

Fundamentals of awk



Andrew Mallett

@theurbanpenguin | www.theurbanpenguin.com

Danny's Plans with awk



- Why awk?
- Review of the basics
- Formatting the employee file
- The problem with `/usr/bin/lastlog` ...

Why awk.....?

- Headers with BEGIN
- Footers with END
- Conditions and Loops
- Internal Function

```
$ cat users.awk
BEGIN { FS=":" ; print "Username" }
{ print $1 }
END { print "Total users = " NR }
$ awk -f users.awk /etc/passwd
```

Printing the login name from /etc/passwd

Here we use an awk file with the three code blocks

Note the internal variables **FS** and **NR**

Execute with awk and **-f**

```
$ cat users.awk
BEGIN { FS=":" ; print "Username" }
$3 > 499 { print $1 }
END { print "Total users = " NR }
$ awk -f users.awk /etc/passwd
```

Print where the UID is greater than 499

Adding the range in front of the main block we can use field values to control the display of row

```
$ cat users.awk
BEGIN { FS=":" ; print "Username" }
/^root/{ print $1 ; count++ }
END { print "Total users = " count }
$ awk -f users.awk /etc/passwd
```

Counting returned rows

However the **NR** variable still shows all lines in the password file

We add the **count** variable to show processed lines

Semi-colon is used to separate commands



Demo Time: Getting to know awk

Danny thinks he can simplify
formatting the employee file

```
$ awk -F",," { print toupper($1), tolower($2), $3 } employees  
$ sed 's@\([^,]*\)@U\1L\2@' employees
```

Choosing the right tool

See how more readable the awk code is compared to the previous solution with sed



Demo Time: Review functions in awk

/usr/bin/lastlog

Danny has identified the limitations of the lastlog command and, subsequently, he would like to be able to more easily review user accounts that have not been active. He feels awk will be the solution.

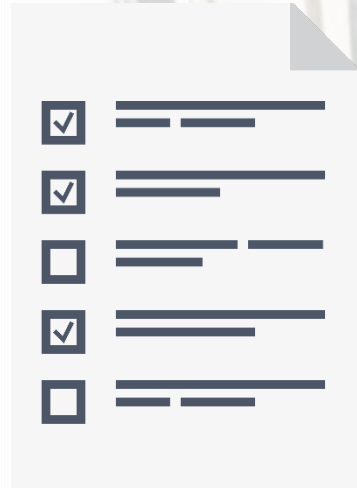
lastlog.awk

```
!(/Never logged in/ || /^Username/ || /^root/) {  
count++  
if ( NF == 8 )  
    printf "%8s %2s %3s %4s\n", $1,$5,$4,$8  
else  
    printf "%8s %2s %3s %4s\n", $1,$6,$5,$9}  
END {print "===== "  
print "Total Number of Users Processed: ", count}
```



Demo Time: Implement a Better lastlog

Summary



Danny awarded employee of the month
awk basics

Reviewed the employee formatting

Produced a better lastlog

Next up: Parsing XML data from
flat files with awk