

Ex01 Lihi Cohen Screenshots

```
hadoop@ip-172-31-30-47:~/course
2025-11-23 18:32:40,977 INFO mapreduce.Job: map 100% reduce 0%
2025-11-23 18:32:43,991 INFO mapreduce.Job: map 100% reduce 67%
2025-11-23 18:32:44,994 INFO mapreduce.Job: map 100% reduce 100%
2025-11-23 18:32:46,003 INFO mapreduce.Job: Job job_1763921376629_0002 completed successfully
2025-11-23 18:32:46,066 INFO mapreduce.Job: Counters: 56
  File System Counters
    FILE: Number of bytes read=1188898
    FILE: Number of bytes written=6971293
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=7215145
    HDFS: Number of bytes written=889817
    HDFS: Number of read operations=42
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=6
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed reduce tasks=1
    Launched map tasks=9
    Launched reduce tasks=3
    Data-local map tasks=4
    Rack-local map tasks=5
    Total time spent by all maps in occupied slots (ms)=169501920
    Total time spent by all reduces in occupied slots (ms)=75718080
    Total time spent by all map tasks (ms)=28945
    Total time spent by all reduce tasks (ms)=6465
    Total vcore-milliseconds taken by all map tasks=28945
    Total vcore-milliseconds taken by all reduce tasks=6465
    Total megabyte-milliseconds taken by all map tasks=169501920
    Total megabyte-milliseconds taken by all reduce tasks=75718080
  Map-Reduce Framework
    Map input records=138644
    Map output records=1168215
    Map output bytes=9000130
    Map output materialized bytes=1839521
    Input split bytes=1125
    Combine input records=0
    Combine output records=0
    Reduce input groups=77928
    Reduce shuffle bytes=1839521
    Reduce input records=1168215
    Reduce output records=77928
    Spilled Records=2336430
    Shuffled Maps =27
    Failed Shuffles=0
    Merged Map outputs=27
    GC time elapsed (ms)=130
    CPU time spent (ms)=12440
```

```

hadoop@ip-172-31-30-47:~/course
2025-11-23 18:37:04,605 INFO mapreduce.Job: map 100% reduce 100%
2025-11-23 18:37:04,610 INFO mapreduce.Job: Job job_1763921376629_0003 completed successfully
2025-11-23 18:37:04,669 INFO mapreduce.Job: Counters: 55
  File System Counters
    FILE: Number of bytes read=1236921
    FILE: Number of bytes written=8096298
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=7215145
    HDFS: Number of bytes written=889817
    HDFS: Number of read operations=57
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=12
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Launched map tasks=9
    Launched reduce tasks=6
    Data-local map tasks=5
    Rack-local map tasks=4
    Total time spent by all maps in occupied slots (ms)=177225984
    Total time spent by all reduces in occupied slots (ms)=139384512
    Total time spent by all map tasks (ms)=30264
    Total time spent by all reduce tasks (ms)=11901
    Total vcore-milliseconds taken by all map tasks=30264
    Total vcore-milliseconds taken by all reduce tasks=11901
    Total megabyte-milliseconds taken by all map tasks=177225984
    Total megabyte-milliseconds taken by all reduce tasks=139384512
  Map-Reduce Framework
    Map input records=138644
    Map output records=1168215
    Map output bytes=9000130
    Map output materialized bytes=1930346
    Input split bytes=1125
    Combine input records=0
    Combine output records=0
    Reduce input groups=77928
    Reduce shuffle bytes=1930346
    Reduce input records=1168215
    Reduce output records=77928
    Spilled Records=2336430
    Shuffled Maps =54
    Failed Shuffles=0
    Merged Map outputs=54
    GC time elapsed (ms)=138
    CPU time spent (ms)=15300
    Physical memory (bytes) snapshot=6759374848
    Virtual memory (bytes) snapshot=135780823040
    Total committed heap usage (bytes)=7239368704

```

```

2025-11-23 19:25:50,423 INFO streaming.StreamJob: Output directory: /user/hduser/gutenberg-output
[hadoop@ip-172-31-30-47 course]$ hadoop fs -cat /user/hduser/gutenberg-output/*
the      74369
and      42768
of        34588
[hadoop@ip-172-31-30-47 course]$

```