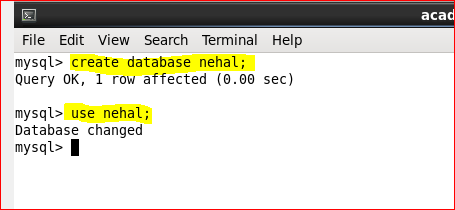
Problem Statement:

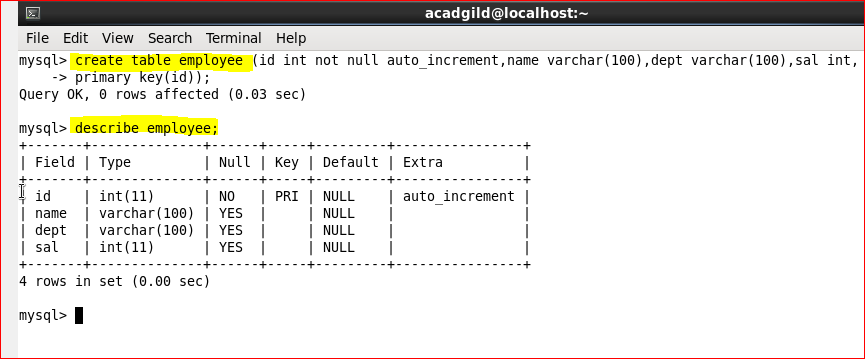
● Explain and perform Sqoop Incremental Import operation to load data from Mysql To HDFS using Sqoop

● Explain the procedures performed, Share the screenshots of commands and results for the same.

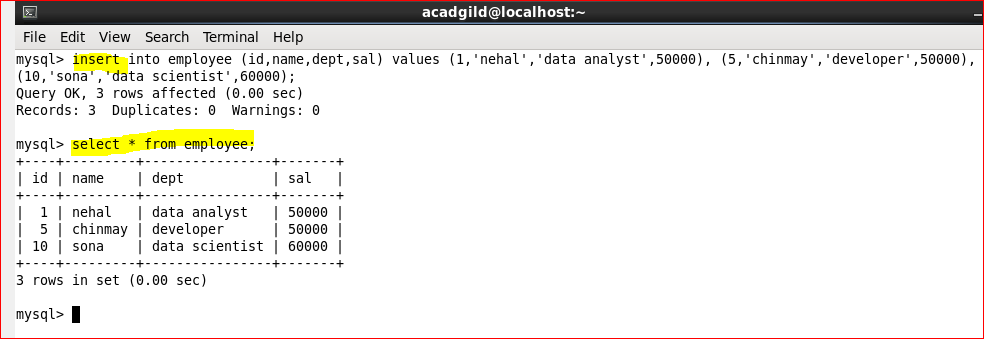
Create a database by using the below command:



Create a table **employee**by using the below command:

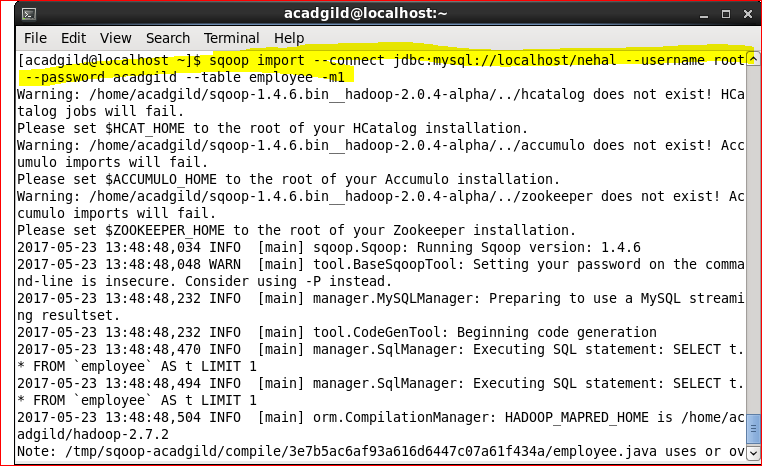


A table has been created with name **employee**and with the columns **id, name, dept and sal.**Let’s now try inserting some sample data into the created table by using the below command:



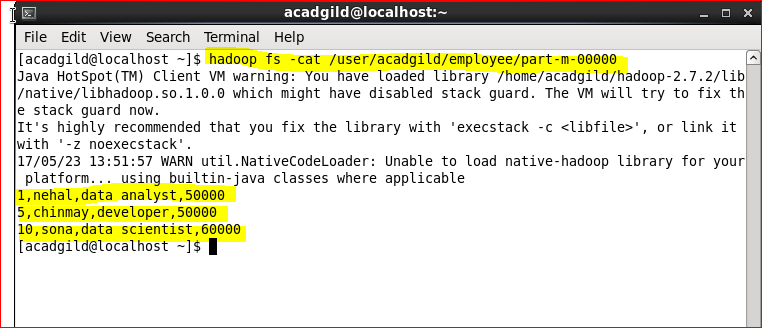
We have successfully created a table in MySQL, and we will now import the same into HDFS by using Sqoop.Here we are connecting to MySQL through JDBC connectors and using the database **nehal.**Here it is necessary to specify the **MySQL‘s username and password**and the **table name.**

Here ‘-m’ specifies the number of map task that can be run simultaneously and ‘m1’ means that only one map task can run.

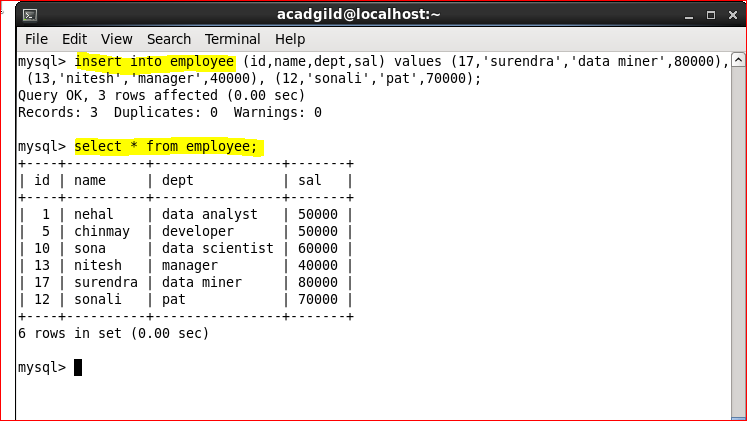


If we do not use **-m1**at the end of the statement, for each record in the MySQL table we will get separate files in the HDFS.

Now the data in RDBMS has been successfully imported into HDFS. By default, the files will be stored here: **/user/acadgild/employee/part-m-00000**file.



Let’s manually insert few extra values in nehal/employee table.



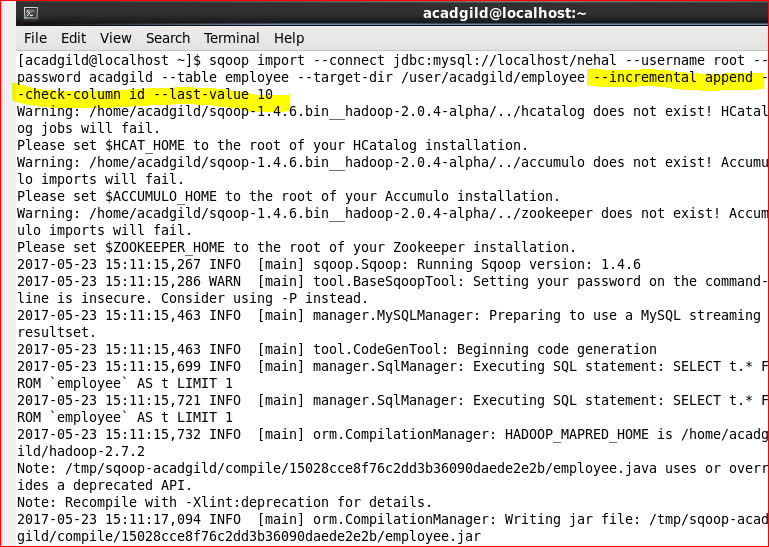
Now, the following command with little few extra syntax will help to feed only the new values in the table**employee.**

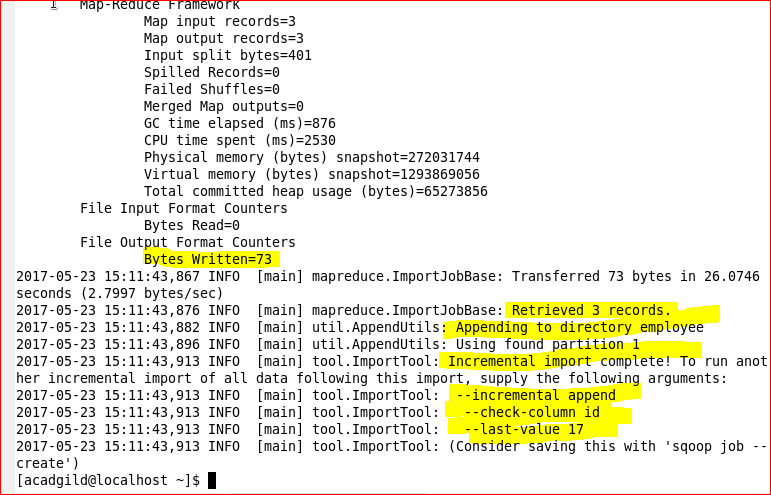
The following syntax is used for the incremental option in Sqoop import command.

--incremental <mode>

--check-column <column name>

--last value <last check column value>

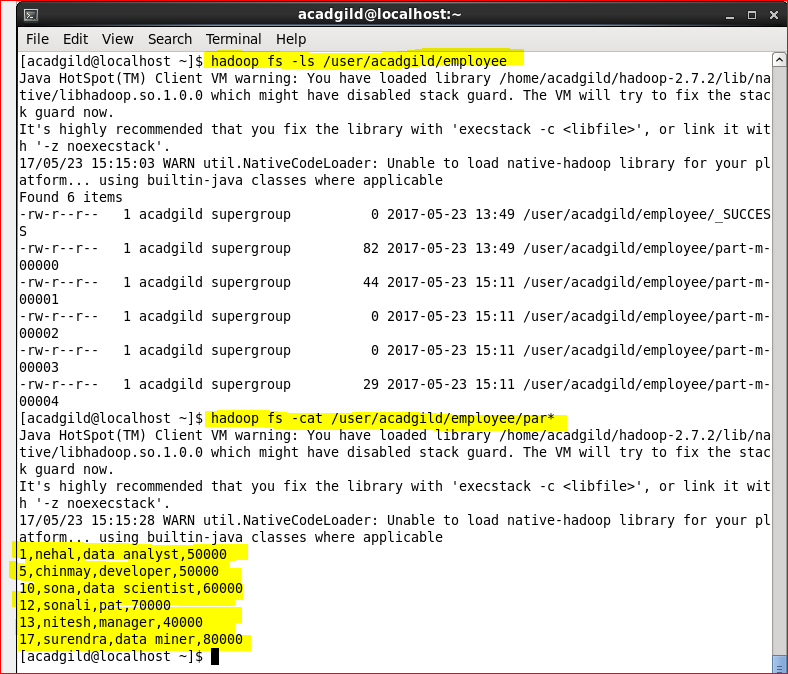




3 more records have been retrieved and the incremental import is now complete.

Along with message for next incremental import, you need to give last value as**17.**

Let’s check and confirm the new data inside HDFS.



This is how incremental import is done every time for any number of new rows.