Assignmnet_3

In previous assignments I created a file in my HDFS called word-count.txt and added few lines into it.

```
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/hadoop/word-count.txt
18/11/14 01:03:44 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
My Name is Shashank Dixit. I am learning Big Data Hadoop and Spark. My mentor is
 Magesh.
```

I utilized the same file for this assignment as well.

1) Used the suggested jar to calculate wordmedian that gave the output as 4 .

```
[acadgild@localhost sbin]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exam
ples-2.6.5.jar wordmedian /user/acadgild/hadoop/word-count.txt /output
18/11/14 01:11:22 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/11/14 01:11:27 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/11/14 01:11:32 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool in
terface and execute your application with ToolRunner to remedy this.
18/11/14 01:11:36 INFO input.FileInputFormat: Total input paths to process : 1
18/11/14 01:11:36 INFO mapreduce.JobSubmitter: number of splits:1
18/11/14 01:11:38 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1542135100691_0001
18/11/14 01:11:41 INFO impl.YarnClientImpl: Submitted application application_1542135100691_0001
18/11/14 01:11:42 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1542135100691_0001/
18/11/14 01:11:42 INFO mapreduce.Job: Running job: job_1542135100691_0001
^[[B18/11/14 01:13:00 INFO mapreduce.Job: Job job_1542135100691_0001 running in uber mode : false
18/11/14 01:13:00 INFO mapreduce.Job:  map 0% reduce 0%
18/11/14 01:13:41 INFO mapreduce.Job:  map 100% reduce 0%
18/11/14 01:14:28 INFO mapreduce.Job:  map 100% reduce 100%
18/11/14 01:14:34 INFO mapreduce.Job: Job job_1542135100691_0001 completed successfully
18/11/14 01:14:39 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=76
                FILE: Number of bytes written=215243
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=211
                HDFS: Number of bytes written=28
                HDFS: Number of read operations=6
```

```
                       Total Committed Heap Usage (bytes)=176004480
          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=89
          File Output Format Counters
                       Bytes Written=28
18/11/14 01:14:40 INFO mapred.ClientServiceDelegate: Application state is completed. FinalApplicationStatus
cting to job history server
The median is: 4
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost sbin]$ hadoop
```

Also checked the generated output file which gave below output:

```
[acadgild@localhost sbin]$ hadoop fs -cat /output/part-r-00000
18/11/14 01:38:48 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
1       1
2       5
3       2
4       2
6       4
7       1
8       2
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost sbin]$
```

2) Executed the same jar to calculate the wordmean:

```
[acadgild@localhost sbin]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exam
ples-2.6.5.jar wordmean /user/acadgild/hadoop/word-count.txt /output_wordmean
```

Output Shown on screen:

```
18/11/14 01:45:12 INFO mapred.ClientServiceDelegate: Application state is completed. FinalApplicationStatus=SUCCEEDED. Redire
cting to job history server
The mean is: 4.235294117647059
You have new mail in /var/spool/mail/acadgild
```

Output from file

```
[acadgild@localhost sbin]$ hadoop fs -cat /output_wordmean/part-r-00000
18/11/14 01:53:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java c
asses where applicable
count   17
length  72
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost sbin]$
```

3) Executed the same jar to calculate the wordstandarddeviation:

```
[acadgild@localhost sbin]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-exam
ples-2.6.5.jar wordstandarddeviation /user/acadgild/hadoop/word-count.txt /output_sd
```

Output Shown on screen:

```
                        Bytes Read=89
             File Output Format Counters
                    Bytes Written=30
 18/11/14 02:00:23 INFO mapred.ClientServiceDelegate: Application state
 cting to job history server
 The standard deviation is: 2.2629863601966287
 You have new mail in /var/spool/mail/acadgild
 [acadgild@localhost sbin]$ █
```

Output from file:

```
 18/11/14 02:00:23 INFO mapred.ClientServiceDelegate: Application stat
 cting to job history server
 The standard deviation is: 2.2629863601966287
 You have new mail in /var/spool/mail/acadgild
 [acadgild@localhost sbin]$ hadoop fs -cat /output_sd/part-r-00000
 18/11/14 02:01:48 WARN util.NativeCodeLoader: Unable to load native-l
 asses where applicable
 count   17
 length  72
 square  392
 You have new mail in /var/spool/mail/acadgild
 [acadgild@localhost sbin]$ █
```