

PRÁCTICA 1

WEB SCRAPING

TIPOLOGÍA Y CICLO DE VIDA

DE LOS DATOS

AULA 1



MÁSTER EN CIENCIA DE DATOS

PABLO LOMBAO VÁZQUEZ

10.Noviembre.2018

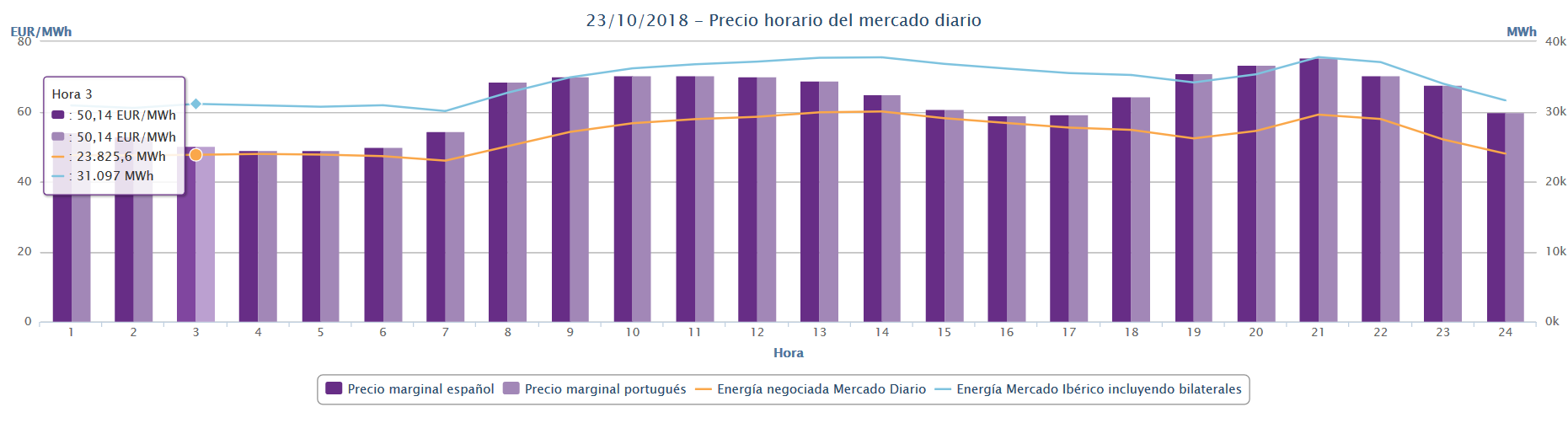


# Precios horarios de la energía en el mercado Ibérico

**Descripción:**

El dataset se trata de un conjunto de datos históricos con el precio marginal horario de la energía (en EUR/MWh) tanto en España como en Portugal, así como los datos de compra y venta de energía en el mercado diario (en MW) basados en las estimaciones de consumo realizadas por OMIE y REE.En este caso se ha extraído para un periodo que abarca desde Junio de 2018 hasta Octubre del mismo año.

**Imagen:**



**Figura 1.** Precio horario del mercado diario para el día 23/10/2018

**Contexto:**

Los datos se encuentran en el contexto del mercado energético para pequeños consumidores, aunque solo las tarifas de más de 10kWh pueden acceder a estos precios, se trata de los precios públicos y reales que emite la entidad reguladora OMIE para pequeños consumidores. Además, incluye la compra y venta de energía llevada a cabo por esta entidad.

**Contenido:**

Los campos incluidos en el dataset son los siguientes:

* Datetime: fecha y hora para la que se hace efectivo el valor
* Precio marginal PDBC de la energía horaria para España
* Precio marginal PDBC de la energía horaria para Portugal (suele coincidir)
* Energía total en MWh negociada en España
* Energía total en MWh vendida en España
* Energía total en MWh negociada en Portugal
* Energía total en MWh vendida en Portugal
* Energía total comparada en MWh para el Mercado Ibérico tras casaciones
* Energía total vendida en MWh para el Mercado Ibérico tras casaciones

El dataset se compone de 3408 instancias de estos 8 atributos que abarcan los precios horarios (24 por día) históricos desde el 1 de junio de 2018 hasta el 20 de octubre del mismo año (fecha última de ejecución).

**Agradecimientos:**

Los datos han sido recolectados desde la página del Operador del Mercado Ibérico de Energía (OMIE): hhtps://www.omie.es.

**Inspiración:**

Entre otras cosas, con estos datos podemos:

* Estudiar las variaciones en el precio de la energía entre estaciones, meses, días de la semana, horas del día, etc.
* Clusterizar zonas horarias para estudiar la diferencia y margen de las horas pico, valle y llano estipuladas por las tarifas más básicas.
* Realizar predicciones a futuro sobre los datos históricos de compra y venta de MWh en los mercados español y portugués.
* Estudiar la diferencia entre precio pagado por el pequeño consumidor medio y precio *real* de la compra de energía por parte de OMIE.
* Realizar Test A/B sobre datos en ES y PT.

**Licencia:**

El código y los datos se publicarían bajo una **licencia CC BY-NC-SA 4.0**, es decir, licencia Creative Commons Non-Commercial y Share-Alike. Bajo esta licencia tenemos las siguientes condiciones:

1. El material se puede copiar, redistribuir, adaptar y modificar en cualquier medio o formato, fomentando la colaboración y la filosofía open-source.
2. Se incluye el término de la *atribución,* para que alreplicar o usar el contenido se deba citar a la fuente original.
3. El término Share Alike de la licencia nos garantiza que cualquier modificación o uso de este material se publique bajo la misma licencia, en aras de promover la colaboración.
4. El término Non Commercial determina que el material no puede ser usado para fines comerciales y por tanto su ámbito se reduce puramente al académico.