

EXPLORANDO LA RELACIÓN ENTRE LAS CARACTERÍSTICAS DE LOS VEHÍCULOS Y SUS PRECIOS

AUTOR: PABLO HERNÁN DÍAZ



CONTENIDO

- Presentación del problema específico, motivación y audiencia**
- Resumen de los Datos**
- Preguntas de interés**
- Análisis exploratorio de los datos (EDA)**
- Objetivo y Modelos Predictivos**
- Insights**
- Recomendaciones**
- Conclusiones**

PRESENTACIÓN DEL PROBLEMA ESPECÍFICO, MOTIVACIÓN Y AUDIENCIA

Presentación del problema específico

Un grupo de concesionarias de automóviles desea predecir los precios minoristas sugeridos (MSRP) de los vehículos nuevos en función de sus características, como el año de fabricación, popularidad, la potencia del motor y la eficiencia de combustible.

Motivación:

La toma de decisiones en la industria automotriz es fundamental para el éxito de las concesionarias y fabricantes de vehículos. La información precisa y detallada sobre la relación entre las características de los automóviles y su precio minorista sugerido (MSRP) es esencial para la toma de decisiones estratégicas. El análisis de estos datos proporciona insights que pueden mejorar significativamente las operaciones y la rentabilidad de las concesionarias.

Audiencia:

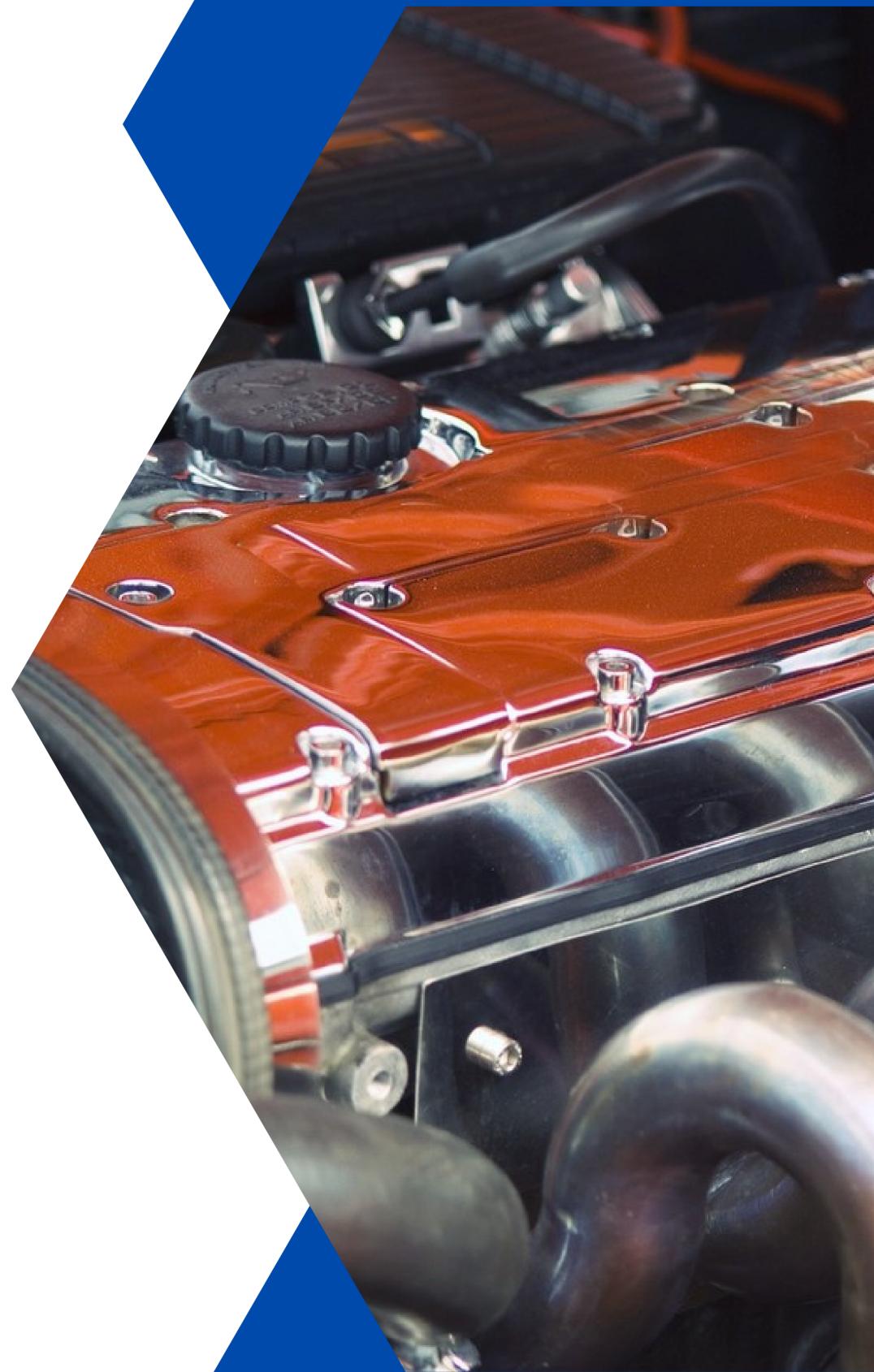
Esta presentación está dirigida a las concesionarias de automóviles que desean comprender mejor cómo las características de los vehículos influyen en su valor en el mercado, en especial la relación del valor de los vehículos en función a la potencia y el consumo de combustible. El público objetivo incluye a gerentes, analistas y profesionales de marketing de concesionarias de automóviles que buscan tomar decisiones más informadas sobre la adquisición, precios y estrategias de ventas de vehículos.

Resumen de los Datos

<https://www.kaggle.com/datasets/CooperUnion/cardataset>

El conjunto de datos "Car Features and MSRP" es un recurso útil para analizar la relación entre los atributos de los automóviles y su precio minorista sugerido (MSRP). Fue recopilado por Jeffrey C. Schlimmer del Departamento de Ciencias de la Computación de la Universidad de California, Irvine, y contiene información detallada sobre 10,193 automóviles de diferentes marcas y modelos. Los datos recopilados son de automóviles producidos entre 1990 y 2017 y han sido ampliamente utilizados en la investigación de aprendizaje automático y la toma de decisiones en el mercado automotriz. El conjunto de datos consta de 16 columnas que incluyen información sobre el precio, la marca, el modelo, el año, el tipo de cuerpo, la transmisión, la eficiencia de combustible en ciudad y carretera, el tamaño del motor, la potencia y el número de cilindros, entre otras características.

#	Column	Non-Null Count	Dtype	Description
0	Make	11914	non-null object	1-Make: Marca del automóvil.
1	Model	11914	non-null object	2-Model: Modelo del automóvil.
2	Year	11914	non-null int64	3-Year: Año de fabricación del automóvil.
3	Engine Fuel Type	11911	non-null object	4-Engine Fuel Type: Tipo de combustible del motor.
4	Engine HP	11845	non-null float64	5-Engine HP: Potencia del motor en caballos de fuerza.
5	Engine Cylinders	11884	non-null float64	6-Engine Cylinders: Número de cilindros en el motor.
6	Transmission Type	11914	non-null object	7-Transmission Type: Tipo de transmisión del automóvil.
7	Driven_Wheels	11914	non-null object	8-Driven_Wheels: Configuración de las ruedas motrices (tracción trasera, delantera o en las cuatro ruedas).
8	Number of Doors	11908	non-null float64	9-Number of Doors: Número de puertas del automóvil.
9	Market Category	8172	non-null object	10-Market Category: Categoría de mercado del automóvil (ejemplo: sedán, SUV, deportivo, etc.).
10	Vehicle Size	11914	non-null object	11-Vehicle Size: Tamaño del vehículo (pequeño, mediano, grande).
11	Vehicle Style	11914	non-null object	12-Vehicle Style: Estilo del vehículo (ejemplo: sedán, coupé, camioneta, etc.).
12	highway MPG	11914	non-null int64	12-Highway MPG: Consumo de combustible en autopista (millas por galón).
13	city mpg	11914	non-null int64	14-City MPG: Consumo de combustible en ciudad (millas por galón).
14	Popularity	11914	non-null int64	15-Popularity: Popularidad del automóvil basada en el número de búsquedas en Internet y ventas.
15	MSRP	11914	non-null int64	16-MSRP: Precio sugerido de venta al público del automóvil.



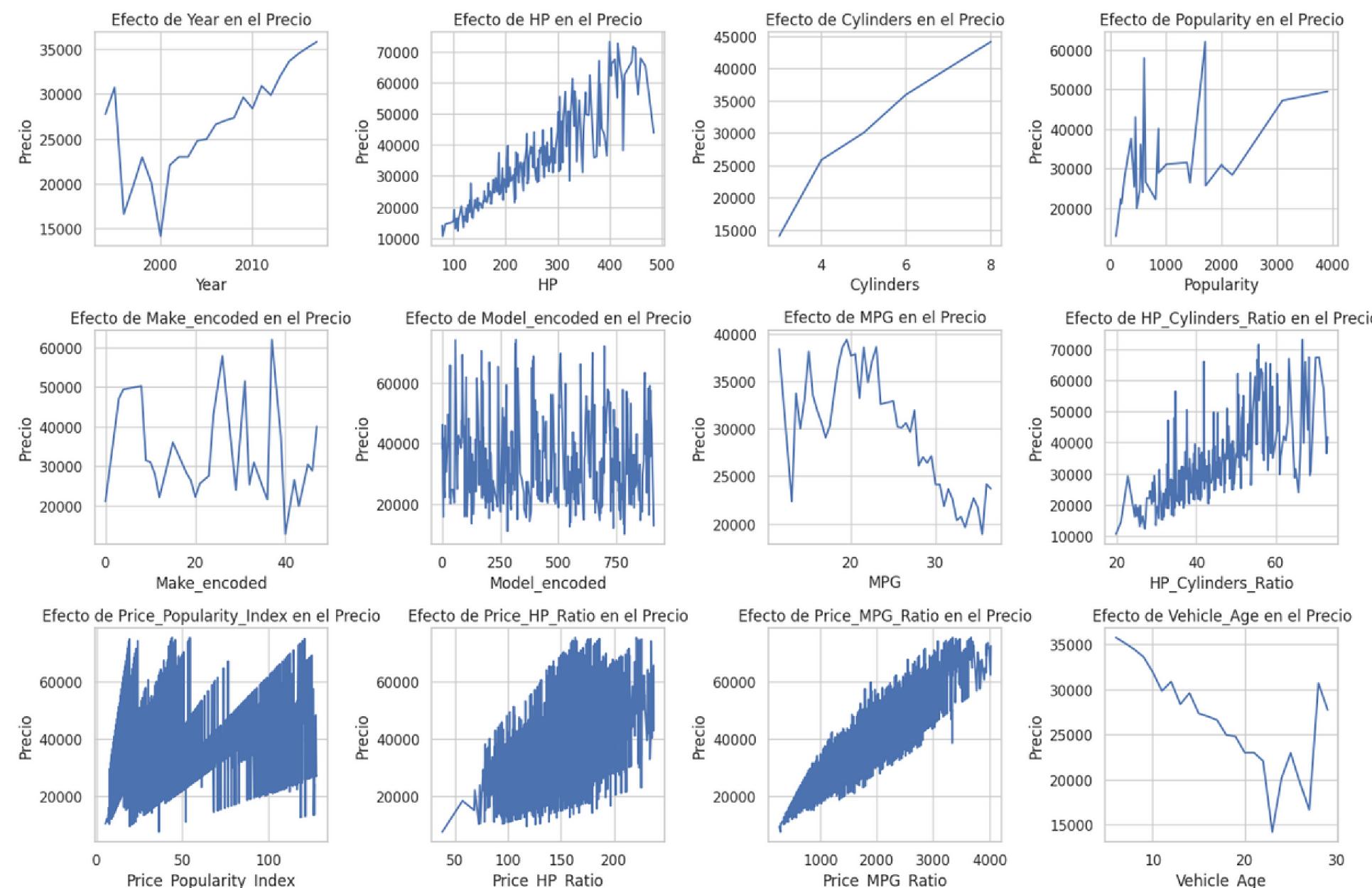
PREGUNTAS DE INTERES

- 01 ¿Cuál es la relación entre las características de los vehículos y sus precios?
- 02 ¿Qué características tienen el mayor impacto en el precio de un vehículo?
- 03 ¿Cómo se comparan los precios de los vehículos de diferentes marcas?
- 04 ¿Cómo ha evolucionado el precio de los vehículos a lo largo del tiempo?

ANÁLISIS EXPLORATORIO DE LOS DATOS (EDA)

¿Cuál es la relación entre las características de los vehículos y sus precios?

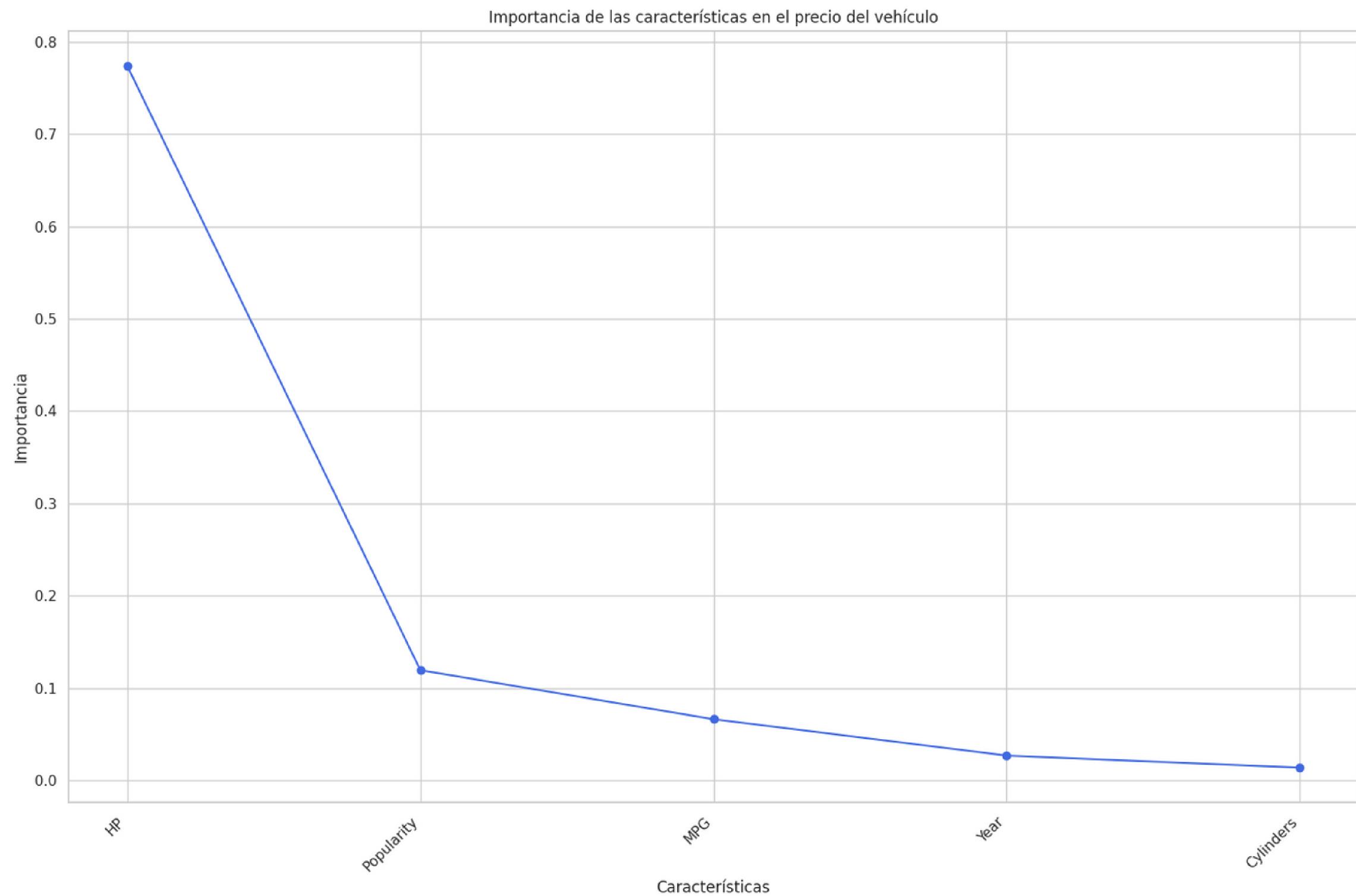
- Las características de los vehículos, como la potencia (HP) y la popularidad, tienen un impacto significativo en sus precios.
- Mayor potencia (HP) y popularidad tienden a aumentar el precio de un vehículo.
- Existe una relación inversa entre el rendimiento de combustible y el precio; los vehículos menos eficientes en términos de consumo tienden a ser más caros



ANÁLISIS EXPLORATORIO DE LOS DATOS (EDA)

¿Qué características tienen el mayor impacto en el precio de un vehículo?

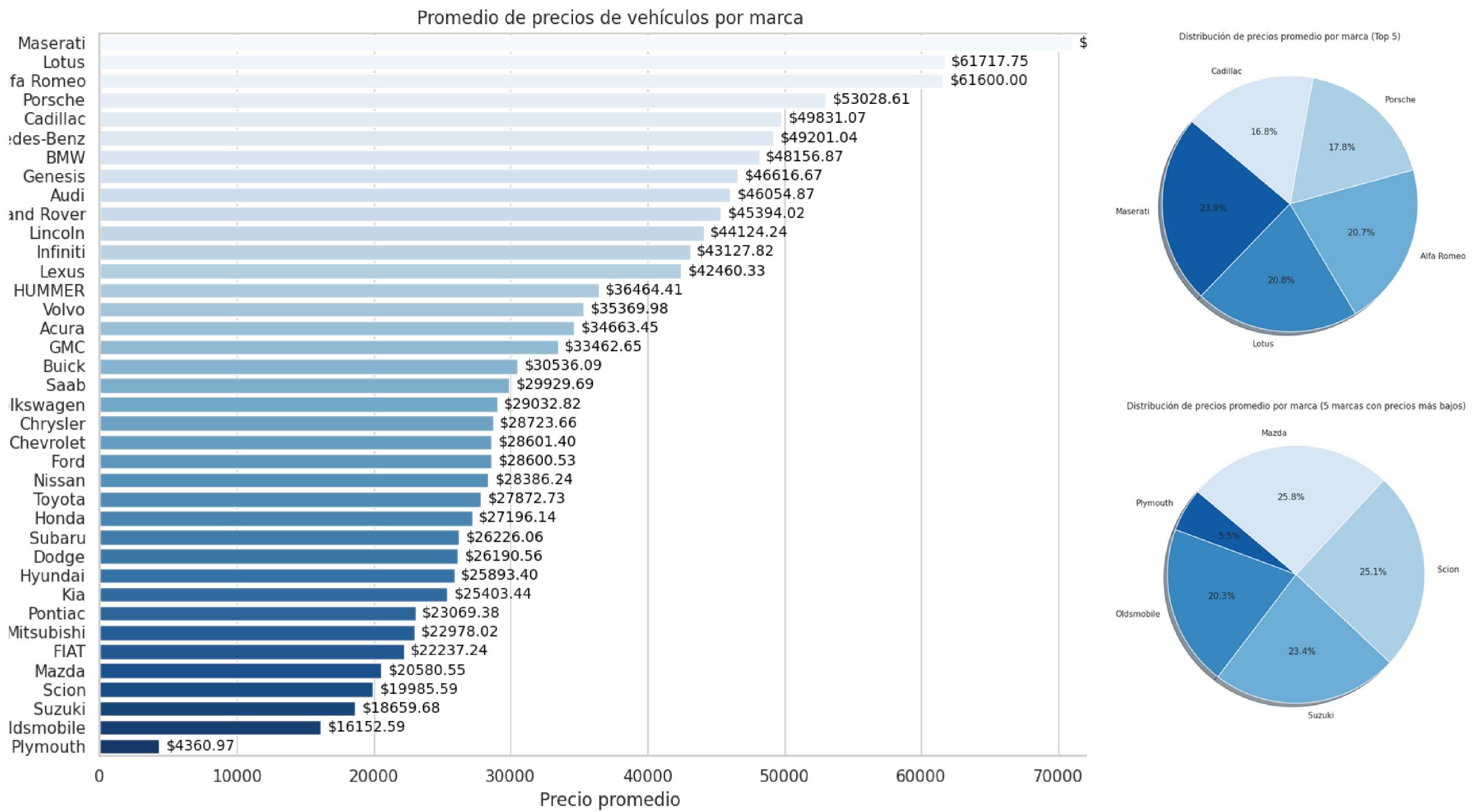
Se realizó mapa de calor en la cual se observó las características que tiene mayor impacto en el precio son "HP" y "Popularity", siendo "HP" la de mayor impacto



ANÁLISIS EXPLORATORIO DE LOS DATOS (EDA)

¿Cómo se comparan los precios de los vehículos de diferentes marcas?

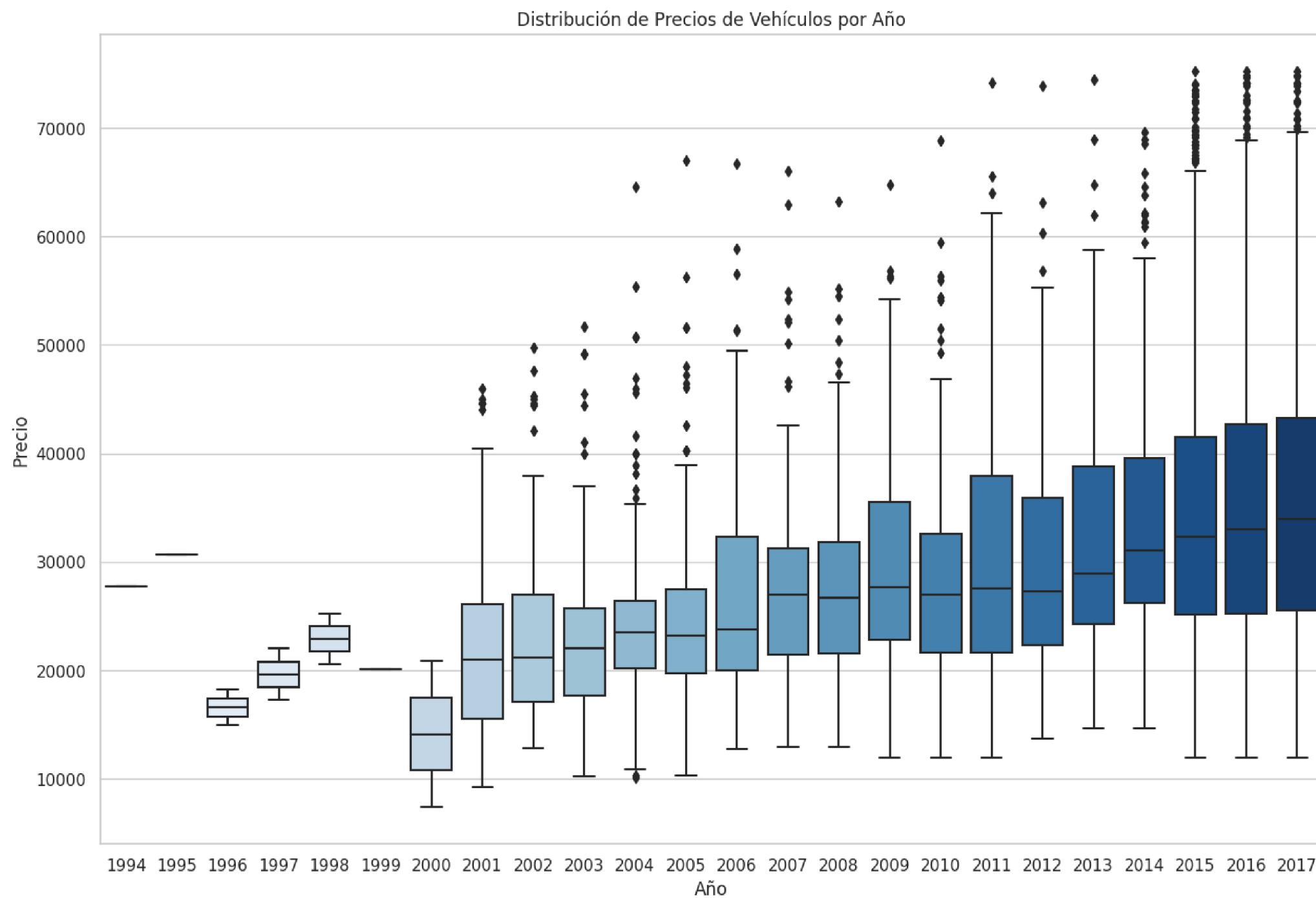
- Las marcas de vehículos se pueden clasificar en función de sus precios promedio.
- Las marcas con una mayor producción tienden a tener precios más bajos.
- Las marcas de mayor producción se sitúan en el rango de precios promedio o por debajo de la media.



ANÁLISIS EXPLORATORIO DE LOS DATOS (EDA)

¿Cómo se comparan los precios de los vehículos de diferentes marcas?

- Los precios se mantuvieron relativamente estables entre 1994 y 2000.
- En el año 2001, los precios aumentaron significativamente y se mantuvieron altos.
- En los últimos 4 años, los precios se estabilizaron nuevamente.
- A largo plazo, los precios de los vehículos tienden a aumentar con el tiempo.



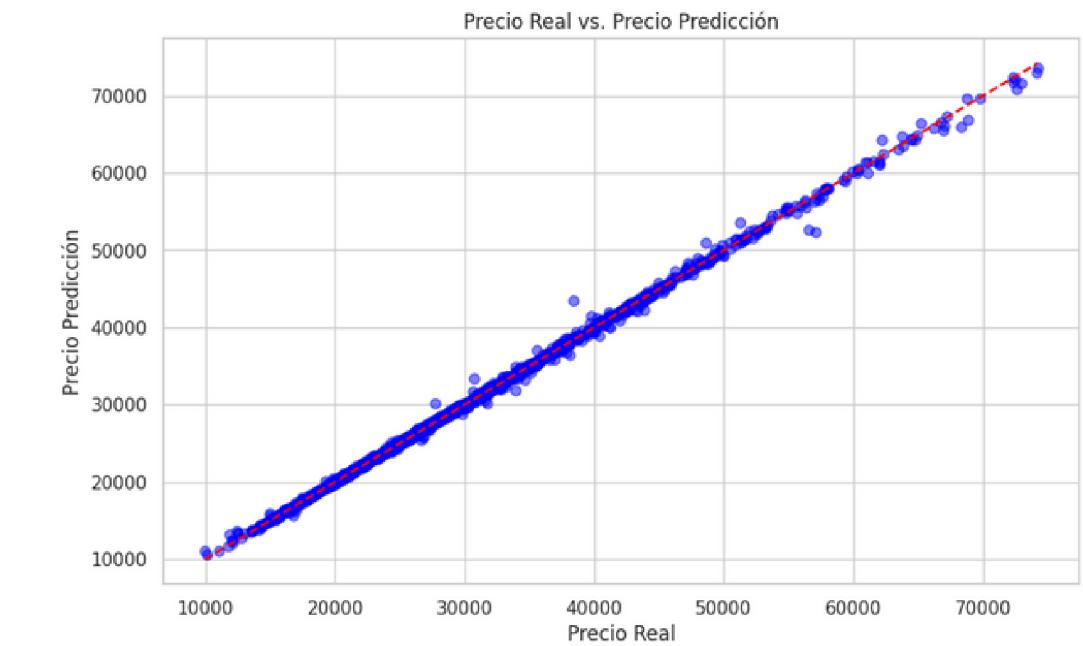
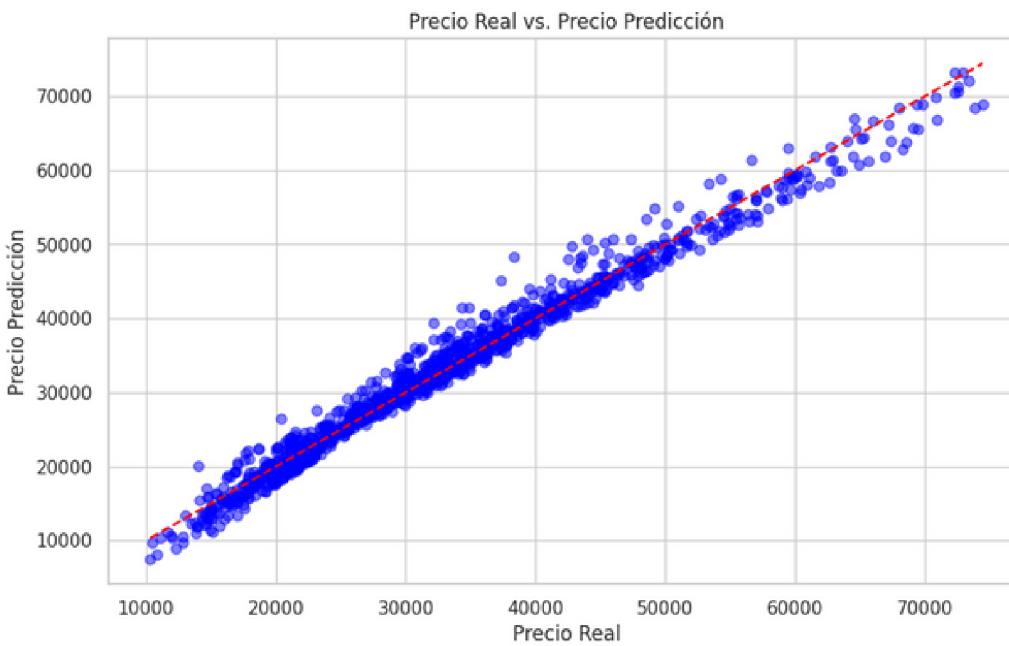
OBJETIVO Y MODELOS PREDICTIVOS

Objetivo

El objetivo principal es utilizar modelos de regresión para predecir los precios de los vehículos en función de sus características.

Modelos Predictivos

- Se emplearon varios algoritmos, incluyendo la Regresión Lineal y el Random Forest, para lograr este propósito.
- Rendimiento de los Modelos:
 - El modelo Random Forest, en sus diferentes variantes, demostró un rendimiento excepcionalmente alto con puntajes R² cercanos a 0.999, lo que indica un ajuste muy cercano a los datos reales.
 - La Regresión Lineal, aunque respetable, tuvo un rendimiento ligeramente inferior con un puntaje R² de aproximadamente 0.979.



INSIGHTS

- **Relación entre Características y Precios:**
 - La potencia (HP) y la popularidad tienen un impacto significativo en los precios de los vehículos. Esto sugiere que los consumidores valoran los vehículos más potentes y populares, lo que justifica precios más altos.
- **Impacto de la Eficiencia de Combustible:**
 - Existe una relación inversa entre la eficiencia de combustible y el precio de los vehículos. Los vehículos menos eficientes en consumo tienden a ser más caros. Esto indica que la eficiencia de combustible es un factor importante para los compradores conscientes del consumo de combustible.
- **Comparación de Precios entre Marcas:**
 - Las marcas de vehículos se pueden clasificar en función de sus precios promedio. Las marcas con una mayor producción tienden a ofrecer precios más bajos, lo que puede influir en las estrategias de adquisición y precios de la concesionaria.
- **Evolución de los Precios a lo Largo del Tiempo:**
 - Los precios de los vehículos se mantuvieron relativamente estables entre 1994 y 2000, pero experimentaron un aumento significativo en 2001. Este cambio puede deberse a factores económicos o de mercado. Es importante estar atento a las tendencias temporales.
- **Rendimiento de los Modelos de Predicción:**
 - Los modelos de Random Forest, especialmente aquellos con optimización de hiperparámetros, demostraron un rendimiento sobresaliente en la predicción de precios de vehículos. Estos modelos tienen un alto ajuste a los datos reales, lo que los hace ideales para establecer precios competitivos.
- **Generalización y Robustez de los Modelos:**
 - Tanto la Regresión Lineal como el Random Forest demostraron una capacidad sólida de generalización. Esto significa que los modelos son capaces de realizar predicciones precisas y consistentes en datos no vistos durante el entrenamiento, lo que respalda la toma de decisiones basadas en datos.

RECOMENDACIONES

- **Estrategia de Precios:**
 - Dado que las características como la potencia (HP) y la popularidad tienen un impacto significativo en los precios, considera ajustar tus estrategias de precios en función de estos factores. Los vehículos con características destacadas en estas áreas pueden justificar precios más altos.
- **Diversificación del Inventory:**
 - La relación entre el rendimiento de combustible y el precio sugiere que los vehículos menos eficientes en consumo tienden a ser más caros. Esto puede influir en las decisiones de compra de los clientes conscientes del consumo de combustible. Considera diversificar tu inventario con opciones de alta eficiencia de combustible para atraer a un público más amplio.
- **Seguimiento de las Tendencias del Mercado:**
 - El análisis de la evolución de los precios a lo largo del tiempo indica que los precios tienden a aumentar con el tiempo. Mantente al tanto de las tendencias actuales del mercado y ajusta tus estrategias de precios en consecuencia.
- **Marcas y Producción:**
 - Analiza las marcas de vehículos que ofreces y su producción. Las marcas con una producción más alta tienden a ofrecer precios más competitivos. Considera la posibilidad de ampliar tu oferta de marcas populares y con una amplia gama de modelos.
- **Modelos de Predicción:**
 - Utiliza modelos de predicción, como Random Forest, para estimar con precisión los precios de los vehículos en función de sus características. Estos modelos pueden ayudarte a establecer precios competitivos y tomar decisiones informadas sobre la adquisición de inventario.
- **Optimización de Hiperparámetros:**
 - Si optas por utilizar modelos de Random Forest, considera la optimización de hiperparámetros para mejorar aún más la precisión de las predicciones. Esto puede implicar ajustar los parámetros del modelo para obtener un rendimiento óptimo.

CONCLUSIONES

En resumen, los análisis y modelos aplicados a los datos de vehículos proporcionan información valiosa para las concesionarias de automóviles. Se ha identificado la relación entre las características de los vehículos y sus precios, lo que puede ayudar en la toma de decisiones sobre precios, adquisiciones y estrategias de ventas. Además, los modelos de predicción, especialmente el Random Forest, son altamente precisos y generalizables, lo que los convierte en herramientas efectivas para estimar precios de vehículos en función de sus atributos. En general, este proyecto proporciona una respuesta que supera las expectativas en términos de valor y precisión en la toma de decisiones en la industria automotriz.