

**Nombre:** Diego Pastor Bonet

**URL:** <https://js-ozpdsy.stackblitz.io>

**GitHub:** <https://github.com/diegopastorbonet/08MBID---Visualizacion-de-Datos>

# Trabajo Práctico

## 1. Objetivo

El objetivo general de este trabajo es avanzar en la comprensión del dataset utilizado en las asignaturas de Minería de Datos y Estadística Avanzada mediante la visualización de gráficas. Los datos originales provienen de la red eléctrica de España, estos datos son de dominio público y están a disposición de cualquiera.

A continuación, se tiene un enlace al portal ENTSO-E Transparency platform, donde se pueden encontrar y descargar los ficheros CSV utilizados.

<https://transparency.entsoe.eu/dashboard/show>

Cabe destacar que los datos han sido limpiados y tratados en la asignatura de Minería de datos, por lo que no se corresponden exactamente con los originales encontrados en el enlace. Los datos consisten en registros a cada hora de datos de precio y generación de energía en la ciudad de Valencia en el período 01/01/2015 – 31/12/2018, además de datos atmosféricos correspondientes al mismo período de tiempo.

Se escogió este dataset en un principio con la intención de descubrir alguna relación entre la condición atmosférica y el precio de la energía en cada momento, de manera que se pueda extraer alguna información útil al respecto. Es por ello por lo que el objetivo principal de este trabajo práctico es el análisis del dataset mediante visualizaciones gráficas con el fin de deducir información que nos permita ahondar en la comprensión del precio de la energía en España y cómo afectan las condiciones atmosféricas en sus variaciones.

## 2. Audiencia

La audiencia de este trabajo es cualquier persona interesada en el funcionamiento del sistema de precios de la energía de España.

El trabajo va especialmente dirigido a aquellos con interés en las características que hagan que el precio suba o baje para poder aprovecharse del sistema de precios.

No es necesario ningún tipo de conocimiento previo para leer este trabajo.

## 3. Cuadro de Mando

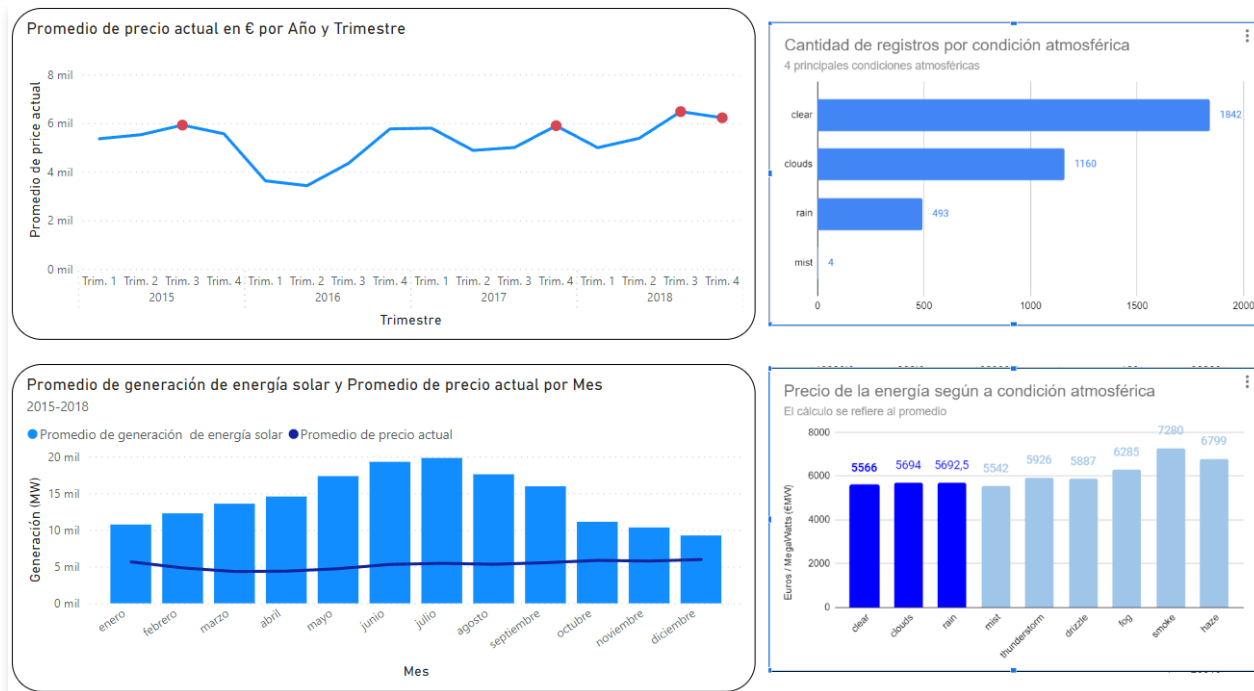


Figura 1: Cuadro de Mando completo.

En el cuadro de mando tenemos cuatro figuras, cada una con unas características determinadas. Destaca la combinación de colores azul en varios tonos y rojo, este último se ha utilizado para resaltar la información relevante y dirigir la atención de la audiencia. Tenemos a la izquierda dos gráficos correspondientes a series temporales, con una forma más alargada para poder leer adecuadamente el eje horizontal. Por otra parte, en el lado derecho tenemos dos gráficos con variables categóricas, con una forma más cuadrada. En el siguiente apartado se desgrena detalladamente cada gráfico por separado, haciendo hincapié en el motivo de su selección y en los elementos que contiene.

## 4. Visualizaciones

### 4.1 Power BI

Para las primeras dos visualizaciones se ha escogido la herramienta Power BI de Microsoft. Se ha escogido este software por su facilidad a la hora de mostrar datos según una jerarquía de fechas marcada, además de por sus capacidades al personalizar las visualizaciones resaltando la información relevante.

Como punto adicional, en Power BI es posible interactuar con los elementos gráficos, por lo que resulta muy útil en el caso que estamos tratando. AL final de esta parte se contará con la posibilidad de acceder a los gráficos creados en Power BI a través de un enlace para interactuar con ellos.

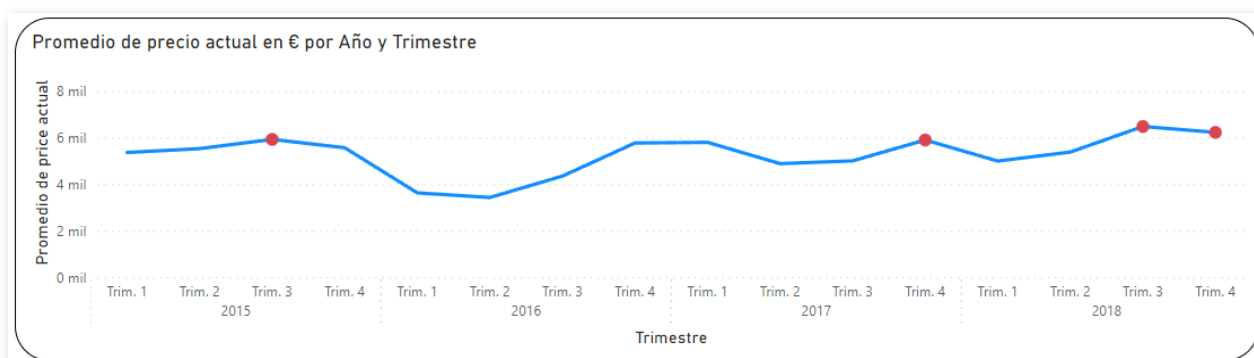


Figura 2: Precio de la energía vs Tiempo.

Este primer gráfico responde a la pregunta más fundamental en este tema: ¿Cómo varía el precio de la

energía con el tiempo? Se trata de un gráfico de líneas en el que se muestra el valor promedio del precio de la energía en España en cada trimestre transcurrido.

Se ha escogido un gráfico de líneas para mostrar la evolución del precio de la energía en el tiempo y comparar los valores más altos con los más bajos. Se destacan en marcadores rojos los 4 valores más elevados, que suelen caer en el tercer o cuarto trimestre del año. Estos marcadores son redondos y destacan en el gráfico por lo que un espectador con daltonismo no tendría problema en identificarlos. La forma se muestra de forma alargada para poder mostrar adecuadamente toda la información del eje horizontal.

De esta visualización se puede extraer por lo tanto que en la segunda mitad del año los precios son en general más elevados, siendo bastante homogéneos en general a lo largo de los años.

Cabe destacar la jerarquía desglosada del tiempo en el eje X, donde se puede apreciar el valor asociado a cada trimestre con su año correspondiente de una forma muy sencilla, característica propia de Power BI que sale a relucir en visualizaciones temporales como esta.

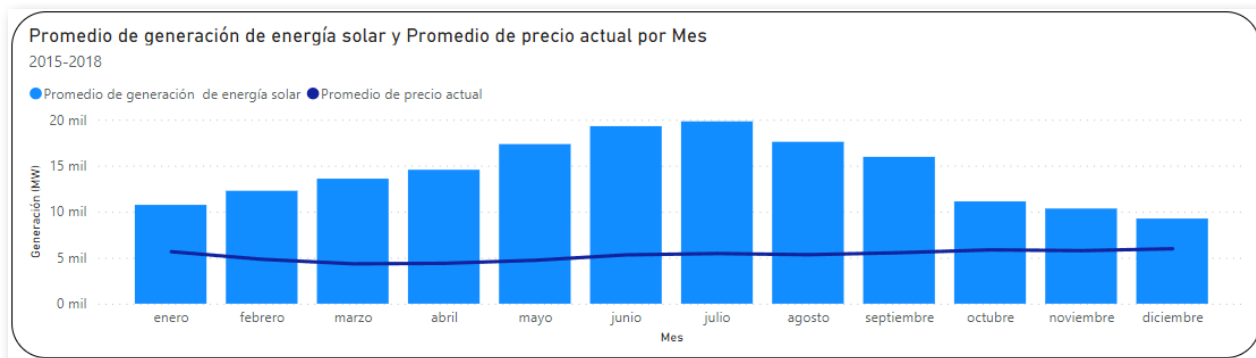


Figura 3: Generación de energía solar y precio de la energía vs Mes.

En este gráfico se responde a una nueva pregunta: ¿Cómo varía el precio de la energía en función de la generación de energía solar? Se trata de un gráfico de barras y líneas combinadas en el que se muestra el valor promedio del precio de la energía en España (línea) junto con la generación de energía solar por mes (barras). Se ha escogido mostrar las dos visualizaciones en el mismo eje puesto que no se presentan dificultades con la escala de los dos atributos del dataset, estando ambos en el rango de los miles.

Se ha escogido este tipo de gráfico para mostrar la evolución del precio de la energía en el tiempo y compararlo con la generación de energía solar, con el objetivo de ver cómo afecta ésta última en el precio. De nuevo se representa de forma alargada para mostrara adecuadamente la información del eje horizontal.

Uno podría esperar que una mayor producción de energía solar conlleva una reducción del precio de la energía puesto que esta energía es más barata de producir. Es chocante por lo tanto observar que el precio es aproximadamente constante a lo largo del año.

Al estar contando con Power BI para estas dos visualizaciones, es posible interactuar con ellas y extraer información muy interesante. Mediante el siguiente enlace podemos acceder a un informe publicado con las credenciales de la Universidad Internacional de Valencia en el que podemos interactuar además con un selector de fechas para especificar los datos a obtener:

<https://app.powerbi.com/groups/me/reports/be7bac4d-9780-4ee1-99e4-743663d74673/ReportSection>

## 4.2 Google Sheets

Para las siguientes visualizaciones se ha escogido la herramienta Google Sheets. Se ha escogido este software por su simplicidad y su facilidad de uso en general.

En esta herramienta es posible editar mínimamente los gráficos, con la complicación justa y necesaria para representar los gráficos deseados y mostrar la información pertinente con muy poca dificultad y en muy poco tiempo.

Además, al compartir las gráficas desde Google Sheets es posible interactuar con ellas desde la visualización web, mostrándose los valores de las columnas al pasar el cursor.

## Cantidad de registros por condición atmosférica

4 principales condiciones atmosféricas

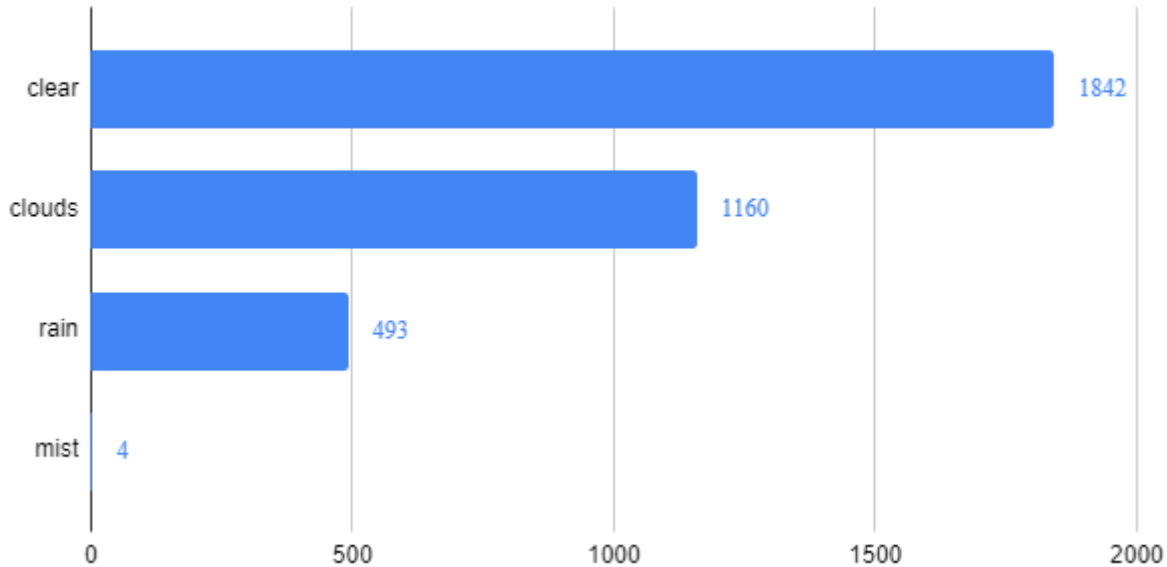


Figura 4: Cantidad de registros por condición atmosférica.

En este gráfico tenemos la distribución de registros por condición atmosférica. Se trata de un gráfico de barras sencillo en el que se ordena cada condición por número de registros para hacer comparaciones. Se muestran datos de las 4 principales condiciones atmosféricas por número de registros.

Se ha escogido un gráfico de barras horizontales por su facilidad en el momento de escribir los valores de las categorías en el eje vertical sin importar la longitud de las palabras.

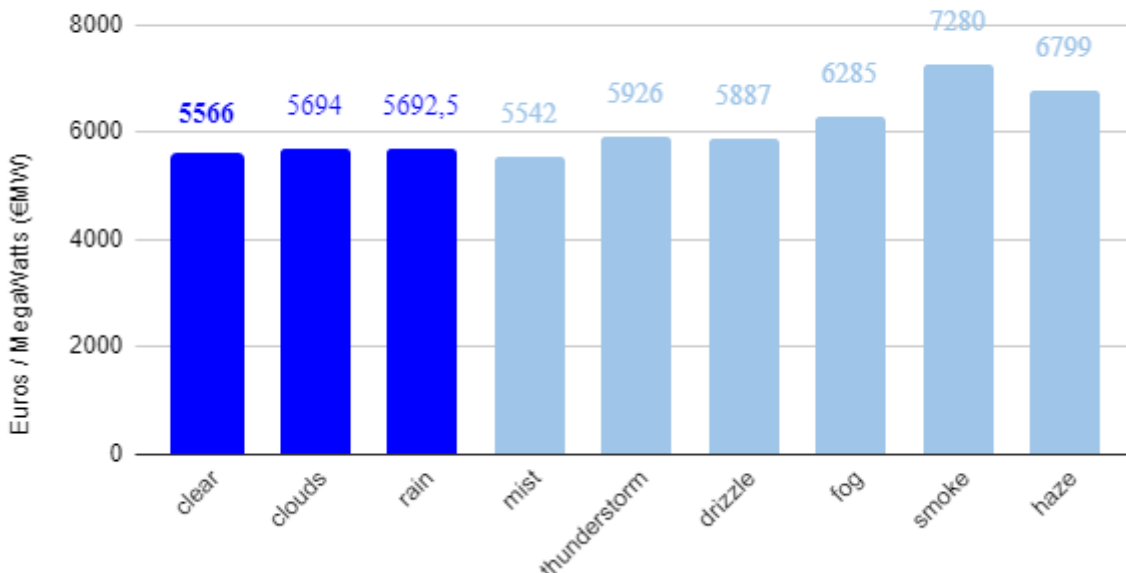
De esta visualización se puede extraer por lo tanto que la mayor parte del año está soleado, nublado o, con menor frecuencia, lluvioso. Las condiciones más extrañas como tormentas son tan poco frecuentes que no se aprecian en el gráfico, se ha mostrado el caso de bruma (mist) para hacer la comparación, pues el resto de las condiciones son incluso menos frecuentes.

Ha sido de gran utilidad la capacidad de Google Sheets de mostrar el número de registros por categoría de forma automática.

Se han colocado las etiquetas de datos para aportar los datos numéricos exactos de cada categoría.

## Precio de la energía según a condición atmosférica

El cálculo se refiere al promedio



### *Figura 5: Precio promedio por condición atmosférica.*

En este gráfico tenemos el promedio del precio del MW por condición atmosférica. Se trata de un gráfico de columnas sencillo.

Se ha escogido un gráfico de columnas por su facilidad en el momento de comparar valores absolutos. Se han destacado en un tono más oscuro los tres elementos más importantes por cantidad de registros, vistos en la visualización anterior. Esto se ha hecho con la intención de poner impedimentos a espectadores con daltonismo.

Cabe señalar que las categorías señaladas, correspondientes a la gran mayoría de registros, son además los que muestran un precio más bajo, junto con bruma (mist). Es importante resaltar que estos valores son muy uniformes, el precio promedio es muy similar en los tres, pudiendo ser la causa de la homogeneidad del precio a lo largo del año que vimos en la segunda visualización.

Ha sido de gran utilidad la capacidad de Google Sheets de mostrar el promedio de cada categoría de manera muy sencilla.

De nuevo se han colocado las etiquetas de los datos para conocer el valor exacto de cada categoría.

## **5. Mensaje y conclusión**

Las cuatro visualizaciones anteriores nos han contado una historia: El precio de la energía en España se mantuvo relativamente constante a lo largo del período 2015 – 2018. El precio no se veía afectado por las variaciones de producción de energía solar, muy fluctuante a lo largo del año. Finalmente, la inmensa mayor parte del año estamos bajo unas condiciones climáticas que se corresponden con precios muy uniformes, que acompañan a la poca variación de los precios.

Queda patente que la fuente de la producción de energía no es el único factor determinante en el precio final de la energía y tampoco lo es el clima. Es de suponer que existan mecanismos muy complejos que regulen este precio y lo mantengan relativamente constante con independencia del momento del año y de la condición climática.

Estas visualizaciones han desempeñado un papel fundamental en la comprensión del dataset utilizado. Es de gran utilidad obtener conclusiones de un vistazo a estos gráficos.