***Assignment3 YARN mean,median,standarddeviation***

***Input file sample.txt is present in /hadoopdata/hdfs***

Using put command file is copied from local to HDFS

[acadgild@localhost mapreduce]$ hdfs dfs -ls /hadoopdata/hdfs/sa\*

18/06/20 14:01:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

-rw-r--r-- 2 acadgild supergroup 0 2018-05-19 14:02 /hadoopdata/hdfs/sagar.txt

-rw-r--r-- 1 acadgild supergroup 136 2018-06-20 13:39 **/hadoopdata/hdfs/sampletext.txt**

[acadgild@localhost ~]$ cat sampletext.txt

Hi This is robot learning Hadoop

The current session is about YARN

YARN has 6 components

RM NM scheduler appmanager appmaster container

***All out put files for three programmes are present in /hadoopdata folder***

drwxr-xr-x - acadgild supergroup 0 2018-06-20 13:48 /hadoopdata/wordmean

drwxr-xr-x - acadgild supergroup 0 2018-06-20 13:49 /hadoopdata/wordmedian

drwxr-xr-x - acadgild supergroup 0 2018-06-20 13:50 /hadoopdata/wordstandarddeviation

Below are the commands executed for processing the sampletext file using the jar present in folder/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce

1. Wordmean

[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /hadoopdata/hdfs/sampletext.txt /hadoopdata/wordmean

1. Wordmedian

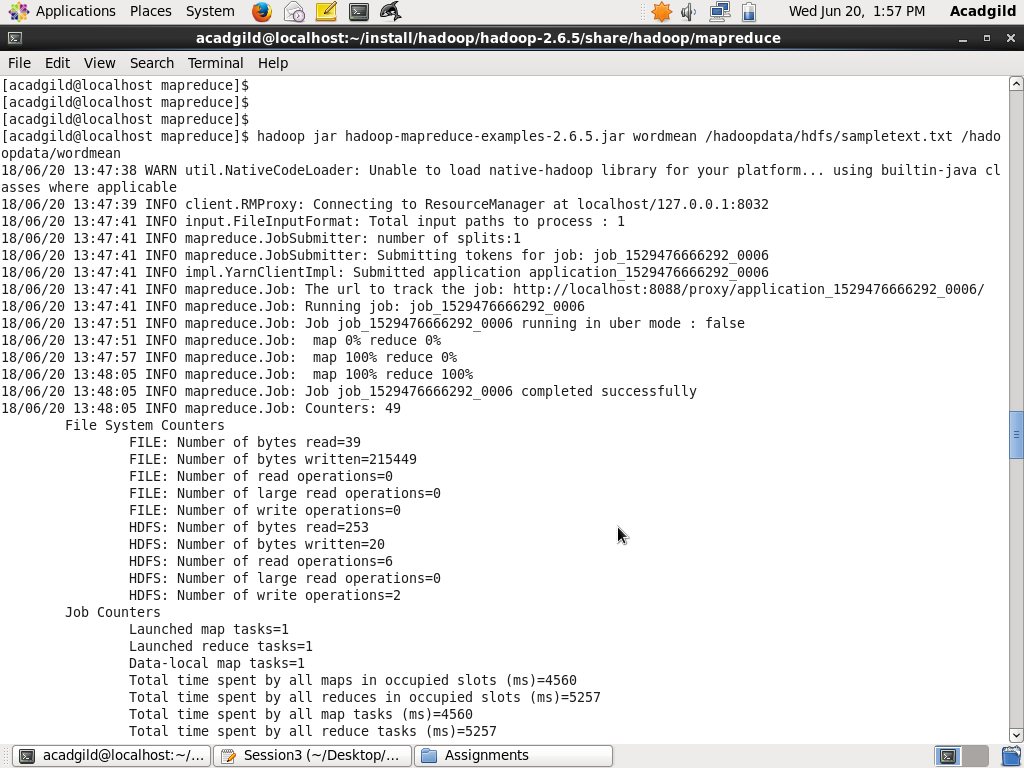
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /hadoopdata/hdfs/sampletext.txt /hadoopdata/wordmedian

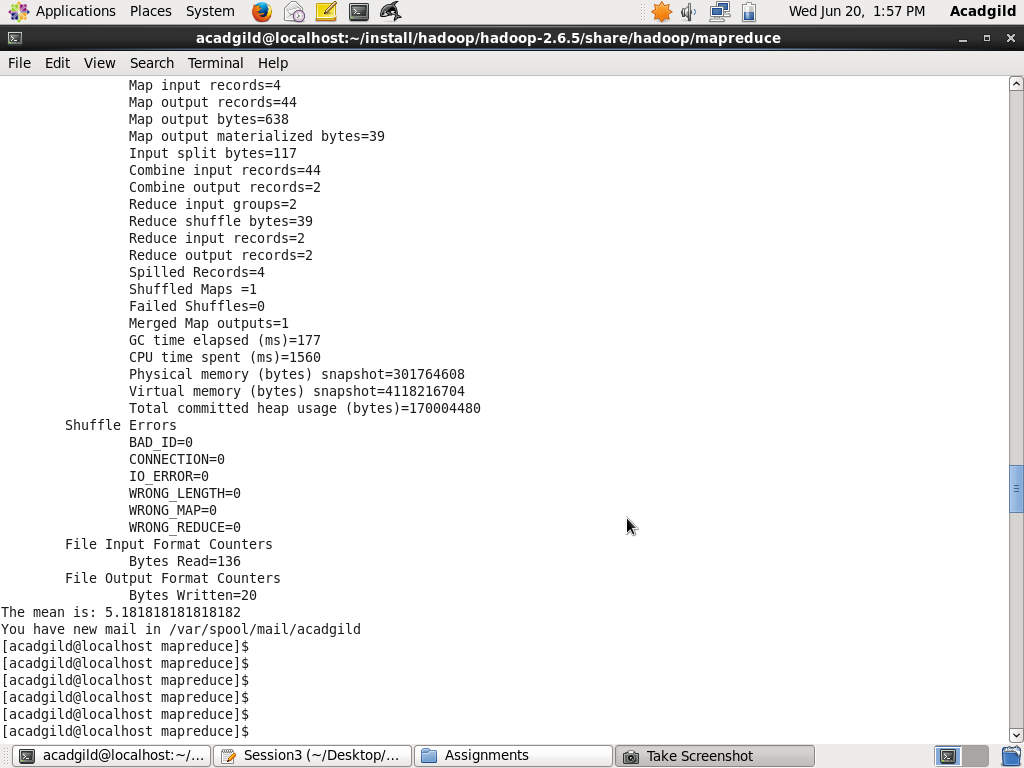
1. WordStandarddeviation

[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /hadoopdata/hdfs/sampletext.txt /hadoopdata/wordstandarddeviation

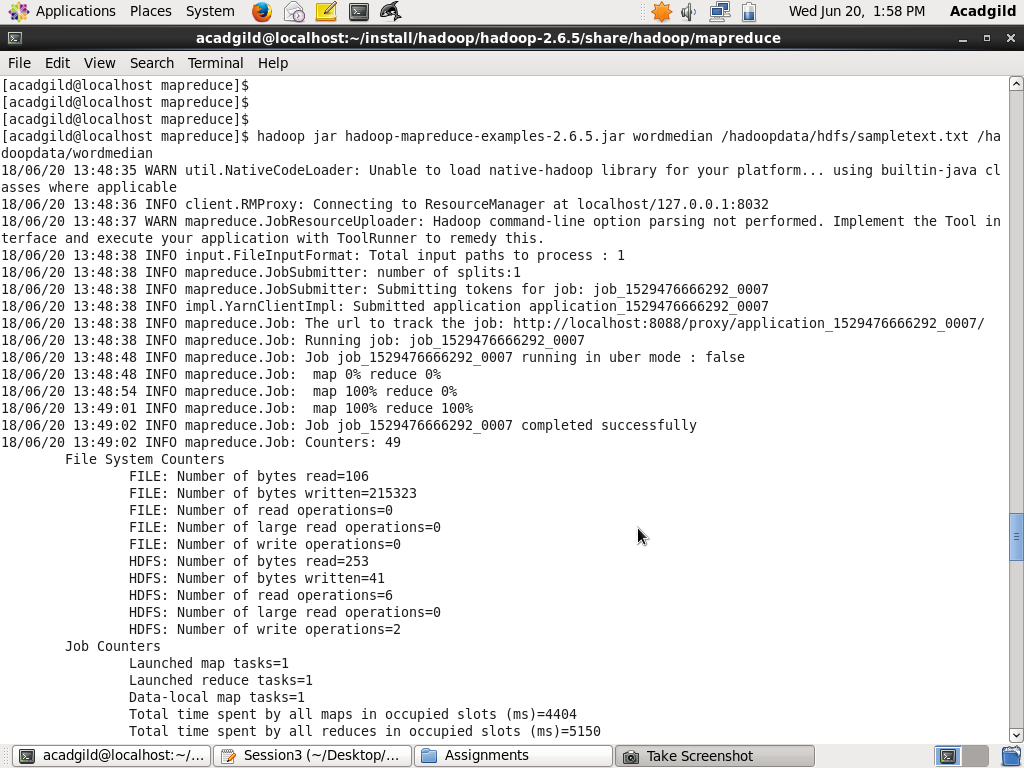
*Screen prints of the programme.*

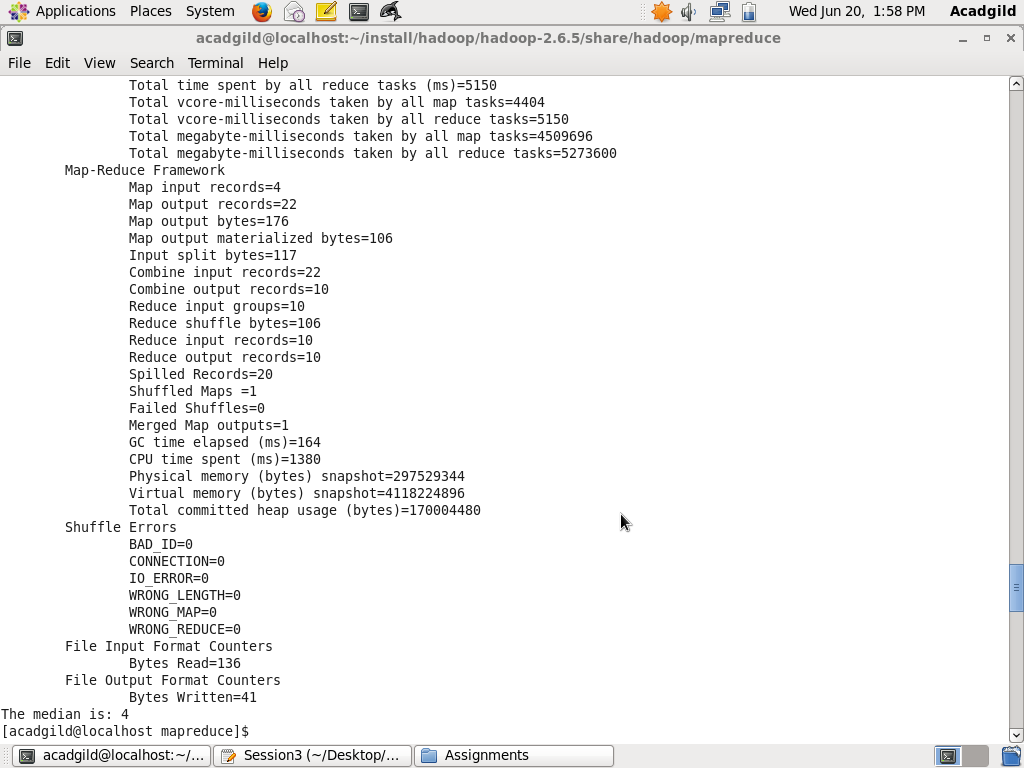
**1.Word Mean**



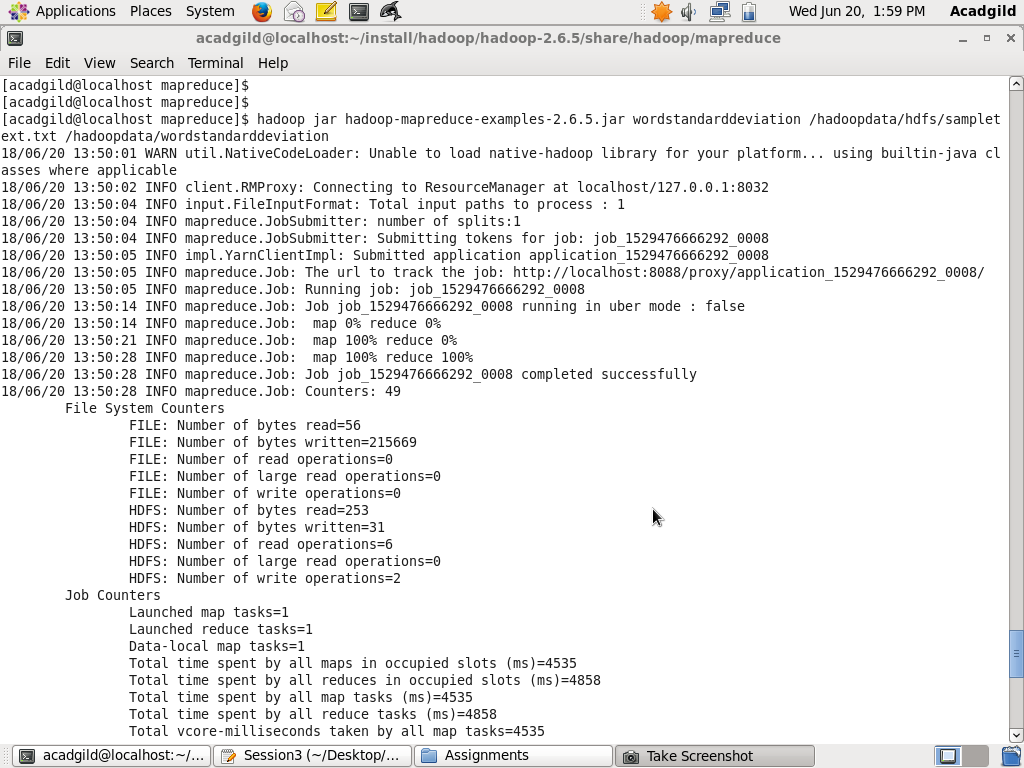


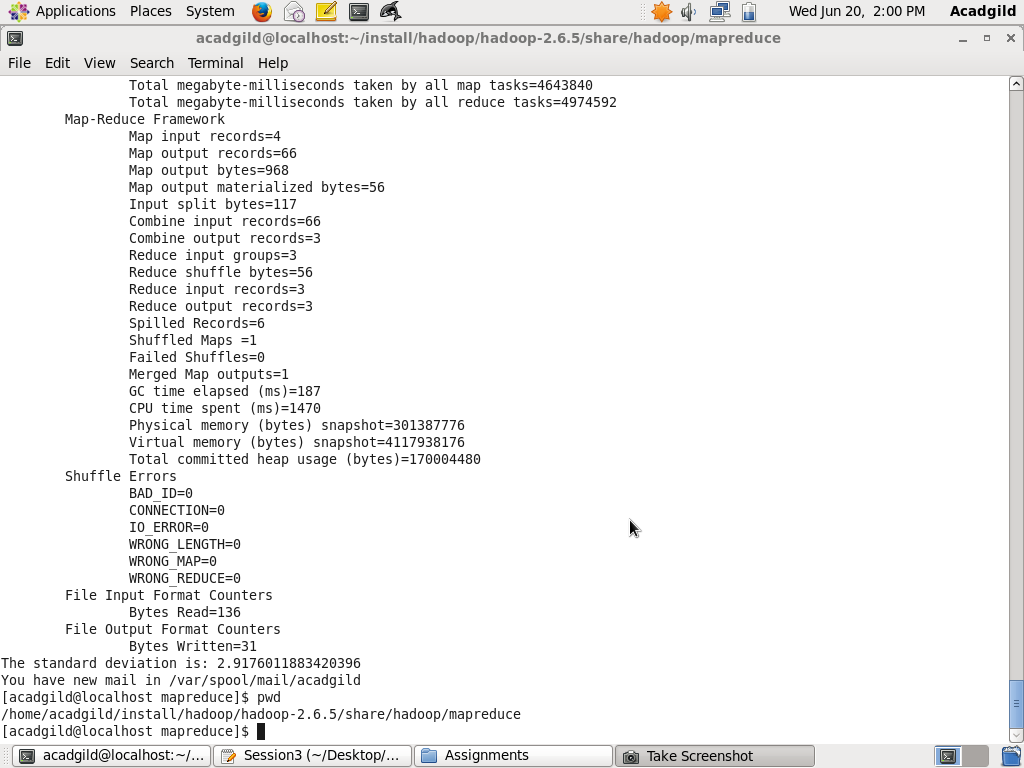
**2. WordMedian**





1. **wordStandarddeviation**





*Copied from terminal with the output files*

**WordMean**

[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /hadoopdata/hdfs/sampletext.txt /hadoopdata/wordmean

18/06/20 13:47:38 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/06/20 13:47:39 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/06/20 13:47:41 INFO input.FileInputFormat: Total input paths to process : 1

18/06/20 13:47:41 INFO mapreduce.JobSubmitter: number of splits:1

18/06/20 13:47:41 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1529476666292\_0006

18/06/20 13:47:41 INFO impl.YarnClientImpl: Submitted application application\_1529476666292\_0006

18/06/20 13:47:41 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application\_1529476666292\_0006/

18/06/20 13:47:41 INFO mapreduce.Job: Running job: job\_1529476666292\_0006

18/06/20 13:47:51 INFO mapreduce.Job: Job job\_1529476666292\_0006 running in uber mode : false

18/06/20 13:47:51 INFO mapreduce.Job: map 0% reduce 0%

18/06/20 13:47:57 INFO mapreduce.Job: map 100% reduce 0%

18/06/20 13:48:05 INFO mapreduce.Job: map 100% reduce 100%

18/06/20 13:48:05 INFO mapreduce.Job: Job job\_1529476666292\_0006 completed successfully

18/06/20 13:48:05 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=39

FILE: Number of bytes written=215449

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=253

HDFS: Number of bytes written=20

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)=4560

Total time spent by all reduces in occupied slots (ms)=5257

Total time spent by all map tasks (ms)=4560

Total time spent by all reduce tasks (ms)=5257

Total vcore-milliseconds taken by all map tasks=4560

Total vcore-milliseconds taken by all reduce tasks=5257

Total megabyte-milliseconds taken by all map tasks=4669440

Total megabyte-milliseconds taken by all reduce tasks=5383168

Map-Reduce Framework

Map input records=4

Map output records=44

Map output bytes=638

Map output materialized bytes=39

Input split bytes=117

Combine input records=44

Combine output records=2

Reduce input groups=2

Reduce shuffle bytes=39

Reduce input records=2

Reduce output records=2

Spilled Records=4

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=177

CPU time spent (ms)=1560

Physical memory (bytes) snapshot=301764608

Virtual memory (bytes) snapshot=4118216704

Total committed heap usage (bytes)=170004480

Shuffle Errors

BAD\_ID=0

CONNECTION=0

IO\_ERROR=0

WRONG\_LENGTH=0

WRONG\_MAP=0

WRONG\_REDUCE=0

File Input Format Counters

Bytes Read=136

File Output Format Counters

Bytes Written=20

**The mean is: 5.181818181818182**

**[acadgild@localhost mapreduce]$ hdfs dfs -ls /hadoopdata/wordmean**

**18/06/20 14:03:40 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 2018-06-20 13:48 /hadoopdata/wordmean/\_SUCCESS**

**-rw-r--r-- 1 acadgild supergroup 20 2018-06-20 13:48 /hadoopdata/wordmean/part-r-00000**

**[acadgild@localhost mapreduce]$ hdfs dfs -cat /hadoopdata/wordmean/part-r-00000**

**18/06/20 14:04:32 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable**

**count 22**

**length 114**

**WORDMEDIAN**

[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /hadoopdata/hdfs/sampletext.txt /hadoopdata/wordmedian

18/06/20 13:48:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/06/20 13:48:36 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/06/20 13:48:37 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.

18/06/20 13:48:38 INFO input.FileInputFormat: Total input paths to process : 1

18/06/20 13:48:38 INFO mapreduce.JobSubmitter: number of splits:1

18/06/20 13:48:38 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1529476666292\_0007

18/06/20 13:48:38 INFO impl.YarnClientImpl: Submitted application application\_1529476666292\_0007

18/06/20 13:48:38 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application\_1529476666292\_0007/

18/06/20 13:48:38 INFO mapreduce.Job: Running job: job\_1529476666292\_0007

18/06/20 13:48:48 INFO mapreduce.Job: Job job\_1529476666292\_0007 running in uber mode : false

18/06/20 13:48:48 INFO mapreduce.Job: map 0% reduce 0%

18/06/20 13:48:54 INFO mapreduce.Job: map 100% reduce 0%

18/06/20 13:49:01 INFO mapreduce.Job: map 100% reduce 100%

18/06/20 13:49:02 INFO mapreduce.Job: Job job\_1529476666292\_0007 completed successfully

18/06/20 13:49:02 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=106

FILE: Number of bytes written=215323

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=253

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)=4404

Total time spent by all reduces in occupied slots (ms)=5150

Total time spent by all map tasks (ms)=4404

Total time spent by all reduce tasks (ms)=5150

Total vcore-milliseconds taken by all map tasks=4404

Total vcore-milliseconds taken by all reduce tasks=5150

Total megabyte-milliseconds taken by all map tasks=4509696

Total megabyte-milliseconds taken by all reduce tasks=5273600

Map-Reduce Framework

Map input records=4

Map output records=22

Map output bytes=176

Map output materialized bytes=106

Input split bytes=117

Combine input records=22

Combine output records=10

Reduce input groups=10

Reduce shuffle bytes=106

Reduce input records=10

Reduce output records=10

Spilled Records=20

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=164

CPU time spent (ms)=1380

Physical memory (bytes) snapshot=297529344

Virtual memory (bytes) snapshot=4118224896

Total committed heap usage (bytes)=170004480

Shuffle Errors

BAD\_ID=0

CONNECTION=0

IO\_ERROR=0

WRONG\_LENGTH=0

WRONG\_MAP=0

WRONG\_REDUCE=0

File Input Format Counters

Bytes Read=136

File Output Format Counters

Bytes Written=41

**The median is: 4**

**[acadgild@localhost mapreduce]$ hdfs dfs -ls /hadoopdata/wordmedian**

**18/06/20 14:05:09 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 2018-06-20 13:49 /hadoopdata/wordmedian/\_SUCCESS**

**-rw-r--r-- 1 acadgild supergroup 41 2018-06-20 13:49 /hadoopdata/wordmedian/part-r-00000**

**[acadgild@localhost mapreduce]$ hdfs dfs -cat /hadoopdata/wordmedian/part-r-00000**

**18/06/20 14:05:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable**

**1 1**

**2 5**

**3 2**

**4 3**

**5 2**

**6 1**

**7 2**

**8 1**

**9 3**

**10 2**

**WORDSTANDARDDEVIATION**

[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /hadoopdata/hdfs/sampletext.txt /hadoopdata/wordstandarddeviation

18/06/20 13:50:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/06/20 13:50:02 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/06/20 13:50:04 INFO input.FileInputFormat: Total input paths to process : 1

18/06/20 13:50:04 INFO mapreduce.JobSubmitter: number of splits:1

18/06/20 13:50:04 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1529476666292\_0008

18/06/20 13:50:05 INFO impl.YarnClientImpl: Submitted application application\_1529476666292\_0008

18/06/20 13:50:05 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application\_1529476666292\_0008/

18/06/20 13:50:05 INFO mapreduce.Job: Running job: job\_1529476666292\_0008

18/06/20 13:50:14 INFO mapreduce.Job: Job job\_1529476666292\_0008 running in uber mode : false

18/06/20 13:50:14 INFO mapreduce.Job: map 0% reduce 0%

18/06/20 13:50:21 INFO mapreduce.Job: map 100% reduce 0%

18/06/20 13:50:28 INFO mapreduce.Job: map 100% reduce 100%

18/06/20 13:50:28 INFO mapreduce.Job: Job job\_1529476666292\_0008 completed successfully

18/06/20 13:50:28 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=56

FILE: Number of bytes written=215669

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=253

HDFS: Number of bytes written=31

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)=4535

Total time spent by all reduces in occupied slots (ms)=4858

Total time spent by all map tasks (ms)=4535

Total time spent by all reduce tasks (ms)=4858

Total vcore-milliseconds taken by all map tasks=4535

Total vcore-milliseconds taken by all reduce tasks=4858

Total megabyte-milliseconds taken by all map tasks=4643840

Total megabyte-milliseconds taken by all reduce tasks=4974592

Map-Reduce Framework

Map input records=4

Map output records=66

Map output bytes=968

Map output materialized bytes=56

Input split bytes=117

Combine input records=66

Combine output records=3

Reduce input groups=3

Reduce shuffle bytes=56

Reduce input records=3

Reduce output records=3

Spilled Records=6

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=187

CPU time spent (ms)=1470

Physical memory (bytes) snapshot=301387776

Virtual memory (bytes) snapshot=4117938176

Total committed heap usage (bytes)=170004480

Shuffle Errors

BAD\_ID=0

CONNECTION=0

IO\_ERROR=0

WRONG\_LENGTH=0

WRONG\_MAP=0

WRONG\_REDUCE=0

File Input Format Counters

Bytes Read=136

File Output Format Counters

Bytes Written=31

**The standard deviation is: 2.9176011883420396**

**You have new mail in /var/spool/mail/acadgild**

**[acadgild@localhost mapreduce]$ pwd**

**/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce**

**[acadgild@localhost mapreduce]$ hdfs dfs -ls /hadoopdata/wordstandarddeviation**

**18/06/20 14:05:53 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 2018-06-20 13:50 /hadoopdata/wordstandarddeviation/\_SUCCESS**

**-rw-r--r-- 1 acadgild supergroup 31 2018-06-20 13:50 /hadoopdata/wordstandarddeviation/part-r-00000**

**You have new mail in /var/spool/mail/acadgild**

**[acadgild@localhost mapreduce]$ hdfs dfs -cat /hadoopdata/wordstandarddeviation/part-r-00000**

**18/06/20 14:06:54 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable**

**count 22**

**length 114**

**square 778**

**You have new mail in /var/spool/mail/acadgild**

**[acadgild@localhost mapreduce]$**

INPUT and OUTFILE screen prints

