

### Práctica. Web Scraping

Esta práctica puede hacerse individualmente.

La práctica se podrá subir al campus virtual antes del 30 de noviembre. Se subirá un fichero comprimido llamado Practica1\_NombreApellidos con una memoria en formato pptx y fichero script, así como cualquier elemento externo como librerías utilizadas.

La memoria debe reflejar como mínimo el problema planteado y la solución propuesta.

El objetivo es el siguiente:

- a) Utilizar al menos las librerías BeautifulSoup y Selenium mediante el lenguaje Python para navegar por una página web, extraer la información elegida y depositarla en un medio de almacenamiento como puede ser un fichero .txt, Excel o base de datos (MySQL o MongoDB).
- b) Además, se realizará una escucha activa en la red social Twitter, investigando las opiniones de un contenido de actualidad, toda la información recogida se procesará con la biblioteca Pandas y se mostrará un cuadro de mando en el que se recoja el número de tweets en una escala temporal que se podrá seleccionar.

Para ello:

- El alumno elegirá libremente la página web / contenido de actualidad a ser tratado.
- Se extraerá la información.
- Se tratará la información para guardar la que queremos.
- Guardamos la información.
- Realizaremos un cuadro de mando.

Se Valora:

- Utilizar Selenium para:
  - Realizar click en algún botón.
  - Escribir texto en alguna caja de texto.
  - Seleccionar opciones de menús desplegables.
  - Seleccionar radiobutton.
  - Otras acciones mediante esta librería.
- Utilizar BeautifulSoup para:
  - Parsear la información.
  - Formatear la información deseada.
- Guardar la información.
- Documentación.
- Utilizar Tweepy para:
  - Escucha en Twitter.
  - Guardado de la información.
- Utilizar Pandas para:
  - Realizar un cuadro de mando.