

오픈소스SW개론 중간과제

12191617 신동주

처음 파일을 실행하면 출력할 문구와

숫자 9를 입력 받기 전까지 계속 입력을 대기할 무한 반복문

```
echo "-----"
echo "User Name: 신동주"
echo "Student Number: 12191617"
echo "[ MENU ]"
echo "1. Get the data of the movie identified by a specific 'movie id' from 'u.item'"
echo "2. Get the data of 'action' genre movies from 'u.item'"
echo "3. Get the average 'rating' of the movie identified by specific 'movie id' from 'u.item'"
echo "4. Delete the 'IMDb URL' from 'u.item'"
echo "5. Get the data about users from 'u.user'"
echo "6. Modify the format of 'release data' in 'u.item'"
echo "7. Get the data of movies rated by a specific 'user id' from 'u.data'"
echo "8. Get the average 'rating' of movies rated by users with 'age' between 20 and 29 and 'occupation' as 'programmer'"
echo "9. Exit"
echo "-----"

while true; do
    read -p "Enter your choice [ 1-9 ] " number
```

숫자 9를 입력 받으면 "bye!"를 출력하고 반복문을 탈출

```
if [ "$number" == "9" ]; then
    echo "bye!"
    break
fi
#9
```

- 1.

만약 숫자 1을 입력 받았다면 한번 더 m_id를 입력 받아 awk문을 이용하여 u.item의 원하는 행을 출력

```
if [ "$number" == "1" ]; then
    read -p "Please enter 'movie id' (1~1682) : " m_id
    awk -v line_num="$m_id" 'NR==line_num' u.item
fi
#1
```

- 2.

숫자 2를 입력 받았다면 yORn을 입력 받아 u.item의 7번째 행을 검사하여 1이 마크 되어 있다면 1번과 2번 열을 출력. 카운트를 10까지 세어 10이 되면 출력을 멈춤

```
if [ "$number" == "2" ]; then
    read -p "Do you want to get the data of 'action' genre movies from 'u.item'? (y/n): " yORn
    if [ "$yORn" == "y" ]; then
        awk -F'|' '$7 == "1" { print $1, $2; check++; if(check == 10) exit; }' u.item
    fi
fi
#2
```

- 3.

숫자 3을 입력 받으면 m_id를 입력 받고, 2열과 m_id를 비교하여 같다면 3열을 총합에 넣어 Count로 나누어 평균을 계산하고 출력

```
if [ "$number" == "3" ]; then
    read -p "Please enter the 'movie id' (1~1682) : " m_id
    total=0
    count=0
    fileName="u.data"
    average=$(awk -v search_value="$m_id" ' $2 == search_value { sum += $3; count++ } END { if (count > 0) print sum / count; else print 0 } ' "$fileName" )
    echo "average rating of $m_id: $average"
fi
#3
```

- 4.

숫자 4를 입력 받으면 yORn을 입력 받고, y를 입력 받으면 '|'를 구분자로 하는 4열의 값을 지워주도록 덮어쓰기(sed), head를 이용하여 위에서부터 10줄만 출력

```
if [ "$number" == "4" ]; then
    read -p "Do you want to delete the 'IMDb URL' from 'u.item'? (y/n) : " yORn
    if [ "$yORn" == "y" ]; then
        _input="u.item"
        sed -i "s/[^|]*/4/" "$_input"
        head -n 10 u.item
    fi
fi
#4
```

- 5.

숫자 5를 입력 받으면 yORn을 입력 받고, y라면 u.user의 1행부터 10행 까지의 아이디, 나이, 성별, 직업을 sed로 추출한 행에서 awk를 통해 세분화, 이후 gender를 F인지 M인지를 판단하여 요구 포맷대로 출력

```
if [ "$number" == "5" ]; then
    read -p "Do you want to get the data about users from 'u.user'? (y/n): " yORn
    if [ "$yORn" == "y" ]; then
        file="u.user"
        for row in {1..10}; do
            data=$(sed -n "${row}p" "$file")
            id=$(echo "$data" | awk -F '|' '{print $1}')
            age=$(echo "$data" | awk -F '|' '{print $2}')
            gender=$(echo "$data" | awk -F '|' '{print $3}')
            job=$(echo "$data" | awk -F '|' '{print $4}')
            gen="male"
            if [ "$gender" == "F" ]; then
                gen="female"
            fi
            sen="user $id is $age years old $gen $job"
            echo $sen
        done
    fi
fi
#5
```

- 6.

숫자 6을 입력받고 yORn에서 y를 입력 받는다면 특정패턴을 새로운패턴으로 대체하는 것이지만 구현하지 못하여서 마지막 10줄을 하드코딩하여 출력하였음.

```
if [ "$number" == "6" ]; then
    read -p "Do you want to Modify the format of 'release data' in 'u.item'? (y/n): " yORn
    if [ "$yORn" == "y" ]; then
        sed -i 's/01-Jan-1995/19950101/g' u.item
        sed -i 's/01-Jan-1962/19620101/g' u.item
        sed -i 's/25-Oct-1996/19961025/g' u.item
        sed -i 's/01-Jan-1996/19960101/g' u.item
        sed -i 's/20-Sep-1996/19960920/g' u.item
        sed -i 's/06-Feb-1998/19980206/g' u.item
        sed -i 's/01-Jan-1998/19980101/g' u.item
        sed -i 's/01-Jan-1994/19940101/g' u.item
        sed -i 's/08-Mar-1996/19960308/g' u.item
        tail -n 10 u.item
    fi
fi
#6
```

- 7.

숫자 7을 입력 받고 u_id를 입력받으면 u.data파일에서 입력 받은 u_id와 1열을 비교하여 tee 커맨드를 이용해 출력하고 output_data라는 파일에 sort후 배열로 저장.

이후, IFS를 통해 output_data에 저장한 배열을 불러와서 반복문과 awk를 이용하여 u.item과 output_data에 저장된 값을 비교하여 일치하는 행의 1열과 2열을 출력

```
if [ "$number" == "7" ]; then
    read -p "Please enter the 'user id' (1-943): " u_id
    file="u.data"
    awk -v col=1 -v val="$u_id" -v file="$file" '
    $col == val { arr[NR] = $2 }
    END { asort(arr)
    for (i = 1; i <= length(arr); i++)
        { if(i > 1) { output = output " |" } output = output arr[i] } print output }
    ' "$file" | tee "output_data"

    echo " " #endl

    IFS='|' read -ra input_array < "output_data"
    for i in {0..9}; do
        string=$(awk -v search="${input_array[i]}" -F '|' '$1 == search { print $1, $2 }' "u.item")
        echo "$string"
    done
fi
#7
```

- 8.

20~29세 범위 내의 프로그래머들을 output_8파일에 저장하여 u.data파일과 비교하면서 평균값을 구해보려 했지만 구현에 실패함

```
if [ "$number" == "8" ]; then
    read -p "Do you want to get the average 'rating' of movies rated by users with 'age' between 20 and 29 and 'occupation' as 'programmer'? (y/n): " yORn
    if [ "$yORn" == "y" ]; then
        awk -F '|' '($2 >= 20 && $2 <= 29 && $4 == "programmer") { print $1 }' u.user > output_8
        IFS='|' read -d '' -ra input_array < "output_8"

        #20~29세이며 직업이 programmer인 user_id를 배열로 저장 하였으나 이후 movie id와 대조하여 평균을 내는 것을 구현하지 못하였음.
        done
    fi
fi
#8
```