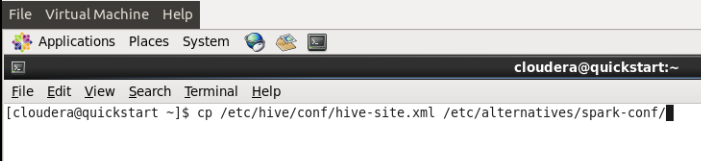
**All three Books, Ratings, and users data Stores in hdfs**

Open New terminal and copy hive-site.xml file in spark-conf

**cp /etc/hive/conf/hive-site.xml /etc/alternatives/spark-conf/**



Spark SQL is the newest component of Spark and provides a SQL like interface.

we will start by launching the Spark shell from the root directory of the provided USB drive:

**bin/spark-shell**

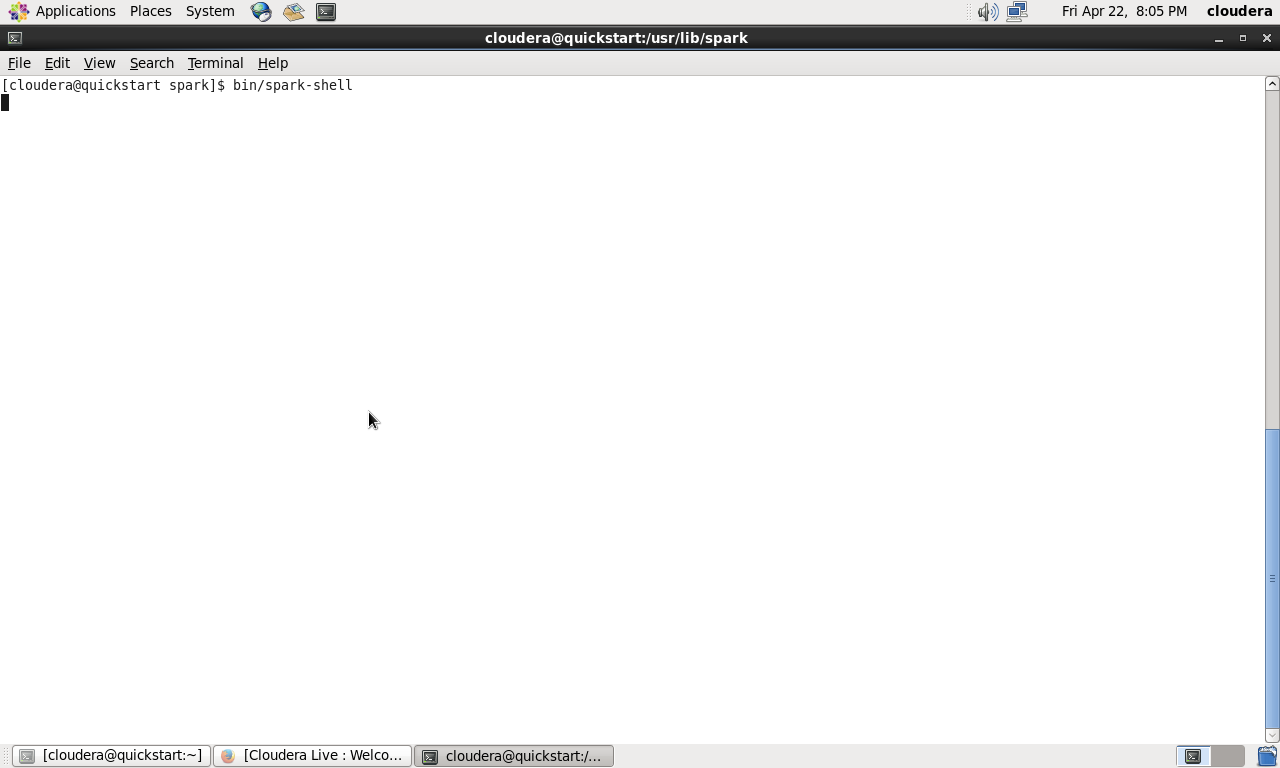
Entering into spark shell:

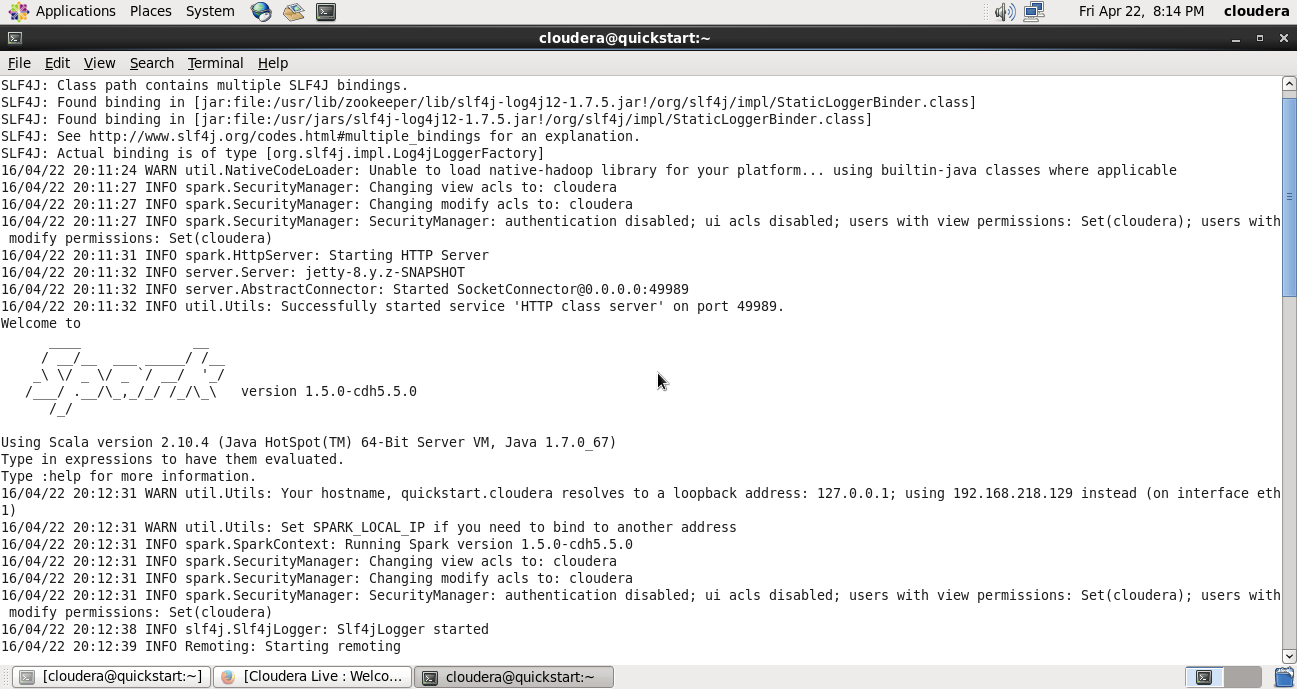
Spark is pre-installed in Quickstart vm CDH 5.3 .0. We need to enter into spark shell by changing present working directory which is shown below

Command: **cd /usr/lib/spark**

Type the following command in terminal.

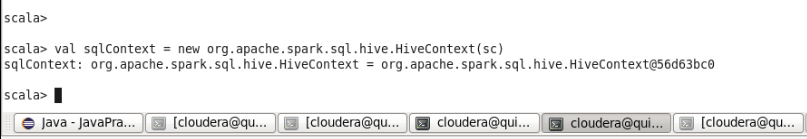
Command: **bin/spark-shell**  You will enter into spark shell and see the following screen as below.



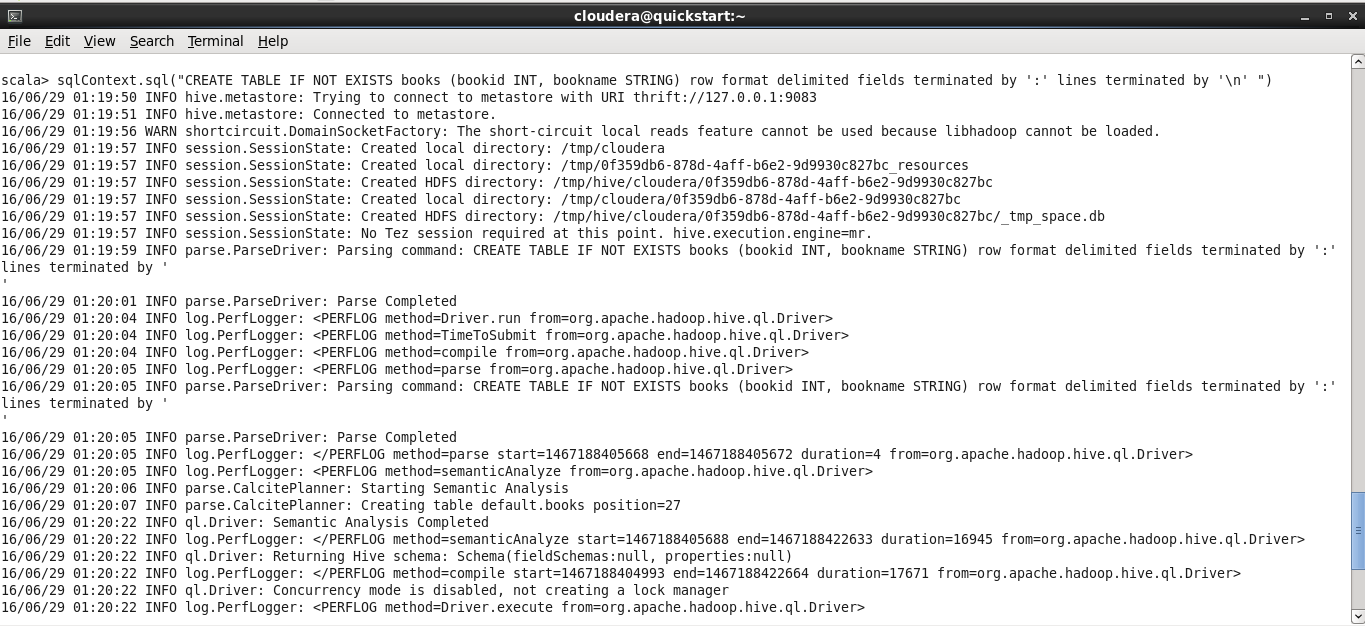


Once you have launched the Spark shell, the next step is to create a SQLContext. A SQLConext wraps the SparkContext.

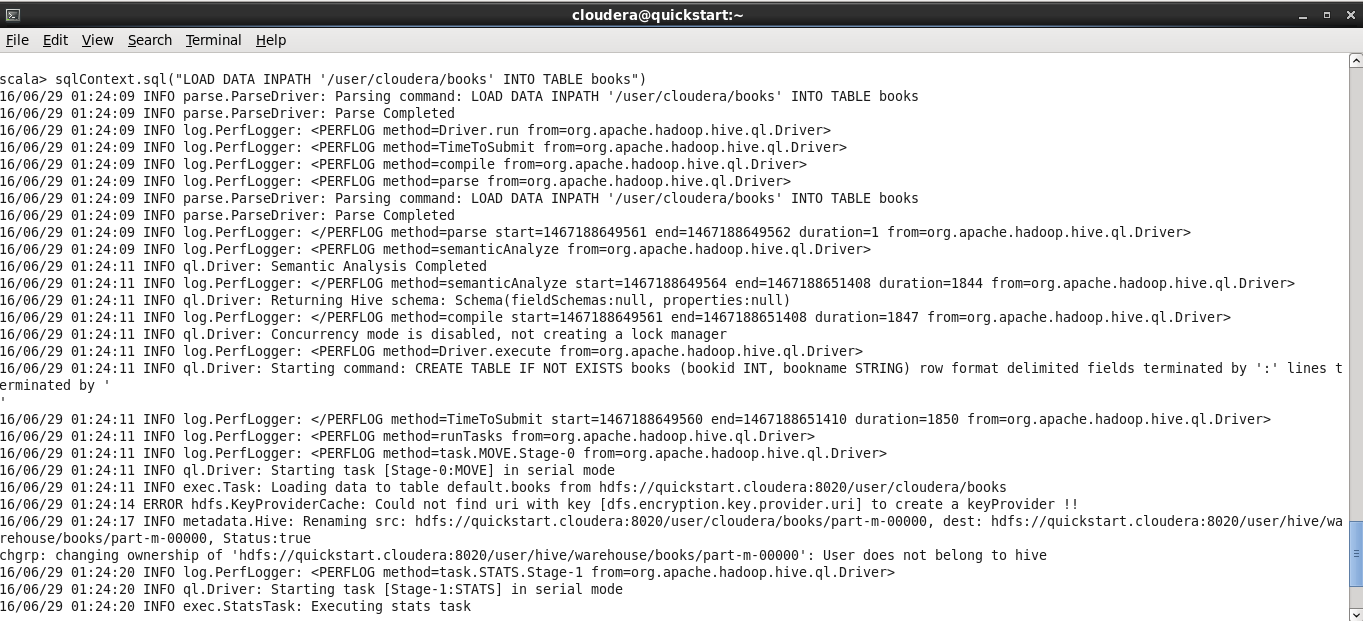
**val sqlContext = new org.apache.spark.sql.hive.HiveContext(sc)**



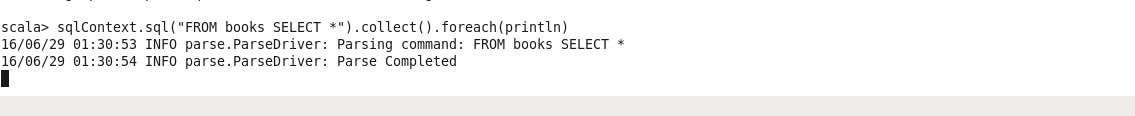
**sqlContext.sql("CREATE TABLE IF NOT EXISTS books (bookid INT, bookname STRING) row format delimited fields terminated by ':' lines terminated by '\n' ")**



**sqlContext.sql("LOAD DATA INPATH '/user/cloudera/books' INTO TABLE books")**



**sqlContext.sql("FROM books SELECT \*").collect().foreach(println)**



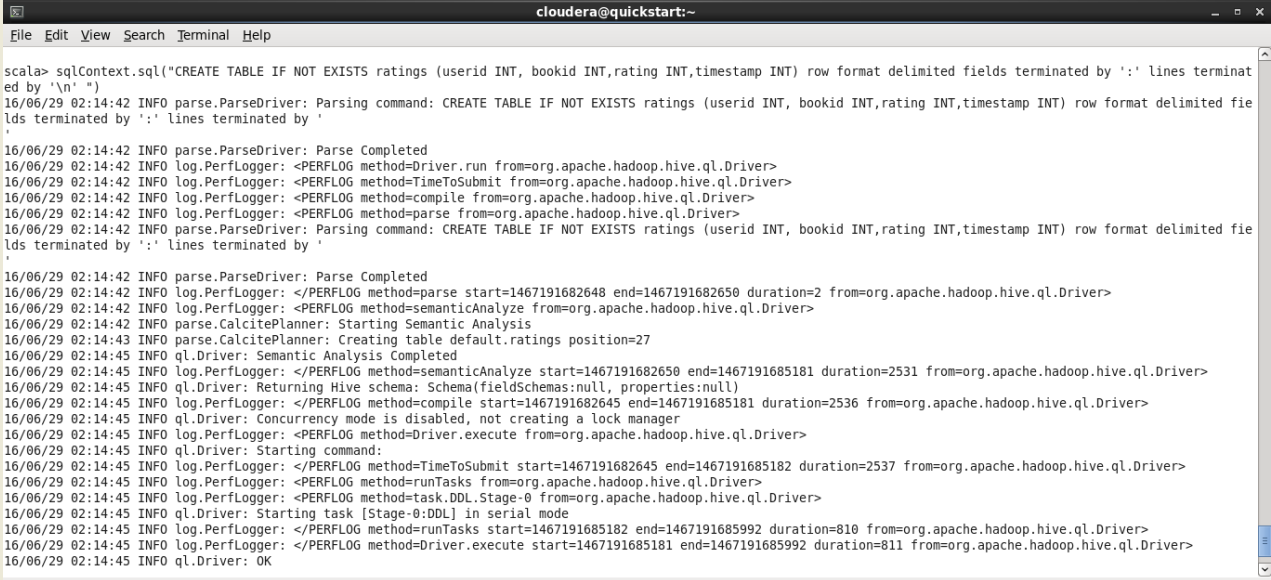


**val sqlContext = new org.apache.spark.sql.hive.HiveContext(sc)**

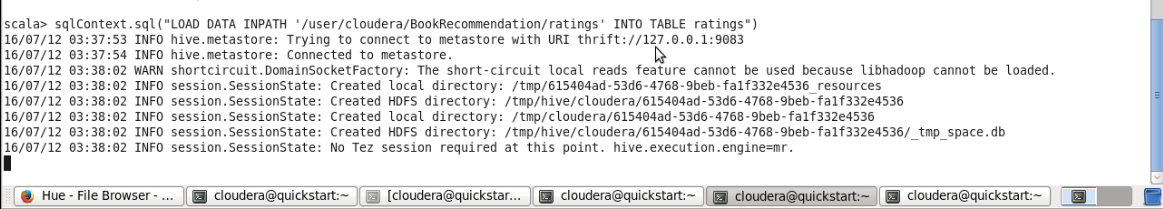


**sqlContext.sql("CREATE TABLE IF NOT EXISTS ratings (userid INT, bookid INT,rating INT,timestamp INT) row format delimited fields terminated by ':' lines terminated by '\n' ")**

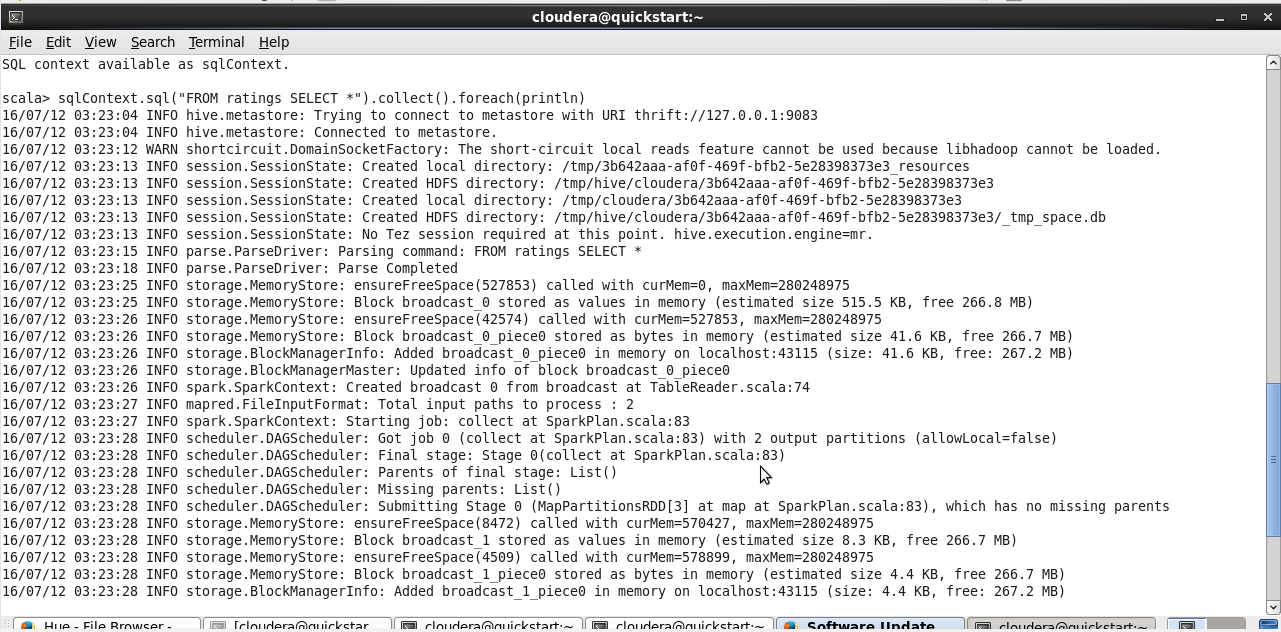




**sqlContext.sql("LOAD DATA INPATH '/user/cloudera/ratings' INTO TABLE ratings")**



**sqlContext.sql("FROM ratings SELECT \*").collect().foreach(println)**

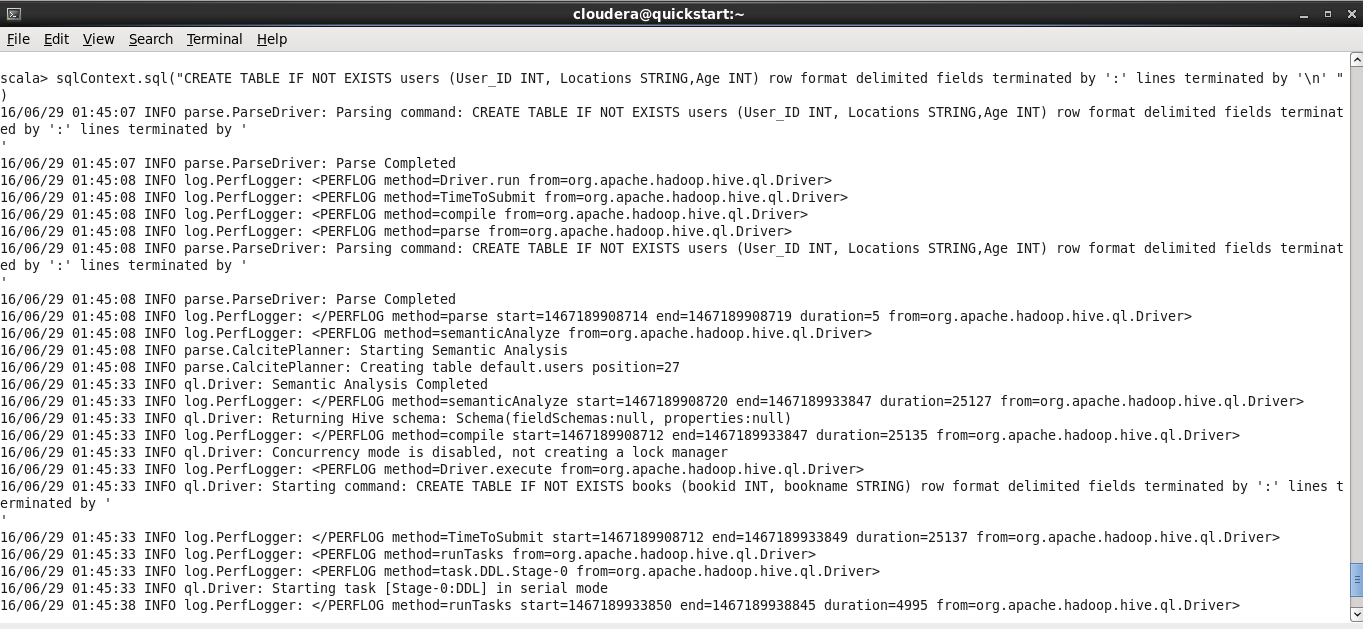


**val sqlContext = new org.apache.spark.sql.hive.HiveContext(sc)**

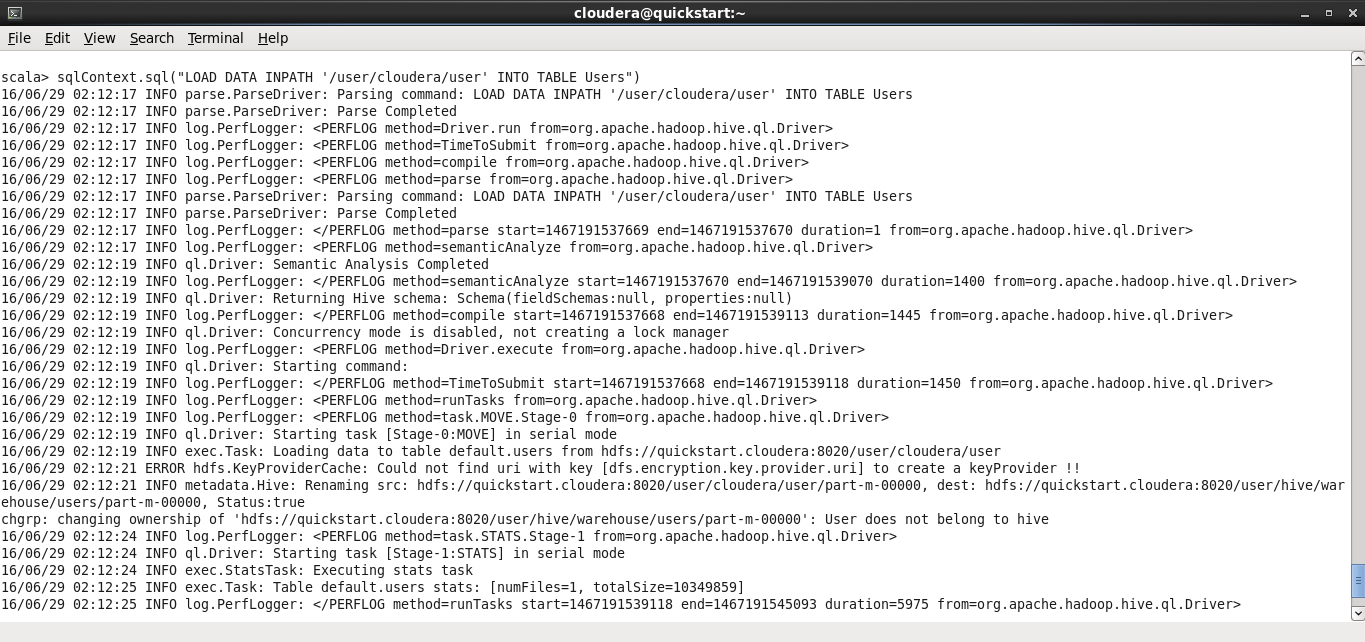


**sqlContext.sql("CREATE TABLE IF NOT EXISTS users (User\_ID INT, Locations STRING,Age INT) row format delimited fields terminated by ':' lines terminated by '\n' ")**





**sqlContext.sql("LOAD DATA INPATH '/user/cloudera/users' INTO TABLE users")**



**sqlContext.sql("FROM users SELECT \*").collect().foreach(println)**

Hive Tables:

