

PRAC 1 – Web Scrapping

1. Título del dataset. Poned un título que sea descriptivo.
Pronóstico del tiempo de Donostia – San Sebastián.
2. Subtítulo del dataset. Agregad una descripción ágil de vuestro conjunto de datos por vuestro subtítulo.
Extracción de información meteorológica en Donostia – San Sebastián para los próximos 7 días.
3. Imagen. Agregad una imagen que identifique vuestro dataset visualmente.

Pronóstico 7 días

Jueves, 1	Viernes, 2	Sábado, 3	Domingo, 4	Lunes, 5	Martes, 6	Miércoles, 7
 15 11 °C Probabilidad de lluvia	 16 8 °C Nublado	 13 8 °C Parcialmente nublado	 19 11 °C Despejado	 14 11 °C Lluvia	 17 11 °C Probabilidad de lluvia	 17 12 °C Parcialmente nublado
60% Posibilidades de precipitación	20% Posibilidades de precipitación	10% Posibilidades de precipitación	0% Posibilidades de precipitación	90% Posibilidades de precipitación	20% Posibilidades de precipitación	10% Posibilidades de precipitación

4. Contexto. ¿Cuál es la materia del conjunto de datos?
El dataset extraído contiene información meteorológica de los próximos 7 días en Donostia – San Sebastián. Concretamente, para cada uno de los días, se extrae la siguiente información:
 - Día de la semana y número del mes
 - Temperatura máxima y mínima
 - Condición meteorológica
 - Porcentaje de lluvia (en %)
 - Probabilidad de lluvia (en texto)
 - Imagen asociada al tiempo
5. Contenido. ¿Qué campos incluye? ¿Cuál es el periodo de tiempo de los datos y cómo se ha recogido?
La información se ha extraído en un archivo de tipo .csv llamado "tiempo.csv" y contiene 6 campos, uno para cada uno de los datos meteorológicos extraídos y un campo donde se indica donde se ha almacenado la imagen asociada al tiempo de cada día. Estos campos son los siguientes:
 - "dia": día de la semana y el número del mes
 - "temp_max_min": temperatura máxima y mínima
 - "condicion": condición meteorológica
 - "porcentaje_lluvia": probabilidad de lluvia (en %)
 - "probabilidad_lluvia": probabilidad de lluvia (en texto)
 - "url_imagen": directorio donde se ha almacenado la imagen asociada

Como se ha comentado anteriormente, los datos recogidos son un pronóstico de los próximos 7 días. La página web no dispone de información sobre cómo o de dónde se han recogido estos datos.

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

El propietario del conjunto de datos es el Ayuntamiento de Donostia / San Sebastián (<http://www.donostia.eus>).

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

Este conjunto de datos puede resultar muy interesante en muchos ámbitos distintos en los que la previsión del tiempo es fundamental. Un ejemplo de aplicación podría ser utilizar esta información para mostrarla en páginas sobre turismo en San Sebastián y que los visitantes puedan organizar su estancia teniendo en cuenta el tiempo de los próximos días. También se podría utilizar con el mismo objetivo en páginas sobre eventos al aire libre, en el cual el tiempo juega un papel muy importante.

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 página web del Ayuntamiento de Donostia / San Sebastián dispone de una licencia de tipo Released Under CC BY-SA 4.0, tal y como se especifica en el apartado de "Licencia" de la página web [1]. A este tipo de licencia se le denomina Atribución-CompartirIgual y tal y como se especifica en la página de Wikipedia sobre las licencias CC [2], consiste principalmente en que:

- El beneficiario de la licencia tiene el derecho de copiar, distribuir, exhibir y representar la obra y hacer obras derivadas siempre y cuando reconozca y cite la obra de la forma especificada por el autor o el licenciente.
- El beneficiario de la licencia tiene el derecho de distribuir obras derivadas bajo una licencia idéntica a la licencia que regula la obra original.

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.

10. Dataset: Dataset en formato CSV.

Bibliografía

- [1] “Donostia.eus - Licencia.” [Online]. Available: <https://www.donostia.eus/taxo.nsf/fwLicencia?ReadForm&idioma=cas>. [Accessed: 09-Nov-2018].
- [2] “Licencias Creative Commons,” *Wikipedia, la enciclopedia libre*. 29-Sep-2018.