[L1-AP] Data Literacy with KNIME Analytics Platform: Basics

You are free to:

Share

copy and redistribute the material in any medium or format

Adapt

remix, transform, and build upon the material for any purpose, even commercially.

You must give **appropriate credit**, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

Course Objectives

- At the end of this course, you will be able to
  1. Outline the uses of KNIME Analytics Platform for data analysis and data science
  2. Describe the advantages of using visual programming for data science
  3. Read data from a file and database
  4. Clean and transform data
  5. Merge, aggregate, and export data
  6. Visualize and report results
  7. Take the L1 - Certification exam

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Who is this Course for?

- People who are just starting their data literacy journey
- Data Analysts, Data Scientists and Data Engineers
- Low code/No code & open source advocates

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Structure of the Course

Session	Topic	Duration
Session 1	Introduction to KNIME Analytics Platform and Data Access	75 min
Session 2	Data Cleaning, Data Transformation, and Workflow Documentation	75 min
Session 3	Data Merging, Data Aggregation, and Data Export	75 min
Session 4	Data Visualization and Reporting	75 min
Session 5	Wrap-up	15 min

- Hands on exercises will be provided at the end of each session

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 1
Introduction to KNIME Analytics Platform and Data Access

Learning Objectives

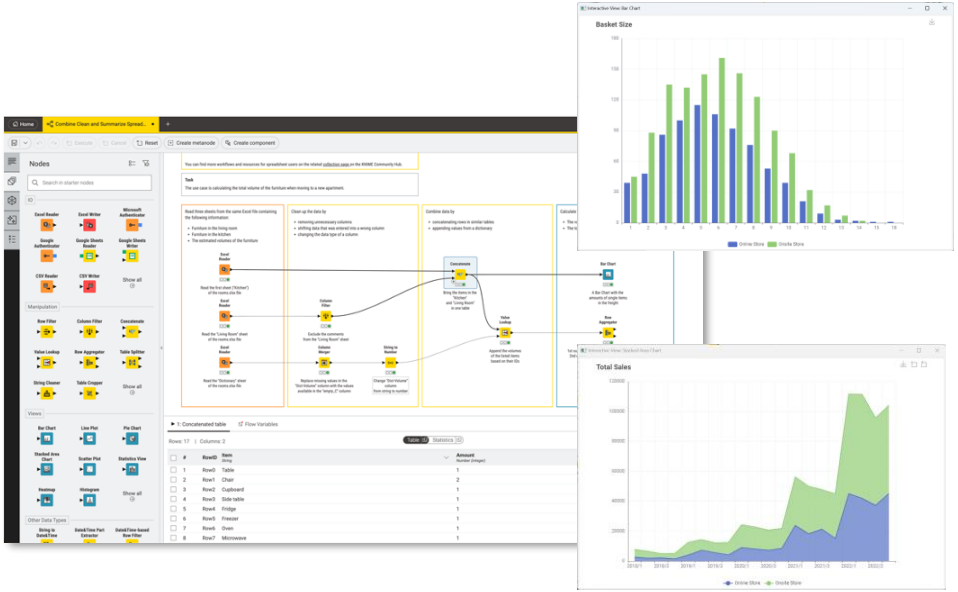
1. Install KNIME Analytics Platform and extensions
2. Navigate and operate the KNIME User Interface
3. Build your first workflow
4. Access and Import workflows from KNIME Community Hub
5. Read different file formats
6. Read data from databases

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Introduction to
KNIME Analytics Platform

What is KNIME Analytics Platform

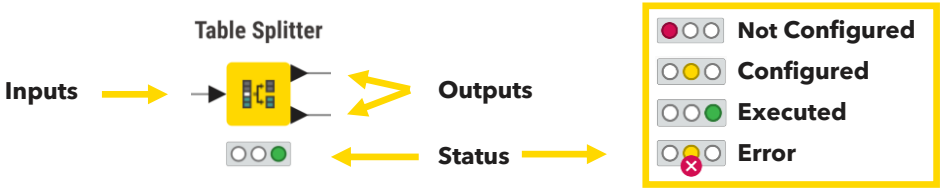
- A tool to make sense of your data
  - Data analysis, Data Science, Data Engineering
  - Transformation, visualization, reporting
- Open source
- Visual programming paradigm
  - No coding required



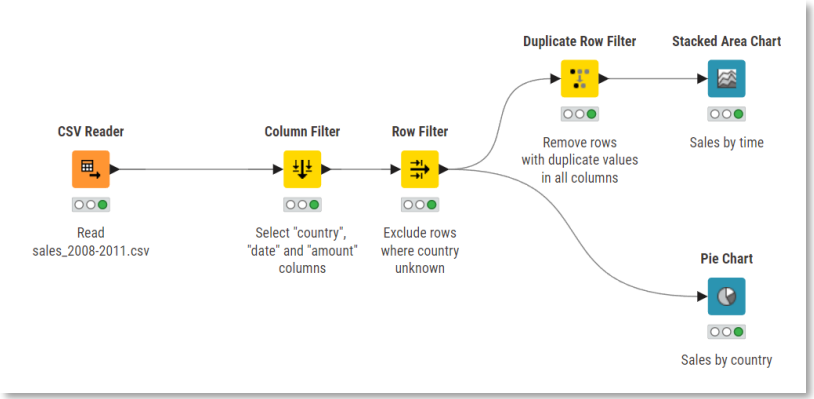
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Nodes and Workflows

- **NODES** perform tasks on data

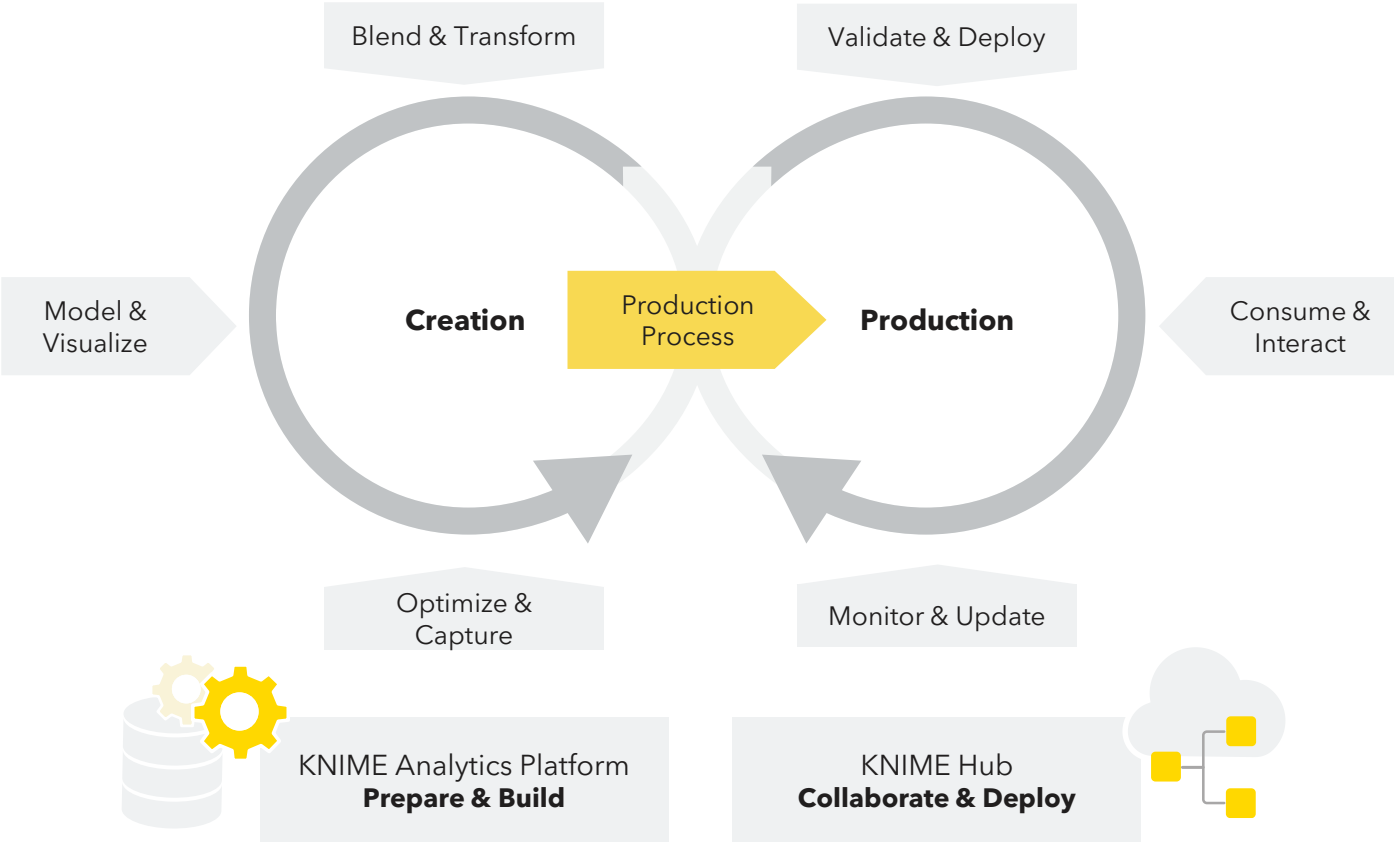


- Nodes are combined to create **WORKFLOWS**



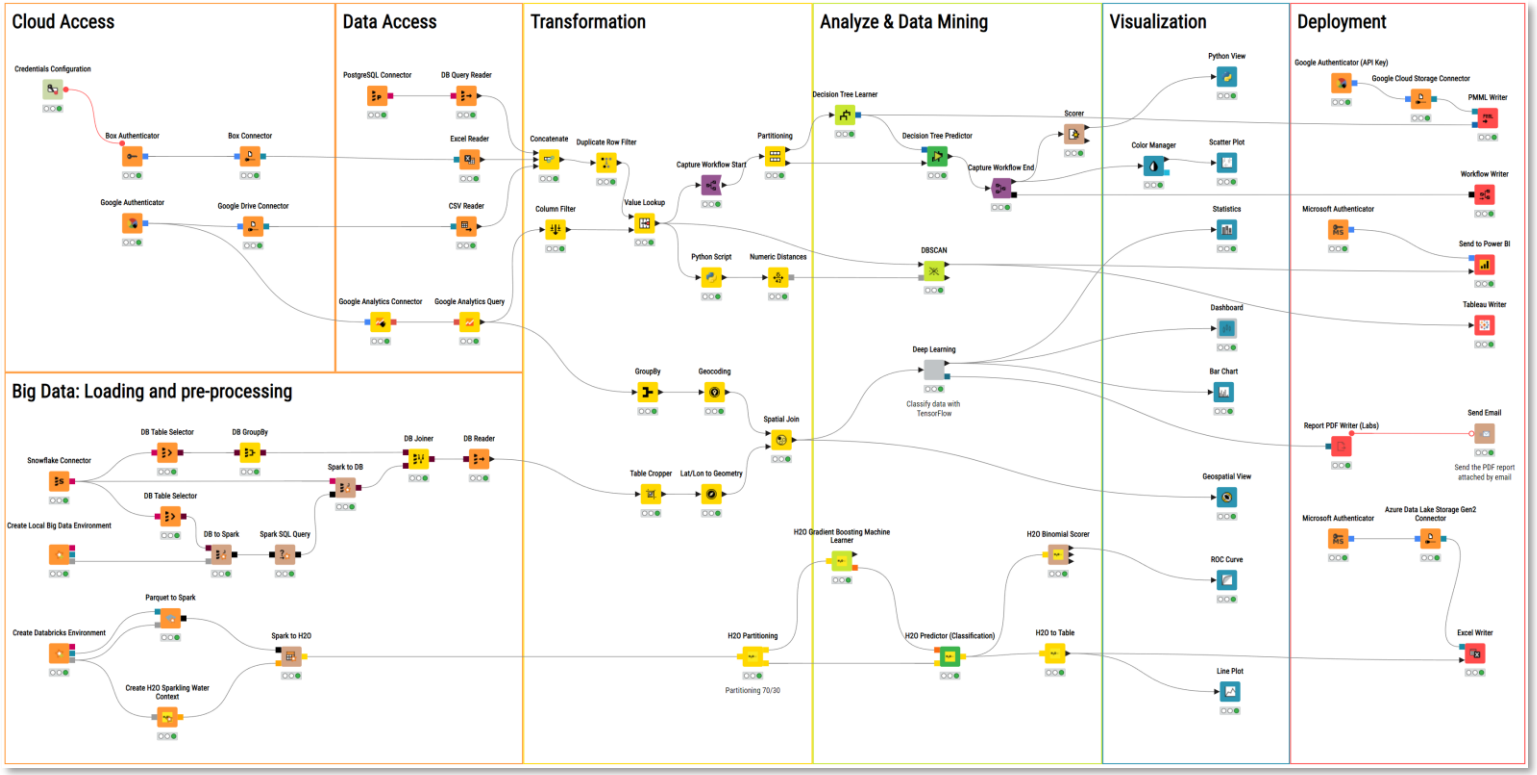
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

The Data Science Life Cycle



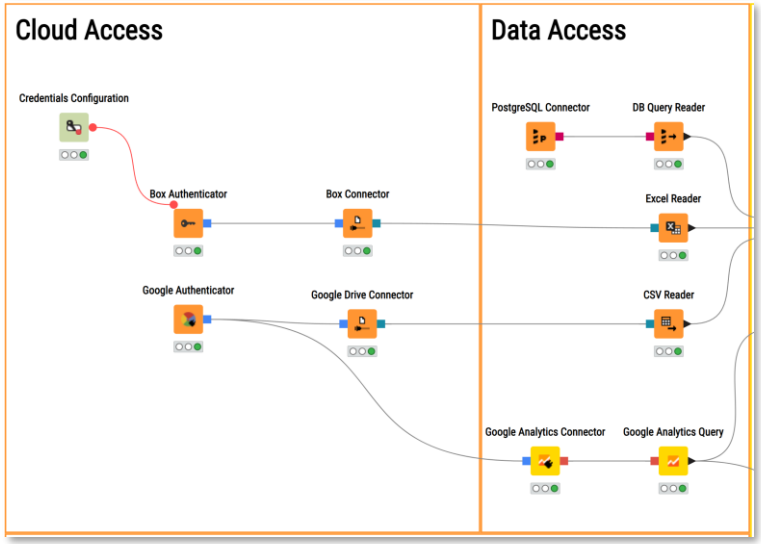
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Over 4000 Nodes included



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

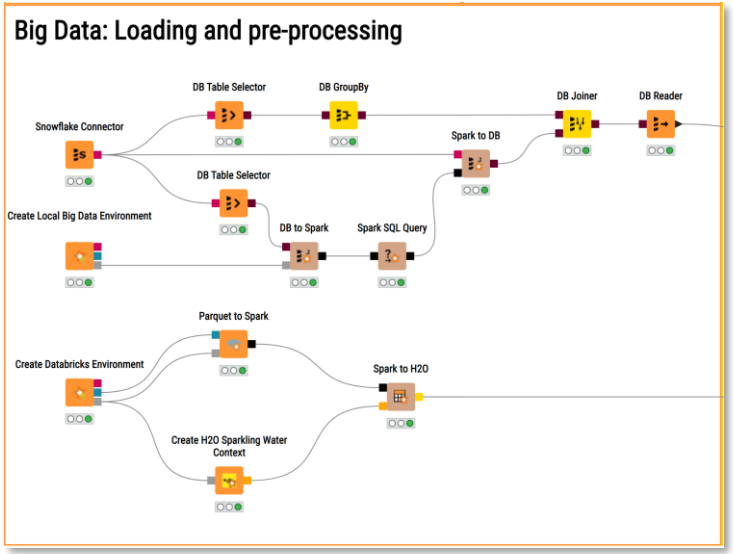
Data Access



- Files
  - CSV, txt, Excel, Word, PDF
  - SAS, SPSS
  - XML, JSON, PMML
  - Images, texts, networks
- Databases
  - MySQL, PostgreSQL, Oracle
  - Theobald
  - any JDBC (DB2, MS SQL Server)
  - Amazon DynamoDB
- Other
  - Twitter, Google
  - Amazon S3, Azure Blob Store
  - Sharepoint, Salesforce
  - Kafka
  - REST, Web services

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

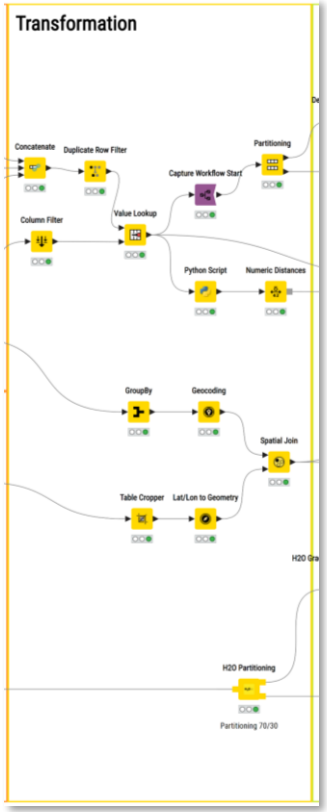
Big Data



- Spark & Databricks
- HDFS support
- Hive
- Impala
- In-database processing

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

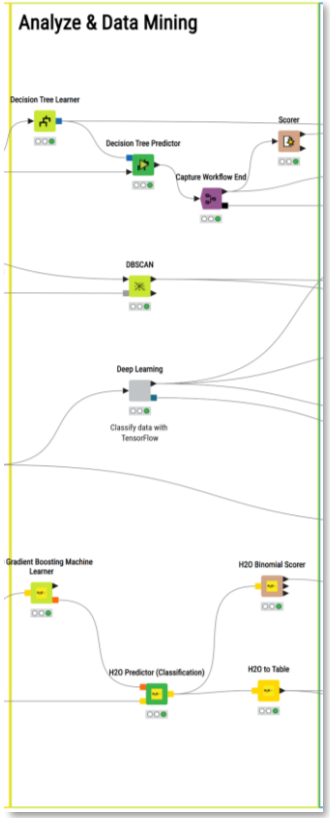
Transformation



- Data Cleaning
  - Row, column, table based
- Data Merging
  - Join, concatenate, append
- Aggregation
  - Grouping, pivoting, binning
- Feature creation and selection

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

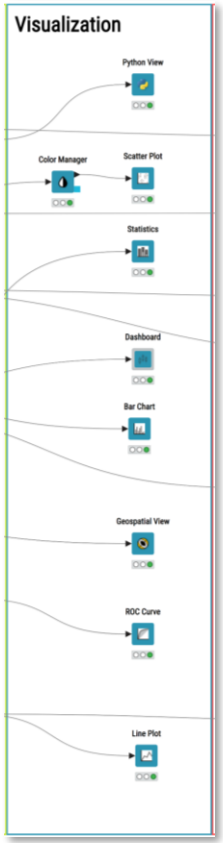
Analyze & Data Mining



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

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

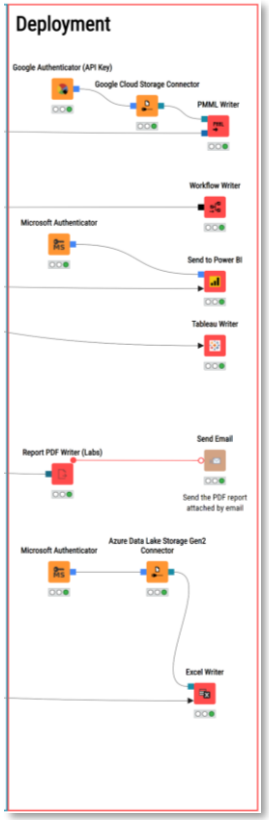
Visualization



- Interactive dashboards
- Many visualization nodes
  - Scatter Plot, Box Plot, Line Plot
  - Networks, ROC Curve, Decision Tree
  - Adding more with each release!
- Misc
  - Tag cloud, geospatial, molecules
- Script-based visualizations
  - R, Python, JavaScript

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Deployment

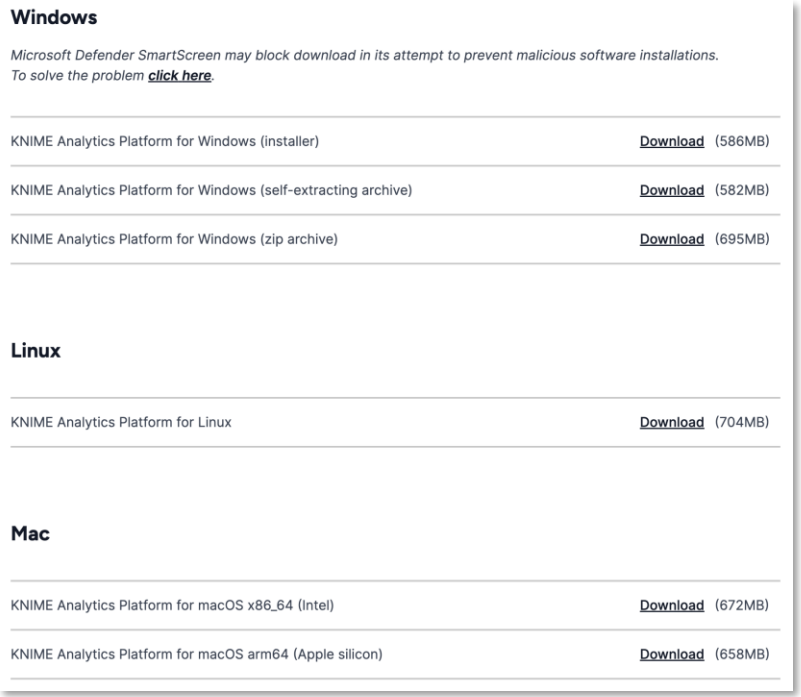


- Database
- Files
  - Excel, CSV, txt
  - XML
  - PMML
  - to: local, KNIME Business Hub, Cloud
- Reporting

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Installation

- Go to knime.com/downloads
- Select the KNIME version for your computer:
  - Mac – Intel or Apple Silicon
  - Windows – 32 or 64 bit
  - Linux
- Download the archive and extract the file, or download the installer package and run it



Windows

*Microsoft Defender SmartScreen may block download in its attempt to prevent malicious software installations. To solve the problem **click here**.*

KNIME Analytics Platform for Windows (installer)	Download	(586MB)
KNIME Analytics Platform for Windows (self-extracting archive)	Download	(582MB)
KNIME Analytics Platform for Windows (zip archive)	Download	(695MB)

Linux

KNIME Analytics Platform for Linux	Download	(704MB)

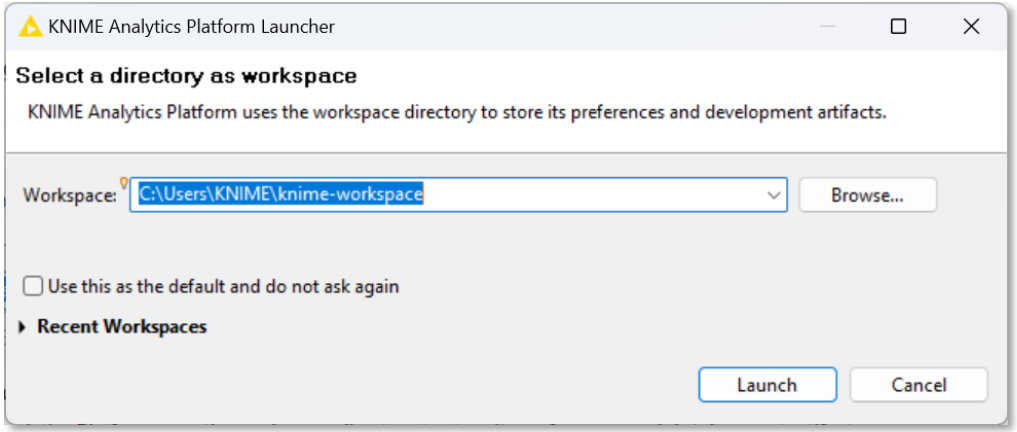
Mac

KNIME Analytics Platform for macOS x86_64 (Intel)	Download	(672MB)
KNIME Analytics Platform for macOS arm64 (Apple silicon)	Download	(658MB)

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

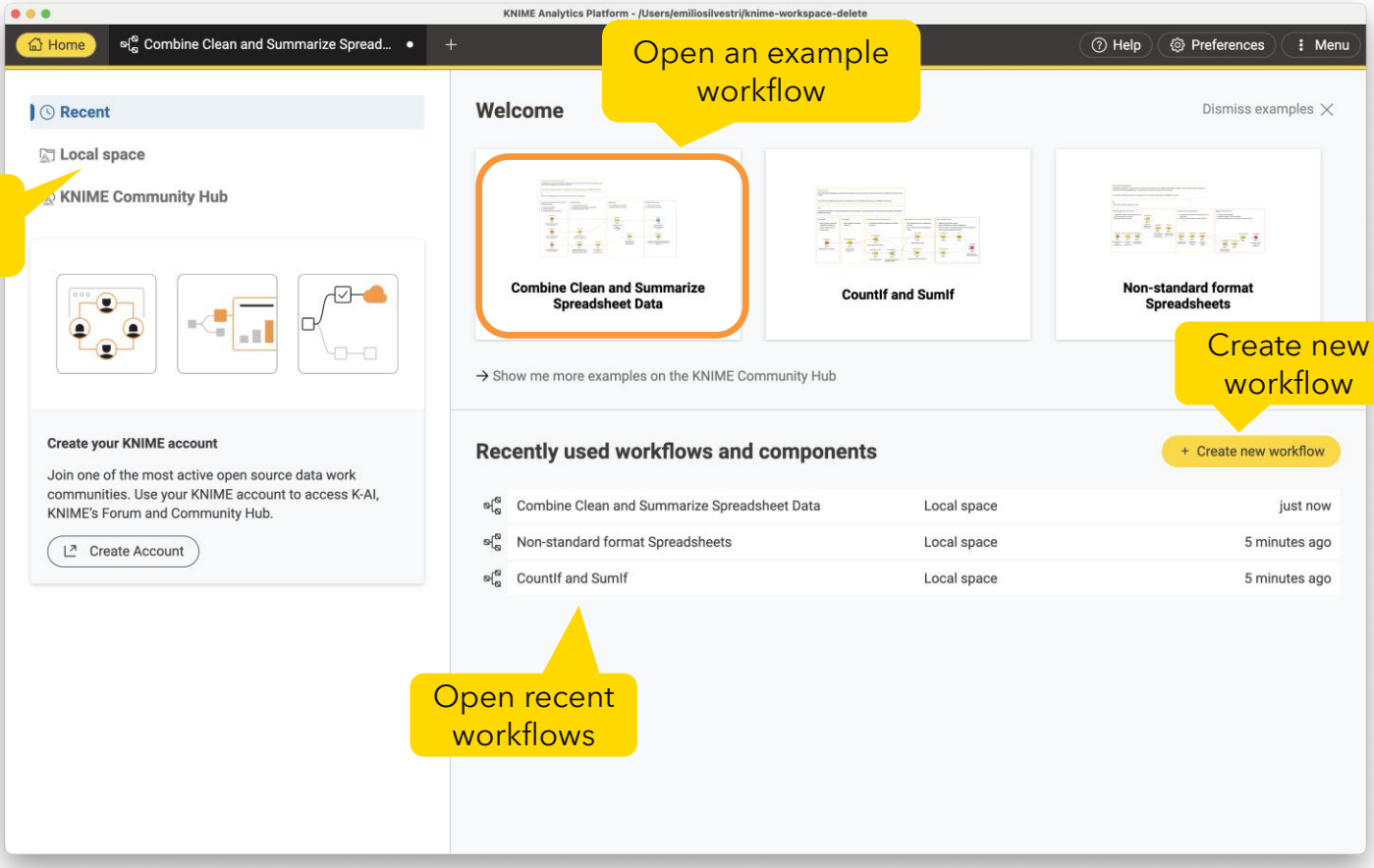
The KNIME Workspace

- Folder on the local computer in which workflows and data files are stored for the current KNIME session
  - Workspaces are portable, just like KNIME



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Tour of the User Interface



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Tour of the User Interface

Description tab

Application tabs

Workflow editor

Node monitor

Home | Combine Clean and Summarize Spread... | Help | Preferences | Menu

Execute | Cancel | Reset | Create metanode | Create component | 55%

Excel Reader

This node reads Excel files (xlsx, xlsm, xlsb, and xls format). It can read a single or multiple files at the same time, however reading only one sheet per file. The supported Excel types that can be read in are string, number, boolean, date, and time but not pictures, diagrams, etc.
The data is read in and converted to the KNIME types string, integer, long, double, boolean, local date, local time, and local date&time. Upon execution, the node will scan the input file to determine number and types of the columns and output a table with the auto-guessed structure and KNIME types.
Formulas can also be read in and reevaluated if desired.The performance of this node is limited (due to the underlying library of the Apache POI project). Reading large files takes a long time and uses a lot of memory (especially files in xlsx format when formula reevaluation is enabled).

The dialog of this node shows a *preview* and the *file content*. While the preview shows the table that will be read in having the settings applied and is updated once a setting is changed, the file content shows the content of the file as it is displayed in Excel. This allows finding row numbers and column names easier to specify different settings.

This node can access a variety of different file systems. More information about file handling in KNIME can be found in the official File Handling Guide.

Ports | Options | Views

1: File Table | Flow Variables

Rows: 9 | Columns: 2 | Table | Statistics

#	RowID	Item String	Amount Number (integer)
1	Row0	Table	1
2	Row1	Chair	2
3	Row2	Cupboard	1
4	Row3	Side table	1
5	Row4	Fridge	1

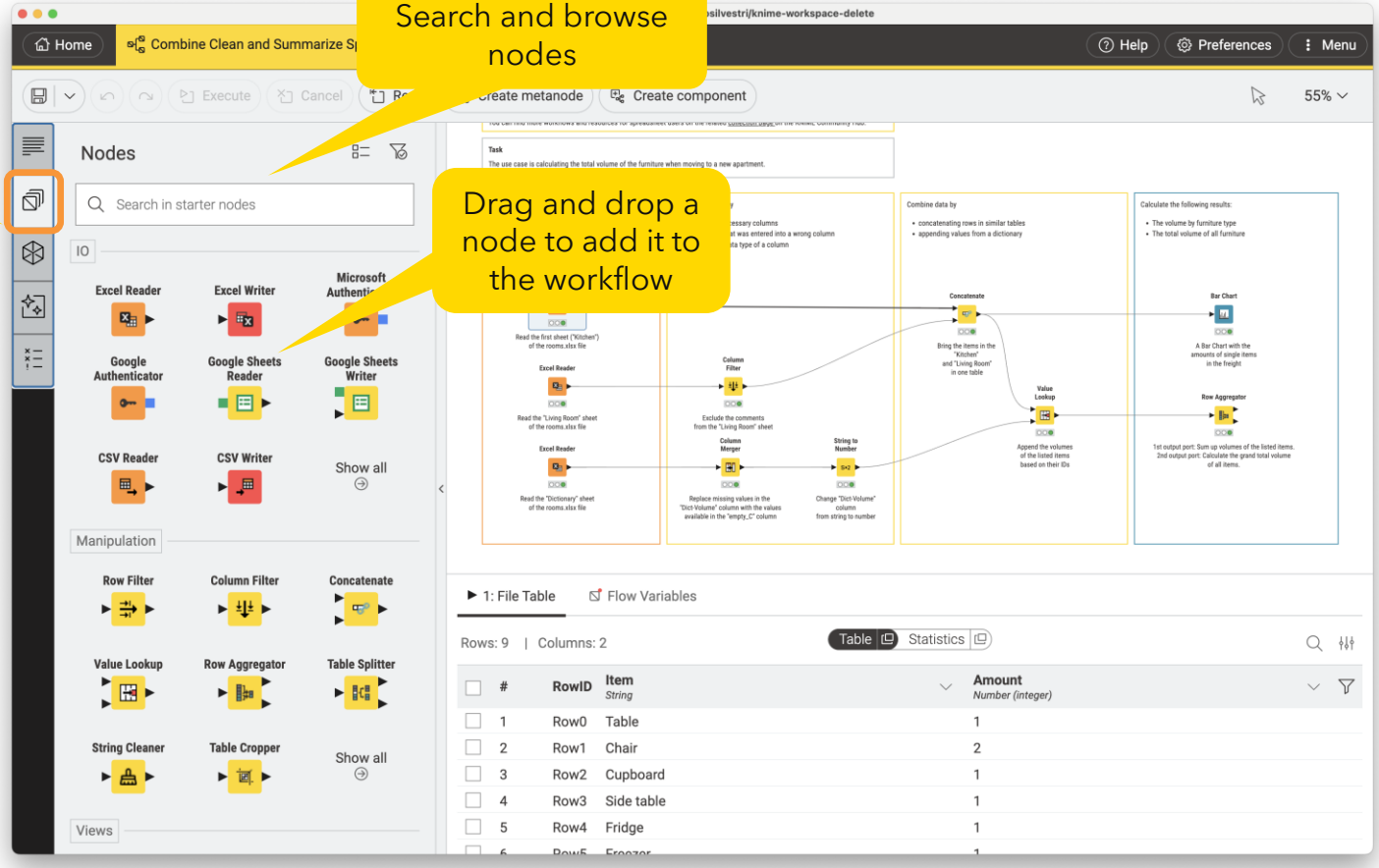
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Tour of the User Interface

Node Repository tab

Search and browse nodes

Drag and drop a node to add it to the workflow



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Tour of the User Interface

Space Explorer tab

Browse workflows in your local space

The space explorer opens in the folder of the current workflow

KNIME Analytics Platform - /Users/emiliosilvestri/knime-workspace-delete

Home | Combine Clean and Summarize Sprea | Help | Preferences | Menu

Execute | Cancel | 55%

Local space

Example Workflows > Basic Examples

- Building a Simple Classifier
- Combine Clean and Summarize Spreadsheet ...
- CountIf and SumIf
- Data Blending
- Non-standard format Spreadsheets
- Visual Analysis of Sales Data

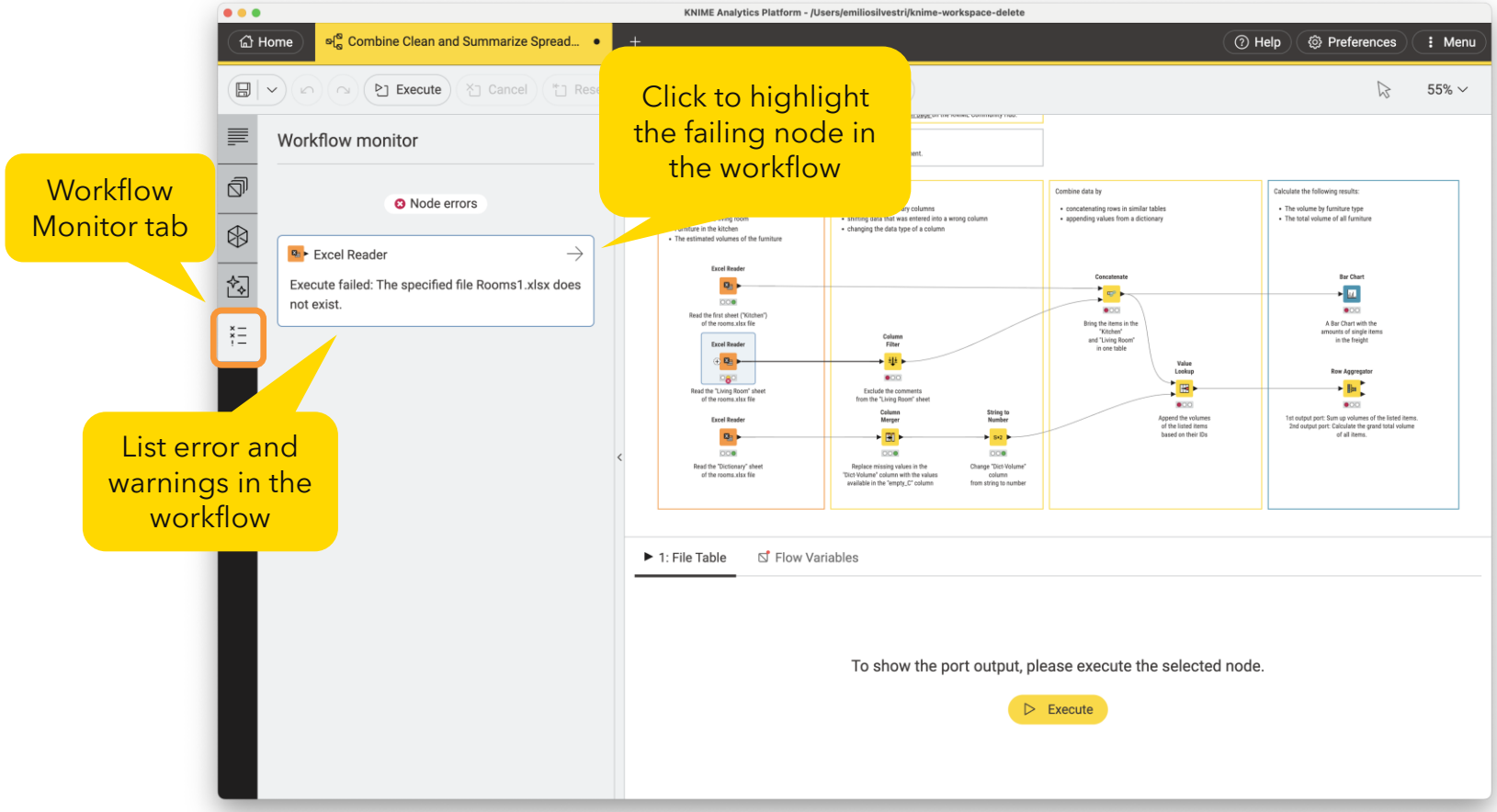
1: File Table | Flow Variables

Rows: 9 | Columns: 2 | Table | Statistics

#	RowID	Item String	Amount Number (integer)
1	Row0	Table	1
2	Row1	Chair	2
3	Row2	Cupboard	1
4	Row3	Side table	1
5	Row4	Fridge	1

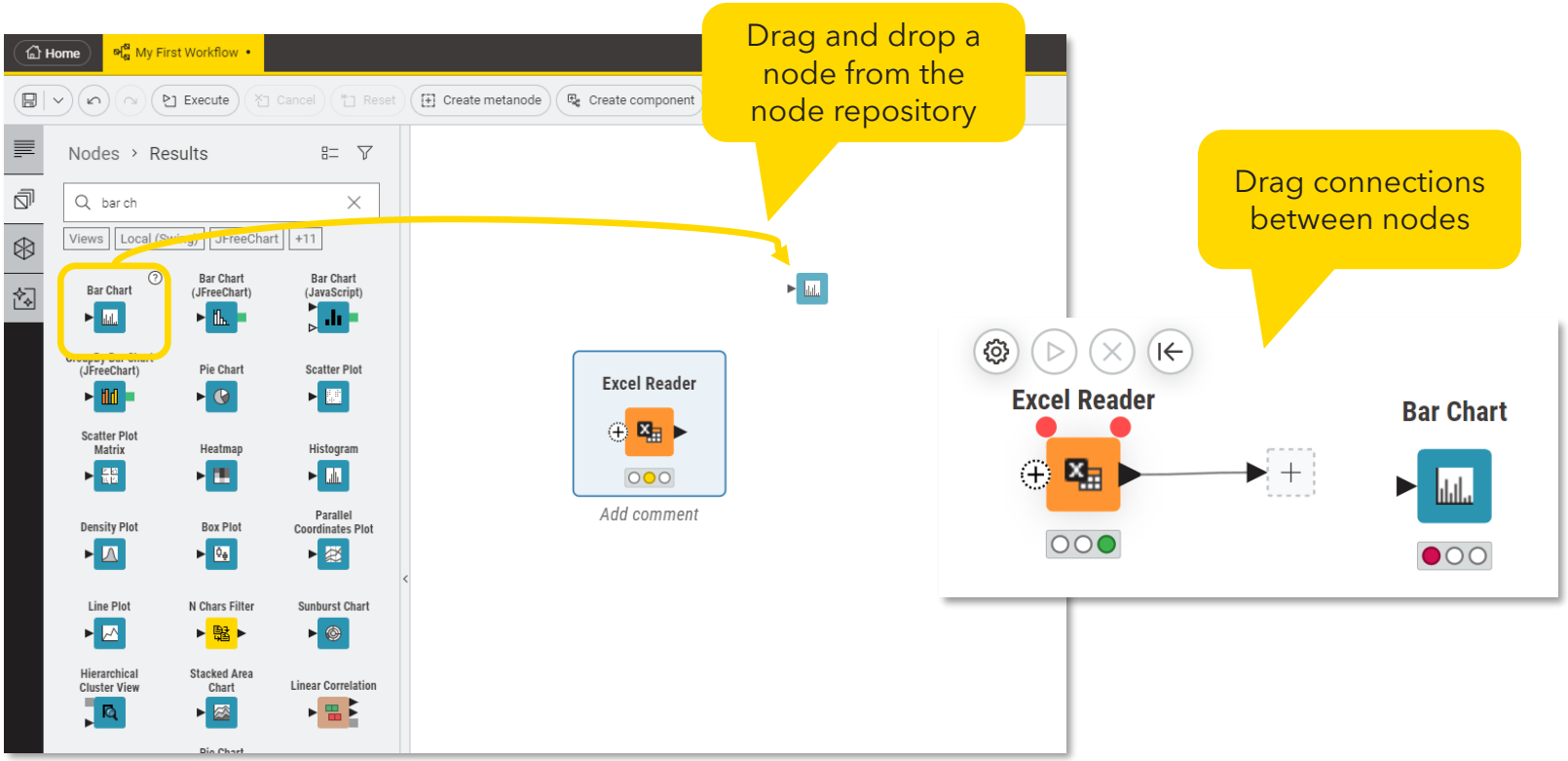
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Tour of the User Interface



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

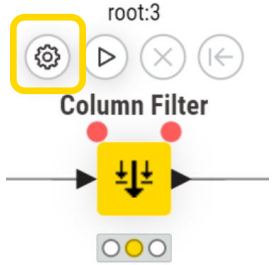
Add and Connect Nodes

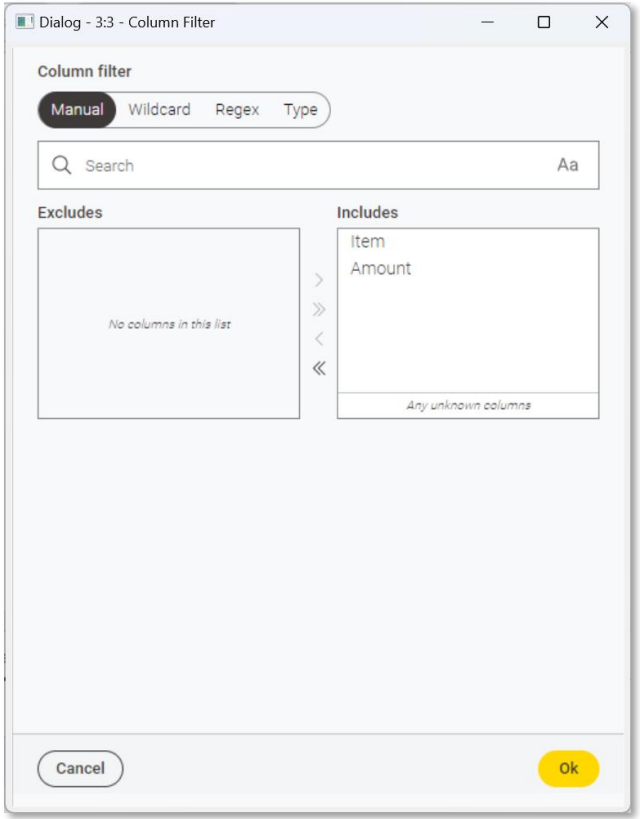


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Configure a Node

- Most nodes require configuration
- Options to open the configuration dialog of a node:
  - Double click the node
  - Right click -> Configure
  - Select the node and press F6
  - Click the Configure button the node action bar

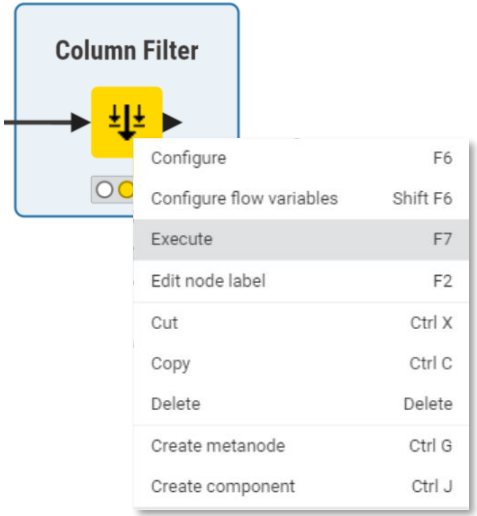


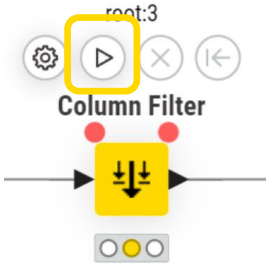


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Execute a Node

- An executed node has carried out its task
- Options to execute a node:
  - Right click -> Execute
  - Select the node and press F7
  - Click the Execute button the node action bar





These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

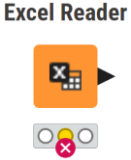
Node Status

- A node can have 4 statuses:

Not Configured:
The node is waiting for configuration or incoming data.

Configured:
The node has been configured correctly and can be executed.

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

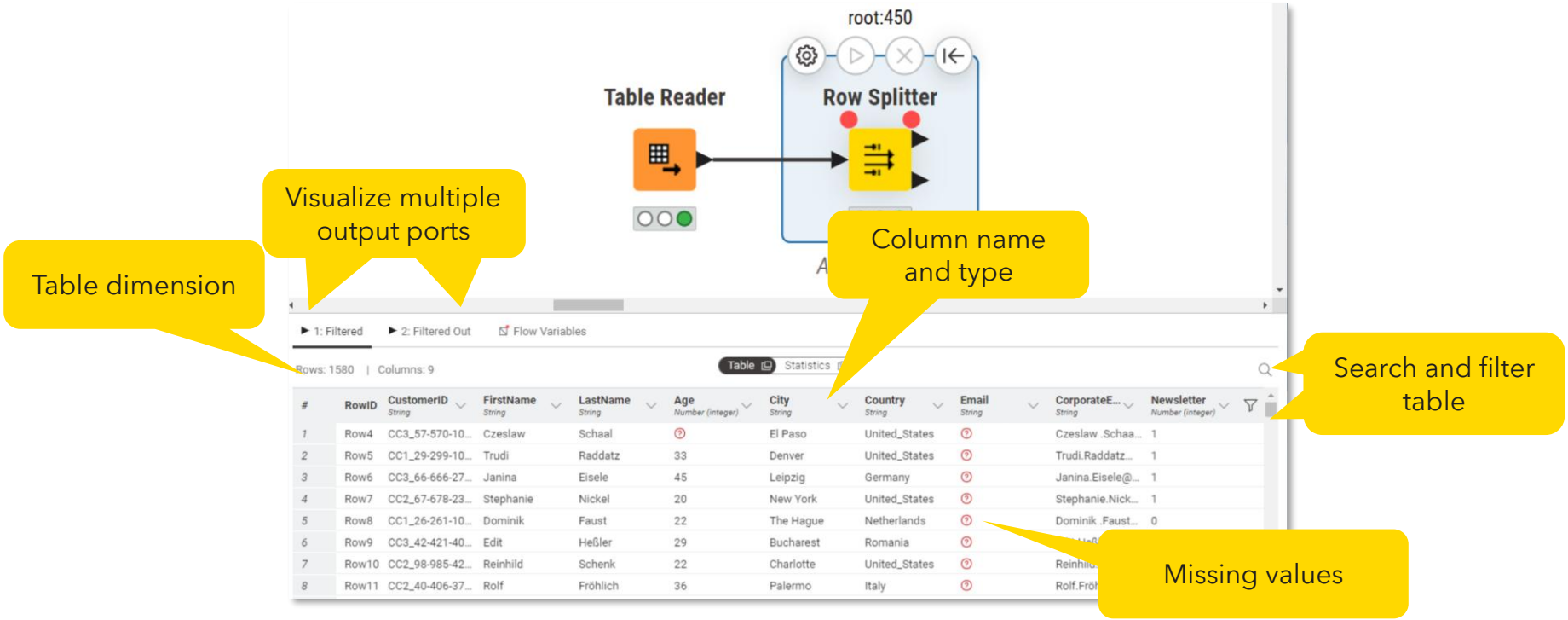


Error:
The node has encountered an error during execution.

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Node Monitor

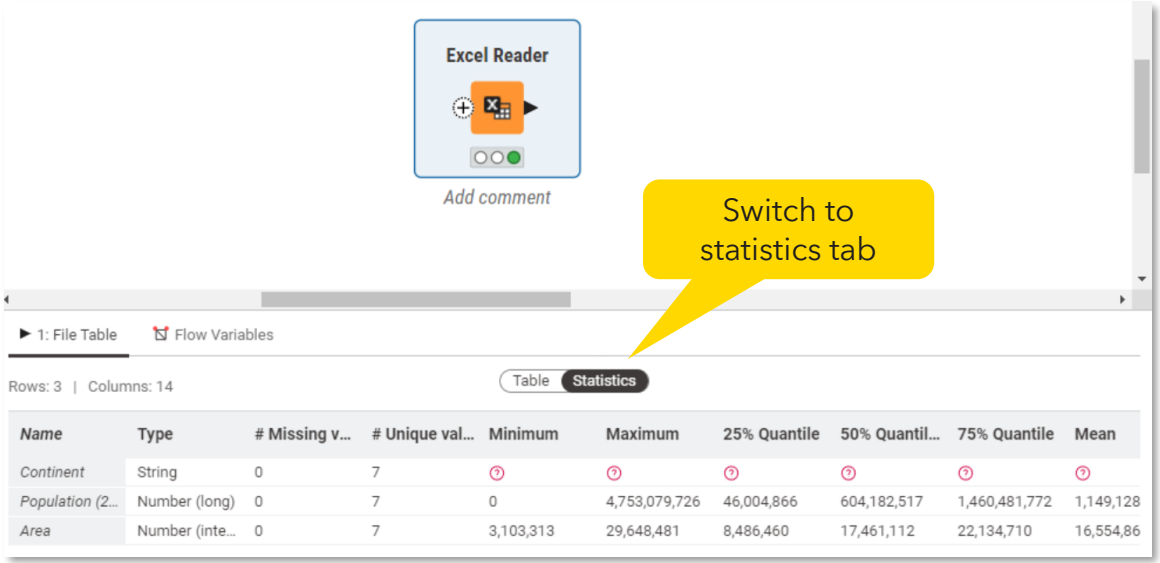
- Select an executed node to inspect its output(s)



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Statistics Tab

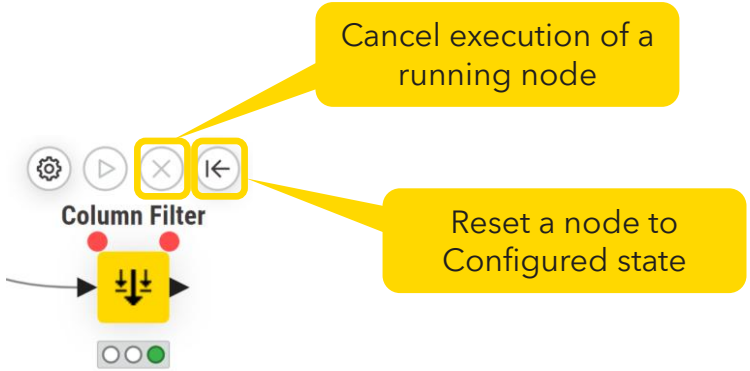
- Get statistics of the selected node output port
  - Missing values
  - Unique values
  - Min and max
  - Quantiles
  - …



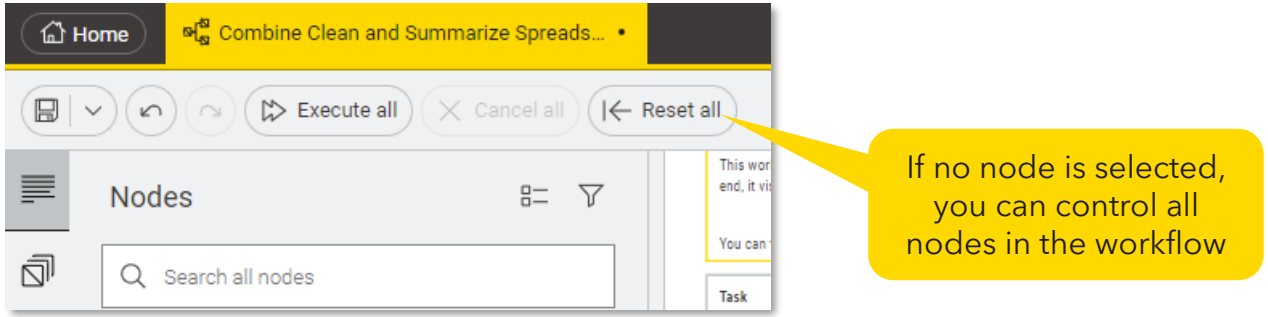
Name	Type	# Missing v...	# Unique val...	Minimum	Maximum	25% Quantile	50% Quantil...	75% Quantile	Mean
Continent	String	0	7	?	?	?	?	?	?
Population (2...	Number (long)	0	7	0	4,753,079,726	46,004,866	604,182,517	1,460,481,772	1,149,128
Area	Number (inte...	0	7	3,103,313	29,648,481	8,486,460	17,461,112	22,134,710	16,554,86

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Cancel and Reset a Node



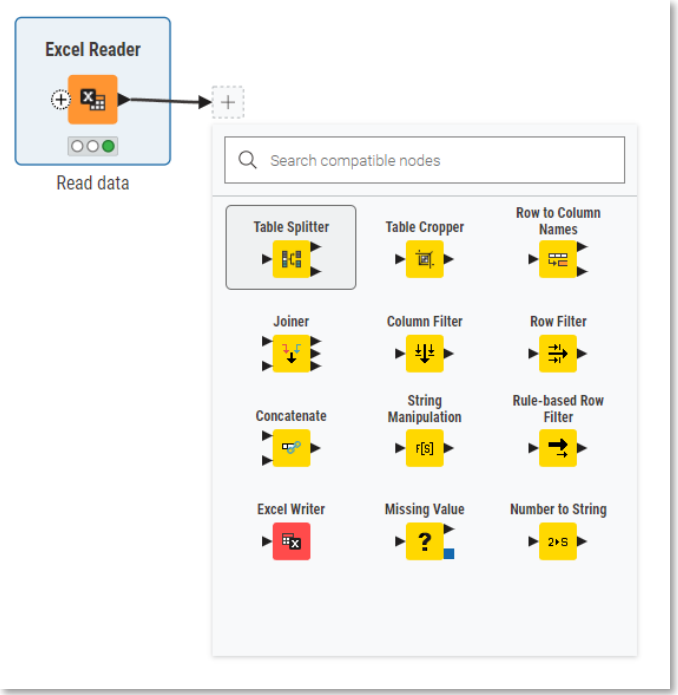
- Node control is also possible from the Workflow Toolbar



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Quick Node Adding Panel

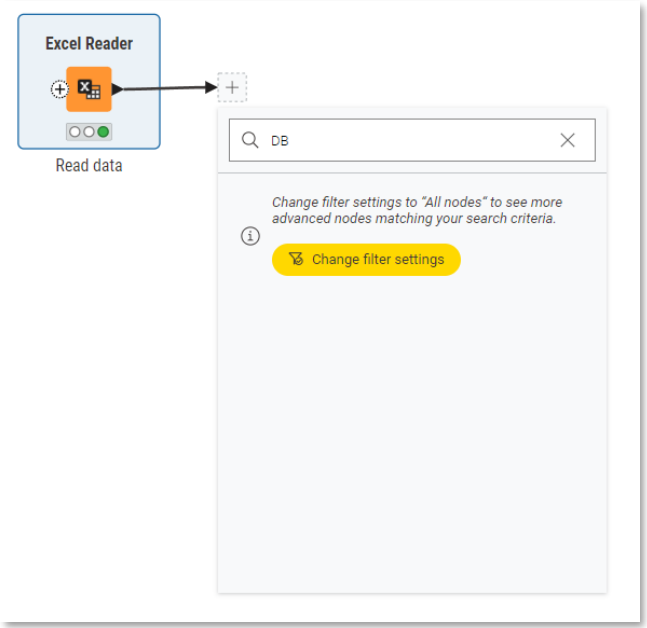
- Drag and drop the output of a node to an empty space
- See recommended nodes to append
- Search for all compatible nodes



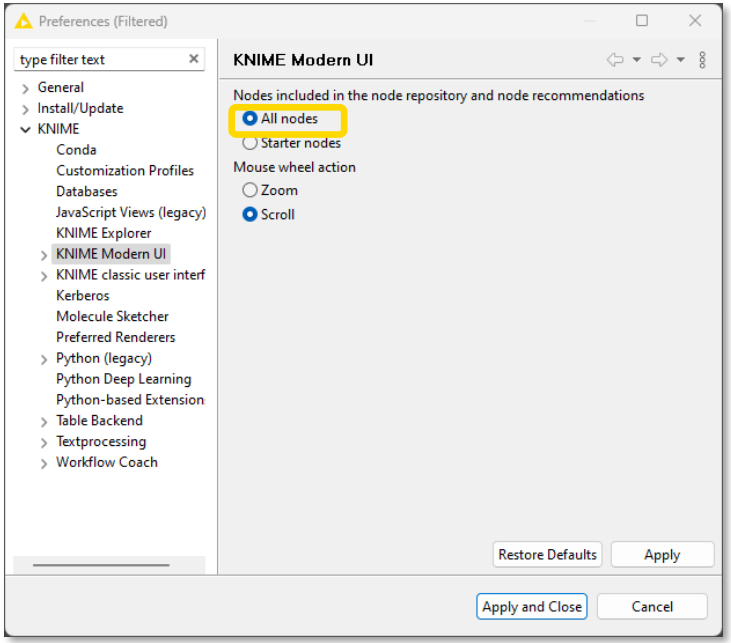
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Starter Perspective

- Click “Change filter settings” and select “All nodes” to see the full list



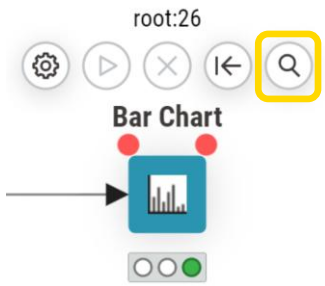
- By default, only a curated set of essential nodes is displayed

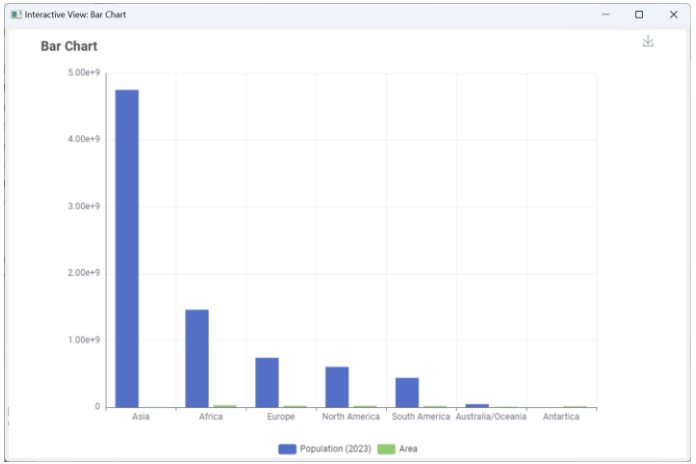


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Node Views

- Open the node view in a new window
- Only available for “blue” nodes





These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME File Extensions

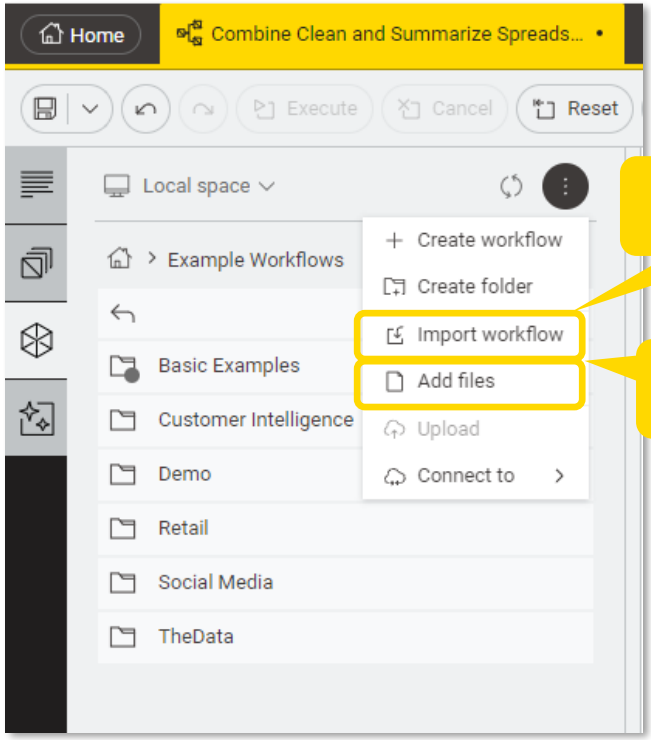
- Dedicated file extensions for workflows and workflow groups associated with KNIME Analytics Platform

- ***.knwf** for KNIME Workflow Files

- ***.knar** for KNIME Archive Files (folder)

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

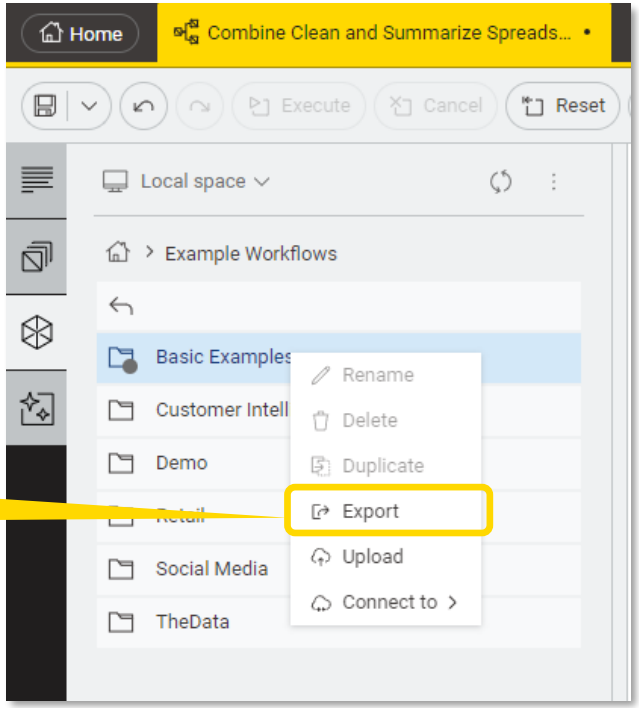
Importing and Exporting Workflows



Import .knwf or .knar files

Import data files

Export to file

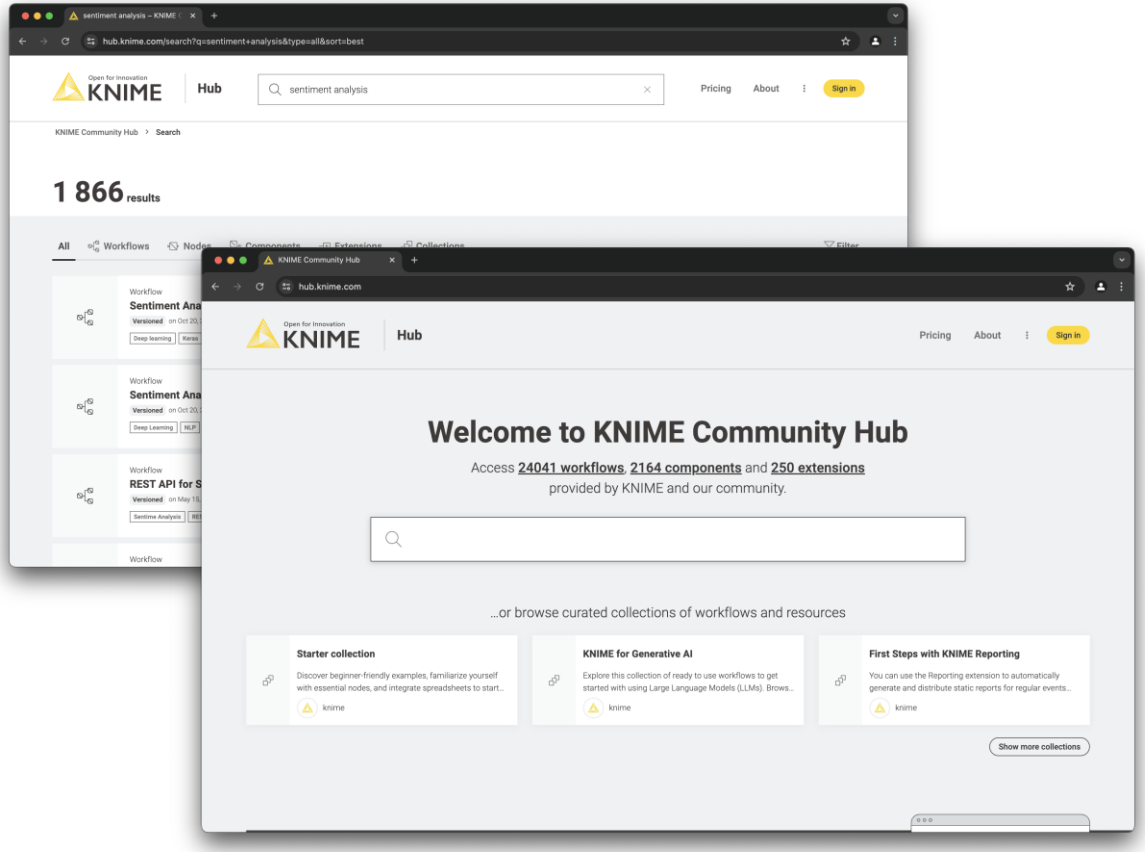


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Community Hub

- Place to search and share
  - Workflows
  - Nodes
  - Components
  - Extensions

- Collaborate in teams
- Execute and deploy workflows

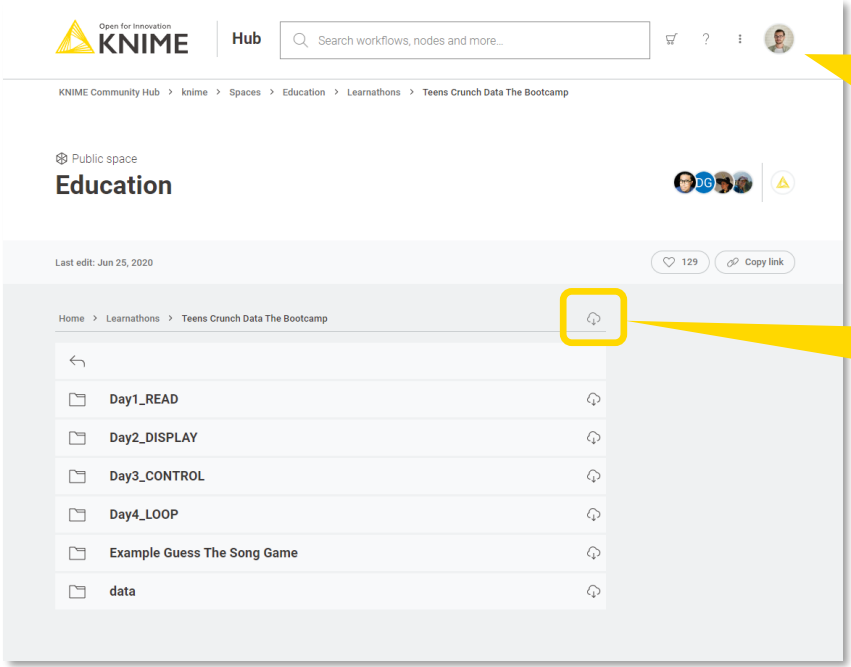


hub.knime.com

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Downloading from KNIME Community Hub

- Sign-in is required to download a folder
- Single workflows can be downloaded without login



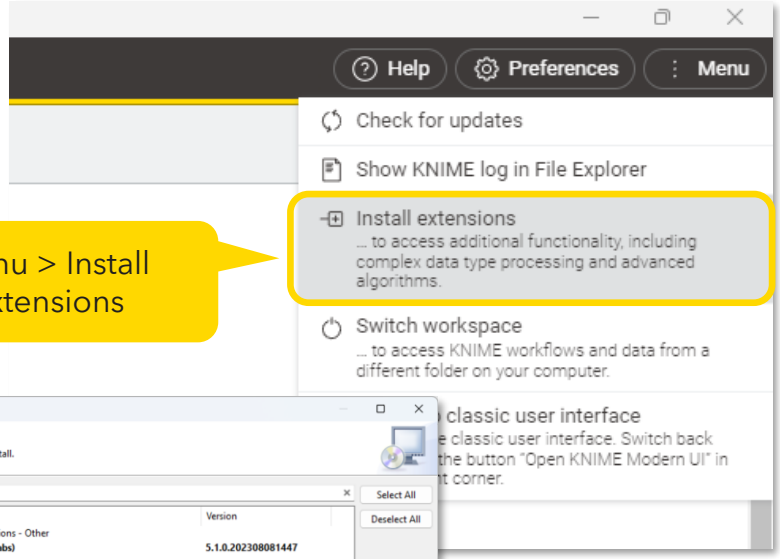
You must be signed in

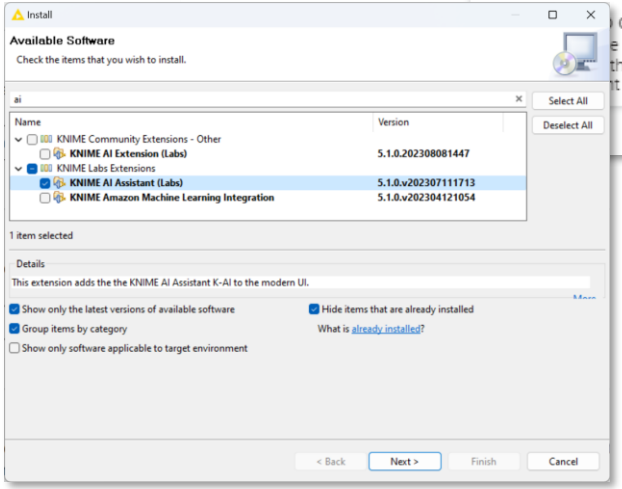
Download full .knar

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Extensions

- Provide additional functionalities
- Add more nodes to the Node Repository
  - Text Processing
  - Image Processing
  - Deep Learning
  - Geospatial Analytics
  - Big Data Connectors
  - and many more

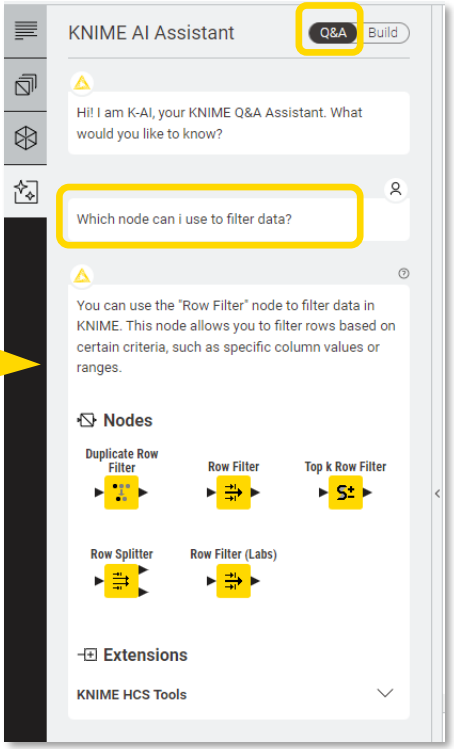




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

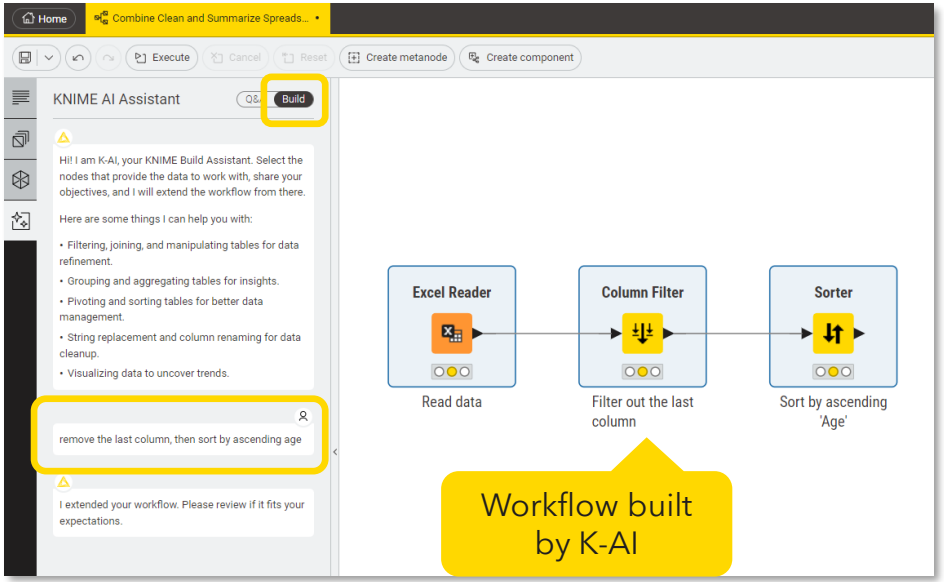
KNIME AI Assistant

- AI-powered chat bot



Get ready-to-be-added recommended nodes

- AI-powered workflow builder

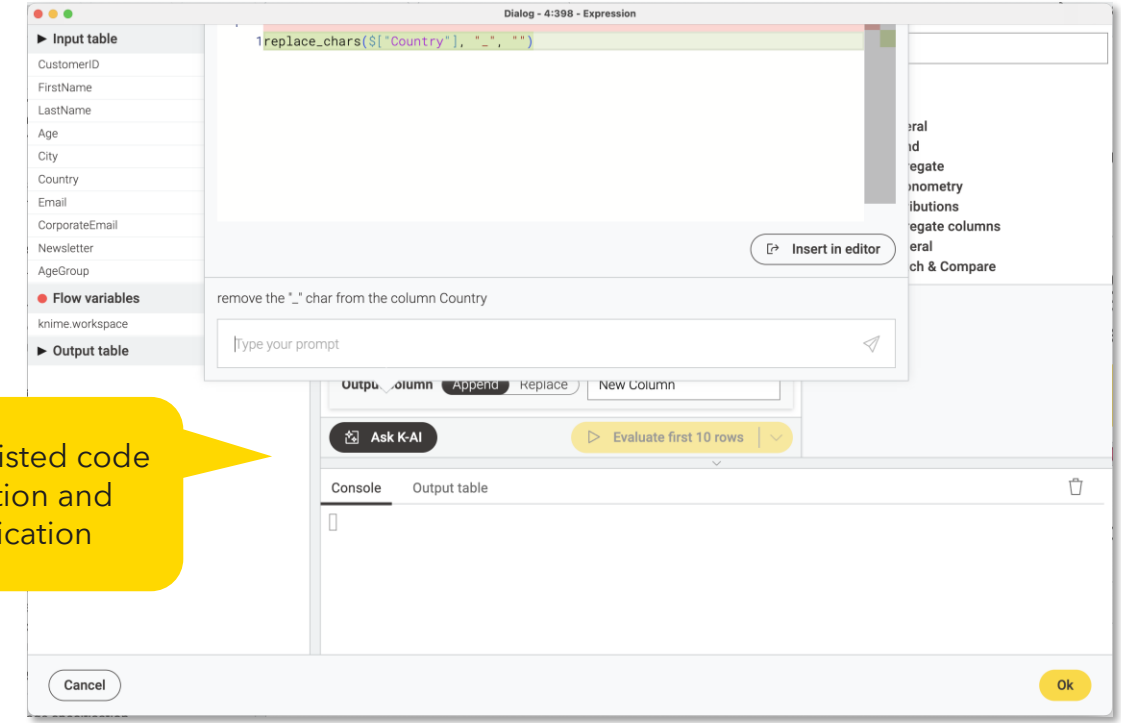


Workflow built by K-AI

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME AI Assistant

- Copilot in scripting nodes



Get AI assisted code generation and modification

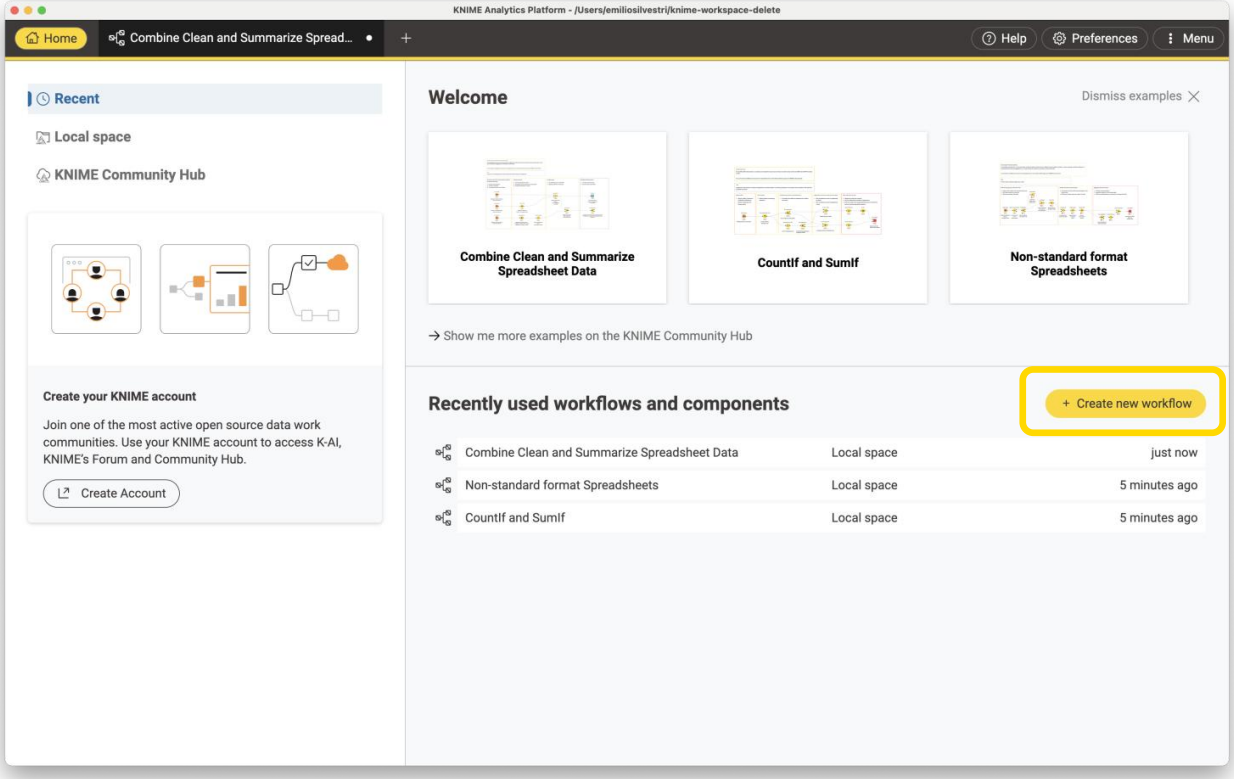
Install the AI Assistant extension

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

Build Your First Workflow

- From the Home page, create a new workflow and give it a name

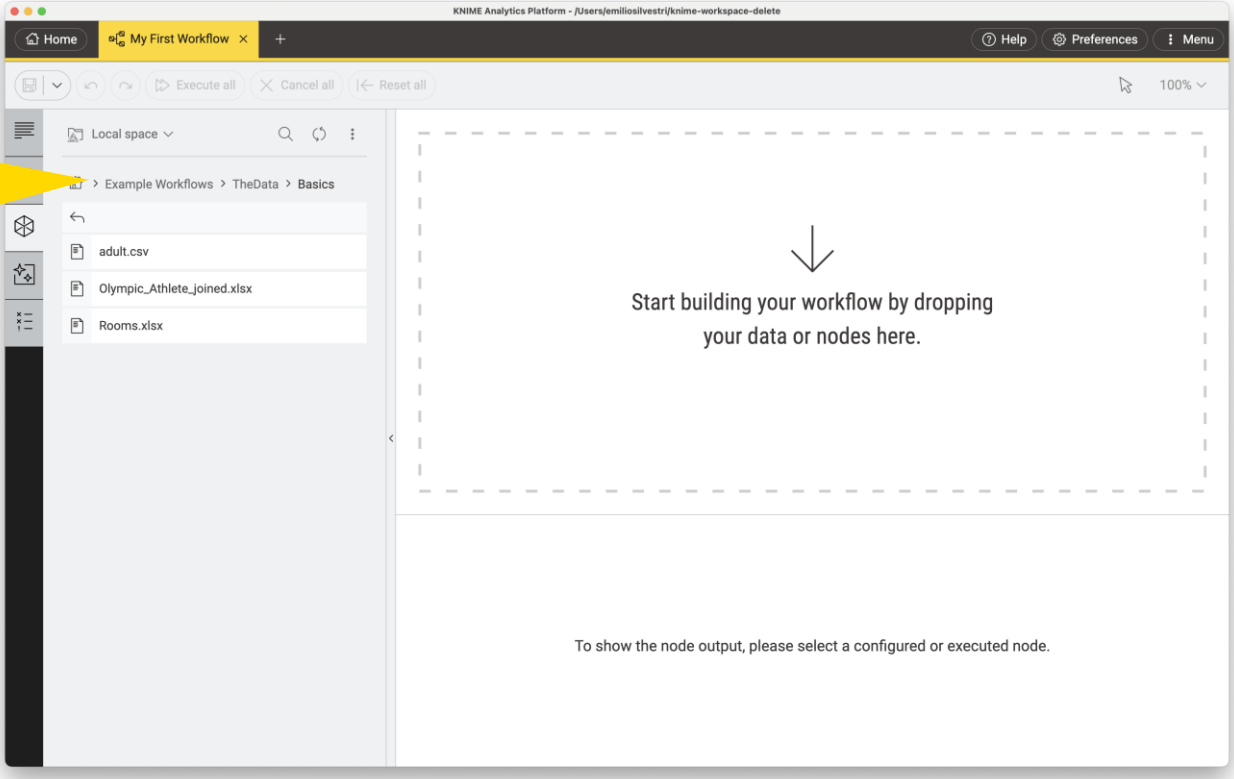


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

- From the side panel navigation, open the Space Explorer tab

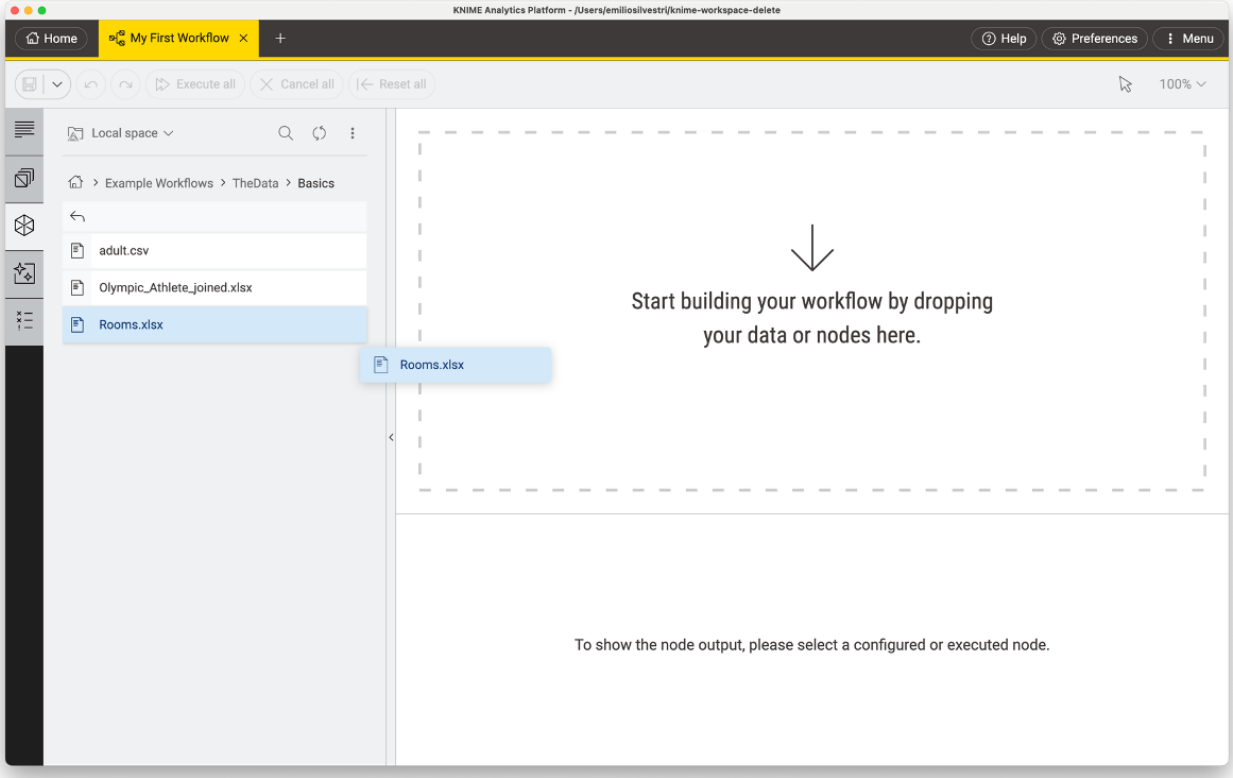
Navigate to Example Workflows > TheData > Basics



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

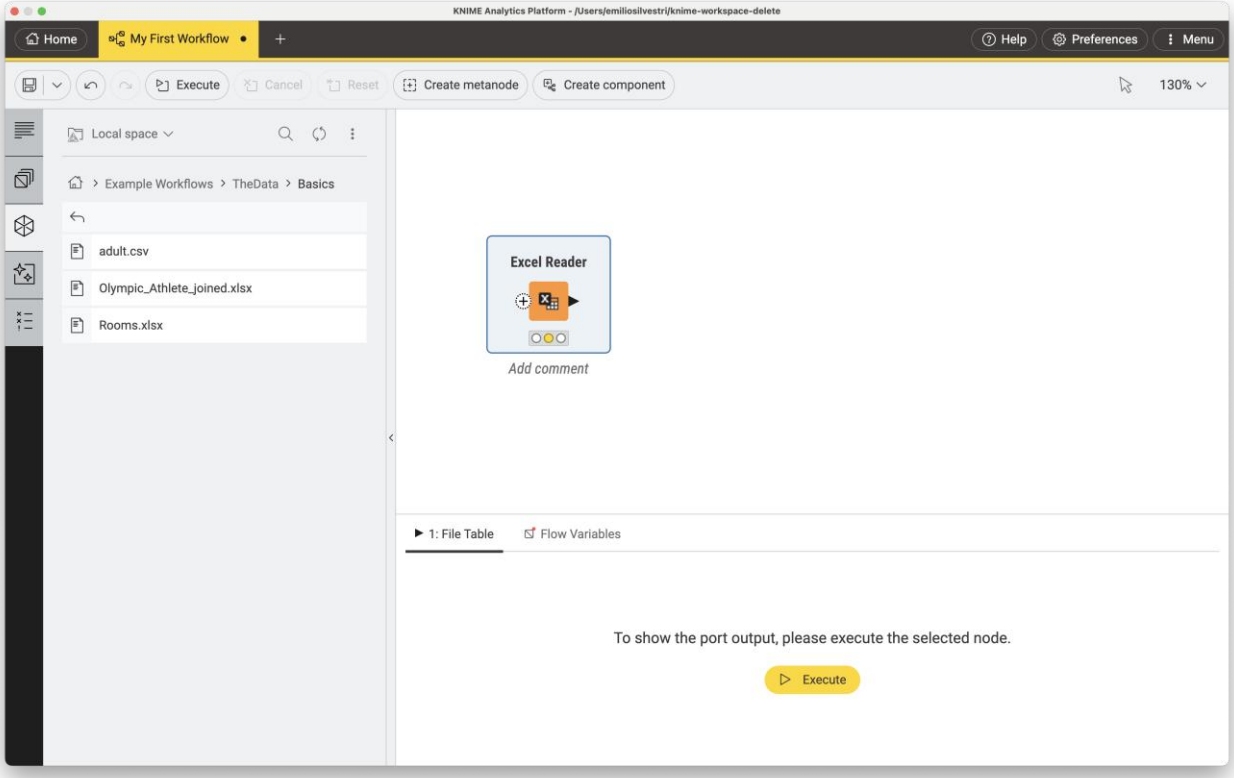
- Drag and drop the file Rooms.xlsx into the empty Workflow Editor



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

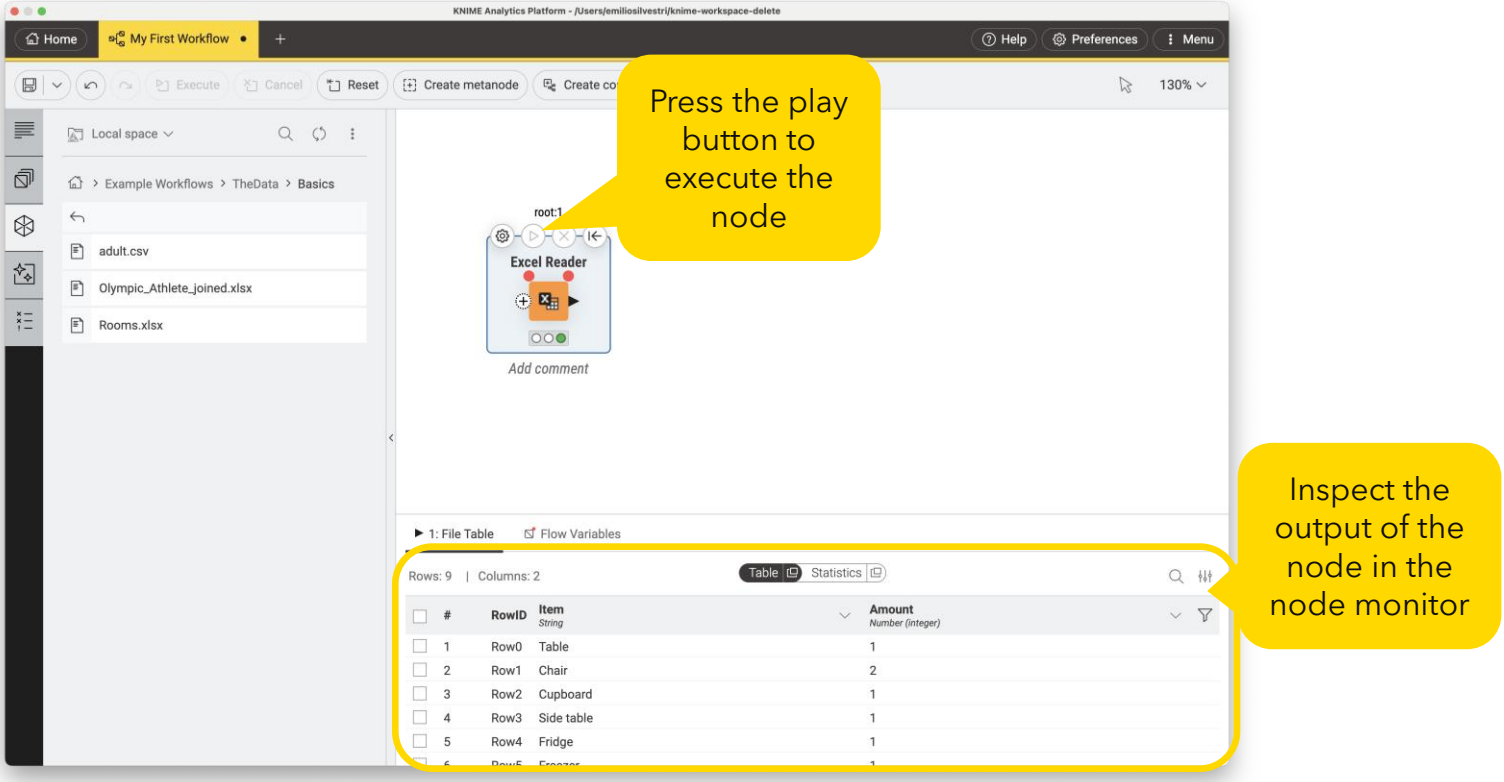
- An Excel Reader node appears in the Workflow Editor



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

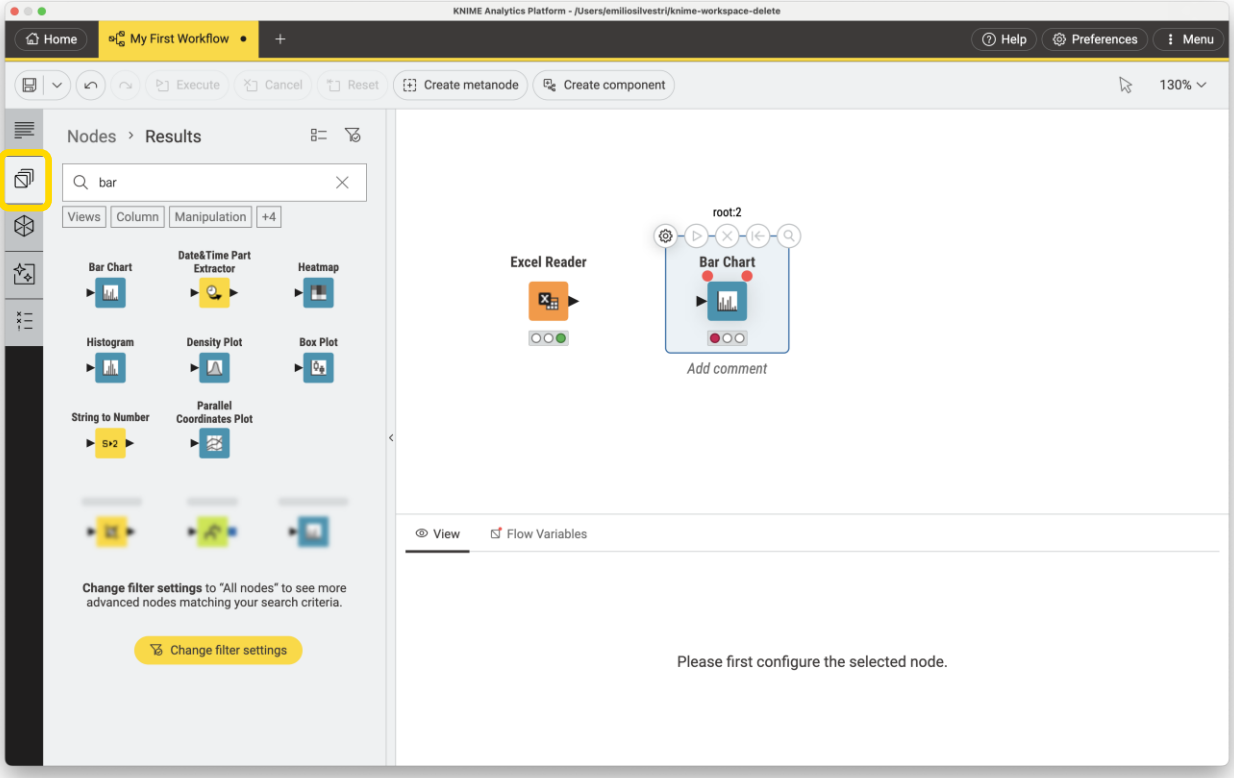
- Select the node and execute it



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

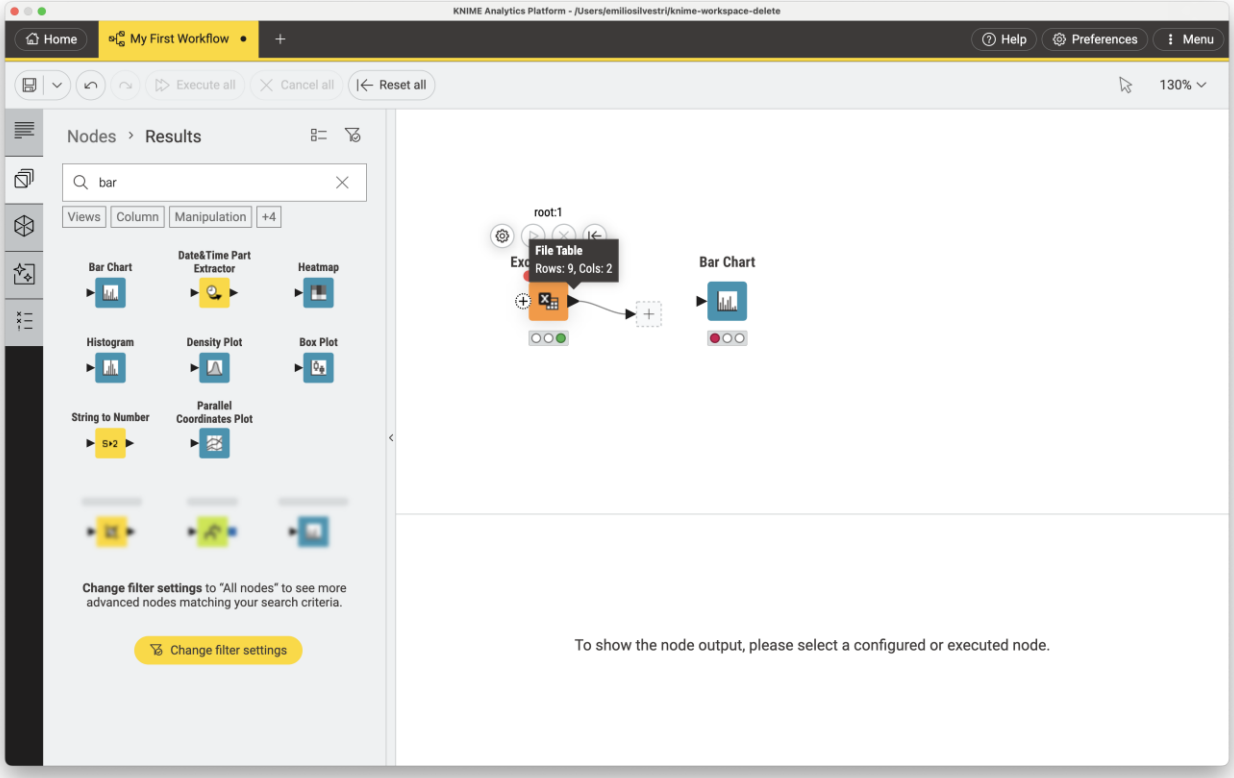
- From the Node Repository tab, drag and drop the Bar Chart node



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

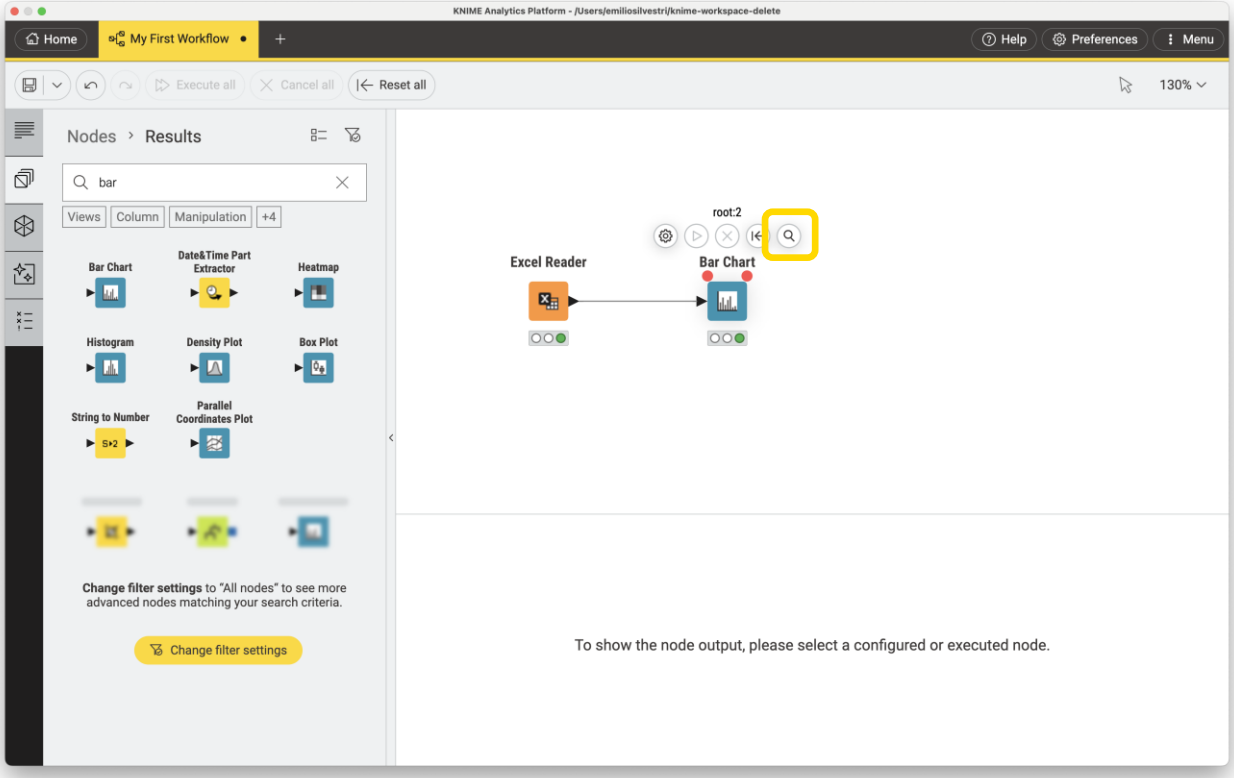
- Drag the connection between the two nodes



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

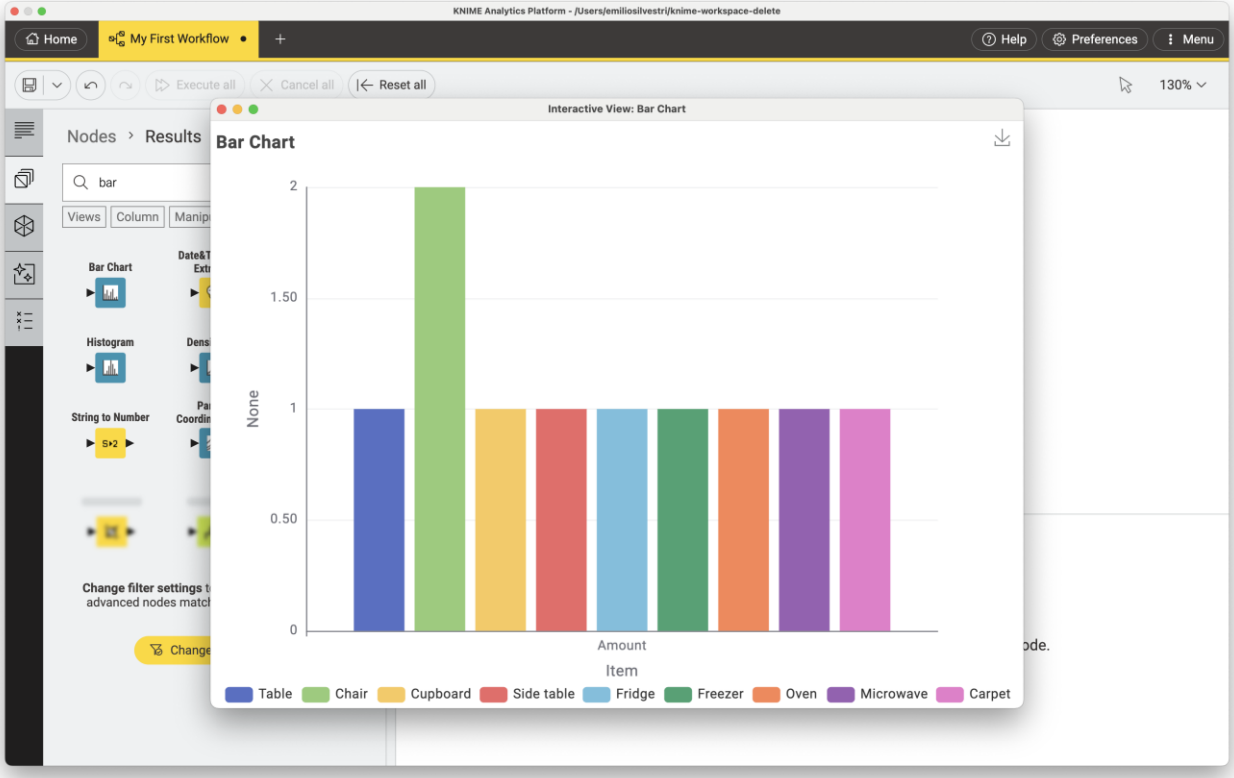
- Hover over the Bar Chart and click the magnifier icon



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Build Your First Workflow

- Have a look at your first data visualization with KNIME Analytics Platform!



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

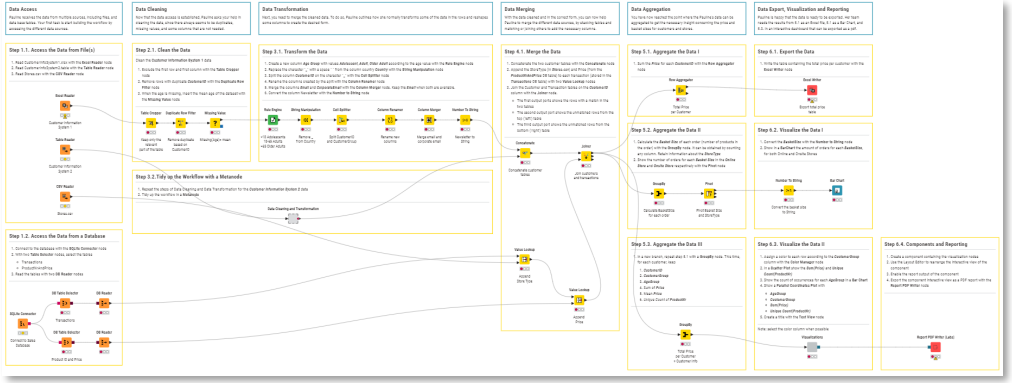
Data Access

Use Case: Customer Transactions Analysis

Use Case

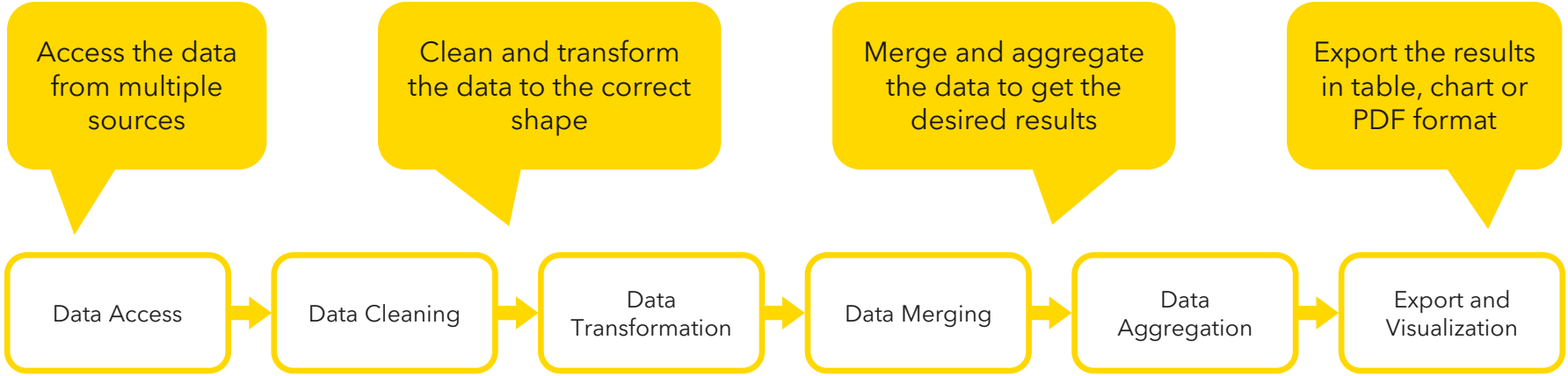
Pauline from Sales would like to write a report and visualize the insights on the customer transaction data her team collects every month.

She has reached out to your data team for help to automate the process and thereby save time.



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Use Case: Customer Transactions Analysis

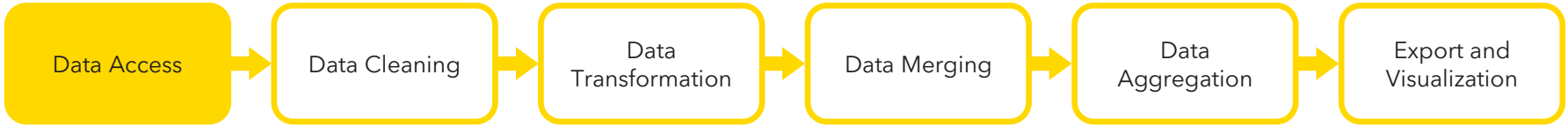


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Data Access

- The beginning of every data process
- Data can be stored in many ways
  - Locally
  - In different data format (.csv, .xls…)
  - On the cloud
  - In a database
  - …
- We need a way to access all of them

Use Case
Pauline receives the data from multiple sources, including files, and data base tables. Your first task is to start building the workflow by accessing the different data sources.

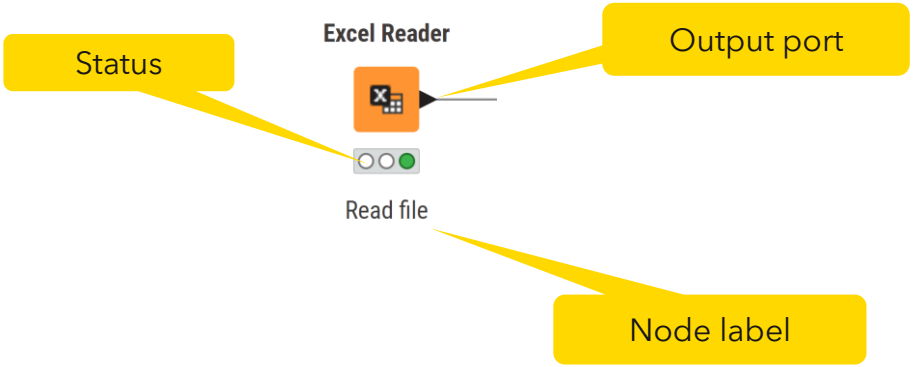


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Data Source Nodes

Typically characterized by:

- Orange color
- By default no input ports, 1-2 output ports
- Many nodes for many data formats
- Support reading from different File Systems



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Excel Reader

- Reads .xls and .xlsx file from Microsoft Excel
- Supports reading from multiple sheets

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Excel Reader

File system

Sheet specific settings

File path

Preview

Dialog - 3:366 - Excel Reader (Customer Information)

File

File and Sheet | Data Area | Advanced | Transformation | Flow Variables | Job Manager Selection | Memory Policy

Input Location

Read from: Local File System

Mode: File / Files in folder

File: C:\Users\Sanket Joshi\1.1 course update 5.1\Data Sources for WF\CustomerInfoSystem1.xlsx

Browse...

Select Sheet

First with data (default_1)

By name: default_1

By position: 0 (Position start with 0.)

Preview | File Content

Preview with current settings

The suggested column types are based on the first 10000 rows only. See 'Advanced Settings' tab.

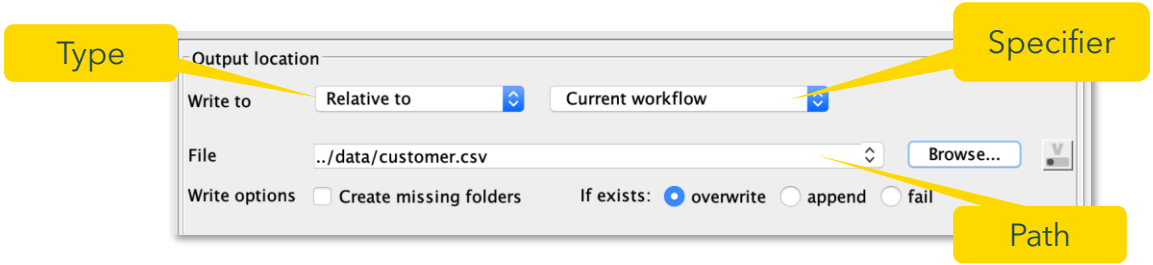
Row ID	S Column	S Custom...	S FirstName	S LastName	I Age	S City	S Country	S Email	S Corpor...	I Newsle...
Row0	Customers	?	?	?	?	?	?	?	?	?
Row1	?	CC2_49-497...	Hans-Wilhelm	Brill	19	Nuernberg	Germany	Hans-Wilhel...	?	0
Row2	?	CC2_50-502...	Deborah	Gilbert	43	Berlin	Germany	Deborah.Gil...	?	0
Row3	?	CC1_27-275...	Ingeburg	Reiter	25	Turin	Italy	Ingeburg.R...	?	0
Row4	?	CC1_65-655...	Ludwina	Sadowski	28	Las Vegas	United_States	Ludwina.Sa...	?	1
Row5	?	CC2_82-823...	Hella	Kittel	17	Sofia	Bulgaria	Hella.Kittel...	?	1
Row6	?	CC2_31-317...	Rüdiger	Ehrlich	25	Minsk	Belarus	Rüdiger .Ehr...	?	1
Row7	?	CC2_50-500...	Annekatrin	Pauly	20	Odessa	Ukraine	Annekatrin....	?	1
Row8	?	CC1_44-449...	Olaf	Gebert	16	Mariupol	Ukraine	Olaf.Gebert...	?	1
Row9	?	CC3_20-206...	Maya	Ruf	54	Palermo	Italy	Maya.Ruf@...	?	1
Row10	?	CC1_43-437...	Samuel	Kracht	48	Nottingham	United_King...	Samuel .Kra...	?	1
Row11	?	CC3_51-519...	Wally	Ruth	24	Dallas	United_States	Wally.Ruth...	?	1
Row12	?	CC1_31-314...	Ron	Pick	27	Palermo	Italy	Ron .Pick@e...	?	1
Row13	?	CC3_55-550...	Ruediger	Gleßner	27	Helsinki	Finland	Ruediger.Gl...	?	1
Row14	?	CC2_28-285...	Eva-Marie	Wenk	48	Krakow	Poland	Eva-Marie....	?	0
Row15	?	CC1_64-646...	Silvia	Heinlein	25	Wandsbek	Germany	Silvia.Heinlei...	?	0
Row16	?	CC1_81-816...	Burghard	Schiffmann	21	Vienna	Austria	Burghard .S...	Burghard .S...	0
Row17	?	CC3_48-484...	Fernando	Baur	29	Zaporizhia	Ukraine	Fernando .B...	?	1
Row18	?	CC1_47-474...	Giuseppina	Nitsch	44	Denver	United_States	Giuseppina....	?	1
Row19	?	CC1_19-191...	Albrecht	Panzer	50	Athens	Greece	Albrecht .Pa...	?	1
Row20	?	CC1_94-940...	Hans-Peter	Weinrich	49	Donetsk	Ukraine	Hans-Peter ...	?	1
Row21	?	CC1_19-196...	Frithjof	Gradl	40	Memphis	United_States	Frithjof .Gra...	?	1
Row22	?	CC1_79-792...	George	Stingl	28	Washington	United_States	George .Stin...	?	0
Row23	?	CC2_87-876...	Lorenzo	Kleinhans	45	Jacksonville	United_States	Lorenzo .Kle...	?	0
Row24	?	CC3_28-286...	Nicole	Niemann	36	Sevastopol	Ukraine	Nicole.Niem...	?	0
Row25	?	CC3_97-974...	Heidelinde	Distler	25	Donetsk	Ukraine	Heidelinde.D...	?	0
Row26	?	CC1_81-815...	Maria-Luise	Ostermeier	26	Gdansk	Poland	Maria-Luise....	?	1
Row27	?	CC2_82-828...	Claudia	Schnelle	37	Toulouse	Franec	Claudia.Sch...	?	1
Row28	?	CC3_64-642...	Blanka	Sandner	44	Gdansk	Poland	Blanka.Sand...	?	1
Row29	?	CC2_92-927...	Jörg-Peter	Mager	17	Helsinki	Finland	Jörg-Peter	?	1
Row30	?	CC3_82-824...	Mattias	Sommerfeld	24	Athens	Greece	Mattias .So...	?	1
Row31	?	CC1_84-848...	Ludwig	Vogel	25	Bristol	United_King...	Ludwig .Vog...	?	1

OK | Apply | Cancel

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Common Settings: File Path

- A path consists of three parts:
  - **Type**: Specifies the file system type - e.g., local, relative, mountpoint, custom URL or connected
  - **Specifier**: Optional string with additional file system specific information - e.g. relative to which location (knime.workflow, LOCAL mountpoint…)
  - **Path**: Specifies the location within the file system

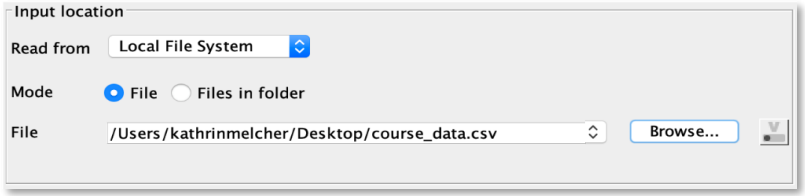


- Examples:
  - (LOCAL, , C:\Users\username\Desktop)
  - (RELATIVE, knime.workflow, file1.csv)
  - (MOUNTPOINT, MOUNTPOINT_NAME, /path/to/file1.csv)
  - (CONNECTED, amazon-s3:eu-west-1, /mybucket/file1.csv)

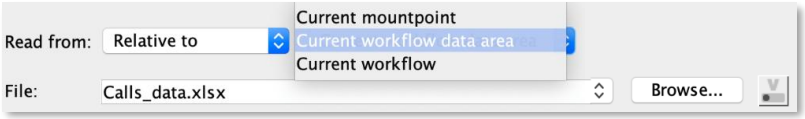
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Common Settings: Four Default File Systems

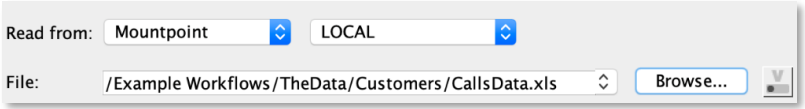
- Local File System



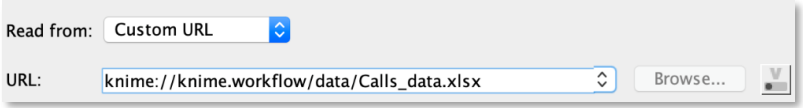
- Relative to …



- Mountpoint



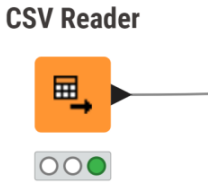
- Custom URL

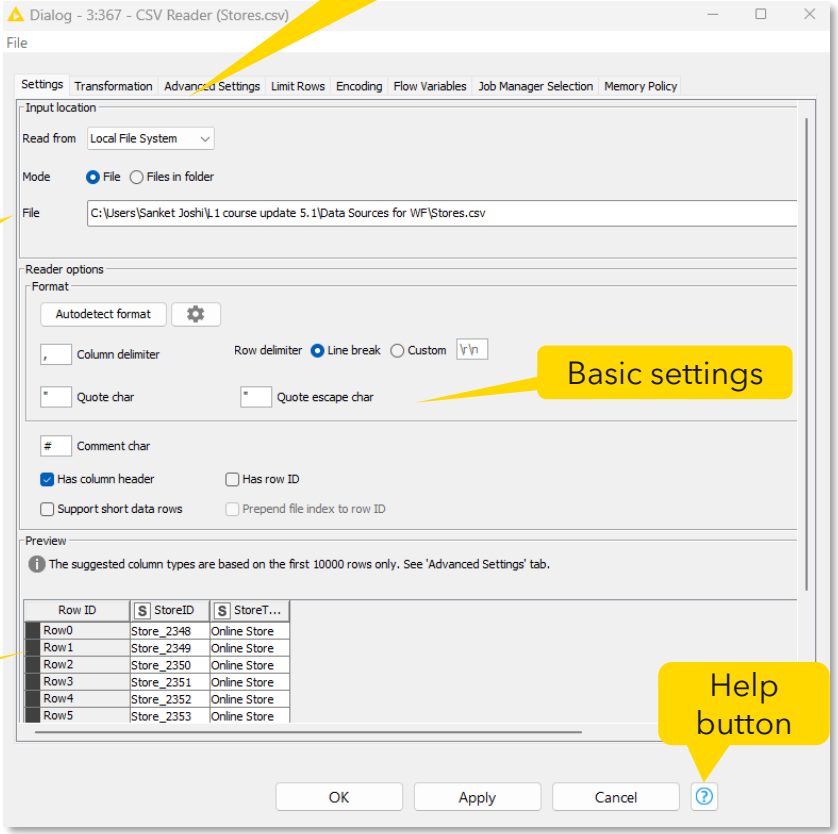


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

CSV Reader

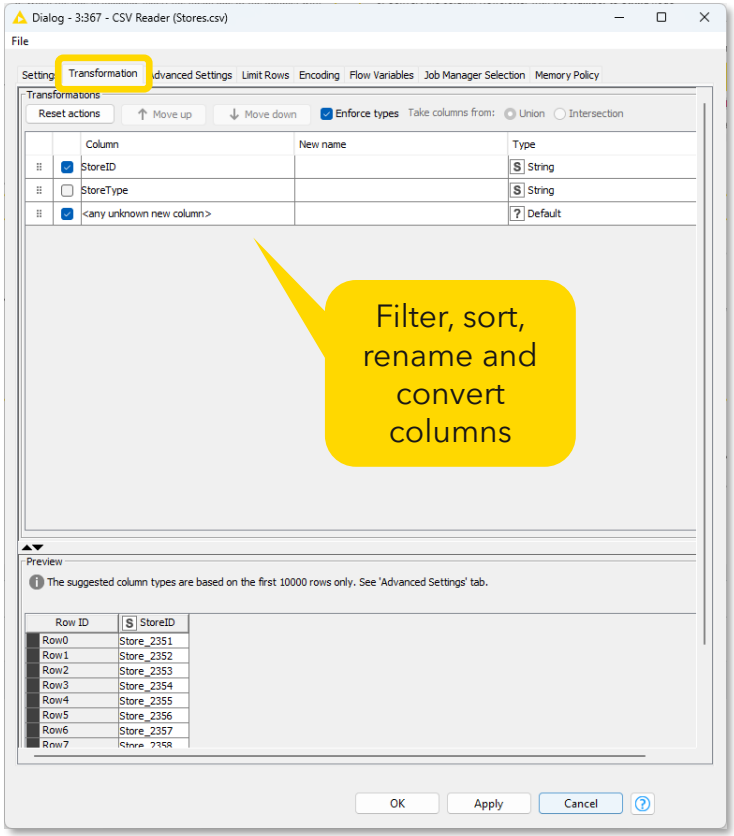
- Reads either one or multiple .csv and .txt files
- Further tabs to
  - Select columns
  - Limit the rows
  - Handle quotes
  - Select encoding

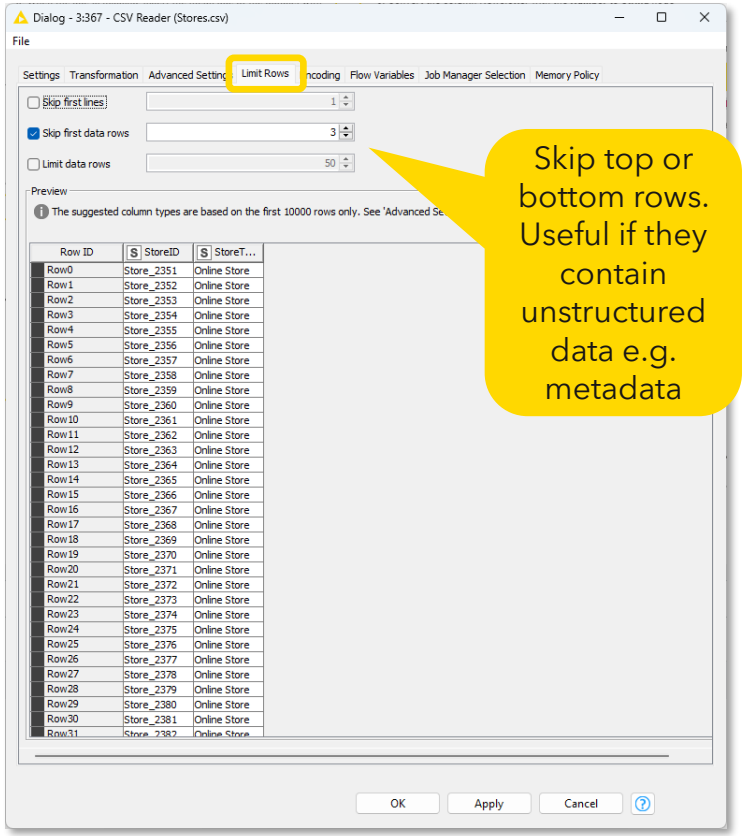




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

CSV Reader

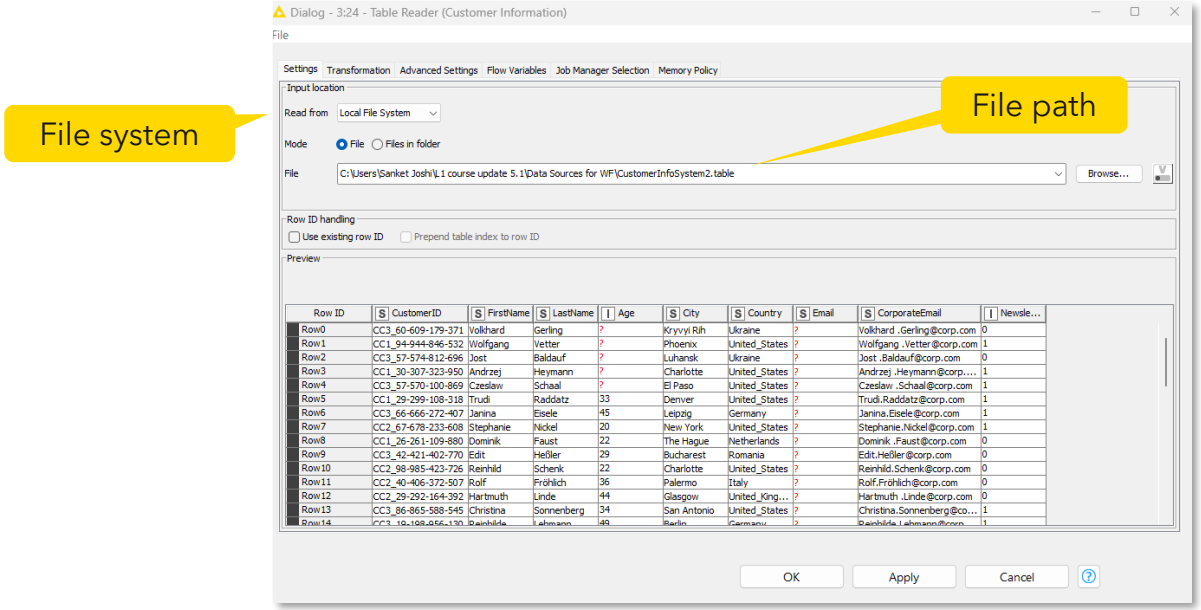




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Table Reader

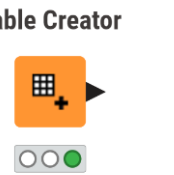
- Reads tables from the native KNIME Format
- Maximum performance, minimum configuration

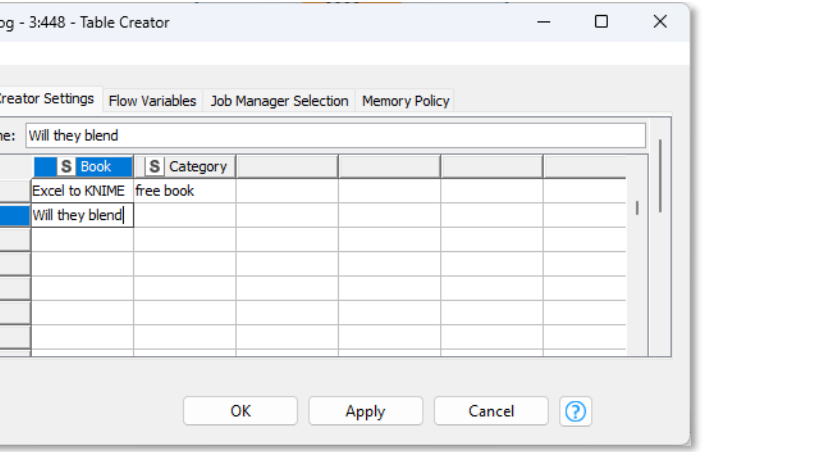


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Table Creator

- Create data tables manually
- Enter data in a spreadsheet-like configuration window

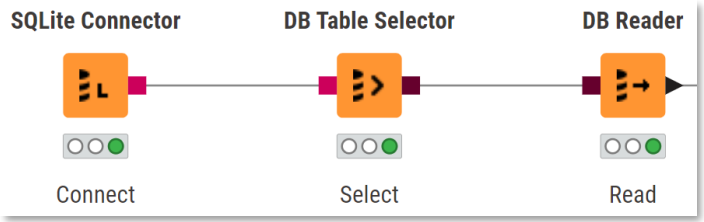




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Database Connectivity

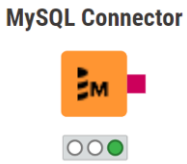
- Read data from any JDBC enabled database
- Write your own SQL or define query using dedicated nodes

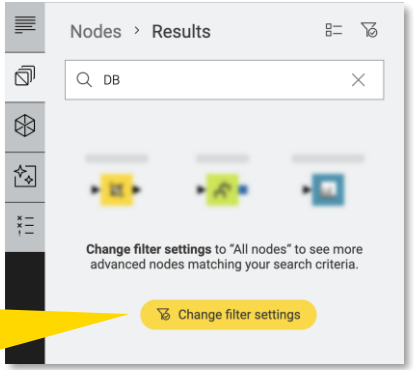


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

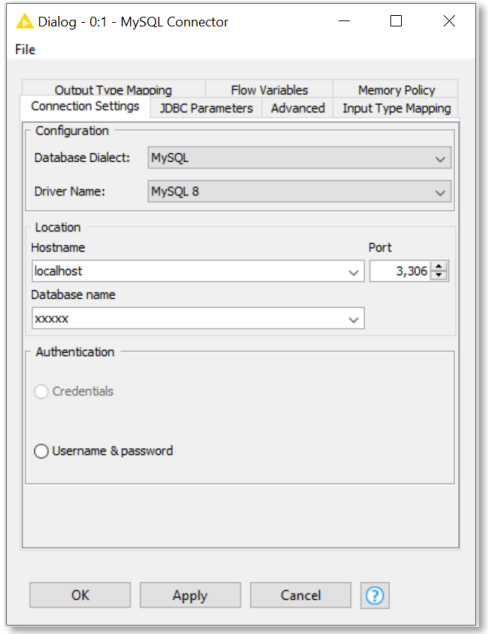
Database Connectors

- Native: PostgreSQL, MySQL, MS SQL Server, SQLite
- DB Connector (e.g., DB2, HANA)
- Big Data: HIVE and Impala



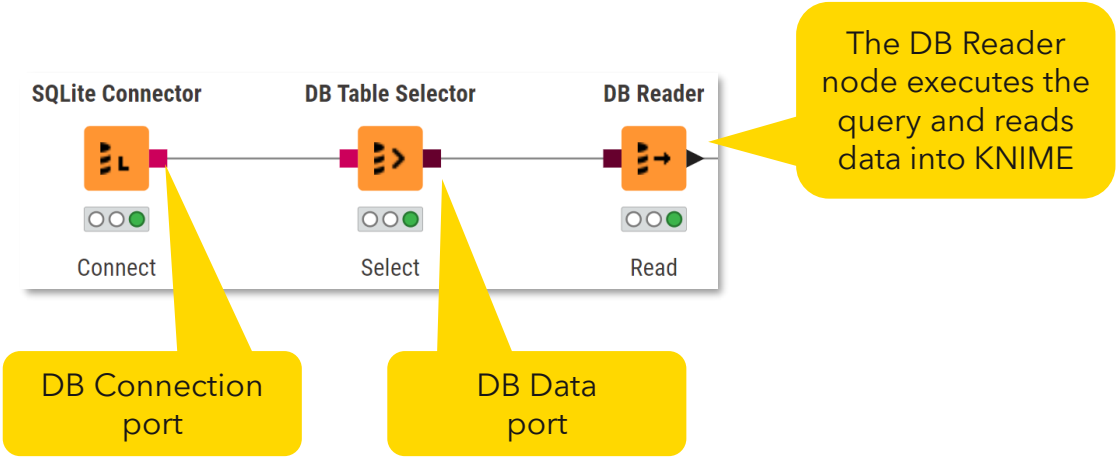


Change the filter settings to see these nodes



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

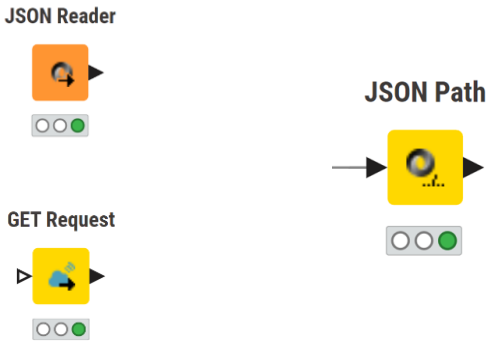
Database Query

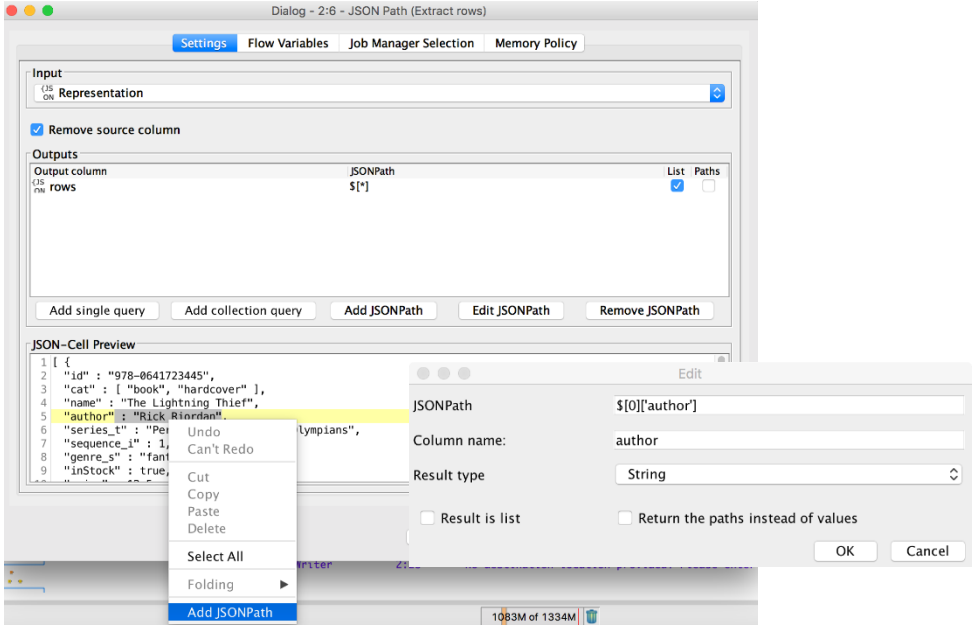


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Read JSON format

- Use the JSON Reader (or GET Request) node to get a JSON cell
- Use the JSON Path node to query the JSON file and extract parameters
  - Editor window simplifies construction of JSON queries by auto-generating them

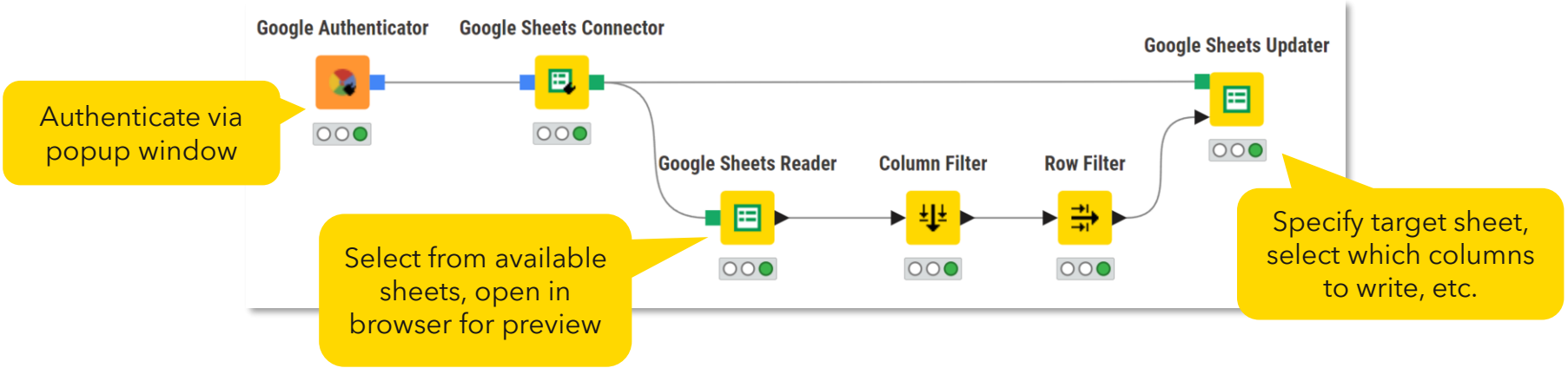




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Google Sheets

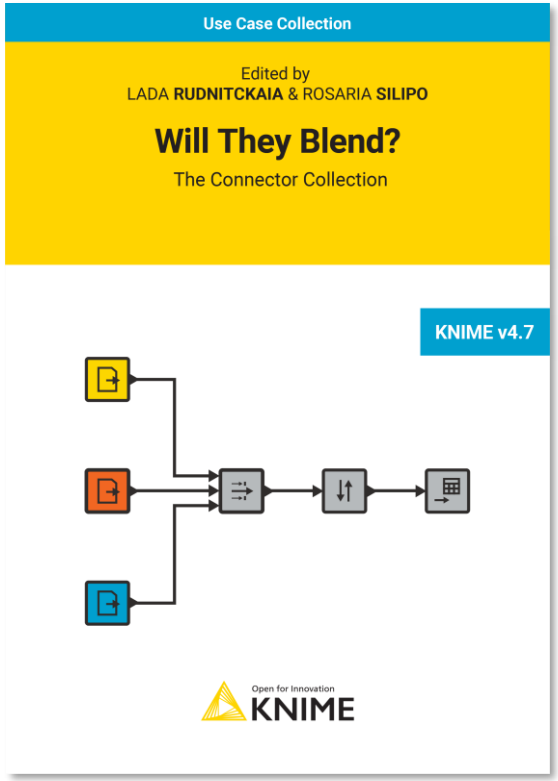
- Select from available sheets on Google Drive
- Transform data in KNIME, or enrich with new data
- Create new sheet or update existing sheets
  - Allows to read from / write to specific range of sheets (e.g. A1:G10)



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Other Useful Data Sources

- KNIME Analytics Platform provides many more options to access data:
  - Azure Data Lake Storage
  - Snowflake Connector
  - SMB Connector (e.g. Samba and Windows Server)
  - Python/R Source nodes
  - Tika Parser – extracts textual data from 200+ file types



Download the free book

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 1: Summary

- Now you are able to:
  1. Install KNIME Analytics Platform and extensions
  2. Navigate and operate the KNIME User Interface
  3. Build you first workflow
  4. Access and Import workflows from KNIME Community Hub
  5. Read different file formats
  6. Read data from databases

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Exercises

Exercises Session 1

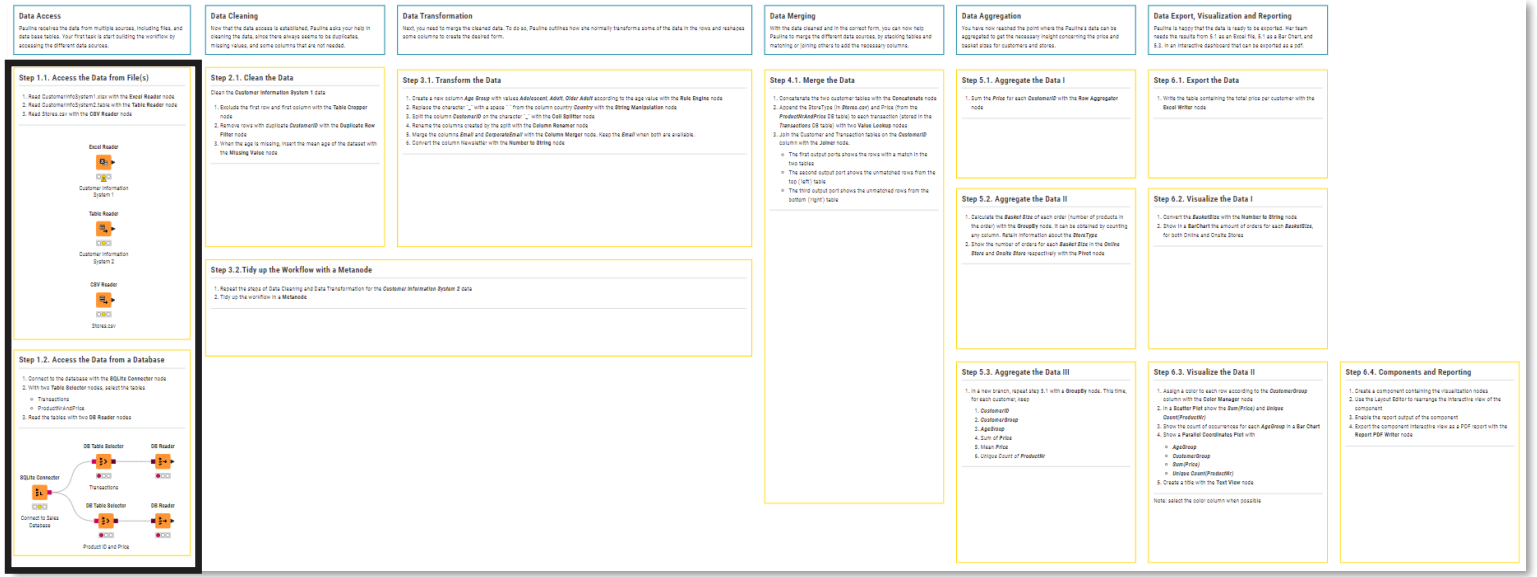
- **Step 00**
  - Download and install KNIME Analytics Platform
  - Get familiar with the User Interface and create your first workflow
  - Download and import the training workflows for this course
  - Open the ***Customer Transactions Analysis - Exercise*** workflow and complete today's steps

You can download the training workflows from the KNIME Community Hub
hub.knime.com/knime/spaces/Education/latest/Courses

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Exercises Session 1

- **Step 1.1** – Read customers and stores data from .xlsx, .table, and .csv files
- **Step 1.2** – Read transactions and products data from database



- Compare your results with the *Customer Transactions Analysis - Solution* workflow

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 2
Data Cleaning, Data Transformation, and Workflow Documentation

Learning Objectives

1. Filter rows and columns
2. Transform values in cells
3. Transform tables
4. List the best practices to organize and document workflows

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

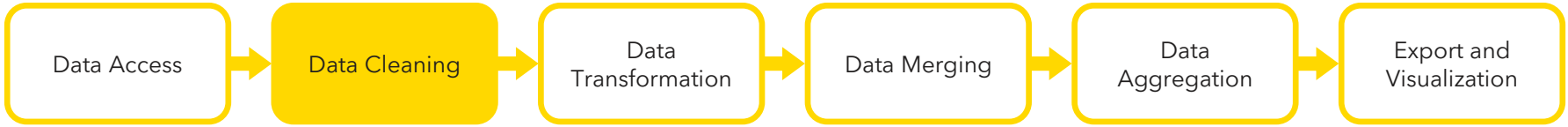
Data Cleaning

Data Cleaning

- Data are rarely clean
- Remove not useful data
- Remove repeated data
- Handle missing values

Use Case

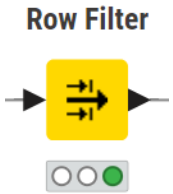
Now that the data access is established, Pauline asks your help in cleaning the data, since there seems to be duplicates, missing values, and some columns that are not needed.

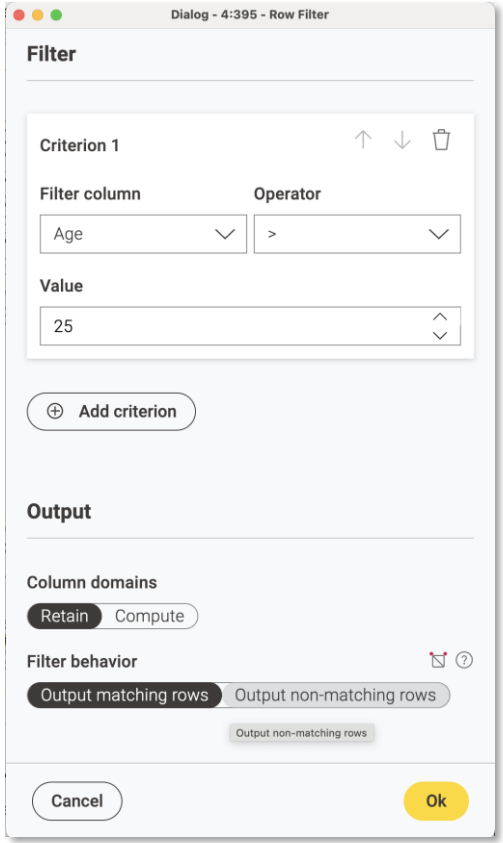


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Row Filter

- Row filtering with include and exclude options according to certain criteria
  - Certain values or patterns in the selected column
  - Row number
  - Row ID

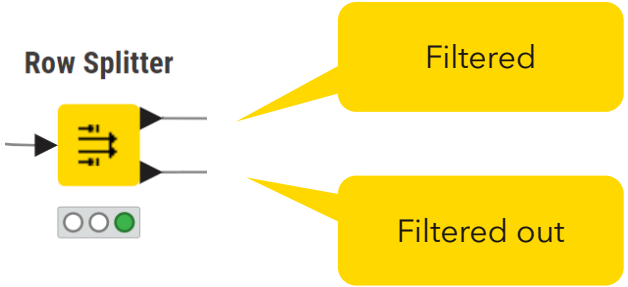


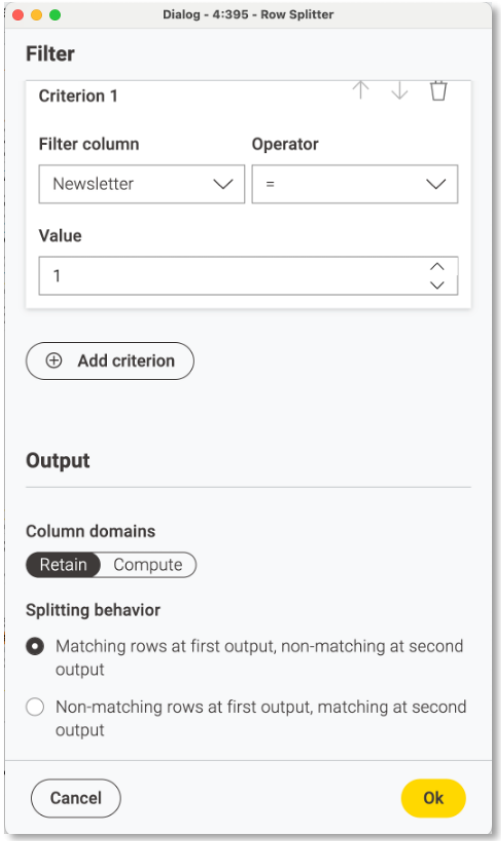


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Row Splitter

- Same configuration as Row Filter
- Filtered out rows are available in second output port

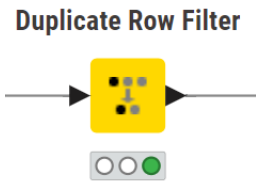


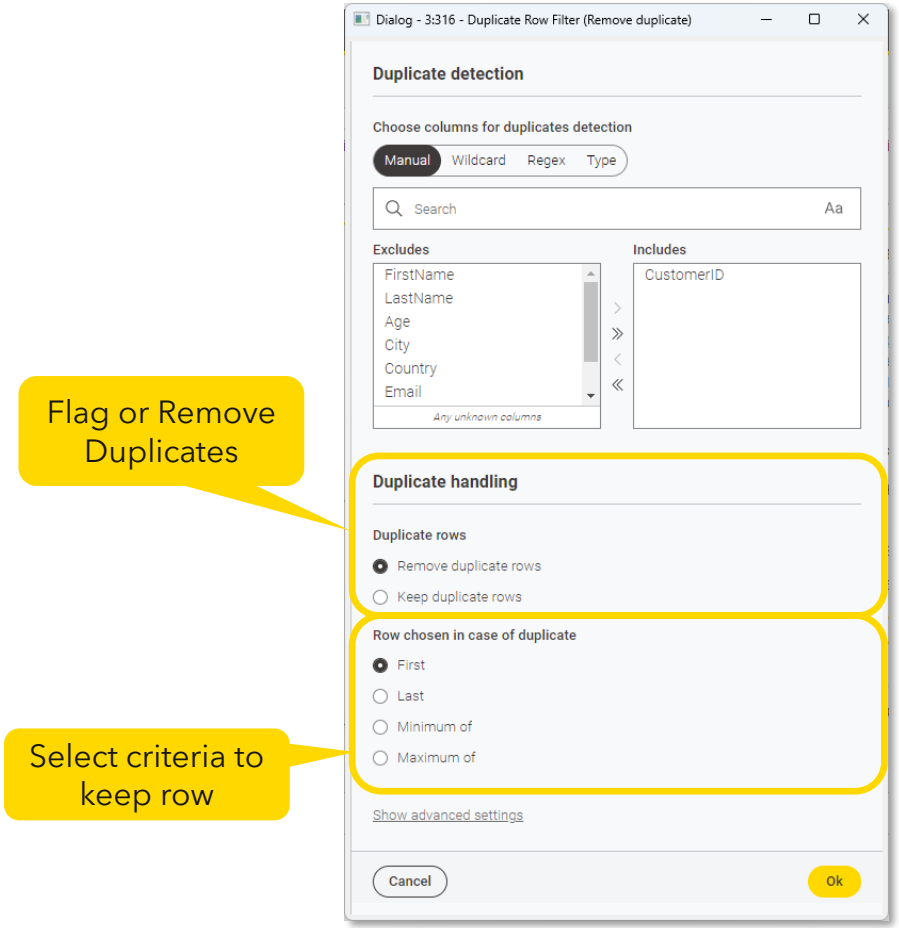


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Duplicate Row Filter

- Detect duplicate rows and apply a selected treatment
  - Select columns to check for duplicates
  - Provide options for treating duplicated values

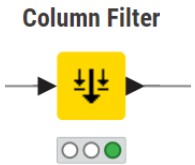




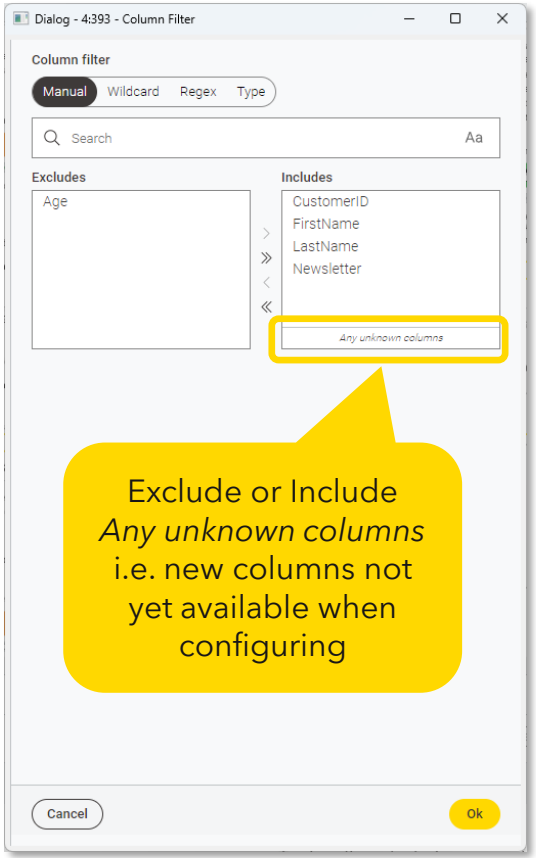
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Column Filter

- Remove columns from table



CustomerID	FirstName	LastName	Age	Newsletter
?	?	?	?	?
6589	Peter	Parker	31	yes
6768	Bruce	Banner	32	no
6925	Natasha	Romanoff	34	no

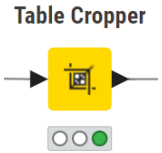


Exclude or Include *Any unknown columns* i.e. new columns not yet available when configuring

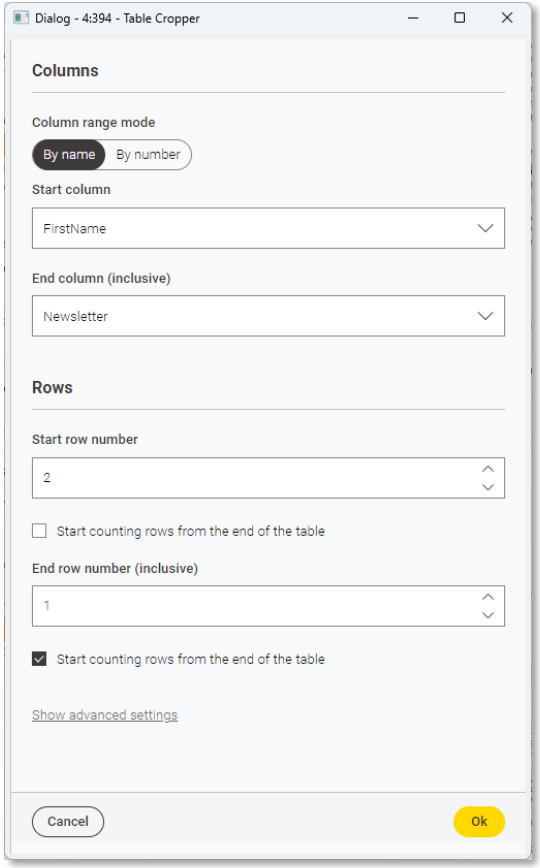
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Table Cropper

- Retain a selected, contiguous portion of a table
  - Select starting and ending rows and columns



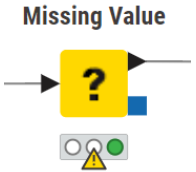
CustomerID	FirstName	LastName	Age	Newsletter
?	?	?	?	?
6589	Peter	Parker	31	yes
6768	Bruce	Banner	32	no
6925	Natasha	Romanoff	34	no

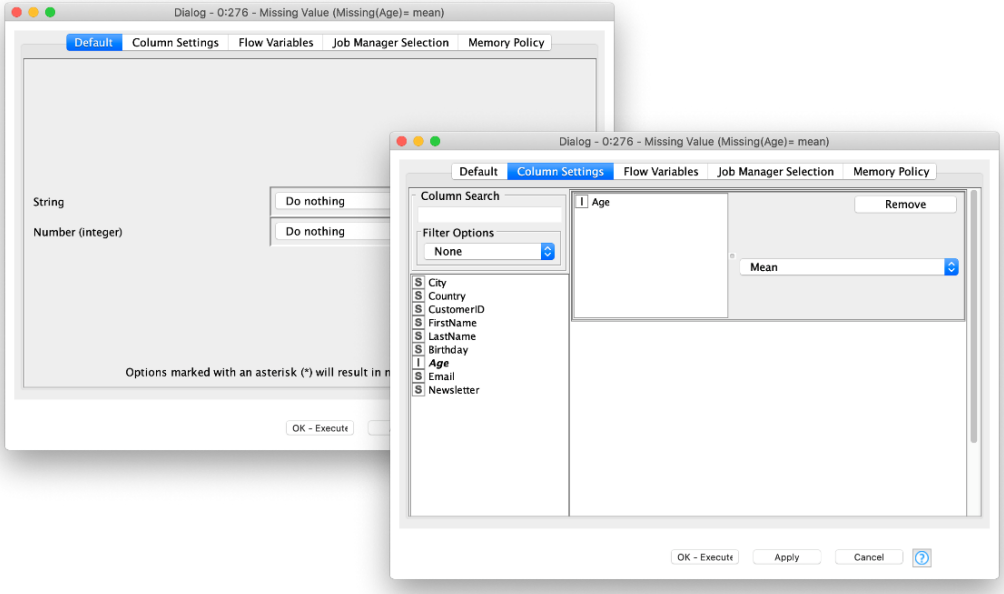


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Missing Value

- Define how to handle missing values for all columns of a given type
  - Affect all columns that are not explicitly mentioned in the second tab
- Define how to handle missing values for each available column





These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

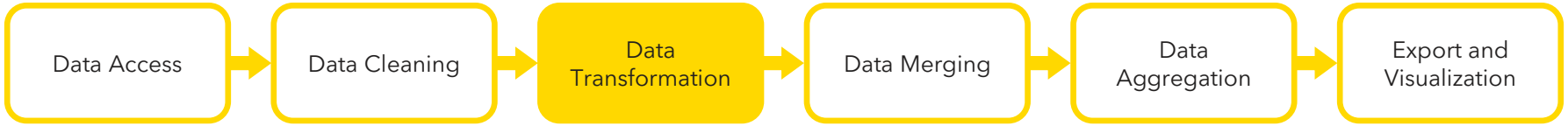
Data Transformation

Data Transformation

- The core of the data pipeline
- Extract more information
  - Define rules and mathematical operations
  - Transform at cell or row level
- Transform the data to the desired shape
  - Rename and resort table columns
  - Split and merge columns
  - Convert data types

Use Case

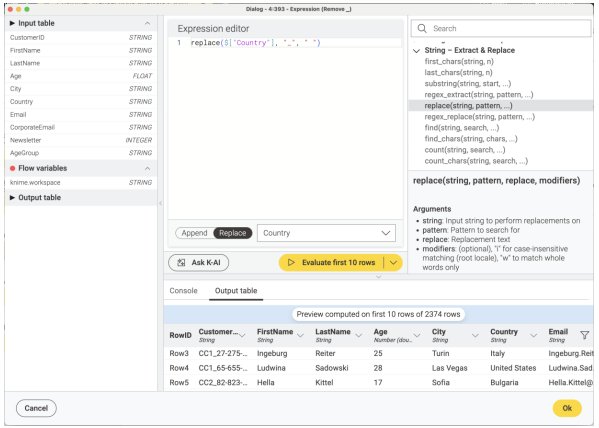
Next, you need to merge the cleaned data. To do so, Pauline outlines how she normally transforms some of the data in the rows and reshapes some columns to create the desired form.



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Expression

- String operations
  - Capitalize
  - Join strings
  - …
- Math operations
  - Round
  - Sum
  - Max
  - …
- Custom logic with if-clause



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Expression

Edit the expression

Select columns from input table

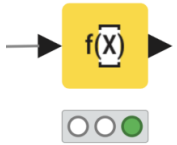
Select a function from the list

Read the documentation

K-AI copilot

Preview of the output

Expression



Dialog - 4:393 - Exp

▶ Input table

CustomerID	STRING
FirstName	STRING
LastName	STRING
Age	FLOAT
City	STRING
Country	STRING
Email	STRING
CorporateEmail	STRING
Newsletter	INTEGER
AgeGroup	STRING

● Flow variables

knime.workspace	STRING

▶ Output table

Expression editor

```
1  replace($["Country"], "_", " ")
```

Append | Replace | Country

Ask K-AI | Evaluate first 10 rows

Search

String – Extract & Replace

- first_chars(string, n)
- last_chars(string, n)
- substring(string, start, ...)
- regex_extract(string, pattern, ...)
- replace(string, pattern, ...)
- regex_replace(string, pattern, ...)
- find(string, search, ...)
- find_chars(string, chars, ...)
- count(string, search, ...)
- count_chars(string, search, ...)

replace(string, pattern, replace, modifiers)

Arguments

- **string**: Input string to perform replacements on
- **pattern**: Pattern to search for
- **replace**: Replacement text
- **modifiers**: (optional), "i" for case-insensitive matching (root locale), "w" to match whole words only

Console | Output table

Preview computed on first 10 rows of 2374 rows

RowID	Customer... String	FirstName String	LastName String	Age Number (dou...	City String	Country String	Email String
Row3	CC1_27-275-...	Ingeburg	Reiter	25	Turin	Italy	Ingeburg.Rei
Row4	CC1_65-655-...	Ludwina	Sadowski	28	Las Vegas	United States	Ludwina.Sad
Row5	CC2_82-823-...	Hella	Kittel	17	Sofia	Bulgaria	Hella.Kittel@

Cancel | Ok

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Expression with K-AI Copilot

Dialog - 4:394 - Expression (<18 Adolescents)

```
1 if(condition_1, value_1, additional_conditions, value_if_all_false)
1 if($["Age"] < 18, "Adolescent", $["Age"] <= 65, "Adult", "Older Adult")
```

▶ Input table
CustomerID
FirstName
LastName
Age
City
Country
Email
CorporateEmail
Newsletter

● Flow variables
knime.workspace

▶ Output table

K-AI provides edits (green) to the existing expression (red)

Accept the edits replace the expression in the editor

Insert in editor

Create a new column AgeGroup with values Adolescent, Adult, Older Adult according to the age value: <18 Adolescents, 18-65 Adults, >65 Older Adults

Type your prompt

Write a prompt to describe what you expect

Output column Append Replace AgeGroup

Ask K-AI

Evaluate first 10 rows

Console Output table

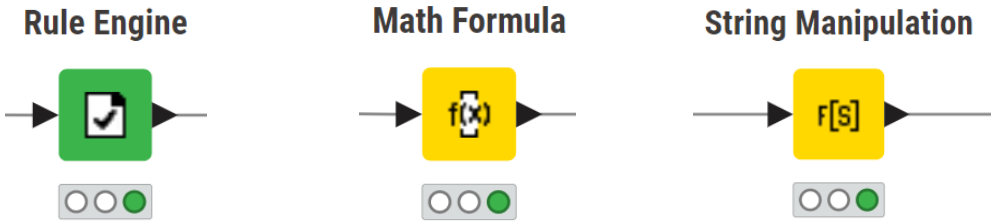
Install the AI Assistant extension

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Expressions: Legacy nodes

- Rule engine
- Math Formula
- String Manipulation

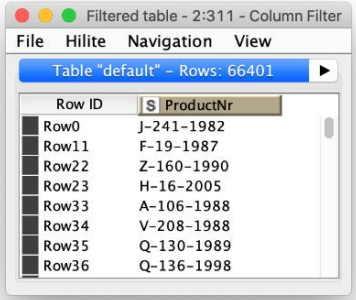
- No copilot support!

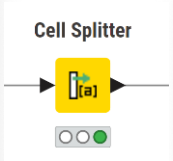


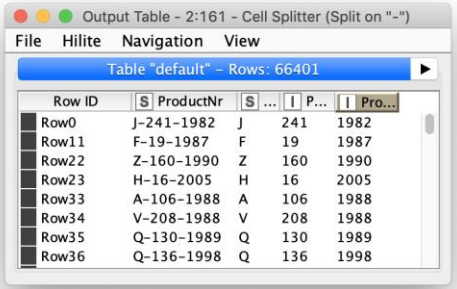
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

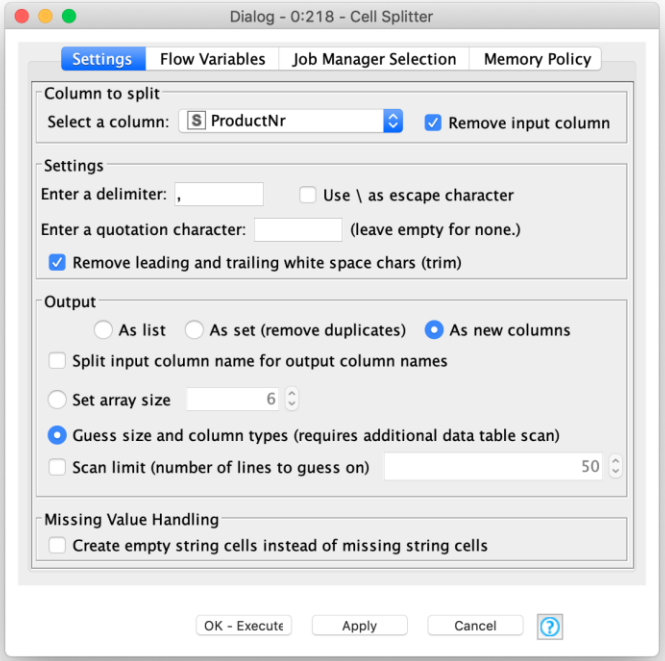
Cell Splitter

- Split the content of one column into many columns based on a delimiter





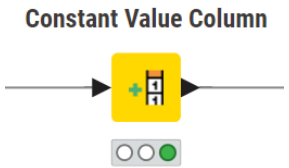


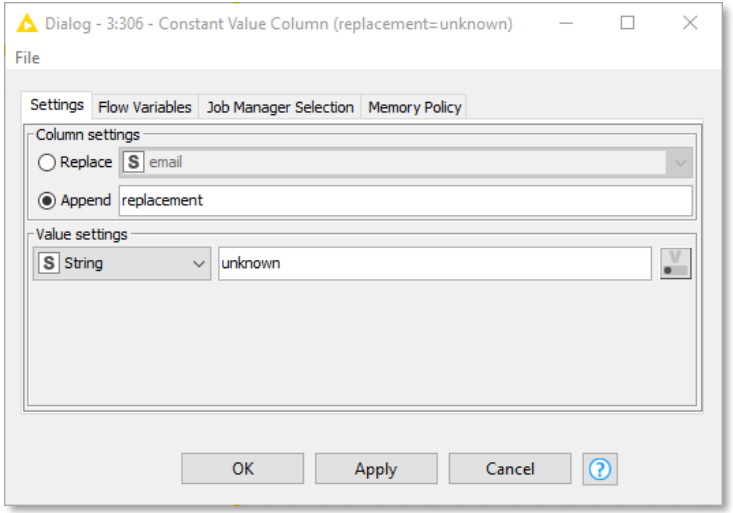


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Constant Value Column

- Add or replace a column with a single constant value
- Can be used to add an empty column



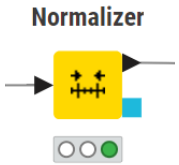


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Normalizer

- Normalize values of numeric columns
  - Min-Max Normalization
  - Z-Score (Gaussian) Normalization
  - Decimal Scaling

User	Age	Income
user1	34	30000
user2	20	45000
user3	59	20000
user4	60	100000

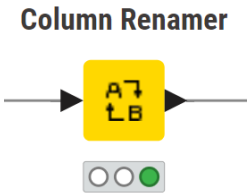


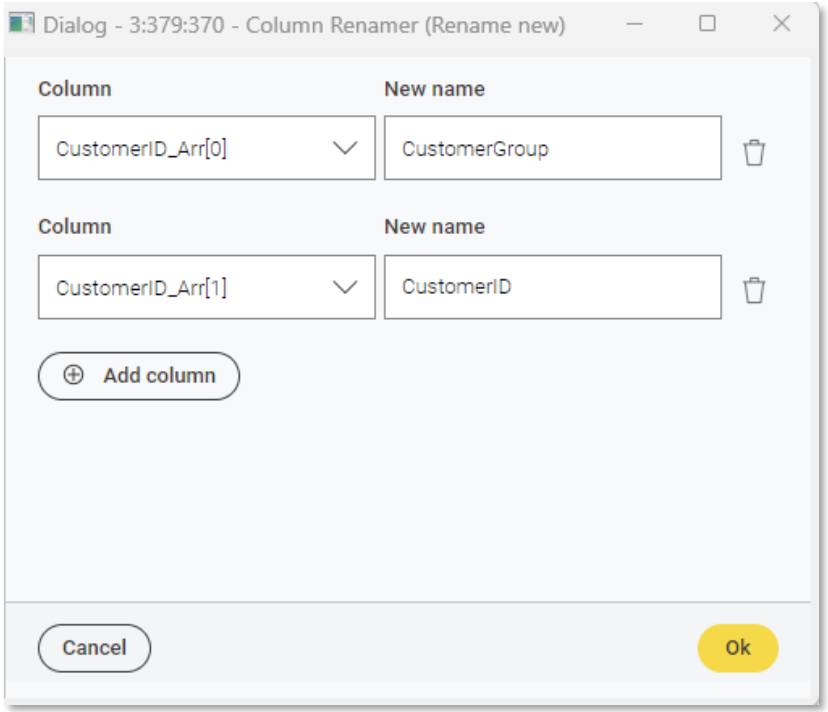
User	Age	Income
user1	0.35	0.125
user2	0	0.312
user3	0.975	0
user4	1	1

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Column Renamer

- Change name of one or more columns

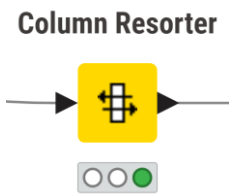


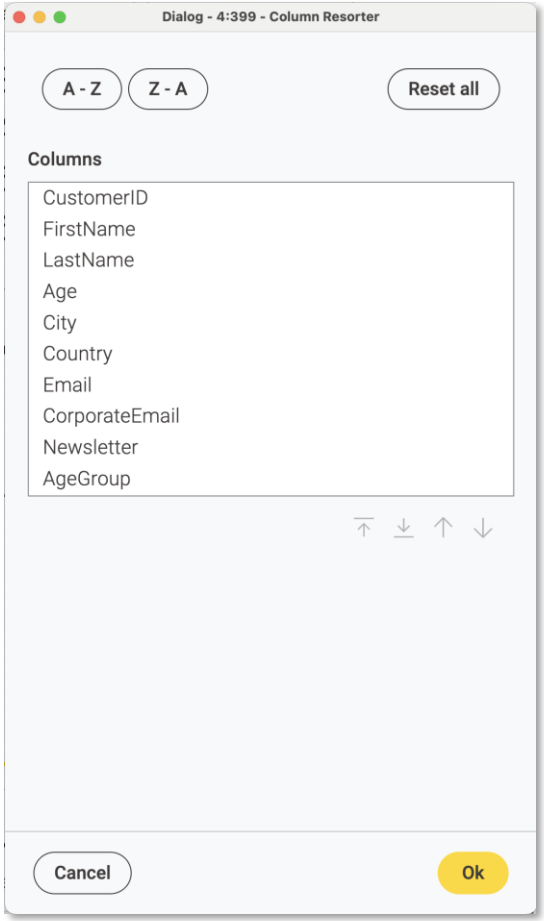


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Column Resorter

- Change the order of the columns in a table

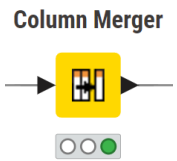


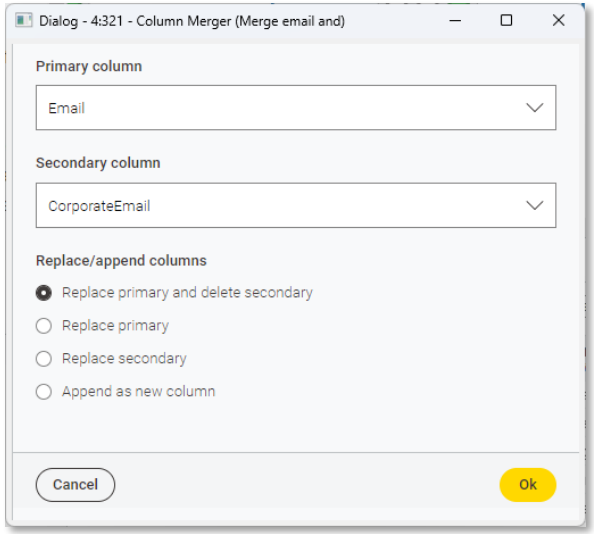


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Column Merger

- Merge two columns into one by choosing the cell that is not missing





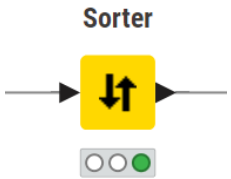
Email	Corporate Email
peter.parker@gmail.com	?
?	blackw@marvel.com
bbanner@gmail.com	thehulk@marvel.com

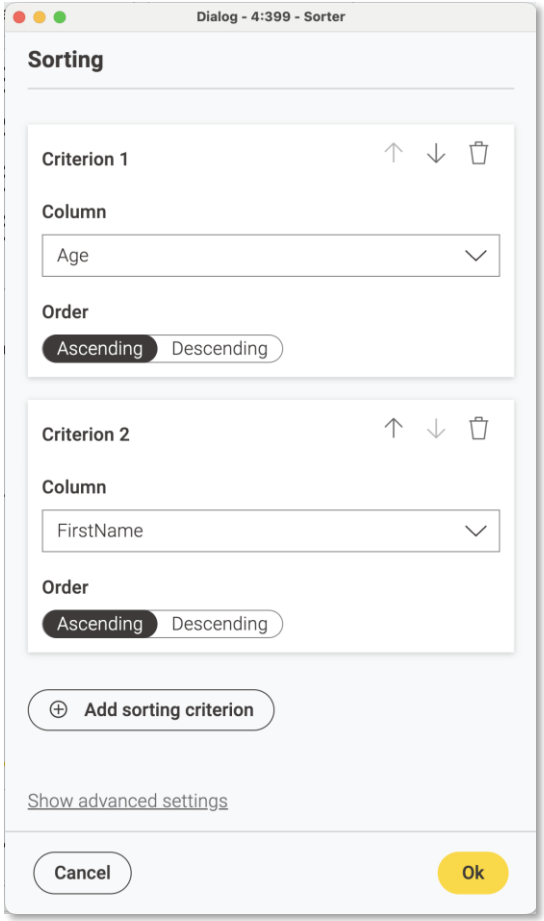
Email
peter.parker@gmail.com
blackw@marvel.com
bbanner@gmail.com

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Sorter

- Sort the rows based on the values of the selected column(s), either
  - ascending or
  - descending

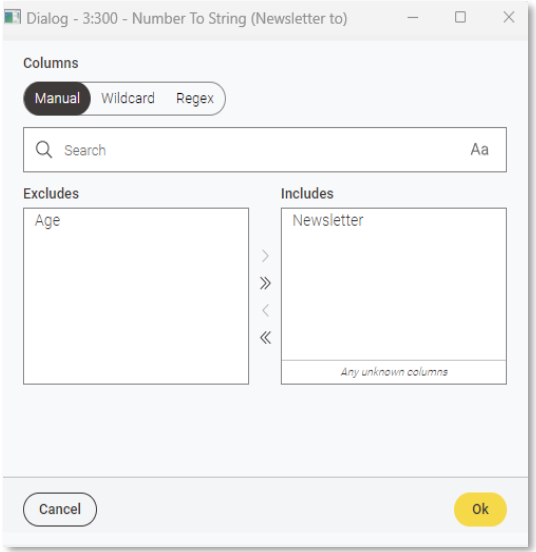




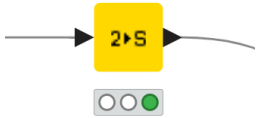
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

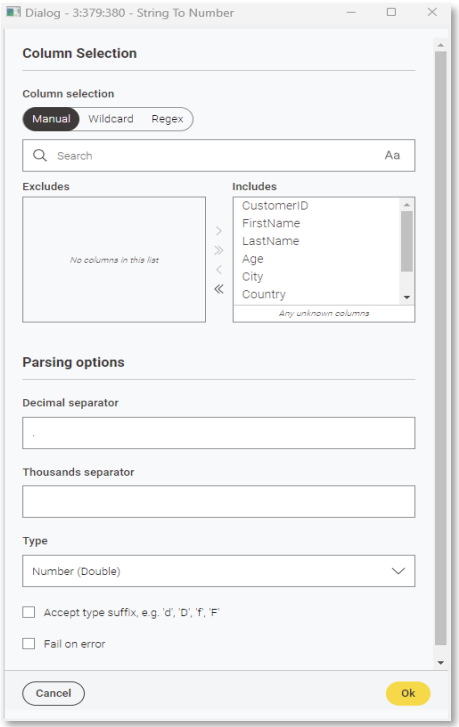
Type Conversion

- Change the data type of the selected columns

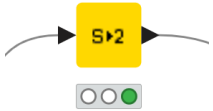


Number To String





String To Number



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Workflow Organization

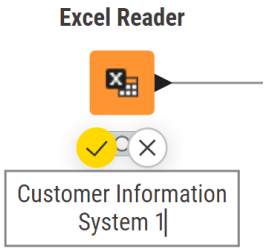
Workflow Organization – Good Practices

- Document your workflow
  - Node labels
  - Workflow annotations
- Hide complexity
  - Metanodes
  - Components

Use Case
You want to outline the workflow to Pauline, but it is messy and hard to read. Keep the workflow tidy and documented – for you and your collaborators.

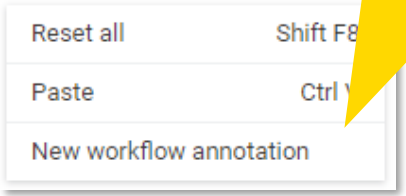
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

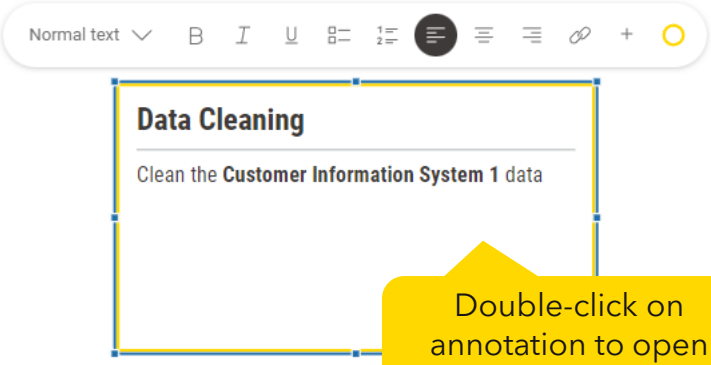
Node Labels & Workflow Annotations



Double-click to change the node label

Right click in the workflow and select New Workflow Annotation



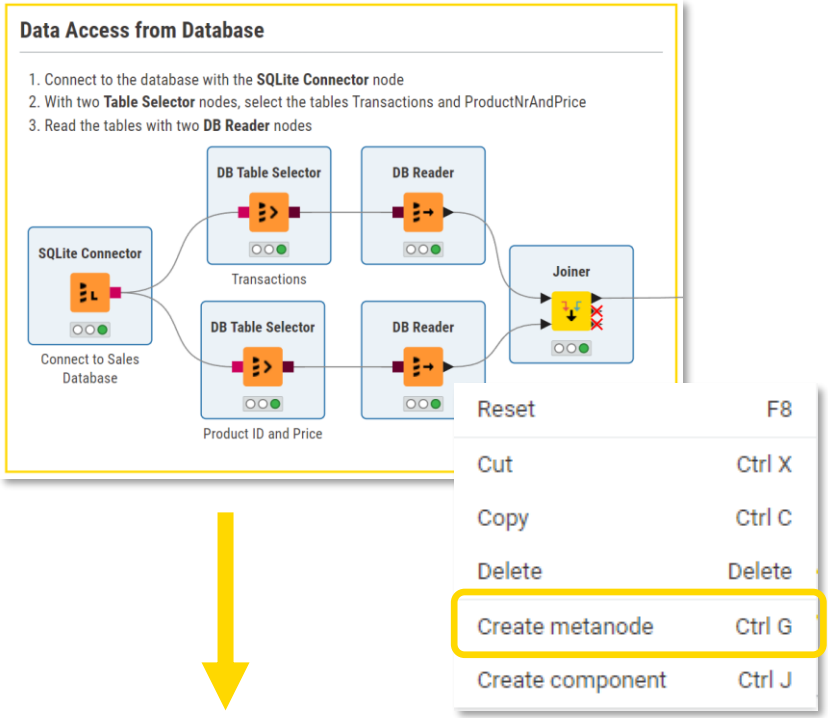


Double-click on annotation to open the editor

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Metanodes

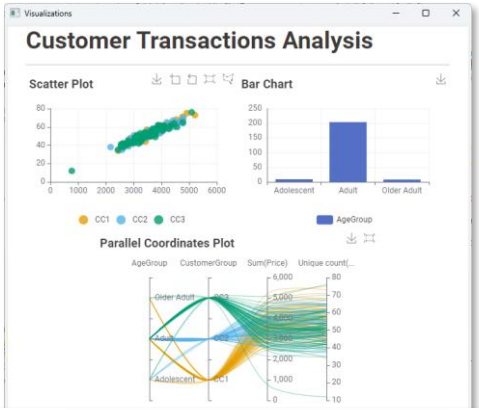
- Multiple nodes into a single node-like entity
- Organize workflow by task
- Hide complexity & improve readability
- Right click -> Create Metanode...

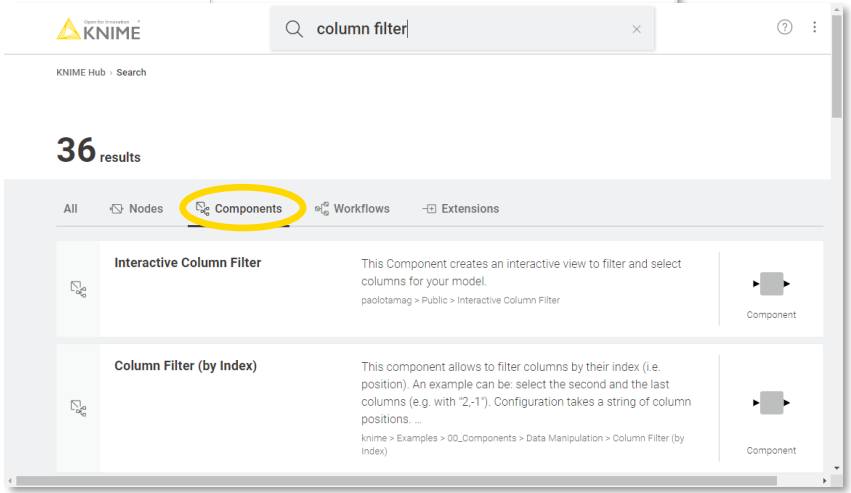


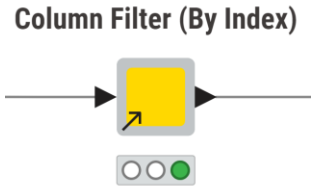
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Components

- Encapsulate multiple nodes into a single node-like entity (just like metanodes)
- More complex than metanodes
  - Can be configured as any KNIME nodes
  - Can create composite views
  - Can be accessed and shared on the KNIME Community Hub







These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 2: Summary

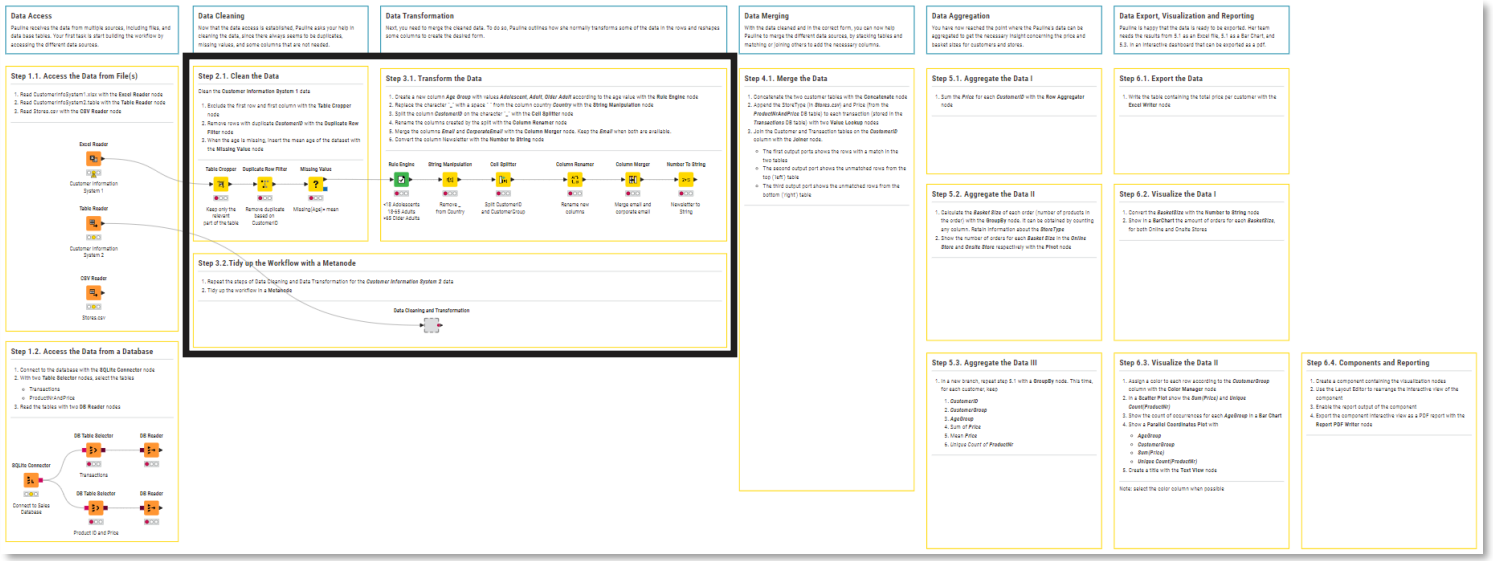
- Now you are able to:
  1. Filter rows and columns
  2. Transform values in cells
  3. Transform tables
  4. List the best practices to organize and document workflows

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Exercises

Exercises Session 2

- **Step 2.1** – Remove duplicated and unused data
- **Step 3.1** – Transform data and table format
- **Step 3.2** – Repeat the above steps and tidy up with a metanode



- Compare your results with the *Customer Transactions Analysis - Solution* workflow

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 3
Data Merging, Data Aggregation, and Data Export

Learning Objectives

1. Concatenate multiple tables
2. Join tables in multiple ways
3. Aggregate table values
4. Write data to various file formats

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

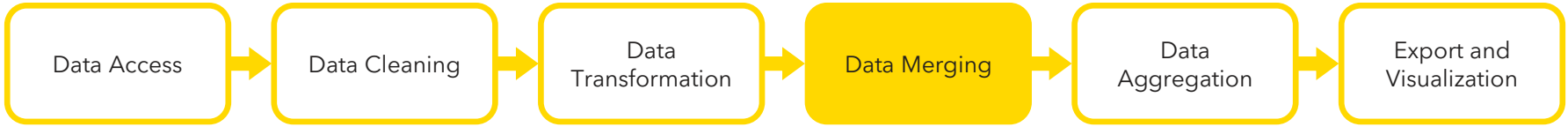
Data Merging

Data Merging

- Put together data from different tables
  - Append new data at the bottom of a table
  - Lookup information from a dictionary table
  - Join information from two tables by defining matching criteria

Use Case

With the data cleaned and in the correct shape, you can now help Pauline to merge the different data sources, by stacking tables and matching or joining others to add the necessary columns.

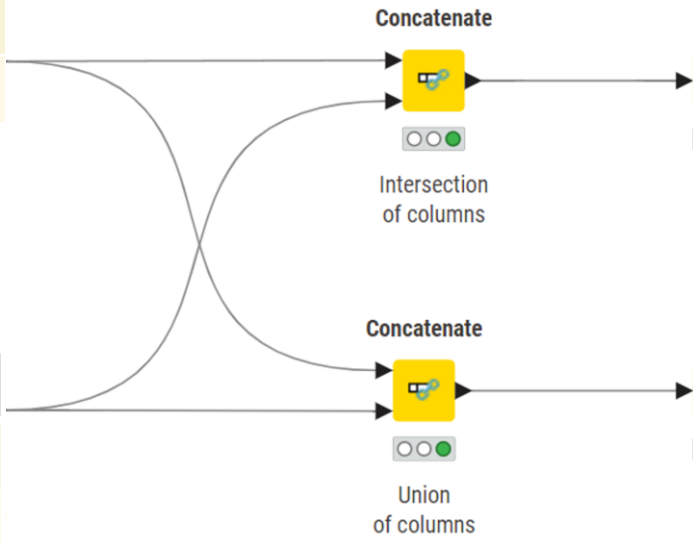


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Concatenate two tables

1	2	4
1	35	ab
2	39	cd

1	3	4
3	Yes	ab
4	No	ef



1	4
1	ab
2	cd
3	ab
4	ef

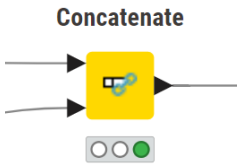
1	2	3	4
1	35	?	ab
2	39	?	cd
3	?	Yes	ab
4	?	No	ef

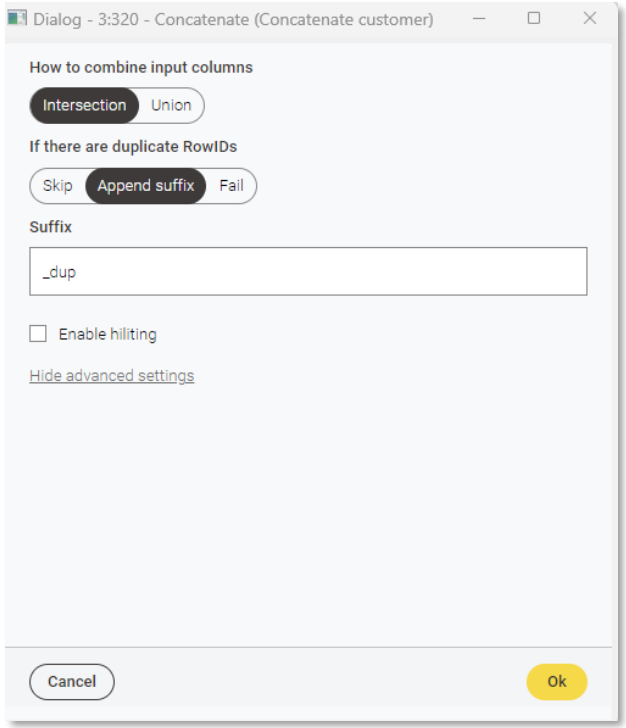
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Concatenate

Combine rows from 2 or more tables with shared columns

- Handles duplicate row keys gracefully
- Take the union or intersection of columns

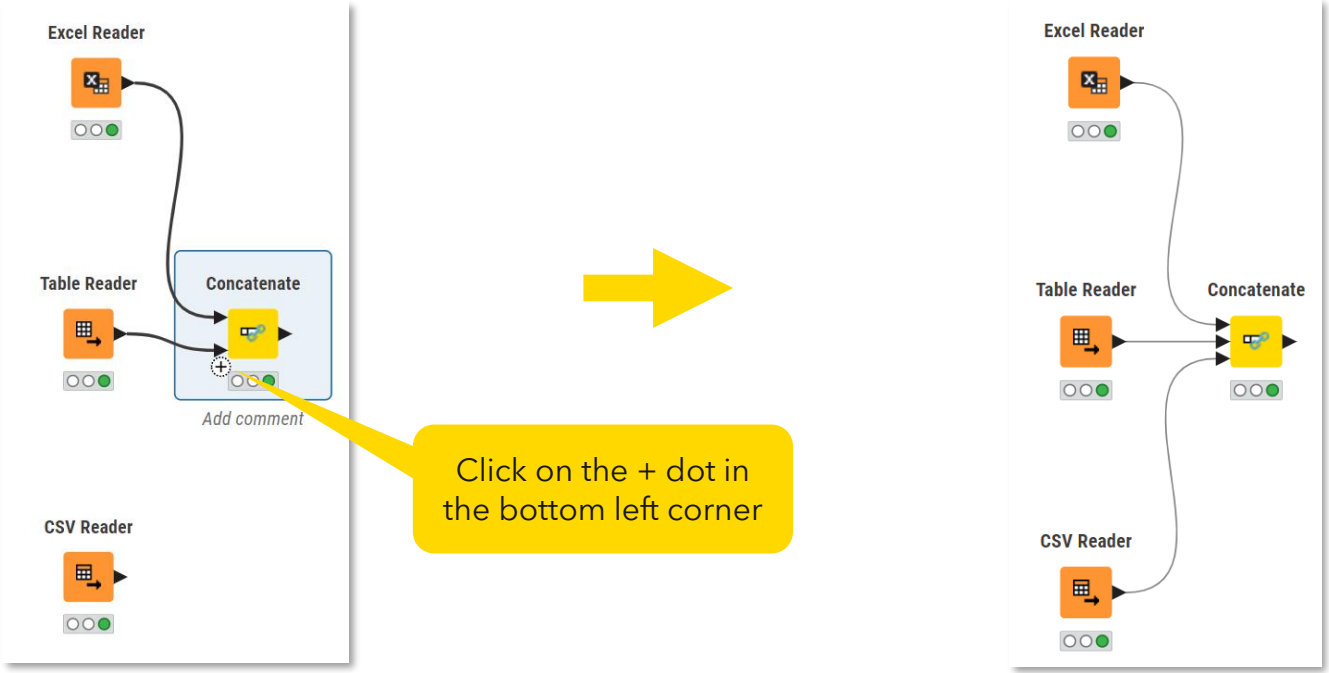




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Dynamic Ports

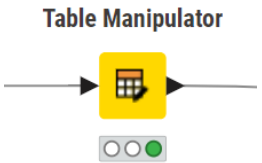
- Add and remove node ports based on your needs, e.g. in order to concatenate three or more tables

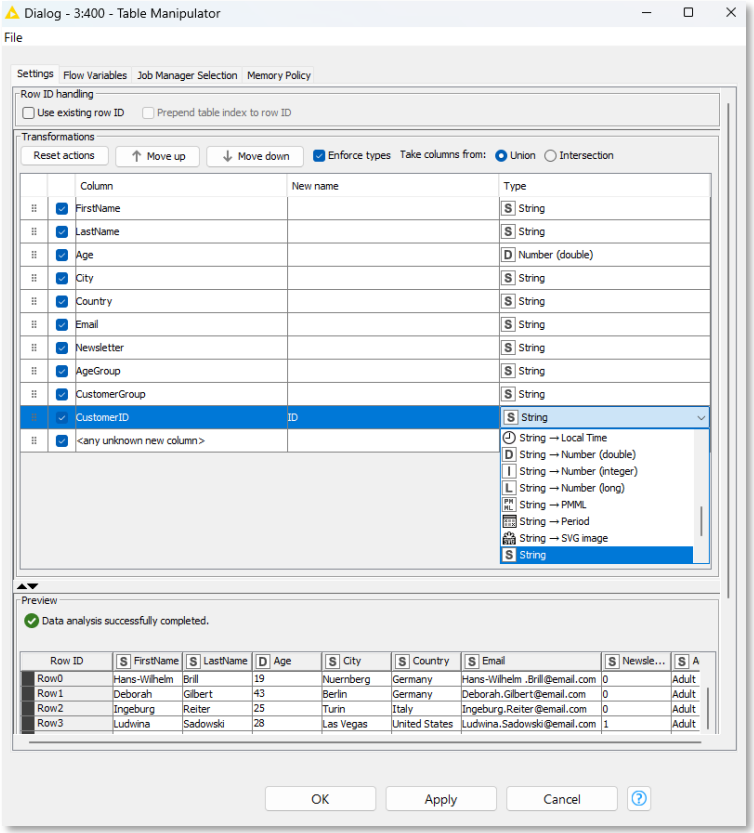


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Table Manipulator

- Multi-functions node
  - Concatenation of multiple files/tables
  - Column filtering
  - Column sorting
  - Column renaming
  - Column type mapping

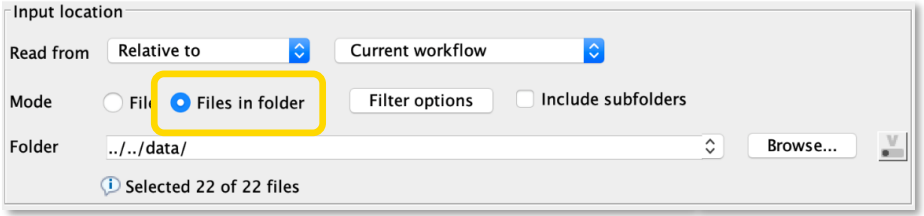


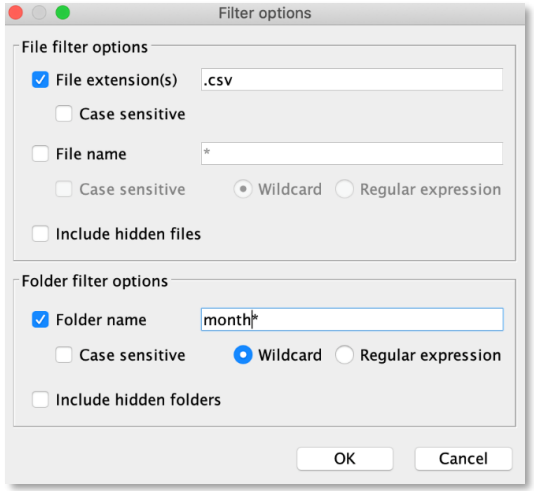


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Read Multiple Files

- Built-in option of reader nodes
- Concatenate content of multiple files in a single table
  - Option to include subfolder
  - Option to define filter criterions





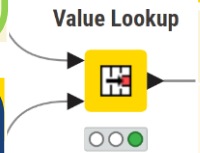
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Value Lookup

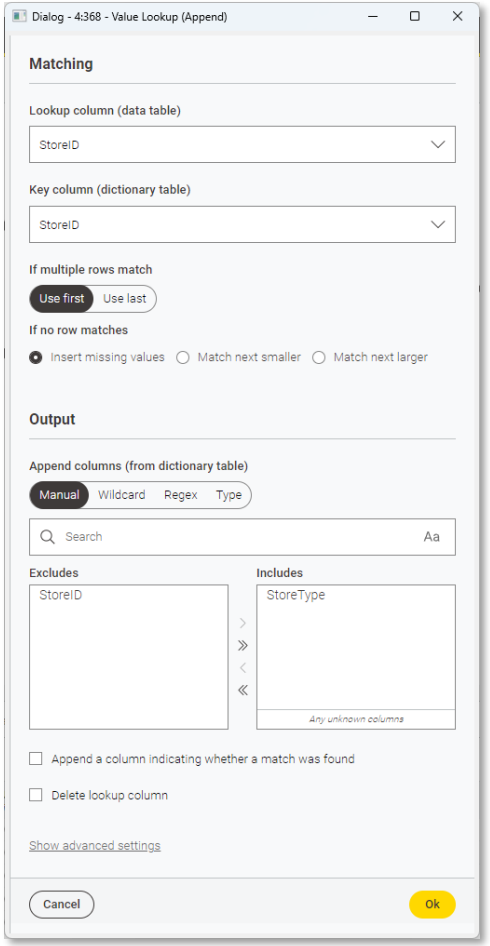
- Append one or more columns from a dictionary table
  - Match on lookup column present in both tables

OrderNr	StoreID
23893756	Store_2394
23893794	Store_2395
23893799	Store_2392

StoreID	StoreType
Store_2392	Online
Store_2393	Online
Store_2394	Onsite
Store_2395	Onsite



OrderNr	StoreID	StoreType
23893756	Store_2394	Onsite
23893794	Store_2395	Onsite
23893799	Store_2392	Online



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Joining Columns of Data

Left Table

CustomerKey	OrderDate	OrderID
22	2019-09-23	#23444
24	2019-09-30	#23457
15	2019-10-07	#28985
10	2091-10-13	#29999

Right Table

CustomerKey	DoB	City	Gender
17	1974-02-23	Berlin	F
65	2001-05-25	Stuttgart	F
35	1988-08-05	Cologne	M
15	1983-07-20	Hamburg	M
10	1993-01-13	Berlin	M

Join by CustomerKey

Inner Join

CustomerKey	OrderDate	OrderID	DoB	City	Gender
15	2019-10-07	#28985	1983-07-20	Hamburg	M
10	2091-10-13	#29999	1993-01-13	Berlin	M

Left Outer Join

CustomerKey	OrderDate	OrderID	DoB	City	Gender
22	2019-09-23	#23444	?	?	?
24	2019-09-30	#23457	?	?	?
15	2019-10-07	#28985	1983-07-20	Hamburg	M
10	2091-10-13	#29999	1993-01-13	Berlin	M

Right Outer Join

CustomerKey	OrderDate	OrderID	DoB	City	Gender
17	?	?	1974-02-23	Berlin	F
65	?	?	2001-05-25	Stuttgart	F
35	?	?	1988-08-05	Cologne	M
15	2019-10-07	#28985	1983-07-20	Hamburg	M
10	2091-10-13	#29999	1993-01-13	Berlin	M

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Joining Columns of Data

Left Table

CustomerKey	OrderDate	OrderID
22	2019-09-23	#23444
24	2019-09-30	#23457
15	2019-10-07	#28985
10	2091-10-13	#29999

Right Table

CustomerKey	DoB	City	Gender
17	1974-02-23	Berlin	F
65	2001-05-25	Stuttgart	F
35	1988-08-05	Cologne	M
15	1983-07-20	Hamburg	M
10	1993-01-13	Berlin	M

Join by CustomerKey

Full Outer Join

CustomerKey	OrderDate	OrderID	DoB	City	Gender
17	?	?	1974-02-23	Berlin	F
65	?	?	2001-05-25	Stuttgart	F
35	?	?	1988-08-05	Cologne	M
15	2019-10-07	#28985	1983-07-20	Hamburg	M
10	2091-10-13	#29999	1993-01-13	Berlin	M
22	2019-09-23	#23444	?	?	?
24	2019-09-30	#23457	?	?	?

Missing values in the left table

Missing values in the right table

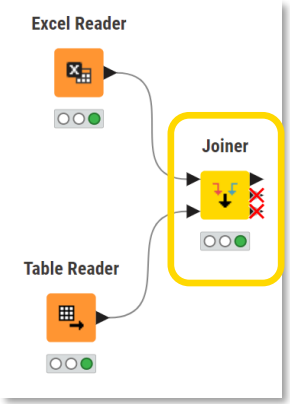
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Joiner

- Combine columns from two different tables
  - Top input port: “Left” data table
  - Bottom input port: “Right” data table

- Three Outputs:
  - Top port: Resulting joined table
  - Middle port: Unmatched rows from the left input table (top input port)
  - Bottom port: Unmatched rows from the right input table (bottom input port)

- By default the two bottom output ports are deactivated

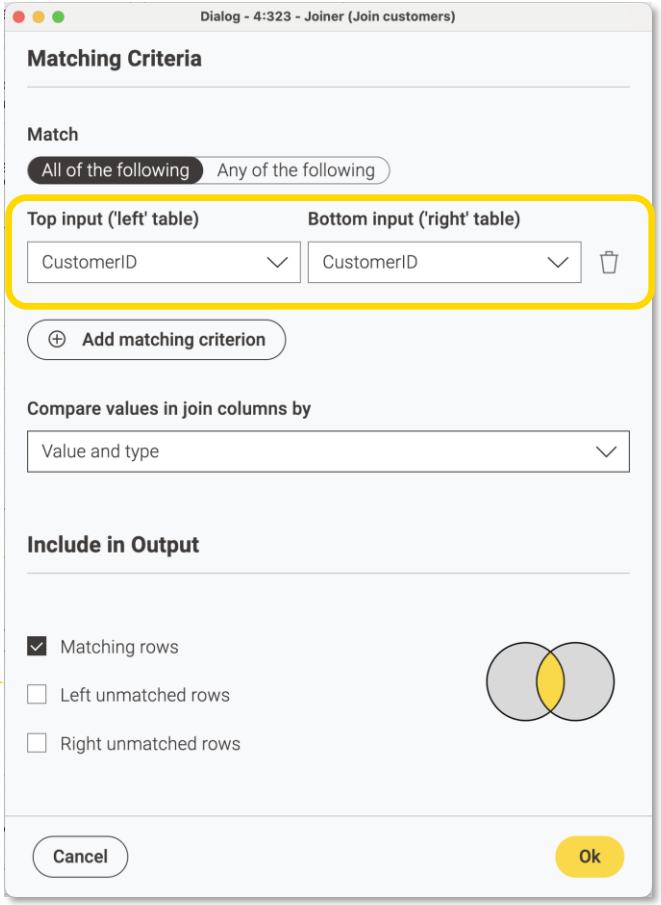


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Joiner Configuration – Linking Rows

Multiple joining columns are allowed. Consider *all* or *any* of them to define a match.

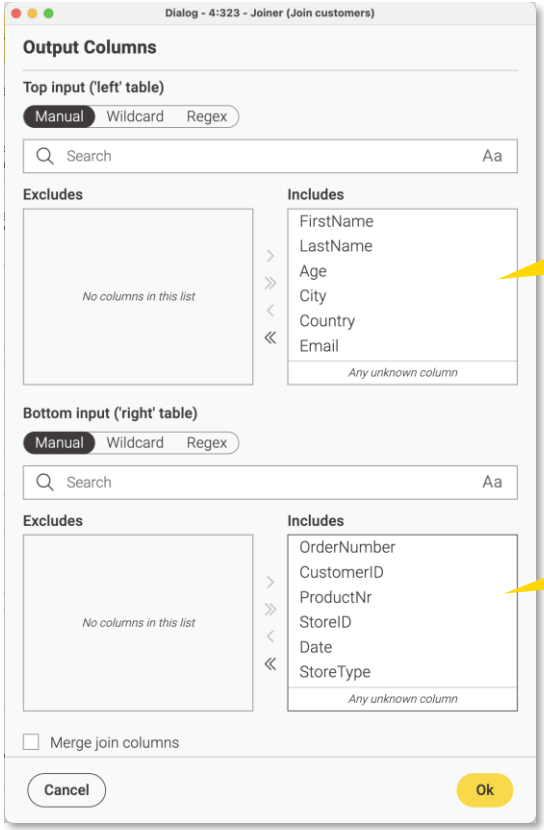
Select the rows which should be included in the joined table



The Joiner node matches all the rows with identical values in the selected column i.e. duplicate rows in case of multiple match

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Joiner Configuration – Column Selection



Columns from top table for joined table

Columns from bottom table for joined table

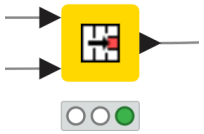
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Value Lookup vs Joiner Node

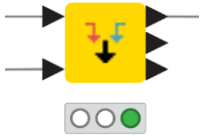
- "Simplified" join operation
- Only one matching ("lookup") column
- Handle multiple and no matches from one table
- Filter columns from dictionary table

- Complete join operation
- Multiple matching ("join") columns
- Handle unmatched rows from both tables
- Duplicate rows in case of multiple match
- Filter columns from both tables

Value Lookup



Joiner



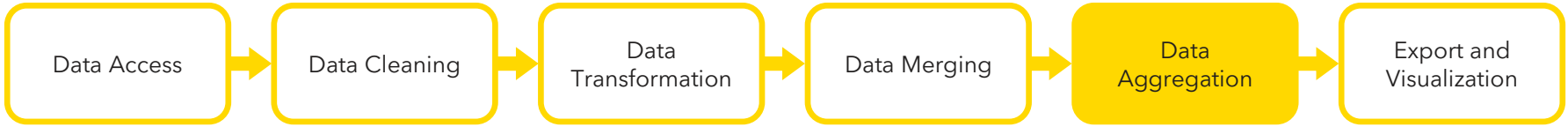
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Data Aggregation

Data Aggregation

- Create a summarized version of the data
- Aggregate by one or more categories
- Use one or more aggregation methods

Use Case
You have now reached the point where Pauline's data can be aggregated to get the necessary insight concerning the price and basket sizes for customers and stores.



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Data Aggregation

Product ID	Category	# Ordered Items
P 1	Clothing	2
P 2	Home	3
P 3	Clothing	1
P 4	Clothing	5
P 5	Electronics	7
P 6	Electronics	5

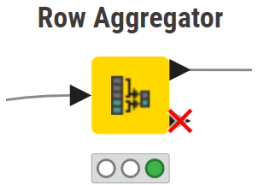
Group	Sum(# Ordered Items)
Clothing	8
Home	3
Electronics	12

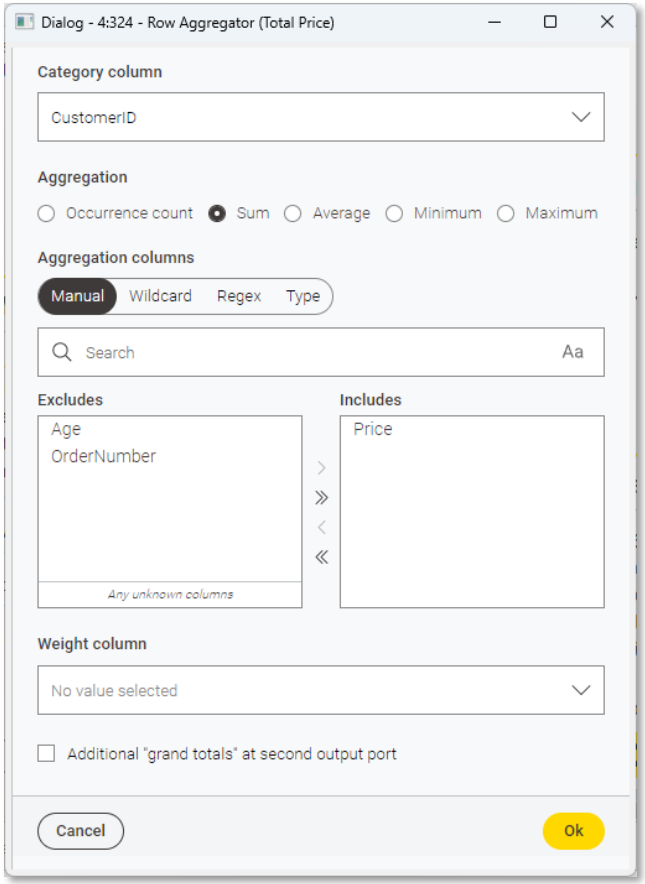
Aggregated on Category (group)
by Sum (aggregation method)

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Row Aggregator

- Straightforward aggregation
  - Select one category(group) column
  - Apply one aggregation on multiple columns
  - (Optional) Weight column

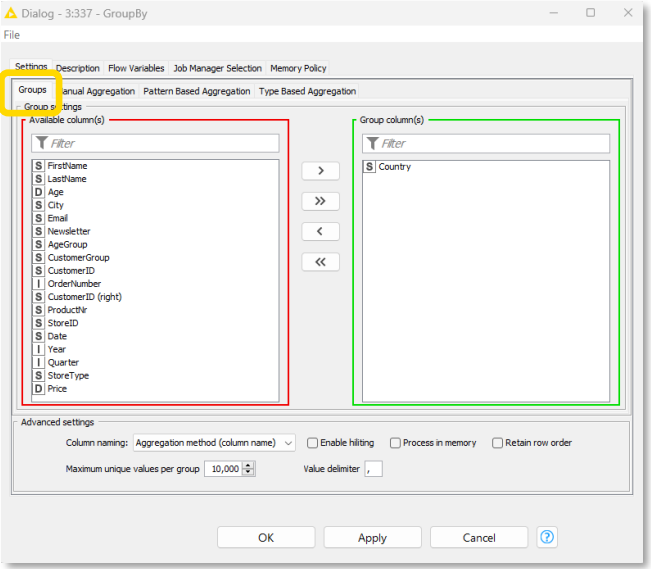


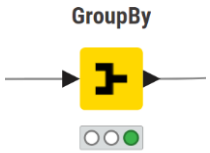


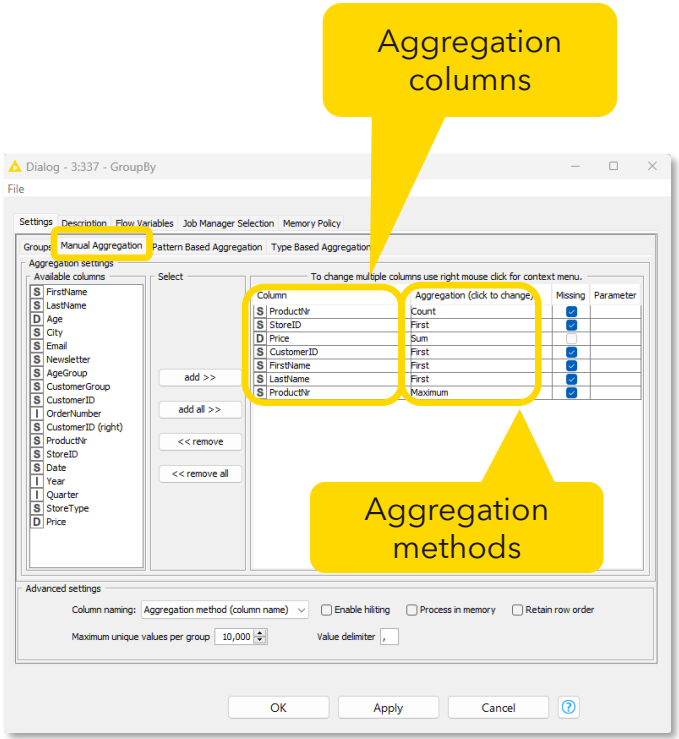
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

GroupBy

- More complex aggregation
  - Select one or more category (group) columns
  - One or more aggregations on multiple columns







These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Pivoting

Product ID	Store	Category	# Ordered Items
P 1	Online	Clothing	2
P 2	Onsite	Home	3
P 3	Onsite	Clothing	1
P 4	Online	Clothing	5
P 5	Online	Electronics	7
P 6	Online	Electronics	5

Aggregation: Count

Category	Online	Onsite
Clothing	2	1
Home	0	1
Electronics	2	0

Aggregation: Sum (# Ordered Items)

Category	Online	Onsite
Clothing	7	1
Home	0	3
Electronics	12	0

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Pivoting

Product ID	Store	Category	# Ordered Items
P 1	Online	Clothing	2
P 2	Onsite	Home	3
P 3	Onsite	Clothing	1
P 4	Online	Clothing	5
P 5	Online	Electronics	7
P 6	Online	Electronics	5

Aggregation: Sum (# Ordered Items)

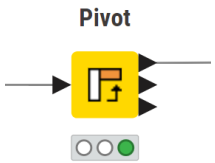
Category	Online	Onsite
Clothing	7	1
Home	0	3
Electronics	12	0

Group - Pivot - Aggregate

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

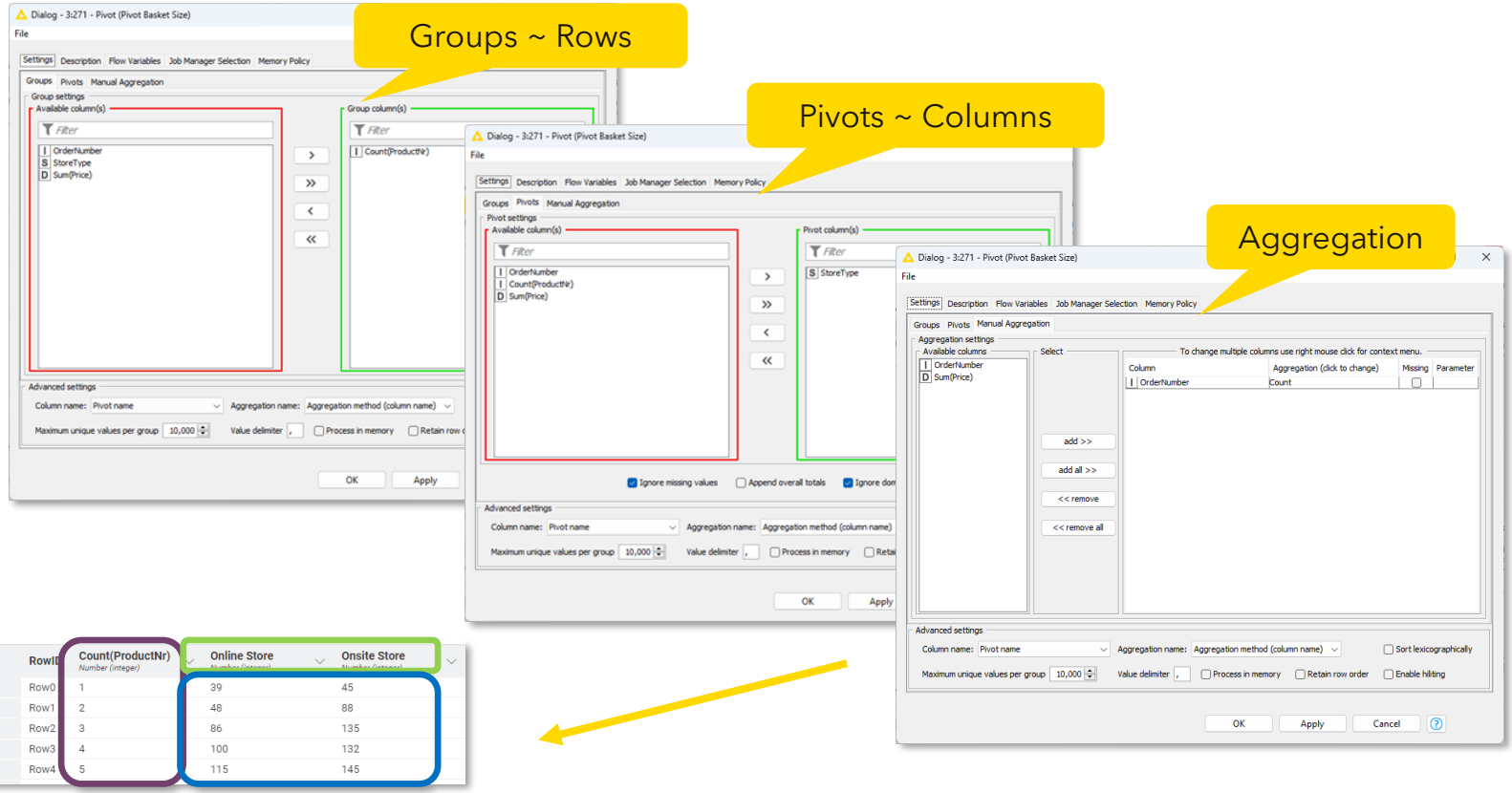
Pivot

- Perform pivoting on selected columns for grouping and pivoting
- Values of group columns become unique rows
- Values of the pivot columns become unique columns for each set of column combination together with each aggregation
- Many aggregation methods are provided (similar to GroupBy)



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Pivot



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Unpivoting

ID	City	Product 1	Product 2
234	Berlin	Pear	Apple
235	London	Nuts	Pear
236	Boston	Rice	Grapes
237	Paris	Pasta	Apple

ID	City	Product
234	Berlin	Pear
234	Berlin	Apple
235	London	Nuts
235	London	Pear
236	Boston	Rice
236	Boston	Grapes
237	Paris	Pasta
237	Paris	Apple

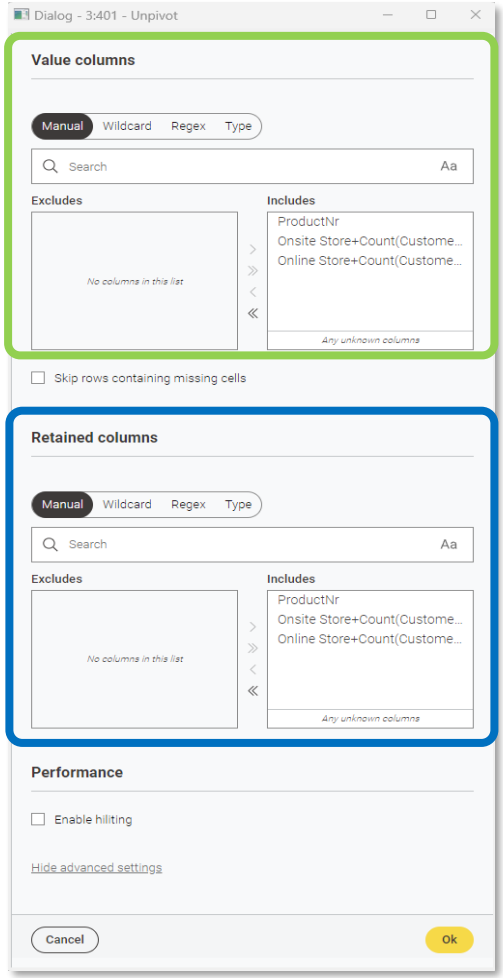
Value Column
Retained Columns

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Unpivot

- Rotate the value columns to rows
- Duplicate the remaining columns
- Append them to each corresponding row

Value Column
Retained Columns



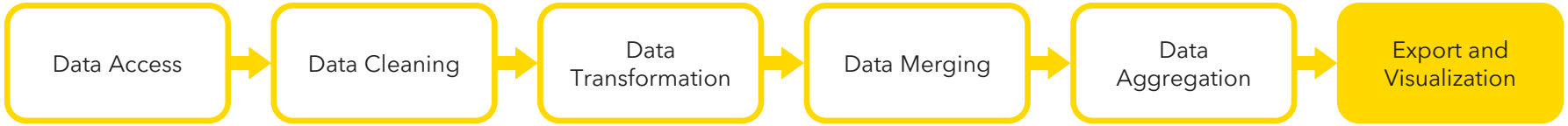
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Data Export

Data Export

- Your data pipeline does not have to end in KNIME Analytics Platform
- Export the data to store it or continue the computation somewhere else
- Different file formats (same as reader nodes)
  - CSV
  - Excel
  - Table
  - DB
  - …

Use Case
Pauline is happy that the data is ready to be exported. Her team needs the results as an Excel file for further processing.

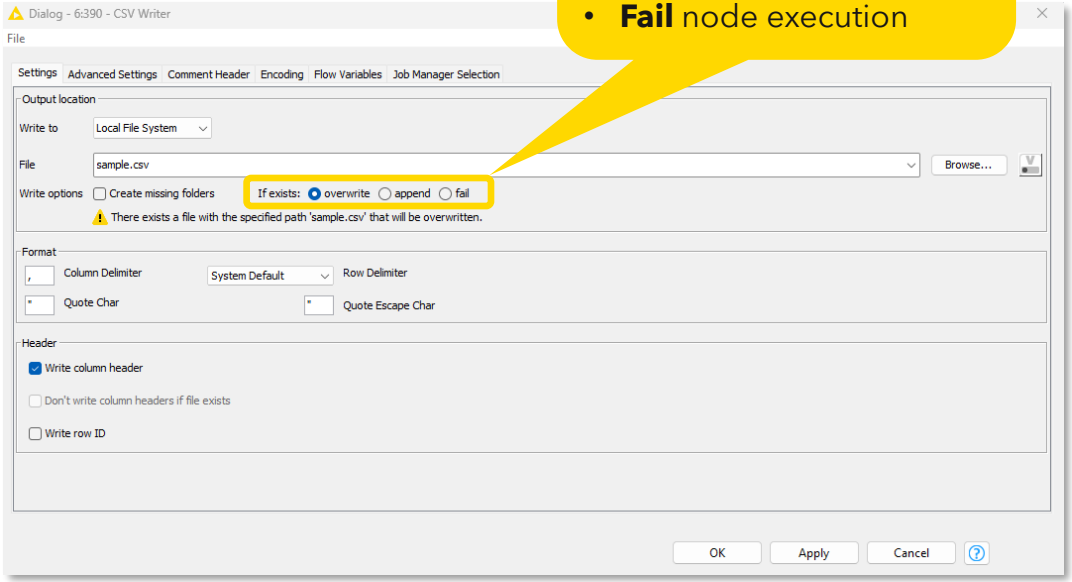


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

CSV Writer

- Write table to a .csv file
- Select delimiter and quote chars

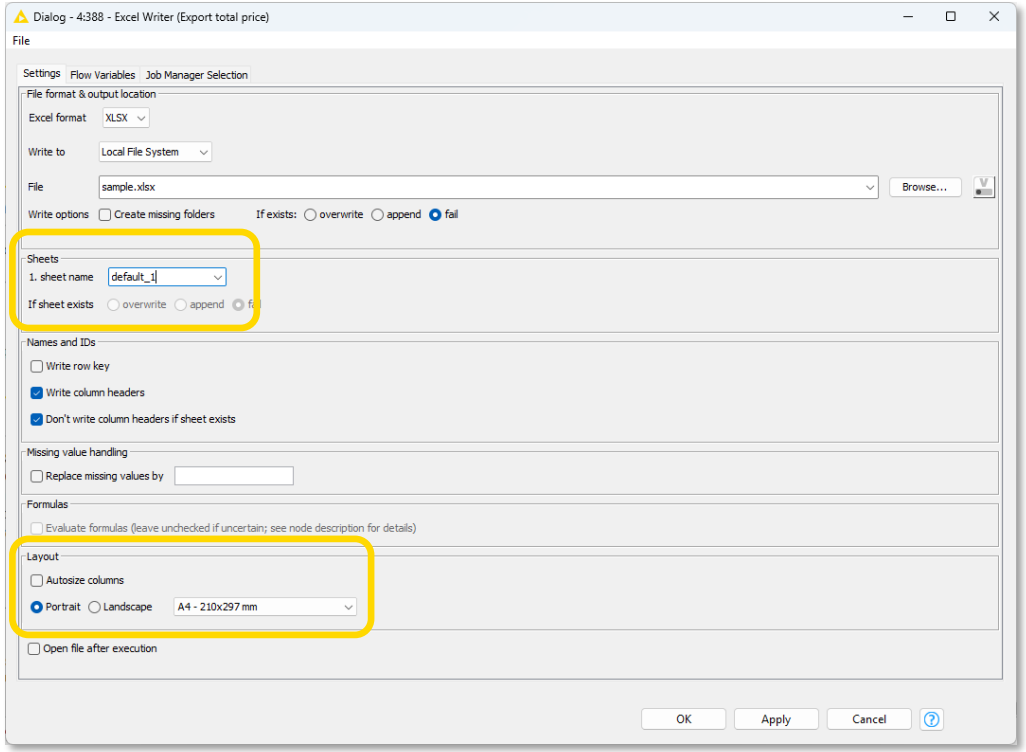
If file exists:
- **Overwrite** with new file
- **Append** rows to the exsisting file
- **Fail** node execution



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Excel Writer

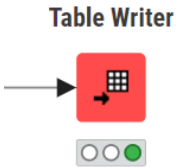
- Write table to a .xlsx or .xls file
- Select sheet and layout

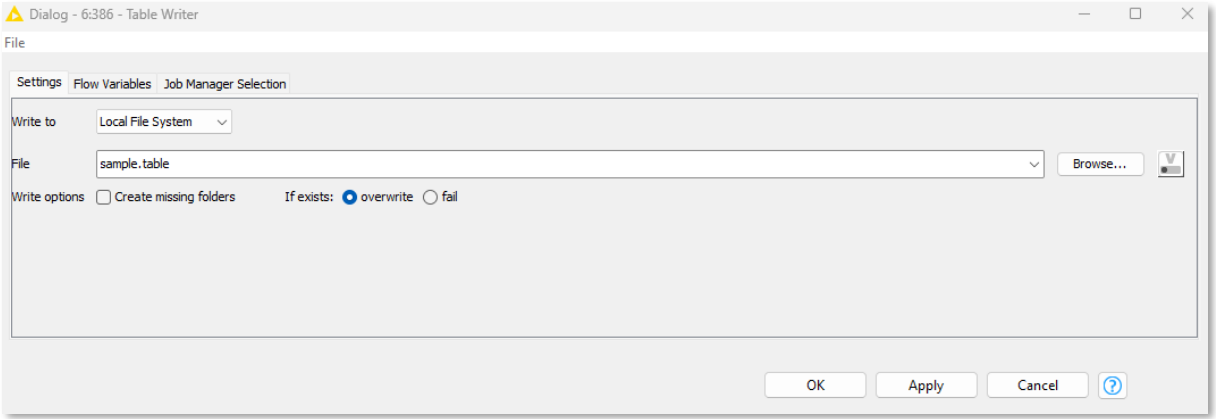


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Table Writer

- Write table to a KNIME format .table file
- Minimal configuration



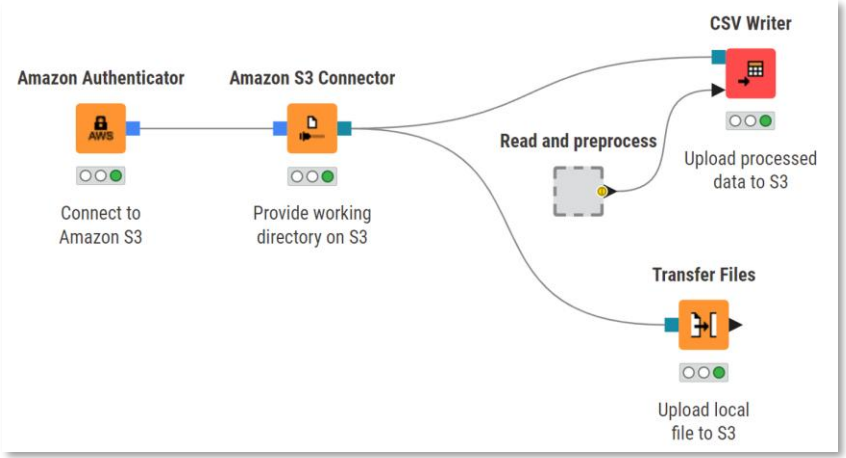


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Write Files to a Remote File System

- Upload data to remote file systems
  - Write processed data directly with a writer node
  - Upload local files with the Transfer Files node

- Supported file systems
  - Microsoft Azure
  - Google
  - Amazon
  - Databricks
  - Big Data file systems (HDFS, httpFS, …)
  - On-premise (e.g., SSH, FTP, …)

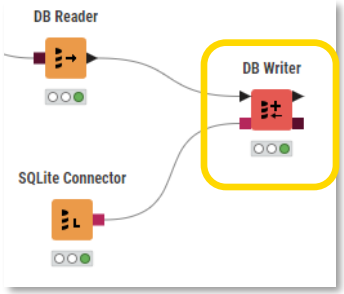


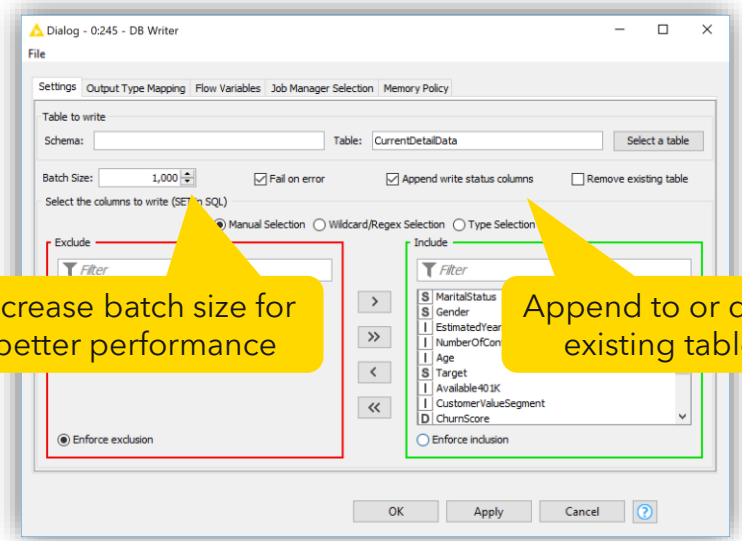
KNIME File Handling Guide

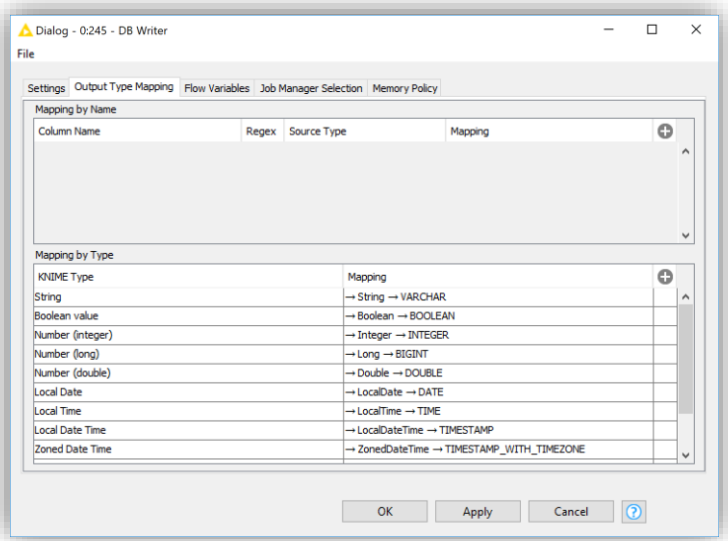
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

DB Writer

- Writes data from a KNIME data table directly into a database table





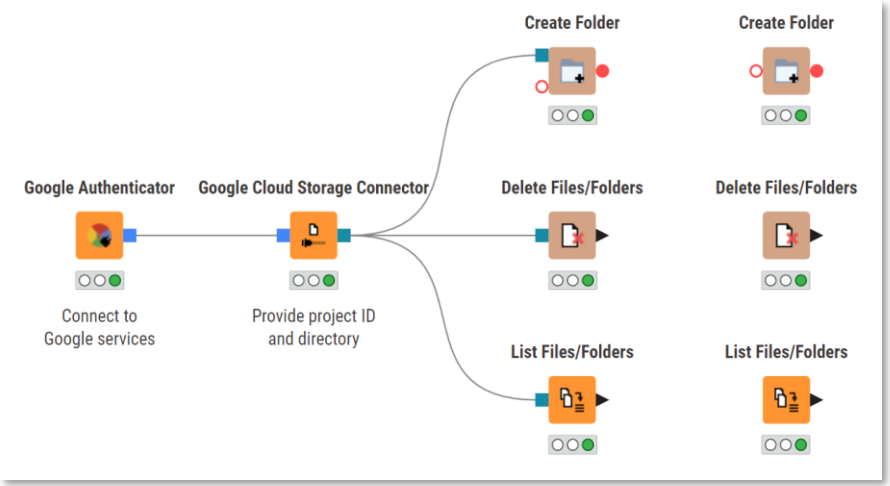


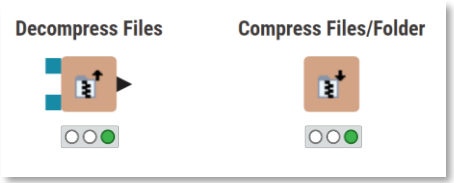
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Other Utility Nodes

Can be used with local and remote file systems

- Create a folder
- Delete files or folders
- List all files in a folder
- Compress/Decompress files





These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 3: Summary

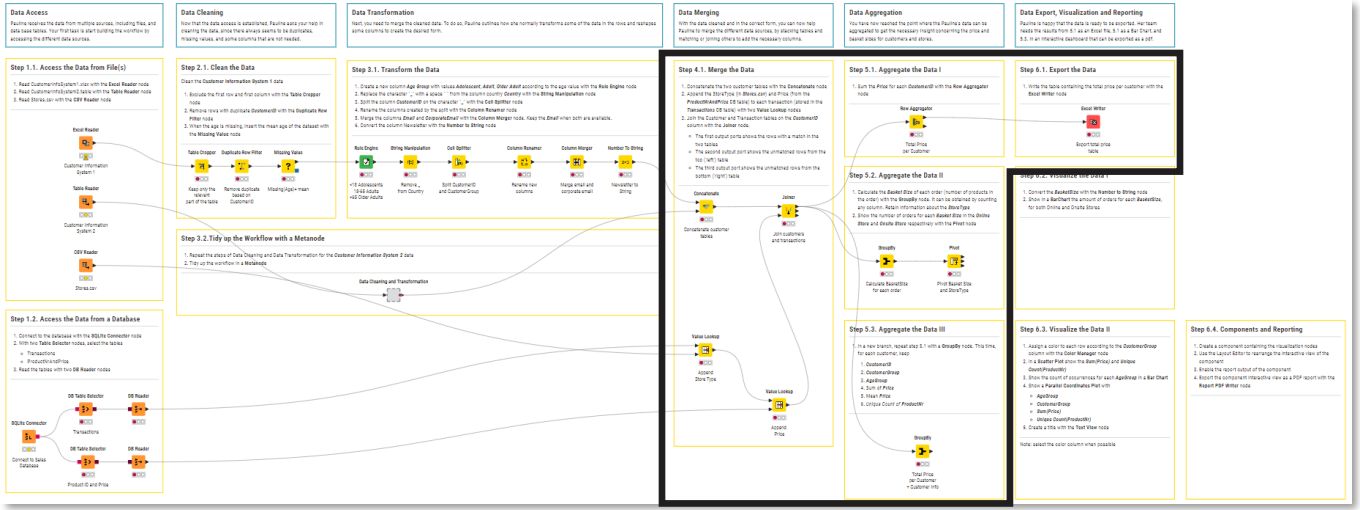
- Now you are able to:
  1. Concatenate multiple tables
  2. Join tables in multiple ways
  3. Aggregate table values
  4. Write data to various file formats

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Exercises

Exercises Session 3

- **Step 4.1** – Concatenate and join customer, products and transactions
- **Step 5.1-3** - Aggregate the data per customer and order
- **Step 6.1** – Write the data to an Excel file



- Compare your results with the *Customer Transactions Analysis - Solution* workflow

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Session 4
Data Visualization and Reporting

Learning Objectives

1. Visualize data with various plots
2. Create a dashboard with customized layout
3. Produce a custom report

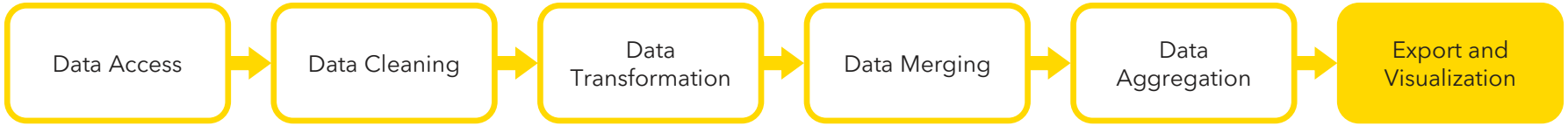
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Data Visualization

Data Visualization

- Transform data tables to communicate their results visually
- Help data understanding and get more insights
- Enable identifying patterns are trends that are hidden in tabular format

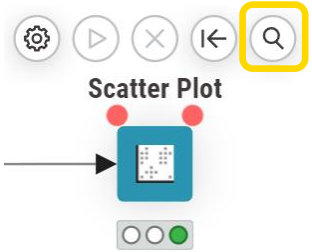
Use Case
Pauline is happy that the data is processed and wants to produce different plots to get more insights.



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Scatter Plot

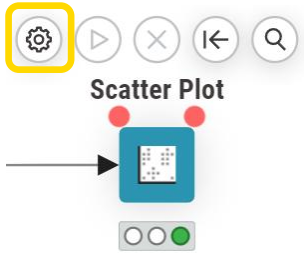
- Plot different columns on X and Y axes
- Display data including color information
- Produce an interactive view and an image
- Select data points and publishe selection to other views

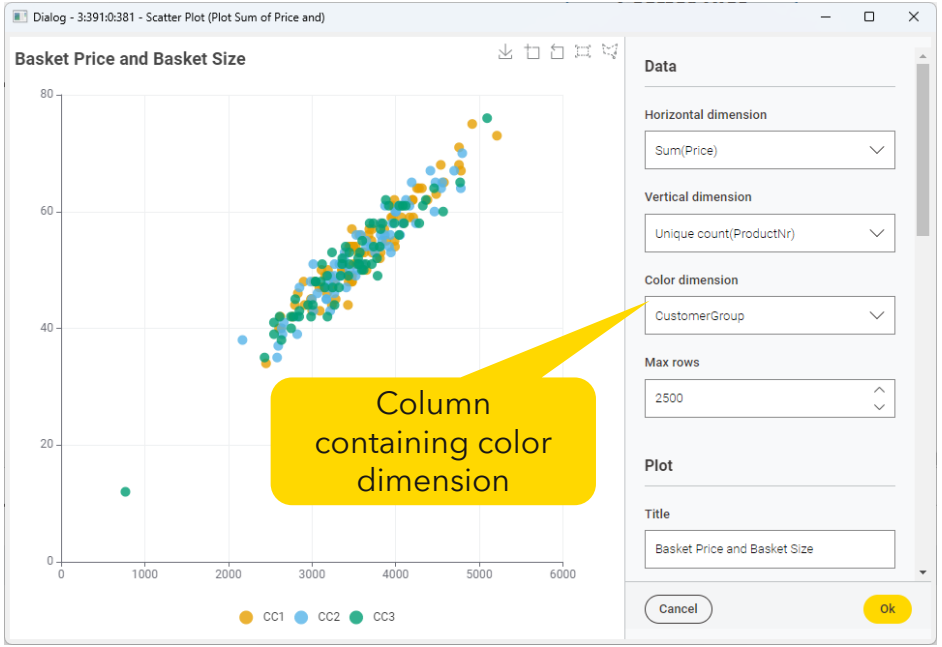


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Scatter Plot

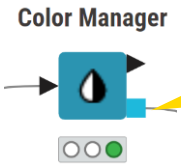
- Configure and see real-time plot preview



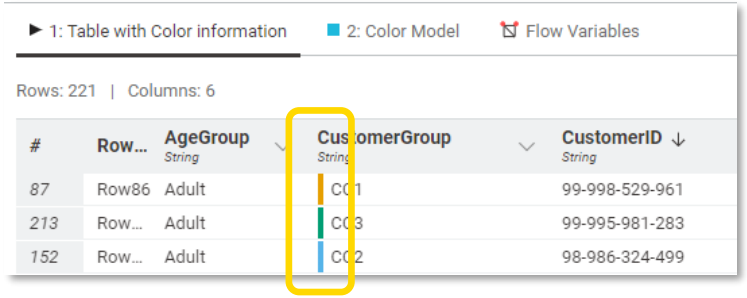


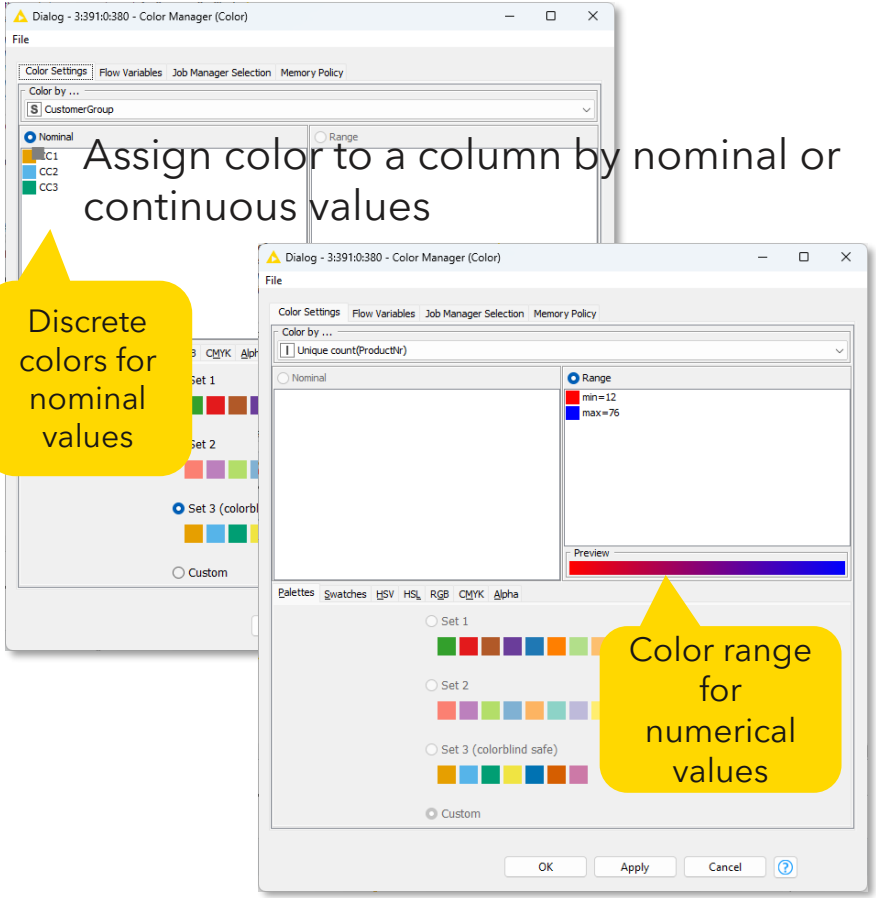
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Color Manager



The Color Model can be reused by the Color Appender node





Assign color to a column by nominal or continuous values

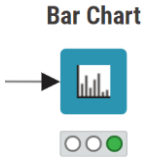
Discrete colors for nominal values

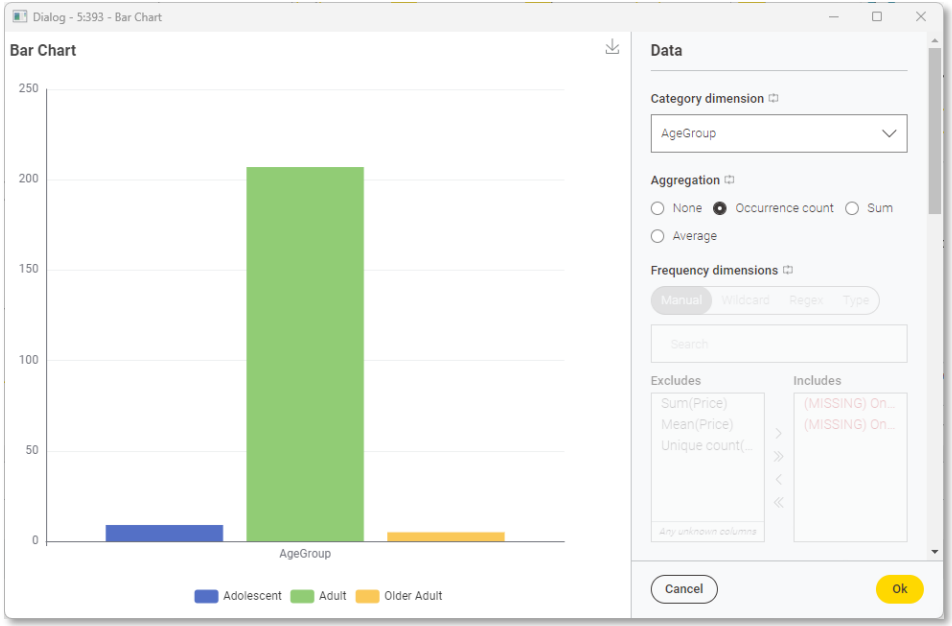
Color range for numerical values

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Bar Chart

- Show numerical values across categories
- Vertical or horizontal bars
- Bars can be grouped or stacked

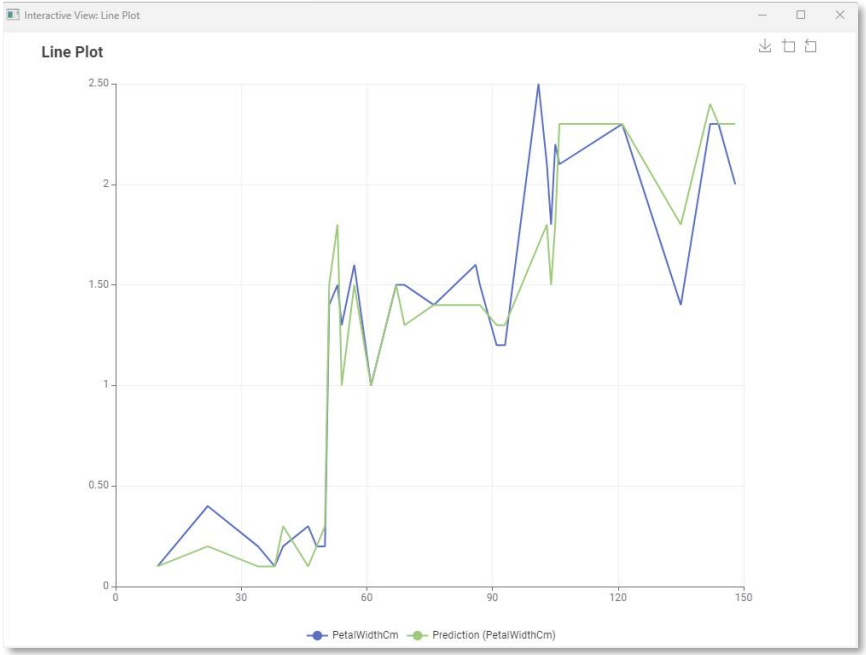




These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Line Plot

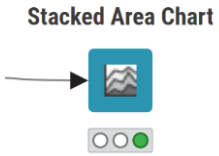
- Plot sequence of values, e.g. over time
- Useful to identify trends, also between groups

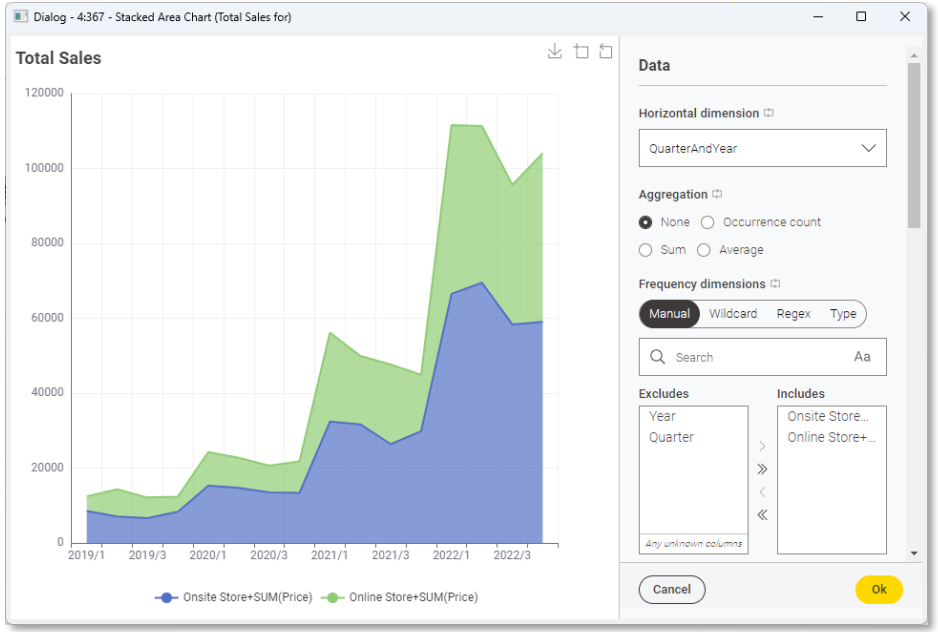


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Stacked Area Chart

- Visualize numerical values from multiple columns as stacked areas
- Great for plotting distributions over time



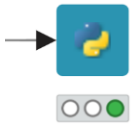


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

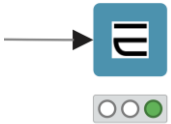
Script-based View Nodes

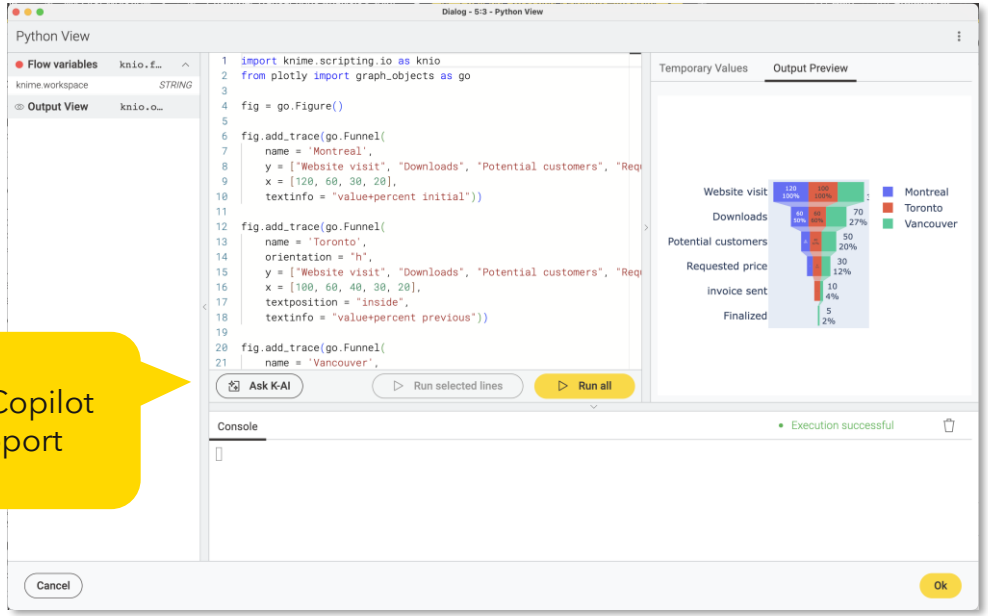
- **Python**, **R** and **JavaScript** support
- Use your favorite libraries e.g. **ECharts** or ggplot2

Python View



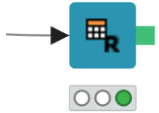
Generic ECharts View





K-AI Copilot support

R View (Table)



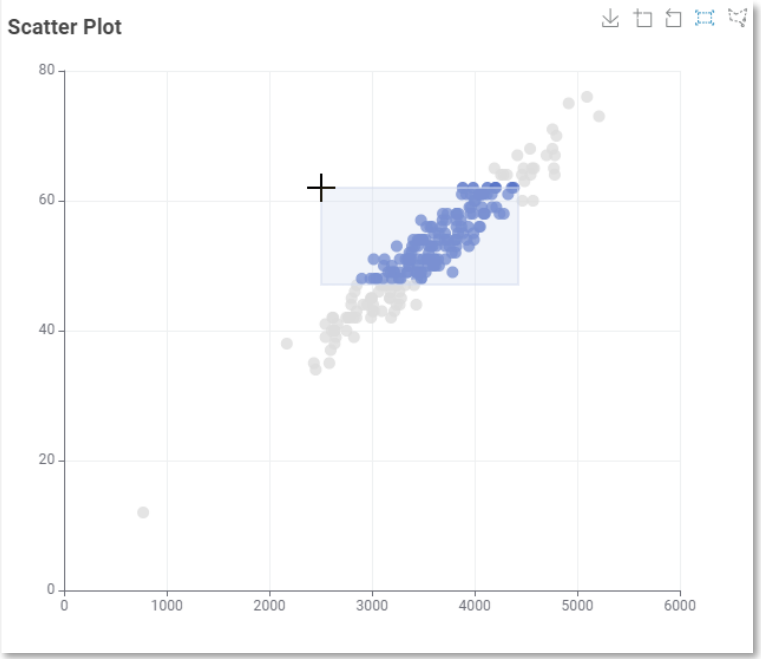
Generic JavaScript View (JavaScript)

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Components

Selection in Views

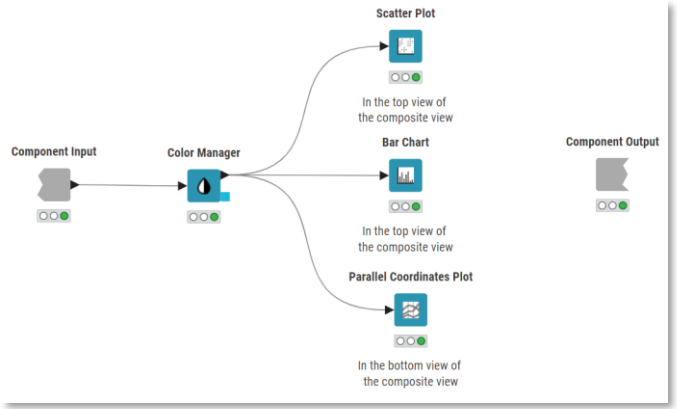
- Interactivity allows you to select data points in single views
  - Selection is propagated to other views
  - Highlight selected rows or filter them

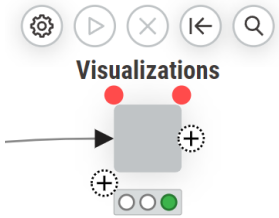


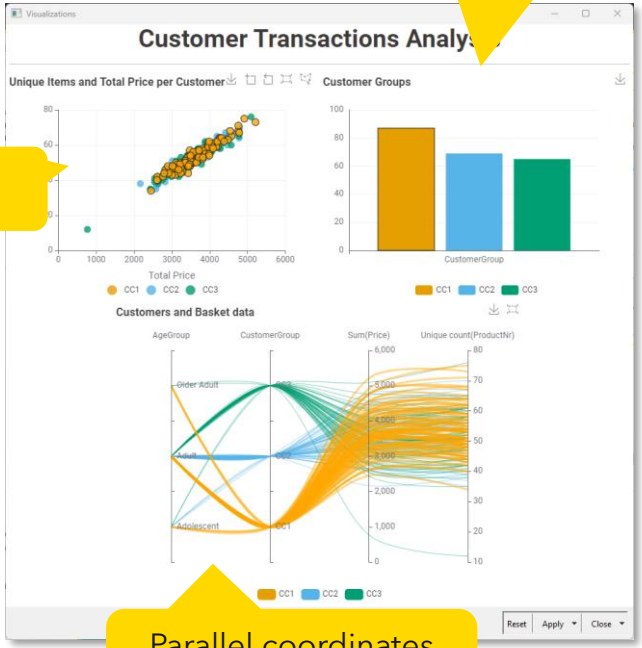
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Components Composite View

- Multiple view nodes can be combined in Components
- Selections are propagated to all other views
- Building blocks of *Data apps*



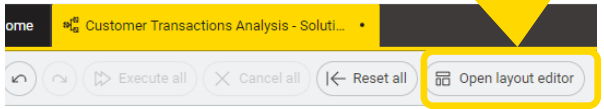




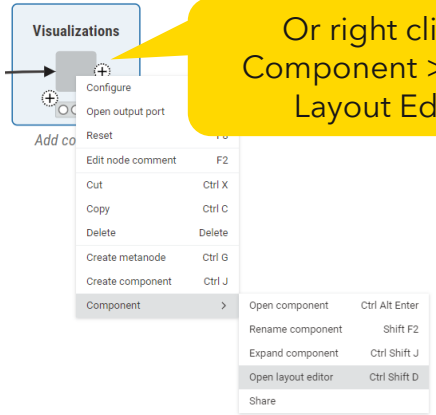
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Composite View Layout

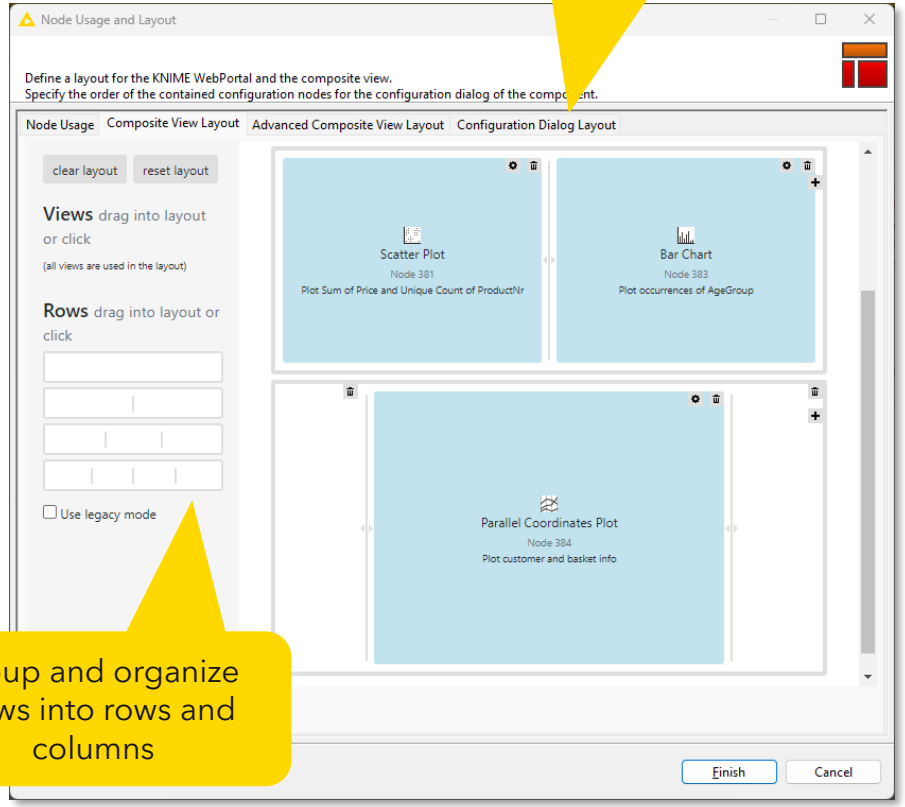
Open layout editor from inside component



Or right click > Component > Open Layout Editor



Rearrange single views



Group and organize views into rows and columns

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

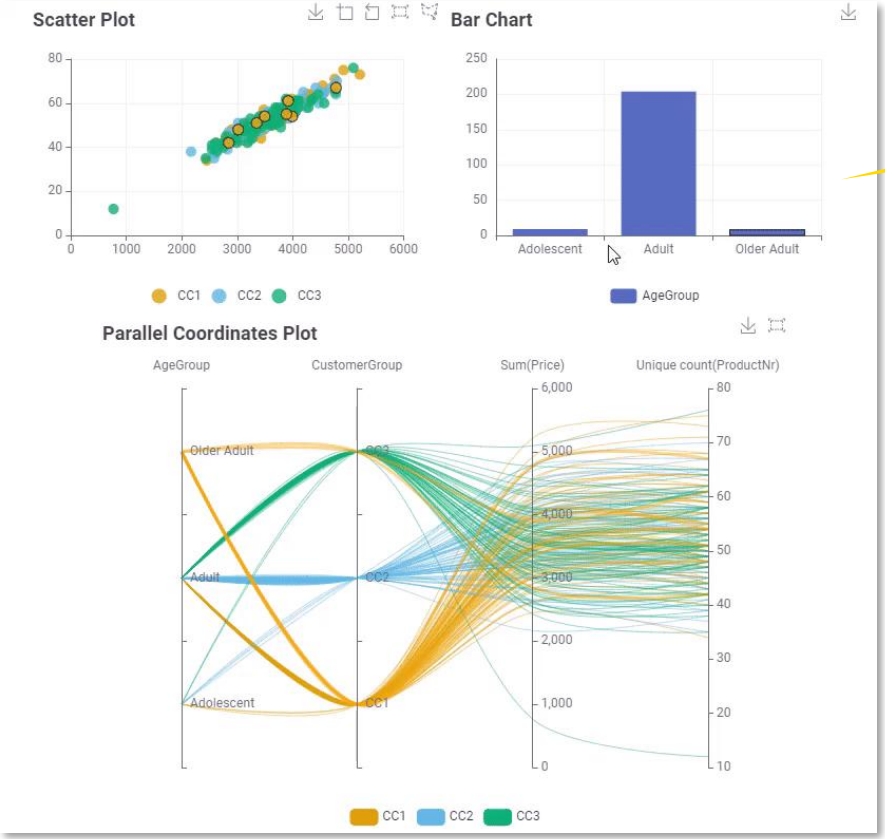
Interactivity across Charts

Publish = propagate selection to other charts

Subscribe to = receive selection from other charts

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Interactivity across Charts



Selection is propagated to other views

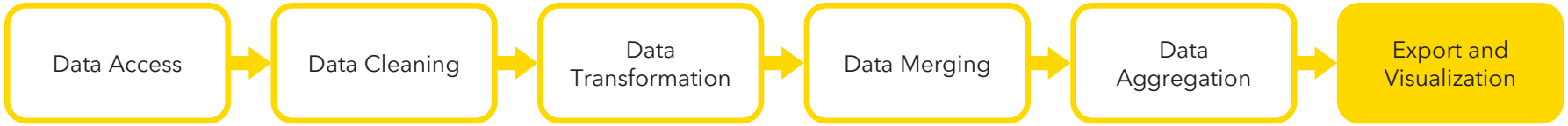
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Reporting

Reporting

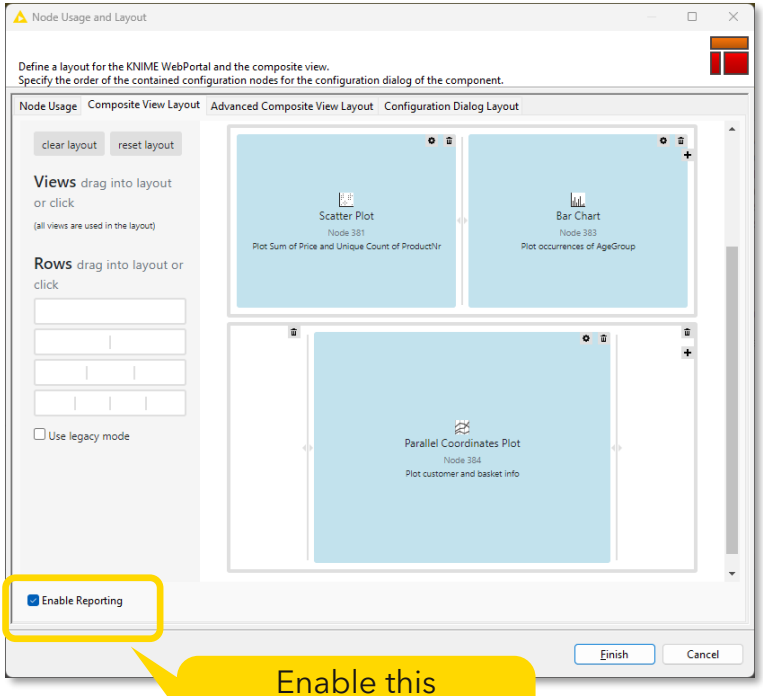
- Easily communicate and share insights
- Export component composite view to PDF

Use Case
Pauline wants to present the insights to her team in the form of a report.



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

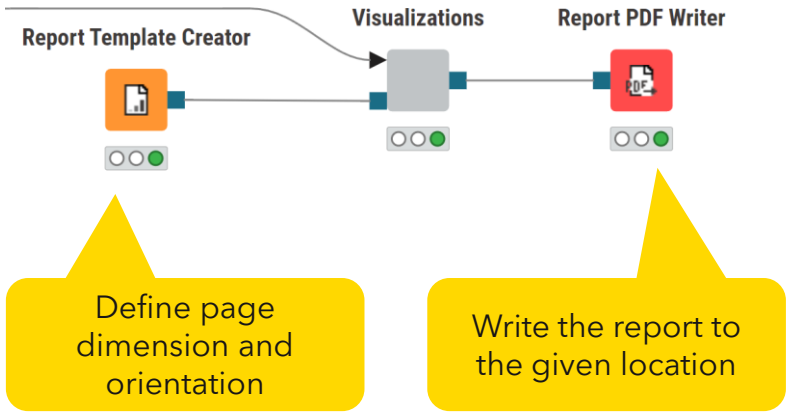
Enable Reporting from Component

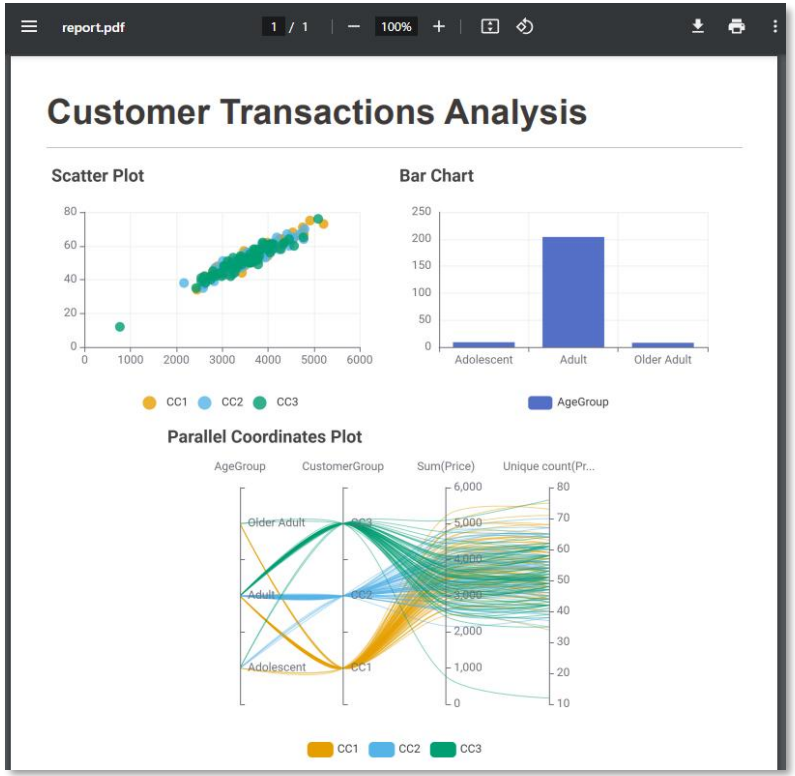


Enable this component to create a report

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Create a PDF Report



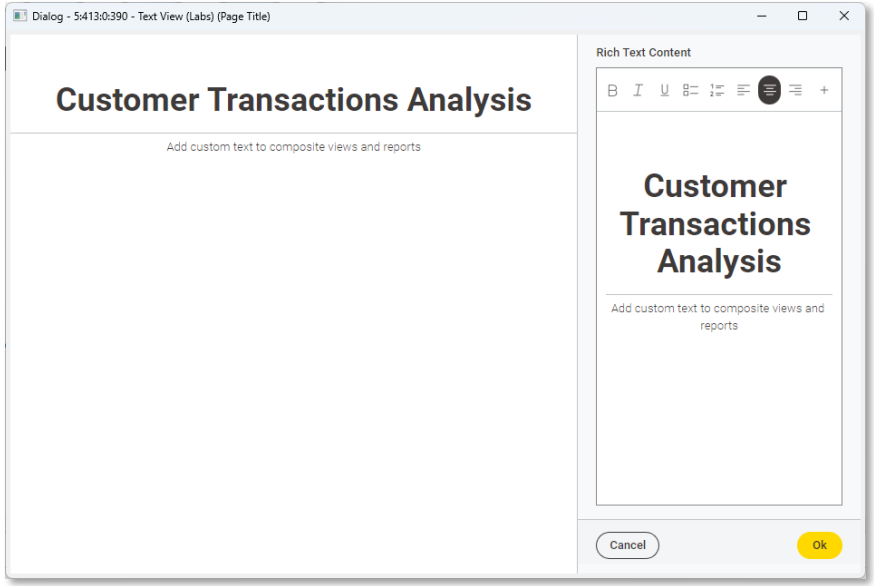


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Text View

- Add custom text to composite view and reports
- Editing option similar to workflow annotations
- Creates a view that can be moved in the Composite View Layout

Text View



These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

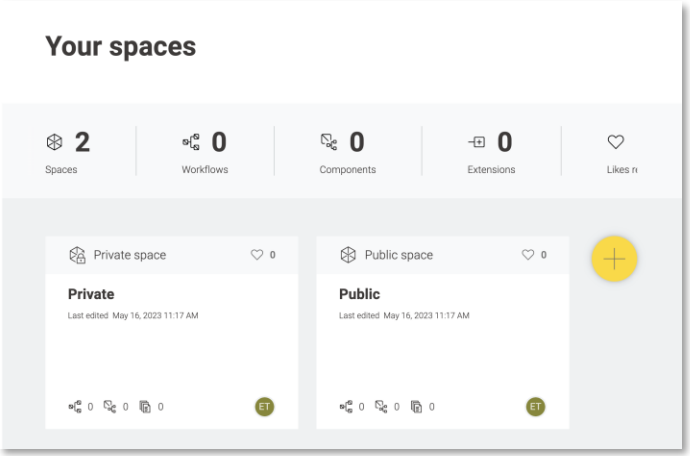
Session 4: Summary

- Now you are able to:
  1. Visualize data with various plots
  2. Create a dashboard with customized layout
  3. Produce a custom report

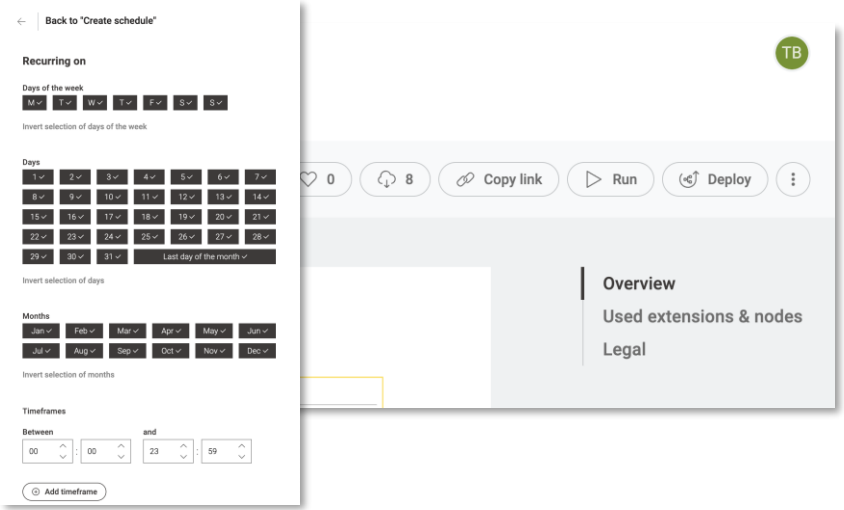
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Next: KNIME Community Hub

- Personal plan (free)
  - Browse public workflows
  - Up to 10GB to store your workflows
  - Share with the KNIME Community



- Team Plan (paid subscription)
  - Collaborate with your team in private spaces
  - Run workflows remotely
  - Schedule workflow execution

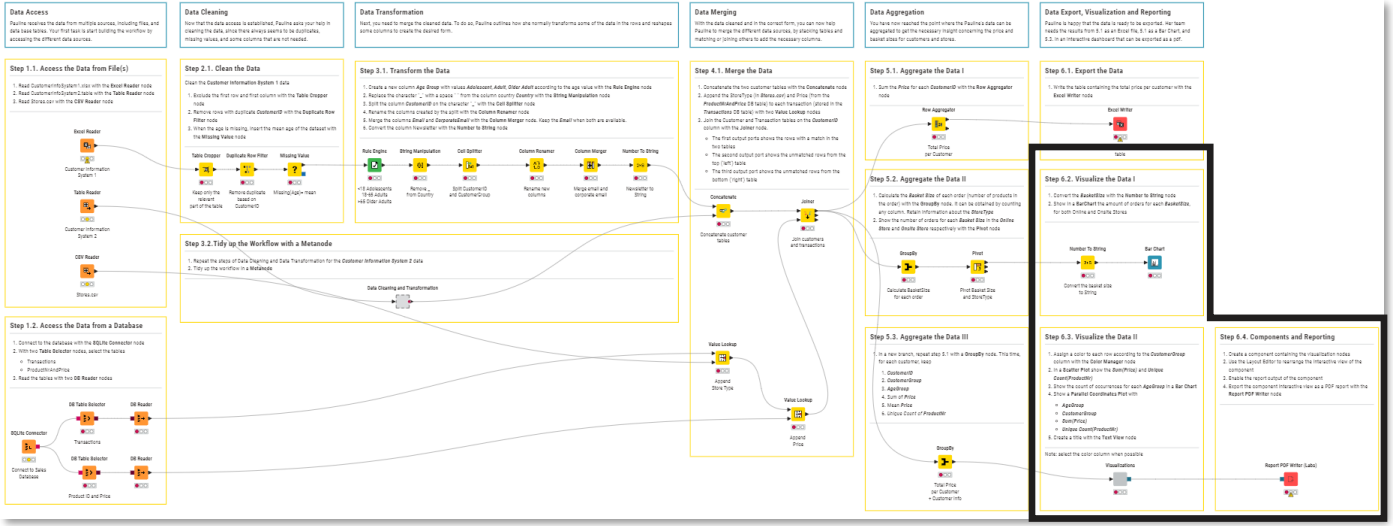


These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Exercises

Exercises Session 4

- **Step 6.2** – Visualize in a single Bar Chart
- **Step 6.3** – Visualize in multiple views
- **Step 6.4** – Create an interactive dashboard and export to PDF report



- Compare your results with the *Customer Transactions Analysis - Solution* workflow

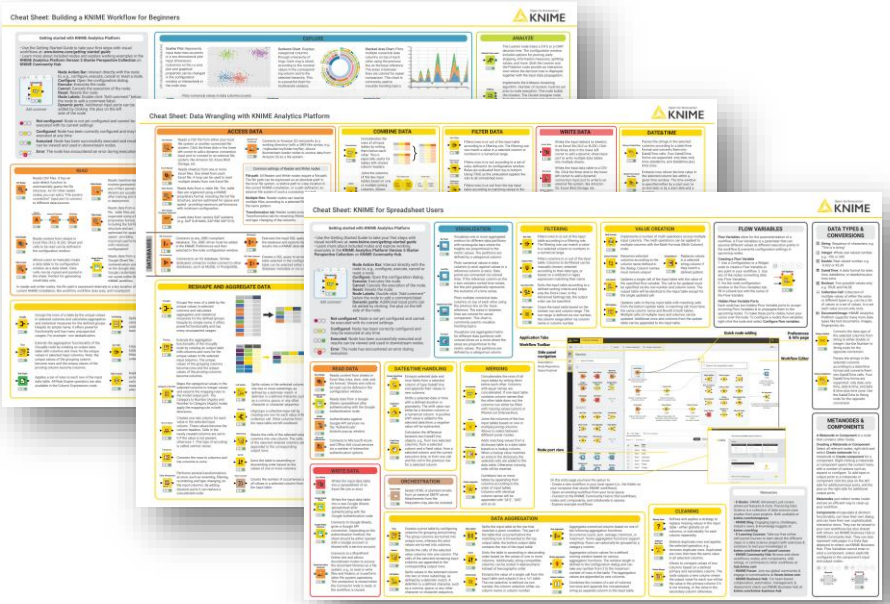
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Course Objectives

- Now you are able to
  1. Outline the uses of KNIME Analytics Platform for data analysis and data science
  2. Describe the advantages of using visual programming for data science
  3. Read data from a file and database
  4. Clean and transform data
  5. Merge, aggregate, and export data
  6. Visualize and report results
  7. Take the L1 - Certification exam

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Press - Cheat Sheets



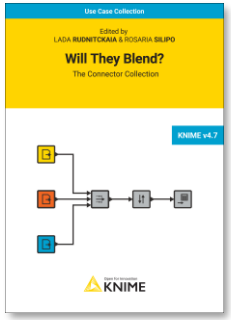
- Many cheat sheets available for:
  - Beginners
  - Spreadsheet Users
  - Machine Learning
  - Data Wrangling
  - Orchestration
  - and many more…

All cheat sheets are available for download at:

knime.com/cheat-sheets

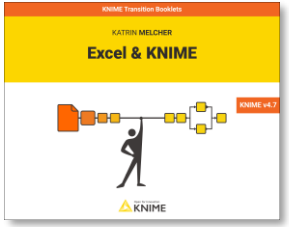
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Press - Books



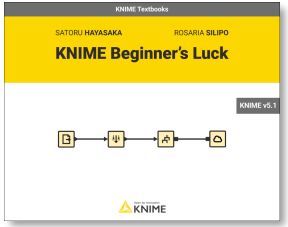
Use Case Collections

Collections of classic and more innovative use cases around specific topics



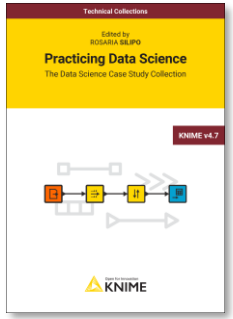
Transition Booklets

Offer an easy onboarding into KNIME from other tools



Textbooks

Well structured schoolbooks with plenty examples



Technical Collections

Collections of specific technical topics to keep you up to date

All KNIME Press books are available for download at:

knime.com/knimepress

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Beginner's Luck Book

The **KNIME Beginner's Luck** is the ideal starting point for those who want to begin working with KNIME

All KNIME Press books are available for download at:

knime.com/knimepress

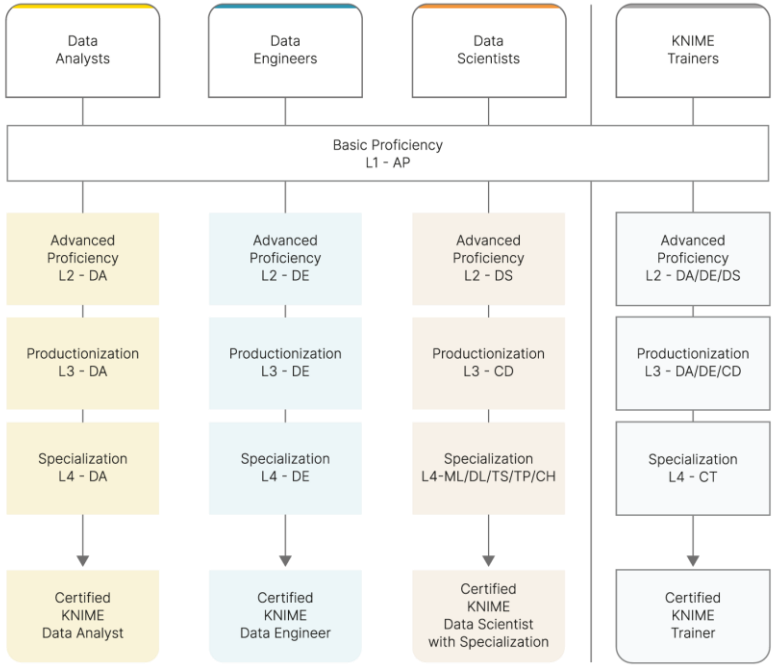
These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Learning Paths

- Courses from level **L1** to level **L4**
- Various professional profiles

- **Self-paced courses**
  - Videos and exercises at your own pace and for free

- **Instructor led courses**
  - Scheduled sessions and guided exercises in paid courses



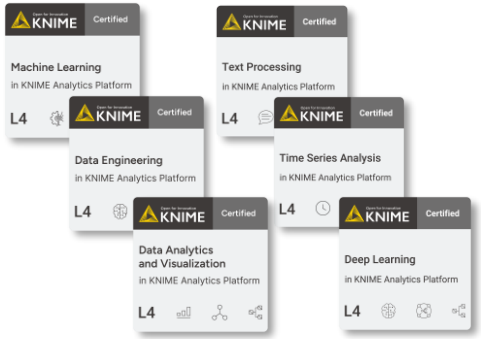
Find the **next course** at:

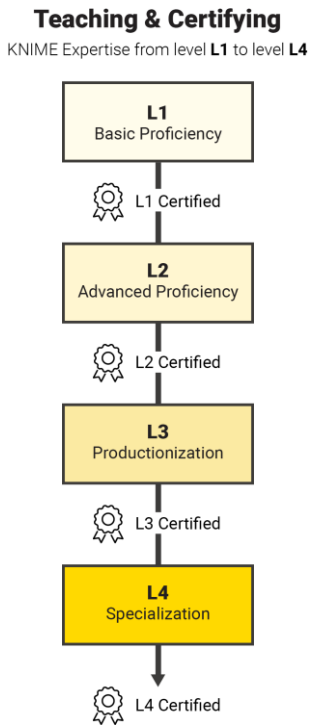
knime.com/learning

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Certification Exams

- From level **L1** to level **L4**
- Online examination
- Measure your expertise with KNIME software and data skills
- Get and share your badge!





Get **certified** at:

knime.com/certification-program

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Just KNIME It! Challenges

- The best way to keep on learning
- Weekly challenges to test your knowledge
- Easy, medium and hard challenges for any level
- Discuss the solution with the community
- Post your solution and climb the Leaderboard

Find the **challenges** at:

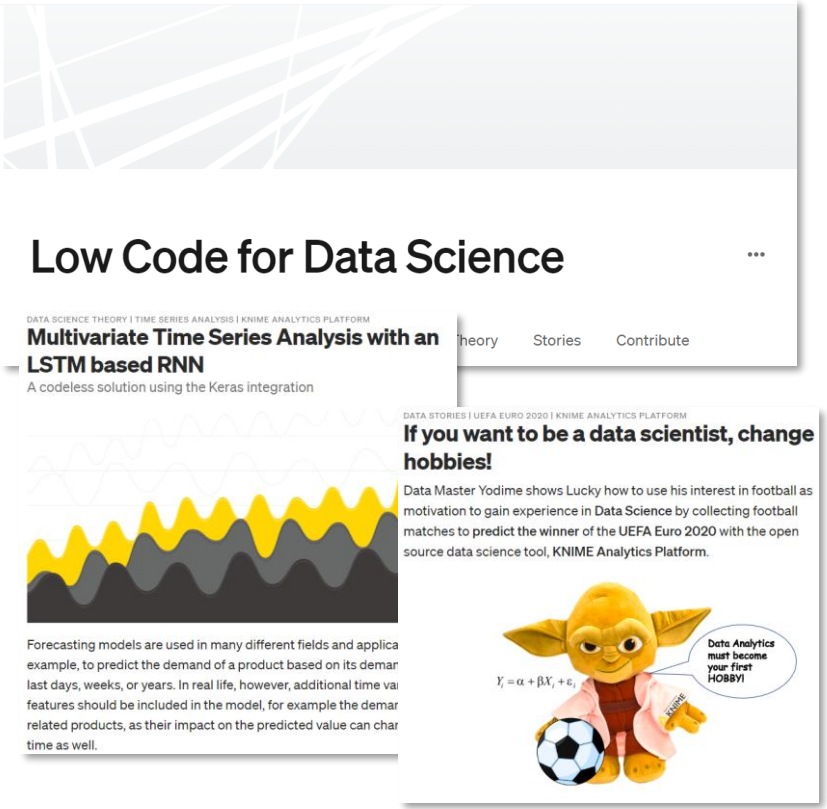
knime.com/just-knime-it

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Community Journal

- Daily content on
  - data stories
  - data science theory
  - getting started with KNIME
  - ...and more
- For the community by the community

- Share your data story with the community
  - Contributions are always welcome!



Low Code for Data Science community journal on Medium

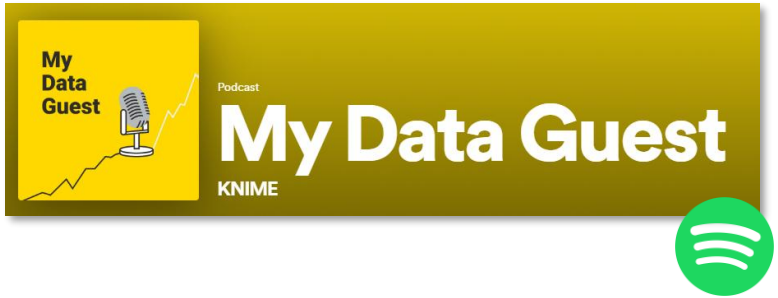
medium.com/low-code-for-advanced-data-science

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME TV and My Data Guest

- KNIME TV
  - Webinars
  - Panel discussions
  - Educational videos
  - Data Science topics
  - Tutorials

- My Data Guest - Podcast
  - Interviews with data professionals in the KNIME community



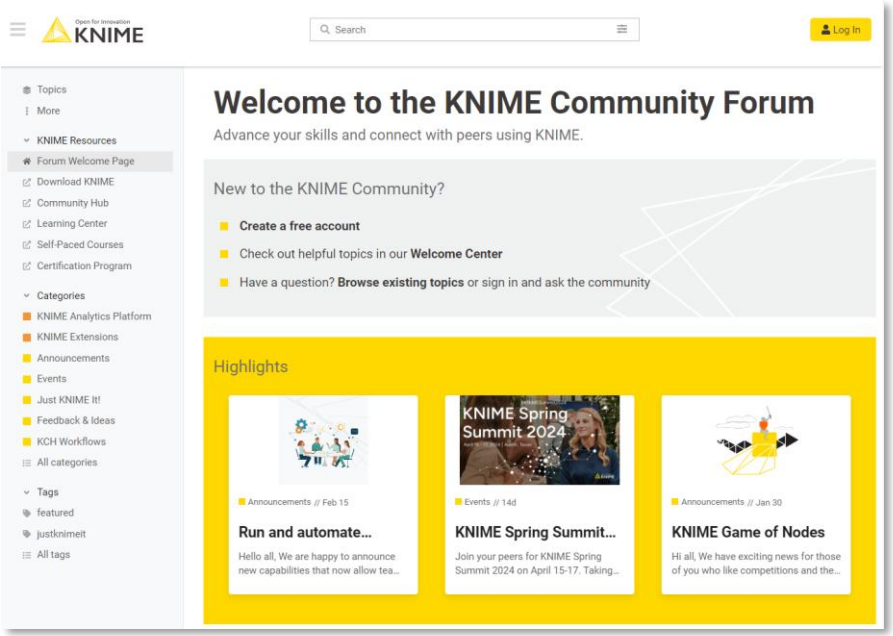
Find **KNIME multimedia content** at:

https://www.youtube.com/@KNIMETV and **Spotify**

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

KNIME Forum

- Round-the-clock community support
- Specialized categories
  - KNIME Analytics Platform
  - KNIME Extensions
  - KNIME Business Hub
  - …
- Multilingual support groups
  - French
  - Italian
  - Spanish
  - …



Visit the **KNIME Forum** at:

https://forum.knime.com/

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

What do you think about us?

- We provide most of our software and learning material for free
- Do you like our stuff? Let everybody know!

May 10, 2023

"No code, easy to use and easy to innovate"

Mar 10, 2023

"Best platform to access, blend, analyze, and visualize data, without any coding."

Mar 07, 2023

"Easy to Learn for non-code Users"

Mar 07, 2023

"Very powerful software for analytics and provide instant support."

Write a review on G2:

www.g2.com/products/knime-analytics-platform/take_survey

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Stay Connected with KNIME

Blog:
knime.com/blog

KNIME Self-Paced Courses:
knime.com/knime-self-paced-courses

Forum:
forum.knime.com

Email:
education@knime.com

KNIME Hub:
hub.knime.com

Medium Journal:
medium.com/low-code-for-advanced-data-science

Follow us on social media:

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0

Thank You!

These slides are a derivative of *KNIME Course Material* of KNIME AG used under CC BY 4.0