Job Output

```
[rs6785@login-1-1 ~]$ hadoop jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/
hadoop-streaming.jar -D mapreduce.job.reduces=1 -files hdfs://dumbo/user/rs6785/
class2/python_code/StringCountMapper.py,hdfs://dumbo/user/rs6785/class2/python_c
ode/StringCountReducer.py -mapper "StringCountMapper.py" -reducer "StringCountRe
ducer.py" -input /user/rs6785/class2/SampleInput.txt -output /user/rs6785/class2
/output
packageJobJar: [] [/opt/cloudera/parcels/CDH-5.15.2-1.cdh5.15.2.p0.3/jars/hadoop
-streaming-2.6.0-cdh5.15.2.jar] /tmp/streamjob6834463529048293190.jar tmpDir=nul
l
19/09/13 19:53:52 INFO mapreduce.JobSubmissionFiles: Permissions on staging dire
ctory /user/rs6785/.staging are incorrect: rwxrwx---. Fixing permissions to corr
ect value rwx------
19/09/13 19:53:52 INFO mapred.FileInputFormat: Total input paths to process : 1
19/09/13 19:53:53 INFO mapreduce.JobSubmitter: number of splits:2
19/09/13 19:53:53 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_15
64775008355_5502
19/09/13 19:53:53 INFO impl.YarnClientImpl: Submitted application application_15
64775008355_5502
19/09/13 19:53:53 INFO mapreduce.Job: The url to track the job: http://babar.es.
its.nyu.edu:8088/proxy/application_1564775008355_5502/
19/09/13 19:53:53 INFO mapreduce.Job: Running job: job_1564775008355_5502
19/09/13 19:53:58 INFO mapreduce.Job: Job job_1564775008355_5502 running in uber
 mode : false
19/09/13 19:53:58 INFO mapreduce.Job:  map 0% reduce 0%
19/09/13 19:54:02 INFO mapreduce.Job:  map 50% reduce 0%
19/09/13 19:54:03 INFO mapreduce.Job:  map 100% reduce 0%
19/09/13 19:54:08 INFO mapreduce.Job:  map 100% reduce 100%
19/09/13 19:54:09 INFO mapreduce.Job: Job job_1564775008355_5502 completed succe
ssfully
19/09/13 19:54:09 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=89
                FILE: Number of bytes written=474952
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=632
                HDFS: Number of bytes written=35
                HDFS: Number of read operations=9
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Launched map tasks=2
```

```
        Job Counters
                Launched map tasks=2
                Launched reduce tasks=1
                Data-local map tasks=2
                Total time spent by all maps in occupied slots (ms)=18208
                Total time spent by all reduces in occupied slots (ms)=14568
                Total time spent by all map tasks (ms)=4552
                Total time spent by all reduce tasks (ms)=2428
                Total vcore-milliseconds taken by all map tasks=4552
                Total vcore-milliseconds taken by all reduce tasks=2428
                Total megabyte-milliseconds taken by all map tasks=18644992
                Total megabyte-milliseconds taken by all reduce tasks=14917632
        Map-Reduce Framework
                Map input records=3
                Map output records=12
                Map output bytes=105
                Map output materialized bytes=132
                Input split bytes=198
                Combine input records=0
                Combine output records=0
                Reduce input groups=4
                Reduce shuffle bytes=132
                Reduce input records=12
                Reduce output records=4
                Spilled Records=24
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=349
                CPU time spent (ms)=6650
                Physical memory (bytes) snapshot=2014629888
                Virtual memory (bytes) snapshot=11202093056
                Total committed heap usage (bytes)=4000317440
        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=434
        File Output Format Counters
                Bytes Written=35
```
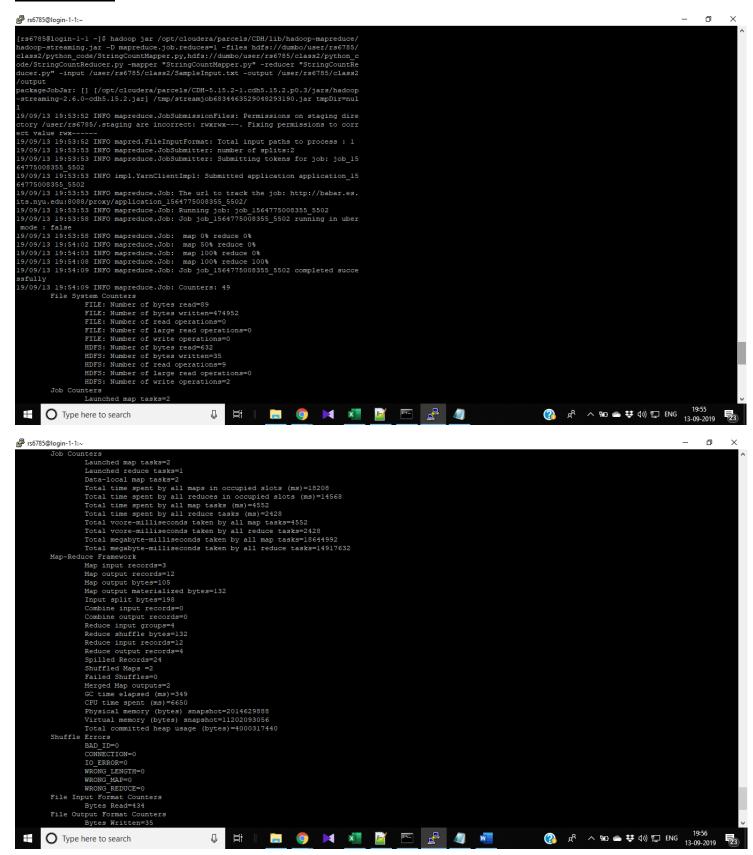
Screenshot that the program ran successfully and the output

```
                Input split bytes=198
                Combine input records=0
                Combine output records=0
                Reduce input groups=4
                Reduce shuffle bytes=132
                Reduce input records=12
                Reduce output records=4
                Spilled Records=24
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=349
                CPU time spent (ms)=6650
                Physical memory (bytes) snapshot=2014629888
                Virtual memory (bytes) snapshot=11202093056
                Total committed heap usage (bytes)=4000317440
        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=434
        File Output Format Counters
                Bytes Written=35
19/09/13 19:54:09 INFO streaming.StreamJob: Output directory: /user/rs6785/class
2/output
[rs6785@login-1-1 ~]$ hdfs dfs -ls /user/rs6785/class2/output
Found 2 items
-rw-r--r--+  3 rs6785 users           0 2019-09-13 19:54 /user/rs6785/class2/outp
ut/_SUCCESS
-rw-r--r--+  3 rs6785 users          35 2019-09-13 19:54 /user/rs6785/class2/outp
ut/part-00000
[rs6785@login-1-1 ~]$ hdfs dfs -ls /user/rs6785/class2/output/part-00000
-rw-r--r--+  3 rs6785 users          35 2019-09-13 19:54 /user/rs6785/class2/outp
ut/part-00000
[rs6785@login-1-1 ~]$ hdfs dfs -cat /user/rs6785/class2/output/part-00000
Chicago 1
Dec     3
Java    0
hackathon       2
[rs6785@login-1-1 ~]$
```