**Hive Data Manipulation**

DML (Data Manipulation Language) commands in Hive are used for inserting and querying the data from hive tables once the structure and architecture of the database has been defined using the DDL commands listed above.

Data can be loaded into Hive tables using –

* LOAD command
* Insert command

**Syntax for Load Command in Hive**

load data local inapth ‘path of the file’ into table table\_name;

The file is a ‘|’ delimited file where each row  can be inserted as a table record.

First let’s create a table student based on the contents in the file –

* The **ROW FORMAT DELIMITED** must appear before any of the other clauses, with the exception of the STORED AS … clause.
* The clause **ROW FORMAT DELIMITED FIELDS TERMINATED BY '|** means I character will be used as field separator by hive.
* The clause **LINES TERMINATED BY ‘\n'**means that the line delimiter will be new line.
* The clause **LINES TERMINATED BY ‘\n'** and **STORED AS …** do not require the ROW FORMAT DELIMITED keywords.

If the keyword LOCAL is not specified, then Hive will need absolute URI of the file. However, if local is specified then it assumes the following rules -

* It will assume it’s an HDFS path and will try to search for the file in HDFS.
* If the path is not absolute, then hive will try to locate the file in the /user/ in HDFS.

Using the OVERWRITE keyword while importing means the data will be ingested i.e. it will delete old data and put new data otherwise it would just append the new data. The contents of the target table will be deleted and replaced by the files referred to by file path; otherwise the files referred by file path will be added to the table.