Data Analytics Foundations

Online Workshop 1

Course structure

Five modules over 8 weeks:

1. Introduction to data analytics
2. Know your data
3. Data pre-processing
4. Data exploration and visualisation
5. Clustering

Assessment

Assignment: Data exploration and preparation

- Explore a real dataset
- Apply ideas and techniques covered in the course
- Identify interesting patterns, problems with the data, scope for further exploration...
- Due two weeks after end of teaching (**Sunday, May 1, 2022**)

KNIME

- KNIME installation
- KNIME exploration

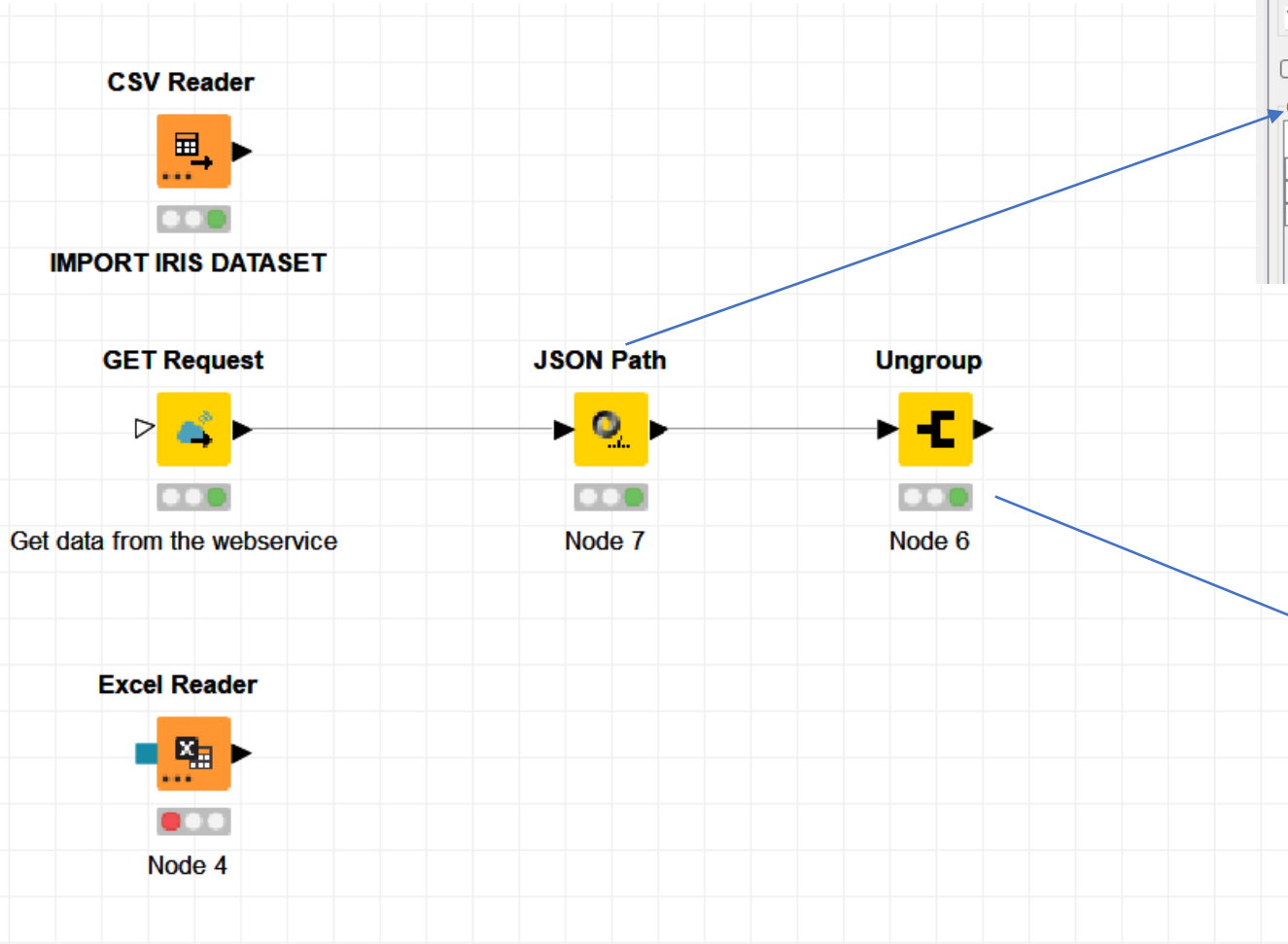# KNIME Exercise

Read data from local drive
- CSV and Excel datasets from local drive using "Reader" KNIME node

Read data from web service
- using "Get Request" node
- E.g. URL: https://api.btcturk.com/api/v2/ticker
- Add JSON Path and Ungroup nodes to extract the results

KNIME Exercise

Workflow 1

CSV Reader

IMPORT IRIS DATASET

GET Request

Get data from the webservice

JSON Path

Node 7

Ungroup

Node 6

Excel Reader

Node 4

Settings | Flow Variables | Memory Policy

Input

body

☐ Remove source column

Outputs

Output column	JSONPath	List	Paths
S pairs	$['data'][*]['pair']	☑	☐
D averages	$['data'][*]['average']	☑	☐
S numeratorSymbols	$['data'][*]['numeratorSymbol']	☑	☐

---

ion ○ Wildcard/Regex Selection

Include

🔽 Filter

[...] pairs
[...] averages
[...] numeratorSymbols

○ Enforce inclusion