

Running Word Mean using Map Reduce Programming Paradigm:

```
[acadgild@localhost ~]$ hdfs dfs -ls /hadoopdata1/wordmeanjob/
18/05/15 21:32:22 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r-- 1 acadgild supergroup 104 2018-05-15 08:20 /hadoopdata1/wordmeanjob/wordmeaninput.txt
```

I created a directory wordmeanjob in HDFS and added a file wordmeaninput.txt which serves as input to the map reduce program.

```
[acadgild@localhost ~]$ hdfs dfs -cat /hadoopdata1/wordmeanjob/wordmeaninput.txt
18/05/15 21:32:52 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
HI I am from Siemens
I will become a BigData Analyst
This is my wordmean logic.
it is a simple program.
```

Contents of my input file is given by above screenshot.

```
[acadgild@localhost ~]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmean /hadoopdata1/wordmeanjob/wordmeaninput.txt /hadoopdata1/wordmeanoutput/
18/05/15 21:35:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/15 21:35:37 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/15 21:35:39 INFO input.FileInputFormat: Total input paths to process : 1
18/05/15 21:35:39 INFO mapreduce.JobSubmitter: number of splits:1
18/05/15 21:35:40 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526399321514_0002
18/05/15 21:35:41 INFO impl.YarnClientImpl: Submitted application application_1526399321514_0002
18/05/15 21:35:41 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526399321514_0002/
18/05/15 21:35:41 INFO mapreduce.Job: Running job: job_1526399321514_0002
18/05/15 21:35:59 INFO mapreduce.Job: Job job_1526399321514_0002 running in uber mode : false
18/05/15 21:35:59 INFO mapreduce.Job:  map 0% reduce 0%
18/05/15 21:36:12 INFO mapreduce.Job:  map 100% reduce 0%
18/05/15 21:36:25 INFO mapreduce.Job:  map 100% reduce 100%
18/05/15 21:36:26 INFO mapreduce.Job: Job job_1526399321514_0002 completed successfully
18/05/15 21:36:26 INFO mapreduce.Job: Counters: 49
```

Above screenshot is running wordmean mapreduce job from hadoop-mapreduce-examples-2.6.5.jar which is present in acadgildvm.

```

Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=219
CPU time spent (ms)=2640
Physical memory (bytes) snapshot=300507136
Virtual memory (bytes) snapshot=4123336704
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=104
File Output Format Counters
Bytes Written=19
The mean is: 3.9523809523809526

```

Finally mean is given by map reduce job.

WordMedian with Mapreduce:

```

[acadgild@localhost ~]$ hdfs dfs -put wordmedianinput.txt /hadoopdata1/wordmedianjob/
18/05/15 22:08:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hdfs dfs -ls /hadoopdata1/wordmedianjob/
18/05/15 22:08:45 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--  1 acadgild supergroup          118 2018-05-15 22:08 /hadoopdata1/wordmedianjob/wordmedianinput.txt

```

I created wordmedianjob directory and created wordmedianinput.txt in HDFS.

```

[acadgild@localhost ~]$ hdfs dfs -cat /hadoopdata1/wordmedianjob/wordmedianinput.txt
18/05/15 22:09:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Hi this is Bangalore
Today election result is announced.
Nobody got majority in the election.
we have to wait and see.

```

Contents of the input file are shown in above screenshot.

```
[acadgild@localhost ~]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian
/hadoopdata1/wordmedianjob/wordmedianinput.txt /hadoopdata1/wordmedianoutput
18/05/15 22:11:47 WARN util.NativeCodeLoader: Unable to load native-hadoop library
for your platform... using builtin-java classes where applicable
18/05/15 22:11:49 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/15 22:11:51 WARN mapreduce.JobResourceUploader: Hadoop command-line option
parsing not performed. Implement the Tool interface and execute your application
with ToolRunner to remedy this.
18/05/15 22:11:51 INFO input.FileInputFormat: Total input paths to process : 1
18/05/15 22:11:52 INFO mapreduce.JobSubmitter: number of splits:1
18/05/15 22:11:52 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_15
26399321514_0003
18/05/15 22:11:52 INFO impl.YarnClientImpl: Submitted application application_15
26399321514_0003
18/05/15 22:11:53 INFO mapreduce.Job: The url to track the job: http://localhost
:8088/proxy/application_1526399321514_0003/
18/05/15 22:11:53 INFO mapreduce.Job: Running job: job_1526399321514_0003
18/05/15 22:12:04 INFO mapreduce.Job: Job job_1526399321514_0003 running in uber
mode : false
18/05/15 22:12:04 INFO mapreduce.Job: map 0% reduce 0%
18/05/15 22:12:12 INFO mapreduce.Job: map 100% reduce 0%
```

Above screenshot shows running wordmedian mapreduce job from hadoop mapreduceexamples jar file.

```
Spilled Records=14
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=157
CPU time spent (ms)=1900
Physical memory (bytes) snapshot=300429312
Virtual memory (bytes) snapshot=4118220800
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=118
File Output Format Counters
Bytes Written=28
The median is: 4
```

Finally after running mapreduce job Median value is shown in the above screenshot.

Output generated by map reduce job is stored in a file.output will look like in the below screenshot.

```
[acadgild@localhost ~]$ hdfs dfs -cat /hadoopdata1/wordmeanoutput/part-r-00000
18/05/15 22:19:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
count      21
length     83
_
```

WordStandardDeviation using Map Reduce:

Input file used in the wordmedian problem is used here.

Running mapreduce for WrodStandardDeviation looks like below screenshot:

```
[acadgild@localhost ~]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordstandarddeviation /hadoopdata1/wordmedianjob/wordmedianinput.txt /hadoopdata1/wordsdo
output
18/05/15 22:36:48 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/15 22:36:50 INFO client.RMPProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/15 22:36:52 INFO input.FileInputFormat: Total input paths to process : 1
18/05/15 22:36:52 INFO mapreduce.JobSubmitter: number of splits:1
18/05/15 22:36:53 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526399321514_0004
18/05/15 22:36:53 INFO impl.YarnClientImpl: Submitted application application_1526399321514_0004
18/05/15 22:36:53 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526399321514_0004/
18/05/15 22:36:53 INFO mapreduce.Job: Running job: job_1526399321514_0004
18/05/15 22:37:05 INFO mapreduce.Job: Job job_1526399321514_0004 running in uber mode : false
18/05/15 22:37:05 INFO mapreduce.Job:  map 0% reduce 0%
18/05/15 22:37:14 INFO mapreduce.Job:  map 100% reduce 0%
18/05/15 22:37:25 INFO mapreduce.Job:  map 100% reduce 100%
18/05/15 22:37:25 INFO mapreduce.Job: Job job_1526399321514_0004 completed successfully
```

```

        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)=181
        CPU time spent (ms)=2140
        Physical memory (bytes) snapshot=297521152
        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=118
    File Output Format Counters
        Bytes Written=30
The standard deviation is: 2.5350155531821277

```

Output is given by Standard deviation Map Reduce Job.

Contents of the output file is like below:

```

[acadgild@localhost ~]$ hdfs dfs -cat /hadoopdata1/wordsdoutput/part-r-00000
18/05/15 22:41:23 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
count    21
length   97
square   583

```