OSS
Project # 1 Bash Shell Programming

https://www.gnu.org/software/sed/manual/

Jinman Jung

jmjung@inha.ac.kr

Project Overview

- Make your script program for handling a input file in BASH
- Requirement
  - 1. Get the data of the movie identified by a specific 'movie id' from 'u.item'
  - 2. Get the data of ‘action’ genre movies from 'u.item’
  - 3. Get the average 'rating’ of the movie identified by specific 'movie id' from 'u.data’
  - 4. Delete the ‘IMDb URL’ from ‘u.item’
  - 5. Get the data about users from 'u.user’
  - 6. Modify the format of 'release date' in 'u.item’
  - 7. Get the data of movies rated by a specific 'user id' from 'u.data'
  - 8. Get the average 'rating' of movies rated by users with 'age' between 20 and 29 and 'occupation' as 'programmer'
  - 9. Exit

```
$ ./test.sh u.item u.data u.user
-------------------------
User Name: fos
Student Number: 00000000
[ MENU ]
1. Get the data of the movie identified by a specific
'movie id' from 'u.item'
2. Get the data of action genre movies from 'u.item’
3. Get the average 'rating’ of the movie identified by
specific 'movie id' from 'u.data’
4. Delete the ‘IMDb URL’ from ‘u.item
5. Get the data about users from 'u.user’
6. Modify the format of 'release date' in 'u.item’
7. Get the data of movies rated by a specific 'user id'
from 'u.data'
8. Get the average 'rating' of movies rated by users with
'age' between 20 and 29 and 'occupation' as 'programmer'
9. Exit
-------------------------
Enter your choice [ 1-9 ] 3

Please enter the 'movie id’(1~1682):1

average rating of 1: 3.87832

Enter your choice [ 1-9 ]
```

whoami

Input file (u.item)

- The genre of the movie is marked as 1
- 1682 rows

movie id | movie title | release date | video release date |IMDb URL |Genre...

| unknown | Action | Adventure | Animation | Children's | Comedy | Crime | Documentary | Drama | Fantasy | Film-Noir | Horror | Musical | Mystery | Romance | Sci-Fi | Thriller | War | Western |

Genre

1|Toy Story (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Toy%20Story%20(1995)|0|0|0|1|1|1|0|0|0|0|0|0|0|0|0|0|0
2|GoldenEye (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?GoldenEye%20(1995)|0|1|1|0|0|0|0|0|0|0|0|0|0|0|1|0|0
3|Four Rooms (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Four%20Rooms%20(1995)|0|0|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0
4|Get Shorty (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Get%20Shorty%20(1995)|0|1|0|0|0|1|0|0|1|0|0|0|0|0|0|0|0
5|Copycat (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Copycat%20(1995)|0|0|0|0|0|0|1|0|1|0|0|0|0|0|0|1|0|0
6|Shanghai Triad (Yao a yao yao dao waipo qiao) (1995)|01-Jan-1995||http://us.imdb.com/Title?Yao+a+yao+yao+dao+waipo+qiao+(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0
7|Twelve Monkeys (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Twelve%20Monkeys%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|1|0|0
8|Babe (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Babe%20(1995)|0|0|0|0|1|1|0|0|1|0|0|0|0|0|0|0|0
9|Dead Man Walking (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Dead%20Man%20Walking%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0
10|Richard III (1995)|22-Jan-1996||http://us.imdb.com/M/title-exact?Richard%20III%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|1|0
11|Seven (Se7en) (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Se7en%20(1995)|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|1|0|0
12|Usual Suspects, The (1995)|14-Aug-1995||http://us.imdb.com/M/title-exact?Usual%20Suspects,%20The%20(1995)|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|1|0|0
13|Mighty Aphrodite (1995)|30-Oct-1995||http://us.imdb.com/M/title-exact?Mighty%20Aphrodite%20(1995)|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0
14|Postino, Il (1994)|01-Jan-1994||http://us.imdb.com/M/title-exact?Postino,%20Il%20(1994)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|1|0|0|0|0
15|Mr. Holland's Opus (1995)|29-Jan-1996||http://us.imdb.com/M/title-exact?Mr.%20Holland's%20Opus%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
16|French Twist (Gazon maudit) (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Gazon%20maudit%20(1995)|0|0|0|0|0|1|0|0|0|0|0|0|0|0|1|0|0|0|0
17|From Dusk Till Dawn (1996)|05-Feb-1996||http://us.imdb.com/M/title-exact?From%20Dusk%20Till%20Dawn%20(1996)|0|1|0|0|0|1|1|0|0|0|0|1|0|0|0|0|1|0|0
18|White Balloon, The (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Badkonake%20Sefid%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0
19|Antonia's Line (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Antonia%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0
20|Angels and Insects (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Angels%20and%20Insects%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|1|0|0|0|0

Input file (u.data, u.user)

u.data

100000 rows

user id | movie id | rating | timestamp

196	242	3	881250949
186	302	3	891717742
22	377	1	878887116
244	51	2	880606923
166	346	1	886397596
298	474	4	884182806
115	265	2	881171488
253	465	5	891628467
305	451	3	886324817

u.user

943 rows

user id | age | gender | occupation | zip code

1|24|M|technician|85711
2|53|F|other|94043
3|23|M|writer|32067
4|24|M|technician|43537
5|33|F|other|15213
6|42|M|executive|98101
7|57|M|administrator|91344
8|36|M|administrator|05201

1. Get the data of the movie identified by a specific 'movie id' from 'u.item'

- Please enter the 'movie id’(1~1682)
- Print the data of the movie identified by the input 'movie id’.
- The data can be found in ‘u.item'

```
$ ./test.sh u.item u.data u.user
--------------------------
User Name: fos
Student Number: 00000000
[ MENU ]
1. Get the data of the movie identified by a specific 'movie id' from 'u.item'
2. Get the data of action genre movies from 'u.item’
3. Get the average 'rating’ of the movie identified by specific 'movie id' from 'u.data’
4. Delete the ‘IMDb URL’ from ‘u.item
5. Get the data about users from 'u.user’
6. Modify the format of 'release date' in 'u.item’
7. Get the data of movies rated by a specific 'user id' from 'u.data'
8. Get the average 'rating' of movies rated by users with 'age' between 20 and 29 and 'occupation' as 'programmer'
9. Exit
--------------------------
Enter your choice [ 1-9 ] 1

Please enter 'movie id'(1~1682):1

1|Toy Story (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Toy%20Story%20(1995)|0|0|0|1|1|1|0|0|0|0|0|0|0|0|0|0|0

Enter your choice [ 1-9 ]
```

whoami

2. Get the data of ‘action’ genre movies from 'u.item’

- Do you want to get the data of ‘action’ genre movies from 'u.item’?(y/n)
- Prints 10 data of action genre movies from 'u.item’
- Sort ascending by ‘movie id’
- Print format
  - ‘movie id’ ‘movie title’

```
Enter your choice [ 1-9 ] 2

Do you want to get the data of ‘action’ genre movies
from 'u.item’?(y/n):y

2 GoldenEye (1995)
4 Get Shorty (1995)
17 From Dusk Till Dawn (1996)
21 Muppet Treasure Island (1996)
22 Braveheart (1995)
24 Rumble in the Bronx (1995)
27 Bad Boys (1995)
28 Apollo 13 (1995)
29 Batman Forever (1995)
33 Desperado (1995)

Enter your choice [ 1-9 ]
```

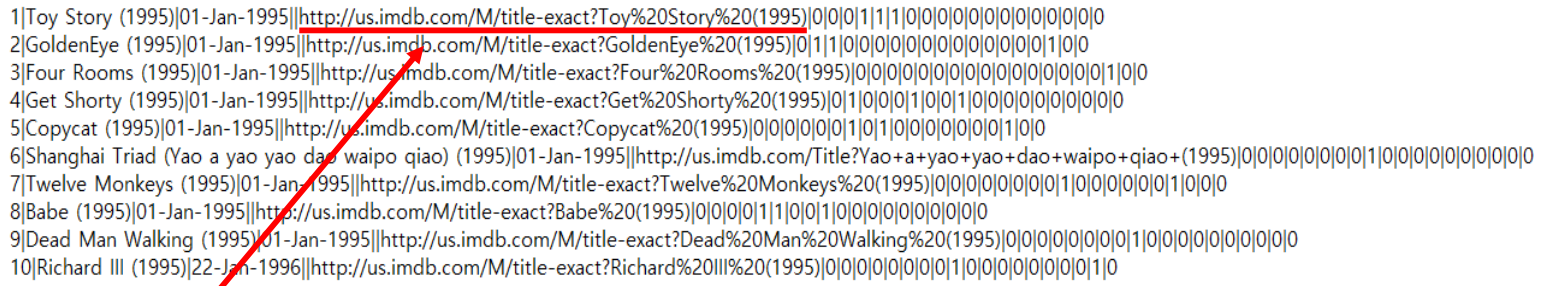
3. Get the average 'rating’ of the movie identified by specific 'movie id' from 'u.data’

- Please enter the 'movie id’(1~1682)
- Print the average ‘rating’ of the movie identified by input 'movie id’.
- The data for 'rating' can be found in 'u.user’
- Round the ‘average rating’ to six decimal places and print it with up to five decimal places
  - Ex) 3.878318 -> 3.87832
- Print format
  - average rating of ‘movie id’: ‘average rating’

```
Enter your choice [ 1-9 ] 3
Please enter the 'movie id’(1~1682):1
average rating of 1: 3.87832
```

4. Delete the ‘IMDb URL’ from ‘u.item’

- Do you want to delete the ‘IMDb URL’ from ‘u.item’?(y/n)
- Get the data from 'u.item,’ delete the 'IMDb URL,' and then print the data
- Print 10 lines only




```
Enter your choice [ 1-9 ] 4

Do you want to delete the ‘IMDb URL’ from ‘u.item’?(y/n):y

1|Toy Story (1995)|01-Jan-1995|||0|0|0|1|1|1|0|0|0|0|0|0|0|0|0|0|0|0
2|GoldenEye (1995)|01-Jan-1995|||0|1|1|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0
3|Four Rooms (1995)|01-Jan-1995|||0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0
4|Get Shorty (1995)|01-Jan-1995|||0|1|0|0|0|1|0|0|1|0|0|0|0|0|0|0|0|0
5|Copycat (1995)|01-Jan-1995|||0|0|0|0|0|0|1|0|1|0|0|0|0|0|0|1|0|0
6|Shanghai Triad (Yao a yao yao dao waipo qiao) (1995)|01-Jan-
1995|||0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
7|Twelve Monkeys (1995)|01-Jan-1995|||0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|1|0|0|0
8|Babe (1995)|01-Jan-1995|||0|0|0|0|1|1|0|0|1|0|0|0|0|0|0|0|0|0
9|Dead Man Walking (1995)|01-Jan-1995|||0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
10|Richard III (1995)|22-Jan-1996|||0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|1|0
```

5. Get the data about users from 'u.user’

- Do you want to get the data about users from ‘u.user’?(y/n)
- Print data about users from u.user
- Print format
  - user ‘user id’ is ‘age’ years old ‘gender’ ‘occupation’
- Print 10 lines only

```
Enter your choice [ 1-9 ] 5

Do you want to get the data about users from
‘u.user’?(y/n):y

user 1 is 24 years old male technician
user 2 is 53 years old female other
user 3 is 23 years old male writer
user 4 is 24 years old male technician
user 5 is 33 years old female other
user 6 is 42 years old male executive
user 7 is 57 years old male administrator
user 8 is 36 years old male administrator
user 9 is 29 years old male student
user 10 is 53 years old male lawyer
```

6. Modify the format of 'release date' in 'u.item’

- Do you want to Modify the format of ‘release data’ in ‘u.item’?(y/n)
- Get the data from 'u.item,’ modify the format of u.item's video release date to YYYYMMDD format, and then print the data
- Ex) 01-Jan-1995 -> 19950101
- Print only the last 10 lines (movie id: 1673 ~ 1682)

6. Modify the format of 'release date' in 'u.item’

1673|Mirage (1995)|01-Jan-1995||http://us.imdb.com/M/title-exact?Mirage%20(1995)|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0
1674|Mamma Roma (1962)|01-Jan-1962||http://us.imdb.com/M/title-exact?Mamma%20Roma%20(1962)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1675|Sunchaser, The (1996)|25-Oct-1996||http://us.imdb.com/M/title-exact?Sunchaser,%20The%20(1996)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|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|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1677|Sweet Nothing (1995)|20-Sep-1996||http://us.imdb.com/M/title-exact?Sweet%20Nothing%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1678|Mat' i syn (1997)|06-Feb-1998||http://us.imdb.com/M/title-exact?Mat%27+i+syn+(1997)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1679|B. Monkey (1998)|06-Feb-1998||http://us.imdb.com/M/title-exact?B%2E+Monkey+(1998)|0|0|0|0|0|0|0|0|0|0|0|0|0|0|1|0|1|0|0
1680|Sliding Doors (1998)|01-Jan-1998||http://us.imdb.com/Title?Sliding+Doors+(1998)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|1|0|0|0|0
1681|You So Crazy (1994)|01-Jan-1994||http://us.imdb.com/M/title-exact?You%20So%20Crazy%20(1994)|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0
1682|Scream of Stone (Schrei aus Stein) (1991)|08-Mar-1996||http://us.imdb.com/M/title-exact?Schrei%20aus%20Stein%20(1991)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0

```
Enter your choice [ 1-9 ] 6

Do you want to Modify the format of ‘release data’ in ‘u.item’?(y/n):y

1673|Mirage (1995)|19950101|http://us.imdb.com/M/title-exact?Mirage%20(1995)|0|1|0|0|0|0|0|0|0|0|0|0|0|0|0|1|0|0
1674|Mamma Roma (1962)|19620101||http://us.imdb.com/M/title-exact?Mamma%20Roma%20(1962)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1675|Sunchaser, The (1996)|19961025||http://us.imdb.com/M/title-exact?Sunchaser,%20The%20(1996)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1676|War at Home, The (1996)|19960101||http://us.imdb.com/M/title-exact?War%20at%20Home%2C%20The%20%281996%29|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1677|Sweet Nothing (1995)|19960920||http://us.imdb.com/M/title-exact?Sweet%20Nothing%20(1995)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1678|Mat' i syn (1997)|19980206||http://us.imdb.com/M/title-exact?Mat%27+i+syn+(1997)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
1679|B. Monkey (1998)|19980206||http://us.imdb.com/M/title-exact?B%2E+Monkey+(1998)|0|0|0|0|0|0|0|0|0|0|0|0|0|0|1|0|1|0|0
1680|Sliding Doors (1998)|19980101||http://us.imdb.com/Title?Sliding+Doors+(1998)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|1|0|0|0|0
1681|You So Crazy (1994)|19940101||http://us.imdb.com/M/title-exact?You%20So%20Crazy%20(1994)|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0|0|0|0
1682|Scream of Stone (Schrei aus Stein) (1991)|19960308||http://us.imdb.com/M/title-
exact?Schrei%20aus%20Stein%20(1991)|0|0|0|0|0|0|0|0|1|0|0|0|0|0|0|0|0|0
```

7. Get the data of movies rated by a specific 'user id' from 'u.data'

- Please enter the ‘user id’(1~943)
- Print the data for movies rated by the user identified by the input 'user id’
- Print format
  - Print “movie id” of all movies rated by the user identified by the input "user id“
  - Sort ascending by ‘movie id’
    - ‘movie id’|’movie id’|’movie id’|’movie id’…
  - Print “movie id” and ‘title’ of 10 movies rated by the user identified by the input "user id“
  - Sort ascending by ‘movie id’
    - ‘movie id’|’movie title’

```
Enter your choice [ 1-9 ] 7

Please enter the ‘user id’(1~943):12

4|15|28|50|69|71|82|88|96|97|98|127|132|133|143|15
7|159|161|168|170|172|174|191|195|196|200|202|203|
204|215|216|228|238|242|276|282|300|318|328|381|39
2|402|416|471|480|591|684|708|735|753|754

4|Get Shorty (1995)
15|Mr. Holland's Opus (1995)
28|Apollo 13 (1995)
50|Star Wars (1977)
69|Forrest Gump (1994)
71|Lion King, The (1994)
82|Jurassic Park (1993)
88|Sleepless in Seattle (1993)
96|Terminator 2: Judgment Day (1991)
97|Dances with Wolves (1990)
```

8. Get the average 'rating' of movies rated by users with 'ag e' between 20 and 29 and 'occupation' as 'programmer'

- Do you want to get the average 'rating' of movies rated by users with 'age' between 20 and 29 and 'occupation' as 'programmer'?(y/n)
- Print all average ratings of movies rated by programmers in their 20s(age: 20~29)
- Calculate the average 'rating' based only on ratings given by programmers in their 20s(age: 20~29)
- Round the ‘average rating’ to six decimal places and print it with up to five decimal places
- Print in this format
  - Print all in ascending order based on ‘movie id’
  - ‘movie id’ ‘average rating’

```
Enter your choice [ 1-9 ] 8

Do you want to get the average 'rating' of
movies rated by users with 'age' between 20 and
29 and 'occupation' as 'programmer'?(y/n):y

1 4.29412
2 3
3 3.5
4 3.7
5 3.25
7 4.22222
8 3.5
9 4.1
10 4
11 4.3125
.
.
.
1512 3
1513 2
1518 4
1531 3
1552 2
1597 1
1600 4
1621 1
1655 2
```

9. Exit

- The menu selection should continue to repeat until option 9 is chosen
- When option 9, 'Exit,' is selected, it will print 'Bye' and then exit the program

```
$ ./test.sh u.item u.data u.user
--------------------------
User Name: fos
Student Number: 00000000
[ MENU ]
1. Get the data of the movie identified by a
specific 'movie id' from 'u.item'
2. Get the data of action genre movies from
'u.item'
3. Get the average 'rating' of the movie
identified by specific 'movie id' from 'u.data'
4. Delete the 'IMDb URL' from 'u.item
5. Get the data about users from 'u.user'
6. Modify the format of 'release date' in 'u.item'
7. Get the data of movies rated by a specific
'user id' from 'u.data'
8. Get the average 'rating' of movies rated by
users with 'age' between 20 and 29 and
'occupation' as 'programmer'
9. Exit
--------------------------
Enter your choice [ 1-9 ] 1

Please enter 'movie id'(1~1682):1

1|Toy Story (1995)|01-Jan-
1995||http://us.imdb.com/M/title-
exact?Toy%20Story%20(1995)|0|0|0|1|1|1|0|0|0|0|0|0
|0|0|0|0|0|0

Enter your choice [ 1-9 ] 3

Please enter the 'movie id'(1~1682):1

average rating of 1: 3.87832

Enter your choice [ 1-9 ] 9
Bye!
```

Additional requirements

- Do not use Python and other libraries such as pandas
- Be sure to print your name and student id.
- Upload the code to GitHub and submit the GitHub link

Hint

- Use awk for problem 1 to 3
- Use sed for problem 4 to 6
- Awk can designate a delimeter using the -F option
  - Ex) cat file | awk –F\| ‘{print $0}’ => delimeter : “|”
- Awk can use shell variables inside awk using the -v option
  - ex)

```
$ temp_a=tmp_a
$ seq 3 | awk -v a=$temp_a '{print $1, a}'
1 tmp_a
2 tmp_a
3 tmp_a
ubuntu@ubuntu-VirtualBox:~/openSW/movie_data_editor
```

- The action after the ' END' pattern is executed after the action before the ‘END' pattern has been executed for all records
- The ‘END' pattern can be useful when calculating the average rating

```
$ seq 3 | awk '{print $1} {sum+=$1} END {print sum}'
1
2
3
6
```

Submission

- Submission Files
  - Submit your script file named proj1_*SID*_*ENGNAME*.sh
    - *SID*: student number, *ENGNAME*: your name
    - Ex) prj1_12345678_honggildong.sh
  - Your submission also should contain README.doc (.hwp, .pdf) file for detailed description
    - For readme file, please describe them in detail in a report format.
  - Upload the code to GitHub and submit the GitHub link
    - Submit it as a text file
- Due date
  - 11/05(Sun) 23:59
  - Your files should be submitted through the Hi portal
- TA will verify your submissions using the another auto-script and copy checking tools
- Enjoy BASH!