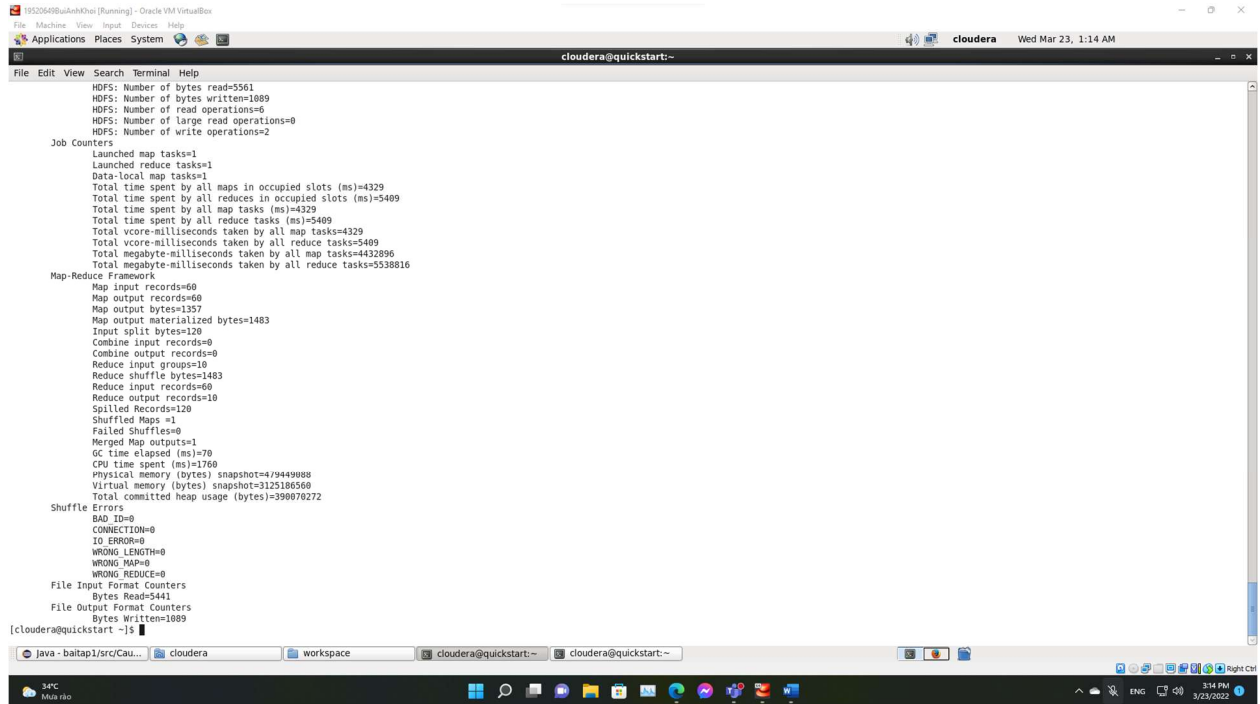


# Assignment 1

Họ tên: Bùi Anh Khôi

Lớp: DS200.M21

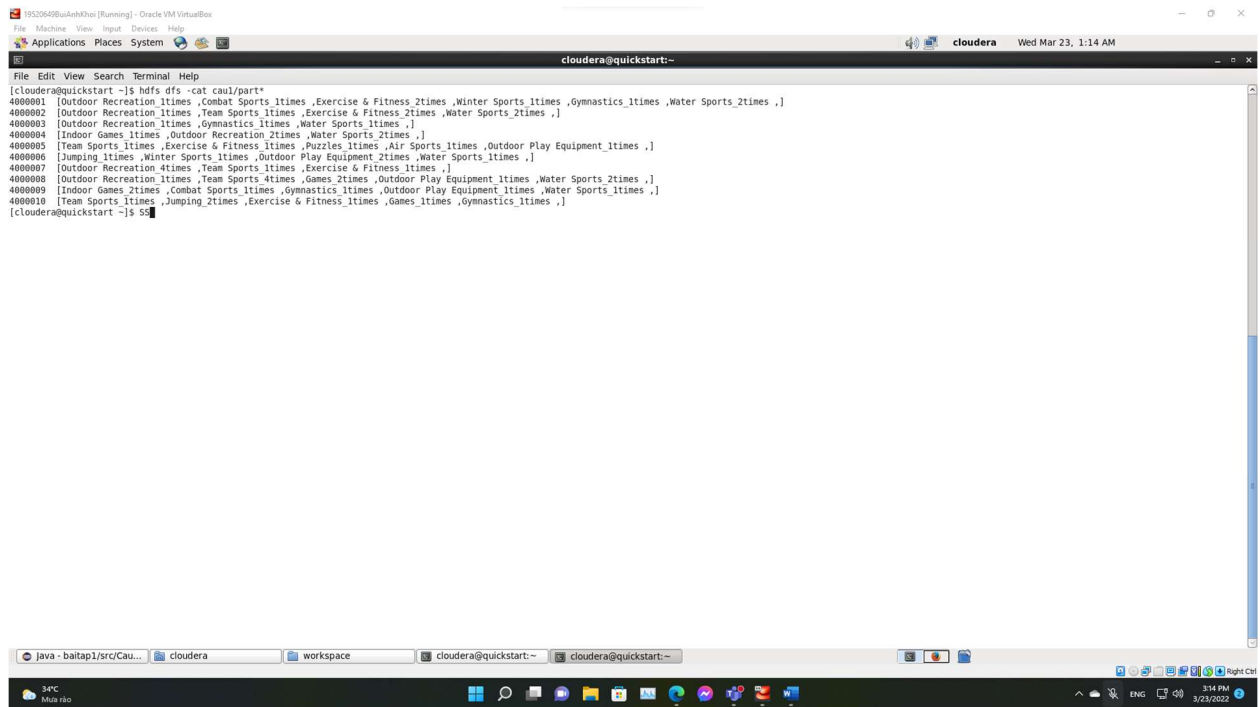
## Câu 1



The screenshot shows a terminal window titled "cloudera@quickstart:~" with the following output:

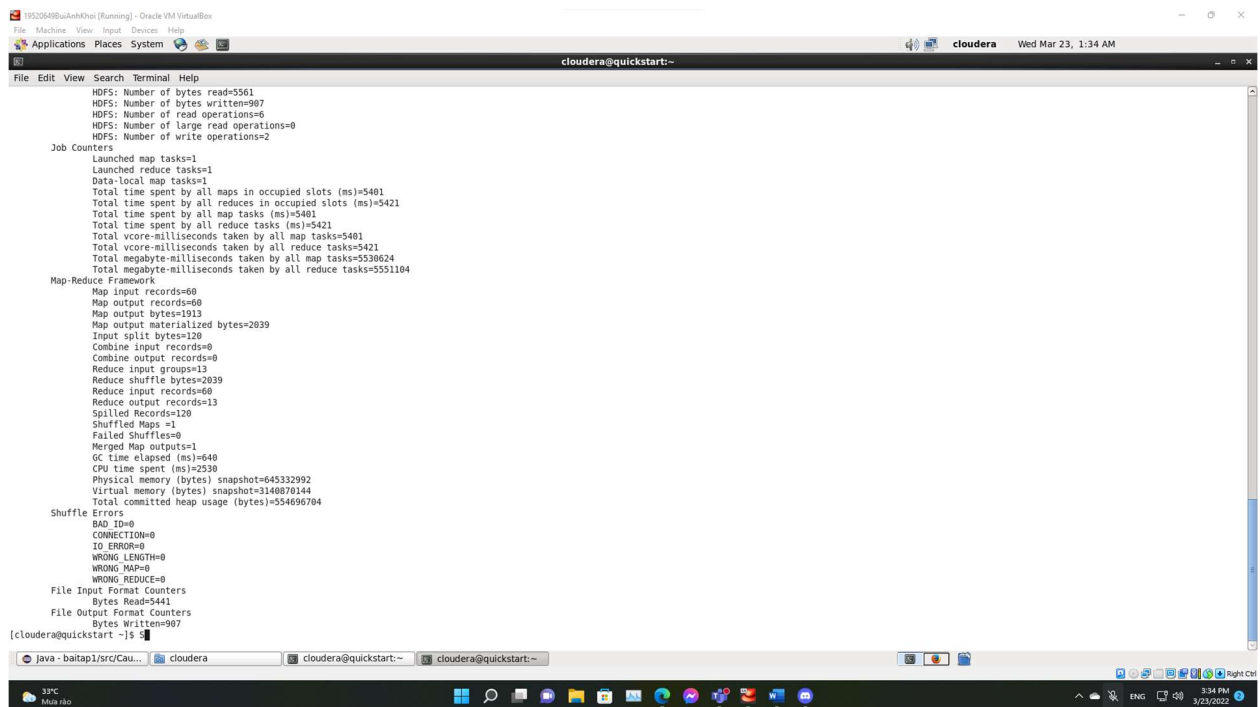
```
File Edit View Search Terminal Help
HDFS: Number of bytes read=5561
HDFS: Number of bytes written=1089
HDFS: Number of read operations=0
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Launched reduce tasks=1
  Data-local map tasks=1
  Total time spent by all maps in occupied slots (ms)=4329
  Total time spent by all reduces in occupied slots (ms)=5409
  Total time spent by all map tasks (ms)=4329
  Total time spent by all reduce tasks (ms)=5409
  Total vcore-milliseconds taken by all map tasks=4329
  Total vcore-milliseconds taken by all reduce tasks=5409
  Total megabyte-milliseconds taken by all map tasks=4432896
  Total megabyte-milliseconds taken by all reduce tasks=5538816
Map-Reduce Framework
  Map input records=60
  Map output records=60
  Map output bytes=1357
  Map output materialized bytes=1483
  Input split bytes=120
  Combine input records=0
  Combine output records=0
  Reduce input groups=10
  Reduce shuffle bytes=1483
  Reduce input records=60
  Reduce output records=10
  Spilled Records=120
  Shuffled Maps =1
  Failed Shuffles=0
  Merged Map outputs=1
  GC time elapsed (ms)=70
  CPU time spent (ms)=1760
  Physical memory (bytes) snapshot=479449088
  Virtual memory (bytes) snapshot=3125186560
  Total committed heap usage (bytes)=390070272
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=5441
File Output Format Counters
  Bytes Written=1089
[cloudera@quickstart ~]$
```

Hình 1 Chạy lệnh submit map reduce Câu 1

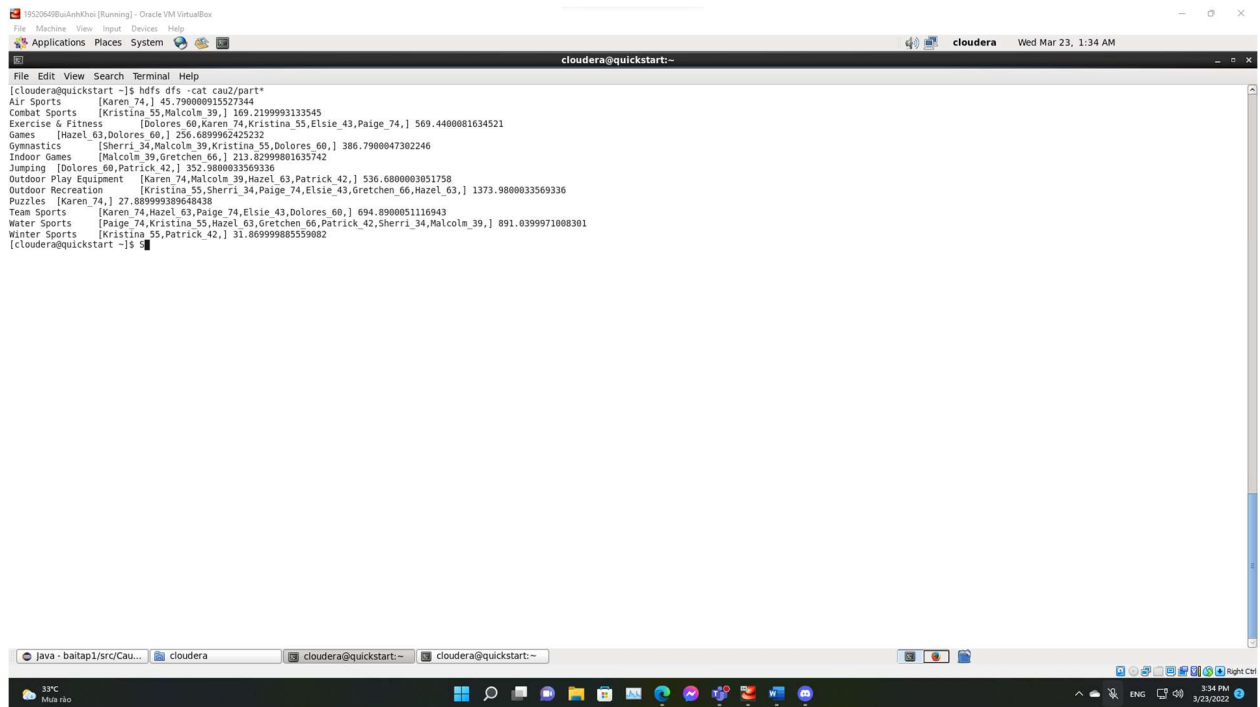


Hình 2 Chạy lệnh hiển thị file output Câu 1

## Câu 2



Hình 3 Chạy lệnh submit map reduce Câu 2



```
195206408BuAnhKhai [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System cloudera Wed Mar 23, 1:34 AM
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -cat cau2/part*
Air Sports      [Karen 74,] 45.7900000915527344
Combat Sports   [Kristina 55,Malcolm 39,] 169.2199993133545
Exercise & Fitness [Dolores 60,Karen 74,Kristina 55,Elsie 43,Paige 74,] 569.4400081634521
Games           [Hazel 63,Dolores 60,] 256.6899962425232
Gymnastics      [Sherri 34,Malcolm 39,Kristina 55,Dolores 60,] 386.7900047302246
Indoor Games    [Malcolm 39,Gretchen 66,] 213.82999801635742
Jumping [Dolores 60,Patrick 42,] 352.9000033569336
Outdoor Play Equipment [Karen 74,Malcolm 39,Hazel 63,Patrick 42,] 536.6800003051758
Outdoor Recreation [Kristina 55,Sherri 34,Paige 74,Elsie 43,Gretchen 66,Hazel 63,] 1373.9000033569336
Puzzles [Karen 74,] 27.809999389648438
Team Sports     [Karen 74,Hazel 63,Paige 74,Elsie 43,Dolores 60,] 694.8900051116943
Water Sports    [Paige 74,Kristina 55,Hazel 63,Gretchen 66,Patrick 42,Sherri 34,Malcolm 39,] 891.0399971008301
Winter Sports   [Kristina 55,Patrick 42,] 31.869999885559082
[cloudera@quickstart ~]$
```

Hình 4 Chạy lệnh hiển thị file output Câu 2