

1. Contexto

Debido al cambio climático, se prevén fuertes épocas de sequía con lo que la capacidad de los embalses es cada vez un dato más interesante, ya que puede permitirnos ver cuán rápido se va consumiendo el agua almacenada. En la página www.embalses.net hay una recopilación de datos de todos los embalses de España, incluyendo capacidad total y variación semanal (las actualizaciones se hacen semanalmente, generalmente los lunes), con lo que es un buen site para hacer esa comprobación de almacenamiento de agua

2. Definir un título para el dataset

Estado de almacenamiento de agua en los embalses en España

3. Descripción del dataset

Para cada uno de los pantanos, semanalmente se recogen los datos de almacenamiento. Tenemos la capacidad total y la cantidad actual en hm³; y la variación porcentual desde la última captura de datos. En este caso esa captura se realiza semanalmente

4. Representación gráfica



5. Contenido.

Es un data set muy sencillo; tan sólo tiene cinco columnas :

Embalse : nombre del embalse

Capacidad : capacidad total en hm³

Embalsada : dato actual en hm³

Variación : Variación respecto de la medida de la lectura anterior en %

Fecha datos : es la fecha en que se ha ejecutado la captura de datos

El intervalo de actualización de las páginas es semanal, los lunes

Dado el poco margen tan sólo he recogido los datos de la semana pasada y esta. Para hacer el ejercicio completo hay que ir ejecutando la captura de datos semanalmente

El contenido del fichero robots.txt nos da como única limitación el acceso a la carpeta contenidos, que no hemos necesitado. Por tanta, ahí no ha habido ningún problema

User-Agent: *

Disallow: /contenidos/

El tamaño de la página es de 356.000 resultados, teniendo en cuenta que incluye página de foros. Por tanto es un tamaño medio sin dificultades de eficiencia



En cuanto al dueño de la web, la respuesta al whois es

```
{
  "domain_name": [
    "EMBALES.NET",
    "embales.net"
  ],
  "registrar": "ENOM, INC.",
  "whois_server": "WHOIS.ENOM.COM",
  "referral_url": null,
  "updated_date": "2017-03-10 20:36:19",
  "creation_date": "2005-12-05 17:45:51",
  "expiration_date": [
    "2022-12-05 17:45:51",
    "2022-12-05 17:45:00"
  ],
  "name_servers": [
    "DNS1.NAME-SERVICES.COM",
    "DNS2.NAME-SERVICES.COM",
    "DNS3.NAME-SERVICES.COM",
    "DNS4.NAME-SERVICES.COM",
```

Pare poder realizar el código python, he tenido que estudiar la página. Para ver la tabla donde los datos están guardados, he cogido el dato de un pantano cualquiera, y con el botón derecho le doy a inspeccionar en el menú que sale, y vemos que los datos están recogidos en una tabla incluido en la clase “ResultadoCampo”, de forma que esa es la tabla que tenemos que buscar en la web

Para saber qué páginas hemos de rastrear, así como el tamaño de la página comprobamos cuales son las páginas que cuelgan de la principal con un

```
for link in soup.find_all('a'):  
    print(link.get('href'))
```

Siendo **soup** el contenido de la página principal www.embalses.net, y nos da el siguiente resultado :

```
/  
#  
https://www.embalses.net  
https://www.embalses.net/cuencas.php  
https://www.embalses.net/comunidades.php  
https://www.embalses.net/provincias.php  
https://www.embalses.net/pluviometros-5.html  
https://www.embalses.net/cuenca-17-cantabrico-occidental.html  
https://www.embalses.net/cuenca-16-cantabrico-oriental.html  
https://www.embalses.net/cuenca-11-cataluna-interna.html  
https://www.embalses.net/cuenca-2-duero.html  
https://www.embalses.net/cuenca-5-ebro.html  
https://www.embalses.net/cuenca-10-galicia-costa.html  
https://www.embalses.net/cuenca-19-guadalete-barbate.html  
https://www.embalses.net/cuenca-4-guadalquivir.html  
https://www.embalses.net/cuenca-6-guadiana.html  
https://www.embalses.net/cuenca-7-jucar.html  
https://www.embalses.net/cuenca-9-med-andaluza.html  
https://www.embalses.net/cuenca-8-mino-sil.html  
https://www.embalses.net/cuenca-12-pais-vasco-interna.html  
https://www.embalses.net/cuenca-1-segura.html  
https://www.embalses.net/cuenca-3-tajo.html  
https://www.embalses.net/cuenca-18-tinto-odiel-y-piedras.html  
https://www.embalses.net/comunidad-1-andalucia.html  
https://www.embalses.net/comunidad-2-aragon.html  
https://www.embalses.net/comunidad-3-asturias.html  
https://www.embalses.net/comunidad-6-canarias.html  
https://www.embalses.net/comunidad-7-cantabria.html  
https://www.embalses.net/comunidad-9-castilla-y-leon.html  
https://www.embalses.net/comunidad-8-castilla-la-mancha.html  
https://www.embalses.net/comunidad-10-cataluna.html  
https://www.embalses.net/comunidad-13-comunidad-de-madrid.html  
https://www.embalses.net/comunidad-17-comunidad-valenciana.html  
https://www.embalses.net/comunidad-11-extremadura.html  
https://www.embalses.net/comunidad-12-galicia.html  
https://www.embalses.net/comunidad-4-islas-baleares.html  
https://www.embalses.net/comunidad-16-la-rioja.html  
https://www.embalses.net/comunidad-15-navarra.html
```

<https://www.embalses.net/comunidad-5-pais-vasco.html>
<https://www.embalses.net/comunidad-14-region-de-murcia.html>
None
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-44-guadalajara.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-40-lugo.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-2-salamanca.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-26-huesca.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-36-oreense.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-24-navarra.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-27-tarragona.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-23-teruel.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-41-pontevedra.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-15-leon.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-4-caceres.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-2-salamanca.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-29-cantabria.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-4-caceres.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-4-caceres.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-26-huesca.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-9-cadiz.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-36-oreense.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-20-lleida.html>
<https://www.embalses.net/pantano--.html>
<https://www.embalses.net/provincia-9-cadiz.html>
None
</pluviometros-0.html>
</pluviometros-1.html>
</pluviometros-2.html>
</pluviometros-3.html>

/pluviometros-4.html
/pluviometros-5.html
<https://www.embalses.net/estacion-meteorologica-740.html>
<https://www.embalses.net/provincia-22-alava.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-857.html>
<https://www.embalses.net/provincia-29-cantabria.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-277.html>
<https://www.embalses.net/provincia-38-vizcaya.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-741.html>
<https://www.embalses.net/provincia-39-guipuzcoa.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-760.html>
<https://www.embalses.net/provincia-29-cantabria.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-481.html>
<https://www.embalses.net/provincia-24-navarra.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-457.html>
<https://www.embalses.net/provincia-22-alava.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-732.html>
<https://www.embalses.net/provincia-39-guipuzcoa.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-729.html>
<https://www.embalses.net/provincia-24-navarra.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-762.html>
<https://www.embalses.net/provincia-29-cantabria.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-471.html>
<https://www.embalses.net/provincia-26-huesca.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-276.html>
<https://www.embalses.net/provincia-39-guipuzcoa.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-469.html>
<https://www.embalses.net/provincia-26-huesca.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-470.html>
<https://www.embalses.net/provincia-26-huesca.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-476.html>
<https://www.embalses.net/provincia-24-navarra.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-474.html>
<https://www.embalses.net/provincia-24-navarra.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-730.html>
<https://www.embalses.net/provincia-24-navarra.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-742.html>
<https://www.embalses.net/provincia-39-guipuzcoa.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-764.html>
<https://www.embalses.net/provincia-29-cantabria.html#pluviometros>
<https://www.embalses.net/estacion-meteorologica-748.html>
<https://www.embalses.net/provincia-22-alava.html#pluviometros>
/suscripciones/boletin-web.php
<https://www.embalses.net>
<https://www.embalses.net>
[whatsapp://send?text=https%3A%2F%2Fwww.embalses.net%2F](https://www.embalses.net)
<https://www.facebook.com/Embalsesnet-105019187595640/>
https://twitter.com/Embalses_net

/contactar.php

Por tanto vemos que al contenido se puede llegar desde diferentes sitios (Comunidad, Cuenca, Provincia). Para este estudio hemos elegido hacerlo por comunidad. En cualquier caso, tenemos siempre los mismos datos. En caso de querer variarlo, habría que cambiar la condición de filtro en la llamada a la función cap_Embalses

6. Agradecimientos

Al Ministerio para la Transición Ecológica, por hacer la recopilación de todos los datos

7. Inspiración

Tal y como he comentado al principio, debido a las sequías que según las teorías del cambio climático pronostican, es muy importante comprobar la profundidad y la velocidad de la variación de estos recursos de los embalses para poder tomar acciones paliativas lo antes posible.

8. Licencia

En este caso la licencia escogida sería Released Under CC0: Public Domain License, ya que los datos son de dominio público y distribuidos por el Ministerio para la Transición Ecológica, por tanto la persona que ha asociado una obra a este documento ha dedicado la misma al dominio público, liberándola de forma mundial y en la medida que lo permita la ley, de todos sus derechos de propiedad intelectual, incluyendo todos los derechos conexos.

Puede copiar, modificar, distribuir la obra y hacer comunicación pública, incluso para fines comerciales, sin pedir permiso.

COMPONENTES DEL GRUPO

Esta práctica ha sido realizada íntegramente por Jaime de la Torre