Vidyavardhini's College of Engineering & Technology
Department of Artificial Intelligence and Data Science

Experiment No. 2
Use of Sqoop tool
Date of Performance:
Date of Submission:

Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

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 imports tables 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"
```

Vidyavardhini's College of Engineering & Technology
Department of Artificial Intelligence and Data Science

4. sqoop – version

This command displays a 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 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
```

CONCLUSION:

We learned about multiple types of sqoop commands and their usage.