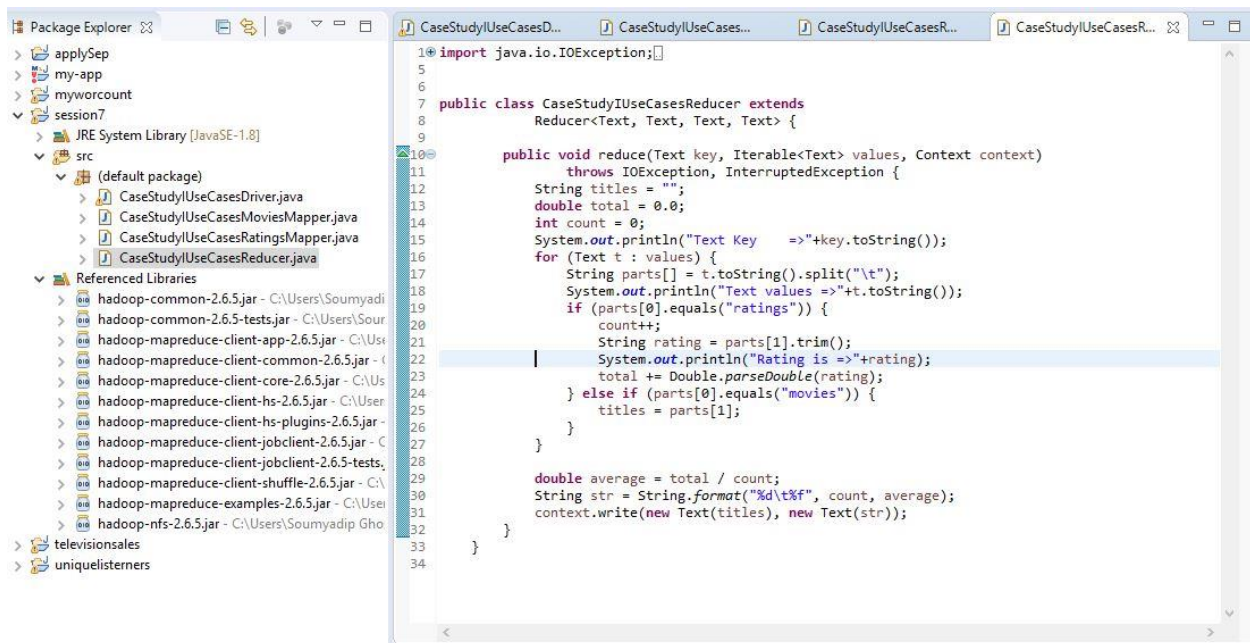
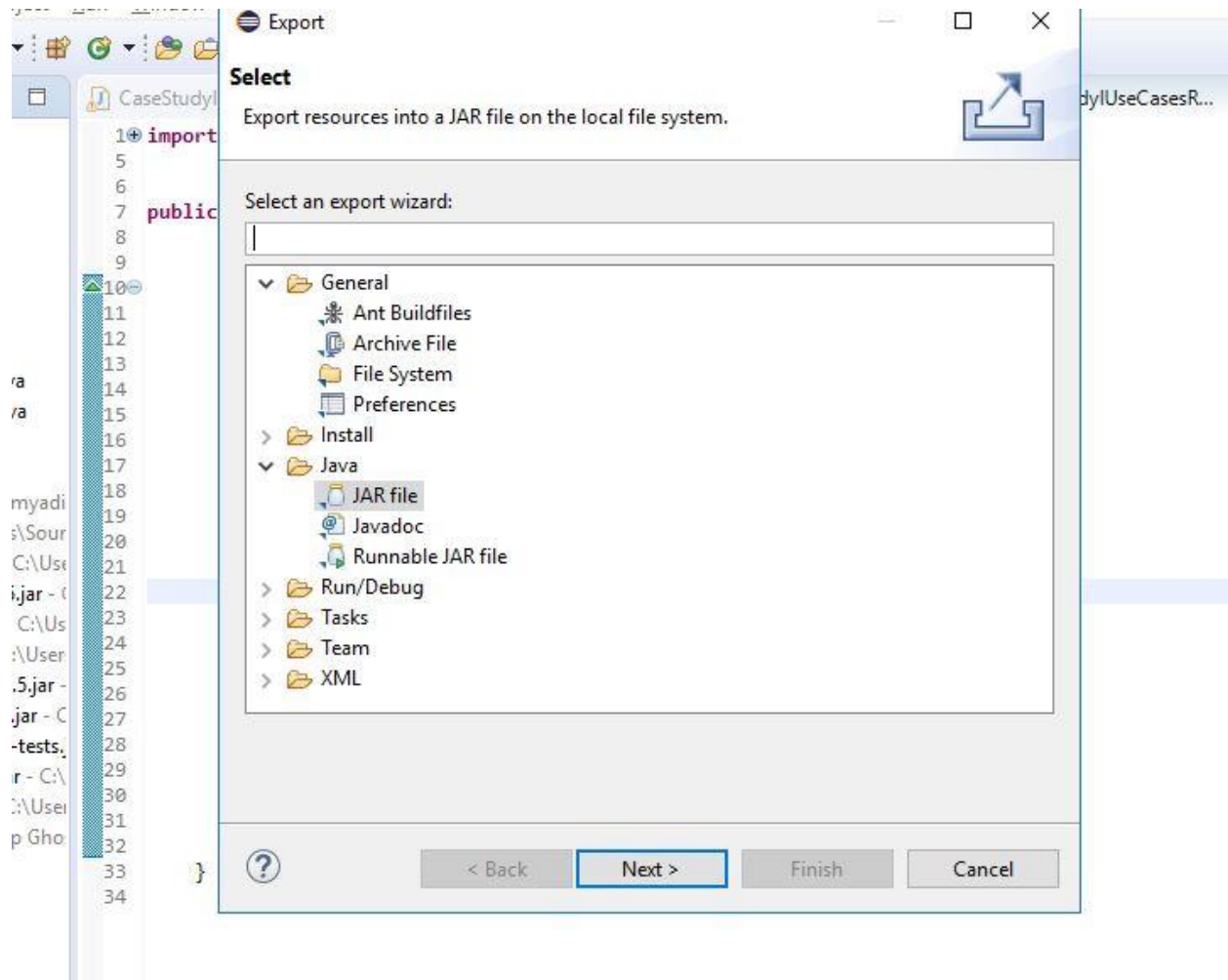


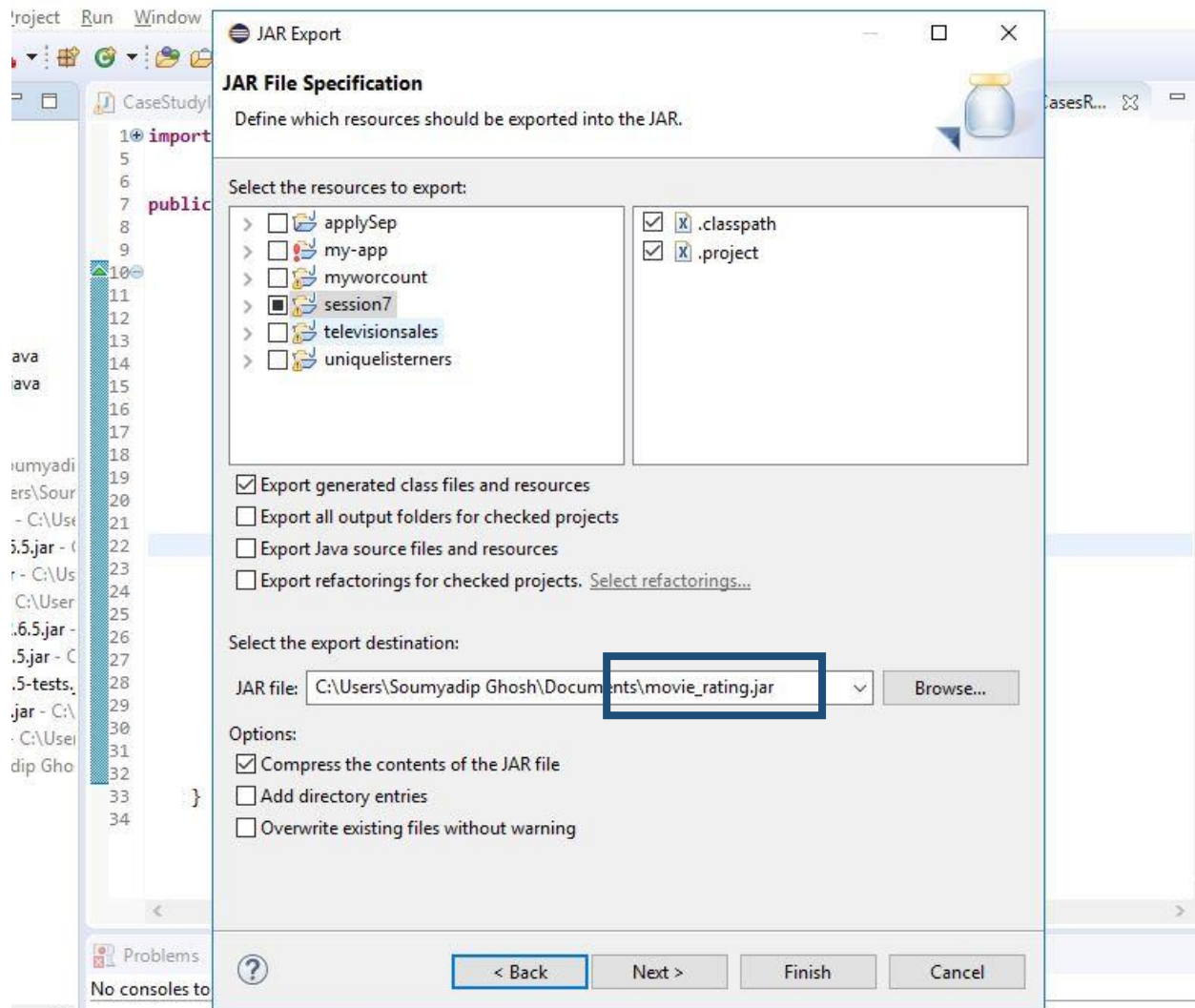
Adding source code and necessary jars :



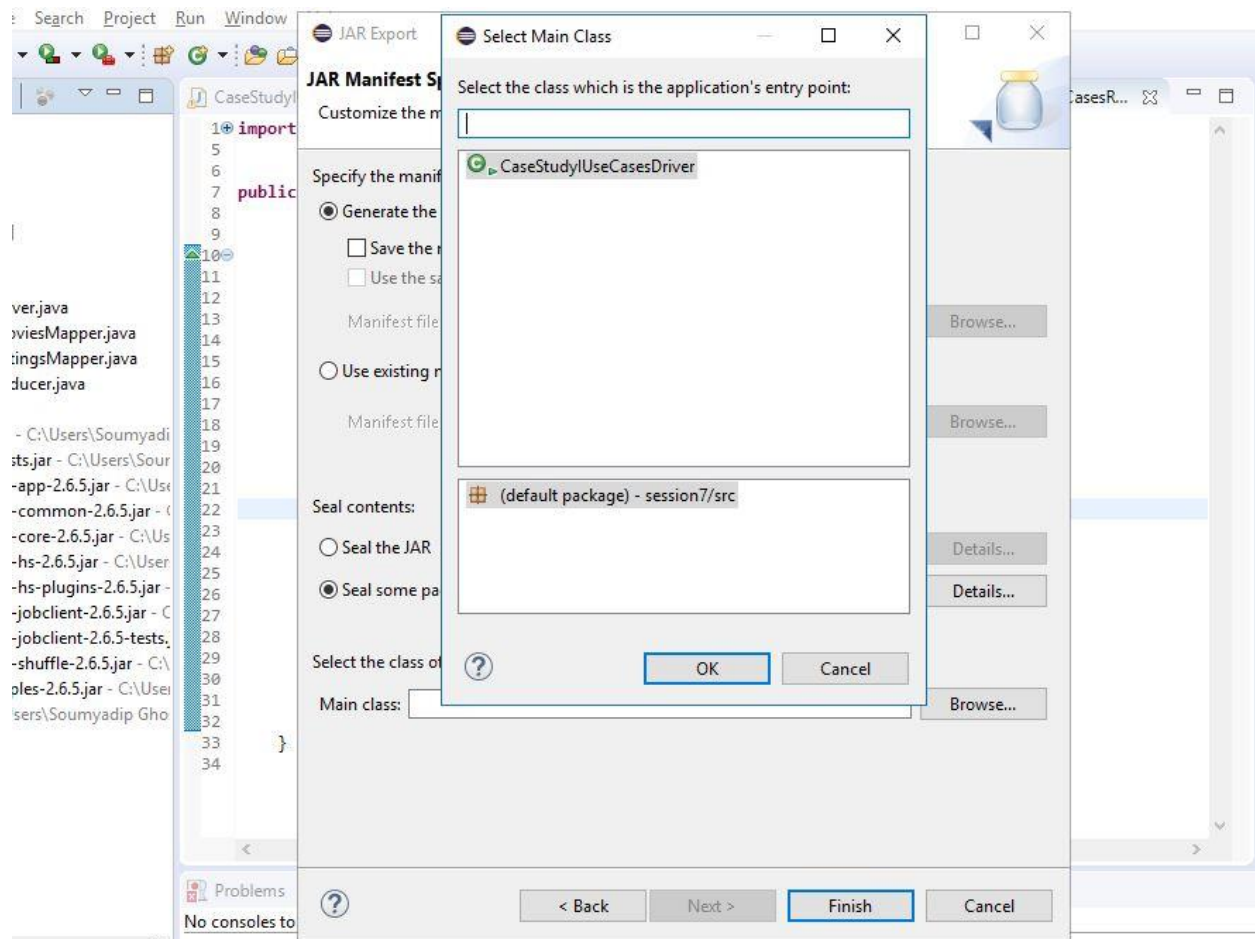
Exporting jar file :



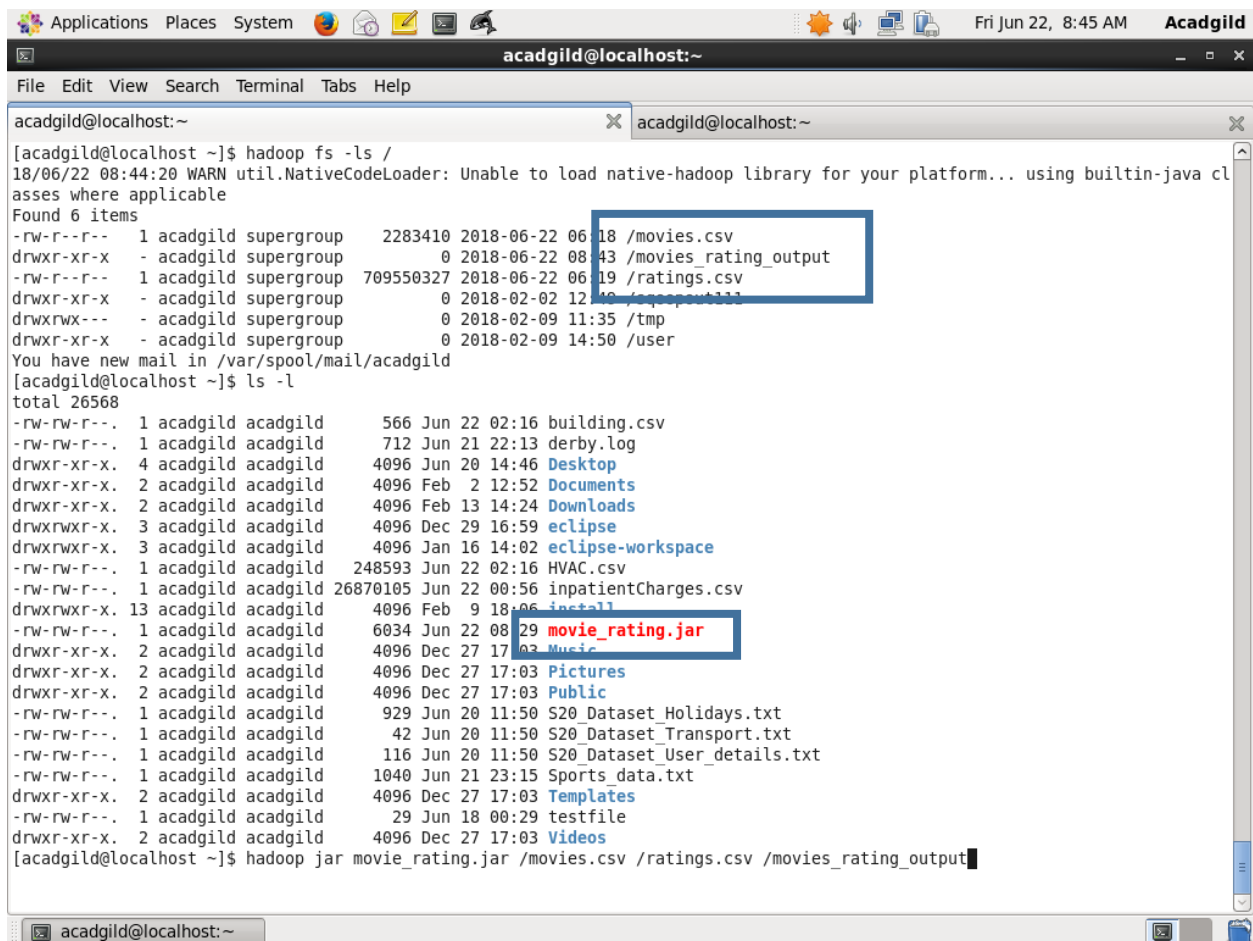
Add name for the jar file and click next :



Select driver class as the main class and click on finish :



Running job:



The screenshot shows a terminal window titled 'acadgild@localhost:~'. The user has executed two commands: 'hadoop fs -ls /' and 'ls -l'. The first command lists files in the Hadoop file system, showing a directory structure with files like 'movies.csv', 'movies_rating_output', and 'ratings.csv'. The second command lists files in the local file system, showing a directory structure with files like 'building.csv', 'derby.log', 'Desktop', 'Documents', 'Downloads', 'eclipse', 'eclipse-workspace', 'HVAC.csv', 'inpatientCharges.csv', 'install', 'movie_rating.jar', 'Music', 'Pictures', 'Public', 'S20_Dataset_Holidays.txt', 'S20_Dataset_Transport.txt', 'S20_Dataset_User_details.txt', 'Sports_data.txt', 'Templates', 'testfile', and 'Videos'. The 'movie_rating.jar' file is highlighted with a red box. The terminal window also shows a warning message about the native-hadoop library and a notification about new mail.

```
acadgild@localhost:~  
[acadgild@localhost ~]$ hadoop fs -ls /  
18/06/22 08:44:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl  
asses where applicable  
Found 6 items  
-rw-r--r-- 1 acadgild supergroup 2283410 2018-06-22 06:18 /movies.csv  
drwxr-xr-x - acadgild supergroup 0 2018-06-22 08:43 /movies_rating_output  
-rw-r--r-- 1 acadgild supergroup 709550327 2018-06-22 06:19 /ratings.csv  
drwxr-xr-x - acadgild supergroup 0 2018-02-02 12:40 /tmp  
drwxrwx--- - acadgild supergroup 0 2018-02-09 11:35 /tmp  
drwxr-xr-x - acadgild supergroup 0 2018-02-09 14:50 /user  
You have new mail in /var/spool/mail/acadgild  
[acadgild@localhost ~]$ ls -l  
total 26568  
-rw-rw-r-- 1 acadgild acadgild 566 Jun 22 02:16 building.csv  
-rw-rw-r-- 1 acadgild acadgild 712 Jun 21 22:13 derby.log  
drwxr-xr-x 4 acadgild acadgild 4096 Jun 20 14:46 Desktop  
drwxr-xr-x 2 acadgild acadgild 4096 Feb 2 12:52 Documents  
drwxr-xr-x 2 acadgild acadgild 4096 Feb 13 14:24 Downloads  
drwxrwxr-x 3 acadgild acadgild 4096 Dec 29 16:59 eclipse  
drwxrwxr-x 3 acadgild acadgild 4096 Jan 16 14:02 eclipse-workspace  
-rw-rw-r-- 1 acadgild acadgild 248593 Jun 22 02:16 HVAC.csv  
-rw-rw-r-- 1 acadgild acadgild 26870105 Jun 22 00:56 inpatientCharges.csv  
drwxrwxr-x 13 acadgild acadgild 4096 Feb 9 18:06 install  
-rw-rw-r-- 1 acadgild acadgild 6034 Jun 22 08:29 movie_rating.jar  
drwxr-xr-x 2 acadgild acadgild 4096 Dec 27 17:03 Music  
drwxr-xr-x 2 acadgild acadgild 4096 Dec 27 17:03 Pictures  
drwxr-xr-x 2 acadgild acadgild 4096 Dec 27 17:03 Public  
-rw-rw-r-- 1 acadgild acadgild 929 Jun 20 11:50 S20_Dataset_Holidays.txt  
-rw-rw-r-- 1 acadgild acadgild 42 Jun 20 11:50 S20_Dataset_Transport.txt  
-rw-rw-r-- 1 acadgild acadgild 116 Jun 20 11:50 S20_Dataset_User_details.txt  
-rw-rw-r-- 1 acadgild acadgild 1040 Jun 21 23:15 Sports_data.txt  
drwxr-xr-x 2 acadgild acadgild 4096 Dec 27 17:03 Templates  
-rw-rw-r-- 1 acadgild acadgild 29 Jun 18 00:29 testfile  
drwxr-xr-x 2 acadgild acadgild 4096 Dec 27 17:03 Videos  
[acadgild@localhost ~]$ hadoop jar movie_rating.jar /movies.csv /ratings.csv /movies_rating_output
```

Applications Places System Fri Jun 22, 8:56 AM Acadgild

acadmild@localhost:~

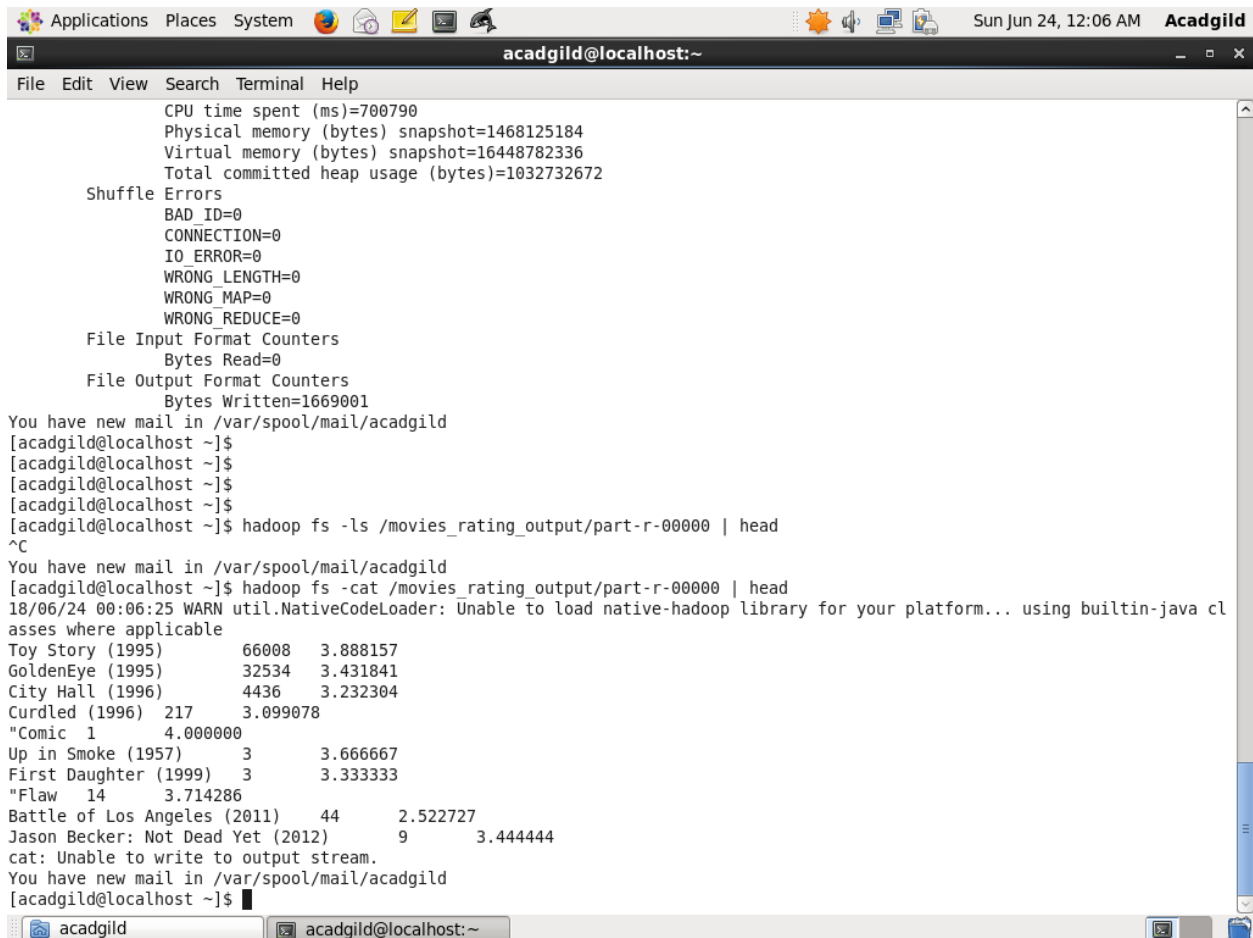
File Edit View Search Terminal Tabs Help

acadmild@localhost:~

```
-rw-rw-r--. 1 acadgild acadgild 116 Jun 20 11:50 S20_Dataset_User_details.txt
-rw-rw-r--. 1 acadgild acadgild 1040 Jun 21 23:15 Sports_data.txt
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Templates
-rw-rw-r--. 1 acadgild acadgild 29 Jun 18 00:29 testfile
drwxr-xr-x. 2 acadgild acadgild 4096 Dec 27 17:03 Videos
[acadgild@localhost ~]$ hadoop jar movie_rating.jar /movies.csv /ratings.csv /movies_rating_output
18/06/22 08:46:06 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/06/22 08:46:07 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/06/22 08:46:09 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/06/22 08:46:10 INFO input.FileInputFormat: Total input paths to process : 1
18/06/22 08:46:10 INFO input.FileInputFormat: Total input paths to process : 1
18/06/22 08:46:10 INFO mapreduce.JobSubmitter: number of splits:7
18/06/22 08:46:10 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1529599366963_0002
18/06/22 08:46:11 INFO impl.YarnClientImpl: Submitted application application_1529599366963_0002
18/06/22 08:46:11 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1529599366963_0002/
18/06/22 08:46:11 INFO mapreduce.Job: Running job: job_1529599366963_0002
18/06/22 08:46:24 INFO mapreduce.Job: Job job_1529599366963_0002 running in uber mode : false
18/06/22 08:46:24 INFO mapreduce.Job: map 0% reduce 0%
18/06/22 08:47:23 INFO mapreduce.Job: map 3% reduce 0%
18/06/22 08:47:28 INFO mapreduce.Job: map 6% reduce 0%
18/06/22 08:47:29 INFO mapreduce.Job: map 16% reduce 0%
18/06/22 08:47:32 INFO mapreduce.Job: map 24% reduce 0%
18/06/22 08:47:35 INFO mapreduce.Job: map 27% reduce 0%
18/06/22 08:47:38 INFO mapreduce.Job: map 30% reduce 0%
18/06/22 08:47:41 INFO mapreduce.Job: map 31% reduce 0%
18/06/22 08:47:42 INFO mapreduce.Job: map 34% reduce 0%
18/06/22 08:47:45 INFO mapreduce.Job: map 36% reduce 0%
18/06/22 08:47:49 INFO mapreduce.Job: map 37% reduce 0%
18/06/22 08:47:52 INFO mapreduce.Job: map 39% reduce 0%
18/06/22 08:47:56 INFO mapreduce.Job: map 40% reduce 0%
18/06/22 08:47:59 INFO mapreduce.Job: map 45% reduce 0%
18/06/22 08:48:28 INFO mapreduce.Job: map 49% reduce 0%
18/06/22 08:48:31 INFO mapreduce.Job: map 52% reduce 0%
18/06/22 08:48:32 INFO mapreduce.Job: map 53% reduce 0%
18/06/22 08:48:34 INFO mapreduce.Job: map 57% reduce 0%
```

acadmild@localhost:~

Output of map reduce :



```
acacgild@localhost:~
File Edit View Search Terminal Help
CPU time spent (ms)=700790
Physical memory (bytes) snapshot=1468125184
Virtual memory (bytes) snapshot=16448782336
Total committed heap usage (bytes)=1032732672
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=0
File Output Format Counters
Bytes Written=1669001
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$
[acacgild@localhost ~]$
[acacgild@localhost ~]$
[acacgild@localhost ~]$
[acacgild@localhost ~]$ hadoop fs -ls /movies_rating_output/part-r-00000 | head
^C
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$ hadoop fs -cat /movies_rating_output/part-r-00000 | head
18/06/24 00:06:25 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Toy Story (1995)          66008    3.888157
GoldenEye (1995)         32534    3.431841
City Hall (1996)         4436     3.232304
Curdled (1996)  217    3.099078
"Comic 1 4.000000
Up in Smoke (1957)       3          3.666667
First Daughter (1999)    3          3.333333
"Flaw 14 3.714286
Battle of Los Angeles (2011) 44    2.522727
Jason Becker: Not Dead Yet (2012) 9    3.444444
cat: Unable to write to output stream.
You have new mail in /var/spool/mail/acacgild
[acacgild@localhost ~]$
```

First column lists the movie which have been rated

Second column is the number of times users have rated

And third column is the average rating

Example toy story has been rated 66008 times with an average rating of 3.88