Exercise Sheet 05

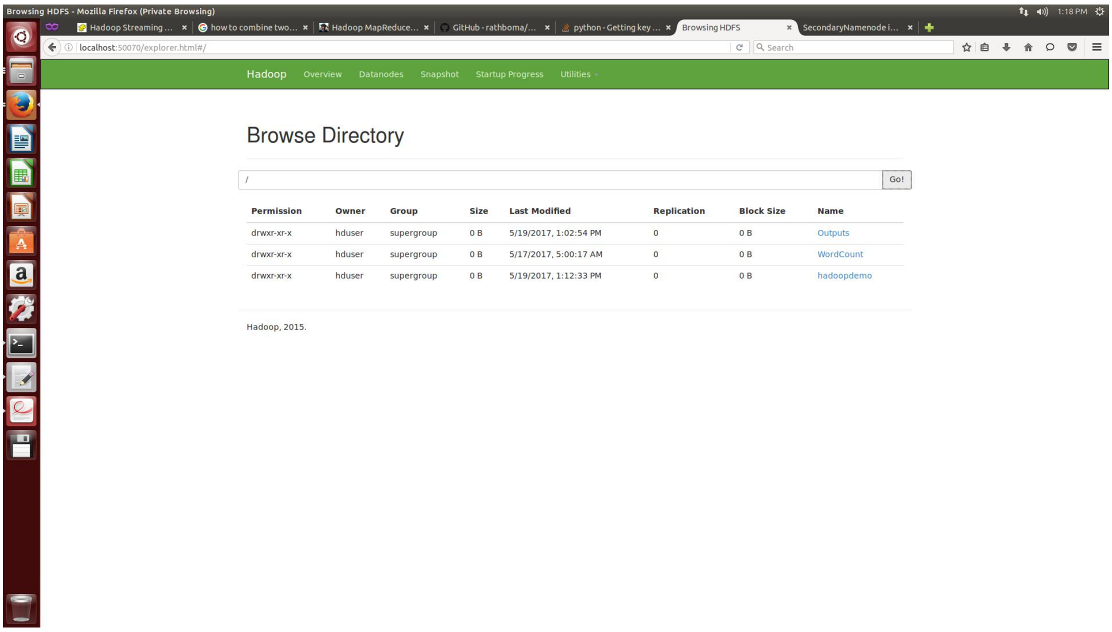
Syed Khalid Ahmed

276970

Exercise 1:

Exercise 1.1:

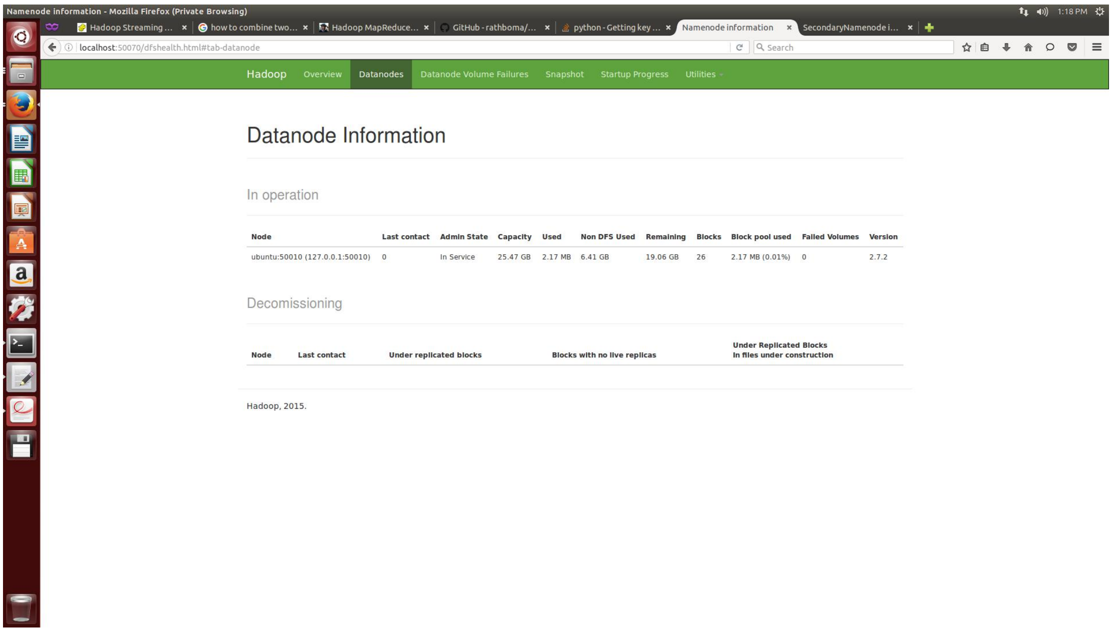
I have attached the snapshots of the working hadoop.



Browse Directory

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	hduser	supergroup	0 B	5/19/2017, 1:02:54 PM	0	0 B	Outputs
drwxr-xr-x	hduser	supergroup	0 B	5/17/2017, 5:00:17 AM	0	0 B	WordCount
drwxr-xr-x	hduser	supergroup	0 B	5/19/2017, 1:12:33 PM	0	0 B	hadoopdemo

Hadoop, 2015.



Datanode Information

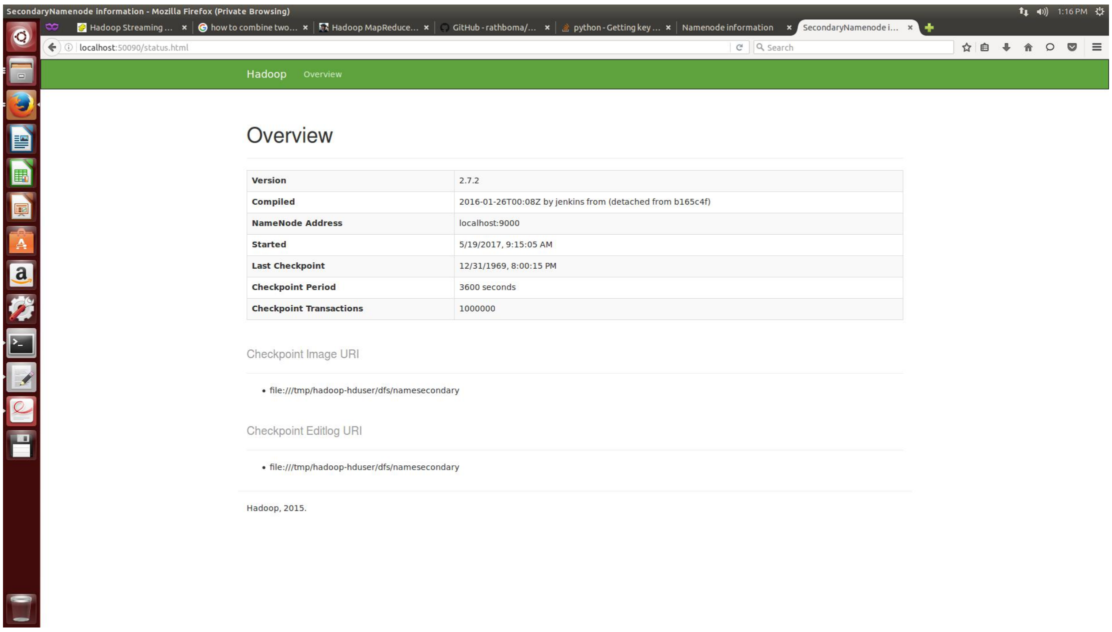
In operation

Node	Last contact	Admin State	Capacity	Used	Non DFS Used	Remaining	Blocks	Block pool used	Failed Volumes	Version
ubuntu:50010 (127.0.0.1:50010)	0	In Service	25.47 GB	2.17 MB	6.41 GB	19.06 GB	26	2.17 MB (0.01%)	0	2.7.2

Decomissioning

Node	Last contact	Under replicated blocks	Blocks with no live replicas	Under Replicated Blocks In files under construction

Hadoop, 2015.



Overview

Version	2.7.2
Compiled	2016-01-26T00:08Z by jenkins from (detached from b165c4f)
NameNode Address	localhost:9000
Started	5/19/2017, 9:15:05 AM
Last Checkpoint	12/31/1969, 8:00:15 PM
Checkpoint Period	3600 seconds
Checkpoint Transactions	1000000

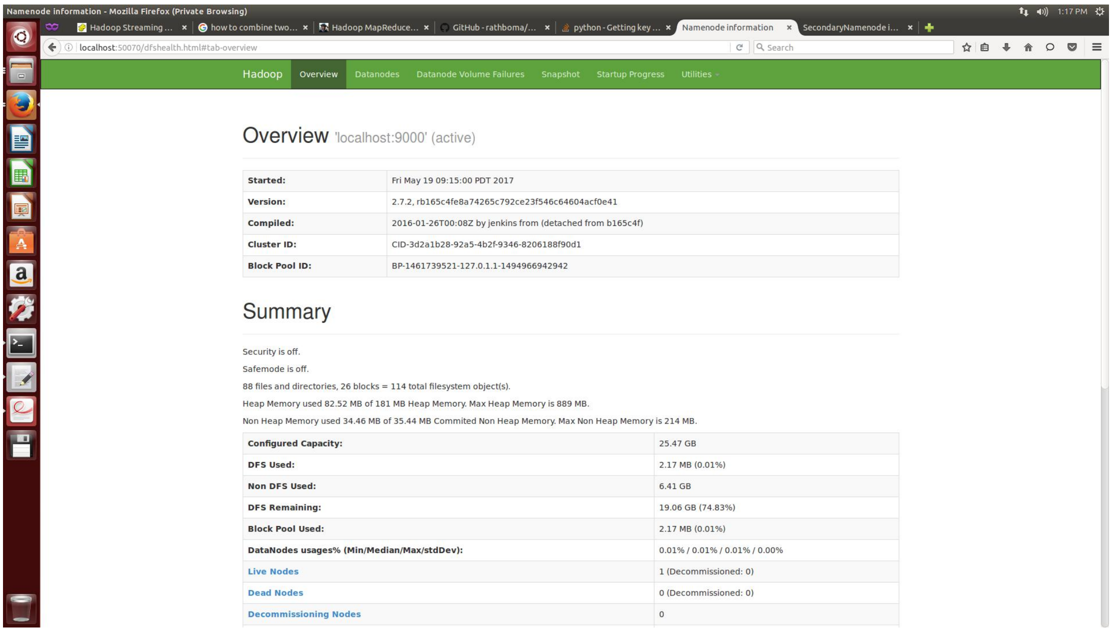
Checkpoint Image URI

- file:///tmp/hadoop-hduser/dfs/namesecondary

Checkpoint Editlog URI

- file:///tmp/hadoop-hduser/dfs/namesecondary

Hadoop, 2015.



Overview 'localhost:9000' (active)

Started:	Fri May 19 09:15:00 PDT 2017
Version:	2.7.2, rb165c4fe8a74265c792ce23f546c64604acf0e41
Compiled:	2016-01-26T00:08Z by jenkins from (detached from b165c4f)
Cluster ID:	CID-3d2a1b28-92a5-4b2f-9346-8206188f90d1
Block Pool ID:	BP-1461739521-127.0.1.1-1494966942942

Summary

Security is off.

Safemode is off.

88 files and directories, 26 blocks = 114 total filesystem object(s).

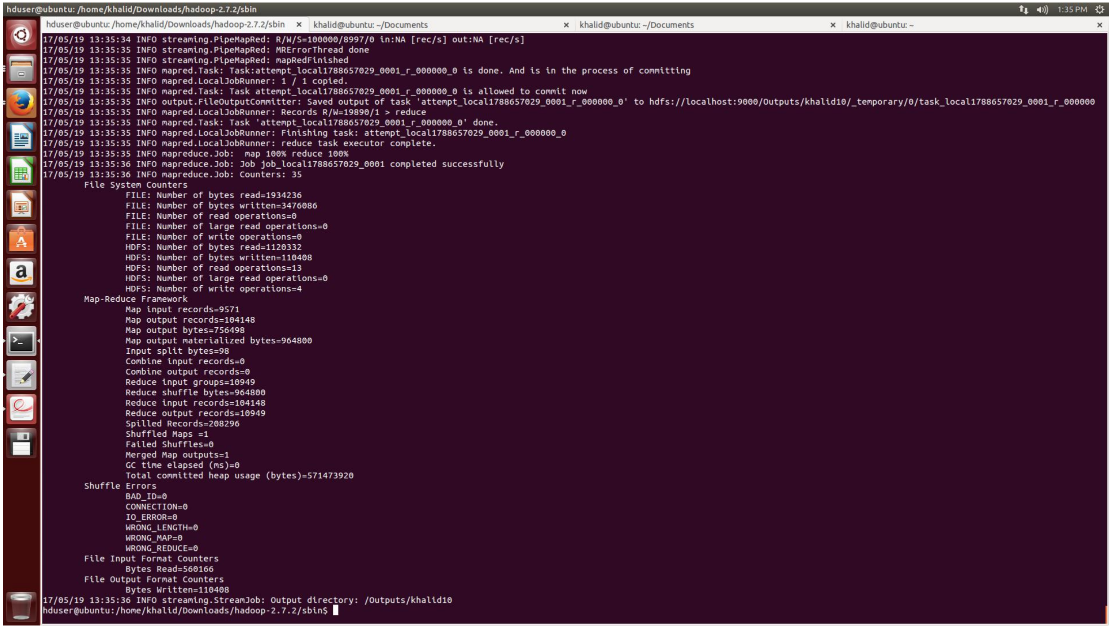
Heap Memory used 82.52 MB of 181 MB Heap Memory. Max Heap Memory is 889 MB.

Non Heap Memory used 34.46 MB of 35.44 MB Commited Non Heap Memory. Max Non Heap Memory is 214 MB.

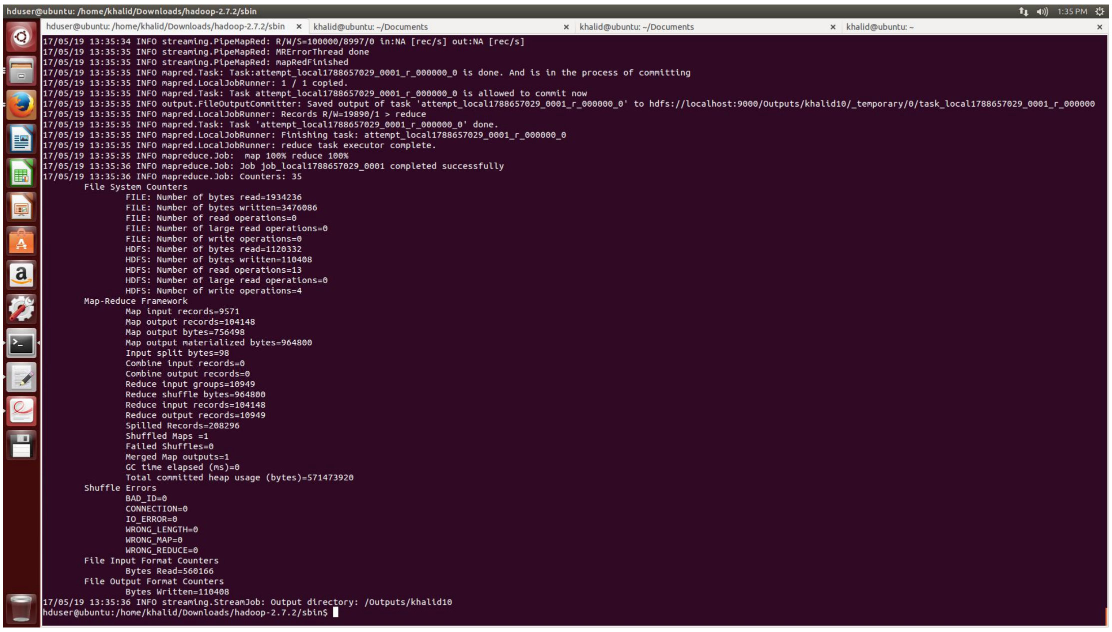
Configured Capacity:	25.47 GB
DFS Used:	2.17 MB (0.01%)
Non DFS Used:	6.41 GB
DFS Remaining:	19.06 GB (74.83%)
Block Pool Used:	2.17 MB (0.01%)
DataNodes usages% (Min/Median/Max/stdDev):	0.01% / 0.01% / 0.01% / 0.00%
Live Nodes	1 (Decommissioned: 0)
Dead Nodes	0 (Decommissioned: 0)
Decommissioning Nodes	0

Exercise 1.2 : Warmup Exercise

I have attached the code files in a separate folder. Below are the screenshots


hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
1:35 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

```
17/05/19 13:35:34 INFO streaming.PipeMapRed: R/W/S=100000/8997/0 in:NA [rec/s] out:NA [rec/s]
17/05/19 13:35:35 INFO streaming.PipeMapRed: MRErrorThread done
17/05/19 13:35:35 INFO streaming.PipeMapRed: mapRedFinished
17/05/19 13:35:35 INFO mapred.Task: Task:attempt_local1788657029_0001_r_000000_0 is done. And is in the process of committing
17/05/19 13:35:35 INFO mapred.LocalJobRunner: 1 / 1 copied.
17/05/19 13:35:35 INFO mapred.Task: Task attempt_local1788657029_0001_r_000000_0 is allowed to commit now
17/05/19 13:35:35 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1788657029_0001_r_000000_0' to hdfs://localhost:9000/Outputs/khalid10/_temporary/0/task_local1788657029_0001_r_000000
17/05/19 13:35:35 INFO mapred.LocalJobRunner: Records R/W=19890/1 > reduce
17/05/19 13:35:35 INFO mapred.Task: Task 'attempt_local1788657029_0001_r_000000_0' done.
17/05/19 13:35:35 INFO mapred.LocalJobRunner: Finishing task: attempt_local1788657029_0001_r_000000_0
17/05/19 13:35:35 INFO mapred.LocalJobRunner: reduce task executor complete.
17/05/19 13:35:35 INFO mapreduce.Job:  map 100% reduce 100%
17/05/19 13:35:36 INFO mapreduce.Job: Job job_local1788657029_0001 completed successfully
17/05/19 13:35:36 INFO mapreduce.Job: Counters: 35
        File System Counters
                FILE: Number of bytes read=1934236
                FILE: Number of bytes written=3476086
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=1120332
                HDFS: Number of bytes written=110408
                HDFS: Number of read operations=13
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=4
        Map-Reduce Framework
                Map input records=9571
                Map output records=104148
                Map output bytes=756498
                Map output materialized bytes=964800
                Input split bytes=98
                Combine input records=0
                Combine output records=0
                Reduce input groups=10949
                Reduce shuffle bytes=964800
                Reduce input records=104148
                Reduce output records=10949
                Spilled Records=208296
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=0
                Total committed heap usage (bytes)=571473920
        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=560166
        File Output Format Counters
                Bytes Written=110408
17/05/19 13:35:36 INFO streaming.StreamJob: Output directory: /Outputs/khalid10
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$
```


hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
1:35 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

```
17/05/19 13:35:34 INFO streaming.PipeMapRed: R/W/S=100000/8997/0 in:NA [rec/s] out:NA [rec/s]
17/05/19 13:35:35 INFO streaming.PipeMapRed: MRErrorThread done
17/05/19 13:35:35 INFO streaming.PipeMapRed: mapRedFinished
17/05/19 13:35:35 INFO mapred.Task: Task:attempt_local1788657029_0001_r_000000_0 is done. And is in the process of committing
17/05/19 13:35:35 INFO mapred.LocalJobRunner: 1 / 1 copied.
17/05/19 13:35:35 INFO mapred.Task: Task attempt_local1788657029_0001_r_000000_0 is allowed to commit now
17/05/19 13:35:35 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1788657029_0001_r_000000_0' to hdfs://localhost:9000/Outputs/khalid10/_temporary/0/task_local1788657029_0001_r_000000
17/05/19 13:35:35 INFO mapred.LocalJobRunner: Records R/W=19890/1 > reduce
17/05/19 13:35:35 INFO mapred.Task: Task 'attempt_local1788657029_0001_r_000000_0' done.
17/05/19 13:35:35 INFO mapred.LocalJobRunner: Finishing task: attempt_local1788657029_0001_r_000000_0
17/05/19 13:35:35 INFO mapred.LocalJobRunner: reduce task executor complete.
17/05/19 13:35:35 INFO mapreduce.Job:  map 100% reduce 100%
17/05/19 13:35:36 INFO mapreduce.Job: Job job_local1788657029_0001 completed successfully
17/05/19 13:35:36 INFO mapreduce.Job: Counters: 35
        File System Counters
                FILE: Number of bytes read=1934236
                FILE: Number of bytes written=3476086
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=1120332
                HDFS: Number of bytes written=110408
                HDFS: Number of read operations=13
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=4
        Map-Reduce Framework
                Map input records=9571
                Map output records=104148
                Map output bytes=756498
                Map output materialized bytes=964800
                Input split bytes=98
                Combine input records=0
                Combine output records=0
                Reduce input groups=10949
                Reduce shuffle bytes=964800
                Reduce input records=104148
                Reduce output records=10949
                Spilled Records=208296
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=0
                Total committed heap usage (bytes)=571473920
        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=560166
        File Output Format Counters
                Bytes Written=110408
17/05/19 13:35:36 INFO streaming.StreamJob: Output directory: /Outputs/khalid10
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$
```

Exercise 2:

I have used the hadoop streaming api to perform the mapreduce operation.

The mapper and reduce code files for each question is in the respective directories.

Part 1.

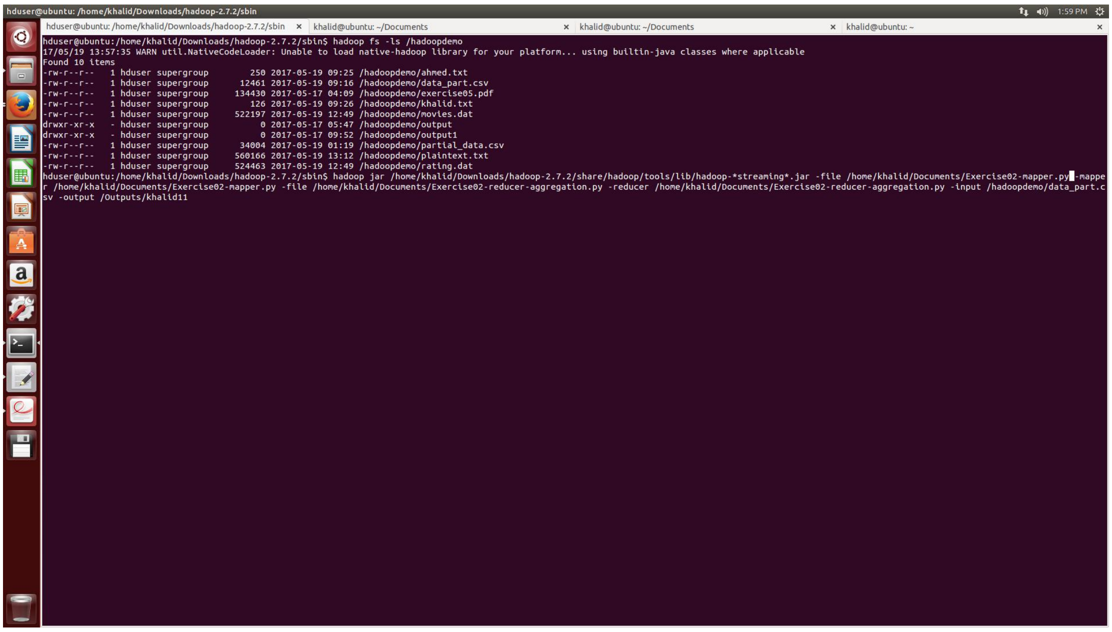
1) First I uploaded the relevant dataset to the HDFS using the following command.

hadoop fs -put /home/khalid/Downloads/ml-10M100K/*.dat /hadoopdemo

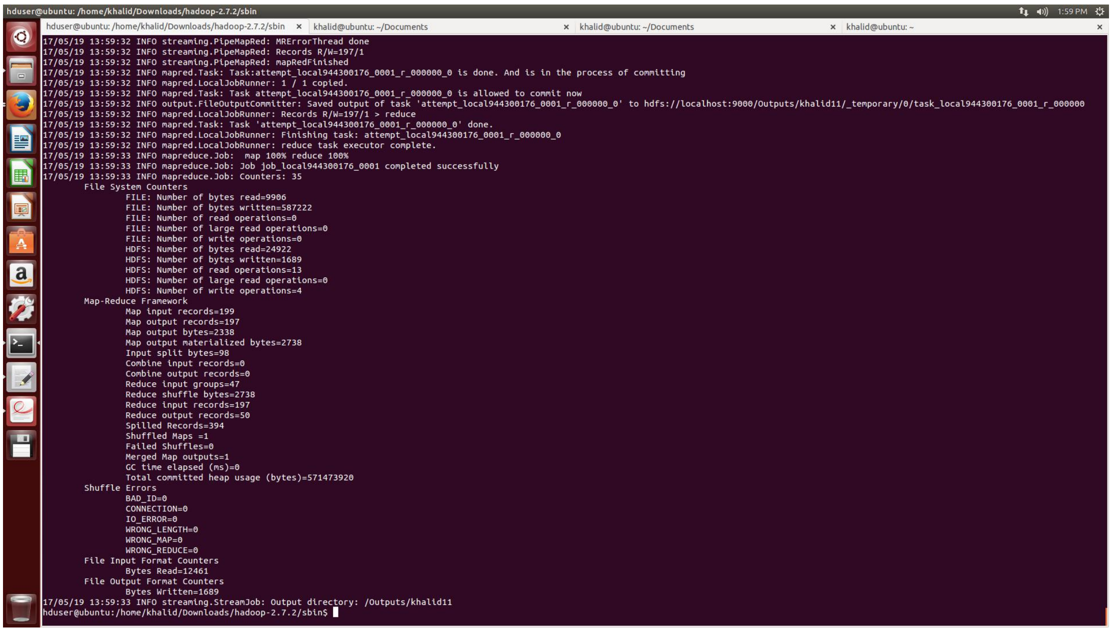
2) Then I ran the mapreduce job with the following command

hadoop jar /home/khalid/Downloads/hadoop-2.7.2/share/hadoop/tools/lib/hadoop-*streaming*.jar -file /home/khalid/Documents/Exercise02-mapper-2.py -mapper /home/khalid/Documents/Exercise02-mapper-2.py -file /home/khalid/Documents/Exercise02-reducer-2.py -reducer /home/khalid/Documents/Exercise02-reducer-2.py -input /hadoopdemo/FlightData.csv   -output /Outputs/khalid13

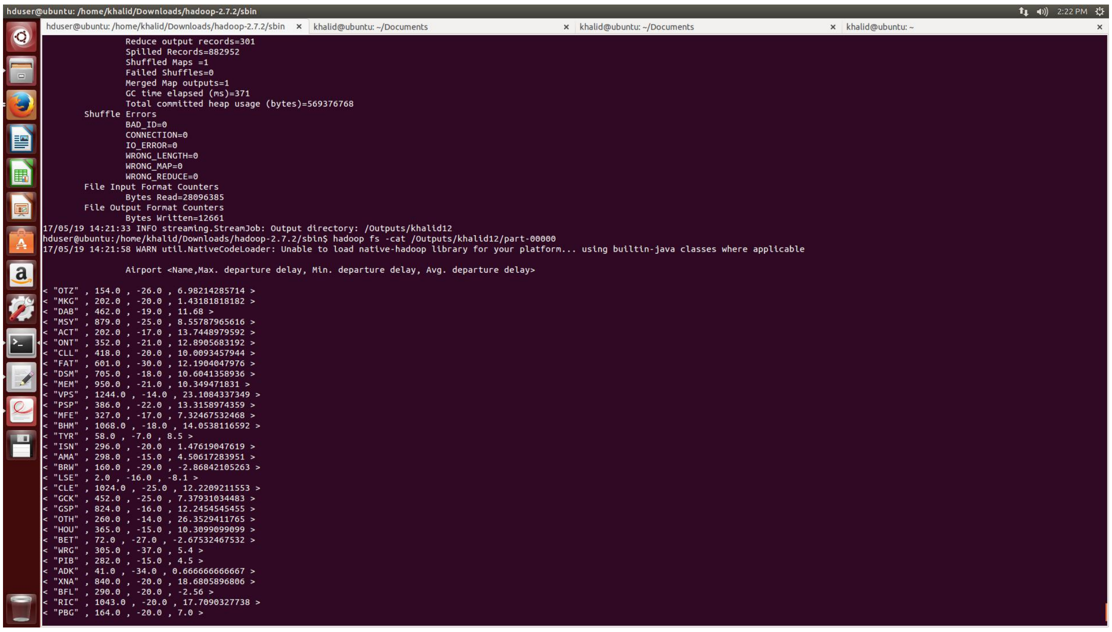
I have attached the screenshots for the code.


hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~
1:59 PM

```
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$ hadoop fs -ls /hadoopdemo
17/05/19 13:57:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 10 items
-rw-r--r--   1 hduser supergroup        250 2017-05-19 09:25 /hadoopdemo/ahmed.txt
-rw-r--r--   1 hduser supergroup      12461 2017-05-19 09:16 /hadoopdemo/data_part.csv
-rw-r--r--   1 hduser supergroup     134430 2017-05-17 04:09 /hadoopdemo/exercise05.pdf
-rw-r--r--   1 hduser supergroup        126 2017-05-19 09:26 /hadoopdemo/khalid.txt
-rw-r--r--   1 hduser supergroup     522197 2017-05-19 12:49 /hadoopdemo/movies.dat
drwxr-xr-x   - hduser supergroup          0 2017-05-17 05:47 /hadoopdemo/output
drwxr-xr-x   - hduser supergroup          0 2017-05-17 09:52 /hadoopdemo/output1
-rw-r--r--   1 hduser supergroup      34004 2017-05-19 01:19 /hadoopdemo/partial_data.csv
-rw-r--r--   1 hduser supergroup     560166 2017-05-19 13:12 /hadoopdemo/plaintext.txt
-rw-r--r--   1 hduser supergroup     524463 2017-05-19 12:49 /hadoopdemo/rating.dat
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$ hadoop jar /home/khalid/Downloads/hadoop-2.7.2/share/hadoop/tools/lib/hadoop-*streaming*.jar -file /home/khalid/Documents/Exercise02-mapper.py -mapper /home/khalid/Documents/Exercise02-mapper.py -file /home/khalid/Documents/Exercise02-reducer-aggregation.py -reducer /home/khalid/Documents/Exercise02-reducer-aggregation.py -input /hadoopdemo/data_part.csv -output /Outputs/khalid11
```


hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
1:59 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

```
17/05/19 13:59:32 INFO streaming.PipeMapRed: MRErrorThread done
17/05/19 13:59:32 INFO streaming.PipeMapRed: Records R/W=197/1
17/05/19 13:59:32 INFO streaming.PipeMapRed: mapRedFinished
17/05/19 13:59:32 INFO mapred.Task: Task:attempt_local944300176_0001_r_000000_0 is done. And is in the process of committing
17/05/19 13:59:32 INFO mapred.LocalJobRunner: 1 / 1 copied.
17/05/19 13:59:32 INFO mapred.Task: Task attempt_local944300176_0001_r_000000_0 is allowed to commit now
17/05/19 13:59:32 INFO output.FileOutputCommitter: Saved output of task 'attempt_local944300176_0001_r_000000_0' to hdfs://localhost:9000/Outputs/khalid11/_temporary/0/task_local944300176_0001_r_000000
17/05/19 13:59:32 INFO mapred.LocalJobRunner: Records R/W=197/1 > reduce
17/05/19 13:59:32 INFO mapred.Task: Task 'attempt_local944300176_0001_r_000000_0' done.
17/05/19 13:59:32 INFO mapred.LocalJobRunner: Finishing task: attempt_local944300176_0001_r_000000_0
17/05/19 13:59:32 INFO mapred.LocalJobRunner: reduce task executor complete.
17/05/19 13:59:33 INFO mapreduce.Job:  map 100% reduce 100%
17/05/19 13:59:33 INFO mapreduce.Job: Job job_local944300176_0001 completed successfully
17/05/19 13:59:33 INFO mapreduce.Job: Counters: 35
        File System Counters
                FILE: Number of bytes read=9906
                FILE: Number of bytes written=587222
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=24922
                HDFS: Number of bytes written=1689
                HDFS: Number of read operations=13
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=4
        Map-Reduce Framework
                Map input records=199
                Map output records=197
                Map output bytes=2338
                Map output materialized bytes=2738
                Input split bytes=98
                Combine input records=0
                Combine output records=0
                Reduce input groups=47
                Reduce shuffle bytes=2738
                Reduce input records=197
                Reduce output records=50
                Spilled Records=394
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=0
                Total committed heap usage (bytes)=571473920
        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=12461
        File Output Format Counters
                Bytes Written=1689
17/05/19 13:59:33 INFO streaming.StreamJob: Output directory: /Outputs/khalid11
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$
```

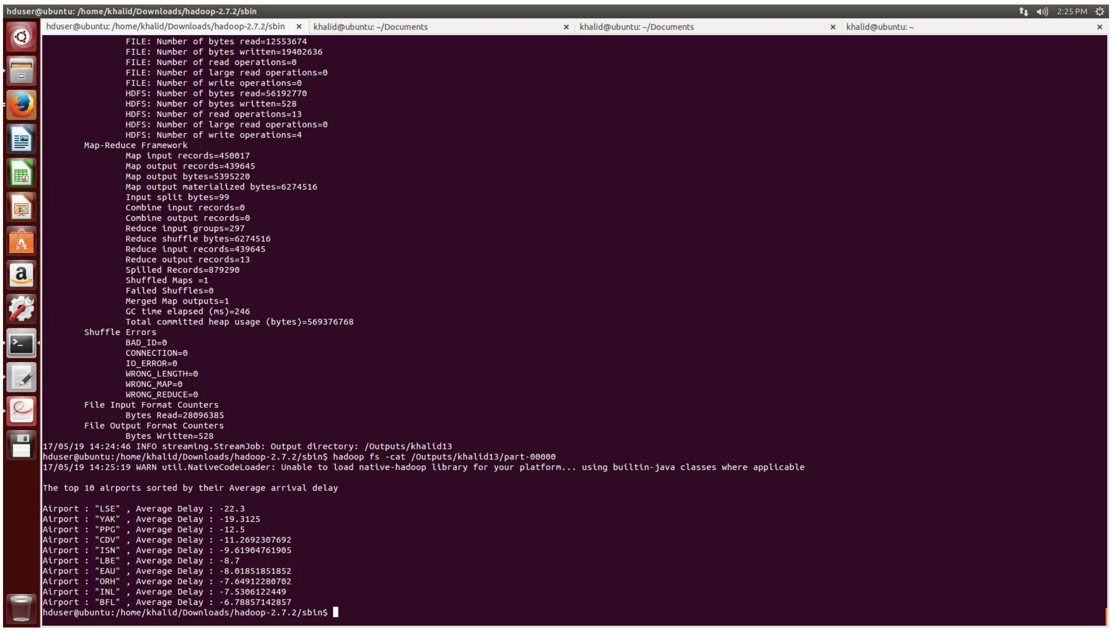

hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
2:22 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

```
                Reduce output records=301
                Spilled Records=882952
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=371
                Total committed heap usage (bytes)=569376768
        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=28096385
        File Output Format Counters
                Bytes Written=12661
17/05/19 14:21:33 INFO streaming.StreamJob: Output directory: /Outputs/khalid12
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$ hadoop fs -cat /Outputs/khalid12/part-00000
17/05/19 14:21:58 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

                Airport <Name,Max. departure delay, Min. departure delay, Avg. departure delay>

< "OTZ" , 154.0 , -26.0 , 6.98214285714 >
< "MKG" , 202.0 , -20.0 , 1.43181818182 >
< "DAB" , 462.0 , -19.0 , 11.68 >
< "MSY" , 879.0 , -25.0 , 8.55787965616 >
< "ACT" , 202.0 , -17.0 , 13.7448979592 >
< "ONT" , 352.0 , -21.0 , 12.8905683192 >
< "CLL" , 418.0 , -20.0 , 10.0093457944 >
< "FAT" , 601.0 , -30.0 , 12.1904047976 >
< "DSM" , 705.0 , -18.0 , 10.6041358936 >
< "MEM" , 950.0 , -21.0 , 10.349471831 >
< "VPS" , 1244.0 , -14.0 , 23.1084337349 >
< "PSP" , 386.0 , -22.0 , 13.3158974359 >
< "MFE" , 327.0 , -17.0 , 7.32467532468 >
< "BHM" , 1068.0 , -18.0 , 14.0538116592 >
< "TYR" , 58.0 , -7.0 , 8.5 >
< "ISN" , 296.0 , -20.0 , 1.47619047619 >
< "AMA" , 298.0 , -15.0 , 4.50617283951 >
< "BRW" , 160.0 , -29.0 , -2.86842105263 >
< "LSE" , 2.0 , -16.0 , -8.1 >
< "CLE" , 1024.0 , -25.0 , 12.2209211553 >
< "GCK" , 452.0 , -25.0 , 7.37931034483 >
< "GSP" , 824.0 , -16.0 , 12.2454545455 >
< "OTH" , 260.0 , -14.0 , 26.3529411765 >
< "HOU" , 365.0 , -15.0 , 10.3099099099 >
< "BET" , 72.0 , -27.0 , -2.67532467532 >
< "WRG" , 305.0 , -37.0 , 5.4 >
< "PIB" , 282.0 , -15.0 , 4.5 >
< "ADK" , 41.0 , -34.0 , 0.666666666667 >
< "XNA" , 840.0 , -20.0 , 18.6805896806 >
< "BFL" , 290.0 , -20.0 , -2.56 >
< "RIC" , 1043.0 , -20.0 , 17.7090327738 >
< "PBG" , 164.0 , -20.0 , 7.0 >
```

Part 2:



hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
2:25 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

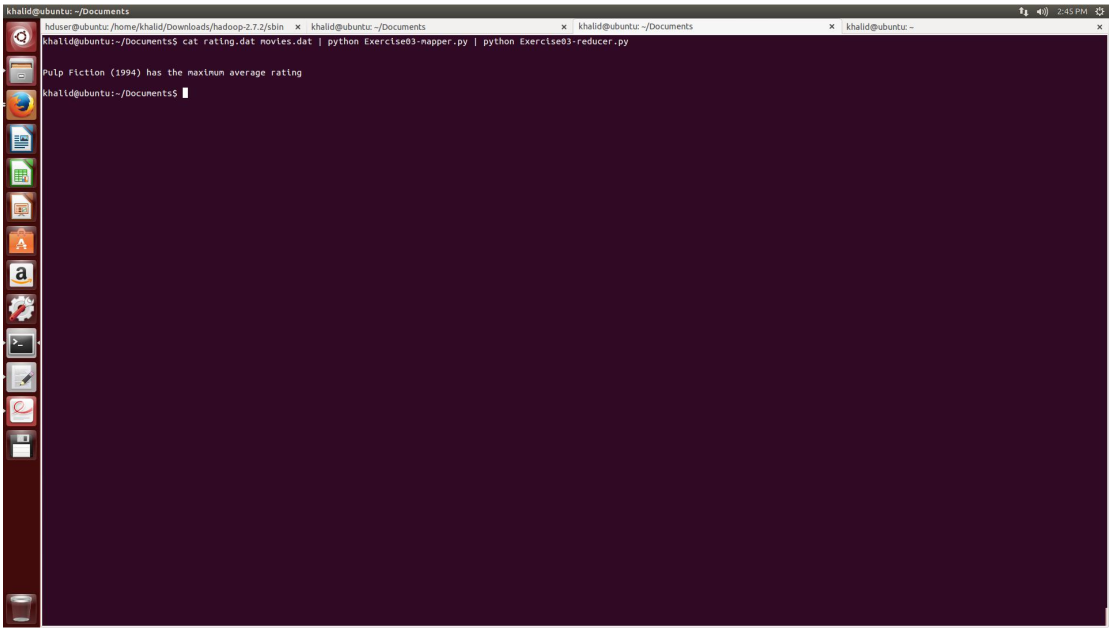
```
                FILE: Number of bytes read=12553674
                FILE: Number of bytes written=19402636
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=56192770
                HDFS: Number of bytes written=528
                HDFS: Number of read operations=13
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=4
        Map-Reduce Framework
                Map input records=450017
                Map output records=439645
                Map output bytes=5395220
                Map output materialized bytes=6274516
                Input split bytes=99
                Combine input records=0
                Combine output records=0
                Reduce input groups=297
                Reduce shuffle bytes=6274516
                Reduce input records=439645
                Reduce output records=13
                Spilled Records=879290
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=246
                Total committed heap usage (bytes)=569376768
        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=28096385
        File Output Format Counters
                Bytes Written=528
17/05/19 14:24:46 INFO streaming.StreamJob: Output directory: /Outputs/khalid13
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$ hadoop fs -cat /Outputs/khalid13/part-00000
17/05/19 14:25:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

The top 10 airports sorted by their Average arrival delay

Airport : "LSE" , Average Delay : -22.3
Airport : "YAK" , Average Delay : -19.3125
Airport : "PPG" , Average Delay : -12.5
Airport : "CDV" , Average Delay : -11.2692307692
Airport : "ISN" , Average Delay : -9.61904761905
Airport : "LBE" , Average Delay : -8.7
Airport : "EAU" , Average Delay : -8.01851851852
Airport : "ORH" , Average Delay : -7.64912280702
Airport : "INL" , Average Delay : -7.5306122449
Airport : "BFL" , Average Delay : -6.78857142857
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$
```

Exercise 3:

Part 1.



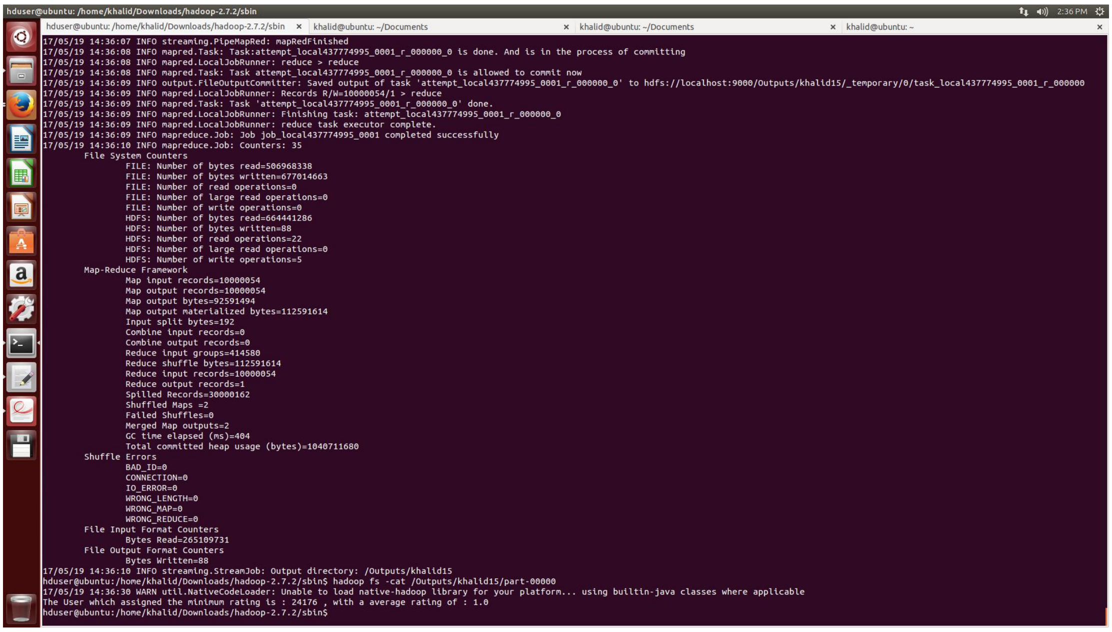
khalid@ubuntu: ~/Documents
2:45 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

```
khalid@ubuntu:~/Documents$ cat rating.dat movies.dat | python Exercise03-mapper.py | python Exercise03-reducer.py


Pulp Fiction (1994) has the maximum average rating

khalid@ubuntu:~/Documents$
```

<u>Part 2.</u>



hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
2:36 PM
hduser@ubuntu: /home/khalid/Downloads/hadoop-2.7.2/sbin
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~/Documents
khalid@ubuntu: ~

```
17/05/19 14:36:07 INFO streaming.PipeMapRed: mapRedFinished
17/05/19 14:36:08 INFO mapred.Task: Task:attempt_local437774995_0001_r_000000_0 is done. And is in the process of committing
17/05/19 14:36:08 INFO mapred.LocalJobRunner: reduce > reduce
17/05/19 14:36:08 INFO mapred.Task: Task attempt_local437774995_0001_r_000000_0 is allowed to commit now
17/05/19 14:36:09 INFO output.FileOutputCommitter: Saved output of task 'attempt_local437774995_0001_r_000000_0' to hdfs://localhost:9000/Outputs/khalid15/_temporary/0/task_local437774995_0001_r_000000
17/05/19 14:36:09 INFO mapred.LocalJobRunner: Records R/W=10000054/1 > reduce
17/05/19 14:36:09 INFO mapred.Task: Task 'attempt_local437774995_0001_r_000000_0' done.
17/05/19 14:36:09 INFO mapred.LocalJobRunner: Finishing task: attempt_local437774995_0001_r_000000_0
17/05/19 14:36:09 INFO mapred.LocalJobRunner: reduce task executor complete.
17/05/19 14:36:09 INFO mapreduce.Job: Job job_local437774995_0001 completed successfully
17/05/19 14:36:10 INFO mapreduce.Job: Counters: 35
        File System Counters
                FILE: Number of bytes read=506968338
                FILE: Number of bytes written=677014663
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=664441286
                HDFS: Number of bytes written=88
                HDFS: Number of read operations=22
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=5
        Map-Reduce Framework
                Map input records=10000054
                Map output records=10000054
                Map output bytes=92591494
                Map output materialized bytes=112591614
                Input split bytes=192
                Combine input records=0
                Combine output records=0
                Reduce input groups=414580
                Reduce shuffle bytes=112591614
                Reduce input records=10000054
                Reduce output records=1
                Spilled Records=30000162
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=404
                Total committed heap usage (bytes)=1040711680
        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=265109731
        File Output Format Counters
                Bytes Written=88
17/05/19 14:36:10 INFO streaming.StreamJob: Output directory: /Outputs/khalid15
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$ hadoop fs -cat /Outputs/khalid15/part-00000
17/05/19 14:36:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
The User which assigned the minimum rating is : 24176 , with a average rating of : 1.0
hduser@ubuntu:/home/khalid/Downloads/hadoop-2.7.2/sbin$
```