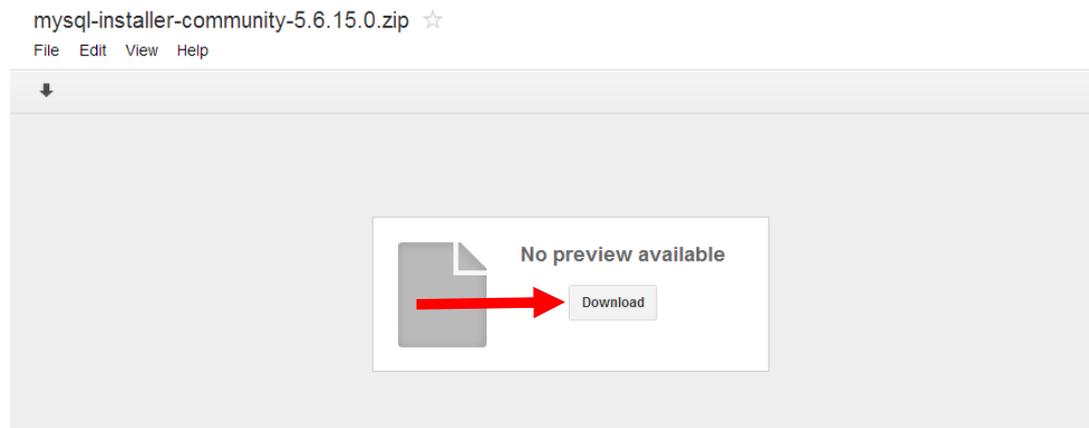


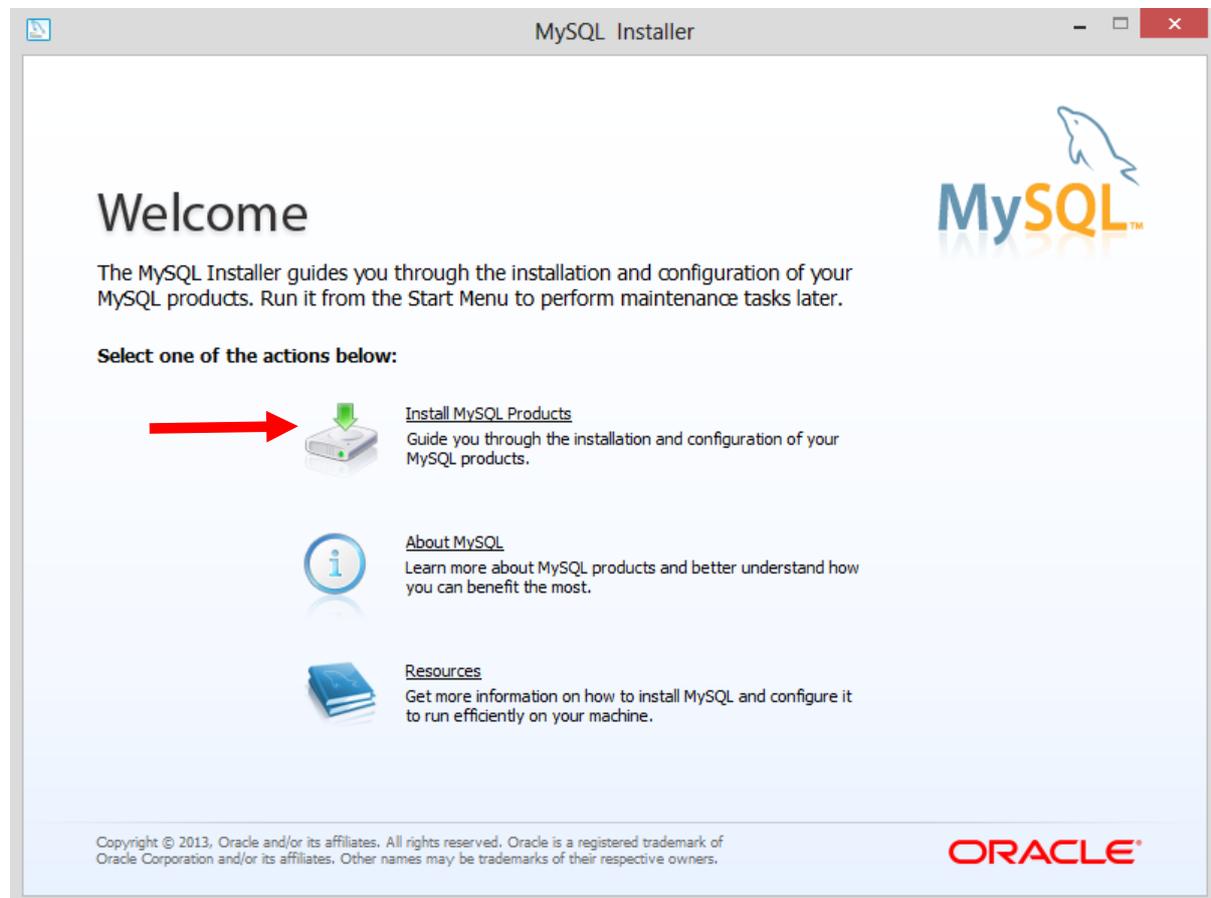
# MySQL to HDFS – Using Sqoop

1 - Download **MySQL Installer** and Unzip it:

<https://drive.google.com/file/d/0B2-rICGKD40NangwRGdLUXg2REE/edit?usp=sharing>

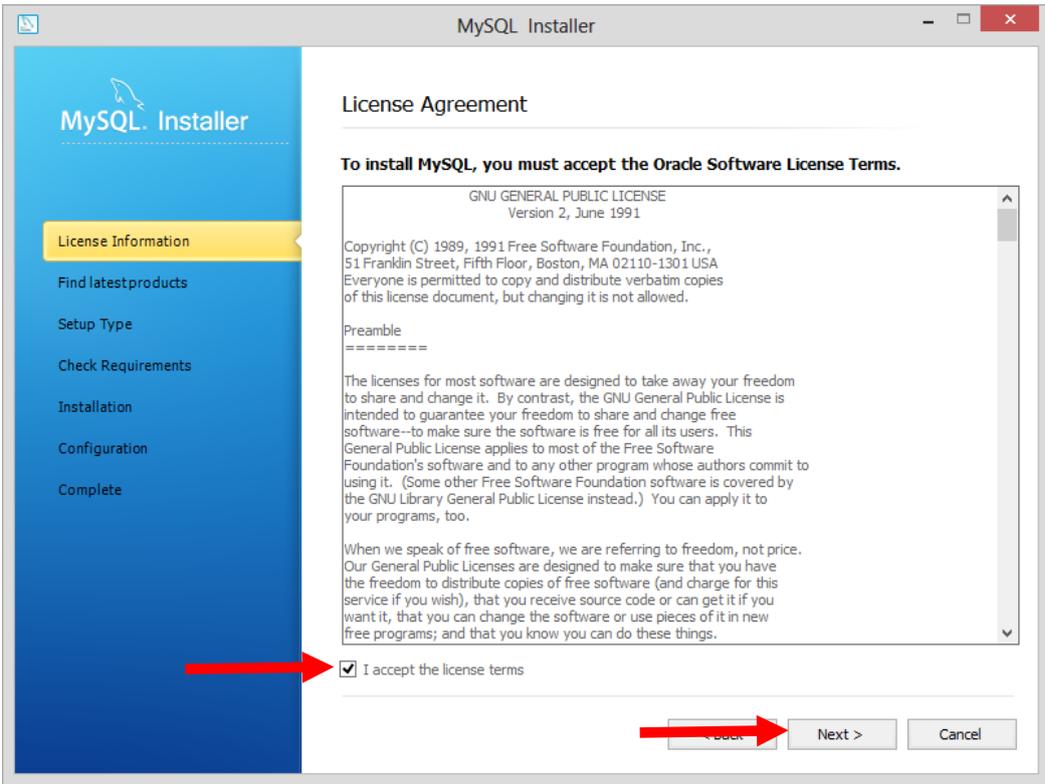


2 - Double click the extracted file and click on **Install MySQL Products**:

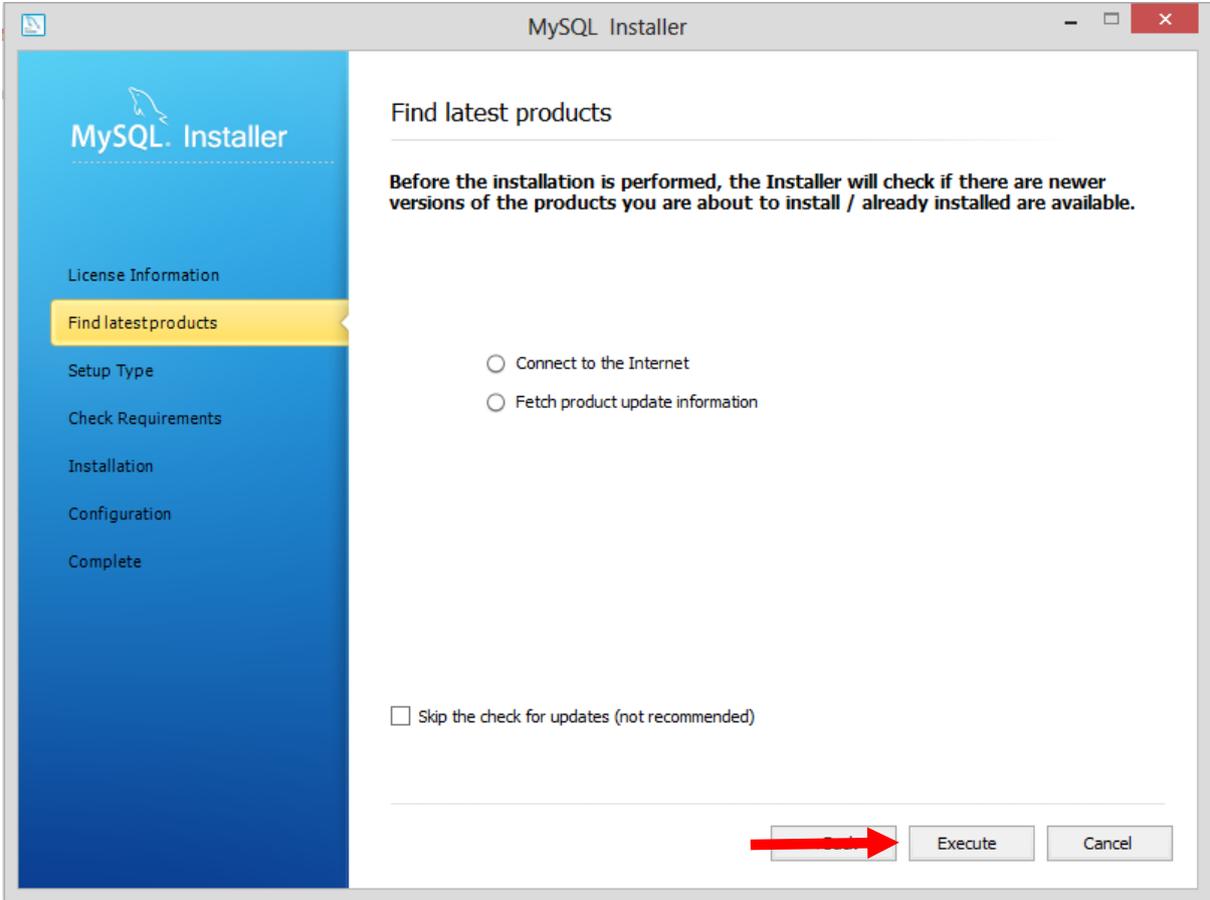


# MySQL to HDFS – Using Sqoop

3 - Click Next:

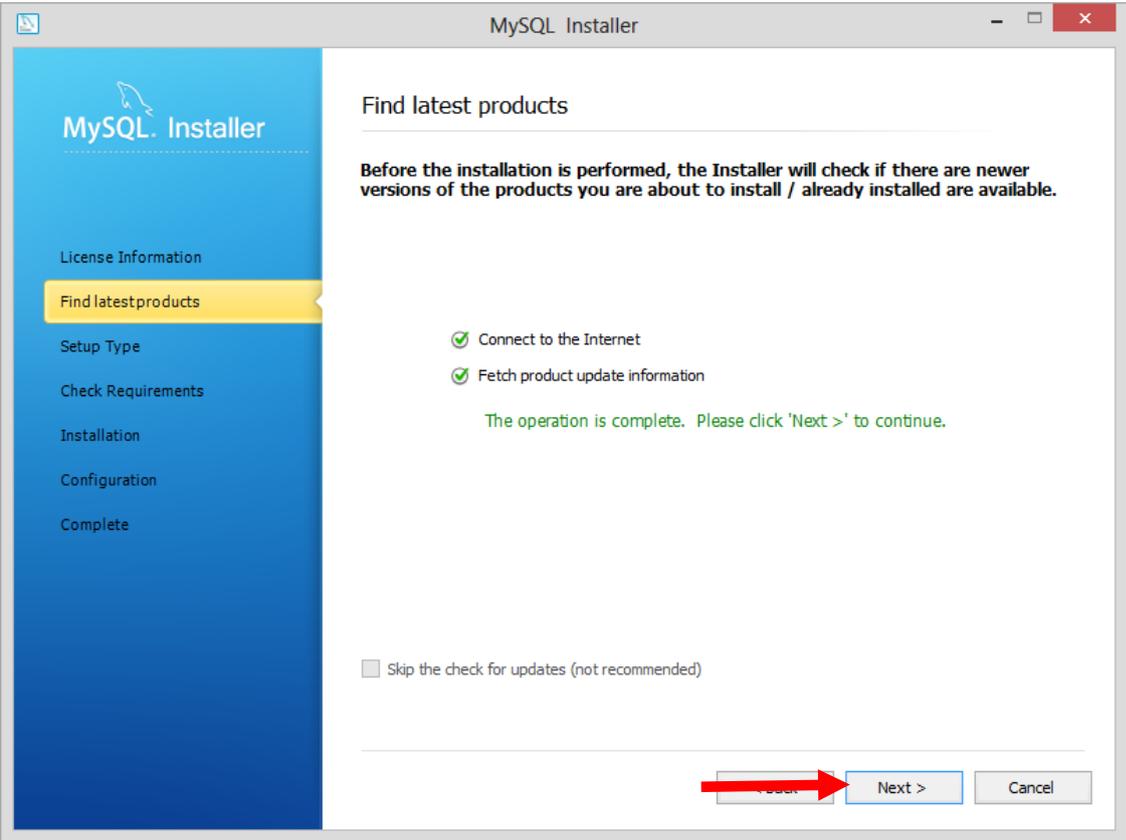


4 - Click Execute:

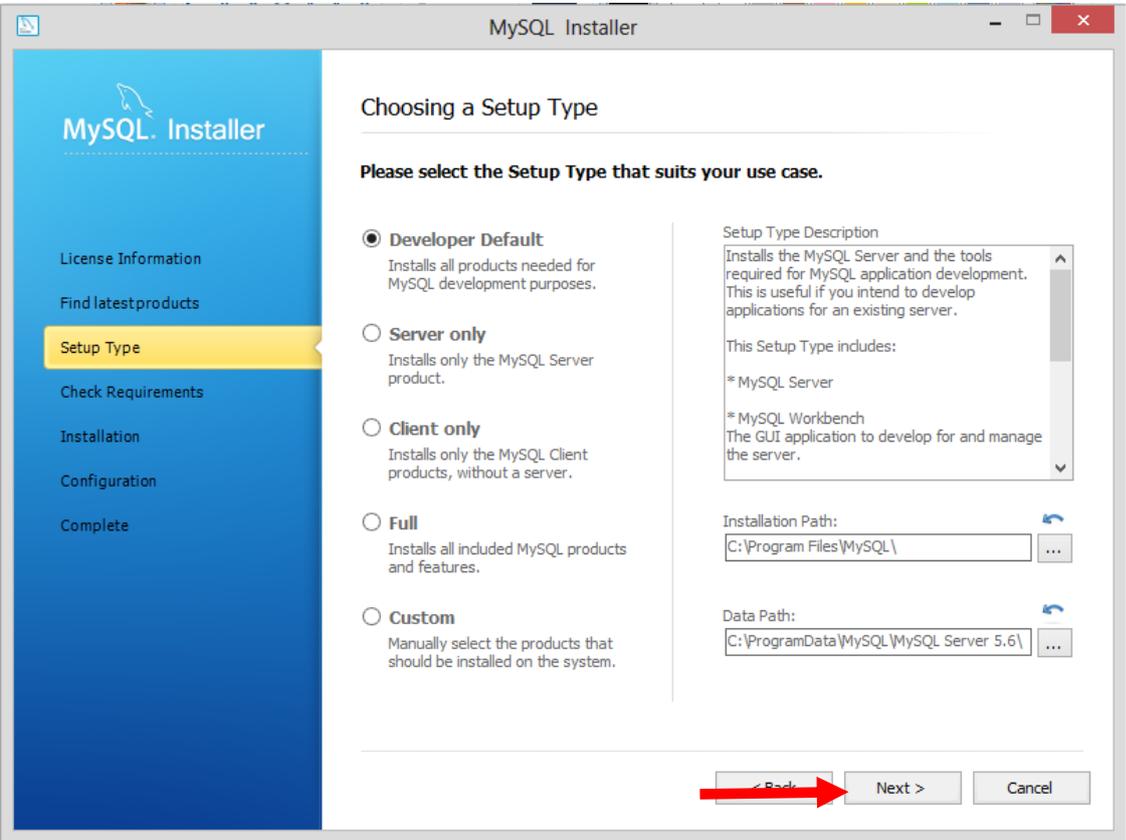


# MySQL to HDFS – Using Sqoop

5 - Click Next:

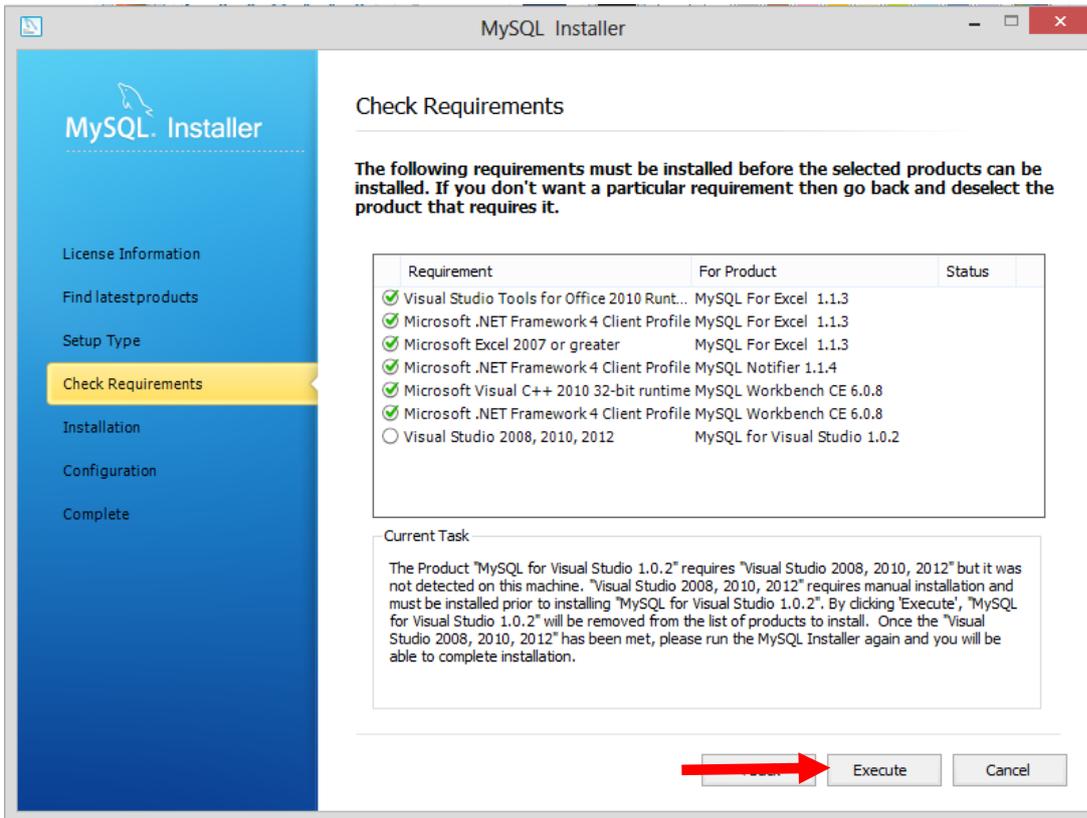


6 - Click Next:

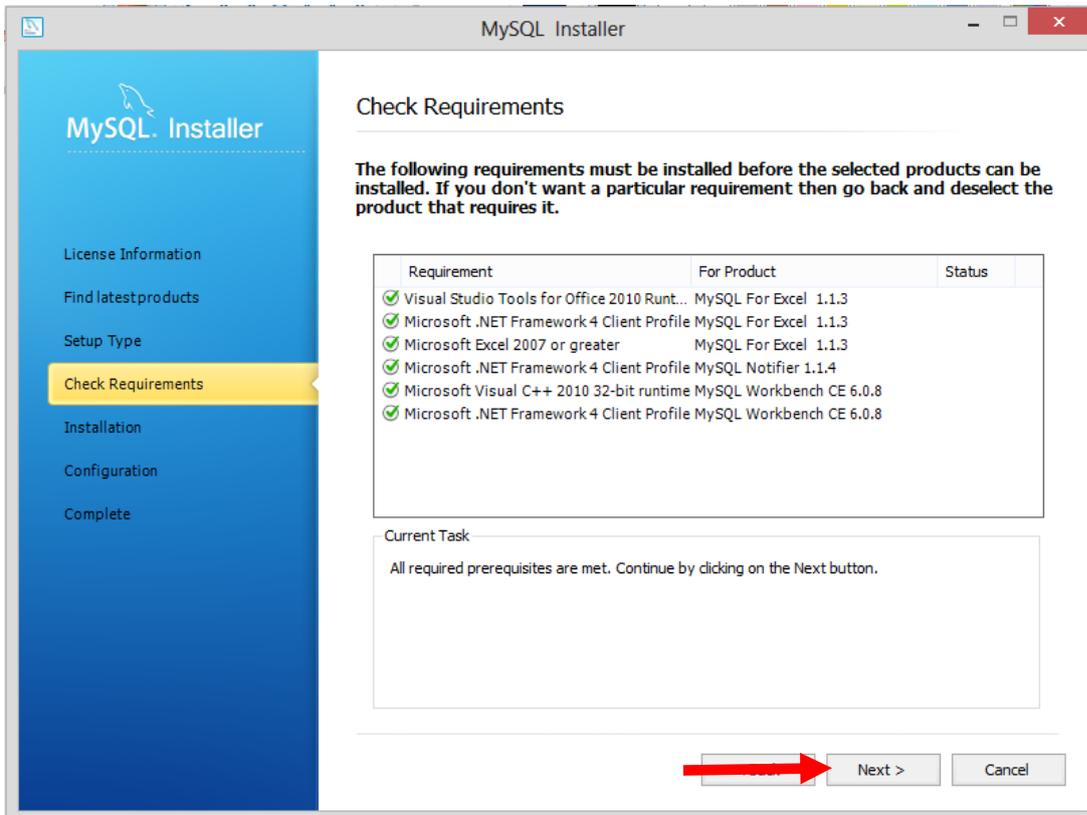


# MySQL to HDFS – Using Sqoop

7 - Click Execute:

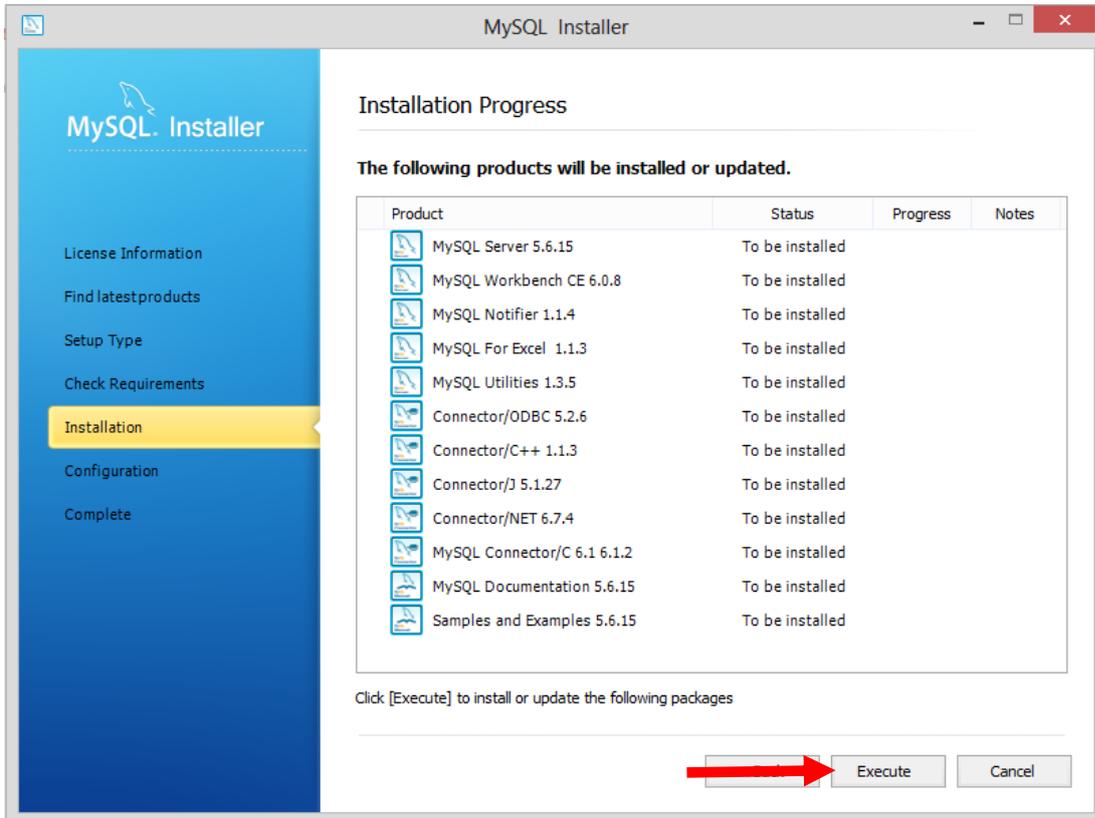


8 - Click Next:

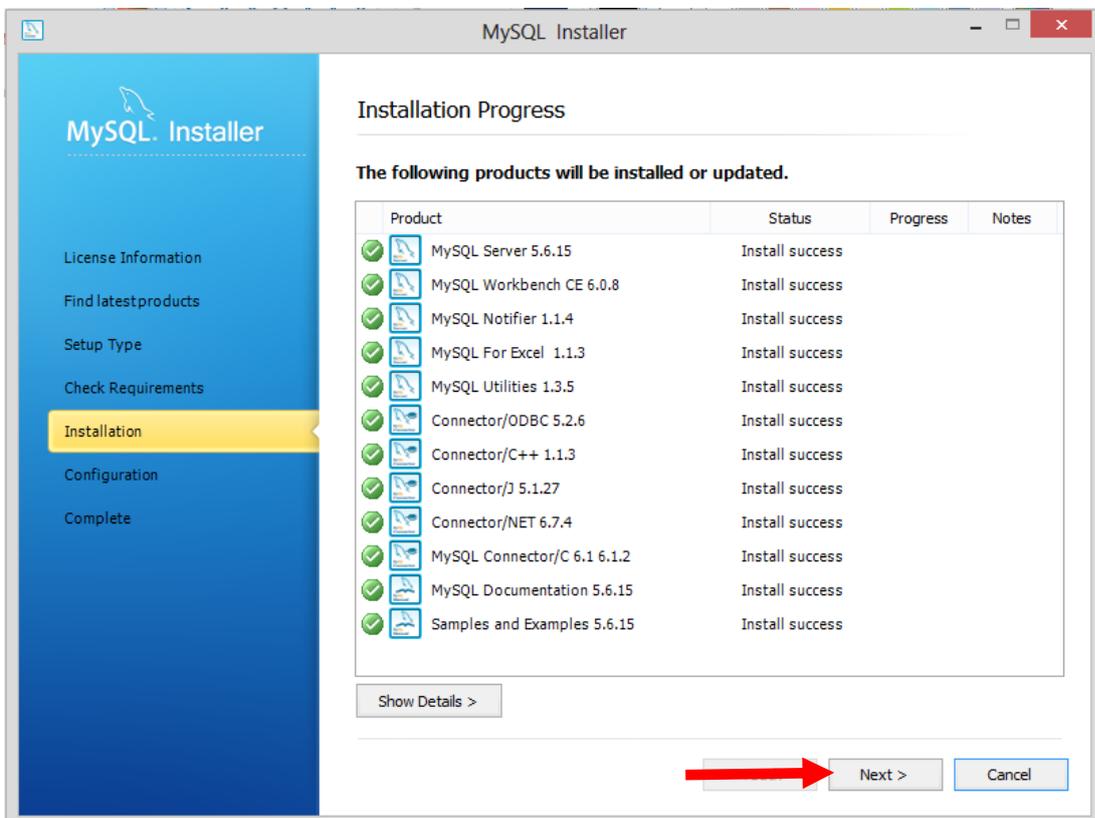


# MySQL to HDFS – Using Sqoop

9 - Click Execute:

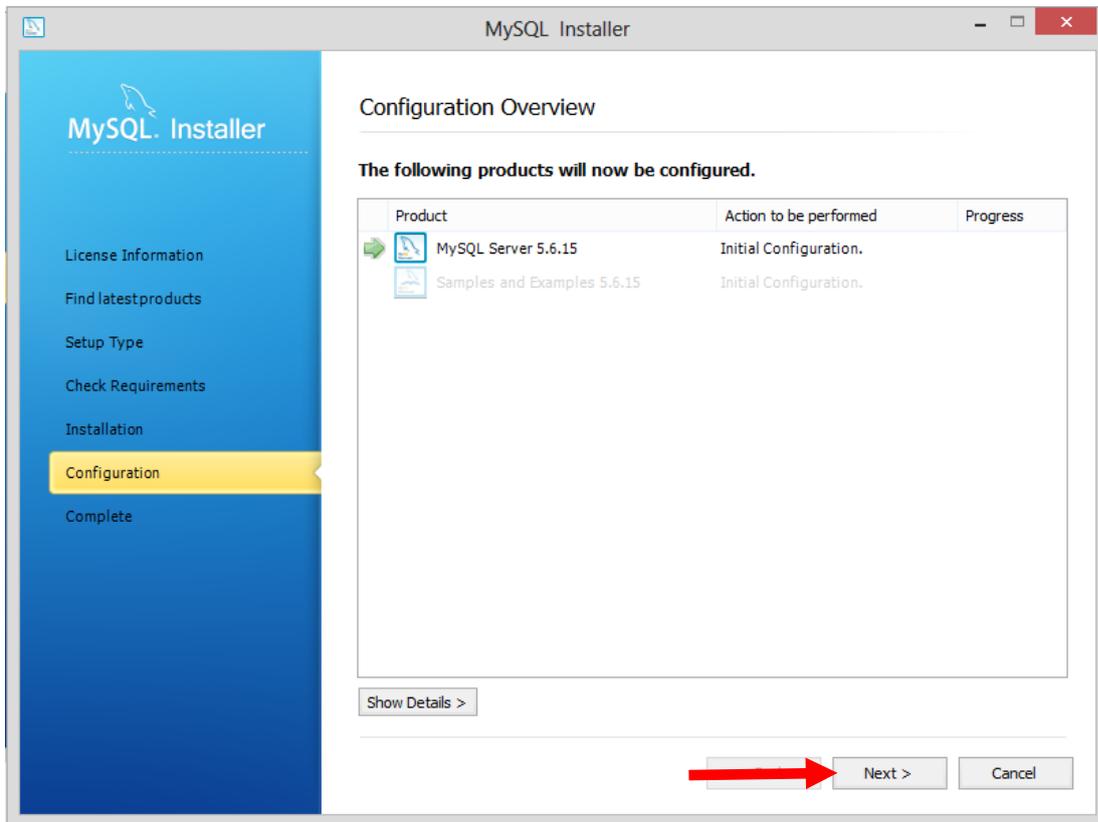


10 - Click Next:

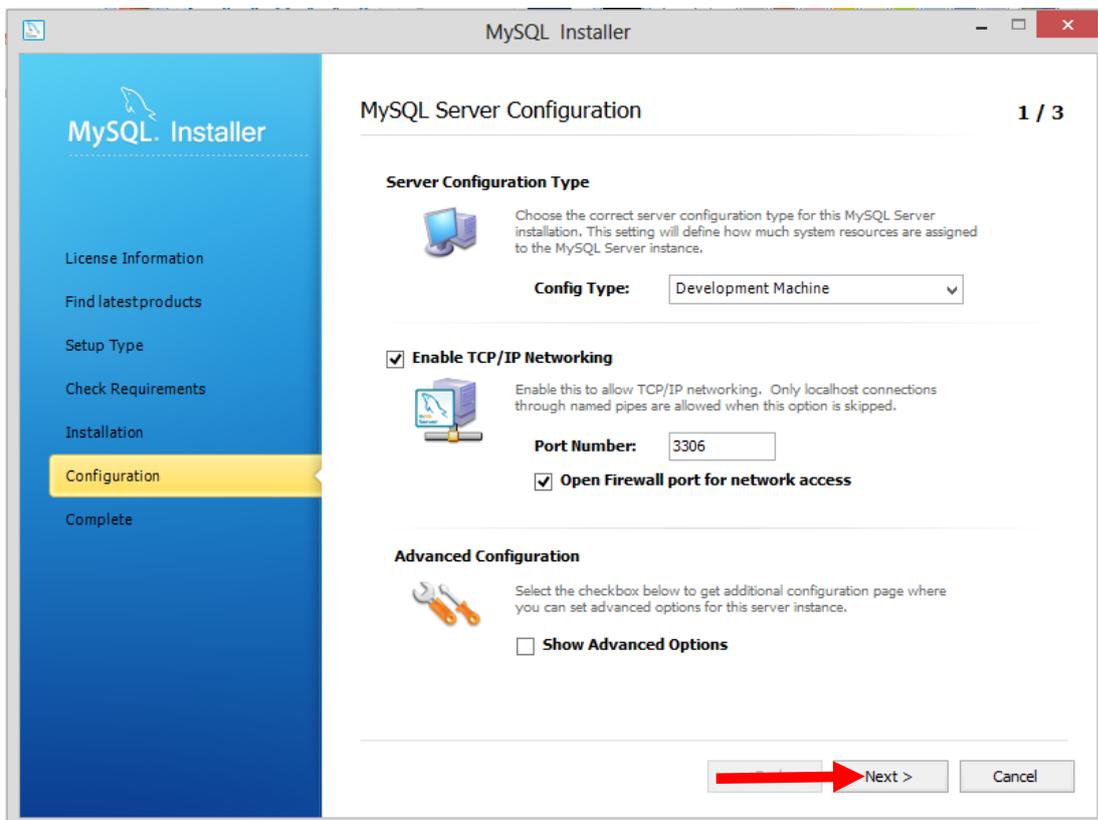


# MySQL to HDFS – Using Sqoop

11 - Click Next:



12 - Click Next:

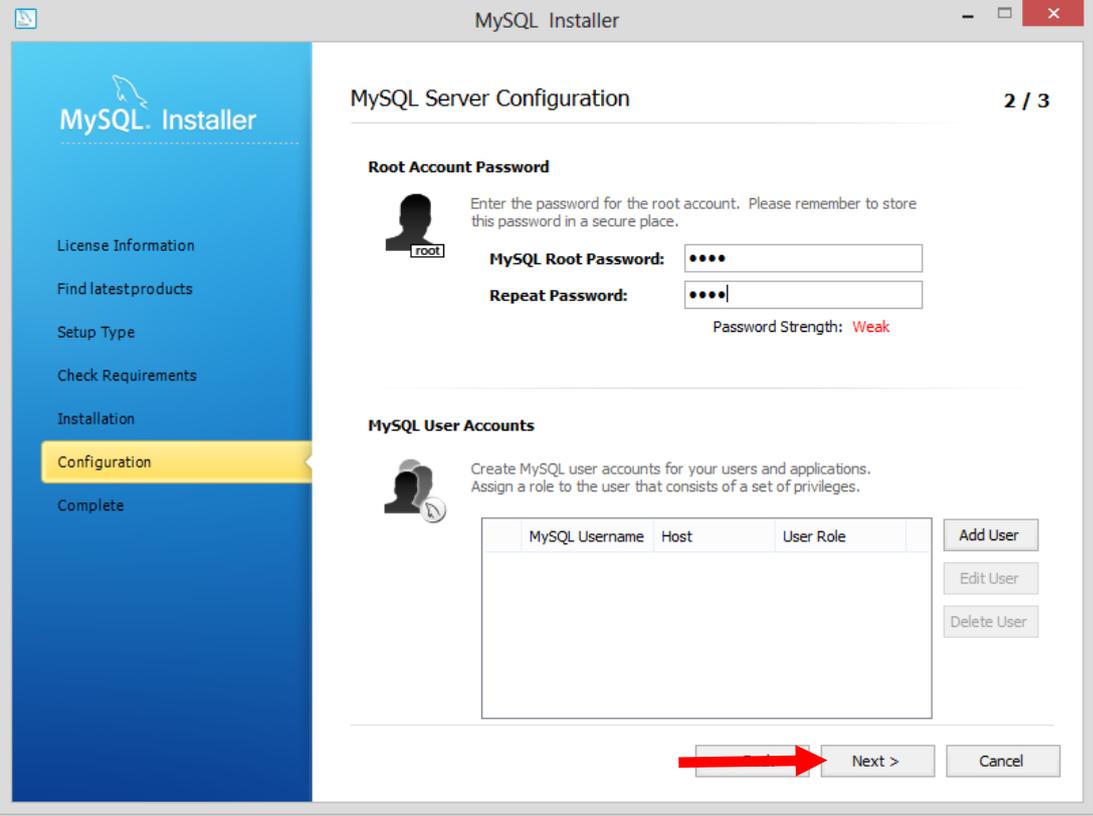


# MySQL to HDFS – Using Sqoop

13 - Enter in **MySQL Root Password** → root

Enter in **Repeat Password** → root

Click **Next**:



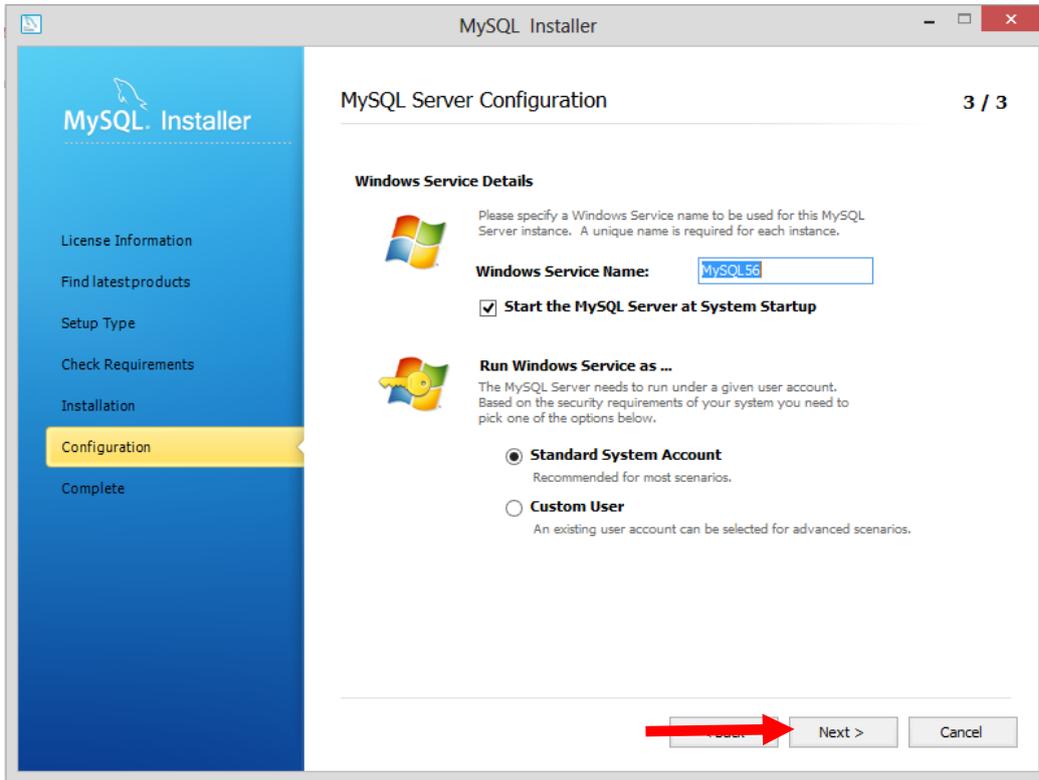
The screenshot shows the MySQL Installer window at the 'MySQL Server Configuration' step (2 / 3). The left sidebar lists the installation steps: License Information, Find latest products, Setup Type, Check Requirements, Installation, Configuration (highlighted), and Complete. The main area is divided into two sections:

- Root Account Password:** A section for setting the root password. It includes a user icon labeled 'root', a prompt to enter the password, and two input fields: 'MySQL Root Password' and 'Repeat Password'. Both fields contain four dots. Below the fields, the password strength is indicated as 'Weak' in red text.
- MySQL User Accounts:** A section for creating user accounts. It includes a user icon and a prompt to create accounts and assign roles. Below this is a table with columns for 'MySQL Username', 'Host', and 'User Role'. To the right of the table are three buttons: 'Add User', 'Edit User', and 'Delete User'.

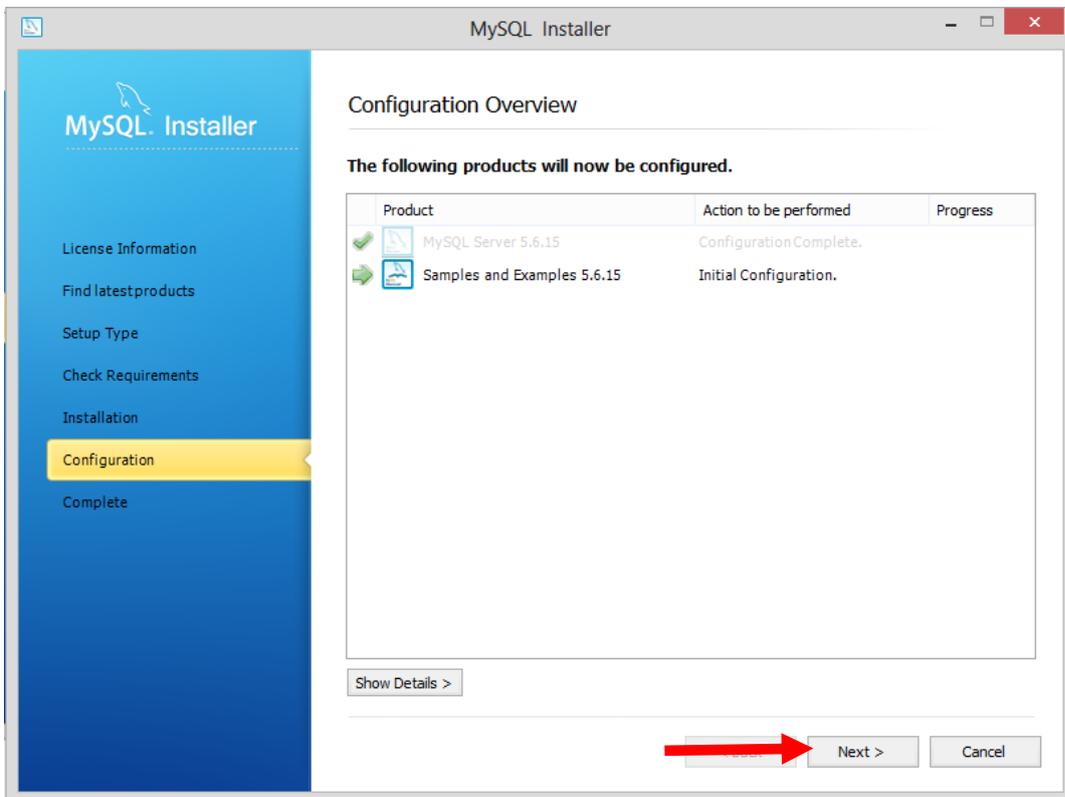
At the bottom right, there are 'Next >' and 'Cancel' buttons. A red arrow points to the 'Next >' button.

# MySQL to HDFS – Using Sqoop

14 - Click Next:

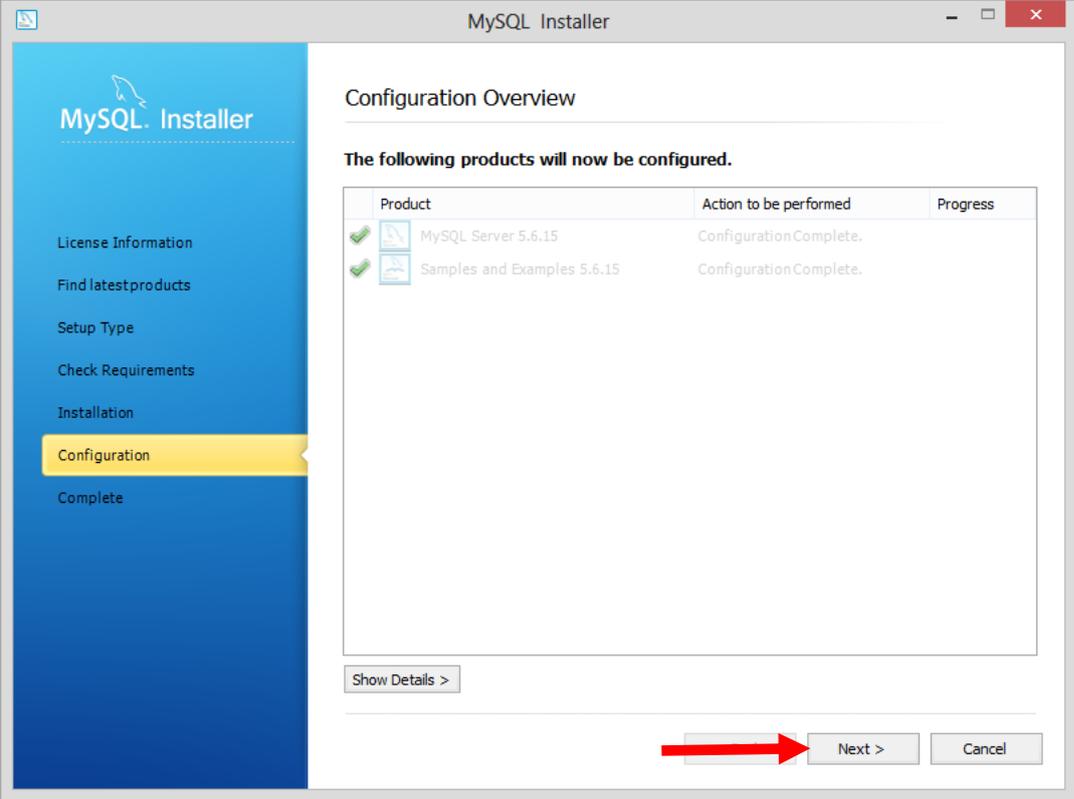


15 - Click Next:



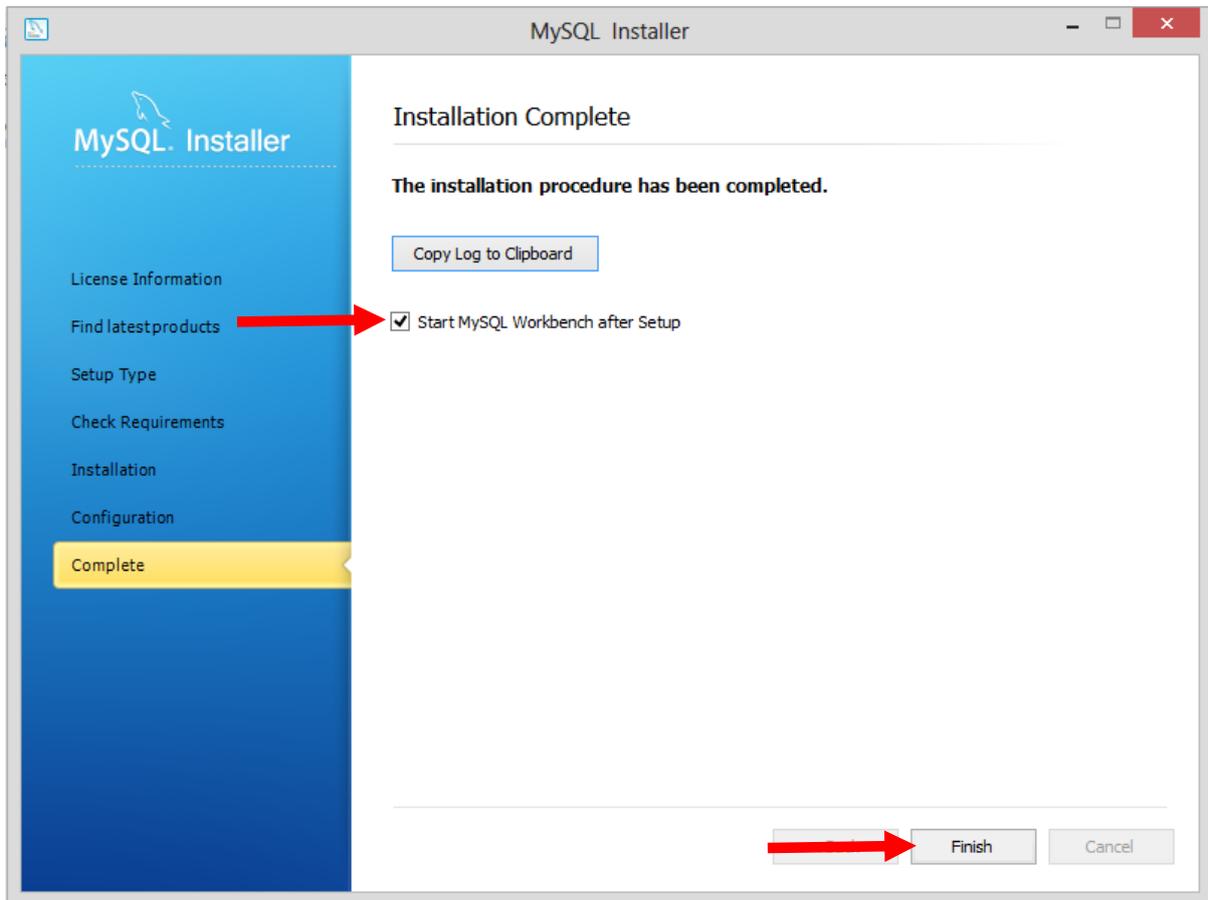
# MySQL to HDFS – Using Sqoop

16 - Click Next:



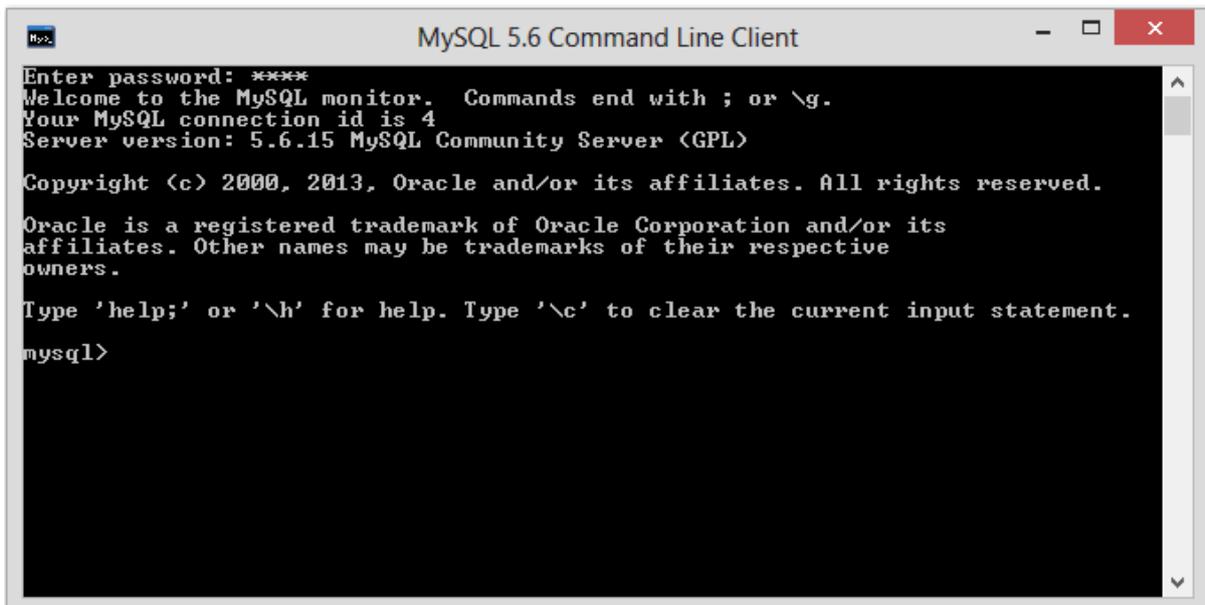
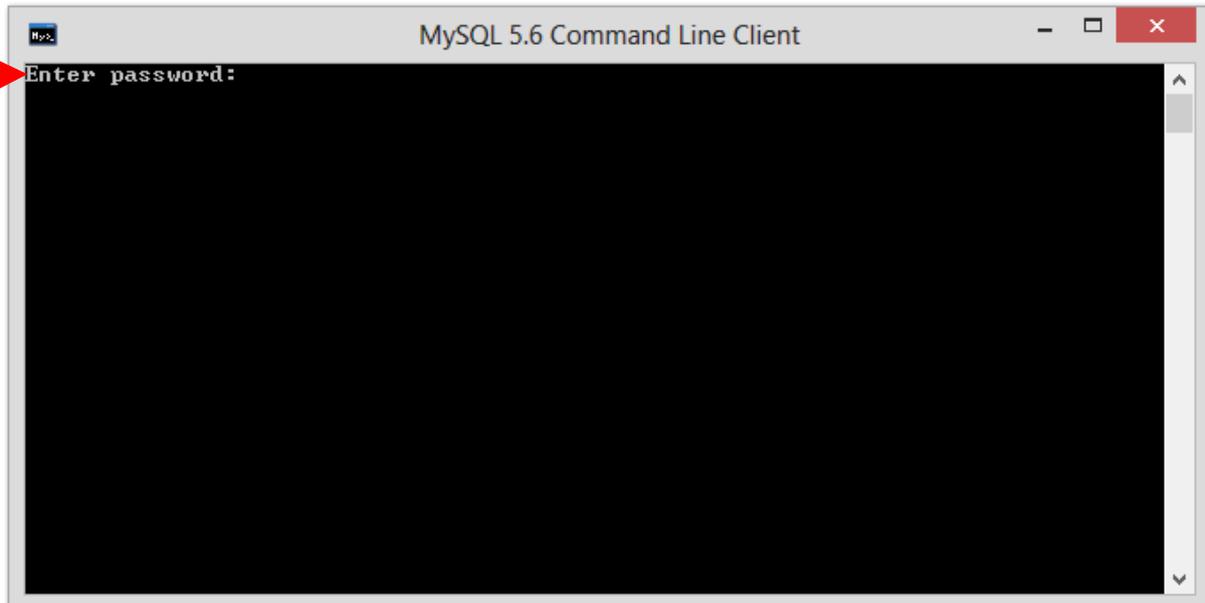
# MySQL to HDFS – Using Sqoop

17 - Uncheck the check-box (Start MySQL Workbench after Setup) and Click Finish:



## MySQL to HDFS – Using Sqoop

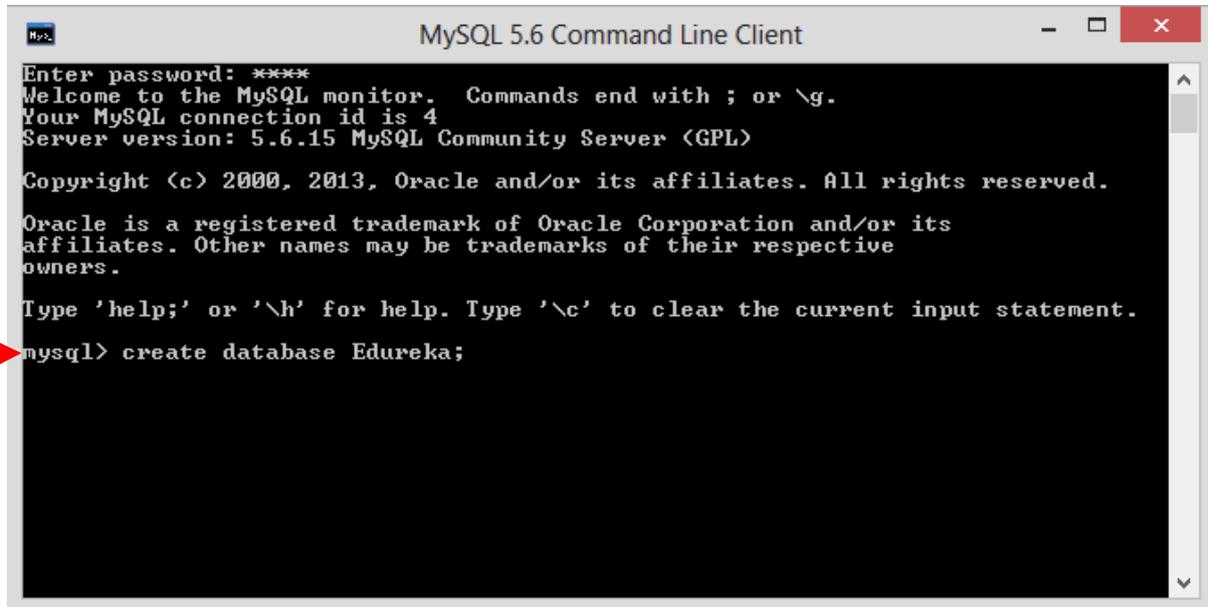
18 - Open MySQL 5.6 Command Line Client (You will find it in startup menu) and enter the Password as root



# MySQL to HDFS – Using Sqoop

19 - Create a Database named Edureka by executing the below command:

**Command:** create database Edureka;

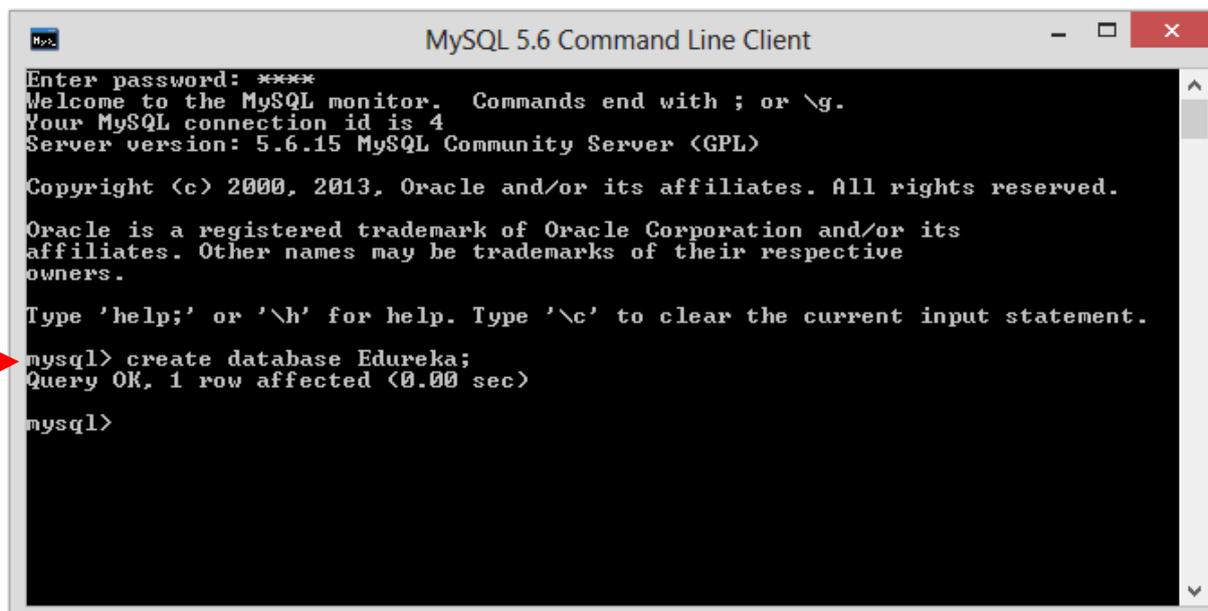


```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, 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> create database Edureka;
```



```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, 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.

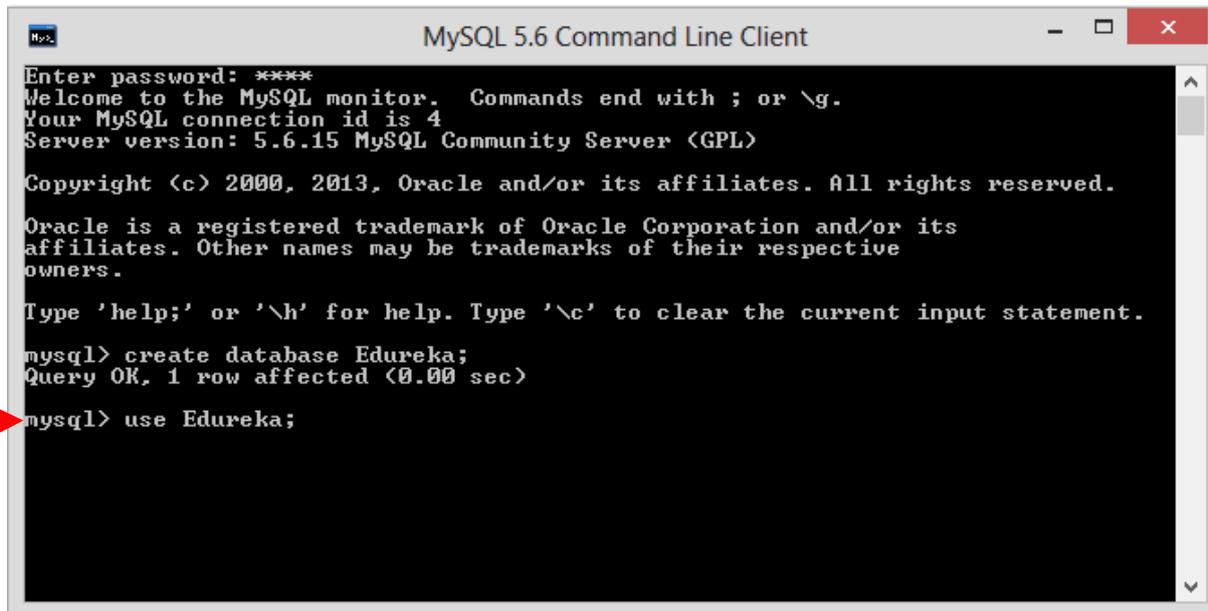
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
mysql> create database Edureka;
Query OK, 1 row affected (0.00 sec)

mysql>
```

# MySQL to HDFS – Using Sqoop

20 - Use Database named Edureka by executing the below command:

Command: use Edureka;

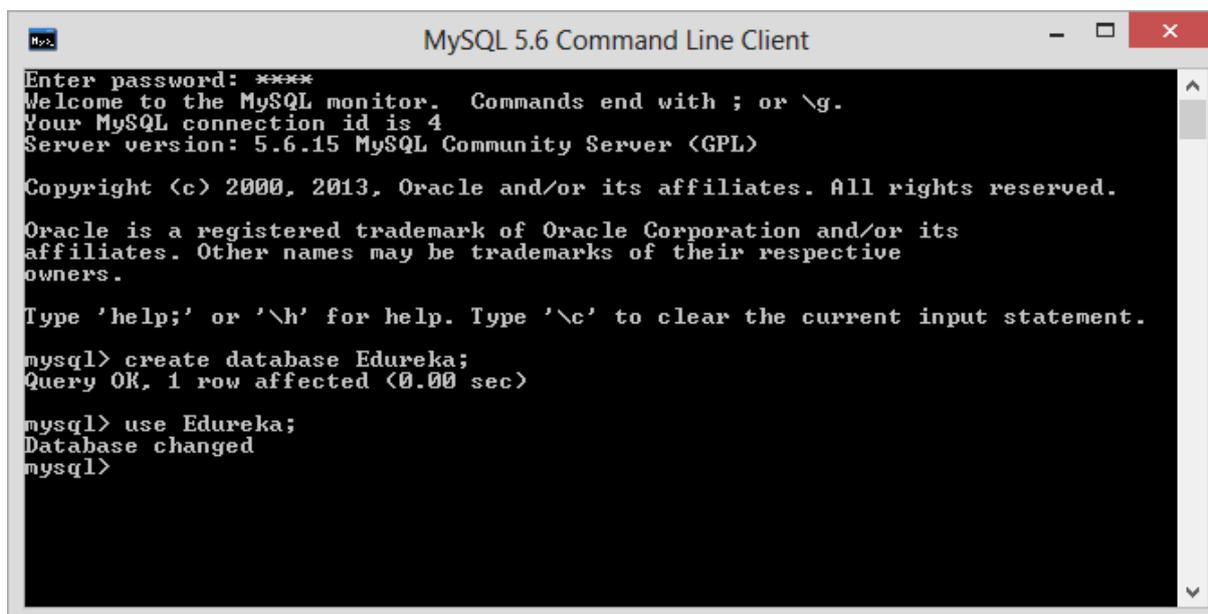


```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, 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> create database Edureka;
Query OK, 1 row affected (0.00 sec)
mysql> use Edureka;
```



```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

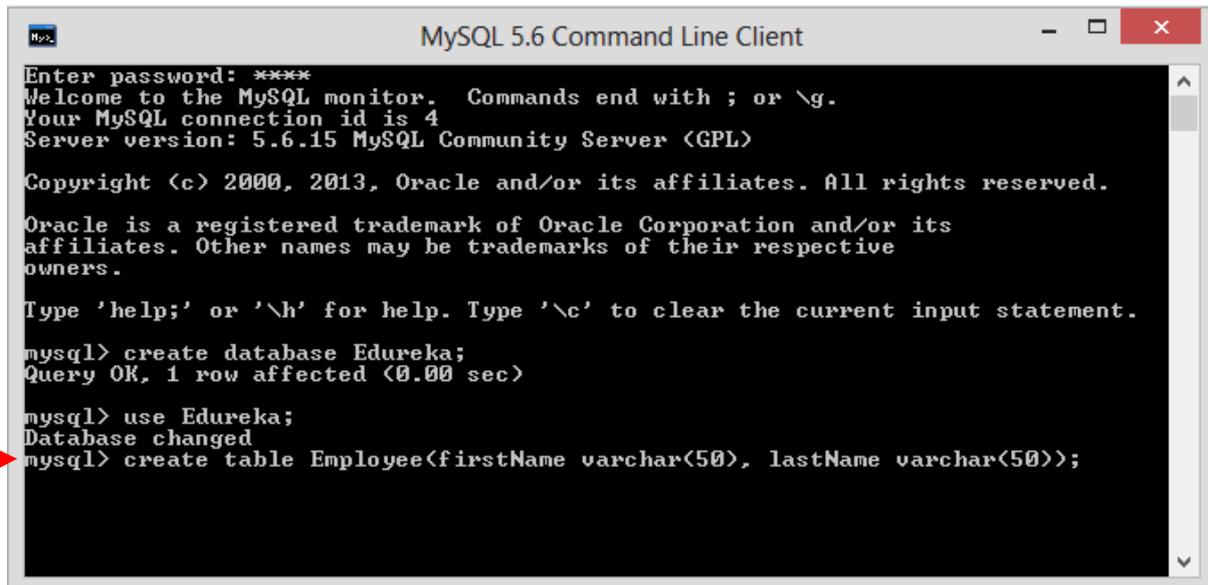
Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
mysql> create database Edureka;
Query OK, 1 row affected (0.00 sec)
mysql> use Edureka;
Database changed
mysql>
```

# MySQL to HDFS – Using Sqoop

21 - Create a Table named **Employee** by executing the below command:

**Command:** create table Employee(firstName varchar(50), lastName varchar(50));



```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

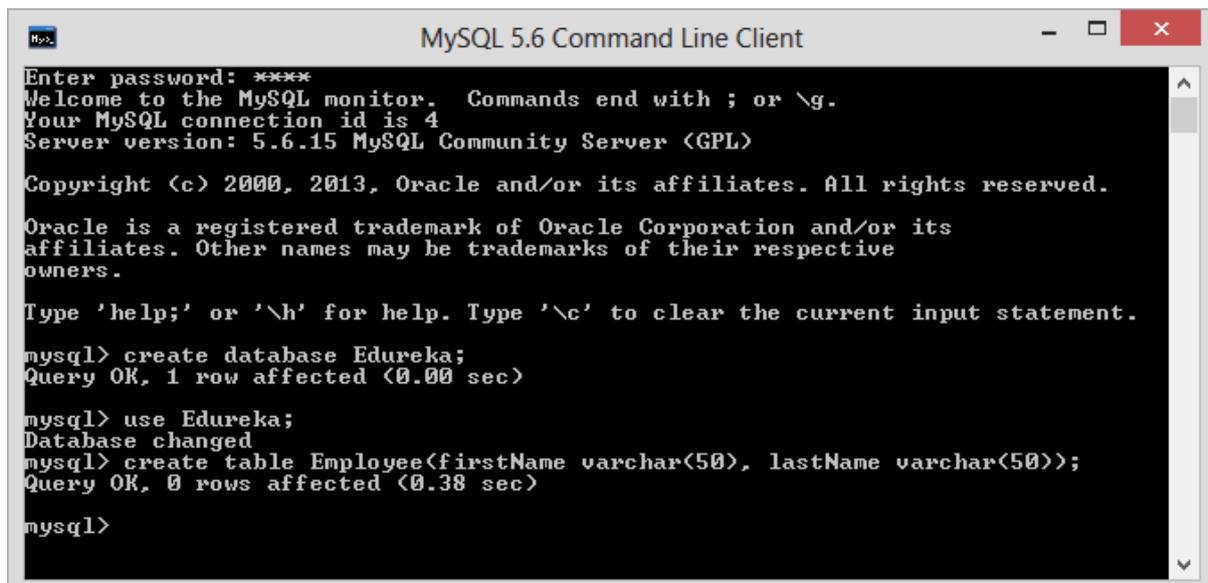
Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

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

mysql> use Edureka;
Database changed
mysql> create table Employee(firstName varchar(50), lastName varchar(50));
```

A red arrow points to the final command in the screenshot.



```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, 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> create database Edureka;
Query OK, 1 row affected (0.00 sec)

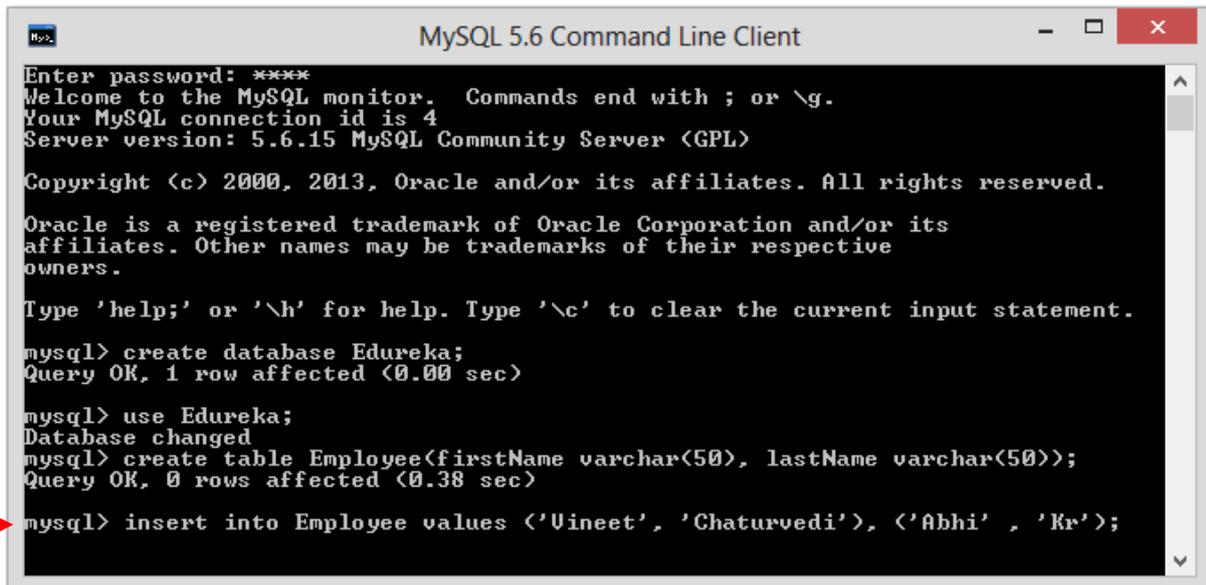
mysql> use Edureka;
Database changed
mysql> create table Employee(firstName varchar(50), lastName varchar(50));
Query OK, 0 rows affected (0.38 sec)

mysql>
```

# MySQL to HDFS – Using Sqoop

22 - Insert values in table Employee by executing the below command:

**Command:** insert into Employee values ('Vineet', 'Chaturvedi'), ('Abhi', 'Kr');



```
MySQL 5.6 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

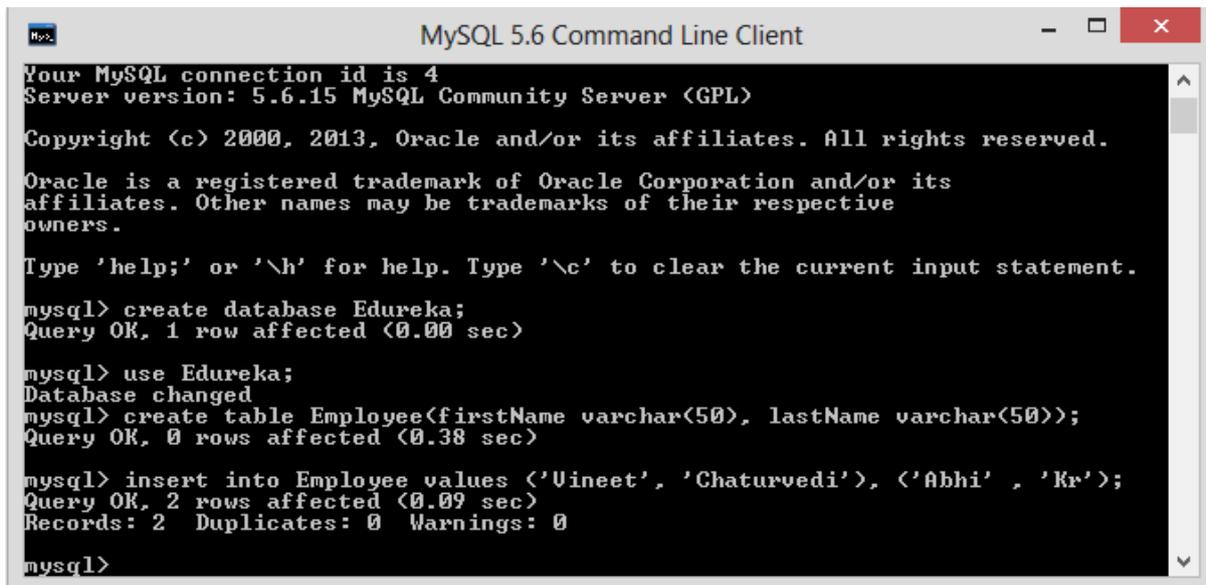
Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

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

mysql> use Edureka;
Database changed
mysql> create table Employee(firstName varchar(50), lastName varchar(50));
Query OK, 0 rows affected (0.38 sec)

mysql> insert into Employee values ('Vineet', 'Chaturvedi'), ('Abhi', 'Kr');
```



```
MySQL 5.6 Command Line Client
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)

Copyright (c) 2000, 2013, 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> create database Edureka;
Query OK, 1 row affected (0.00 sec)

mysql> use Edureka;
Database changed
mysql> create table Employee(firstName varchar(50), lastName varchar(50));
Query OK, 0 rows affected (0.38 sec)

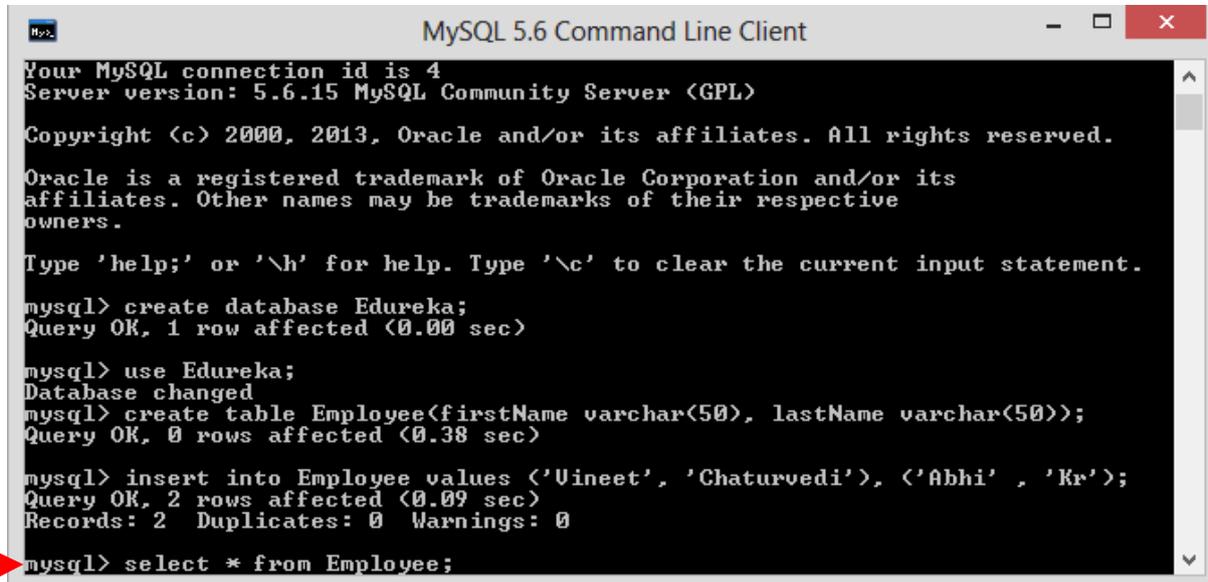
mysql> insert into Employee values ('Vineet', 'Chaturvedi'), ('Abhi', 'Kr');
Query OK, 2 rows affected (0.09 sec)
Records: 2  Duplicates: 0  Warnings: 0

mysql>
```

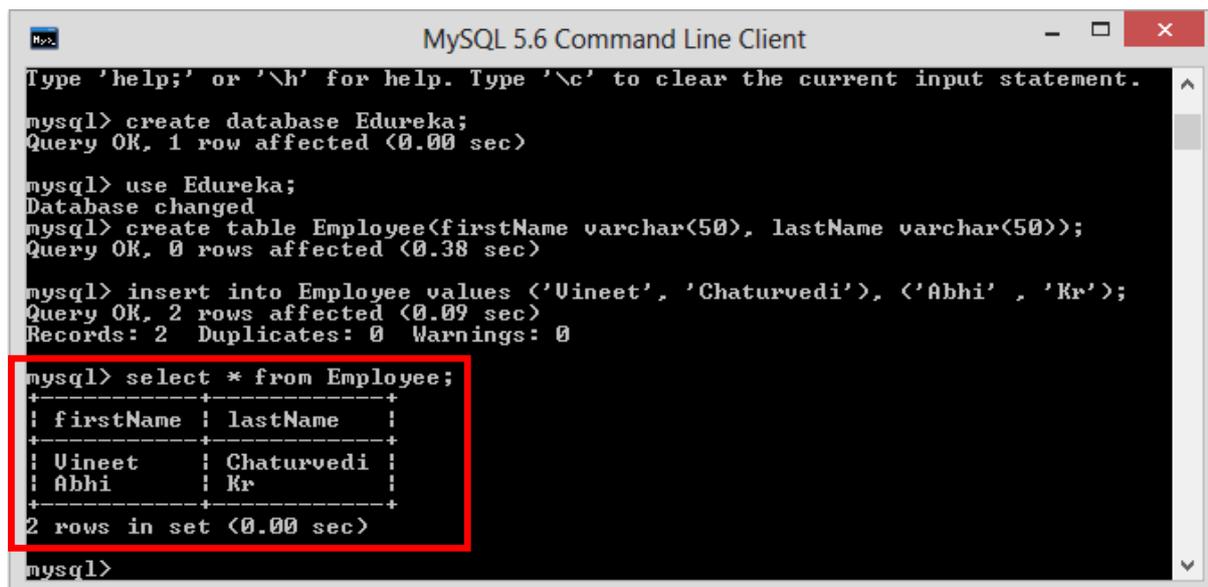
# MySQL to HDFS – Using Sqoop

23 - Check the rows present in table Employee by executing the below command:

Command: select \* from Employee;



```
MySQL 5.6 Command Line Client
Your MySQL connection id is 4
Server version: 5.6.15 MySQL Community Server (GPL)
Copyright (c) 2000, 2013, 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> create database Edureka;
Query OK, 1 row affected (0.00 sec)
mysql> use Edureka;
Database changed
mysql> create table Employee(firstName varchar(50), lastName varchar(50));
Query OK, 0 rows affected (0.38 sec)
mysql> insert into Employee values ('Vineet', 'Chaturvedi'), ('Abhi', 'Kr');
Query OK, 2 rows affected (0.09 sec)
Records: 2 Duplicates: 0 Warnings: 0
mysql> select * from Employee;
```

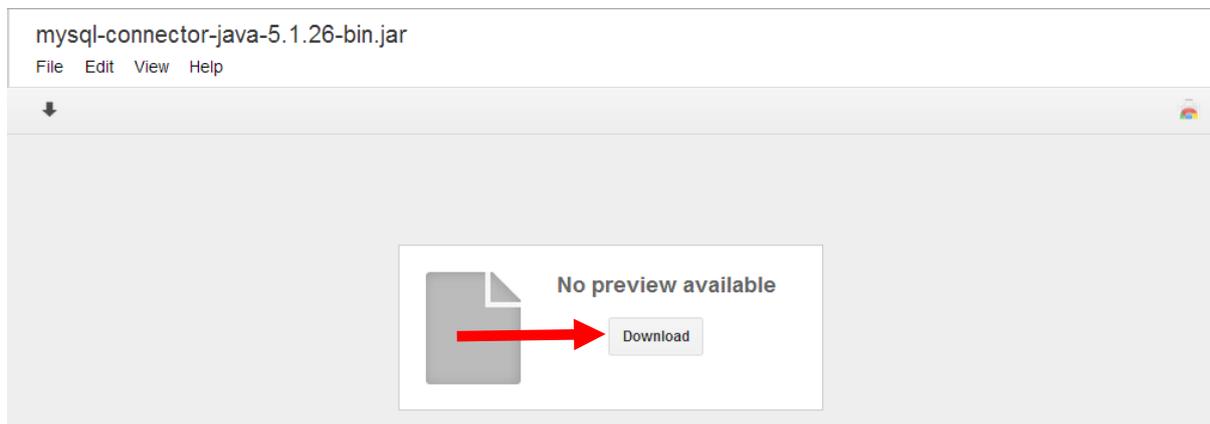


```
MySQL 5.6 Command Line Client
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
mysql> create database Edureka;
Query OK, 1 row affected (0.00 sec)
mysql> use Edureka;
Database changed
mysql> create table Employee(firstName varchar(50), lastName varchar(50));
Query OK, 0 rows affected (0.38 sec)
mysql> insert into Employee values ('Vineet', 'Chaturvedi'), ('Abhi', 'Kr');
Query OK, 2 rows affected (0.09 sec)
Records: 2 Duplicates: 0 Warnings: 0
mysql> select * from Employee;
+-----+-----+
| firstName | lastName |
+-----+-----+
| Vineet    | Chaturvedi |
| Abhi      | Kr        |
+-----+-----+
2 rows in set (0.00 sec)
mysql>
```

# MySQL to HDFS – Using Sqoop

24 - Download MySQL connector using the below link:

<https://drive.google.com/file/d/0B2-rICGKD40NSVJ3QlpWSk95OFE/edit?usp=sharing>



25 - Open Cloudera cdh3 and move MySQL connector to Cloudera cdh3 (To Desktop) using FileZilla.

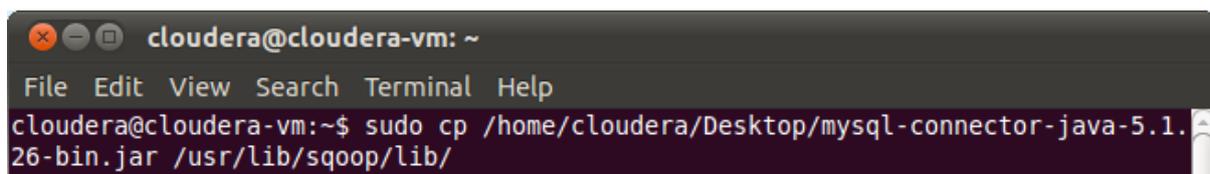
Use the below link to understand how to move a file from Windows to cloudera cdh3 vm.

<http://www.edureka.in/blog/transfer-files-windows-cloudera-demo-vm/>

26 - Once the MySQL connector is present on Cloudera Cdh3 Desktop, move it to the lib folder of sqoop by executing the below command:

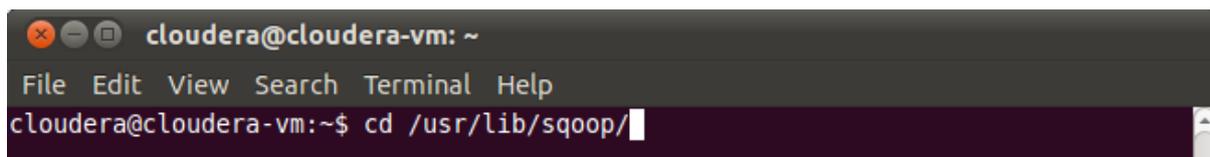
**Command:**

```
sudo cp /home/cloudera/Desktop/mysql-connector-java-5.1.26-bin.jar /usr/lib/sqoop/lib
```



27 - Change the directory to Sqoop by executing the below command:

**Command:** cd /usr/lib/sqoop



# MySQL to HDFS – Using Sqoop

28 - Open Command Prompt (CMD) on Windows and check the IPv4 Address by executing the

below command:

Command: ipconfig

```
C:\Users\User>ipconfig

Windows IP Configuration

Wireless LAN adapter Local Area Connection* 13:

    Media State . . . . . : Media disconnected
    Connection-specific DNS Suffix  . :

Wireless LAN adapter Local Area Connection* 11:

    Media State . . . . . : Media disconnected
    Connection-specific DNS Suffix  . :

Wireless LAN adapter Wi-Fi:

    Connection-specific DNS Suffix  . :
    Link-local IPv6 Address . . . . . : fe80::30fb:d292:86bf:1473%24
    IPv4 Address. . . . . : 192.168.1.149
    Subnet Mask . . . . . : 255.255.255.0
    Default Gateway . . . . . : 192.168.1.1

Ethernet adapter Bluetooth Network Connection:

    Media State . . . . . : Media disconnected
    Connection-specific DNS Suffix  . :

Ethernet adapter Ethernet:

    Media State . . . . . : Media disconnected
    Connection-specific DNS Suffix  . :

Ethernet adapter VMware Network Adapter VMnet1:

    Connection-specific DNS Suffix  . :
    Link-local IPv6 Address . . . . . : fe80::41ae:2459:93df:f6a5%20
    IPv4 Address. . . . . : 192.168.243.1
    Subnet Mask . . . . . : 255.255.255.0
    Default Gateway . . . . . :
```

29 - Grant all privileges to root@your\_ipv4\_address by executing the below command

(in MySQL 5.6 Command Line Client):

Required items for the command:

Ip – Find out the IPv4 address of your system using the above step. In my case it is

192.168.243.1

Command:

grant all privileges on \*.\* to [root@192.168.243.1](#) IDENTIFIED BY 'root' WITH GRANT OPTION;

```
mysql> grant all privileges on *.* to root@192.168.243.1 IDENTIFIED BY 'root' WI
TH GRANT OPTION;
Query OK, 0 rows affected (0.06 sec)
```

# MySQL to HDFS – Using Sqoop

30 - Import the table Employee present in MySQL database to hdfs by executing the below

command:

**Required items for the command:**

**IPv4 Address** – Your IPv4 address. In my case it is 192.168.243.1

**Database Name** – Edureka

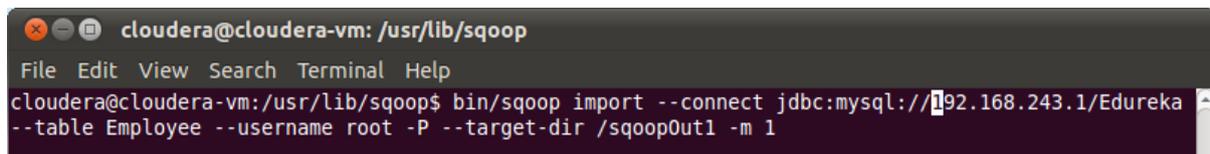
**Table Name** – Employee

**Username** – root

**Output Directory** – Could be any. I have used sqoopOut1

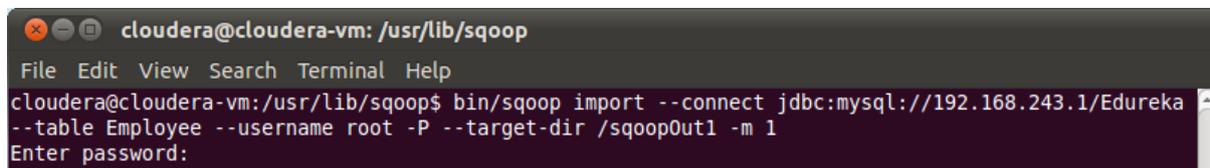
**Command:**

```
bin/sqoop import --connect jdbc:mysql://192.168.243.1/Edureka --table Employee --username root -P --target-dir /sqoopOut1 -m 1
```

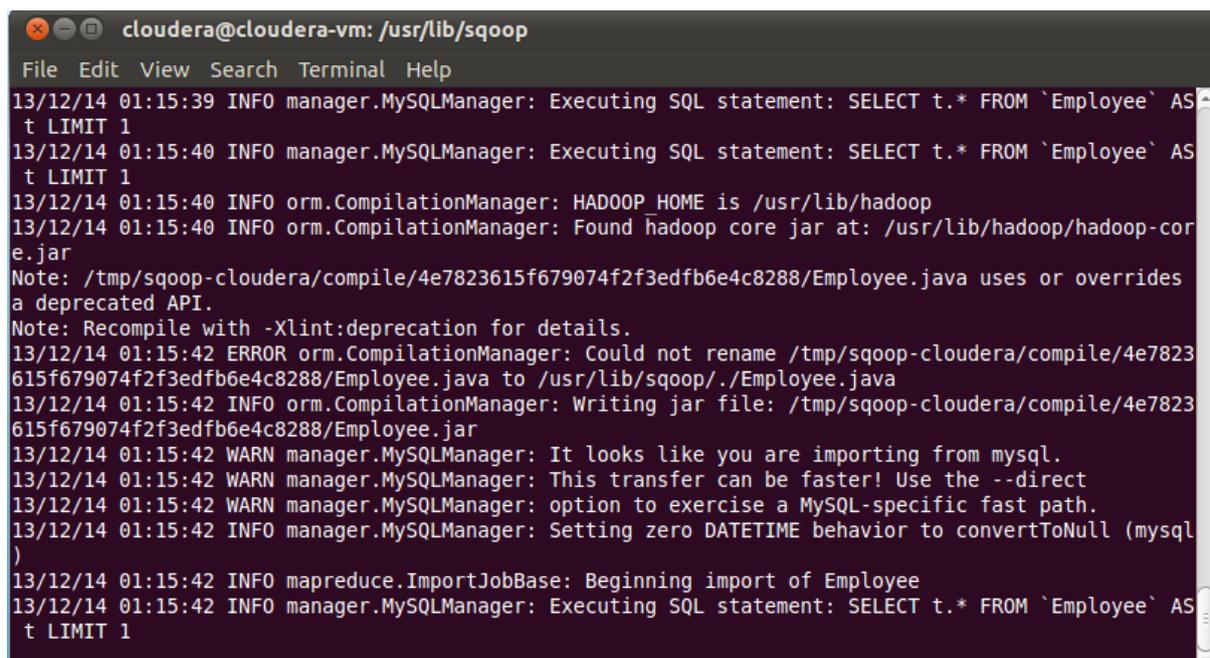


```
cloudera@cloudera-vm: /usr/lib/sqoop
File Edit View Search Terminal Help
cloudera@cloudera-vm: /usr/lib/sqoop$ bin/sqoop import --connect jdbc:mysql://192.168.243.1/Edureka --table Employee --username root -P --target-dir /sqoopOut1 -m 1
```

31 - Enter the Password - root



```
cloudera@cloudera-vm: /usr/lib/sqoop
File Edit View Search Terminal Help
cloudera@cloudera-vm: /usr/lib/sqoop$ bin/sqoop import --connect jdbc:mysql://192.168.243.1/Edureka --table Employee --username root -P --target-dir /sqoopOut1 -m 1
Enter password:
```



```
cloudera@cloudera-vm: /usr/lib/sqoop
File Edit View Search Terminal Help
13/12/14 01:15:39 INFO manager.MySQLManager: Executing SQL statement: SELECT t.* FROM `Employee` AS t LIMIT 1
13/12/14 01:15:40 INFO manager.MySQLManager: Executing SQL statement: SELECT t.* FROM `Employee` AS t LIMIT 1
13/12/14 01:15:40 INFO orm.CompilationManager: HADOOP_HOME is /usr/lib/hadoop
13/12/14 01:15:40 INFO orm.CompilationManager: Found hadoop core jar at: /usr/lib/hadoop/hadoop-core.jar
Note: /tmp/sqoop-cloudera/compile/4e7823615f679074f2f3edfb6e4c8288/Employee.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
13/12/14 01:15:42 ERROR orm.CompilationManager: Could not rename /tmp/sqoop-cloudera/compile/4e7823615f679074f2f3edfb6e4c8288/Employee.java to /usr/lib/sqoop/./Employee.java
13/12/14 01:15:42 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/4e7823615f679074f2f3edfb6e4c8288/Employee.jar
13/12/14 01:15:42 WARN manager.MySQLManager: It looks like you are importing from mysql.
13/12/14 01:15:42 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
13/12/14 01:15:42 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
13/12/14 01:15:42 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
13/12/14 01:15:42 INFO mapreduce.ImportJobBase: Beginning import of Employee
13/12/14 01:15:42 INFO manager.MySQLManager: Executing SQL statement: SELECT t.* FROM `Employee` AS t LIMIT 1
```

# MySQL to HDFS – Using Sqoop

```

cloudera@cloudera-vm: /usr/lib/sqoop
File Edit View Search Terminal Help
13/12/13 04:14:33 INFO mapred.JobClient: map 100% reduce 0%
13/12/13 04:14:34 INFO mapred.JobClient: Job complete: job_201312130137_0001
13/12/13 04:14:34 INFO mapred.JobClient: Counters: 12
13/12/13 04:14:34 INFO mapred.JobClient: Job Counters
13/12/13 04:14:34 INFO mapred.JobClient: SLOTS_MILLIS_MAPS=12768
13/12/13 04:14:34 INFO mapred.JobClient: Total time spent by all reduces waiting after reservi
ng slots (ms)=0
13/12/13 04:14:34 INFO mapred.JobClient: Total time spent by all maps waiting after reserving
slots (ms)=0
13/12/13 04:14:34 INFO mapred.JobClient: Launched map tasks=1
13/12/13 04:14:34 INFO mapred.JobClient: SLOTS_MILLIS_REDUCE=0
13/12/13 04:14:34 INFO mapred.JobClient: FileSystemCounters
13/12/13 04:14:34 INFO mapred.JobClient: HDFS_BYTES_READ=87
13/12/13 04:14:34 INFO mapred.JobClient: FILE_BYTES_WRITTEN=59336
13/12/13 04:14:34 INFO mapred.JobClient: HDFS_BYTES_WRITTEN=26
13/12/13 04:14:34 INFO mapred.JobClient: Map-Reduce Framework
13/12/13 04:14:34 INFO mapred.JobClient: Map input records=2
13/12/13 04:14:34 INFO mapred.JobClient: Spilled Records=0
13/12/13 04:14:34 INFO mapred.JobClient: Map output records=2
13/12/13 04:14:34 INFO mapred.JobClient: SPLIT_RAW_BYTES=87
13/12/13 04:14:34 INFO mapreduce.ImportJobBase: Transferred 26 bytes in 24.4973 seconds (1.0613 by
tes/sec)
13/12/13 04:14:34 INFO mapreduce.ImportJobBase: Retrieved 2 records.

```

32 - Open the Browser and go to the below URL:

URL: <http://localhost:50070/dfshealth.jsp>

Click on **Browse the filesystem**

**NameNode 'localhost.localdomain:8020'**

<b>Started:</b>	Fri Dec 13 23:12:34 PST 2013
<b>Version:</b>	0.20.2-cdh3u0, r81256ad0f2e4ab2bd34b04f53d25a6c23686dd14
<b>Compiled:</b>	Sat Mar 26 00:14:04 UTC 2011 by root
<b>Upgrades:</b>	There are no upgrades in progress.

[Browse the filesystem](#)  
[NameNode Logs](#)

**Cluster Summary**

31 files and directories, 18 blocks = 49 total. Heap Size is 31.32 MB / 966.69 MB (3%)

<b>Configured Capacity</b>	:	7.49 GB
<b>DFS Used</b>	:	196 KB
<b>Non DFS Used</b>	:	3.6 GB
<b>DFS Remaining</b>	:	3.89 GB
<b>DFS Used%</b>	:	0 %
<b>DFS Remaining%</b>	:	51.98 %
<b>Live Nodes</b>	:	1
<b>Dead Nodes</b>	:	0
<b>Decommissioning Nodes</b>	:	0
<b>Number of Under-Replicated Blocks</b>	:	0

# MySQL to HDFS – Using Sqoop

33 - Click on sqoopOut1 directory:

Contents of directory /

Goto :  go

Name	Type	Size	Replication	Block Size	Modification Time	Permission	Owner	Group
sqoopOut	dir				2013-12-13 04:14	rwxr-xr-x	cloudera	supergroup
sqoopOut1	dir				2013-12-14 01:15	rwxr-xr-x	cloudera	supergroup
tmp	dir				2011-04-01 17:25	rwxrwxrwx	hue	supergroup
user	dir				2011-04-08 16:48	rwxr-xr-x	hue	supergroup
var	dir				2013-12-13 01:38	rwxr-xr-x	mapred	supergroup

[Go back to DFS home](#)

## Local logs

34 - Click on part-m-00000 file:

Contents of directory /sqoopOut1

Goto :  go

[Go to parent directory](#)

Name	Type	Size	Replication	Block Size	Modification Time	Permission	Owner	Group
_SUCCESS	file	0 KB	1	64 MB	2013-12-14 01:15	rw-r--r--	cloudera	supergroup
_logs	dir				2013-12-14 01:15	rwxr-xr-x	cloudera	supergroup
part-m-00000	file	0.03 KB	1	64 MB	2013-12-14 01:15	rw-r--r--	cloudera	supergroup

35 - Below is the data that was imported from MySQL database:

Goto :  go

[Go back to dir listing](#)  
[Advanced view/download options](#)

```
Vineet, Chaturvedi
Abhi, Kr
```

**Congratulations! You have Successfully Imported Data from MySQL Database to HDFS Using Sqoop..!**