

ASSIGNMENT 15.5

Implement incremental import in sqoop

CREATION OF MYSQL TABLE

```
mysql> create table emp(id int not null auto_increment, name varchar(10), salary int, primary key(id));
Query OK, 0 rows affected (0.03 sec)
```

INSERTING VALUES INTO MYSQL TABLE

```
mysql> insert into emp values(1,'navya',20000),(2,'sriya',52000),(3,'adithi',56000);
Query OK, 3 rows affected (0.00 sec)
Records: 3 Duplicates: 0 Warnings: 0
```

DISPLAYING CONTENTS

```
mysql> select * from emp;
+----+-----+-----+
| id | name  | salary |
+----+-----+-----+
| 1  | navya | 20000  |
| 2  | sriya | 52000  |
| 3  | adithi | 56000  |
+----+-----+-----+
3 rows in set (0.00 sec)
```

SQOOP IMPORT

```
[training@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/movielens --username training --password training --table emp --m 1 --target-dir employee
17/04/25 20:08:17 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
17/04/25 20:08:18 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
17/04/25 20:08:18 INFO tool.CodeGenTool: Beginning code generation
17/04/25 20:08:18 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `emp` AS t LIMIT 1
17/04/25 20:08:19 INFO orm.CompilationManager: HADOOP_HOME is /usr/lib/hadoop
```

DISPLAY CONTENTS

```
[training@localhost ~]$ hadoop dfs -cat employee/part-m-00000
1,navya,20000
2,sriya,52000
3,adithi,56000
[training@localhost ~]$
```

ADDING MORE DATA

```
mysql> insert into emp values(4,'sriman',209000),(5,'keerthana',50000),(6,'shruthi',5000);
Query OK, 3 rows affected (0.00 sec)
Records: 3  Duplicates: 0  Warnings: 0
```

DISPLAYING CONTENTS

```
mysql> select * from emp
-> ;
+----+-----+-----+
| id | name   | salary |
+----+-----+-----+
| 1  | navya  | 20000  |
| 2  | sriya  | 52000  |
| 3  | adithi | 56000  |
| 4  | sriman | 209000 |
| 5  | keerthana | 50000 |
| 6  | shruthi | 5000  |
+----+-----+-----+
6 rows in set (0.00 sec)
```

INCREMENTAL IMPORT IN SQOOP

```
[training@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/movielens --username training --password training --table emp --m 1 --target-dir employee --incremental append --check-column id --last-value 3
17/04/25 20:20:37 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
```

```
17/04/25 20:20:54 INFO mapreduce.ImportJobBase: Retrieved 3 records.
17/04/25 20:20:54 INFO util.AppendUtils: Appending to directory employee
17/04/25 20:20:54 INFO util.AppendUtils: Using found partition 1
17/04/25 20:20:54 INFO tool.ImportTool: Incremental import complete! To run
another incremental import of all data following this import, supply the fol
lowing arguments:
17/04/25 20:20:54 INFO tool.ImportTool: --incremental append
17/04/25 20:20:54 INFO tool.ImportTool: --check-column id
17/04/25 20:20:54 INFO tool.ImportTool: --last-value 6
17/04/25 20:20:54 INFO tool.ImportTool: (Consider saving this with 'sqoop jo
> --create')
```

OUTPUT

```
[training@localhost ~]$ hadoop dfs -cat employee/part-m-00001
4,sriman,209000
5,keerthana,50000
6,shruthi,5000
```

```
[training@localhost ~]$ hadoop dfs -cat employee/par*
1,navya,20000
2,sriya,52000
3,adithi,56000
4,sriman,209000
5,keerthana,50000
6,shruthi,5000
[training@localhost ~]$
```
