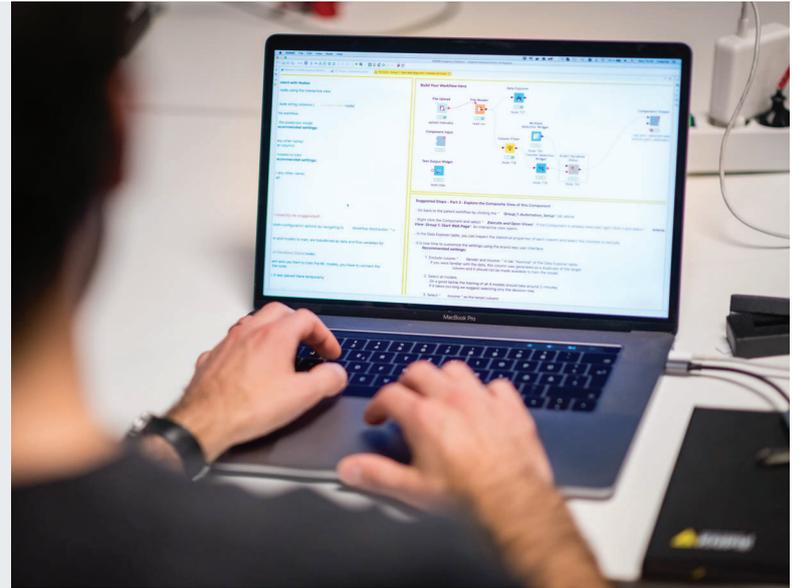




Automated Workflow Testing and Validation

Once a workflow enters production it's very important to regularly make sure that it's still doing what it's supposed to do. This is particularly true when something in the environment changes, e.g. installing a new version of KNIME or making other updates. Doing this manually for many workflows is time consuming and prone to errors. Automating these tests saves a lot of time, helps problems to be caught early, and avoids errors by systematically checking the workflow(s).



Automating the validation of workflows in KNIME requires a few simple steps. The key piece is provided by some infrastructure and a set of nodes that KNIME has developed to test KNIME Analytics Platform itself: the KNIME Testing Framework. There is a lot of functionality, but one of the most important nodes is the Table Difference Checker. Its job is simple - compare the contents of two KNIME tables and produce an error if they aren't identical. To test a workflow, read in the standard data with the appropriate reader, feed that data through the workflow, and use the Table Difference Checker to compare the generated results to a table containing the expected results - the so-called "golden table" (Fig. 1).

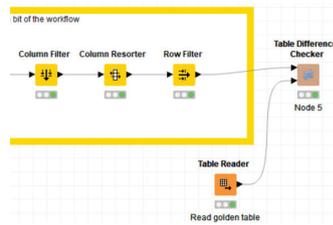


Fig. 1: The Table Difference Checker enables easy comparison of the results from executing a workflow with intended results.

To test a workflow with the KNIME Testing Framework, find the workflow in the KNIME Explorer that needs to be tested, right click, and select "Run as workflow test". This will run the workflow test automatically and display a report containing information about any nodes that failed (Fig. 2). Multiple production workflows can also be tested together. Just put them in a single workflow group in the KNIME Explorer before running the workflow tests. By putting workflow(s) on KNIME Server, it's possible to completely automate this process by scheduling the execution of these tests.

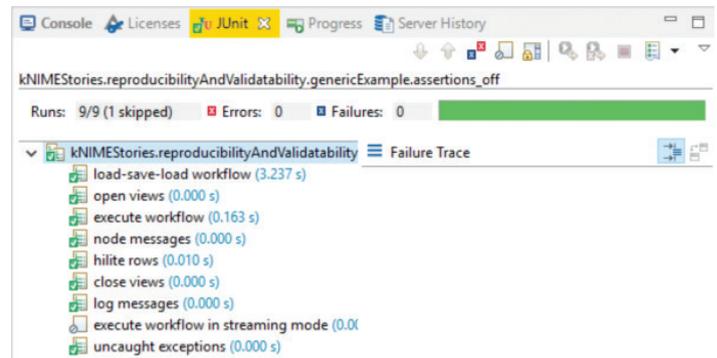


Fig. 2: The report generated when the tests for our workflow have been run. If any pieces of the workflow had failed, the green bar on top of the report would be red and there would be additional information available in the pane at the bottom.

The workflow (Fig. 3) represents the full testable version of a sample workflow. It's also possible to allow the workflow to be automatically tested without editing or requiring the workflow to be saved with loaded test data. This is contained in the blue boxes. The reference input data is loaded and compared with the golden table if the user does not make any changes to the configuration of the Table Reader. If everything is working as it should, the "Run tests if necessary" component will switch to the green, executed state. If there are problems, this node will report an error, which will be visible in the report if this workflow is run as a test.

To use this workflow, download it from the KNIME Hub and adjust the "interesting bit of the workflow" to a workflow snippet of interest.

Results:

By using the freely available KNIME Testing Framework, it's possible to:

- Be certain that workflows are still doing what they are supposed to do
- Avoid laborious, time-consuming, and manual testing
- Run automated tests for multiple workflows at once

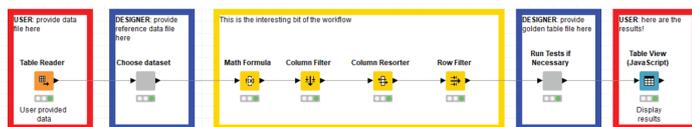


Fig. 3: The testable version of a sample workflow. Components in the blue boxes allow automated testing. Nodes in the red boxes indicate the end user interaction points. The workflow in the middle is exchangeable, depending on what needs to be tested.

The KNIME Testing Framework makes it possible to automatically ensure that workflows are doing what they are supposed to be doing. The Table Difference Checker allows users to easily compare current results to a golden table. It is easy to test multiple workflows at once, saving time and helping detect problems early. With KNIME Server, execution of these tests can be scheduled - allowing the process to be fully automated.

