```
root@hadoop-master:~# hadoop jar /shared_volume/HadoopFileStatus.jar /user/root/input purchases.txt achats.txt
2549 bytes
File Name: purchases.txt
File Size: 2549
File owner: root
File permission: rw-r--r--
File Replication: 2
File Block Size: 134217728
Block offset: 0
Block length: 2549
Block hosts: hadoop-slave1 hadoop-slave2
root@hadoop-master:~#
```

```
root@hadoop-master:~# hadoop jar /shared_volume/ReadHDFS.jar /user/root/input/purchases.txt
2012-01-01      09:00   San Jose        Men's Clothing  214.05  Amex
2012-01-01      09:00   Fort Worth      Women's Clothing        153.57  Visa
2012-01-01      09:00   San Diego       Music   66.08   Cash
2012-01-01      09:00   Pittsburgh      Pet Supplies    493.51  Discover
2012-01-01      09:00   Omaha   Children's Clothing     235.63  MasterCard
2012-01-01      09:00   Stockton        Men's Clothing  247.18  MasterCard
2012-01-01      09:00   Austin  Cameras 379.6   Visa
2012-01-01      09:00   New York        Consumer Electronics    296.8   Cash
2012-01-01      09:00   Corpus Christi  Toys    25.38   Discover
2012-01-01      09:00   Fort Worth      Toys    213.88  Visa
2012-01-01      09:00   Las Vegas       Video Games     53.26   Visa
2012-01-01      09:00   Newark  Video Games     39.75   Cash
2012-01-01      09:00   Austin  Cameras 469.63  MasterCard
2012-01-01      09:00   Greensboro      DVDs    290.82  MasterCard
2012-01-01      09:00   San Francisco   Music   260.65  Discover
```

```
root@hadoop-master:~# hadoop jar /shared_volume/ReadHDFS.jar /user/root/input/bonjour.txt "m"
Bonjour tout le monde !Mon message personnalis??????
root@hadoop-master:~#
```

```
PS C:\Users\ahmed\Downloads\Bigdatalabs\hadoop_lab> docker cp "C:\hadoop_project\ReadHDFS.jar" hadoop-master:/shared_volume/ReadHDFS.jar
Successfully copied 7.68kB to hadoop-master:/shared_volume/ReadHDFS.jar
PS C:\Users\ahmed\Downloads\Bigdatalabs\hadoop_lab> docker cp "C:\hadoop_project\WriteHDFS.jar" hadoop-master:/shared_volume/WriteHDFS.jar^C
PS C:\Users\ahmed\Downloads\Bigdatalabs\hadoop_lab> docker cp "C:\hadoop_project\WriteHDFS.jar" hadoop-master:/shared_volume/WriteHDFS.jar
Successfully copied 9.22kB to hadoop-master:/shared_volume/WriteHDFS.jar
PS C:\Users\ahmed\Downloads\Bigdatalabs\hadoop_lab> docker cp "C:\hadoop_project\WordCount.jar" hadoop-master:/shared_volume/WordCount.jar
>>
Successfully copied 12.8kB to hadoop-master:/shared_volume/WordCount.jar
PS C:\Users\ahmed\Downloads\Bigdatalabs\hadoop_lab>
```

```
root@hadoop-master:~# hadoop jar /shared_volume/WordCount.jar /user/root/input/purchases.txt /user/root/output/wordcount
2025-10-07 17:44:19,349 INFO client.RMProxy: Connecting to ResourceManager at hadoop-master/172.19.0.2:8032
2025-10-07 17:44:19,638 INFO client.AHSProxy: Connecting to Application History server at localhost/127.0.0.1:10200
2025-10-07 17:44:20,097 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and e
xecute your application with ToolRunner to remedy this.
2025-10-07 17:44:20,179 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/root/.staging/job_175985
2401135_0002
2025-10-07 17:44:20,670 INFO input.FileInputFormat: Total input files to process : 1
2025-10-07 17:44:20,902 INFO mapreduce.JobSubmitter: number of splits:1
2025-10-07 17:44:21,017 INFO Configuration.deprecation: yarn.resourcemanager.system-metrics-publisher.enabled is deprecated. Instead, use yarn.s
ystem-metrics-publisher.enabled
2025-10-07 17:44:21,244 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1759852401135_0002
2025-10-07 17:44:21,247 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-10-07 17:44:21,560 INFO conf.Configuration: resource-types.xml not found
2025-10-07 17:44:21,560 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.

2025-10-07 17:44:23,797 INFO impl.YarnClientImpl: Submitted application application_1759852401135_0002
2025-10-07 17:44:23,901 INFO mapreduce.Job: The url to track the job: http://hadoop-master:8088/proxy/application_1759852401135_0002/
2025-10-07 17:44:23,902 INFO mapreduce.Job: Running job: job_1759852401135_0002
2025-10-07 17:44:44,594 INFO mapreduce.Job: Job job_1759852401135_0002 running in uber mode : false
2025-10-07 17:44:44,614 INFO mapreduce.Job:  map 0% reduce 0%
2025-10-07 17:45:14,896 INFO mapreduce.Job:  map 100% reduce 0%
2025-10-07 17:45:25,977 INFO mapreduce.Job:  map 100% reduce 100%
2025-10-07 17:45:29,059 INFO mapreduce.Job: Job job_1759852401135_0002 completed successfully
2025-10-07 17:45:29,930 INFO mapreduce.Job: Counters: 54
        File System Counters
                FILE: Number of bytes read=1637
                FILE: Number of bytes written=446681
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=2669
                HDFS: Number of bytes written=1142
                HDFS: Number of read operations=8
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
```

```
        File System Counters
                FILE: Number of bytes read=1637
                FILE: Number of bytes written=446681
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=2669
                HDFS: Number of bytes written=1142
                HDFS: Number of read operations=8
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1

                Peak Reduce Virtual memory (bytes)=261
        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=2549
        File Output Format Counters
                Bytes Written=1142
root@hadoop-master:~#
```

RAM 3.17 GB   CPU 1.13%   Disk: 19.17 GB used (limit 1006.85 GB)

```
                    bytes written=1142
root@hadoop-master:~# hdfs dfs -cat /user/root/output/wordcount/part-r-00000
09:00    22
09:01    22
09:02    6
108.69   1
117.81   1
135.96   1
136.9    1
15.41    1
153.57   1
16.78    1
170.2    1
182.05   1
189.01   1
2012-01-01      50
205.06   1
```

```
root@hadoop-master:~# find / -name 'hadoop-streaming*.jar' 2>/dev/null
/usr/local/hadoop/share/hadoop/tools/lib/hadoop-streaming-3.2.0.jar
/usr/local/hadoop/share/hadoop/tools/sources/hadoop-streaming-3.2.0-sources.jar
/usr/local/hadoop/share/hadoop/tools/sources/hadoop-streaming-3.2.0-test-sources.jar
```

```
root@hadoop-master:~# hadoop jar /usr/local/hadoop/share/hadoop/tools/lib/hadoop-streaming-3.2.0.jar \
>     -files /shared_volume/mapper.py,/shared_volume/reducer.py \
>     -mapper "python3 mapper.py" \
>     -reducer "python3 reducer.py" \
>     -input /user/root/input/purchases.txt \
>     -output /user/root/output/wordcount_python
packageJobJar: [/tmp/hadoop-unjar1827008077805825182/] [] /tmp/streamjob3683874952715688917.jar tmpDir=null
2025-10-07 18:00:52,964 INFO client.RMProxy: Connecting to ResourceManager at hadoop-master/172.19.0.2:8032
2025-10-07 18:00:53,298 INFO client.AHSProxy: Connecting to Application History server at localhost/127.0.0.1:10200
2025-10-07 18:00:53,379 INFO client.RMProxy: Connecting to ResourceManager at hadoop-master/172.19.0.2:8032
2025-10-07 18:00:53,379 INFO client.AHSProxy: Connecting to Application History server at localhost/127.0.0.1:10200
2025-10-07 18:00:54,385 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/root/.staging/job_175985
2401135_0003
2025-10-07 18:00:56,011 INFO mapred.FileInputFormat: Total input files to process : 1
2025-10-07 18:00:56,311 INFO mapreduce.JobSubmitter: number of splits:2
2025-10-07 18:00:56,401 INFO Configuration.deprecation: yarn.resourcemanager.system-metrics-publisher.enabled is deprecated. Instead, use yarn.s
```

```
2025-10-07 18:00:56,678 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1759852401135_0003
2025-10-07 18:00:56,682 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-10-07 18:00:57,075 INFO conf.Configuration: resource-types.xml not found
2025-10-07 18:00:57,078 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2025-10-07 18:00:57,989 INFO impl.YarnClientImpl: Submitted application application_1759852401135_0003
2025-10-07 18:00:58,066 INFO mapreduce.Job: The url to track the job: http://hadoop-master:8088/proxy/application_1759852401135_0003/
2025-10-07 18:00:58,068 INFO mapreduce.Job: Running job: job_1759852401135_0003
2025-10-07 18:01:13,092 INFO mapreduce.Job: Job job_1759852401135_0003 running in uber mode : false
2025-10-07 18:01:13,115 INFO mapreduce.Job:  map 0% reduce 0%
2025-10-07 18:01:53,085 INFO mapreduce.Job:  map 100% reduce 0%
2025-10-07 18:02:59,282 INFO mapreduce.Job:  map 100% reduce 100%
2025-10-07 18:03:04,618 INFO mapreduce.Job: Job job_1759852401135_0003 completed successfully
2025-10-07 18:03:10,530 INFO mapreduce.Job: Counters: 54
        File System Counters
                FILE: Number of bytes read=3839
                FILE: Number of bytes written=684015
                FILE: Number of read operations=0


                Peak Map Physical memory (bytes)=272109568
                Peak Map Virtual memory (bytes)=2615517184
                Peak Reduce Physical memory (bytes)=207179776
                Peak Reduce Virtual memory (bytes)=2616299520
        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=3824
        File Output Format Counters
                Bytes Written=1142
2025-10-07 18:03:10,547 INFO streaming.StreamJob: Output directory: /user/root/output/wordcount_python
root@hadoop-master:~# 
```