**PROBLEM - 1**

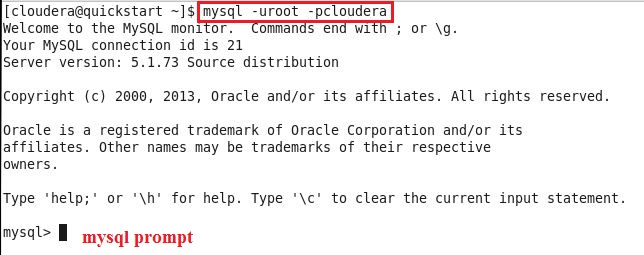
**Import and Export between MYSQL and HDFS using SQOOP**

**Note: Cloudera quickstart VM is used to perform import and export between MYSQL and HDFS**

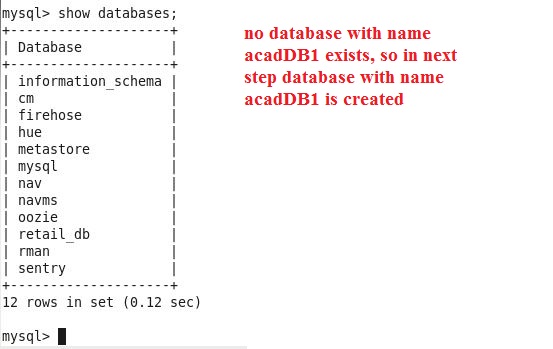
**In Cloudera quickstart VM all daemons are started at the time when we start VM, so there is no need to start all the required daemons manually, like as in acadgild VM we start hadoop daemons with “start-all.sh” command and mysql service with “sudo service mysqld start” command.**

**To import the data from MYSQL into HDFS, below steps are followed:**

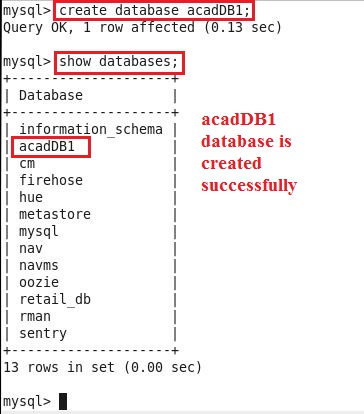
**Step 1: Logged into mysql database using below command:**

****

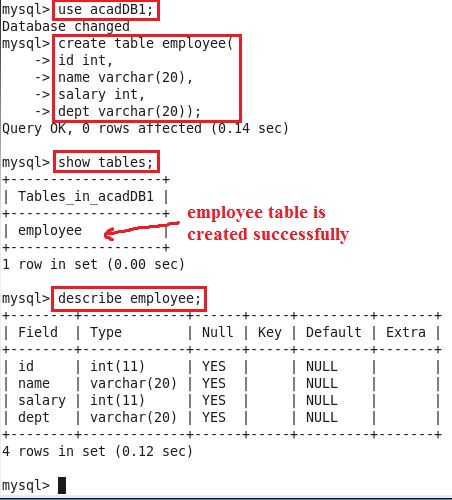
**Step 2: Checked whether database with name “acadDB1” already exists using below command:**

****

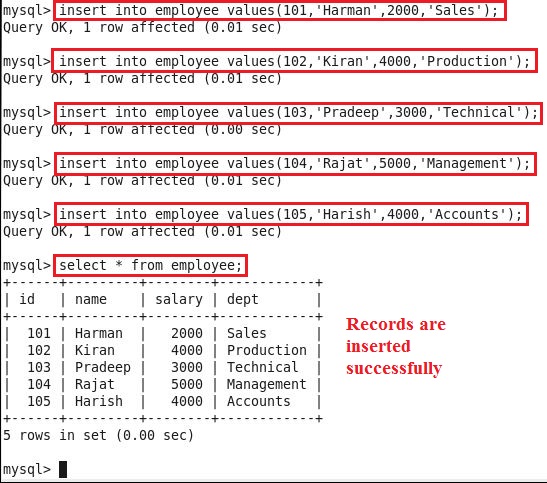
**Step 3: Created Database “acadDB1” using below command:**

****

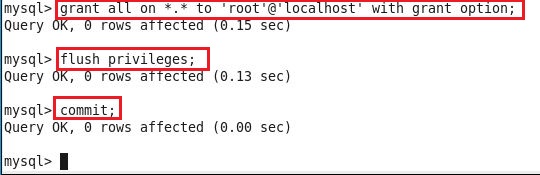
**Step 4: Created table “employee” with following fields inside acadDB1 database:**

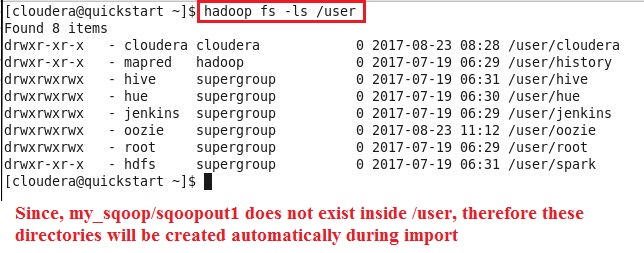
****

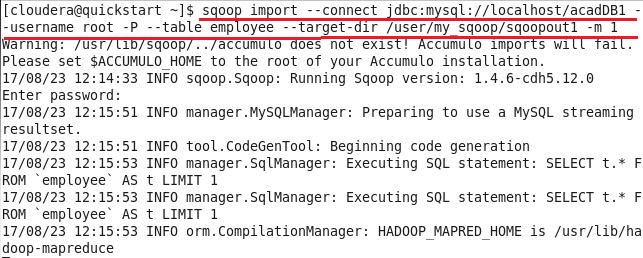
**Step 5: Inserted few records inside employee table:**

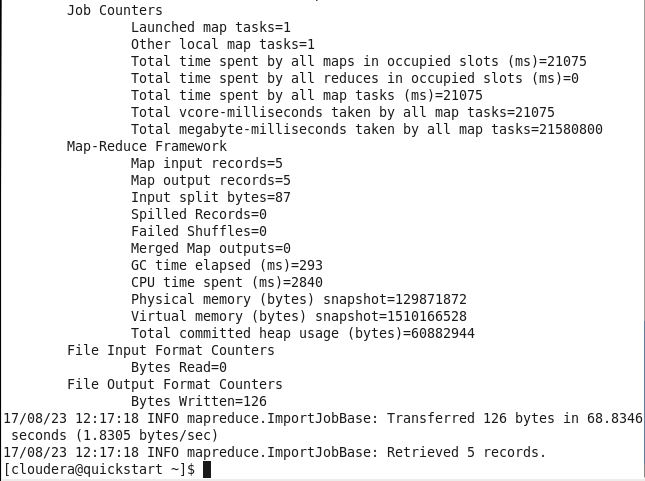
****

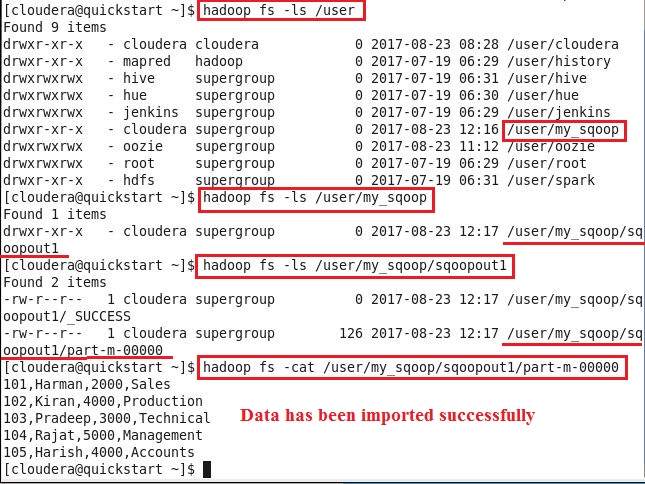
**Step 6: Granted permission to root user to access the database over the network, followed by flushing the privileges (The reload/flush privileges command tells the server to reload the grant tables into memory), and committing all changes to database as follows:**

****

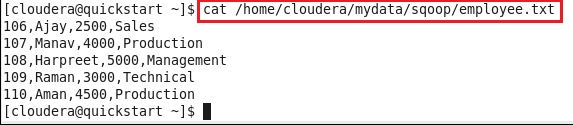
**Step 7: Checked in hdfs, which directories exist inside /user, because we want data should be imported to /user/my\_sqoop/sqoopout1 location, so sqoopout1 should not already exist, else during import error would have been returned if we used already existing directory name:**

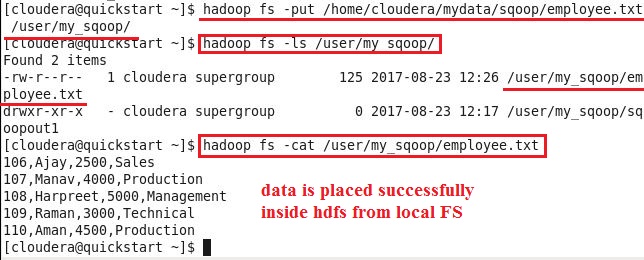
**Step 8: Using sqoop import command, imported the data from employee table of acadDB1 database inside mysql to hdfs at /user/my\_sqoop/sqoopout1 location:**

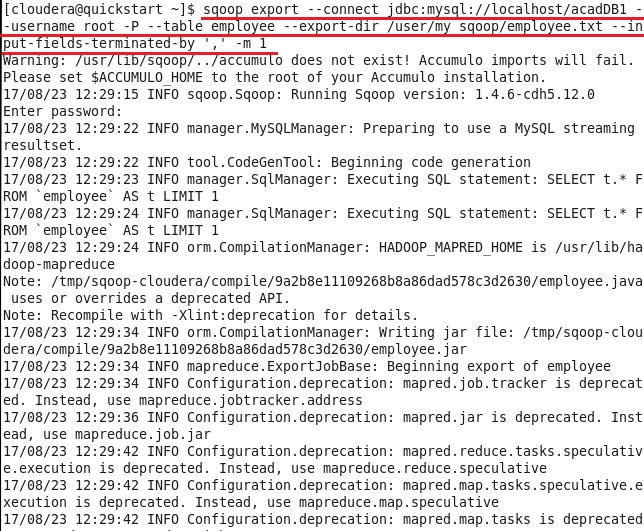
****

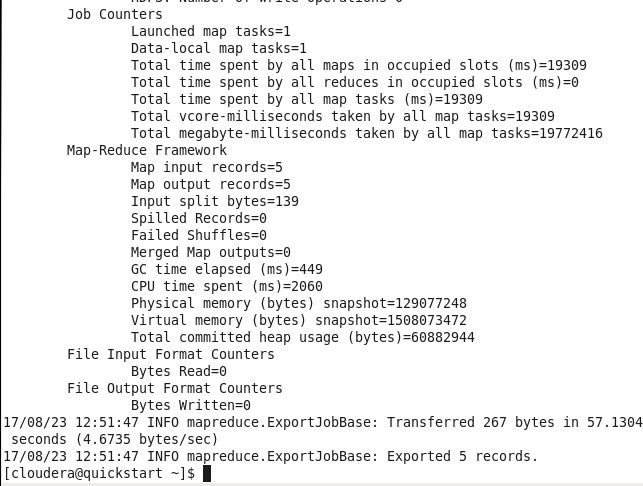
**Step 9: Inside hdfs, checked using below commands whether data imported successfully or not:**

**To export the data from HDFS to MYSQL, below steps are followed:**

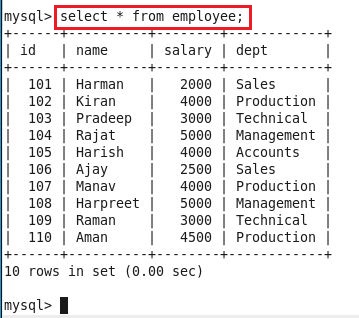
**Step 1: To export the data from hdfs to employee table [created during import statement], created employee.txt file [using gedit employee.txt command] in local file system first with data matching the schema of employee table:**

**Step 2: Put employee.txt file from local FS to hdfs at following location:**

**Step 3: Using export command, exported data from hdfs to mysql:**

****

**Step 4: Above screenshots show that export command ran successfully, so using select statement checked whether data in employee table inside mysql placed properly or not:**

****

**From above screenshot we can see that last 5 rows are placed from hdfs, earlier there were only first 5 rows inside employee table.**

**PROBLEM - 2**

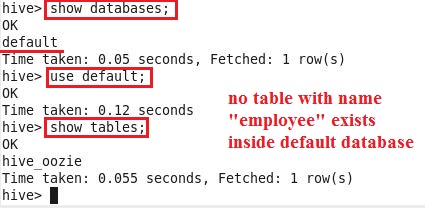
**Import and Export between MYSQL and HIVE using SQOOP**

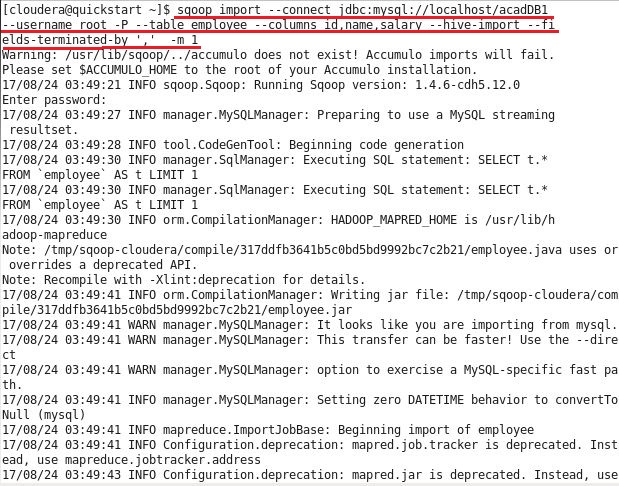
**NOTE: Here, we need to import and export only selected columns from mysql table to hive table and vice-versa using sqoop.**

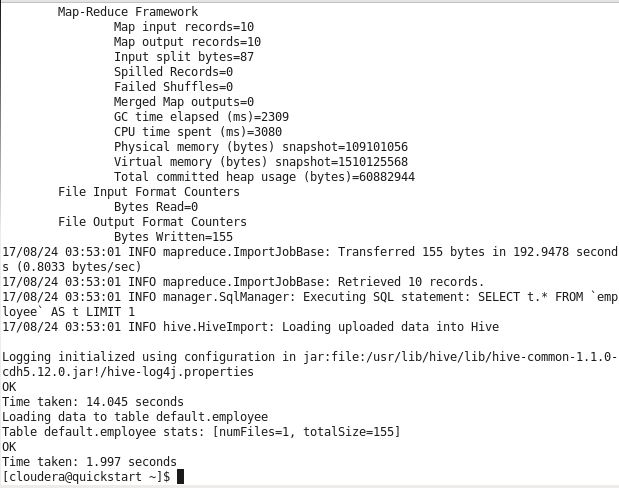
**Employee table created in Problem-1 is considered to import the data from mysql to hive, and empid table is created in mysql to export only selected data from hive to mysql and it’s assumed that grant and flush commands are executed in mysql.**

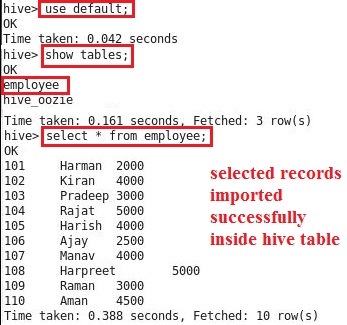
**To import the data from MYSQL into hive, below steps are followed:**

**Step 1: Checked whether employee table already exists inside default database in hive or not**

****

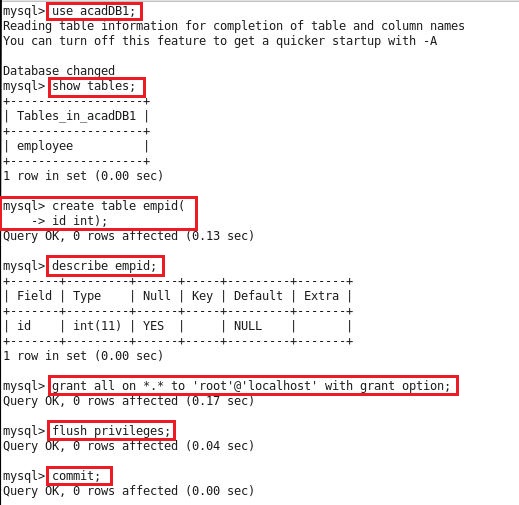
**Step 2: As employee table does not already exist inside hive, so using sqoop import command, imported selected data from mysql table to hive table [table will get created automatically with same name as in mysql]**

****

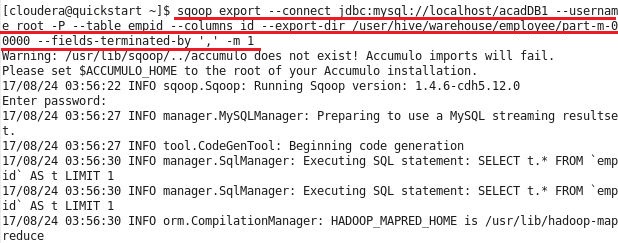
**Step 3: Above screenshot shows that import statement ran successfully, without any error, so checked inside hive whether table with three columns i.e. id, name, salary along with data got created or not inside default database:**

**To export the data from HIVE to MYSQL, below steps are followed:**

**Step 1: Created “empid” table in acadGB1 database in mysql with one column i.e. id (type int), so that only id field values should get exported from hive table to mysql table [NOTE: we need to create table in mysql with the required number of fields, else export command will fail if table will not already exist in mysql]:**

****

**Step 2: To export only selected column i.e. id from employee table in hive to mysql empid table, used below command:**

****

****

**Step 3: Checked in mysql, whether data got successfully exported to mysql table or not:**

****