

AWK



AWK

Created in 1977

Named after creators Aho, Weinberger, Kernighan

Inspired perl, a more common modern language.

Very powerful for 1-line commands

See wikipedia for more background on AWK

<http://en.wikipedia.org/wiki/AWK>

Printing in awk

An AWK program is a series of pattern action pairs:

```
condition { action }
```

The default action is to print the current line (\$0 in awk), so if no action is specified, that is what is performed, if the condition is true.

```
awk '1 {print $0}' mt.vcf
```

Is the same as

```
cat mt.vcf
```

You can also print out particular columns (fields)

\$1 is column 1, \$2 is column 2, etc.

Please use awk to print out the quality scores

If statements in awk

The other half of awk commands is the if statement, outside of the brackets

```
awk '$4 == "A" {print $0}' mt.vcf
```

Which is also the same as:

```
awk '$4 == "A"' mt.vcf
```

What does this do?

```
awk '$2 < 40000' mt.vcf
```

AWK

`awk 'IF {what you want to do }'`

<http://www.catonmat.net/blog/awk-one-liners-explained-part-one/>

awk

- Built for parsing large text files and tables
- `awk '{ print FNR "\t" $0 }'`
`allsnps.txt > numbered.txt`
- `awk '$4 ~ /^[AGCT]$/ && $5 ~`
`/^[AGCT]$/'` `ler_snps_indels.txt`
 - Look only as snps – no indels!
 - Also: `grep -v 'INDEL' ler_snps_indels.txt`
- <http://www.catonmat.net/blog/awk-one-liners-explained-part-one/>

Awk homework for tuesday

- Please use awk to create a SNP table, with no INDELS, and a minimum quality score of 100
- Submit this SNP table to D2L. Also keep track of a complete list of commands that you ran to create that SNP file.

And, to learn more awk, read:

- <http://www.catonmat.net/blog/awk-one-liners-explained-part-one/>