

## **Procesamiento de información numérica recibida en PDF, por Emiliano Rodríguez**

Todos los meses recibía información en pdf que tenía que procesar para verificar la misma. Había probado convertirla con el sitio [ilovepdf.com](https://ilovepdf.com) pero no la convertía bien porque no se había generado a partir de un excel sino que era una salida de un sistema anticuado.

Esta tarea era sumamente manual y consistía en tener en parte del monitor el PDF con la información que me devolvía el organismo y compararla visualmente con la que yo había enviado.

Por ese motivo, busqué alguna librería que me permitiera “leer” cada renglón con Python y almacenar correctamente los datos en formato numérico, para luego poder comparar ya normalmente la información.

Fue bastante complicado y en este recorrido aprendí a usar expresiones regulares, para que el Python supiera como clasificar la información de cada renglón.

Luego de varios meses de haber realizado el curso, sigo usando esta herramienta y me ahorro mucho tiempo. Lo comparto porque si bien este es un uso muy específico, cualquier otra persona lo puede adaptar para leer sus pdfs y procesar la información que recibe.