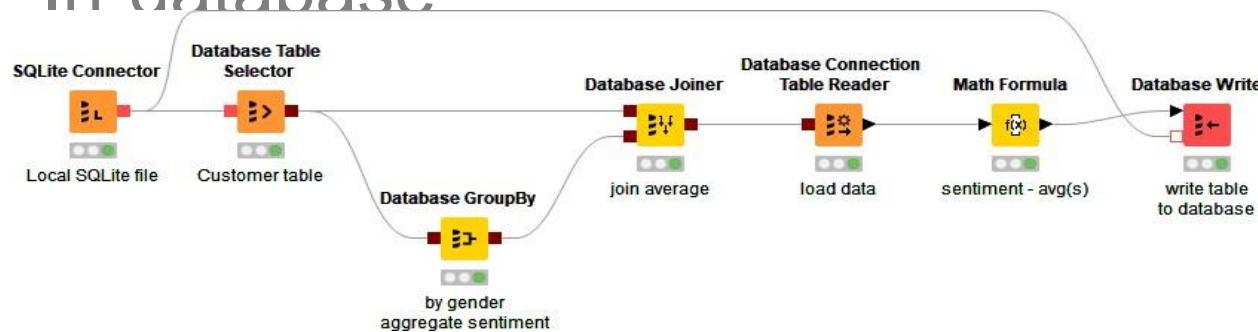
# Microsoft Excel, CSV, etc.

- Reading and combining data from different formats
  - XLSX, CSV, text, table, pmml, xml, json, pdf, word,



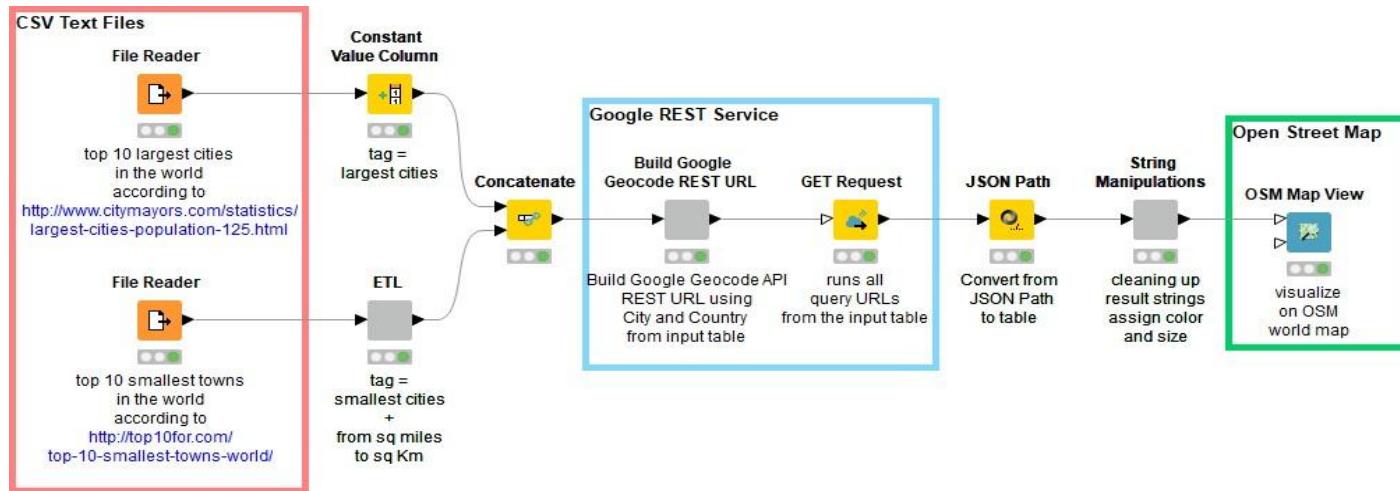
# Databases

- Connect via JDBC
- Read /write to databases
- In-database



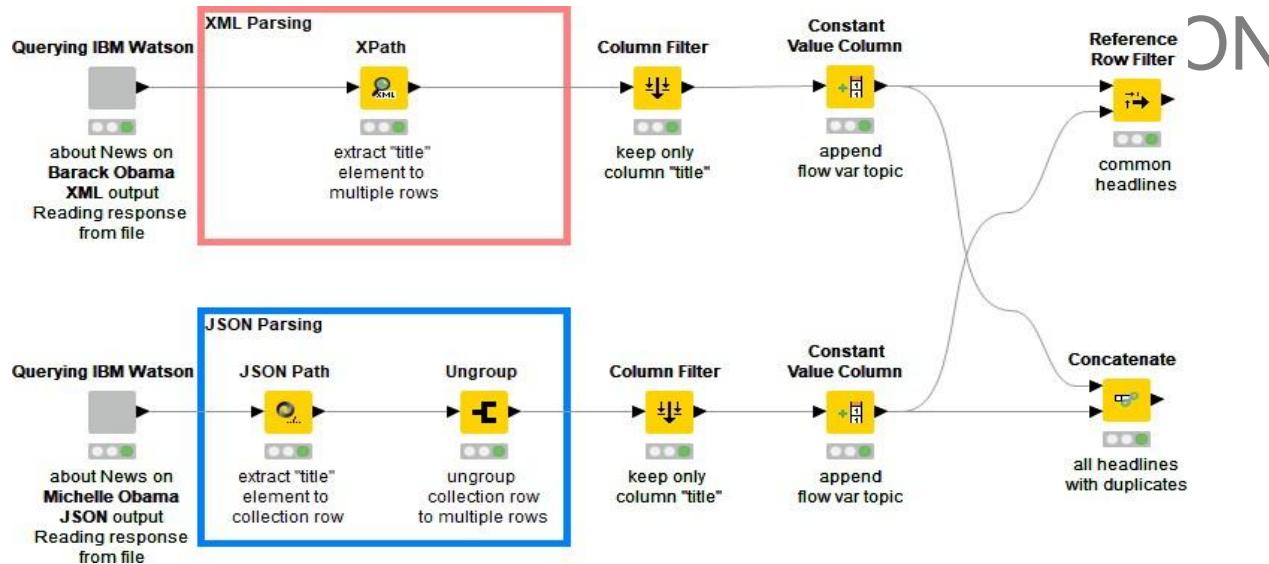
# REST Web Services

- Connect to any RESTful service (GET, POST)
- Load data from services



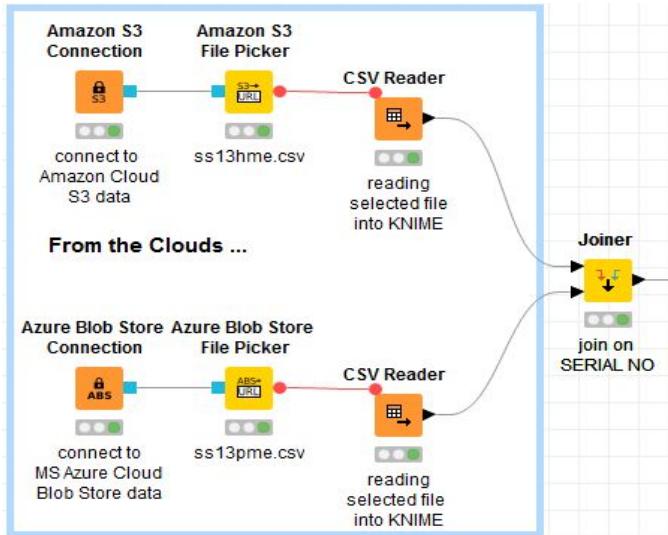
# XML & JSON

- Parse remote or locale XML and JSON data

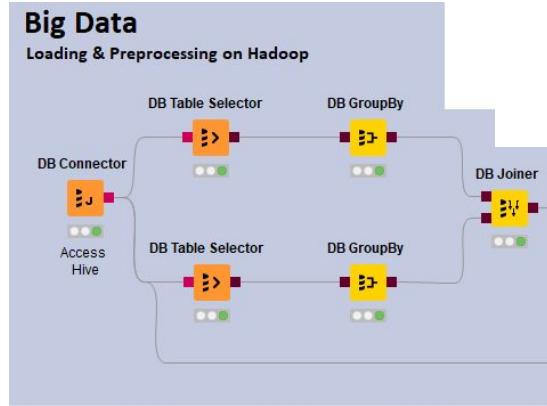


# Amazon S3 & Microsoft Azure Blob Storage

- Read files directly from Amazon S3 or Azure Blob Store.



# Big Data



- Spark
- Hadoop
- HDFS support
- Hive
- Impala
- In-database processing

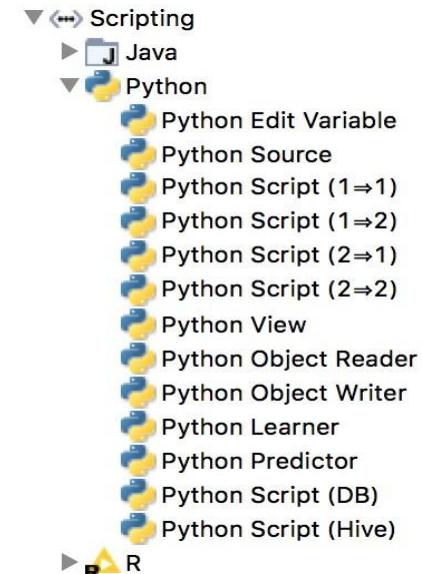
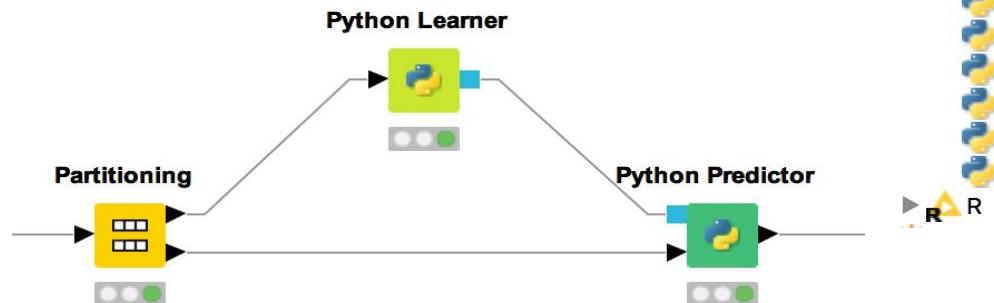


---

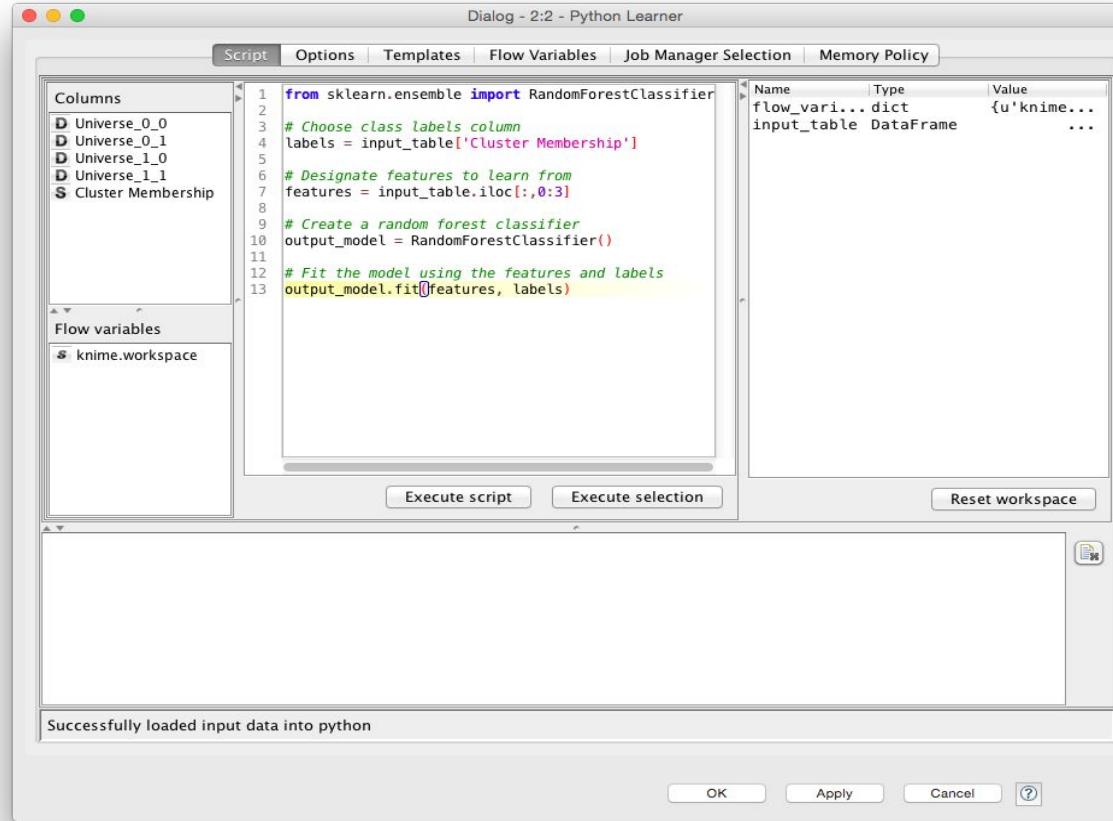
## 3. Integration

# Python Integration

- Run Python inside KNIME
- Works with existing installations
- UI modeled after R integration

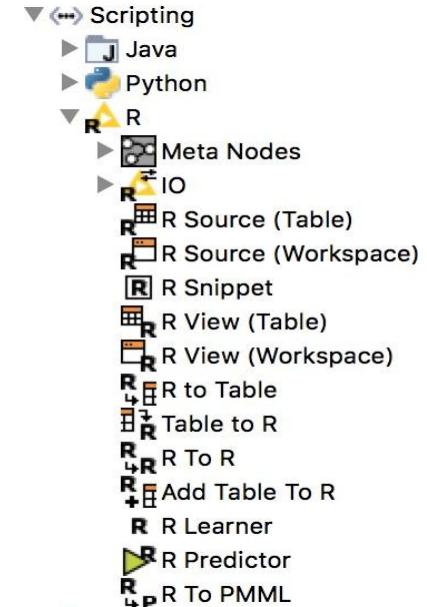


# Python Scripting UI

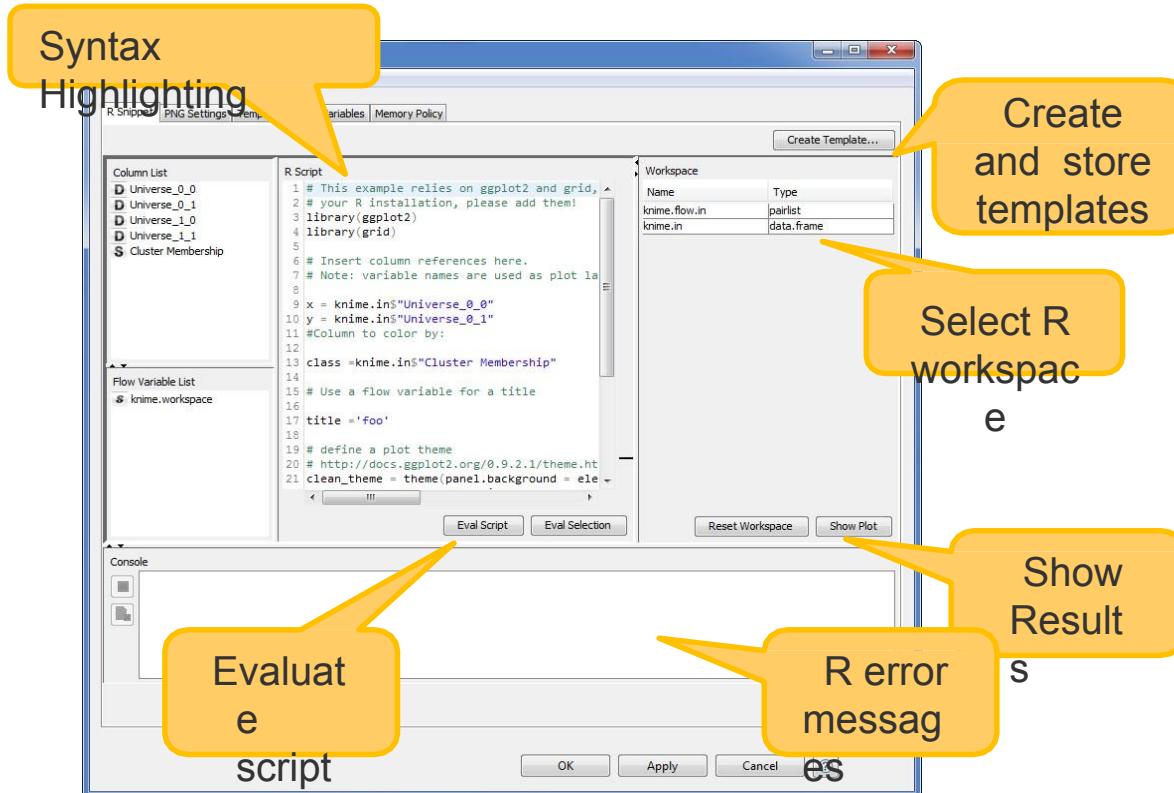


# R Integration

- Run R inside KNIME.
- Works with existing R installations.
- Nodes for many tasks, but all with similiar interfaces.

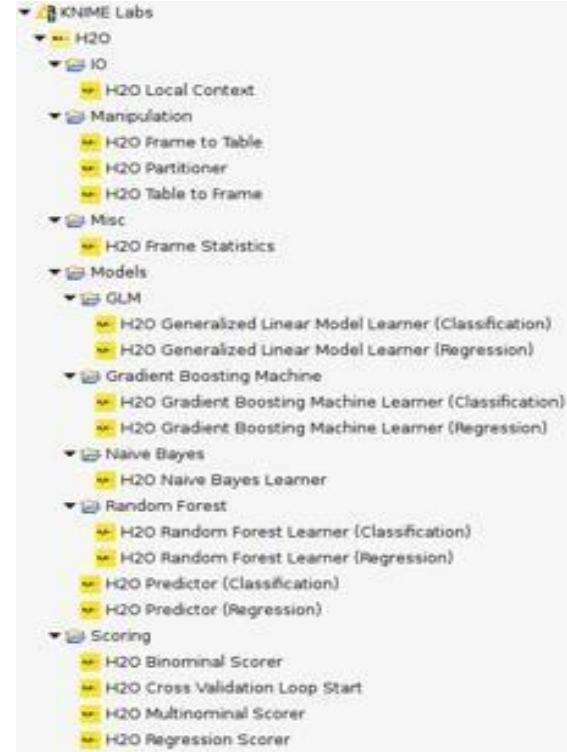


# R Integration UI



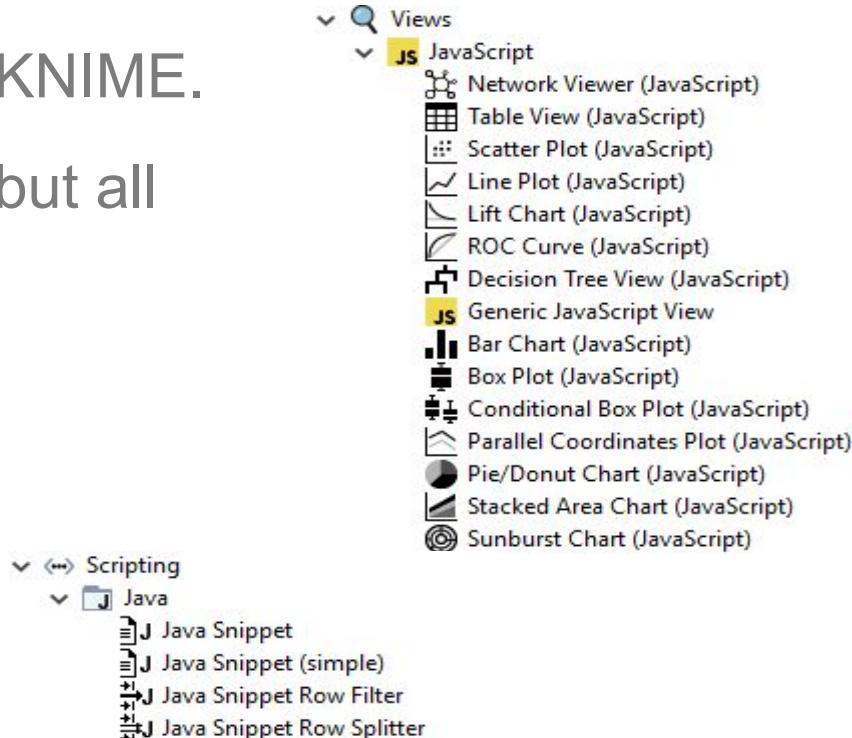
# H2O Integration

- H2O - Machine Learning and Analytics library
- KNIME – H2O integration provides nodes for machine learning and scoring.



# Java Script Integration

- Run Java Script inside KNIME.
- Nodes for many tasks, but all with similiar interfaces.
- Visualization - create interactive Views

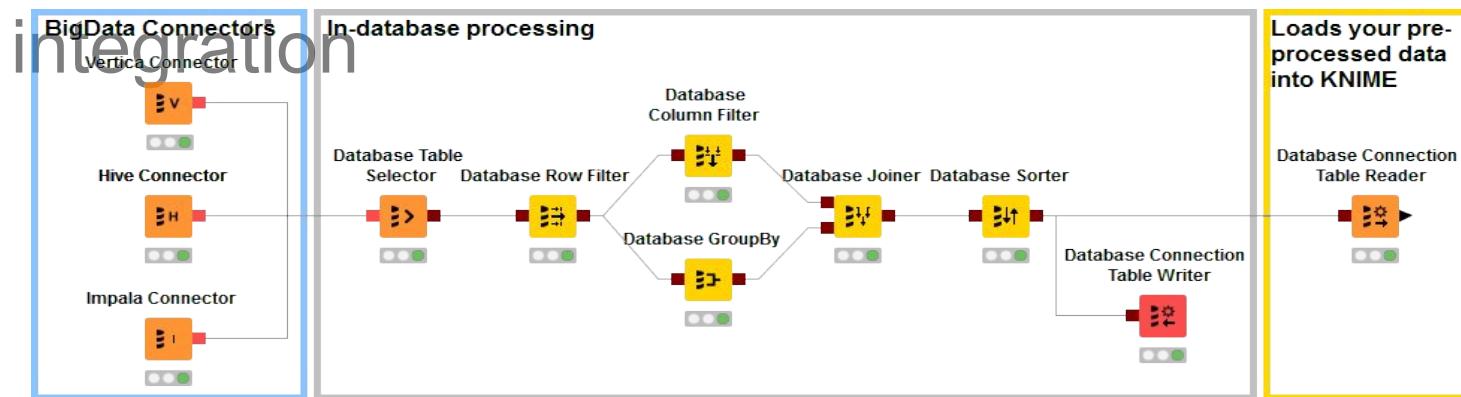


---

## 4. Big Data

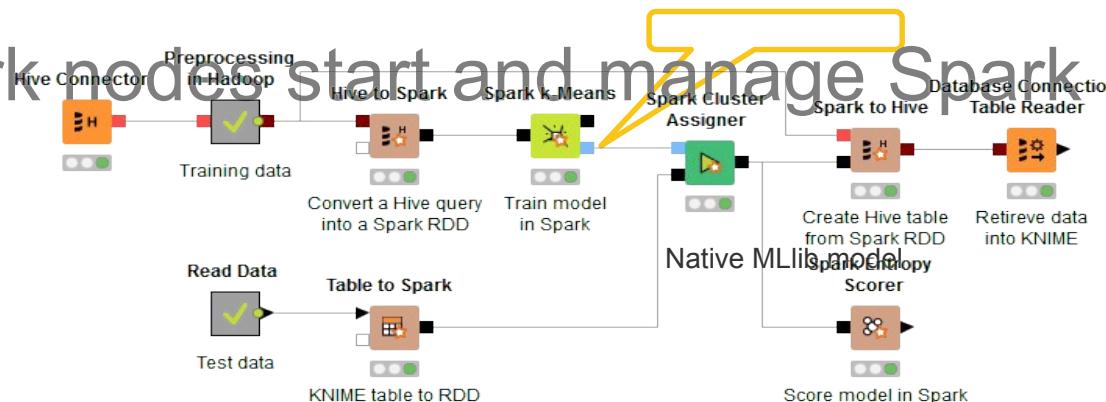
# KNIME Big Data Connectors

- Preconfigured connectors
  - Hive
  - Cloudera Impala
- Extends the open source database

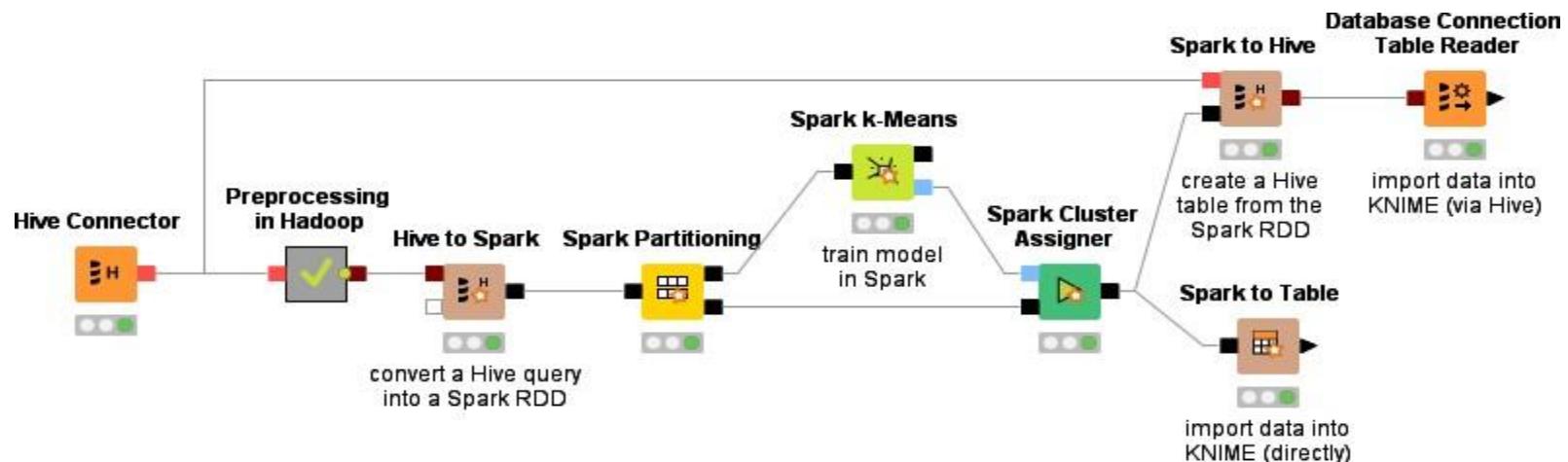


# KNIME Extension for Apache Spark

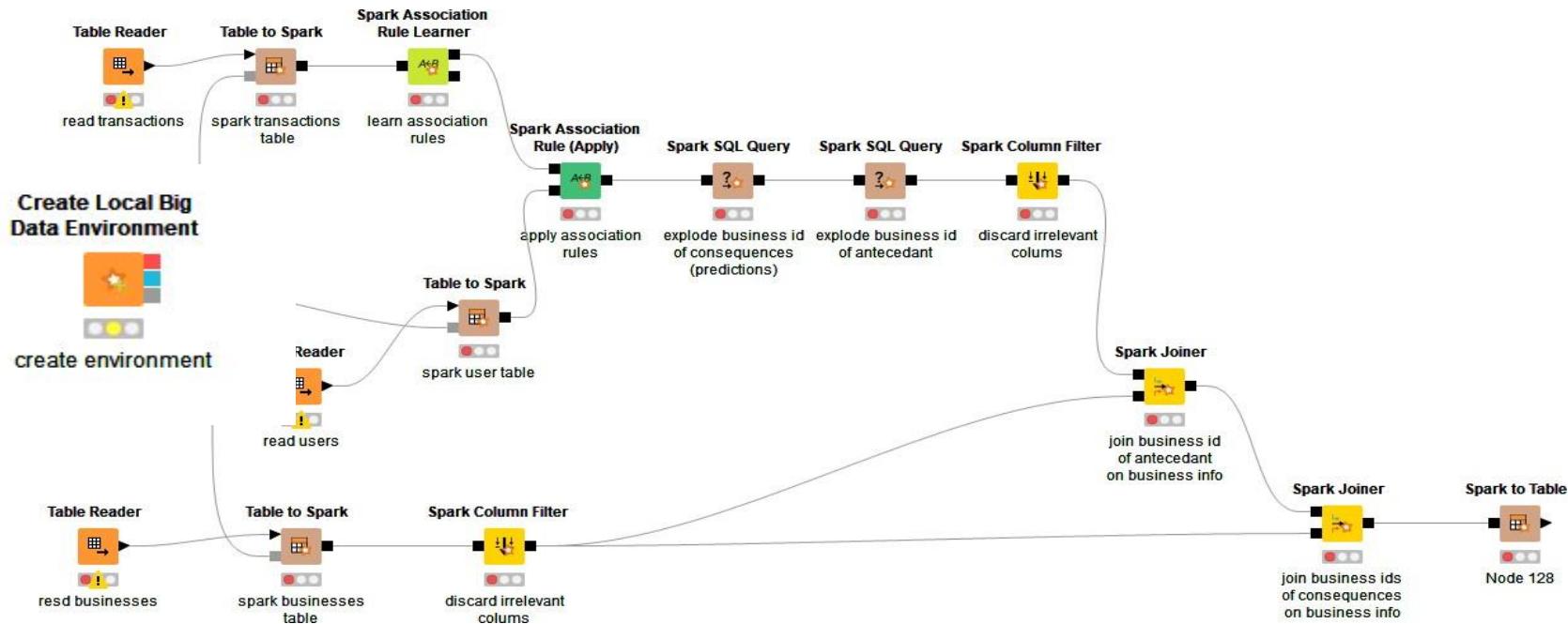
- Advanced Analytics on Hadoop
- Based on Spark MLlib
- Native MLlib model learning and prediction
- Spark nodes start and manage Spark jobs



# Machine Learning for Apache Spark



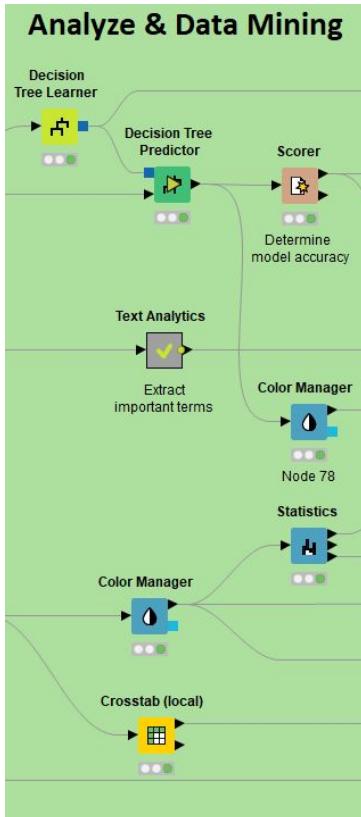
# Test Big Data Nodes – Create Local Big Data Environment



---

## 5. Advanced Analytics

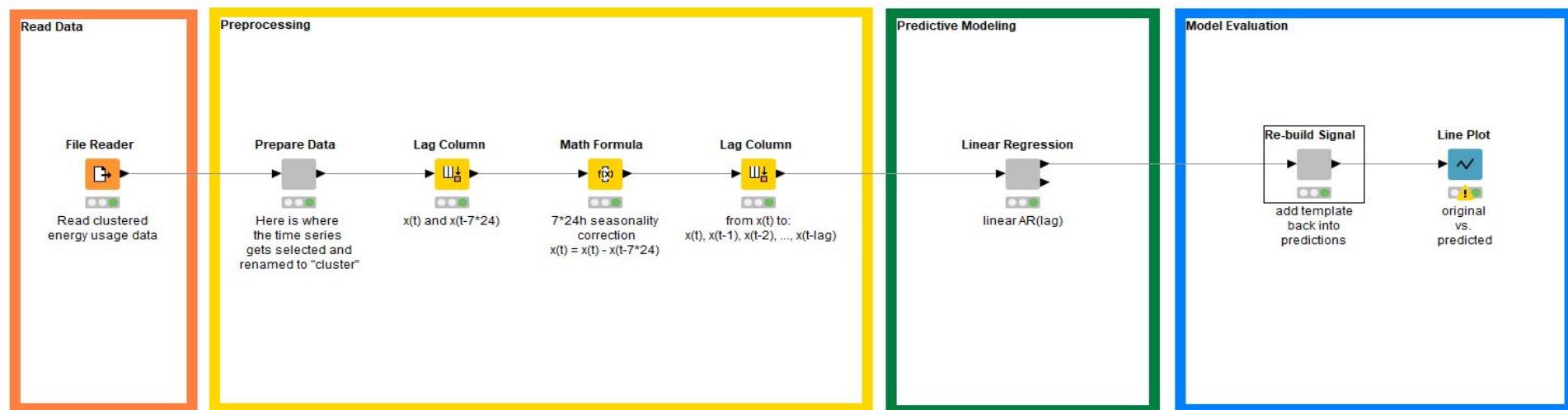
# Analysis & Data Mining



- Regression
  - Linear, logistic
- Classification
  - Decision tree, ensembles, SVM, MLP, Naïve Bayes
- Clustering
  - k-means, DBSCAN, hierarchical
- Validation
  - Cross-validation, scoring, ROC
- Deep Learning
  - Keras, DL4J
- External
  - R, Python, Weka, H2O, Keras

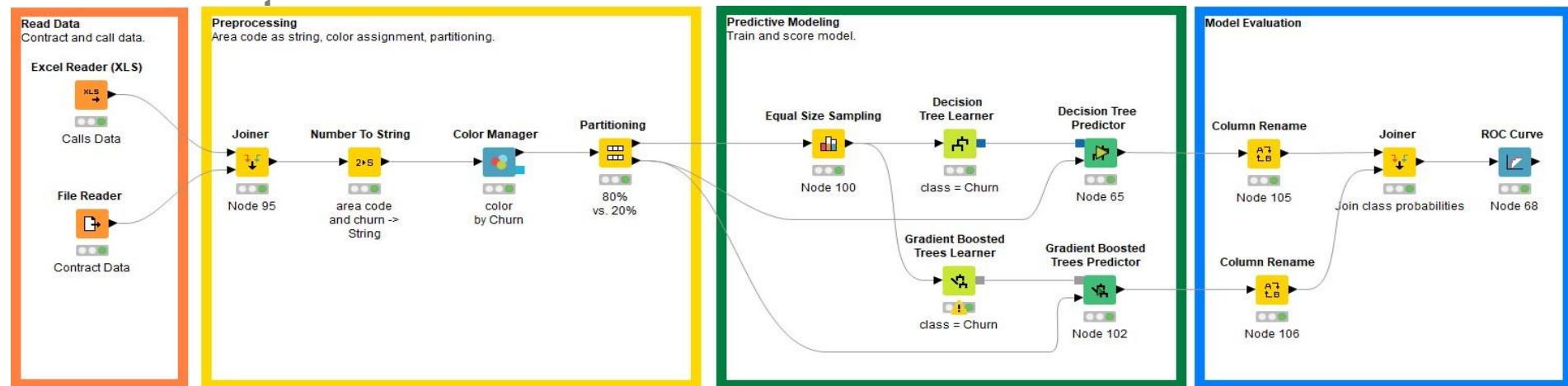
# Energy Usage Prediction

- Data: aggregated energy usage data
- Objective: predict energy usage of next hour



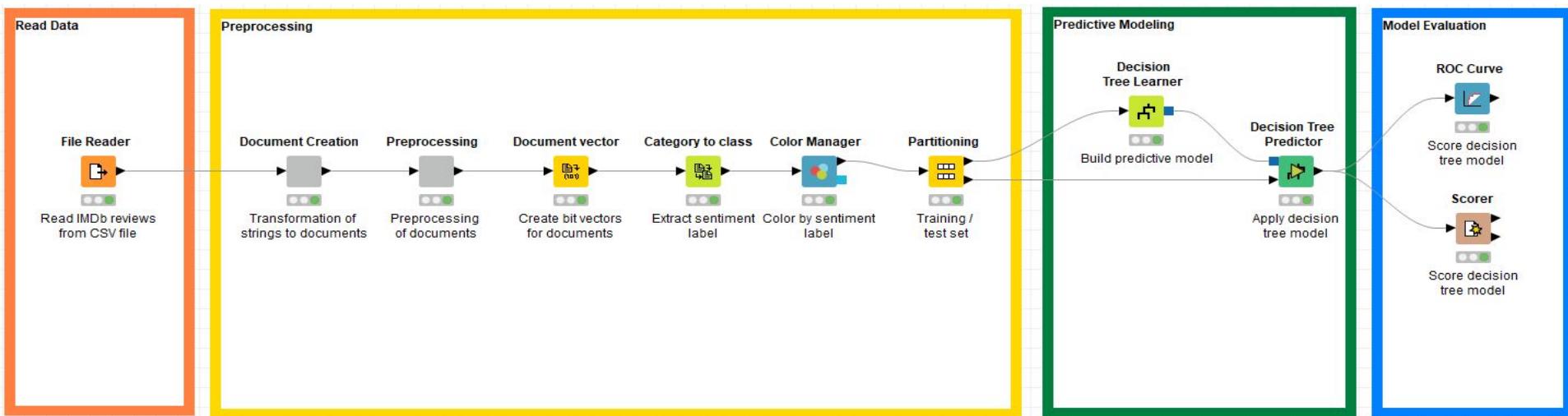
# Churn Prediction

- Data: customer contracts and call data
- Objective: predict churn of



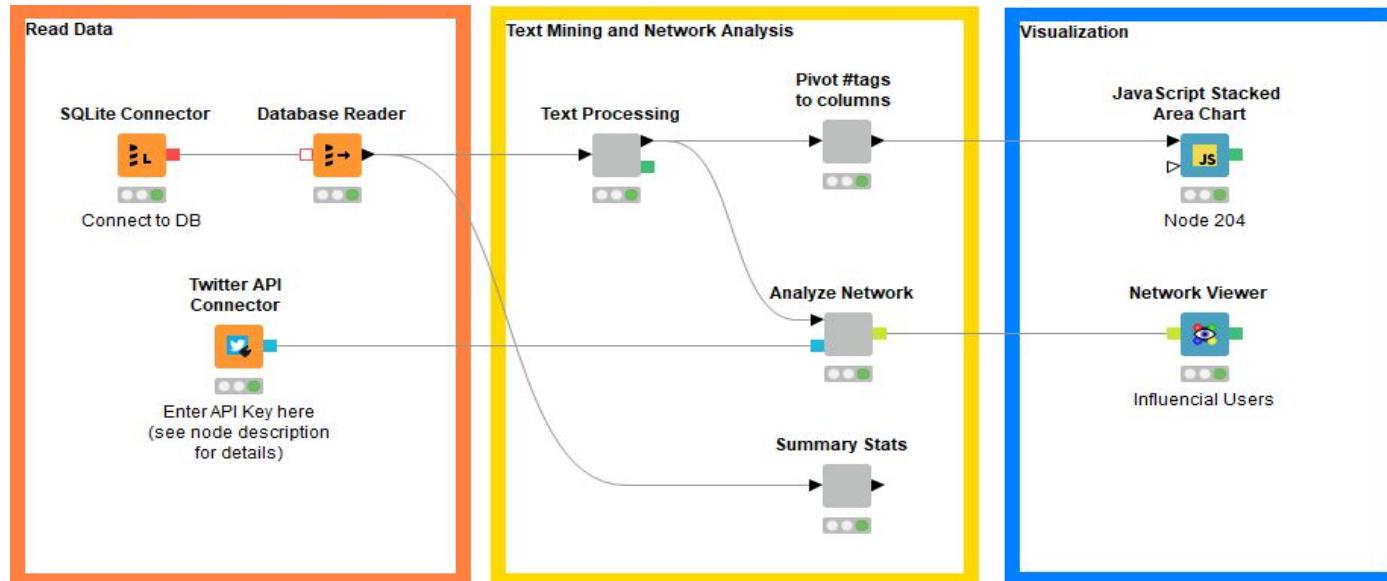
# Sentiment Analysis

- Data: IMDB movie reviews
- Objective: predict sentiment of review



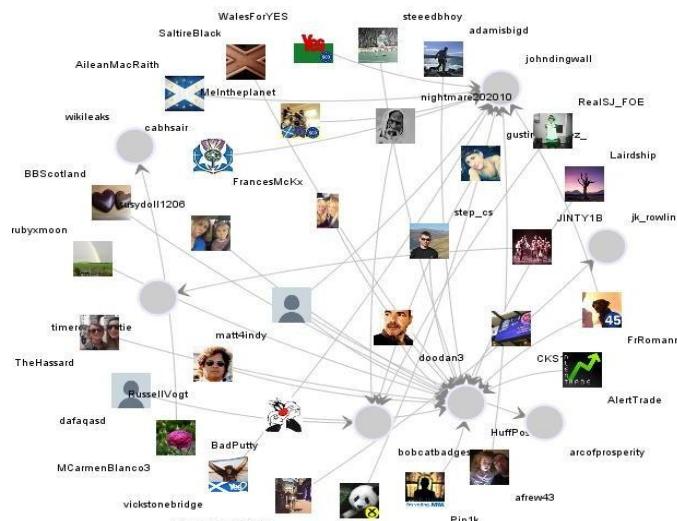
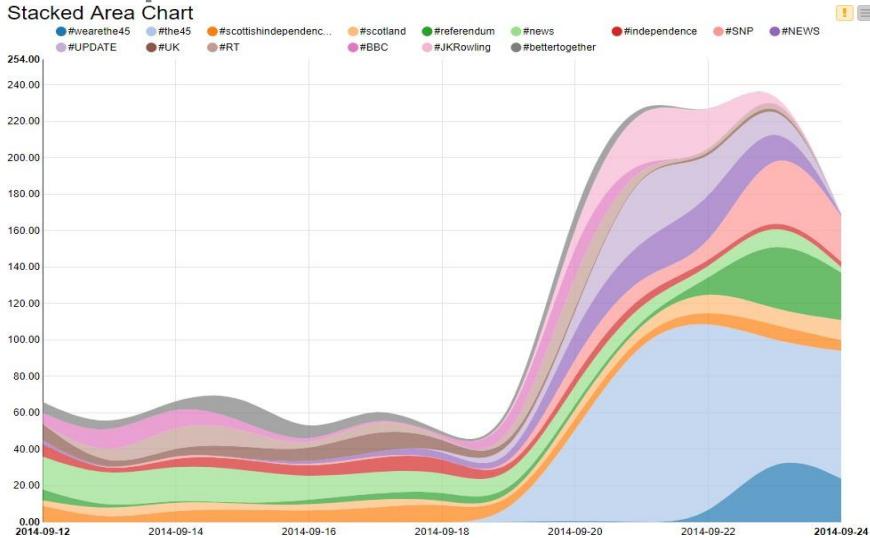
# Twitter Analysis

- Data: Tweets about the Scotland referendum
- Objective: identify upcoming topics, user



# Twitter Analysis

- Method: Twitter API,  
visualization
  - Topic river



---

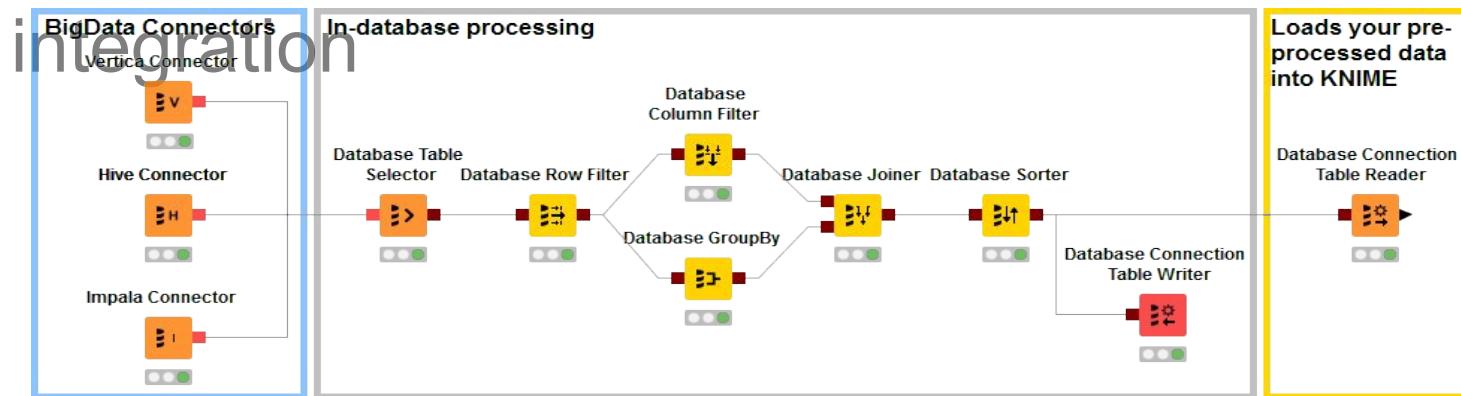
## 6. KNIME Commercial

---

# The KNIME Commercial Extensions

# KNIME Big Data Connectors

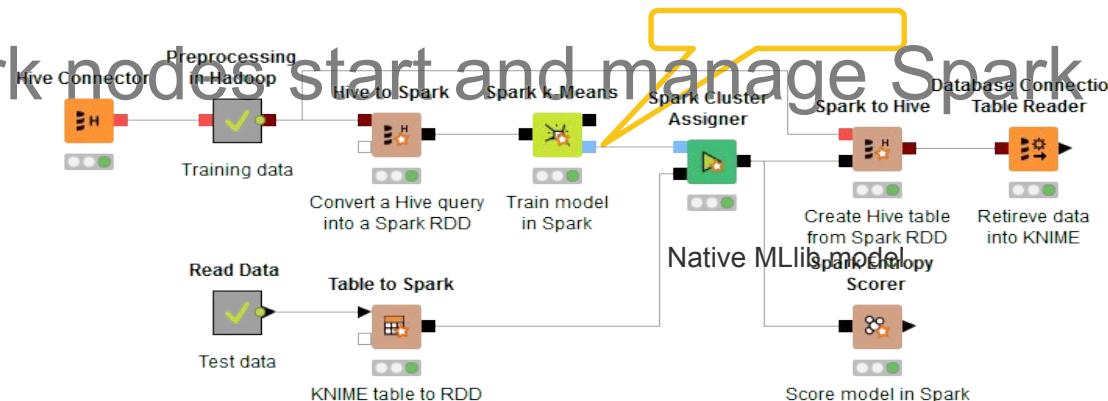
- Preconfigured connectors
  - Hive
  - Cloudera Impala
- Extends the open source database



# KNIME Spark Executor



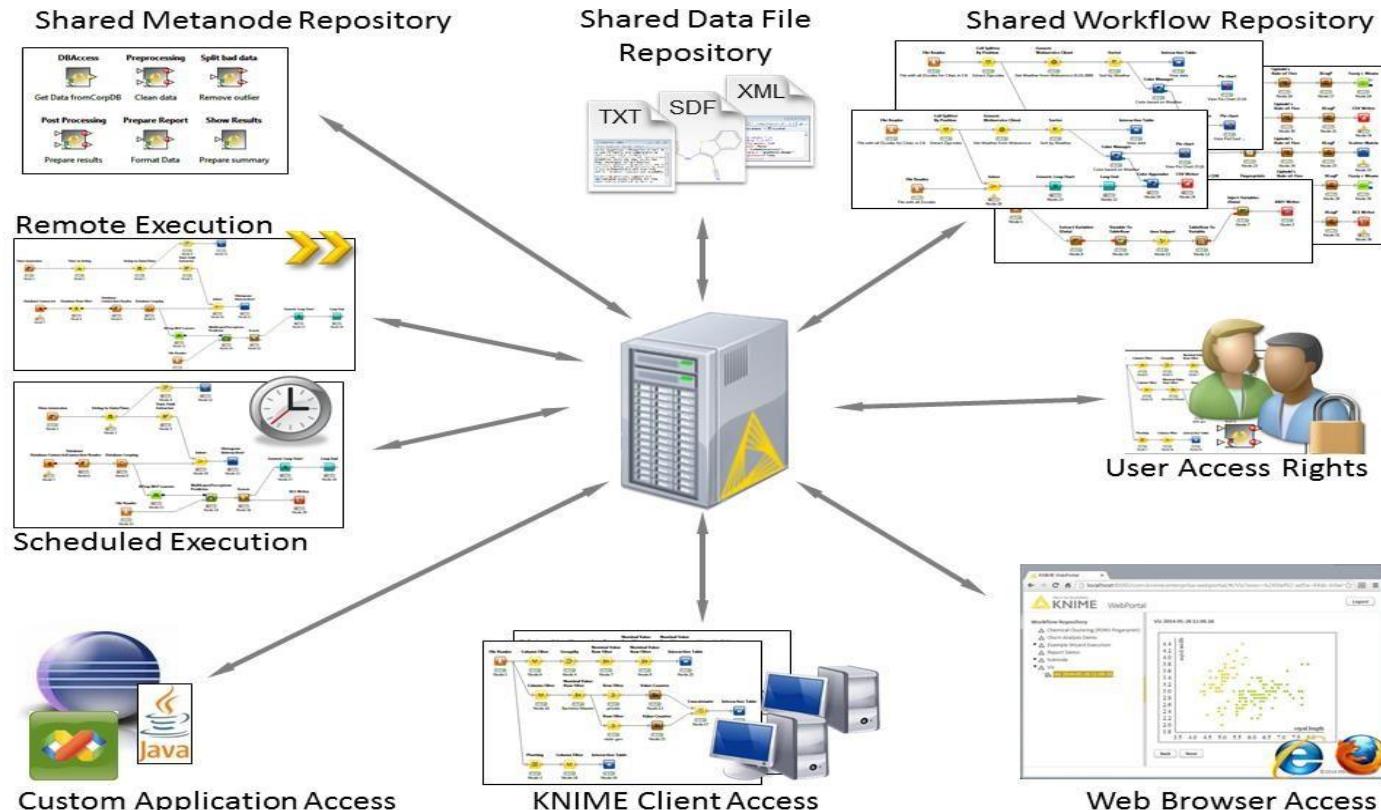
- Advanced Analytics on Hadoop
- Based on Spark MLlib
- Native MLlib model learning and prediction
- Spark nodes start and manage Spark jobs



---

# Why do I need KNIME Server?

# KNIME Server



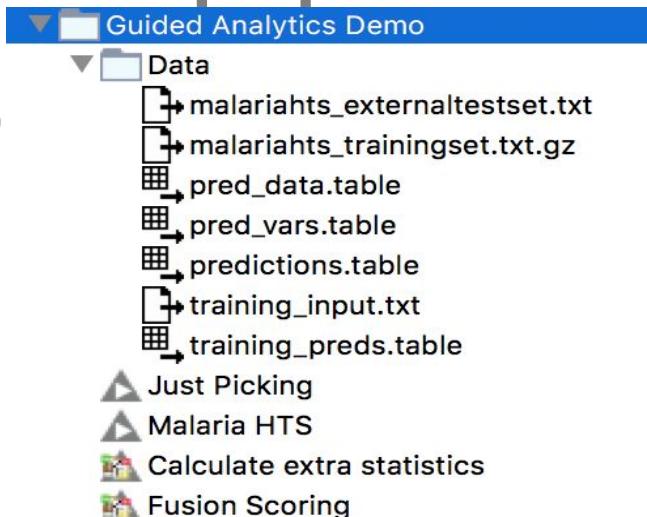
# KNIME Server Provides

---

- Collaboration
- Deployment
- Access Rights
- Scheduling
- REST API
- Guided Analytics

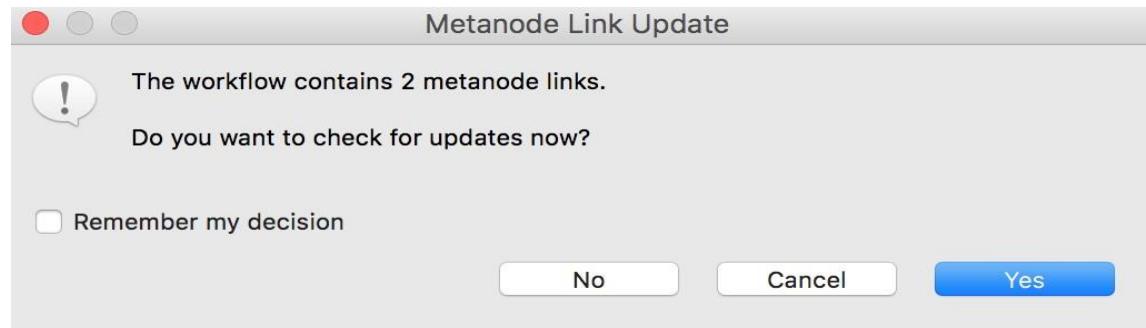
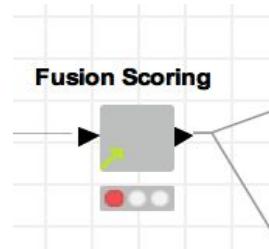
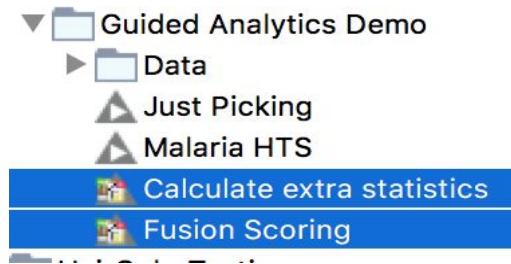
# Server features: Secure workflows and Data files

- Share workflows
- Directory- and workflow-level possible, including LDAP gro



# Server features: Shared metanodes

- Metanodes can be stored on the server and used in workflows as links



# Server features: Remote and scheduled execution

Execute workflow on server

Standard job options | Scheduling options 

Reset before Execution  
 Discard Workflow Job after Execution

Notification Settings  
 Notify upon completion  
thorsten.meinl@knime.com, god@heaven.ly

Report formats to include  
 pdf  html  doc  xls  ppt  ps  odp  odt  ods

Custom job name (default: workflow name plus execution time)

Execute workflow on server

Standard job options | Scheduling options 

Schedule job  
First execution      
 Last execution

Repeat execution  
Repeat every  hours

Repeat job on selected days and months

Days of week	Days of month	Months
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6
<input type="checkbox"/> 7	<input type="checkbox"/> 8	<input type="checkbox"/> 9
<input type="checkbox"/> 10	<input type="checkbox"/> 11	<input type="checkbox"/> 12
<input type="checkbox"/> 13	<input type="checkbox"/> 14	<input type="checkbox"/> 15
<input type="checkbox"/> 16	<input type="checkbox"/> 17	<input type="checkbox"/> 18
<input type="checkbox"/> 19	<input type="checkbox"/> 20	<input type="checkbox"/> 21
<input type="checkbox"/> 22	<input type="checkbox"/> 23	<input type="checkbox"/> 24
<input type="checkbox"/> 25	<input type="checkbox"/> 26	<input type="checkbox"/> 27
<input type="checkbox"/> 28	<input type="checkbox"/> 29	<input type="checkbox"/> 30
<input type="checkbox"/> 31	<input checked="" type="checkbox"/> Last	<input type="button" value="Toggle all"/>

Skip execution if previous job is still running  
 Disable schedule  
Next execution at Di 31.05.16 01:00

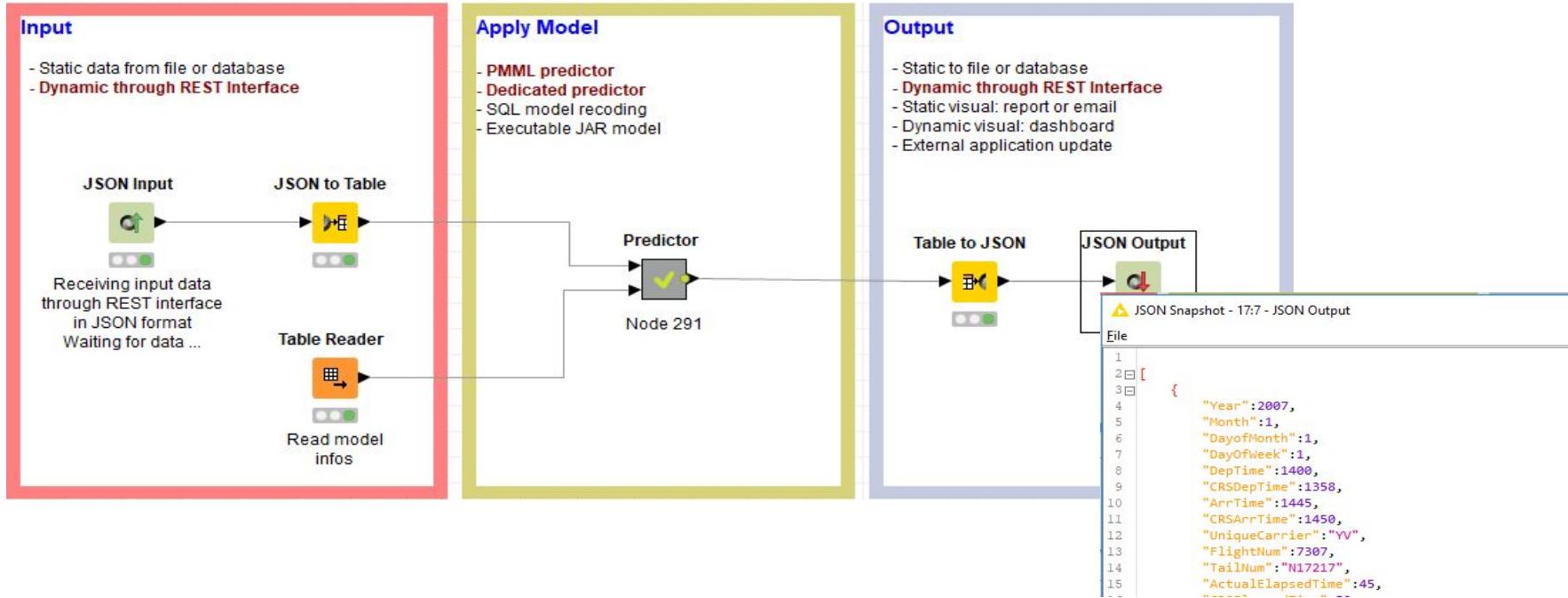
# Server features: Workflow diff

- Versions can be compared to each other:

Node Settings Comparison					
RDKit Descriptor Calculation (88)			RDKit Descriptor Calculation (88)		
Name	Type	Value	Name	Type	Value
<b>▼ Node Settings</b>					
► input_column_Inter...	sub-config		► input_column_Inter...	sub-config	
input_column	string	smiles	input_column	string	smiles
► selectedDescriptor...	sub-config		► selectedDescriptor...	sub-config	
▼ selectedDescriptors	sub-config		▼ selectedDescriptors	sub-config	
array-size	int	22	array-size	int	24
0	string	slogp	0	string	slogp
1	string	LabuteASA	1	string	LabuteASA
2	string	TPSA	2	string	TPSA
3	string	AMW	3	string	AMW
4	string	ExactMW	4	string	ExactMW
5	string	NumLipinskiHBA	5	string	NumLipinskiHBA
6	string	NumLipinskiHBD	6	string	NumLipinskiHBD
7	string	NumRotatableBonds	7	string	NumRotatableBonds
8	string	NumAmideBonds	8	string	NumHBD
9	string	NumHeteroAtoms	9	string	NumHBA
10	string	NumHeavyAtoms	10	string	NumAmideBonds
11	string	NumRings	11	string	NumHeteroAtoms
12	string	NumAromaticRings	12	string	NumHeavyAtoms

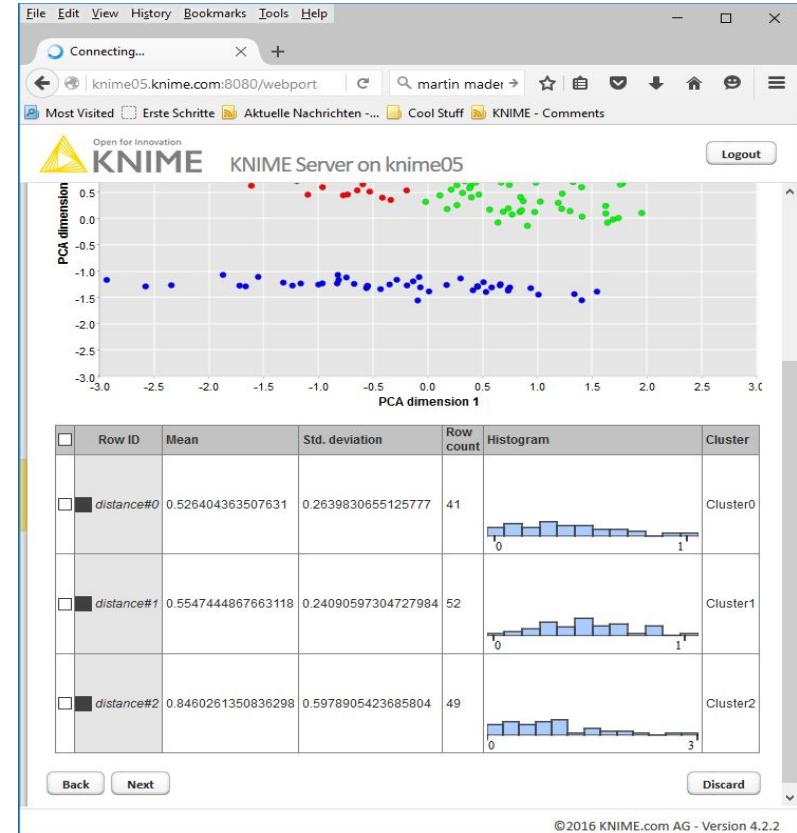
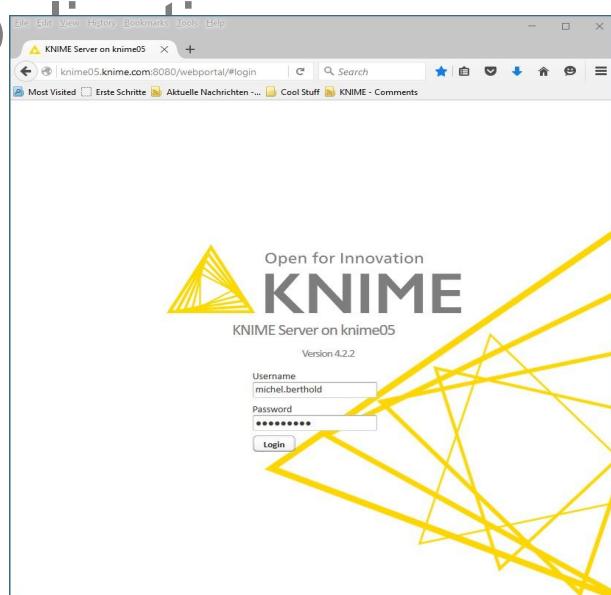
# Server features: Web services

- Deploy workflows as REST-based web



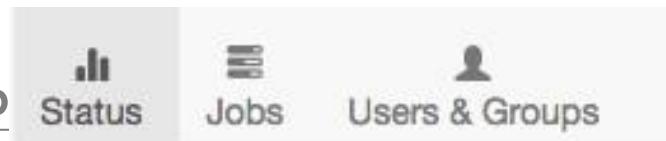
# KNIME Server features: Web portal

- Allows workflows to be used as web apps



# KNIME Server features: User management

- User Authentication – DB based authentication
  - Recommended for small groups with no LDAP access
  - The WebPortal can be used to add/remove users
- User Authentication – LDAP authentication
  - LDAP integration
  - Support for Kerberos and LDAP



# KNIME Software in the Cloud



- Get started quickly
- Get closer to your data
- Connectors for AWS S3 & Azure Blob Store
- Scale to fit your workloads

# KNIME Server Summary

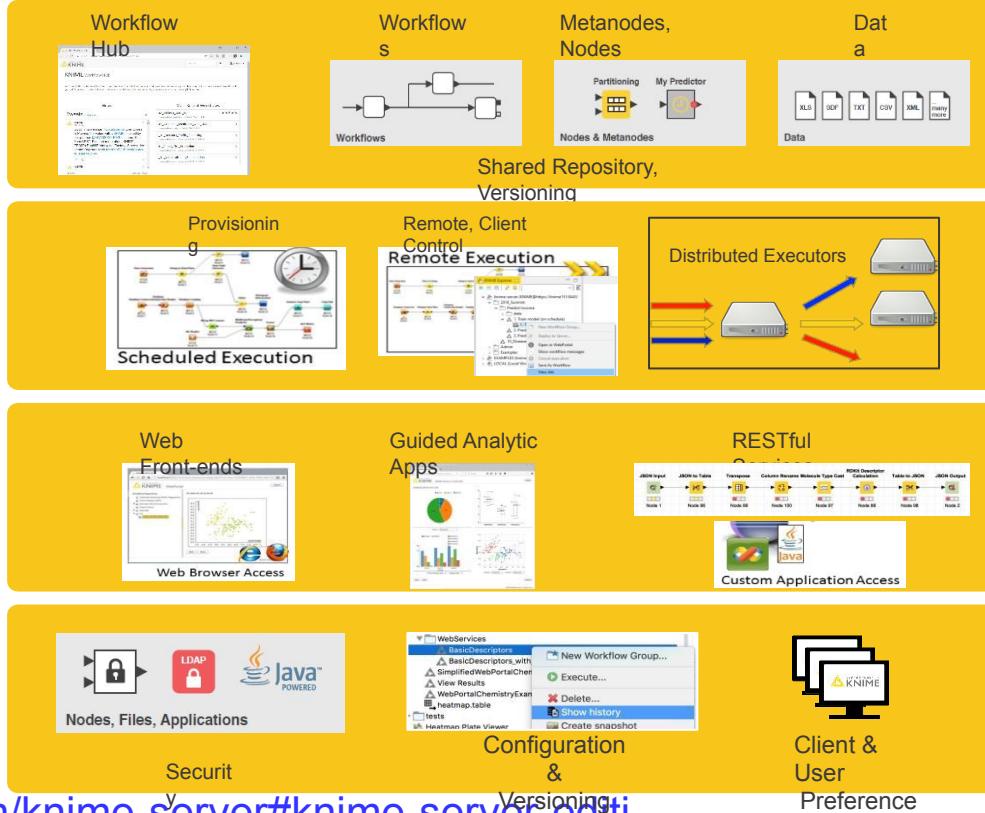
Collaboration

Automation

Deployment

Management

<https://www.knime.com/knime-server#knime-server-editions>



Guided Analytics for Everybody  
Marketing and Sales Decisions

Financial Decisions

Product R&D

Decisions Manufacturing Decisions

IT Decisions

---

## 7. KNIME GUI

# Overview

---

- Installing KNIME Analytics Platform
- The KNIME Workspace
- The KNIME File Extensions
- The KNIME Workbench
  - Workflow editor
  - Explorer
  - Node Repository
  - Description
- Installing new features

# Install KNIME Analytics Platform

- Select the KNIME version for your computer:
  - Mac
  - Windows – 32 or 64 bit
  - Linux
- Download archive and extract the file, or download installer package and run it

Windows	
KNIME Analytics Platform for Windows (installer) <i>The installer adds an icon to the desktop and suggests suitable memory settings</i>	<a href="#">32 Bit</a> (393.38 MB) <a href="#">64 Bit</a> (396.38 MB)
KNIME Analytics Platform for Windows (self-extracting archive) <i>The self-extracting archive only creates a folder holding the KNIME installation</i>	<a href="#">32 Bit</a> (396.87 MB) <a href="#">64 Bit</a> (400.72 MB)
KNIME Analytics Platform for Windows (zip archive)	<a href="#">32 Bit</a> (466.11 MB) <a href="#">64 Bit</a> (470.07 MB)

Linux	
KNIME Analytics Platform for Linux	<a href="#">64 Bit</a> (417.21 MB)

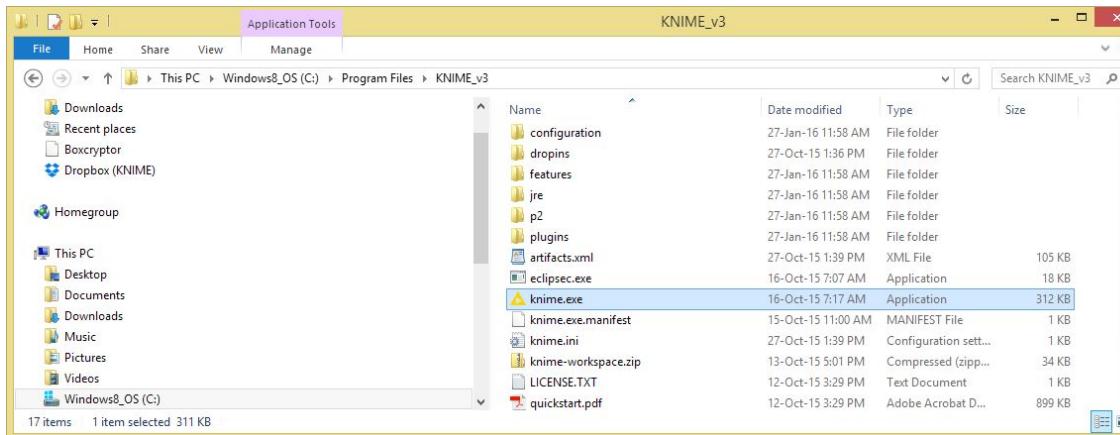
Mac	
KNIME Analytics Platform for Mac OSX (10.11 and above)	<a href="#">64 Bit</a> (388.44 MB)

# Start KNIME Analytics Platform

- Use the shortcut created by the installer

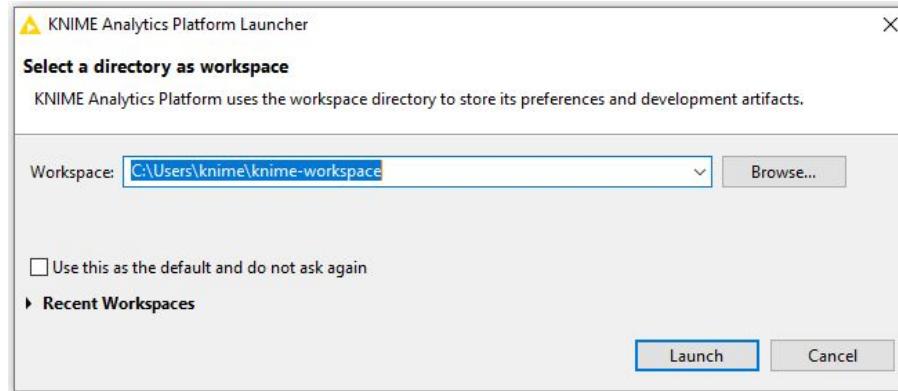


- Or go to the installation directory and launch KNIME via the knime.exe

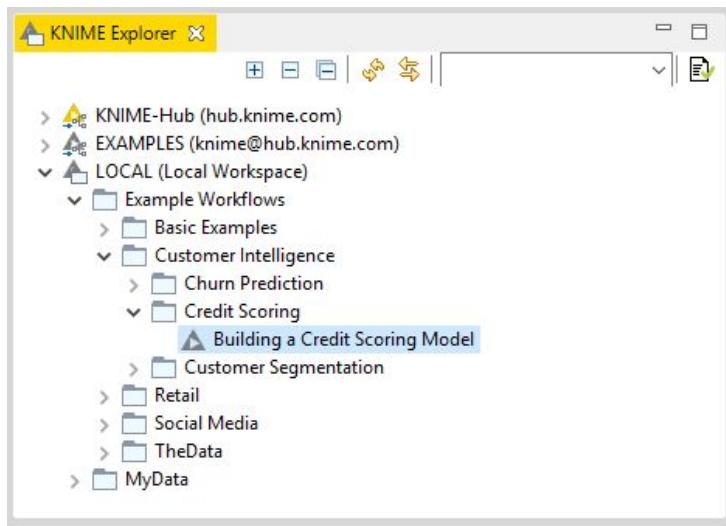


# The KNIME Workspace

- The workspace is the **folder/directory** in which workflows (and potentially data files) are stored for the current KNIME session.
- Workspaces are portable (just like KNIME)



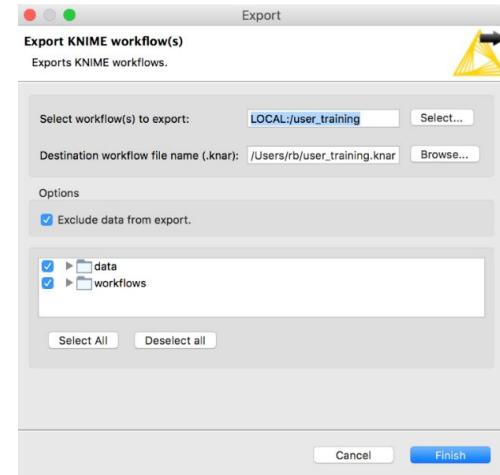
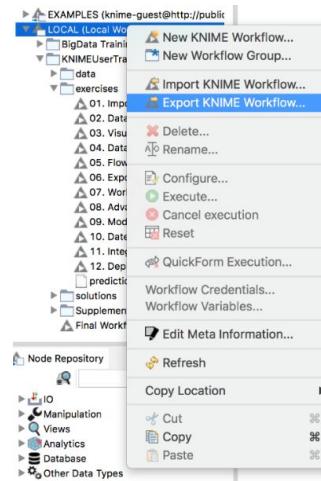
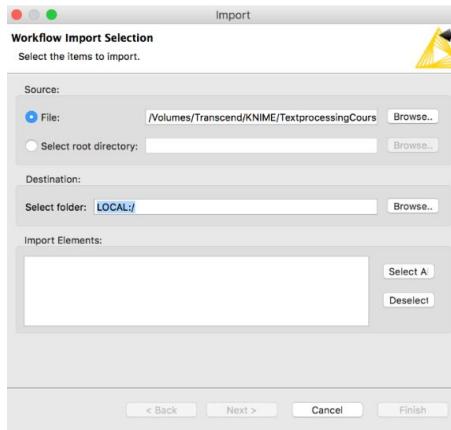
# KNIME Explorer



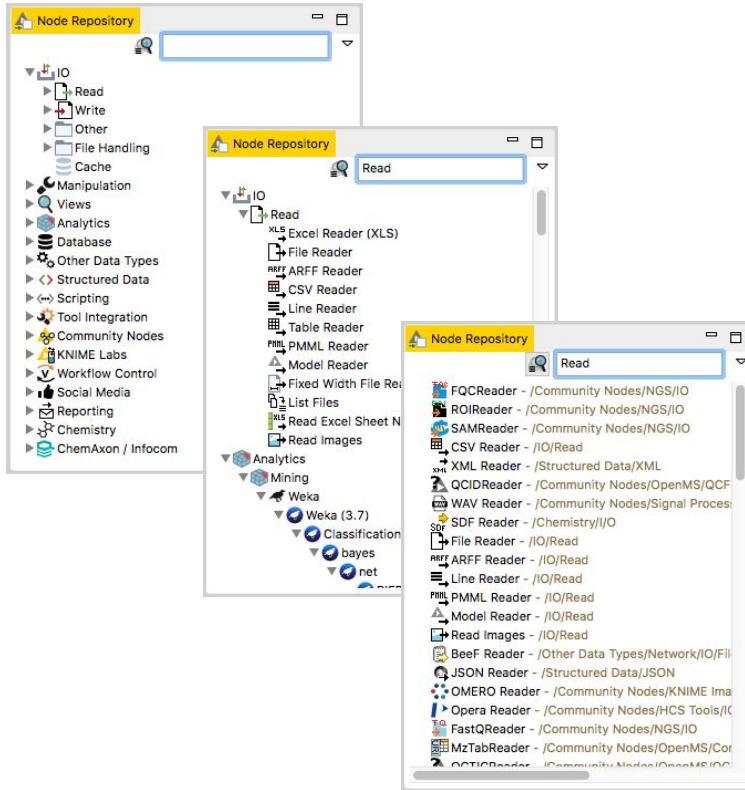
- In LOCAL you can access your own workflow projects.
- The Explorer toolbar on the top has a search box and buttons to
  - select the workflow displayed in the active editor
  - refresh the view
- The KNIME Explorer can contain 4 types of content:
  - Workflows
  - Workflow groups
  - Data files
  - Shared Components

# Creating New Workflows, Importing and Exporting

- Right-click in KNIME Explorer to create new workflow or workflow group or to import workflow
- Right-click on workflow or workflow group to export

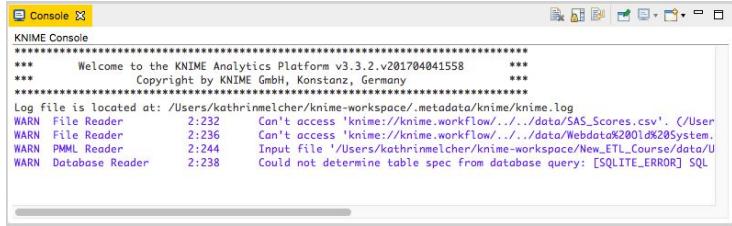


# Node Repository

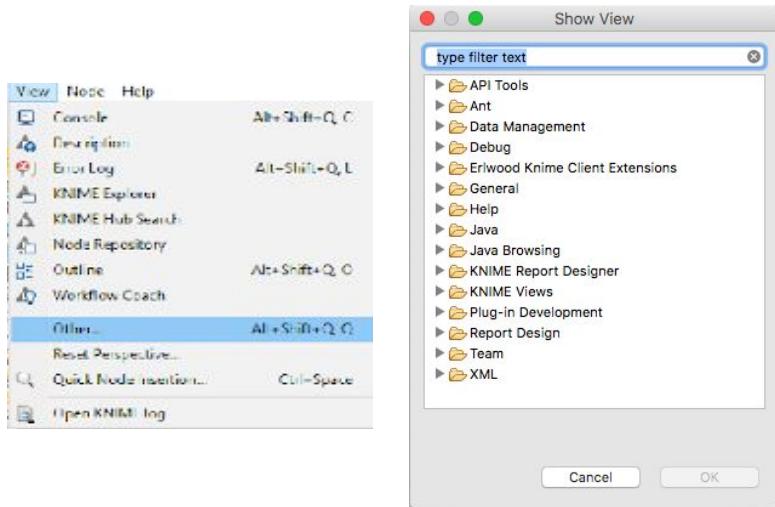


- The Node Repository lists all KNIME nodes
- The search box has 2 modes
  - Standard Search – exact match of node name
  - Fuzzy Search – finds the most similar node name
- Nodes can be added by drag and drop from the Node Repository to the Workflow Editor.

# Console and Other Views

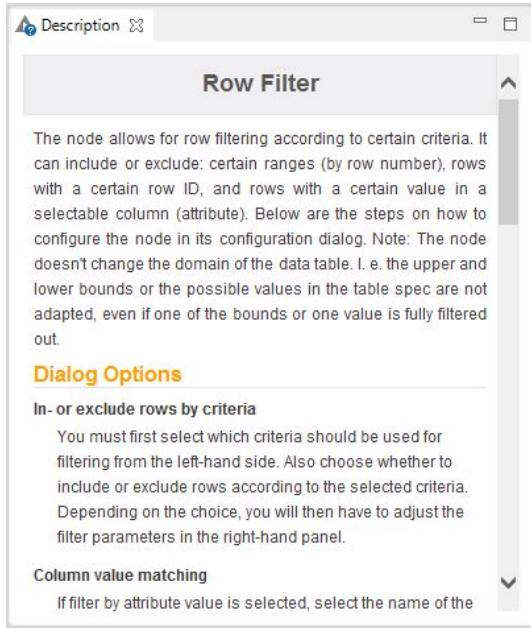


- Console view prints out error and warning messages about what is going on under the hood



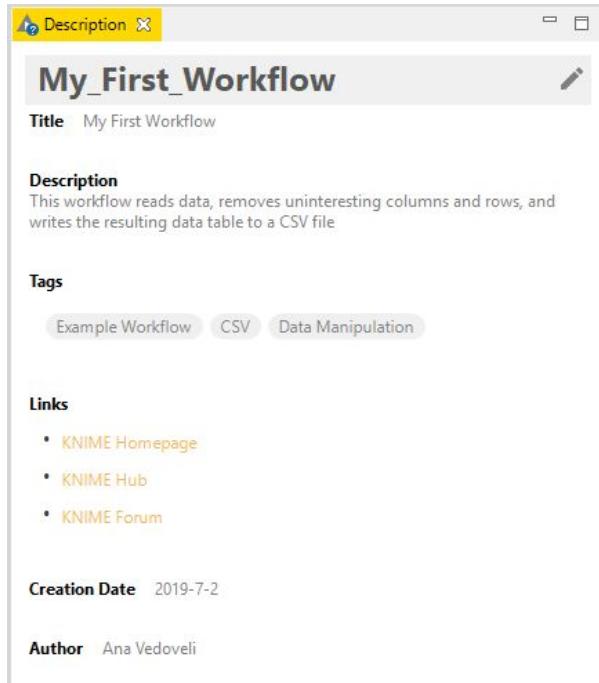
- Click on View and select Other... to add different views
  - Node Monitor, Licenses, etc.

# Description



- The Description window gives information about:
  - Node Functionality
  - Input & Output
  - Node Settings
  - Ports
  - References to literature

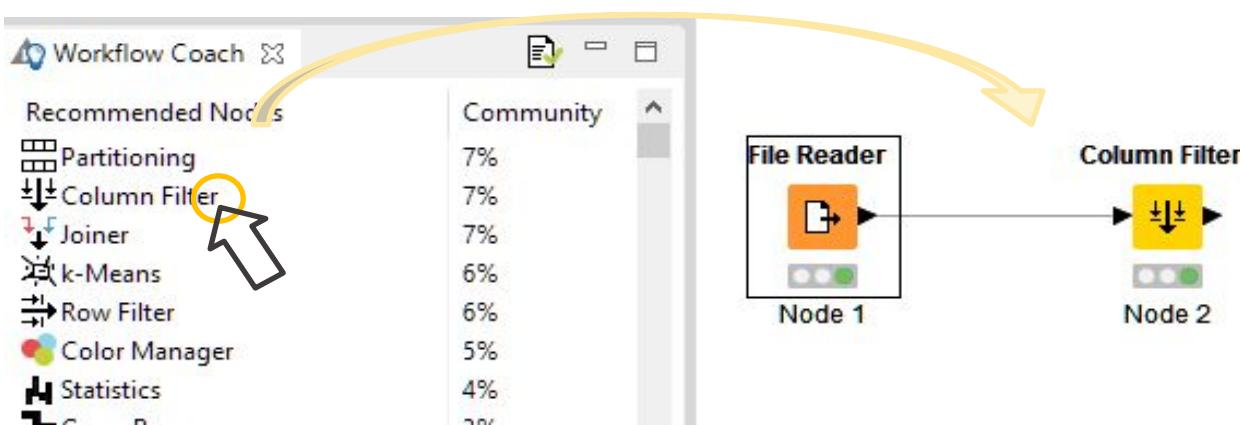
# Workflow Description



- When selecting the workflow, the Description window gives information about the workflow's:
  - Title
  - Description
  - Associated Tags and Links
  - Creation Date
  - Author

# Workflow Coach

- Node recommendation engine
  - Gives hints about which node use next in the workflow
  - Based on KNIME communities' usage statistics
  - Based on own KNIME workflows



# Tool Bar



The buttons in the toolbar can be used for the active workflow. The most important buttons:

- Execute selected and executable nodes (F7)
- Execute all executable nodes
- Execute selected nodes and open first view
- Cancel all selected, running nodes (F9)
- Cancel all running nodes

# KNIME File Extensions

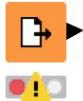
- Dedicated file extensions for Workflows and Workflow groups associated with KNIME Analytics Platform
- **\*.knwf** for KNIME Workflow Files
- **\*.knar** for KNIME Archive Files



# More on Nodes...

A node can have 3 states:

File Reader



## Not Configured:

The node is waiting for configuration or incoming data.

File Reader



## Configured:

The node has been configured correctly, and can be executed.

File Reader

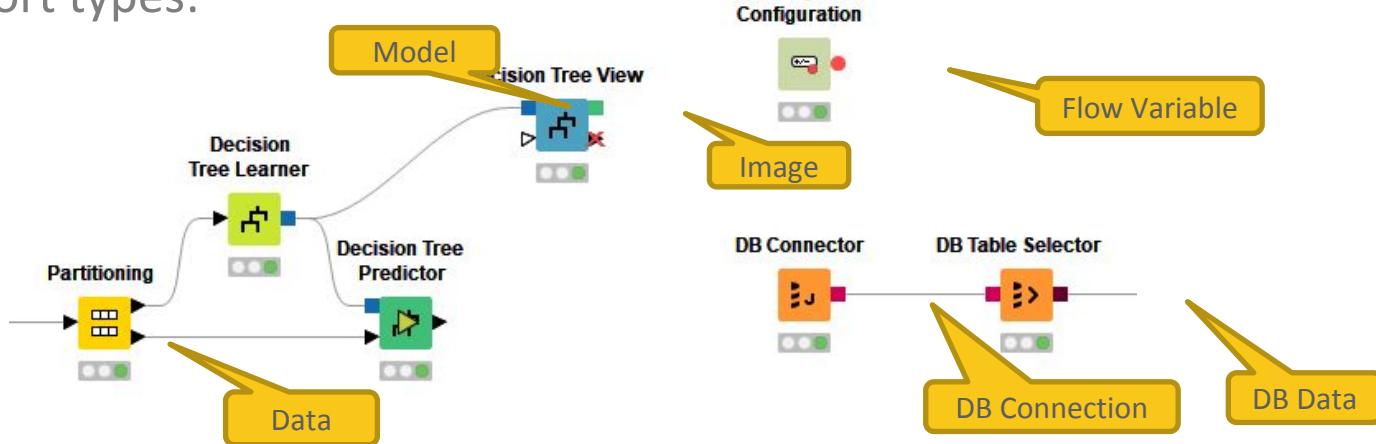


## Executed:

The node has been successfully executed. Results may be viewed and used in downstream nodes.

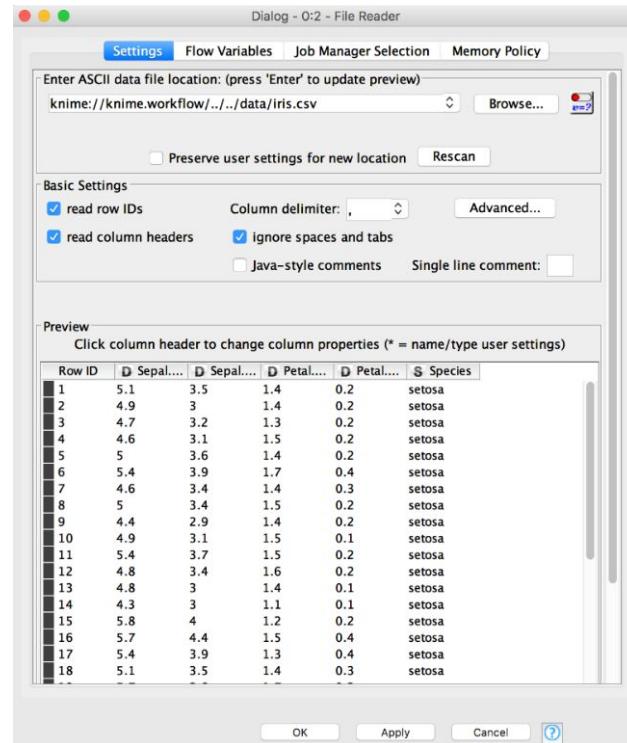
# Inserting and Connecting Nodes

- Insert nodes into workspace by dragging them from Node Repository or by double-clicking in Node Repository
- Connect nodes by left-clicking output port of Node A and dragging the cursor to (matching) input port of Node B
- Common port types:



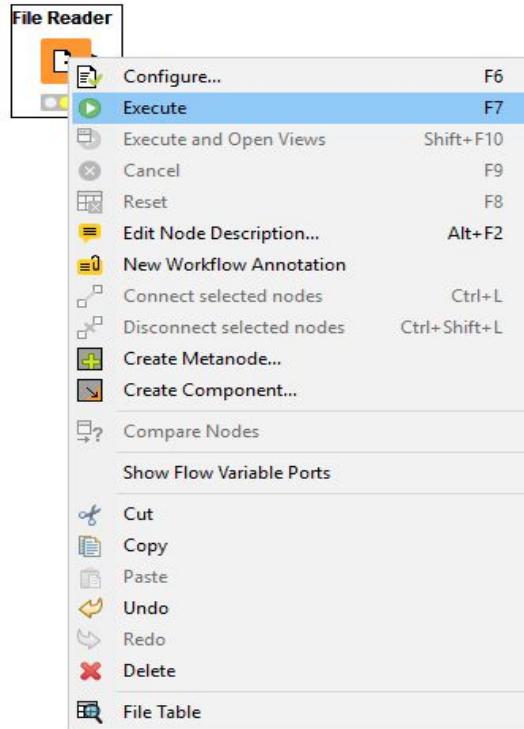
# Node Configuration

- Most nodes require configuration
- To access a node configuration window:
  - Double-click the node
  - Right-click -> Configure



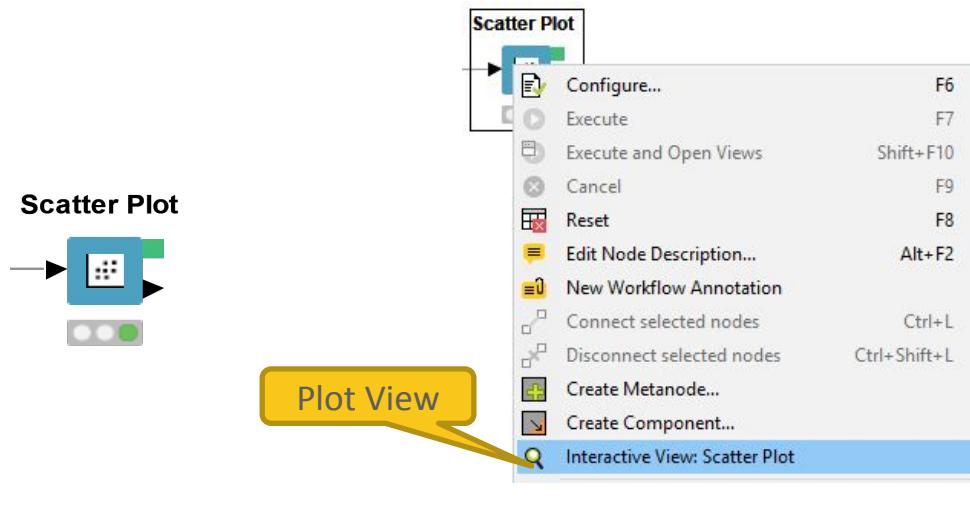
# Node Execution

- Right-click node
- Select Execute in the context menu
- If execution is successful, status shows green light
- If execution encounters errors, status shows red light

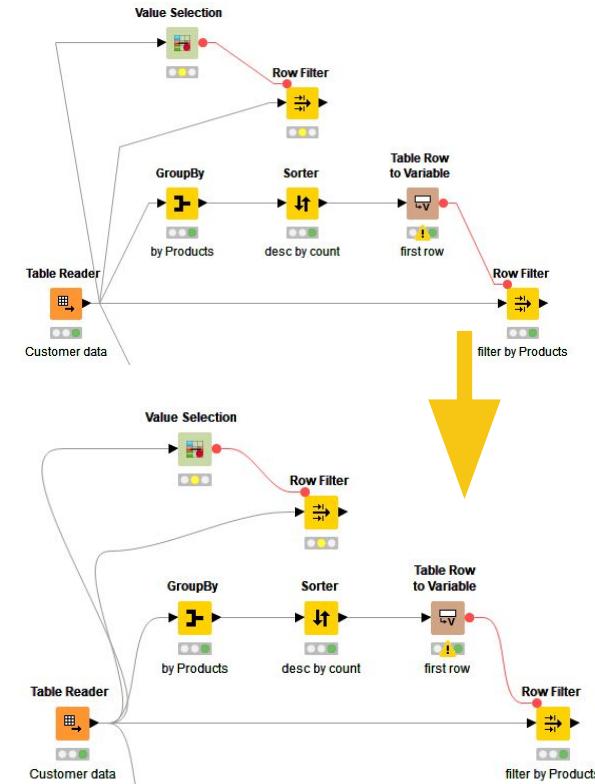
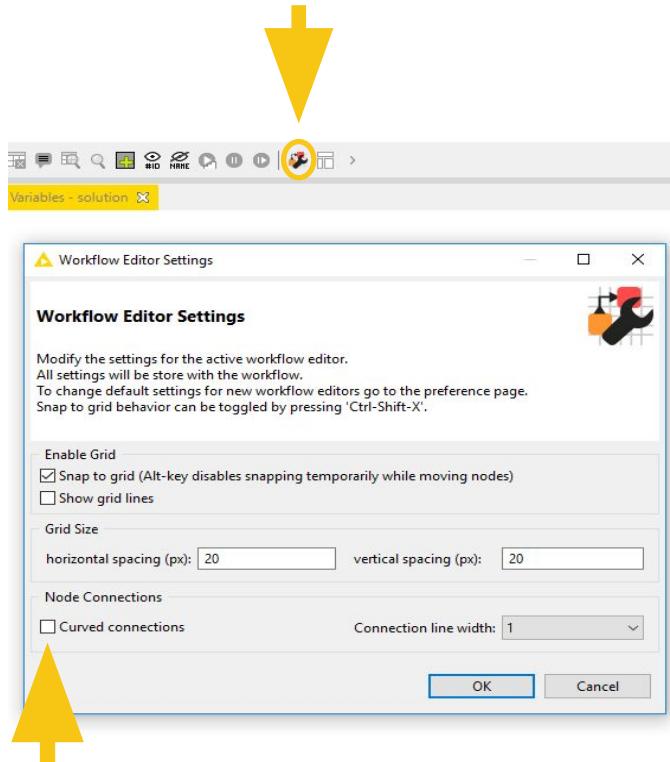


# Node Views

- Right-click node
- Select Views in context menu
- Select output port to inspect execution results

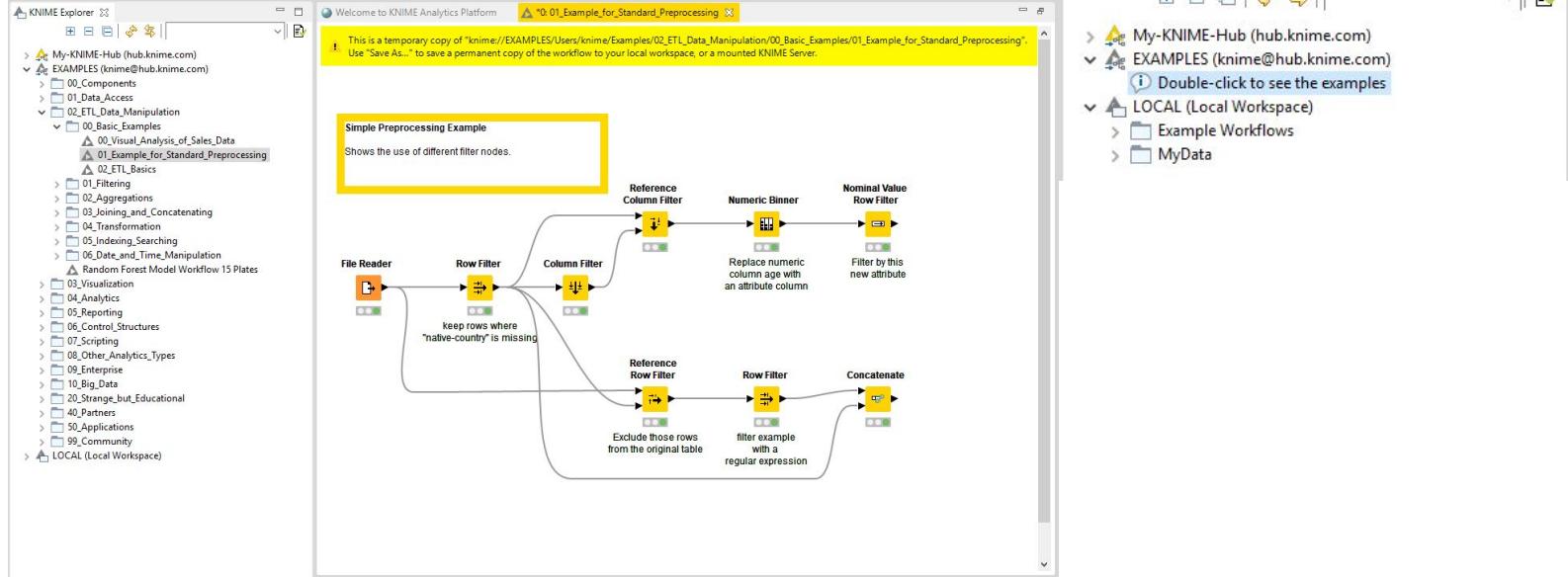


# Curved Connections!



# Getting Started: KNIME Example Server

- Connect via KNIME Explorer to a public repository with large selection of example workflows for many, many applications
- Workflows also available on KNIME Hub



# Sharing Workflows

How to use the KNIME Hub

# KNIME Hub

Welcome to the **KNIME Hub**

The place to find and collaborate on KNIME workflows and nodes. Here you can find solutions for your data science questions.

Search workflows, nodes and more...

**Training a Churn Predictor**

This workflow is an example of how to build a basic PNaive Bayes model for a churn prediction using a Decision Tree.

**External Resources**

→ Churn Prediction  
→ Extensibility  
→ Nodes  
→ KNIME Excel Support  
by KNIME AG, Zürich, Switzerland  
v1.0.0 or later

**Scatter Plot**

A scatter plot using a JavaScript based charting library. The view can be accessed either via the "Interactive view" action on the executed node or in KNIME Server web port page.

The configuration of the node lets you choose the size of a sample to display and to enable certain controls, which are then available in the view. This includes the ability to zoom the displayed area, pan x and y or the possibility to set a limit. Enabling or disabling these controls via the configuration allows making the first usage of the benefits when used in a web portal/remote execution where the end user has no access to the workflow itself.

Since missing values as well as NaN (not a number) or infinite values cannot be displayed in the view, they will be omitted with a corresponding warning message.

Additionally a static SVG image can be rendered, which is then made available at the first output port.

Note: this node is currently under development. Future versions of the node might have more or changed functionality.

**Ports**

**Input Ports**

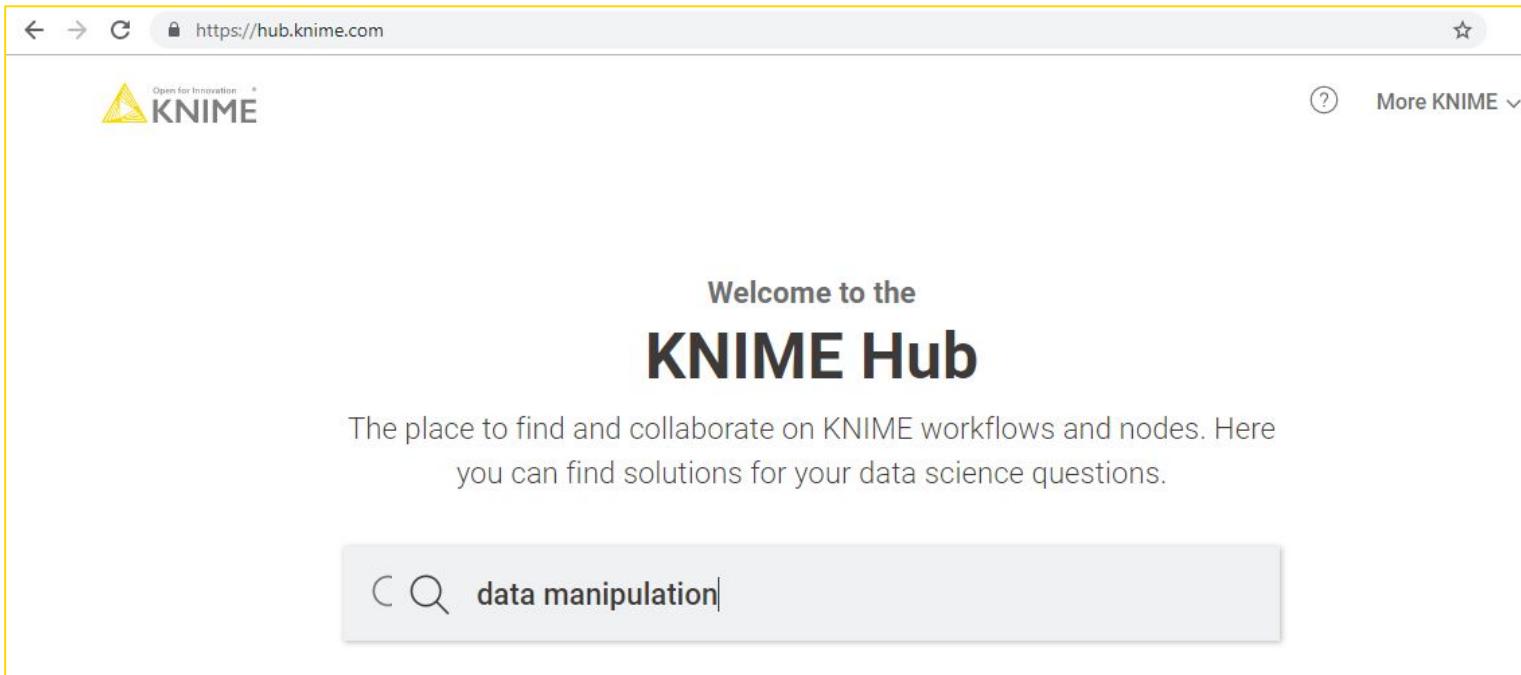
- Type Data  
Data table with data to display.

**Output Ports**

- Type Image  
SVG image rendered by the JavaScript implementation of the scatter plot.
- Type Data  
Data table containing the input data appended with a column, that represents the selection made in the scatter plot view.

A place to share knowledge about  
Workflows and Nodes  
<https://hub.knime.com>

# The KNIME Hub

A screenshot of the KNIME Hub homepage. The page has a white background with a yellow border. At the top, there is a header bar with a back arrow, forward arrow, refresh button, a lock icon indicating https://hub.knime.com, and a star icon. Below the header is the KNIME logo (a yellow triangle with the word "KNIME" next to it) and a "More KNIME" dropdown menu. The main content area starts with a "Welcome to the" message, followed by a large bold heading "KNIME Hub". Below the heading, there is a descriptive text: "The place to find and collaborate on KNIME workflows and nodes. Here you can find solutions for your data science questions." At the bottom of the main content area is a search bar containing a magnifying glass icon, the letter "C", and the text "data manipulation".

Welcome to the

# KNIME Hub

The place to find and collaborate on KNIME workflows and nodes. Here you can find solutions for your data science questions.

C Q data manipulation

# Searching Nodes and Workflows

The screenshot shows the KNIME search interface with the following details:

- Header:** KNIME Open for Innovation logo, search bar containing "data manipulation", and a "More KNIME" dropdown.
- Search Results:** A large bold number **3140** results.
- Filter Options:** Buttons for All, Nodes, and Workflows.
- Result Cards:** Three cards are visible, each with a yellow dashed border around its title and description.
  - String Manipulation:** Streamable node. Description: Manipulates strings like search and replace, capitalize or remove leading and trailing white spaces. Examples: To remove leading and trailing blanks from a column with name c0 you would use the expres...  
Category: Manipulator.
  - String Manipulation, Math Formula and Rule Engine Example:** Workflow. Description: This workflow shows three different data manipulation operations, namely: - creating three categories of people based on their weekly work hours with the Rule Engine node - rounding up people's age to ...  
Tags: ETL, data manipulation, string manipulation, strings, numbers, math, math formula, data transformation, data wrangling, rules, rule engine.  
Category: Manipulator.
  - String Manipulation (Variable):** Node. Description: Manipulates or defines values of variables like search and replace, capitalize or remove leading and trailing white spaces. Examples: To remove leading and trailing blanks from a variable with name c0...  
Category: Manipulator.

# Opening a Workflow from the Hub

Open for Innovation **KNIME**

Search workflows, nodes and more...

More KNIME ▾

## String Manipulation, Math Formula and Rule Engine Example

**String Manipulation, Math Formula and Rule Engine Example**

This workflow shows three different data manipulation operations, namely:

- creating three categories of people based on their weekly work hours with the Rule Engine node
- rounding up people's age to the nearest 10 with the Math Formula node
- replacing hyphens with " " characters in the native country column

```
graph LR; A[File Reader] --> B[Rule Engine]; B --> C[Math Formula]; C --> D[String Manipulation]
```

The diagram illustrates a sequential workflow. It starts with a 'File Reader' node (orange square with a file icon) labeled 'Read adult data'. An arrow points to a 'Rule Engine' node (green square with a checkmark icon) labeled 'New column for work status:  
- unemployed  
- employed part-time  
- employed full-time'. Another arrow points to a 'Math Formula' node (yellow square with a calculator icon) labeled 'Round up age to the nearest 10'. A final arrow points to a 'String Manipulation' node (yellow square with a text icon) labeled 'Replace "-" with " "' in native country values'.

Maarit  
KNIME Team Member

Basic Silver Anniversary First Share 5 more

hosted by

KNIME

**Open workflow** or download workflow

By downloading the workflow, you agree to our [terms and conditions](#).

CC-BY-4.0

Short Link

<https://knime.me/w/pyg3vLc9BL4sJPd>

# Open Workflow in KNIME Analytics Platform

The screenshot shows the KNIME Analytics Platform interface. The top menu bar includes File, Edit, View, Node, and Help. The left sidebar contains the KNIME Explorer (with entries for My-KNIME-Hub, EXAMPLES, and LOCAL), Node Repository (with categories like IO, Manipulation, Views, Analytics, DB, etc.), and a search bar. The main workspace displays a workflow titled "3: 02\_StringManipulation\_MathFormula\_RuleEngine". A yellow box highlights the title and a descriptive text block:

This workflow shows three different data manipulation operations, namely:

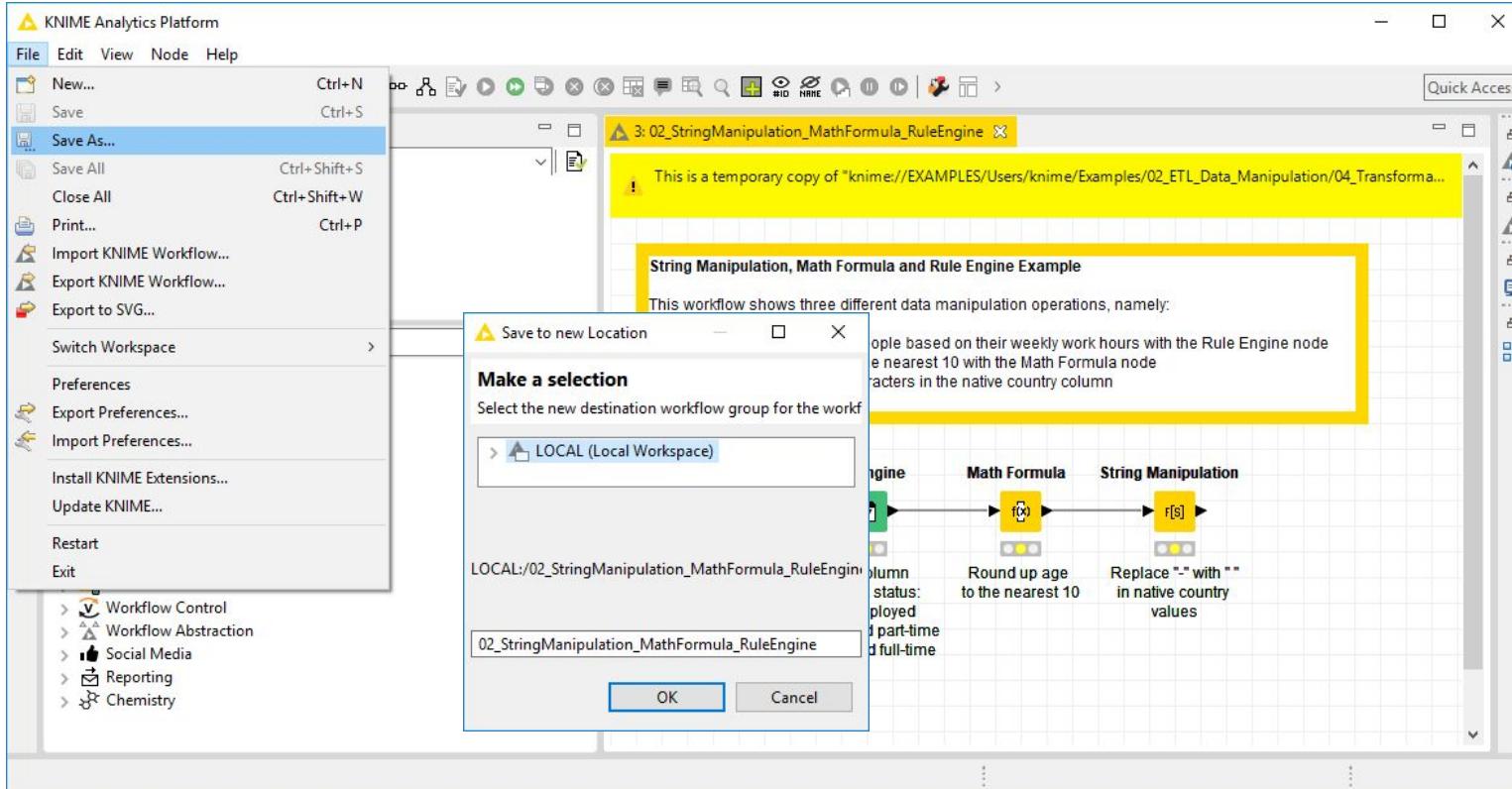
- creating three categories of people based on their weekly work hours with the Rule Engine node
- rounding up people's age to the nearest 10 with the Math Formula node
- replacing hyphens with "" characters in the native country column

The workflow diagram consists of four nodes connected by arrows:

- File Reader (Read adult data)
- Rule Engine (New column for work status:
  - unemployed
  - employed part-time
  - employed full-time)
- Math Formula (Round up age to the nearest 10)
- String Manipulation (Replace "-" with "" in native country values)

```
graph LR; FR[File Reader] --> RE[Rule Engine]; RE --> MF[Math Formula]; MF --> SM[String Manipulation]
```

# Saving the Workflow



# Edit the Workflow

The image shows two windows side-by-side. On the left is the KNIME Hub browser window, displaying a page about the 'Row Filter' node. The right window is the KNIME Analytics Platform interface, showing a workflow titled 'String Manipulation, Math Formula and Rule Engine Example'. A yellow callout box labeled 'Drag & Drop' points from the bottom left towards the workflow diagram.

**Row Filter**

The node allows for row filtering based on various criteria. It can include or exclude rows based on specific conditions (e.g., row number), rows with a certain value in a selectable column, or rows where a certain value in a selectable column is within a range. The node also includes steps for capturing configuration dialog parameters. Note: The node does not filter the domain of the data table. I.e. the upper and lower possible values in the table spec are not adapted, even if one value is fully filtered out.

**Manipulator**

**KNIME Analytics Platform**

**Workflow Overview:**

- File Reader**: Read adult data.
- Rule Engine**: New column for work status:
  - unemployed
  - employed part-time
  - employed full-time
- Math Formula**: Round up age to the nearest 10.
- String Manipulation**: Replace "-" with "" in native country values.

This workflow shows three different data manipulation operations, namely:

- creating three categories of people based on their weekly work hours with the Rule Engine node
- rounding up people's age to the nearest 10 with the Math Formula node
- replacing hyphens with "" characters in the native country column

# Sharing the Workflow

1. Save your Edits

2. Connect to KNIME Hub

This workflow shows three different data manipulation operations, namely:

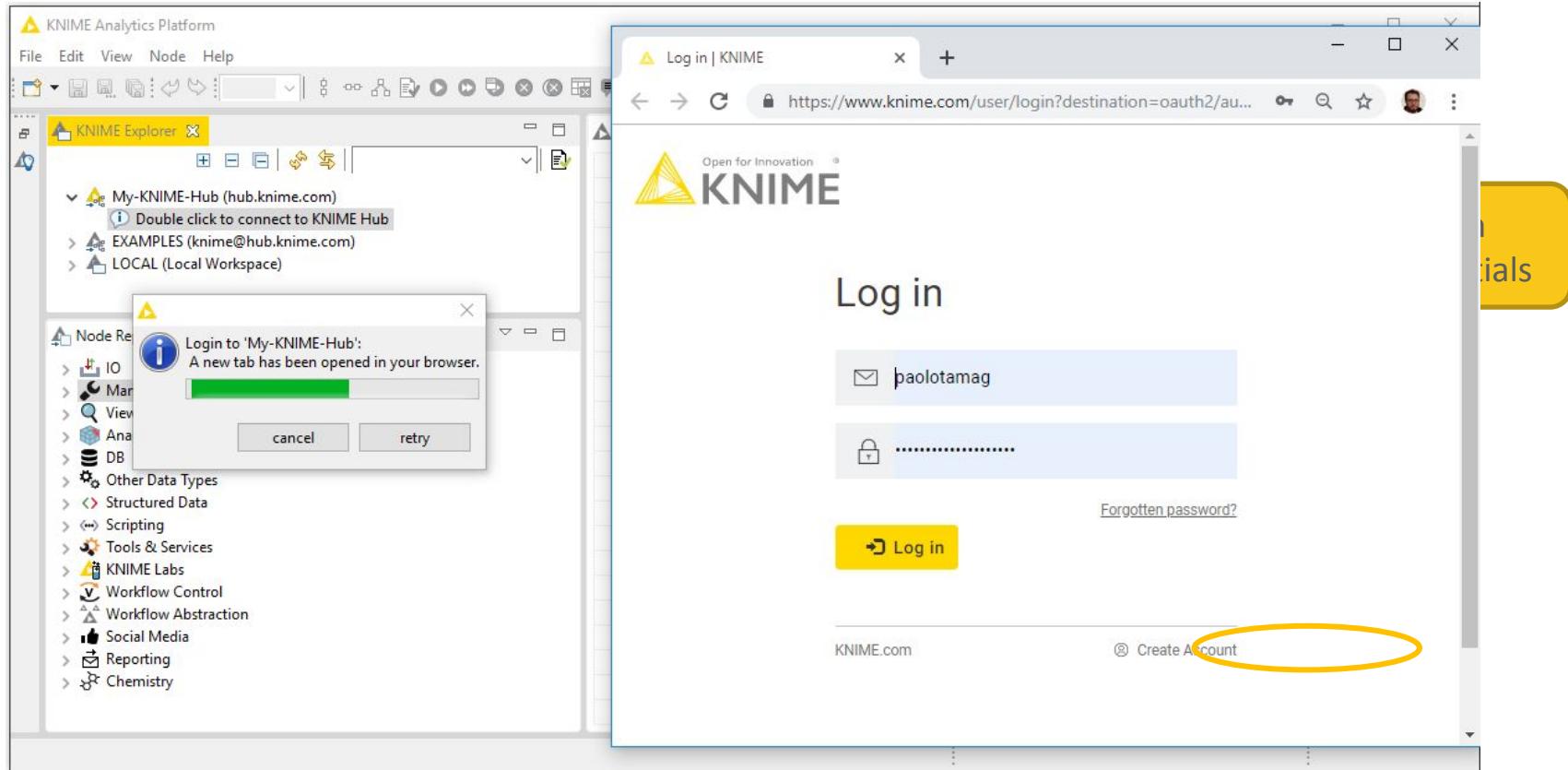
- creating three categories of people based on their weekly work hours with the Rule Engine node
- rounding up people's age to the nearest 10 with the Math Formula node
- replacing hyphens with " " characters in the native country column

File Reader      Row Filter      Rule Engine      Math Formula      String Manipulation

Read adult data      filter data      New column for work status:  
- unemployed  
- employed part-time  
- employed full-time      Round up age to the nearest 10      Replace "-" with " " in native country values

```
graph LR; FR[File Reader] --> RF[Row Filter]; RF --> RE[Rule Engine]; RE --> MF[Math Formula]; MF --> SM[String Manipulation]
```

# Log in the Hub



# Publish your Workflow

The screenshot shows the KNIME Analytics Platform interface. On the left, the KNIME Explorer panel displays a tree structure with nodes like 'My-KNIME-Hub' and 'LOCAL (Local Workspace)'. A yellow callout bubble labeled '2. Drag & Drop' has an arrow pointing to the 'My-KNIME-Hub' node. In the center, a workflow titled '02\_StringManipulation\_MathFormula\_RuleEngine' is shown with a 'File Reader' node connected to a 'Row Filter' node. A yellow callout bubble labeled '1. Edit Metadata' has an arrow pointing to the edit icon in the 'Description' panel on the right. The 'Description' panel contains fields for Title, Description, Tags, Links, Creation Date, and Author, all currently empty or showing placeholder text.

KNIME Analytics Platform

File Edit View Node Help

Quick Access

KNIME Explorer

My-KNIME-Hub (paolotamag@hub.knime.com)

Drag and drop your workflows here to share them publicly on KNIME Hub or double-click to learn more.

EXAMPLES (knime@hub.knime.com)

LOCAL (Local Workspace)

02\_StringManipulation\_MathFormula\_RuleEngine

String Manipulation, Math Formula

This workflow shows three different ways of manipulating strings:

- creating three categories of people
- rounding up people's age to the nearest integer
- replacing hyphens with " " characters

File Reader → Row Filter

Read adult data

filter data

Description

02\_StringManipulation\_Math...

Title Hub Example

Description No description has been set yet.

Tags No tags have been added yet.

Links No links have been added yet.

Creation Date 2018-9-3

Author

Node Repository

2. Drag & Drop

1. Edit Metadata

# Open your Workflow in the Hub

KNIME Analytics Platform

File Edit View Node Help

Quick Access

KNIME Explorer

My-KNIME-Hub (paolotamag@hub.knime.com)

- 02\_String
- EXAMPLES
- LOCAL (Loc)

Open > as Local Copy in KNIME Hub

New Workflow Group...  
Delete...  
Rename...  
Refresh  
Copy Location  
Show Meta Information  
Disconnect  
Compare  
Cut Ctrl+X  
Copy Ctrl+C  
Paste Ctrl+V

String Manipulation, Math Formula and Rule Engine Example

This workflow shows three different data manipulation operations, namely:

- creating three categories of people based on their weekly work hours with the Rule Engine node
- rounding up people's age to the nearest 10 with the Math Formula node
- replacing hyphens with " " characters in the native country column

File Reader → Row Filter → Rule Engine → Math Formula → String Manipulation

Read adult data  
filter data  
New column for work status:  
- unemployed  
- employed part-time  
- employed full-time  
Round up age to the nearest 10  
Replace "-" with " "  
in native country values

Node Repository

IO

```
graph LR; FR[File Reader] --> RF[Row Filter]; RF --> RE[Rule Engine]; RE --> MF[Math Formula]; MF --> SM[String Manipulation];
```

Read adult data  
filter data  
New column for work status:  
- unemployed  
- employed part-time  
- employed full-time  
Round up age to the nearest 10  
Replace "-" with " "  
in native country values

# Open your Workflow in the Hub

The screenshot shows a web browser window for the KNIME Hub at [https://hub.knime.com/paolotamag/space/02\\_StringManipulation\\_MathFormula\\_RuleEngine](https://hub.knime.com/paolotamag/space/02_StringManipulation_MathFormula_RuleEngine). The page title is "Hub Example - KNIME Hub". The main content area displays a workflow titled "String Manipulation, Math Formula and Rule Engine Example". The workflow consists of five nodes connected sequentially: "File Reader" (Read adult data), "Row Filter" (filter data), "Rule Engine" (New column for work status:  
- unemployed  
- employed part-time  
- employed full-time), "Math Formula" (Round up age to the nearest 10), and "String Manipulation" (Replace "." with "" in native country values). A yellow box highlights the "String Manipulation" node. To the right of the workflow, there is a profile for "Paolotamag" (KNIME Team Member) with a picture, a "Regular" badge, a "Silver Anniversary" badge, a "Thoughtful" badge, and a "9 more" badge. Below the profile is a "hosted by" section with the KNIME logo. On the far right, there is a "Download workflow" button, a "CC-BY-4.0" license link, and a "Short Link" button with the URL <https://knime.w/JJIKY8rqMm21jkH>, which is circled in yellow.

# Hot Keys (for Future Reference)

Task	Hot key	Description
Node Configuration	F6	opens the configuration window of the selected node
	F7	executes selected configured nodes
Node Execution	Shift + F7	executes all configured nodes
	Shift + F10	executes all configured nodes and opens all views
	F9	cancels selected running nodes
	Shift + F9	cancels all running nodes
	Ctrl + L	connects selected nodes
Node Connections	Ctrl + Shift + L	disconnects selected nodes
	Ctrl + Shift + Arrow	moves the selected node in the arrow direction
Move Nodes and Annotations	Ctrl + Shift + PgUp/PgDown	moves the selected annotation in the front or in the back of all overlapping annotations
	F8	resets selected nodes
Workflow Operations	Ctrl + S	saves the workflow
	Ctrl + Shift + S	saves all open workflows
	Ctrl + Shift + W	closes all open workflows
	Shift + F12	opens metanode wizard
Metanode		

# Stay connected with KNIME

---



**Blog:** [knime.com/blog](http://knime.com/blog)



**Forum:** [forum.knime.com](http://forum.knime.com)



**KNIME Hub:**  
[hub.knime.com](http://hub.knime.com)



**KNIME E-Learning Course:**  
[www.knime.com/e-learning-course](http://www.knime.com/e-learning-course)

Follow us on social media:





# Let's Talk

Cevi Herdian  
Data-Driven  
Strategist

[itsmecivi.github.io](https://itsmecivi.github.io)

