

Práctica 1. Web Scraping.

El objetivo de esta actividad será la creación de un dataset a partir de los datos contenidos al web. Tenéis que indicar las siguientes características del dataset general:

1. Título del dataset. **Narcan**
2. Subtítulo del dataset. *"Como evitar muerte por sobredosis heroína"*
3. Contexto. Debido al aumento del consumo de heroína en USA se ha extendido el uso del Narcan con la intención de poder evitar las muertes por sobredosis. El proposito de esta base de datos es recoger información para poder comparar como de extendido esta su uso, donde podría estar disponible y la concienciación de la población en este medicamento como medida para evitar la muerte por sobredosis.
4. Contenido. El archivo contiene 1000 registros extraídos utilizando R, los campos extraídos son: texto, si el texto esta interrumpido (truncated), la fecha cuando fue creado, status, replies, favoritos, si esta retuiteado, el computo de retuits y el nombre que aparece en pantalla.
5. Inspiración. ¿Por qué es interesante este conjunto de datos? ¿Qué preguntas le gustaría responder la comunidad? El interés radica en la posibilidad de evaluar la concienciación social del medicamento Narcan para evitar las muertes por sobredosis.
6. Licencia. Seleccionad una de estas licencias y decid porqué la habéis seleccionado. Seleccionaría, **Released Under CC BY-SA 4.0 License**. Porque permite copiar y redistribuir el material en cualquier medio y formato. Permite además remezclarlo, transformarlo y crear material nuevo para cualquier proposito incluso comercial. Para lo cual habría que nombrar al autor y si el contenido se remezcla, transforma o se crea nuevo material se debe distribuir bajo esta misma licencia
7. Código: En adjunto
8. Dataset: En adjunto

Contexto

Debido al aumento del consumo de heroína en USA se ha extendido el uso del Narcan entre la población con la intención de poder evitar las muertes por sobredosis. El proposito de esta base de datos es recoger información para poder comparar como de extendido esta su uso, donde podría estar disponible y la concienciación de la población en este medicamento como medida para evitar la muerte por sobredosis.

Contenido

9. El archivo contiene 1,000 registros extraídos utilizando R, los campos extraídos son: texto, si el texto esta interrumpido (truncated), la fecha cuando fue creado, status, replies, favoritos, si esta retuiteado, el computo de retuits y el nombre que aparece en pantalla.