



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

Experiment No. 2
Use of Sqoop tool
Date of Performance: 24/7/23
Date of Submission: 31/7/23



AIM: To install SQOOP and execute basic commands of Hadoop ecosystem component Sqoop.

THEORY:

Installation and configuration of SQOOP

1) Download SQOOP from <https://sqoop.apache.org>

2) Unzip and Install SQOOP

After Downloading the SQOOP, we need to Unzip the sqoop-1.4.7.bin_hadoop-2.6.0.tar.gz file.

3) Create a folder and move the final extracted file in it.

4) Set up the environment variables

a. Set SQOOP_HOME

b. Set up path variable

5) Configure SQOOP

Basic SQOOP commands:

1. List Table

This command lists the particular table of the database in MYSQL server.

```
sqoop list - tables --connect jdbc:mysql://localhost/payment --username gatner
```

2. Target directory

This command import table in a specific directory in HDFS. -m denotes mapper argument. They have an integer value.

```
$ sqoop import --connect jdbc:mysql://localhost/inventory --username jony -table inventory --m 1 --target-dir/inv
```

3. sqoop-eval

This command runs quickly SQL queries of the respective database.

```
$ sqoop eval --connect --query "SQLQuery"
```



4. sqoop – version

This command displays version of the sqoop.

```
$ sqoop version      sqoop {revnumber}
```

5. sqoop-job

This command allows us to create a job, the parameters that are created can be invoked at any time. They take options like (–create,–delete,–show,–exit).

```
$ sqoop job --create --import --connect --table
```

6. code gen

This Sqoop command creates java class files which encapsulate the imported records. All the java files are recreated, and new versions of a class are generated. They generate code to interact with database records. Retrieves a list of all the columns and their datatypes.

```
$ sqoop codegen --connect -table
```

7. List Database

This Sqoop command lists have all the available database in the RDBMS server.

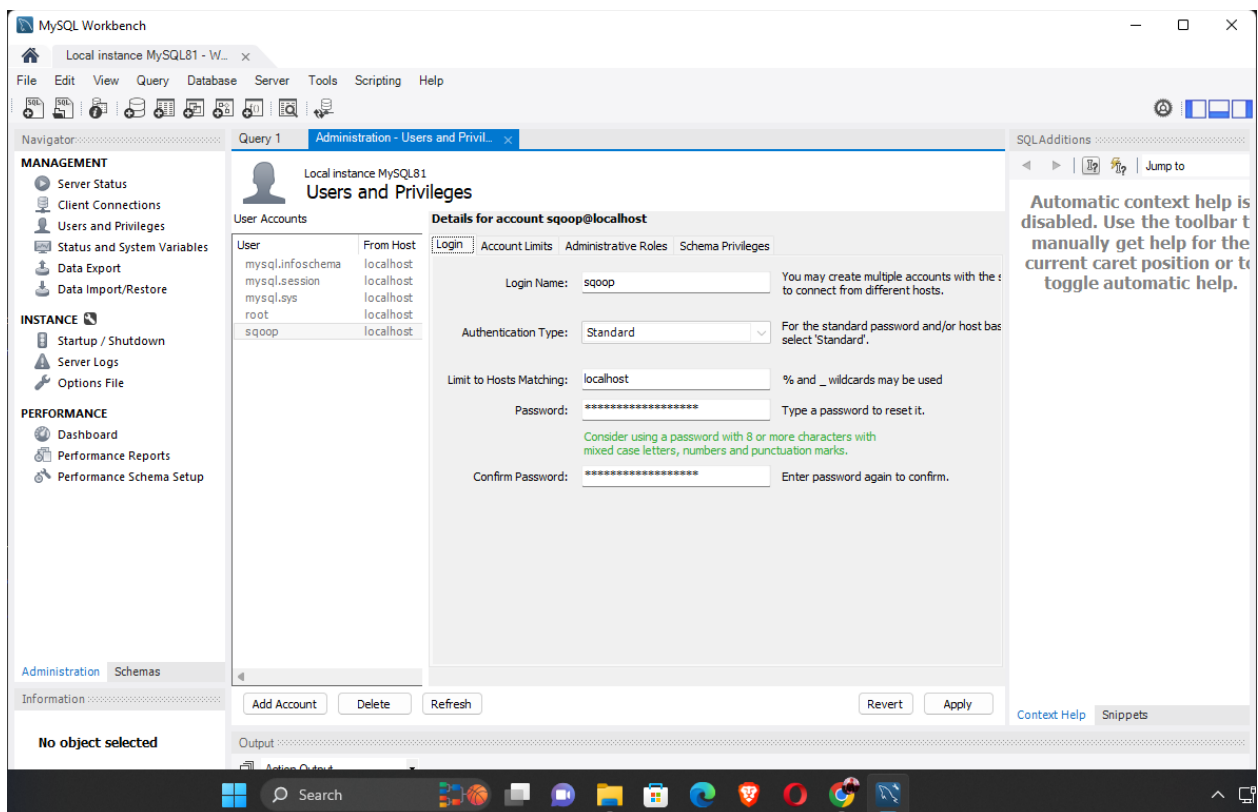
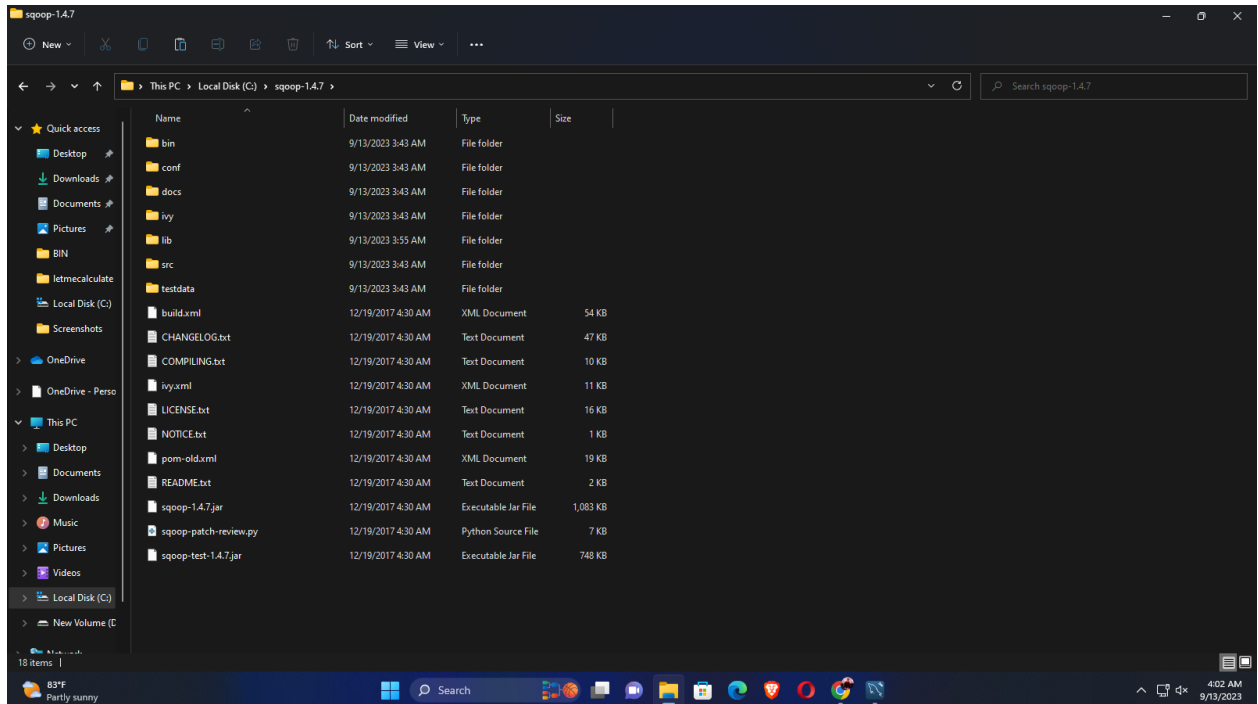
```
>$ sqoop list - database -- connect
```



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

OUTPUT:





Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

MySQL Workbench

Local instance MySQL81 - W...

File Edit View Query Database Server Tools Scripting Help

Navigator: Query 1 Administration - Users and Priv...

MANAGEMENT

- Server Status
- Client Connections
- Users and Privileges
- Status and System Variables
- Data Export
- Data Import/Restore

INSTANCE

- Startup / Shutdown
- Server Logs
- Options File

PERFORMANCE

- Dashboard
- Performance Reports
- Performance Schema Setup

Administration Schemas

Information

No object selected

Local instance MySQL81

Users and Privileges

User Accounts

User	From Host
mysql.infoschema	localhost
mysql.session	localhost
mysql.sys	localhost
root	localhost
sqoop	localhost

Details for account sqoop@localhost

Login Account Limits Administrative Roles Schema Privileges

Global Privileges

- ☒ ALTER
- ☒ ALTER ROUTINE
- ☒ CREATE
- ☒ CREATE ROUTINE
- ☒ CREATE TEMPORARY TABLES
- ☒ CREATE USER
- ☒ CREATE VIEW
- ☒ DELETE
- ☒ DROP
- ☒ EVENT
- ☐ EXECUTE
- ☐ FILE
- ☒ GRANT OPTION
- ☒ INDEX
- ☒ INSERT
- ☒ LOCK TABLES
- ☐ PROCESS
- ☐ REFERENCES

Revoke All Privileges

Add Account Delete Refresh Revert Apply

Output

Context Help Snippets

MySQL Workbench

Local instance MySQL81 - W...

File Edit View Query Database Server Tools Scripting Help

Navigator: Query 1 Administration - Users and Priv...

MANAGEMENT

- Server Status
- Client Connections
- Users and Privileges
- Status and System Variables
- Data Export
- Data Import/Restore

INSTANCE

- Startup / Shutdown
- Server Logs
- Options File

PERFORMANCE

- Dashboard
- Performance Reports
- Performance Schema Setup

Administration Schemas

Information

No object selected

Local instance MySQL81

Users and Privileges

User Accounts

User	From Host
mysql.infoschema	localhost
mysql.session	localhost
mysql.sys	localhost
root	localhost
sqoop	localhost

Details for account sqoop@localhost

Login Account Limits Administrative Roles Schema Privileges

Schema Privileges

Schema	Privileges
%_bigdata%	ALTER, ALTER ROUTINE, CREATE, CREATE ROUTINE, CREATE TEMPORARY TABLES, CREATE VIEW, DE

Revoke All Privileges Delete Entry Add Entry...

The user 'sqoop@localhost' will have the following access rights to schemas matching '%_bigdata%':

Object Rights

- ☒ SELECT
- ☒ INSERT
- ☒ UPDATE
- ☒ DELETE
- ☒ EXECUTE
- ☒ SHOW VIEW

DDL Rights

- ☒ CREATE
- ☒ ALTER
- ☒ REFERENCES
- ☒ INDEX
- ☒ CREATE VIEW
- ☒ CREATE ROUTINE
- ☒ ALTER ROUTINE
- ☒ EVENT
- ☒ DROP
- ☒ TRIGGER

Other Rights

- ☒ GRANT OPTION
- ☒ CREATE TEMPORARY TABLES
- ☒ LOCK TABLES

The REFERENCES privilege currently is unused.

Add Account Delete Refresh Revert Apply

Output

Action Output

#	Time	Action
---	------	--------

Message Duration / Fetch

Context Help Snippets



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

The screenshot shows the MySQL Workbench interface with the 'Users and Privileges' window open. The window displays a list of users on the left, including 'hive', 'mysql.infoschema', 'mysql.session', 'mysql.sys', 'root', and 'sqoop'. The main pane shows the details for the 'hive' user, including a table of privileges for the schema '%_bigdata%'. The privileges listed are ALTER, ALTER ROUTINE, CREATE, CREATE ROUTINE, CREATE TEMPORARY TABLES, and CREATE VIEW. The 'Object Rights' section shows a list of permissions with checkboxes for SELECT, INSERT, UPDATE, DELETE, EXECUTE, and SHOW VIEW. The 'DDL Rights' section shows a list of permissions with checkboxes for CREATE, ALTER, REFERENCES, INDEX, CREATE VIEW, CREATE ROUTINE, ALTER ROUTINE, EVENT, DROP, and TRIGGER. The 'Other Rights' section shows a list of permissions with checkboxes for GRANT OPTION, CREATE TEMPORARY TABLES, and LOCK TABLES. The 'Output' pane at the bottom shows the results of the queries, indicating that 0 rows were affected for both queries.

```
MySQL 8.1 Command Line Client
Enter password: ****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 16
Server version: 8.1.0 MySQL Community Server - GPL

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

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> grant all privileges on test_bigdata.* to 'sqoop'@'localhost';
Query OK, 0 rows affected (0.00 sec)

mysql> grant all privileges on test_bigdata.* to 'hive'@'localhost';
Query OK, 0 rows affected (0.00 sec)

mysql>
```



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

```
Command Prompt
Microsoft Windows [Version 10.0.22000.2295]
(c) Microsoft Corporation. All rights reserved.

C:\Users\admin>echo %SQOOP_HOME%
C:\sqoop-1.4.7

C:\Users\admin>sqoop list-databases --connect jdbc:mysql://localhost/ --username sqoop -P
Warning: HBASE_HOME and HBASE_VERSION not set.
Warning: HCAT_HOME not set
Warning: HCATALOG_HOME does not exist HCatalog imports will fail.
Please set HCATALOG_HOME to the root of your HCatalog installation.
Warning: ACCUMULO_HOME not set.
Warning: ZOOKEEPER_HOME not set.
Warning: HBASE_HOME does not exist HBase imports will fail.
Please set HBASE_HOME to the root of your HBase installation.
Warning: ACCUMULO_HOME does not exist Accumulo imports will fail.
Please set ACCUMULO_HOME to the root of your Accumulo installation.
Warning: ZOOKEEPER_HOME does not exist Accumulo imports will fail.
Please set ZOOKEEPER_HOME to the root of your Zookeeper installation.
2023-09-13 04:22:22,757 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
Enter password:
2023-09-13 04:22:26,809 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
Loading class `com.mysql.jdbc.Driver'. This is deprecated. The new driver class is `com.mysql.cj.jdbc.Driver'. The drive
r is automatically registered via the SPI and manual loading of the driver class is generally unnecessary.
mysql
information_schema
performance_schema
sys
C:\Users\admin>
```

```
Command Prompt
No such sqoop tool: list. See 'sqoop help'.

C:\Users\admin>sqoop list-tables --connect jdbc:mysql://localhost/ --username sqoop -P
Warning: HBASE_HOME and HBASE_VERSION not set.
Warning: HCAT_HOME not set
Warning: HCATALOG_HOME does not exist HCatalog imports will fail.
Please set HCATALOG_HOME to the root of your HCatalog installation.
Warning: ACCUMULO_HOME not set.
Warning: ZOOKEEPER_HOME not set.
Warning: HBASE_HOME does not exist HBase imports will fail.
Please set HBASE_HOME to the root of your HBase installation.
Warning: ACCUMULO_HOME does not exist Accumulo imports will fail.
Please set ACCUMULO_HOME to the root of your Accumulo installation.
Warning: ZOOKEEPER_HOME does not exist Accumulo imports will fail.
Please set ZOOKEEPER_HOME to the root of your Zookeeper installation.
2023-09-13 04:25:49,023 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
Enter password:
2023-09-13 04:25:53,985 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
Loading class `com.mysql.jdbc.Driver'. This is deprecated. The new driver class is `com.mysql.cj.jdbc.Driver'. The drive
r is automatically registered via the SPI and manual loading of the driver class is generally unnecessary.

C:\Users\admin>
```

CONCLUSION:

The experiment centered on the installation and utilization of Sqoop, a pivotal component within the Hadoop ecosystem. It successfully demonstrated Sqoop's capabilities, including connecting to various databases, importing and exporting data between Hadoop and relational databases, and performing data transformations during the process. Sqoop's parallel data transfer and seamless integration with Hadoop components were showcased. This experiment emphasized Sqoop's role in bridging the gap between Hadoop's distributed storage and relational databases, making it an indispensable tool for organizations managing diverse data sources. Proficiency in Sqoop equips data professionals with the essential skills to streamline data workflows and maximize the potential of big data projects.



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering
