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

File to execute wordmedian,wordmean, wordstandarddeviation:-

```
[acadgild@localhost mapreduce]$ hadoop fs -cat /hadoopdata/hdfs/file1.txt
18/02/21 16:55:17 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Hello,

Welcome to acadgild!This program is regarding to training on Big Data Hadoop and
 Spark training.

Thank you all!
```

Executing word mean

Executed jar files for word mean using the command as highlighted in the below screenshot and wordmean of the file "*File1.txt*" is **5.26.**

```
acadgild@localhost:~/install/hadoop/hadoop-2.6.5/share/hado

File   Edit   View   Search   Terminal   Help
Usage: wordmean <in> <out>
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar w
ordmean /hadoopdata/hdfs/file1.txt /hadoopdata/hdfs-output
18/02/21 17:19:14 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
18/02/21 17:19:20 INFO client.RMProxy: Connecting to ResourceManager at localhos
t/127.0.0.1:8032
18/02/21 17:19:27 INFO input.FileInputFormat: Total input paths to process : 1
18/02/21 17:19:29 INFO mapreduce.JobSubmitter: number of splits:1
18/02/21 17:19:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_15
19210873612_0004
18/02/21 17:19:32 INFO impl.YarnClientImpl: Submitted application application_15
19210873612_0004
18/02/21 17:19:33 INFO mapreduce.Job: The url to track the job: http://localhost
:8088/proxy/application_1519210873612_0004/
18/02/21 17:19:33 INFO mapreduce.Job: Running job: job_1519210873612_0004
18/02/21 17:20:28 INFO mapreduce.Job: Job job_1519210873612_0004 running in uber
 mode : false
18/02/21 17:20:28 INFO mapreduce.Job:   map 0% reduce 0%
18/02/21 17:21:12 INFO mapreduce.Job:   map 100% reduce 0%
18/02/21 17:21:53 INFO mapreduce.Job:   map 100% reduce 100%
18/02/21 17:21:57 INFO mapreduce.Job: Job job_1519210873612_0004 completed succe
ssfully
18/02/21 17:21:58 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=39
                FILE: Number of bytes written=215445
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=233
                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
```

File   Edit   View   Search   Terminal   Help

```
        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)=39561
                Total time spent by all reduces in occupied slots (ms)=37722
                Total time spent by all map tasks (ms)=39561
                Total time spent by all reduce tasks (ms)=37722
                Total vcore-milliseconds taken by all map tasks=39561
                Total vcore-milliseconds taken by all reduce tasks=37722
                Total megabyte-milliseconds taken by all map tasks=40510464
                Total megabyte-milliseconds taken by all reduce tasks=38627328
        Map-Reduce Framework
                Map input records=5
                Map output records=38
                Map output bytes=551
                Map output materialized bytes=39
                Input split bytes=112
                Combine input records=38
                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)=944
                CPU time spent (ms)=8660
                Physical memory (bytes) snapshot=292208640
                Virtual memory (bytes) snapshot=4120752128
                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=121
        File Output Format Counters
                Bytes Written=20
The mean is: 5.2631578947368425
```

Executing WordMedian:

Executed jar files for word median using the command as highlighted in the below screenshot and wordmean of the file "*File1.txt*" is **5.**

```
                    acadgild@localhost:~/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce       _ □ ☑

File Edit View Search Terminal Help
asses where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost mapreduce]$ hadoop jar hadoop-mapreduce-examples-2.6.5.jar wordmedian /hadoopdata/hdfs/file1.txt /hadoopd
ata/hdfs/results-output
18/02/21 17:26:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/02/21 17:26:36 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/02/21 17:26:41 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool in
terface and execute your application with ToolRunner to remedy this.
18/02/21 17:26:44 INFO input.FileInputFormat: Total input paths to process : 1
18/02/21 17:26:45 INFO mapreduce.JobSubmitter: number of splits:1
18/02/21 17:26:46 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1519210873612_0005
18/02/21 17:26:48 INFO impl.YarnClientImpl: Submitted application application_1519210873612_0005
18/02/21 17:26:49 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1519210873612_0005/
18/02/21 17:26:49 INFO mapreduce.Job: Running job: job_1519210873612_0005
18/02/21 17:27:42 INFO mapreduce.Job: Job job_1519210873612_0005 running in uber mode : false
18/02/21 17:27:42 INFO mapreduce.Job:  map 0% reduce 0%
18/02/21 17:28:27 INFO mapreduce.Job:  map 100% reduce 0%
18/02/21 17:29:10 INFO mapreduce.Job:  map 100% reduce 100%
18/02/21 17:29:12 INFO mapreduce.Job: Job job_1519210873612_0005 completed successfully
18/02/21 17:29:14 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=96
                FILE: Number of bytes written=215311
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=233
                HDFS: Number of bytes written=37
                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)=38992
                Total time spent by all reduces in occupied slots (ms)=38903
                Total time spent by all map tasks (ms)=38992
```

File Edit View Search Terminal Help

```
            Total vcore-milliseconds taken by all map tasks=38992
            Total vcore-milliseconds taken by all reduce tasks=38903
            Total megabyte-milliseconds taken by all map tasks=39927808
            Total megabyte-milliseconds taken by all reduce tasks=39836672
        Map-Reduce Framework
            Map input records=5
            Map output records=19
            Map output bytes=152
            Map output materialized bytes=96
            Input split bytes=112
            Combine input records=19
            Combine output records=9
            Reduce input groups=9
            Reduce shuffle bytes=96
            Reduce input records=9
            Reduce output records=9
            Spilled Records=18
            Shuffled Maps =1
            Failed Shuffles=0
            Merged Map outputs=1
            GC time elapsed (ms)=694
            CPU time spent (ms)=8760
            Physical memory (bytes) snapshot=293756928
            Virtual memory (bytes) snapshot=4117938176
            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=121
        File Output Format Counters
            Bytes Written=37
The median is: 5
```

Executing WordStandardDeviation:

Executed jar files for word standarddeviation using the command as highlighted in the below screenshot and wordmean of the file "*File1.txt*" is **2.93.**

```
                Total vcore-milliseconds taken by all map tasks=57993
                Total vcore-milliseconds taken by all reduce tasks=57034
                Total megabyte-milliseconds taken by all map tasks=59384832
                Total megabyte-milliseconds taken by all reduce tasks=58402816
        Map-Reduce Framework
                Map input records=5
                Map output records=57
                Map output bytes=836
                Map output materialized bytes=56
                Input split bytes=112
                Combine input records=57
                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)=976
                CPU time spent (ms)=9860
                Physical memory (bytes) snapshot=296062976
                Virtual memory (bytes) snapshot=4117938176
                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=121
        File Output Format Counters
                Bytes Written=31
The standard deviation is: 2.935124946020775
```