

# Assignment Day 5

## Input Dataset:

**hadoop fs -cat /dataset/musicdata.txt**

(Displays content of Input Dataset)

## ScreenShot:

```
[acadgild@localhost ~]$ hadoop fs -cat /dataset/musicdata.txt
18/07/03 05:45:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
111115|222|0|1|0
111113|225|1|0|0
111117|223|0|1|1
111115|225|1|0|0
[acadgild@localhost ~]$
```

## Task 1:

Write a Map Reduce program to find the number of unique listeners in the data set.

**Ans:**

## Run a Job:

**hadoop jar task1.jar /dataset/musicdata.txt /hadoopdata/hdfs/out/hands-on/o1Jul18/Assignment5/task1**

(To run a MR job using created jar file.)

## Output Directory:

**hadoop fs -ls /hadoopdata/hdfs/out/hands-on/o1Jul18/Assignment5/task1**

(Listing contents of output directory)

**hadoop fs -cat /hadoopdata/hdfs/out/hands-on/o1Jul18/Assignment5/task1/part-r-00000**

(Displaying content of Output file)

**Note:** Program files are properly documented for a detailed description of each instruction used within the program.

## ScreenShot:

```
[acadgild@localhost Assignment5]$ hadoop jar task1.jar /dataset/musicdata.txt /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task1
18/07/03 05:47:40 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/03 05:47:42 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/03 05:47:43 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/07/03 05:47:44 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 05:47:44 INFO mapreduce.JobSubmitter: number of splits:1
18/07/03 05:47:44 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530539200724_0039
18/07/03 05:47:44 INFO impl.YarnClientImpl: Submitted application application_1530539200724_0039
18/07/03 05:47:44 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530539200724_0039/
18/07/03 05:47:44 INFO mapreduce.Job: Running job: job_1530539200724_0039
18/07/03 05:47:56 INFO mapreduce.Job: Job job_1530539200724_0039 running in uber mode : false
18/07/03 05:47:56 INFO mapreduce.Job: map 0% reduce 0%
18/07/03 05:48:04 INFO mapreduce.Job: map 100% reduce 0%
18/07/03 05:48:13 INFO mapreduce.Job: map 100% reduce 100%
18/07/03 05:48:13 INFO mapreduce.Job: Job job_1530539200724_0039 completed successfully
18/07/03 05:48:13 INFO mapreduce.Job: Counters: 49
File System Counters
  FILE: Number of bytes read=30
  FILE: Number of bytes written=215411
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=179
  HDFS: Number of bytes written=21
  HDFS: Number of read operations=6
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Launched reduce tasks=1
  Data-local map tasks=1
  Total time spent by all maps in occupied slots (ms)=6179
  Total time spent by all reduces in occupied slots (ms)=6554
  Total time spent by all map tasks (ms)=6179
  Total time spent by all reduce tasks (ms)=6554

[acadgild@localhost ~]$ hadoop fs -ls /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task1
18/07/03 05:53:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2018-07-03 05:48 /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task1/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 21 2018-07-03 05:48 /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task1/part-r-000000
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task1/part-r-000000
18/07/03 05:53:13 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
111113
111115
111117
[acadgild@localhost ~]$
```

## Task 2:

Write a Map Reduce program to calculate number of times a song was heard fully.

Ans:

### Run a Job:

```
hadoop jar task2.jar /dataset/musicdata.txt /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2
```

(To run a MR job using created jar file.)

## Output Directory:

**hadoop fs -ls** /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2

(Listing contents of output directory)

**hadoop fs -cat** /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2/part-r-00000

(Displaying content of Output file)

**Note:** Program files are properly documented for a detailed description of each instruction used within the program.

## ScreenShot:

```
[acadgild@localhost Assignment5]$ hadoop jar task2.jar /dataset/musicdata.txt /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2
18/07/03 05:54:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/03 05:54:35 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/03 05:54:37 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/07/03 05:54:37 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 05:54:37 INFO mapreduce.JobSubmitter: number of splits:1
18/07/03 05:54:38 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530539200724_0040
18/07/03 05:54:38 INFO impl.YarnClientImpl: Submitted application application_1530539200724_0040
18/07/03 05:54:38 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530539200724_0040/
18/07/03 05:54:38 INFO mapreduce.Job: Running job: job_1530539200724_0040
18/07/03 05:54:49 INFO mapreduce.Job: Job job_1530539200724_0040 running in uber mode : false
18/07/03 05:54:49 INFO mapreduce.Job:  map 0% reduce 0%
18/07/03 05:54:58 INFO mapreduce.Job:  map 100% reduce 0%
18/07/03 05:55:08 INFO mapreduce.Job:  map 100% reduce 100%
18/07/03 05:55:08 INFO mapreduce.Job: Job job_1530539200724_0040 completed successfully
18/07/03 05:55:08 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=12
    FILE: Number of bytes written=215375
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=179
    HDFS: Number of bytes written=46
    HDFS: Number of read operations=6
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
  Job Counters
    Launched map tasks=1
    Launched reduce tasks=1
    Data-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=5838
    Total time spent by all reduces in occupied slots (ms)=6911
    Total time spent by all map tasks (ms)=5838
    Total time spent by all reduce tasks (ms)=6911

[acadgild@localhost ~]$ hadoop fs -ls /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2
18/07/03 05:56:11 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup          0 2018-07-03 05:55 /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2/_SUCCESS
-rw-r--r-- 1 acadgild supergroup        46 2018-07-03 05:55 /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2/part-r-00000
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task2/part-r-00000
18/07/03 05:56:25 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Number of times a song was heard fully is : 1
[acadgild@localhost ~]$
```

## **Task 3:**

Write a Map Reduce program to calculate the number of times a song was shared.

**Ans:**

**Run a Job:**

```
hadoop jar task3.jar /dataset/musicdata.txt /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3
```

(To run a MR job using created jar file.)

**Output Directory:**

```
hadoop fs -ls /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3
```

(Listing contents of output directory)

```
hadoop fs -cat /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3/part-r-00000
```

(Displaying content of Output file)

**Note:** Program files are properly documented for a detailed description of each instruction used within the program.

**ScreenShot:**

```
[acadgild@localhost Assignment5]$ hadoop jar task3.jar /dataset/musicdata.txt /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3
18/07/03 06:01:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/03 06:01:28 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/03 06:01:29 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/07/03 06:01:30 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 06:01:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530539200724_0041
18/07/03 06:01:30 INFO impl.YarnClientImpl: Submitted application application_1530539200724_0041
18/07/03 06:01:31 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530539200724_0041/
18/07/03 06:01:31 INFO mapreduce.Job: Running job: job_1530539200724_0041
18/07/03 06:01:43 INFO mapreduce.Job: Job job_1530539200724_0041 running in uber mode : false
18/07/03 06:01:44 INFO mapreduce.Job: map 0% reduce 0%
18/07/03 06:01:52 INFO mapreduce.Job: map 100% reduce 0%
18/07/03 06:02:03 INFO mapreduce.Job: map 100% reduce 100%
18/07/03 06:02:03 INFO mapreduce.Job: Job job_1530539200724_0041 completed successfully
18/07/03 06:02:03 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=18
    FILE: Number of bytes written=215387
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=179
    HDFS: Number of bytes written=41
    HDFS: Number of read operations=6
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
  Job Counters
    Launched map tasks=1
    Launched reduce tasks=1
    Data-local map tasks=1
    Total time spent by all maps in occupied slots (ms)=6520
    Total time spent by all reduces in occupied slots (ms)=7290
    Total time spent by all map tasks (ms)=6520
    Total time spent by all reduce tasks (ms)=7290
```

```
[acadgild@localhost ~]$ hadoop fs -ls /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3
18/07/03 06:02:38 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2018-07-03 06:02 /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 41 2018-07-03 06:02 /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3/part-r-00000
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /hadoopdata/hdfs/out/hands-on/01Jul18/Assignment5/task3/part-r-00000
18/07/03 06:03:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Number of times a song was shared is : 2
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

\*\*\*\*\*

End

\*\*\*\*\*