

Ejercicios de curación y limpieza de datos

Ejercicio 3:

- a. Carga el fichero SAFI_ejercicio3_openrefine.csv para realizar el ejercicio 3.
- b. Trata de crear un facet de tipo texto para el campo “start”. ¿Aparece algo? ¿Qué errores podrías corregir de esta forma y cómo?
- c. Trata ahora de crear un facet de tiempo timeline sobre el campo “start”. ¿Se puede?
- d. Transforma el tipo del campo “start” en fecha (“date”). Repite ahora la pregunta c. ¿Cuáles son las fechas más y menos frecuentes?
- e. Modifica el valor del campo “end” de la primera fila, e introduce como fecha el 31 de febrero de 2018 (no modifiques la hora). Trata de convertir ahora el campo “end” al tipo “date”. ¿Se puede? Si es así, crea un facet de tipo timeline sobre este campo, y observa el efecto de este error.
- f. Modifica el campo “interview_date” en la primera fila, e introduce un valor que no sea una fecha (por ejemplo, un string con el contenido “no soy una fecha”). Trata de convertir ahora el campo “interview_date” al tipo “date”. ¿Se puede? Si es así, crea un facet de tipo timeline sobre este campo, y observa el efecto de este error.
- g. Intenta modificar el campo “start” de la primera fila con un valor que no sea de tipo fecha, después de haber realizado el punto d. ¿Se puede? ¿Qué sucede?