**ASSIGNMENT - 27.6**

**Problem Statement:**

Using the below given two datasets you need to give a demo on the below joins in hive.

* Bucket Map join
* Sort-Merge Bucket join
* Sort-Merge Bucket Map join
* Left semi join

**Input:**

You can download the datasets from the below link:

<https://drive.google.com/open?id=0ByJLBTmJojjzSWhfXzI5NGh4RGM>

Dataset description of dataset1.csv is as follows:

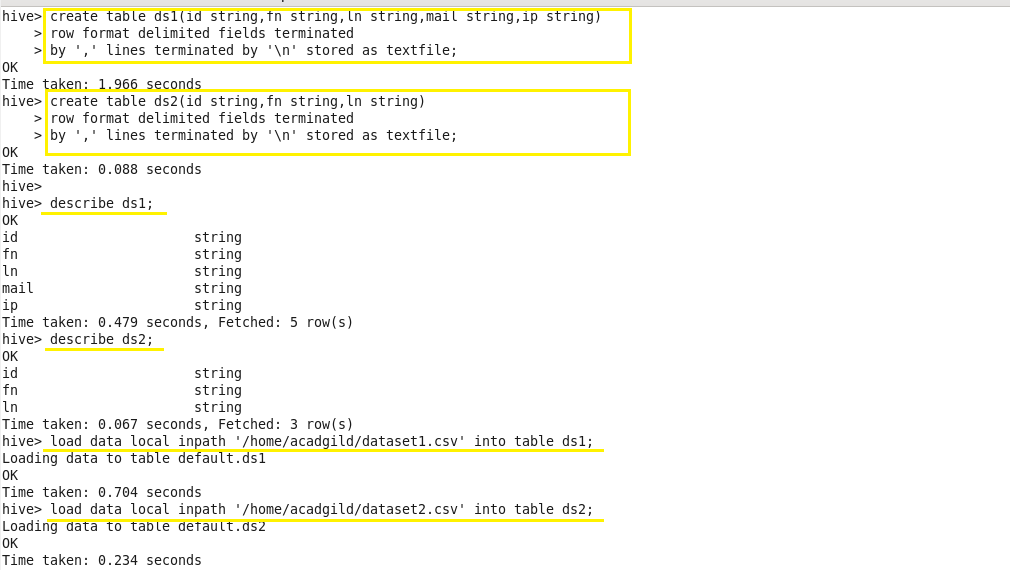
id,first\_name,last\_name,email,gender,ip\_address

Dataset description of tdataset2.csv is as follows:

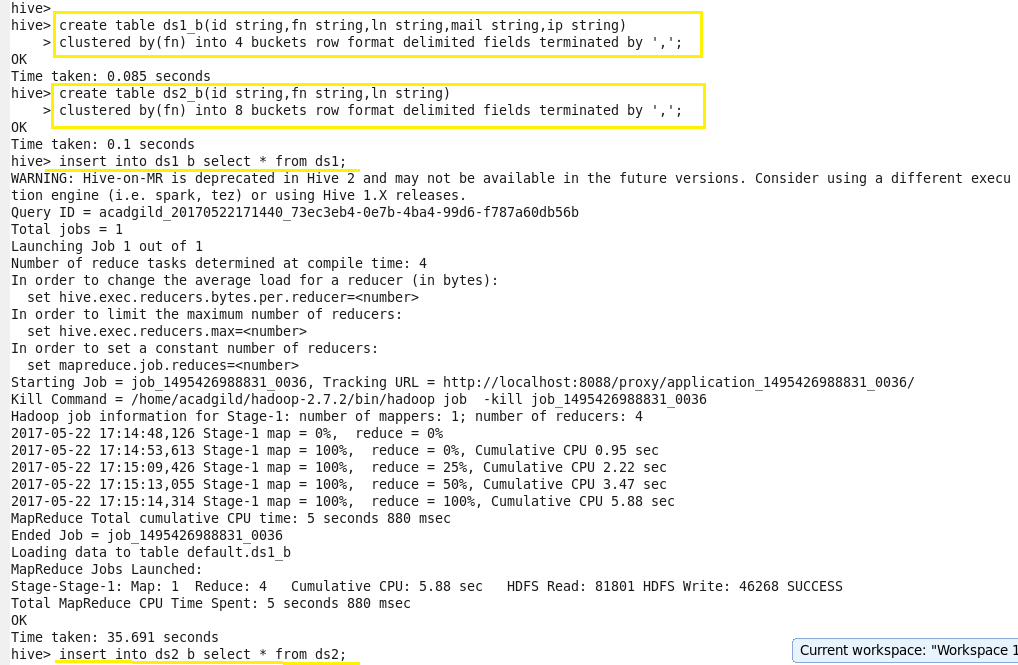
id,first\_name,last\_name

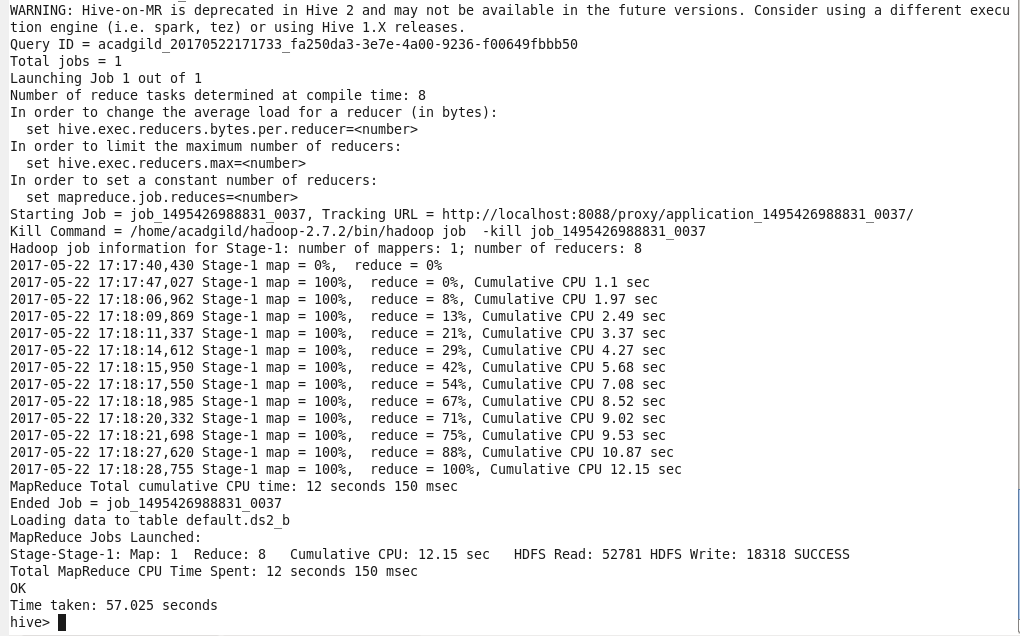
**Output:**

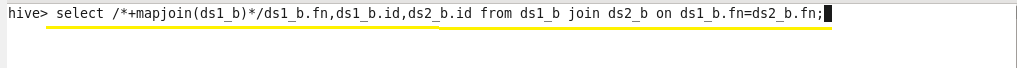
Creating tables ds1 & ds2 and loading dataset1 and dataset2 into the tables.

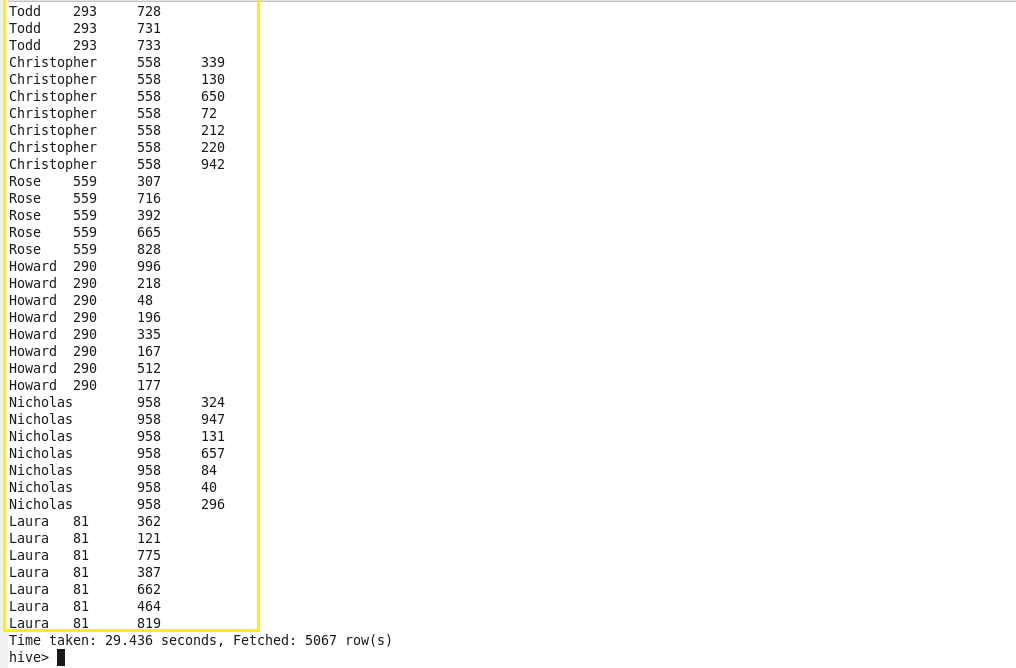


* **Bucket Map join:**

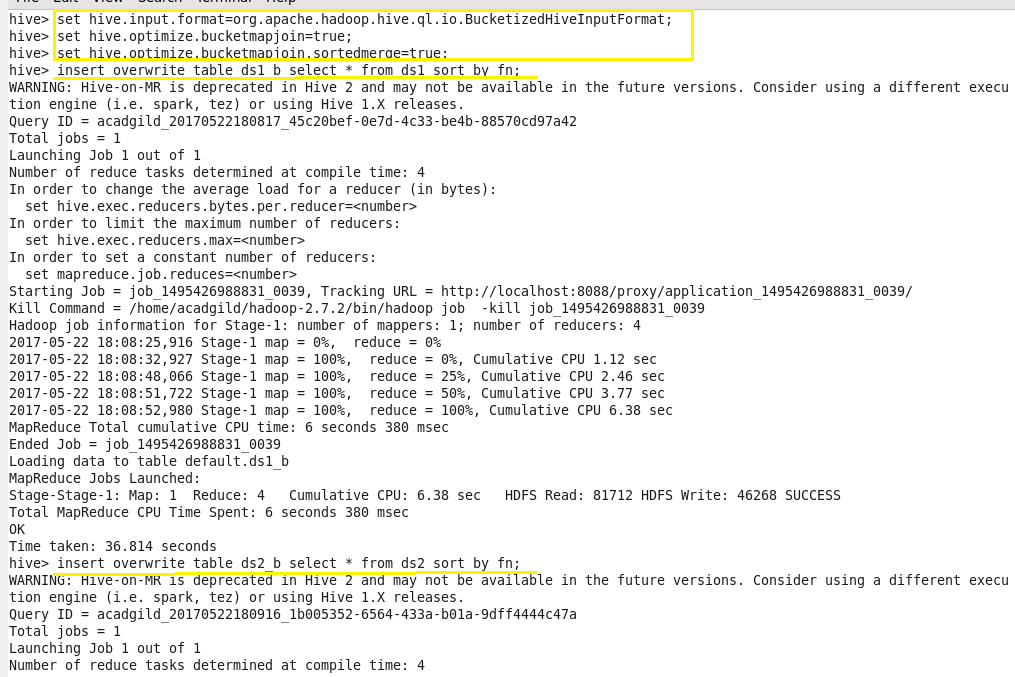


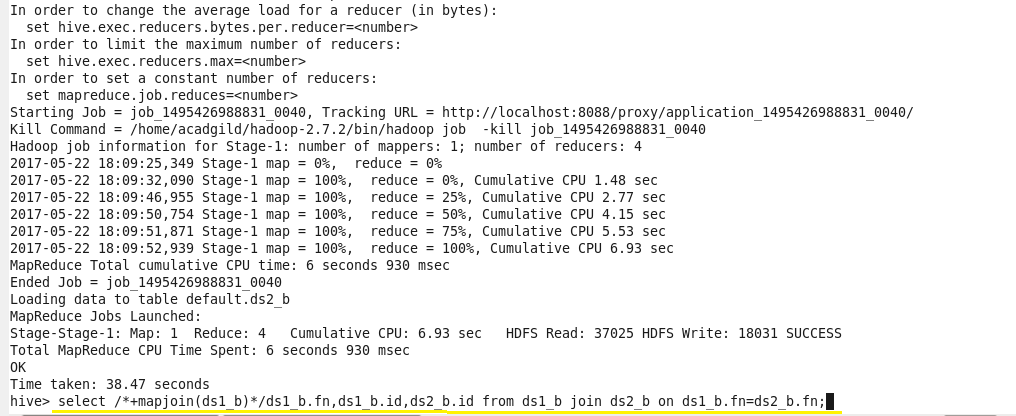


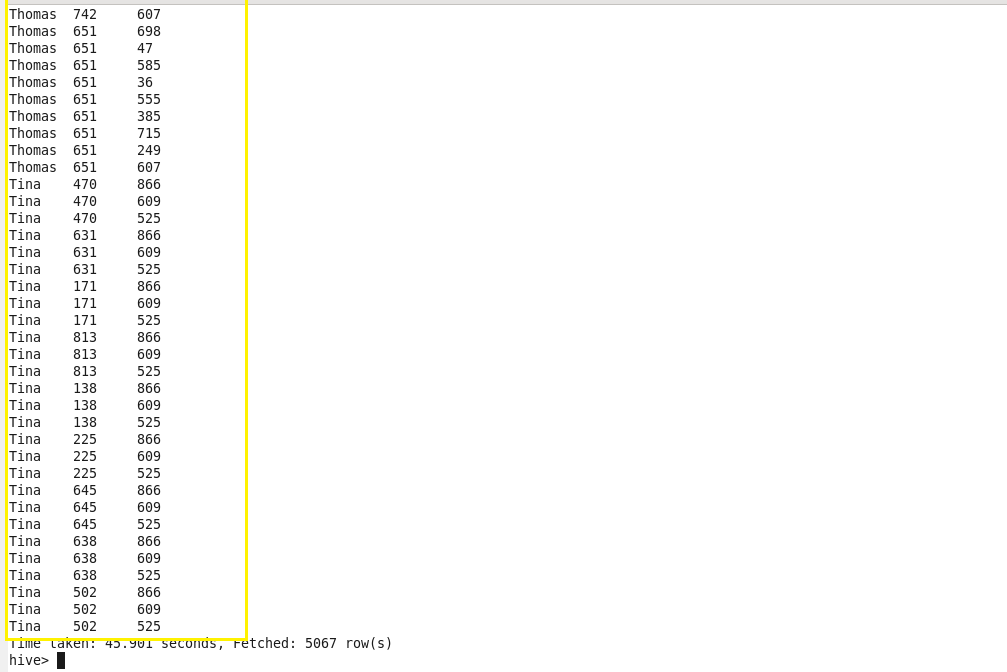




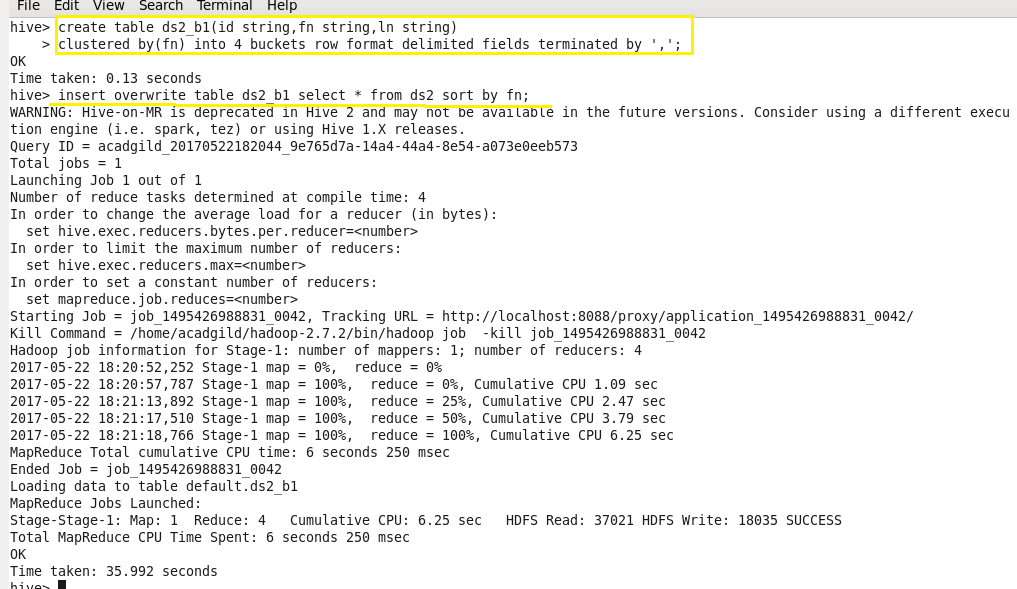
* **Sort-Merge Bucket join:**

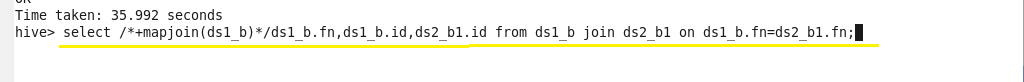






* **Sort-Merge Bucket Map join:**







* **Left semi join:**

