

Activities Terminal Tue 17:20 hduser@sreeraj: ~

```
(49538,Life With Boys: Season 1: The Big Kiss Off with Boys,2011,3.451207,1368)
(49540,Life With Boys: Season 1: Social Death with Boys,2011,3.451207,1368)
(49541,Life With Boys: Season 1: This Time the Problem Is with Dad and Not with Boys,2011,3.451207,1368)
(49542,Life With Boys: Season 1: In the Principal's Office with Boys,2011,3.451207,1368)
(49543,El Fuente: 5994 MP,2013,3.5,471)
(49544,El Fuente: 56 MP,2013,3.3,465)
(49545,El Fuente: 50 MP,2013,3.3,465)
(49546,Bo Burnham: what.,2013,4.1,3614)
(49547,Life With Boys: Season 1: A Perfect Life with Boys,2011,3.451207,1368)
(49548,Life With Boys: Season 1: Wrestling with Boys,2011,3.451207,1368)
(49549,Life With Boys: Season 1,2011,4.1,2563)
(49550,El Fuente: 30 MP,2013,3.3,471)
(49551,El Fuente: 2997 MP,2013,3.2,471)
(49552,El Fuente: 25 MP,2013,3.3,465)
(49553,El Fuente: 24 MP,2013,3.5,484)
(49554,Max Steel,2013,4.1,2563)
(49555,Saving Santa,2013,3.8,5038)
(49556,Lillyhammer: Season 1 (Recap),2013,4.2,194)
(49557,Shinobli Grl1,2012,2,0,2563)
(49558,GLOW: The Story of the Gorgeous Ladies of Wrestling,2012,3.7,4593)
(49559,Mitt (Trailer),2013,3.0,138)
(49560,My Hope America with Billy Graham: The Cross,2013,3.451207,1706)
(49561,The Square (Trailer),2014,3.6,154)
(49562,My Hope America with Billy Graham,2013,3.9,2563)
(49563,My Hope America with Billy Graham: Defining Moments,2013,3.451207,1791)
(49564,My Hope America with Billy Graham: Lose to Gain,2013,3.451207,1406)
(49565,American Addict,2013,3.5,5377)
(49566,My Hope America with Billy Graham,2013,3.9,2563)
(49567,El Fuente: 2997 MP10,2013,2.7,470)
(49568,El Fuente: 25 MP10,2013,2.9,464)
(49569,El Fuente: 24 MP10,2013,2.8,484)
(49570,El Fuente: 23976 MP10,2013,2.9,484)
(49571,The Short Game (Trailer),2013,4.1,156)
(49572,El Fuente: 5994 MP10,2013,2.8,471)
(49573,El Fuente: 50 MP10,2013,2.9,464)
(49574,El Fuente: 30 MP10,2013,2.8,470)
(49575,Greg Filtzsimmons: Life on Stage,2013,3.3,3671)
(49576,Dave Foley: Relatively Well,2013,3.2,3446)
(49577,Barbie: Life in the Dreamhouse: Barbie Life in the Dreamhouse: Best of Family,2013,3.451207,1390)
(49578,Barbie: Life in the Dreamhouse: Barbie Life in the Dreamhouse: Best of Friends,2013,3.451207,1458)
(49579,Transformers Prime Beast Hunters: Predacons Rising,2013,4.2,3950)
(49580,Underground: The Julian Assange Story,2012,3.7,5665)
(49581,Curious George: A Very Monkey Christmas,2009,3.8,3438)
(49582,Mumble's White Christmas,1996,2.4,1350)
(49583,Lady Gaga #38; The Muppets' Holiday Spectacular,2013,3.1,3496)
(49584,Sunset Strip,2012,3.0,5770)
(49585,Silver Bells,2013,3.5,5287)
(49586,Winter Wonderland,2013,2.8,1812)
(49587,Top Gear: Series 19: Africa Special,2013,3.451207,6822)
(49588,Fireplace For Your Home: Crackling Fireplace with Music,2010,3.451207,3610)
(49589,Kate Plus El8ht,2010,2.7,2563)
(49590,Kate Plus El8ht: Season 1,2010,2.7,2563)
grunt> grp_data = GROUP movies all;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(movies);[]
```



Activities Terminal Tue 17:28 ●  
hduser@sreeraj: ~

```
File Edit View Search Terminal Help
(75,Apocalypse Now,1979,4.8,8825)
(76,Leviathan,1989,3.2,5877)
(77,Marvin's Room,1996,3.5,5901)
(78,Love and a .45,1994,3.3,6125)
(79,The Mark of Zorro,1920,3.1,6433)
(80,Little Buddha,1993,3.4,8455)
(81,Like Water for Chocolate,1992,3.8,6320)
(82,The Mikado,1939,3.1,5448)
(83,Lionheart,1990,3.5,6301)
(84,Minnie and Moskowitz,1971,3.2,6928)
(85,The Hunt for Red October,1990,4.6,8106)
(86,The Hindenburg,1975,3.4,7546)
(87,I Spit on Your Grave,1978,3.1,6084)
(88,Hellbound: Hellraiser II,1988,3.4,5606)
(89,Highlander 3: The Final Dimension,1994,3.2,5943)
(90,Heart and Souls,1993,3.8,6196)
(91,House of Whipplecord,1974,2.5,6144)
(92,The Great Outdoors,1988,3.7,5424)
(93,Heaven's Gate,1980,3.0,13143)
(94,House on Haunted Hill,1959,3.6,4491)
(95,The Golden Child,1986,3.5,5669)
(96,The Hunted,1995,3.4,6665)
(97,The Great Waldo Pepper,1975,3.5,6467)
(98,Godzilla: King of the Monsters,1956,3.5,4828)
(99,Highlander 2: Renegade Version,1991,3.1,6585)
(100,High Noon,1952,3.9,5087)
(101,The Hunchback of Notre Dame,1923,3.5,6750)
(102,Heathers,1989,3.7,6193)
(103,Godzilla's Revenge,1969,3.3,4165)
(104,Hard Target,1993,3.4,5771)
(105,Henry: Portrait of a Serial Killer,1986,3.2,4938)
(106,Godzilla vs. Mothra,1964,3.5,5285)
(107,The Grifters,1990,3.4,6628)
(108,An Ideal Husband,1999,3.5,5879)
(109,A Fish Called Wanda,1988,3.7,6477)
(110,Gentleman's Agreement,1947,3.7,7104)
(111,The Prophecy,1995,3.6,5864)
(112,Flirting with Disaster,1996,3.2,5565)
(113,Ghostbusters,1984,3.8,6308)
(114,Far and Away,1992,3.6,8396)
(115,Fargo,1996,3.8,5889)
(116,Georgia,1995,2.9,650)
(117,Fascination,1979,2.7,4980)
(118,Fresh,1994,4.6,6829)
(119,The Frighteners,1996,3.6,6667)
(120,Go West,1929,3.5,4181)
(121,Ferris Bueller's Day Off,1986,3.9,6177)
(122,The Ghost and the Darkness,1996,3.7,6594)
(123,Four Rooms,1995,3.6,5872)
(124,The First Wives' Club,1996,3.6,6139)
(125,Operation Condor,1991,3.6,5435)
(126,Gallipoli,1981,3.7,6691)
(127,Fallen Angels,1995,3.5,5939)
grunt> grp_data = GROUP duplicate all;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(duplicate);
```

```

Activities Terminal 10:17:51 ~
hduser@sreeraj: ~

File Edit View Search Terminal Help
2023-02-28 17:29:13,710 [pool-35-thread-1] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt_local723010045_0012_r_000000_0' to hdfs://localhost:54310/tmp/temp555996968/tmp-63268740/_temporary/0/task_local723010045_0012_r_000000
2023-02-28 17:29:13,711 [pool-35-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2023-02-28 17:29:13,711 [pool-35-thread-1] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local723010045_0012_r_000000_0' done.
2023-02-28 17:29:13,711 [pool-35-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local723010045_0012_r_000000_0
2023-02-28 17:29:13,711 [Thread-415] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce task executor complete.
2023-02-28 17:29:13,968 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Running Jobs are [job_local723010045_0012]
2023-02-28 17:29:18,476 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,477 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,478 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,483 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-02-28 17:29:18,525 [main] INFO org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-02-28 17:29:04 2023-02-28 17:29:18 GROUP_BY,DISTINCT

Success!
Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs
job_local1616556853_0011 1 1 n/a n/a n/a n/a n/a n/a movies DISTINCT
job_local723010045_0012 1 1 n/a n/a n/a n/a n/a n/a cnt_data,grp_data GROUP_BY,COMBINER hdfs://localhost:54310/tmp/temp555996968/tmp-63268740
0,
Input(s):
Successfully read 49589 records (157178564 bytes) from: "/assignment/movies_cleaned_pandas.csv"
Output(s):
Successfully stored 1 records (156610535 bytes) to: "hdfs://localhost:54310/tmp/temp555996968/tmp-63268740"

Counters:
Total records written : 1
Total bytes written : 156610535
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1616556853_0011 -> job_local723010045_0012,
job_local723010045_0012

2023-02-28 17:29:18,527 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,528 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,529 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,533 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,534 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,536 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 17:29:18,538 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-02-28 17:29:18,539 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-02-28 17:29:18,544 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-02-28 17:29:18,544 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapReduceUtil - Total input paths to process : 1
(all,49589)
grunt> release = ORDER movies by year desc;
grunt> dump release[]

```

Activities Terminal Tue 18:57 ● hduser@sreeraj: ~

```
(49540,Life With Boys: Season 1: Social Death with Boys,2011,3.451207,1368)
(49541,Life With Boys: Season 1: This Time the Problem Is with Dad and Not with Boys,2011,3.451207,1368)
(49542,Life With Boys: Season 1: In the Principal's Office with Boys,2011,3.451207,1368)
(49543,El Fuente: 5994 MP,2013,3.5,471)
(49544,Life With Boys: Season 1: Monkey Talk with Boys,2011,3.451207,1368)
(49545,El Fuente: 58 MP,2013,3.3,465)
(49546,Bo Burnham: what.,2013,4.1,364)
(49547,Life With Boys: Season 1: A Perfect Life with Boys,2011,3.451207,1368)
(49548,Life With Boys: Season 1: Wrestling with Boys,2011,3.451207,1369)
(49549,Life With Boys: Season 1,2011,4.1,2563)
(49550,El Fuente: 38 MP,2013,3.3,471)
(49551,El Fuente: 2997 MP,2013,3.2,471)
(49552,El Fuente: 23 MP,2013,3.3,465)
(49553,El Fuente: 24 MP,2013,3.5,484)
(49554,Man Steel,2013,4.1,2563)
(49555,Saving Santa,2013,3.8,5038)
(49556,Lillyhammer: Season 1 (Recap),2013,4.2,194)
(49557,Shinobi Girl,2012,2.6,2963)
(49558,GLOW: The Story of the Gorgeous Ladies of Wrestling,2012,3.7,4593)
(49559,Mitt (Trailer),2013,3.0,138)
(49560,My Hope America with Billy Graham: The Cross,2013,3.451207,1706)
(49561,The Square (Trailer),2014,3.6,154)
(49562,My Hope America with Billy Graham,2013,3.9,2563)
(49563,My Hope America with Billy Graham: Defining Moments,2013,3.451207,1791)
(49564,My Hope America with Billy Graham: Lose to Gññ,2013,3.451207,1406)
(49565,American Addict,2013,3.5,3377)
(49566,My Hope America with Billy Graham,2013,3.9,2563)
(49567,El Fuente: 2997 MP10,2013,2.7,470)
(49568,El Fuente: 25 MP10,2013,2.9,464)
(49569,El Fuente: 24 MP10,2013,2.8,484)
(49570,El Fuente: 23976 MP10,2013,2.9,484)
(49571,The Short Game (Trailer),2013,2.1,156)
(49572,El Fuente: 5994 MP10,2013,2.8,471)
(49573,El Fuente: 30 MP10,2013,2.9,464)
(49574,El Fuente: 30 MP10,2013,2.8,470)
(49575,Greg Fitzsimmons: Life on Stage,2013,3.3,3671)
(49576,Dave Foley: Relatively Well,2013,3.2,3446)
(49577,Barbie: Life in the Dreamhouse: Barbie Life in the Dreamhouse: Best of Family,2013,3.451207,1390)
(49578,Barbie: Life in the Dreamhouse: Barbie Life in the Dreamhouse: Best of Friends,2013,3.451207,1458)
(49579,Transformers Prime Beast Hunters: Predacons Rising,2013,4.2,3950)
(49580,Underground: The Julian Assange Story,2012,3.7,5665)
(49581,Elves Design A Very Monkey Christmas,2009,3.8,3438)
(49582,Murphy's White Christmas,1996,2.4,1358)
(49583,Lady Gaga &#38; The Muppets: Holiday Spectacular,2013,3.1,3496)
(49584,Lady Gaga &#38; The Muppets: Holiday Spectacular,2013,3.1,3496)
(49585,Sunset Strip,2012,3.0,3776)
(49586,Winter Wonderland,2013,2.8,1812)
(49587,Top Gear Series 19: Africa Special,2013,3.451207,6822)
(49588,Fireplace For Your Home: Crackling Fireplace with Music,2010,3.451207,3610)
(49589,Kate Plus El8ht,2018,2.7,2563)
(49590,Kate Plus El8ht: Season 1,2018,2.7,2563)
grunt> rating1 = ORDER movies by rating desc;
grunt> limit_data = LIMIT rating1 5;
grunt> col_data = FOREACH limit_data generate movie,year,rating;
grunt> dump col_data]
```

```

Activities Terminal Tue 19:07 ●
hduser@sreeraj: ~

File Edit View Search Terminal Help
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-02-28 19:01:42 2023-02-28 19:02:07 ORDER_BY,LIMIT

Success!

Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReduceTime Alias Feature Outputs
job_local1215734529_0008 1 1 n/a n/a n/a n/a n/a rating2 ORDER_BY,COMBINER
job_local1643337318_0007 1 1 n/a n/a n/a n/a n/a n/a rating2 SAMPLER
job_local610562426_0006 1 0 n/a n/a n/a 0 0 0 movies MAP_ONLY
job_local891114359_0009 1 1 n/a n/a n/a n/a n/a n/a col_data,rating2 hdfs://localhost:54310/tmp/temp1296218285/tmp-77010805,
Input(s):
Successfully read 49589 records (54199664 bytes) from: "/assignment/movies_cleaned_pandas.csv"

Output(s):
Successfully stored 3 records (120702602 bytes) in: "hdfs://localhost:54310/tmp/temp1296218285/tmp-77010805"

Counters:
Total records written : 3
Total bytes written : 120702602
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local610562426_0006 -> job_local1643337318_0007,
job_local1643337318_0007 -> job_local1215734529_0008,
job_local1215734529_0008 -> job_local891114359_0009,
job_local891114359_0009

2023-02-28 19:02:07,004 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,005 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,006 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,009 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,010 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,011 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,015 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,017 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,018 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,022 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,023 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,024 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 19:02:07,027 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-02-28 19:02:07,033 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-02-28 19:02:07,033 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-02-28 19:02:07,033 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapReduceUtil - Total input paths to process : 1
(Lagegl,2007,1.4)
(Sun Yaar Calll Maar,2007,1.4)
(Lagegl,2007,1.4)

grunt> grp_data = GROUP movies by year;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(movies) as abc;
grunt> order_data = ORDER cnt_data by abc desc;
grunt> dump order_data |
```

```

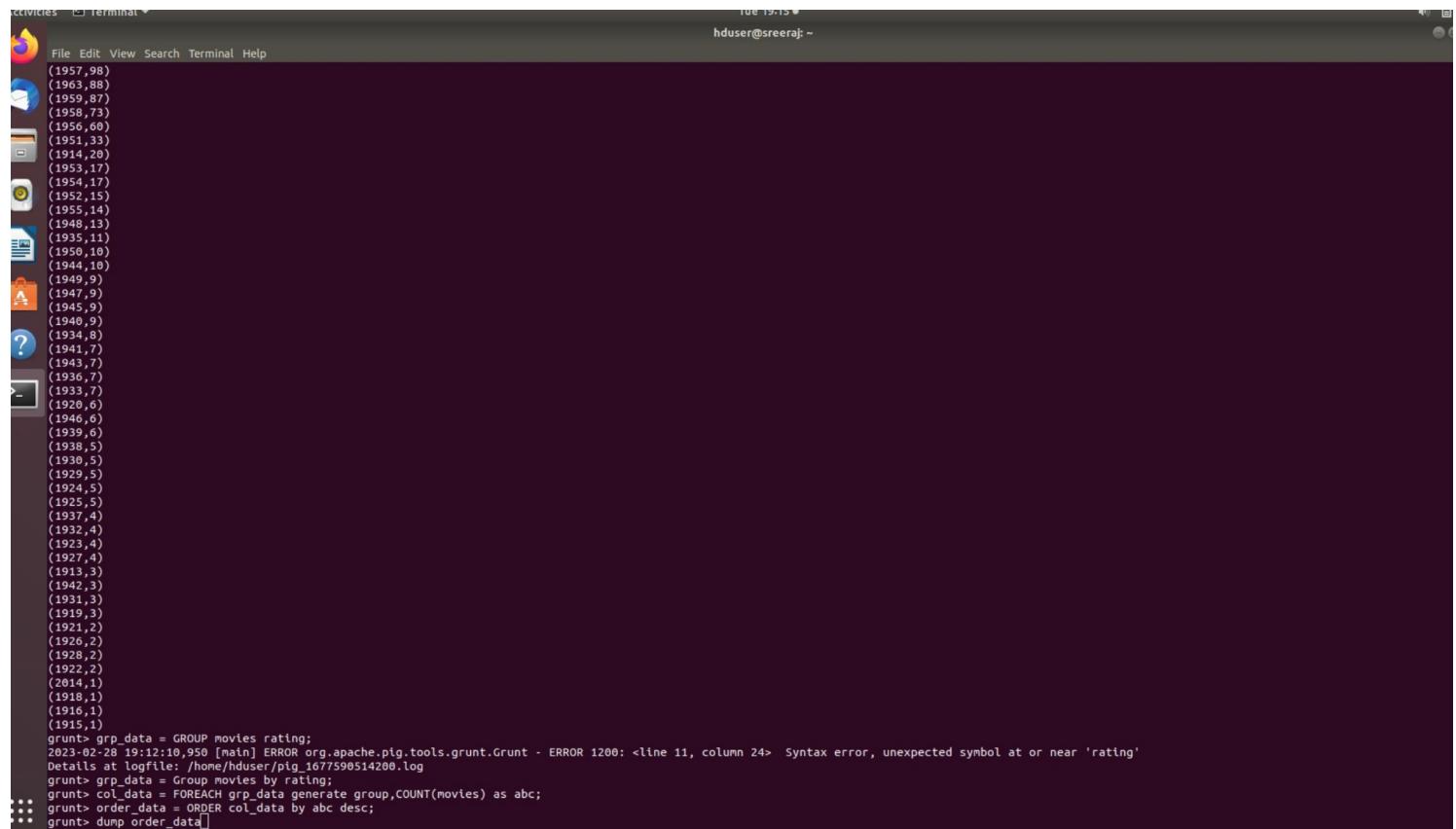
Activities Terminal ▾ Tue 19:01
hduser@sreeraj: ~

Success!
Job Stats (time in seconds):
JobID Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReduceTime Alias Feature Outputs
job_local139225806_0004 1 n/a n/a n/a n/a n/a n/a rating1 ORDER_BY,COMBINER
job_local1833232606_0002 1 0 n/a n/a n/a 0 0 0 movies MAP_ONLY
job_local2114517198_0005 1 1 n/a n/a n/a n/a n/a n/a col_data,rating1
job_local830189117_0003 1 1 n/a n/a n/a n/a n/a n/a rating1 SAMPLER
hdfs://localhost:54310/tmp/temp1296218285/tmp-819418076,
Input(s):
Successfully read 49589 records (21992399 bytes) from: "/assignment/movies_cleaned_pandas.csv"
Output(s):
Successfully stored 5 records (69177397 bytes) in: "hdfs://localhost:54310/tmp/temp1296218285/tmp-819418076"
Counters:
Total records written : 5
Total bytes written : 69177397
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
Job DAG:
job_local1833232606_0002 -> job_local830189117_0003,
job_local830189117_0003 -> job_local139225806_0004,
job_local139225806_0004 -> job_local2114517198_0005,
job_local2114517198_0005

2023-02-28 18:57:40,875 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,877 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,878 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,882 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,883 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,884 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,891 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,893 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,894 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,895 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,896 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,902 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-02-28 18:57:40,905 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-02-28 18:57:40,908 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-02-28 18:57:40,916 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-02-28 18:57:40,916 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(Breaking Bad: Season 5,2012,4,5)
(Arrested Development (Trailer),2013,4,5)
(Blackfish,2013,4,5)
(Sherlock,2010,4,5)
(Orange Is the New Black: Season 1,2013,4,5)
grunt> rating2 = ORDER movies by rating asc;
grunt> limit data = LIMIT rating2 3;
grunt> col_data = FOREACH limit_data generate movie,year,rating;
grunt> dump col data

```

```
Activities Terminal
File Edit View Search Terminal Help
2023-02-28 19:00:18,530 [math] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLib - Total input paths to process : 1
(2011, 551)
(2010, 5107)
(2009, 4451)
(2012, 4339)
(2008, 3358)
(2007, 2892)
(2006, 2416)
(2005, 1937)
(2003, 1399)
(2004, 1381)
(1999, 1181)
(2001, 1173)
(2002, 1117)
(2013, 981)
(2000, 902)
(1998, 843)
(1997, 788)
(1996, 688)
(1995, 592)
(1993, 563)
(1994, 517)
(1992, 479)
(1996, 478)
(1989, 421)
(1991, 364)
(1985, 334)
(1988, 334)
(1984, 303)
(1986, 287)
(1987, 288)
(1967, 279)
(1983, 278)
(1978, 231)
(1974, 178)
(1968, 173)
(1973, 168)
(1972, 166)
(1982, 153)
(1978, 141)
(1979, 140)
(1977, 136)
(1975, 134)
(1971, 131)
(1964, 126)
(1969, 124)
(1962, 124)
(1966, 123)
(1961, 119)
(1976, 118)
(1981, 112)
(1988, 107)
(1965, 104)
(1966, 103)
(1957, 98)
(1963, 96)
```



A screenshot of a Linux desktop environment, specifically Ubuntu, showing a terminal window. The terminal window is titled 'Terminal' and has the command 'hduser@sreeraj: ~' at the top. The window contains a list of movie ratings and a Grunt shell session. The list of movie ratings is as follows:

```
(1957,98)
(1963,88)
(1959,87)
(1958,73)
(1956,60)
(1951,33)
(1914,20)
(1953,17)
(1954,17)
(1952,15)
(1955,14)
(1948,13)
(1935,11)
(1950,10)
(1944,10)
(1949,9)
(1947,9)
(1945,9)
(1946,9)
(1934,8)
(1941,7)
(1943,7)
(1936,7)
(1933,7)
(1929,6)
(1946,6)
(1939,6)
(1938,5)
(1930,5)
(1929,5)
(1924,5)
(1925,5)
(1937,4)
(1932,4)
(1923,4)
(1927,4)
(1913,3)
(1942,3)
(1931,3)
(1919,3)
(1921,2)
(1926,2)
(1928,2)
(1922,2)
(2014,1)
(1918,1)
(1916,1)
(1915,1)
```

The Grunt shell session shows the following commands and errors:

```
grunt> grp_data = GROUP movies rating;
2023-02-28 19:12:10,950 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line 11, column 24>  Syntax error, unexpected symbol at or near 'rating'
Details at logfile: /home/hduser/pig_1677596514206.log
grunt> grp_data = Group movies by rating;
grunt> col_data = FOREACH grp_data generate group,COUNT(movies) as abc;
grunt> order_data = ORDER col_data by abc desc;
grunt> dump order_data]
```

```

Activities Terminal ~
Wed 13:15 ●
hduser@sreeraj: ~

File Edit View Search Terminal Help
2023-03-01 13:04:53,881 [pool-52-thread-1] INFO org.apache.hadoop.mapred.Task - Task:attempt_local243410976_0017_r_000000_0 is done. And is in the process of committing
2023-03-01 13:04:53,885 [pool-52-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2023-03-01 13:04:53,886 [pool-52-thread-1] INFO org.apache.hadoop.mapred.Task - Task attempt_local243410976_0017_r_000000_0 is allowed to commit now
2023-03-01 13:04:53,983 [pool-52-thread-1] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt_local243410976_0017_r_000000_0' to hdfs://localhost:54310/tmp/te
rp-795517020/tmp-1936471947/_temporary/0/task_local243410976_0017_r_000000
2023-03-01 13:04:53,984 [pool-52-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2023-03-01 13:04:53,985 [pool-52-thread-1] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local243410976_0017_r_000000_0' done.
2023-03-01 13:04:53,985 [pool-52-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local243410976_0017_r_000000_0
2023-03-01 13:04:53,985 [Thead-646] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce task executor complete.
2023-03-01 13:04:54,294 [main] INFO org.apache.pig.backend.hadoop.executionengine.MapReduceLayer.MapReduceLauncher - Running jobs are [job_local243410976_0017]
2023-03-01 13:04:58,802 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:04:58,804 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:04:58,805 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:04:58,808 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-03-01 13:04:58,808 [main] INFO org.apache.pig.tools.pigstats.ScriptStatistics:
Success!
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-03-01 13:04:52 2023-03-01 13:04:58 GROUP_BY,FILTER

Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs
job_local243410976_0017 1 1 n/a n/a n/a n/a n/a n/a cnt_data,filter_data.grp_data,movies GROUP_BY,COMBINER hdfs://localhost:54310/tmp/temp-795517020/tmp-1936471947,
17628/tmp-1936471947,
Input(s):
Successfully read 49589 records (305643588 bytes) from: "/assignment/movies_cleaned_pandas.csv"
Output(s):
Successfully stored 1 records (232187509 bytes) in: "hdfs://localhost:54310/tmp/temp-795517020/tmp-1936471947"
Counters:
Total records written : 1
Total bytes written : 232187509
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local243410976_0017

2023-03-01 13:04:58,810 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:04:58,811 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:04:58,811 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:04:58,813 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-03-01 13:04:58,815 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-03-01 13:04:58,821 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-03-01 13:04:58,821 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(all,2779)
grunt> filter_data = FILTER movies by year==2008;
grunt> grp_data = GROUP filter_data all;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(filter_data);
grunt> dump cnt_data

```

```

Activities Terminal ~
Wed 13:20 ~
hduser@sreeraj: ~

File Edit View Search Terminal Help
s[1,9],movies[-1,-1],filter_data[17,14],cnt_data[19,11],grp_data[18,11] C: cnt_data[19,11],grp_data[18,11] R: cnt_data[19,11]
2023-03-01 13:15:13,879 [pool-55-thread-1] INFO org.apache.hadoop.mapred.Task - Task:attempt_local549847120_0018_r_000000_0 is done. And is in the process of committing
2023-03-01 13:15:13,883 [pool-55-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2023-03-01 13:15:13,886 [pool-55-thread-1] INFO org.apache.hadoop.mapred.Task - Task attempt_local549847120_0018_r_000000_0 is allowed to commit now
2023-03-01 13:15:13,901 [pool-55-thread-1] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt_local549847120_0018_r_000000_0' to hdfs://localhost:54310/tmp/temp-795517020/tmp1012756810/_temporary/0/task_local549847120_0018_r_000000_0
2023-03-01 13:15:13,902 [pool-55-thread-1] INFO org.apache.hadoop.mapred.Task - LocalJobRunner - reduce > reduce
2023-03-01 13:15:13,902 [pool-55-thread-1] INFO org.apache.hadoop.mapred.Task - Task attempt_local549847120_0018_r_000000_0 is done.
2023-03-01 13:15:13,903 [pool-55-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local549847120_0018_r_000000_0
2023-03-01 13:15:13,903 [Thread-689] INFO org.apache.hadoop.mapred.local.JobRunner - reduce task executor complete.
2023-03-01 13:15:14,164 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Running jobs are [job_local549847120_0018]
2023-03-01 13:15:18,671 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:15:18,672 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:15:18,673 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:15:18,676 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-03-01 13:15:18,676 [main] INFO org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-03-01 13:15:11 2023-03-01 13:15:18 GROUP_BY,FILTER
Success!
Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs
job_local549847120_0018 1 n/a n/a n/a n/a n/a n/a n/a n/a cnt_data,filter_data,grp_data,movies GROUP_BY,COMBINER hdfs://localhost:54310/tmp/temp-795517020/tmp1012756810,
Input(s):
Successfully read 49589 records (324565762 bytes) from: "/assignment/movies_cleaned_pandas.csv"
Output(s):
Successfully stored 1 records (243699099 bytes) in: "hdfs://localhost:54310/tmp/temp-795517020/tmp1012756810"
Counters:
Total records written : 1
Total bytes written : 243699099
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
Job DAG:
job_local549847120_0018

2023-03-01 13:15:18,678 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:15:18,680 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:15:18,681 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 13:15:18,682 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - already initialized
2023-03-01 13:15:18,684 [main] WARN org.apache.pig.data.SchemaTupleBackend SchemaTupleBackend has already been initialized
2023-03-01 13:15:18,685 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-03-01 13:15:18,685 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(all,3358)
grunt> grp_data = GROUP filter_data by year;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(filter_data);
grunt> dump cnt_data

```

```
Activities Terminal wed 12:58 ~
hduser@sreeraj: ~
File Edit View Search Terminal Help
(49540, Life With Boys: Season 1: Social Death with Boys,2011,3.4512068806066765,1368)
(49541, Life With Boys: Season 1: This Time the Problem Is with Dad and Not with Boys,2011,3.4512068806066765,1368)
(49542, Life With Boys: Season 1: In the Principal's Office with Boys,2011,3.4512068806066765,1368)
(49543, El Fuente: 599 MP,2013,3.5,471)
(49544, Life With Boys: Season 1: Monkey Talk with Boys,2011,3.4512068806066765,1368)
(49545, Bo Burnham: what.,2013,4.1,3614)
(49546, El Fuente: 50 MP,2013,3.3,465)
(49547, Life With Boys: Season 1: A Perfect Life with Boys,2011,3.4512068806066765,1368)
(49548, Life With Boys: Season 1: Wrestling with Boys,2011,3.4512068806066765,1369)
(49549, Life With Boys: Season 1,2011,4.1,2563)
(49550, El Fuente: 30 MP,2013,3.3,471)
(49551, El Fuente: 2997 MP,2013,3.2,471)
(49552, El Fuente: 25 MP,2013,3.3,465)
(49553, El Fuente: 20 MP,2013,3.5,484)
(49554, Max Steel,2013,4.1,2563)
(49555, Saving Santa,2013,3.8,5038)
(49556, Lilyhammer: Season 1 (Recap),2013,4.2,194)
(49557, Shinobi Girl,2012,2.6,2563)
(49558, GLOW: The Story of the Gorgeous Ladies of Wrestling,2012,3.7,4593)
(49559, Mitt (Trailer),2013,3.0,138)
(49560, My Hope America with Billy Graham: The Cross,2013,3.4512068806066765,1706)
(49561, The Square (Trailer),2014,3.6,154)
(49562, My Hope America with Billy Graham,2013,3.9,2563)
(49563, My Hope America with Billy Graham: Defining Moments,2013,3.4512068806066765,1791)
(49564, My Hope America with Billy Graham: Lose to Gain,2013,3.4512068806066765,1400)
(49565, American Addict,2013,3.5,5377)
(49566, My Hope America with Billy Graham,2013,3.9,2563)
(49567, El Fuente: 2997 MP10,2013,2.7,470)
(49568, El Fuente: 25 MP10,2013,2.9,464)
(49569, El Fuente: 24 MP10,2013,2.8,484)
(49570, El Fuente: 23976 MP10,2013,2.9,484)
(49571, The Short Game (Trailer),2013,4.1,156)
(49572, El Fuente: 599 MP10,2013,2.8,471)
(49573, El Fuente: 50 MP10,2013,2.9,464)
(49574, El Fuente: 30 MP10,2013,2.8,470)
(49575, Greg Fitzsimmons: Life on Stage,2013,3.3,3671)
(49576, Dave Foley: Relatively Well,2013,3.2,3446)
(49577, Barbie: Life in the Dreamhouse: Barbie Life in the Dreamhouse: Best of Family,2013,3.4512068806066765,1398)
(49578, Barbie: Life in the Dreamhouse: Barbie Life in the Dreamhouse: Best of Friends,2013,3.4512068806066765,1458)
(49579, Transformers Prime Beast Hunters: Predacons Rising,2013,4.2,3950)
(49580, Underground: The Julian Assange Story,2012,3.7,5665)
(49581, Curious George: A Very Monkey Christmas,2009,3.8,3438)
(49582, Mumpfie's White Christmas,1996,2.4,1358)
(49583, Lady Gaga &38; The Muppets' Holiday Spectacular,2013,3.1,3496)
(49584, Sunset Strip,2012,3.0,5778)
(49585, Silver Bells,2013,3.5,5287)
(49586, Winter Wonderland,2013,2.8,1812)
(49587, Top Gear: Series 19: Africa Special,2013,3.4512068806066765,6822)
(49588, Fireplace For Your Home: Crackling Fireplace with Music,2010,3.4512068806066765,3610)
(49589, Kate Plus E18ht,2010,2.7,2563)
(49590, Kate Plus E18ht: Season 1,2010,2.7,2563)
grunt> filter_data = FILTER movies by year==2008 and rating>'3';
grunt> grp_data = GROUP filter_data all;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(filter_data);
grunt> dump cnt_data]
```

```

Activities Terminal
hduser@sreeraj: ~

File Edit View Search Terminal Help
2023-03-01 12:39:01,577 [pool-9-thread-1] INFO org.apache.hadoop.mapred.Task - Task:attempt_local1462791640_0002_r_000000_0 is done. And is in the process of committing
2023-03-01 12:39:01,587 [pool-9-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2023-03-01 12:39:01,611 [pool-9-thread-1] INFO org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task 'attempt_local1462791640_0002_r_000000_0' to hdfs://localhost:54310/tmp/te
mp-795517020/tmp-106606030/_temporary/0/task_local1462791640_0002_r_000000
2023-03-01 12:39:01,613 [pool-9-thread-1] INFO org.apache.hadoop.mapred.Task - Task 'attempt_local1462791640_0002_r_000000_0' done.
2023-03-01 12:39:01,614 [pool-9-thread-1] INFO org.apache.hadoop.mapred.LocalJobRunner - Finishing task: attempt_local1462791640_0002_r_000000_0
2023-03-01 12:39:01,614 [Thread-58] INFO org.apache.hadoop.mapred.LocalJobRunner - reduce task executor complete.
2023-03-01 12:39:02,024 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Running jobs are [job_local1462791640_0002]
2023-03-01 12:39:05,529 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:39:05,531 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:39:05,532 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:39:05,540 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-03-01 12:39:05,541 [main] INFO org.apache.pig.tools.pigstats.MapReduceLauncher - Script Statistics:
HadoopVersion PigVersion Userid StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-03-01 12:38:59 2023-03-01 12:39:05 GROUP_BY,FILTER
Success!
Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReducetime Alias Feature Outputs
job_local1462791640_0002 1 1 n/a n/a n/a n/a n/a n/a n/a n/a n/a cnt_data,filter_data,grp_data,movies GROUP_BY,COMBINER hdfs://localhost:54310/tmp/te
mp-795517020/tmp-106606030,
Input(s):
Successfully read 49589 records (45247256 bytes) from: "/assignment/movies_cleaned_pandas.csv"
Output(s):
Successfully stored 1 records (30426101 bytes) in: "hdfs://localhost:54310/tmp/temp-795517020/tmp-106606030"
Counters:
Total records written : 1
Total bytes written : 30426101
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
Job DAG:
job_local1462791640_0002

2023-03-01 12:39:05,542 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:39:05,544 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:39:05,546 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:39:05,551 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-03-01 12:39:05,553 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-03-01 12:39:05,560 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-03-01 12:39:05,566 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(all,3278)
grunt> order_data = ORDER movies by duration desc;
grunt> limit_data = LIMIT order_data 1;
grunt> col_data = FOREACH limit_data generate movie,year,rating,duration;
grunt> dump col_data

```

Activities Terminal ▾

Wed 12:48 ●  
hduser@sreeraj: ~

```

File Edit View Search Terminal Help
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-03-01 12:43:56 2023-03-01 12:44:22 ORDER_BY,LIMIT

Success!

Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReduceTime Alias Feature Outputs
job_local1092562571_0003 1 0 n/a n/a n/a 0 0 0 movies MAP_ONLY
job_local1553724115_0005 1 1 n/a n/a n/a n/a n/a n/a order_data ORDER_BY,COMBINER
job_local2064503341_0006 1 1 n/a n/a n/a n/a n/a n/a col_data,order_data hdfs://localhost:54310/tmp/temp-795517020/tmp598698491,
Input(s):
Successfully read 49589 records (32084715 bytes) from: "/assignment/movies_cleaned_pandas.csv"

Output(s):
Successfully stored 1 records (83510978 bytes) in: "hdfs://localhost:54310/tmp/temp-795517020/tmp598698491"

Counters:
Total records written : 1
Total bytes written : 83510978
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1092562571_0003      -> job_local1092562571_0003
job_local1092562571_0004      -> job_local1553724115_0005
job_local1553724115_0005      -> job_local2064503341_0006
job_local2064503341_0006

2023-03-01 12:44:22,579 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,581 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,582 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,582 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,587 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,587 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,588 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,588 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,594 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,594 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,597 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,597 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,602 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,603 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,604 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:44:22,607 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-03-01 12:44:22,607 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-03-01 12:44:22,617 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-03-01 12:44:22,617 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapReduceUtil - Total input paths to process : 1
(Example: 8 Hour 23, 976, 2010, 3, 2, 28813)
grunt> order_data = ORDER
movies grp_data col_data order_data limit_data filter_data cnt_data
grunt> order_data = ORDER movies by rating asc;
grunt> limit_data = LIMIT order_data 1;
grunt> col_data = FOREACH limit_data generate movie,year,rating,duration;
grunt> dump col_data

```

```
Activities Terminal Wed 12:59 •
hduser@sreeraj: ~

File Edit View Search Terminal Help
(37423,Cirque du Soleil: Worlds Away,2012,4.0,5471)
(37404,The Central Park Five,2012,4.0,7135)
(42789,Skins: Vol. 7,2013,4.0,2563)
(8558,Better Off Ted: Season 2,2009,4.0,2563)
(42816,Moto 4: The Movie,2012,4.0,3187)
(37322,Beauty &#38; the Beast: Season 1,2012,4.0,2563)
(37310,Beauty &#38; the Beast,2012,4.0,2563)
(4600,MythBusters: Collection 2,2006,4.0,2563)
(8496,Dollhouse: Season 2,2009,4.0,2563)
(37289,WWE: Best of Raw and SmackDown 2011,2012,4.0,2563)
(37276,DreamWorks How to Train Your Dragon Legends,2011,4.0,2563)
(43074,B Gata H Kei - Yamada's First Time,2010,4.0,2563)
(37236,DreamWorks How to Train Your Dragon Legends,2011,4.0,2563)
(37233,America in PrimeTime,2011,4.0,2563)
(37209,The Sapphires,2012,4.0,5923)
(49084,Homicide Hunter: Lt. Joe Kenda,2011,4.0,2563)
(2697,Katt Williams: The Pimp Chronicles: Pt. 1,2006,4.0,2714)
(37146,Don't Trust the B---- in Apartment 23: Season 2,2012,4.0,2563)
(37146,Private Practice: Season 6,2012,4.0,2563)
(43154,Melissa &#38; Joey: Season 3,2013,4.0,2563)
(43181,Ben 10: Ultimate Alien: Season 1,2010,4.0,2563)
(4579,The Rape of Europa,2007,4.0,6979)
(37135,Crooked Arrows,2012,4.0,6305)
(37117,WWE: Best of Raw and SmackDown 2011,2012,4.0,2563)
(37054,Paranormal Witness: Season 1,2011,4.0,2563)
(37045,Paranormal Witness: Season 1,2011,4.0,2563)
(43201,Ben 10: Ultimate Alien,2010,4.0,2563)
(36963,Pair of Kings: Season 3,2012,4.0,2563)
(36936,Comanche Moon: Road to Lonesome Dove,2008,4.0,2563)
(43329,Tanked,2011,4.0,2563)
(36911,Silenced,2011,4.0,7489)
(8474,Ghost Whisperer: Season 5,2009,4.0,2563)
(43358,Generator Rex: Season 1,2010,4.0,2563)
(36780,Them Idiots! Whirled Tour,2012,4.0,7334)
(36692,The Electric Daisy Carnival Experience,2011,4.0,6960)
(8466,Damages: Season 3,2010,4.0,2563)
(36415,The Marinovich Project,2011,4.0,4813)
(36390,Ethos,2011,4.0,4120)
(36324,George Gently: Series 4,2011,4.0,2563)
(43443,Generator Rex,2010,4.0,2563)
(36222,WWE: The Ladder Match 2: Crash &#38; Burn,2011,4.0,2563)
(36216,WWE: The Ladder Match 2: Crash &#38; Burn,2011,4.0,2563)
(36192,Damages: Season 5,2012,4.0,2563)
(8437,The Mighty Macs,2009,4.0,5929)
(8436,The 5th Quarter,2010,4.0,5878)
(36133,Secrets of the Sun: Nova,2012,4.0,3191)
(36097,Brother White,2012,4.0,5466)
(8433,Greek: Chapter Three,2008,4.0,2563)
(36003,WWE's Greatest Rivalries: Shawn Michaels vs Bret Hart,2011,4.0,7525)
(8431,Masterpiece Classic: The Diary of Anne Frank,2009,4.0,6019)
(4180,Metalocalypse: Season 1,2006,4.0,2563)
(35982,Brooklyn Castle,2012,4.0,6133)
(35887,Dallas (2012),2012,4.0,2563)
grunt> order_data2 = ORDER filter_data by rating asc;
grunt> dump order data]
```

```

activities Terminal ▾ Wed 12:57 ●
hduser@seeraj:~ - 

File Edit View Search Terminal Help
2023-03-01 12:49:17,286 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100% complete
2023-03-01 12:49:17,322 [main] INFO org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:
HadoopVersion PigVersion UserId StartedAt FinishedAt Features
2.9.2 0.17.0 hduser 2023-03-01 12:48:53 2023-03-01 12:49:17 ORDER_BY,LIMIT
Success!
Job Stats (time in seconds):
JobId Maps Reduces MaxMapTime MinMapTime AvgMapTime MedianMapTime MaxReduceTime MinReduceTime AvgReduceTime MedianReduceTime Alias Feature Outputs
job_local1128513831_0010 1 n/a n/a n/a n/a n/a n/a n/a col_data,order_data hdfs://localhost:54310/tmp/temp-795517020/tmp442166848,
job_local1416230018_0009 1 1 n/a n/a n/a n/a n/a n/a order_data ORDER_BY,COMBINER
job_local367042287_0007 1 0 n/a n/a n/a n/a 0 0 0 movies MAP_ONLY
job_local919730706_0008 1 1 n/a n/a n/a n/a n/a n/a n/a order_data SAMPLER

Input(s):
Successfully read 49589 records (65851652 bytes) from: "/assignment/movies_cleaned_pandas.csv"

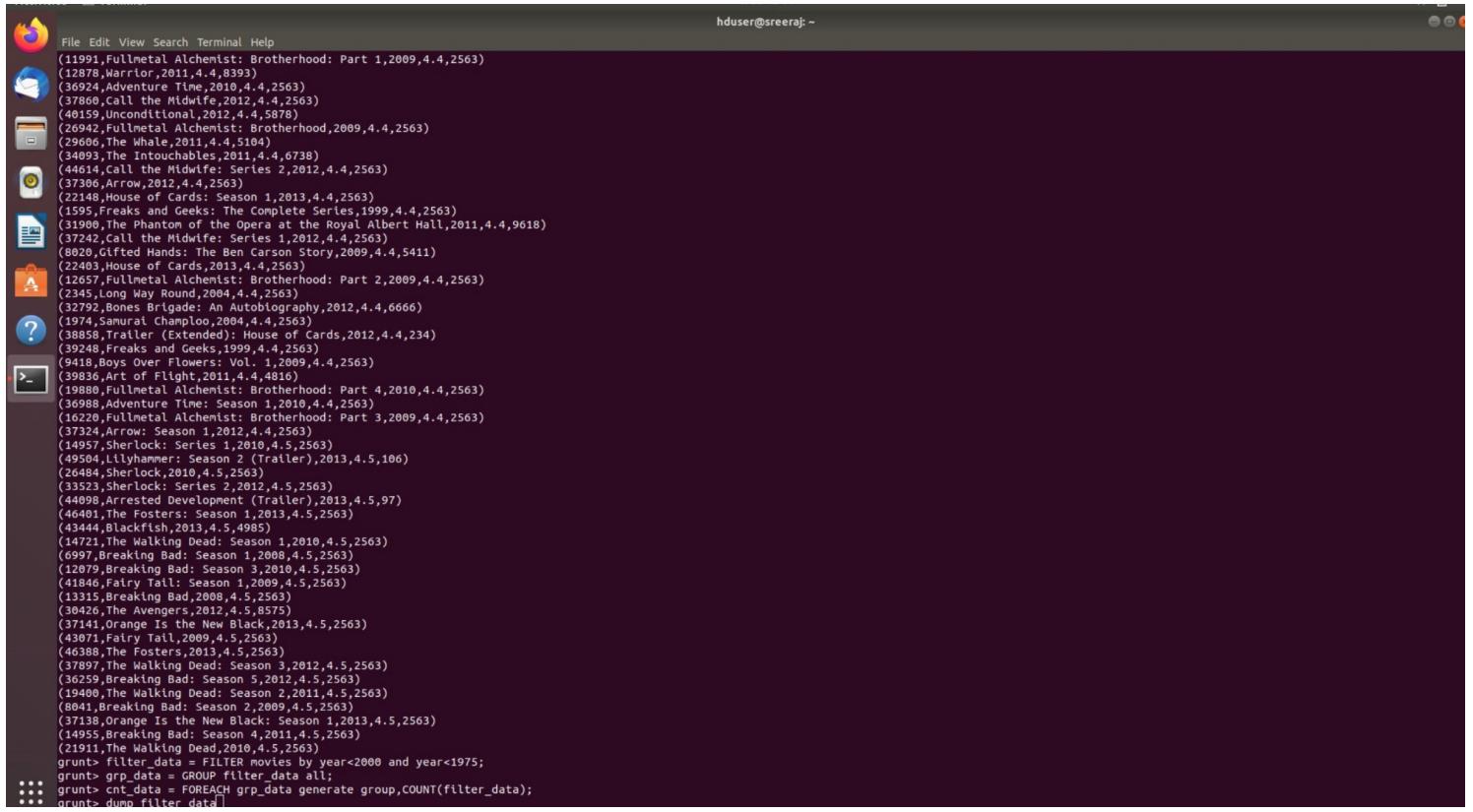
Output(s):
Successfully stored 1 records (136595843 bytes) in: "hdfs://localhost:54310/tmp/temp-795517020/tmp442166848"

Counters:
Total records written : 1
Total bytes written : 136595843
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local367042287_0007 -> job_local919730706_0008,
job_local919730706_0008 -> job_local1416230018_0009,
job_local1416230018_0009 -> job_local1128513831_0010,
job_local1128513831_0010

2023-03-01 12:49:17,324 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,325 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,326 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,329 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,330 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,331 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,334 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,335 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,336 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,339 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,341 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,342 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 12:49:17,344 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-03-01 12:49:17,346 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-03-01 12:49:17,350 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-03-01 12:49:17,351 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(Lagegl, 2007, 1, 4, 2563)
grunt> filter_data = FILTER movies by rating>'4' and year>2005;
grunt> order_data = ORDER filter_data by rating desc;
grunt> dump order data;

```



hduser@sreeraj: ~

```
(11991,Fullmetal Alchemist: Brotherhood: Part 1,2009,4.4,2563)
(36924, Adventure Time,2010,4.4,2563)
(37860, Call the Midwife,2012,4.4,2563)
(44614, Call the Midwife: Series 2,2012,4.4,2563)
(26942, Fullmetal Alchemist: Brotherhood,2009,4.4,2563)
(29696, The Whale,2011,4.4,5164)
(34893, The Intouchables,2011,4.4,6738)
(44614, Call the Midwife: Series 1,2012,4.4,2563)
(37396, Arrow,2012,4.4,2563)
(22148, House of Cards: Season 1,2013,4.4,2563)
(1595, Freaks and Geeks: The Complete Series,1999,4.4,2563)
(31980, The Phantom of the Opera at the Royal Albert Hall,2011,4.4,9618)
(37242, Call the Midwife: Series 1,2012,4.4,2563)
(8829, Gifted Hands: The Ben Carson Story,2009,4.4,5411)
(22403, House of Cards,2013,4.4,2563)
(12657, Fullmetal Alchemist: Brotherhood: Part 2,2009,4.4,2563)
(2345, Long Way Round,2004,4.4,2563)
(32792, Bones Brigade: An Autobiography,2012,4.4,6666)
(1974, Samurai Champloo,2004,4.4,2563)
(38858, Trailer (Extended): House of Cards,2012,4.4,234)
(39248, Freaks and Geeks,1999,4.4,2563)
(9418, Boys Over Flowers: Vol. 1,2009,4.4,2563)
(39836, Art of Flight,2011,4.4,4816)
(19880, Fullmetal Alchemist: Brotherhood: Part 4,2010,4.4,2563)
(36988, Adventure Time: Season 1,2010,4.4,2563)
(16220, Fullmetal Alchemist: Brotherhood: Part 3,2009,4.4,2563)
(37324, Arrow: Season 1,2012,4.4,2563)
(14957, Sherlock: Series 1,2010,4.5,2563)
(49584, Lilyhammer: Season 2 (Trailer),2013,4.5,106)
(26484, Sherlock,2010,4.5,2563)
(33523, Sherlock: Series 2,2012,4.5,2563)
(44498, Arrested Development (Trailer),2013,4.5,97)
(46481, The Fosters: Season 1,2013,4.5,2563)
(43444, Blackfish,2013,4.5,4985)
(14721, The Walking Dead: Season 1,2010,4.5,2563)
(6997, Breaking Bad: Season 1,2008,4.5,2563)
(12079, Breaking Bad: Season 3,2010,4.5,2563)
(41846, Fairy Tail: Season 1,2009,4.5,2563)
(13315, Breaking Bad,2008,4.5,2563)
(36426, The Avengers,2012,4.5,8575)
(37141, Orange Is the New Black,2013,4.5,2563)
(43071, Fairy Tail,2009,4.5,2563)
(46388, The Fosters,2013,4.5,2563)
(37897, The Walking Dead: Season 3,2012,4.5,2563)
(36259, Breaking Bad: Season 5,2012,4.5,2563)
(19400, The Walking Dead: Season 2,2011,4.5,2563)
(8641, Breaking Bad: Season 2,2009,4.5,2563)
(37138, Orange Is the New Black: Season 1,2013,4.5,2563)
(14955, Breaking Bad: Season 4,2011,4.5,2563)
(21911, The Walking Dead,2010,4.5,2563)
grunt> filter_data = FILTER movies by year>2000 and year<1975;
grunt> grp_data = GROUP filter_data all;
grunt> cnt_data = FOREACH grp_data generate group,COUNT(filter_data);
grunt> dump filter_data[]
```

```

HadoopVersion PigVersion     UserId  StartedAt      FinishedAt      Features
2.9.2  0.17.0  hduser  2023-03-01 11:45:59  2023-03-01 11:46:05  GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId  Maps   Reduces  MaxMapTime    MinMapTime    AvgMapTime    MedianMapTime  MaxReduceTime  MinReduceTime  AvgReduceTime  MedianReducetime   AliasF
eature Outputs
job_local1745821394_0003      1       1      n/a        n/a        n/a        n/a        n/a        n/a        n/a        cnt_data8,filter_data8.grp_data,movies  GROUP_BY,COMBI
NER   hdfs://localhost:54310/tmp/temp-1481442308/tmp-1287522238,

Input(s):
Successfully read 49589 records (58216816 bytes) from: "/luminar/movies_cleaned_pandas.csv"

Output(s):
Successfully stored 1 records (35985077 bytes) in: "hdfs://localhost:54310/tmp/temp-1481442308/tmp-1287522238"

Counters:
Total records written : 1
Total bytes written : 35985077
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1745821394_0003

2023-03-01 11:46:05,194 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 11:46:05,197 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 11:46:05,199 [main] INFO org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2023-03-01 11:46:05,205 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2023-03-01 11:46:05,206 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2023-03-01 11:46:05,210 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input files to process : 1
2023-03-01 11:46:05,210 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(all,3278)
grunt> filter_data14 = FILTER movies by year<2000 and year>1975 and rating>'3.5';
grunt> grp_data14 = GROUP filter_data14 all;
grunt> cnt_data14 = FOREACH grp_data14 generate group,COUNT(filter_data14);
grunt> dump cnt_data14

```