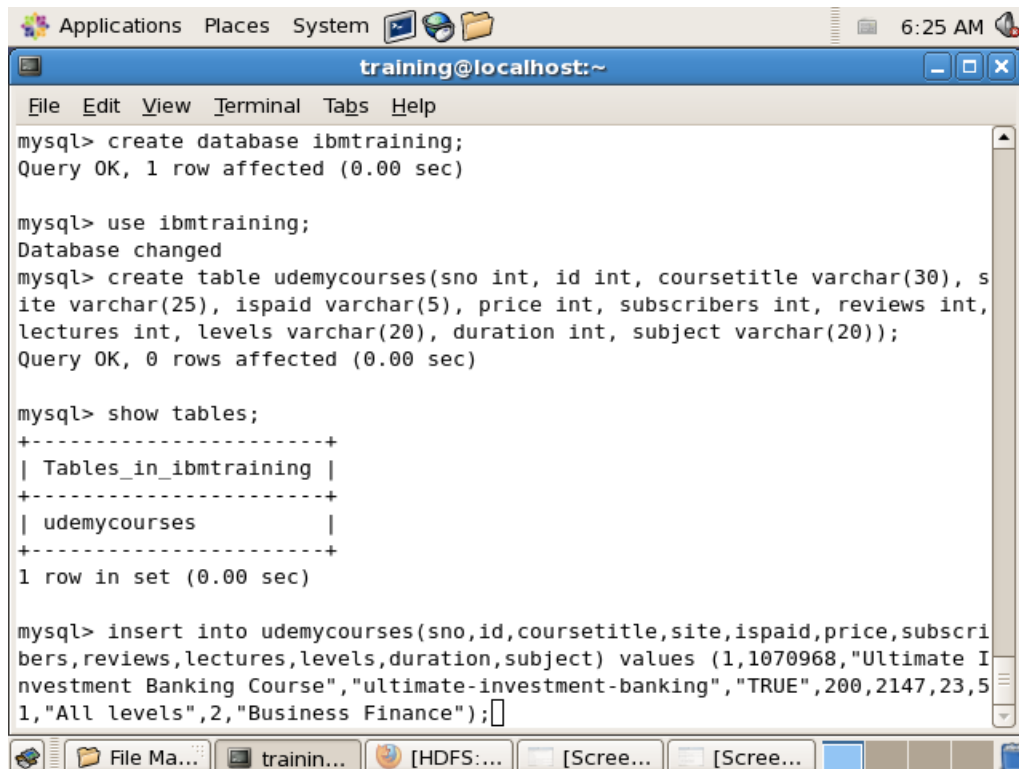


SQOOP

Import Command

1. Open the terminal, type `mysql -u root` and hit enter.
2. Use a database and create table inside that and insert values. We have inserted 3 rows.
3. Exit by `ctrl+z`. Write `sqoop import` command to import table from MySQL database to hadoop HDFS.



The screenshot shows a terminal window titled "training@localhost:~" with a menu bar (File, Edit, View, Terminal, Tabs, Help) and a system bar (Applications, Places, System, 6:25 AM). The terminal displays the following MySQL commands and their outputs:

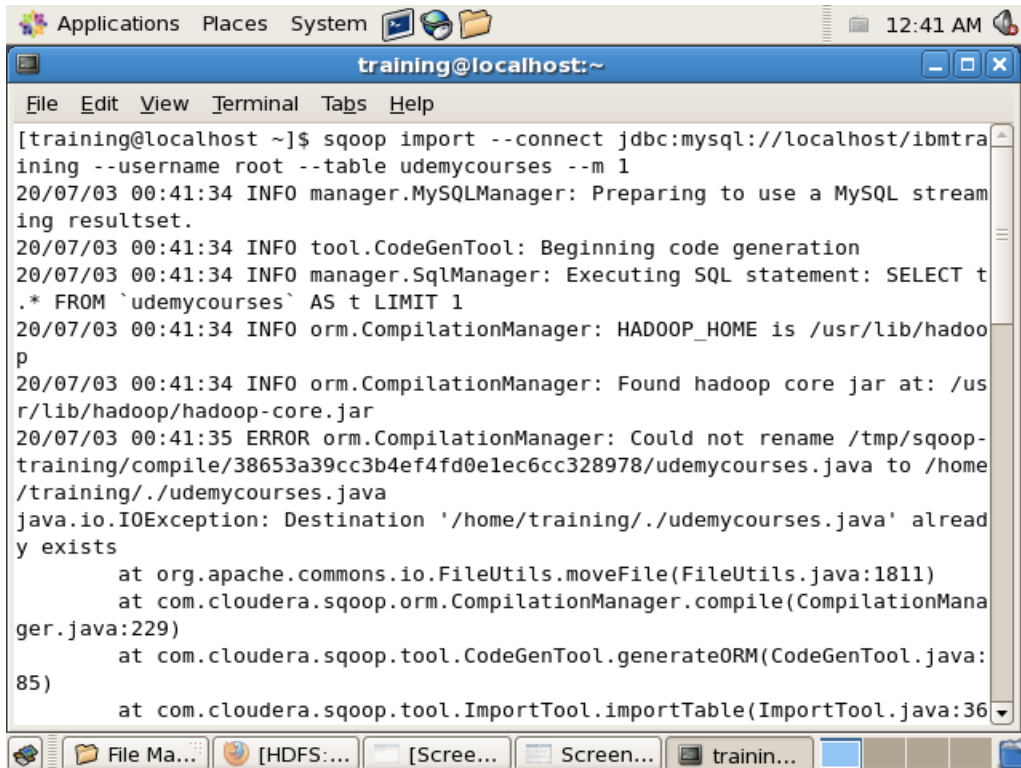
```
mysql> create database ibmtraining;
Query OK, 1 row affected (0.00 sec)

mysql> use ibmtraining;
Database changed
mysql> create table udemycourses(sno int, id int, coursetitle varchar(30), site varchar(25), ispaid varchar(5), price int, subscribers int, reviews int, lectures int, levels varchar(20), duration int, subject varchar(20));
Query OK, 0 rows affected (0.00 sec)

mysql> show tables;
+-----+
| Tables_in_ibmtraining |
+-----+
| udemycourses          |
+-----+
1 row in set (0.00 sec)

mysql> insert into udemycourses(sno,id,coursetitle,site,ispaid,price,subscribers,reviews,lectures,levels,duration,subject) values (1,1070968,"Ultimate Investment Banking Course","ultimate-investment-banking","TRUE",200,2147,23,51,"All levels",2,"Business Finance");
```

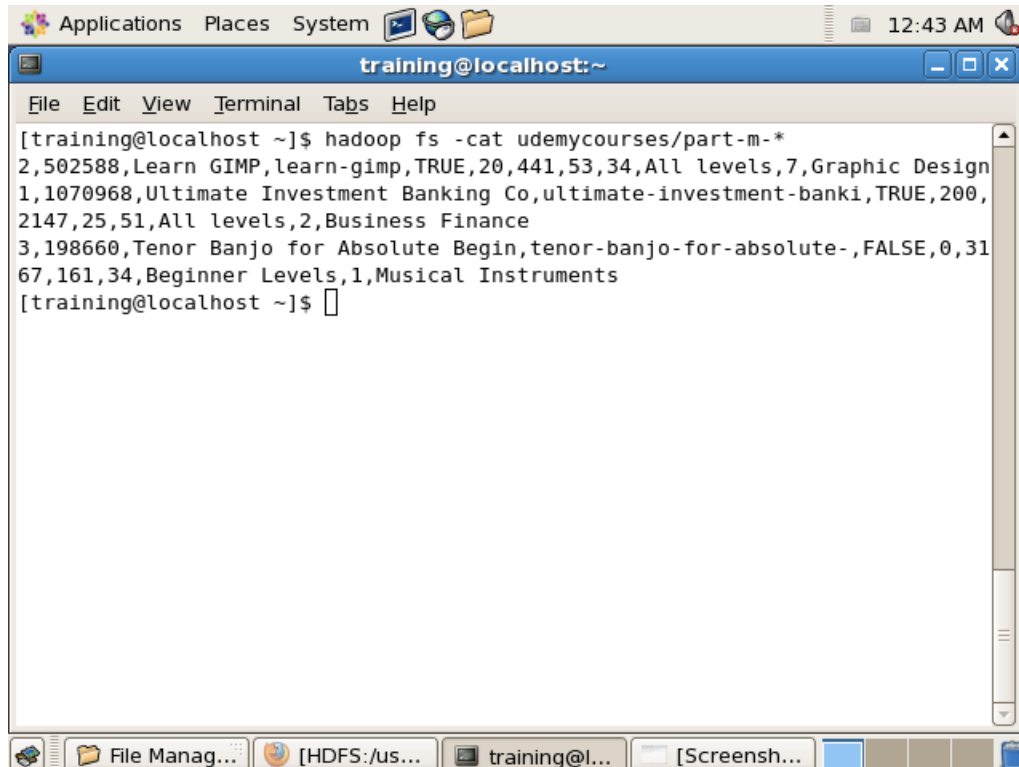
The terminal window also shows a taskbar at the bottom with icons for File Manager, a terminal window titled "trainin...", and several window titles starting with "[HDFS:..." and "[Scree...".



A terminal window titled "training@localhost:~" is shown. The window has a menu bar with "File", "Edit", "View", "Terminal", "Tabs", and "Help". The terminal output shows the execution of the command `sqoop import --connect jdbc:mysql://localhost/ibmtraining --username root --table udemycourses --m 1`. It displays several informational messages from the Cloudera Sqoop tools, including preparing a MySQL stream, beginning code generation, and executing a SQL statement. An error message follows: `java.io.IOException: Destination '/home/training/./udemycourses.java' already exists`. The stack trace indicates the error occurred in `org.apache.commons.io.FileUtils.moveFile`, `com.cloudera.sqoop.orm.CompilationManager.compile`, `com.cloudera.sqoop.tool.CodeGenTool.generateORM`, and `com.cloudera.sqoop.tool.ImportTool.importTable`. The window's taskbar at the bottom shows icons for "File Ma...", "[HDFS:...", "[Scree...", "Screen...", and "trainin...".

```
[training@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/ibmtraining --username root --table udemycourses --m 1
20/07/03 00:41:34 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
20/07/03 00:41:34 INFO tool.CodeGenTool: Beginning code generation
20/07/03 00:41:34 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `udemycourses` AS t LIMIT 1
20/07/03 00:41:34 INFO orm.CompilationManager: HADOOP_HOME is /usr/lib/hadoop
20/07/03 00:41:34 INFO orm.CompilationManager: Found hadoop core jar at: /usr/lib/hadoop/hadoop-core.jar
20/07/03 00:41:35 ERROR orm.CompilationManager: Could not rename /tmp/sqoop-training/compile/38653a39cc3b4ef4fd0e1ec6cc328978/udemycourses.java to /home/training/./udemycourses.java
java.io.IOException: Destination '/home/training/./udemycourses.java' already exists
    at org.apache.commons.io.FileUtils.moveFile(FileUtils.java:1811)
    at com.cloudera.sqoop.orm.CompilationManager.compile(CompilationManager.java:229)
    at com.cloudera.sqoop.tool.CodeGenTool.generateORM(CodeGenTool.java:85)
    at com.cloudera.sqoop.tool.ImportTool.importTable(ImportTool.java:36)
```

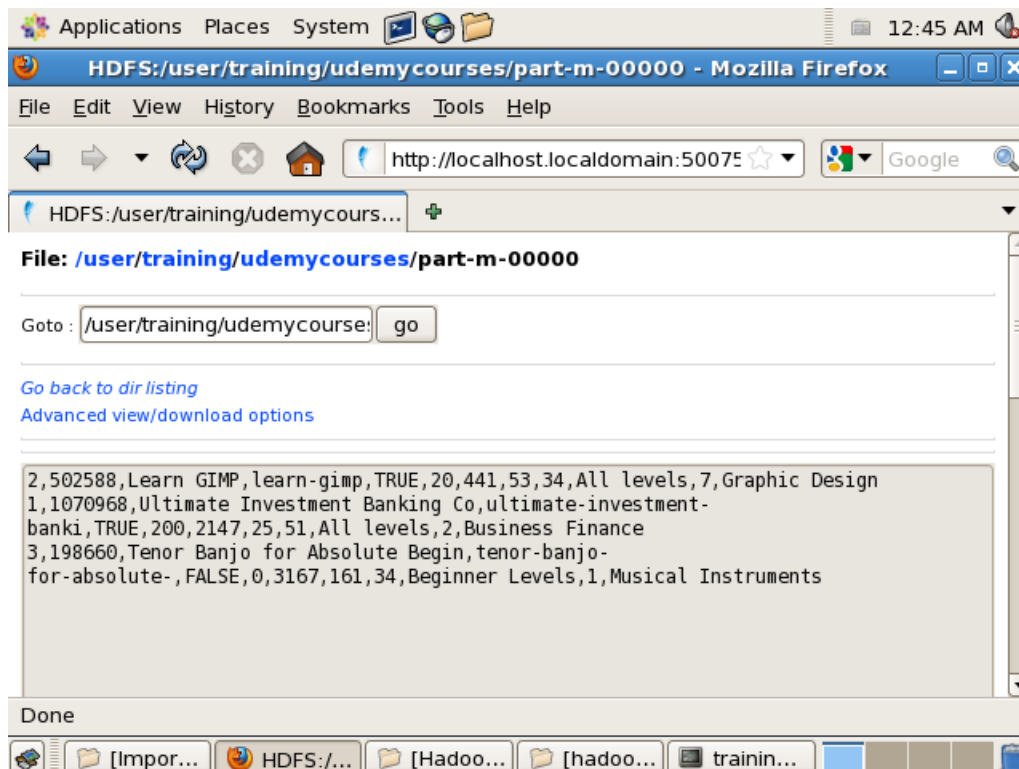
To verify imported data :



A terminal window titled 'training@localhost:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help). The command 'hadoop fs -cat udemycourses/part-m-*' has been executed, displaying the following output:

```
[training@localhost ~]$ hadoop fs -cat udemycourses/part-m-*
2,502588,Learn GIMP,learn-gimp,TRUE,20,441,53,34,All levels,7,Graphic Design
1,1070968,Ultimate Investment Banking Co,ultimate-investment-banki,TRUE,200,
2147,25,51,All levels,2,Business Finance
3,198660,Tenor Banjo for Absolute Begin,tenor-banjo-for-absolute-,FALSE,0,31
67,161,34,Beginner Levels,1,Musical Instruments
[training@localhost ~]$
```

The window's taskbar at the bottom shows icons for File Manager, HDFS, training@l..., and a screenshot utility.



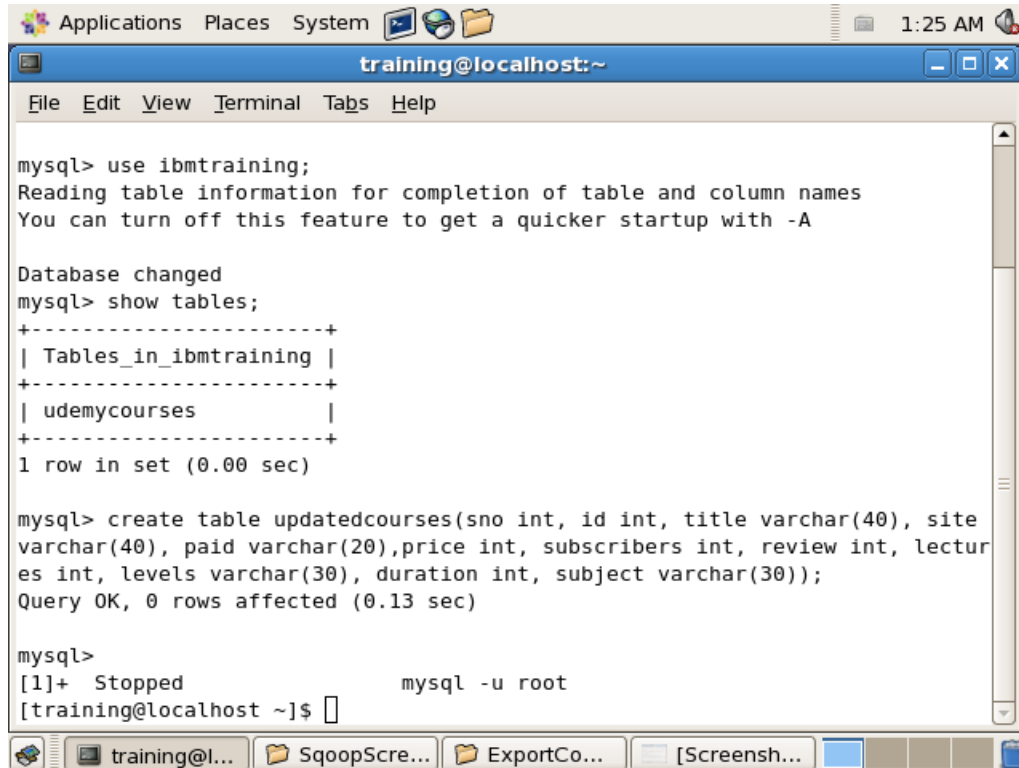
A Mozilla Firefox browser window titled 'HDFS:/user/training/udemycourses/part-m-00000 - Mozilla Firefox' with a menu bar (File, Edit, View, History, Bookmarks, Tools, Help). The address bar shows 'http://localhost.localdomain:50075'. The page content displays the file path 'File: /user/training/udemycourses/part-m-00000' and a 'Goto' field with the path '/user/training/udemycourse:' and a 'go' button. Below this, there are links for 'Go back to dir listing' and 'Advanced view/download options'. The main content area shows the same data as the terminal window:

```
2,502588,Learn GIMP,learn-gimp,TRUE,20,441,53,34,All levels,7,Graphic Design
1,1070968,Ultimate Investment Banking Co,ultimate-investment-
banki,TRUE,200,2147,25,51,All levels,2,Business Finance
3,198660,Tenor Banjo for Absolute Begin,tenor-banjo-
for-absolute-, FALSE,0,3167,161,34,Beginner Levels,1,Musical Instruments
```

The status bar at the bottom indicates 'Done' and shows taskbar icons for [Impor..., HDFS:/..., [Hadoo..., [hadoo..., and trainin....

Export Command

Load data from HDFS to RDBMS file system.



The screenshot shows a terminal window titled 'training@localhost:~' with a menu bar (File, Edit, View, Terminal, Tabs, Help) and a system bar at the top (Applications, Places, System, 1:25 AM). The terminal displays the following MySQL session:

```
mysql> use ibmtraining;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> show tables;
+-----+
| Tables_in_ibmtraining |
+-----+
| udemycourses           |
+-----+
1 row in set (0.00 sec)

mysql> create table updatedcourses(sno int, id int, title varchar(40), site
varchar(40), paid varchar(20), price int, subscribers int, review int, lectur
es int, levels varchar(30), duration int, subject varchar(30));
Query OK, 0 rows affected (0.13 sec)

mysql>
[1]+  Stopped                  mysql -u root
[training@localhost ~]$
```

The terminal window has a taskbar at the bottom with icons for 'training@l...', 'SqoopScre...', 'ExportCo...', and '[Screensh...']'.

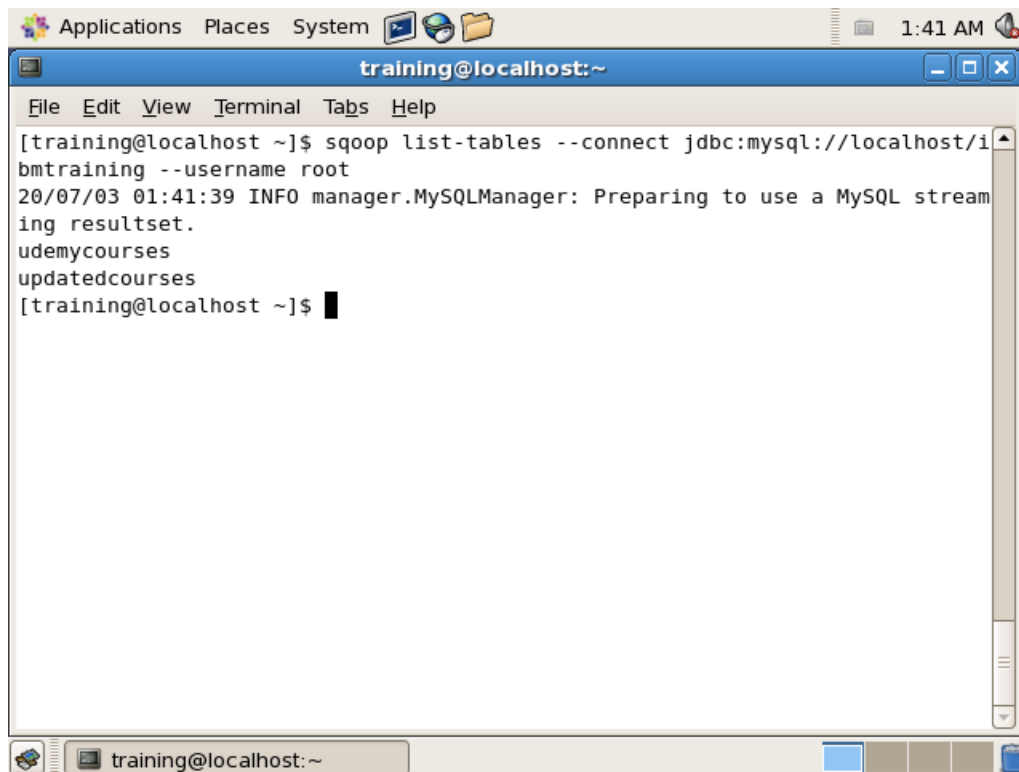
```
Applications Places System 1:35 AM
training@localhost:~
File Edit View Terminal Tabs Help
[training@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/ibmtraining --username root --table updatedcourses --export-dir /user/training/udemycourses
20/07/03 01:34:38 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
20/07/03 01:34:38 INFO tool.CodeGenTool: Beginning code generation
20/07/03 01:34:39 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `updatedcourses` AS t LIMIT 1
20/07/03 01:34:39 INFO orm.CompilationManager: HADOOP_HOME is /usr/lib/hadoop
20/07/03 01:34:39 INFO orm.CompilationManager: Found hadoop core jar at: /usr/lib/hadoop/hadoop-core.jar
20/07/03 01:34:41 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-training/compile/940e2bf2a2ad54dcc938eec0ddc5f083/updatedcourses.jar
20/07/03 01:34:41 INFO mapreduce.ExportJobBase: Beginning export of updatedcourses
20/07/03 01:34:43 INFO input.FileInputFormat: Total input paths to process : 1
20/07/03 01:34:43 INFO input.FileInputFormat: Total input paths to process : 1
20/07/03 01:34:44 INFO mapred.JobClient: Running job: job_202007030114_0001
20/07/03 01:34:45 INFO mapred.JobClient: map 0% reduce 0%
```

```
Applications Places System 1:36 AM
training@localhost:~
File Edit View Terminal Tabs Help
| Tables_in_ibmtraining |
+-----+
| udemycourses          |
| updatedcourses        |
+-----+
2 rows in set (0.00 sec)

mysql> select * from updatedcourses;
+-----+-----+-----+-----+-----+-----+-----+
| sno | id      | title                                     | site          | paid | price | subscribers | review | lectures | levels | duration | subject |
+-----+-----+-----+-----+-----+-----+-----+
| 2   | 502588 | Learn GIMP                             | learn-gimp    | TRUE | 20    | 441         | 53     | 34       | All levels |          |          |
| 7   |         | Graphic Design                         |               |      |       |             |        |          |           |          |          |
| 1   | 1070968 | Ultimate Investment Banking Co         | ultimate-investment-bank | TRUE | 200   | 2147        | 25     | 51       | All levels |          |          |
+-----+-----+-----+-----+-----+-----+-----+
```

List Tables - Sqoop Command

To list out the tables of a particular database in MySQL database server using Sqoop.



The screenshot shows a Linux desktop environment with a terminal window titled 'training@localhost:~'. The terminal displays the following command and output:

```
[training@localhost ~]$ sqoop list-tables --connect jdbc:mysql://localhost/i
bmtraining --username root
20/07/03 01:41:39 INFO manager.MySQLManager: Preparing to use a MySQL stream
ing resultset.
udemycourses
updatedcourses
[training@localhost ~]$
```

The output lists two tables: 'udemycourses' and 'updatedcourses'. The terminal window has a menu bar with 'File', 'Edit', 'View', 'Terminal', 'Tabs', and 'Help'. The system tray at the bottom shows the time as 1:41 AM and the date as 20/07/03.