



Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

---

Experiment No. 2
Use of Sqoop tool
Date of Performance: 24/07/2023
Date of Submission: 31/07/2023



# Vidyavardhini's College of Engineering & Technology

## Department of Computer Engineering

**Aim:** To install SQOOP and execute basic commands of Hadoop eco system 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.
  - Set up the environment variables
  - Set SQOOP\_HOME
- 4) 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 gartner
```

#### 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 SQL queries of the respective database.

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

#### 4. sqoop – version

This command displays a version of the sqoop.



# Vidyavardhini's College of Engineering & Technology

## Department of Computer Engineering

---

```
$ 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 data types.

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

### 7. List Database

This Sqoop command lists all the available databases in the RDBMS server.

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

Sqoop is a command-line interface application for transferring data between relational databases and Hadoop.

### Conclusion:

Sqoop is a crucial component of the Hadoop ecosystem, and this experiment explored its installation and basic command execution. Sqoop facilitates the easy transfer of data between relational databases and Hadoop, which makes it possible to integrate and analyze data effectively. The installation process included downloading Sqoop, setting up the necessary directories, and configuring environment variables. Additionally, some fundamental Sqoop commands were demonstrated, such as creating Java class files for data interaction, setting up tasks, running SQL queries, specifying destination folders, displaying tables and databases, and checking the version.