

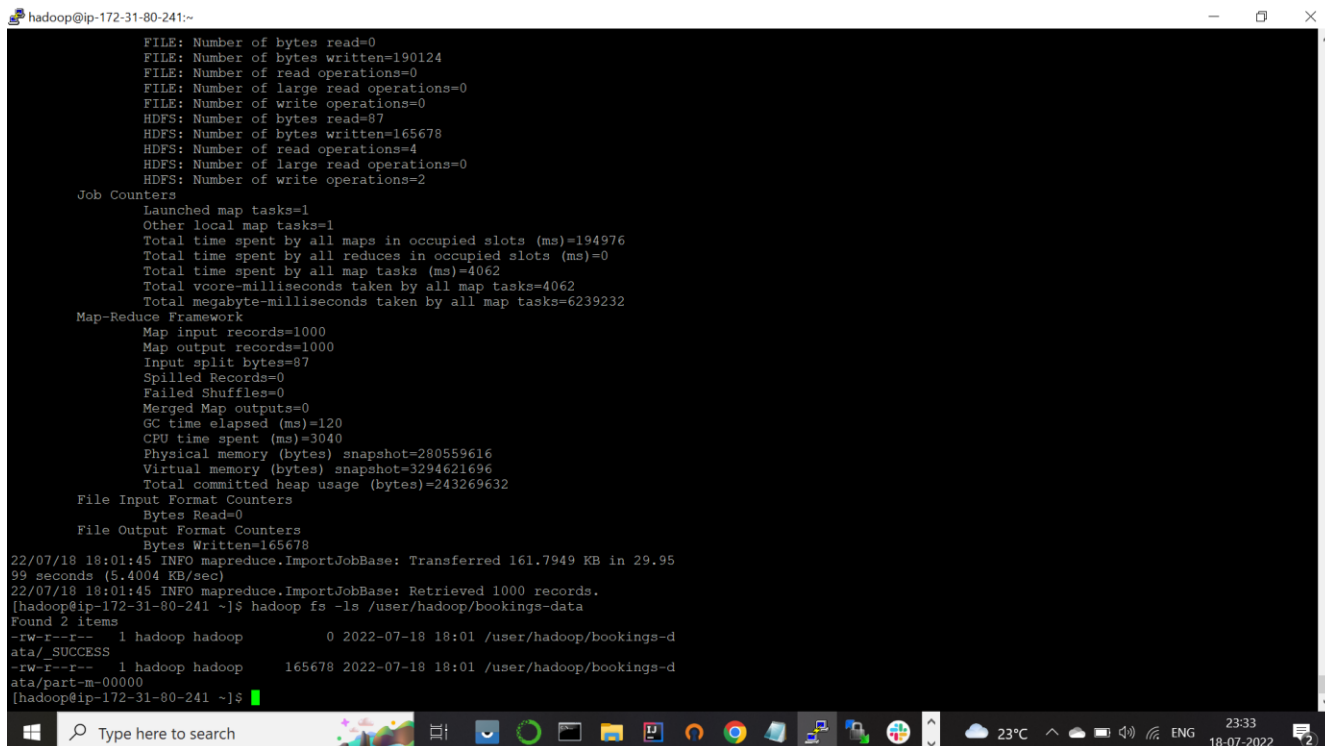
## Load data from AWS RDS to Hadoop

1. Command to check if the directory is not present already:

```
hadoop fs -rm -r /user/hadoop/bookings-data
```

2. Command to import data from AWS RDS to Hadoop:

```
sqoop import \  
--connect \  
jdbc:mysql://upgraddetest.cyaielc9bmnf.us- \  
east-1.rds.amazonaws.com/testdatabase \  
--table bookings \  
--username student --password STUDENT123 \  
--target-dir /user/hadoop/bookings-data \  
--m 1
```



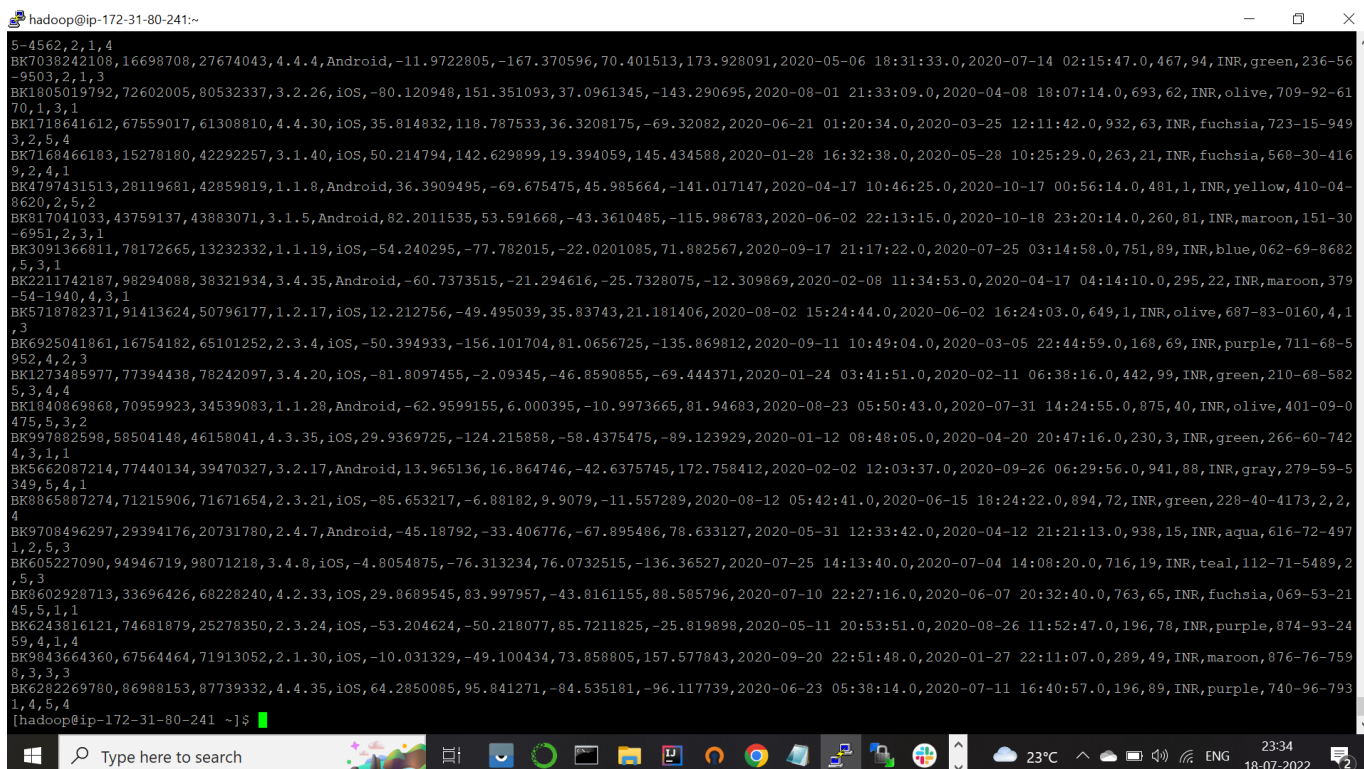
```
hadoop@ip-172-31-80-241:~$ sqoop import \  
--connect jdbc:mysql://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \  
--table bookings \  
--username student --password STUDENT123 \  
--target-dir /user/hadoop/bookings-data \  
--m 1  
  
FILE: Number of bytes read=0  
FILE: Number of bytes written=190124  
FILE: Number of read operations=0  
FILE: Number of large read operations=0  
FILE: Number of write operations=0  
HDFS: Number of bytes read=87  
HDFS: Number of bytes written=165678  
HDFS: Number of read operations=4  
HDFS: Number of large read operations=0  
HDFS: Number of write operations=2  
  
Job Counters  
  Launched map tasks=1  
  Other local map tasks=1  
  Total time spent by all maps in occupied slots (ms)=194976  
  Total time spent by all reduces in occupied slots (ms)=0  
  Total time spent by all map tasks (ms)=4062  
  Total vcore-milliseconds taken by all map tasks=4062  
  Total megabyte-milliseconds taken by all map tasks=6239232  
  
Map-Reduce Framework  
  Map input records=1000  
  Map output records=1000  
  Input split bytes=87  
  Spilled Records=0  
  Failed Shuffles=0  
  Merged Map outputs=0  
  GC time elapsed (ms)=120  
  CPU time spent (ms)=3040  
  Physical memory (bytes) snapshot=280559616  
  Virtual memory (bytes) snapshot=3294621696  
  Total committed heap usage (bytes)=243269632  
  
File Input Format Counters  
  Bytes Read=0  
  File Output Format Counters  
    Bytes Written=165678  
22/07/18 18:01:45 INFO mapreduce.ImportJobBase: Transferred 161.7949 KB in 29.95  
39 seconds (5.4004 KB/sec)  
22/07/18 18:01:45 INFO mapreduce.ImportJobBase: Retrieved 1000 records.  
[hadoop@ip-172-31-80-241 ~]$ hadoop fs -ls /user/hadoop/bookings-data  
Found 2 items  
-rw-r--r-- 1 hadoop hadoop 0 2022-07-18 18:01 /user/hadoop/bookings-d  
ata/_SUCCESS  
-rw-r--r-- 1 hadoop hadoop 165678 2022-07-18 18:01 /user/hadoop/bookings-d  
ata/part-m-00000  
[hadoop@ip-172-31-80-241 ~]$
```

### 3. Command to view the imported data

```
hadoop fs -ls /user/hadoop/bookings-data
```

```
hadoop fs -cat /user/hadoop/bookings-data/part-m-00000
```

### Screenshot of the data:



```
hadoop@ip-172-31-80-241:~$ hadoop fs -cat /user/hadoop/bookings-data/part-m-00000
5-4562,2,1,4
BK7038242108,16698708,27674043,4.4.4,Android,-11.9722805,-167.370596,70.401513,173.928091,2020-05-06 18:31:33.0,2020-07-14 02:15:47.0,467,94,INR,green,236-56
-9503,2,1,3
BK1805019792,72602005,80532337,3.2.26,iOS,-80.120948,151.351093,37.0961345,-143.290695,2020-08-01 21:33:09.0,2020-04-08 18:07:14.0,693,62,INR,olive,709-92-61
70,1,3,1
BK1718641612,67559017,61308810,4.4.30,iOS,35.814832,118.787533,36.3208175,-69.32082,2020-06-21 01:20:34.0,2020-03-25 12:11:42.0,932,63,INR,fuchsia,723-15-949
3,2,5,4
BK7168466183,15278180,42292257,3.1.40,iOS,50.214794,142.629899,19.394059,145.434588,2020-01-28 16:32:38.0,2020-05-28 10:25:29.0,263,21,INR,fuchsia,568-30-416
9,2,4,1
BK4797431513,28119681,42859819,1.1.8,Android,36.3909495,-69.675475,45.985664,-141.017147,2020-04-17 10:46:25.0,2020-10-17 00:56:14.0,481,1,INR,yellow,410-04-
8620,2,5,2
BK817041033,43759137,43883071,3.1.5,Android,82.2011535,53.591668,-43.3610485,-115.986783,2020-06-02 22:13:15.0,2020-10-18 23:20:14.0,260,81,INR,maroon,151-30
-6951,2,3,1
BK3091366811,78172665,13232332,1.1.19,iOS,-54.240295,-77.782015,-22.0201085,71.882567,2020-09-17 21:17:22.0,2020-07-25 03:14:58.0,751,89,INR,blue,062-69-8682
,5,3,1
BK2211742187,98294088,38321934,3.4.35,Android,-60.7373515,-21.294616,-25.7328075,-12.309869,2020-02-08 11:34:53.0,2020-04-17 04:14:10.0,295,22,INR,maroon,379
-54-1940,4,3,1
BK5718782371,91413624,50796177,1.2.17,iOS,12.212756,-49.495039,35.83743,21.181406,2020-08-02 15:24:44.0,2020-06-02 16:24:03.0,649,1,INR,olive,687-83-0160,4,1
,3
BK6925041861,16754182,65101252,2.3.4,iOS,-50.394933,-156.101704,81.0656725,-135.869812,2020-09-11 10:49:04.0,2020-03-05 22:44:59.0,168,69,INR,purple,711-68-5
952,4,2,3
BK1273485977,77394438,78242097,3.4.20,iOS,-81.8097455,-2.09345,-46.8590855,-69.444371,2020-01-24 03:41:51.0,2020-02-11 06:38:16.0,442,99,INR,green,210-68-582
5,3,4,4
BK1840869868,70959923,34539083,1.1.28,Android,-62.9599155,6.000395,-10.9973665,81.94683,2020-08-23 05:50:43.0,2020-07-31 14:24:55.0,875,40,INR,olive,401-09-0
475,5,3,2
BK997882598,58504148,46158041,4.3.35,iOS,29.9369725,-124.215858,-58.4375475,-89.123929,2020-01-12 08:48:05.0,2020-04-20 20:47:16.0,230,3,INR,green,266-60-742
4,3,1,1
BK5662087214,77440134,39470327,3.2.17,Android,13.965136,16.864746,-42.6375745,172.758412,2020-02-02 12:03:37.0,2020-09-26 06:29:56.0,941,88,INR,gray,279-59-5
349,5,4,1
BK8865887274,71215906,71671654,2.3.21,iOS,-85.653217,-6.88182,9.9079,-11.557289,2020-08-12 05:42:41.0,2020-06-15 18:24:22.0,894,72,INR,green,228-40-4173,2,2,
4
BK9708496297,29394176,20731780,2.4.7,Android,-45.18792,-33.406776,-67.895486,78.633127,2020-05-31 12:33:42.0,2020-04-12 21:21:13.0,938,15,INR,aqua,616-72-497
1,2,5,3
BK605227090,94946719,98071218,3.4.8,iOS,-4.8054875,-76.313234,76.0732515,-136.36527,2020-07-25 14:13:40.0,2020-07-04 14:08:20.0,716,19,INR,teal,112-71-5489,2
,5,3
BK8602928713,33696426,68228240,4.2.33,iOS,29.8689545,83.997957,-43.8161155,88.585796,2020-07-10 22:27:16.0,2020-06-07 20:32:40.0,763,65,INR,fuchsia,069-53-21
45,5,1,1
BK6243816121,74681879,25278350,2.3.24,iOS,-53.204624,-50.218077,85.7211825,-25.819898,2020-05-11 20:53:51.0,2020-08-26 11:52:47.0,196,78,INR,purple,874-93-24
59,4,1,4
BK9843664360,67564464,71913052,2.1.30,iOS,-10.031329,-49.100434,73.858805,157.577843,2020-09-20 22:51:48.0,2020-01-27 22:11:07.0,289,49,INR,maroon,876-76-759
8,3,3,3
BK2822269780,86988153,87739332,4.4.35,iOS,64.2850085,95.841271,-84.535181,-96.117739,2020-06-23 05:38:14.0,2020-07-11 16:40:57.0,196,89,INR,purple,740-96-793
1,4,5,4
hadoop@ip-172-31-80-241 ~]$
```