4. Problem Statement

Task 1:

Execute WordMedian , WordMean , WordStandardDeviation programs using

hadoop-mapreduce-examples-2.9.0.jar file present in your AcadGild VM.
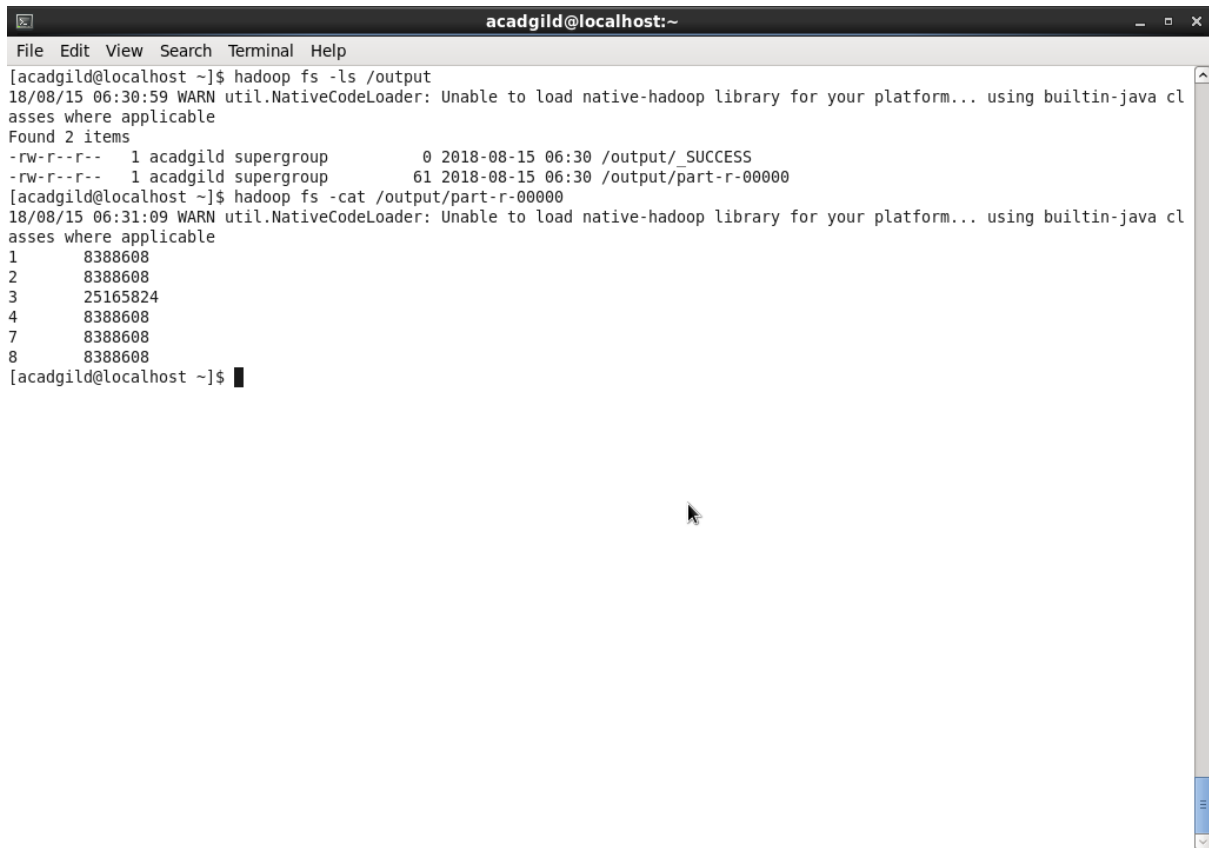
WordMedian

This can be executed using command

->hadoop jar install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples.jar wordmedian /file327.txt /output

```
acadgild@localhost:~                                                    _  □  ✕

File  Edit  View  Search  Terminal  Help

[acadgild@localhost ~]$ hadoop jar install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wor
dmedian /file327.txt /output
18/08/15 06:28:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/08/15 06:28:37 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/08/15 06:28:39 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool in
terface and execute your application with ToolRunner to remedy this.
18/08/15 06:28:39 INFO input.FileInputFormat: Total input paths to process : 1
18/08/15 06:28:39 INFO mapreduce.JobSubmitter: number of splits:3
18/08/15 06:28:40 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1534293639900_0005
18/08/15 06:28:40 INFO impl.YarnClientImpl: Submitted application application_1534293639900_0005
18/08/15 06:28:40 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1534293639900_0005/
18/08/15 06:28:40 INFO mapreduce.Job: Running job: job_1534293639900_0005
18/08/15 06:28:53 INFO mapreduce.Job: Job job_1534293639900_0005 running in uber mode : false
18/08/15 06:28:53 INFO mapreduce.Job:  map 0% reduce 0%
18/08/15 06:29:07 INFO mapreduce.Job:  map 3% reduce 0%


                                    Acadgild
```
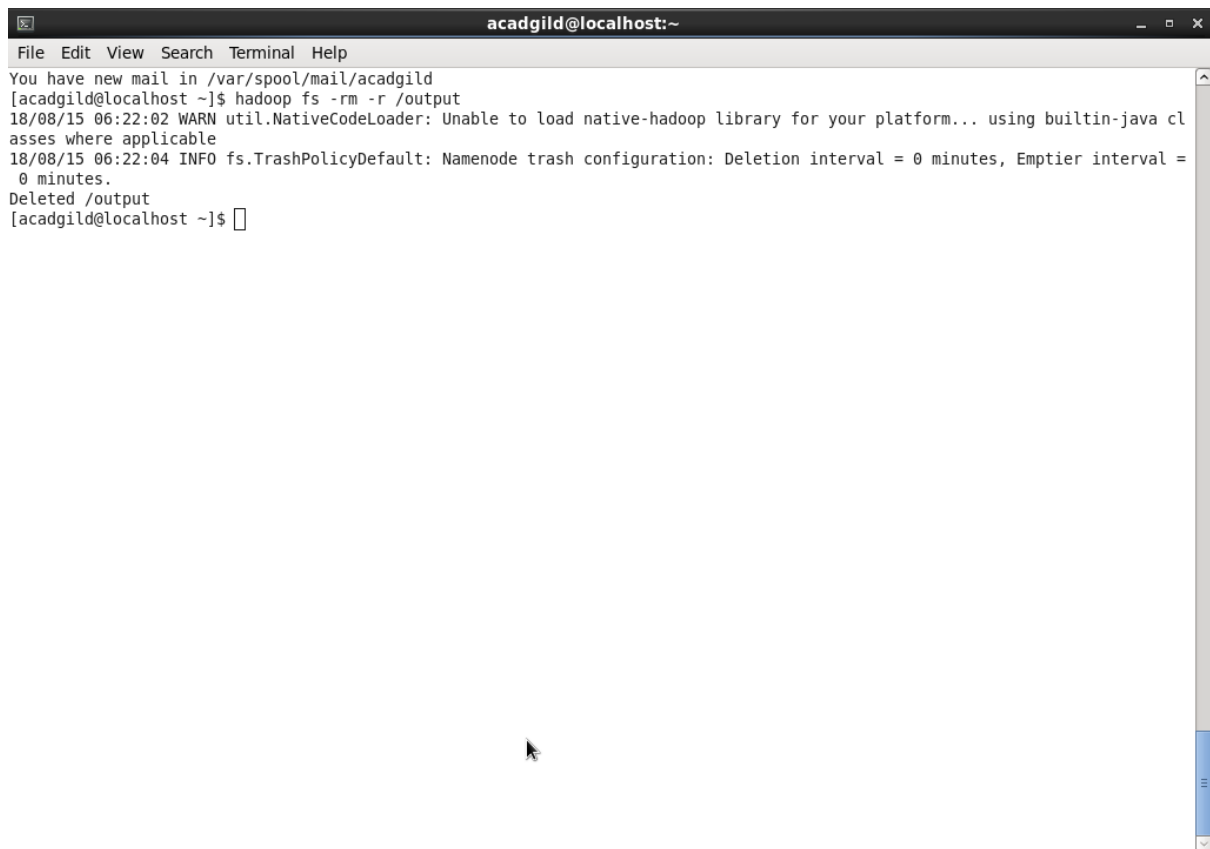
wordmedian is the class present in hadoop-mapreduce-examples.jar

/file327.txt is file having some text

/output would be the directory where output would be generated after this job has been completed

```
acadgild@localhost:~

File   Edit   View   Search   Terminal   Help
[acadgild@localhost ~]$ hadoop fs -ls /output
18/08/15 06:30:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-08-15 06:30 /output/_SUCCESS
-rw-r--r--   1 acadgild supergroup         61 2018-08-15 06:30 /output/part-r-00000
[acadgild@localhost ~]$ hadoop fs -cat /output/part-r-00000
18/08/15 06:31:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
1       8388608
2       8388608
3       25165824
4       8388608
7       8388608
8       8388608
[acadgild@localhost ~]$
```

Remove the output folder or else we would get exception as "Output Directory Already Exists"
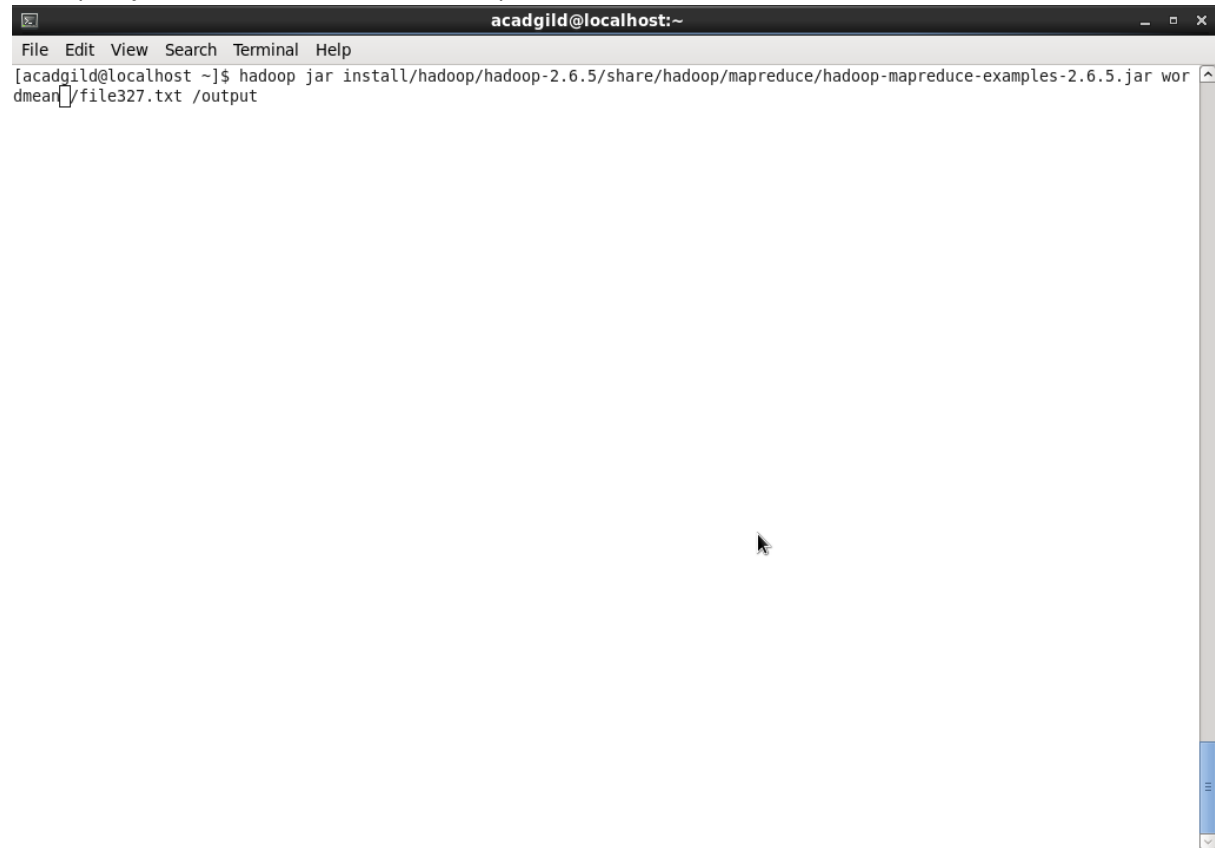
```
acadgild@localhost:~

File  Edit  View  Search  Terminal  Help

You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -rm -r /output
18/08/15 06:22:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/08/15 06:22:04 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval =
 0 minutes.
Deleted /output
[acadgild@localhost ~]$ 
```

WordMean

This can be executed using command

->hadoop jar install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples.jar wordmean /file327.txt /output

WordMean

```
acadgild@localhost:~

File  Edit  View  Search  Terminal  Help

[acadgild@localhost ~]$ hadoop jar install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wor
dmean /file327.txt /output
18/08/15 06:24:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/08/15 06:24:27 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/08/15 06:24:29 INFO input.FileInputFormat: Total input paths to process : 1
18/08/15 06:24:29 INFO mapreduce.JobSubmitter: number of splits:3
18/08/15 06:24:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1534293639900_0004
18/08/15 06:24:30 INFO impl.YarnClientImpl: Submitted application application_1534293639900_0004
18/08/15 06:24:30 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1534293639900_0004/
18/08/15 06:24:30 INFO mapreduce.Job: Running job: job_1534293639900_0004
18/08/15 06:24:43 INFO mapreduce.Job: Job job_1534293639900_0004 running in uber mode : false
18/08/15 06:24:43 INFO mapreduce.Job:  map 0% reduce 0%
18/08/15 06:24:57 INFO mapreduce.Job:  map 2% reduce 0%
18/08/15 06:25:00 INFO mapreduce.Job:  map 4% reduce 0%
18/08/15 06:25:03 INFO mapreduce.Job:  map 6% reduce 0%
```
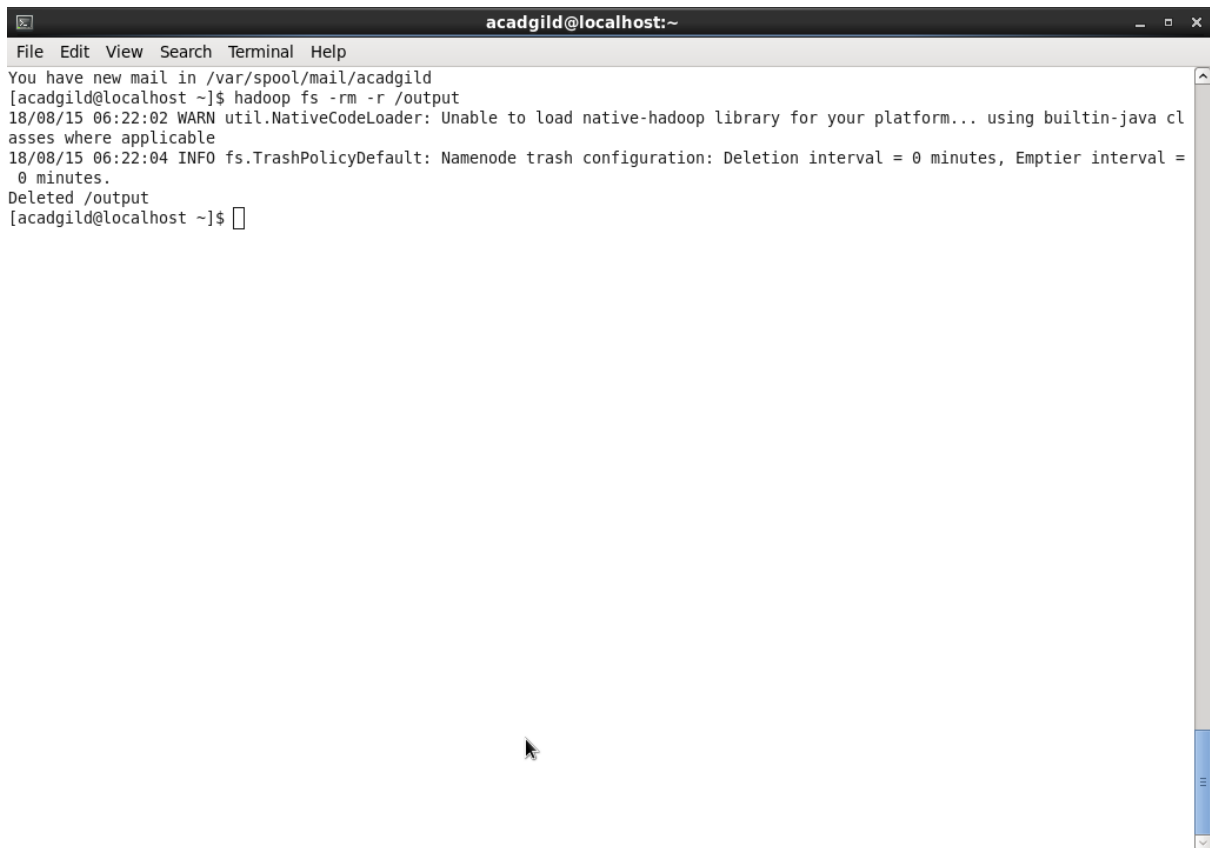
wordmean is the class present in hadoop-mapreduce-examples.jar

/file327.txt is file having some text

/output would be the directory where output would be generated after this job has been completed

```
[acadgild@localhost ~]$ hadoop fs -ls /output
18/08/15 06:27:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-08-15 06:26 /output/_SUCCESS
-rw-r--r--   1 acadgild supergroup         32 2018-08-15 06:26 /output/part-r-00000
[acadgild@localhost ~]$ hadoop fs -cat /output/part-r-00000
18/08/15 06:27:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
count   67108864
length  260046848
[acadgild@localhost ~]$
```

Remove the output folder or else we would get exception as "Output Directory Already Exists"

```
                            acadgild@localhost:~                        _  □  ✕

 File  Edit  View  Search  Terminal  Help

You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -rm -r /output
18/08/15 06:22:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/08/15 06:22:04 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval =
 0 minutes.
Deleted /output
[acadgild@localhost ~]$ ▯
```

WordStandardDeviation

This can be executed using command
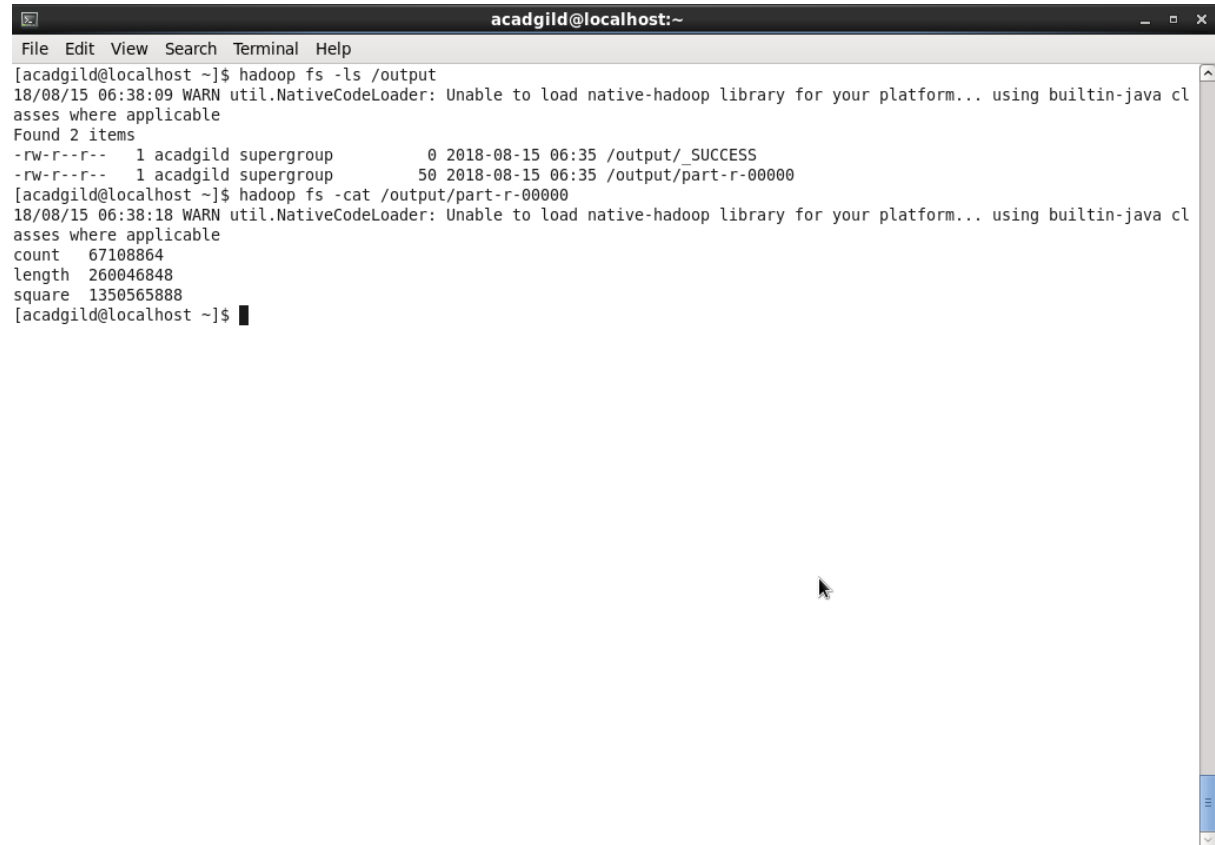
->hadoop jar install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/hadoop-mapreduce-
examples.jar wordStandardDeviation /file327.txt /output

wordStandardDeviation is the class present in Hadoop-mapreduce-examples.jar

/file327.txt is file having some text

/output would be the directory where output would be generated after this job has been completed

```
acadgild@localhost:~
File  Edit  View  Search  Terminal  Help
[acadgild@localhost ~]$ hadoop fs -ls /output
18/08/15 06:38:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup          0 2018-08-15 06:35 /output/_SUCCESS
-rw-r--r--   1 acadgild supergroup         50 2018-08-15 06:35 /output/part-r-00000
[acadgild@localhost ~]$ hadoop fs -cat /output/part-r-00000
18/08/15 06:38:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
count   67108864
length  260046848
square  1350565888
[acadgild@localhost ~]$
```