DataStage has three processing stages that can join tables based on the values of key columns: Lookup, Join and Merge. In this post, we discuss when to choose which stage, the difference between these stages, and development references when we use those stages.

Use the Lookup stage when:

* Having a small reference dataset.
* Doing a range lookup.
* Validating a row (If there is no corresponding entry in a lookup table to the key’s values, you can output the row in the reject link).

Use the Join stage when:

* Joining large tables (you will run out of RAM with the Lookup stage).
* Doing outer joins (left, right, full outer).
* Joining multiple tables with the same keys.

Use the Merge stage when:

* Multiple update and reject links are needed (e.g. Combining a master data set with one or more update datasets).