

——-WORDCount.txt File

```
[acadgild@localhost mapreduce]$ hadoop fs -cat  
useracadgild/hadoop/wordcount.txt
```

```
18/09/28 02:44:22 WARN util.NativeCodeLoader: Unable to load native-hadoop  
library for your platform... using builtin-java classes where applicable
```

hi how are you

i am fine

how are you

are you ok

You have new mail in *varspool/mail/acadgild*

——-Wordcount Program

```
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-  
2.6.5.jar wordcount useracadgild/hadoop/wordcount.txt  
useracadgild/hadoop/out/output.txt
```

```
[acadgild@localhost mapreduce]$ hadoop fs -cat  
useracadgild/hadoop/out/output.txt/part-r-00000
```

```
18/09/27 16:38:27 WARN util.NativeCodeLoader: Unable to load native-hadoop  
library for your platform... using builtin-java classes where applicable
```

am 1

are 3

fine 1

hi 1

how 2

i 1

ok 1

you 3

—Word Median Program

Input :

hi how are you

i am fine

how are you

are you ok

Expected OUtput :

Calculates the Key value Pair

Key - length of the word

value - number of word with the length

1 1

2 3

3 8

4 1

Median = $1+3+8+1 = 13 / 4$ (count) = 3

```
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian useracadgild/hadoop/wordcount.txt  
useracadgild/hadoop/out/median
```

18/09/28 02:31:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/09/28 02:31:44 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/09/28 02:31:46 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.

18/09/28 02:31:47 INFO input.FileInputFormat: Total input paths to process : 1

18/09/28 02:31:48 INFO mapreduce.JobSubmitter: number of splits:1

18/09/28 02:31:48 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1538032559808_0002

18/09/28 02:31:49 INFO impl.YarnClientImpl: Submitted application application_1538032559808_0002

18/09/28 02:31:50 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1538032559808_0002/

18/09/28 02:31:50 INFO mapreduce.Job: Running job: job_1538032559808_0002

18/09/28 02:32:08 INFO mapreduce.Job: Job job_1538032559808_0002

running in uber mode : false

18/09/28 02:32:08 INFO mapreduce.Job: map 0% reduce 0%

18/09/28 02:32:22 INFO mapreduce.Job: map 100% reduce 0%

18/09/28 02:32:35 INFO mapreduce.Job: map 100% reduce 100%

18/09/28 02:32:36 INFO mapreduce.Job: Job job_1538032559808_0002
completed successfully

18/09/28 02:32:37 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=46

FILE: Number of bytes written=215231

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=171

HDFS: Number of bytes written=16

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)=10703

Total time spent by all reduces in occupied slots (ms)=11126

Total time spent by all map tasks (ms)=10703

Total time spent by all reduce tasks (ms)=11126

Total vcore-milliseconds taken by all map tasks=10703

Total vcore-milliseconds taken by all reduce tasks=11126

Total megabyte-milliseconds taken by all map tasks=10959872

Total megabyte-milliseconds taken by all reduce tasks=11393024

MapReduce Framework

Map input records=4

Map output records=13

Map output bytes=104

Map output materialized bytes=46

Input split bytes=121

Combine input records=13

Combine output records=4

Reduce input groups=4

Reduce shuffle bytes=46

Reduce input records=4

Reduce output records=4

Spilled Records=8

Shuffled Maps =1

Failed Shuffles=0

Merged Map outputs=1

GC time elapsed (ms)=316

CPU time spent (ms)=2920

Physical memory (bytes) snapshot=305090560

Virtual memory (bytes) snapshot=4118216704

Total committed heap usage (bytes)=202379264

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=50

File Output Format Counters

Bytes Written=16

The median is: 3

You have new mail in *varspool/mail/acadgild*

```
[acadgild@localhost mapreduce]$ hadoop fs -ls useracadgild/hadoop/out/median
```

```
18/09/28 02:32:56 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 2018-09-28 02:32
useracadgild/hadoop/out/median/_SUCCESS
```

```
-rw-r—r— 1 acadgild supergroup 16 2018-09-28 02:32
useracadgild/hadoop/out/median/part-r-00000
```

```
[acadgild@localhost mapreduce]$ hadoop fs -cat
useracadgild/hadoop/out/median/part-r-00000
```

```
18/09/28 02:33:35 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
```

```
1 1
```

```
2 3
```

```
3 8
```


4 1

You have new mail in *varspool/mail/acadgild*

——-WORDMEAN——-

Input :

hi how are you

i am fine

how are you

are you ok

Expected OUtput :

Calculates the Key value Pair

Key - length of the word

value - number of word with the length

23

38

4 1

Mean = $3+8 = 11 / 4$ (count) = 2.69

```
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean useracadgild/hadoop/wordcount.txt  
useracadgild/hadoop/out/mean
```

18/09/28 05:01:05 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/09/28 05:01:07 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/09/28 05:01:11 INFO input.FileInputFormat: Total input paths to process : 1

18/09/28 05:01:11 INFO mapreduce.JobSubmitter: number of splits:1

18/09/28 05:01:12 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1538032559808_0005

18/09/28 05:01:13 INFO impl.YarnClientImpl: Submitted application application_1538032559808_0005

18/09/28 05:01:13 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1538032559808_0005/

18/09/28 05:01:13 INFO mapreduce.Job: Running job: job_1538032559808_0005

18/09/28 05:01:34 INFO mapreduce.Job: Job job_1538032559808_0005 running in uber mode : false

18/09/28 05:01:34 INFO mapreduce.Job: map 0% reduce 0%

18/09/28 05:01:49 INFO mapreduce.Job: map 100% reduce 0%

18/09/28 05:02:04 INFO mapreduce.Job: map 100% reduce 100%

18/09/28 05:02:05 INFO mapreduce.Job: Job job_1538032559808_0005 completed successfully

18/09/28 05:02:06 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=39

FILE: Number of bytes written=215477

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=171

HDFS: Number of bytes written=19

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)=11340

Total time spent by all reduces in occupied slots (ms)=12757

Total time spent by all map tasks (ms)=11340

Total time spent by all reduce tasks (ms)=12757

Total vcore-milliseconds taken by all map tasks=11340

Total vcore-milliseconds taken by all reduce tasks=12757

Total megabyte-milliseconds taken by all map tasks=11612160

Total megabyte-milliseconds taken by all reduce tasks=13063168

MapReduce Framework

Map input records=4

Map output records=26

Map output bytes=377

Map output materialized bytes=39

Input split bytes=121

Combine input records=26

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)=408

CPU time spent (ms)=3430

Physical memory (bytes) snapshot=304586752

Virtual memory (bytes) snapshot=4118228992

Total committed heap usage (bytes)=202379264

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=50

File Output Format Counters

Bytes Written=19

The mean is: 2.6923076923076925

```
[acadgild@localhost mapreduce]$ hadoop fs -cat  
useracadgild/hadoop/out/mean/part-r-0000018/09/28 05:06:42 WARN  
util.NativeCodeLoader: Unable to load native-hadoop library for your
```

platform... using builtin-java classes where applicable

count 13

length 35

————STANDARD DEVIATION ————

Input :

hi how are you

i am fine

how are you

are you ok

Expected OUtput :

Calculates the Key value Pair

Key - length of the word

value - number of word with the length

1 1

23

38

4 1

Median = $1+3+8+1 = 13/4 = 3$

Mean = $3+8 = 11 / 4$ (count) = 2.69

Standard deviation : 0.7216024245882193

```
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation useracadgild/hadoop/wordcount.txt useracadgild/hadoop/out/deviation
```

18/09/28 05:13:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/09/28 05:13:03 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/09/28 05:13:07 INFO input.FileInputFormat: Total input paths to process : 1

18/09/28 05:13:07 INFO mapreduce.JobSubmitter: number of splits:1

18/09/28 05:13:08 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1538032559808_0006

18/09/28 05:13:09 INFO impl.YarnClientImpl: Submitted application application_1538032559808_0006

18/09/28 05:13:09 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1538032559808_0006/

18/09/28 05:13:09 INFO mapreduce.Job: Running job: job_1538032559808_0006

18/09/28 05:13:30 INFO mapreduce.Job: Job job_1538032559808_0006

running in uber mode : false

18/09/28 05:13:30 INFO mapreduce.Job: map 0% reduce 0%

18/09/28 05:13:44 INFO mapreduce.Job: map 100% reduce 0%

18/09/28 05:14:00 INFO mapreduce.Job: map 100% reduce 100%

18/09/28 05:14:01 INFO mapreduce.Job: Job job_1538032559808_0006
completed successfully

18/09/28 05:14:02 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=56

FILE: Number of bytes written=215681

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=171

HDFS: Number of bytes written=30

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)=11304

Total time spent by all reduces in occupied slots (ms)=12972

Total time spent by all map tasks (ms)=11304

Total time spent by all reduce tasks (ms)=12972

Total vcore-milliseconds taken by all map tasks=11304

Total vcore-milliseconds taken by all reduce tasks=12972

Total megabyte-milliseconds taken by all map tasks=11575296

Total megabyte-milliseconds taken by all reduce tasks=13283328

MapReduce Framework

Map input records=4

Map output records=39

Map output bytes=572

Map output materialized bytes=56

Input split bytes=121

Combine input records=39

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)=371

CPU time spent (ms)=3280

Physical memory (bytes) snapshot=306839552

Virtual memory (bytes) snapshot=4117938176

Total committed heap usage (bytes)=202379264

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=50

File Output Format Counters

Bytes Written=30

The standard deviation is: 0.7216024245882193

```
[acadgild@localhost mapreduce]$ hadoop fs -cat  
useracadgild/hadoop/out/deviation/part-r-00000
```

```
18/09/28 05:15:00 WARN util.NativeCodeLoader: Unable to load native-hadoop  
library for your platform... using builtin-java classes where applicable
```

```
count 13
```

```
length 35
```

```
square 101
```

```
You have new mail in varspool/mail/acadgild
```