```
[cloudera@quickstart ~]$ ls
cloudera-manager  Downloads                    kerberos  Pictures
workspace
cm_api.py         eclipse                      lib       Public
Desktop           enterprise-deployment.json   Music     Templates
Documents         express-deployment.json      parcels   Videos
[cloudera@quickstart ~]$ cd workspace
[cloudera@quickstart workspace]$ ls
[cloudera@quickstart WordCount]$ curl http://www.gutenberg.org/
files/100/100-0.txt | perl -pe 's/^\xEF\xBB\xBF//' > 100-0.txt
  % Total    % Received % Xferd  Average Speed   Time     Time
Time  Current
                                 Dload  Upload   Total    Spent
Left  Speed
100 5622k  100 5622k    0      0   390k       0  0:00:14  0:00:14
--:--:--   742k
[cloudera@quickstart WordCount]$ ls
100-0.txt  bin  conf  hadoop.log  output  src
[cloudera@quickstart WordCount]$ cd ..
[cloudera@quickstart workspace]$ cd ..
[cloudera@quickstart ~]$ pwd
/home/cloudera
[cloudera@quickstart ~]$ hadoop fs -put workspace/WordCount/
100-0.txt
[cloudera@quickstart ~]$ hadoop jar WordCount.jar
edu.stanford.cs246.wordcount.WordCount 100-0.txt output
[100-0.txt, output]
[100-0.txt, output]
21/03/09 05:15:24 INFO client.RMProxy: Connecting to ResourceManager
at /0.0.0.0:8032
21/03/09 05:15:25 INFO input.FileInputFormat: Total input paths to
process : 1
21/03/09 05:15:25 INFO mapreduce.JobSubmitter: number of splits:1
21/03/09 05:15:25 INFO mapreduce.JobSubmitter: Submitting tokens for
job: job_1615290818956_0001
21/03/09 05:15:25 INFO impl.YarnClientImpl: Submitted application
application_1615290818956_0001
21/03/09 05:15:26 INFO mapreduce.Job: The url to track the job:
http://quickstart.cloudera:8088/proxy/
application_1615290818956_0001/
21/03/09 05:15:26 INFO mapreduce.Job: Running job:
job_1615290818956_0001
21/03/09 05:15:35 INFO mapreduce.Job: Job job_1615290818956_0001
running in uber mode : false
21/03/09 05:15:35 INFO mapreduce.Job:  map 0% reduce 0%
21/03/09 05:15:43 INFO mapreduce.Job:  map 100% reduce 0%
21/03/09 05:15:52 INFO mapreduce.Job:  map 100% reduce 100%
21/03/09 05:15:52 INFO mapreduce.Job: Job job_1615290818956_0001
completed successfully
21/03/09 05:15:52 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=11671983
                FILE: Number of bytes written=23567909
                FILE: Number of read operations=0
```

```
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=5757232
                HDFS: Number of bytes written=829099
                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)=6158
                Total time spent by all reduces in occupied slots
(ms)=6359
                Total time spent by all map tasks (ms)=6158
                Total time spent by all reduce tasks (ms)=6359
                Total vcore-seconds taken by all map tasks=6158
                Total vcore-seconds taken by all reduce tasks=6359
                Total megabyte-seconds taken by all map
tasks=6305792
                Total megabyte-seconds taken by all reduce
tasks=6511616
        Map-Reduce Framework
                Map input records=170607
                Map output records=1044475
                Map output bytes=9583027
                Map output materialized bytes=11671983
                Input split bytes=120
                Combine input records=0
                Combine output records=0
                Reduce input groups=76517
                Reduce shuffle bytes=11671983
                Reduce input records=1044475
                Reduce output records=76517
                Spilled Records=2088950
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=133
                CPU time spent (ms)=4360
                Physical memory (bytes) snapshot=361603072
                Virtual memory (bytes) snapshot=3008335872
                Total committed heap usage (bytes)=226365440
        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=5757112
        File Output Format Counters
```

```
                Bytes Written=829099
[cloudera@quickstart ~]$ hadoop fs -ls output
Found 2 items
-rw-r--r--   1 cloudera cloudera          0 2021-03-09 05:15 output/
_SUCCESS
-rw-r--r--   1 cloudera cloudera     829099 2021-03-09 05:15 output/
part-r-00000
[cloudera@quickstart ~]$ hadoop fs -cat output/part\* | head
        82975
"       49
"Air,"  1
"Defects,"      1
"Information     1
"Neither 1
"Plain   2
"Project 5
"Right   1
"so      1
cat: Unable to write to output stream.
[cloudera@quickstart ~]$
```