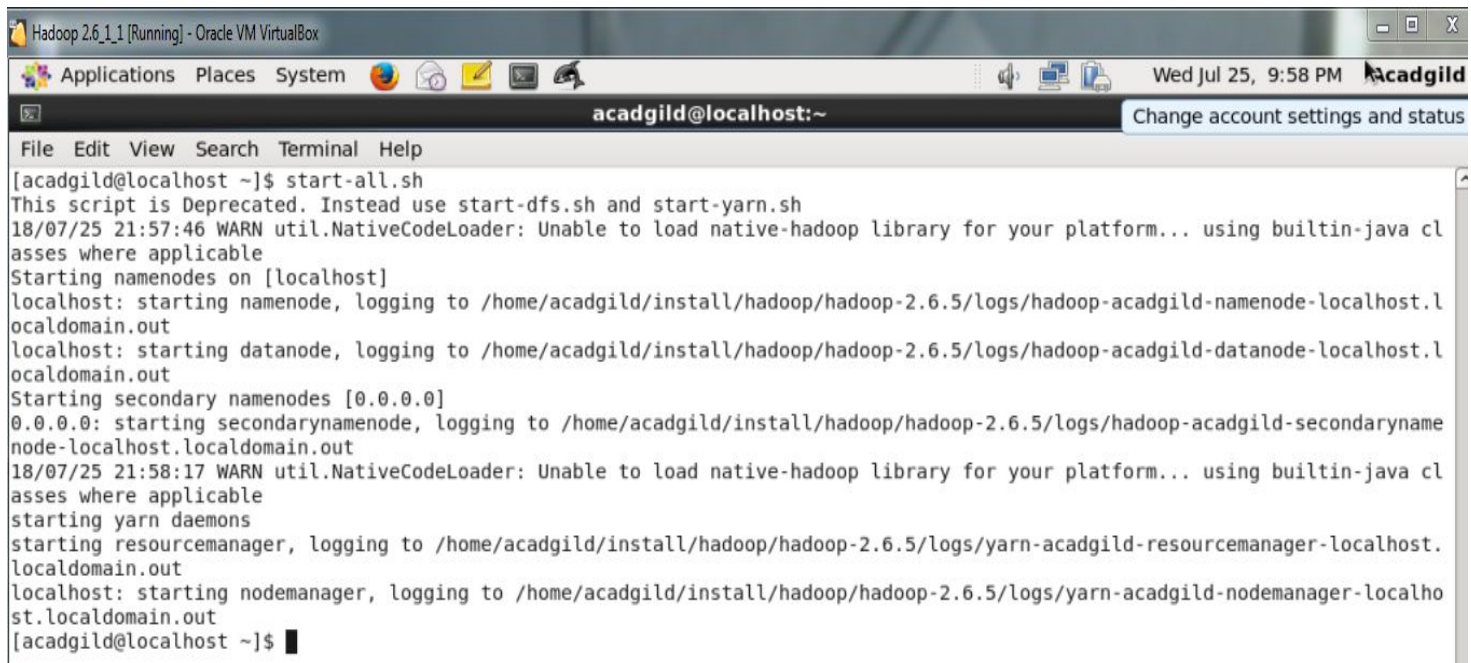


Task 1: Execute WordMedian, WordMean, WordStandardDeviation programs using hadoop-mapreduce-examples-2.9.0.jar file present in your Acadgild VM.

1. Start hadoop daemons in VM, using start-all.sh script.

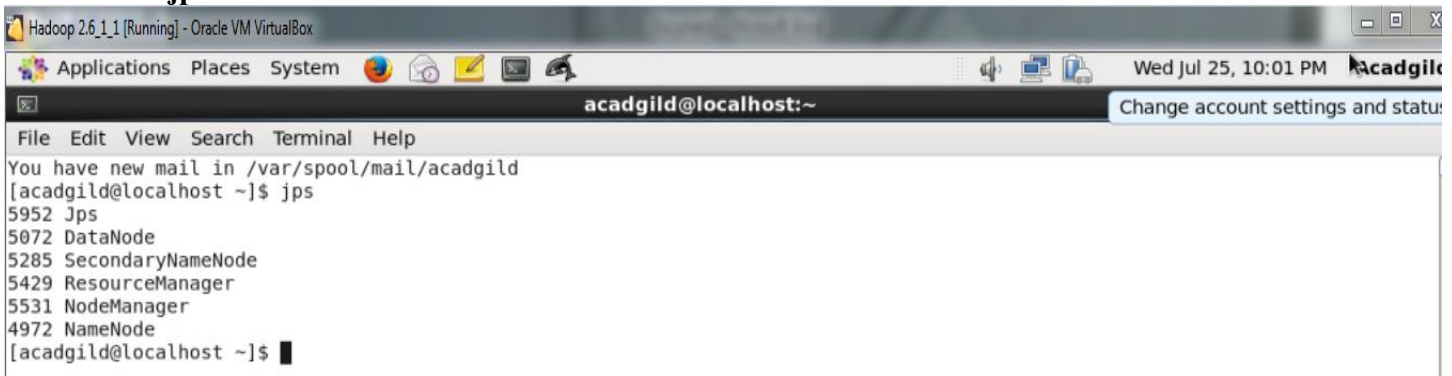
Ex: **\$start-all.sh**



```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
18/07/25 21:57:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-namenode-localhost.l
ocaldomain.out
localhost: starting datanode, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-datanode-localhost.l
ocaldomain.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/hadoop-acadgild-secondaryname
node-localhost.localdomain.out
18/07/25 21:58:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
starting yarn daemons
starting resourcecemanager, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-resourcemanager-localhost.
localdomain.out
localhost: starting nodemanager, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-nodemanager-localho
st.localdomain.out
[acadgild@localhost ~]$
```

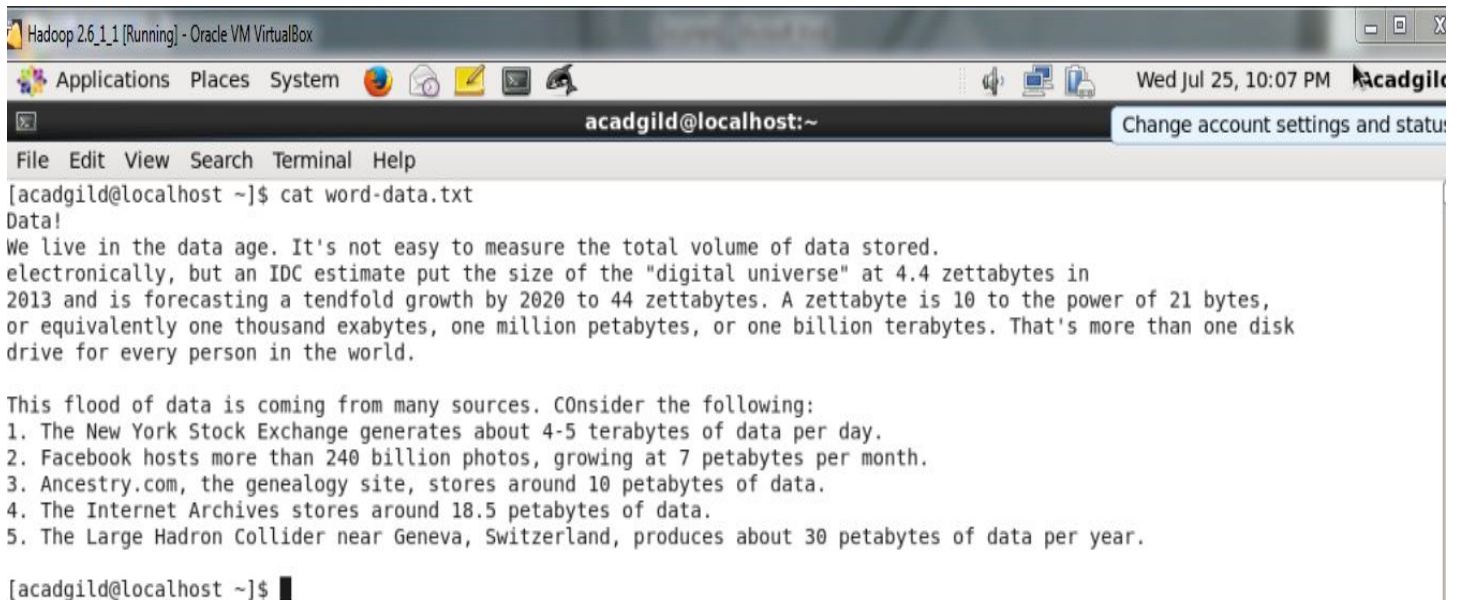
2. Now checks what are all the daemons running in hadoop. This can be done by JPS command.

Ex: **\$jps**



```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System
acadgild@localhost:~
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ jps
5952 Jps
5072 DataNode
5285 SecondaryNameNode
5429 ResourceManager
5531 NodeManager
4972 NameNode
[acadgild@localhost ~]$
```

3. Now create a text file and add some data into that text file.
Ex: **\$cat >> word-data.txt <EOF> =====>** Here <EOF> is and end of file, this can be supplied by Ctrl+d.
4. After adding the data verify contents of the data present in the file **word-data.txt**.
Ex: **\$cat word-data.txt**



Hadoop 2.6.1.1 [Running] - Oracle VM VirtualBox

Applications Places System Wed Jul 25, 10:07 PM acadgild

acadgild@localhost:~

File Edit View Search Terminal Help

```
[acadgild@localhost ~]$ cat word-data.txt
Data!
We live in the data age. It's not easy to measure the total volume of data stored.
electronically, but an IDC estimate put the size of the "digital universe" at 4.4 zettabytes in
2013 and is forecasting a tendfold growth by 2020 to 44 zettabytes. A zettabyte is 10 to the power of 21 bytes,
or equivalently one thousand exabytes, one million petabytes, or one billion terabytes. That's more than one disk
drive for every person in the world.

This flood of data is coming from many sources. Consider the following:
1. The New York Stock Exchange generates about 4-5 terabytes of data per day.
2. Facebook hosts more than 240 billion photos, growing at 7 petabytes per month.
3. Ancestry.com, the genealogy site, stores around 10 petabytes of data.
4. The Internet Archives stores around 18.5 petabytes of data.
5. The Large Hadron Collider near Geneva, Switzerland, produces about 30 petabytes of data per year.
```

5. From the above picture, we can see that the file word-data.txt contains some words present.

6. Now copy this file into hadoop.

Ex: **\$hadoop fs -put word-data.txt /word-data.txt**



Hadoop 2.6.1.1 [Running] - Oracle VM VirtualBox

Applications Places System Wed Jul 25, 10:13 PM acadgild

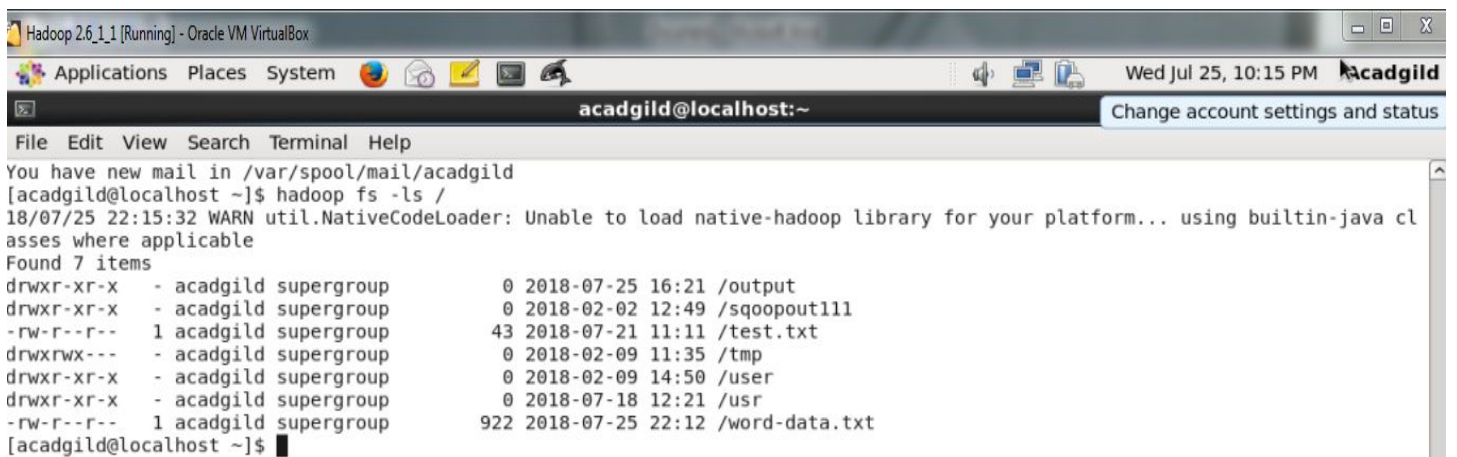
acadgild@localhost:~

File Edit View Search Terminal Help

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -put word-data.txt /word-data.txt
18/07/25 22:12:37 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
[acadgild@localhost ~]$
```

7. We can see whether the file is present in hadoop or not using -ls option.

Ex: **\$hadoop fs -ls /**



Hadoop 2.6.1.1 [Running] - Oracle VM VirtualBox

Applications Places System Wed Jul 25, 10:15 PM acadgild

acadgild@localhost:~

File Edit View Search Terminal Help

```
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -ls /
18/07/25 22:15:32 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 7 items
drwxr-xr-x - acadgild supergroup 0 2018-07-25 16:21 /output
drwxr-xr-x - acadgild supergroup 0 2018-02-02 12:49 /sqoopout111
-rw-r--r-- 1 acadgild supergroup 43 2018-07-21 11:11 /test.txt
drwxrwx--- - acadgild supergroup 0 2018-02-09 11:35 /tmp
drwxr-xr-x - acadgild supergroup 0 2018-02-09 14:50 /user
drwxr-xr-x - acadgild supergroup 0 2018-07-18 12:21 /usr
-rw-r--r-- 1 acadgild supergroup 922 2018-07-25 22:12 /word-data.txt
[acadgild@localhost ~]$
```

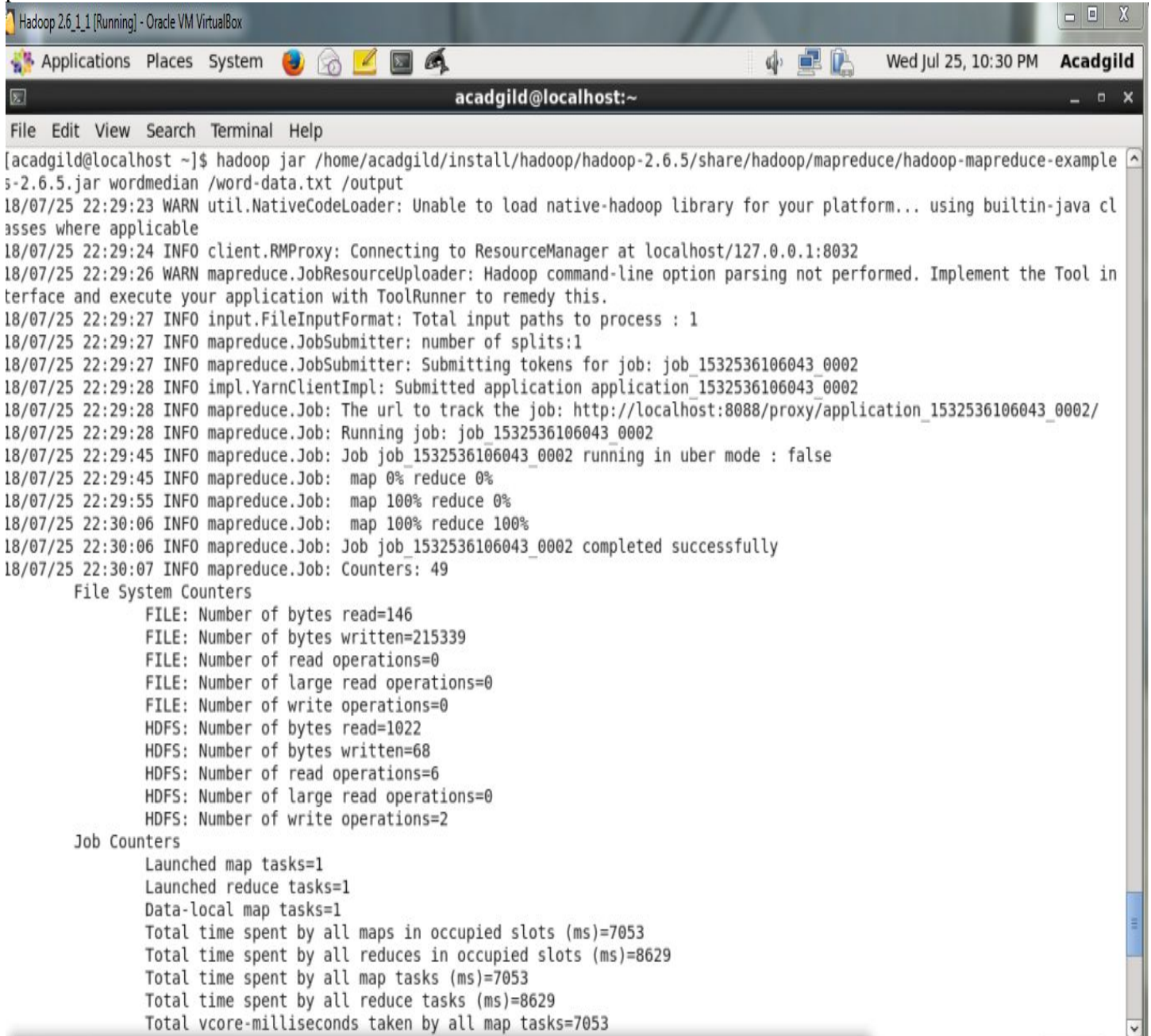
8. Yes. The file word-data.txt was copied. We can observe the last line **/word-data.txt**.

1. WordMedian: A map/reduce program that counts the median length of the words in the input file.

1. Run the jar file (**hadoop-mapreduce-examples-2.6.5.jar**) which contains the Word Median class on the file (**word-data.txt**).

Ex: **\$hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.7/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmedian /word-data.txt /output**

Here **/output** is the hadoop directory, which contains the SUCCESS/FAILURE status of the mapreduce process.

A screenshot of a terminal window titled "Hadoop 2.6_1.1 [Running] - Oracle VM VirtualBox". The terminal shows the execution of the command: `hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmedian /word-data.txt /output`. The output includes various log messages from Hadoop, such as warnings about native code loading and information about job submission and progress. The job is identified as `job_1532536106043_0002`. The terminal also displays "File System Counters" and "Job Counters" at the bottom, providing detailed statistics on the execution. The window's title bar indicates the date and time as "Wed Jul 25, 10:30 PM" and the user as "Acadgild".

```
[acadgild@localhost ~]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmedian /word-data.txt /output
18/07/25 22:29:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/25 22:29:24 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/25 22:29:26 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/07/25 22:29:27 INFO input.FileInputFormat: Total input paths to process : 1
18/07/25 22:29:27 INFO mapreduce.JobSubmitter: number of splits:1
18/07/25 22:29:27 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1532536106043_0002
18/07/25 22:29:28 INFO impl.YarnClientImpl: Submitted application application_1532536106043_0002
18/07/25 22:29:28 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1532536106043_0002/
18/07/25 22:29:28 INFO mapreduce.Job: Running job: job_1532536106043_0002
18/07/25 22:29:45 INFO mapreduce.Job: Job job_1532536106043_0002 running in uber mode : false
18/07/25 22:29:45 INFO mapreduce.Job:  map 0% reduce 0%
18/07/25 22:29:55 INFO mapreduce.Job:  map 100% reduce 0%
18/07/25 22:30:06 INFO mapreduce.Job:  map 100% reduce 100%
18/07/25 22:30:06 INFO mapreduce.Job: Job job_1532536106043_0002 completed successfully
18/07/25 22:30:07 INFO mapreduce.Job: Counters: 49
    File System Counters
      FILE: Number of bytes read=146
      FILE: Number of bytes written=215339
      FILE: Number of read operations=0
      FILE: Number of large read operations=0
      FILE: Number of write operations=0
      HDFS: Number of bytes read=1022
      HDFS: Number of bytes written=68
      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)=7053
      Total time spent by all reduces in occupied slots (ms)=8629
      Total time spent by all map tasks (ms)=7053
      Total time spent by all reduce tasks (ms)=8629
      Total vcore-milliseconds taken by all map tasks=7053
```



```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System
acadgild@localhost:~
File Edit View Search Terminal Help
Total time spent by all reduce tasks (ms)=8629
Total vcore-milliseconds taken by all map tasks=7053
Total vcore-milliseconds taken by all reduce tasks=8629
Total megabyte-milliseconds taken by all map tasks=7222272
Total megabyte-milliseconds taken by all reduce tasks=8836096
Map-Reduce Framework
  Map input records=14
  Map output records=157
  Map output bytes=1256
  Map output materialized bytes=146
  Input split bytes=100
  Combine input records=157
  Combine output records=14
  Reduce input groups=14
  Reduce shuffle bytes=146
  Reduce input records=14
  Reduce output records=14
  Spilled Records=28
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=205
  CPU time spent (ms)=2260
  Physical memory (bytes) snapshot=299110400
  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=922
File Output Format Counters
  Bytes Written=68
The median is: 4
[acadgild@localhost ~]$
```

2. We can observe that in the last line The median is: 4.
3. The process status and output of the process can be verified in the directory **/output**.

Ex: **\$hadoop fs -ls /output**

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop fs -ls /output
18/07/25 22:35:20 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 2018-07-25 22:30 /output/_SUCCESS
-rw-r--r--  1 acadgild supergroup        68 2018-07-25 22:30 /output/part-r-000000
[acadgild@localhost ~]$
```

4. You can see that the contents of the **/output/_SUCCESS** states that the process is success.
5. And the actual result of the wordmedian is present in the file **/output/part-r-000000** file. Lets open the file for the result:

Ex: **\$hadoop fs -cat /output/part-r-000000**

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Wed Jul 25, 10:39 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /output/part-r-00000
18/07/25 22:38:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
1      3
2     34
3     28
4     24
5     15
6     13
7      8
8     12
9     10
10     4
11     2
12     2
13     1
15     1
[acadgild@localhost ~]$
```

6. The result of the **wordmedian** will be as shown in the picture.
7. Now remove the **/output** directory for next mapreduce process to store its status. Removing of directory can be done by **rm** option.

Ex: **\$hadoop fs -rm -r /output**

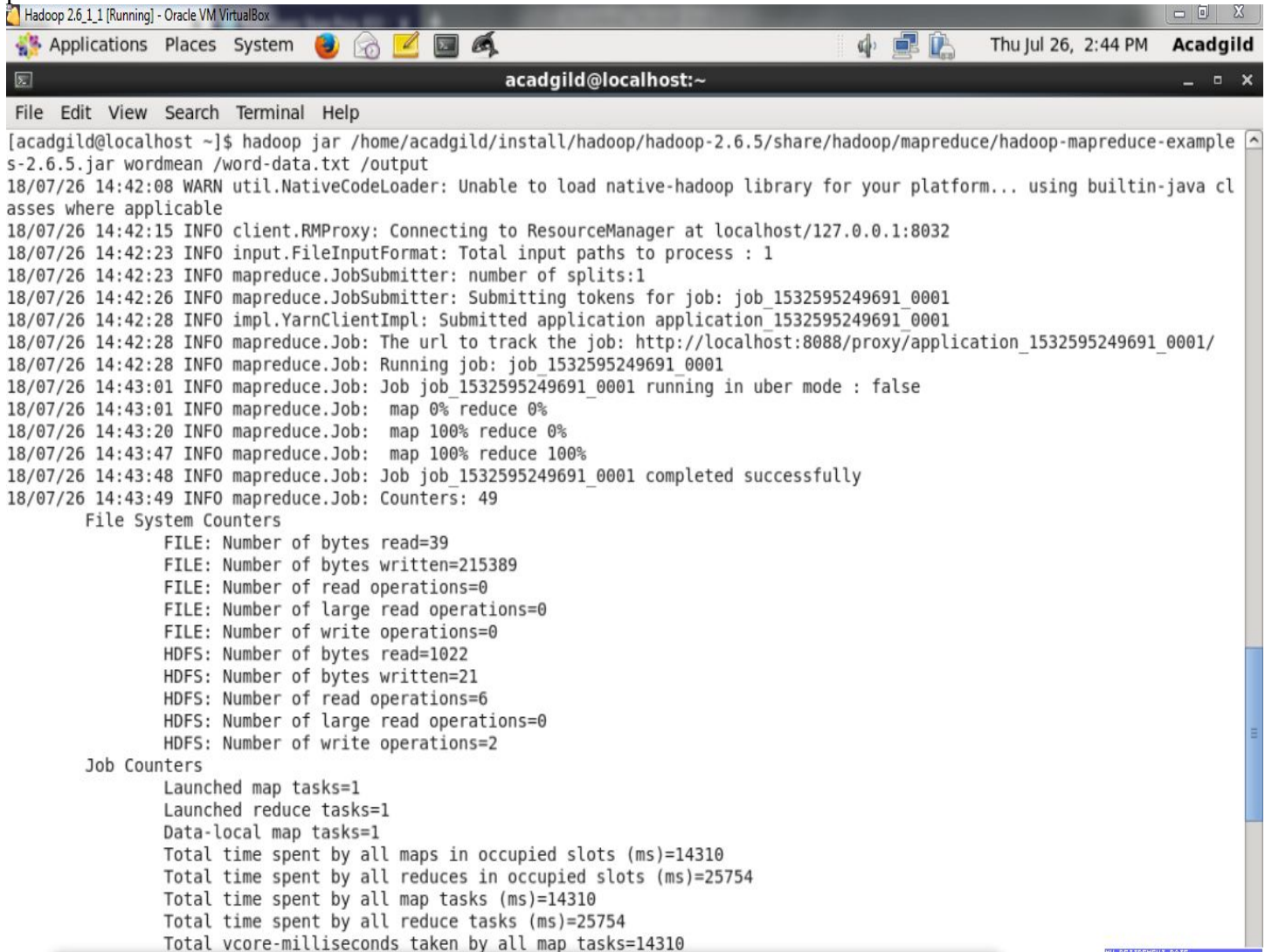
```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Thu Jul 26, 2:33 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop fs -rm -r /output
18/07/26 14:33:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/07/26 14:33:17 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptyier interval =
0 minutes.
Deleted /output
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

2. WordMean: A map/reduce program that counts the average length of the words in the input files.

1. Run the jar file (**hadoop-mapreduce-examples-2.6.5.jar**) which contains the WordMean class on the file (**word-data.txt**).

Ex: **\$hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.7/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmean /word-data.txt /output**

Here **/output** is the hadoop directory, which contains the SUCCESS/FAILURE status of the mapreduce process.



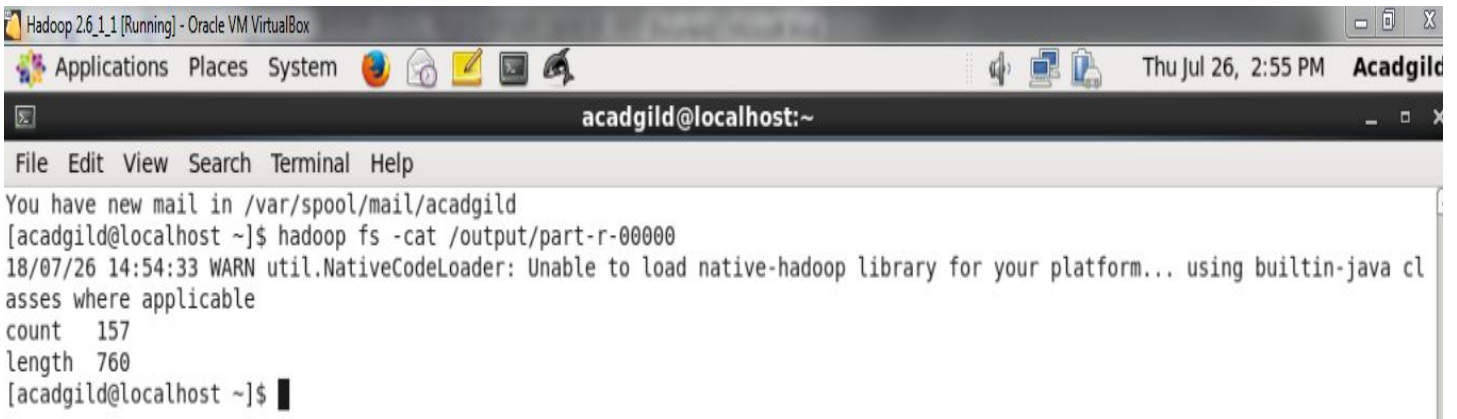
```
[acadgild@localhost ~]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordmean /word-data.txt /output
18/07/26 14:42:08 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/26 14:42:15 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/26 14:42:23 INFO input.FileInputFormat: Total input paths to process : 1
18/07/26 14:42:23 INFO mapreduce.JobSubmitter: number of splits:1
18/07/26 14:42:26 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1532595249691_0001
18/07/26 14:42:28 INFO impl.YarnClientImpl: Submitted application application_1532595249691_0001
18/07/26 14:42:28 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1532595249691_0001/
18/07/26 14:42:28 INFO mapreduce.Job: Running job: job_1532595249691_0001
18/07/26 14:43:01 INFO mapreduce.Job: Job job_1532595249691_0001 running in uber mode : false
18/07/26 14:43:01 INFO mapreduce.Job: map 0% reduce 0%
18/07/26 14:43:20 INFO mapreduce.Job: map 100% reduce 0%
18/07/26 14:43:47 INFO mapreduce.Job: map 100% reduce 100%
18/07/26 14:43:48 INFO mapreduce.Job: Job job_1532595249691_0001 completed successfully
18/07/26 14:43:49 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=39
    FILE: Number of bytes written=215389
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1022
    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)=14310
    Total time spent by all reduces in occupied slots (ms)=25754
    Total time spent by all map tasks (ms)=14310
    Total time spent by all reduce tasks (ms)=25754
    Total vcore-milliseconds taken by all map tasks=14310
```

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Thu Jul 26, 2:48 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
Total vcore-milliseconds taken by all map tasks=14310
Total vcore-milliseconds taken by all reduce tasks=25754
Total megabyte-milliseconds taken by all map tasks=14653440
Total megabyte-milliseconds taken by all reduce tasks=26372096
Map-Reduce Framework
  Map input records=14
  Map output records=314
  Map output bytes=4553
  Map output materialized bytes=39
  Input split bytes=100
  Combine input records=314
  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)=225
  CPU time spent (ms)=3000
  Physical memory (bytes) snapshot=297963520
  Virtual memory (bytes) snapshot=4118228992
  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=922
File Output Format Counters
  Bytes Written=21
The mean is: 4.840764331210191
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

2. We can observe that in the last line The mean is: 4.840764331210191.
3. The process status and output of the process can be verified in the directory **/output**.
Ex: **\$hadoop fs -ls /output**

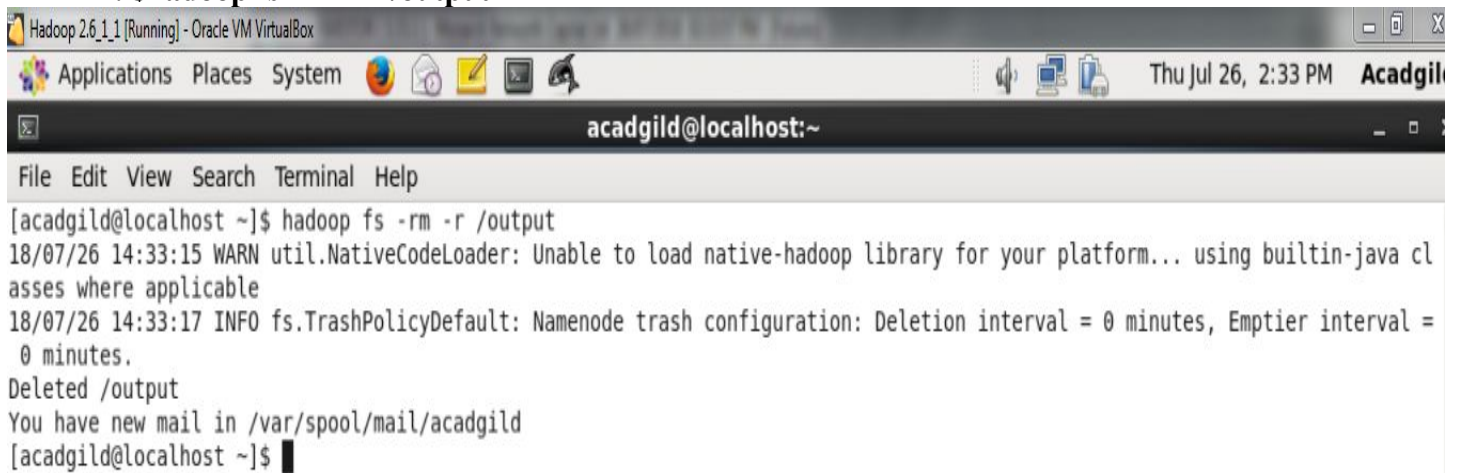
```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Wed Jul 25, 10:35 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop fs -ls /output
18/07/25 22:35:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2018-07-25 22:30 /output/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 68 2018-07-25 22:30 /output/part-r-00000
[acadgild@localhost ~]$
```

4. You can see that the contents of the **/output/_SUCCESS** states that the process is success.
5. And the actual result of the wordmedian is present in the file **/output/part-r-00000** file. Lets open the file for the result:
Ex: **\$hadoop fs -cat /output/part-r-00000**



```
Hadoop 2.6_1.1 [Running] - Oracle VM VirtualBox
Applications Places System
acadmild@localhost:~
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadmild
[acadmild@localhost ~]$ hadoop fs -cat /output/part-r-00000
18/07/26 14:54:33 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
count 157
length 760
[acadmild@localhost ~]$
```

6. The result of the **wordmean** will be as shown in the picture. With **count 157** and **length 760**.
7. Now remove the **/output** directory for next mapreduce process to store its status. Removing of directory can be done by **rm** option.
Ex: **\$hadoop fs -rm -r /output**



```
Hadoop 2.6_1.1 [Running] - Oracle VM VirtualBox
Applications Places System
acadmild@localhost:~
File Edit View Search Terminal Help
[acadmild@localhost ~]$ hadoop fs -rm -r /output
18/07/26 14:33:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/07/26 14:33:17 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval =
0 minutes.
Deleted /output
You have new mail in /var/spool/mail/acadmild
[acadmild@localhost ~]$
```

3. **WordStandardDeviation:** A map/reduce program that counts the standard deviation of the length of the words in the input files.

1. Run the jar file (**hadoop-mapreduce-examples-2.6.5.jar**) which contains the WordStandardDeviation class on the file (**word-data.txt**).

Ex: **\$hadoop jar /home/acadmild/install/hadoop/hadoop-2.6.7/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /word-data.txt /output**

Here **/output** is the hadoop directory, which contains the SUCCESS/FAILURE status of the mapreduce process.

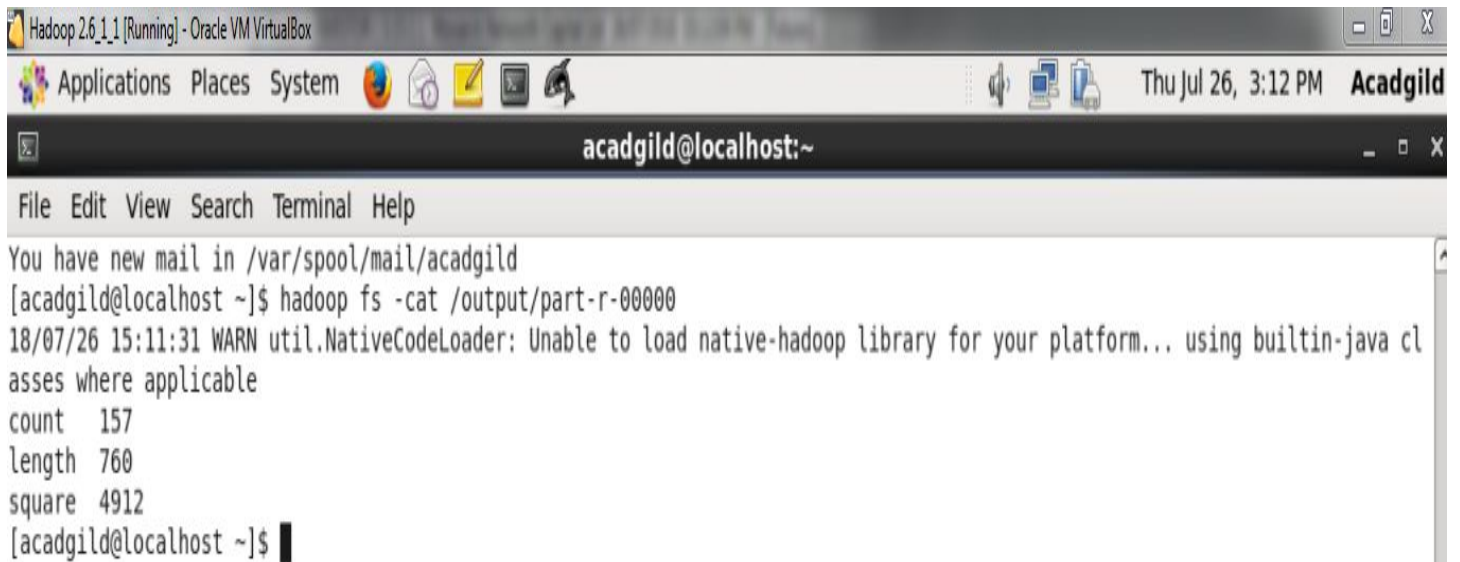

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Thu Jul 26, 3:05 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /word-data.txt /output
18/07/26 15:04:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/26 15:04:20 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/26 15:04:22 INFO input.FileInputFormat: Total input paths to process : 1
18/07/26 15:04:22 INFO mapreduce.JobSubmitter: number of splits:1
18/07/26 15:04:23 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1532595249691_0002
18/07/26 15:04:23 INFO impl.YarnClientImpl: Submitted application application_1532595249691_0002
18/07/26 15:04:24 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1532595249691_0002/
18/07/26 15:04:24 INFO mapreduce.Job: Running job: job_1532595249691_0002
18/07/26 15:04:38 INFO mapreduce.Job: Job job_1532595249691_0002 running in uber mode : false
18/07/26 15:04:38 INFO mapreduce.Job: map 0% reduce 0%
18/07/26 15:04:48 INFO mapreduce.Job: map 100% reduce 0%
18/07/26 15:05:00 INFO mapreduce.Job: map 100% reduce 100%
18/07/26 15:05:01 INFO mapreduce.Job: Job job_1532595249691_0002 completed successfully
18/07/26 15:05:01 INFO mapreduce.Job: Counters: 49
  File System Counters
    FILE: Number of bytes read=56
    FILE: Number of bytes written=215583
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1022
    HDFS: Number of bytes written=33
    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)=7679
    Total time spent by all reduces in occupied slots (ms)=8664
    Total time spent by all map tasks (ms)=7679
    Total time spent by all reduce tasks (ms)=8664
    Total vcore-milliseconds taken by all map tasks=7679
```

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Thu Jul 26, 3:07 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
Total vcore-milliseconds taken by all map tasks=7679
Total vcore-milliseconds taken by all reduce tasks=8664
Total megabyte-milliseconds taken by all map tasks=7863296
Total megabyte-milliseconds taken by all reduce tasks=8871936
Map-Reduce Framework
  Map input records=14
  Map output records=471
  Map output bytes=6908
  Map output materialized bytes=56
  Input split bytes=100
  Combine input records=471
  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)=190
  CPU time spent (ms)=2260
  Physical memory (bytes) snapshot=298254336
  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=922
File Output Format Counters
  Bytes Written=33
The standard deviation is: 2.802431960548696
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

2. We can observe that in the last line The standard deviation is: 2.802431960548696.
3. The process status and output of the process can be verified in the directory **/output**.
Ex: **\$hadoop fs -ls /output**

```
Hadoop 2.6.1_1 [Running] - Oracle VM VirtualBox
Applications Places System Wed Jul 25, 10:35 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ hadoop fs -ls /output
18/07/25 22:35:20 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 2018-07-25 22:30 /output/_SUCCESS
-rw-r--r-- 1 acadgild supergroup 68 2018-07-25 22:30 /output/part-r-00000
[acadgild@localhost ~]$
```

4. You can see that the contents of the **/output/_SUCCESS** states that the process is success.
5. And the actual result of the wordmedian is present in the file **/output/part-r-00000** file. Lets open the file for the result:
Ex: **\$hadoop fs -cat /output/part-r-00000**



The screenshot shows a terminal window titled 'Hadoop 2.6_1.1 [Running] - Oracle VM VirtualBox'. The window has a menu bar with 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The terminal output shows a notification about new mail, followed by the execution of the command 'hadoop fs -cat /output/part-r-00000'. The output of this command is a warning message and three lines of statistics: 'count 157', 'length 760', and 'square 4912'. The prompt '[acadgild@localhost ~]\$' is visible at the bottom.

```
Hadoop 2.6_1.1 [Running] - Oracle VM VirtualBox
Applications Places System Thu Jul 26, 3:12 PM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -cat /output/part-r-00000
18/07/26 15:11:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
count 157
length 760
square 4912
[acadgild@localhost ~]$
```

6. The result of the **wordstandarddeviation** will be as shown in the picture. With **count 157** and **length 760** and **square 4912**.