

Objetivo: Aplicar el concepto de expresiones regulares.

Instrucciones:

1. Investigar sobre la librería re de Python (para expresiones regionales).
2. Elaborar los diagramas de las máquinas de estado finito para el reconocimiento de los patrones correspondientes de cada expresión regular que se vaya a usar.
3. Cargar el archivo BL-Flickr-Images-Book.csv en forma de texto (no usar una librería que cargue directamente las columnas a un dataframe).
4. Utilizar expresiones regulares para determinar los encabezados.
5. Usar expresiones regulares para determinar el contenido de cada columna (puede ser por línea, identificando patrones y combinando de forma algorítmica).
6. Con cada columna bien identificada con sus datos, transfórmelas a un dataframe con sus tipos de datos respectivos, usando las estructuras de Python.

Nota: Cargar en la tarea de Canvas: el programa en python, el documento con la descripción del juego de datos y el listado de las expresiones regulares aplicadas en la notación vista en clase con su máquina de estado finito correspondiente y un pantallazo de cada ejecución de la expresión regular, donde se evidencia que se aplicó correctamente. Asimismo, se presentará según la fecha de finalización de la tarea en Canvas y no hay prórroga.

Referencias:

- https://www.w3schools.com/python/python_regex.asp
- <https://realpython.com/regex-python/>
- <https://realpython.com/regex-python-part-2/>
- <https://www3.ntu.edu.sg/home/ehchua/programming/howto/Regexe.html>
- <https://docs.python.org/3/library/re.html>
- <https://pazciudadana.cl/wp-content/uploads/2019/01/taller-7-limpieza-y-transformacion-de-datos.pdf>
- <https://medium.com/@yashj302/text-cleaning-using-regex-python-f1dded1ac5bd>
- <https://ajaxhispano.com/ask/regex-agarrar-valores-entre-comillas-4572/>
- <https://regex101.com/>