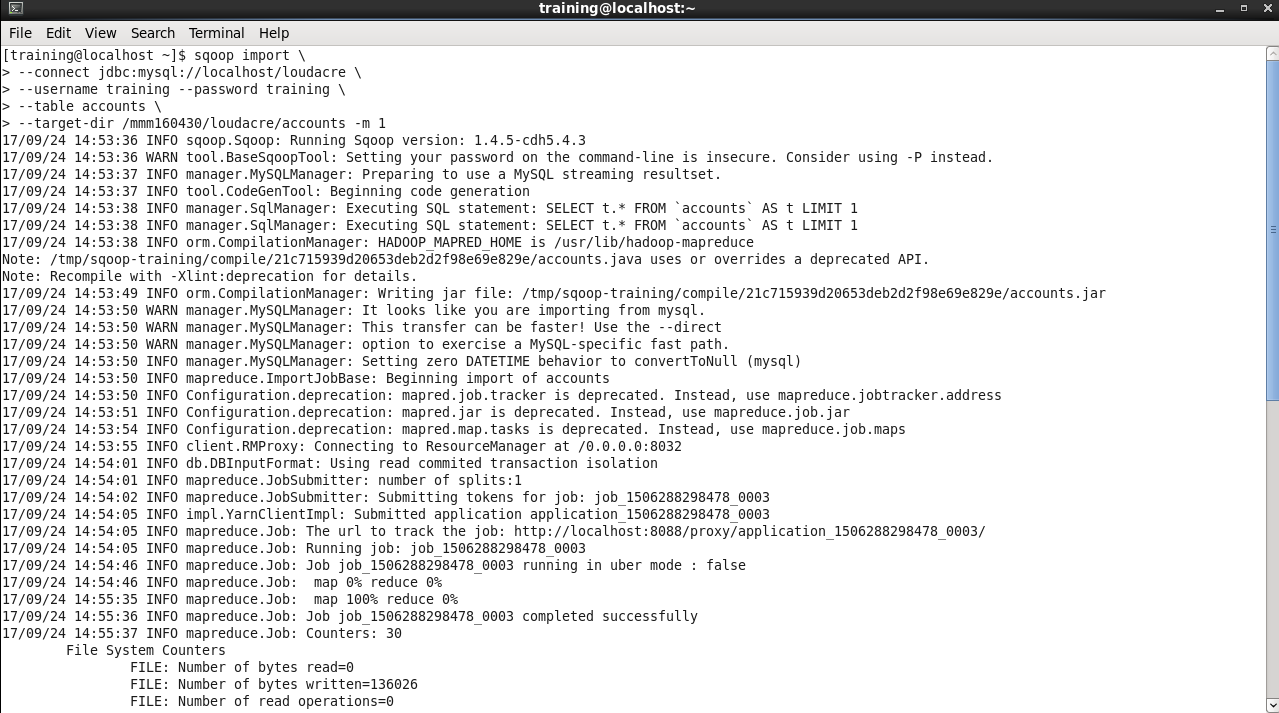
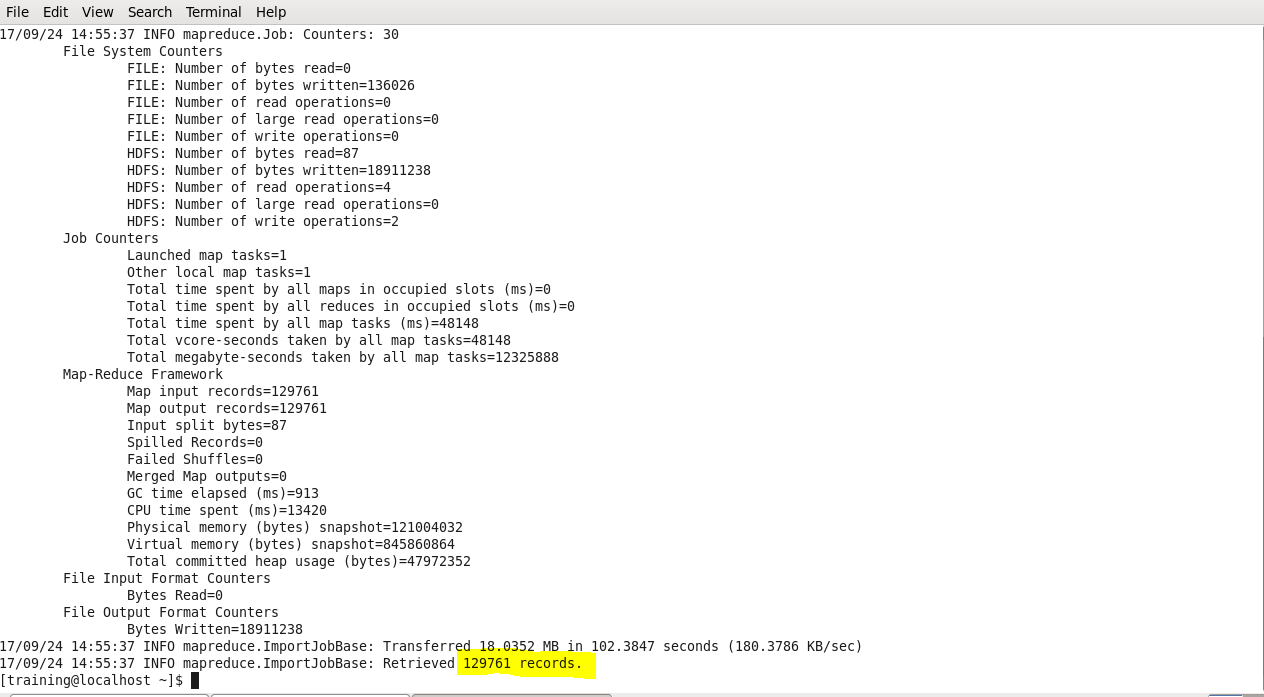
**Big Data Analytics – Assignment 1**

**Submitted by: Mrunmayi Modak**

**PART A**

1. Import table ‘accounts’ into a directory named as your NET-ID/loudacre, without creating sub directories

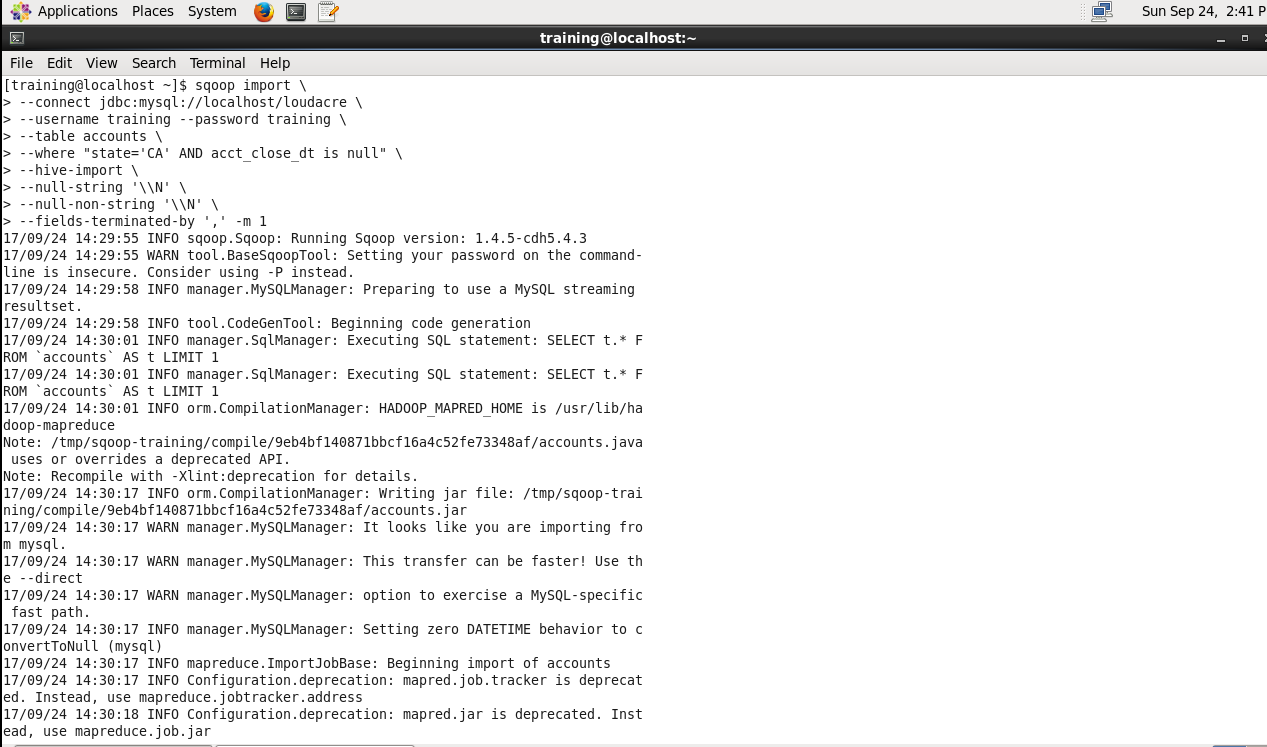


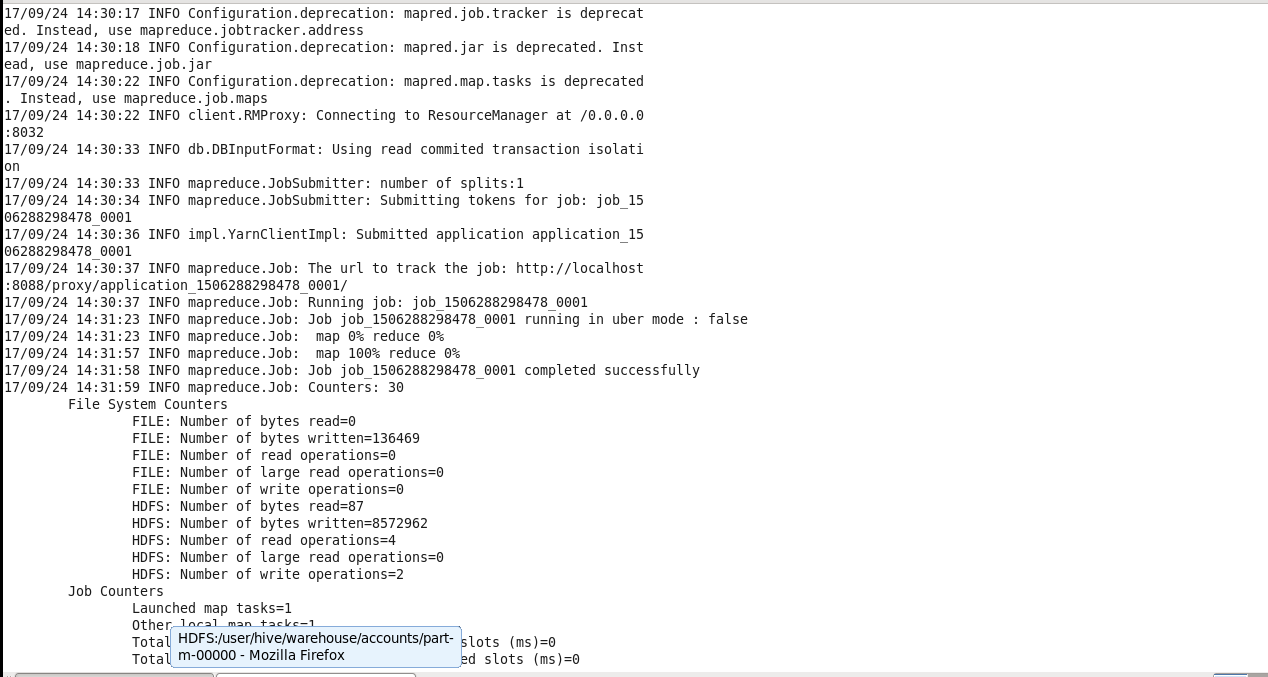


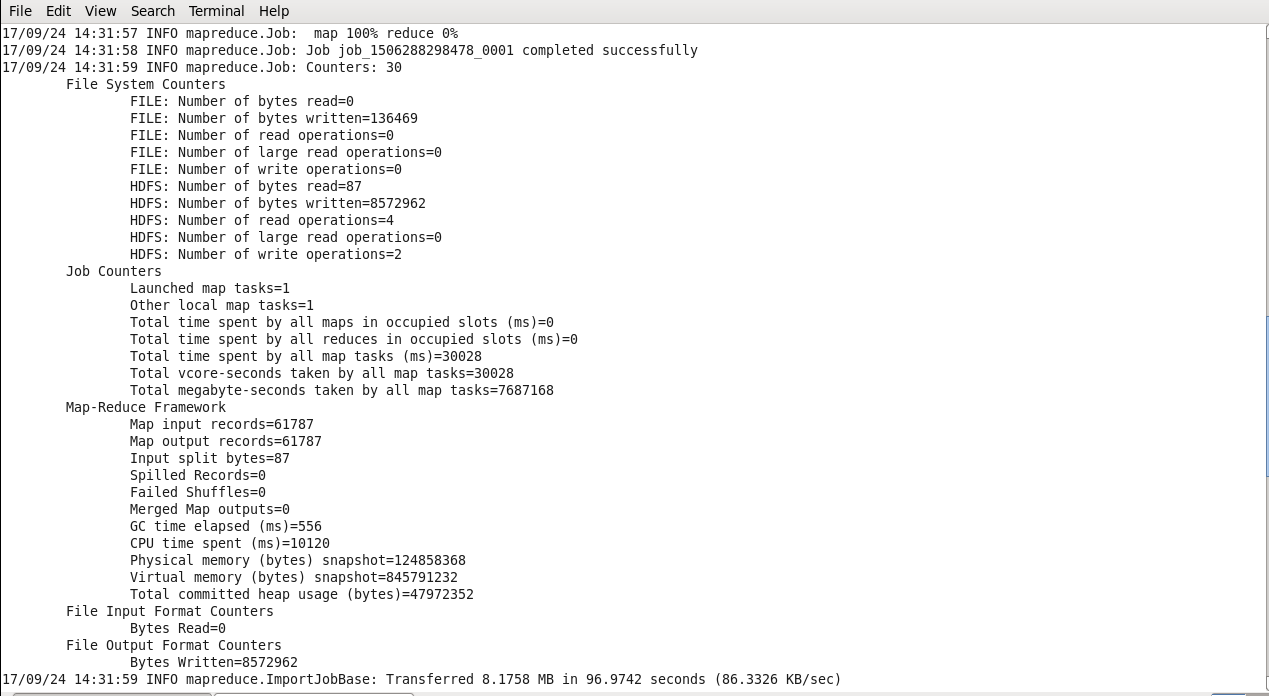
Import command loads 129761 records into the specified location(/mmm160430/loudacre/accounts).

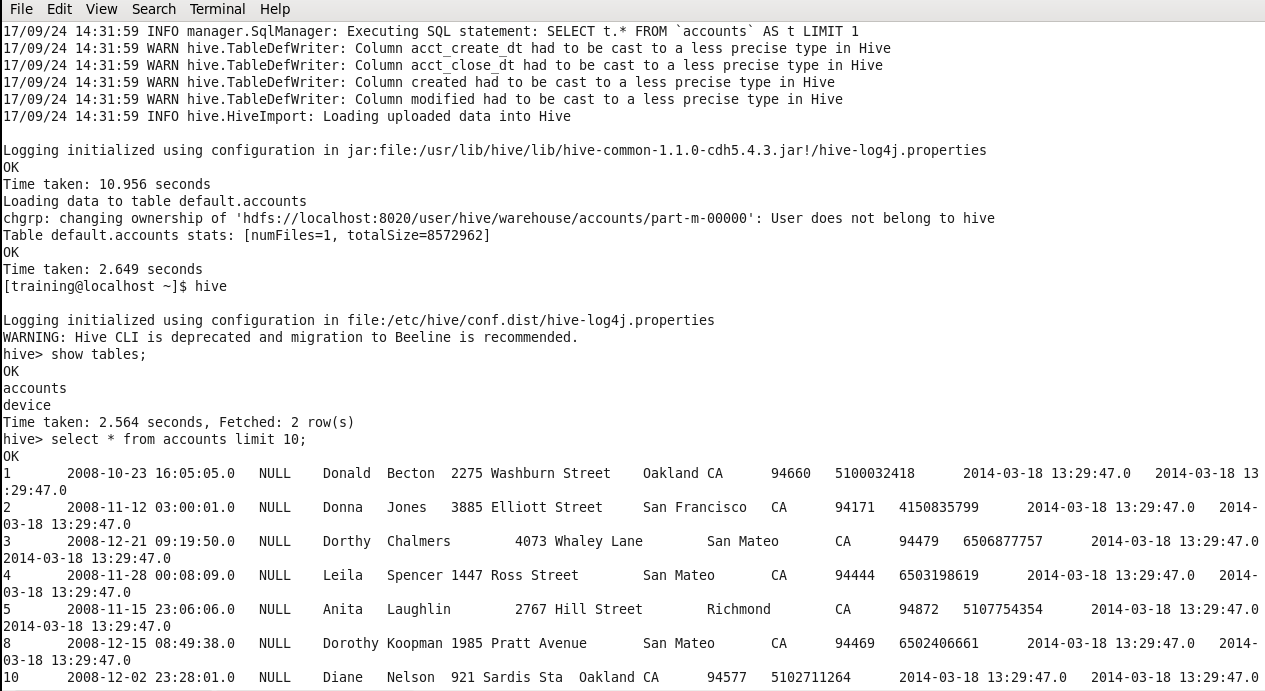
2. Import table ‘accounts’ into hive with the following conditions: a. Nulls are represented as /N b. Fields terminated by ‘,’ c. Accounts that are active and from the state of California(CA)

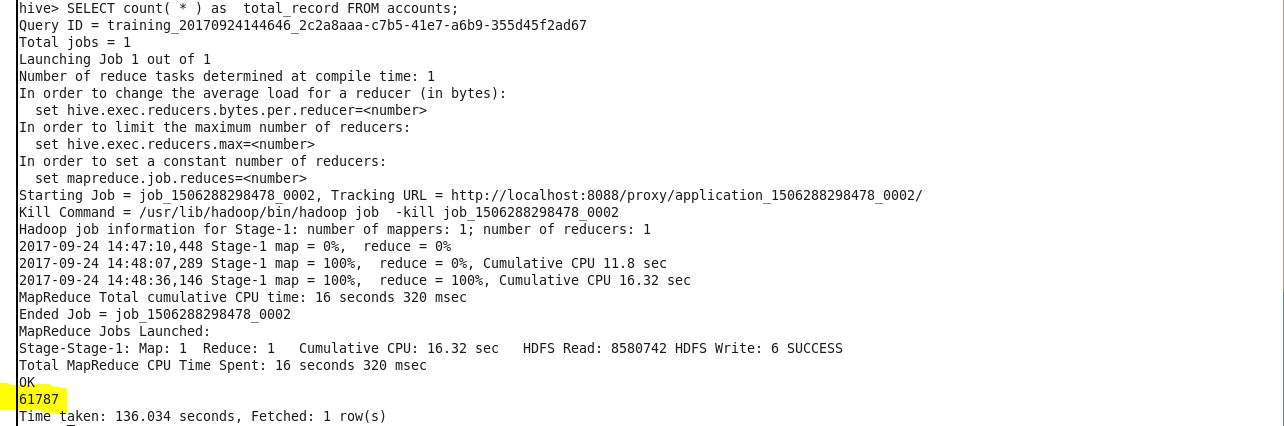
**Importing table into hive with the criterias mentioned above.**







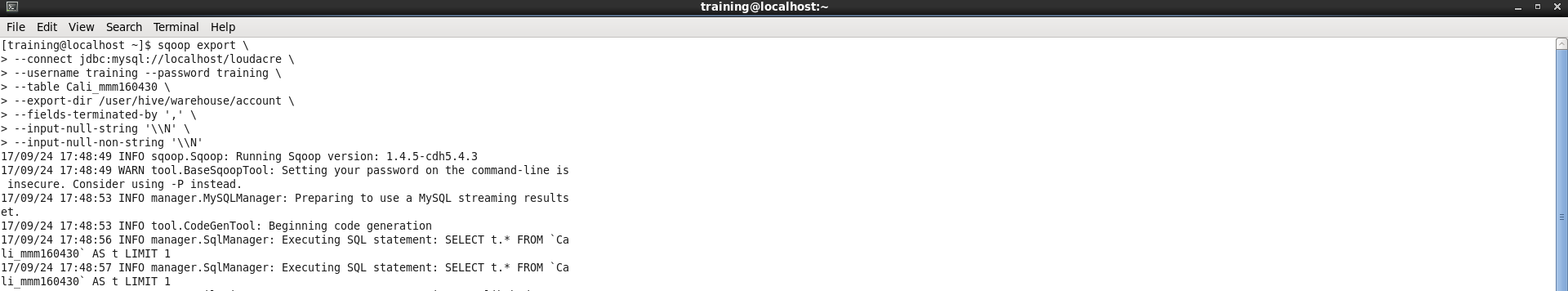


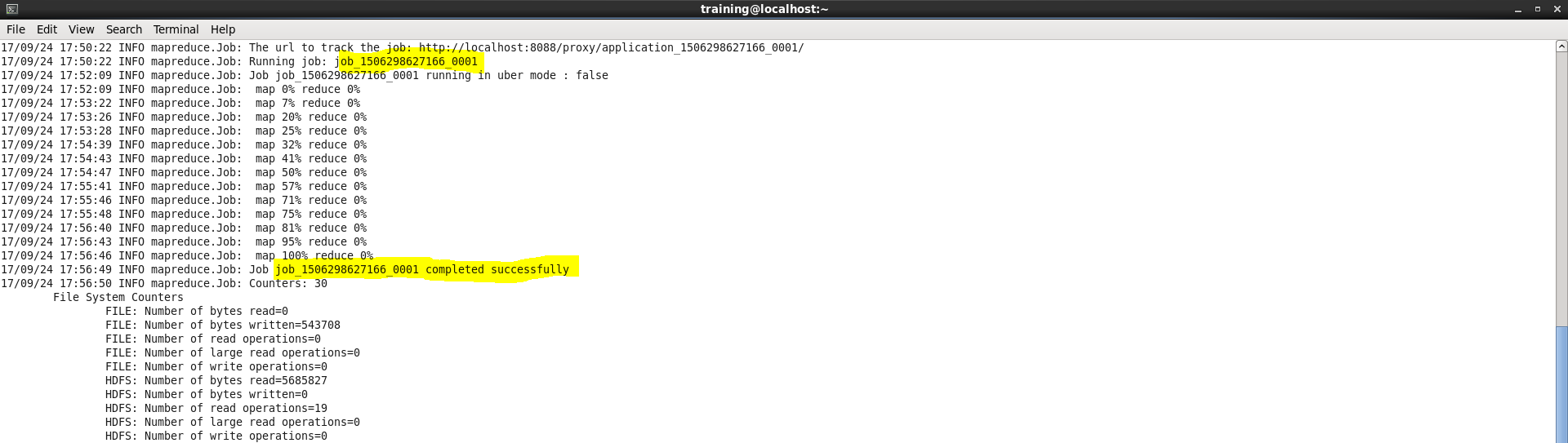


Total **61,787** records match the criteria mentioned.

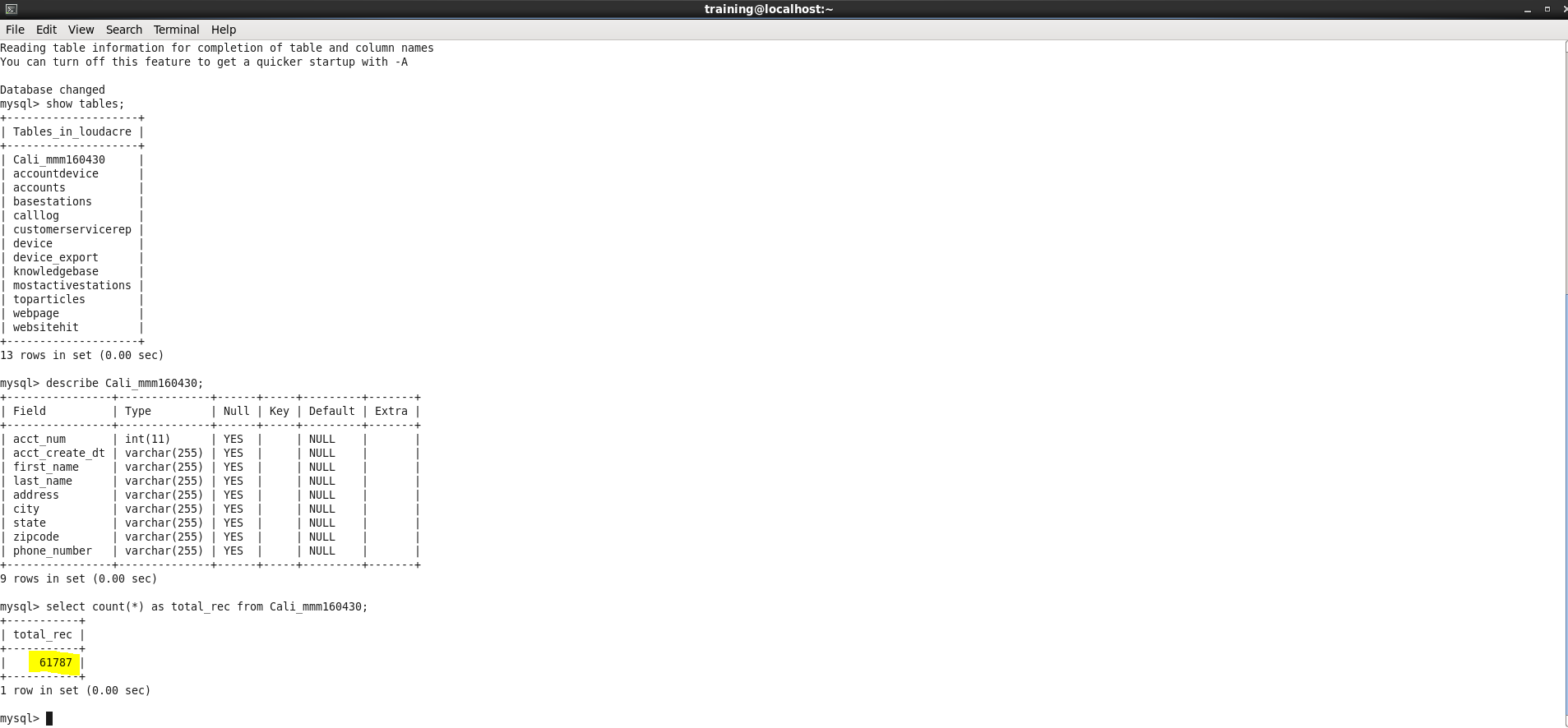
3. Export the contents from hive into a new MySQL table called ‘Cali-‘ followed by your NET-ID a. Do not import the following columns: i. acct\_close\_dt ii. modified iii. Created

Exporting from table account in hive to mysql table Cali\_mmm160430:





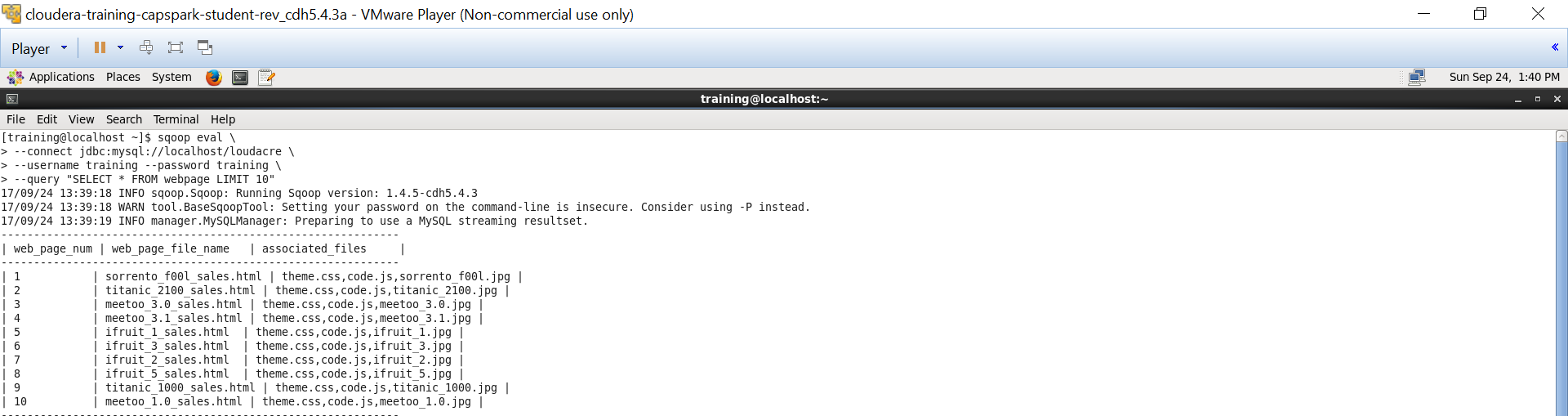
**Checking mysql databse loudacre for table Cali\_mmm16430**



It has expected number of records (61,787).

**PART B**

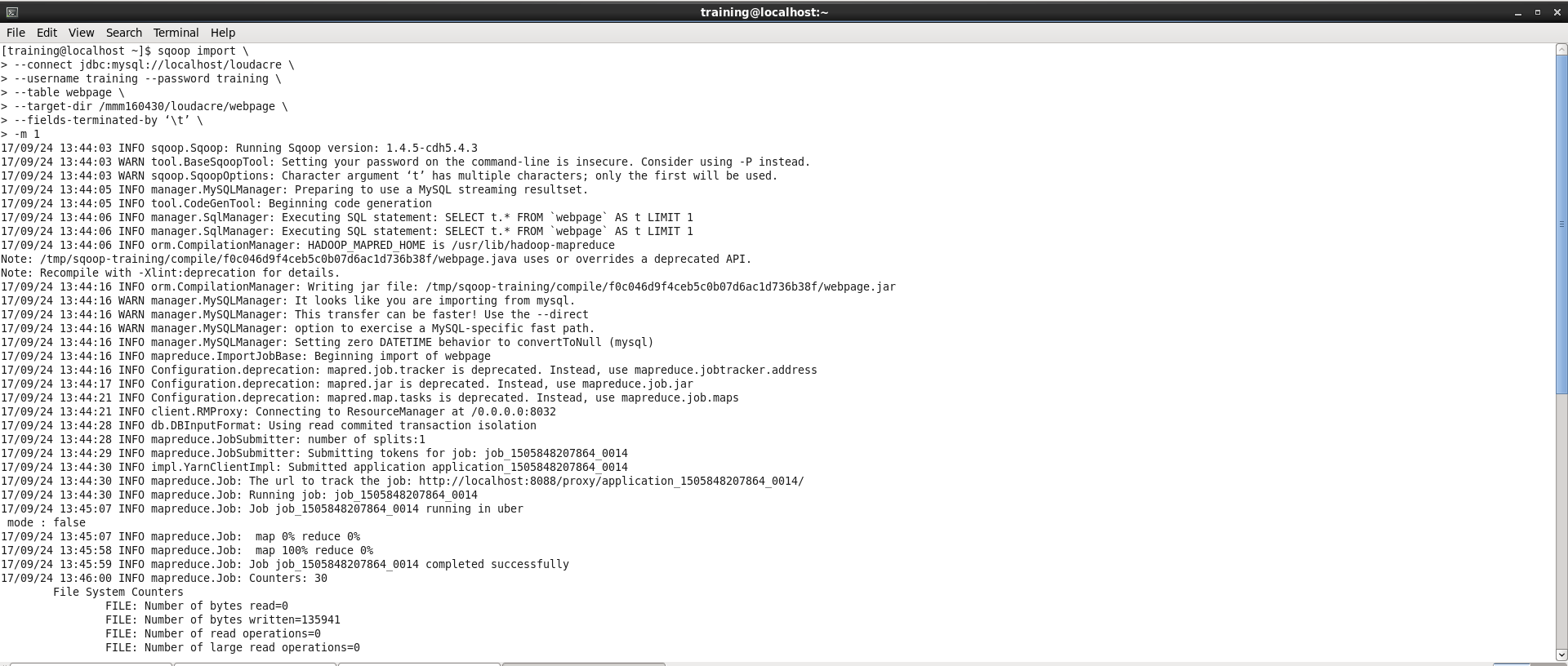
1. Check 10 rows of content on table ‘webpage’ without importing the table.

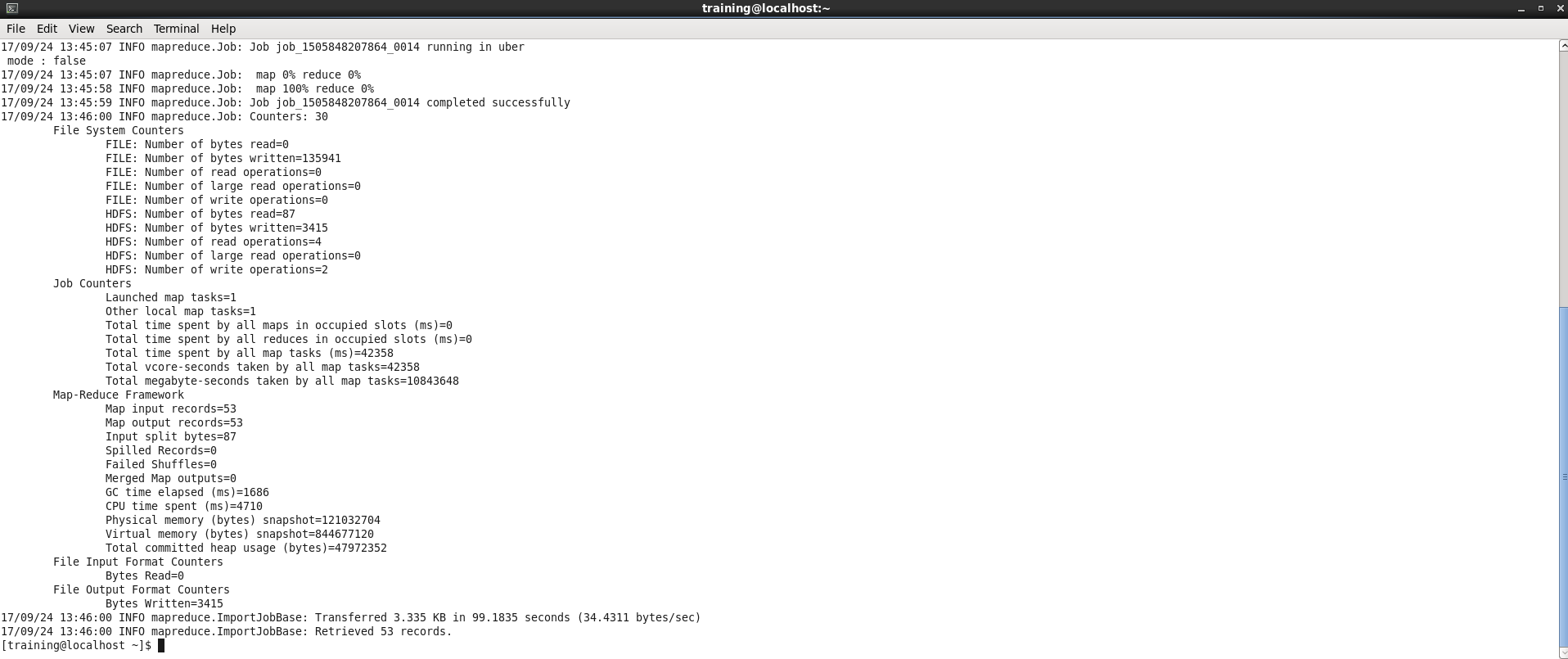


This command fetches the first 10 rows of the table webpage in loudacre database without importing the table.

2. Import table ‘webpage’ into a directory called Your NET-ID/loudacre/webpage a. Fields should be terminated by tab.

The result is split into 2 screenshots below:





After the execution is complete, it retrieves 53 records from table webpage.