

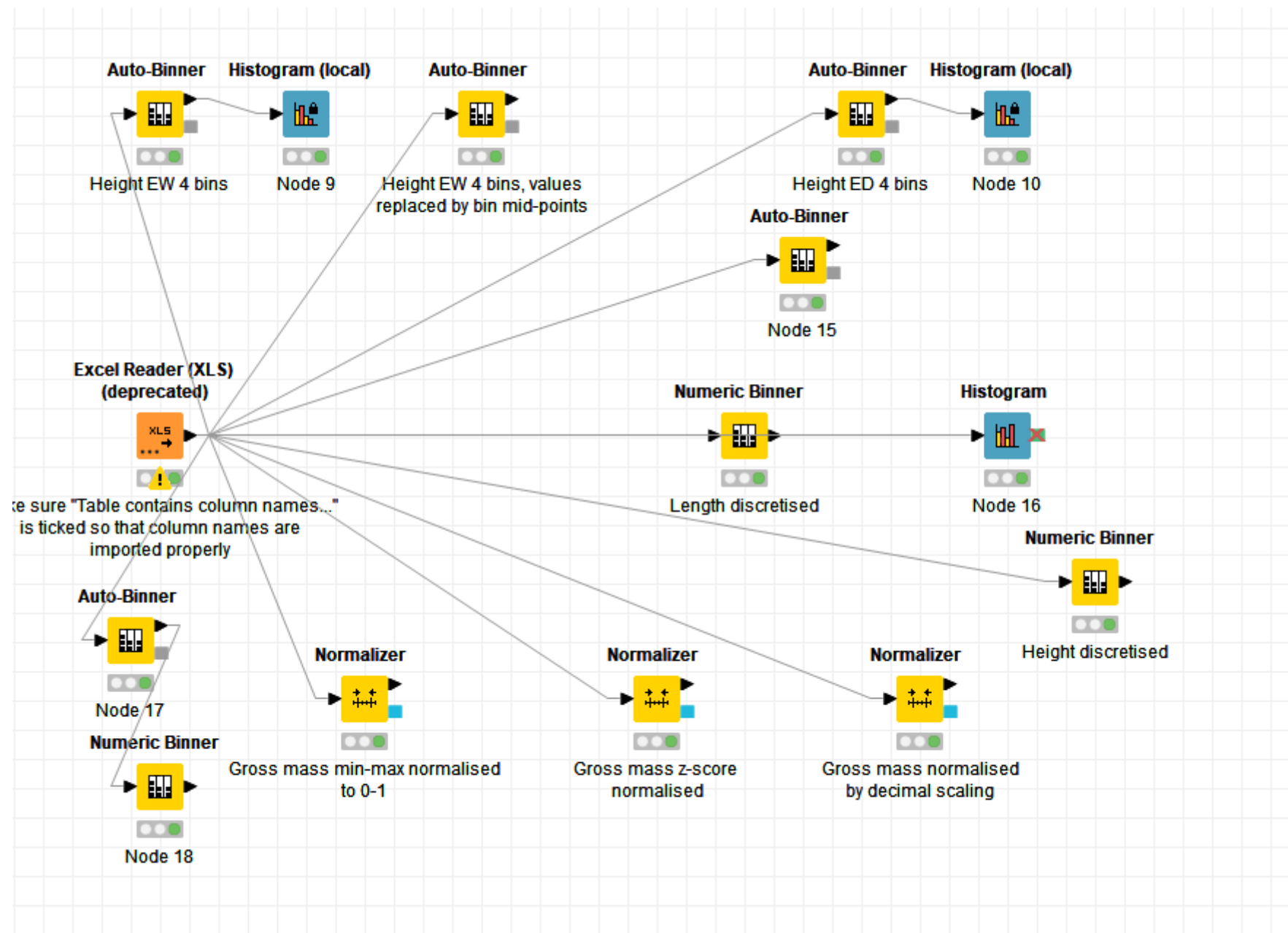
Data Analytics Foundations

Online Workshop 5
Data Preprocessing (2)

Questions from Module 3?

Preprocessing in KNIME

- Now see if you can work out how to do some of the same preprocessing tasks in KNIME, again using the [abalone](#) dataset:
 - Binning: equi-width and equi-depth
 - Min-max normalisation to $[0,1]$
 - Min-max normalisation to something other than $[0,1]$
 - z-score normalisation



Preprocessing in KNIME

- KNIME doesn't have nodes to smooth the data by means or bin boundaries
 - Think about using some KNIME nodes to smooth the data

Q&A

1. When might you want to bin the data? When might you NOT want to bin the data?
2. Why would you apply binarisation? What are the downsides of binarising an attribute?