**Pig Assignment Solution**

**STUDENT NAME:**

1. **Source code for movie\_rating.pig**

**--movie\_rating.pig**

**ratingTable = LOAD '/user/hive/warehouse/datatest/u.data' USING PigStorage('\t') AS (userID:int, itemID:int, rating:double, timestamp:int);**

**DESCRIBE ratingTable;**

**DUMP ratingTable;**

**itemTable = LOAD '/user/hive/warehouse/itemtest/u.item' USING PigStorage('|') AS (itemID:int, movieTitle:chararray, releaseDate:chararray, videoReleaseDate:chararray, imdbURL:chararray, unknown:int, action:int, adventure:int, animation:int, childrens:int, comedy:int, crime:int, documentary:int, drama:int, fantasy:int, filmNoir:int, horror:int, musical:int, mystery:int, romance:int, sciFi:int, thriller:int, war:int, western:int);**

**DESCRIBE itemTable;**

**DUMP itemTable;**

**ratingByItem = GROUP ratingTable by itemID;**

**DUMP ratingByItem;**

**movieInfo = FOREACH itemTable GENERATE itemID, movieTitle, releaseDate, imdbURL;**

**DUMP movieInfo;**

**DESCRIBE ratingByItem;**

**trimmedRatingTable = FOREACH ratingTable GENERATE itemID, rating;**

**DUMP trimmedRatingTable;**

**ratingInfo = GROUP trimmedRatingTable BY itemID;**

**DUMP ratingInfo;**

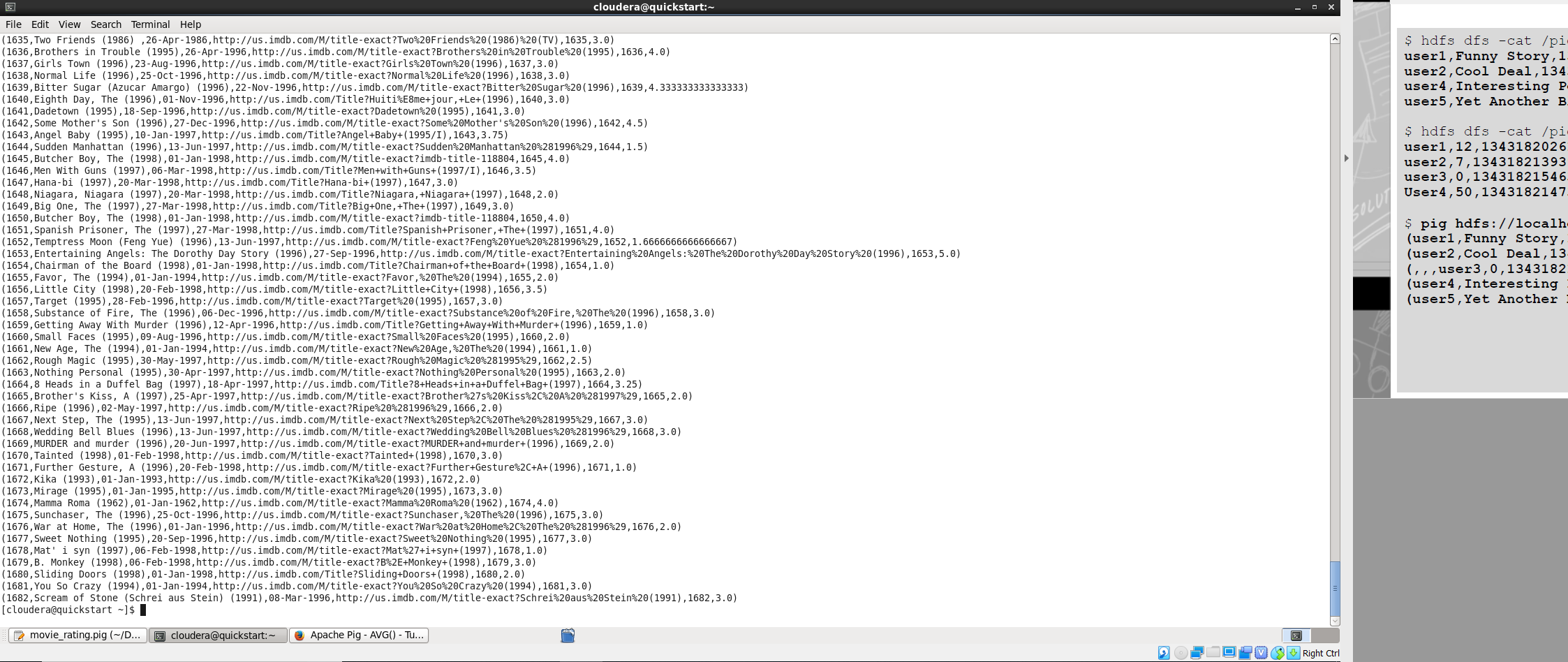
**ratingAverages = FOREACH ratingInfo GENERATE group AS itemID, AVG(trimmedRatingTable.rating) AS avgRating;**

**DUMP ratingAverages;**

**completeMovieList = JOIN movieInfo by itemID FULL OUTER, ratingAverages by itemID;**

**DUMP completeMovieList;**

1. **Sample run with output**



1. **Run the program on command line and place the screenshot of successful run right here.**
2. **A screen shot of the JobHistoryServer when you ran the job. Click on the Job, get into the details, then take a screen shot.**
3. **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** |  |  |  |  |
| **495** |  |  |  |  |