Data Partitioning and Modeling

The data was partitioned into train and test datasets.
The **train data set** was used to create the decision tree model.
The trained model was then applied to **the test dataset.**
This is important because a train model needs to be **evaluated/validated** using **data it has not seen before or touched during training phase.** Such data comes from the test set

When partitioning the data using sampling, it is important to set the random seed because this ensures that the result is **reproduce-able after multiple iterations** of the KNIME workflow and student's result **matches** the model answer

A screenshot of the resulting decision tree can be seen below: