

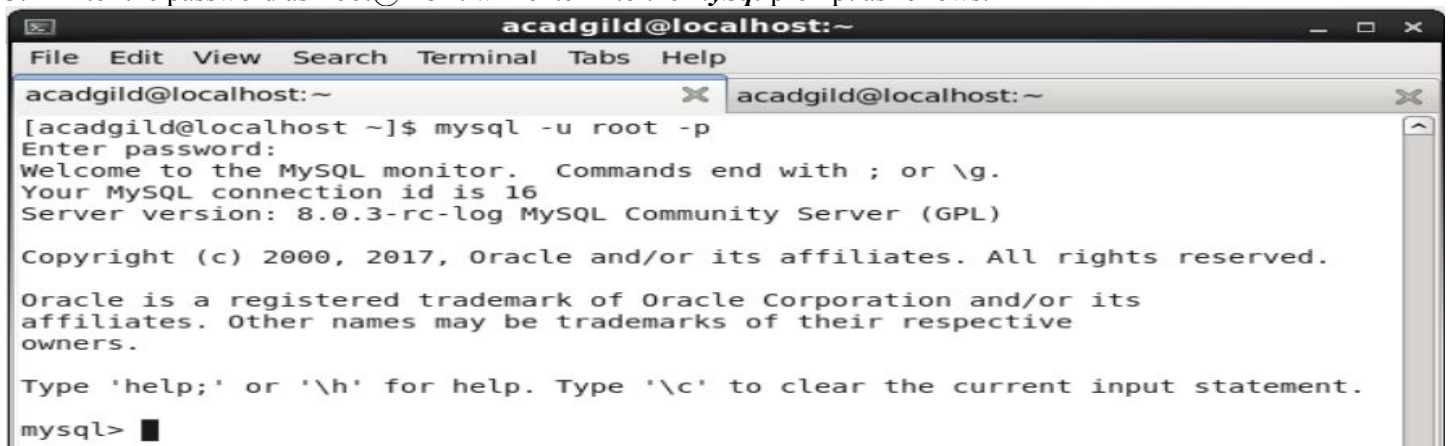
Session 6: Data Ingestion tool Sqoop & Introduction to Case Study 1.

1. Start the Hadoop daemons in the VM with the following command:
\$start-all.sh
2. Check all the hadoop daemons started in the VM with the following command:
\$jps



```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ jps  
3122 NameNode  
3572 ResourceManager  
3415 SecondaryNameNode  
3225 DataNode  
3677 NodeManager  
9421 Jps  
[acadgild@localhost ~]$
```

1. Now start the *mysql* with the following command:
\$mysql -u root -p
2. It will prompt the user to enter the password:
Enter password:
3. Enter the password as *Root@123* it will enter into the *mysql* prompt as follows:



```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
[acadgild@localhost ~]$ mysql -u root -p  
Enter password:  
Welcome to the MySQL monitor.  Commands end with ; or \g.  
Your MySQL connection id is 16  
Server version: 8.0.3-rc-log MySQL Community Server (GPL)  
  
Copyright (c) 2000, 2017, Oracle and/or its affiliates. All rights reserved.  
  
Oracle is a registered trademark of Oracle Corporation and/or its  
affiliates. Other names may be trademarks of their respective  
owners.  
  
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.  
mysql>
```

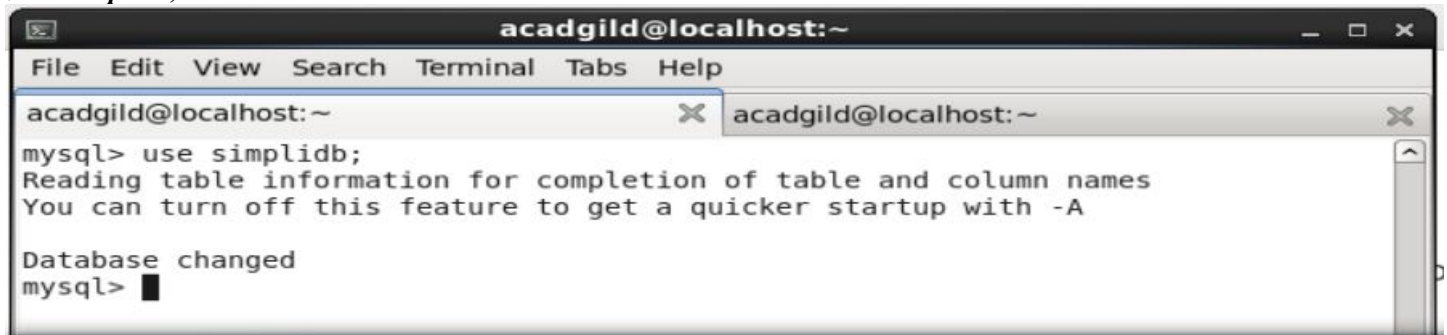
4. Now enter the command to show data base as follows:
mysql> show databases;
5. It will show the list of databases as follows:



```
mysql>  
mysql> show databases;  
+-----+  
| Database |  
+-----+  
| information_schema |  
| metastore |  
| mysql |  
| oozie |  
| performance_schema |  
| simplidb |  
| sys |  
+-----+  
7 rows in set (0.01 sec)  
  
mysql>
```

6. Now select the *simplidb* schema database as follows:

\$use simplidb;

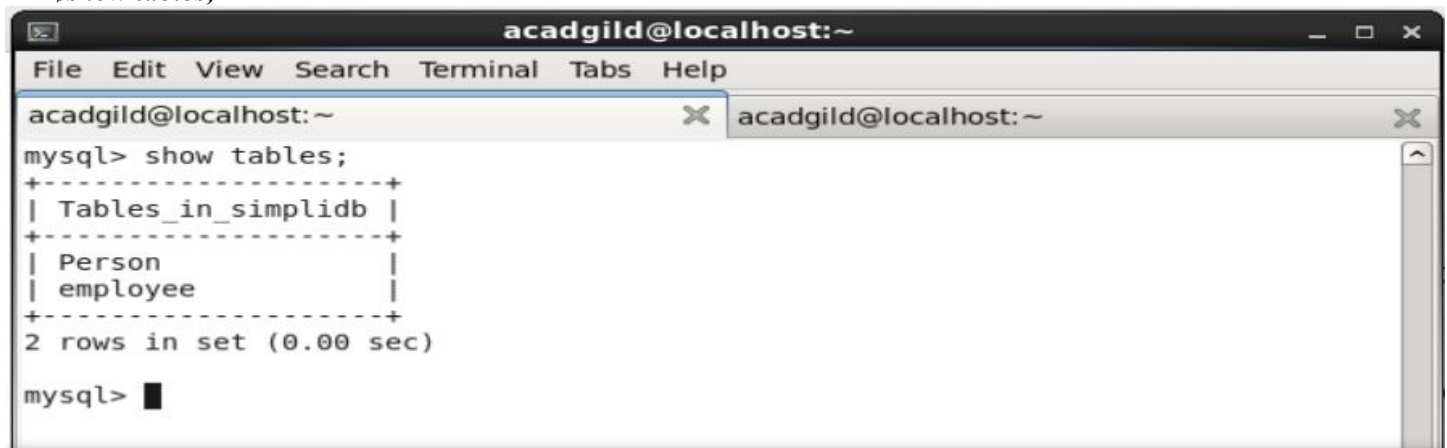


```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
mysql> use simplidb;  
Reading table information for completion of table and column names  
You can turn off this feature to get a quicker startup with -A  
  
Database changed  
mysql> █
```

7. We can observe that the message Database changed for selecting the database simplidb.

8. Now list all the tables present in the *simplidb* database, as follows:

\$show tables;



```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
mysql> show tables;  
+-----+  
| Tables_in_simplidb |  
+-----+  
| Person             |  
| employee            |  
+-----+  
2 rows in set (0.00 sec)  
  
mysql> █
```

9. Now we can observe that there are two tables with **Person** and *employee* tables present in the database *simplidb*.

10. Now for the demo from the beginning, I am going to drop the table employee to start from the beginning.

11. To drop the table we can use the command:

mysql> drop table employee;



```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
mysql> select * from employee;  
+-----+-----+  
| firstname | lastname |  
+-----+-----+  
| Prema     | Vardhan  |  
| Prema     | Vardhan  |  
+-----+-----+  
2 rows in set (0.00 sec)  
  
mysql> drop table employee;  
Query OK, 0 rows affected (0.03 sec)  
  
mysql> █
```

12. We can see the message Query OK, 0 rows affected (0.03 sec) for deletion of an employee table.

13. Now we can create the table again with the following command:

mysql> create table employee (firstname varchar(10), lastname varchar(10));

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
mysql> create table employee (firstname varchar(10), lastname varchar(10));  
Query OK, 0 rows affected (0.05 sec)  
  
mysql> show tables;  
+-----+  
| Tables_in_simplidb |  
+-----+  
| Person              |  
| employee             |  
+-----+  
2 rows in set (0.00 sec)  
  
mysql> █
```

14. We can see that the table is create with the name employee.

15. Now we can add records into the table employee using the following command:

\$insert into employee values ("Prema", "Vardhan");

\$insert into employee values ("Bhaskar", "Reddy");

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
mysql> insert into employee values ("Prema","Vardhan");  
Query OK, 1 row affected (0.01 sec)  
  
mysql> insert into employee values ("Bhaskar","Reddy");  
Query OK, 1 row affected (0.01 sec)  
  
mysql> █
```

16. Now list the records present in the employee table as follows:

*\$select * from employee;*

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~  
mysql> insert into employee values ("Prema","Vardhan");  
Query OK, 1 row affected (0.01 sec)  
  
mysql> insert into employee values ("Bhaskar","Reddy");  
Query OK, 1 row affected (0.01 sec)  
  
mysql> select * from employee;  
+-----+-----+  
| firstname | lastname |  
+-----+-----+  
| Prema     | Vardhan  |  
| Bhaskar   | Reddy    |  
+-----+-----+  
2 rows in set (0.00 sec)  
  
mysql> █
```

17. Now we can observe that there are two records present in the employee table.

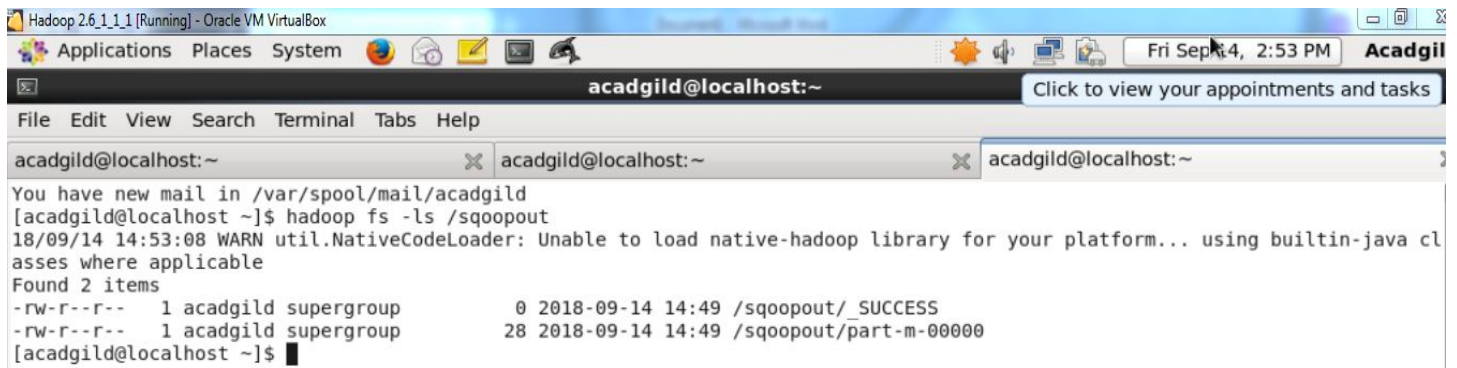
18. Now we can use **sqoop** tool to import the table **employee** into hadoop file system using the following command:

\$sqoop import --connect jdbc:mysql://localhost/simplidb --table employee --username root --password Root@123 --target-dir /sqoopout -m 1


```
Hadoop 2.6.1_1.1 [Running] - Oracle VM VirtualBox
Applications Places System
acacgild@localhost:~
Click to view your appointments and tasks
File Edit View Search Terminal Tabs Help
acacgild@localhost:~
acacgild@localhost:~
acacgild@localhost:~
[acacgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/simplidb --table employee --username root --password Root@123 --target-dir /sqoopout -m 1
Warning: /home/acacgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acacgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/09/14 14:48:32 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
18/09/14 14:48:32 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
18/09/14 14:48:32 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/09/14 14:48:32 INFO tool.CodeGenTool: Beginning code generation
Fri Sep 14 14:48:32 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
18/09/14 14:48:33 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
18/09/14 14:48:33 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
18/09/14 14:48:33 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acacgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acacgild/compile/2351bc3e7021e1a3527484628ce8d743/employee.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/09/14 14:48:36 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acacgild/compile/2351bc3e7021e1a3527484628ce8d743/employee.jar
18/09/14 14:48:36 WARN manager.MySQLManager: It looks like you are importing from mysql.
18/09/14 14:48:36 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
18/09/14 14:48:36 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
18/09/14 14:48:36 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
18/09/14 14:48:36 INFO mapreduce.ImportJobBase: Beginning import of employee
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acacgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acacgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/09/14 14:48:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java class
Hadoop 2.6.1_1.1 [Running] - Oracle VM VirtualBox
Applications Places System
acacgild@localhost:~
Click to view your appointments and tasks
File Edit View Search Terminal Tabs Help
acacgild@localhost:~
acacgild@localhost:~
acacgild@localhost:~
FILE: Number of bytes written=127705
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=87
HDFS: Number of bytes written=28
HDFS: Number of read operations=4
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=5806
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=5806
Total vcore-milliseconds taken by all map tasks=5806
Total megabyte-milliseconds taken by all map tasks=5945344
Map-Reduce Framework
Map input records=2
Map output records=2
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=128
CPU time spent (ms)=2340
Physical memory (bytes) snapshot=170725376
Virtual memory (bytes) snapshot=2086739968
Total committed heap usage (bytes)=93323264
File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=28
18/09/14 14:49:03 INFO mapreduce.ImportJobBase: Transferred 28 bytes in 25.3753 seconds (1.1034 bytes/sec)
18/09/14 14:49:03 INFO mapreduce.ImportJobBase: Retrieved 2 records.
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$
```

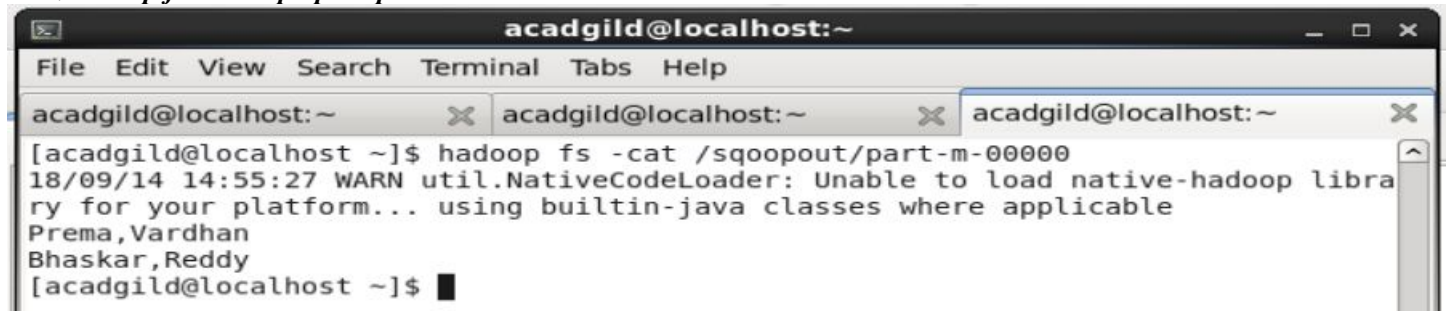
19. Now the output can be present in the `/sqoopout` folder present in the file system:

Shadoop fs -ls /sqoopout



```
Hadoop 2.6.1_1.1 [Running] - Oracle VM VirtualBox
Applications Places System
acadgild@localhost:~
File Edit View Search Terminal Tabs Help
acadgild@localhost:~ acadgild@localhost:~ acadgild@localhost:~
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -ls /sqoopout
18/09/14 14:53:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2018-09-14 14:49 /sqoopout/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 28 2018-09-14 14:49 /sqoopout/part-m-00000
[acadgild@localhost ~]$
```

20. Now check the output present in the mapper output file: `/sqoopout/part-m-00000`. As follows:
\$hadoop fs -cat /sqoopout/part-m-00000

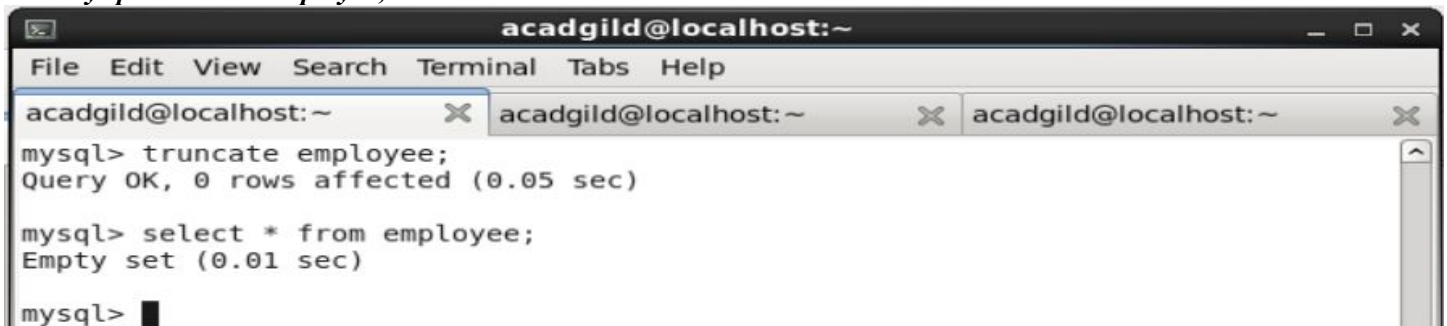


```
acacgild@localhost:~
File Edit View Search Terminal Tabs Help
acacgild@localhost:~ acadgild@localhost:~ acadgild@localhost:~
[acadgild@localhost ~]$ hadoop fs -cat /sqoopout/part-m-00000
18/09/14 14:55:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Prema,Vardhan
Bhaskar,Reddy
[acadgild@localhost ~]$
```

21. With this we can observe that the table is imported from mysql to hadoop.

Task 1: Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table.

1. The above demo is used to import the table ***employee*** from the mysql to sqoop.
2. Now we are exporting the same employee records into the employee table in mysql.
3. To do that first we need to create a new database, or we can use an existing database also. In the above example we used the database as ***simplidb***.
4. Create table named ***employee*** for us the table employee is also present in the database ***simplidb***.
5. Now truncate if any records present the employee table to make table empty, as follows:
mysql> truncate employee;



```
acacgild@localhost:~
File Edit View Search Terminal Tabs Help
acacgild@localhost:~ acadgild@localhost:~ acadgild@localhost:~
mysql> truncate employee;
Query OK, 0 rows affected (0.05 sec)

mysql> select * from employee;
Empty set (0.01 sec)

mysql>
```

6. We also checked that if there are any records present, we got an Empty set represents no records present in the table
7. Now give a command, ***describe <tablename>*** to show the various fields and types of it. This will help in comparing the type of data present inside HDFS which is ready to be mapped.


```

acadgild@localhost:~
File Edit View Search Terminal Tabs Help
acadgild@localhost:~ X acadgild@localhost:~ X acadgild@localhost:~ X
mysql> describe employee;
+-----+-----+-----+-----+-----+-----+
| Field      | Type      | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| firstname  | varchar(10) | YES  |     | NULL    |       |
| lastname   | varchar(10) | YES  |     | NULL    |       |
+-----+-----+-----+-----+-----+-----+
2 rows in set (0.01 sec)

mysql> █

```

8. Now the file in the HDFS must have the same format as that of MYSQL table, to enable the mapping of the data. The file will be like this:

```

acadgild@localhost:~
File Edit View Search Terminal Tabs Help
acadgild@localhost:~ X acadgild@localhost:~ X acadgild@localhost:~ X
[acadgild@localhost ~]$ hadoop fs -cat /sqoop_mysql/input.txt
18/09/14 17:07:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Prema,Vardhan
Bhaskar,Reddy
[acadgild@localhost ~]$ █

```

9. Now we have a file named *input.txt* in HDFS with two records of *prema*, *Vardhan* and *Bhaskar*, *Reddy* with first name and second name separated.
10. Now export the *input.txt* file from HDFS to MySQL, as follows:

\$sqoop export --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table employee --export-dir /sqoop_mysql/ -m 1

```

Hadoop 2.6.1_1.1 [Running] - Oracle VM VirtualBox
Applications Places System
acadgild@localhost:~
File Edit View Search Terminal Tabs Help
acadgild@localhost:~ X acadgild@localhost:~ X acadgild@localhost:~ X
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table employee --export-dir /sqoop_mysql/ -m 1
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/09/14 17:19:19 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
18/09/14 17:19:19 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
18/09/14 17:19:19 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/09/14 17:19:19 INFO tool.CodeGenTool: Beginning code generation
Fri Sep 14 17:19:19 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.
18/09/14 17:19:20 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
18/09/14 17:19:20 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
18/09/14 17:19:20 INFO orm.CompilationManager: HADOOP MAPRED HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/7b648cc833083a27f67b6f7f60c16ab6/employee.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/09/14 17:19:22 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/7b648cc833083a27f67b6f7f60c16ab6/employee.jar
18/09/14 17:19:22 INFO mapreduce.ExportJobBase: Beginning export of employee
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/09/14 17:19:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/09/14 17:19:23 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/09/14 17:19:24 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative

```

```

FILE: Number of bytes read=0
FILE: Number of bytes written=127537
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=155
HDFS: Number of bytes written=0
HDFS: Number of read operations=4
HDFS: Number of large read operations=0
HDFS: Number of write operations=0
Job Counters
  Launched map tasks=1
  Data-local map tasks=1
  Total time spent by all maps in occupied slots (ms)=5499
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=5499
  Total vcore-milliseconds taken by all map tasks=5499
  Total megabyte-milliseconds taken by all map tasks=5630976
Map-Reduce Framework
  Map input records=2
  Map output records=2
  Input split bytes=124
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=103
  CPU time spent (ms)=2150
  Physical memory (bytes) snapshot=178573312
  Virtual memory (bytes) snapshot=2084880384
  Total committed heap usage (bytes)=93323264
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=0
18/09/14 17:19:47 INFO mapreduce.ExportJobBase: Transferred 155 bytes in 22.4289 seconds (6.9107 bytes/sec)
18/09/14 17:19:47 INFO mapreduce.ExportJobBase: Exported 2 records.
[acadgild@localhost ~]$

```

11. The message at the end represents “Exported 2 records” into the simplidb database into the table employee.

12. Now we will verify the list of records present in the employee table as:

*mysql> select * from employee;*

```

acadgild@localhost: ~
File Edit View Search Terminal Tabs Help
acadgild@localhost: ~
mysql> select * from employee;
+-----+-----+
| firstname | lastname |
+-----+-----+
| Prema    | Vardhan  |
| Bhaskar   | Reddy    |
+-----+-----+
2 rows in set (0.00 sec)

mysql>

```

13. Here we can observe that the two record are exported from HDFS to mysql.

Task 2: Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table with parameter `person_id = 3`.

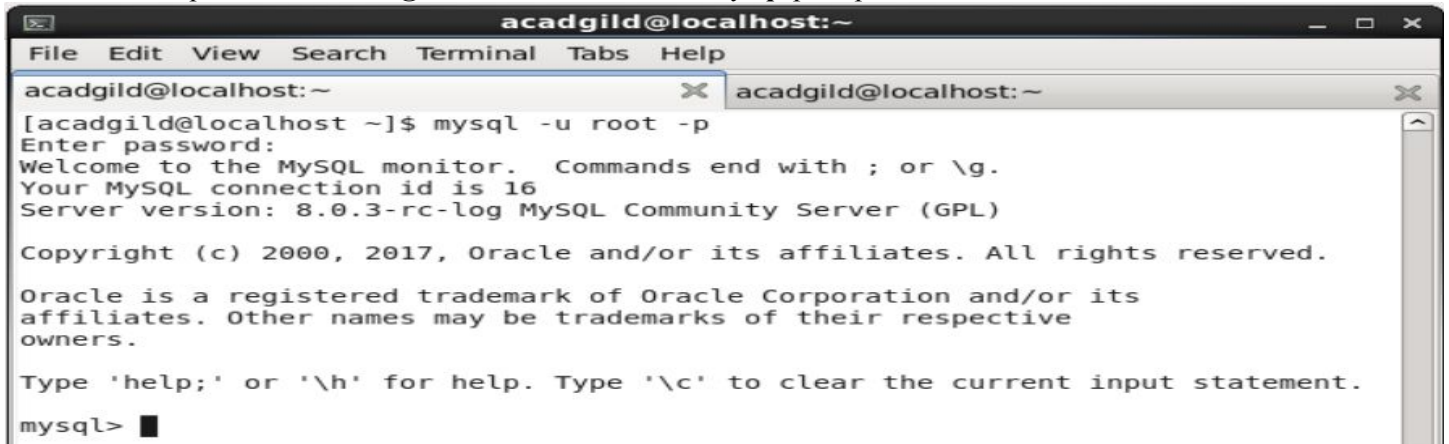
1. Now start the *mySql* with the following command:

\$mysql -u root -p

2. It will prompt the user to enter the password:

Enter password:

3. Enter the password as ***Root@123*** it will enter into the *mySql* prompt as follows:

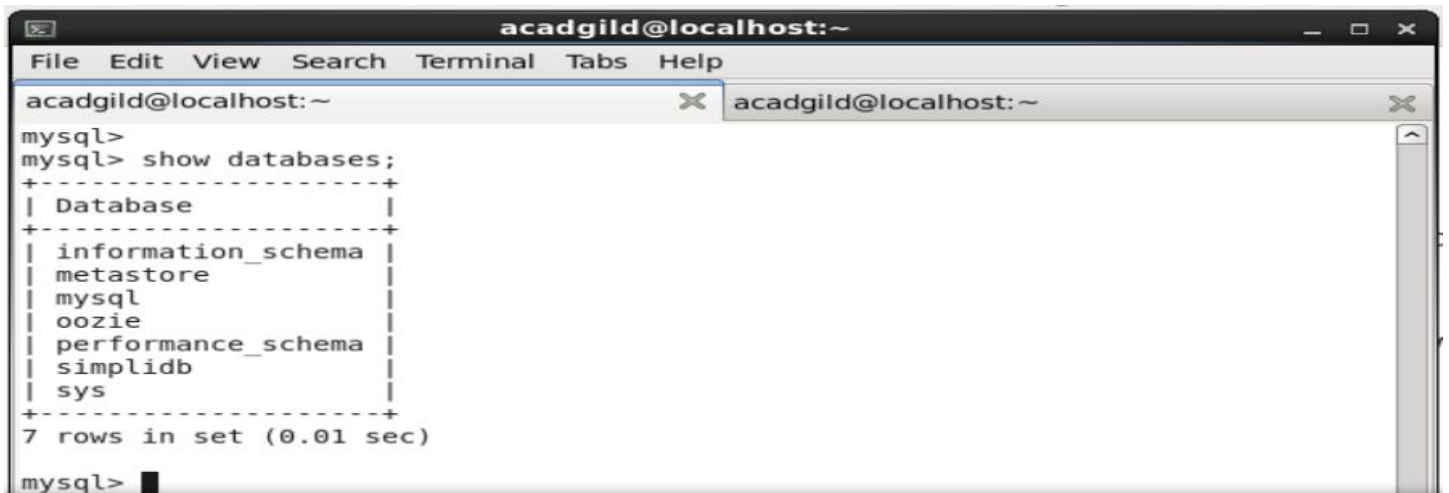
A terminal window titled 'acadgild@localhost:~' showing the MySQL login process. The user enters 'mysql -u root -p' and the password 'Root@123'. The MySQL prompt 'mysql>' is shown at the bottom.

```
acadgild@localhost:~  
[acadgild@localhost ~]$ mysql -u root -p  
Enter password:  
Welcome to the MySQL monitor.  Commands end with ; or \g.  
Your MySQL connection id is 16  
Server version: 8.0.3-rc-log MySQL Community Server (GPL)  
  
Copyright (c) 2000, 2017, Oracle and/or its affiliates. All rights reserved.  
  
Oracle is a registered trademark of Oracle Corporation and/or its  
affiliates. Other names may be trademarks of their respective  
owners.  
  
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.  
mysql>
```

4. Now enter the command to show data base as follows:

mysql> show databases;

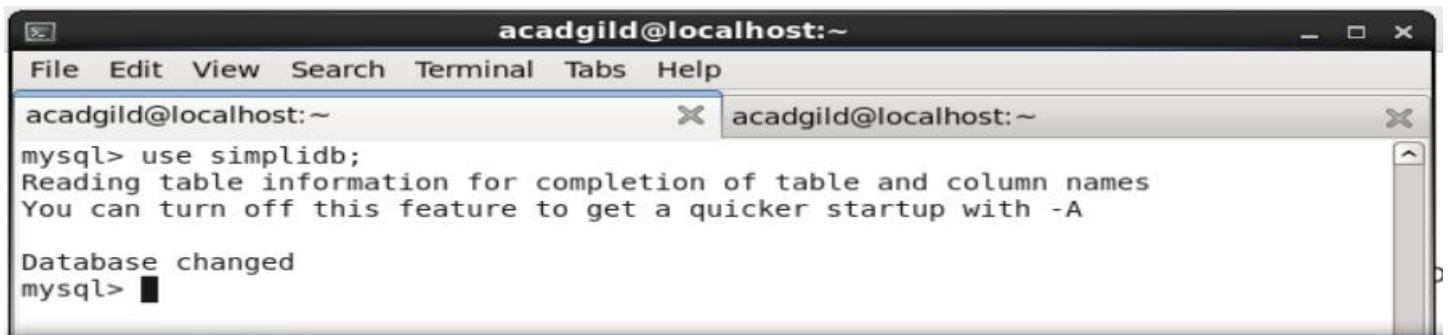
5. It will show the list of databases as follows:

A terminal window titled 'acadgild@localhost:~' showing the output of the 'show databases;' command. It lists seven databases: information_schema, metastore, mysql, oozie, performance_schema, simplidb, and sys.

```
mysql>  
mysql> show databases;  
+-----+  
| Database |  
+-----+  
| information_schema |  
| metastore |  
| mysql |  
| oozie |  
| performance_schema |  
| simplidb |  
| sys |  
+-----+  
7 rows in set (0.01 sec)  
mysql>
```

6. Now select the ***simplidb*** schema database as follows:

mysql> use simplidb;

A terminal window titled 'acadgild@localhost:~' showing the output of the 'use simplidb;' command. It displays 'Database changed' and the MySQL prompt 'mysql>'.

```
mysql> use simplidb;  
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>
```

7. We can observe that the message Database changed for selecting the database simplidb.

8. Now list all the tables present in the ***simplidb*** database, as follows:

mysql> show tables;


```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~  
mysql> show tables;  
+-----+  
| Tables_in_simplidb |  
+-----+  
| Person  
| employee  
+-----+  
2 rows in set (0.00 sec)  
  
mysql> █
```

9. Now we can observe that there are two tables with **Person** and **employee** tables present in the database **simplidb**.

10. Print the list of records present in the Person table with the following command:

mysql> select * from Person;

```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~  
mysql> mysql> select * from Person;  
+-----+  
| person_id | lname | fname | area | city |  
+-----+  
| 1 | Rizvi | Syed | Whitefield | Bangalore |  
| 2 | Rizvi | Syed | Whitefield | Bangalore |  
| 3 | Rizvi | Syed | Whitefield | Bangalore |  
| 4 | Rizvi | Syed | Whitefield | Bangalore |  
| 5 | Rizvi | Syed | Whitefield | Bangalore |  
| 6 | Rizvi | Syed | Whitefield | Bangalore |  
| 7 | Rizvi | Syed | Whitefield | Bangalore |  
| 8 | Rizvi | Syed | Whitefield | Bangalore |  
| 9 | Rizvi | Syed | Whitefield | Bangalore |  
| 10 | Rizvi | Syed | Whitefield | Bangalore |  
+-----+  
10 rows in set (0.00 sec)  
  
mysql> █
```

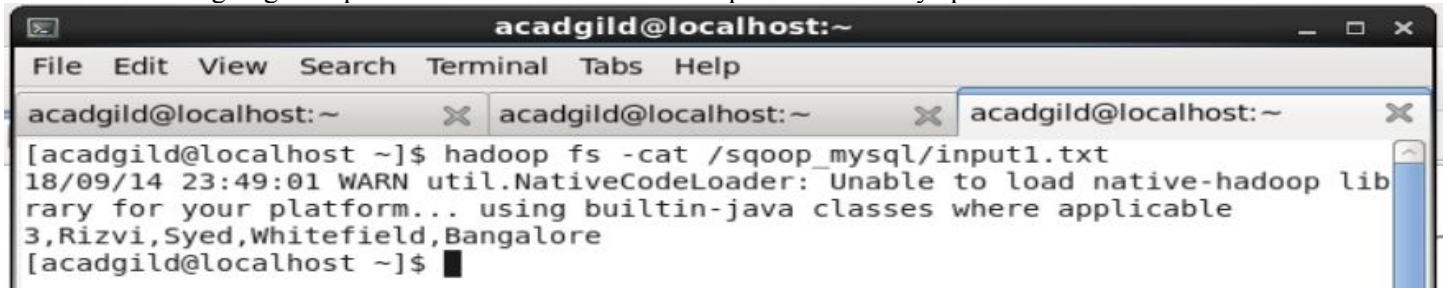
11. Now delete a record from the table **Person**;

mysql> delete from Person where person_id=3;

```
mysql> delete from Person where person_id=3;  
Query OK, 1 row affected (0.00 sec)  
  
mysql> select * from Person;  
+-----+  
| person_id | lname | fname | area | city |  
+-----+  
| 1 | Rizvi | Syed | Whitefield | Bangalore |  
| 2 | Rizvi | Syed | Whitefield | Bangalore |  
| 4 | Rizvi | Syed | Whitefield | Bangalore |  
| 5 | Rizvi | Syed | Whitefield | Bangalore |  
| 6 | Rizvi | Syed | Whitefield | Bangalore |  
| 7 | Rizvi | Syed | Whitefield | Bangalore |  
| 8 | Rizvi | Syed | Whitefield | Bangalore |  
| 9 | Rizvi | Syed | Whitefield | Bangalore |  
| 10 | Rizvi | Syed | Whitefield | Bangalore |  
+-----+  
9 rows in set (0.00 sec)  
  
mysql> █
```

12. And here we can observe that after deletion, we used command *select * from Person;* to display all the records present in the Person table. But the record with person_id=3 is deleted.

1. Now in the hadoop file system we can create a file named with */sqoop_mysql/input1.txt* which contains a record that we are going to export to the table Person which is present in the mysql database.

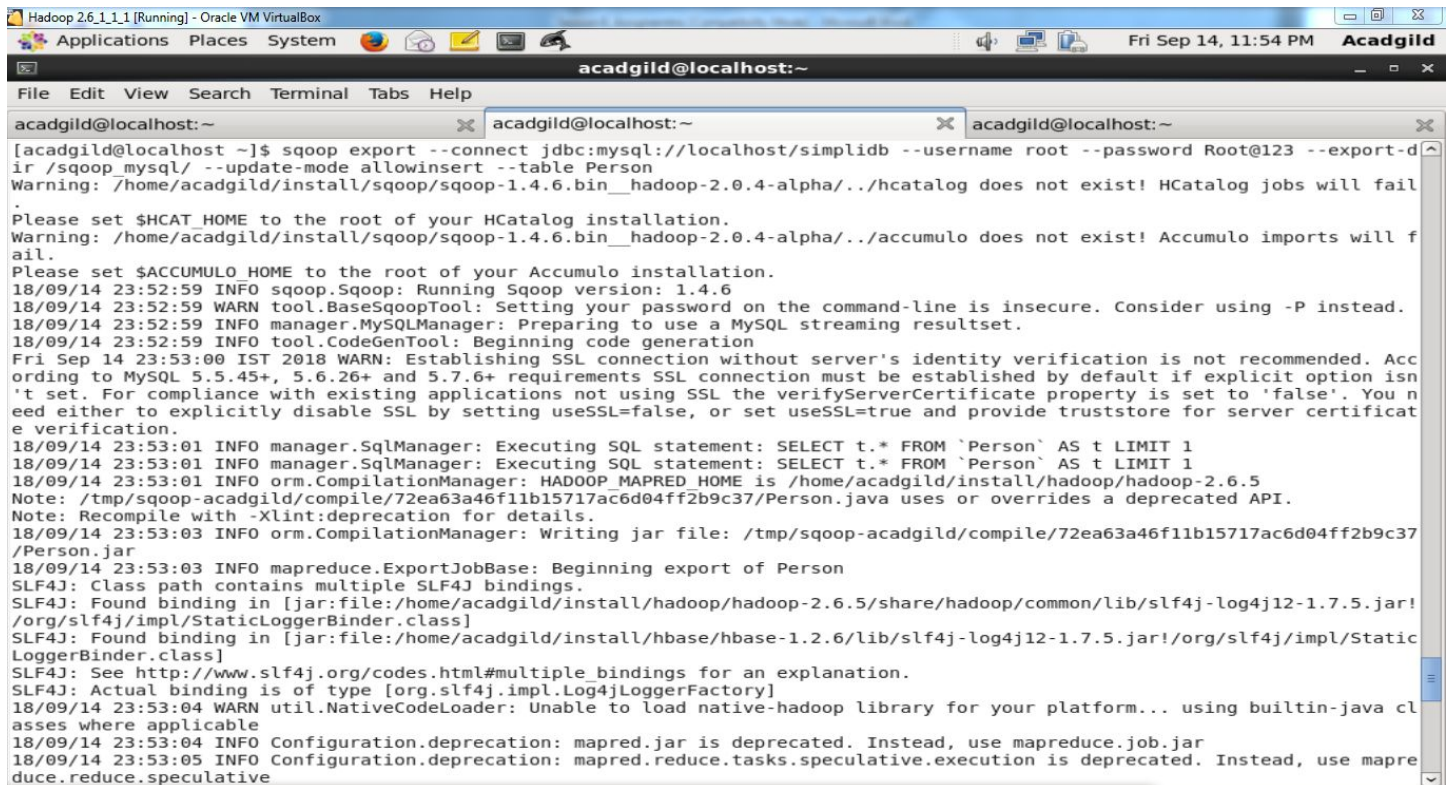


```
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~ acadgild@localhost:~  
[acadgild@localhost ~]$ hadoop fs -cat /sqoop_mysql/input1.txt  
18/09/14 23:49:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
3,Rizvi,Syed,Whitefield,Bangalore  
[acadgild@localhost ~]$
```

2. We can observe that the record 3, *Rizvi, Syed, Whitefield, Bangalore* is present in the directory */sqoop_mysql/input1.txt* file.

3. Now export the file to update the record into the Person table with the following command:

\$sqoop export --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --export-dir /sqoop_mysql/ --update-mode allowinsert --table Person



```
Hadoop 2.6.1_1.1 [Running] - Oracle VM VirtualBox  
Applications Places System  
acadgild@localhost:~  
File Edit View Search Terminal Tabs Help  
acadgild@localhost:~ acadgild@localhost:~ acadgild@localhost:~  
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --export-dir /sqoop_mysql/ --update-mode allowinsert --table Person  
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./hcatalog does not exist! HCatalog jobs will fail.  
Please set $HCAT_HOME to the root of your HCatalog installation.  
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/./accumulo does not exist! Accumulo imports will fail.  
Please set $ACCUMULO_HOME to the root of your Accumulo installation.  
18/09/14 23:52:59 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6  
18/09/14 23:52:59 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.  
18/09/14 23:52:59 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.  
18/09/14 23:52:59 INFO tool.CodeGenTool: Beginning code generation  
Fri Sep 14 23:53:00 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificate verification.  
18/09/14 23:53:01 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Person` AS t LIMIT 1  
18/09/14 23:53:01 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Person` AS t LIMIT 1  
18/09/14 23:53:01 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5  
Note: /tmp/sqoop-acadgild/compile/72ea63a46f1b15717ac6d04ff2b9c37/Person.java uses or overrides a deprecated API.  
Note: Recompile with -Xlint:deprecation for details.  
18/09/14 23:53:03 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/72ea63a46f1b15717ac6d04ff2b9c37/Person.jar  
18/09/14 23:53:03 INFO mapreduce.ExportJobBase: Beginning export of Person  
SLF4J: Class path contains multiple SLF4J bindings.  
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]  
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]  
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.  
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]  
18/09/14 23:53:04 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
18/09/14 23:53:04 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar  
18/09/14 23:53:05 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative
```

```

FILE: Number of bytes written=510176
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=669
HDFS: Number of bytes written=0
HDFS: Number of read operations=19
HDFS: Number of large read operations=0
HDFS: Number of write operations=0
Job Counters
  Launched map tasks=4
  Data-local map tasks=4
  Total time spent by all maps in occupied slots (ms)=79999
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=79999
  Total vcore-milliseconds taken by all map tasks=79999
  Total megabyte-milliseconds taken by all map tasks=81918976
Map-Reduce Framework
  Map input records=1
  Map output records=1
  Input split bytes=561
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=1365
  CPU time spent (ms)=10500
  Physical memory (bytes) snapshot=668778496
  Virtual memory (bytes) snapshot=8344002560
  Total committed heap usage (bytes)=369098752
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=0
18/09/14 23:53:43 INFO mapreduce.ExportJobBase: Transferred 669 bytes in 38.1442 seconds (17.5387 bytes/sec)
18/09/14 23:53:43 INFO mapreduce.ExportJobBase: Exported 1 records.
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$

```

4. We can observe that at the end the message says 'Exported 1 Records' show the record is updated in the Person table.
5. We can see the Person table to check the update with the following command:
mysql>select * from Person;

```

mysql> delete from Person where person_id=3;
Query OK, 1 row affected (0.00 sec)

mysql> select * from Person;
+-----+-----+-----+-----+-----+
| person_id | lname | fname | area | city |
+-----+-----+-----+-----+-----+
| 1 | Rizvi | Syed | Whitefield | Bangalore |
| 2 | Rizvi | Syed | Whitefield | Bangalore |
| 4 | Rizvi | Syed | Whitefield | Bangalore |
| 5 | Rizvi | Syed | Whitefield | Bangalore |
| 6 | Rizvi | Syed | Whitefield | Bangalore |
| 7 | Rizvi | Syed | Whitefield | Bangalore |
| 8 | Rizvi | Syed | Whitefield | Bangalore |
| 9 | Rizvi | Syed | Whitefield | Bangalore |
| 10 | Rizvi | Syed | Whitefield | Bangalore |
+-----+-----+-----+-----+-----+
9 rows in set (0.00 sec)

mysql> select * from Person;
+-----+-----+-----+-----+-----+
| person_id | lname | fname | area | city |
+-----+-----+-----+-----+-----+
| 1 | Rizvi | Syed | Whitefield | Bangalore |
| 2 | Rizvi | Syed | Whitefield | Bangalore |
| 3 | Rizvi | Syed | Whitefield | Bangalore |
| 4 | Rizvi | Syed | Whitefield | Bangalore |
| 5 | Rizvi | Syed | Whitefield | Bangalore |
| 6 | Rizvi | Syed | Whitefield | Bangalore |
| 7 | Rizvi | Syed | Whitefield | Bangalore |
| 8 | Rizvi | Syed | Whitefield | Bangalore |
| 9 | Rizvi | Syed | Whitefield | Bangalore |
| 10 | Rizvi | Syed | Whitefield | Bangalore |
+-----+-----+-----+-----+-----+
10 rows in set (0.00 sec)

mysql>

```

6. Here we can observe that before update the table contains only 9 rows, and after update, the table contains 10 rows with updating record person_id=3.