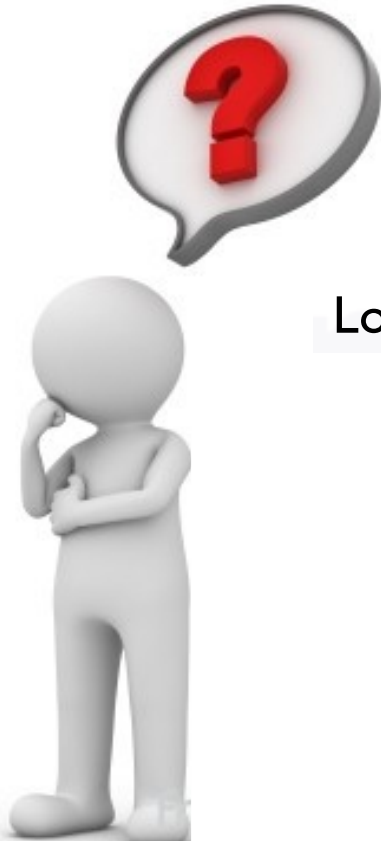




Lookup Stage



Lookup stage means what?

Lookup Stage:



The Lookup stage is a processing stage.

- It is used to perform lookup operations on a data set
- The most common use for a lookup is to map short codes in the input dataset onto expanded information from a lookup table which is then joined to the incoming data and output
- Lookups can also be used for validation of a row. If there is no corresponding entry in a lookup table to the key's values, the row is rejected.
- The Lookup stage is one of three stages that join tables based on the values of key columns.
- The other two are: Join stage and Merge stage

Lookup Stage:



- The Lookup stage can have a reference link, a single input link, a single output link, and a single rejects link.
- Lookup stages do not require data on the input link or reference links to be sorted.
- Large in-memory look up tables will degrade performance because of their paging requirements.
- The optional reject link carries source records that do not have corresponding entry in the input lookup tables.

Partitioning in a Lookup stage:



- Set to Entire partitioning to ensure that the data being looked up in the lookup table is in the same partition as the input data referencing it.
- Partition it in the same way as the input data (although this implies sorting of the data).

Lets us know the difference between
Normal and Sparse Lookup.



IntelliPaat

Normal Vs Sparse Lookup:



Normal Lookup :-

- Reference data needs to be in memory
- Might provide poor performance if the reference data is huge as it has to put all the data in memory.
- Normal Lookup can have more than one reference link.
- Normal lookup can be used with any database.

Normal Vs Sparse Lookup:



Sparse Lookup :-

- Sparse Lookup directly hits the database.
 - If the input stream data is less and reference data is more like 1:100 or more in such cases sparse lookup is better.
 - It can have only one reference link.
 - It can be used with databases.
- Sparse lookup sends individual sql statements for every incoming row.

This Lookup type option can be found in Oracle or DB2 stages. Default is Normal.

Thank You

Email us – support@intellipaat.com

Visit us - <https://intellipaat.com>