

Minería De Datos

INFORMACIÓN SOBRE VEHÍCULOS ELÉCTRICOS

Diego Cordero Contreras

Enrique Albalade Prieto

Lucía De Ancos Villa

Pablo Del Hoyo Abad

Mohamed Essalhi Ahamyan

Índice:

1. Introducción.....	3
2. Alcance	3
3. Descripción de los datos.....	3
3.1 Análisis exploratorio de datos.....	3
3.2 Análisis detallados	5
4. Trabajos Similares	6
5. Objetivo.....	6
6. Hipótesis.....	6
7. Enriquecimiento de los datos	7
8. Metodología seguida	7
9. Anexo.....	8
9.1 Diccionario de Datos Principal.....	8
9.2 Diccionario de Datos Enriquecimiento	10
10. Participación.....	11
11. Bibliografía.....	11

1. Introducción

Los vehículos eléctricos han surgido con fuerza en los últimos años, principalmente debido a su colaboración con el medio ambiente. Estos vehículos utilizan una fuente energética que no contamina como la electricidad para cargar sus baterías, al contrario que otros coches de combustión que necesitan diésel o gasolina.

Es necesario expandir su uso a lo largo del mundo y en este trabajo queremos encontrar información sobre los diferentes motivos que pueden llevar a un comprador a decantarse por el vehículo eléctrico, esta información podría ser de utilidad para gobiernos o empresas que quieran ayudar al medio ambiente mediante la mejora de las características que buscan los conductores.

2. Alcance

Con este proyecto hemos tenido que fijar un alcance de datos, es decir, una serie de países y regiones de las que los tomamos. No podemos tomar los datos de todos los países mundialmente debido a que no todos los países hacen públicos este tipo de información.

Por este motivo, hemos votado un “alcance mínimo”, es decir, una lista de regiones y países de los que al menos debemos tener datos para considerar válido un dataset.

Estos son:

- Norte América
- Europa
- Asia central

3. Descripción de los datos

3.1 Análisis exploratorio de datos

Queremos trabajar con datos que contengan información de los siguientes campos:

- Información técnica sobre vehículos eléctricos
- Venta de vehículos eléctricos en el mercado europeo y estadounidense
- Puntos de recarga a nivel mundial instalados
- Precios de la gasolina, diésel y electricidad en Europa y USA
- Nivel económico de los mercados que estamos investigando

Todos estos datos los obtendremos en una escala temporal que abarcará desde 2015 hasta 2023, periodo que corresponde con el auge de la venta de vehículos eléctricos.

Para saber qué es lo que buscamos con estos datos, podemos realizar un análisis exploratorio de estos. Vayamos pasando por todos los tipos de datos:

Información técnica sobre vehículos eléctricos:

Preguntas clave:

- ¿Cuál es la tendencia en el número de modelos de vehículos eléctricos disponibles en el mercado?

- ¿Cuál es la capacidad de la batería promedio?
- ¿Cuál es el alcance promedio de los vehículos eléctricos y su relación con el precio de venta?
- ¿Cómo ha evolucionado la eficiencia energética a lo largo de los años?

Venta de vehículos eléctricos en el mercado europeo y estadounidense:

Preguntas clave:

- ¿Cómo ha sido la evolución de las ventas de vehículos eléctricos en Europa y Estados Unidos en los últimos años?
- ¿Cuáles son los modelos más populares en cada región?
- ¿Existen patrones estacionales en las ventas o a niveles de polución en el aire?

Puntos de recarga a nivel mundial instalados:

Preguntas clave:

- ¿Cuál es la distribución geográfica de los puntos de recarga?
- ¿Cómo ha aumentado el número de puntos de recarga en el tiempo?
- ¿Qué tipos de puntos de recarga son más comunes (rápidos, semirápidos, lentos)?
- ¿Cuál es la cobertura de puntos de recarga en áreas urbanas y rurales?

Precios de la gasolina, diésel y electricidad en Europa y USA:

Preguntas clave:

- ¿Cómo han variado los precios de la gasolina, diésel y electricidad en Europa y Estados Unidos en los últimos años?
- ¿Existe una correlación entre los precios de los combustibles fósiles y la adopción de vehículos eléctricos?
- ¿Cómo afectan los impuestos y las políticas gubernamentales a los precios?
- ¿Cuál es el costo promedio de cargar un vehículo eléctrico en comparación con llenar un tanque de combustible?

Nivel económico de los mercados que estamos investigando:

Preguntas clave:

- ¿Cuánta diferencia hay entre el SMI dentro de la comunidad europea y con respecto a USA?
- ¿Existe una correlación entre el nivel económico y la adopción de vehículos eléctricos?
- ¿Cuál es la disposición de los consumidores a invertir en tecnología de vehículos eléctricos en relación con su nivel de ingresos?
- ¿Cómo impacta la economía en las políticas de incentivos gubernamentales para vehículos eléctricos?

Solo una vez podamos responder a estas preguntas con los datos que encontremos, podremos proceder con la resolución de las hipótesis concretadas.

3.2 Análisis detallados

Ventas por modelo

Este dataset ha sido obtenido del portal especializado en automoción MarkLines. Si bien para acceder a los mismo es necesario registrarse, este proceso es totalmente gratuito. En este conjunto de datos se muestra, para cada año desde el 2017 hasta el 2022 el número de unidades vendidas de diferentes modelos de coches. El dataset es central para responder a las hipótesis planteadas puesto que es el que nos indica cuántos vehículos han sido en diferentes zonas del mundo.

Precios de la gasolina

En este dataset, podemos consultar los precios de la gasolina en una amplia lista de países. Estos precios serían el precio medio de los años 2015 al 2023 y nos ayudarán a la investigación de la hipótesis sobre si el número de coches eléctricos vendidos aumenta en una relación inversa al precio de la gasolina.

Rendimiento de los vehículos eléctricos

Este dataset corresponde a una base de datos pública en la que se encuentran todos los coches eléctricos comerciales, es decir, que se pueden comprar en 2023, con sus características técnicas.

Dentro de este dataset podremos sacar de cada modelo de coche su batería, su PVP de fábrica y muchos datos que nos ayudarán con todas las hipótesis en general.

Puntos de Carga de Vehículos Eléctricos

El conjunto de datos de los puntos de carga indica por países y por año los puntos de carga de los vehículos eléctricos en los principales países del mundo. En este archivo, principalmente podremos diferenciar entre dos tipos de puntos de carga, los de carga rápida y los de carga lenta.

Sin contar el tipo del punto de carga, el año, el país y la cantidad de puntos de carga, el resto de los parámetros son utilizados por la página web como parámetros de control sin relevancia para nuestro estudio.

PIB per cápita

Este archivo nos servirá a la hora de manejar información relacionada con el producto interior bruto y su evolución temporal en diferentes países del mundo (en concreto 196), además de ciertos territorios más amplios como Centroamérica, la zona sur de Europa o las islas del Caribe, entre muchas otras.

En sí los datos se estructuran de la siguiente manera: cada una de las filas hace referencia a un país o zona comentada anteriormente; mientras que cada una de las columnas se refiere a un año determinado desde 1980 hasta 2028, siendo los valores de las últimas 5 previsiones futuras.

4. Trabajos Similares

Existen distintos trabajos que ya han desarrollado estudios similares identificando características que llaman la atención a los compradores de coches eléctricos

- [Identifying early adopters in Germany](#)
- [Relating consumers' information and willingness to buy electric vehicles](#)

Sin embargo, estos estudios, aun siendo completos en muchos aspectos, los consideramos mejorables en muchos otros. En el primer estudio, se limita a Alemania, además de ser de 2014 y con el crecimiento de los vehículos eléctricos ha quedado obsoleto. En el segundo estudio, limitándose a la India, trata únicamente aspectos demográficos como pueden ser la edad, el sexo o el nivel de estudios.

En nuestro caso nos enfocaremos en aspectos relacionados con la economía ya que es el aspecto principal para que un consumidor pueda elegir este tipo de vehículos por encima de otros.

5. Objetivo

El objetivo principal de nuestro estudio es conocer que motivos económicos pueden llevar a un cliente a elegir un vehículo eléctrico. Como ya hemos mencionado, estos datos pueden ser interesantes tanto para gobiernos como empresas con tal de incrementar la presencia de estos vehículos.

Por tanto, debemos tener en cuenta distintos indicadores económicos para hacer lo más completo posible nuestro análisis.

6. Hipótesis

Las hipótesis que se han planteado de cara al estudio son las siguientes:

1. [Se venden más vehículos por el aumento del rendimiento de los coches eléctricos](#)

Aun sin ser un indicador económico el rendimiento de los vehículos eléctricos es interesante conocer si el rendimiento de estos vehículos merma su número de ventas.

2. [Se venden más vehículos por el aumento de los precios del carburante](#)

Tendremos en cuenta los precios de la gasolina y diésel para conocer si el abaratamiento de los costes asociados al vehículo produce un mayor número de ventas

3. [Se venden más vehículos por el aumento de los puntos de carga](#)

En este caso, tampoco es un indicador económico, sin embargo, queremos comprobar que la facilidad a la hora de recargar un vehículo eléctrico aumenta sus ventas

4. [Se venden más vehículos en países con más PIB per cápita](#)

Este es el indicador económico principal, el poder adquisitivo de los clientes hace que aumente el número de ventas y queremos comprobar si se corrobora dicha hipótesis

7. Enriquecimiento de los datos

Llegados a este punto, es importante dejar claro que, al no haber utilizado ninguno de los dataset disponibles en el campus virtual, todos los que utilizaremos más adelante fueron específicamente buscados por nosotros. De esa forma, resulta difícil diferenciar entre aquella información básica y la que enriquecerá a la primera.

No obstante, además de los datasets que son esenciales para responder a las hipótesis, tenemos en mente algunos más que nos permitirán valorar los datos anteriores. Esto se conoce como el enriquecimiento de datos, tomar distintos puntos de datos sin procesar y combinarlos con otros dando lugar a una base de datos más grande.

Uno de los datasets que hemos seleccionado contiene el precio por megavatio/hora de cada país europeo desde el año 2015 al 2023, consideramos que puede ser útil a la hora de afirmar o refutar alguna de las hipótesis propuestas

Además, hemos añadido un conjunto de datos que nos indica el nivel de estudios por países. Con este dataset podríamos probar si se puede afirmar que existe relación con el nivel de estudios.

Por último, el dataset de emisiones de CO₂, el cual podría llevarnos a ver si los países que más vehículos eléctricos compran poseen niveles de polución menores de CO₂.

8. Metodología seguida

Inicialmente, entre todos los integrantes del equipo tratamos de elegir un tema principal sobre el que centrar este proyecto. Tras discutirlo en las primeras sesiones de clase que tuvimos, comprobar que existían suficientes datasets y estar todos de acuerdo; optamos por la temática actual relacionada con vehículos eléctricos.

Una vez establecido el objetivo final y fijadas las hipótesis a demostrar, procedimos a investigar posibles fuentes desde las cuales descargar los datos, de tal forma que al poco tiempo formalizamos las direcciones a emplear.

Con dichas direcciones, comenzamos la recolección de información utilizando la técnica conocida como Web Scraping, a través de la cual ejecutando código programado (en nuestro caso, en el lenguaje Python) se automatiza esa tarea. Después, tomamos la decisión de almacenar los archivos construidos en una base de datos sql común a los 5 miembros del equipo para contar con un repositorio donde guardar el resultado de los tratamientos aplicados a los datos.

9. Anexo

9.1 Diccionario de Datos Principal

Ventas por modelo

Nombre del Campo	Tipo de Dato	Descripción
Group	String	Grupo empresarial al que pertenece el fabricante de coches
Maker/Brand	String	Fabricante del coche eléctrico
Segment	String	El segmento al que pertenece el modelo del coche
Model	String	Modelo del coche eléctrico
PowerTrain	String	Mecanismo que hace que el coche pueda moverse.
Country	String	País donde ocurrieron las ventas del coche eléctrico
Años (2017-2022)	Int	Año al que se refiere las ventas del coche eléctrico.

Precio gasolina y diésel

Nombre del Campo	Tipo de Dato	Descripción
País	String	Región de donde se han obtenido los datos
Gasolina	Float	Precio medio en euros de la gasolina contando impuestos en el año especificado
Diesel	Float	Precio medio en euros del diésel contando impuestos en el año especificado
Años (2015-2023)	Int	Indica el año en el que se tomaron los precios

Rendimiento de los Vehículos

Nombre del Campo	Tipo de Dato	Descripción
Brand	String	Marca del coche analizado
Model	String	Modelo del coche analizado
AccelSec	Float	Segundos que el coche tarda en alcanzar, desde 0, los 100 kilómetros por hora
TopSpeed_Kmh	Int	Velocidad máxima, en kmh, que alcanza el vehículo
Range_km	Int	Número máximo de km alcanzables con una batería según el ciclo WLTP
Efficiency_WhKm	Int	Eficiencia energética del vehículo, específicamente la cantidad de energía en vatios-hora (Wh) que se requiere para recorrer una distancia de un kilómetro.
FastCharge_KmH	Int	Indica la cantidad de kilómetros de autonomía que el coche eléctrico puede

		obtener por hora de carga rápida en una estación de carga de alta potencia.
RapidCharge	Boolean	Indica si un coche dispone o no de carga rápida (Como por ejemplo el tesla Supercharger)
PowerTrain	String	Indica a qué ruedas traccionan en el modelo.
PlugType	String	Tipo de enchufe que tiene el vehículo
BodyStyle	String	Tipo de carrocería de la que dispone el vehículo
Segment	String	Indica al segmento al que pertenec el vehículo
Seat	Int	Número de asientos/plazas del que dispone el vehículo
PriceEuro	Int	Precio de venta recomendado por el fabricante en euros

PIB per cápita

Nombre del Campo	Tipo de Dato	Descripción
GDP per capita, current prices (U.S. dollars per capita)	String	País o región del mundo a la que pertenecen los valores de PIB de cada fila determinada
Años (1980-2028)	Float	Producto interior bruto per capita para un año determinado (las filas marcarán la zona). Serán o bien datos históricos (desde 1980 y hasta 2023), o previsiones futuras (desde 2023 y hasta 2028)

Puntos de Carga

Nombre del Campo	Tipo de Dato	Descripción
region	String	País o zona donde se están tomando los datos
category	String	Es un parámetro de control que utiliza la página web. En nuestro caso siempre será 'Historical'
parameter	String	Es un parámetro de control que utiliza la página web. En nuestro caso siempre será 'EV charging points'
mode	String	Es un parámetro de control que utiliza la página web. En nuestro caso siempre será 'EV'
powertrain	String	Nos indica el tipo de cargador. Puede tomar dos valores: 'Publicly available fast' para un cargador rápido o 'Publicly available slow' para un cargador lento
year	Int	Nos mostrará la información del año indicado
unit	String	Es un parámetro de control que utiliza la página web. En nuestro caso siempre será 'charging points'

value	Int	Nos indica el número de puntos de carga para vehículos eléctricos
-------	-----	---

9.2 Diccionario de Datos Enriquecimiento

Precio luz

Nombre del Campo	Tipo de Dato	Descripción
País	String	País de donde se han obtenido los datos
Código	String	Precio medio en euros de la gasolina contando impuestos en el año especificado
Fecha (UTC)	String	Fecha en tiempo universal coordinado
Fecha (local)	String	Fecha local del país
Precio	Float	Precio de la electricidad medido en (euro/MWhe) en la fecha y país especificado

Nivel de Estudios

Nombre del Campo	Tipo de Dato	Descripción
Country	String	País de donde se han obtenido los datos
Tertiary Total	Int	Total de personas con estudios terciarios
Tertiary Total Men	Int	Total de hombres con estudios terciarios
Tertiary Total Women	Int	Total de mujeres con estudios terciarios
Below Upper Secondary Total	Int	Total de personas con estudios por debajo de secundaria
Below Upper Secondary Men	Int	Total de hombres con estudios por debajo de secundaria
Below Upper Secondary Women	Int	Total de mujeres con estudios por debajo de secundaria
Upper Secondary Total	Int	Total de personas con estudios por encima de secundaria
Upper Secondary Men	Int	Total de hombres con estudios por encima de secundaria
Upper Secondary Women	Int	Total de mujeres con estudios por encima de secundaria

Nivel de CO2

Nombre del Campo	Tipo de Dato	Descripción
Substance	String	Indica el tipo de sustancia que se esta investigando.
EDGAR Country Code	String	Identificador del país
Country	String	Nombre del país
Años	Int	Años desde 1970 a 2021 indicando la producción de CO2 en MT por año

10. Participación

Alumno	Participación (%)
Diego Cordero Contreras	20
Enrique Albalade Prieto	20
Lucía De Ancos Villa	20
Pablo Del Hoyo Abad	20
Mohamed Essalhi Ahamyran	20

11. Bibliografía

Dataset	Enlace
Rendimiento de los vehículos	Kaggle
Precio gasolina	Trading economics
Precio diésel	The global economy
Precio luz	Ember Climbase
Puntos de Carga	IEA
PIB per cápita	IMF
Ventas por modelo de coche	Marklines
Nivel CO2	Our world in data
Nivel de estudios	World population reviews

Github: <https://github.com/Mohamed11302/MineriaDeDatosYSistemasMultiagentes>