

## Load data from AWS RDS to Hadoop

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

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

```
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=165678
23/10/23 21:00:58 INFO mapreduce.ImportJobBase: Transferred 161.7949 KB in 19.3177 seconds (8.3755 KB/sec)
23/10/23 21:00:58 INFO mapreduce.ImportJobBase: Retrieved 1000 records.
[hadoop@ip-172-31-32-203 ~]$
```

### Command to view the imported data

```
hadoop fs -ls /home/hadoop/bookings_data
```

### Screenshot of the data

```
[hadoop@ip-172-31-32-203 ~]$ hadoop fs -ls /home/hadoop/
Found 6 items
drwxr-xr-x - hadoop hadoop 0 2023-10-23 21:02 /home/hadoop/booking_data_csv
drwxr-xr-x - hadoop hadoop 0 2023-10-23 21:00 /home/hadoop/bookings_data
drwxr-xr-x - hadoop hadoop 0 2023-10-23 20:55 /home/hadoop/clickstream_checkpoint
drwxr-xr-x - hadoop hadoop 0 2023-10-23 20:55 /home/hadoop/clickstream_data
drwxr-xr-x - hadoop hadoop 0 2023-10-23 20:58 /home/hadoop/clickstream_data_flatten
drwxr-xr-x - hadoop hadoop 0 2023-10-23 21:02 /home/hadoop/datetime_aggregation
[hadoop@ip-172-31-32-203 ~]$
```

Later the data is converted to csv for easier usage and is stored in booking\_data\_csv:

```
[hadoop@ip-172-31-32-203 ~]$ hadoop fs -ls /home/hadoop/booking_data_csv
Found 2 items
-rw-r--r-- 1 hadoop hadoop 0 2023-10-23 21:02 /home/hadoop/booking_data_csv/_SUCCESS
-rw-r--r-- 1 hadoop hadoop 171678 2023-10-23 21:02 /home/hadoop/booking_data_csv/part-00000-d00df237-2c73-4fcf-80d3-3a71d5676377-c000.csv
[hadoop@ip-172-31-32-203 ~]$ hadoop fs -cat /home/hadoop/booking_data_csv/part-00000-d00df237-2c73-4fcf-80d3-3a71d5676377-c000.csv | wc -l
1000
[hadoop@ip-172-31-32-203 ~]$ hadoop fs -cat /home/hadoop/booking_data_csv/part-00000-d00df237-2c73-4fcf-80d3-3a71d5676377-c000.csv | head
BK8968087150,51811359,15055660,2.2.14,Android,-49.4319655,103.917851,-58.8043875,146.477367,2020-06-23T19:33:10.000Z,2020-06-06T09:02:10.000Z,534,83,INR,black,054-38-4479,4,3,3
BK629851904,31663218,60872180,3.4.1,iOS,-83.5408405,175.80085,86.20705,128.367238,2020-05-23T12:22:04.000Z,2020-08-09T19:02:56.000Z,126,67,INR,lime,796-39-6801,3,2,4
BK1797410350,86869399,94276051,4.1.36,iOS,-67.8930645,55.234128,-51.1079,-31.07475,2020-05-19T14:14:32.000Z,2020-08-23T18:38:39.000Z,297,63,INR,olive,748-73-1579,1,3,3
BK5788246325,58230837,45457227,2.4.27,Android,13.707887,113.499943,54.3812915,-18.437751,2020-03-24T01:30:15.000Z,2020-05-19T11:16:45.000Z,932,32,INR,white,558-80-6346,3,2,2
BK8342703255,84232510,86494681,4.1.34,Android,-6.091461,-114.649789,22.8449505,70.137827,2020-08-03T19:10:52.000Z,2020-03-24T08:25:40.000Z,260,7,INR,blue,068-72-1637,3,3,3
BK6015582453,11981042,35862658,2.4.39,iOS,-18.910034,-70.193103,-10.182921,173.877213,2020-07-17T05:33:48.000Z,2020-04-30T04:54:27.000Z,907,53,INR,purple,102-10-5639,3,2,3
BK4529355854,60071878,78022360,2.1.9,iOS,1.215274,-56.014903,35.152876,104.324905,2020-01-02T01:48:40.000Z,2020-02-16T04:28:55.000Z,547,17,INR,teal,866-83-4349,2,3,4
BK9720088219,14327312,94427067,3.1.2,Android,-55.4822225,173.362256,65.0121265,51.390751,2020-04-10T15:11:07.000Z,2020-01-20T21:17:42.000Z,259,33,INR,maroon,572-73-6526,3,3,2
BK7157532607,46407210,43160003,1.3.4,Android,46.005843,-16.826146,7.6126015,-156.428577,2020-06-09T05:56:31.000Z,2020-03-19T01:53:16.000Z,787,21,INR,olive,667-23-5880,2,2,3
BK5014871433,65861573,64708618,1.3.28,iOS,-29.565326,64.843709,84.068109,-49.820835,2020-08-14T20:43:42.000Z,2020-06-03T09:39:59.000Z,586,5,INR,fuchsia,255-52-5654,5,5,1
```