STUDENT NAME: LEE YU XIAN
STUDENT NUMBER: 7322164

```
bigdata@bigdata-VirtualBox:~/Desktop$ javac -cp $HADOOP_CLASSPATH
solution2.javabigdata@bigdata-VirtualBox:~/Desktop$ jar cvf solution2.jar
solution2*.classadded manifest
adding: solution2.class(in = 1485) (out= 811)(deflated 45%)
adding: solution2$MapReducer.class(in = 1820) (out= 797)(deflated 56%)
adding: solution2$TokenizerMapper.class(in = 1896) (out= 826)(deflated
56%)
bigdata@bigdata-VirtualBox:~/Desktop$ $HADOOP_HOME/bin/hadoop fs -rmr
task2outputrmr: DEPRECATED: Please use 'rm -r' instead.
21/10/12 01:46:28 INFO fs.TrashPolicyDefault: Namenode trash
configuration: Deletion interval = 0 minutes, Emptier interval = 0
minutes.
Deleted task2output
bigdata@bigdata-VirtualBox:~/Desktop$ $HADOOP_HOME/bin/hadoop jar
/home/bigdata/Desktop/solution2.jar solution2 task2.txt task2output
21/10/12 01:46:33 INFO client.RMProxy: Connecting to ResourceManager at
bigdata-VirtualBox/10.0.2.15:8032
21/10/12 01:46:33 WARN mapreduce.JobResourceUploader: Hadoop command-line
option parsing not performed. Implement the Tool interface and execute
your application with ToolRunner to remedy this.
21/10/12 01:46:34 INFO input.FileInputFormat: Total input paths to
process : 1
21/10/12 01:46:34 INFO mapreduce.JobSubmitter: number of splits:1
21/10/12 01:46:34 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1633958939443_0006
21/10/12 01:46:34 INFO impl.YarnClientImpl: Submitted application
application_1633958939443_0006
21/10/12 01:46:34 INFO mapreduce.Job: The url to track the job:
http://bigdata-VirtualBox:8088/proxy/application_1633958939443_0006/
21/10/12 01:46:34 INFO mapreduce.Job: Running job: job_1633958939443_0006
21/10/12 01:46:38 INFO mapreduce.Job: Job job_1633958939443_0006 running
in uber mode : false
21/10/12 01:46:38 INFO mapreduce.Job:  map 0% reduce 0%
21/10/12 01:46:41 INFO mapreduce.Job:  map 100% reduce 0%
21/10/12 01:46:45 INFO mapreduce.Job:  map 100% reduce 100%
21/10/12 01:46:46 INFO mapreduce.Job: Job job_1633958939443_0006
completed successfully
21/10/12 01:46:46 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=110
                FILE: Number of bytes written=237765
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=649
                HDFS: Number of bytes written=84
                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)=1333
                Total time spent by all reduces in occupied slots (ms)=1432
                Total time spent by all map tasks (ms)=1333
                Total time spent by all reduce tasks (ms)=1432
                Total vcore-milliseconds taken by all map tasks=1333
                Total vcore-milliseconds taken by all reduce tasks=1432
                Total megabyte-milliseconds taken by all map tasks=1364992
                Total megabyte-milliseconds taken by all reduce tasks=1466368
        Map-Reduce Framework
                Map input records=24
                Map output records=24
                Map output bytes=264
                Map output materialized bytes=110
                Input split bytes=109
                Combine input records=24
                Combine output records=8
                Reduce input groups=8
                Reduce shuffle bytes=110
                Reduce input records=8
                Reduce output records=8
                Spilled Records=16
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=70
                CPU time spent (ms)=690
                Physical memory (bytes) snapshot=441081856
                Virtual memory (bytes) snapshot=3820204032
                Total committed heap usage (bytes)=332398592
        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=540
        File Output Format Counters
                Bytes Written=84
bigdata@bigdata-VirtualBox:~/Desktop$ $HADOOP_HOME/bin/hadoop fs -ls
task2output
Found 2 items
-rw-r--r--   1 bigdata supergroup          0 2021-10-12 01:46
task2output/_SUCCESS
-rw-r--r--   1 bigdata supergroup         84 2021-10-12 01:46
task2output/part-r-00000
bigdata@bigdata-VirtualBox:~/Desktop$ $HADOOP_HOME/bin/hadoop fs -cat
task2output/part-r-00000
AAB123      121
AKS047      169
ASD654      69
```

```
PAK051      63
PKR856      93
PUS835      93
UPS234      148
UPS963      130
```