

ANÁLISIS DEL MERCADO DE VEHÍCULOS USADOS

Saúl Arranz Herrero

THE  BRIDGE

ÍNDICE

1. Introducción
2. Hipótesis de partida
3. Análisis exploratorio
4. Análisis de variables
5. Contraste de hipótesis



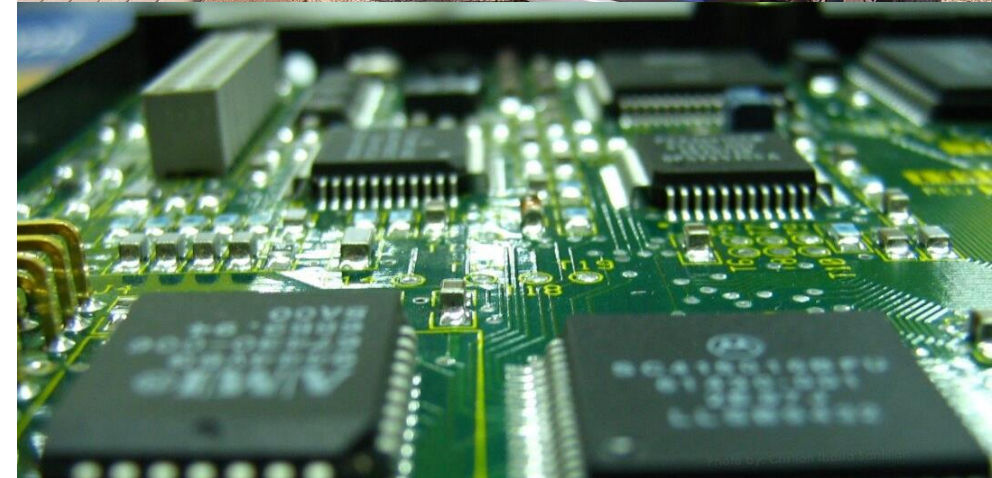
1. INTRODUCCIÓN

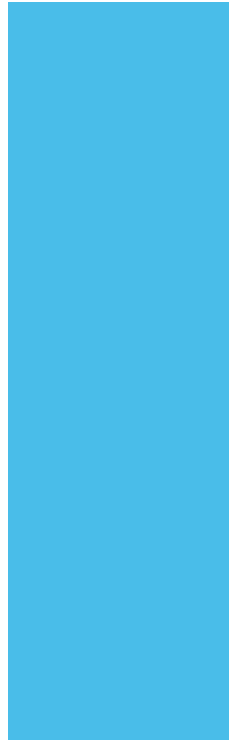
La situación económica y social actual ha provocado mucha incertidumbre en el mercado de la automoción:

- ¿Prohibición vehículos de combustión en 2030?
- Crisis de componentes → largos tiempos de espera para conseguir un coche nuevo
- Cambios en la movilidad → car sharing
- Crisis derivada de la pandemia

Todo esto hace que el mercado de segunda mano haya tomado mucho protagonismo en los últimos meses/años.

Por estas razones vamos a realizar un estudio sobre el precio de los vehículos de segunda mano a la venta en España





2.HIPÓTESIS DE PARTÍDA

Dado el contexto del que hemos hablado en la diapositiva anterior, definimos las hipótesis:

- Los vehículos con más km son los más baratos

- La marca de coches más barata es Dacia

- Los vehículos híbridos son más caros que los Diésel

- La media de precios es la misma para todas las comunidades autónomas

- Los vehículos vendidos por profesionales son más caros que los vendidos por particulares

- Los vehículos más potentes son los más caros



3. ANÁLISIS EXPLORATORIO



Los datos se han obtenido de la base de datos
<https://www.kaggle.com/datasets/datamarket/venta-de-coches>

Los datos son legibles y el dataset se carga sin problema en Visual Studio

make	model	version	price
SEAT	Toledo	SEAT Toledo 4p.	950
CITROEN	C1	CITROEN C1 PureTech 60KW 82CV Feel 5p.	6200
FORD	Transit Connect	FORD Transit Connect Van 1.5 TDCi 100cv Ambien...	7851



`df.shape =`
`(50000, 21)`

observaciones suficientes para el análisis que queremos



No existen elementos duplicados y no hay gran cantidad de elementos nulos



El dataset consta de las siguientes variables:

- power
- doors
- shift
- url
- color
- company
- potos
- make
- model
- version
- Price
- price financed
- fuel
- year
- kms
- is professional:
- vendedor
- dealer
- province
- country
- publish date
- insert date



3. ANÁLISIS EXPLORATORIO



Creación de una nueva columna “Comunidad Autónoma”



Se eliminan las columnas que no se van a necesitar para el análisis, quedan:

- Marca
- Precio
- Combustible
- Año
- Kilómetros
- Potencia
- Vendedor profesional
- Comunidad Autónoma



Se completan los valores nulos de las columnas Marca, Modelo y Potencia

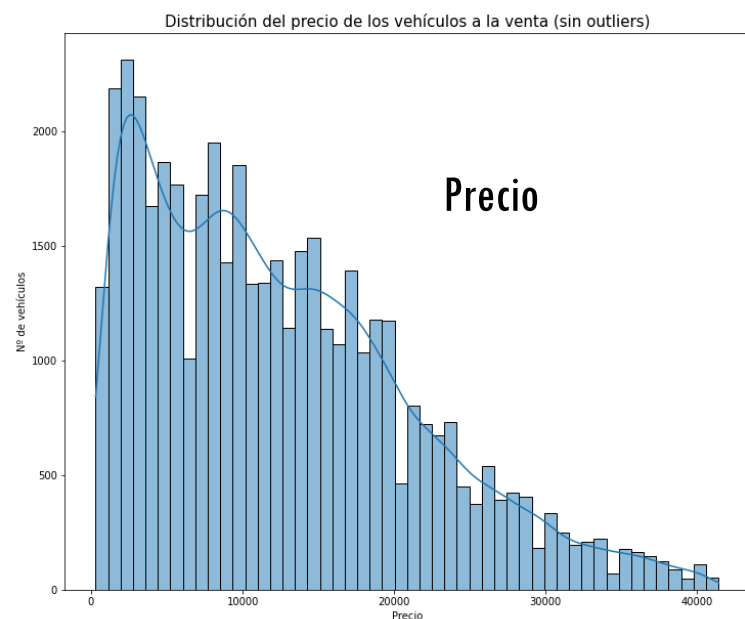


Finalmente se eliminan los outliers:

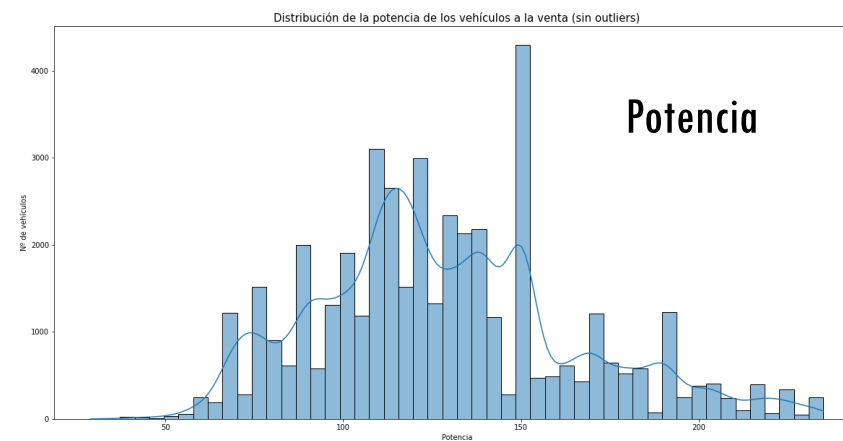
- Vehículos exclusivos
- Vehículos con muchos kms



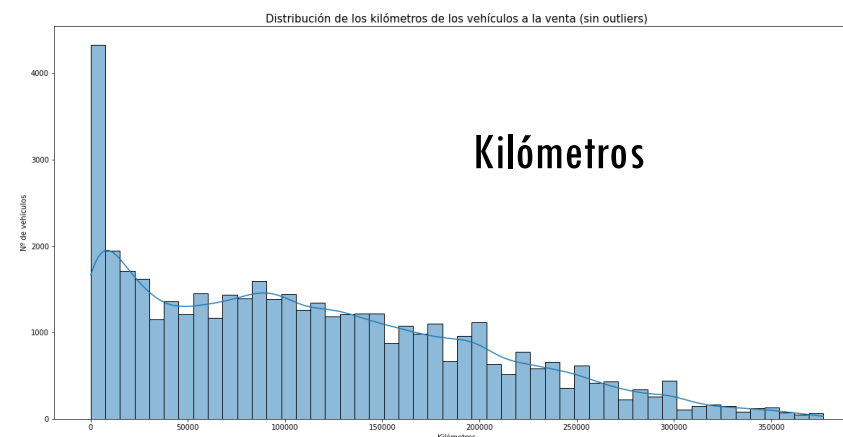
4. ANÁLISIS DE VARIABLES



La mayor parte de las observaciones se encuentran en la primera mitad de gráfica

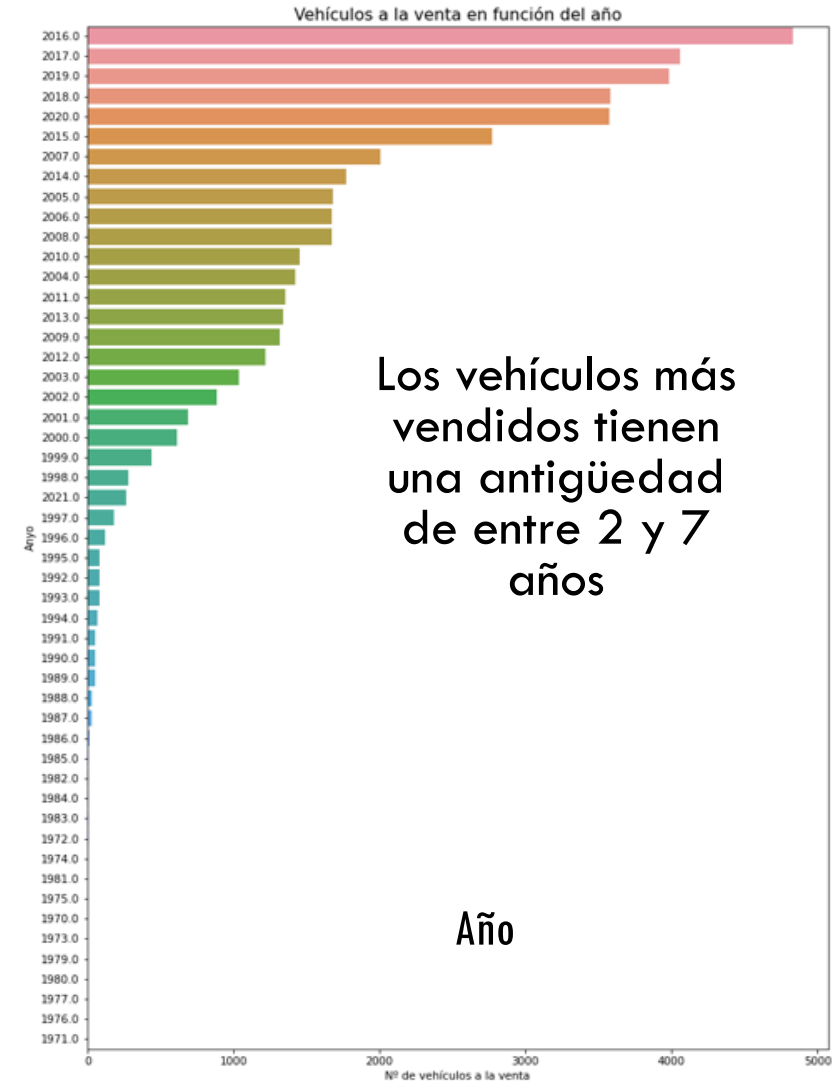


Los datos de potencia están más concentrados en la zona central sin llegar a ser una dsit. Normal

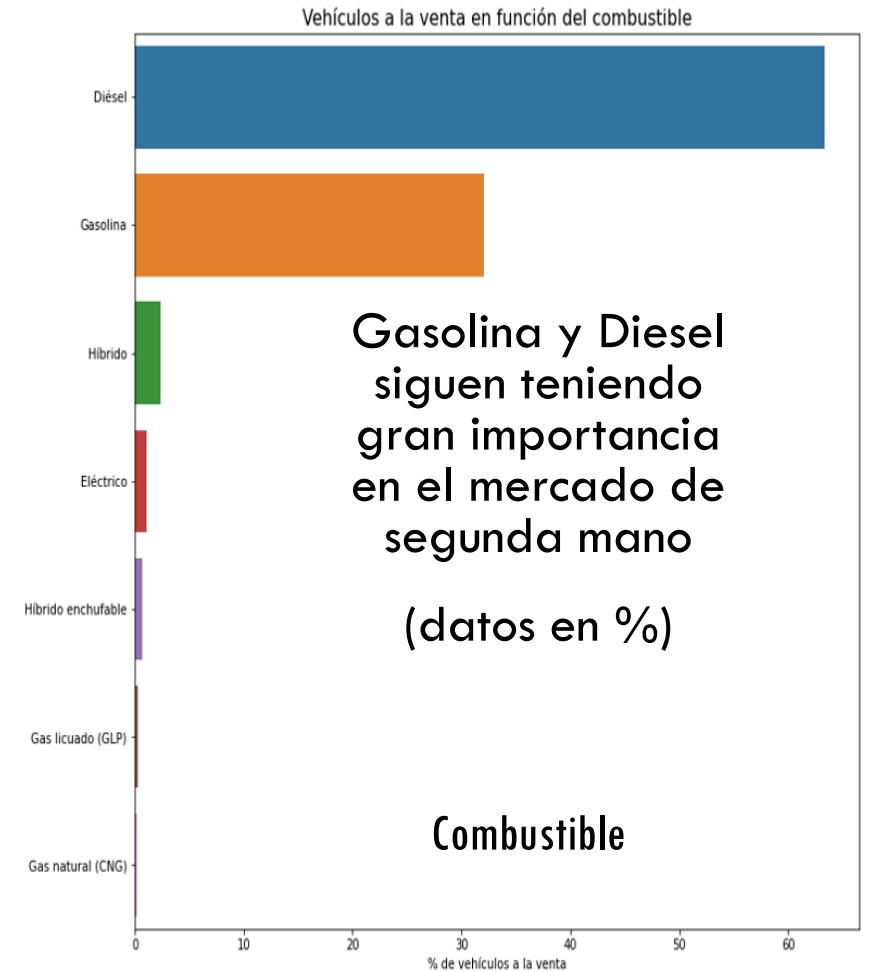
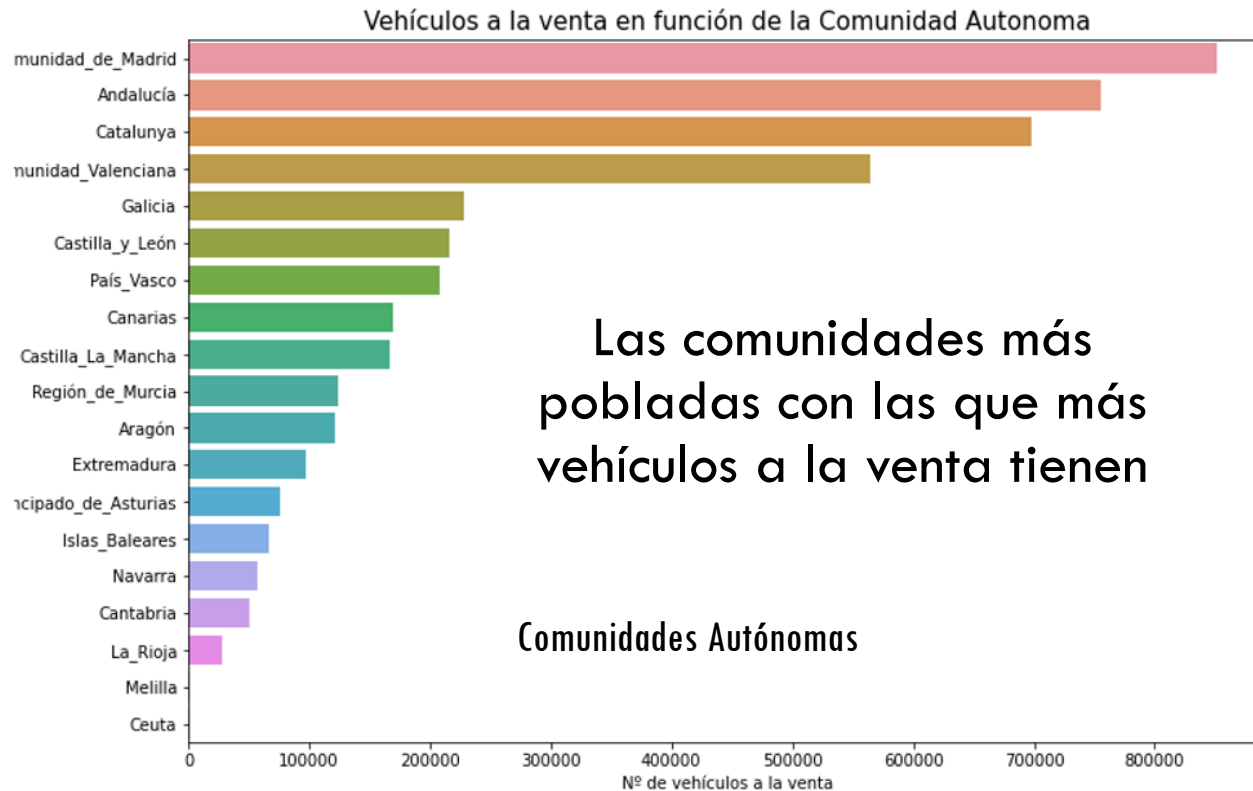


La mayoría de los datos de kilómetros, como los de precio, se concentran la primera mitad de la tabla

4. ANÁLISIS DE VARIABLES



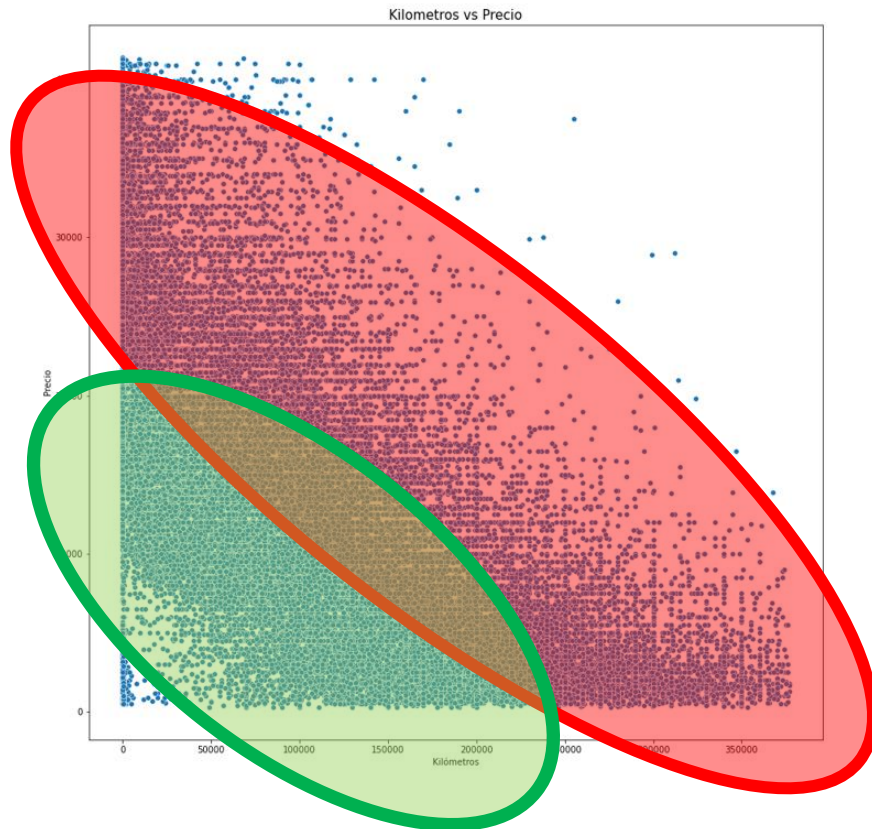
4. ANÁLISIS DE VARIABLES



Los vehículos con más km son los más baratos → **confirmada**

5.CONTRASTE DE HIPÓTESIS

↑ Kilómetros → ↓ Precio



* Diesel
* Gasolina



5.CONTRASTE DE HIPÓTESIS

La marca de coches más barata es Dacia



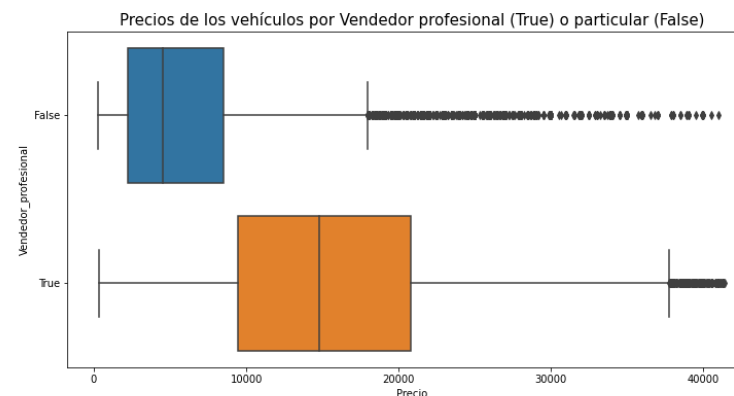
Marca	Precio
DAEWOO	1264
TATA	2516
MG	2757
ROBER	3313
DAIHATSU	3500
SAAB	3896
CHRISLER	4194
LANCIA	6156
.	
.	
.	
DACIA	8401



Marca	Kilometros
DACIA	77467
DAEWOO	117902
DAIHATSU	137000
LANCIA	138567
MG	95347
ROVER	174288
SAAB	182858
TATA	78291

Marca	DACIA	DAEWOO
Anyo		
1982	0	0
1986	0	0
1989	0	0
1992	0	0
1993	0	0
1994	0	0
1995	0	0
1996	0	0
1997	0	2
1998	0	1
1999	0	3
2000	0	7
2001	0	7
2002	0	5
2003	0	5
2004	0	11
2005	0	1
2006	3	0
2007	5	0
2008	3	0
2009	6	0
2010	6	0
2011	7	0
2012	10	0
2013	15	0
2014	38	0
2015	32	0
2016	40	0
2017	31	0
2018	31	0
2019	17	0
2020	18	0
2021	2	0
All	264	42

Los vehículos vendidos por profesionales son más caros que los vendidos por particulares

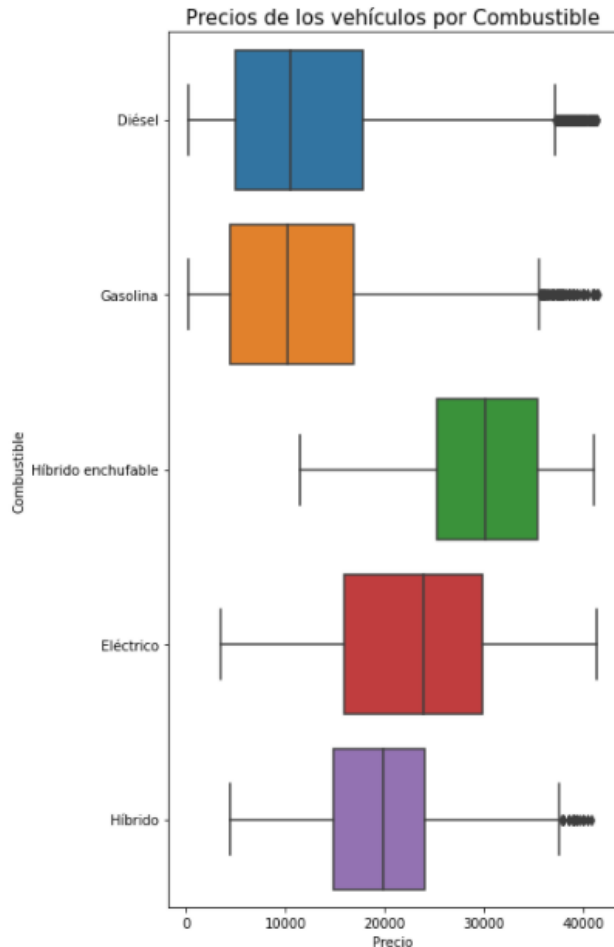


Vendedor_profesional	Precio
False	11633
True	18238

Combustible	Diésel	Eléctrico	Gasolina	Hibrido	Hibrido enchufable	All
Vendedor_profesional						
False	10635	23	4886	92	21	15657
True	17732	464	9494	973	277	28940
All	28367	487	14380	1065	298	44597

5.CONTRASTE DE HIPÓTESIS

Los vehículos híbridos son más caros que los Diésel



20/05/2022



	Híbridos		
	Precio	Kilómetros	Potencia
mean	20228	51100	146
75%	23990	80325	184
max	40795	320000	231

	Diésel		
	Precio	Kilómetros	Potencia
mean	12304	135113	131
75%	17850	195000	150
max	41300	377000	235



Combustible	Precio
Diésel	18687.17915
Eléctrico	22558.33259
Gasolina	14941.81879
Híbrido	20891.95042
Híbrido enchufable	28891.64894

Test Chi2



Todas las variables categóricas son dependientes entre ellas

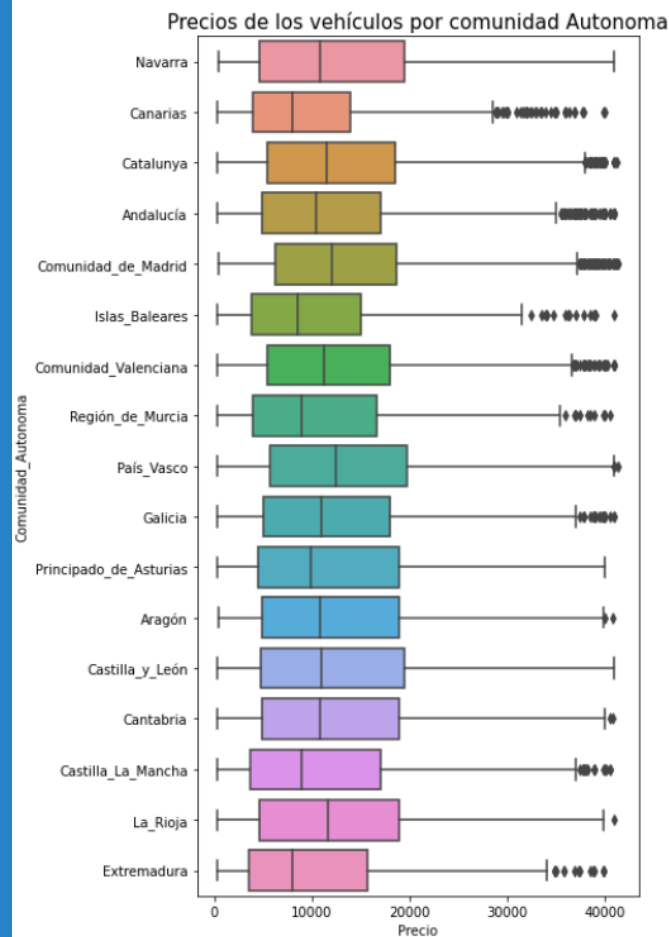
Teorema central del límite+ Anova



	Diésel	Gasolina	Híbrido	Eléctrico	Híbrido enchufable
Diésel	X	X	X	X	X
Gasolina	Distintas	X	X	X	X
Híbrido	Distintas	Distintas	X	X	X
Eléctrico	Distintas	Distintas	Distintas	X	X
Híbrido enchufable	Distintas	Distintas	Distintas	Distintas	X

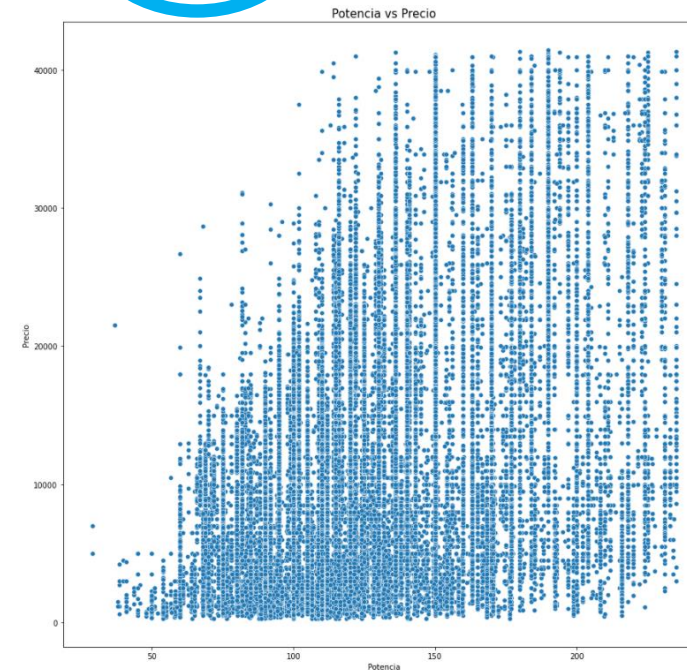
5.CONTRASTE DE HIPÓTESIS

La media de precios es la misma para todas las comunidades autónoma



Comunidad_Autónoma	Precio
Andalucía	11940.1738
Aragón	12633.89127
Canarias	9773.995255
Cantabria	13001.93738
Castilla-La Mancha	11449.61492
Castilla y León	13115.8468
Cataluña	12904.156
Comunidad Valenciana	12661.34199
Comunidad de Madrid	13317.5032
Extremadura	10350.59834
Galicia	12559.05039
Islas Baleares	10345.34342
La Rioja	13048.34783
Navarra	13040.53322
País Vasco	13730.49328
Principado de Asturias	12292.11023
Región de Murcia	11110.58568

Los vehículos más potentes son los más caros





RESUMEN HIPÓTESIS

Dado el contexto del que hemos hablado en la diapositiva anterior, definimos las hipótesis:

- Los vehículos con más km son los más baratos



- La marca de coches más barata es Dacia



- Los vehículos híbridos son más caros que los Diésel



- La media de precios es la misma para todas las comunidades autónomas



- Los vehículos vendidos por profesionales son más caros que los vendidos por particulares



- Los vehículos más potentes son los más caros



¡MUCHAS GRACIAS!

¿PREGUNTAS?

