



# Introducción a UNIX: Trabajo Práctico 2

## Anexos



Daniel Millán  
Evelin Giaroli & Nora Moyano

*Facultad de Ciencias Aplicadas a la Industria, UNCuyo*  
San Rafael 5600, Argentina  
Junio de 2017

### 1. awk (Aho, Weinberger y Kernigan)

**awk**, es un programa de Unix que interpreta programas escritos en el lenguaje de programación AWK. AWK es un lenguaje de programación diseñado para procesar datos basados en texto, ya sean ficheros o flujos de datos.

Debido a su densa notación, frecuentemente es usado para escribir programas de una línea.

AWK fue una de las primeras herramientas en aparecer en Unix (en la versión 3) y ganó popularidad como una manera de añadir funcionalidad a las tuberías de Unix. El lenguaje AWK es estándar en casi todo sistema operativo tipo Unix. AWK es mencionado en las *Single UNIX Specification* (especificaciones básicas de unix) como una de las utilidades necesarias de todo sistema operativo Unix.

**awk** es útil para manipular archivos que contienen columnas de datos de la misma longitud. Como **sed**, **awk** puede ser utilizado directamente mediante la línea de órdenes, o se puede escribir un archivo de secuencia de órdenes o guión de órdenes (**script**) y dejar que **awk** interprete las órdenes desde este archivo. Esto suele ser muy útil cuando realizamos varias operaciones sobre los datos.

Por ejemplo, si tenemos un archivo “nbateams.txt” con los resultados para cada equipo de baloncesto de la NBA durante la temporada 2015, el cual contiene información para cada equipo de los partidos jugados, los goles realizados, los intentos, etc (ver “README\_nbateams.txt”):

Team	GP	FGM	FGA	2FG	2FGM	2FGA	3FG	3PM	3PA
Boston_Celtics	74	39.1	89.3	70.7 %	30.4	63.1	29.3 %	8.7	26.2
Golden_State_Warriors	74	42.9	88.2	64.6 %	30.0	57.0	35.4 %	12.9	31.2
Chicago_Bulls	75	38.6	87.9	75.9 %	30.8	66.7	24.1 %	7.8	21.2
...	...	...	...	...	...	...	...	...	...
Minnesota_Timberwolves	74	37.4	81.0	80.1 %	32.0	64.9	19.9 %	5.4	16.1
Utah_Jazz	75	36.0	80.6	70.9 %	27.6	57.1	29.1 %	8.4	23.5

Por ejemplo si deseamos imprimir en la salida estándar sólo la primera y segunda columna:  
`$ awk '{print $1 $2}' nbateams.txt`

Team	GP
Boston_Celtics	74
Golden_State_Warriors	74
Chicago_Bulls	75
...	...
Minnesota_Timberwolves	74
Utah_Jazz	75

Aquí  $\$n$  representa el  $n$ -ésimo campo o columna de cada línea en el archivo de datos. Si se escribe  $\$0$  se hace referencia a toda la línea.

Pero esto no es todo, podemos hacer aún mucho más con **awk**. Por ejemplo, podemos escribir un guión de órdenes “nbateams.awk” que permita calcular el *Field Goal Percentage* (FGP)\*,

```
$ cat > nbateams.awk
BEGIN { teams = -1 }
{ teams++; FGM=$3; FGA=$4}
{ if (teams>0) print FGM/FGA*100 }
END { print "The number of NBA teams is "teams }
(Ctrl-d)
$ awk -f nbateams.awk nbateams.txt
43.785
48.6395
43.9135
...
46.1728
44.665
The number of NBA teams is 30
```

La cláusula BEGIN se ejecuta una vez al comienzo de la secuencia de órdenes, la cláusula principal una vez por cada línea, la cláusula que calcula FGP sólo si el número de “teams” es mayor que 0 y la cláusula de finalización una vez al final de la secuencia de órdenes.

Para más detalles y la potencia de **awk** ver las páginas del manual (**man awk**) y recordar GETA.

## 2. World power plants

Below main raw information related with the world power plants sourced from Global Energy Observatory [1] is given. The raw data can be found in the file **world-power-plants-list.csv**, where the extension **csv** implies a *comma separated values* file which allows data to be saved in a table structured format.

Source copyrights: <http://www.globalenergyobservatory.org>

Theme: Renewable & Alternative Fuels, Coal, Nuclear, Crude Oil & Refined Products, Natural Gas

Keyword: Gas, Geothermal, Hydro-Electric, Oil, Solar PV Farm, Solar Thermal, Waste, Wind

Modified: April 11 2016 4:47 AM

Publisher: Global Energy Observatory

Data Source: <https://datasource.kapsarc.org/explore/dataset/world-power-plants-list>

---

\*FGP is the percentage of field goals that a team makes. The formula to determine FGP is Field Goals Made/Field Goals Attempted.

## 2.1. Conventions in the electric power industry

In the electric power industry [2], megawatt electrical (MWe or MWe) refers by convention to the electric power produced by a generator, while megawatt thermal or thermal megawatt (MWt, MWt, or MWth, MWth) refers to thermal power produced by the plant. For example, the Embalse nuclear power plant in Argentina uses a fission reactor to generate 2109 MWt (i.e. heat), which creates steam to drive a turbine, which generates 648 MWe (i.e. electricity). Note: The International Bureau of Weights and Measures, which maintains the SI-standard, states that further information about a quantity should not be attached to the unit symbol but instead to the quantity symbol (i.e.,  $P_{\text{thermal}} = 270 \text{ W}$  rather than  $P = 270 \text{ Wth}$ ) and so these units are non-SI.

## Referencias

- [1] Global Energy Observatory, <http://www.globalenergyobservatory.org>, 8-Junio-2017.
- [2] MWe, <https://en.wikipedia.org/wiki/Watt#MWe>, 8-Junio-2017.