

## Load data from AWS RDS to Hadoop

### <Command to import data from AWS RDS to Hadoop>

```
squop import \
--connect jdbc:mysql://upgradtest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
--table bookings \
--username student
--password STUDENT123 \
--target-dir /user/root/bookings_data \
-m1
```

### <Command to view the imported data>

```
hadoop fs -ls /user/root/bookings_data
hadoop fs -cat /user/root/bookings_data/part-m-00000 | wc -l
```

### <Screenshot of the data>

```
[root@ip-10-0-0-133 ~]# clear
[root@ip-10-0-0-133 ~]# squop import \
> --connect jdbc:mysql://upgradtest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
> --table bookings \
> --username student --password STUDENT123 \
> --target-dir /user/root/bookings_data \
> -m1 \
>
Warning: /opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/bin/../lib/squop/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
22/02/13 14:18:20 INFO squop.Squop: Running Squop version: 1.4.6-cdh5.15.1
22/02/13 14:18:20 WARN tool.BaseSquopTool: Setting your password on the command-line is insecure. Consider using -P instead.
22/02/13 14:18:21 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
22/02/13 14:18:21 INFO tool.CodeGenTool: Beginning code generation
22/02/13 14:18:21 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'bookings' AS t LIMIT 1
22/02/13 14:18:21 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'bookings' AS t LIMIT 1
22/02/13 14:18:21 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/11b/hadoop-mapreduce
Note: /tmp/squop-root/compile/5bbaab85996c9daae14b22df85fa2568/bookings.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
22/02/13 14:18:24 INFO orm.CompilationManager: Writing jar file: /tmp/squop-root/compile/5bbaab85996c9daae14b22df85fa2568/bookings.jar

22/02/13 14:18:49 INFO mapreduce.ImportJobBase: Transferred 161.7949 KB in 24.0578 seconds (6.7253 KB/sec)
22/02/13 14:18:49 INFO mapreduce.ImportJobBase: Retrieved 1000 records.
[root@ip-10-0-0-133 ~]# hadoop fs -ls /user/root/bookings_data
Found 2 items
-rw-r--r-- 3 root supergroup 0 2022-02-13 14:18 /user/root/bookings_data/_SUCCESS
-rw-r--r-- 3 root supergroup 165678 2022-02-13 14:18 /user/root/bookings_data/part-m-00000
[root@ip-10-0-0-133 ~]# hadoop fs -cat /user/root/bookings_data/part-m-00000 | wc -l
1000
```