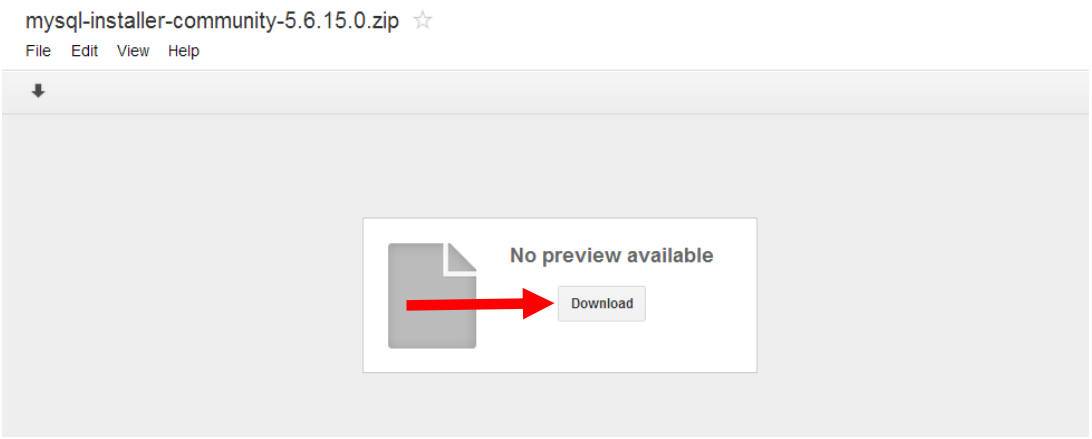
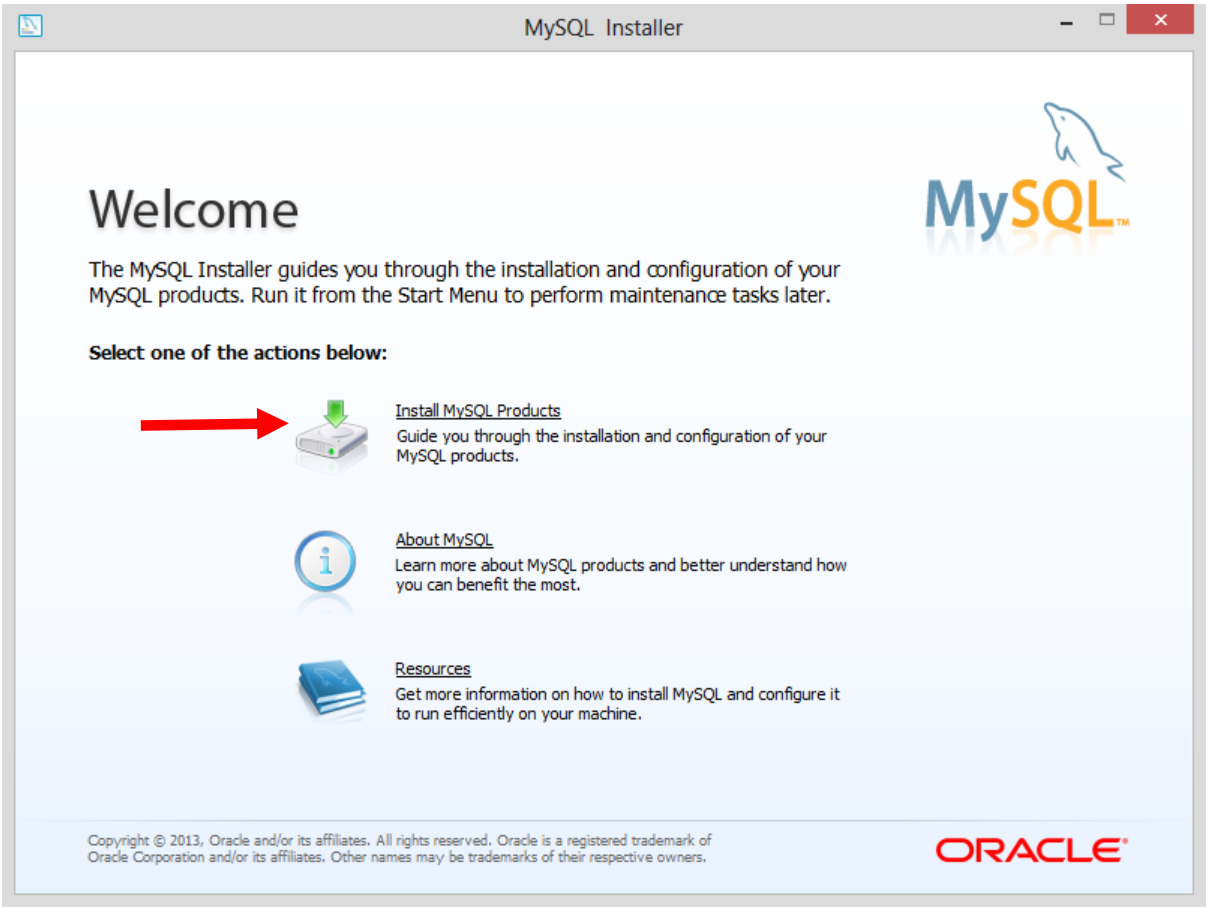
MySQL to HDFS – Using Sqoop

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

https://drive.google.com/file/d/0B2-rlCGKD40NangwRGdLUXg2REE/edit?usp=sharing

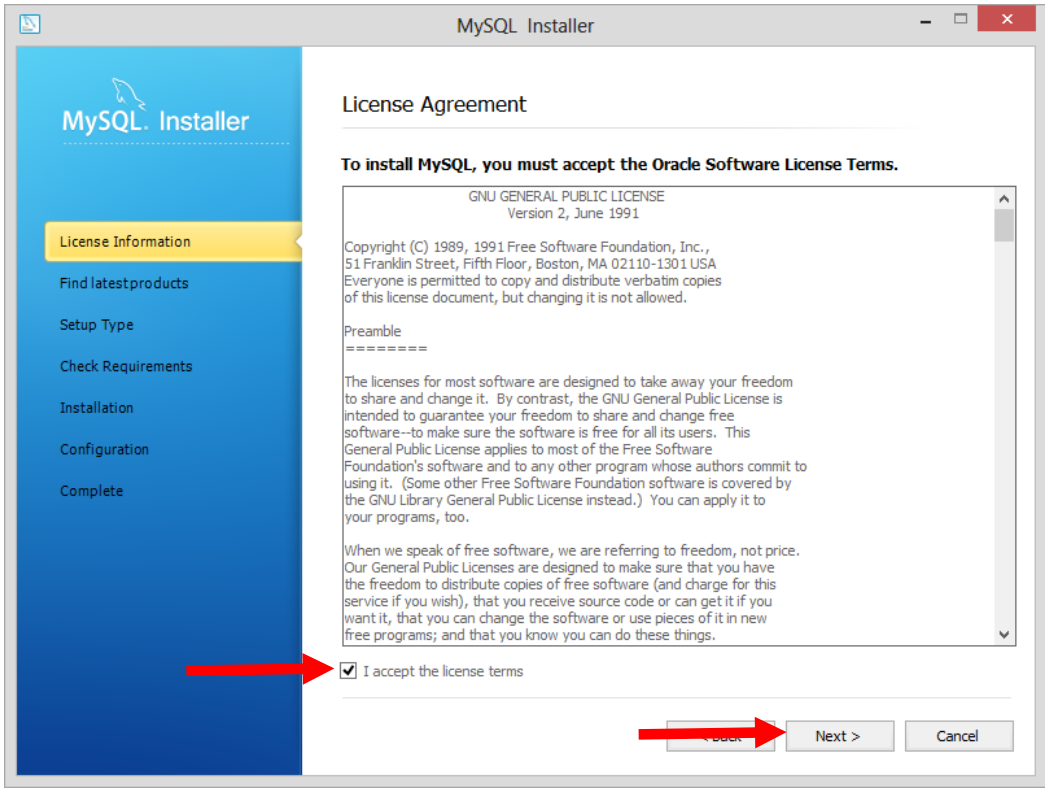


2 - Double click the extracted file and click on **Install MySQL Products**:

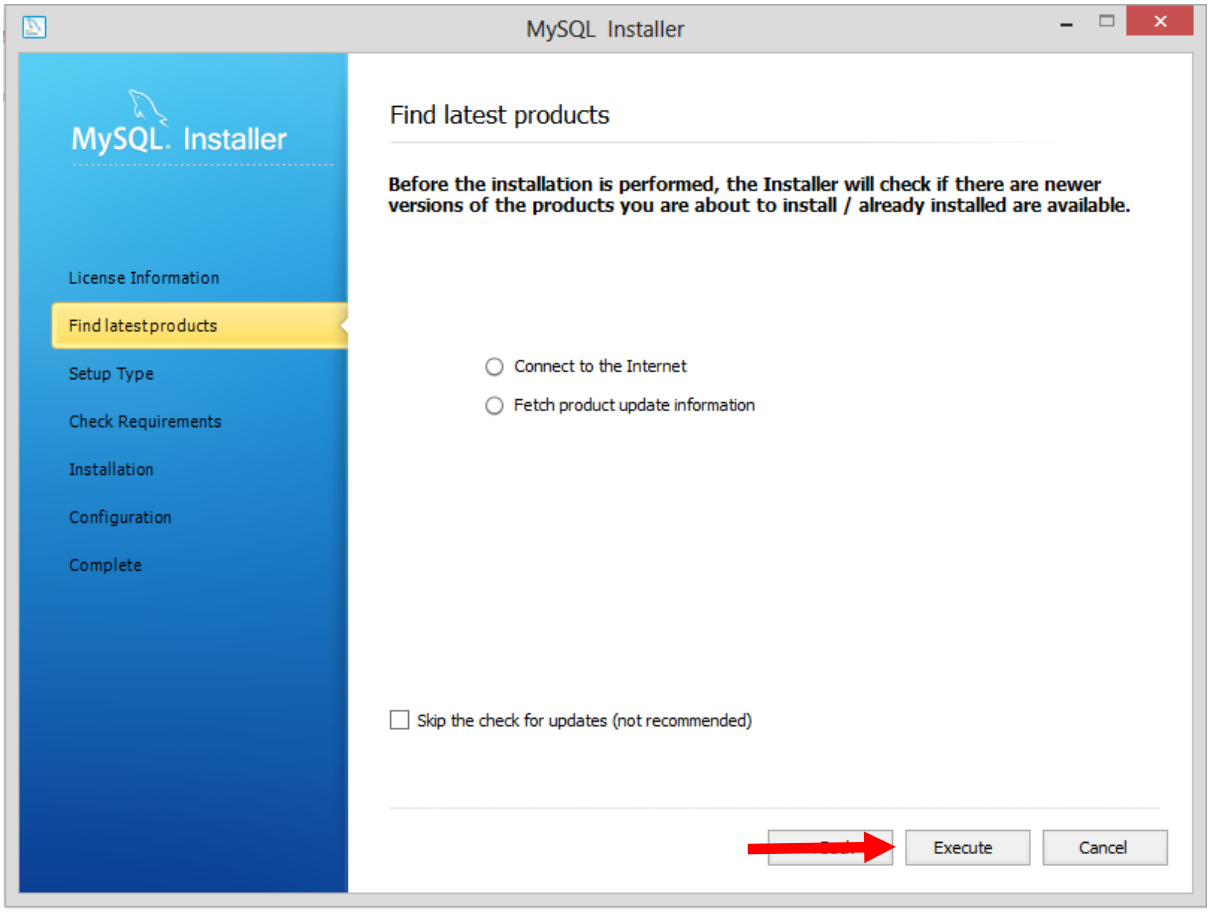


MySQL to HDFS – Using Sqoop

3 - Click **Next**:

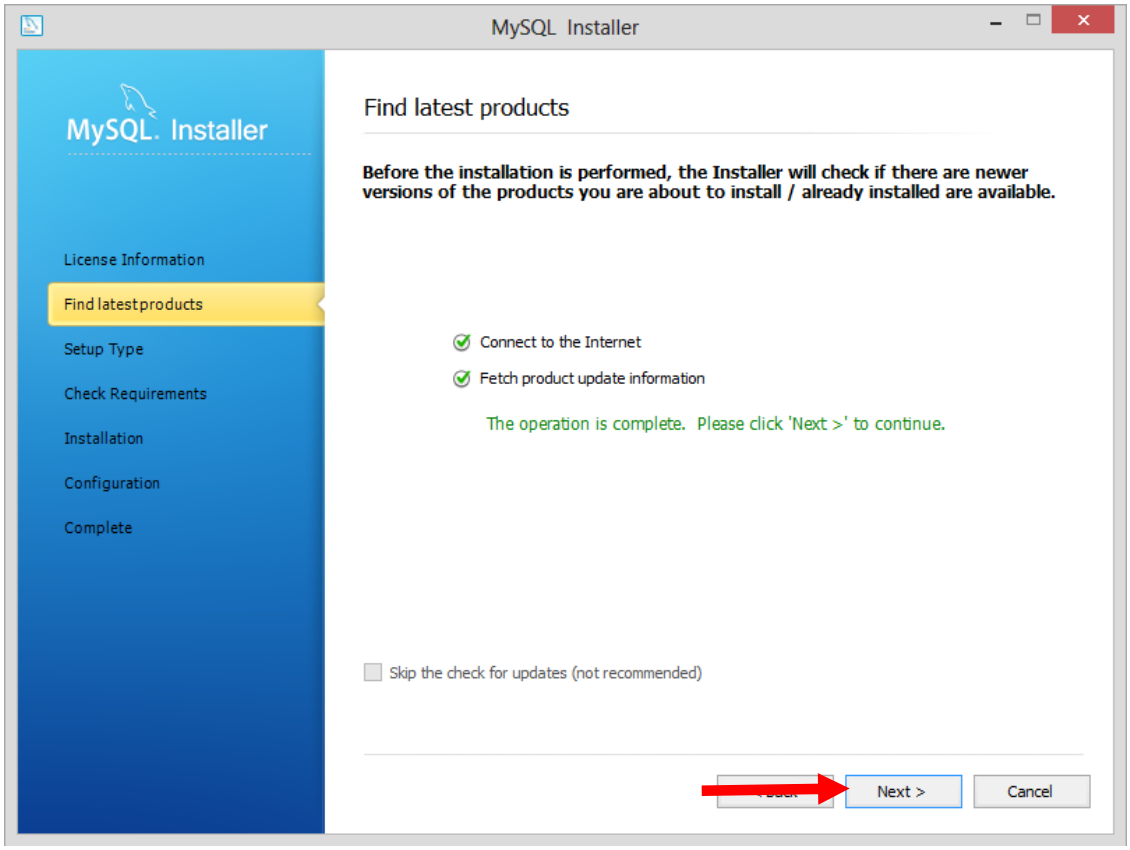


4 - Click **Execute**:

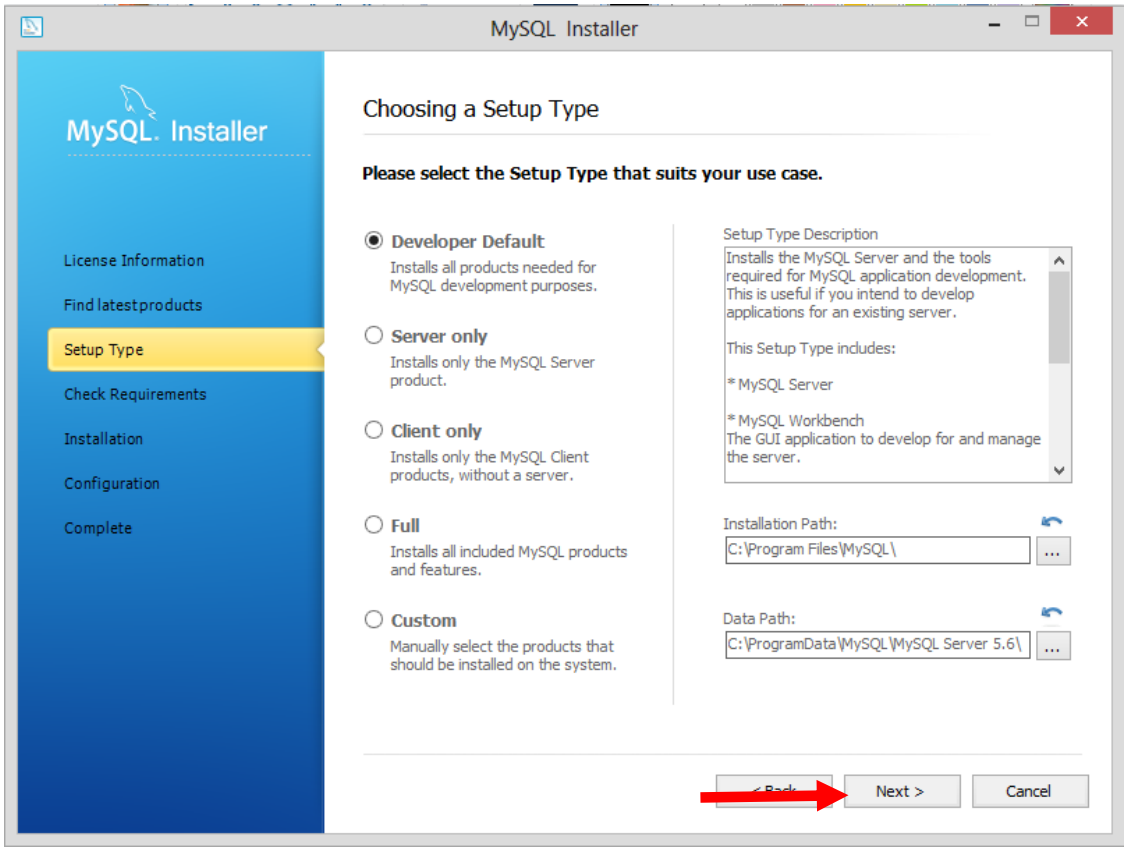


MySQL to HDFS – Using Sqoop

5 - Click **Next**:

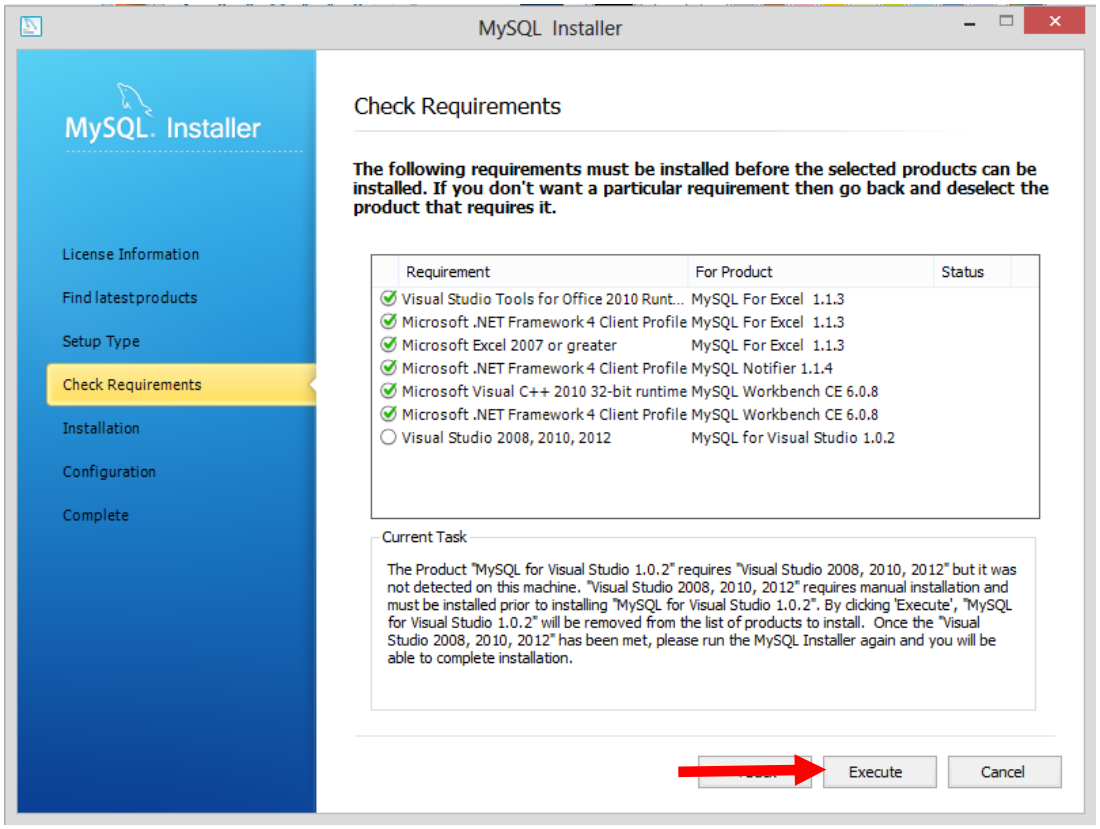


6 - Click **Next**:

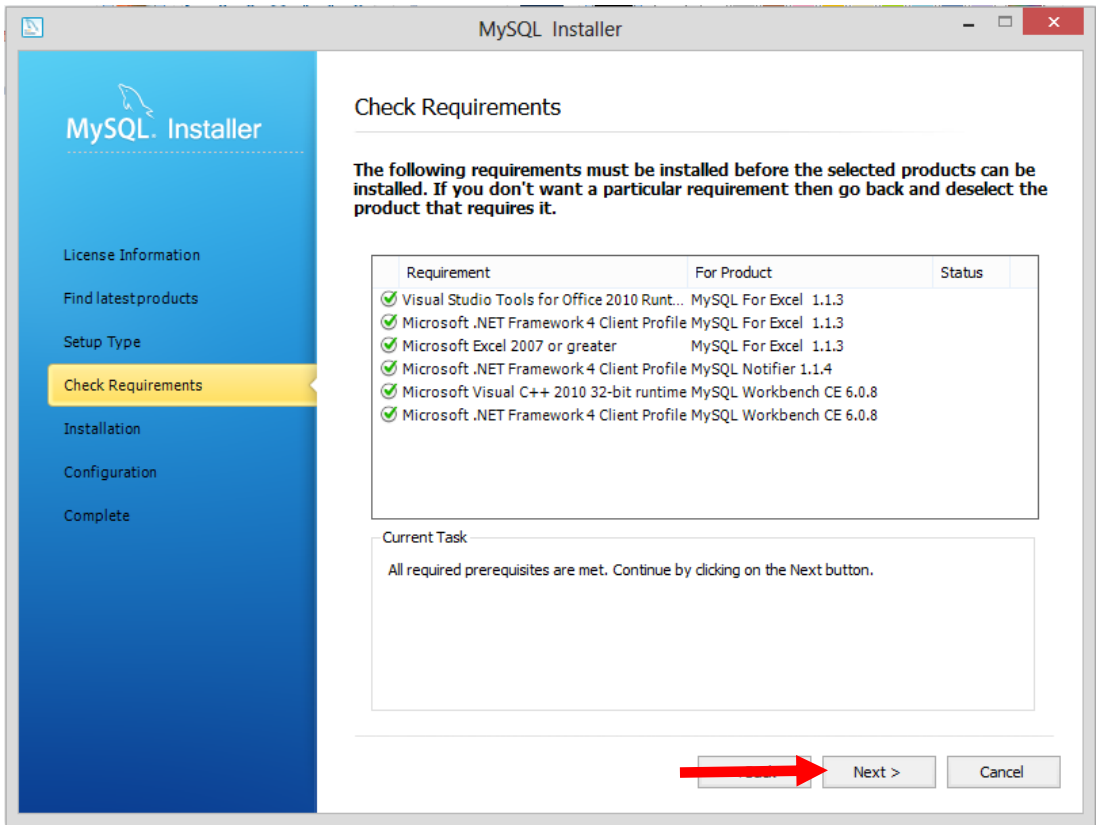


MySQL to HDFS – Using Sqoop

7 - Click **Execute**:

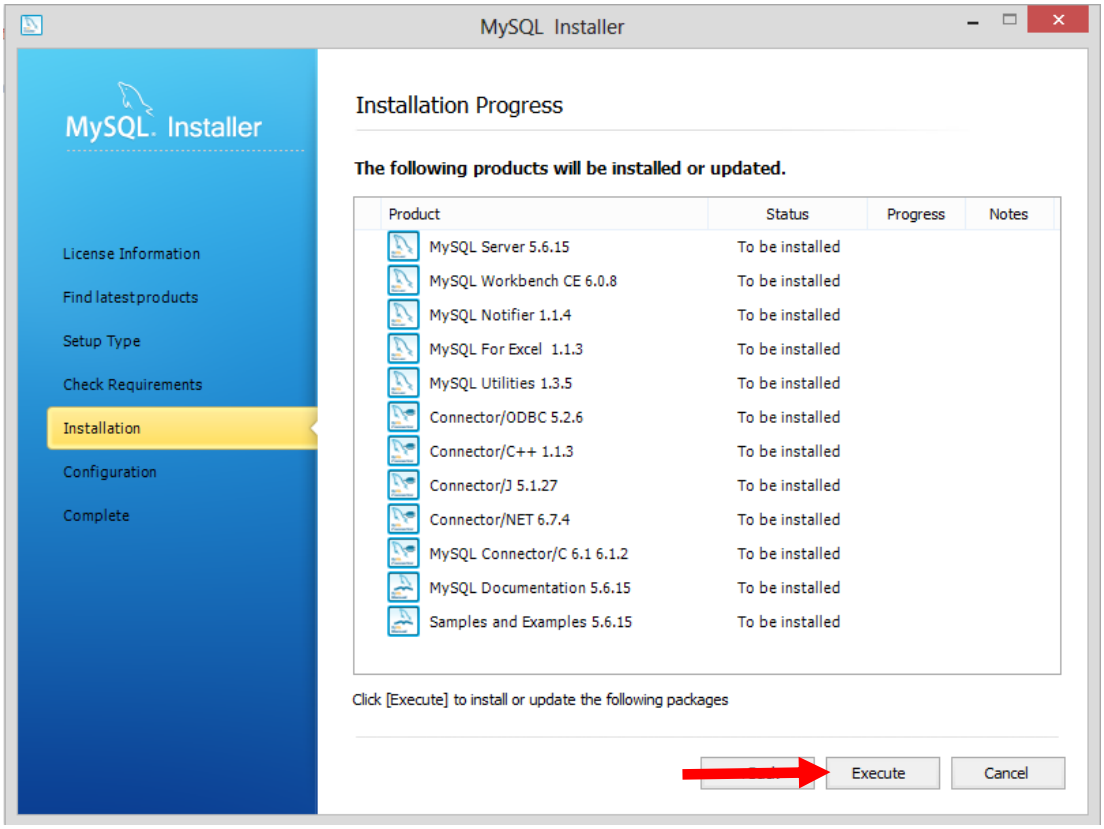


8 - Click **Next**:

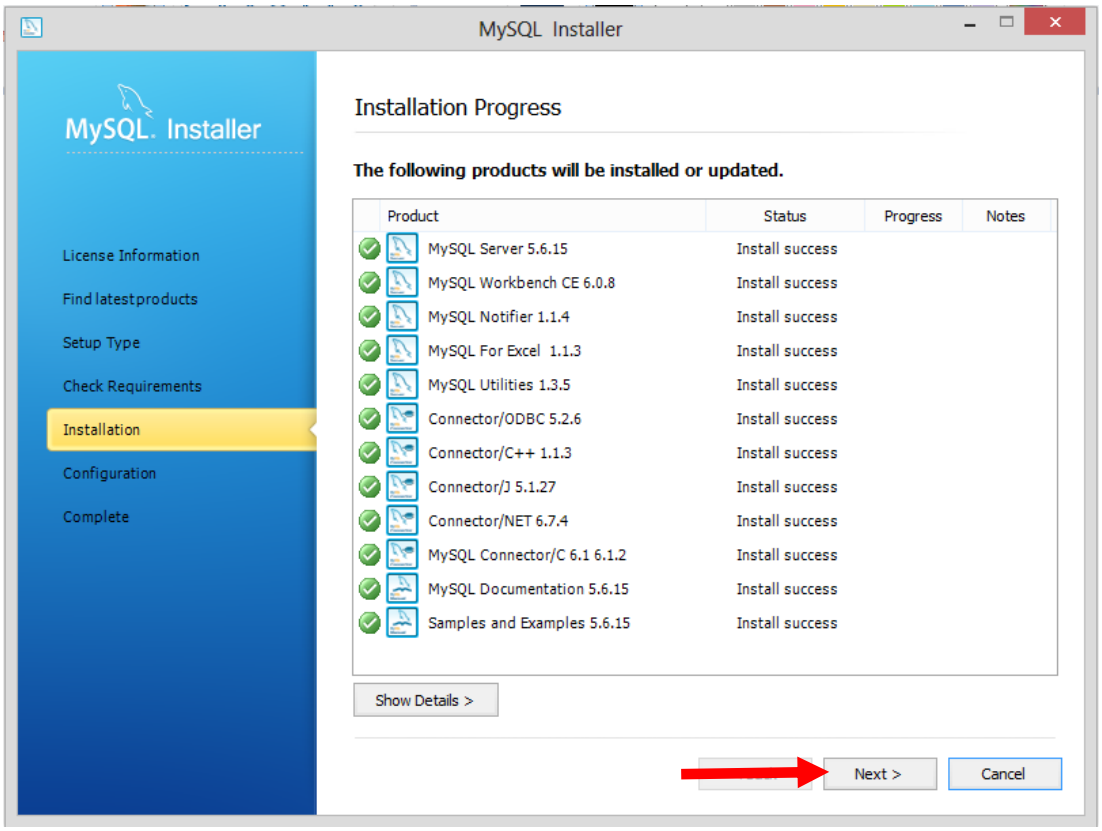


MySQL to HDFS – Using Sqoop

9 - Click **Execute**:

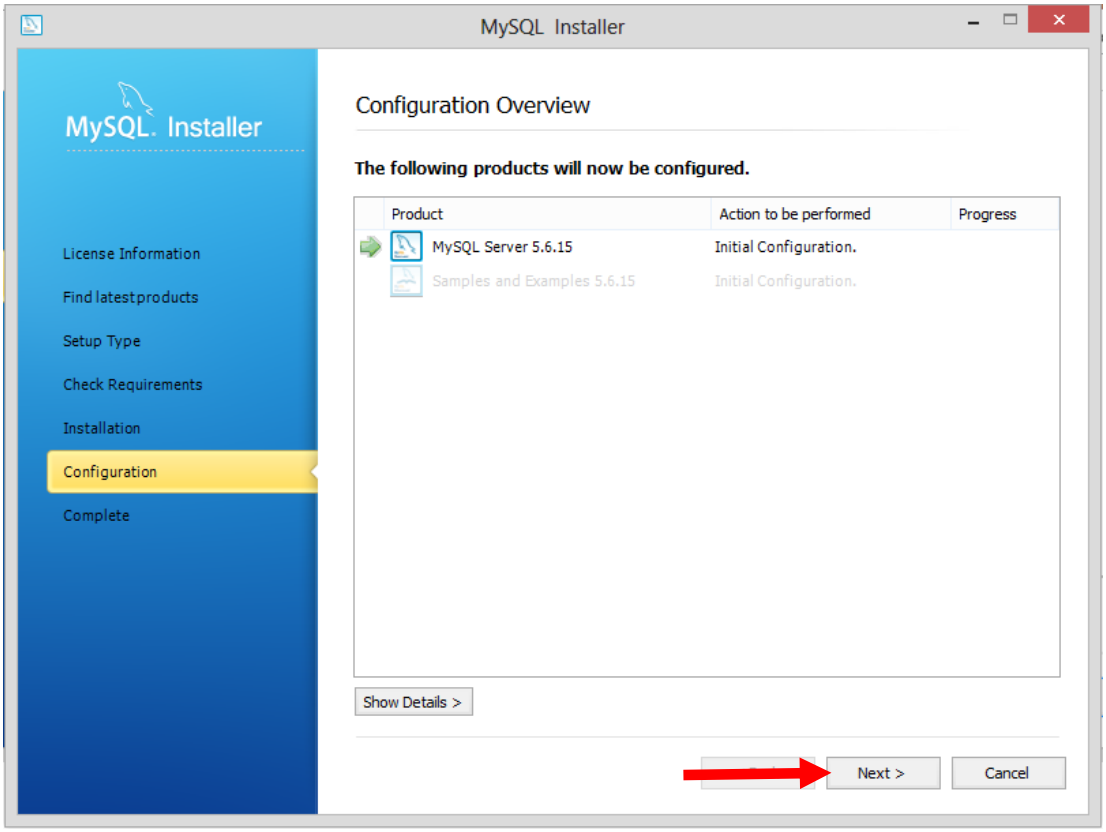


10 - Click **Next**:

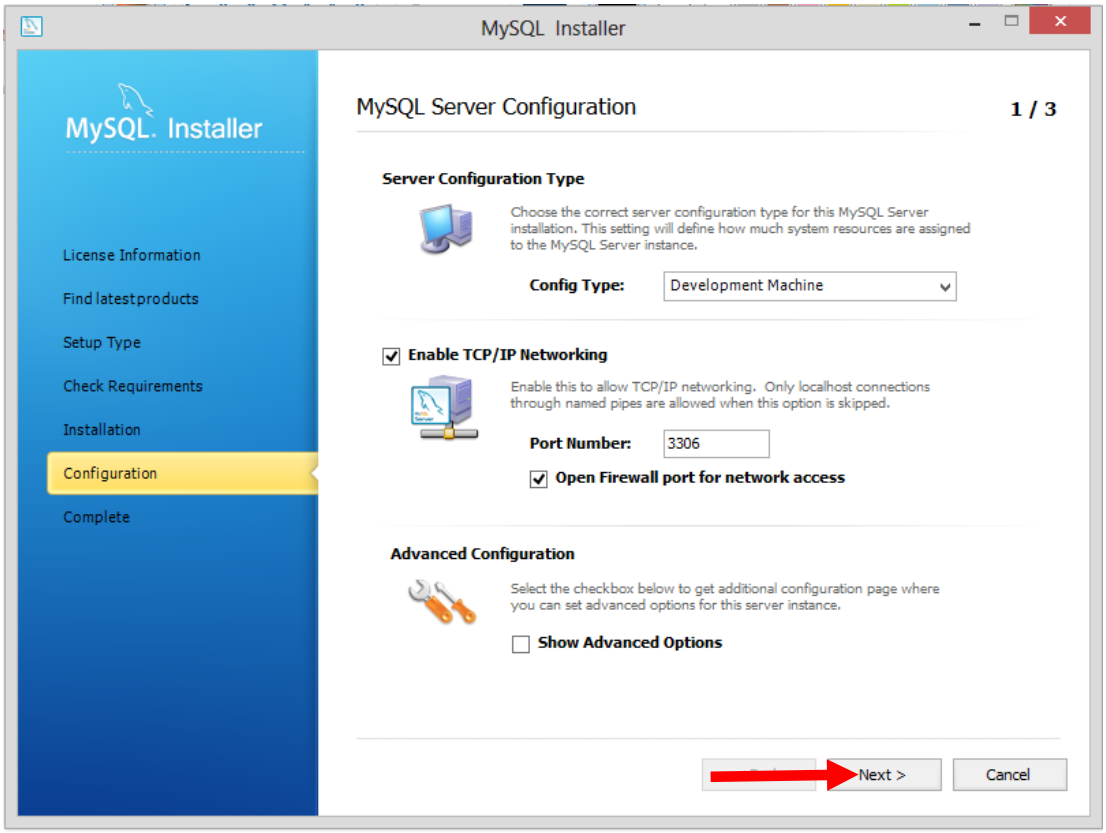


MySQL to HDFS – Using Sqoop

11 - Click **Next**:



12 - Click **Next**:

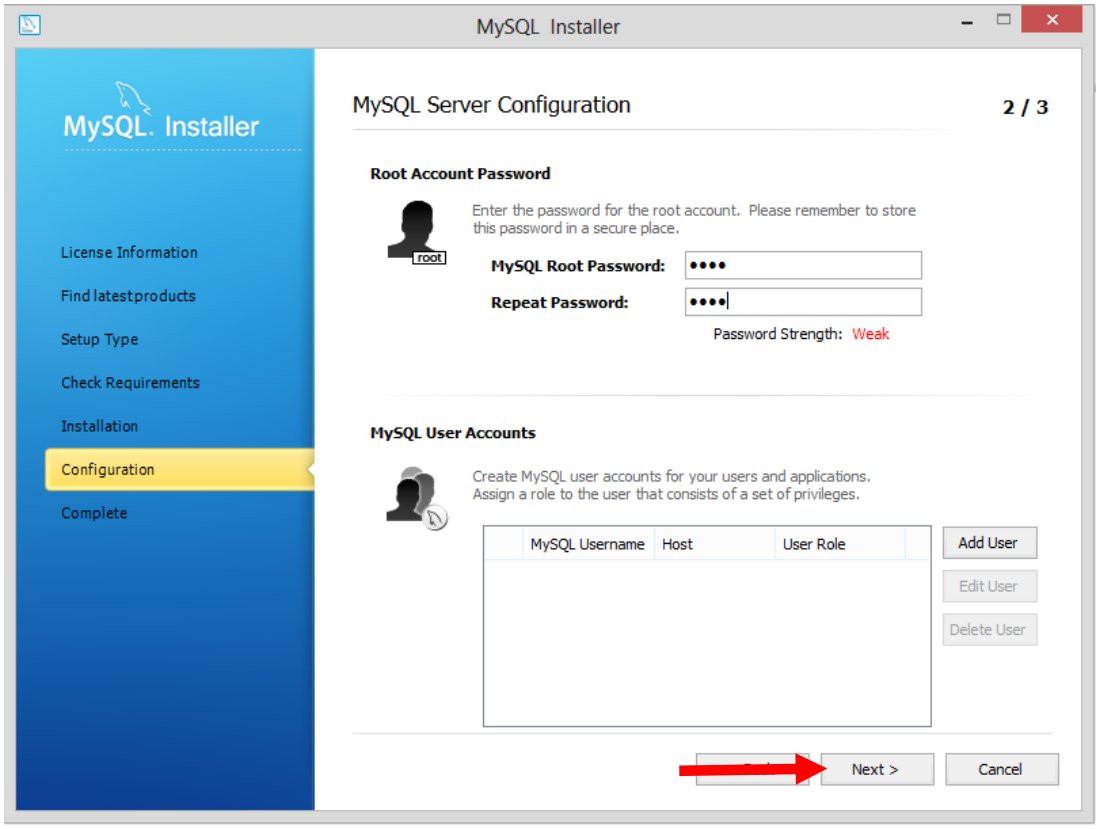


MySQL to HDFS – Using Sqoop

13 - Enter in **MySQL Root Password** –> **root**

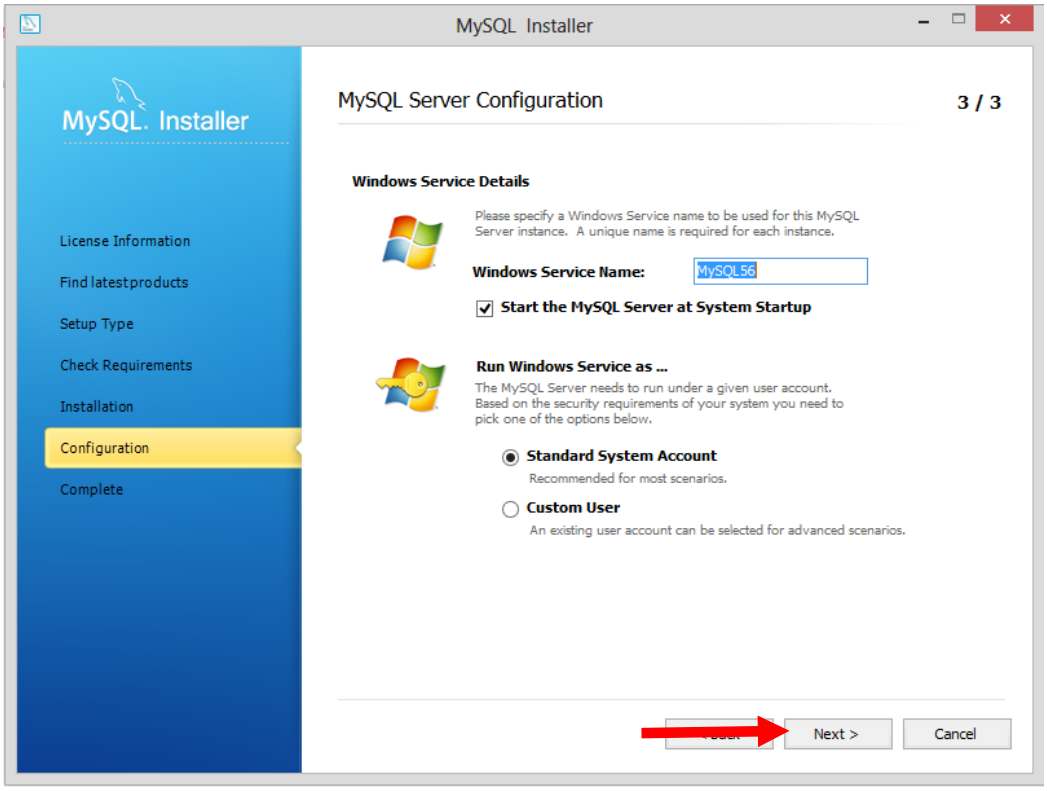
Enter in **Repeat Password** –> **root**

Click **Next**:

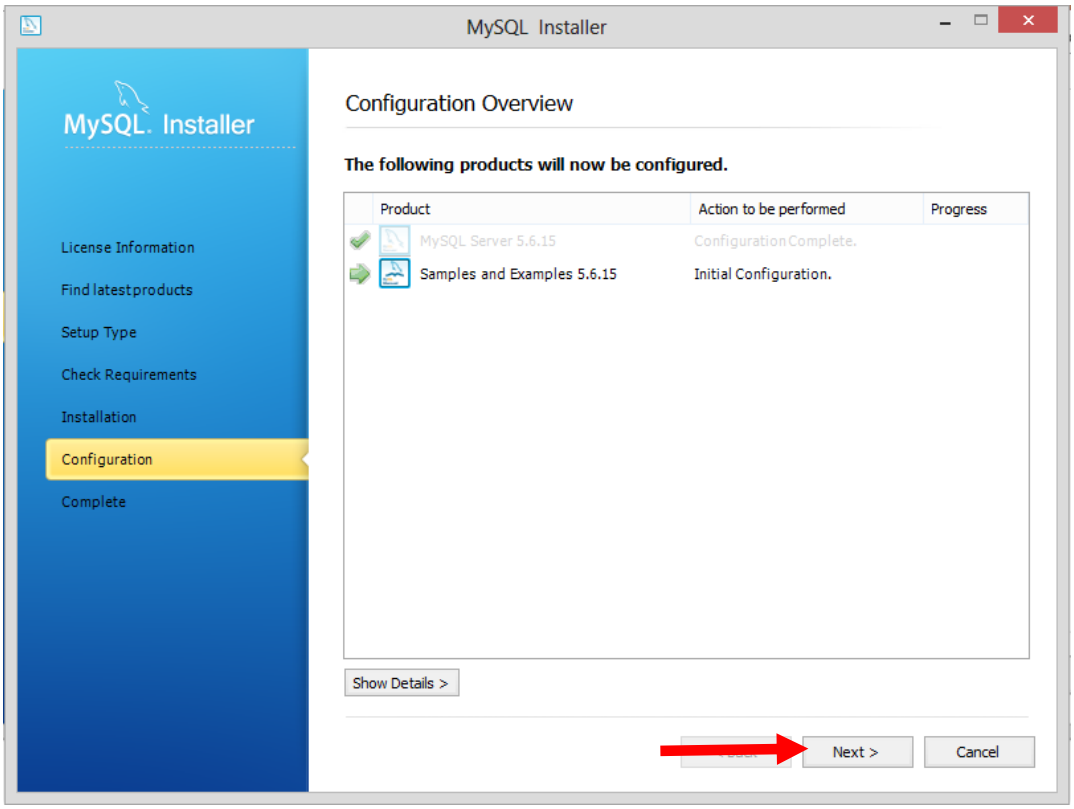


MySQL to HDFS – Using Sqoop

14 - Click **Next:**

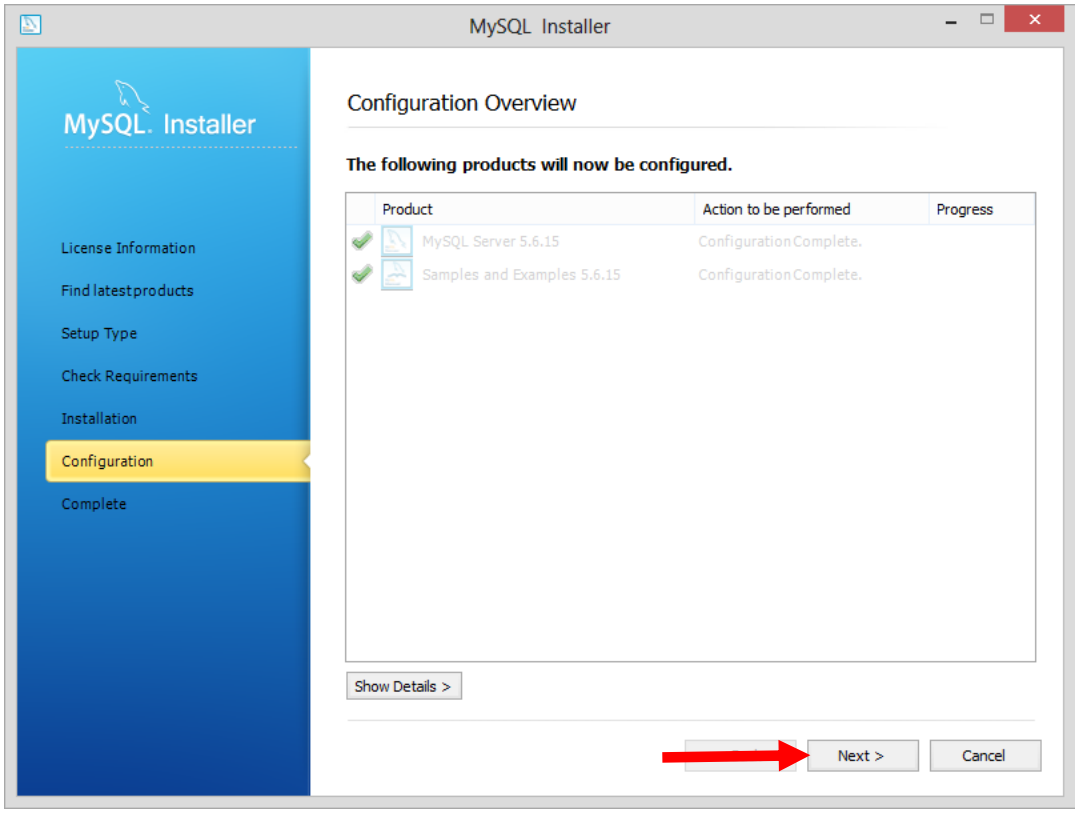


15 - Click **Next**:



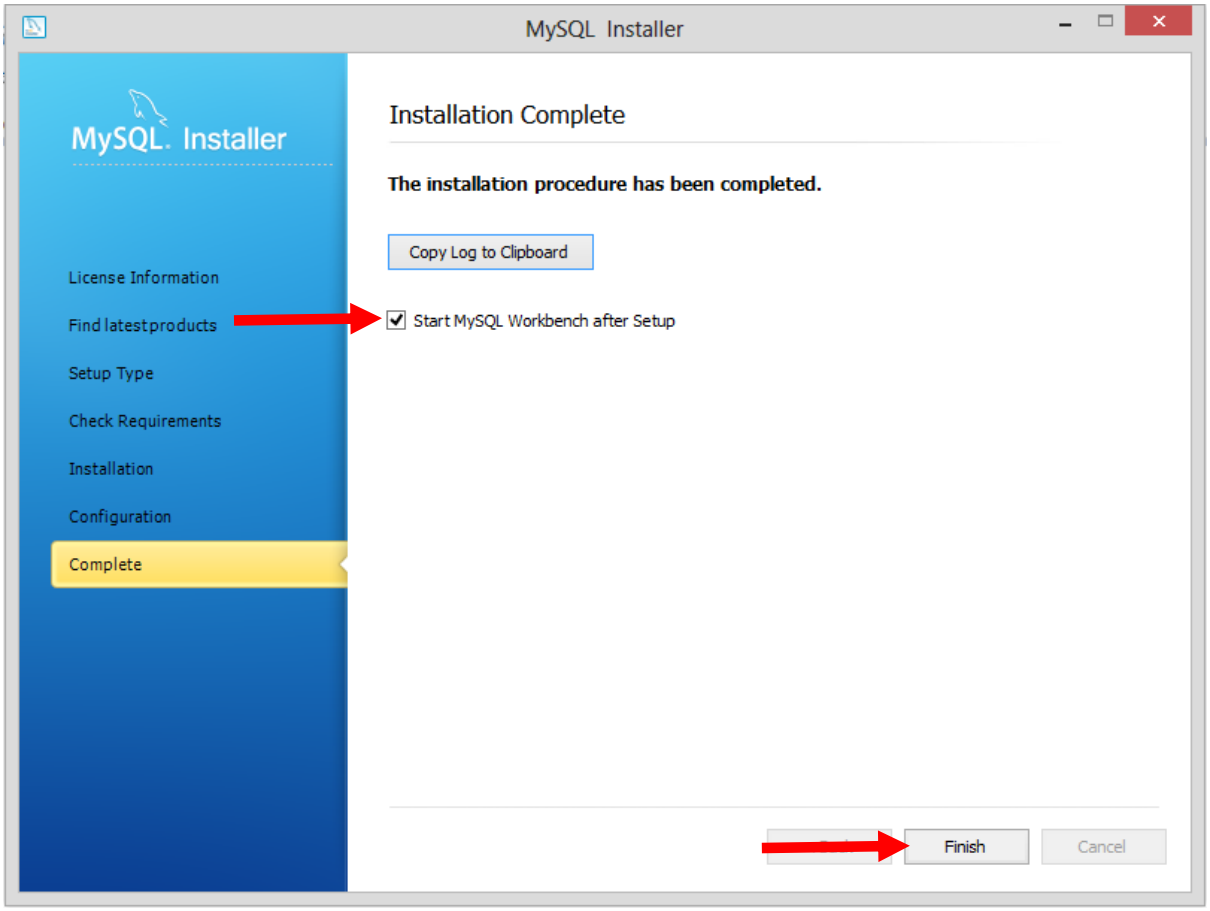
MySQL to HDFS – Using Sqoop

16 - Click **Next**:



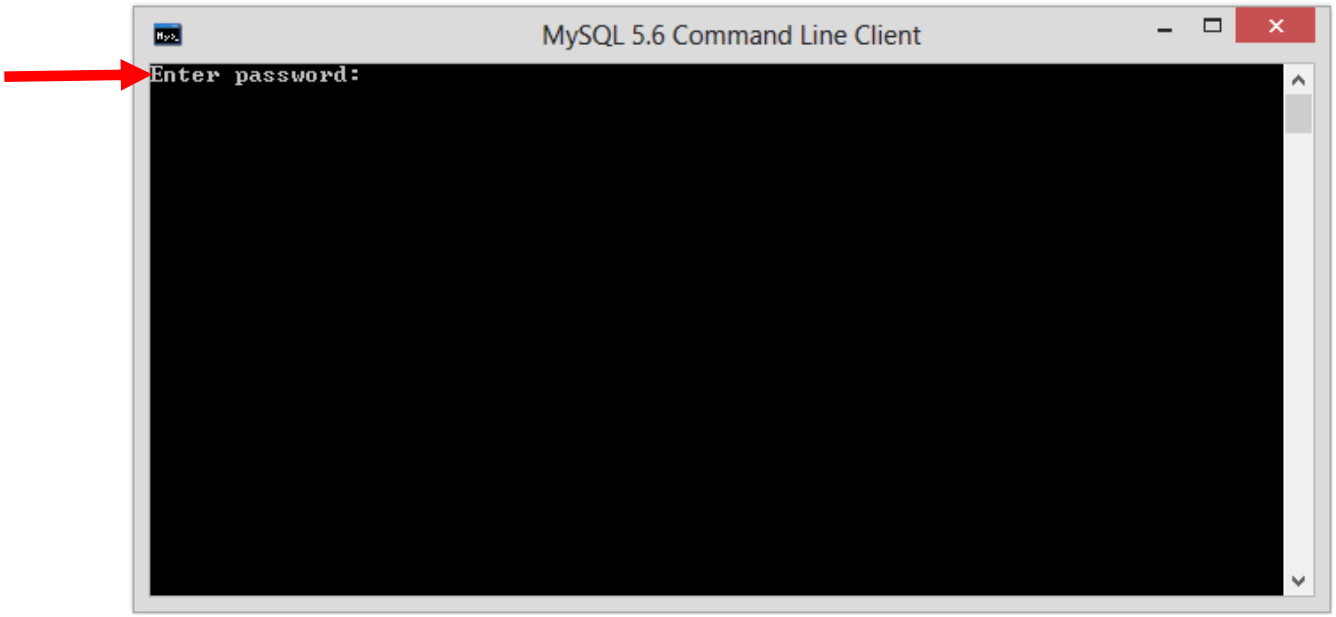
MySQL to HDFS – Using Sqoop

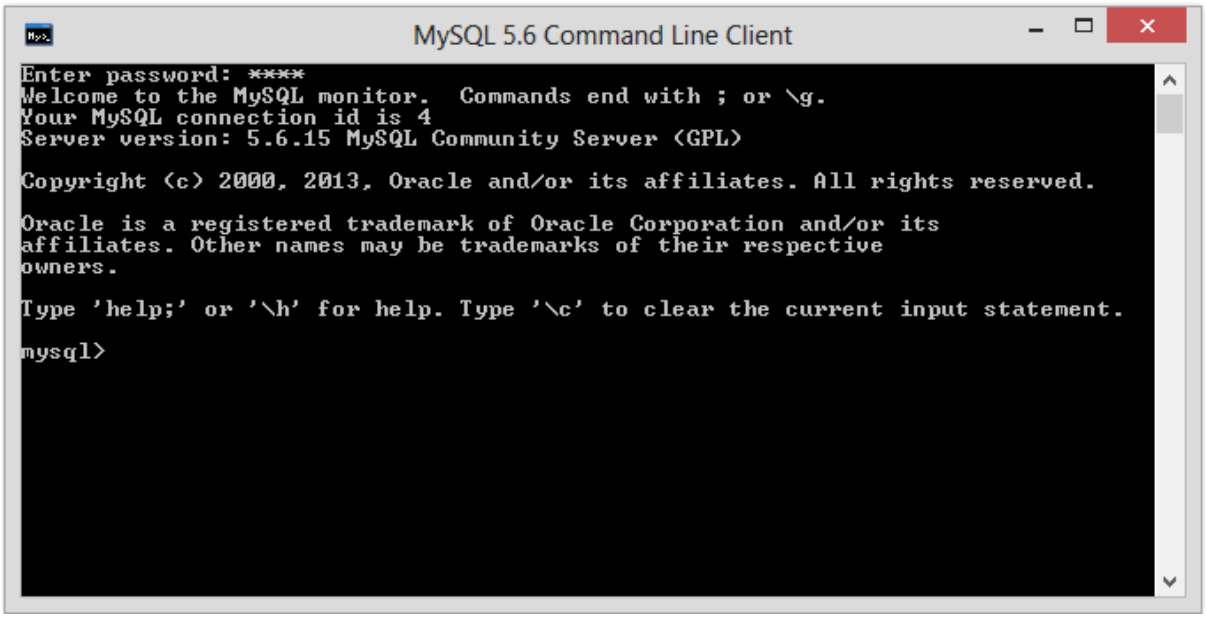
17 - Uncheck the check-box (Start MySQL Workbench after Setup) and Click **Finish**:



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**

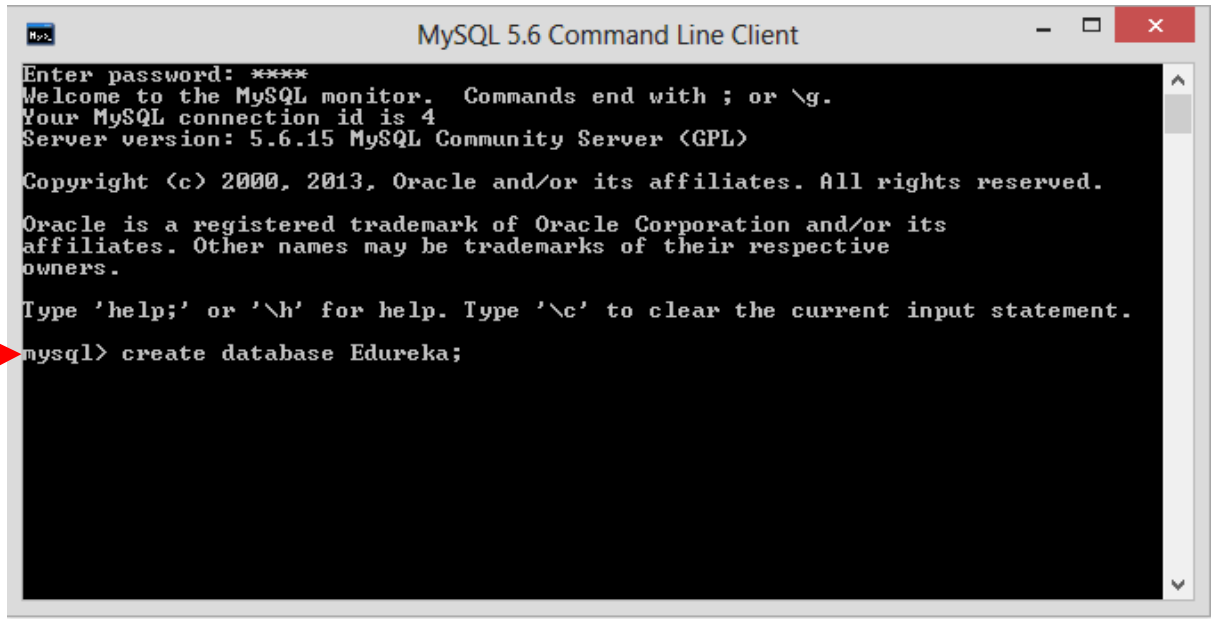


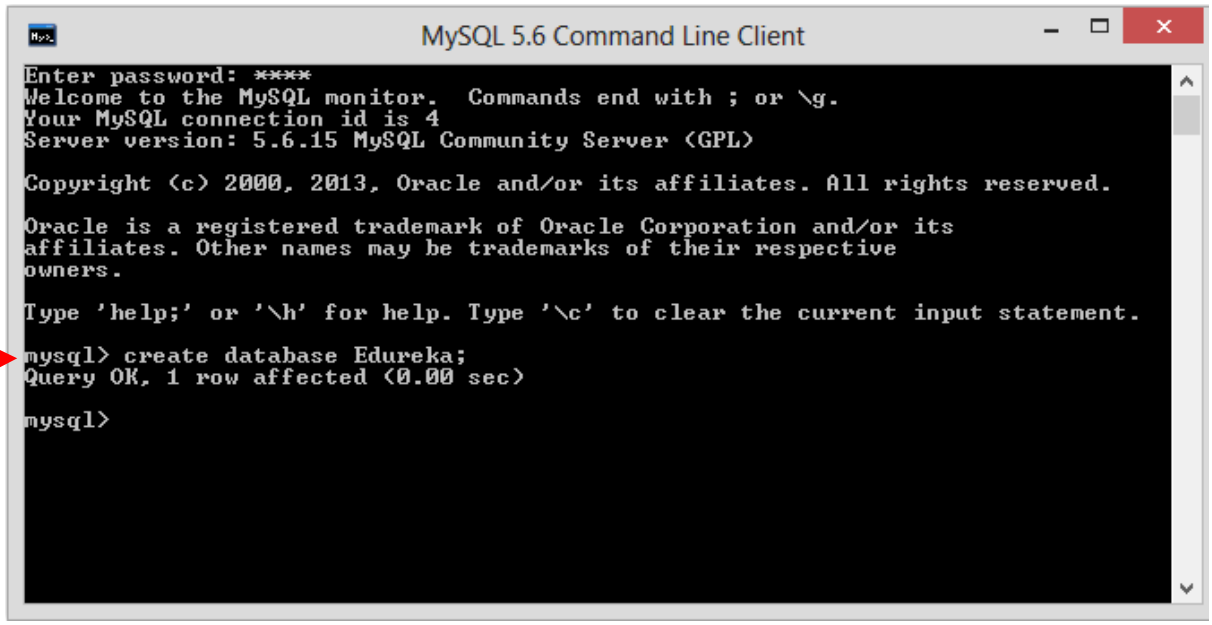


MySQL to HDFS – Using Sqoop

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

Command: create database Edureka;

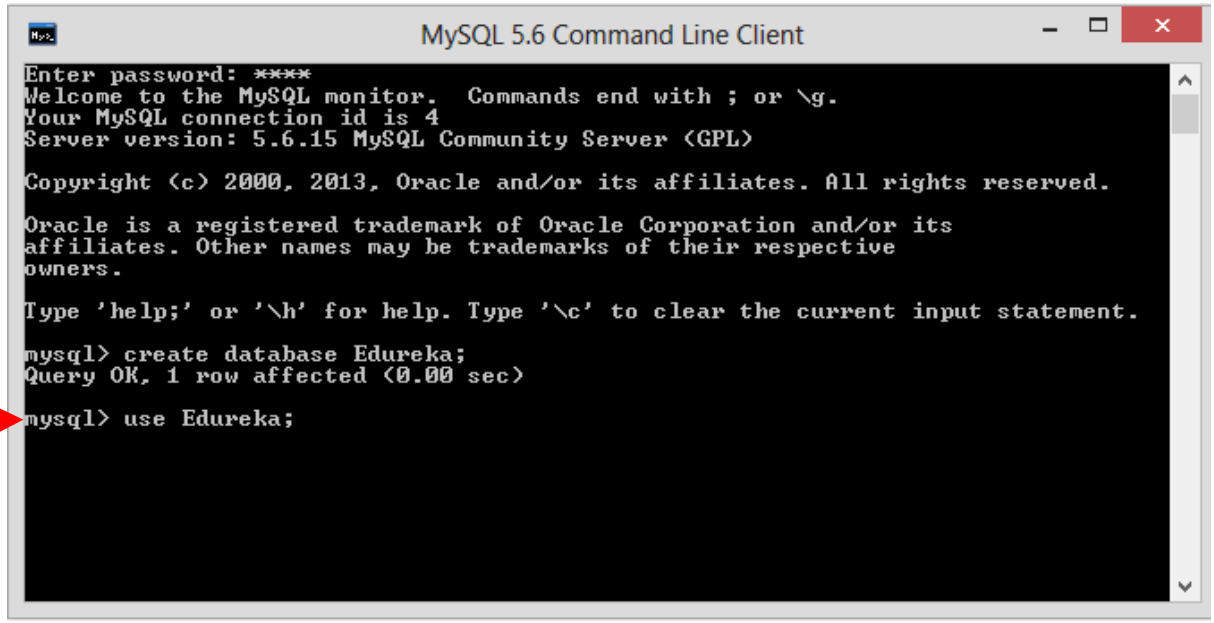


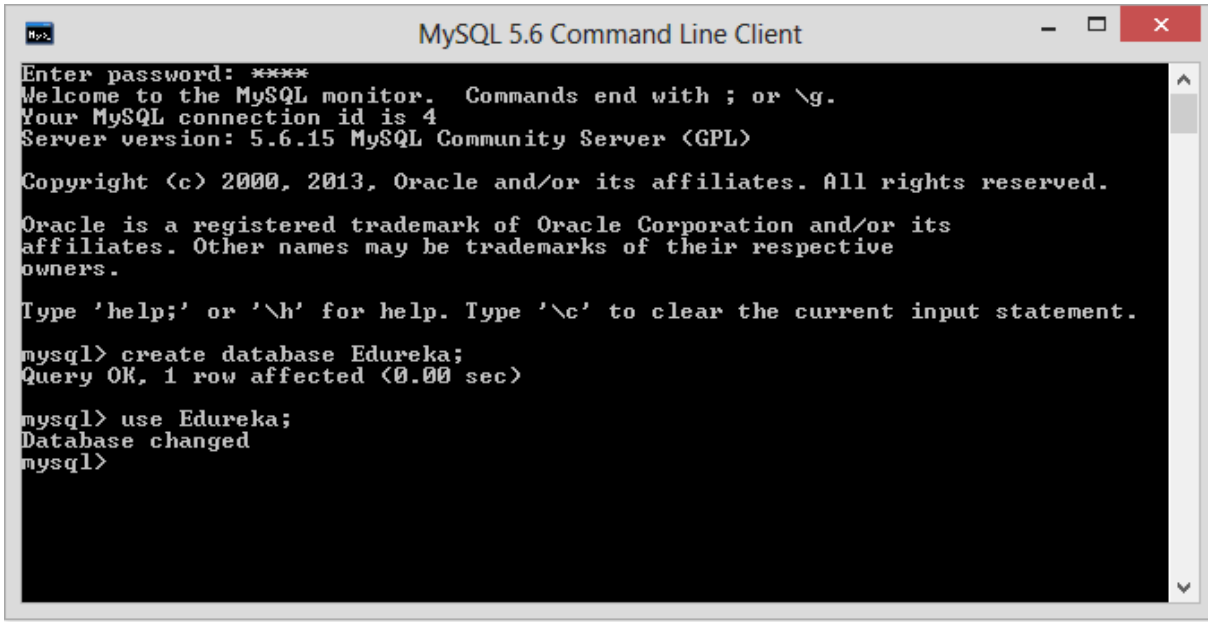


MySQL to HDFS – Using Sqoop

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

Command: use Edureka;

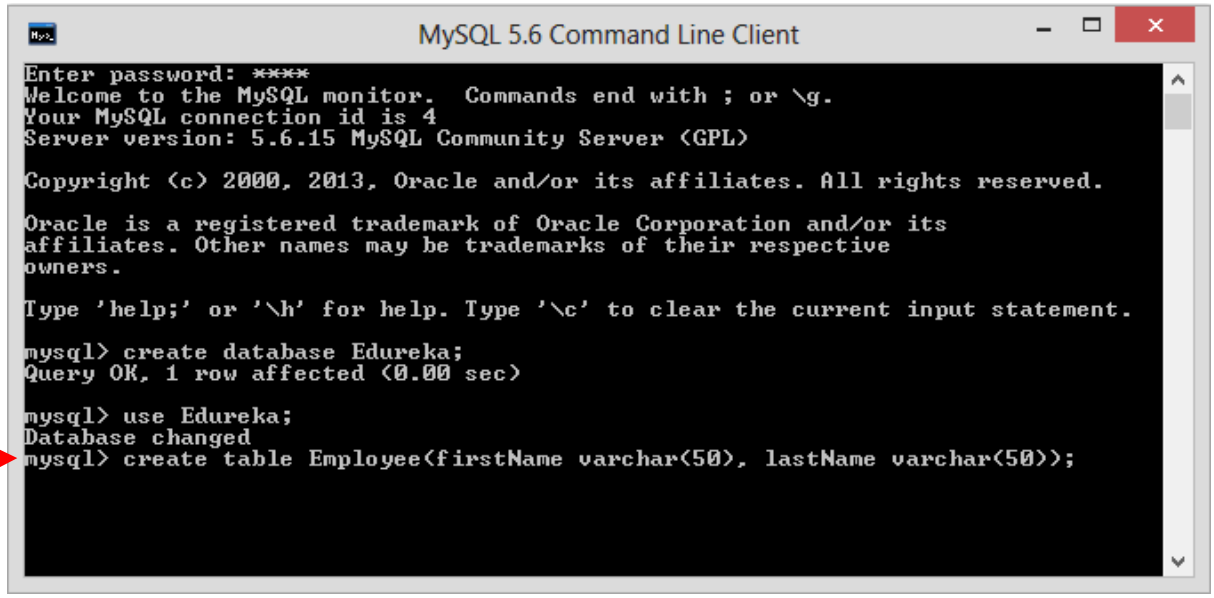


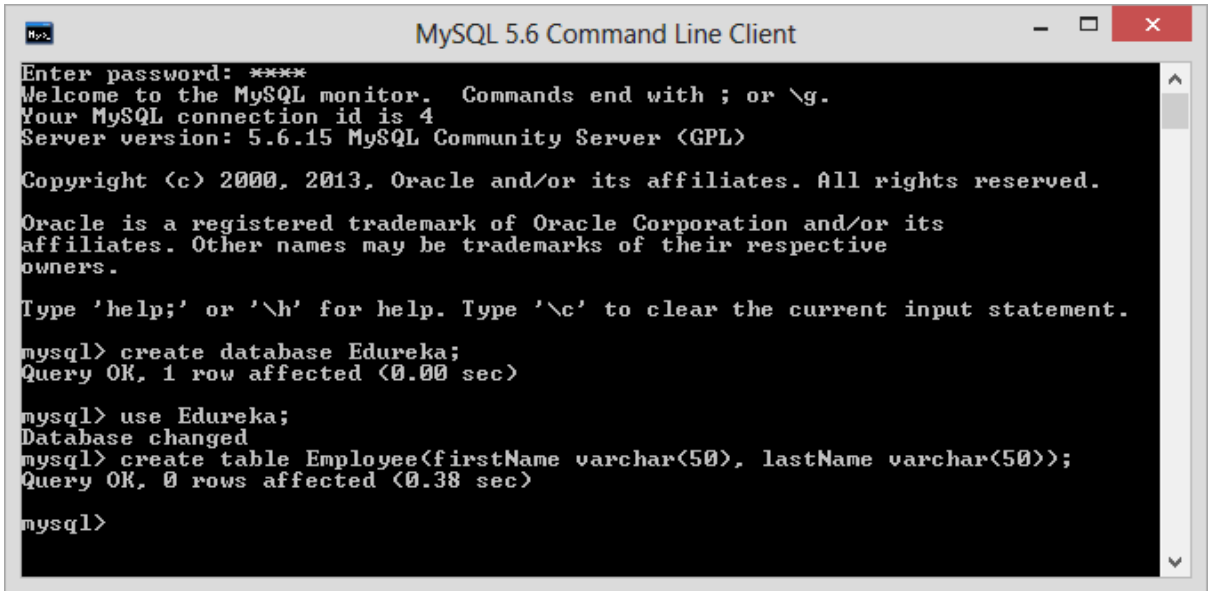


MySQL to HDFS – Using Sqoop

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

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

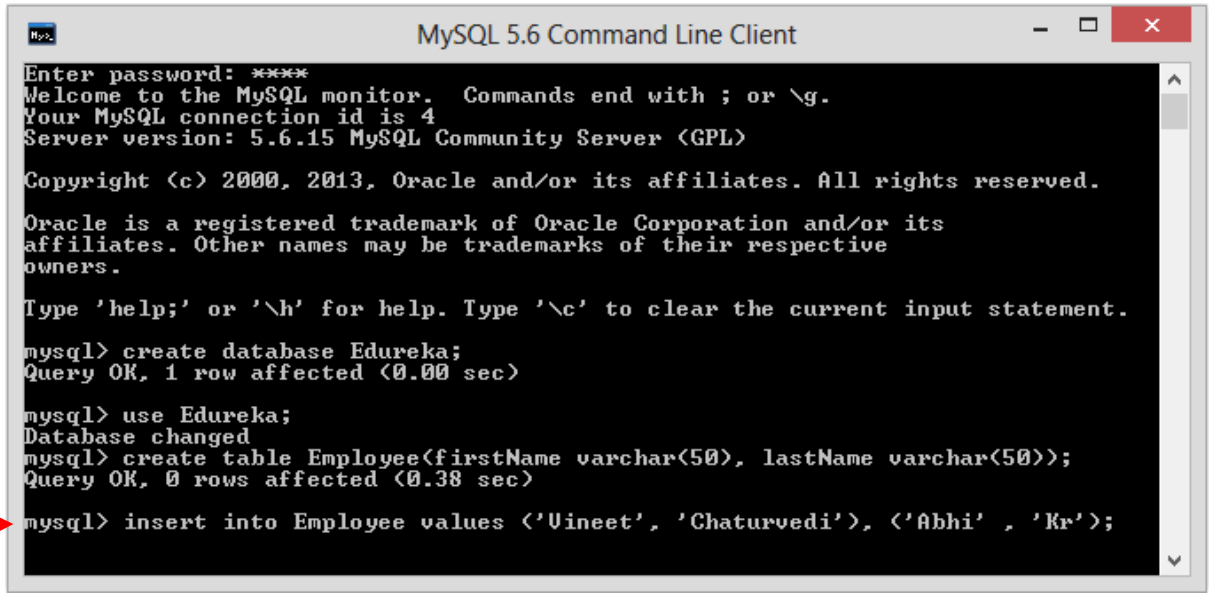


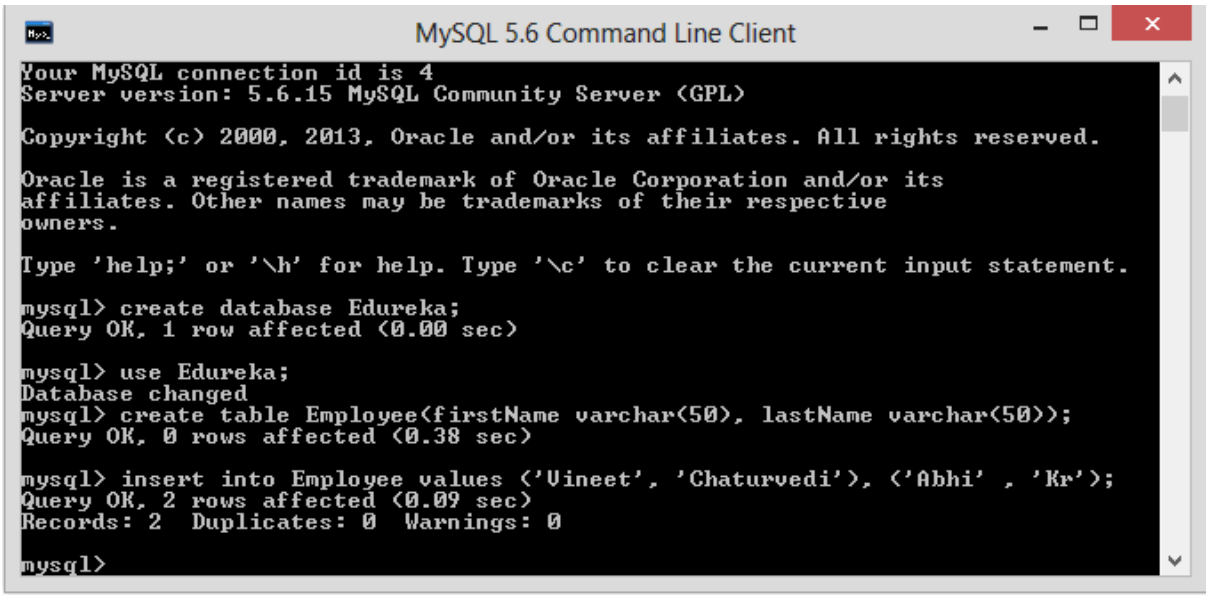


MySQL to HDFS – Using Sqoop

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

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

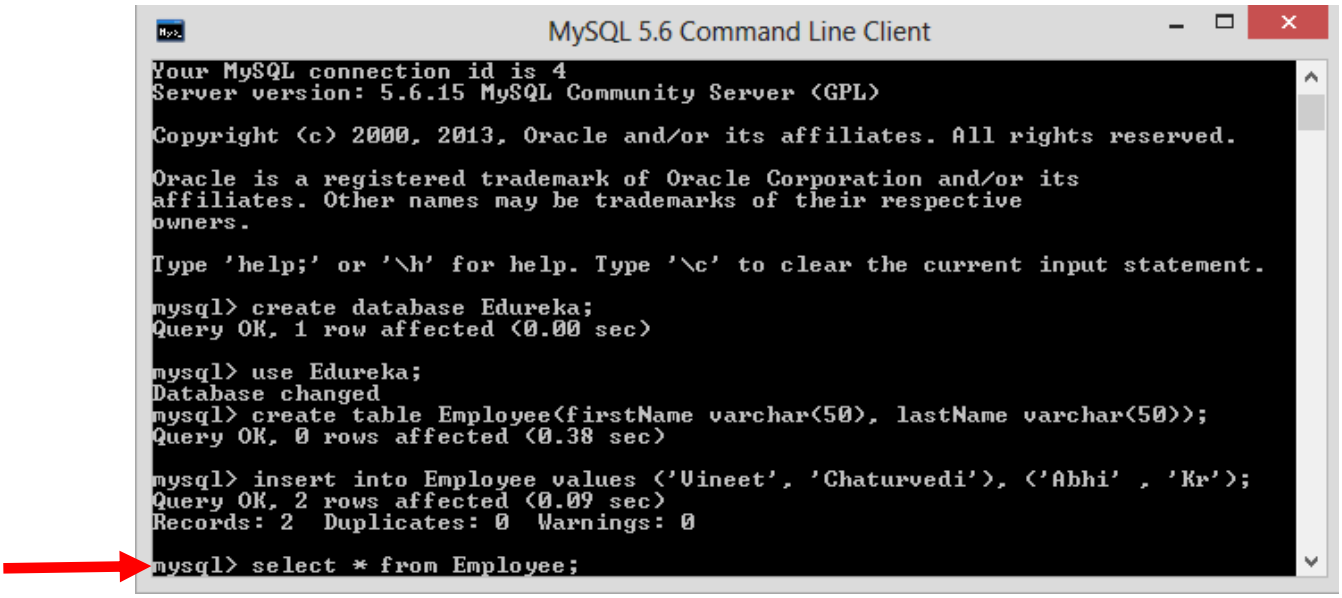


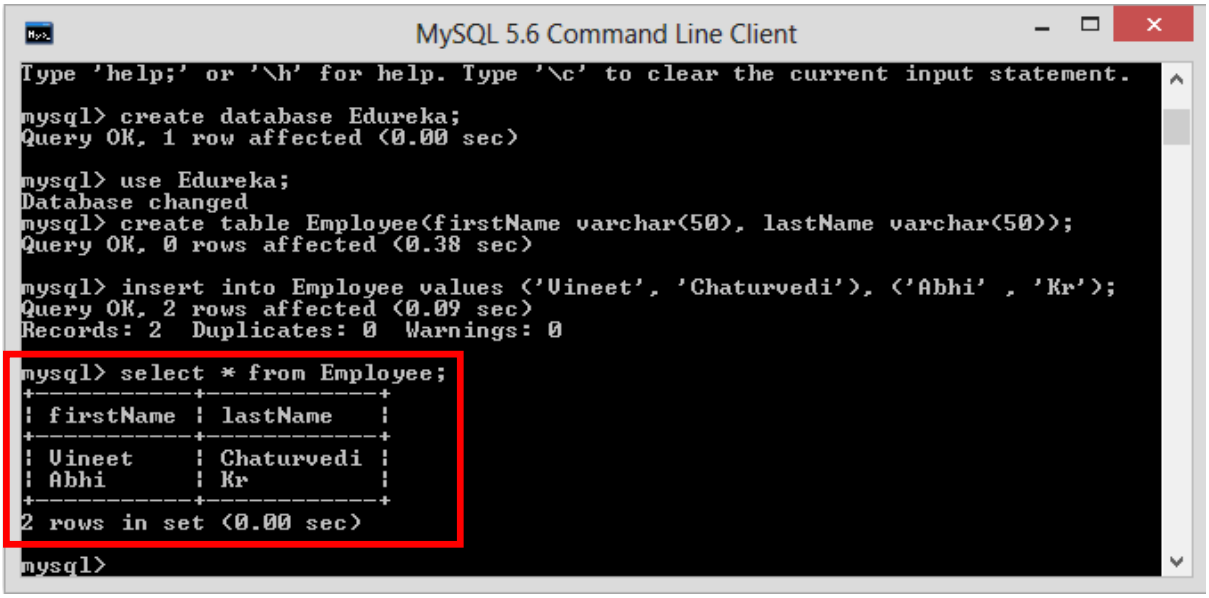


MySQL to HDFS – Using Sqoop

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

Command: select * from Employee;

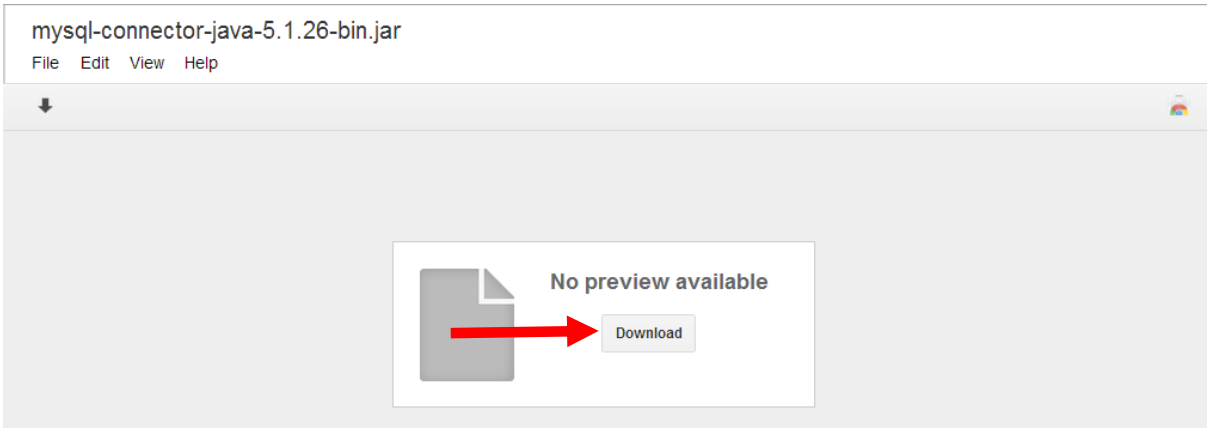




MySQL to HDFS – Using Sqoop

24 - Download MySQL connector using the below link:

https://drive.google.com/file/d/0B2-rlCGKD40NSVJ3QlpWSk95OFE/edit?usp=sharing



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

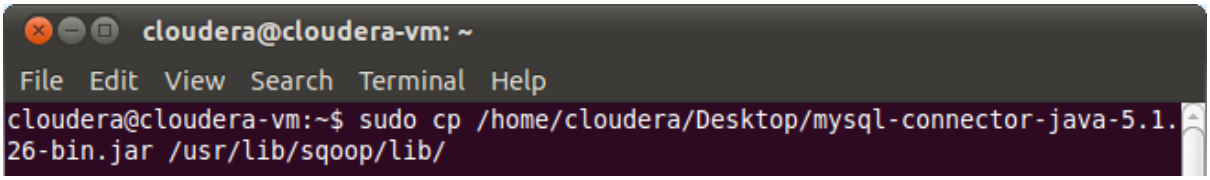
Use the below link to understand how to move a file from Windows to cloudera cdh3 vm.

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:

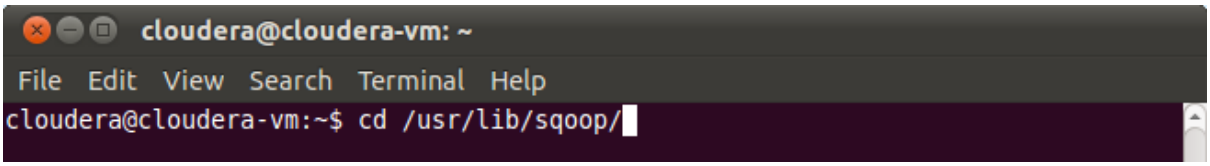
Command:

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:

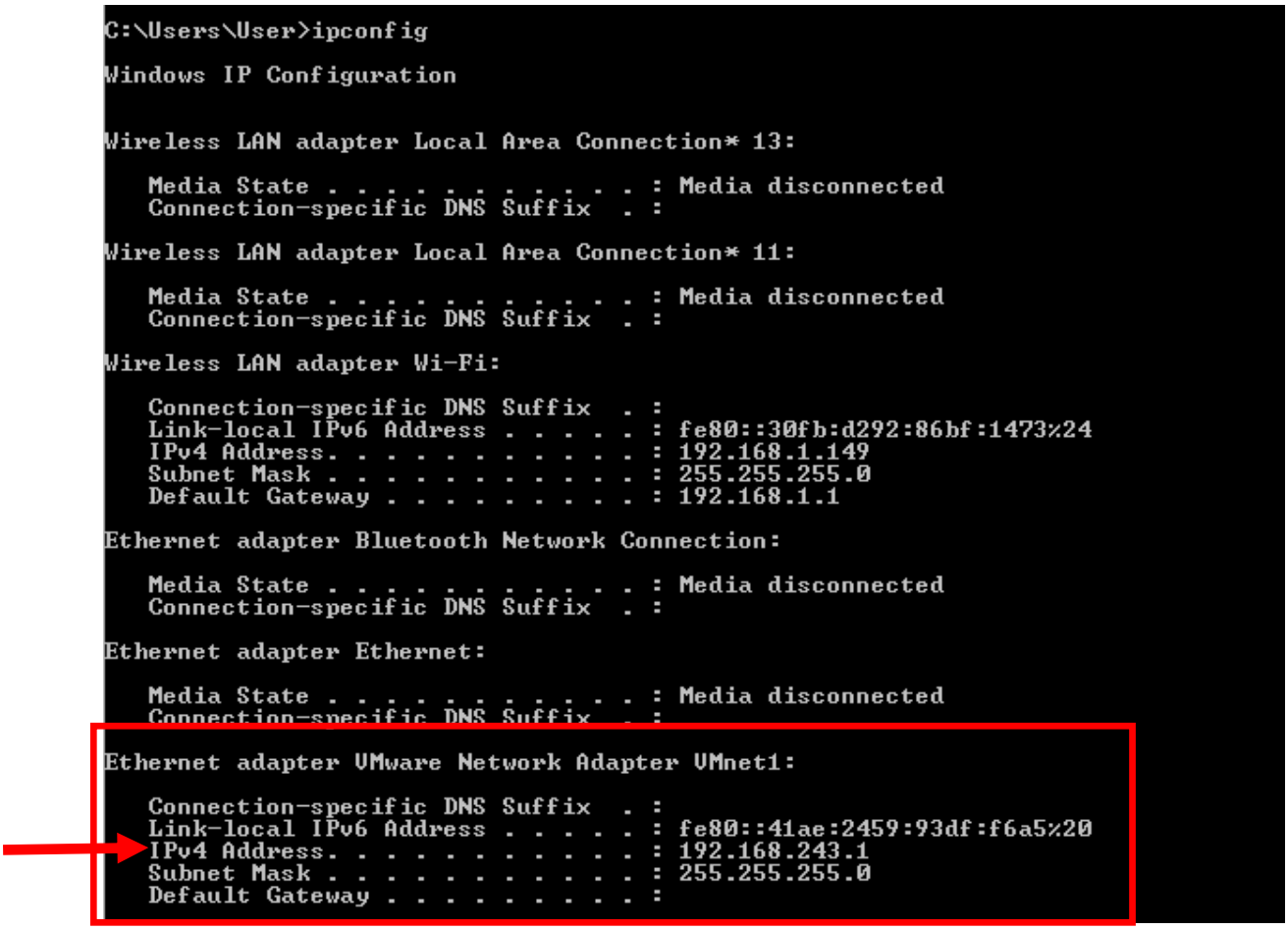
Command: cd /usr/lib/sqoop



MySQL to HDFS – Using Sqoop

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

Command: ipconfig



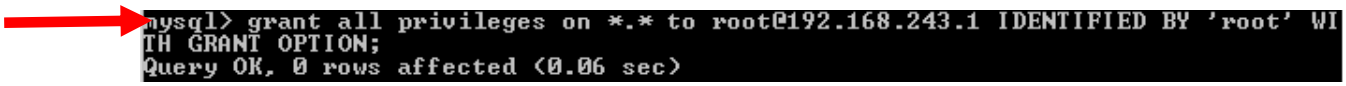
29 - Grant all privileges to root@your_ipv4_address by executing the below command (in **MySQL 5.6 Command Line Client**):

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 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

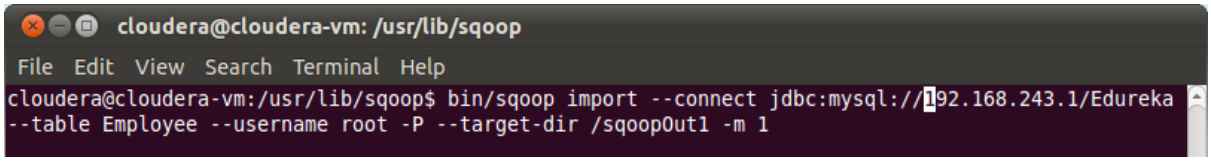
Table Name – Employee

Username – root

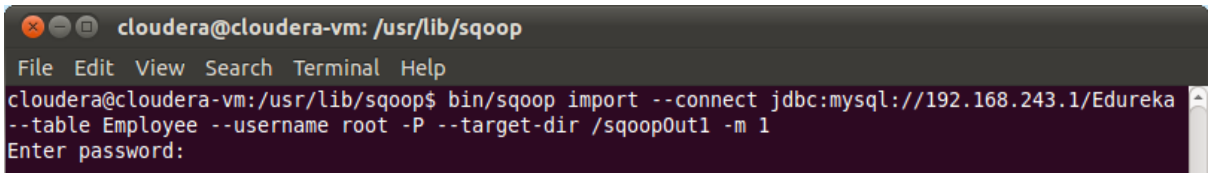
Output Directory – Could be any. I have used sqoopOut1

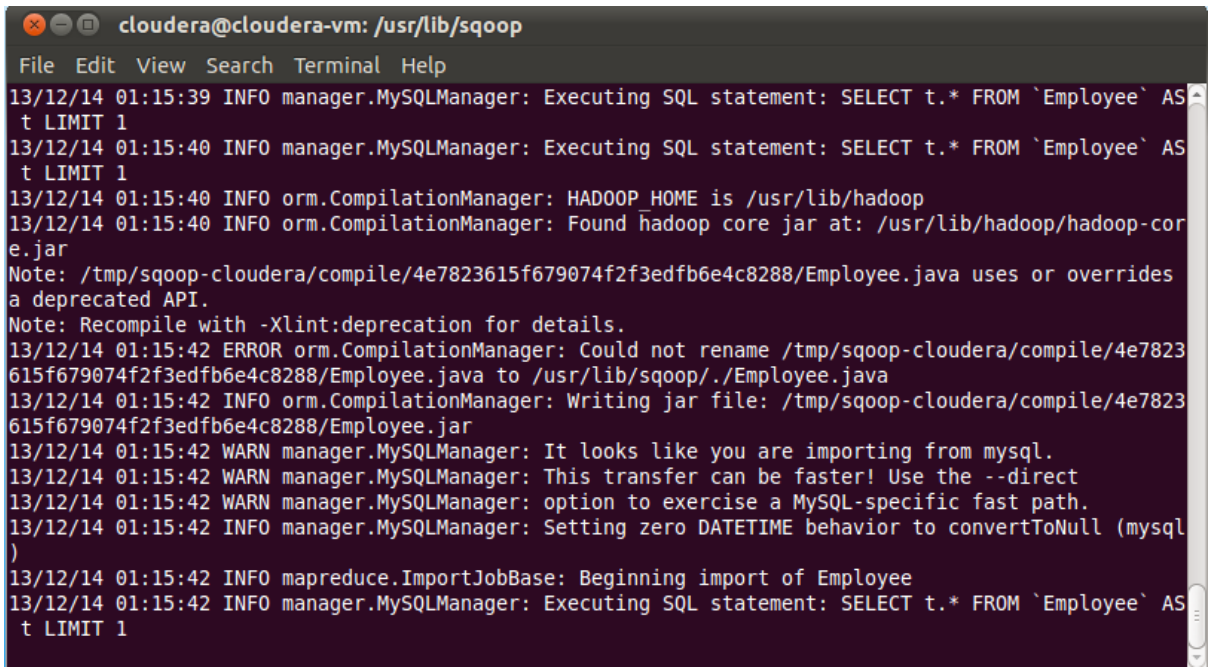
Command:

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



31 - Enter the Password - **root**





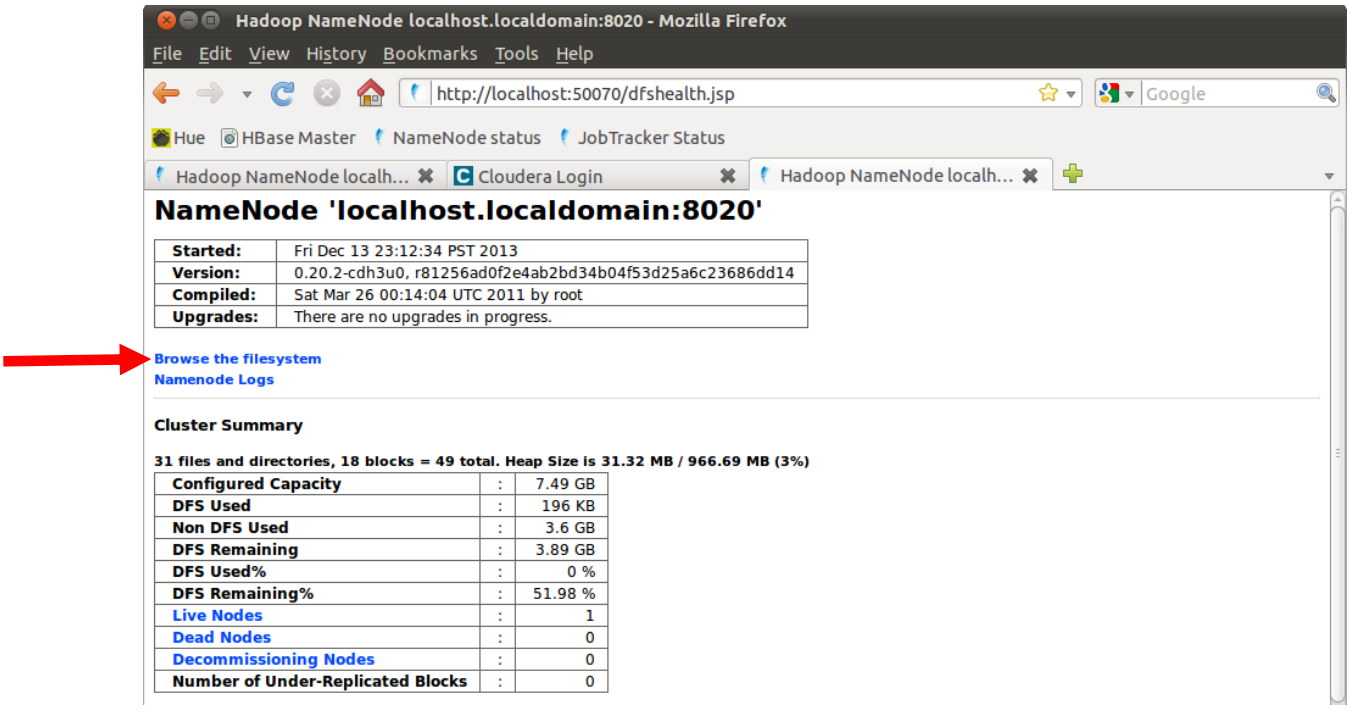
MySQL to HDFS – Using Sqoop

```
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_REDUCES=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:

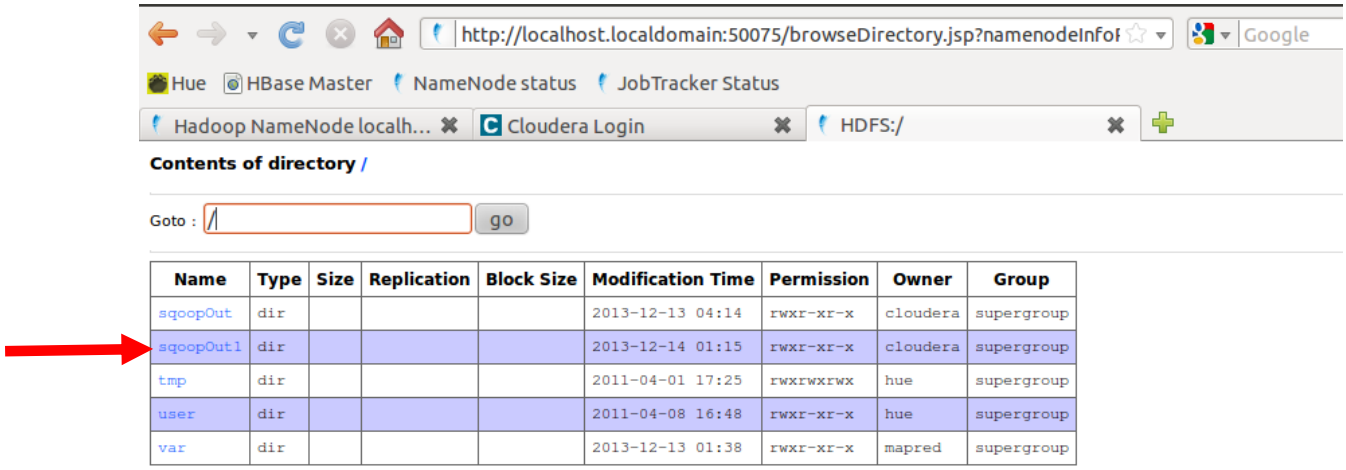
URL: http://localhost:50070/dfshealth.jsp

Click on **Browse the filesystem**

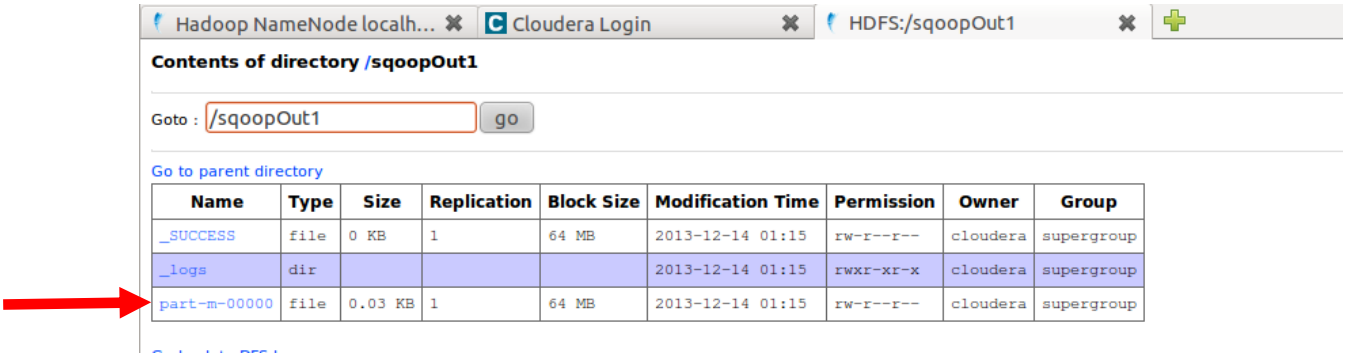


MySQL to HDFS – Using Sqoop

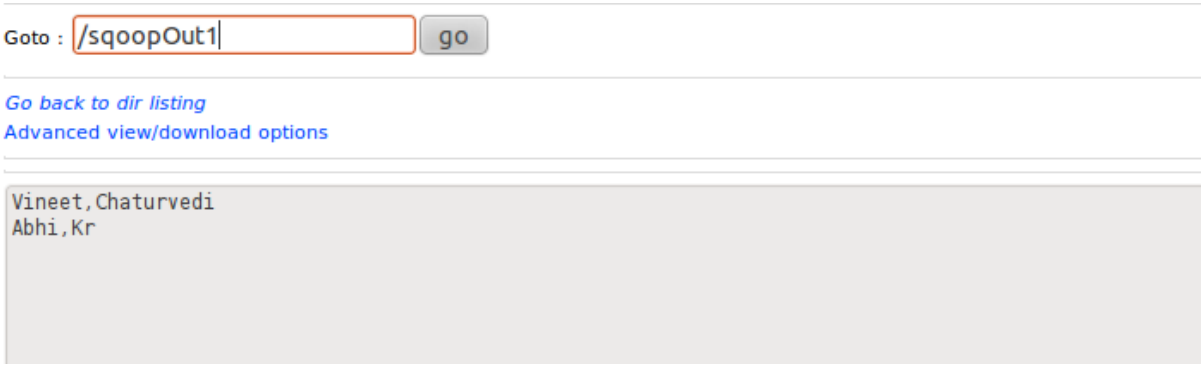
33 - Click on **sqoopOut1 directory**:



34 - Click on **part-m-00000** file:



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



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