Create and Load a Hive-Managed Table

Next, you will create and then load a Hive-managed table with product ratings data.

1. Create a table named `ratings` for storing tab-delimited records using this structure:

Field Name	Field Type
posted	`TIMESTAMP`
cust_id	`INT`
prod_id	`INT`
rating	`TINYINT`
message	`STRING`

2. Show the table description and verify that its fields have the correct order, names, and types:

```
DESCRIBE ratings;
```

3. Next, open a separate terminal window (File -> Open Terminal) so you can run the following shell command. This will populate the table directly by using the `hadoop fs` command to copy product ratings data from 2012 to that directory in HDFS:

```
$ hadoop fs -put $ADIR/data/ratings_2012.txt \
/user/hive/warehouse/ratings
```

   Leave the window open afterwards so that you can easily switch between Hive and the command prompt.

4. Next, verify that Hive can read the data we just added. Run the following query in Hive to count the number of records in this table (the result should be 464):

```
SELECT COUNT(*) FROM ratings;
```

5. Another way to load data into a Hive table is through the `LOAD DATA` command. The next few commands will lead you through the process of copying a local file to HDFS and loading it into Hive. First, copy the 2013 ratings data to HDFS:

```
$ hadoop fs -put $ADIR/data/ratings_2013.txt /dualcore
```

6. Verify that the file is there:

```
$ hadoop fs -ls /dualcore/ratings_2013.txt
```

7. Use the `LOAD DATA` statement in Hive to load that file into the `ratings` table:

```
LOAD DATA INPATH '/dualcore/ratings_2013.txt' INTO
TABLE ratings;
```

8. The `LOAD DATA INPATH` command *moves* the file to the table's directory. Verify that the file is no longer present in the original directory:

```
$ hadoop fs -ls /dualcore/ratings_2013.txt
```

9. Verify that the file is shown alongside the 2012 ratings data in the table's directory:

```
$ hadoop fs -ls /user/hive/warehouse/ratings
```

10. Finally, count the records in the ratings table to ensure that all 21,997 are available:

```
SELECT COUNT(*) FROM ratings;
```