Análisis y predicción del precio de vehículos segunda mano

Maribel Seara

Introducción

El mercado de vehículos de segunda mano siempre ha estado activo en España pero la demanda se ha visto aumentada notablemente tras la pandemia del COVID19 con la escasez de materiales para fabricar coches nuevos. Clientes que tradicionalmente compraban vehículos nuevos se han desplazado al mercado de segunda mano. En consecuencia cabe esperar que los precios hayan aumentado en relación a antes de la pandemia y por el aumento de demanda. Aunque este aumento puede haberse suavizado por el restablecimiento de la fabricación de vehículos nuevos, la reciente crisis tras la invasión de Ucrania ha conllevado un aumento general de los precios, así que se espera encontrar dicho aumento de precio en los anuncios actuales.

Los datos

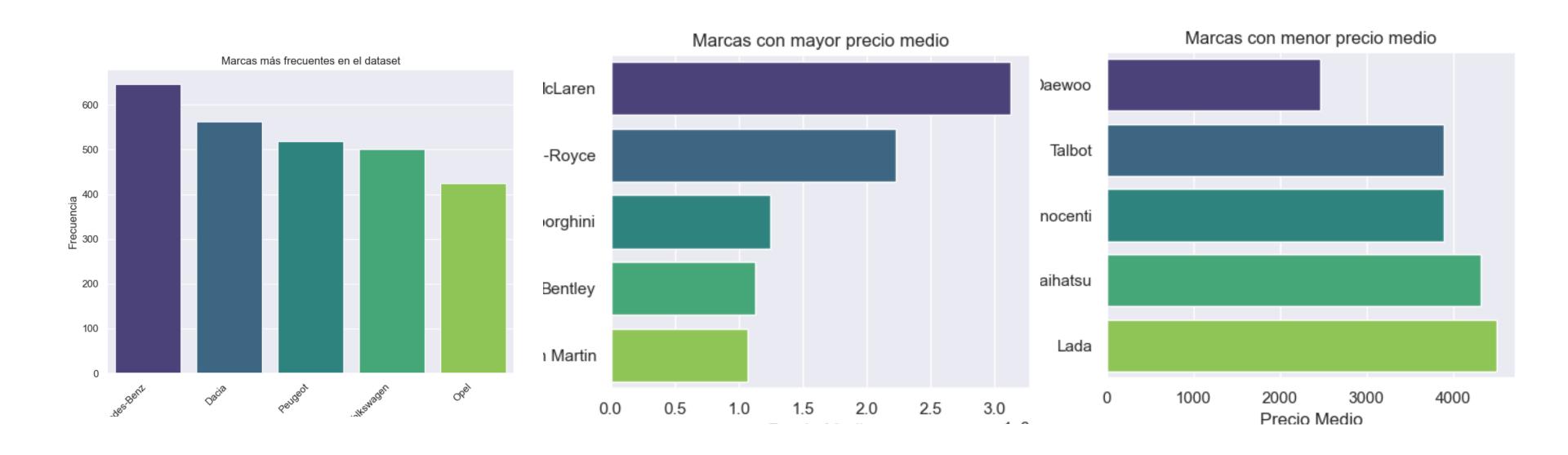
El dataset principal es el resultado de un scrapeado de unos 15000 registros realizado por mí en la web Autoscout24. Este proceso requirió varias semanas ya que cada búsqueda solo devolvía un máximo de 400 resultados y algunos de ellos no se incorporaban por haberse obtenido en búsquedas anteriores.

Para poder realizar análisis comparativo se han utilizado otros dos datasets obtenidos de Kaggle ambos provenientes de scrapeos de la web Autoscout24:

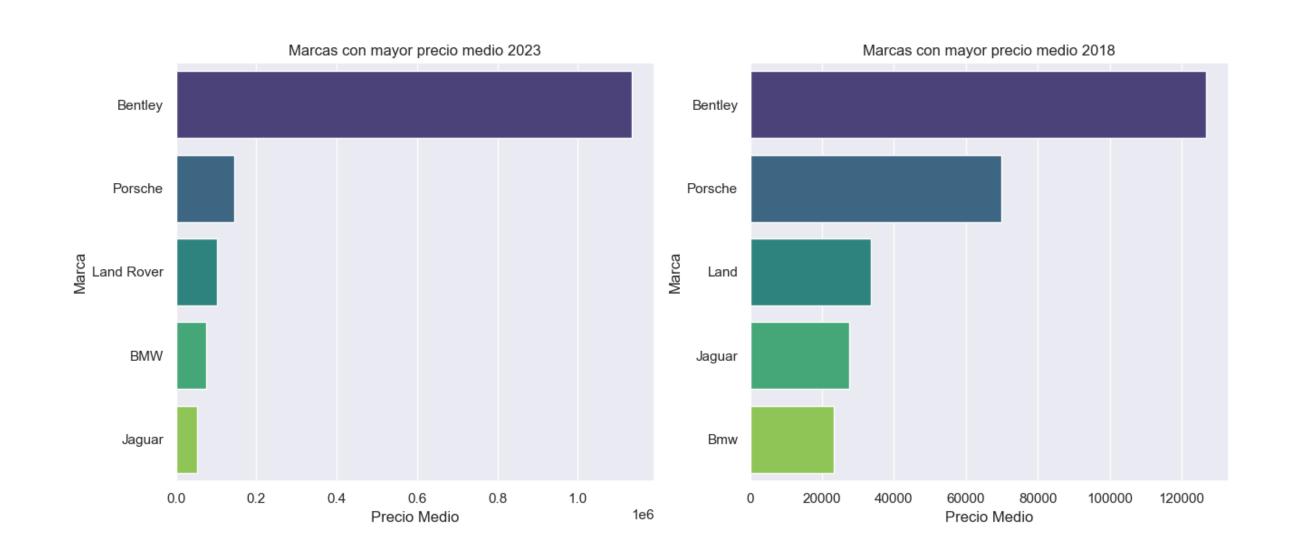
- -Dataset con más de 100.000 registros de 2018 en España.
- -Dataset con más de 250.000 registros de 2023 en Alemania.



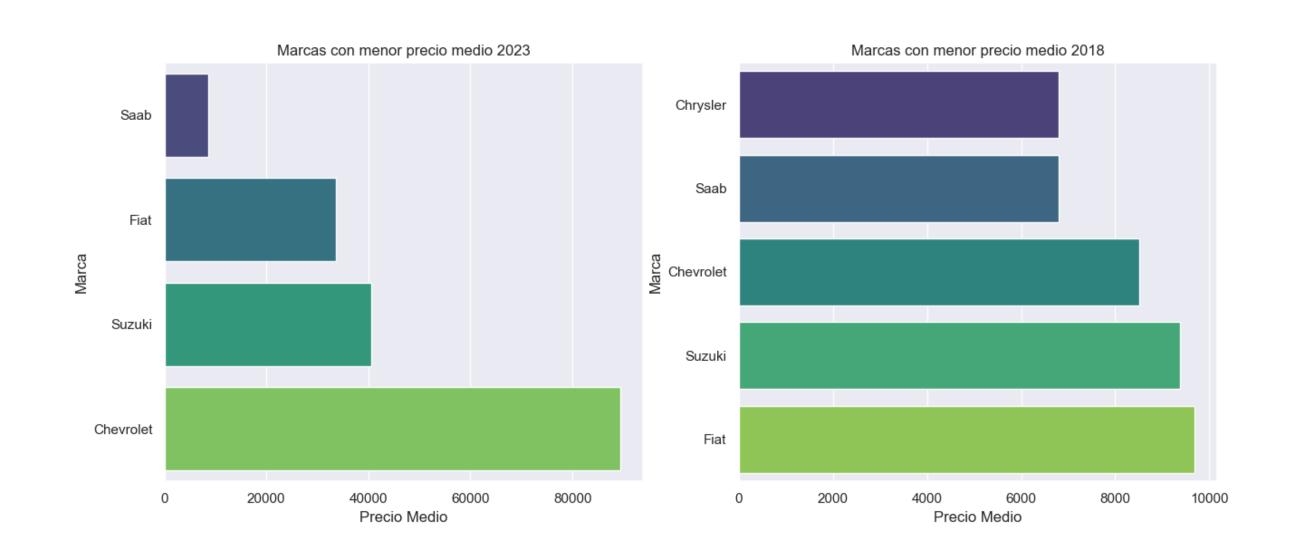
El análisis España 2023



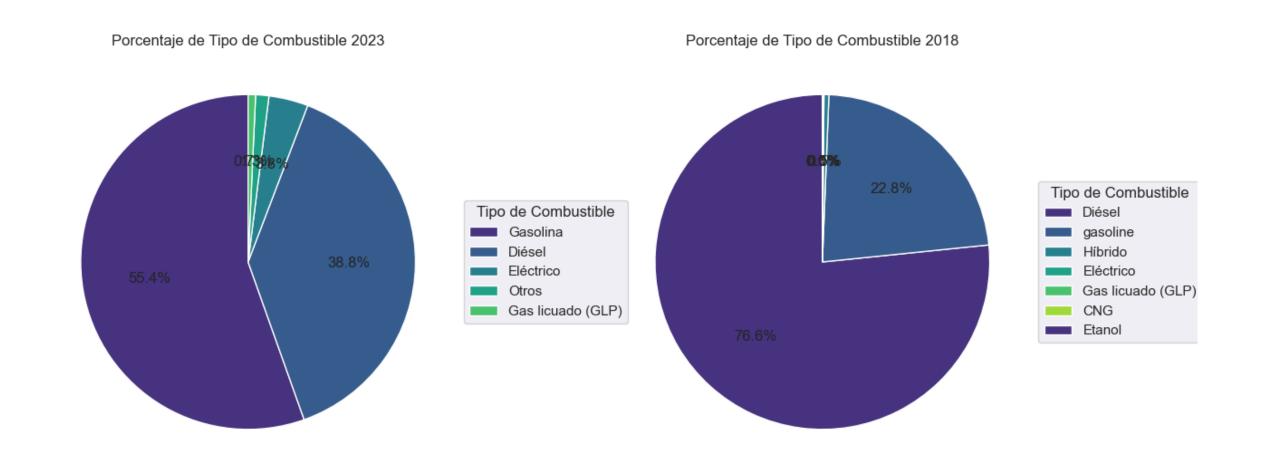
Comparativa con 2018



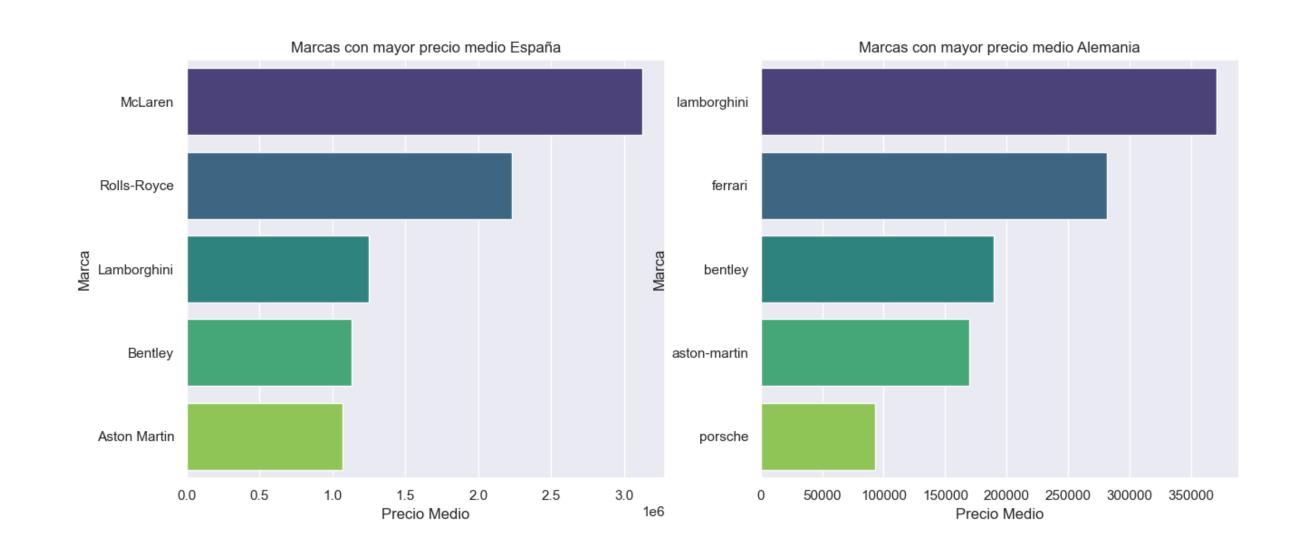
Comparativa con 2018



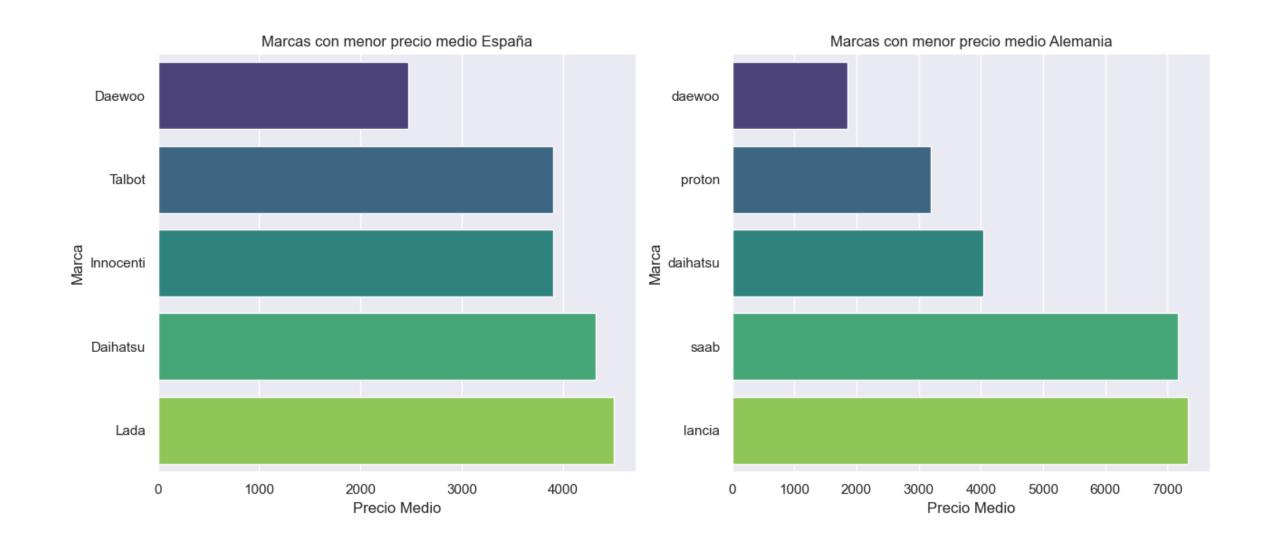
Comparativa con 2018



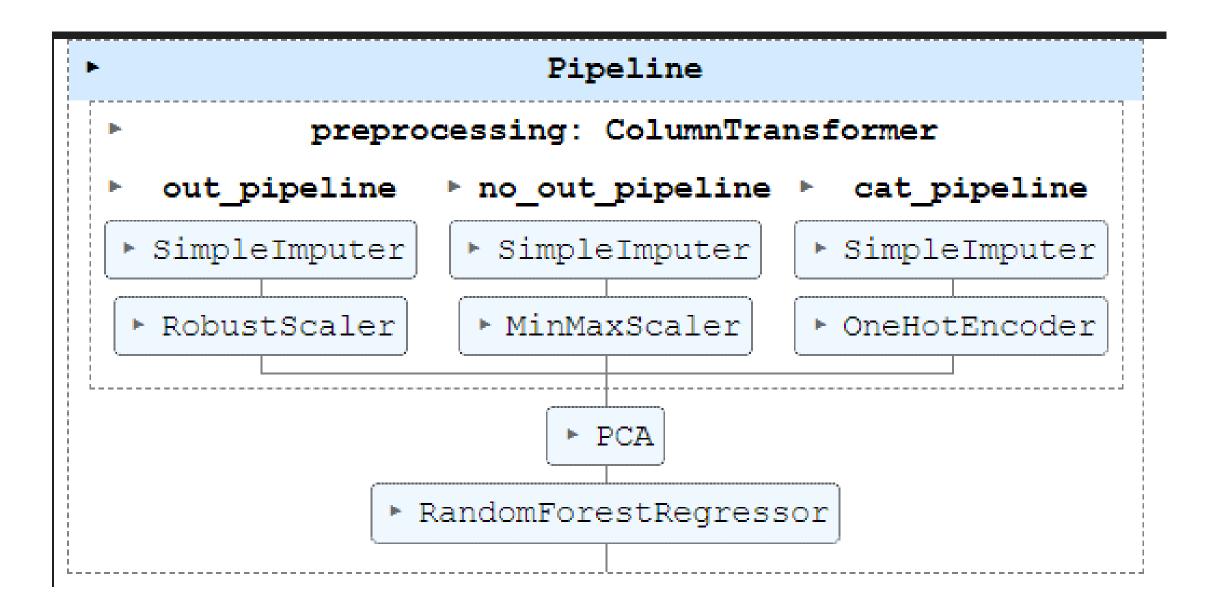
Comparativa con Alemania



Comparativa con Alemania



Machine Learning

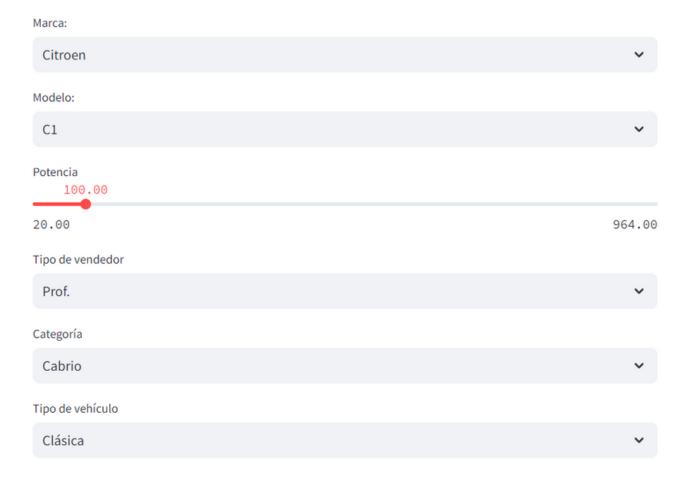


La web

Se trata de una web en local cuyo objetivo es tener un entorno más amigable para interactuar con el modelo y sus predicciones.

Está pensada como un asesor que te recomienda un precio al que anunciar tu coche de segunda mano.

Asesor precios anuncios vehículos segunda mano





Conclusiones

- -El precio de los vehículos en España en general ha aumentado mucho, pero sin más datos es dificil atribuirlo a una causa en concreto, hay muchos factores que pueden afectar.
- -Siempre ha habido una idea general en la población de que los vehículos en Alemania eran más baratos y el análisis parece confirmarlo.
- -Los datos que he recopilado no deben ser los ideales para este tipo de predicción porque la mayoría de modelos daban un resultado por defecto negativo y algunos lo mantenían después de ajustar parámetros.
 - -Afortunadamente con el modelo Random Forest Regressor, ajustando parámetros, hemos podido obtener una puntuación R2 de 0.898

Gracias

