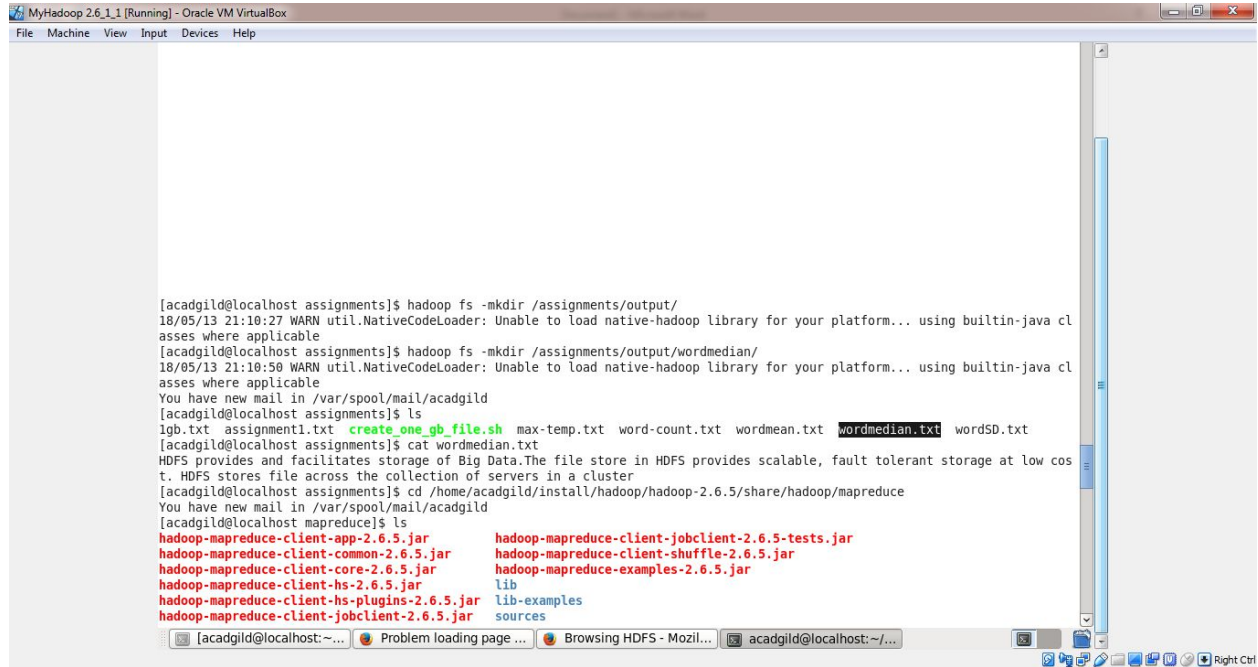


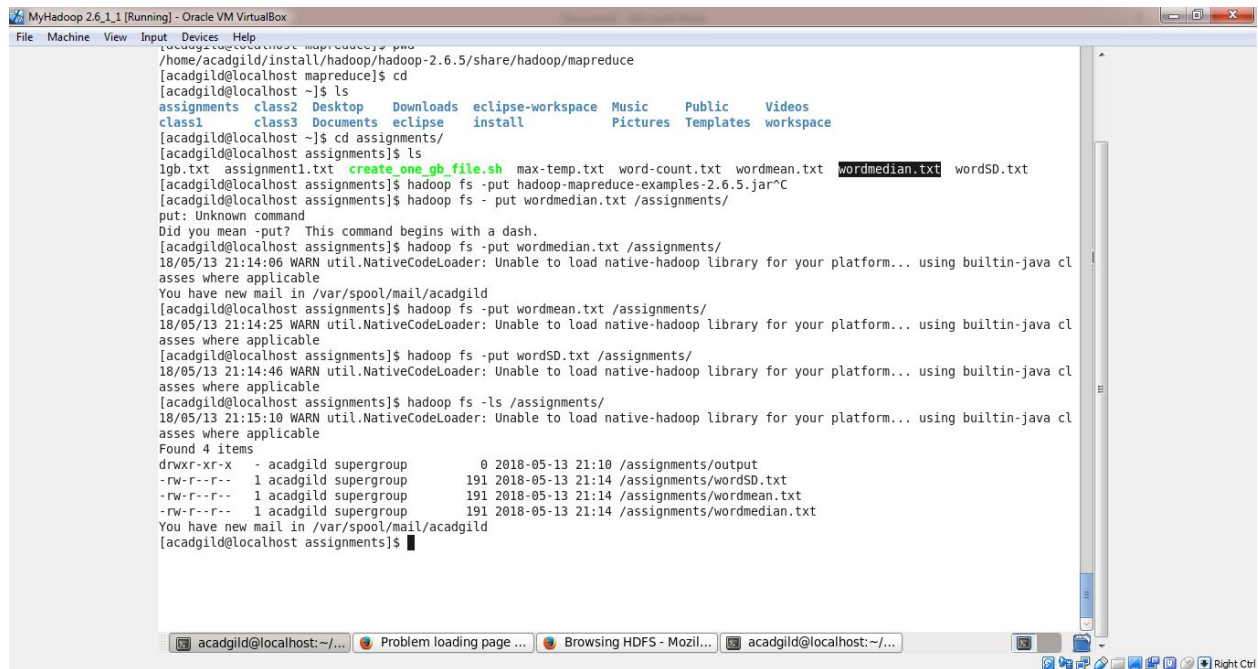
Execute WordMedian , WordMean , WordStandardDeviation programs using hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.

Create files to execute Word mean, Word Median, Word standard deviation programs and move the files to HDFS.



```
MyHadoop 2.6.1.1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

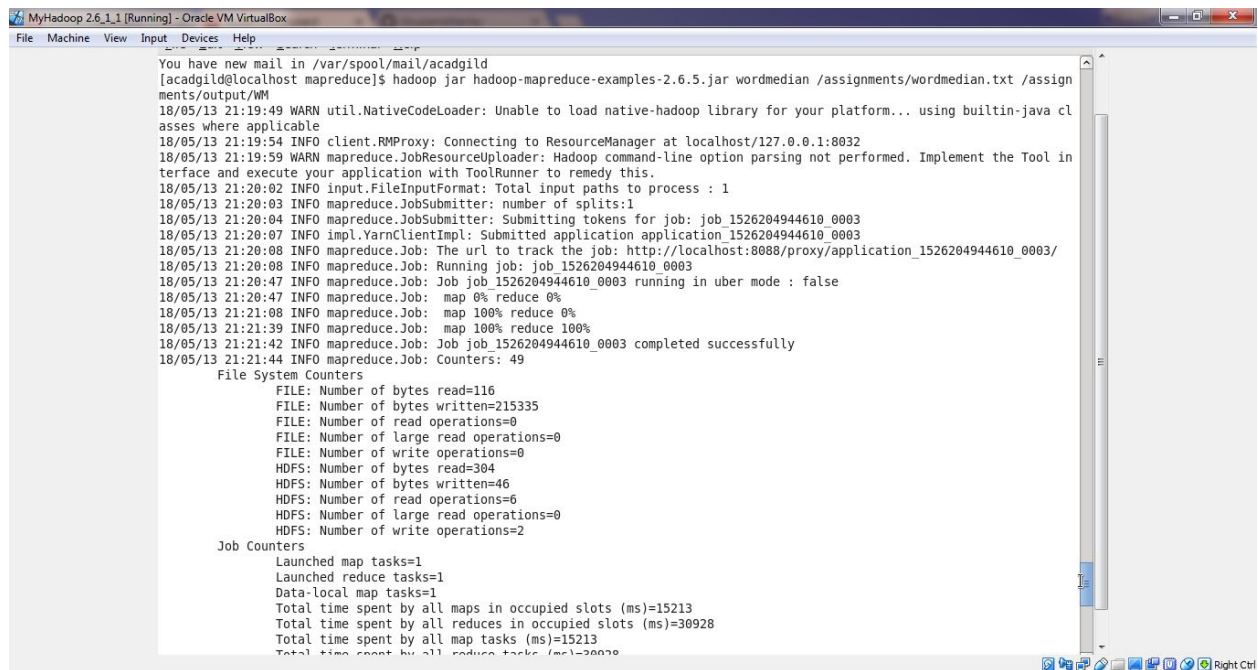
[acadgild@localhost assignments]$ hadoop fs -mkdir /assignments/output/
18/05/13 21:10:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
[acadgild@localhost assignments]$ hadoop fs -mkdir /assignments/output/wordmedian/
18/05/13 21:10:50 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost assignments]$ ls
lgb.txt assignment1.txt create_one_gb_file.sh max-temp.txt word-count.txt wordmean.txt wordmedian.txt wordSD.txt
[acadgild@localhost assignments]$ cat wordmedian.txt
HDFS provides and facilitates storage of Big Data.The file store in HDFS provides scalable, fault tolerant storage at low cos
t. HDFS stores file across the collection of servers in a cluster
[acadgild@localhost assignments]$ cd /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ ls
hadoop-mapreduce-client-app-2.6.5.jar      hadoop-mapreduce-client-jobclient-2.6.5-tests.jar
hadoop-mapreduce-client-common-2.6.5.jar   hadoop-mapreduce-client-shuffle-2.6.5.jar
hadoop-mapreduce-client-core-2.6.5.jar     hadoop-mapreduce-examples-2.6.5.jar
hadoop-mapreduce-client-hs-2.6.5.jar       lib
hadoop-mapreduce-client-hs-plugins-2.6.5.jar lib-examples
hadoop-mapreduce-client-jobclient-2.6.5.jar sources
[acadgild@localhost:~/... Problem loading page ... Browsing HDFS - Mozil... acadgild@localhost:~/...
```



```
MyHadoop 2.6.1.1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

[acadgild@localhost mapreduce]$ pwd
/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce
[acadgild@localhost mapreduce]$ cd
[acadgild@localhost ~]$ ls
assignments class2 Desktop Downloads eclipse-workspace Music Public Videos
class1 class3 Documents eclipse install Pictures Templates workspace
[acadgild@localhost ~]$ cd assignments/
[acadgild@localhost assignments]$ ls
lgb.txt assignment1.txt create_one_gb_file.sh max-temp.txt word-count.txt wordmean.txt wordmedian.txt wordSD.txt
[acadgild@localhost assignments]$ hadoop fs -put hadoop-mapreduce-examples-2.6.5.jar~C
[acadgild@localhost assignments]$ hadoop fs -put wordmedian.txt /assignments/
put: Unknown command
Did you mean -put? This command begins with a dash.
[acadgild@localhost assignments]$ hadoop fs -put wordmedian.txt /assignments/
18/05/13 21:14:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost assignments]$ hadoop fs -put wordmean.txt /assignments/
18/05/13 21:14:25 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
[acadgild@localhost assignments]$ hadoop fs -put wordSD.txt /assignments/
18/05/13 21:14:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
[acadgild@localhost assignments]$ hadoop fs -ls /assignments/
18/05/13 21:15:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 4 items
drwxr-xr-x - acadgild supergroup 0 2018-05-13 21:10 /assignments/output
-rw-r--r-- 1 acadgild supergroup 191 2018-05-13 21:14 /assignments/wordSD.txt
-rw-r--r-- 1 acadgild supergroup 191 2018-05-13 21:14 /assignments/wordmean.txt
-rw-r--r-- 1 acadgild supergroup 191 2018-05-13 21:14 /assignments/wordmedian.txt
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost assignments]$
```

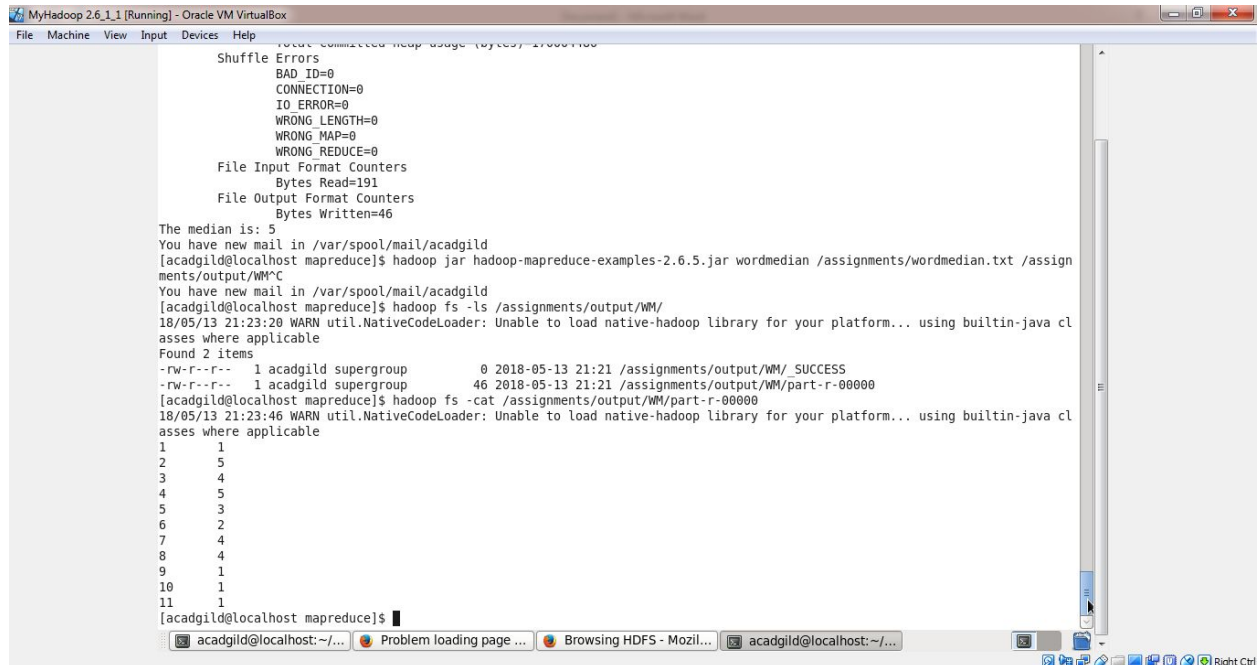
Word Median is a map reduce program which counts the median length of words in the input file.



```
MyHadoop 2.6.1.1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /assignments/wordmedian.txt /assignments/output/WM
18/05/13 21:19:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/13 21:19:54 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/13 21:19:59 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/05/13 21:20:02 INFO input.FileInputFormat: Total input paths to process : 1
18/05/13 21:20:03 INFO mapreduce.JobSubmitter: number of splits:1
18/05/13 21:20:04 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526204944610_0003
18/05/13 21:20:07 INFO impl.YarnClientImpl: Submitted application application_1526204944610_0003
18/05/13 21:20:08 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526204944610_0003/
18/05/13 21:20:08 INFO mapreduce.Job: Running job: job_1526204944610_0003
18/05/13 21:20:47 INFO mapreduce.Job: Job job_1526204944610_0003 running in uber mode : false
18/05/13 21:20:47 INFO mapreduce.Job: map 0% reduce 0%
18/05/13 21:21:08 INFO mapreduce.Job: map 100% reduce 0%
18/05/13 21:21:39 INFO mapreduce.Job: map 100% reduce 100%
18/05/13 21:21:42 INFO mapreduce.Job: Job job_1526204944610_0003 completed successfully
18/05/13 21:21:44 INFO mapreduce.Job: Counters: 49
File System Counters
  FILE: Number of bytes read=116
  FILE: Number of bytes written=215335
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=304
  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)=15213
  Total time spent by all reduces in occupied slots (ms)=30928
  Total time spent by all map tasks (ms)=15213
  Total time spent by all reduce tasks (ms)=30928
```

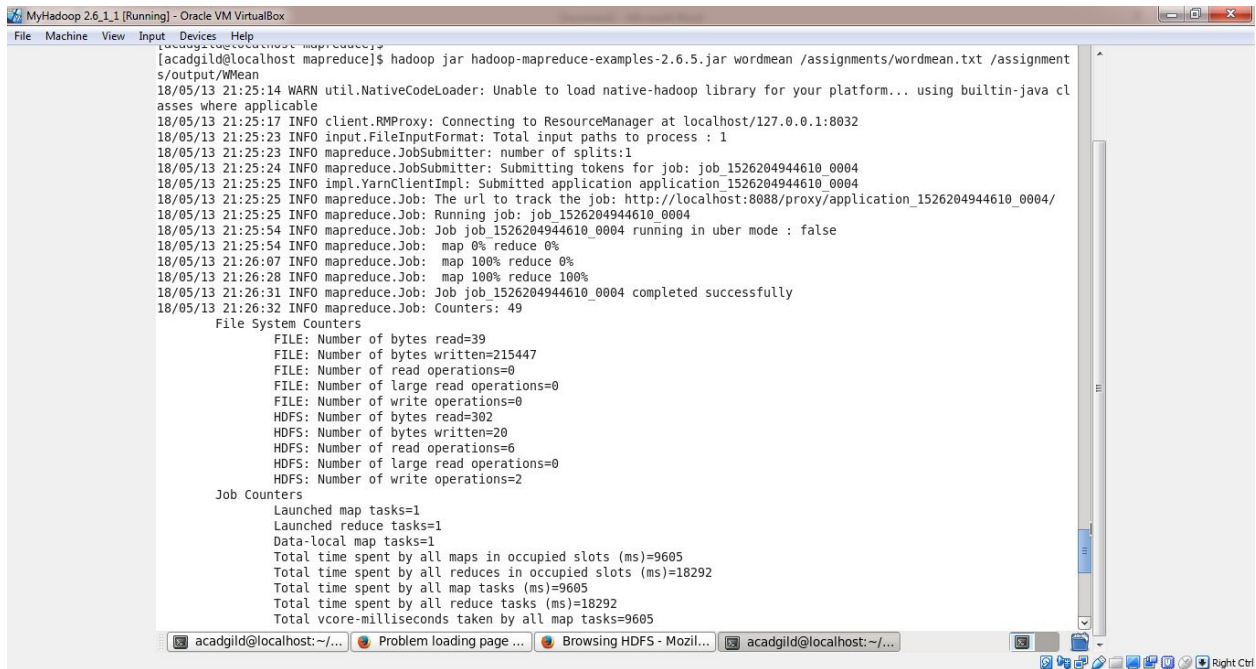
Word median – output:



```
MyHadoop 2.6.1.1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

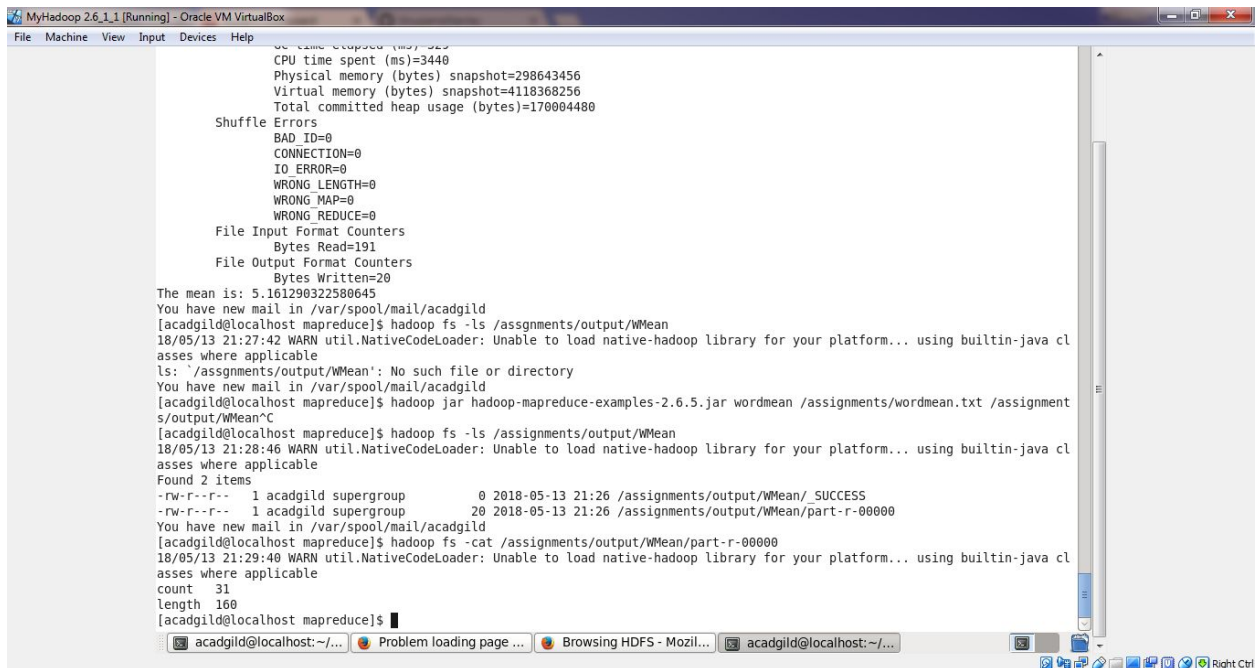
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=191
File Output Format Counters
  Bytes Written=46
The median is: 5
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /assignments/wordmedian.txt /assignments/output/WM^C
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop fs -ls /assignments/output/WM/
18/05/13 21:23:20 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 2018-05-13 21:21 /assignments/output/WM/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 46 2018-05-13 21:21 /assignments/output/WM/part-r-00000
[acadgild@localhost mapreduce]$ hadoop fs -cat /assignments/output/WM/part-r-00000
18/05/13 21:23:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
1
1
2
5
3
4
4
5
5
3
6
2
7
4
8
4
9
1
10
1
11
1
[acadgild@localhost mapreduce]$
```

Word mean is a map reduce program that counts the average length of words in the input file.



```
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /assignments/wordmean.txt /assignment
s/output/WMean
18/05/13 21:25:14 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/05/13 21:25:17 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/13 21:25:23 INFO input.FileInputFormat: Total input paths to process : 1
18/05/13 21:25:23 INFO mapreduce.JobSubmitter: number of splits:1
18/05/13 21:25:24 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526204944610_0004
18/05/13 21:25:25 INFO impl.YarnClientImpl: Submitted application application_1526204944610_0004
18/05/13 21:25:25 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526204944610_0004/
18/05/13 21:25:25 INFO mapreduce.Job: Running job: job_1526204944610_0004
18/05/13 21:25:54 INFO mapreduce.Job: Job job_1526204944610_0004 running in uber mode : false
18/05/13 21:25:54 INFO mapreduce.Job: map 0% reduce 0%
18/05/13 21:26:07 INFO mapreduce.Job: map 100% reduce 0%
18/05/13 21:26:28 INFO mapreduce.Job: map 100% reduce 100%
18/05/13 21:26:31 INFO mapreduce.Job: Job job_1526204944610_0004 completed successfully
18/05/13 21:26:32 INFO mapreduce.Job: Counters: 49
    File System Counters
      FILE: Number of bytes read=39
      FILE: Number of bytes written=215447
      FILE: Number of read operations=0
      FILE: Number of large read operations=0
      FILE: Number of write operations=0
      HDFS: Number of bytes read=302
      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)=9605
      Total time spent by all reduces in occupied slots (ms)=18292
      Total time spent by all map tasks (ms)=9605
      Total time spent by all reduce tasks (ms)=18292
      Total vcore-milliseconds taken by all map tasks=9605
```

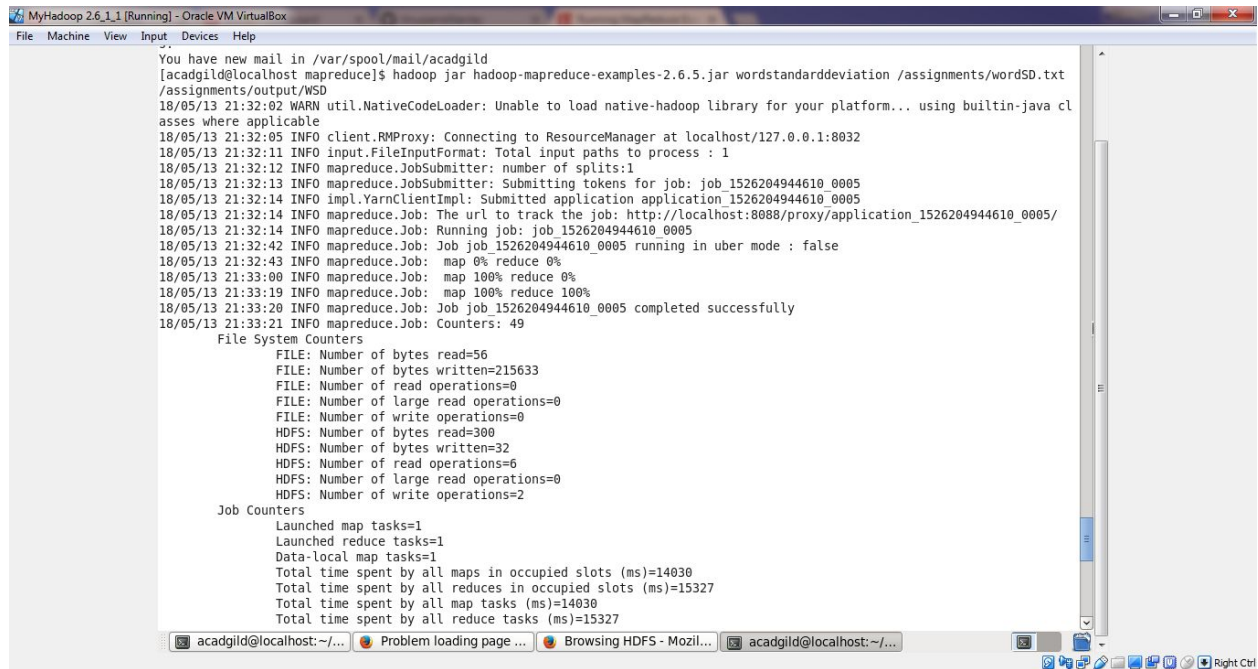
Word mean – output:



```

CPU time spent (ms)=3440
Physical memory (bytes) snapshot=298643456
Virtual memory (bytes) snapshot=4118368256
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=191
File Output Format Counters
  Bytes Written=20
The mean is: 5.161290322580645
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop fs -ls /assignments/output/WMean
18/05/13 21:27:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
ls: '/assignments/output/WMean': No such file or directory
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /assignments/wordmean.txt /assignment
s/output/WMean^C
[acadgild@localhost mapreduce]$ hadoop fs -ls /assignments/output/WMean
18/05/13 21:28:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup          0 2018-05-13 21:26 /assignments/output/WMean/_SUCCESS
-rw-r--r-- 1 acadgild supergroup        20 2018-05-13 21:26 /assignments/output/WMean/part-r-000000
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop fs -cat /assignments/output/WMean/part-r-000000
18/05/13 21:29:40 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
count      31
length    160
[acadgild@localhost mapreduce]$
```

Word Standard deviation is a mapreduce program that counts the standard deviation of length of words in the input files



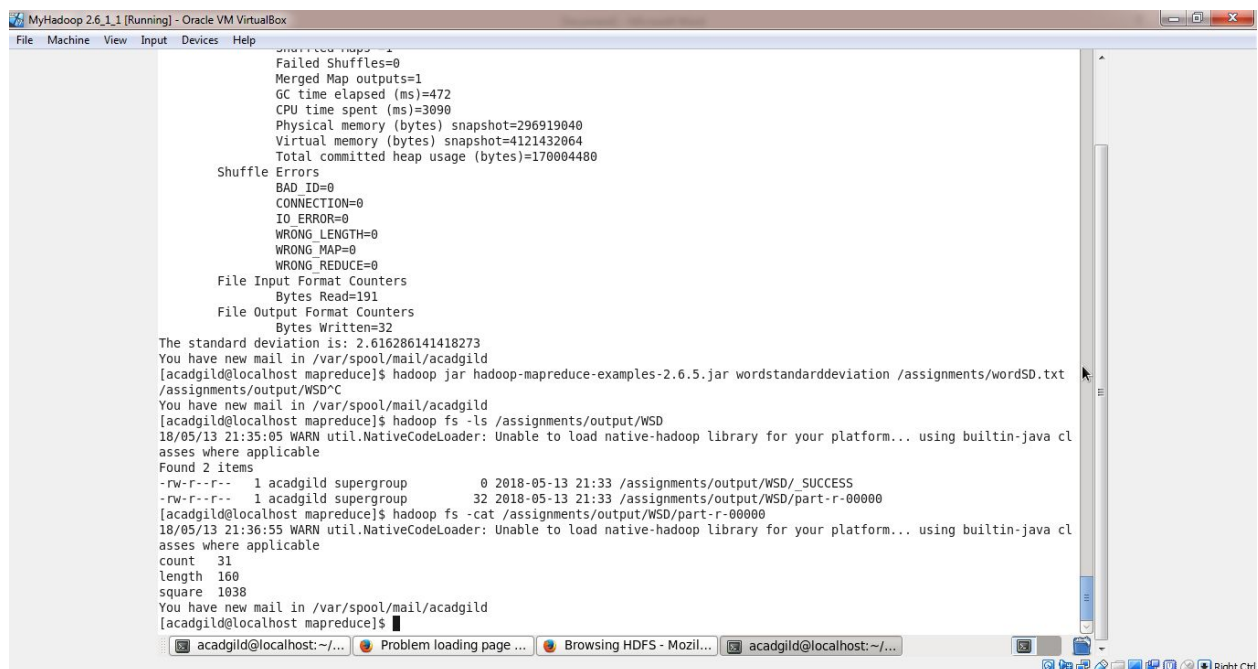
```
MyHadoop 2.6.1.1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /assignments/wordSD.txt /assignments/output/WSDC
18/05/13 21:32:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/13 21:32:05 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/13 21:32:11 INFO input.FileInputFormat: Total input paths to process : 1
18/05/13 21:32:12 INFO mapreduce.JobSubmitter: number of splits:1
18/05/13 21:32:13 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526204944610_0005
18/05/13 21:32:14 INFO impl.YarnClientImpl: Submitted application application_1526204944610_0005
18/05/13 21:32:14 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526204944610_0005/
18/05/13 21:32:14 INFO mapreduce.Job: Running job: job_1526204944610_0005
18/05/13 21:32:42 INFO mapreduce.Job: Job job_1526204944610_0005 running in uber mode : false
18/05/13 21:32:43 INFO mapreduce.Job: map 0% reduce 0%
18/05/13 21:33:00 INFO mapreduce.Job: map 100% reduce 0%
18/05/13 21:33:19 INFO mapreduce.Job: map 100% reduce 100%
18/05/13 21:33:20 INFO mapreduce.Job: Job job_1526204944610_0005 completed successfully
18/05/13 21:33:21 INFO mapreduce.Job: Counters: 49

File System Counters
  FILE: Number of bytes read=56
  FILE: Number of bytes written=215633
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=300
  HDFS: Number of bytes written=32
  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)=14030
  Total time spent by all reduces in occupied slots (ms)=15327
  Total time spent by all map tasks (ms)=14030
  Total time spent by all reduce tasks (ms)=15327
```

Word Standard Deviation- output:



```
MyHadoop 2.6.1.1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help

Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=472
CPU time spent (ms)=3090
Physical memory (bytes) snapshot=296919040
Virtual memory (bytes) snapshot=4121432064
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=191
File Output Format Counters
  Bytes Written=32

The standard deviation is: 2.616286141418273
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /assignments/wordSD.txt /assignments/output/WSDC
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop fs -ls /assignments/output/WSDC
18/05/13 21:35:05 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 2018-05-13 21:33 /assignments/output/WSDC/ SUCCESS
-rw-r--r-- 1 acadgild supergroup 32 2018-05-13 21:33 /assignments/output/WSDC/part-r-000000
[acadgild@localhost mapreduce]$ hadoop fs -cat /assignments/output/WSDC/part-r-000000
18/05/13 21:36:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
count 31
length 160
square 1038
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$
```