

Apache Hadoop:Sqoop Commands

Lab Exercise Workbook

SQOOP Lab exercises

Exercise 1: Sqoop import commands

Task 1: Login to mysql database on your VM.

```
$mysql -h <ip address> -u training -p  
Enter password -      training
```

Task 2: Browse the database

```
Show databases;  
Use moviesdb;  
Show tables;  
Execute command "select * from movies"
```

Task 3: Import the movie table in to HDFS

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-  
1.compute.internal:3306/ncnallan --username ncnallan --password  
Bigdata123 --table movies --driver com.mysql.jdbc.Driver
```

Task 4: Check the imported table in HDFS. Execute the following hdfs commands

```
hdfs dfs -ls /user/<username>  
hdfs dfs -ls /user/<username>/movies  
hdfs dfs -cat /user/<username>/movies/part-m-00000
```

Task 5: Importing without deleting target dir manually

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-  
1.compute.internal:3306/ncnallan --username ncnallan --password
```

SQOOP Lab exercises

```
Bigdata123 --table movies --driver com.mysql.jdbc.Driver --delete-target-dir
```

Task 6: Specifying a target directory

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --table movies --driver com.mysql.jdbc.Driver --target-dir mymovies
```

Task 7: Importing a table by --split-by and custom mappers

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --table movies --driver com.mysql.jdbc.Driver --split-by year --target-dir mymovies -m 6
```

Task 8: Specifying a warehouse dir to import multiple tables

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --table movies --driver com.mysql.jdbc.Driver --warehouse-dir mymovies
```

Task 9: Import all tables in to HDFS

For the **import-all-tables** tool to be useful, the following conditions must be met:

- Each table must have a single-column primary key.

SQOOP Lab exercises

- You must intend to import all columns of each table.
- You must not intend to use non-default splitting column, nor impose any conditions via a WHERE clause

```
sqoop import-all-tables --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --driver com.mysql.jdbc.Driver
```

Task 10: Using a where clause

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --table movies --driver com.mysql.jdbc.Driver --where id=4900
```

Task 11: Import as other formats

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --table movies --driver com.mysql.jdbc.Driver --as-parquetfile
```

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-1.compute.internal:3306/ncnallan --username ncnallan --password Bigdata123 --table movies --driver com.mysql.jdbc.Driver --as-avrodatafile
```

SQOOP Lab exercises

Exercise 2: Sqoop Export command

Task 1: Create a table in database to export

```
shell>mysql -h <ipaddress> -u training -p
mysql>CREATE TABLE employees_export (
  emp_no int(11) NOT NULL,
  birth_date date NOT NULL,
  first_name varchar(14) NOT NULL,
  last_name varchar(16) NOT NULL,
  gender enum('M','F') NOT NULL,
  hire_date date NOT NULL,
  PRIMARY KEY (emp_no));

mysql>exit;
```

Task 2: Copy data to be exported to HDFS

```
hdfs dfs -put employees-hive /user/<username>
sqoop export --connect jdbc:mysql://ip-10-1-1-204.ap-south-
1.compute.internal:3306/ncnallan --username ncnallan --password
Bigdata123 --driver com.mysql.jdbc.Driver --table employees_export
--export-dir /user/<username>/employees
```

Exercise 3: Sqoop Job

```
sqoop job --create mydailyjob -- import --connect jdbc:mysql://ip-10-
1-1-204.ap-south-1.compute.internal:3306/ncnallan --username
ncnallan --password Bigdata123 --table movies --driver
com.mysql.jdbc.Driver --delete-target-dir
```

```
sqoop job --list
```

SQOOP Lab exercises

```
sqoop job --show mydailyjob  
sqoop job --exec mydailyjob
```

Exercise 4: Sqoop import to hive

```
sqoop import --connect jdbc:mysql://ip-10-1-1-204.ap-south-  
1.compute.internal:3306/ncnallan --username ncnallan --password  
Bigdata123 --table movies --driver com.mysql.jdbc.Driver --hive-  
import --hive-database moviesdb
```