**ASSIGNMENT - 27.2**

**Problem Statement:**

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

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

**Input Data Set:**

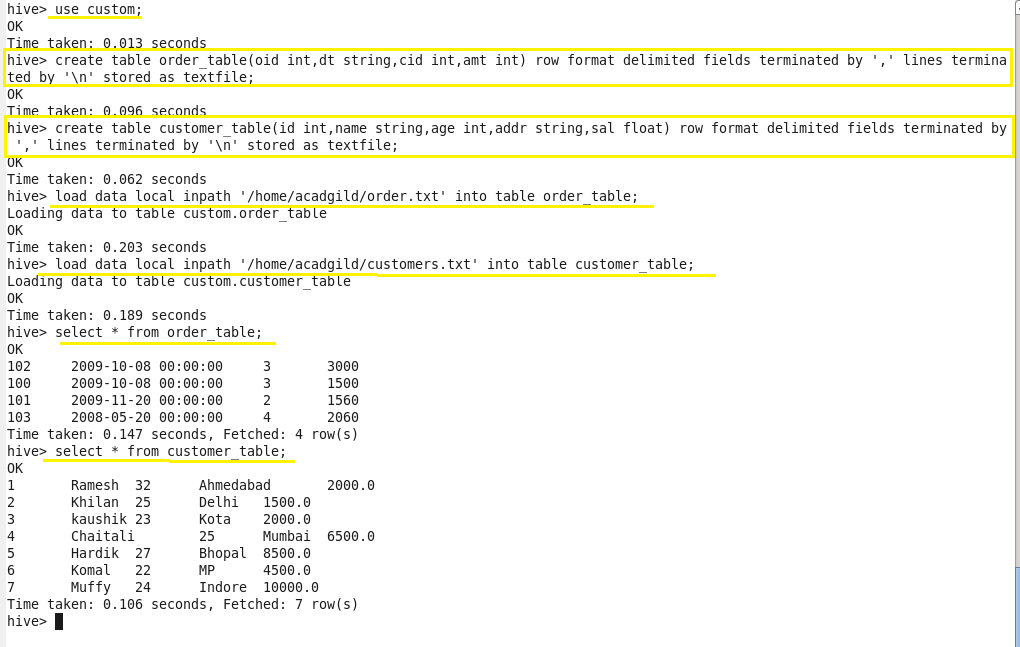
**Link1:** <https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/order.txt>

**Link2:**

<https://github.com/prateekATacadgild/DatasetsForCognizant/blob/master/customers.txt>

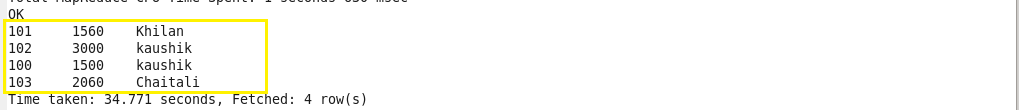
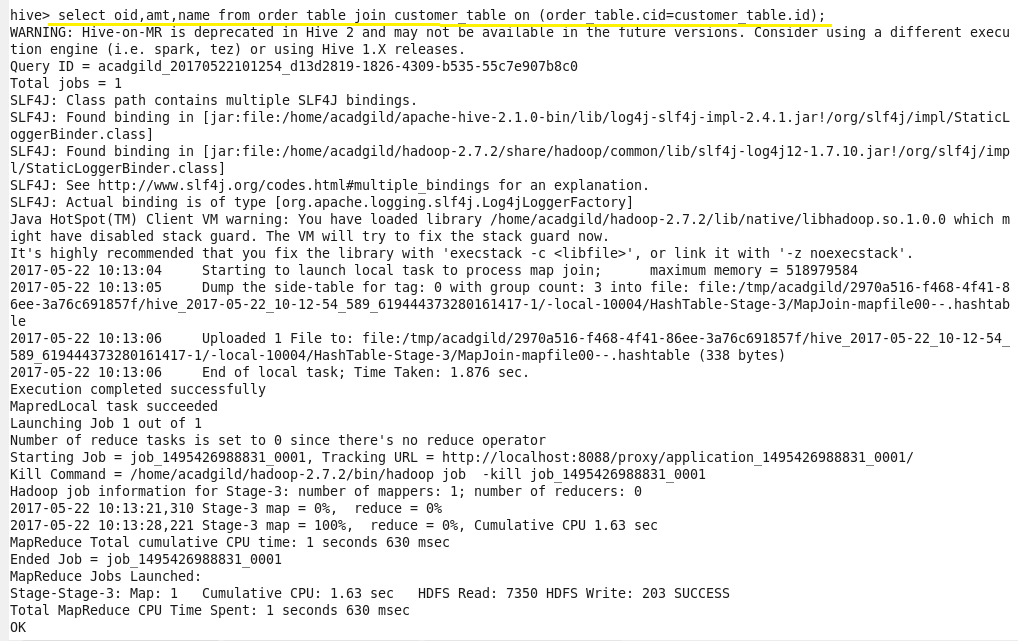
**Solution:**

Here we have created two tables and we have loaded given datasets into these tables.



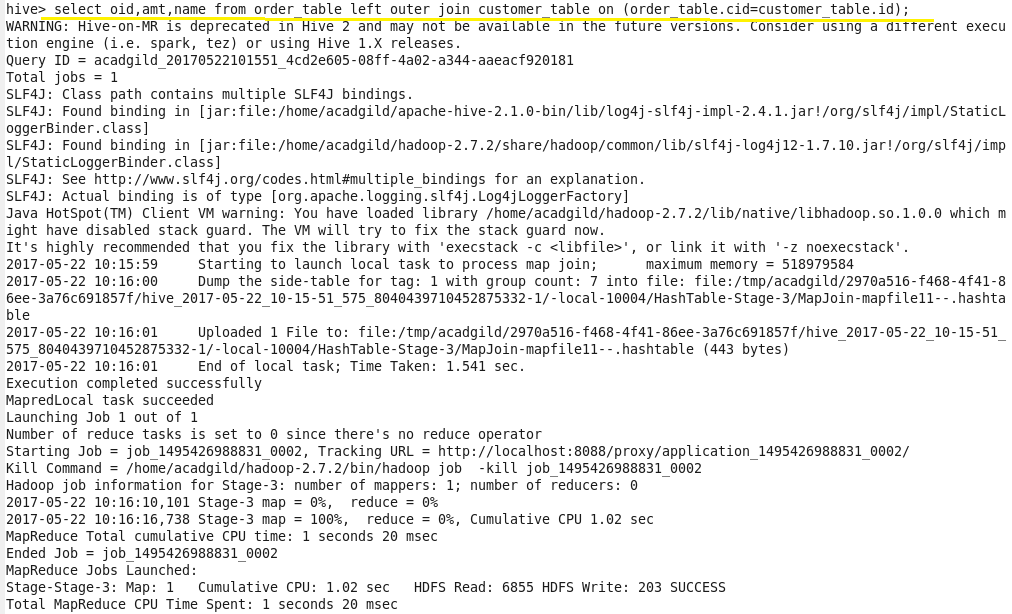
* **Join:**

Select records that have matching values in both tables. Hive takes the last table in the JOIN statement for streaming.



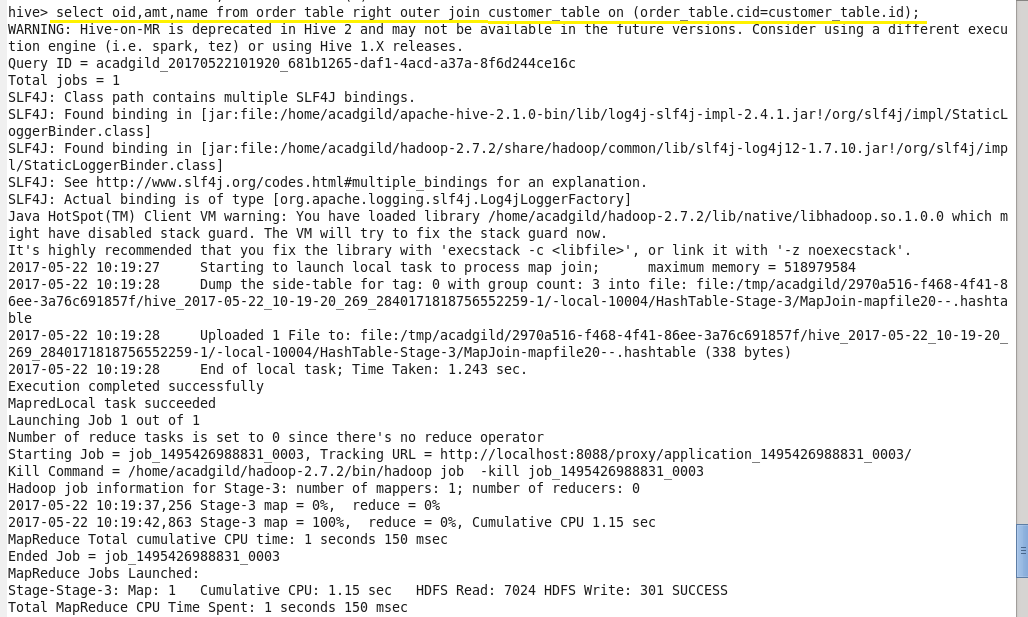
* **Left Outer Join (Left Join):**

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 (Right Join):**

Returns all the values from the right table, plus the matched values from the left table, or NULL in case of no matching join predicate.



* **Full Outer Join (Full Join):**

Selects all records that match either left or right table records.

