Proyecto integrador

Una biblioteca de tu vecindario tiene la intención de informar a sus clientes acerca de las novedades en publicaciones para poder atraerlos según su interés. Actualmente este proceso lo hace una persona de manera manual, buscando información en la página de internet de la editorial con la que más trabaja esta librería, que es Cúspide (www.cuspide.com), y donde se publica, por ejemplo, cuáles son los 100 más vendidos en la semana o cuáles son las novedades.

Como científico de datos, la biblioteca te consulta si sería posible mejorar ese proceso, teniendo la posibilidad por ejemplo, extraer, procesar y guardar esos datos con Python..

Propuesta técnica:

Realizar un Scraping de la página web de cúspide libros, www.cuspide.com, y de allí obtener el listado de los 100 libros más vendidos de la semana. Los datos deben ser: título del libro, url del libro, precio en pesos, precio en dólares, y precio en dólares considerando la cotización del dólar blue de Argentina o en Euros, y fecha. La cotización del dólar blue de Argentina o Euros debe ser obtenida también utilizando scraping de cualquier página web a elección.

Estos datos deben ser reflejados en un archivo .csv con los respectivos campos. Asimismo debe crearse otro archivo de salida como log de errores, por si algún título no se puede scrapear por alguna razón, con la información del título del libro en cuestión, y la url del mismo.

El csv debería quedar aproximadamente así:

