



**UNIVERSIDAD DE TALCA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA CIVIL EN BIOINFORMÁTICA**

**Programación Avanzada
Unidad II - Proyecto parte 2 de 2**

Objetivos

El objetivo es poder utilizar y complementar una serie de herramientas para la manipulación de información desde la terminal. Al término de este trabajo, se espera que el estudiante sea capaz de:

- Utilizar herramientas para descargar archivos remotos.
- Utilizar herramientas para procesar archivos y extraer determinada información.
- Utilizar pipeline para la comunicación entre procesos.
- Escribir scripts en bash para la automatización de tareas.

Pipeline de datos climatológicos

Una estación meteorológica automática (E.M.A.), es una herramienta por la cual se obtienen datos de los parámetros meteorológicos como temperatura, humedad, velocidad y dirección del viento, presión atmosférica, lluvia y otros, leídos por medio de sensores eléctricos. Las lecturas son acondicionadas para luego ser procesadas mediante la tecnología de microcontroladores o microprocesadores, y transmitidas a través de un sistema de comunicación (radio, satélites, teléfono, etc) en forma automática. La estación automática funciona en forma autónoma, las 24 horas, con un sistema de alimentación a través de energía solar (paneles solares), o mediante el uso de la energía eólica. En la figura 1 se muestra una EMA.

Generalmente la información capturada por estas EMA puede ser exportada a archivos utilizando algún formato para su manipulación (planilla de cálculo, CSV¹, etc.) y generación de información en otros formatos (gráficos por ejemplo).

¹<https://es.wikipedia.org/wiki/CSV>



Figura 1: EMA

Se dispone de una gran cantidad de archivos en formato CSV que han sido generados por diferentes EMA. Estos archivos están comprimidos en otro llamado data.zip estando alojado en el servidor <https://srvbioinf1.utalca.cl/~fduran/>. El formato de estos archivos es el siguiente, donde cada columna está separada por “;”:

```
Fecha;Hora;Pyranometer SP-Lite (W/m2);Wind Speed (km/h);Wind Direction;Relative Humidity (%);Temperature (C);Precipitation (mm)
2010-05-01;00:00:00;0.00;0.00;125.71;83.00;9.13;0.00;
2010-05-01;00:15:00;0.00;0.00;108.48;86.13;8.30;0.00;
2010-05-01;00:30:00;0.00;0.00;309.80;88.45;7.79;0.00;
2010-05-01;00:45:00;0.00;0.00;162.20;90.26;7.91;0.00;
2010-05-01;01:00:00;0.00;0.00;39.74;89.77;7.94;0.00;
2010-05-01;01:15:00;0.00;0.00;39.74;89.94;8.18;0.00;
2010-05-01;01:30:00;0.00;0.00;39.74;90.67;7.69;0.00;
...
```

La lista de archivos a procesar se encuentra en el archivo *lista.txt*, por ejemplo:

```
Panguilemodetalle2010-2011.csv
Panguilemodetalle2011-2012.csv
Panguilemodetalle2012-2013.csv
```

Requerimientos

Dada la información anterior se requiere crear uno o varios scripts en BASH que permitan a partir del procesamiento de una serie de archivos CSV la generación de gráficos con estadísticas promedio de temperaturas y humedad relativa del aire por cada año detectado. Cada gráfico corresponderá a un año distinto. Ver los gráficos referenciales en la figura 2.

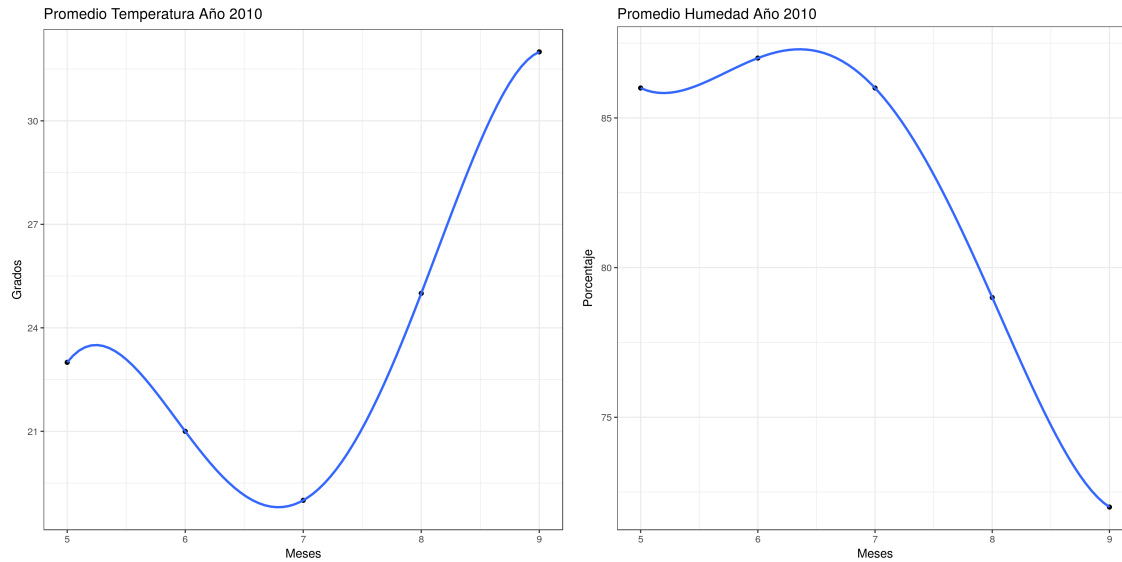


Figura 2: Temperatura y Humedad promedio por cada mes del año 2010.

En términos generales, los pasos a seguir consideran los siguientes:

1. Descargar los archivos CSV a partir de la lista contenida en el archivo *lista.txt*. Se sugiere el uso de *wget*² o *curl*³.
2. Procesar cada archivo para obtener la información que se necesite. Para esto se sugiere el uso de *cat*, *grep*, *awk*, entre otros.
3. Debe crear la siguiente estructura de directorio para almacenar información (entre otros):

```
resultados/
  2010/
    temperatura.csv
    humedad.csv
    graficoHumedad.png
    graficoTemperatura.png
  2011/
    temperatura.csv
    humedad.csv
    graficoHumedad.png
    graficoTemperatura.png
  ...
```

Los archivos CSV deben contener las siguientes columnas, en este ejemplo (*temperatura.csv*), promedios mensuales por el año correspondiente:

```
5;86
6;87
7;86
```

²<https://www.gnu.org/software/wget/>

³<https://curl.haxx.se/docs/manpage.html>

```
8;79
9;72
...
```

4. Se proveerá de un script python en conjunto a un directorio de ejemplos (script.zip) alojado en <https://srvbioinf1.utalca.cl/~fduran> para la generación de los gráficos, al cual Ud le debe pasar los siguientes parámetros:

```
Ruta donde dejar el gráfico resultante.
Nombre del archivo con el gráfico generado, sin extensión.
Nombre del archivo CSV a procesar (Humedad o Temperatura)
Texto para la etiqueta X del gráfico
Texto para la etiqueta Y del gráfico
Texto para el título del gráfico.
```

Por ejemplo:

```
$ python -Wignore generateGraphic.py . graficoHumedad humedad.csv "Meses"
"Porcentaje" "Promedio Humedad Año 2010"

$ python -Wignore generateGraphic.py . graficoTemperatura temperatura.csv "Meses"
"Grados" "Promedio Temperatura Año 2010"
```

* Puede generar los archivos temporales que necesite, recuerde que puede redireccionar la entradas y salidas de los comandos.

Entrega

Debe entregar el o los scripts en BASH necesarios para realizar la tarea. Cada script debe estar bien documentado. Además, debe adjuntar un archivo README que describa el uso del o los scripts.

El proyecto es de 2 alumnos(as). Trabajos similares o copias serán calificados con la nota mínima (1.0).