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

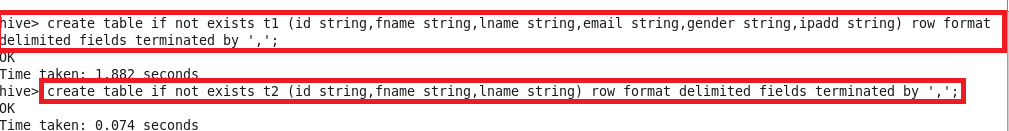
Sort-Merge Bucket Map join

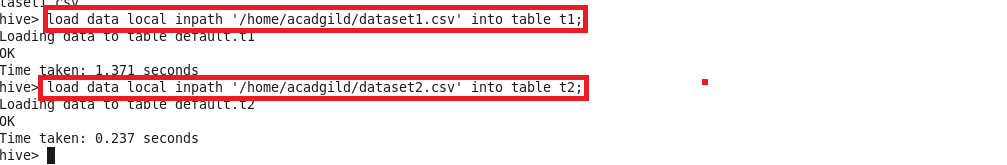
Left semi join

**Solution:**

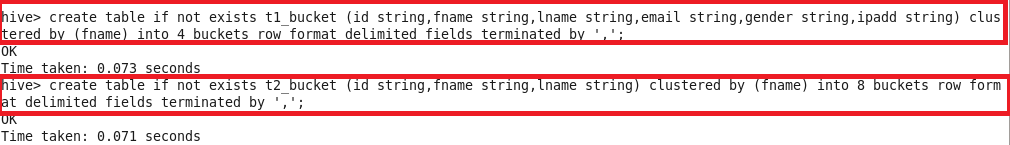
**Bucket Map join**

**Creating two tables and loading the data in these tables:**

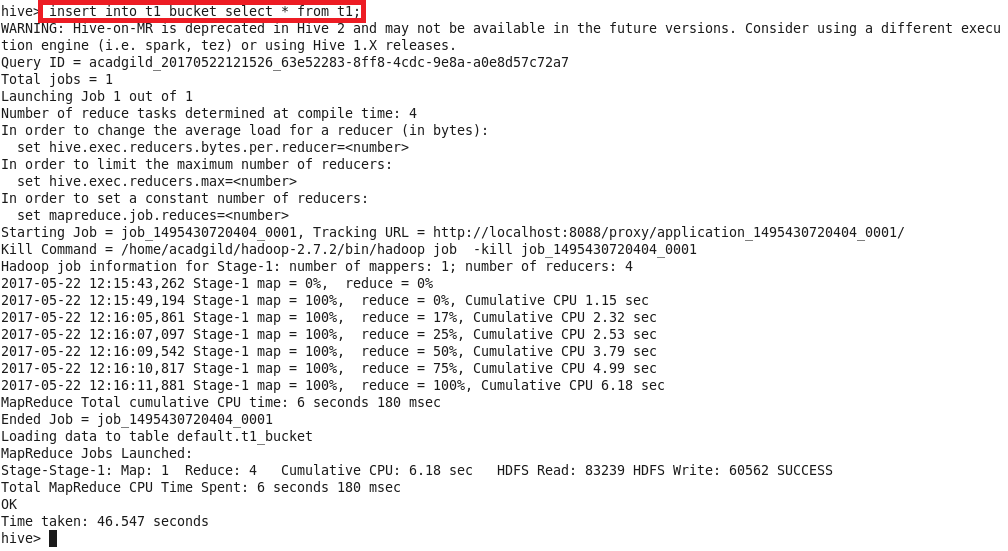
****

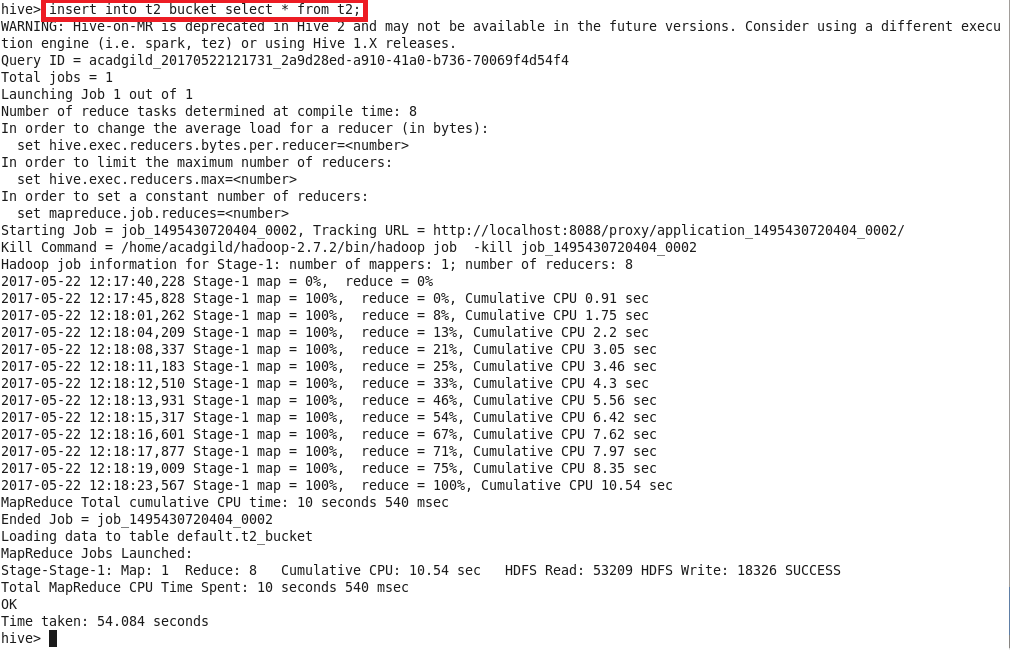
****

**Creating 2 bucket tables with 4 and 8 buckets:**

****

**Loading the data into bucket tables:**

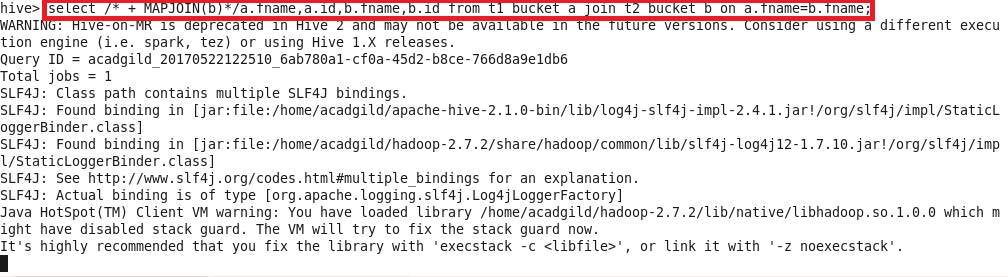
****

****

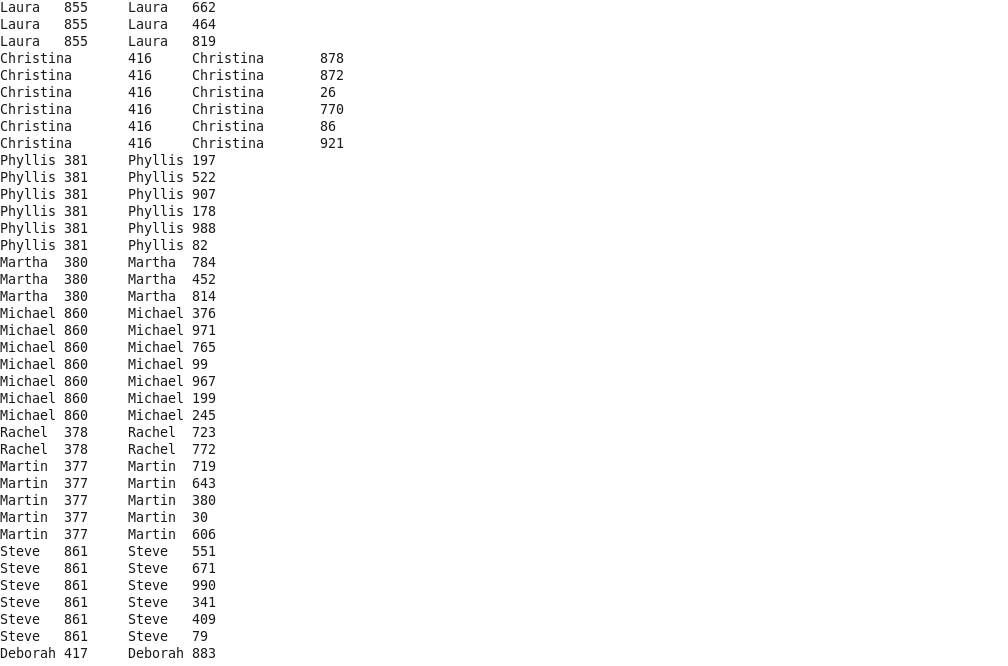
For performing Bucket-Map join, we need to set

set hive.optimize.bucketmapjoin = true

and writing query for bucket map join



**Output:**

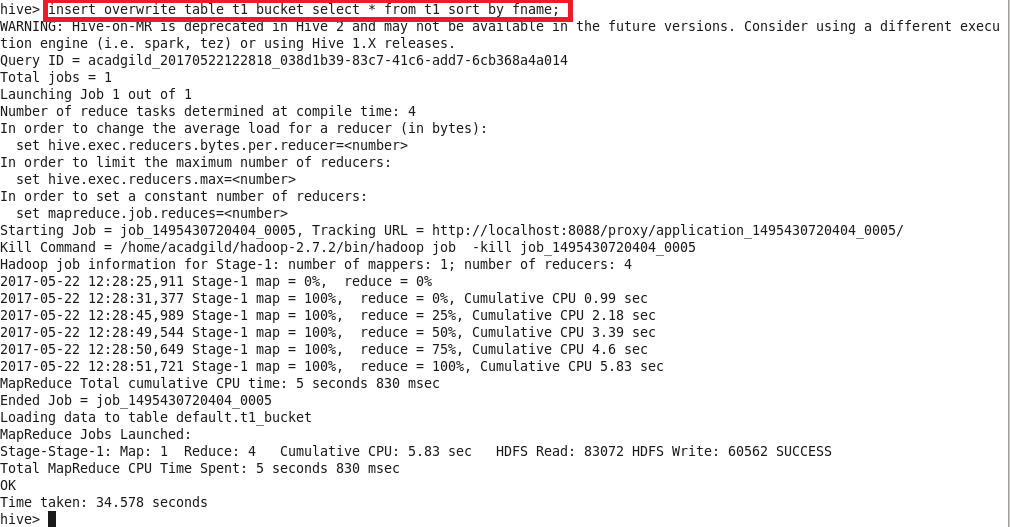
****

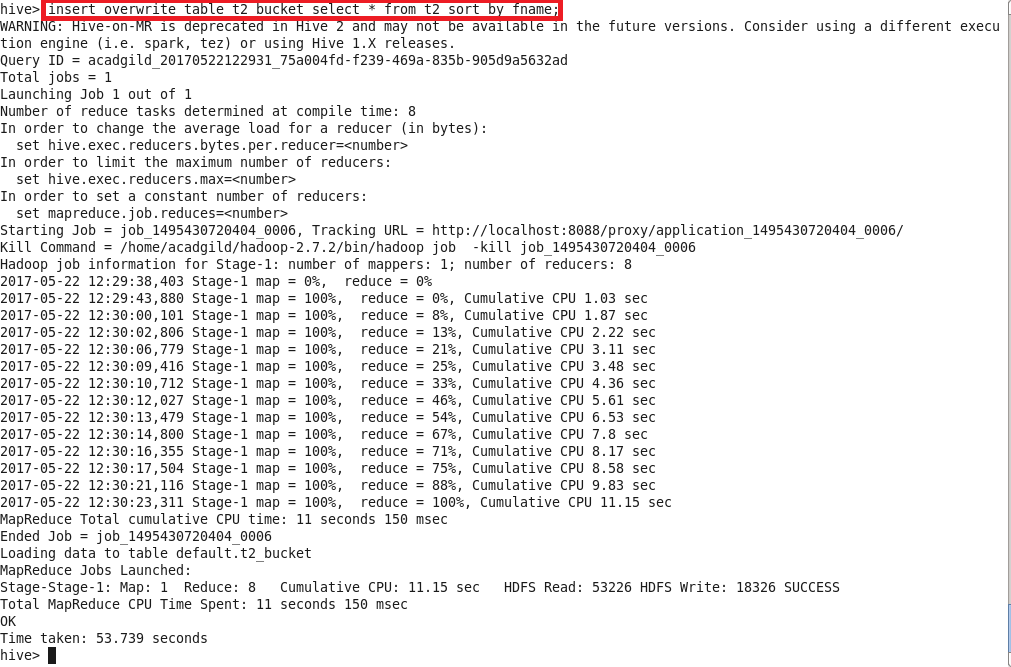
**Sort-Merge Bucket join**

We need to set :

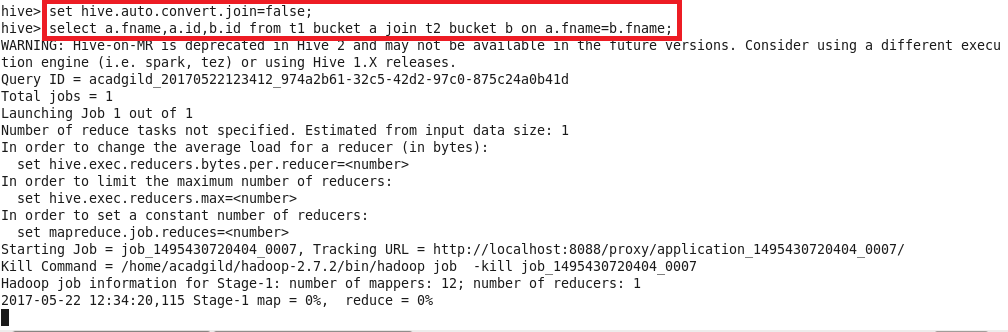


**Sorting the contents of bucket tables:**

****

****

**Writing query for sort merge bucket join:**

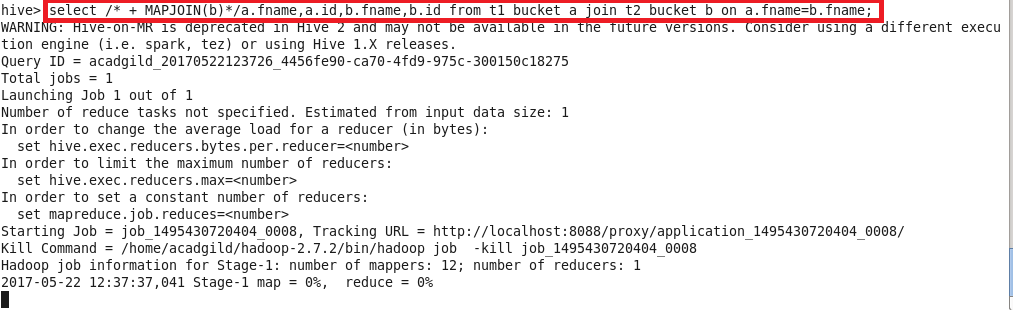
****

**Output:**

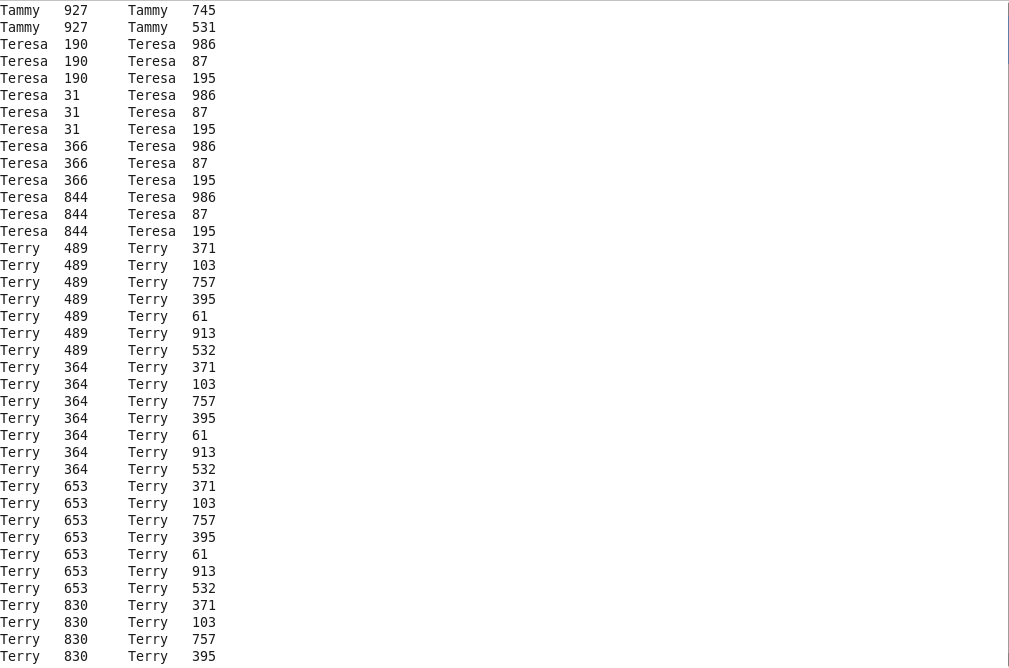
****

**SORT-MERGE BUCKET MAP JOIN:**

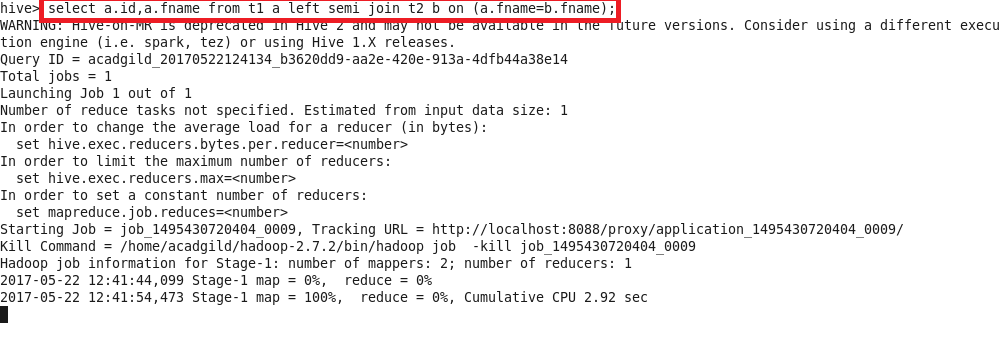
**Using the same bucketed tables previously used,**

****

**Output:**

****

**Left Semi Join:**

****

**Output:**

****