

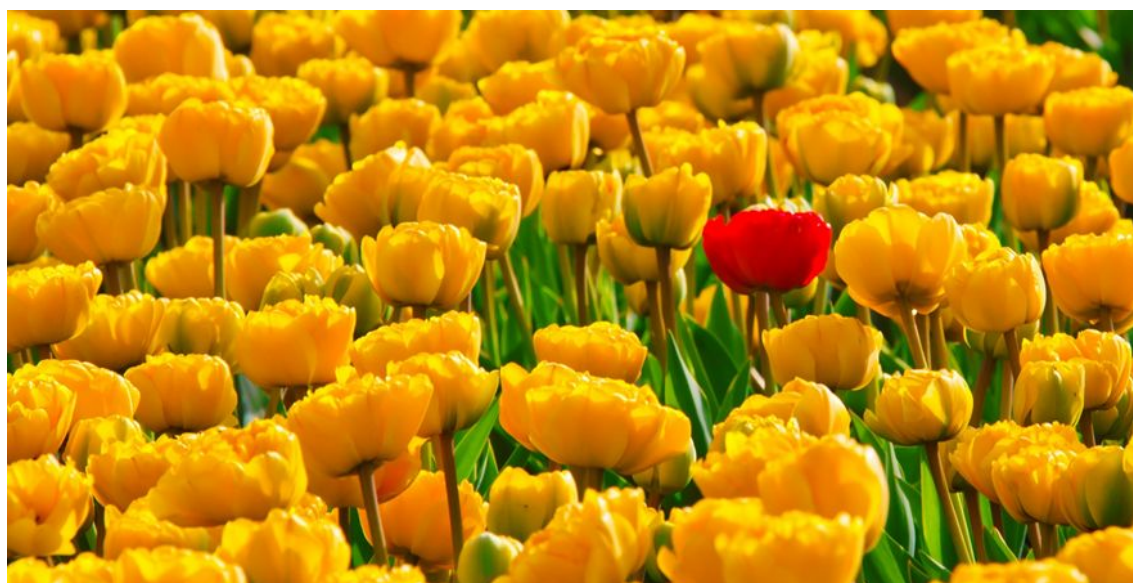
1. Título del dataset. Poned un título que sea descriptivo.

Testimonios de personas afectadas por enfermedades raras en el estado español.

2. Subtítulo del dataset. Agregad una descripción ágil de vuestro conjunto de datos por vuestro subtítulo.

Testimonios de personas afectadas por enfermedades raras en el estado español.

3. Imagen. Agregad una imagen que identifique vuestro dataset visualmente



4. Contexto. ¿Cuál es la materia del conjunto de datos?

Este conjunto de datos explica testimonios o experiencias vividas de personas afectadas por enfermedades raras. Estas personas pueden haber sufrido en primera persona enfermedades raras, o pueden ser familiares que se han visto afectados por la enfermedad. También hay casos donde son personas que han trabajado o han mejorado la calidad de vida de personas con enfermedades raras. Estos testimonios dan apoyo a personas afectadas, y también dan difusión de la problemática de las enfermedades raras.

5. Contenido. ¿Qué campos incluye? ¿Cuál es el periodo de tiempo de los datos y cómo se ha recogido?

Los campos incluyen el título y una descripción del testimonio. El periodo de tiempo es toda la historia de FEDER, y se han recogido mediante su página web <https://www.enfermedades-raras.org/index.php/testimonios>.

6. Agradecimientos. ¿Quién es propietario del conjunto de datos? Includ citas de investigación o análisis anteriores.

El propietario de los datos son los testimonios y FEDER. Les agradecemos el hecho de hacer su testimonio público.

7. Inspiración. ¿Por qué es interesante este conjunto de datos? ¿Qué preguntas le gustaría responder la comunidad?

Conocer a testimonios que están en la misma situación que tú es siempre interesante para personas con enfermedades raras para que se sienten más apoyadas. Para los gobiernos, es relevante visualizar la problemática de personas con enfermedades raras para que hagan las políticas necesarias para resolver a sus necesidades. Las enfermedades raras afectan al 7% de la población mundial según la Organización Mundial de la Salud (OMS) por lo tanto tienen un gran impacto en la sociedad.

Como preguntas le haría cuál es la polaridad del título y de los mensajes, cuál es la subjetividad del título y los mensajes (haría *sentiment analysis*), miraría qué clasificaciones de la OMS están presentes en los testimonios (ICD, ICF y ICHI), crearía una nube de palabras de los testimonios, y finalmente haría una comparativa entre diferentes testimonios indicando cuáles son las palabras más comunes y más diferentes.

8. Licencia. Seleccionad una de estas licencias y decid porqué la habéis seleccionado:

- Released Under CC0: Public Domain License
- Released Under CC BY-NC-SA 4.0 License
- Released Under CC BY-SA 4.0 License
- Database released under Open Database License, individual contents under Database Contents License
- Other (specified above)
- Unknown License

La licencia escogida para la publicación de este conjunto de datos ha sido CC BY-SA 4.0 License. Los motivos que han llevado a la elección de esta licencia tienen que ver con la idoneidad de las cláusulas que esta presenta en relación con el trabajo realizado:

- Se debe proveer el nombre del creador del conjunto de datos generado, indicando los cambios que se han realizado. De esta manera, se reconoce el trabajo ajeno y en qué medida se han realizado aportaciones en relación con el trabajo original.
- Se permite un uso comercial. Esto haría que incrementen las probabilidades de que una empresa utilice los datos generados y realicen trabajos de calidad que reporten cierto reconocimiento al autor original.
- Las contribuciones realizadas a posteriori sobre el trabajo publicado bajo esta licencia deberán distribuirse bajo la misma. Esto hace que el trabajo del autor original continúe distribuyéndose bajo los términos que él mismo planteó.

9. Código: Hay que adjuntar el código con el que habéis generado el dataset, preferiblemente con R o Python, que os ha ayudado a generar el dataset

Se adjunta el código en Python.

10. Dataset: Dataset en formato CSV

Se adjuntan los datos en formato CSV. Los campos se separan con ‘;’.