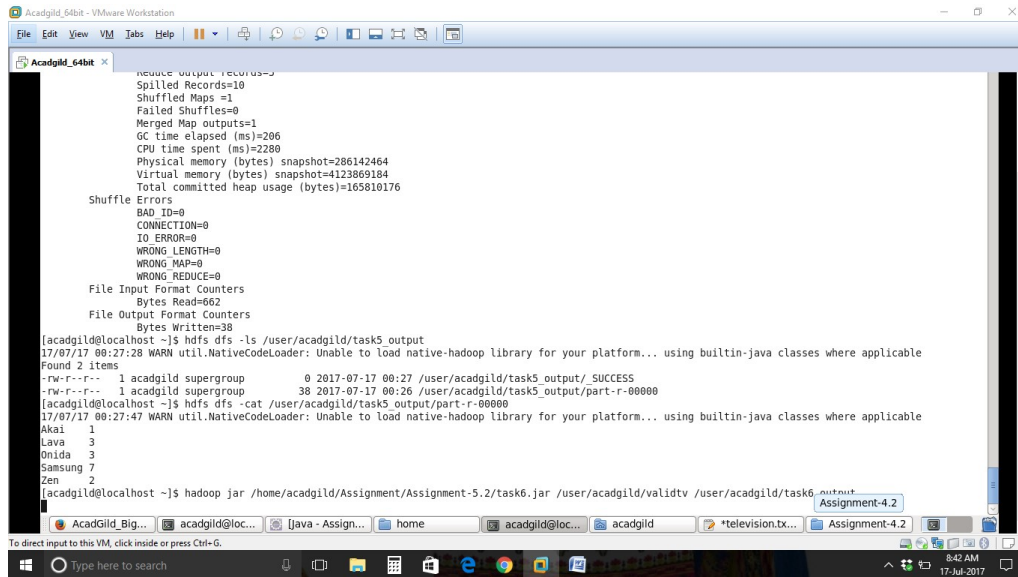


Assignment 5.2 Screenshots

Task 6

Modify Task 3 (refer session 4, assignment 2) to take advantage of Combiner.



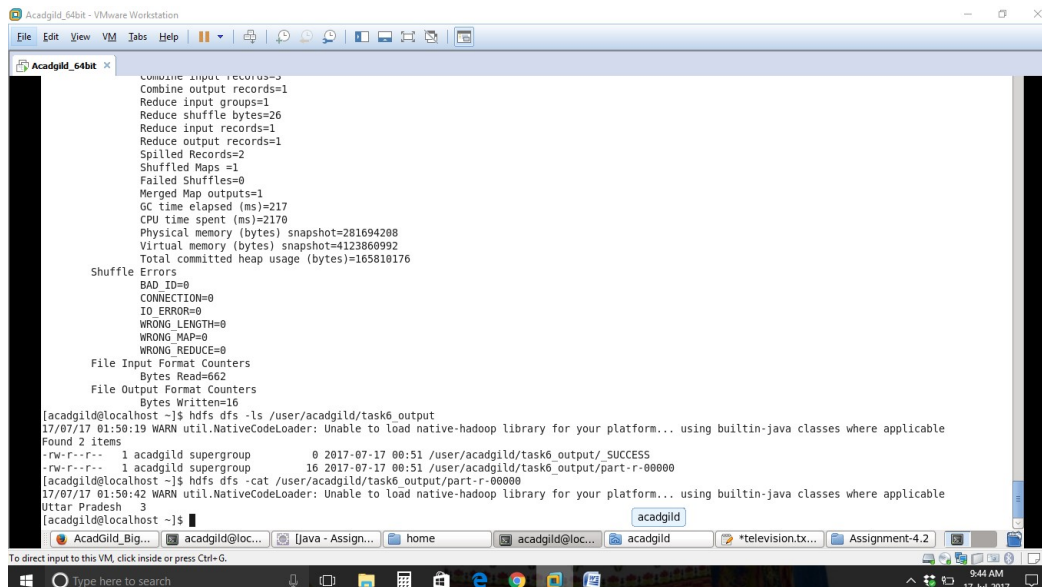
```
Acadgild_64bit x
reduce output records=3
Spilled Records=10
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=206
CPU time spent (ms)=2280
Physical memory (bytes) snapshot=286142464
Virtual memory (bytes) snapshot=4123869184
Total committed heap usage (bytes)=165810176

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=662
File Output Format Counters
Bytes Written=38

[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/task5 output
17/07/17 00:27:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2017-07-17 00:27 /user/acadgild/task5 output/ SUCCESS
-rw-r--r-- 1 acadgild supergroup 38 2017-07-17 00:26 /user/acadgild/task5 output/part-r-00000
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task5 output/part-r-00000
17/07/17 00:27:47 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Aka1 1
Lava 3
Onida 3
Samsung 7
Zen 2
[acadgild@localhost ~]$ hadoop jar /home/acadgild/Assignment/Assignment-5.2/task6.jar /user/acadgild/validtv /user/acadgild/task6 output
```

Output of Task 6



```
Acadgild_64bit x
Combine input records=3
Combine output records=1
Reduce input groups=1
Reduce shuffle bytes=26
Reduce input records=1
Reduce output records=1
Spilled Records=2
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=217
CPU time spent (ms)=2170
Physical memory (bytes) snapshot=281694288
Virtual memory (bytes) snapshot=412386992
Total committed heap usage (bytes)=165810176

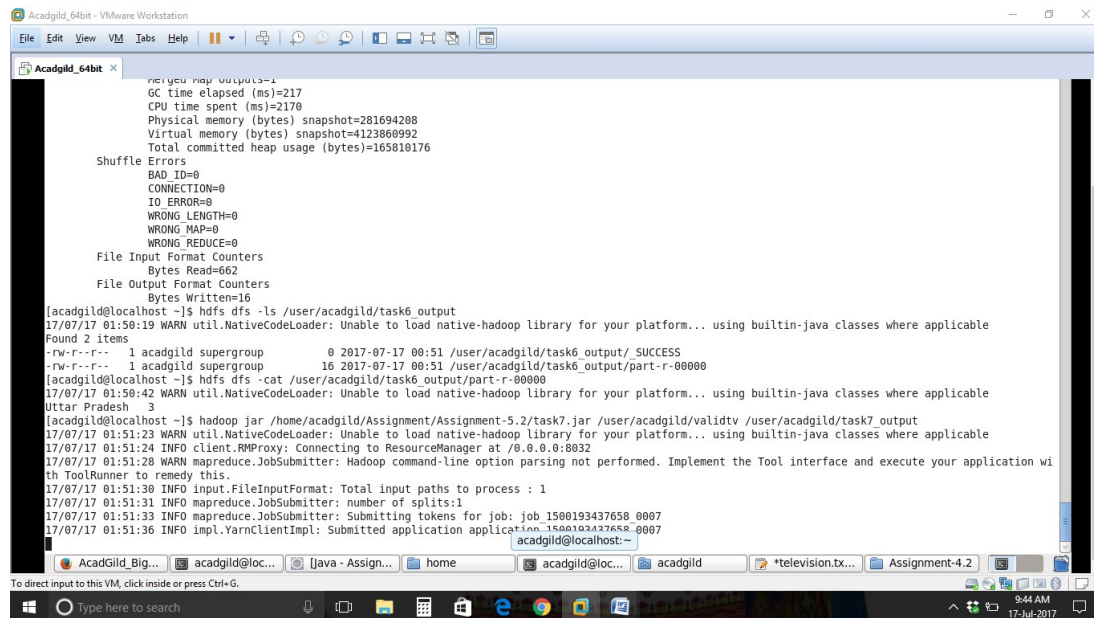
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0

File Input Format Counters
Bytes Read=662
File Output Format Counters
Bytes Written=16

[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/task6 output
17/07/17 01:50:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2017-07-17 00:51 /user/acadgild/task6 output/ SUCCESS
-rw-r--r-- 1 acadgild supergroup 16 2017-07-17 00:51 /user/acadgild/task6 output/part-r-00000
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task6 output/part-r-00000
17/07/17 01:50:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Uttar Pradesh 3
[acadgild@localhost ~]$
```

Task 7

Write a Map Reduce program to view the total sales for each product for every Company corresponding to each size. Make sure that all records for a single company goes to a single reducer and inside every reducer, keys must be sorted in descending order of the size. You may write a custom WritableComparable for this purpose.

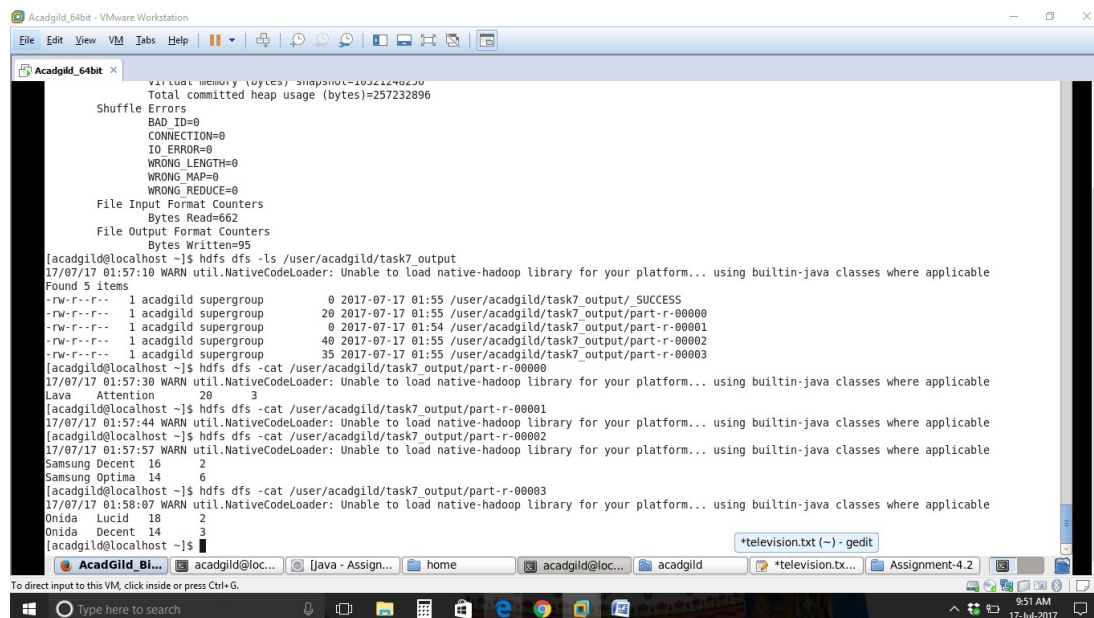


```
merge map outputs=1
GC time elapsed (ms)=217
CPU time spent (ms)=2170
Physical memory (bytes) snapshot=281694208
Virtual memory (bytes) snapshot=4123860992
Total committed heap usage (bytes)=165810176

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=662
File Output Format Counters
Bytes Written=16

[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/task6 output
17/07/17 01:50:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup 0 2017-07-17 00:51 /user/acadgild/task6 output/ SUCCESS
-rw-r--r-- 1 acadgild supergroup 16 2017-07-17 00:51 /user/acadgild/task6 output/part-r-00000
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task6 output/part-r-00000
17/07/17 01:50:42 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Uttar Pradesh 3
[acadgild@localhost ~]$ hadoop jar /home/acadgild/Assignment/Assignment-5.2/task7.jar /user/acadgild/validtv /user/acadgild/task7 output
17/07/17 01:51:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
17/07/17 01:51:24 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
17/07/17 01:51:28 WARN mapreduce.JobSubmitter: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
17/07/17 01:51:30 INFO input.FileInputFormat: Total input paths to process : 1
17/07/17 01:51:31 INFO mapreduce.JobSubmitter: number of splits:1
17/07/17 01:51:33 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1500193437658_0007
17/07/17 01:51:36 INFO impl.YarnClientImpl: Submitted application application_1500193437658_0007
acadgild@localhost:~$
```

Output of Task 7



```
Virtual memory (bytes) snapshot=10221240256
Total committed heap usage (bytes)=257232896

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=662
File Output Format Counters
Bytes Written=95

[acadgild@localhost ~]$ hdfs dfs -ls /user/acadgild/task7 output
17/07/17 01:57:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 5 items
-rw-r--r-- 1 acadgild supergroup 0 2017-07-17 01:55 /user/acadgild/task7 output/ SUCCESS
-rw-r--r-- 1 acadgild supergroup 20 2017-07-17 01:55 /user/acadgild/task7 output/part-r-00000
-rw-r--r-- 1 acadgild supergroup 0 2017-07-17 01:54 /user/acadgild/task7 output/part-r-00001
-rw-r--r-- 1 acadgild supergroup 40 2017-07-17 01:55 /user/acadgild/task7 output/part-r-00002
-rw-r--r-- 1 acadgild supergroup 35 2017-07-17 01:55 /user/acadgild/task7 output/part-r-00003
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task7 output/part-r-00000
17/07/17 01:57:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Lava Attention 20 3
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task7 output/part-r-00001
17/07/17 01:57:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task7 output/part-r-00002
17/07/17 01:57:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Samsung Decent 16 2
Samsung Optima 14 6
[acadgild@localhost ~]$ hdfs dfs -cat /user/acadgild/task7 output/part-r-00003
17/07/17 01:58:07 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Onida Lucid 18 2
Onida Decent 14 3
[acadgild@localhost ~]$
```