**Big Data Analytics - Assignment 1**

- Darshil Sanghvi

NetID: drs170530

**Question 1-**

**Commands:**

[training@localhost ~]$ sqoop import \

> --connect jdbc:mysql://localhost/loudacre \

> --username training –password training \

> --target-dir /drs170530/loudacre/pset1 \

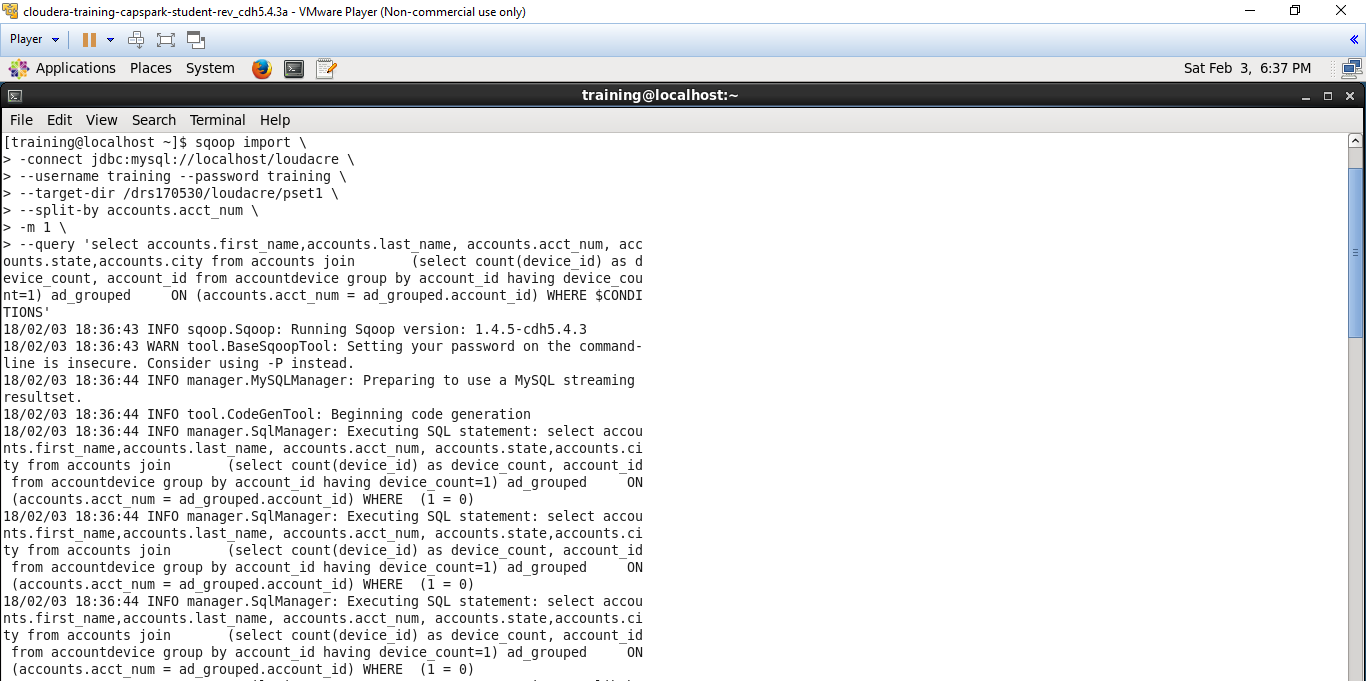
> --split-by accounts.acct\_num \

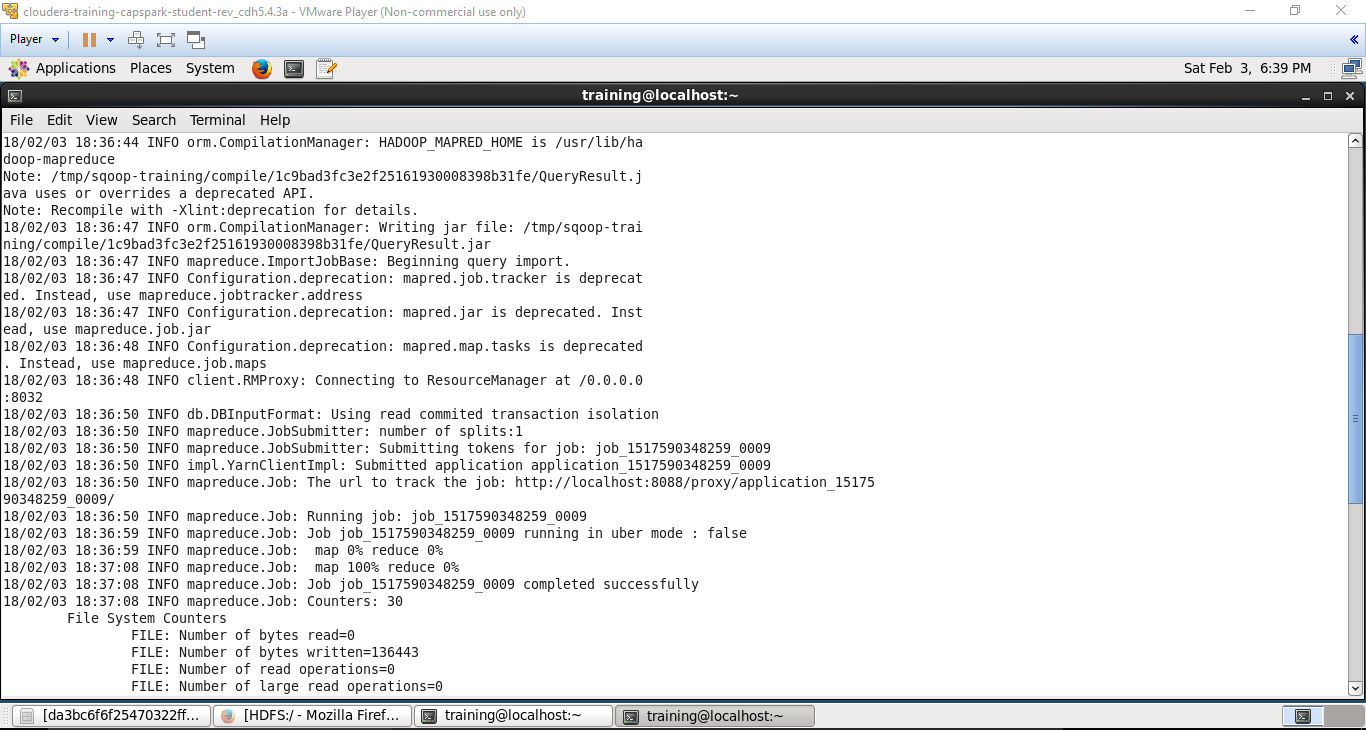
> -m 1 \

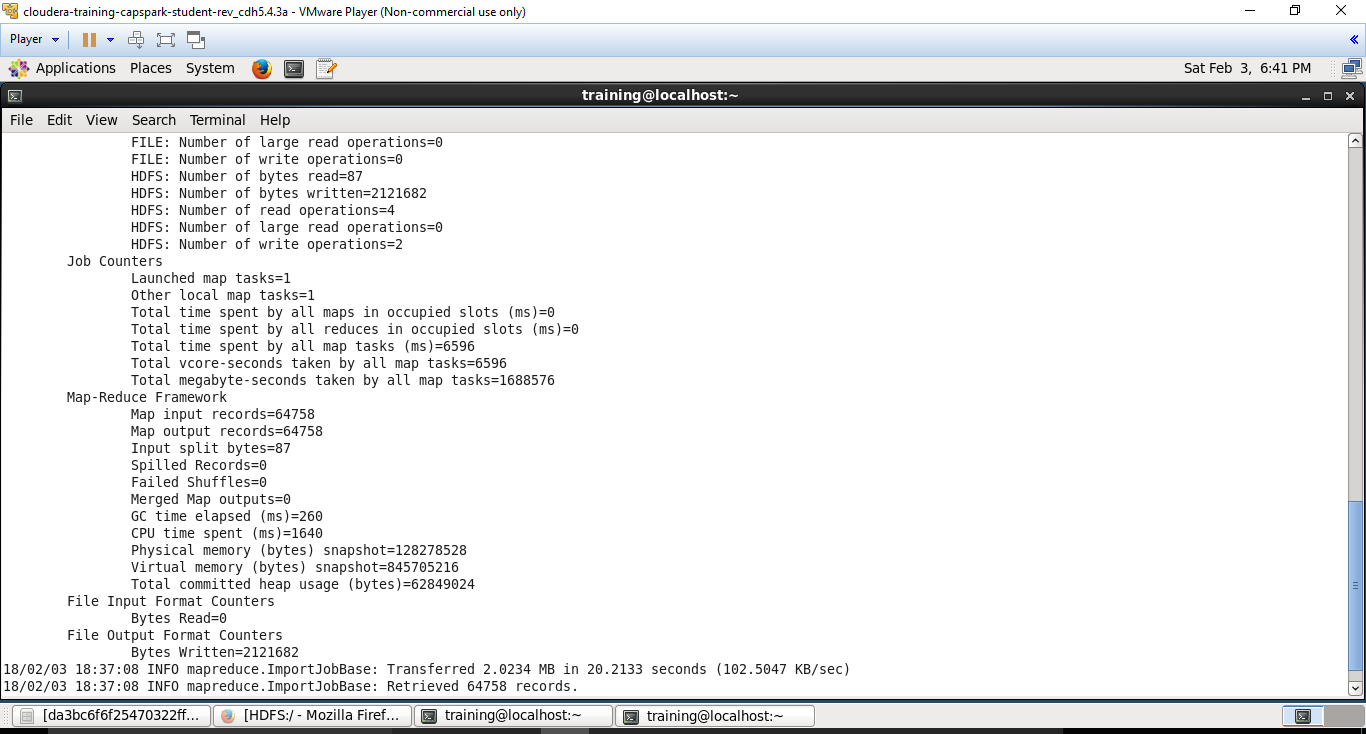
>-- query ‘select accounts.first\_name, accounts.last\_name, accounts.acct\_num, accounts.state, waccounts.city from accounts join (select count(device\_id) as device\_count, account\_id from accountdevice group by account\_id having device\_count=1) ad\_grouped ON

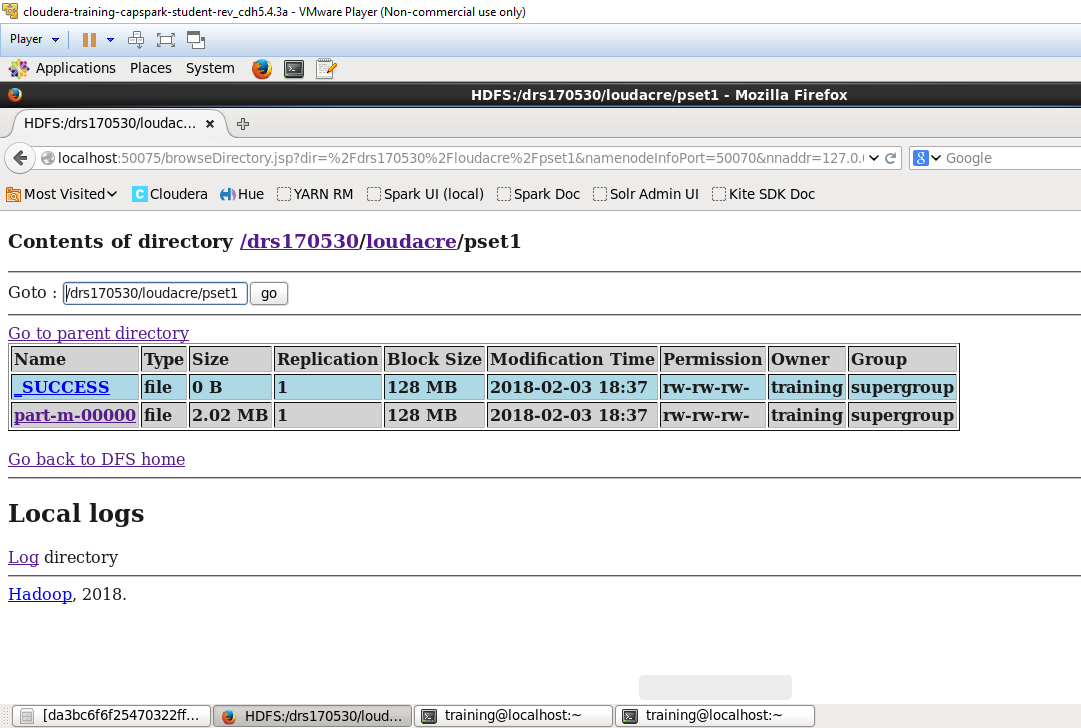
( accounts.acct\_num = ad\_grouped.account\_id ) WHERE $CONDITIONS’

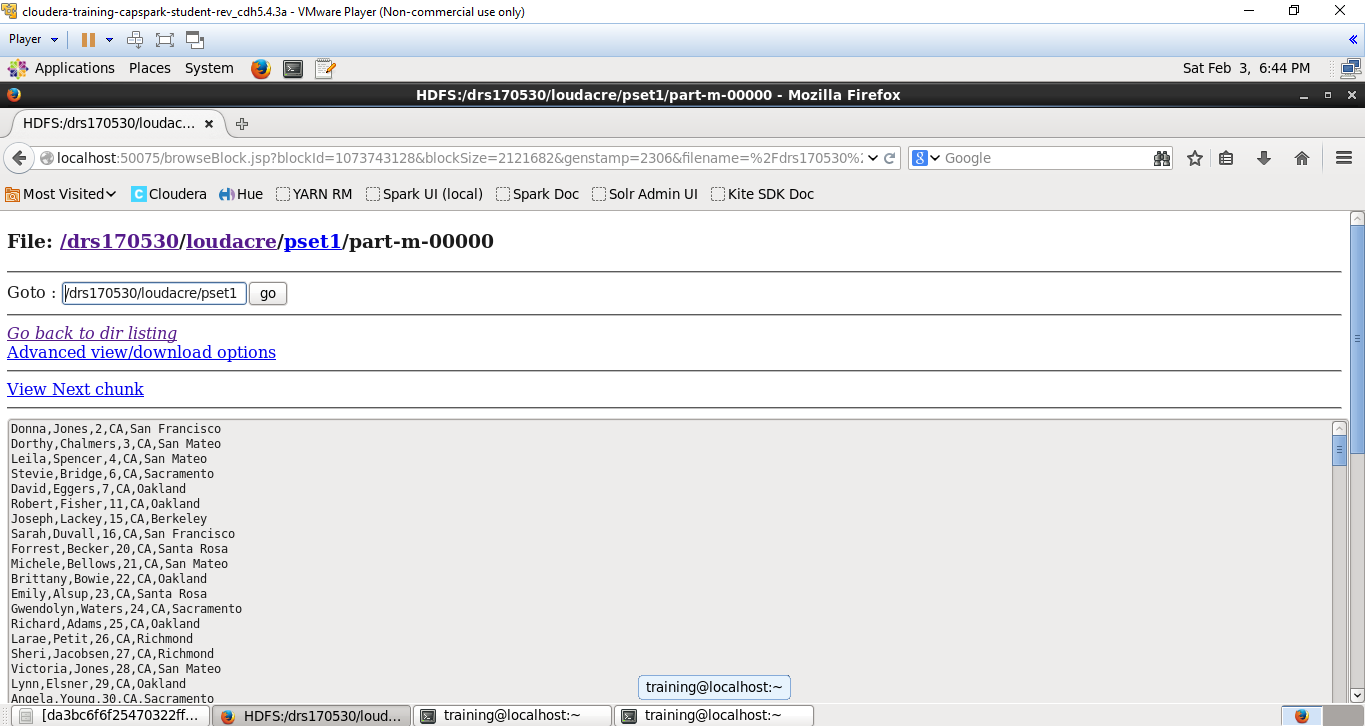
**Snapshots:**

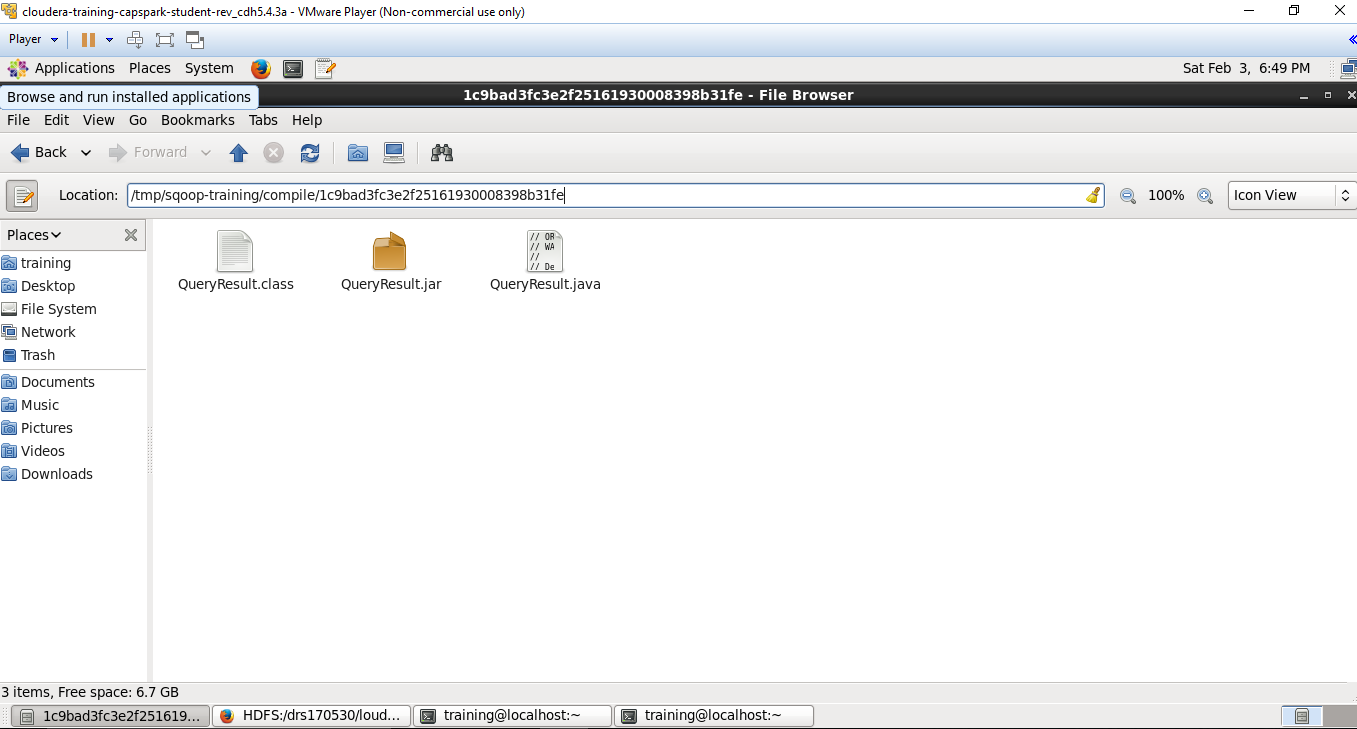
****











-File browser consisting of class and jar files.

--------------------------------End of Question 1 --------------------------------------------------------------------

**Question 2:**

**Commands:**

[training@localhost ~]$ sqoop import \

> --connect jdbc:mysql://localhost/loudacre \

> --username training –password training \

> --target-dir /drs170530/loudacre/pset2 \

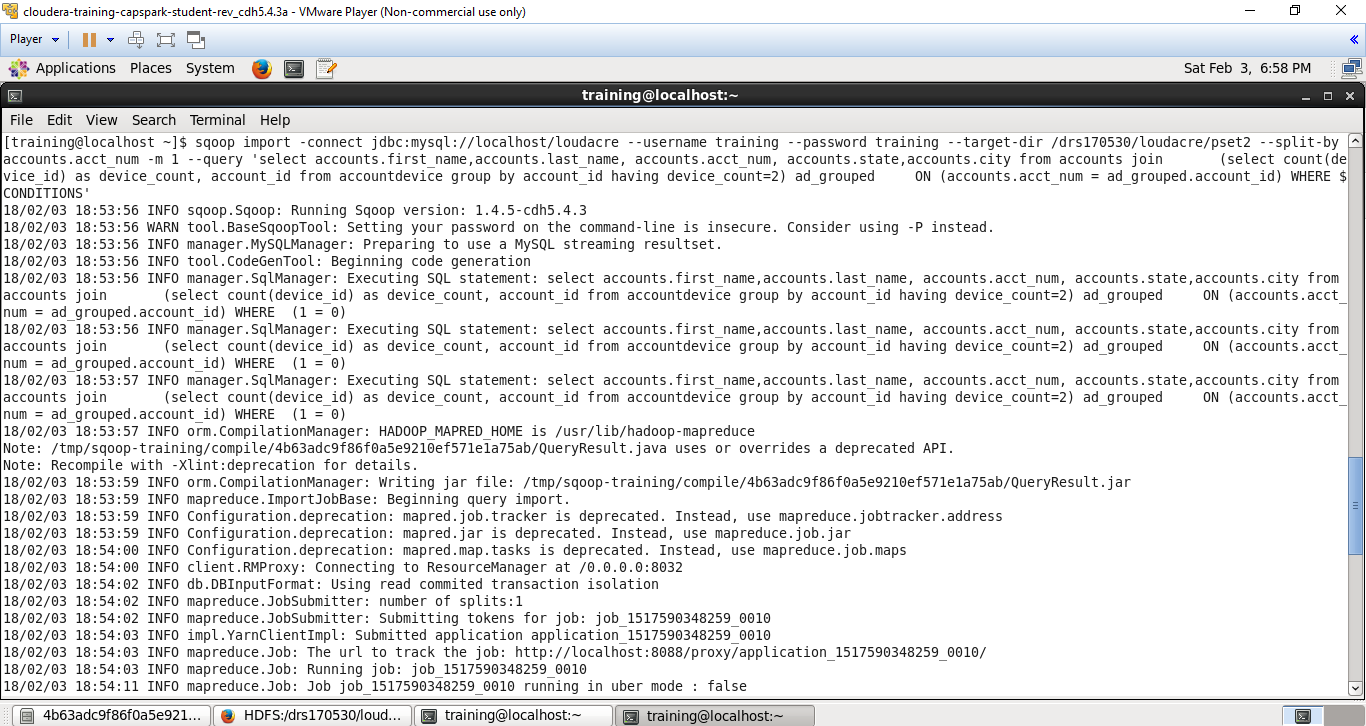
> --split-by accounts.acct\_num \

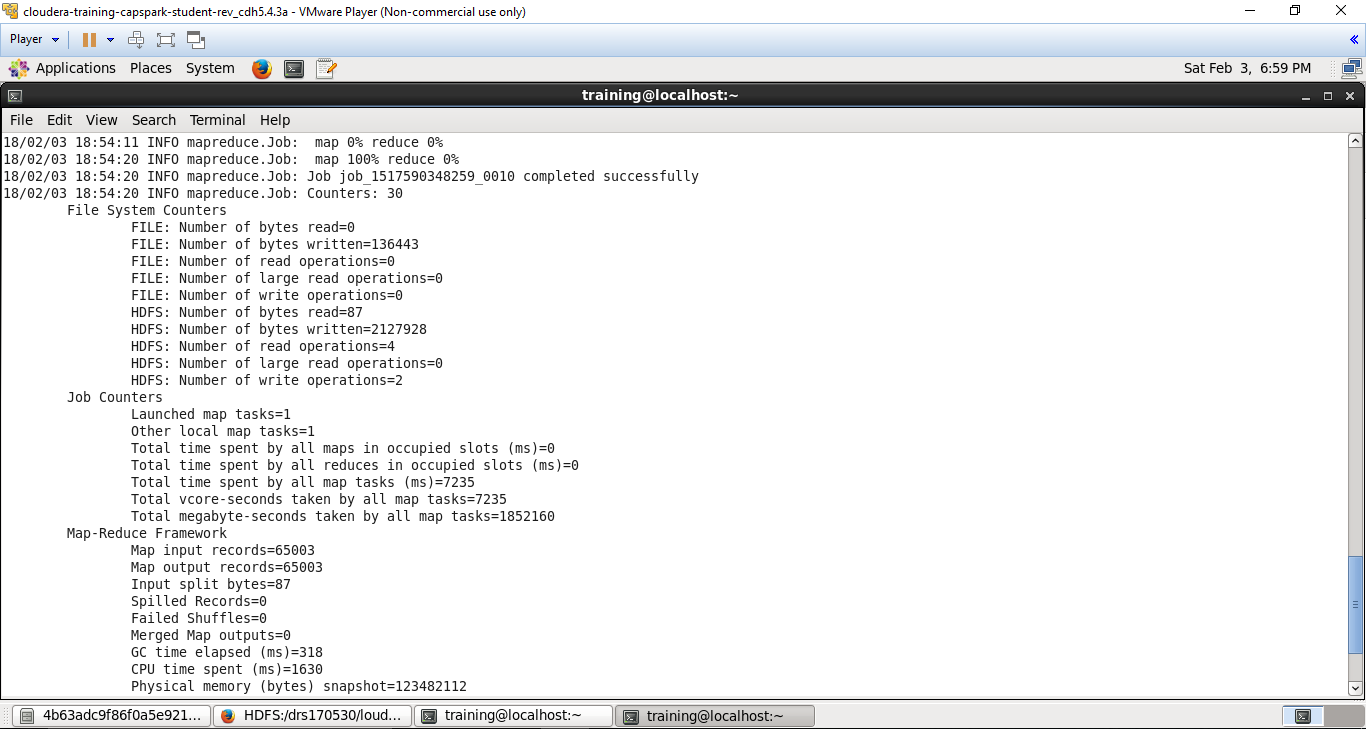
> -m 1 \

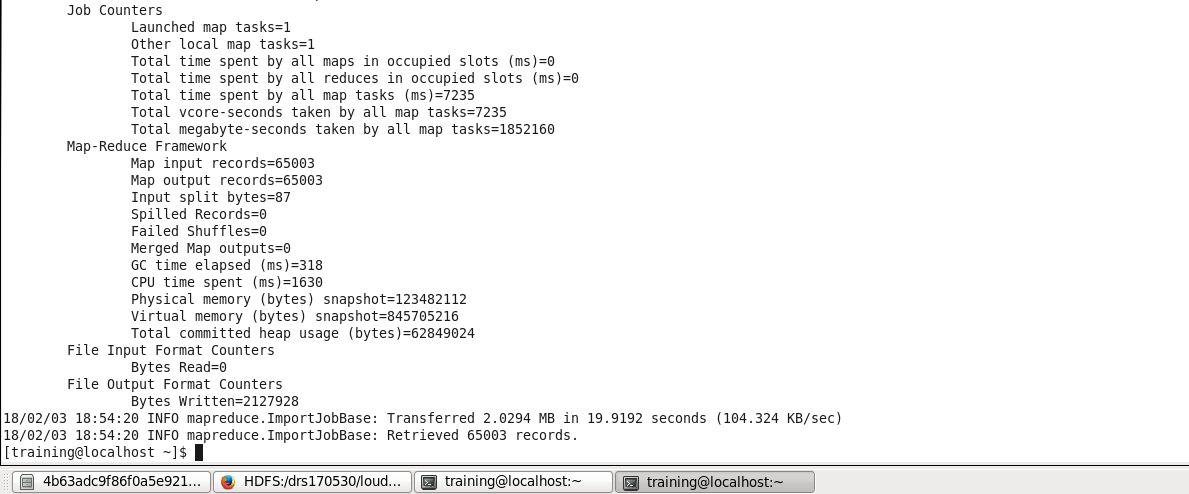
>-- query ‘select accounts.first\_name, accounts.last\_name, accounts.acct\_num, accounts.state, waccounts.city from accounts join (select count(device\_id) as device\_count, account\_id from accountdevice group by account\_id having device\_count=2) ad\_grouped ON

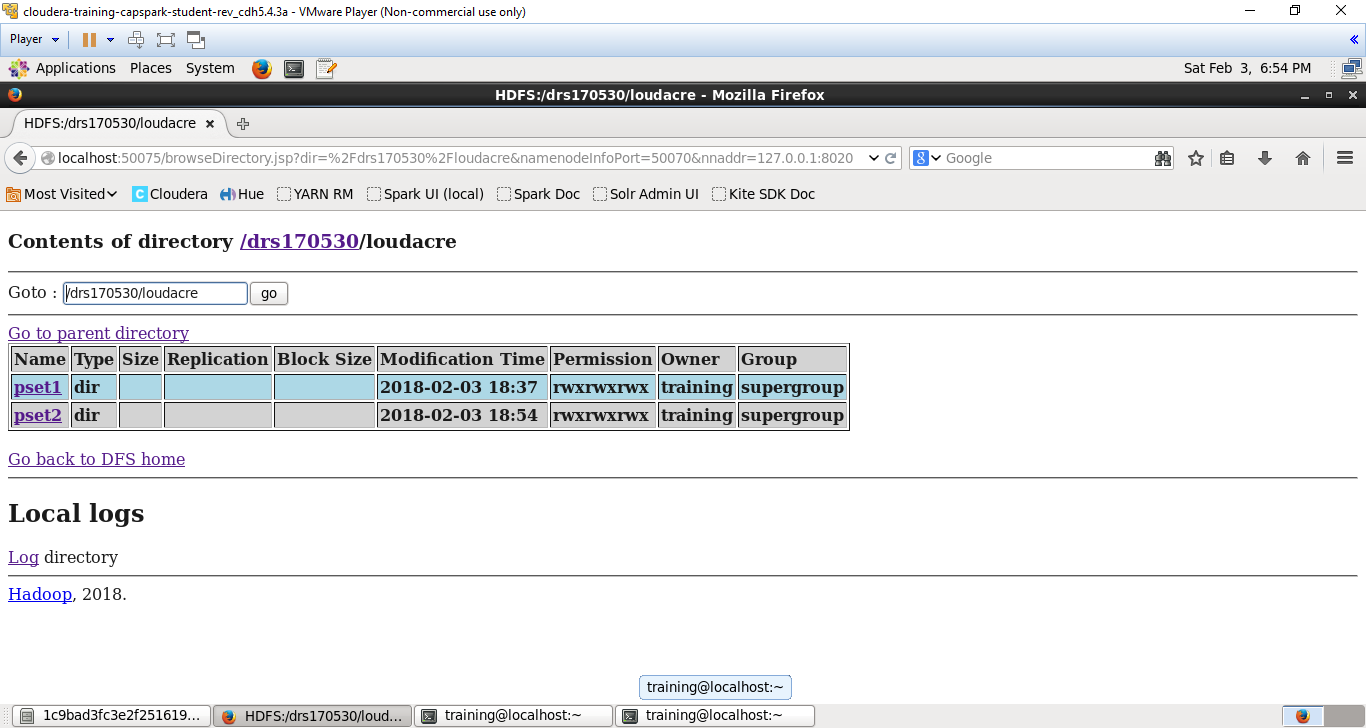
( accounts.acct\_num = ad\_grouped.account\_id ) WHERE $CONDITIONS’

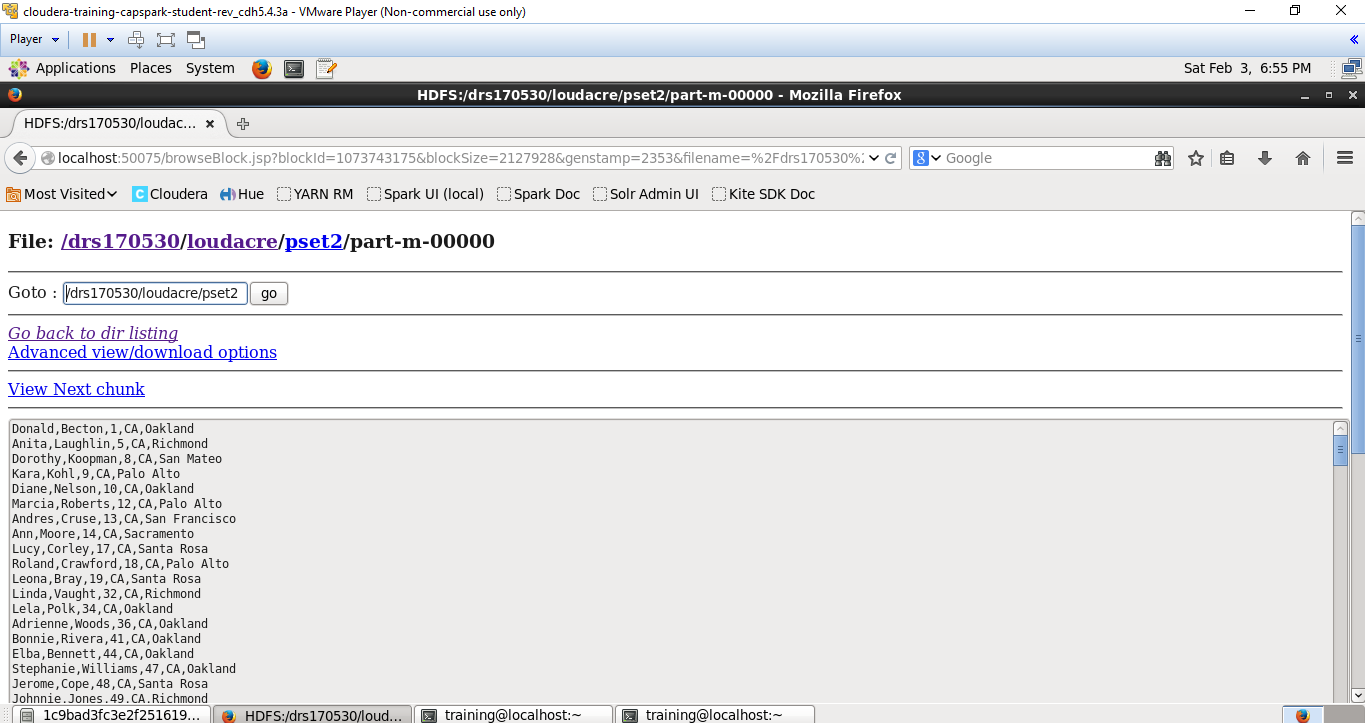
**Snapshots:**



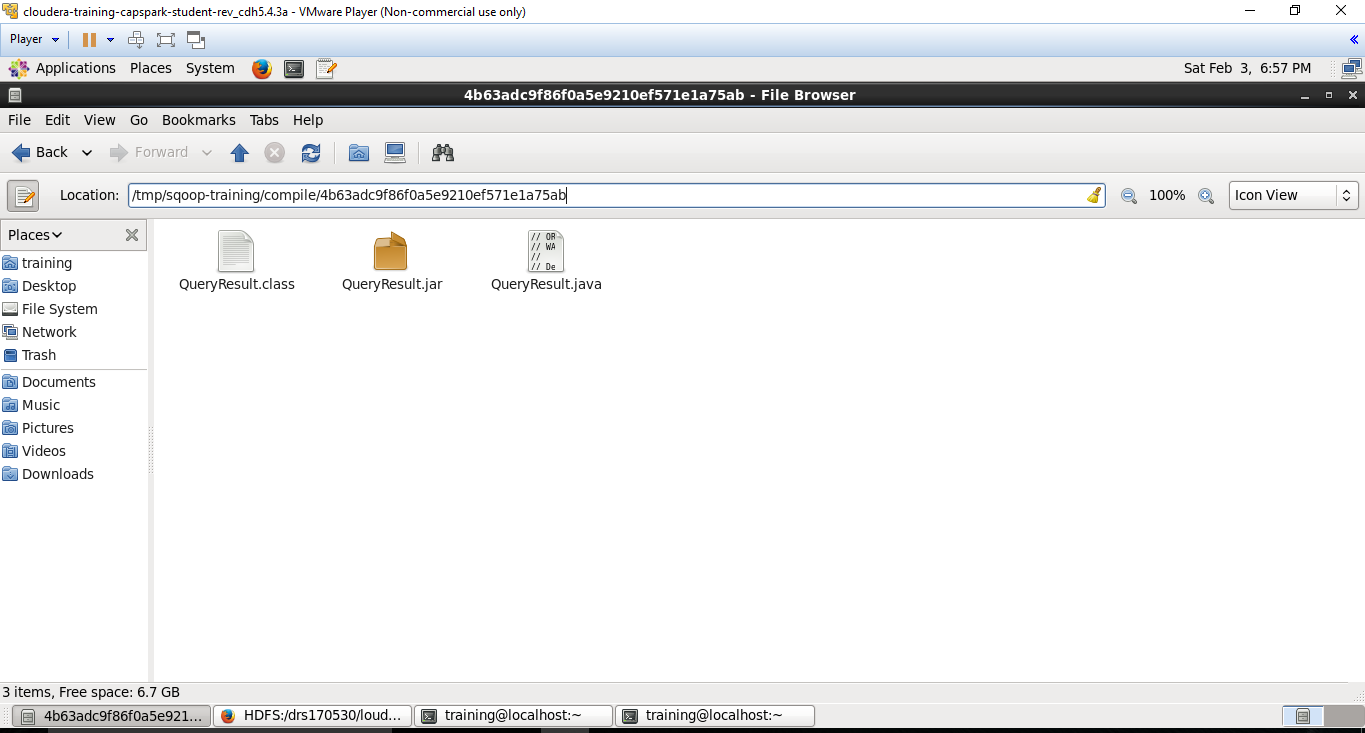








Jar and Class file creation in File browser:



---------------------------End of Question 2 ----------------------------------------

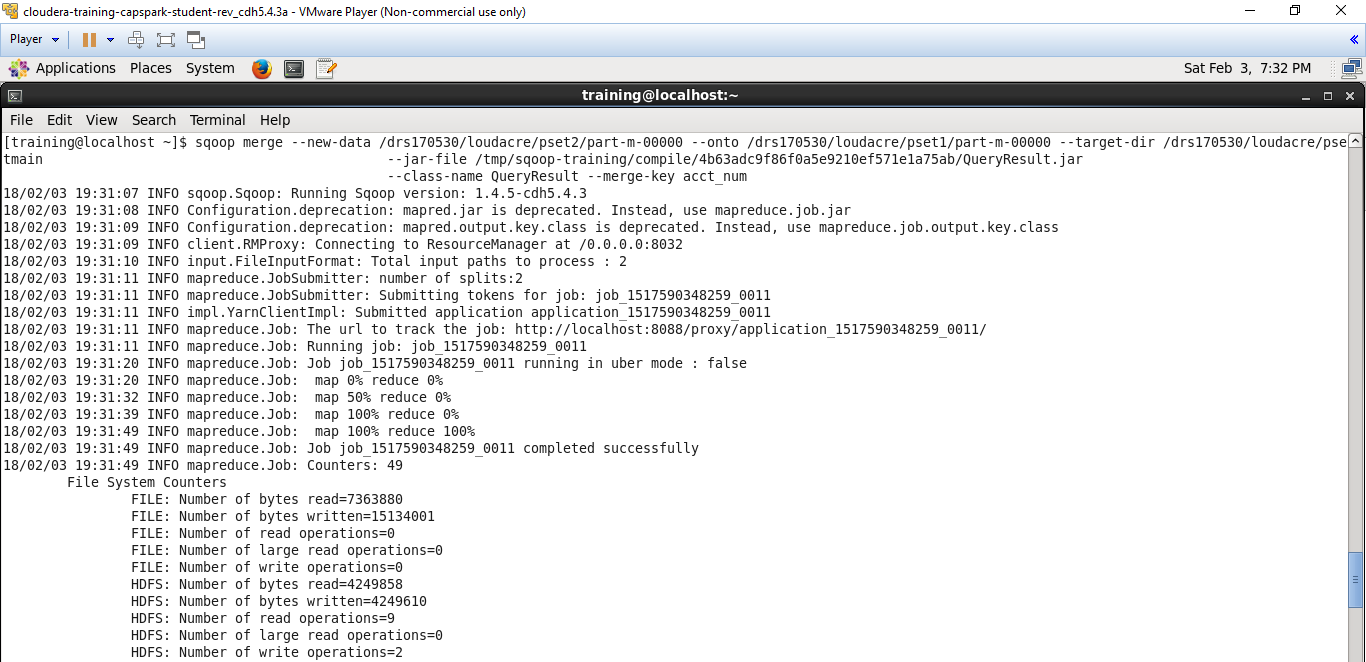
**Question 3:**

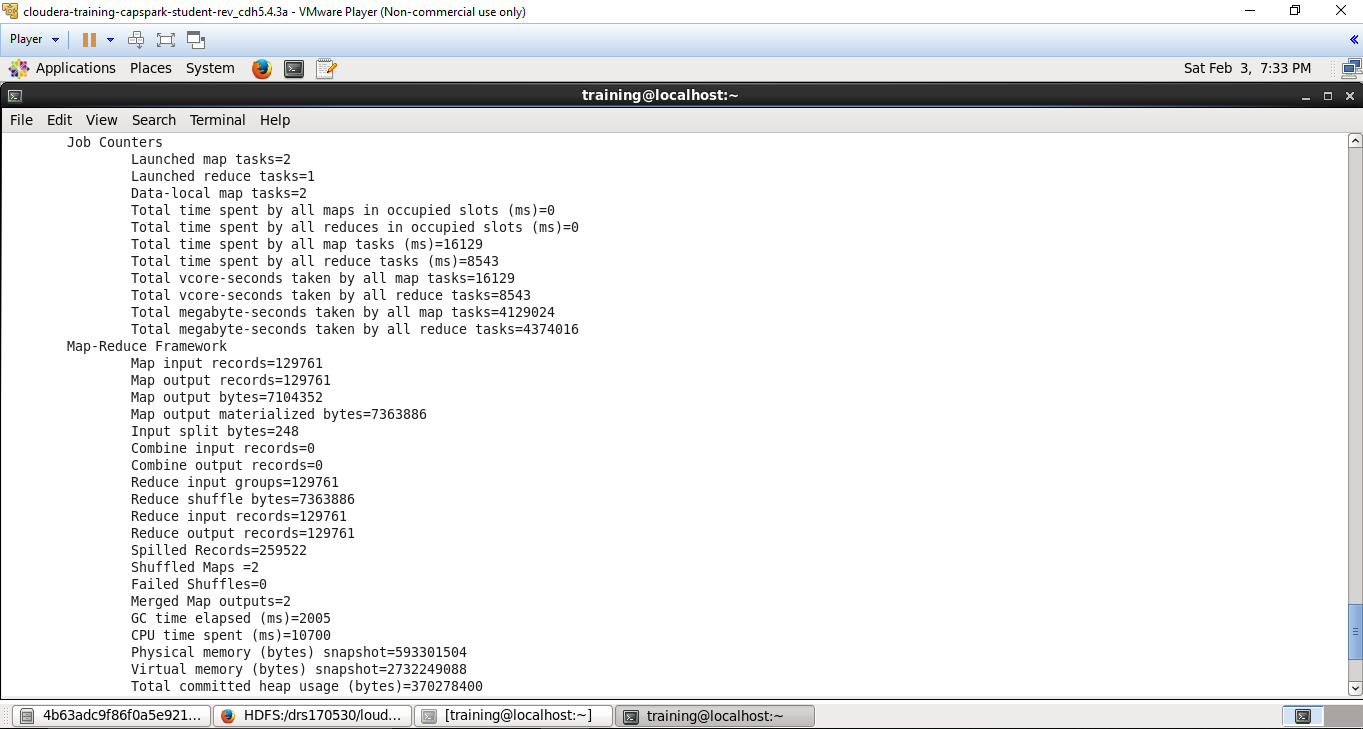
**Commands: sqoop merge –new-data /drs170530/loudacre/pset2/part-m-00000**

**--onto /drs170530/loudacre/pset1/part-m-00000 --target-dir /drs170530/loudacre/psetmain**

**--jar-file /tmp/sqoop-training/compile/4b63adc9f86f0a5e9210ef571e1a75ab/QueryResult.jar**

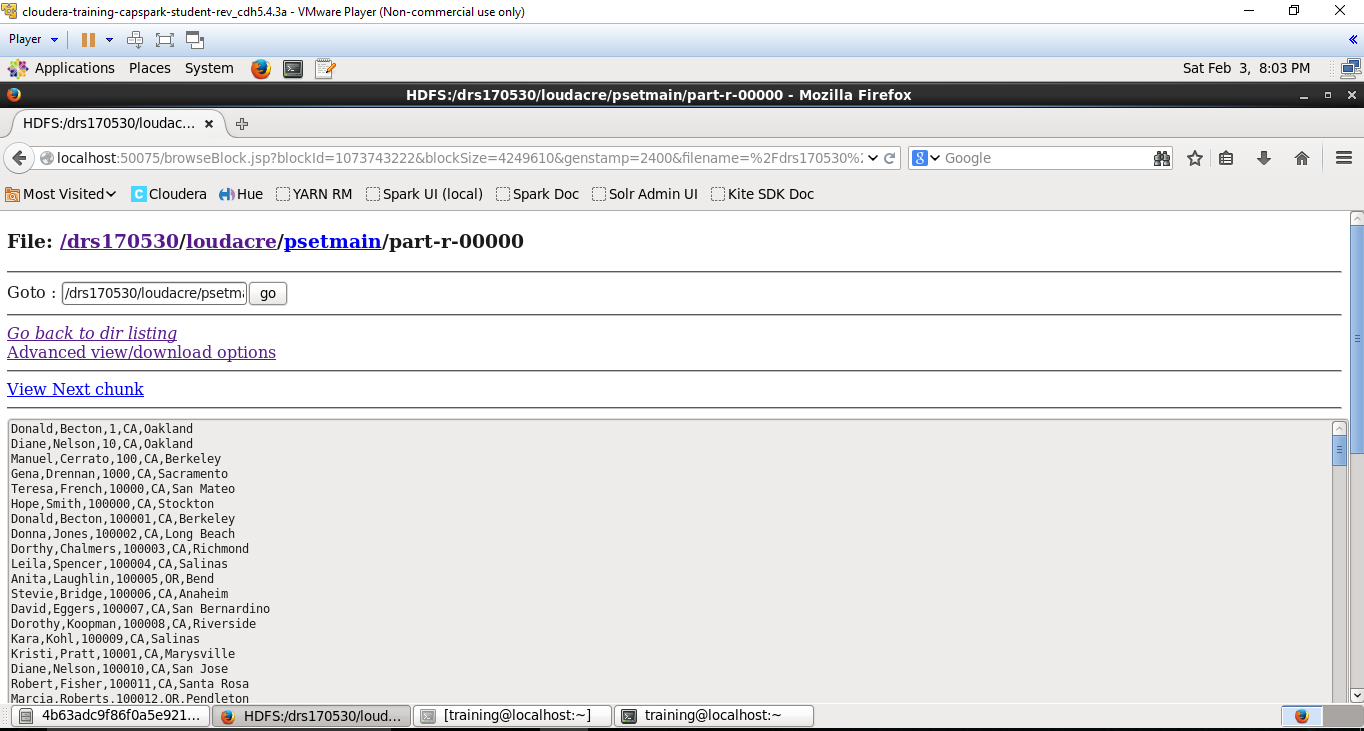
**--class-name QueryResult --merge-key acct\_num**







Data merged into a new table psetmain:



-------------------------------- End of Question 3------------------------------------------------------------

**Question 4:**

**Commands:**

Step 1: Sql command to create a new table in mysql

mysql>use loudacre;

mysql> create table MergedAccts( first\_name varchar(255),last\_name varchar(255), acct\_num int(11),state varchar(255), city varchar(255));

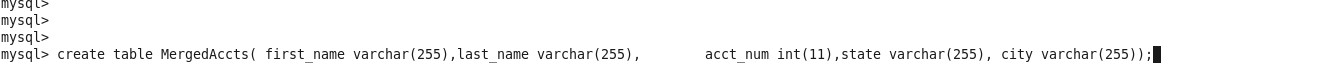
Step 2: sqoop command to export data from target directory of psetmain to table MargedAccts in loudacre database of mysql.

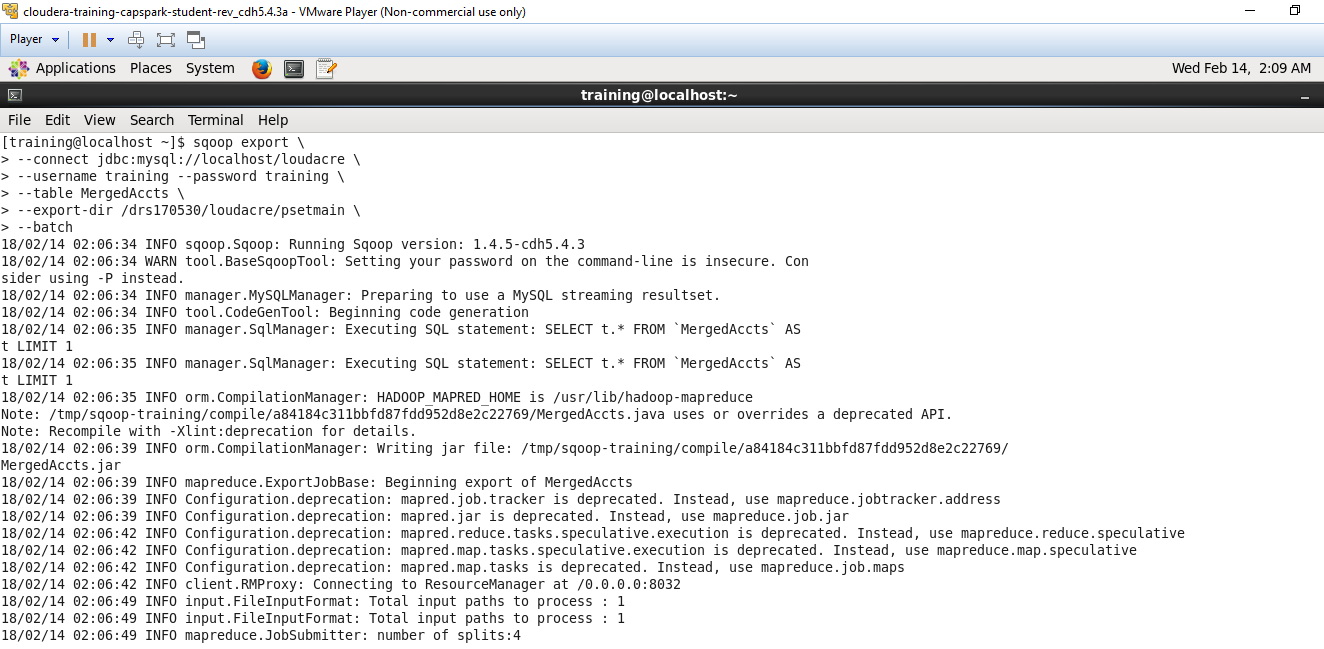
sqoop export \ --connect jdbc:mysql://localhost/loudacre \ --username training \ --password training \ --table MergedAccts \ --export-dir /drs170530/loudacre/psetmain \ --batch

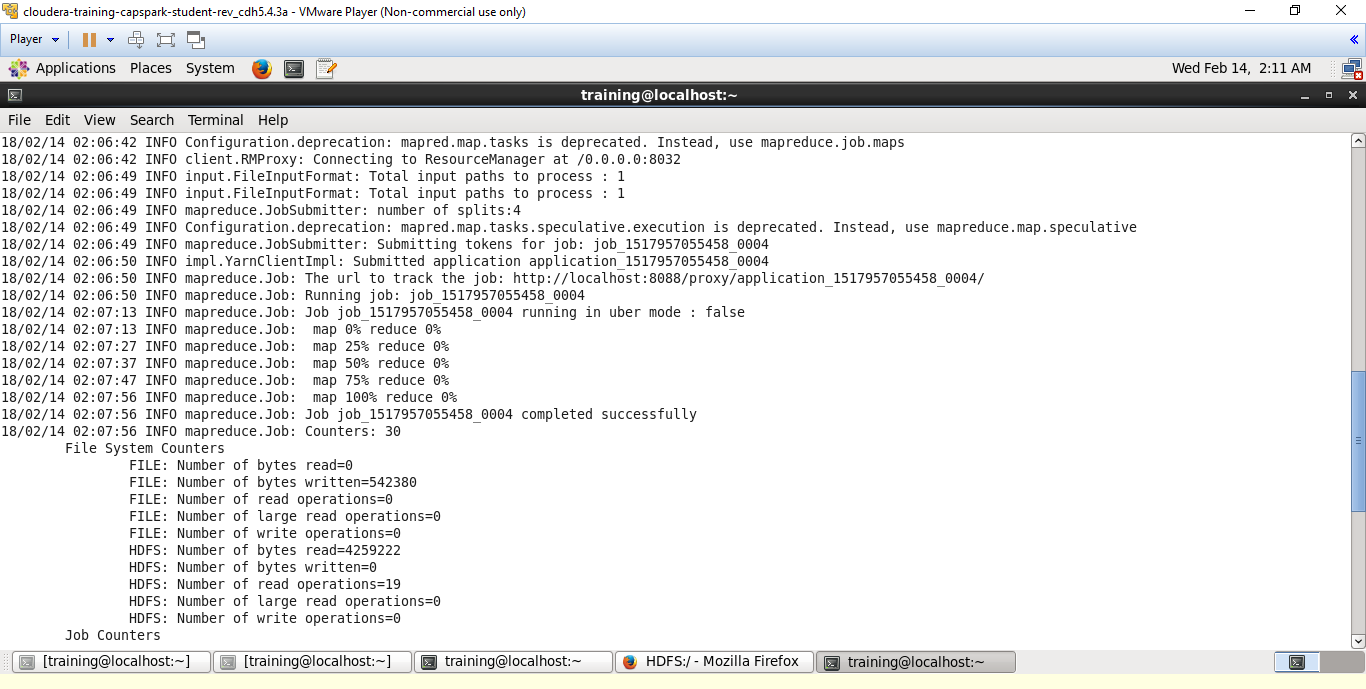
Step 3: sql command to display the newly exported records to table MergedAccts

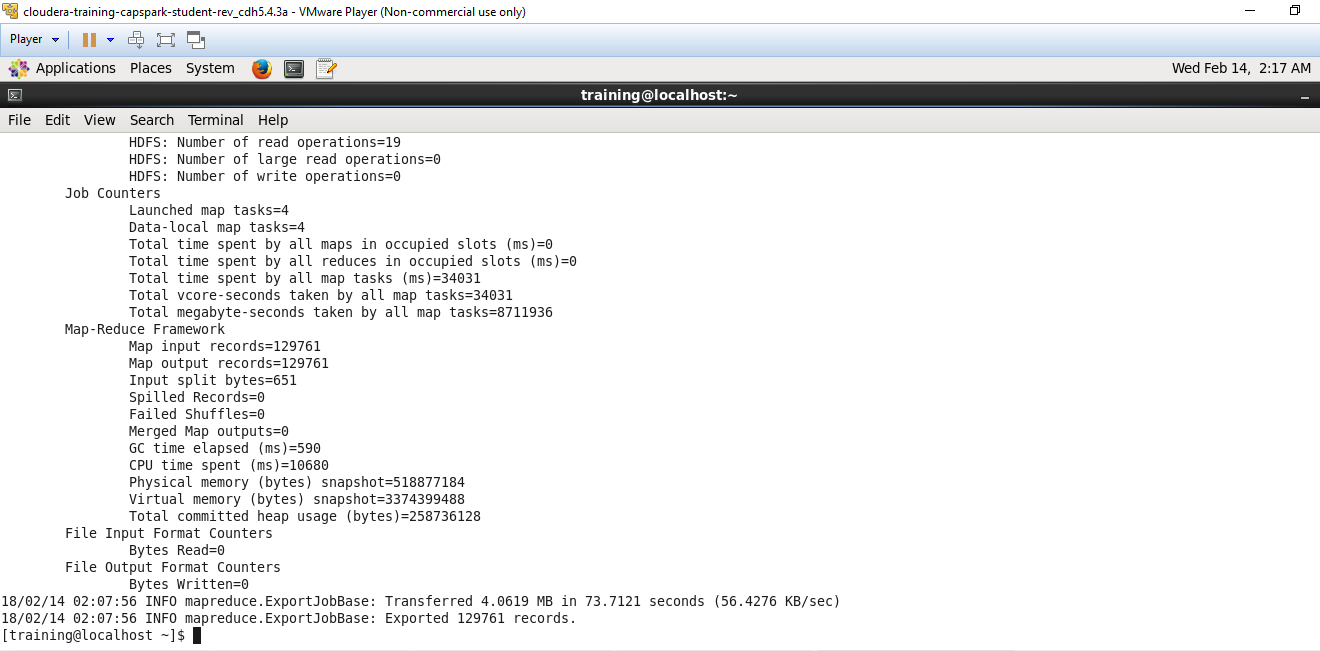
mysql> select \* from MergedAccts;

**Snapshots:**

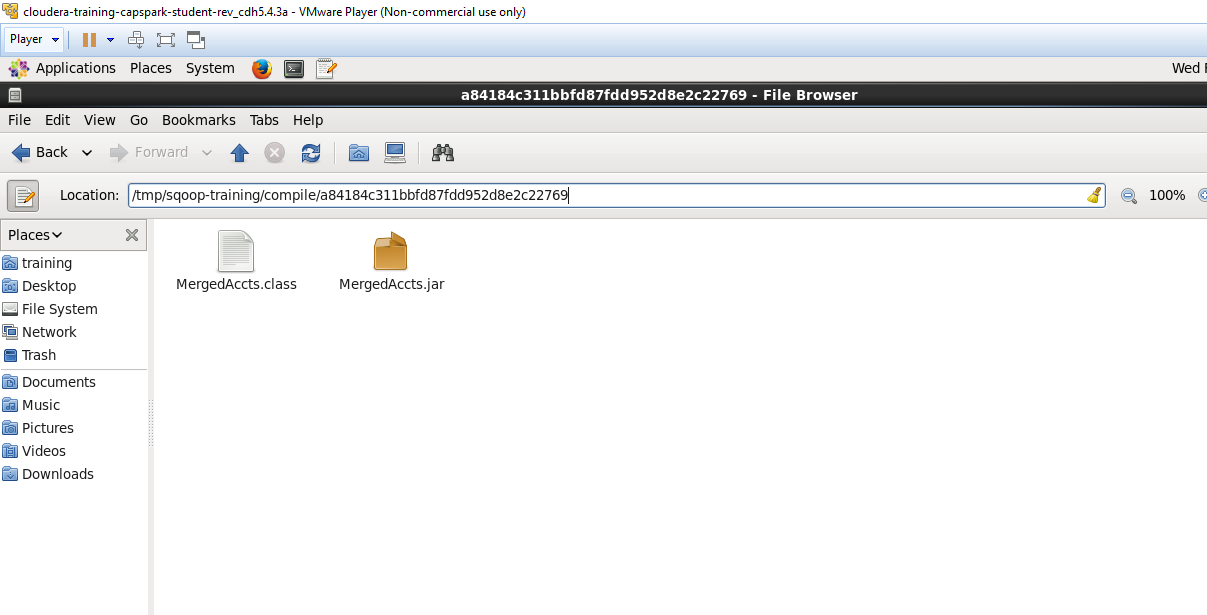




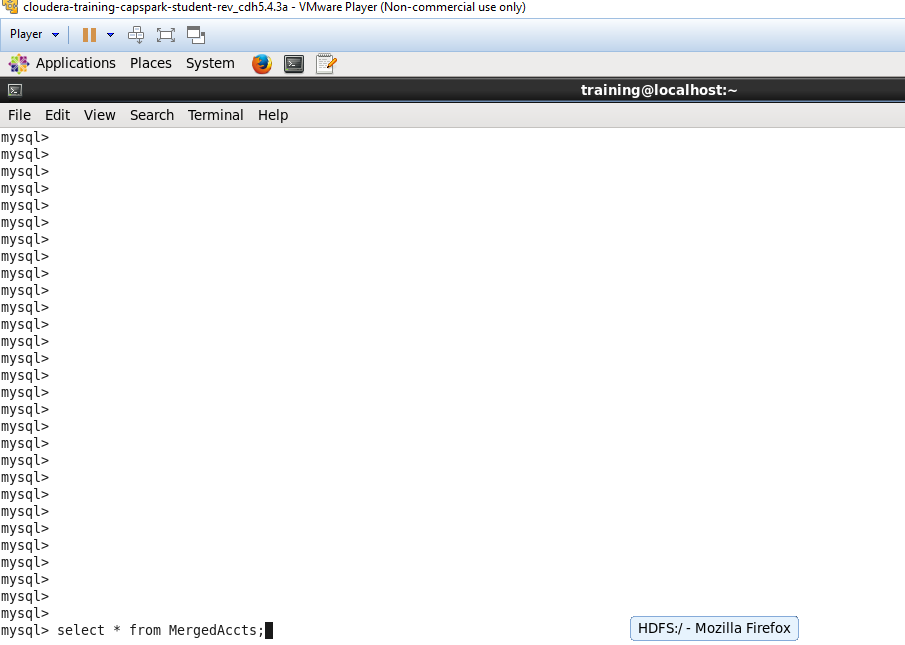




Jar and class files of new export command in file browser:



Querying the newly populated table:





-------------------------------------------------End of Question 4 -----------------------------------------------------