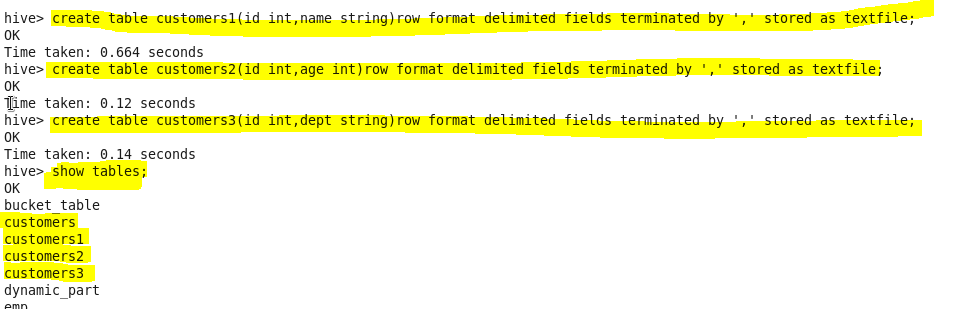
**ASSIGNMENT 27.4**

* **Run parallel jobs in hive**

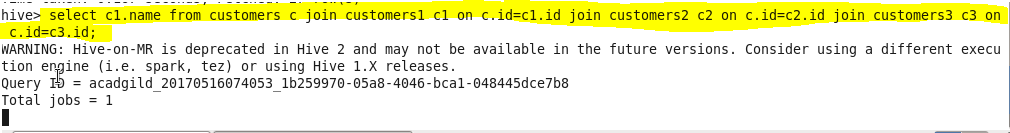
A particular job may consist of some stages that are not dependent on each other and could be executed in parallel, possibly allowing the overall job to complete more quickly.

Hive can convert a query into one or more stages and to save time executes multiple jobs in parallel.

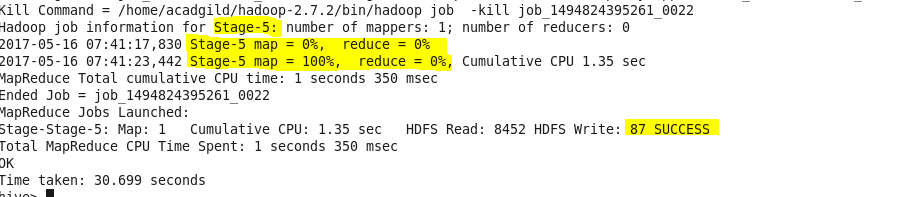
First we create new tables to make joins:



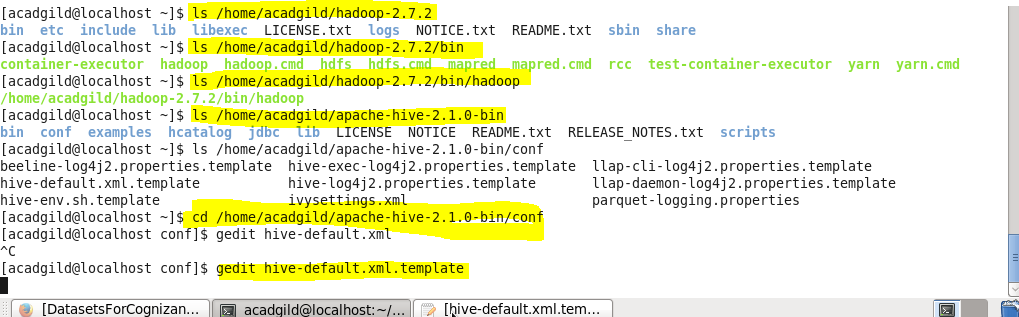
Query to join the tables:

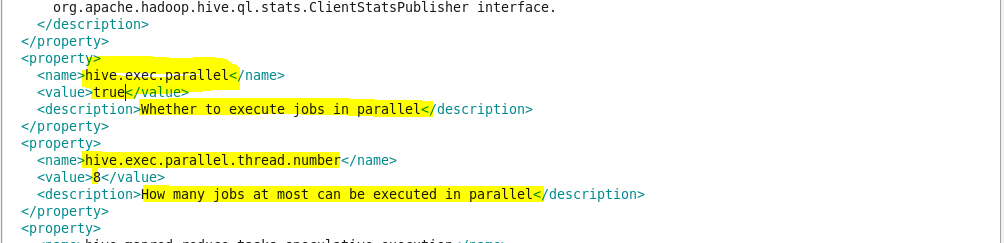


The processing is completed in many stages by the mapper a shown:



If the query is optimized and more stages are run simultaneously, the job may complete much faster.   
However,If a job is running more stages in parallel, it will increase its cluster utilization.





Once property is set save it and restart the hive shell in new terminal.

Shoot the optimized query inside shell to see multiple jobs launching .

