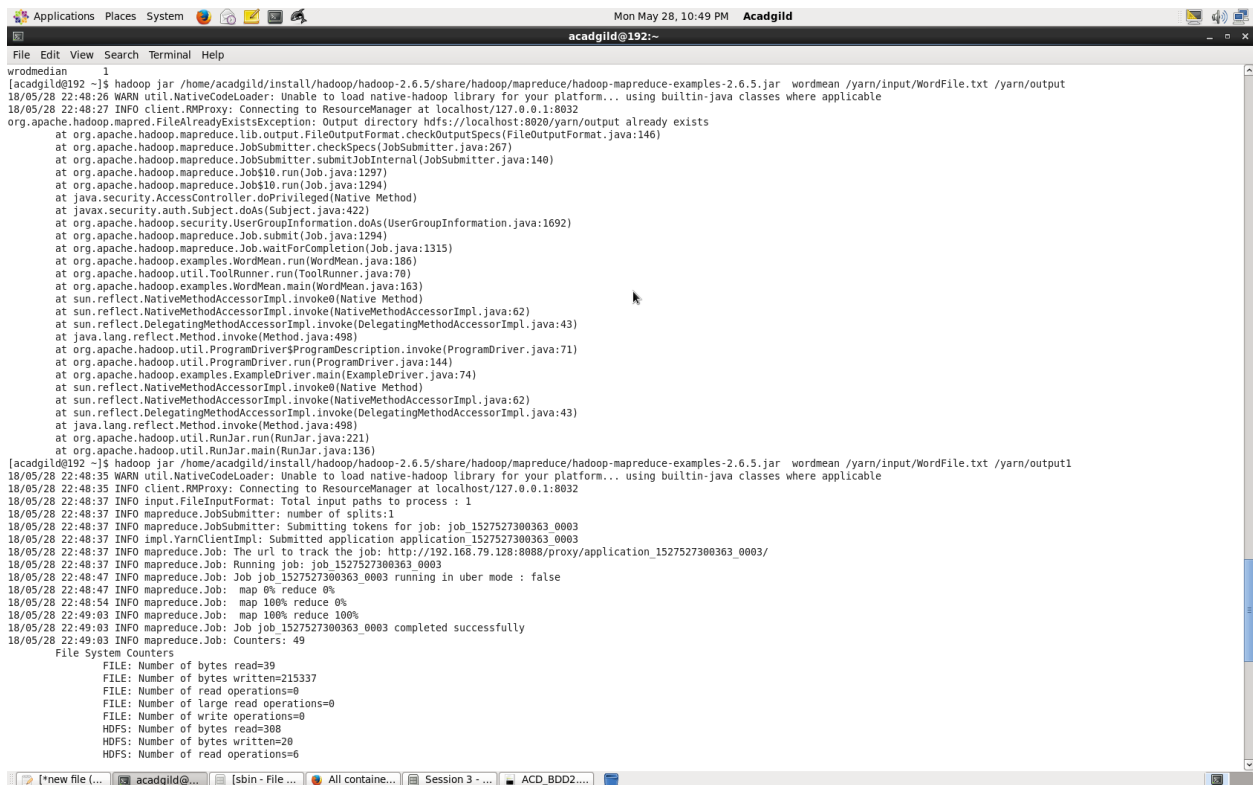## Task:1

Calculate word mean using hadoop-mapreduce-examples-2.6.5.jar

Here I have created one WordFile.txt using vi WordFile.txt from terminal & I have executed job for wordmean from terminal as below

acadgild@192:~

File   Edit   View   Search   Terminal   Help

```
18/05/28 22:49:03 INFO mapreduce.Job: Job job_1527527300363_0003 completed successfully
18/05/28 22:49:03 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=39
                FILE: Number of bytes written=215337
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=308
                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
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=5610
                Total time spent by all reduces in occupied slots (ms)=5663
                Total time spent by all map tasks (ms)=5610
                Total time spent by all reduce tasks (ms)=5663
                Total vcore-milliseconds taken by all map tasks=5610
                Total vcore-milliseconds taken by all reduce tasks=5663
                Total megabyte-milliseconds taken by all map tasks=5744640
                Total megabyte-milliseconds taken by all reduce tasks=5798912
        Map-Reduce Framework
                Map input records=1
                Map output records=68
                Map output bytes=986
                Map output materialized bytes=39
                Input split bytes=110
                Combine input records=68
                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)=193
                CPU time spent (ms)=1620
                Physical memory (bytes) snapshot=328122368
                Virtual memory (bytes) snapshot=4126875648
                Total committed heap usage (bytes)=230424576
        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=198
```
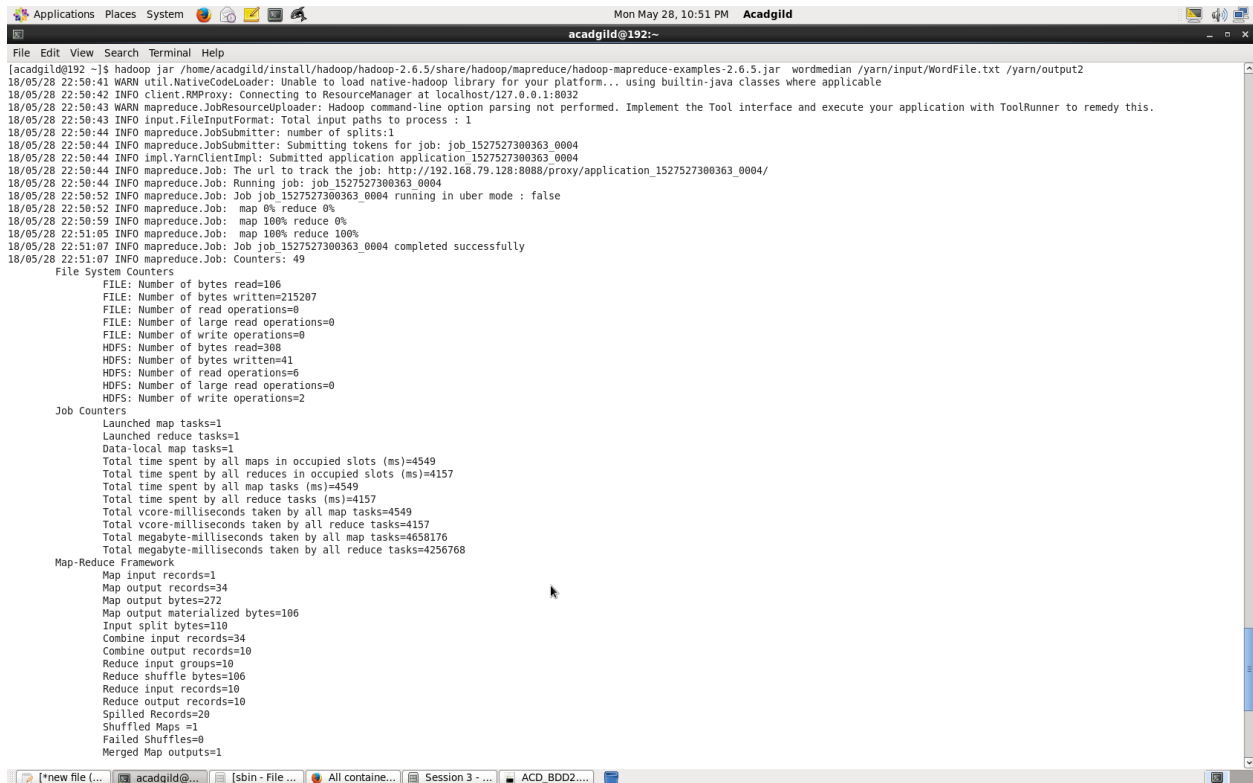
[*new file (... | acadgild@... | [sbin - File ... | All containe... | Session 3 - ... | ACD_BDD2...

Please find the answer for task 1

acadgild@192:~

File   Edit   View   Search   Terminal   Help

```
                HDFS: Number of bytes read=308
                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
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=5610
                Total time spent by all reduces in occupied slots (ms)=5663
                Total time spent by all map tasks (ms)=5610
                Total time spent by all reduce tasks (ms)=5663
                Total vcore-milliseconds taken by all map tasks=5610
                Total vcore-milliseconds taken by all reduce tasks=5663
                Total megabyte-milliseconds taken by all map tasks=5744640
                Total megabyte-milliseconds taken by all reduce tasks=5798912
        Map-Reduce Framework
                Map input records=1
                Map output records=68
                Map output bytes=986
                Map output materialized bytes=39
                Input split bytes=110
                Combine input records=68
                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)=193
                CPU time spent (ms)=1620
                Physical memory (bytes) snapshot=328122368
                Virtual memory (bytes) snapshot=4126875648
                Total committed heap usage (bytes)=230424576
        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=198
        File Output Format Counters
                Bytes Written=20
The mean is: 4.764705882352941
[acadgild@192 ~]$ hadoop fs -cat /yarn/output1/part-r-00000
18/05/28 22:49:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
count   34
length  162
[acadgild@192 ~]$
```

[*new file (... | acadgild@... | [sbin - File ... | All containe... | Session 3 - ... | ACD_BDD2...

Task:2

Calculate word median using hadoop-mapreduce-examples-2.6.5.jar

Please find the answer for Task2

Task:3

Calculate word standard deviation using hadoop-mapreduce-examples-2.6.5.jar

acadgild@192:~

File Edit View Search Terminal Help

```
[acadgild@192 ~]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar  wordstandarddeviation /yarn/input/WordFile.txt /yarn/output3
18/05/28 22:52:38 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/05/28 22:52:39 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/05/28 22:52:40 INFO input.FileInputFormat: Total input paths to process : 1
18/05/28 22:52:40 INFO mapreduce.JobSubmitter: number of splits:1
18/05/28 22:52:40 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1527527300363_0005
18/05/28 22:52:41 INFO impl.YarnClientImpl: Submitted application application_1527527300363_0005
18/05/28 22:52:41 INFO mapreduce.Job: The url to track the job: http://192.168.79.128:8088/proxy/application_1527527300363_0005/
18/05/28 22:52:41 INFO mapreduce.Job: Running job: job_1527527300363_0005
18/05/28 22:52:49 INFO mapreduce.Job: Job job_1527527300363_0005 running in uber mode : false
18/05/28 22:52:49 INFO mapreduce.Job:  map 0% reduce 0%
18/05/28 22:52:54 INFO mapreduce.Job:  map 100% reduce 0%
18/05/28 22:53:00 INFO mapreduce.Job:  map 100% reduce 100%
18/05/28 22:53:01 INFO mapreduce.Job: Job job_1527527300363_0005 completed successfully
18/05/28 22:53:02 INFO mapreduce.Job: Counters: 49
	File System Counters
		FILE: Number of bytes read=56
		FILE: Number of bytes written=215531
		FILE: Number of read operations=0
		FILE: Number of large read operations=0
		FILE: Number of write operations=0
		HDFS: Number of bytes read=308
		HDFS: Number of bytes written=32
		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)=3485
		Total time spent by all reduces in occupied slots (ms)=3886
		Total time spent by all map tasks (ms)=3485
		Total time spent by all reduce tasks (ms)=3886
		Total vcore-milliseconds taken by all map tasks=3485
		Total vcore-milliseconds taken by all reduce tasks=3886
		Total megabyte-milliseconds taken by all map tasks=3568640
		Total megabyte-milliseconds taken by all reduce tasks=3979264
	Map-Reduce Framework
		Map input records=1
		Map output records=102
		Map output bytes=1496
		Map output materialized bytes=56
		Input split bytes=110
		Combine input records=102
		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)=119
```

[*new file (...] acadgild@... [sbin - File ...] All containe... Session 3 - ... ACD_BDD2...

Please find the answer for Task 3

acadgild@192:~

File Edit View Search Terminal Help

```
		HDFS: Number of bytes written=32
		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)=3485
		Total time spent by all reduces in occupied slots (ms)=3886
		Total time spent by all map tasks (ms)=3485
		Total time spent by all reduce tasks (ms)=3886
		Total vcore-milliseconds taken by all map tasks=3485
		Total vcore-milliseconds taken by all reduce tasks=3886
		Total megabyte-milliseconds taken by all map tasks=3568640
		Total megabyte-milliseconds taken by all reduce tasks=3979264
	Map-Reduce Framework
		Map input records=1
		Map output records=102
		Map output bytes=1496
		Map output materialized bytes=56
		Input split bytes=110
		Combine input records=102
		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)=119
		CPU time spent (ms)=1190
		Physical memory (bytes) snapshot=326836224
		Virtual memory (bytes) snapshot=4126875648
		Total committed heap usage (bytes)=230424576
	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=198
	File Output Format Counters
		Bytes Written=32
The standard deviation is: 2.6015699279407034
[acadgild@192 ~]$ hadoop fs -cat /yarn/output3/part-r-00000
18/05/28 22:53:16 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
count	34
length	162
square	1002
[acadgild@192 ~]$ 
```

[*new file (...] acadgild@... [sbin - File ...] All containe... Session 3 - ... ACD_BDD2...