**Session 27**

**Assignment 2**

**Problem Statement:**

Explain the following in brief by using the below datasets and their uses.

**Dataset:**

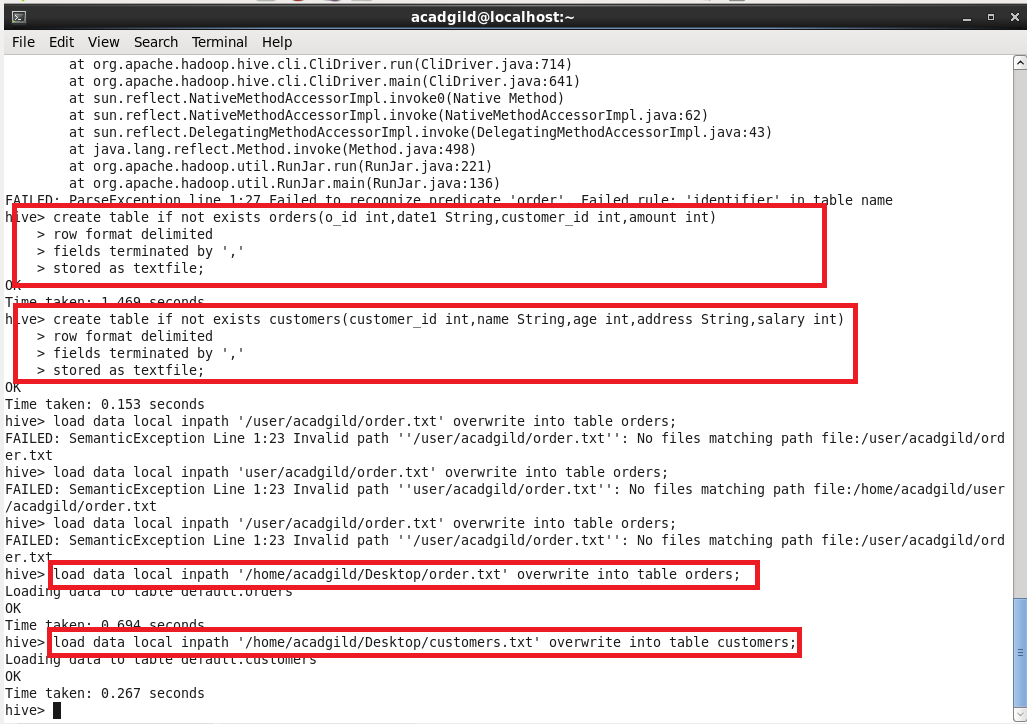
**Link:** order.txt

**Link:** customers.txt

**Terms:**

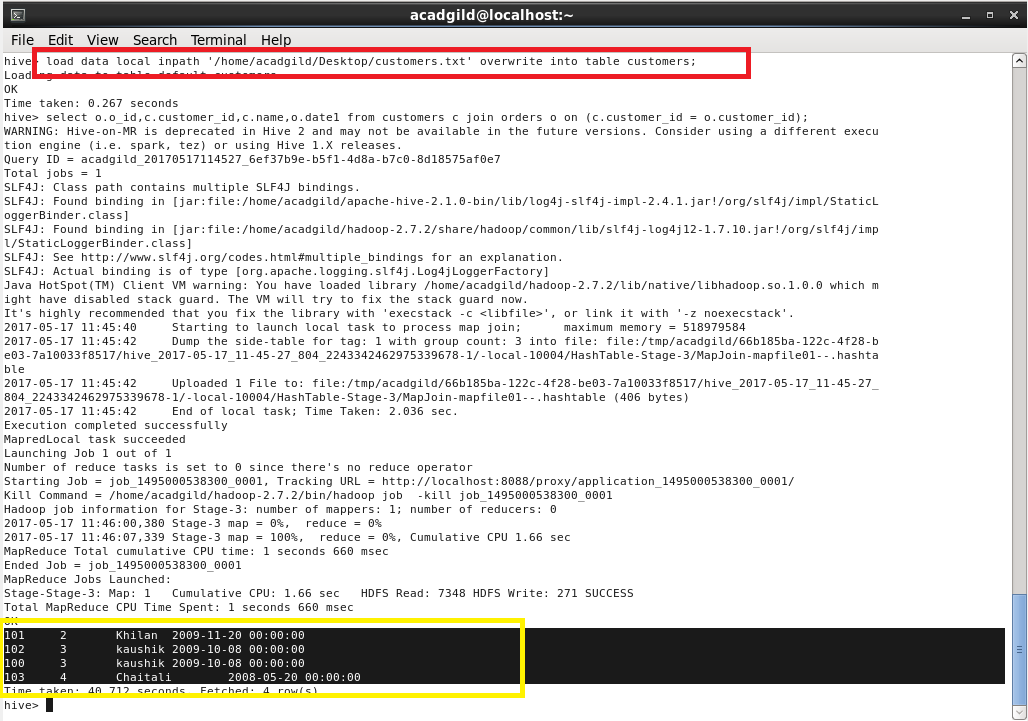
* Join
* Left Outer Join
* Right Outer Join
* Full Outer Join

Creating 2 tables and loading data into the tables



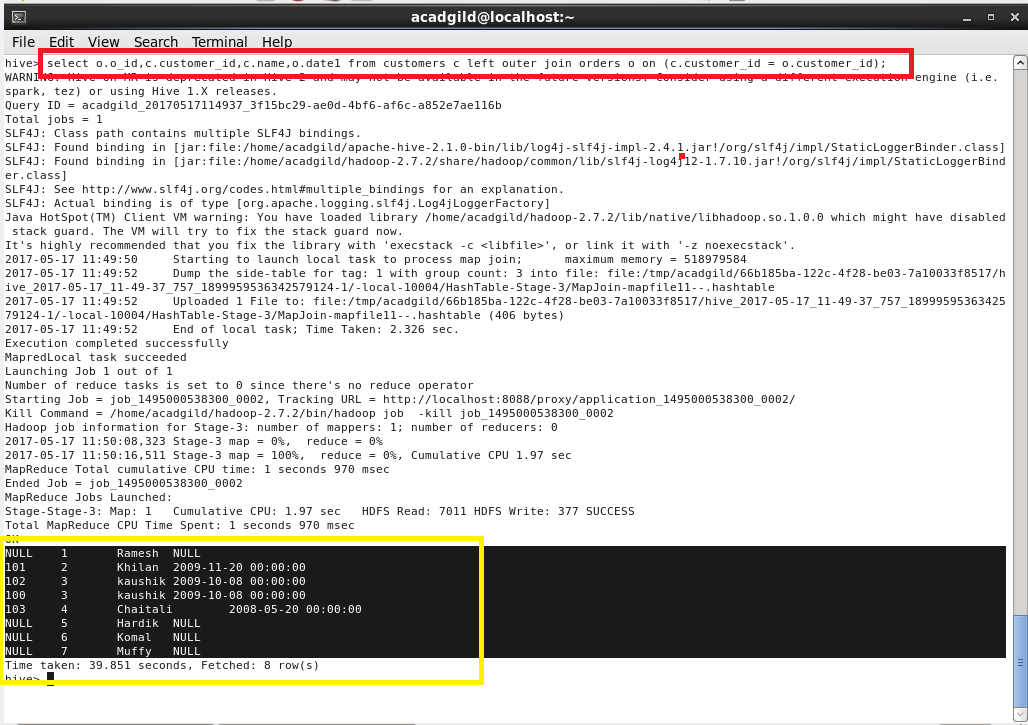
**Join:**

* JOIN clause is used to combine and retrieve the records from multiple tables.
* JOIN is same as INNER JOIN in SQL.



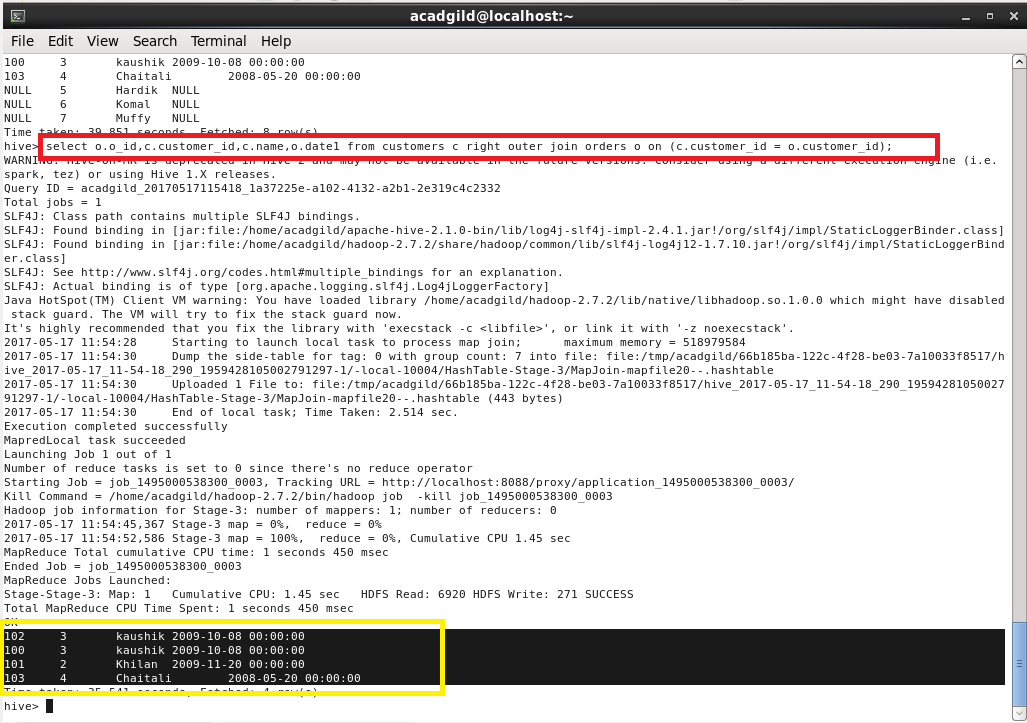
**Left Outer Join:**

* The HiveQL LEFT OUTER JOIN returns all the rows from the left table, even if there are no matches in the right table.
* This means, if the ON clause matches 0 (zero) records in the right table, the JOIN still returns a row in the result, but with NULL in each column from the right table.
* A LEFT JOIN returns all the values from the left table, plus the matched values from the right table, or NULL in case of no matching JOIN predicate.



**Right Outer Join:**

* The HiveQL RIGHT OUTER JOIN returns all the rows from the right table, even if there are no matches in the left table.
* If the ON clause matches 0 (zero) records in the left table, the JOIN still returns a row in the result, but with NULL in each column from the left table.



**Full Outer Join:**

* The HiveQL FULL OUTER JOIN combines the records of both the left and the right outer tables that fulfil the JOIN condition.
* The joined table contains either all the records from both the tables, or fills in NULL values for missing matches on either side.

