

# Modelo Inteligente para la Identificación y Análisis de Curvas de Carga de Vehículos Eléctricos

Mario Rabanal Pérez del Río

Becario en TotalEnergies y estudiante del grado en Ciencia e Ingeniería de Datos

Cátedra TotalEnergies de Analítica de Datos e Inteligencia Artificial



## Introducción

Se dispone de datos temporales relacionados con las sesiones de carga de vehículos eléctricos (GUPPY y TUA), pero sin identificadores específicos que permitan asociarlos directamente a un vehículo único.

El proyecto se centra en aplicar técnicas de aprendizaje automático no supervisado, como clustering, para agrupar las sesiones de carga, con el objetivo de establecer patrones relevantes y avanzar hacia la identificación de vehículos y la evaluación de la salud de sus baterías.

## Objetivos del estudio

El objetivo final de este proyecto sería poder identificar cada vehículo GUPPY por su curva de carga, es decir, que cada vez que se cargara un vehículo, la información obtenida fuera suficiente para poder identificarlos, aunque debido a la limitación de los datos, nos centraremos en si podemos encontrar diferentes grupos de vehículos que comparten características.

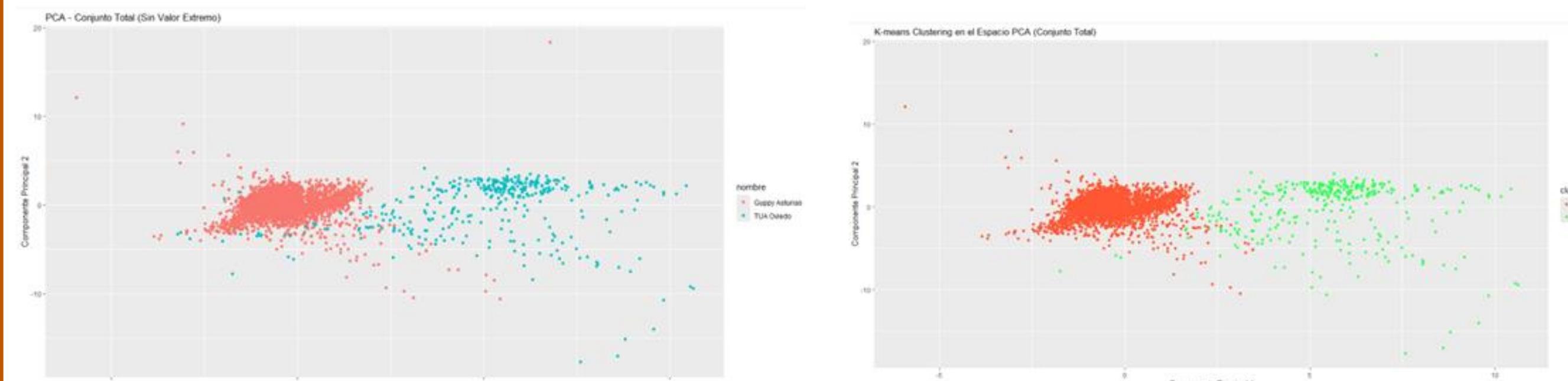
## Métodos

En el proyecto se emplearon técnicas de análisis exploratorio y reducción de dimensionalidad (PCA, T-SNE, UMAP) para visualizar y simplificar los datos relacionados con las sesiones de carga de los vehículos. Estas técnicas permitieron representar los datos en espacios de menor dimensión, facilitando su interpretación y la aplicación de algoritmos de agrupamiento.

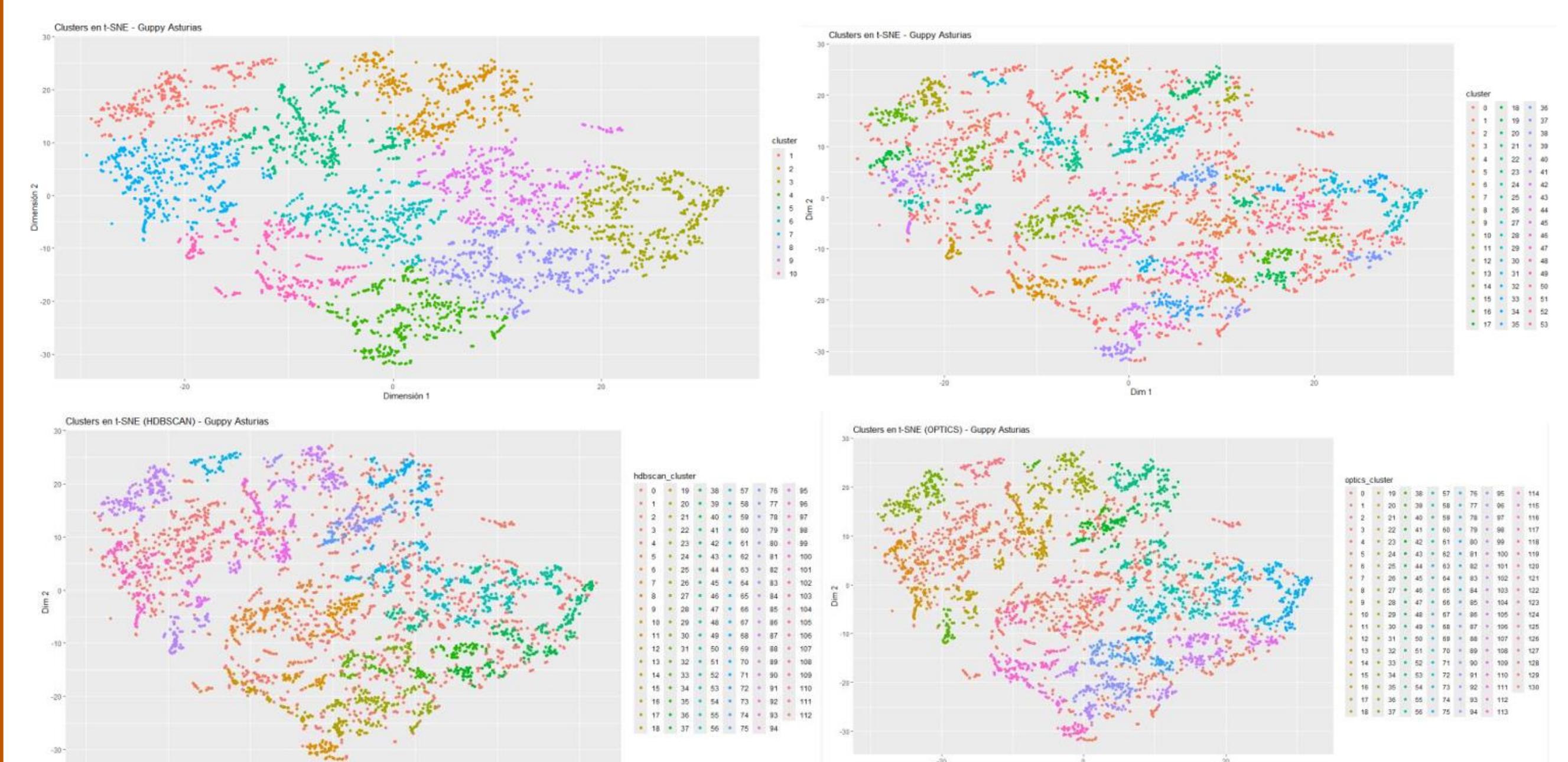
Para el clustering, se utilizaron diversos algoritmos, como k-means, DbScan, HDbscan, Optics y algoritmos aglomerativos. Estos métodos se aplicaron tanto en datos escalados como no escalados, lo que permitió evaluar la influencia de la normalización en la calidad de los grupos formados. Además, se realizaron análisis específicos en diferentes rangos de carga (0-25%, 25-50%, 50-75%, 75-100%) para identificar patrones particulares en las sesiones de carga según el nivel de energía inicial y final.

Finalmente, se integraron los resultados mediante ensambles, combinando múltiples algoritmos y técnicas de reducción de dimensionalidad. Esto permitió mejorar la segmentación de los datos y obtener una visión más robusta de las posibles diferencias entre las curvas de carga de los vehículos analizados.

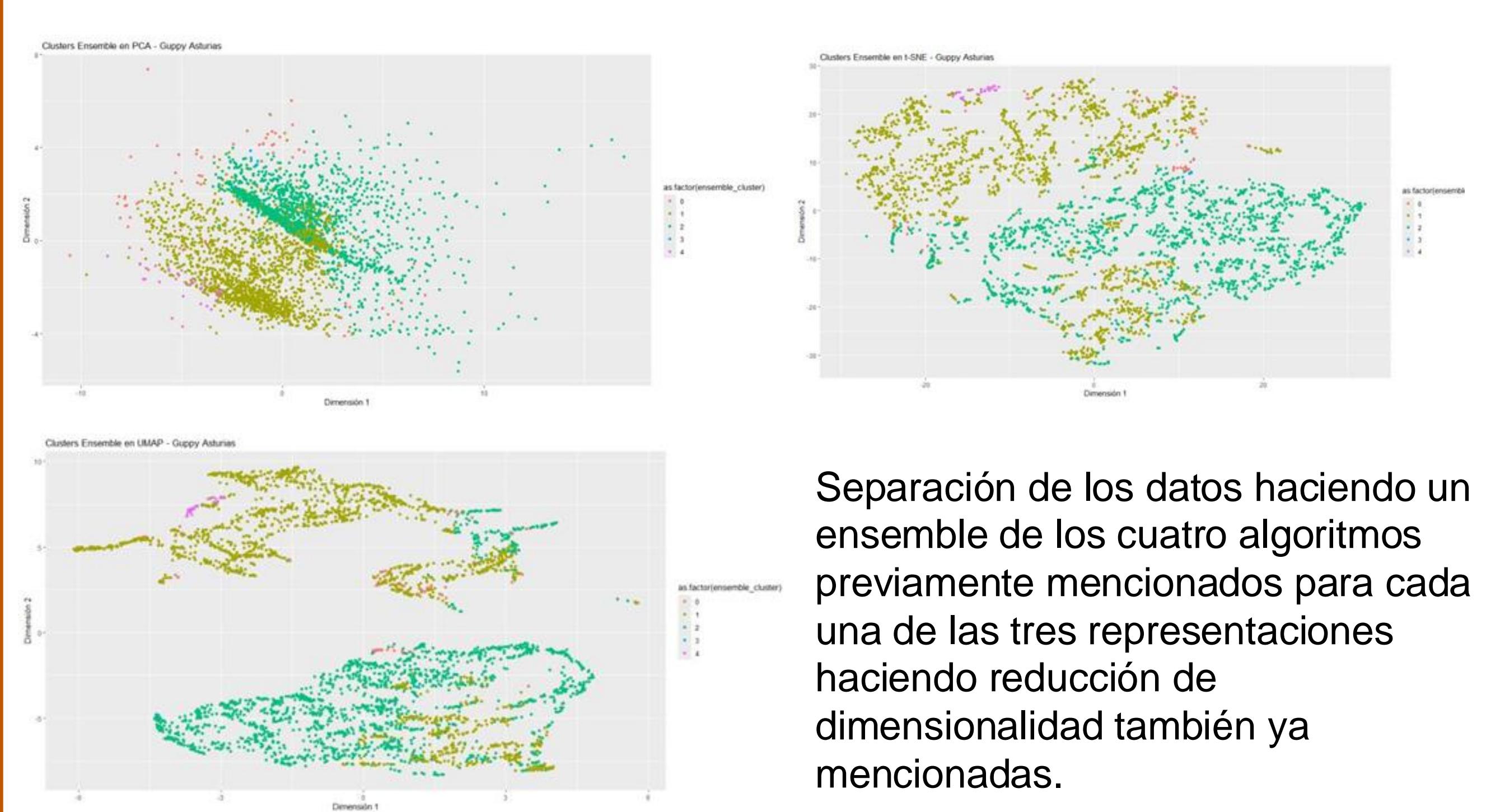
## Figuras y Resultados



Distribución de los datos con PCA frente clusterización con el algoritmo k-means y representación PCA, donde apreciamos buenos resultados, aquí analizamos la separación de TUA frente GUPPY.



La visualización obtenida mediante T-SNE aplicando los algoritmos (K-means, Db-Sacn, HD-Scan y Optics) muestra la distribución de las sesiones de carga en un espacio de menor dimensión, resaltando las relaciones tanto locales como globales entre los datos.



Separación de los datos haciendo un ensemble de los cuatro algoritmos previamente mencionados para cada una de las tres representaciones haciendo reducción de dimensionalidad también ya mencionadas.

## Conclusiones

El estudio ha demostrado que es posible identificar diferencias significativas en las curvas de carga de los vehículos GUPPY mediante técnicas de clustering y reducción de dimensionalidad, como PCA, T-SNE y UMAP. Estos métodos permitieron visualizar patrones latentes en los datos y agrupar las sesiones de carga en función de características compartidas.

Los análisis iniciales, aplicados a rangos de carga (0-25%, 25-50%, 50-75%, 75-100%), evidenciaron que los patrones eran más definidos en algunos intervalos, como 0-25% y 75-100%, mientras que el rango 25-50% presentó mayores dificultades de separación. La escalabilidad de los datos también jugó un papel importante, mostrando que los datos normalizados ofrecían mejores resultados en la segmentación.

La integración de múltiples algoritmos de clustering mediante ensambles proporcionó una representación más robusta y consistente, destacando como una técnica clave para la exploración de diferencias entre modelos de vehículos. Aunque los resultados son prometedores, la falta de identificadores específicos de vehículos limita la validación completa del análisis, ya que no es posible relacionar los grupos obtenidos con características específicas, como modelo, año de fabricación o nivel de degradación de las baterías.

Con la incorporación de identificadores de vehículos, será factible validar los resultados obtenidos, analizar cómo evolucionan las curvas de carga a lo largo del tiempo y desarrollar modelos predictivos y de clasificación capaces de identificar vehículos y anticipar el comportamiento de sus baterías.

## Trabajo futuro

El estudio se ha realizado con datos bastante limitantes, ya que no sabemos qué cargas corresponden a qué vehículo con certeza.

Una vez disponibles los identificadores específicos de los vehículos, el siguiente paso será analizar la evolución temporal de las curvas de carga para cada vehículo, identificando patrones relacionados con la degradación de la batería y su uso a lo largo del tiempo. Esto permitirá desarrollar un modelo predictivo capaz de anticipar el comportamiento futuro de las sesiones de carga y, además, crear un modelo de clasificación que, basándose en los datos de la carga, pueda identificar a qué vehículo corresponde una sesión..