

# ICD2 - Análisis de Sesiones de Estaciones de Carga

## Grupo 3

Jesús Cumplido Almenara  
Ismael Alonso Gomez Calero

20 de marzo de 2022

## 1 Introducción

Nuestro proyecto va a consistir en realizar un análisis de ciencia de datos para una empresa de compañía eléctrica o de vehículos eléctricos. Este análisis consiste específicamente en un estudio sobre el consumo energético en las sesiones de carga realizadas por los usuarios en los puntos de cargas (públicas o domésticas) a nivel mundial.

Este estudio consistirá en obtener diferentes datasets abiertos de sesiones de carga en diferentes países a lo largo de los últimos años (sobre todo países con un gran desarrollo en puntos de carga como son Países Bajos, Reino Unido, Francia y EEUU), limpiar estas sesiones de anomalías o datos inválidos, y realizar un posterior análisis para observar el comportamiento de los datos, como pueden ser:

- Tendencia del número de sesiones de carga en los últimos años.
- Energía consumida media (según el tipo de conector).
- Duración media (según el tipo de conector).
- Frecuencia del inicio y finalización de las sesiones de carga por hora del día (histograma).
- Frecuencia de sesiones de carga por día de la semana (histograma).
- Predicción (forecasting) del número de sesiones de carga en los próximos días en una zona.
- Etc.

Este proyecto tiene el objetivo de entender el comportamiento y avance del desarrollo de los puntos de cargas, que permitirá a diferentes compañías o entidades gubernamentales tomar decisiones de mercado, como por ejemplo la necesidad de desplegar más puntos de carga en una zona debido a su elevada demanda.

## 2 Datos abiertos y modelos de datos

Para obtener una mayor calidad de los datos, se ha decidido obtener datos abiertos de sesiones de cargas de diferentes ciudades de diferentes países, sobre todo, países con gran desarrollo de los puntos de cargas. La gran mayoría de los datos que se han obtenido en este proyecto han sido a partir de las páginas oficiales de datos abiertos de diferentes ciudades. A continuación, se detalla las diferentes fuentes de datos:

- **Palo Alto (EE.UU):** aprox. 260.000 datos abiertos de sesiones de carga desde 2011 a 2020. <https://data.cityofpaloalto.org/dataviews/257812/electric-vehicle-charging-station-usage-july-2011-dec-2020/>
- **Boulder, Colorado (EE.UU):** aprox. 30.000 datos abiertos de sesiones de carga desde 2018 a 2021. [https://open-data.bouldercolorado.gov/datasets/183adc24880b41c4be9fd6a14eb6165f\\_0/explore](https://open-data.bouldercolorado.gov/datasets/183adc24880b41c4be9fd6a14eb6165f_0/explore)
- **Países Bajos:** ejemplo de 10.000 datos abiertos de sesiones de carga desde 2019 a 2020 de diferentes ciudades de Países Bajos, bajo la red de puntos de carga de Elaadnl. <https://platform.elaad.io/download-data/>
- **París (Francia):** aprox. 6.700 datos abiertos de sesiones de carga y de estados en tiempo real de París durante abril y mayo de 2017. <https://opendata.paris.fr/>
- **Dundee (UK):** aprox. 160.000 datos abiertos de sesiones de carga de 2017 a 2018 desde la página web <https://data.dundee.gov.uk/dataset/ev-charging-data>. Por otro lado, dar agradecimientos a dundee city, ya que se ha logrado obtener más datos recientes desde 2021 a 2022 tras solicitarlos por email.
- **Perth & Kinross (UK):** aprox. 66.000 datos abiertos de sesiones de carga de 2016 a 2019. <https://data.pkc.gov.uk/dataset/ev-charging-data>

Tras recoger, todos estos datos que se encuentran en diferentes formatos, el objetivo siguiente es normalizar y limpiar los datos para construir el siguiente modelo de datos común, en formato JSON:

```

{
  "country": <string>, #pais
  "city": <string>, #ciudad
  "connectorType": <string>, #Slow/Fast/Rapid/Ultra-Rapid
  "durationCharge": <float>, #horas de recarga
  "durationSession": <float>, #horas de la sesion de carga
  "energy": <float>, #energia (kWh) consumida
  "tariff": <float>, #tarifa (euros/kWh)
  "cost": <float>, #coste total (euros)
  "meanPower": <float>, #potencia media (kW)
  "maxPower": <float>, #potencia maxima (kW)
  "start":{
    "startDate": <string>, #fecha inicio (dd/mm/yyyy hh:mm:ss)
    "startTimestamp": <integer>, #timestamp de fecha inicio
    "weekDayStart": <integer>, #dia de la semana de inicio (0 a 6)
    "yearStart": <integer>, #year de inicio
    "hourStart": <integer>, #hora de inicio
    "minuteStart": <integer>, #minuto de inicio
  }
  "end":{
    "endDate": <string>, #fecha fin (dd/mm/yyyy hh:mm:ss)
    "endTimestamp": <integer>, #timestamp de fecha fin
    "weekDayEnd": <integer>, #dia de la semana de fin (0 a 6)
    "yearEnd": <integer>, #year de fin
    "hourEnd": <integer>, #hora de fin
    "minuteEnd": <integer>, #minuto de fin
  }
}

```

Tener de manera redundante información sobre la fecha de inicio y fin, nos facilita para posteriormente realizar consultas para ver las sesiones de carga dependiendo del año, día de la semana, hora, etc.