

## Problem Statement:

What are the movie titles that the user has rated?

How many times a movie has been rated by the user?

In question 2 above, what is the average rating given for a movie?

```
[acadgild@localhost ~]$ hadoop fs -mkdir /user/mapreduce
18/07/03 18:14:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -put /home/acadgild/Desktop/data/usecase/movies.csv /user/mapreduce/
18/07/03 18:15:28 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -put /home/acadgild/Desktop/data/usecase/ratings.csv /user/mapreduce/
18/07/03 18:15:41 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[acadgild@localhost ~]$ hadoop fs -ls /user/mapreduce
18/07/03 18:16:13 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 2283410 2018-07-03 18:15 /user/mapreduce/movies.csv
-rw-r--r-- 1 acadgild supergroup 789558327 2018-07-03 18:15 /user/mapreduce/ratings.csv
[acadgild@localhost ~]$
```

```
18/07/03 18:26:00 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/07/03 18:26:02 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/03 18:26:03 WARN mapreduce.JobResourcePilot: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/07/03 18:26:04 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 18:26:04 INFO input.FileInputFormat: Total input paths to process : 1
18/07/03 18:26:04 INFO mapreduce.JobSubmitter: number of splits:7
18/07/03 18:26:04 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1530544942678_0001
18/07/03 18:26:05 INFO impl.YarnClientImpl: Submitted application application_1530544942678_0001
18/07/03 18:26:05 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1530544942678_0001/
18/07/03 18:26:05 INFO mapreduce.Job: Running job: job_1530544942678_0001
18/07/03 18:26:17 INFO mapreduce.Job: Job job_1530544942678_0001 running in uber mode : false
18/07/03 18:26:17 INFO mapreduce.Job: map 0% reduce 0%
18/07/03 18:26:51 INFO mapreduce.Job: map 5% reduce 0%
18/07/03 18:26:52 INFO mapreduce.Job: map 6% reduce 0%
18/07/03 18:26:54 INFO mapreduce.Job: map 11% reduce 0%
18/07/03 18:26:55 INFO mapreduce.Job: map 13% reduce 0%
18/07/03 18:26:57 INFO mapreduce.Job: map 18% reduce 0%
18/07/03 18:26:58 INFO mapreduce.Job: map 19% reduce 0%
18/07/03 18:27:00 INFO mapreduce.Job: map 22% reduce 0%
18/07/03 18:27:01 INFO mapreduce.Job: map 25% reduce 0%
18/07/03 18:27:04 INFO mapreduce.Job: map 29% reduce 0%
18/07/03 18:27:07 INFO mapreduce.Job: map 33% reduce 0%
18/07/03 18:27:10 INFO mapreduce.Job: map 36% reduce 0%
18/07/03 18:27:13 INFO mapreduce.Job: map 38% reduce 0%
18/07/03 18:27:16 INFO mapreduce.Job: map 39% reduce 0%
18/07/03 18:27:17 INFO mapreduce.Job: map 45% reduce 0%
18/07/03 18:27:41 INFO mapreduce.Job: map 52% reduce 0%
18/07/03 18:27:44 INFO mapreduce.Job: map 54% reduce 0%
18/07/03 18:27:47 INFO mapreduce.Job: map 74% reduce 0%
18/07/03 18:27:50 INFO mapreduce.Job: map 76% reduce 0%
18/07/03 18:28:14 INFO mapreduce.Job: map 76% reduce 10%
18/07/03 18:28:16 INFO mapreduce.Job: map 77% reduce 10%
18/07/03 18:28:20 INFO mapreduce.Job: map 84% reduce 10%
18/07/03 18:28:25 INFO mapreduce.Job: map 88% reduce 10%
18/07/03 18:28:26 INFO mapreduce.Job: map 92% reduce 10%
18/07/03 18:28:29 INFO mapreduce.Job: map 95% reduce 10%
18/07/03 18:28:31 INFO mapreduce.Job: map 96% reduce 10%
```

```

18/07/09 00:30:02 INFO mapreduce.Job: map 87% reduce 18%
18/07/09 00:30:05 INFO mapreduce.Job: map 90% reduce 19%
18/07/09 00:30:06 INFO mapreduce.Job: map 91% reduce 19%
18/07/09 00:30:08 INFO mapreduce.Job: map 93% reduce 19%
18/07/09 00:30:09 INFO mapreduce.Job: map 96% reduce 19%
18/07/09 00:30:11 INFO mapreduce.Job: map 98% reduce 19%
18/07/09 00:30:12 INFO mapreduce.Job: map 99% reduce 19%
18/07/09 00:30:15 INFO mapreduce.Job: map 100% reduce 19%
18/07/09 00:30:18 INFO mapreduce.Job: map 100% reduce 29%
18/07/09 00:30:21 INFO mapreduce.Job: map 100% reduce 67%
18/07/09 00:30:30 INFO mapreduce.Job: map 100% reduce 68%
18/07/09 00:30:36 INFO mapreduce.Job: map 100% reduce 69%
18/07/09 00:30:46 INFO mapreduce.Job: map 100% reduce 70%
18/07/09 00:30:52 INFO mapreduce.Job: map 100% reduce 71%
18/07/09 00:31:01 INFO mapreduce.Job: map 100% reduce 72%
18/07/09 00:31:07 INFO mapreduce.Job: map 100% reduce 73%
18/07/09 00:31:13 INFO mapreduce.Job: map 100% reduce 74%
18/07/09 00:31:22 INFO mapreduce.Job: map 100% reduce 75%
18/07/09 00:31:31 INFO mapreduce.Job: map 100% reduce 76%
18/07/09 00:31:48 INFO mapreduce.Job: map 100% reduce 77%
18/07/09 00:31:46 INFO mapreduce.Job: map 100% reduce 78%
18/07/09 00:31:55 INFO mapreduce.Job: map 100% reduce 79%
18/07/09 00:32:05 INFO mapreduce.Job: map 100% reduce 80%
18/07/09 00:32:11 INFO mapreduce.Job: map 100% reduce 81%
18/07/09 00:32:17 INFO mapreduce.Job: map 100% reduce 82%
18/07/09 00:32:26 INFO mapreduce.Job: map 100% reduce 83%
18/07/09 00:32:32 INFO mapreduce.Job: map 100% reduce 84%
18/07/09 00:32:38 INFO mapreduce.Job: map 100% reduce 85%
18/07/09 00:32:44 INFO mapreduce.Job: map 100% reduce 86%
18/07/09 00:32:53 INFO mapreduce.Job: map 100% reduce 87%
18/07/09 00:33:08 INFO mapreduce.Job: map 100% reduce 88%
18/07/09 00:33:06 INFO mapreduce.Job: map 100% reduce 89%
18/07/09 00:33:15 INFO mapreduce.Job: map 100% reduce 90%
18/07/09 00:33:21 INFO mapreduce.Job: map 100% reduce 91%
18/07/09 00:33:27 INFO mapreduce.Job: map 100% reduce 92%
18/07/09 00:33:33 INFO mapreduce.Job: map 100% reduce 93%
18/07/09 00:33:42 INFO mapreduce.Job: map 100% reduce 94%
18/07/09 00:33:48 INFO mapreduce.Job: map 100% reduce 95%
18/07/09 00:33:54 INFO mapreduce.Job: map 100% reduce 96%

```

```

18/07/09 00:34:03 INFO mapreduce.Job: map 100% reduce 97%
18/07/09 00:34:09 INFO mapreduce.Job: map 100% reduce 98%
18/07/09 00:34:15 INFO mapreduce.Job: map 100% reduce 99%
18/07/09 00:34:28 INFO mapreduce.Job: map 100% reduce 100%
18/07/09 00:34:30 INFO mapreduce.Job: Job job_1531075833986_0861 completed successfully
18/07/09 00:34:31 INFO mapreduce.Job: Counters: 58
  File System Counters
    FILE: Number of bytes read=961694264
    FILE: Number of bytes written=1457487398
    FILE: Number of read operations=8
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=711858274
    HDFS: Number of bytes written=1669861
    HDFS: Number of read operations=24
    HDFS: Number of large read operations=6
    HDFS: Number of write operations=2
  Job Counters
    Killed map tasks=1
    Launched map tasks=8
    Launched reduce tasks=1
    Data-local map tasks=8
    Total time spent by all maps in occupied slots (ms)=875751
    Total time spent by all reduces in occupied slots (ms)=382847
    Total time spent by all map tasks (ms)=875751
    Total time spent by all reduce tasks (ms)=382847
    Total vcore-milliseconds taken by all map tasks=875751
    Total vcore-milliseconds taken by all reduce tasks=382847
    Total megabyte-milliseconds taken by all map tasks=896769024
    Total megabyte-milliseconds taken by all reduce tasks=389296128
  Map-Reduce Framework
    Map input records=26070134
    Map output records=26070132
    Map output bytes=442789828
    Map output materialized bytes=494936141
    Input split bytes=2607

```

Total time spent by all reduce tasks (ms)=332047  
 Total vcore-milliseconds taken by all map tasks=875751  
 Total vcore-milliseconds taken by all reduce tasks=332947  
 Total megabyte-milliseconds taken by all map tasks=89676924  
 Total megabyte-milliseconds taken by all reduce tasks=369290128

#### Map-Reduce Framework

Map input records=26876134  
 Map output records=26876132  
 Map output bytes=442789828  
 Map output materialized bytes=494938141  
 Input split bytes=2867  
 Combine input records=0  
 Combine output records=0  
 Reduce input groups=45843  
 Reduce shuffle bytes=494938141  
 Reduce input records=26876132  
 Reduce output records=45843  
 Spilled Records=76775523  
 Shuffled Maps =7  
 Failed Shuffles=0  
 Merged Map outputs=7  
 GC time elapsed (ms)=16019  
 CPU time spent (ms)=327420  
 Physical memory (bytes) snapshot=1449228248  
 Virtual memory (bytes) snapshot=16448745472  
 Total committed heap usage (bytes)=1832732672

#### 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=1869681

You have new mail in /var/spool/mail/acadgild

I

Nature Calls (2012)	12	1.791667	
Babbitt (1934)	1	3.099000	
In Their Sleep (Dens ton someil) (2010)	3	2.666667	
Requiem for a Killer (Requies pour une tueuse) (2011)	1	4.096699	
*Bohème 3	3.666667		
*Caine Mutiny Court-Martial	8	3.437500	
Vamps (2012)	44	2.795455	
Cash on Demand (1962)	2	3.750000	
*Naughty Room	1	3.099000	
*Moth Diaries	39	3.612821	
Doe + Belle (2011)	2	4.099000	
*Invisible War	86	3.738372	
OMG Oh My God! (2012)	61	3.942623	
Picnic (1955)	619	3.644588	
*Strawberry Statement	10	3.400000	
*Twilight Saga: Breaking Dawn - Part 2	1176	2.523384	
Shakespeare High (2011)	3	3.333333	
*Circle 2	1.750000		
Holiday Engagement (2011)	14	3.687143	
*Tiger Brigades	5	2.800000	
*Year One	6	3.003333	
*Big Blonde	2	3.099000	
Walter (1982)	2	4.099000	
10 Years (2011)	95	3.642105	
Price Check (2012)	24	2.937500	
Red Dawn (2012)	261	2.747126	
Starry starry night (Xing kong) (2011)	7	3.057143	
Rise of the Guardians (2012)	1527	3.528195	
*Seafarers	18	2.388889	
*Counterfeiters 6	3.666667		
Kabei: Our Mother (Kâbê) (2008)	5	4.288669	
Samurai Banners (Fûrin kazen) (1969)	4	3.625000	
Murder on Flight 562 (1975)	2	3.099000	
Amor? (2011)	1	3.099000	
*Last Trapper	4	3.375000	
*Octopus	11	3.181818	
Locked Out (Enfermés dehors) (2006)	8	2.687500	
*Fantastic Fear of Everything	114	3.122807	
Fire Over England (1937)	4	2.750000	

```
[acadgill@localhost Use Case]$ hadoop fs -cat /user/masreduce/out/part-r-00000 | head
18/07/09 00:40:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Toy Story (1995)      64000  3.000157
GoldenEye (1995)     32534  3.431841
City Hall (1995)     4436   3.232104
Cuddled (1996)       217    3.099073
*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    1.444444
cat: Unable to write to output stream.
You have new mail in /var/spool/mail/acadgilld
[acadgill@localhost Use Case]$
```