Contents

Assignment 3- Hive & Sqoop 1

Objective: Use Sqoop to transfer data to MySQL from Hadoop or vice versa 1

Import a table from MySQL to HDFS (Choose the parameter as you wish) 1

2 Demonstrate export command. 6

Task 2: Week 6: Hive 8

Assignment 3- Hive & Sqoop

Objective: Use Sqoop to transfer data to MySQL from Hadoop or vice versa

Import a table from MySQL to HDFS (Choose the parameter as you wish)

Well, importing a table from **MYSQL** to **HDFS**, I had to start MYSQL with ***systemctl start mysqld*** then I logged with root ***(MySQL -uroot).*** After that I checked databases in mysql where I got hive databases which used for table. (Figure 1)

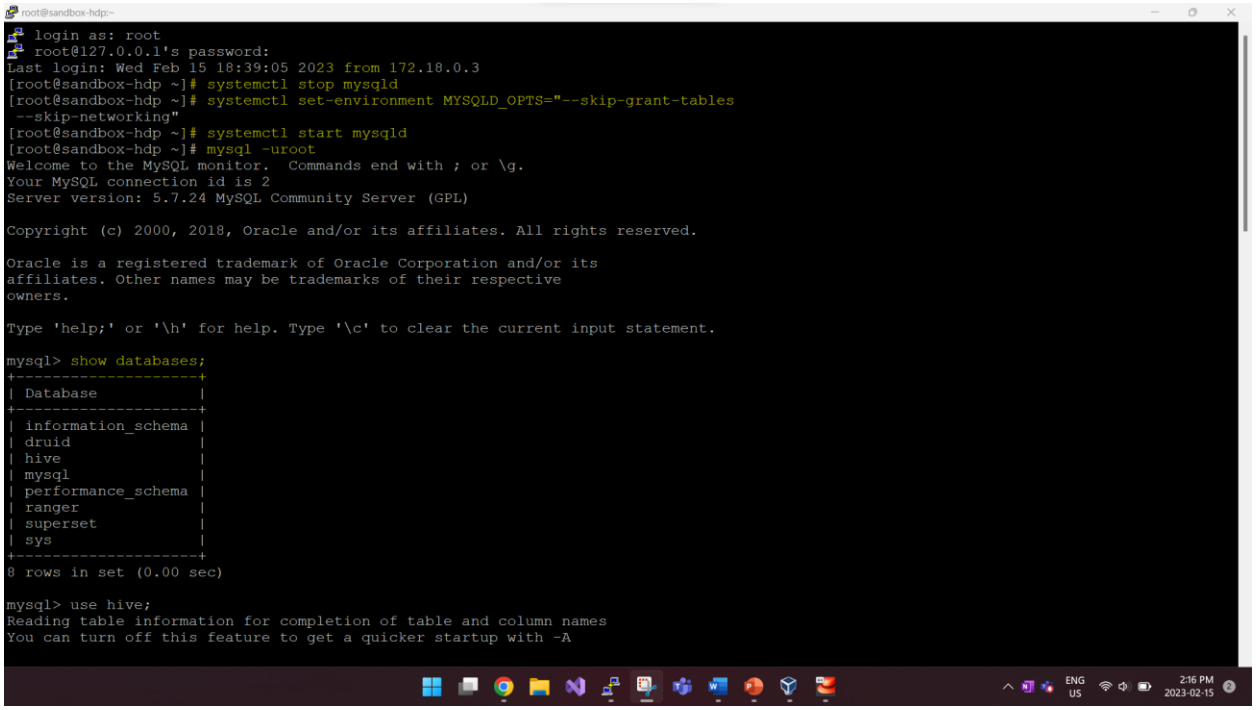


Figure 1 Start Mysql and logged with root user

I used **ROLES** table, and Firstly I noticed that what's in a table, used **select * from ROLES.** (Figure 4)

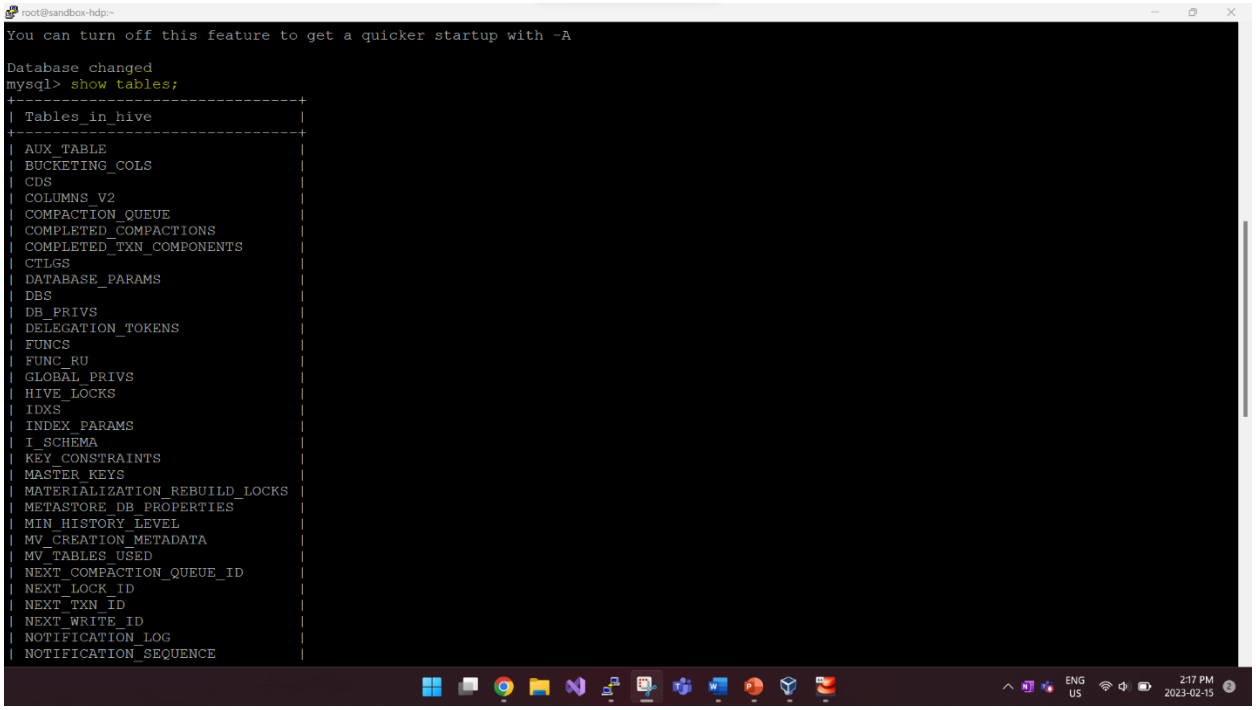


Figure 2 Hive table details

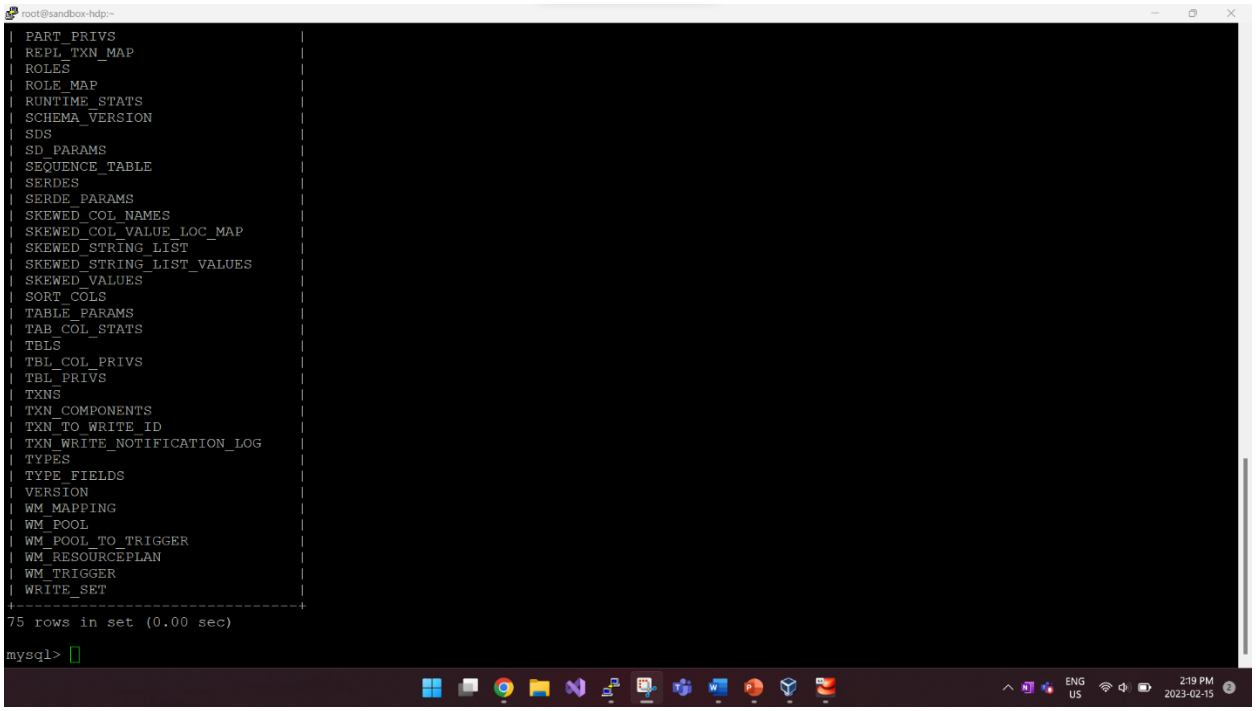


Figure 3 Hive table details

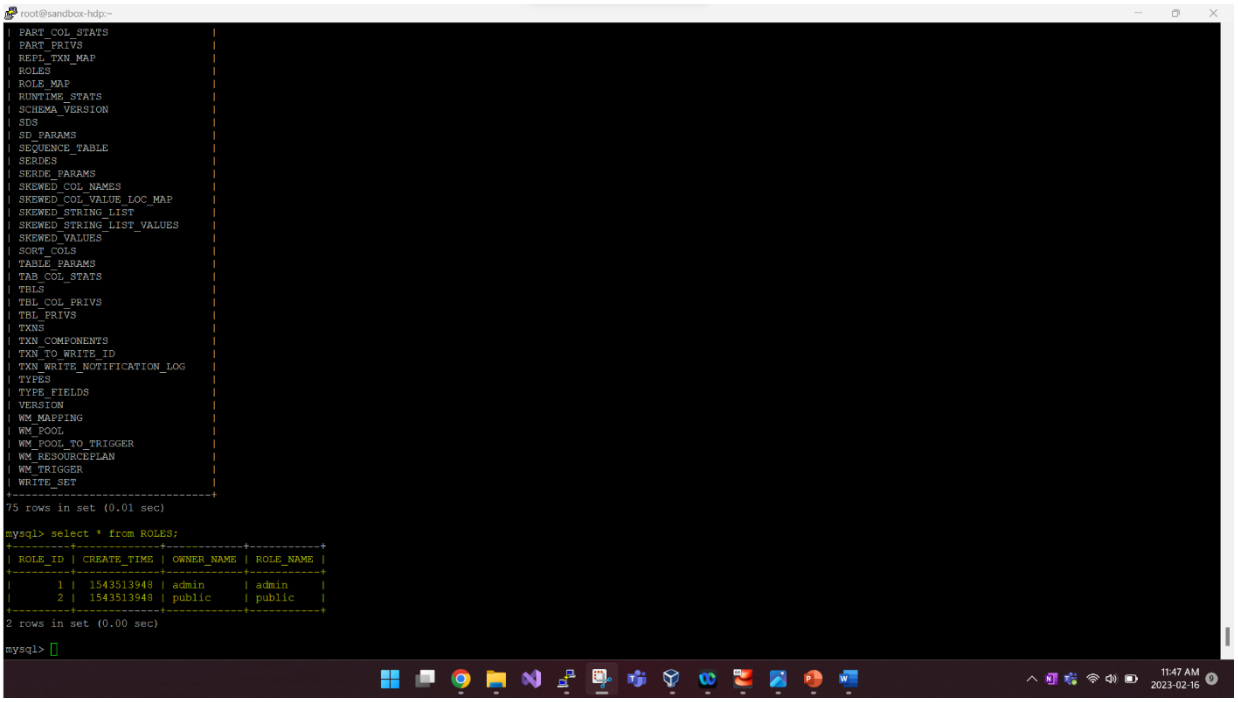


Figure 4 Inside ROLES table

I used ***Hadoop command*** to check all directories where I got ***user directory*** then I made a new ***subdirectory*** in directory which is ***sandeep.*** I used subdirectory to ***import table*** from ***databases*** to ***hdfs.*** (Figure 5)

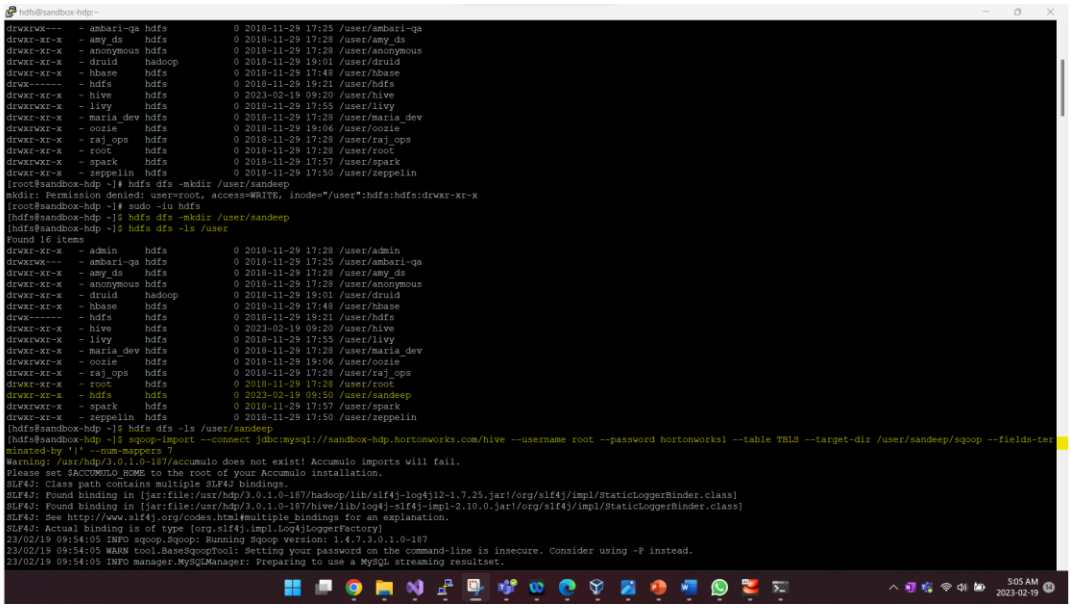


Figure 5 made a subdirectory with sandeep name

After importing a table, I used ***ls command*** to check there is a ***sqoop*** directory where I got it and used again ***ls command*** to check what's in a table, I got all content. (Figure 6)

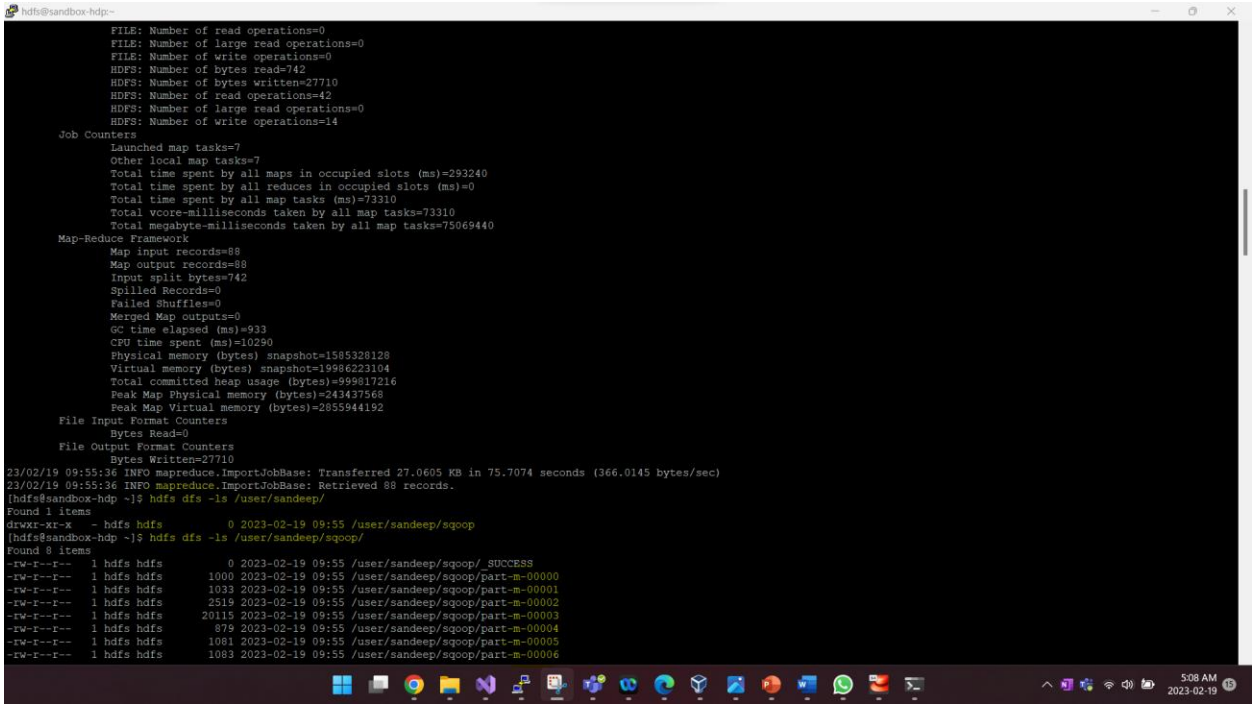


Figure 6 used ls command to chcek the directory and got sqoop

and used again ***ls command*** to check what's in a table, I got all content which I showed in Figure 7. I have checked the content in ***part-m-0000.***

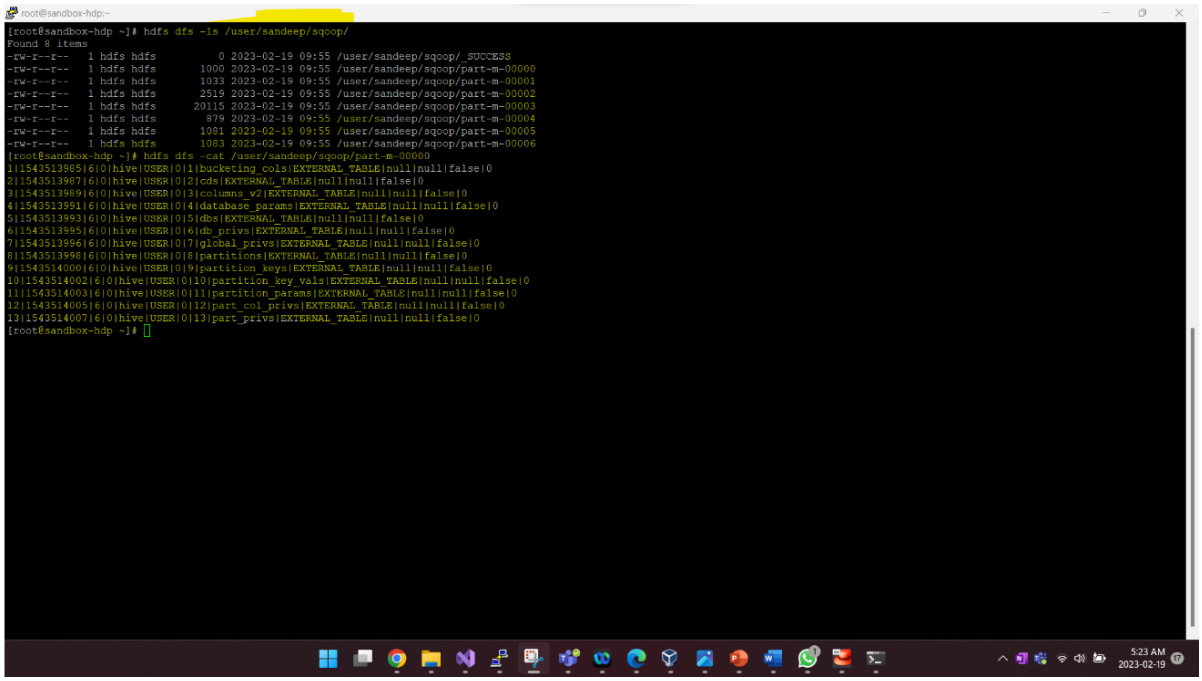


Figure 7 used ls command to check the whole content inside sqoop

2 Demonstrate export command.

I showed a command to export the content from **hdfs to mysql** databases. But before exporting the data from **HDFS to MySql**, I needed to create a table in **mysql** where I had to export all content from **hdfs.** I created a table which is **TBLS_2** from texting directory After creating a table I used export command which I showed in Figure 8

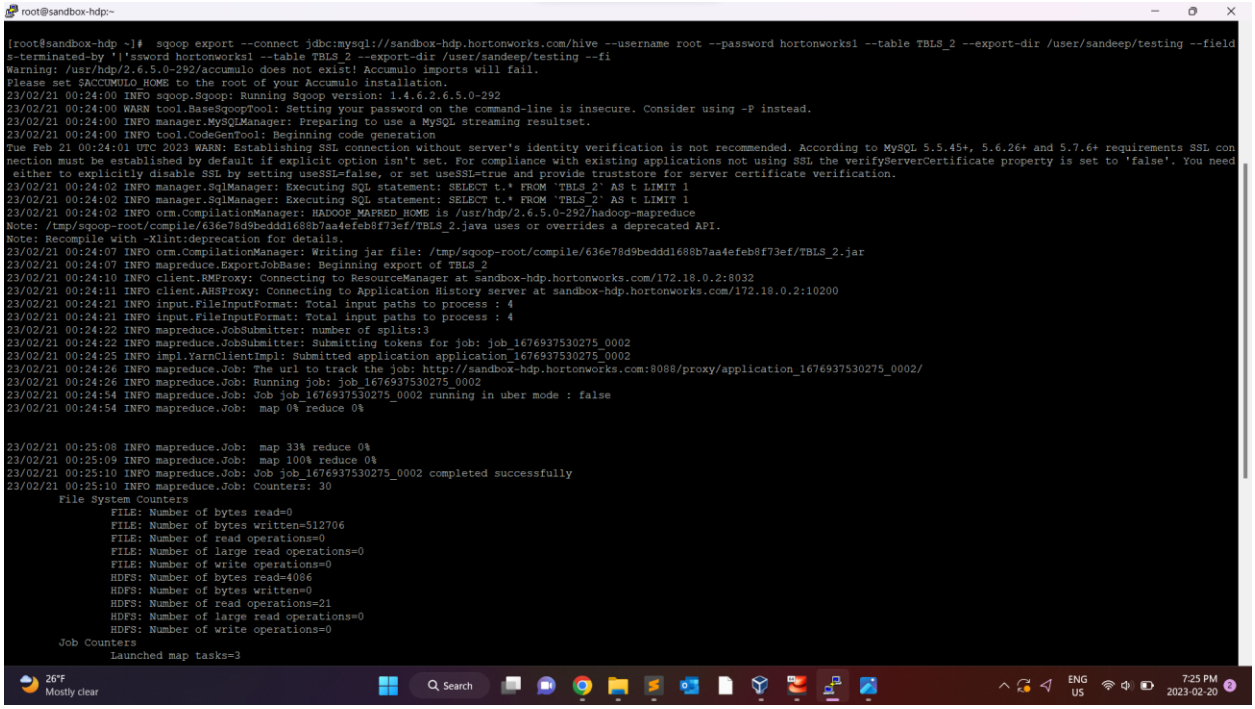


Figure 8 Export command

Export Command> sqoop export –connect jdbc:mysql://sandbox-hdp.hortonworks.com/hive –username root –password hortonworks1 –table TBLS_2 –table TBLS_2 --export-dir /user/Sandeep/testing –fields-terminated-by '|'

I used **hive** databases to send all content from hdfs to mysql where I used TBLS_2 table.

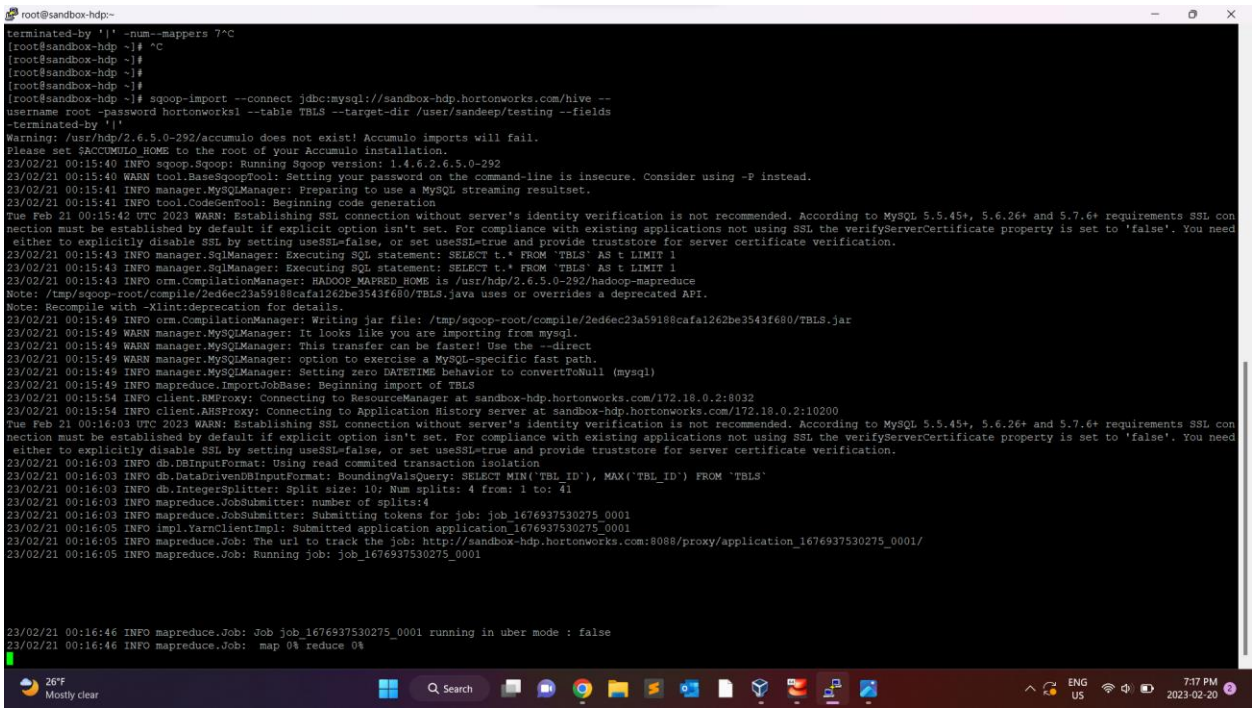


Figure 9 Export command

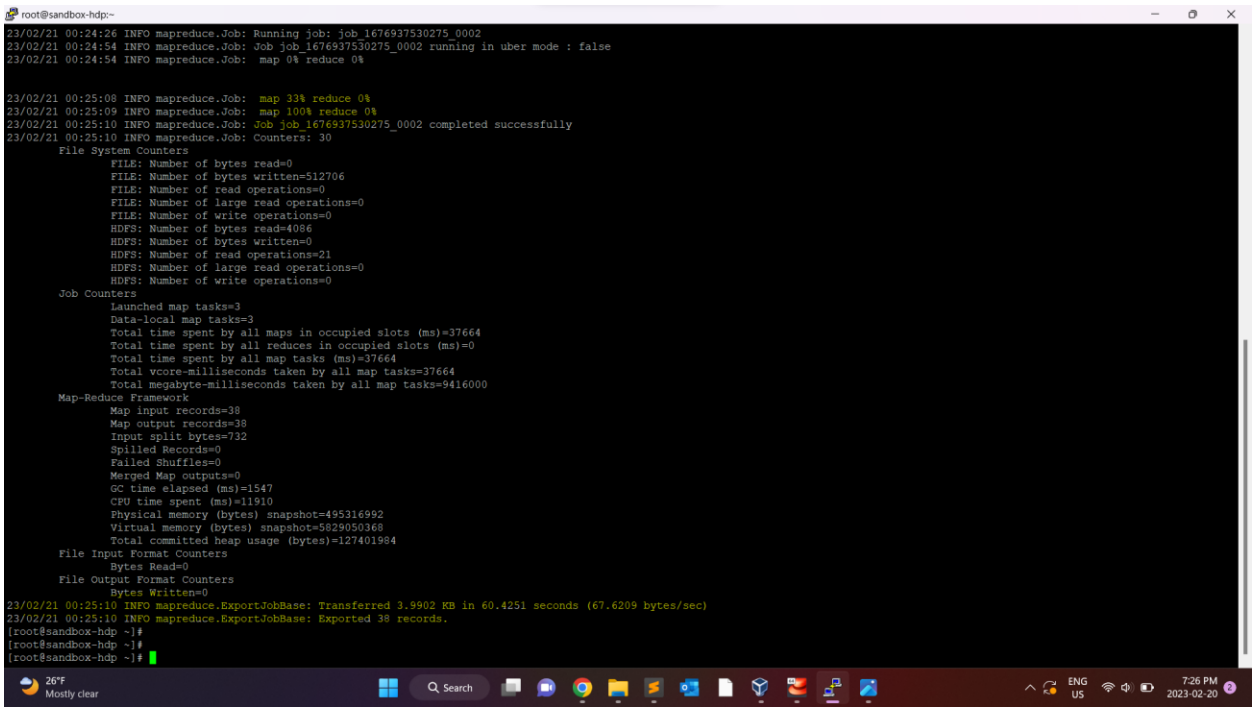
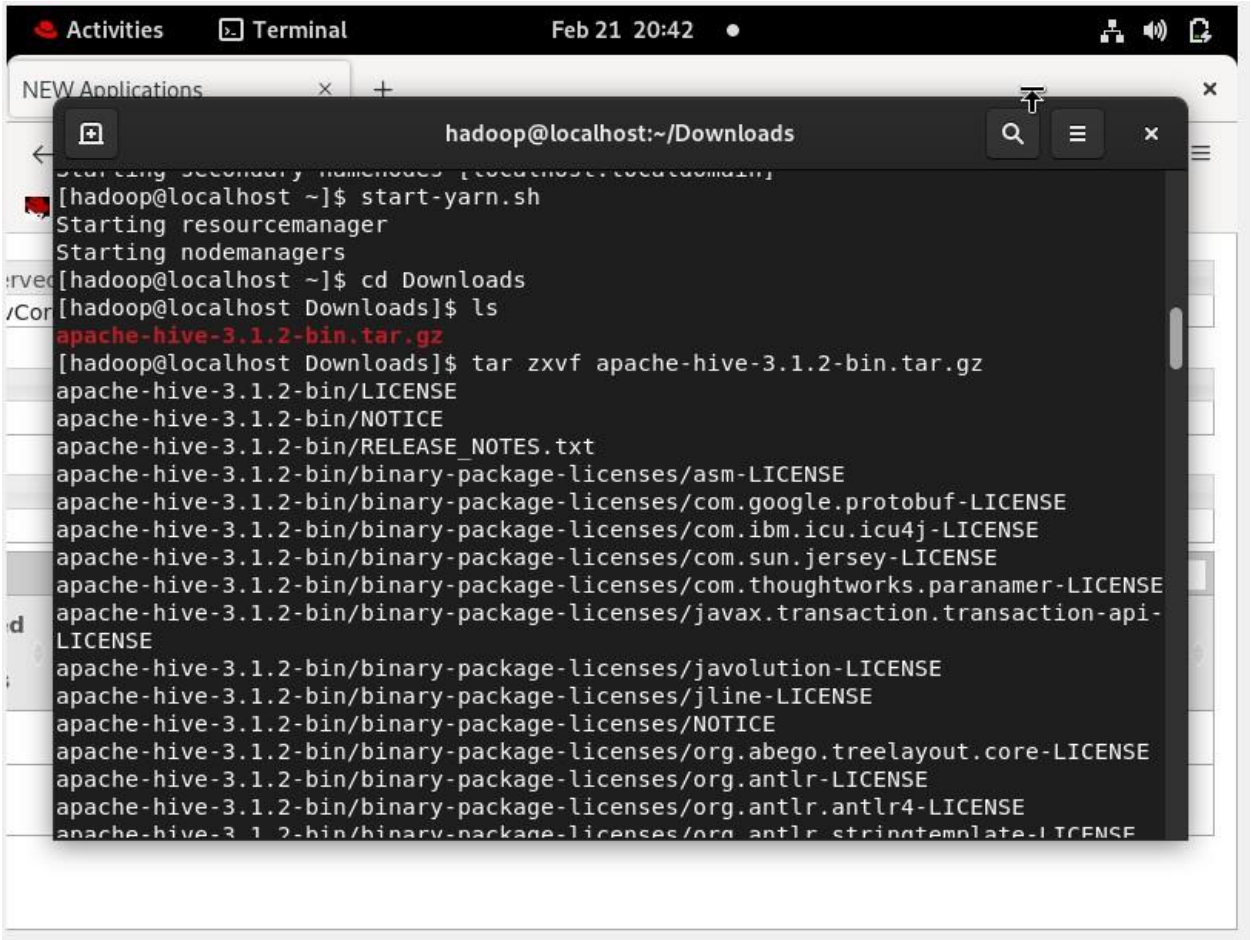


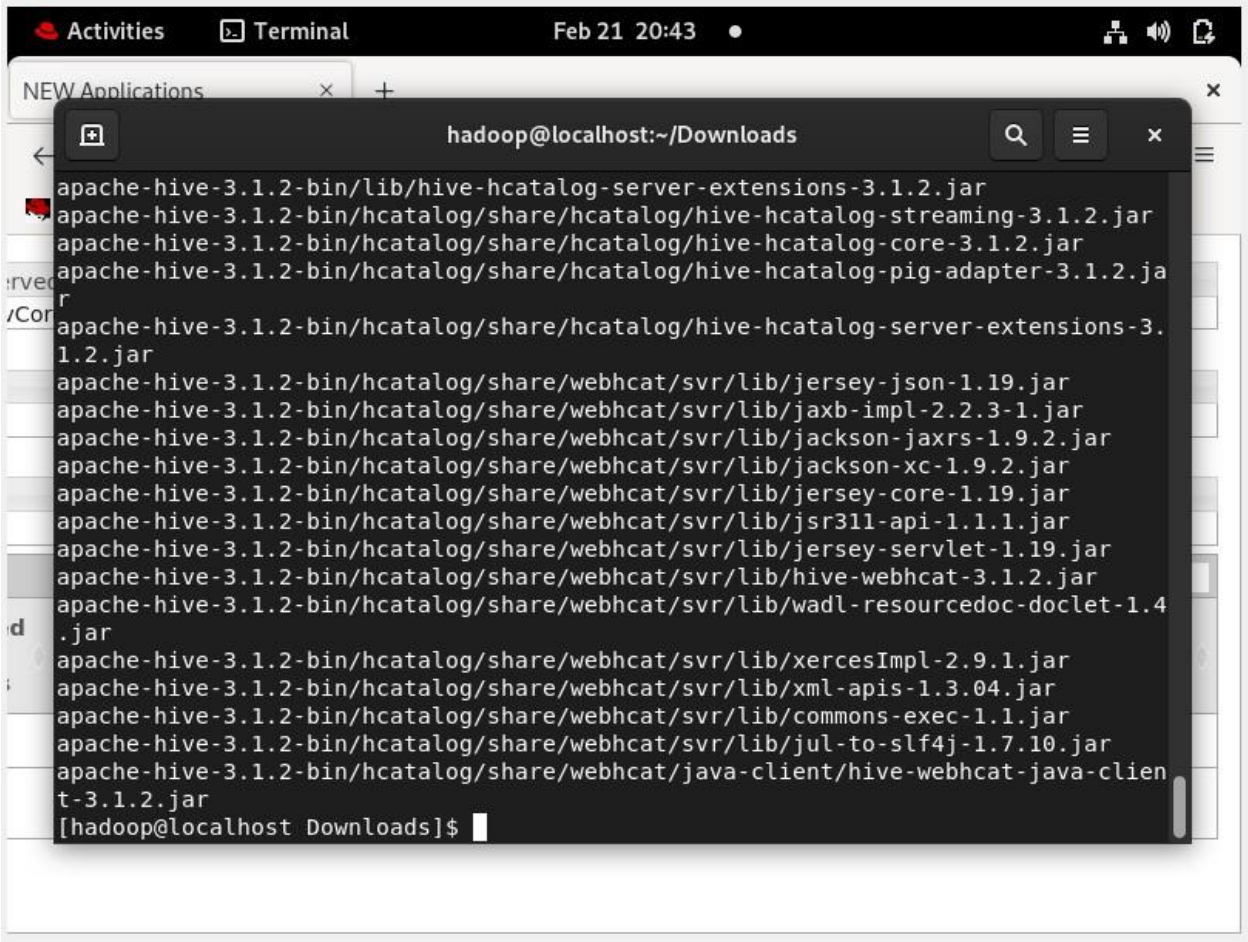
Figure 10 **Exported all data from hdfs to mysql**

Task 2: Week 6: Hive

Upload Book.xml in HDFS then create a table in Hive and parse it.

Download the Apache hive from the Apache website and extract the setup file. Using the tar command that is tar zxvf apache-hive-3.1.2-bin.tar.gz

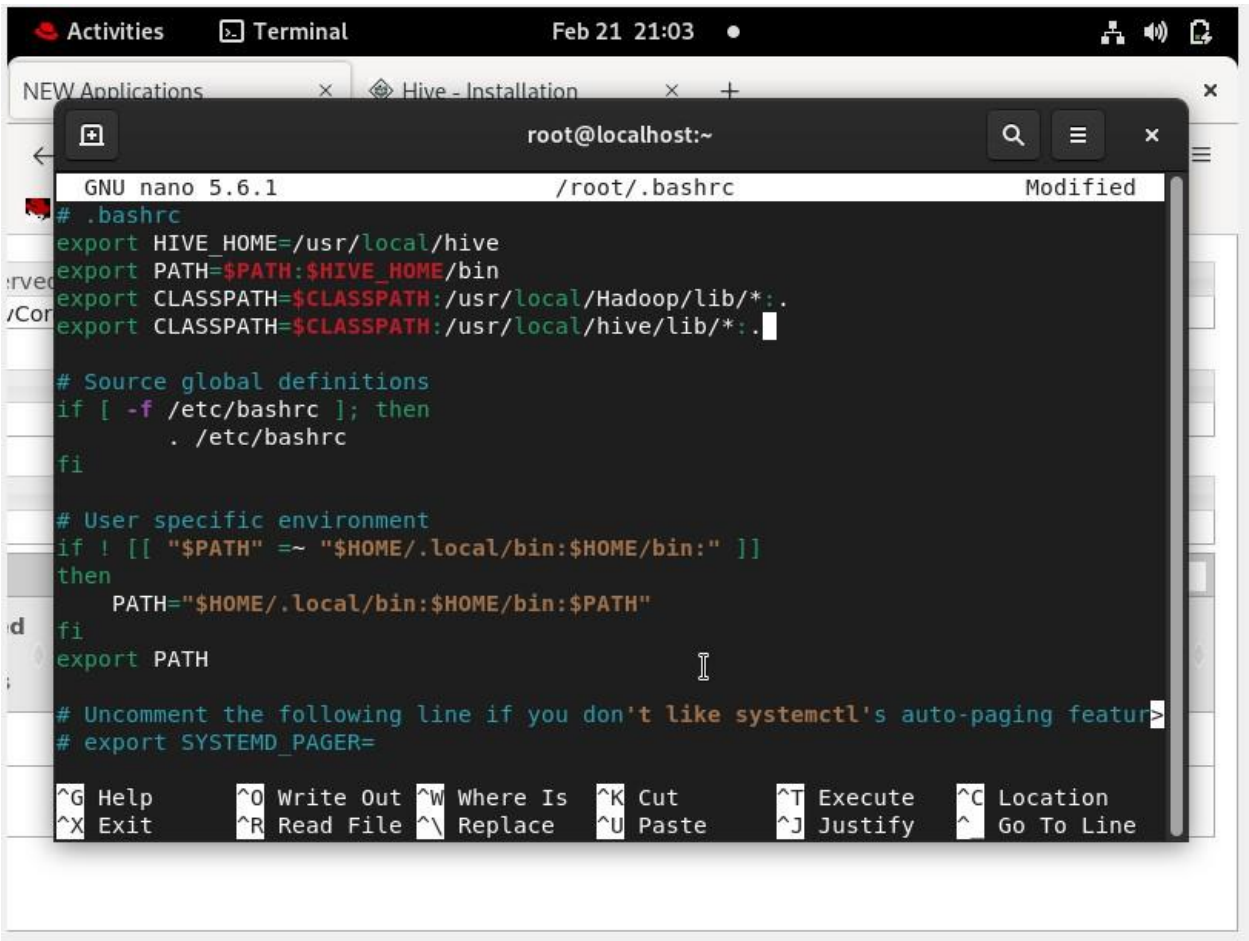

Activities
Terminal
Feb 21 20:42
NEW Applications
hadoop@localhost:~/Downloads
[hadoop@localhost ~]$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
[hadoop@localhost ~]$ cd Downloads
[hadoop@localhost Downloads]$ ls
apache-hive-3.1.2-bin.tar.gz
[hadoop@localhost Downloads]$ tar zxvf apache-hive-3.1.2-bin.tar.gz
apache-hive-3.1.2-bin/LICENSE
apache-hive-3.1.2-bin/NOTICE
apache-hive-3.1.2-bin/RELEASE_NOTES.txt
apache-hive-3.1.2-bin/binary-package-licenses/asm-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.google.protobuf-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.ibm.icu.icu4j-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.sun.jersey-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.thoughtworks.paranamer-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/javax.transaction.transaction-api-
LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/javolution-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/jline-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/NOTICE
apache-hive-3.1.2-bin/binary-package-licenses/org.abego.treelayout.core-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.antlr-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.antlr.antlr4-LICENSE



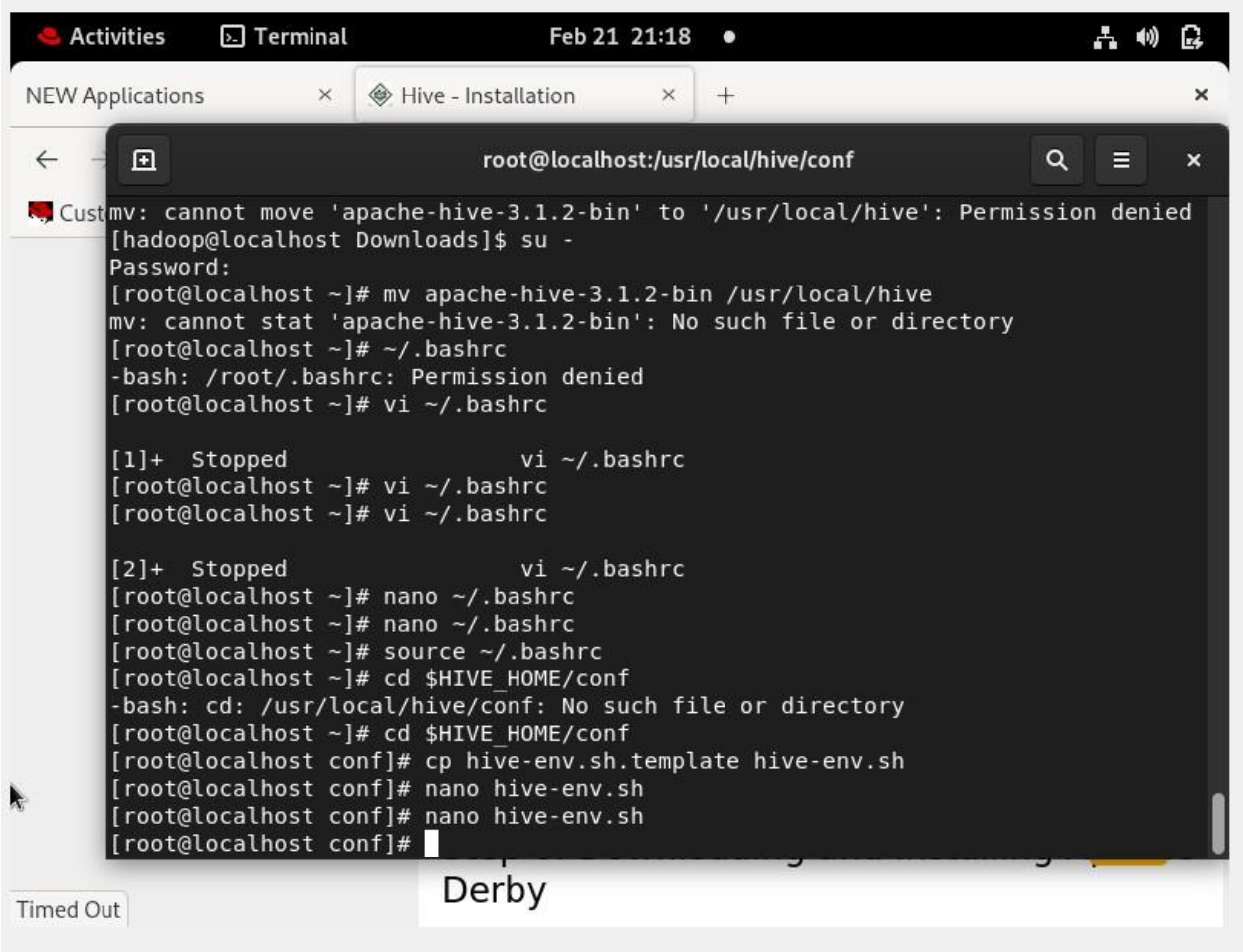
2. Update the hive location in the bashrc file using export command.

Export HIVE Home=/usr/local/hive

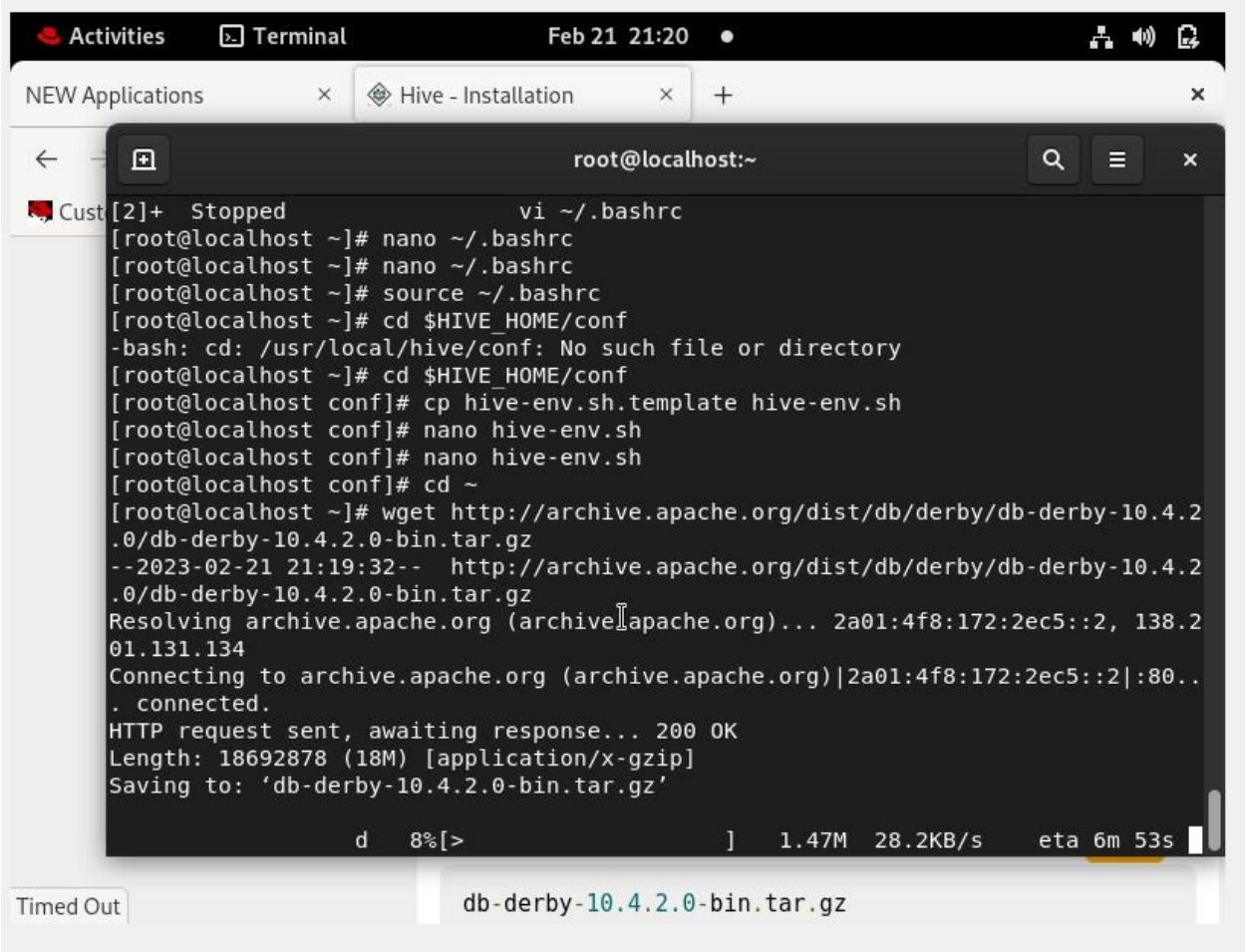
Export PATH=$PATH:$HIVE HOME/bin



3. Update the hive-env.sh file



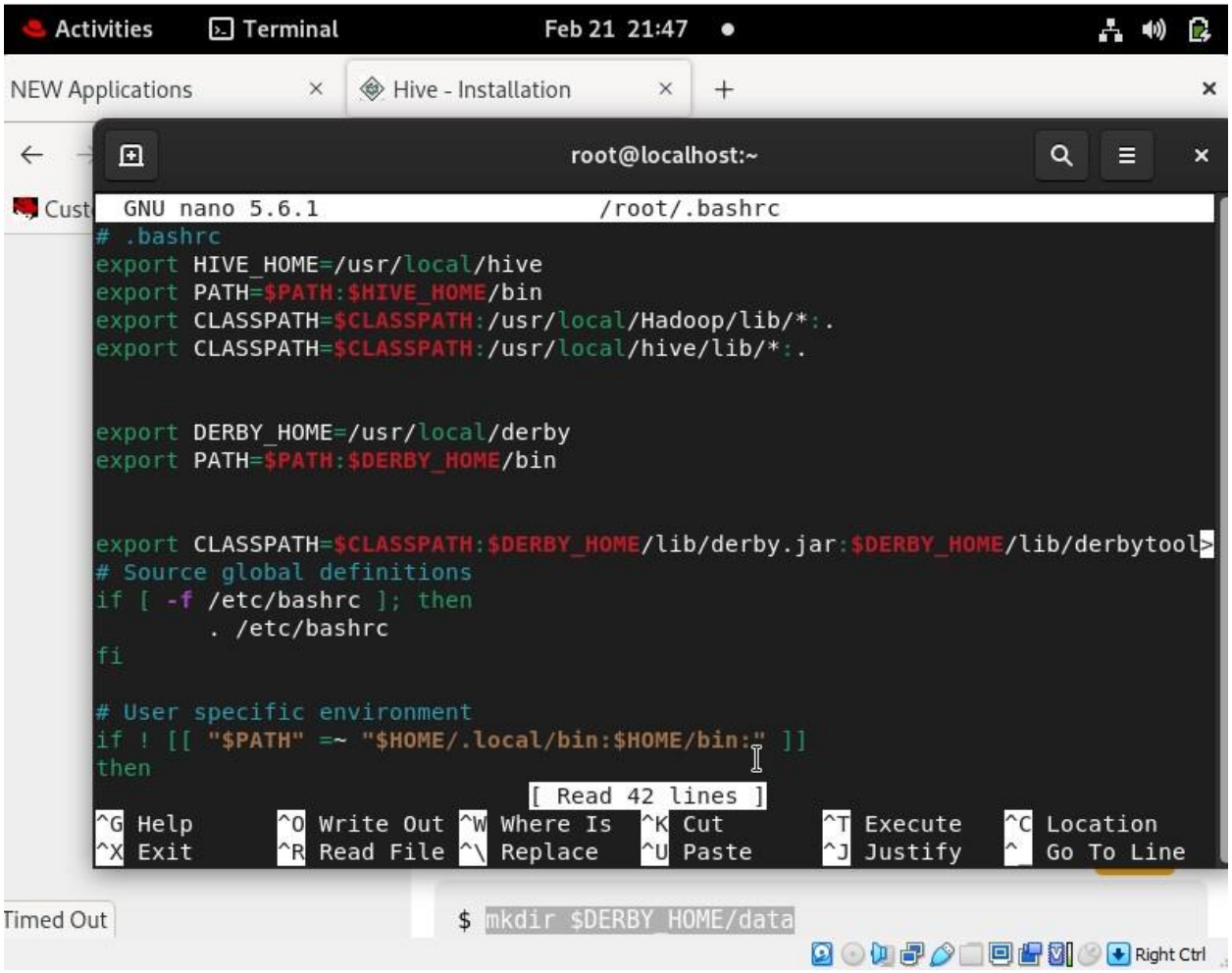
4. Install the derby in the hive.



5. Set the path of derby in the bashrc file using export command.

Export DERBY HOME /usr/local/derby

Export PATH $PTH:$DERBY HOME/bin



6. Create the book table in the hive.

```
harman@harman-VirtualBox: ~
login as: harman
harman@192.168.43.233's password:
Welcome to Ubuntu 22.04.2 LTS (GNU/Linux 5.8.0-43-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

933 updates can be installed immediately.
0 of these updates are security updates.
To see these additional updates run: apt list --upgradable

New release '22.04.1 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Wed Feb 22 05:55:33 2023 from 127.0.0.1
harman@harman-VirtualBox:~$ hivehive> create temporary table if not exists books_xml(line string);
OK
Time taken: 0.932 seconds
hive> load data local inpath 'books.xml' into table book_xml;
Loading data to table default.book_xml
OK
Time taken: 0.935 seconds
hive> Create temporary table if not exists books(
    > author string,
    > title array<string>,
    > genre array<string>,
    > price array<string>,
    > discount array<string>,
    > publishdate array<string>,
    > description array<string>,
    > row format delimited
    > fields terminated by '|';
OK
Time taken: 0.194 seconds
```

Name of students who has participated in the assignment.
Sandeep Kaur
Harmandeep Kaur
Gabriel Adeniyi
Sukhman Kaur