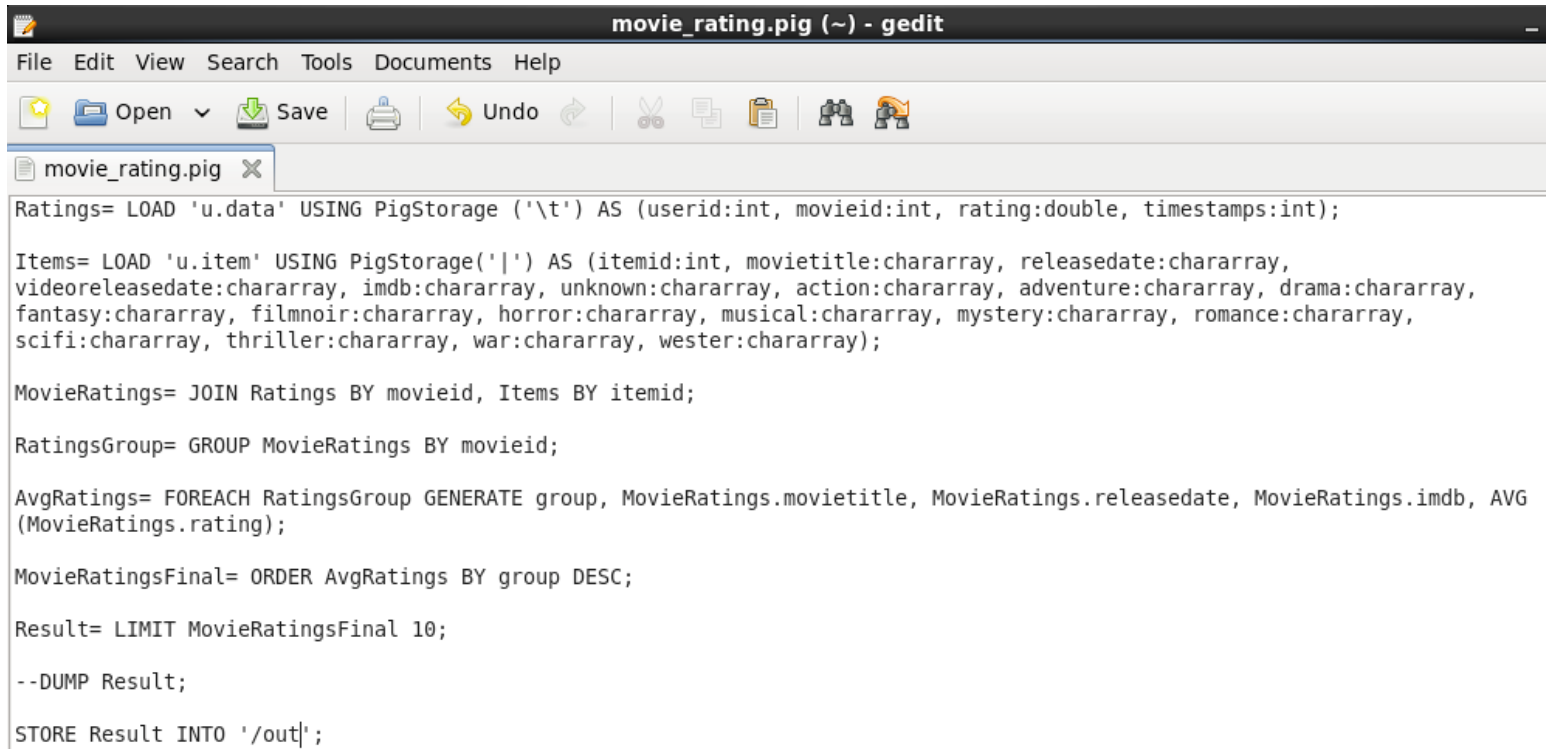


Assignment 8 Solution

STUDENT NAME: Chin-Ting Ko

1) Source code for movie_rating.pig



```
movie_rating.pig (~) - gedit
File Edit View Search Tools Documents Help
Open Save Undo
movie_rating.pig x
Ratings= LOAD 'u.data' USING PigStorage ('\t') AS (userid:int, movieid:int, rating:double, timestamps:int);

Items= LOAD 'u.item' USING PigStorage('|') AS (itemid:int, movietitle:chararray, releasedate:chararray,
videoreleasedate:chararray, imdb:chararray, unknown:chararray, action:chararray, adventure:chararray, drama:chararray,
fantasy:chararray, filmnoir:chararray, horror:chararray, musical:chararray, mystery:chararray, romance:chararray,
scifi:chararray, thriller:chararray, war:chararray, wester:chararray);

MovieRatings= JOIN Ratings BY movieid, Items BY itemid;

RatingsGroup= GROUP MovieRatings BY movieid;

AvgRatings= FOREACH RatingsGroup GENERATE group, MovieRatings.movietitle, MovieRatings.releasedate, MovieRatings.imdb, AVG
(MovieRatings.rating);

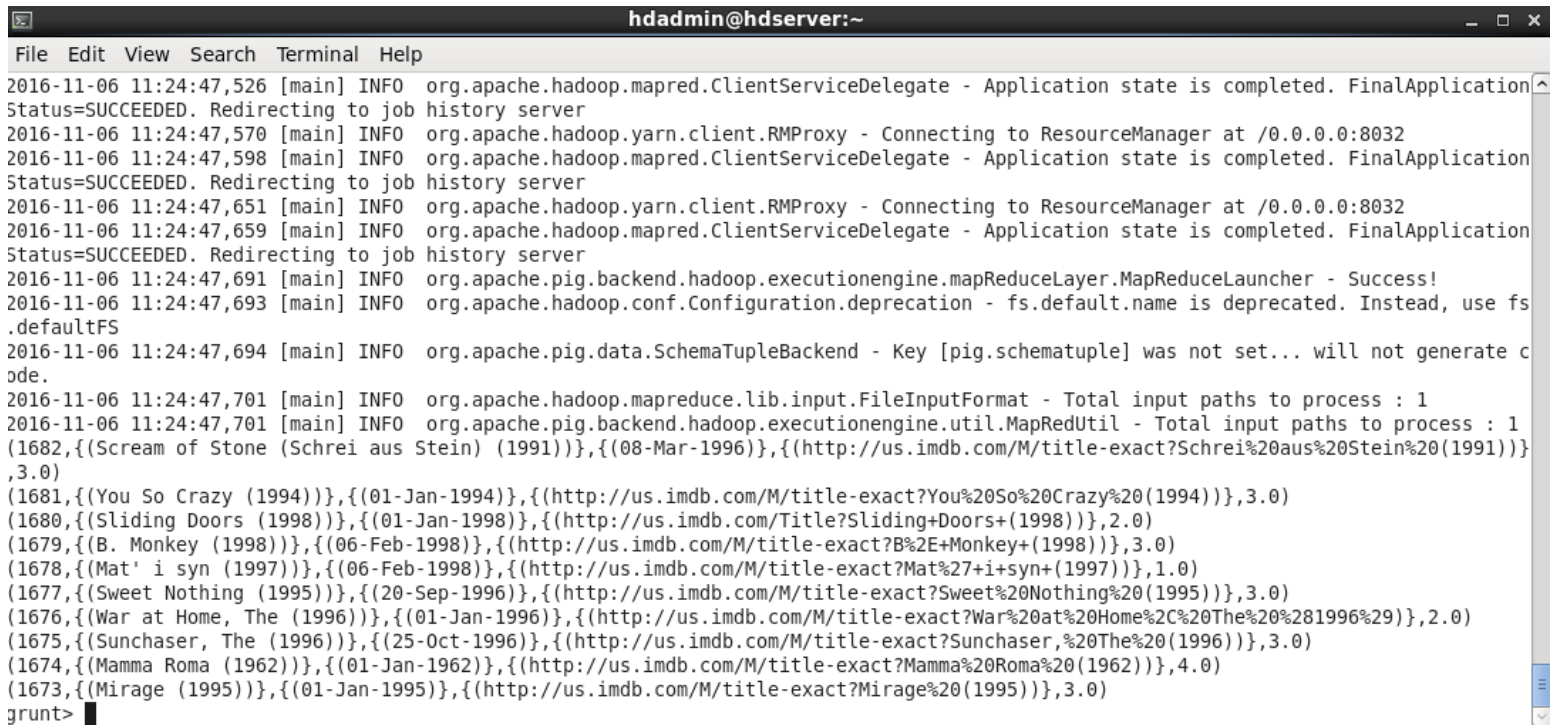
MovieRatingsFinal= ORDER AvgRatings BY group DESC;

Result= LIMIT MovieRatingsFinal 10;

--DUMP Result;

STORE Result INTO '/out|';
```

2) Run the program on command line and place the screenshot of successful run right here.

A screenshot of a terminal window titled 'hdadmin@hdserver:~'. The terminal displays a series of log messages from Hadoop and Pig. The logs indicate that several Hadoop MapReduce applications have completed successfully and are being redirected to the job history server. There are also messages about connecting to the Resource Manager and deprecated configuration options. The final part of the log shows Pig execution output, including a success message and a list of movie titles with their IMDb ratings and release dates. The terminal ends with a 'grunt>' prompt.

```
hdadmin@hdserver:~  
File Edit View Search Terminal Help  
2016-11-06 11:24:47,526 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplication  
Status=SUCCEEDED. Redirecting to job history server  
2016-11-06 11:24:47,570 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032  
2016-11-06 11:24:47,598 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplication  
Status=SUCCEEDED. Redirecting to job history server  
2016-11-06 11:24:47,651 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032  
2016-11-06 11:24:47,659 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplication  
Status=SUCCEEDED. Redirecting to job history server  
2016-11-06 11:24:47,691 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!  
2016-11-06 11:24:47,693 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs  
.defaultFS  
2016-11-06 11:24:47,694 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate c  
ode.  
2016-11-06 11:24:47,701 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1  
2016-11-06 11:24:47,701 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1  
(1682,{{Scream of Stone (Schrei aus Stein) (1991)}},{(08-Mar-1996)}},{(http://us.imdb.com/M/title-exact?Schrei%20aus%20Stein%20(1991))}  
,3.0)  
(1681,{{You So Crazy (1994)}},{(01-Jan-1994)}},{(http://us.imdb.com/M/title-exact?You%20So%20Crazy%20(1994))},3.0)  
(1680,{{Sliding Doors (1998)}},{(01-Jan-1998)}},{(http://us.imdb.com/Title?Sliding+Doors+(1998))},2.0)  
(1679,{{B. Monkey (1998)}},{(06-Feb-1998)}},{(http://us.imdb.com/M/title-exact?B%2E+Monkey+(1998))},3.0)  
(1678,{{Mat' i syn (1997)}},{(06-Feb-1998)}},{(http://us.imdb.com/M/title-exact?Mat%27+i+syn+(1997))},1.0)  
(1677,{{Sweet Nothing (1995)}},{(20-Sep-1996)}},{(http://us.imdb.com/M/title-exact?Sweet%20Nothing%20(1995))},3.0)  
(1676,{{War at Home, The (1996)}},{(01-Jan-1996)}},{(http://us.imdb.com/M/title-exact?War%20at%20Home%2C%20The%20%281996%29)},2.0)  
(1675,{{Sunchaser, The (1996)}},{(25-Oct-1996)}},{(http://us.imdb.com/M/title-exact?Sunchaser,%20The%20(1996))},3.0)  
(1674,{{Mamma Roma (1962)}},{(01-Jan-1962)}},{(http://us.imdb.com/M/title-exact?Mamma%20Roma%20(1962))},4.0)  
(1673,{{Mirage (1995)}},{(01-Jan-1995)}},{(http://us.imdb.com/M/title-exact?Mirage%20(1995))},3.0)  
grunt>
```

3) A screen shot of the JobHistoryServer when you ran the job. Click on the Job, get into the details, then take a screen shot.

MapReduce Job job_1478458998164_0008 - Mozilla Firefox

Welcome to CentOS x MapReduce Job job_14... x

localhost:19888/jobhistory/job/job_1478458998164_0008

hadoop

MapReduce Job job_1478458998164_0008

Logged in as: dr.who

Application

Job

- Overview
- Counters
- Configuration
- Map tasks
- Reduce tasks

Tools

Job Overview

Job Name: [PigLatin:file:///home/hdadmin/movie_rating.pig](#)

User Name: hdadmin

Queue: default

State: SUCCEEDED

Uberized: false

Submitted: Sun Nov 06 11:24:16 PST 2016

Started: Sun Nov 06 11:24:24 PST 2016

Finished: Sun Nov 06 11:24:40 PST 2016

Elapsed: 15sec

Diagnostics:

Average Map Time 5sec

Average Shuffle Time 3sec

Average Merge Time 0sec

Average Reduce Time 0sec

ApplicationMaster			
Attempt Number	Start Time	Node	Logs
1	Sun Nov 06 11:24:18 PST 2016	localhost:8042	logs

Task Type	Total	Complete
Map	1	1
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	1

4) For the movie id 376 and 495, what is the movie title, release date, IMDB URL, average rating, total number of unique users rated, total number of ratings for the movie. Please so please fill the data grid below:

Movie id	Title	Release date	IMDB URL	Average rating
376	Houseguest	01-Jan-1994	http://us.imdb.com/ M/title-exact? Houseguest %20(1994)	2.833
495	Around the world in 80 days	01- Jan-1956	http://us.imdb.com/ M/title-exact? Around %20the %20World %20in %2080%20 Da ys%20(1956)	3.661