# Importing data from MySQL to HDFS using Sqoop

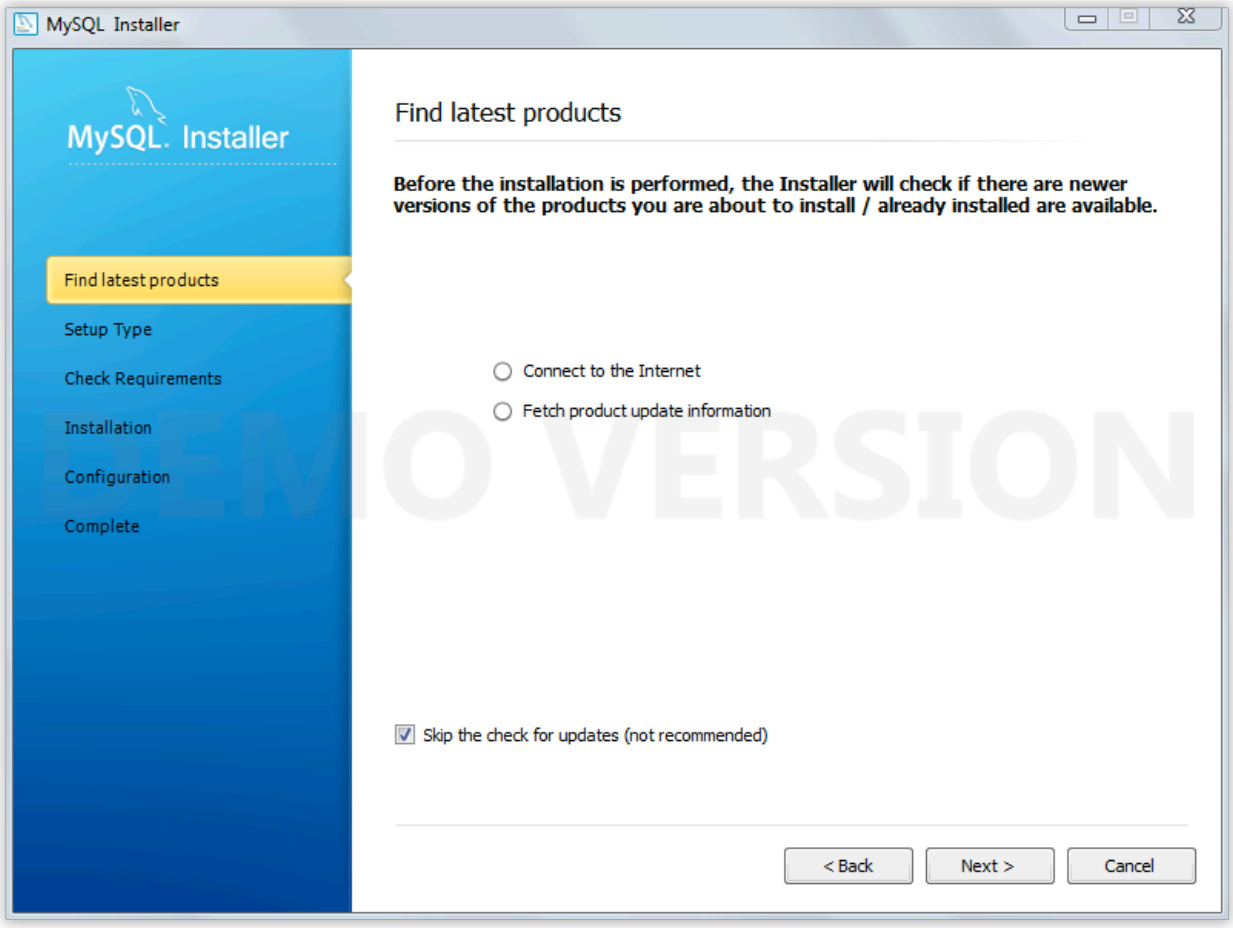
## a) MySQL Installation on Windows

1. Download MySQL Installer and unzip it.

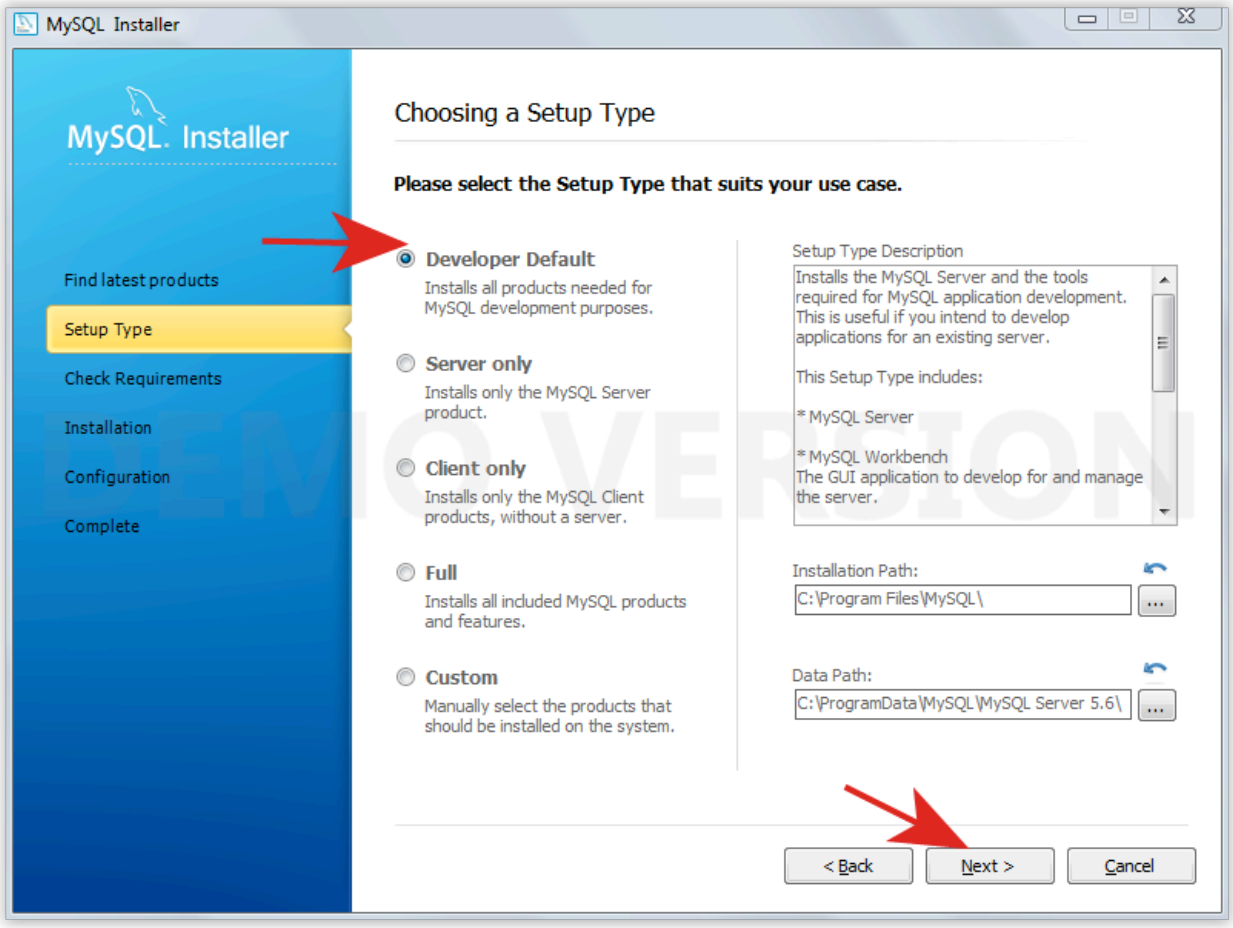
   http://dev.mysql.com/downloads/mysql/

2. Double click on the extracted file and click on **Install Mysql Products.**

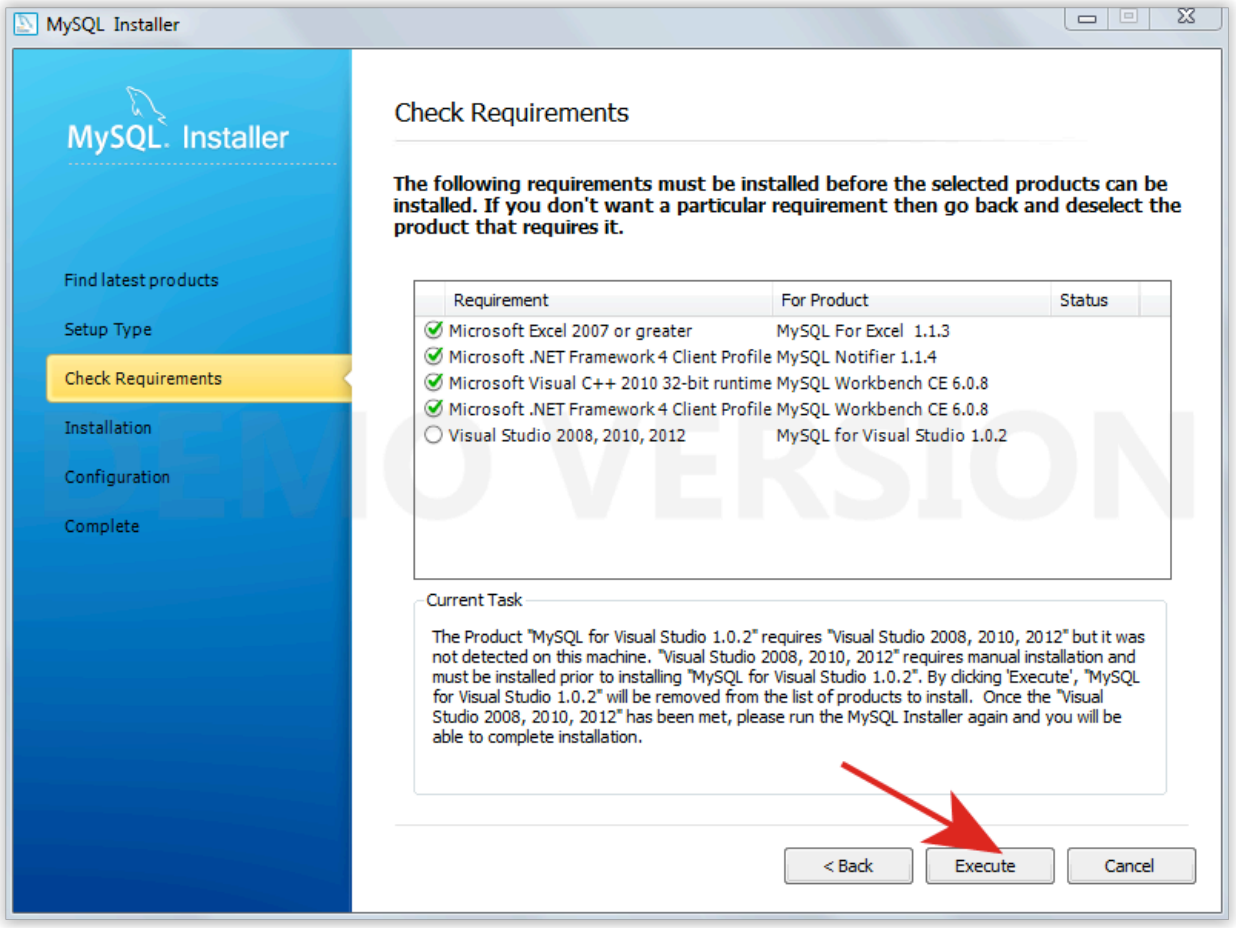
3. Click Next


MySQL Installer
MySQL. Installer
Find latest products
Setup Type
Check Requirements
Installation
Configuration
Complete
Find latest products
Before the installation is performed, the Installer will check if there are newer versions of the products you are about to install / already installed are available.
Connect to the Internet
Fetch product update information
Skip the check for updates (not recommended)
< Back
Next >
Cancel

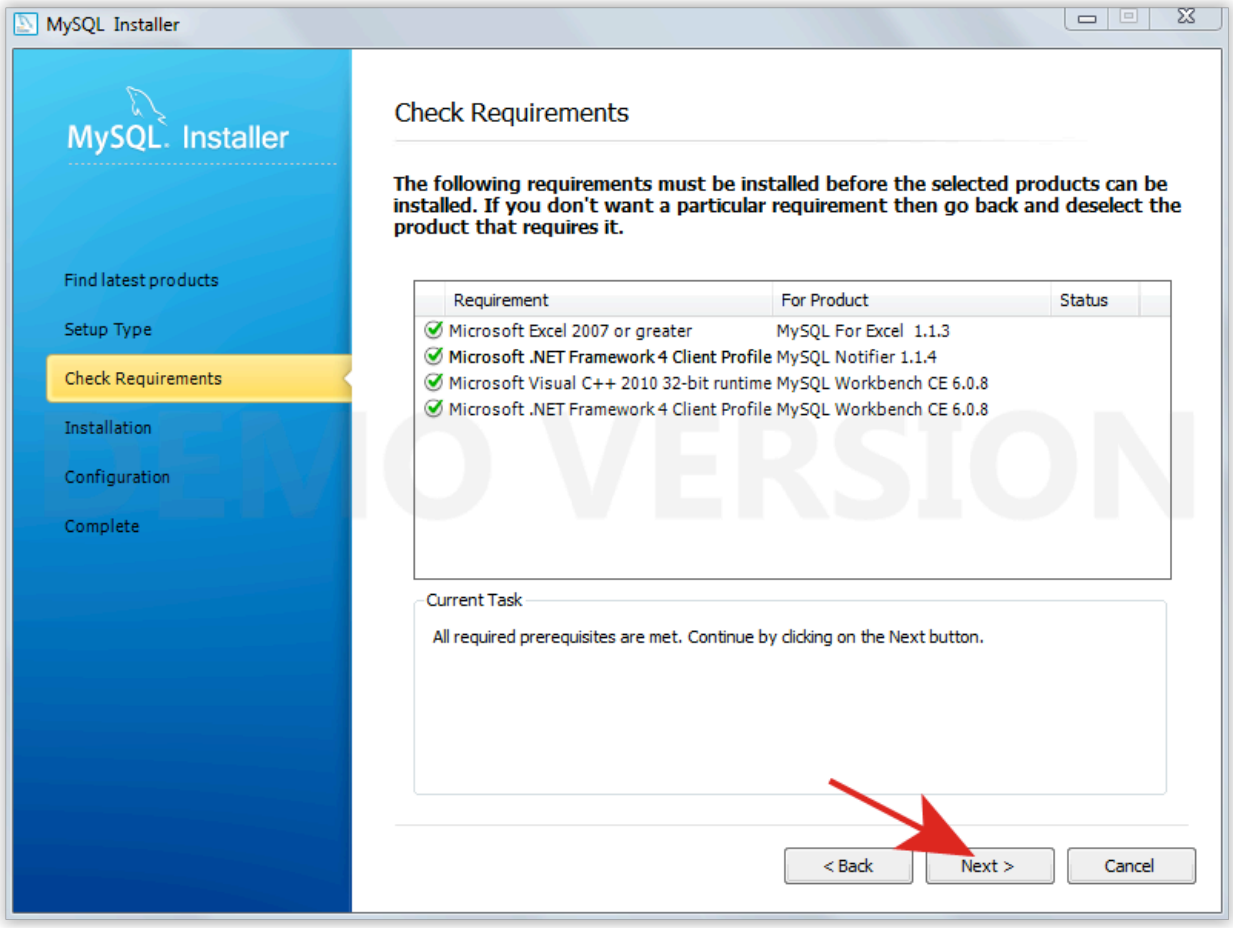
4. Click next.


MySQL Installer
MySQL. Installer
Find latest products
Setup Type
Check Requirements
Installation
Configuration
Complete
Choosing a Setup Type
Please select the Setup Type that suits your use case.
Developer Default
Installs all products needed for MySQL development purposes.
Server only
Installs only the MySQL Server product.
Client only
Installs only the MySQL Client products, without a server.
Full
Installs all included MySQL products and features.
Custom
Manually select the products that should be installed on the system.
Setup Type Description
Installs the MySQL Server and the tools required for MySQL application development. This is useful if you intend to develop applications for an existing server.
This Setup Type includes:
* MySQL Server
* MySQL Workbench
The GUI application to develop for and manage the server.
Installation Path:
C:\Program Files\MySQL\
Data Path:
C:\ProgramData\MySQL\MySQL Server 5.6\
< Back
Next >
Cancel

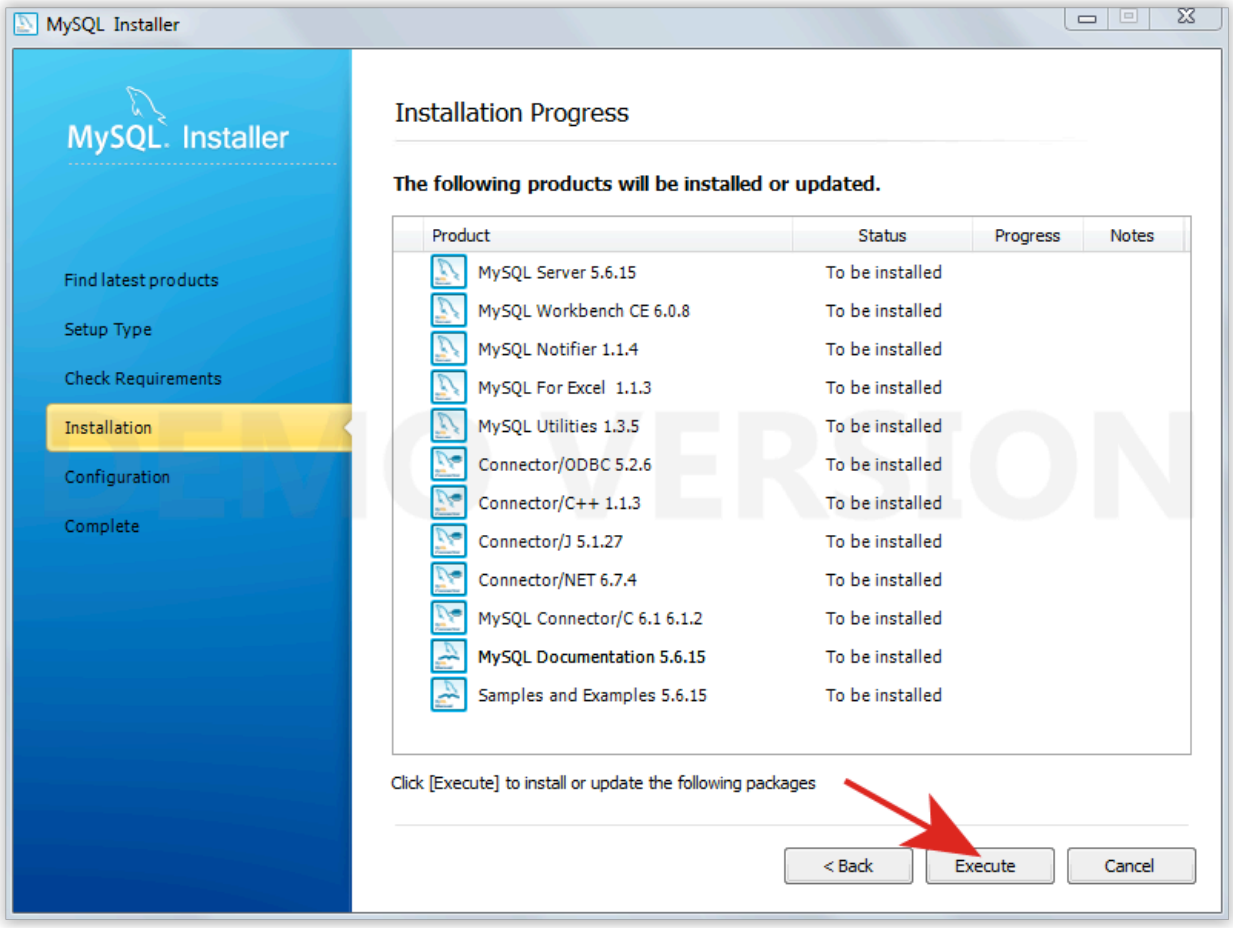
5. Click Execute.



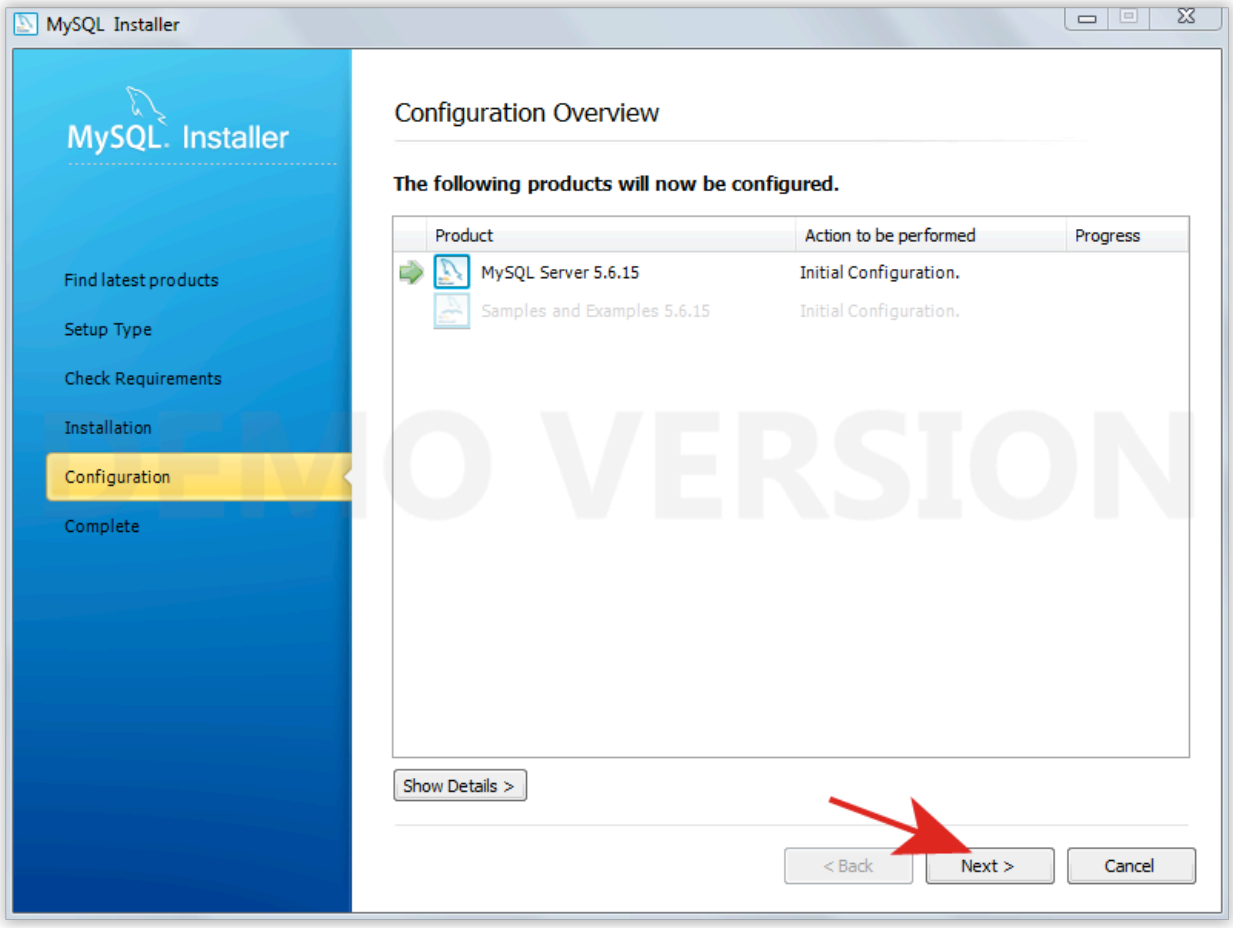
6. Click Next


MySQL Installer
MySQL. Installer
Find latest products
Setup Type
Check Requirements
Installation
Configuration
Complete
Check Requirements
The following requirements must be installed before the selected products can be installed. If you don't want a particular requirement then go back and deselect the product that requires it.
Requirement
For Product
Status
Microsoft Excel 2007 or greater
MySQL For Excel 1.1.3
Microsoft .NET Framework 4 Client Profile
MySQL Notifier 1.1.4
Microsoft Visual C++ 2010 32-bit runtime
MySQL Workbench CE 6.0.8
Microsoft .NET Framework 4 Client Profile
MySQL Workbench CE 6.0.8
Current Task
All required prerequisites are met. Continue by clicking on the Next button.
< Back
Next >
Cancel

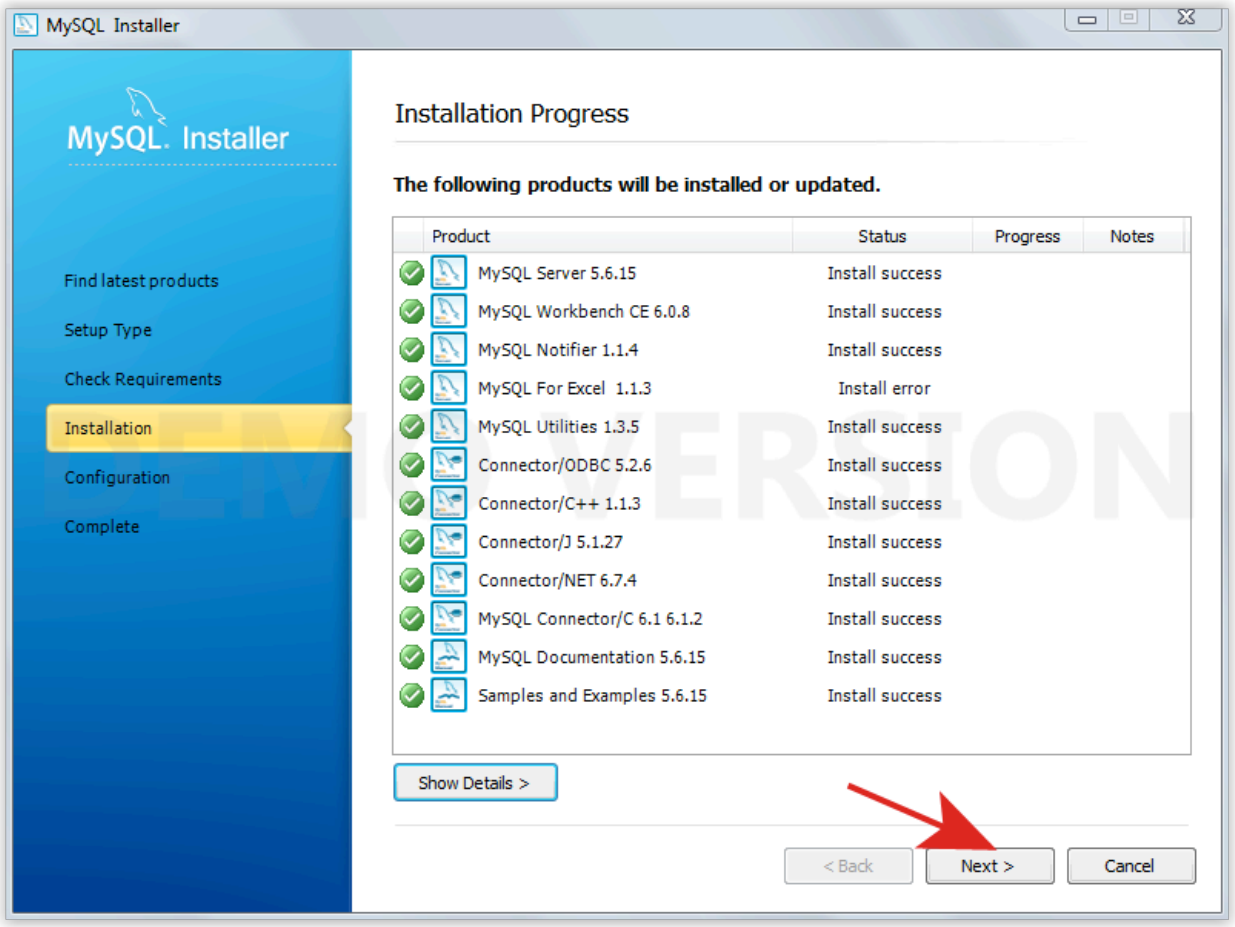
7. Click Execute.



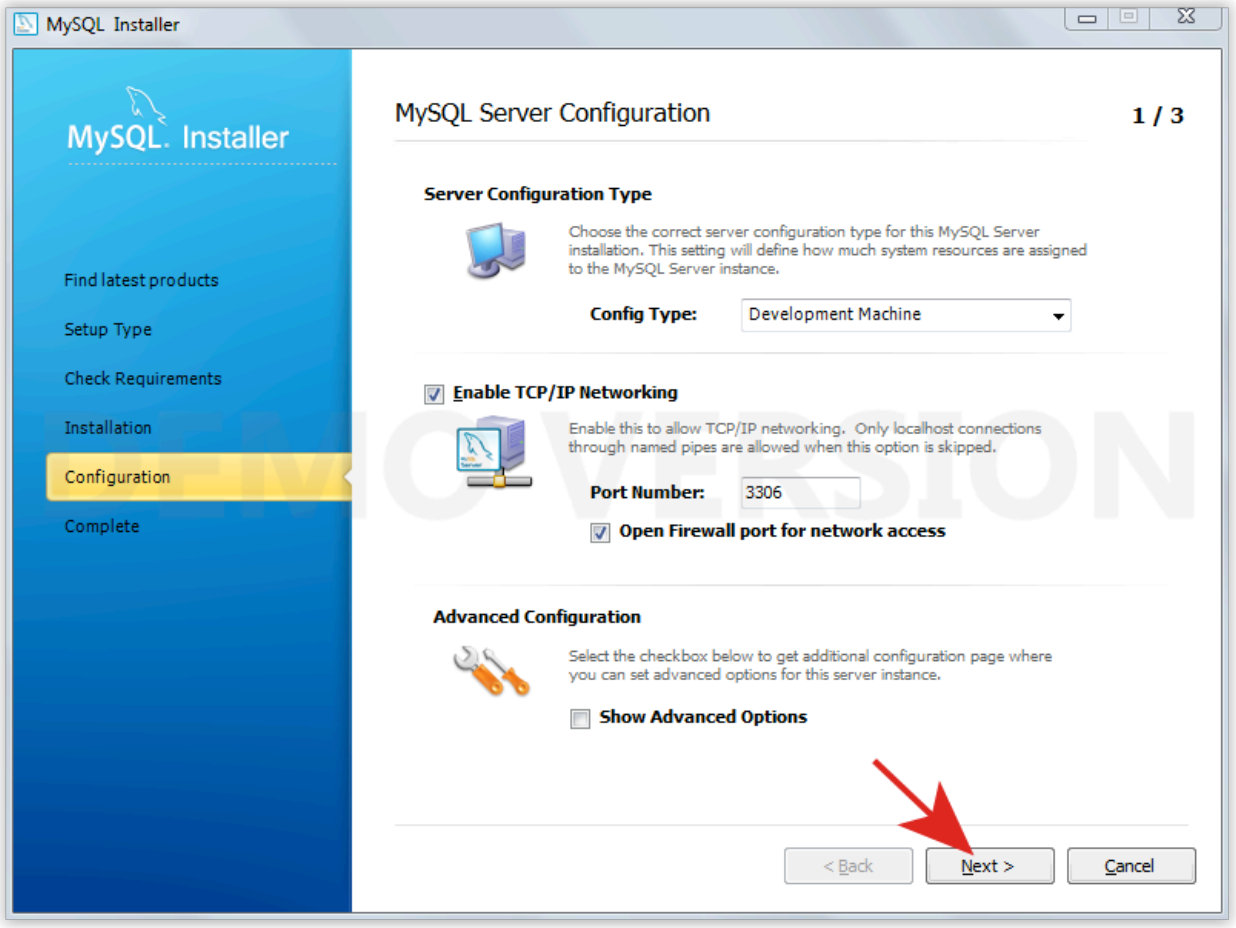
8. Click Next.


MySQL Installer
MySQL. Installer
Find latest products
Setup Type
Check Requirements
Installation
Configuration
Complete
Configuration Overview
The following products will now be configured.
Product
Action to be performed
Progress
MySQL Server 5.6.15
Initial Configuration.
Samples and Examples 5.6.15
Initial Configuration.
Show Details >
< Back
Next >
Cancel

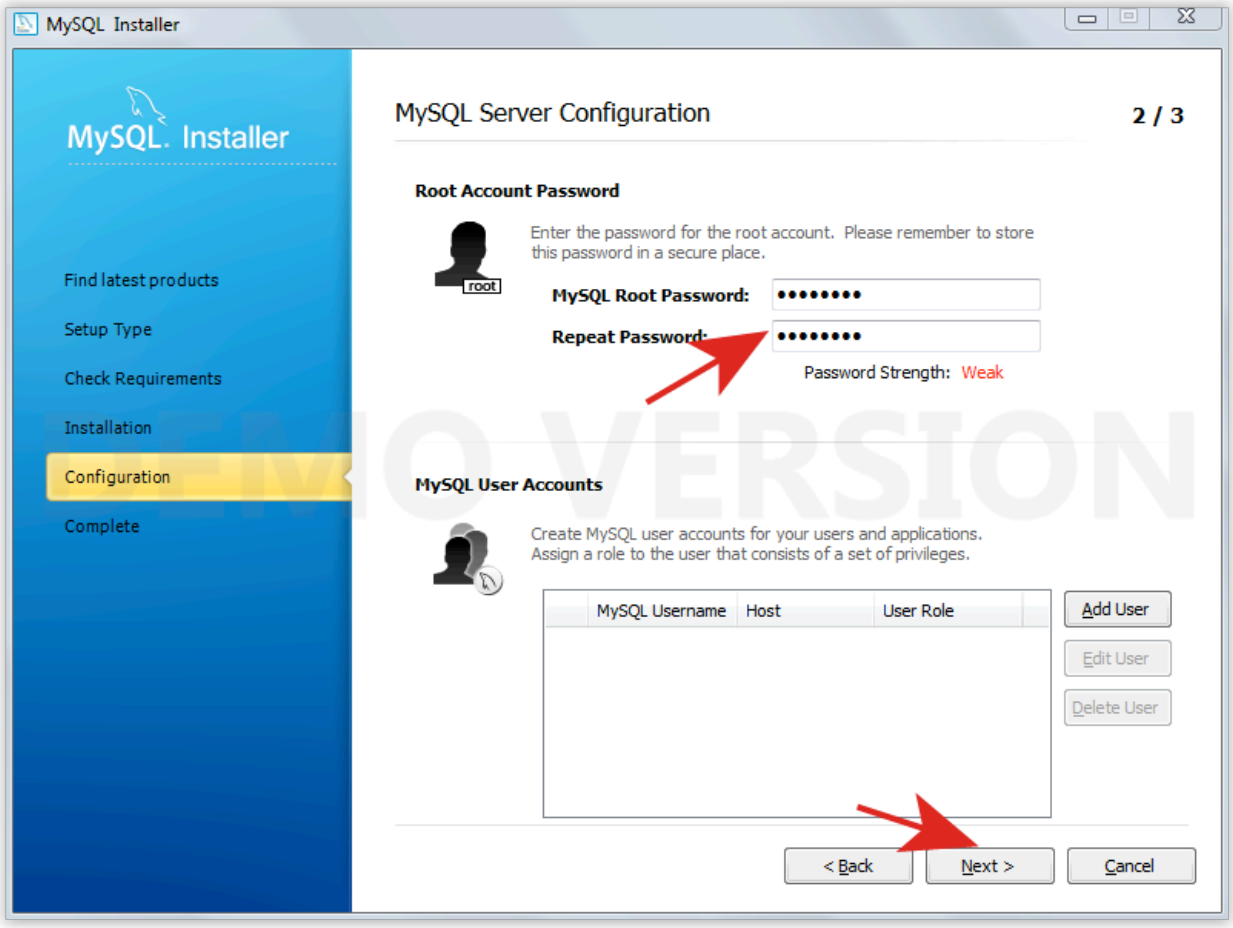
9. Click Next.


MySQL Installer
MySQL. Installer
Find latest products
Setup Type
Check Requirements
Installation
Configuration
Complete
Installation Progress
The following products will be installed or updated.
Product
Status
Progress
Notes
MySQL Server 5.6.15
Install success
MySQL Workbench CE 6.0.8
Install success
MySQL Notifier 1.1.4
Install success
MySQL For Excel 1.1.3
Install error
MySQL Utilities 1.3.5
Install success
Connector/ODBC 5.2.6
Install success
Connector/C++ 1.1.3
Install success
Connector/J 5.1.27
Install success
Connector/NET 6.7.4
Install success
MySQL Connector/C 6.1 6.1.2
Install success
MySQL Documentation 5.6.15
Install success
Samples and Examples 5.6.15
Install success
Show Details >
< Back
Next >
Cancel

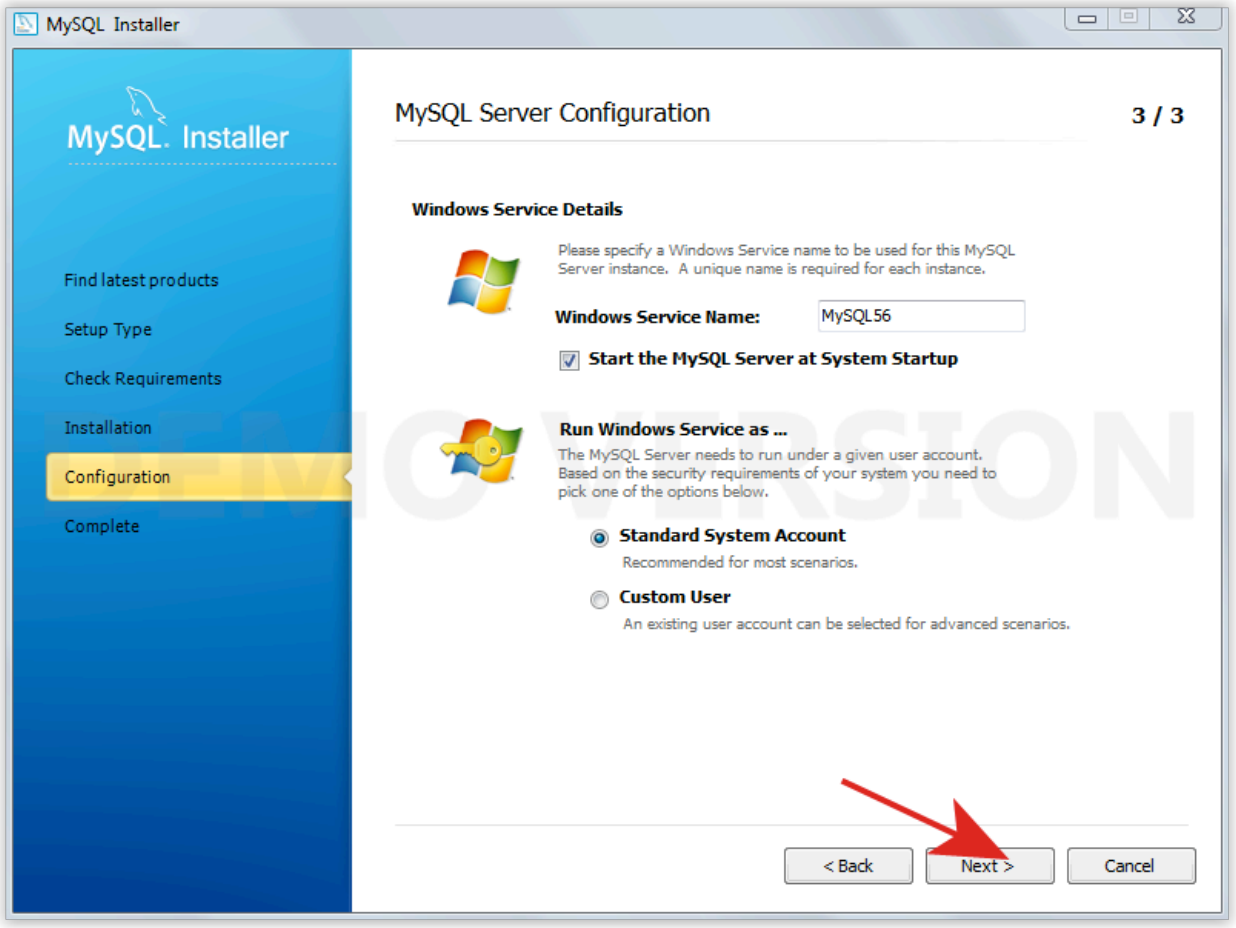
10.Click Next.



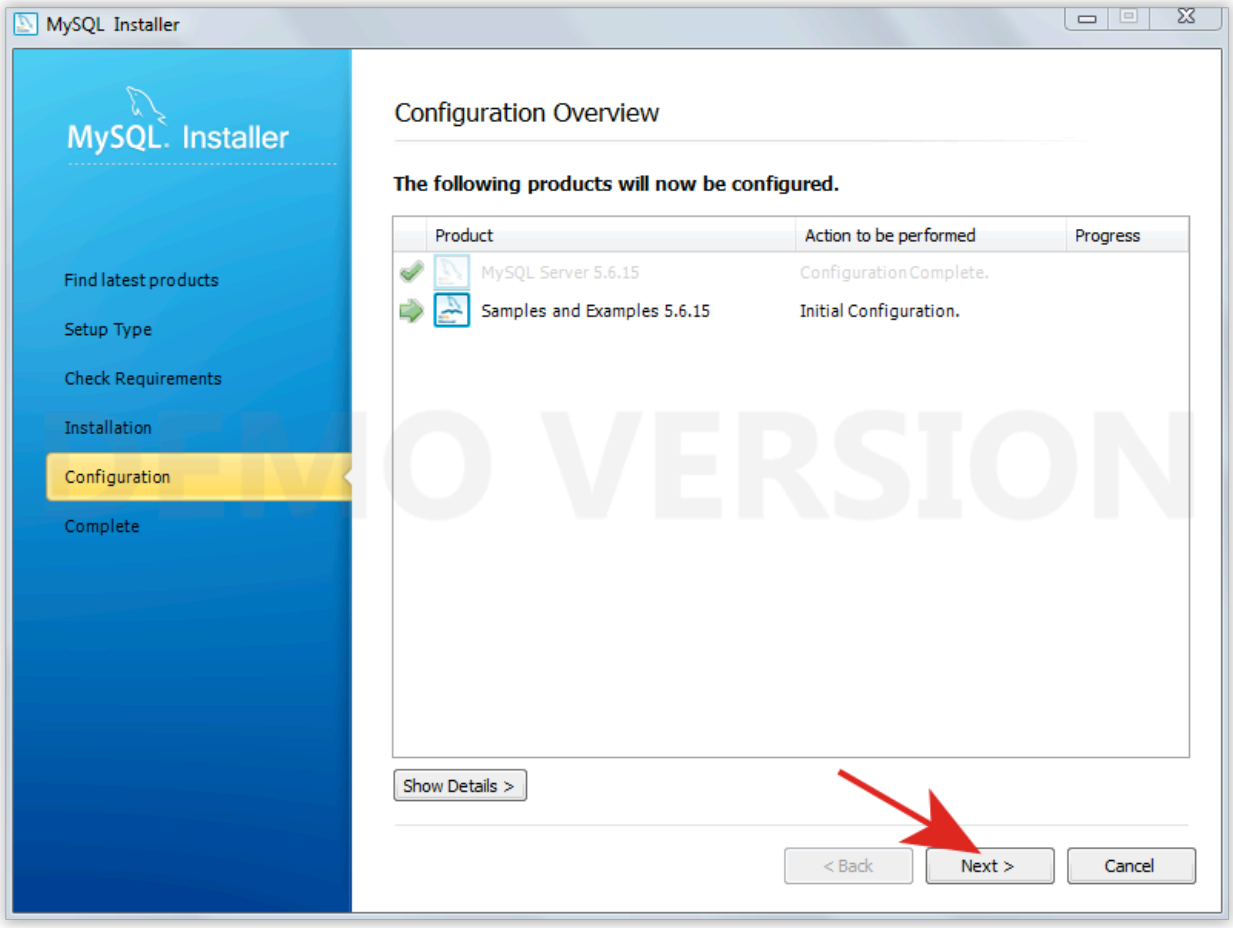
11.Set password for root user and click next.



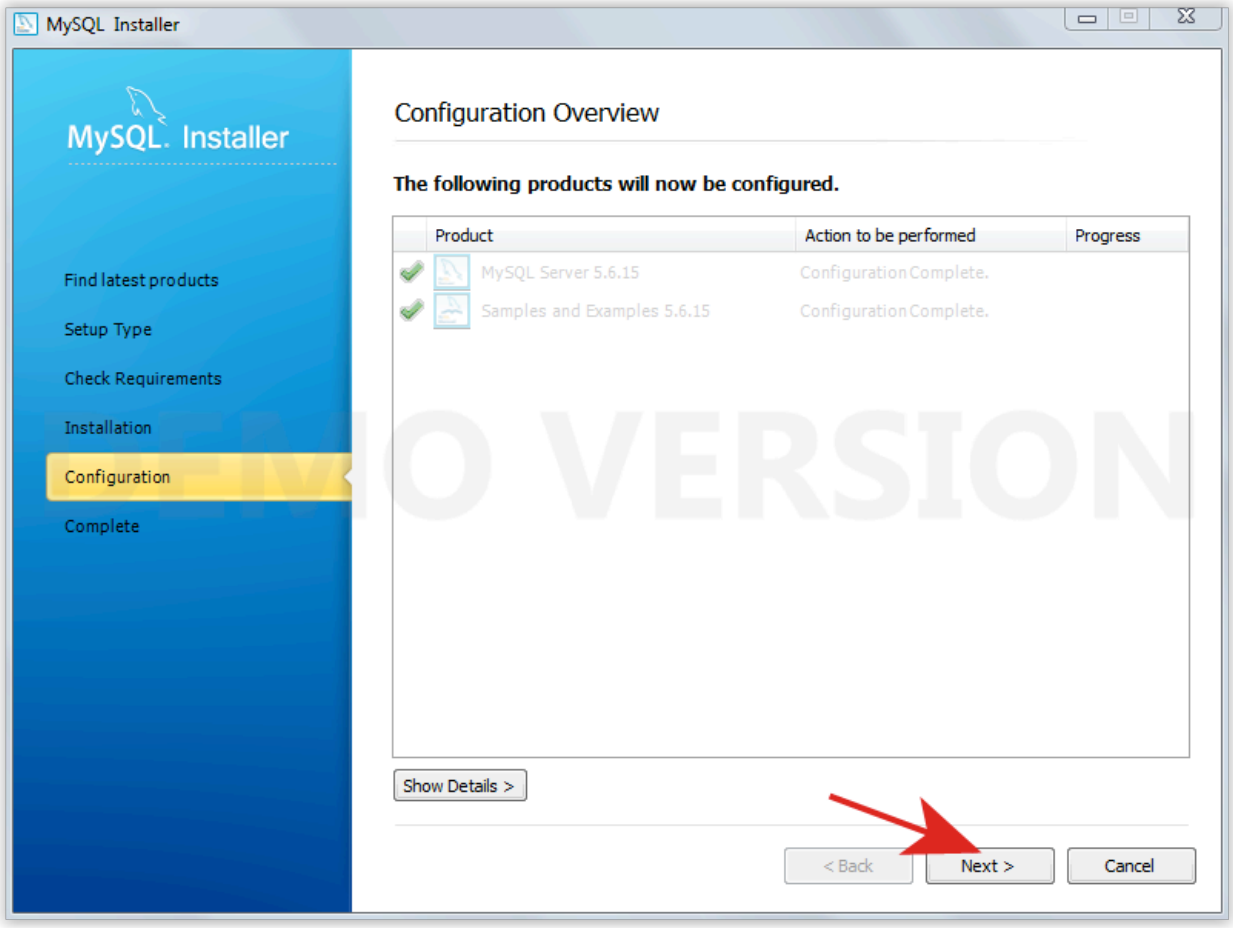
12.Click Next.



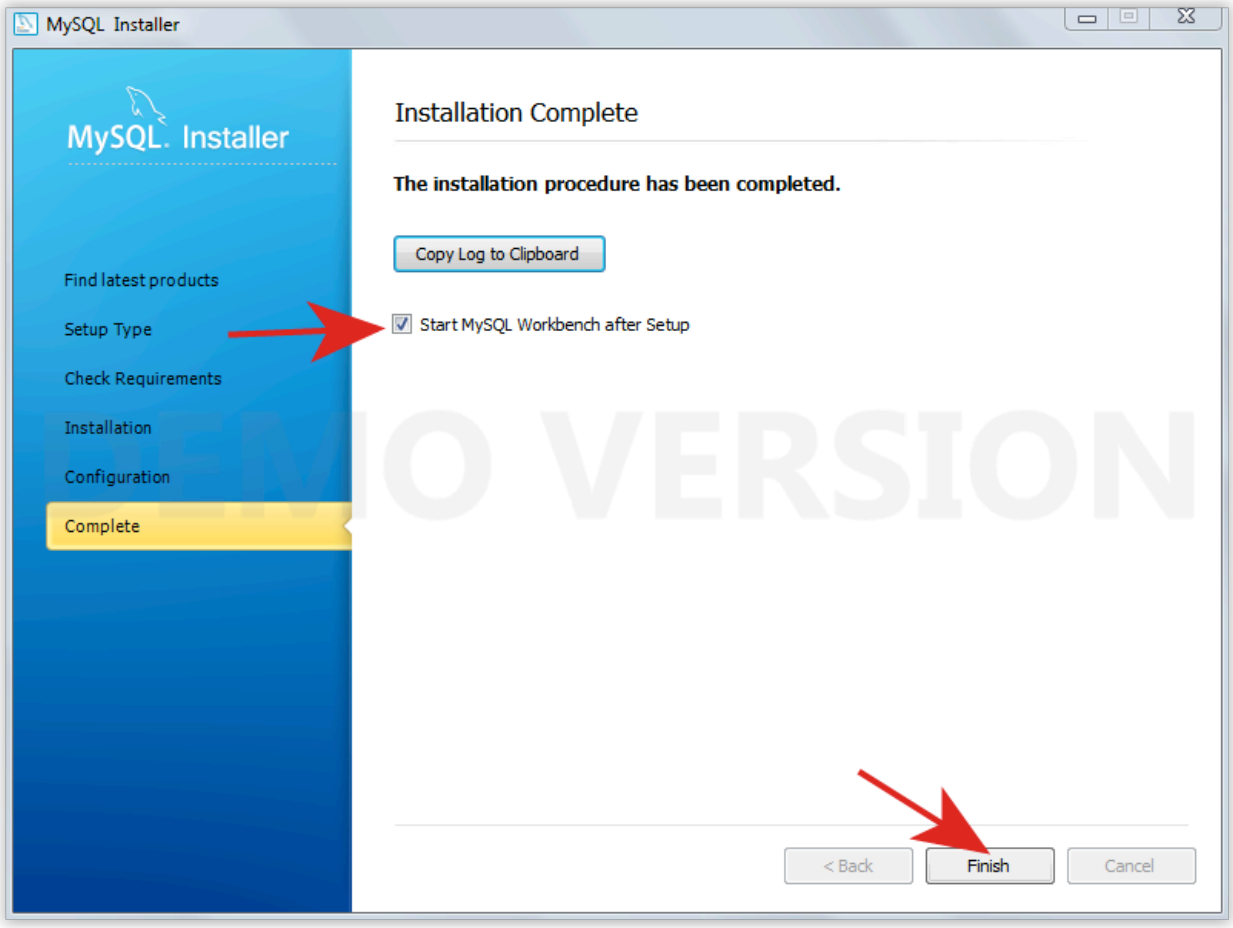
13.Click Next.



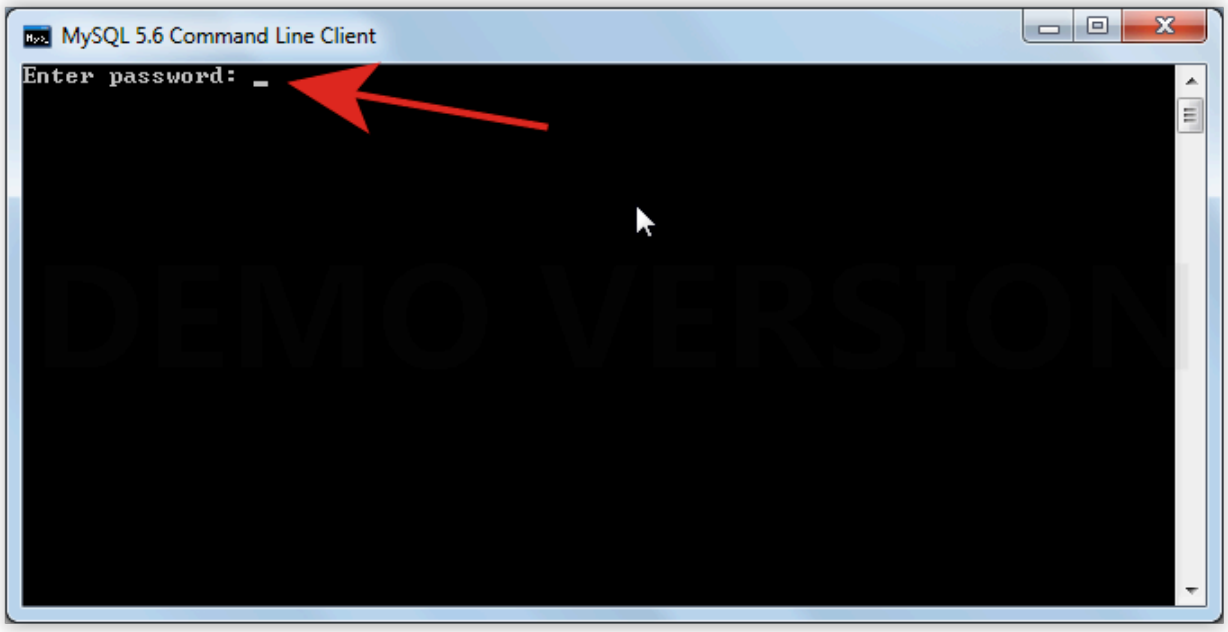
14.Click Next.

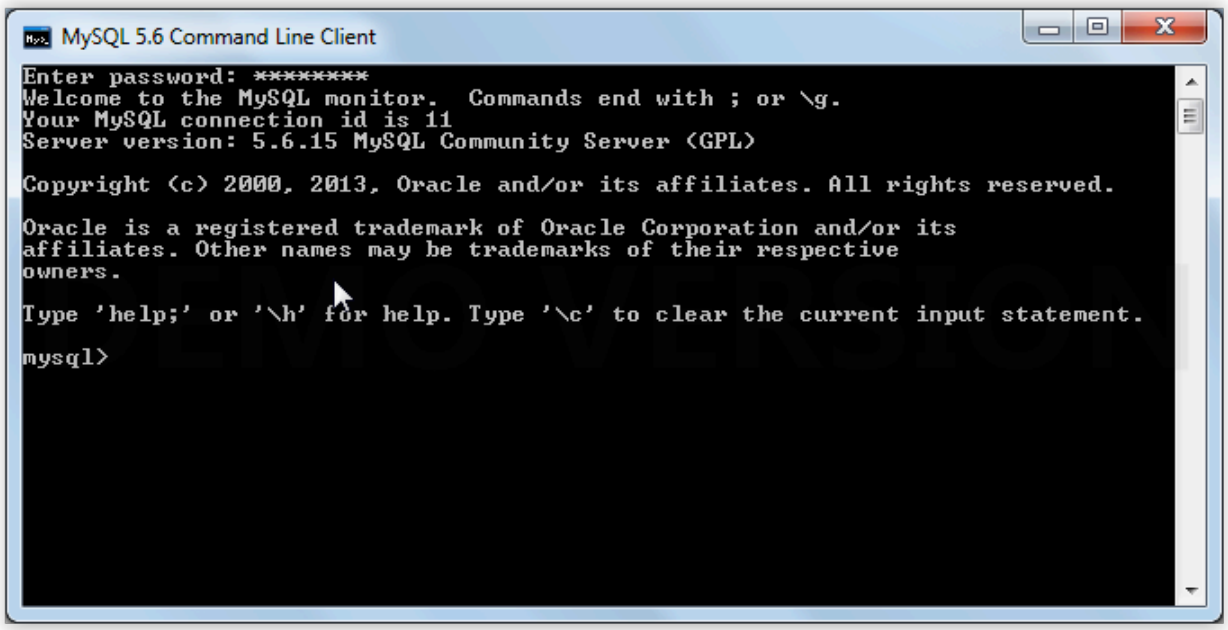


15.Click Next.



16. Open Mysql Command Line Client, give password and mysql command line will open. Now we can create tables, databases etc.

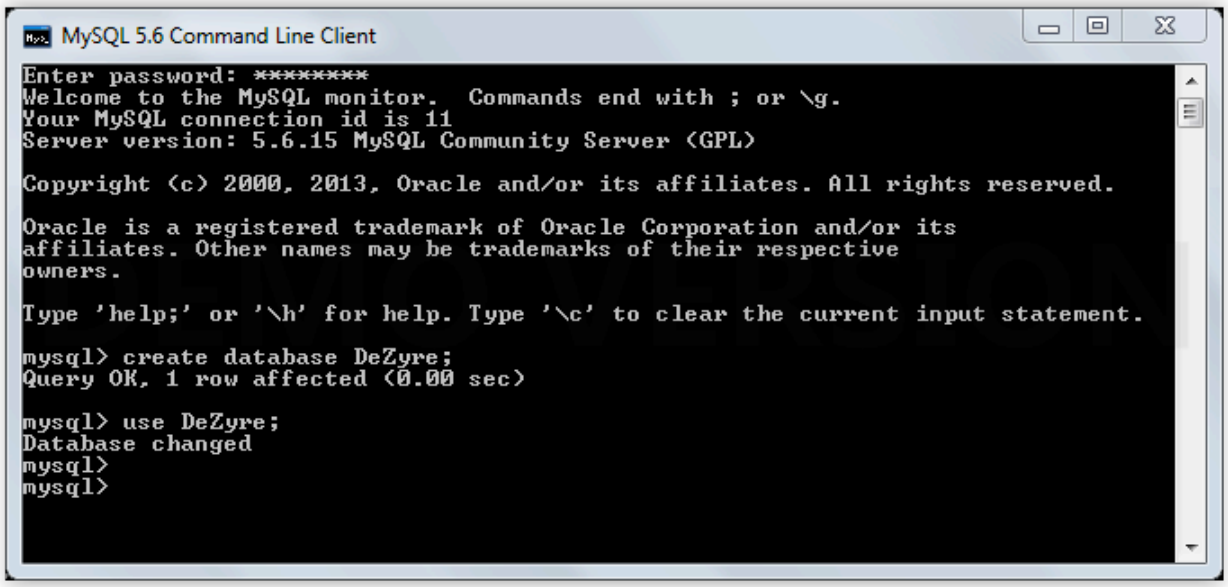


17. Create a new database and use new database using command given below.

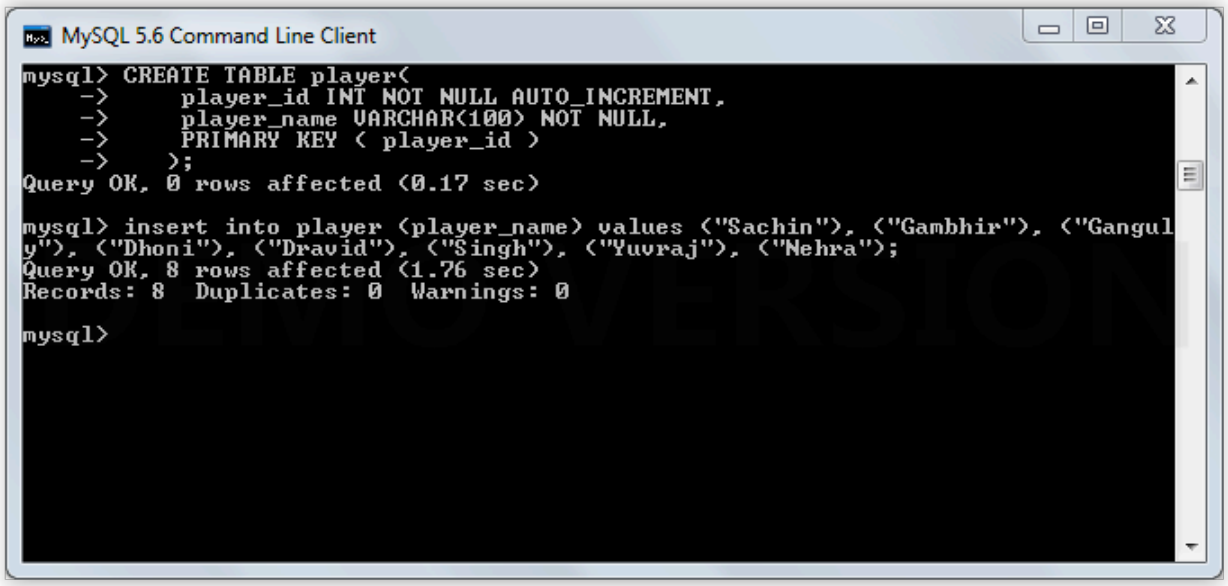
**create database DeZyre;**

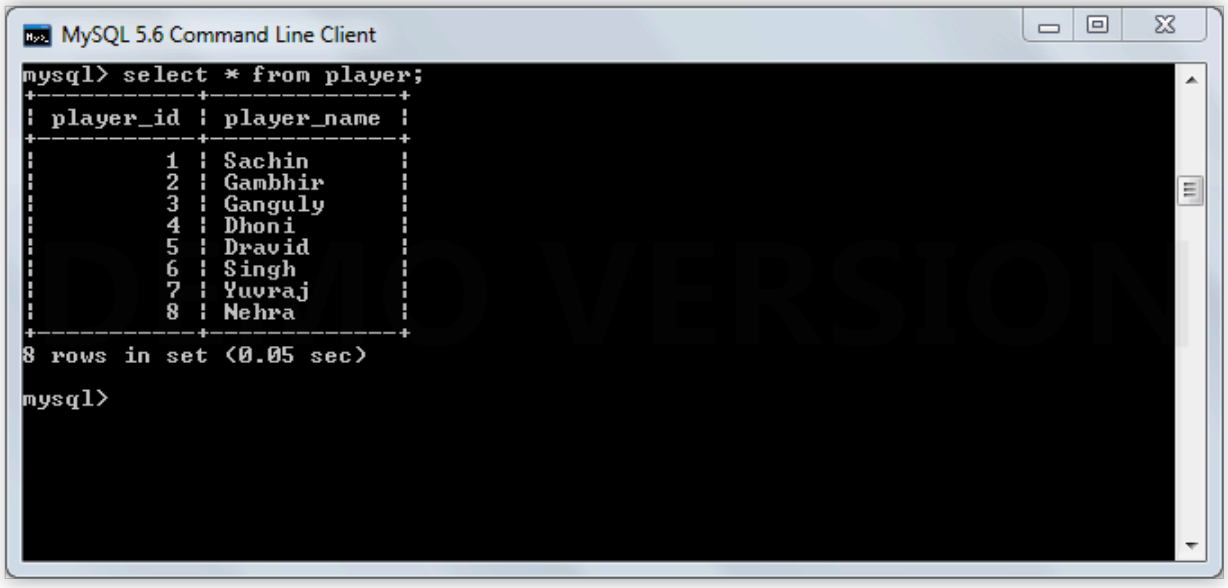
**use DeZyre;**



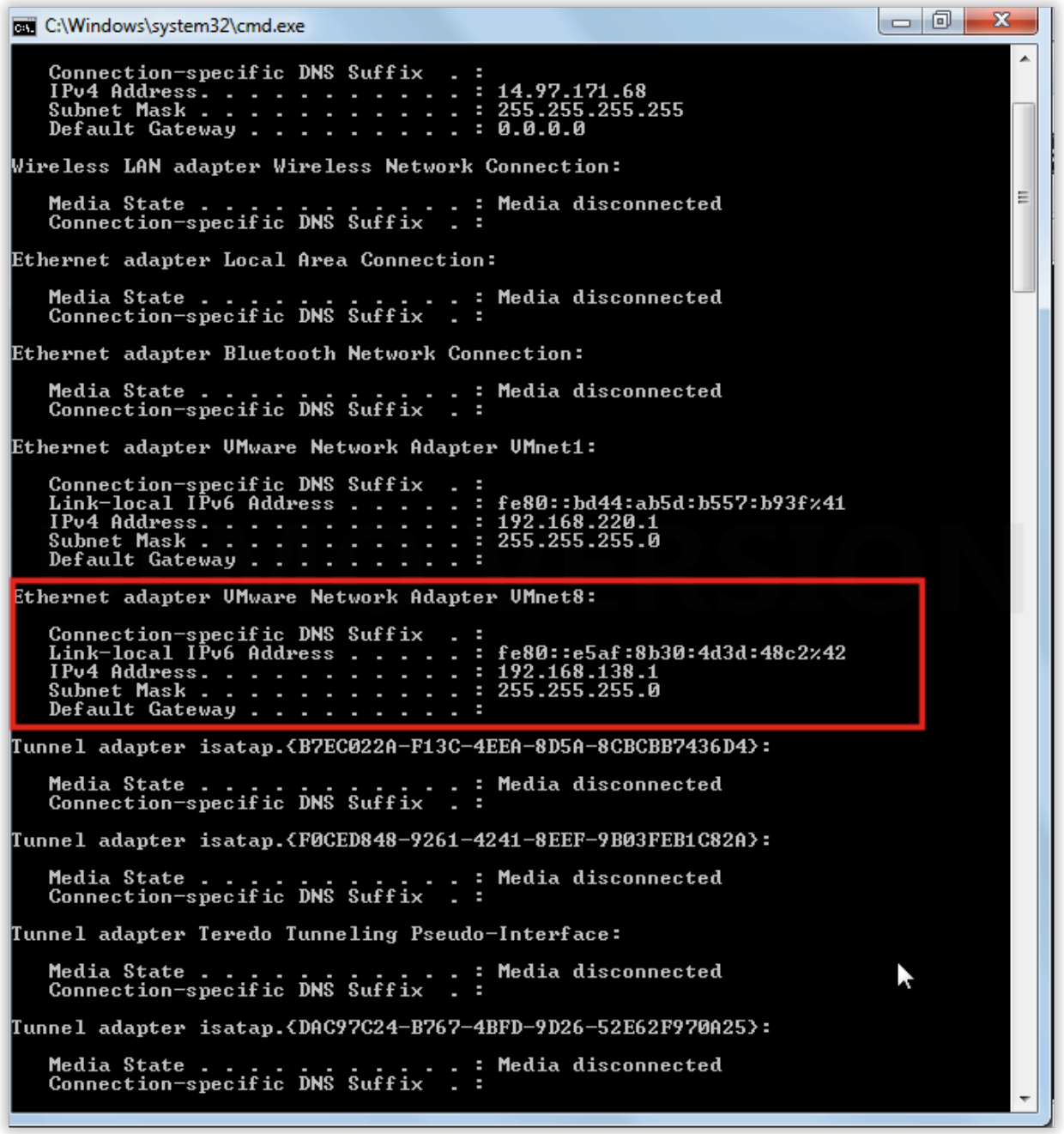

18.Create a table and insert records as given below.

```
CREATE TABLE player(
   player_id INT NOT NULL AUTO_INCREMENT,
   player_name VARCHAR(100) NOT NULL,
   PRIMARY KEY ( player_id )
  );
insert into player (player_name) values ("Sachin"), ("Dravid"), ("Dhoni"), ("Ganguly"),
("Yuvraj"), ("Nehra"), ("Singh");
```


MySQL 5.6 Command Line Client
mysql> CREATE TABLE player(
-> player_id INT NOT NULL AUTO_INCREMENT,
-> player_name VARCHAR(100) NOT NULL,
-> PRIMARY KEY ( player_id )
-> );
Query OK, 0 rows affected (0.17 sec)
mysql> insert into player (player_name) values ("Sachin"), ("Gambhir"), ("Gangul
y"), ("Dhoni"), ("Dravid"), ("Singh"), ("Yuvraj"), ("Nehra");
Query OK, 8 rows affected (1.76 sec)
Records: 8 Duplicates: 0 Warnings: 0
mysql>


MySQL 5.6 Command Line Client
mysql> select * from player;
+-----------+-------------+
| player_id | player_name |
+-----------+-------------+
| 1 | Sachin |
| 2 | Gambhir |
| 3 | Ganguly |
| 4 | Dhoni |
| 5 | Dravid |
| 6 | Singh |
| 7 | Yuvraj |
| 8 | Nehra |
+-----------+-------------+
8 rows in set (0.05 sec)
mysql>

19.Use “ipconfig” command to find out IP of mysql server.



**Importing Data from MySQL to HDFS**

20. Find out IP of the server where you want to run the Sqoop. Grant privilege to that server on MySQL server. Let's say IP of the server where Sqoop will be running is '192.168.138.128'. To grant privilege as shown above open MySQL client and run following command.

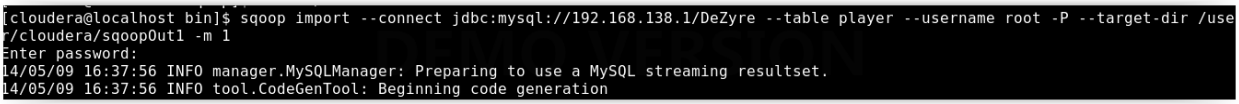
    To get IP of VM machine , Fire command **Ifconfig**

```
[cloudera@localhost ~]$ ifconfig
eth1      Link encap:Ethernet  HWaddr 00:0C:29:FC:15:DD
          inet addr:192.168.2.164  Bcast:192.168.2.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1903 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1680 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:2182743 (2.0 MiB)  TX bytes:136854 (133.6 KiB)

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:107411 errors:0 dropped:0 overruns:0 frame:0
          TX packets:107411 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:156069538 (148.8 MiB)  TX bytes:156069538 (148.8 MiB)

[cloudera@localhost ~]$
```

**grant all privileges on *.* to 'root'@'192.168.2.164' IDENTIFIED BY 'password' WITH GRANT OPTION**

21.For importing data from SQL to HDFS use below command on Cloudera VM.

```
sqoop import --connect jdbc:mysql://192.168.138.1/dezyre --table player --username root -P --target-dir /user/cloudera/sqoopOut1 -m 1
```



22. Once above command is complete you can see data in /user/cloudera/sqoopOut1 directory.

```
[cloudera@localhost ~]$
[cloudera@localhost ~]$ hadoop dfs -ls /user/cloudera/sqoopOut1
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

Found 3 items
-rw-r--r--   3 cloudera cloudera          0 2014-05-09 16:42 /user/cloudera/sqoopOut1/_SUCCESS
drwxr-xr-x   - cloudera cloudera          0 2014-05-09 16:39 /user/cloudera/sqoopOut1/_logs
-rw-r--r--   3 cloudera cloudera         71 2014-05-09 16:42 /user/cloudera/sqoopOut1/part-m-00000
[cloudera@localhost ~]$ hadoop dfs -cat /user/cloudera/sqoopOut1/part-m-00000
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

1,Sachin
2,Gambhir
3,Ganguly
4,Dhoni
5,Dravid
6,Singh
7,Yuvraj
8,Nehra
[cloudera@localhost ~]$
```

**Export data from HDFS to MySQL**

23. Run following command for creating a new directory and creating data.

```
hadoop dfs -mkdir /user/cloudera/sqoopOut2/

vi player.csv

hadoop dfs -put player.csv /user/cloudera/sqoopOut2/

hadoop dfs -cat /user/cloudera/sqoopOut2/player.csv
```

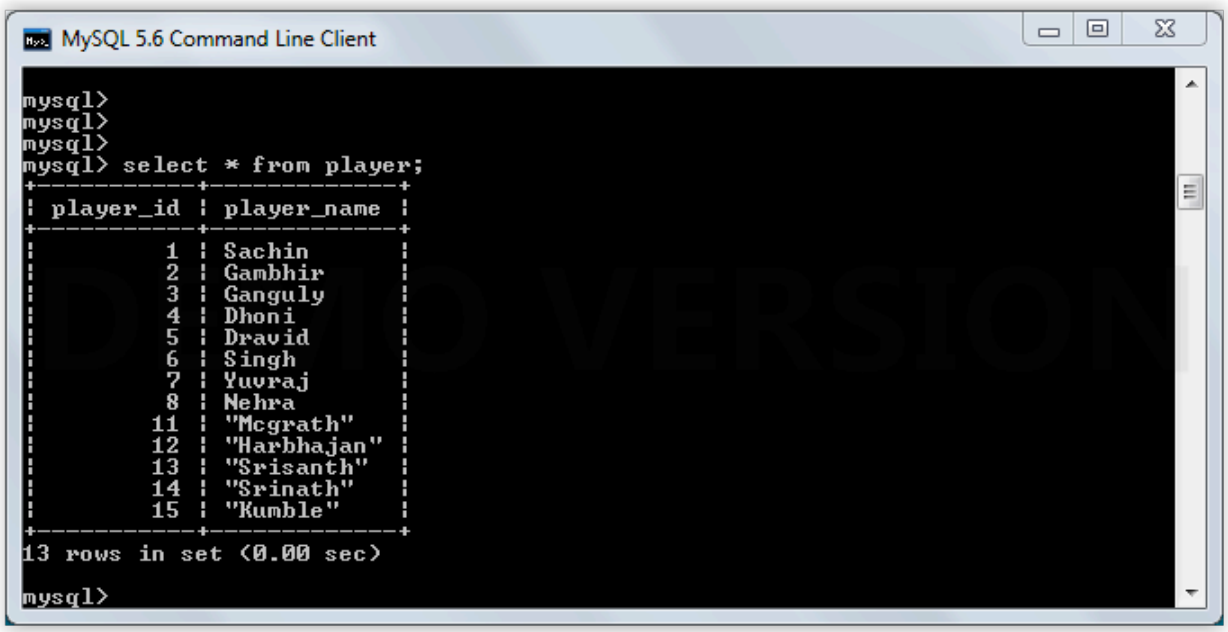
```
[cloudera@localhost ~]$ hadoop dfs -mkdir /user/cloudera/sqoopOut2/
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

[cloudera@localhost ~]$
[cloudera@localhost ~]$
[cloudera@localhost ~]$ vi player.csv
[cloudera@localhost ~]$
[cloudera@localhost ~]$
[cloudera@localhost ~]$ hadoop dfs -put player.csv /user/cloudera/sqoopOut2/
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

[cloudera@localhost ~]$ hadoop dfs -cat /user/cloudera/sqoopOut2/player.csv
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

11,"Mcgrath"
12,"Harbhajan"
13,"Srisanth"
14,"Srinath"
15,"Kumble"
[cloudera@localhost ~]$
```

24.Run following command for exporting data.

**sqoop export --connect jdbc:mysql://192.168.138.1/dezyre --table player --username root -P --export-dir /user/cloudera/sqoopOut2/ -m 1**

```
[cloudera@localhost ~]$ sqoop export --connect jdbc:mysql://192.168.138.1/DeZyre --table player --username root -P --export-dir  /user
/cloudera/sqoopOut2/ -m 1
Enter password:
14/05/09 16:52:36 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
14/05/09 16:52:36 INFO tool.CodeGenTool: Beginning code generation
14/05/09 16:52:38 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `player` AS t LIMIT 1
14/05/09 16:52:38 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `player` AS t LIMIT 1
14/05/09 16:52:38 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-0.20-mapreduce
14/05/09 16:52:38 INFO orm.CompilationManager: Found hadoop core jar at: /usr/lib/hadoop-0.20-mapreduce/hadoop-core.jar
Note: /tmp/sqoop-cloudera/compile/5c71b7218aae07c398f98405e3f966bf/player.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
14/05/09 16:52:50 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/5c71b7218aae07c398f98405e3f966bf/player.j
ar
14/05/09 16:52:50 INFO mapreduce.ExportJobBase: Beginning export of player
14/05/09 16:52:58 WARN mapred.JobClient: Use GenericOptionsParser for parsing the arguments. Applications should implement Tool for th
e same.
14/05/09 16:53:03 INFO input.FileInputFormat: Total input paths to process : 1
14/05/09 16:53:03 INFO input.FileInputFormat: Total input paths to process : 1
14/05/09 16:53:05 INFO mapred.JobClient: Running job: job_201405091237_0002
```

25.See from MySQL client if data exported.




**Import Data to Hive**

***sqoop import --connect jdbc:mysql://192.168.2.1/dezyre --username root --password password --table player --hive-import***