Opensource Software Project #1 Bash Shell Programming Details for commands Report

12211652 Woo Byeong Jae

First, typing '#1 /bin/bash' on the 1st line to tell the system to start th e shell and execute the commands.

The 'clear' command means remove all the code from the shell so that we can see only the menus we're going to choose.

From line 4 to line 17, printing user information and details about the menus that can be executed. The 'echo' command is used for print som ething on the shell, so use that command only for the printing menus.

Line 19, declare flag variable named as 'stop' to use for stopping loop.

From line 20 to line 201 the whole codes constitute a big loop executing commands requested by user until the user input exit-code.

Line 22, printing instruction and accepting the command requested by user. The requested is stored in 'request' variable.

From line 26 to line 31, (also known as command 1) these are codes about extracting specific data among the u.item file. The line 28 request user's input and stored that to 'answer1' variable. And then for the readability, using ' echo "" ' command to line feed.

Line 30 is the key-point, use the 'cat' command to display contents of the specific file and use '|' command meaning pipe to next command.

The 'awk' command is used for extracting some data that satisfying conditions. '-F\|' option meaning that modify the default delimiter to the '|'. The \ is escape character for release the special meaning of specific character.

And the '-v' option allows the conversion external variable into the awk command variable. So we can use the answer1 variable in awk command.

Consequently, divide each row from the u.item file separately by | and extract only the first field is equal to ‘answer1’ variable. And print that line including the valid field. The ‘$0’ means all of the field.

From line 34 to line 53, these are codes extract only the data record including specific data and then display on the shell.

The difference between command 1 is that this command contains the re-confirm process and distinguishes the codes will be executed. Line 38 demonstrate the above content.

If the user inputs ‘y’ meaning yes, the codes related to display data record are executed. Line 42 is the key-point statement. The ‘$7’ is 7-th field among each row of data, meaning ‘action genre’. And this field is either 0 or 1, 1 means that this movie belongs to action genre.

And for the display only the 10 movies, one more command has been added that containing ‘NR’. ‘NR’ can be said as index of row simply. So, using ‘NR<=10’ option to display 10 data row.

Else if the user inputs ‘n’ meaning no, no codes are executed and just break out to the command 2. If the user inputs neither ‘y’ nor ‘n’ the codes request valid inputs to the user again until the inputs is appropriate.

From line 56 to line 63, (also known as command 3) these are codes to get the average data of specific data record. Line 58 request the Id to the user in order to extract valid data among u.data file.

Line 60 is command for getting how much data is valid in the u.data file. And move the vertically sorted data to ‘wc -l’ command using pipe. ‘wc’ stands for word count and ‘-l’ option means return the count of line. Lastly, store the result into ‘quantity’ variable.

Line 61 is command for getting the summation of specific field. The third command means that extract all of the specific field and then add to ‘sum’ variable. After the summation, return the value of ‘sum’ variable to store actual local variable.

Line 62 is command for calculating the result. But if the ‘bc -l’ command is omitted, the value would be displayed in integer format. So that command is required essentially.

From line 66 to line 83, (also known as command 4) these are codes to delete specific field in the all data record. And this command also contains re-confirm process like command 2.

Line 74 is the key-point statement. For only displays 10 rows, the ‘NR<=10’ condition is used. And then using the pipe so that delete the specific field.

In the ‘sed’ command, ‘-E’ option means that regular expression will be used. And ‘s/http[^\|]*\|/\|g’ looks quite complex, but we can examine step by step.

‘/’ is for used separate commands and options. ‘s’ means that changing field command. ‘http[^\|]*\|’ is regular expression, searching specific string satisfying expression condition.

First, start with ‘http’, and the ‘[^\|]*’ means zero or more sequences of expression ‘[^\|]’. ‘^’ means negative, ‘\|’ is ‘|’ character. The ‘\’ character is used for release the special meaning of specific character.

Consequently, start with ‘http’ string and searching sequences until the ‘|’ character is found. And then convert this string as just ‘|’ character.

From line 86 to line 109, (also known as command 5) these are codes to display information of data as reconstructed format.

Line 93 is command for extract 10 data record from u.user file. and then store that data to the temp.txt file.

Line 94 is command for add some string at the first position. The ‘^’ means the start point of row.

Line 95 is also command for add some string after the second word. The ‘\s’ means any word and specify the index at the last option in command.

Line 96 is almost as same as Line 95.

Line 97 and line 98 are commands for change some string not to add.

And above 5 lines, the temp.txt and temp2.txt files are used alternately to prevent overwrite issue.

From line 112 to line 131, (also known as command 6) these are codes to modify specific field as new format.

Line 120 is the key-point statement. The '([0-9]{2})-[a-zA-Z]{3}-([0-9]{4})' means start two digits and three characters and four digits. and grouped the digits using brace. The '\2\1' means second group and first group. In other words, '([0-9]{4})([0-9]{2})'.

From line 134 to line 151, (also known as 7), these are codes to display information about specific data requested by user.

Line 136 requests user's input and stored that to the 'answer7' variable. And line 137 extracts some valid data record among u.data file and then stored that to temp.txt file.

Line 138 formats the data as 'integer | integer | integer …' and then store that to temp2.txt file.

Line 139 removes last '|' character in temp2.txt file and then store that to horizon.txt file.

From line 143 to line 149 these are the codes for internal loop display 10 data record. Line 145 extract first data record in temp.txt file. And line 146 extract specific valid data record that corresponds to extracted value from line 145. And print format as 'first field | second field'

Line 147 delete first data record and then store that to temp2.txt file. And copy the contents of temp2.txt file to the temp.txt file.

From line 154 to line 189, (also known as command 8) these are codes to get average of data field of all data record that satisfying some conditions.

Line 160 extracts some data records that satisfying three conditions among u.user file and then store that to userId.txt file. And get the number of data record in userId.txt file and then store that to 'count' variable.

Line 166 extracts some data records that satisfying specific condition related to contents of userId.txt file among u.data file and then store that to movieList.txt file. This command repeats as much as the value of the ‘count’ variable.

And like command 7, line 167 and line 168 execute that deleting first data record and copying to another txt file.

Line 170 declare ‘count’ variable as the number of data record of u.item file.
From line 172 to line 181, looping from 0 to value of ‘count’ variable.

Line 174 extracts some data records satisfying the specific field is equals to index of loop and then store that to temp.txt file. But the point is there can be nothing stored in temp.txt file. So line 176 represents that if there are no records, omitting the loop and start next loop process.

Line 179 is command for store result data from temp.txt file that average of specific data field and then store that to ‘result’ variable.

Line 180 displays index of loop and ‘result’ variable.

From line 192 to line 196, (also known as command 9) these are codes to exit the program. Print ‘Bye’ and exit the program.

From line 198 to the end, these are codes to control user’s InputException. If the user’s input is not the 1-9 integer, request the input again until the correct answer.

Conclusion

Coding the whole program according to demands is not an easy job. But as reviewing the syntax of each command and working on assignment while has provided a great opportunity to enhance my practical skills. This assignment gives me enormous fulfilling.