Assignment-3

1) Execution of Word Median program:

The below command is used to execute the jar file.

The generalized command is like below:

Hadoop jar <jar location> <Main class name> <Command line arguments/Input data> <output path>

- **a.** Jar location: It is the location of the jar file where it is placed.
- **b.** Main class name: It is the main class name of the jar file.
- **c.** Command line arguments: Data to be processed has to be provided.
- **d.** Output: The output of the process data will be available in this path.

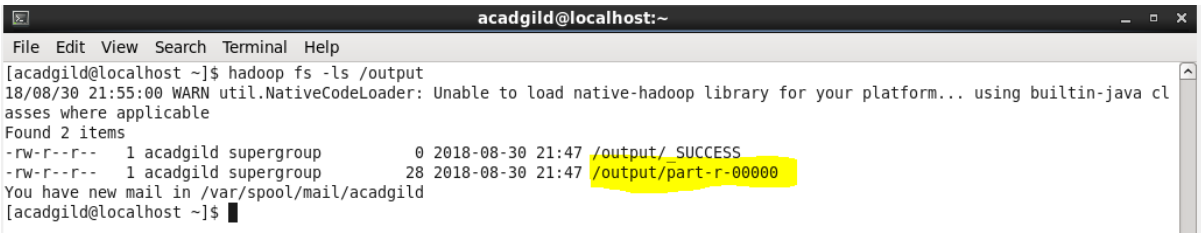
The command used to execute **wordmedian** program is as shown in the below screenshot.

```
[acadgild@localhost ~]$
[acadgild@localhost ~]$ hadoop jar /home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-example
s-2.6.5.jar wordmedian /test.txt /output
```

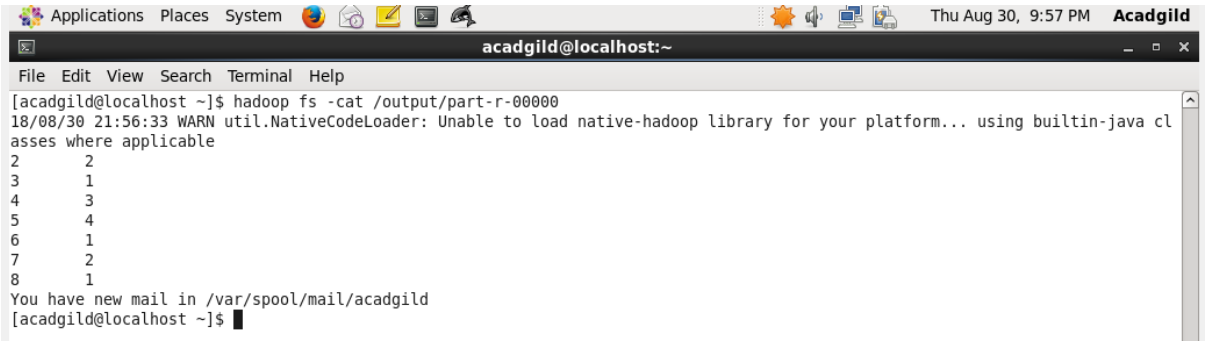
Log displayed on the screen while execution of the program is shown below.

```
18/08/30 21:46:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/08/30 21:46:58 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/08/30 21:46:59 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool in
terface and execute your application with ToolRunner to remedy this.
18/08/30 21:46:59 INFO input.FileInputFormat: Total input paths to process : 1
18/08/30 21:46:59 INFO mapreduce.JobSubmitter: number of splits:1
18/08/30 21:47:00 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1535645471088_0002
18/08/30 21:47:00 INFO impl.YarnClientImpl: Submitted application application_1535645471088_0002
18/08/30 21:47:00 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1535645471088_0002/
18/08/30 21:47:00 INFO mapreduce.Job: Running job: job_1535645471088_0002
18/08/30 21:47:10 INFO mapreduce.Job: Job job_1535645471088_0002 running in uber mode : false
18/08/30 21:47:10 INFO mapreduce.Job:  map 0% reduce 0%
18/08/30 21:47:18 INFO mapreduce.Job:  map 100% reduce 0%
18/08/30 21:47:27 INFO mapreduce.Job:  map 100% reduce 100%
18/08/30 21:47:27 INFO mapreduce.Job: Job job_1535645471088_0002 completed successfully
18/08/30 21:47:27 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=76
                FILE: Number of bytes written=215189
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=176
                HDFS: Number of bytes written=28
                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)=5194
                Total time spent by all reduces in occupied slots (ms)=5466
                Total time spent by all map tasks (ms)=5194
                Total time spent by all reduce tasks (ms)=5466
                Total vcore-milliseconds taken by all map tasks=5194
                Total vcore-milliseconds taken by all reduce tasks=5466
                Total megabyte-milliseconds taken by all map tasks=5318656
acadgild@localhost:~
```

The output of the processed data is stored in a path as highlighted in the below screen shot.



Data in the output file is as shown in the below screenshot.



2) Execution of Word Mean program:

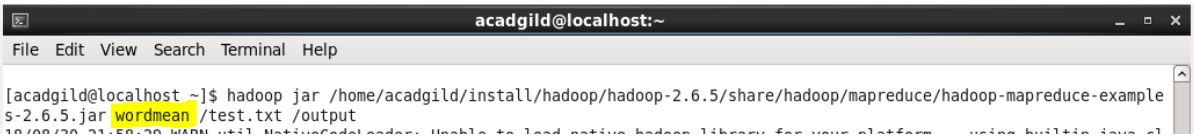
The below command is used to execute the jar file.

The generalized command is like below:

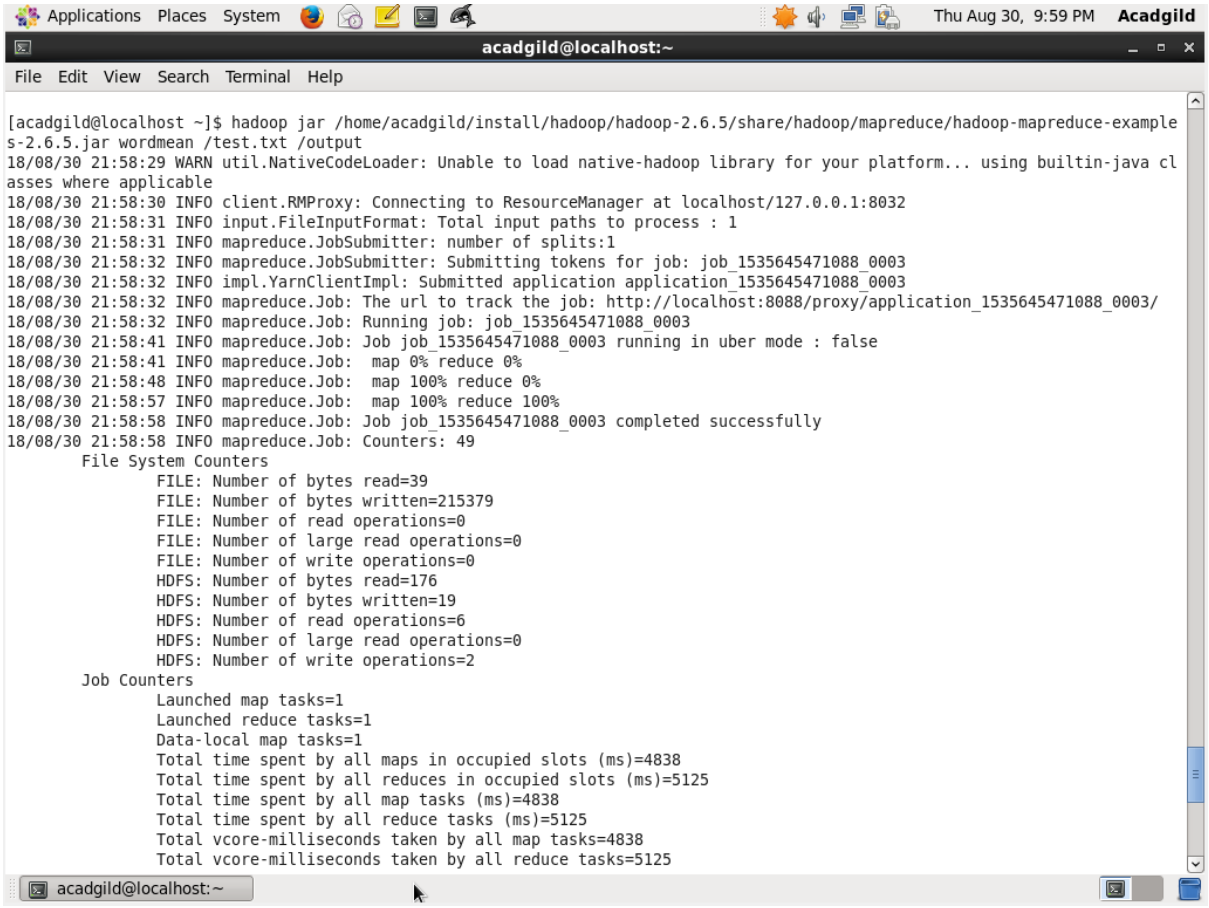
Hadoop jar <jar location> <Main class name> <Command line arguments/Input data> <output path>

a. Jar location: It is the location of the jar file where it is placed.
b. Main class name: It is the main class name of the jar file.
c. Command line arguments: Data to be processed has to be provided.
d. Output: The output of the process data will be available in this path.

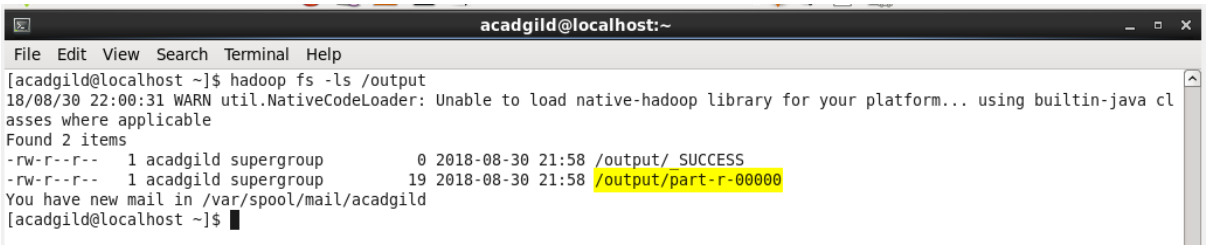
The command used to execute **wordmean** program is as shown in the below screenshot.



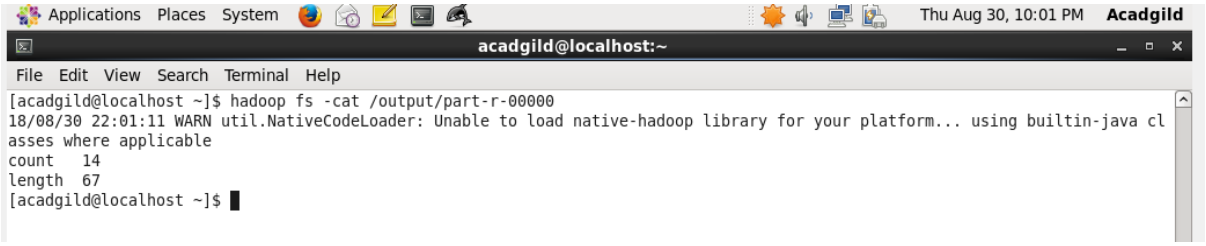
Log displayed on the screen while execution of the program is shown below.



The output of the processed data is stored in a path as highlighted in the below screen shot.



Data in the output file is as shown in the below screenshot.



3) Execution of Word Standard Deviation program:

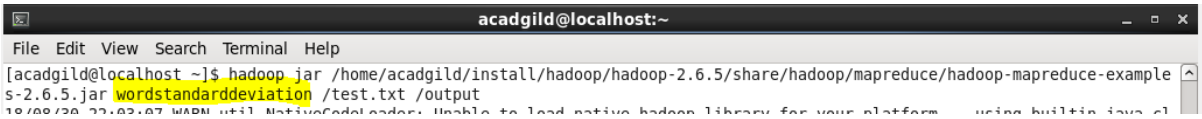
The below command is used to execute the jar file.

The generalized command is like below:

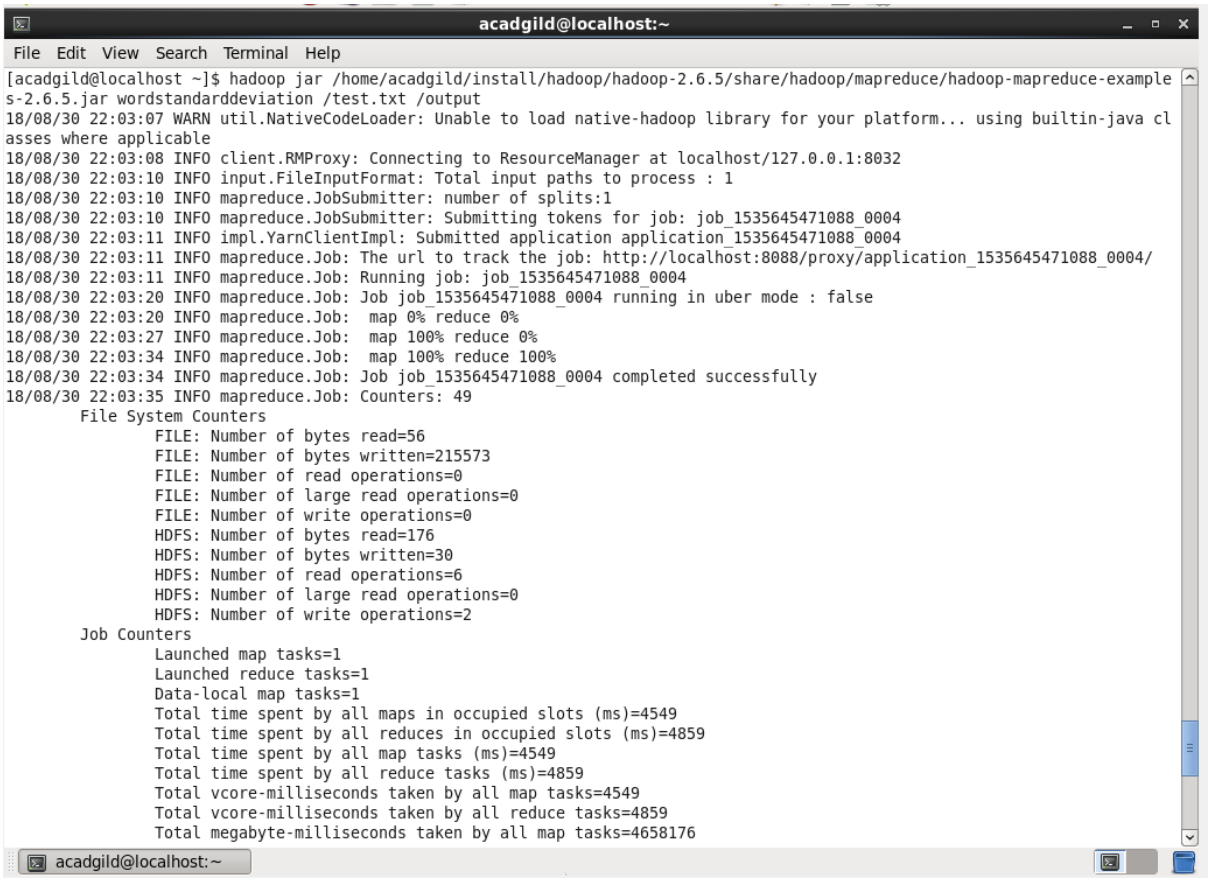
Hadoop jar <jar location> <Main class name> <Command line arguments/Input data> <output path>

a. Jar location: It is the location of the jar file where it is placed.
b. Main class name: It is the main class name of the jar file.
c. Command line arguments: Data to be processed has to be provided.
d. Output: The output of the process data will be available in this path.

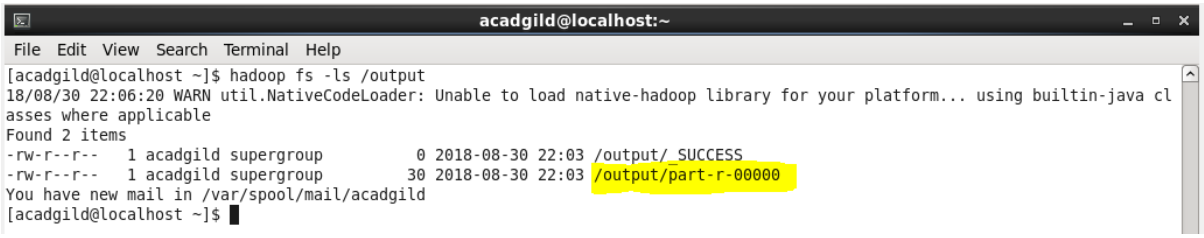
The command used to execute **wordstandarddeviation** program is as shown in the below screenshot.



Log displayed on the screen while execution of the program is shown below.



The output of the processed data is stored in a path as highlighted in the below screen shot.



Data in the output file is as shown in the below screenshot.