

Ejercicios de curación y limpieza de datos

Ejercicio 5:

- a. Carga el fichero SAFI_ejercicio5_openrefine.csv para realizar el ejercicio 5.
- b. Eliminar de la columna “months_lack_food” los corchetes, las comillas y los espacios en blanco de los valores de las filas, dejando únicamente listas de meses separadas por “;”.
- c. Crear un facet de tipo texto con los valores de la columna “months_lack_food” modificada. ¿Cuál es el mes más frecuente? ¿Y el menos frecuente? ¿Podemos corregir los valores erróneos?
- d. Realizar la misma operación que en los apartados b. y c. con las columnas “months_no_water”, “liv_owned”, “res_change”, y “no_food_mitigation”. Pista: para reutilizar un commando GREL, puedes hacer click en la pestaña “History” y usar “Reuse”. Pruébalo.

M1968 - Ciclo de vida de los datos

Máster en Data Science – Universidad de Cantabria Curso 2018-2019

Diego García Saiz