

SolaRadiationAemetscraper

Autores: Eduardo Rafael Martín Martín y David Ramos Díaz

Abril 2018

##1. Descripción

Los datos recogidos en esta actividad práctica es extraída de la web de la agencia estatal de meteorología <http://www.aemet.es> Los datos conforman una tabla con el contenido que se extrae mediante el empleo de web scraping.

##2. Imagen identificativa



##3. Contexto

La información contenida en el apartado de opendata de la web de aemet incluye una gran cantidad de datos de estaciones meteorológicas en España, entre estos datos se puede extraer un informe diario de radiación solar de las distintas estaciones.

##4. Contenido

El conjunto de datos consiste en registros diarios totales y con desglose horario de Radiación global, Radiación solar directa, Radiación solar indirecta, radiación ultravioleta y radiación infrarroja de 35 estaciones meteorológicas.

Incluye el nombre de la estación, índice (**Estación**)

Código de la estación, índice (**Indicativo**)

Tipo de radiación, índice (**tipo**)

Franjas horarias, múltiples índices (Por ejemplo **5 (GL)**, cantidad de radiación global registrada a las 5 de la mañana durante esa hora)

Suma total de cada tipo de radiación, múltiples índices (Por ejemplo suma (GL) suma total diaria de radiación solar directa)

Fecha de recogida de datos, índice(**Fecha**)

##5. Agradecimientos

Los datos se han extraído de la web de AEMET

##6. Inspiración

Los datos obtenidos pueden utilizarse para múltiples funciones, correctamente procesados y ordenados pueden usarse para crear un mapa solar de España, como hace la página web <http://www.adrase.com/>, con esos mismos datos.

Otra posible aplicación podría ser para el organismo del estado encargado de garantizar que no hay ninguna clase de falseamiento de datos de generación eléctrica solar, al poder comparar la radiación solar de una zona, tanto horaria como total, la autoridad encargada podrá verificar que las lecturas son correctas dentro de un margen. Esta aplicación tiene utilidad para todas las formas de generación solar (Los datos de radiación global en la generación termosolar, y los de directa y difusa en los distintos tipos de fotovoltaica)

##7. Licencia

Para este proyecto se ha escogido la licencia GNU General Public License v2.0 que permite utilizarlo tanto de forma comercial como privada.

##8. Código fuente y dataset

Tanto el Código fuente como la información obtenida del dataset están disponibles en el siguiente repositorio de GitHub: <https://github.com/Edurmm/WebScraping>

##9. Bibliografía

Masip, D. (2010). El lenguaje Python. Editorial UOC

Lawson, R. (2015). Web Scraping with Python. Packt Publishing Ltd. Chapter 2. Scraping the Data.

Simon Munzert, Christian Rubba, Peter Meißner, Dominic Nyhuis. (2015). Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. John Wiley & Sons.