

AUTOR: EMANUEL LINCIC

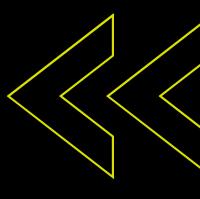


Data Analitycs Segunda Pre-entrega

Profesor: Rubén Alfredo Parente Aun

Tutor: Juan Carlos Borges Pons

Comisión: 53515





DESCRIPCIÓN DE LA TEMÁTICA

No solo en Argentina, sino que en casi todo el mundo es indispensable tener un medio de transporte, ya sea personal o público, y a la hora de querer comprar uno siempre conviene saber los distintos precios de los vehículos que te interesan.

HIPÓTESIS DEL TRABAJO

En nuestra base de datos podremos observar la variación del valor de los distintos autos por las dependiendo de las característica de cada uno.

Por ejemplo, dependiendo el motor y cuantos kilómetros tenga, el valor puede llegar a ser muy distinto entre dos modelos, incluso el color puede cambiar el precio entre dos modelos exactamente iguales.





OBJETIVO DEL PROYECTO

El objetivo principal es analizar los precios entre los diversos modelos de automóviles y sus características únicas del año 2022.

Alcance

Para este proyecto, se contará con vehículos desde el año 1990 y 2022.

USUARIO FINAL Y NIVEL DE APLICACIÓN DE ANALISIS

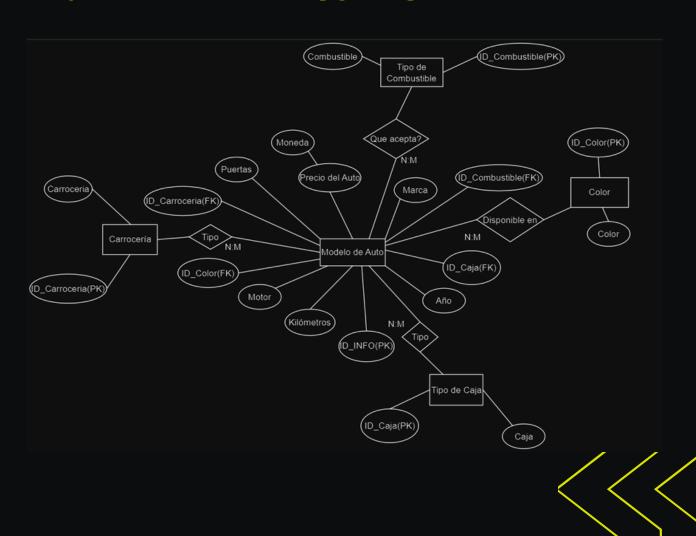
Se busca brindar al usuario la capacidad de realizar una comparación de precios entre los diversos modelos de automóviles. Esto permitirá no solo entender el incremento en el costo de cada vehículo (hasta la fecha), sino también analizar cómo la inflación ha impactado en el mercado automotriz. Al proporcionar esta información detallada y actualizada, el usuario podrá tomar decisiones informadas al momento de comprar o vender un auto, asegurándose de obtener el mejor valor posible en un contexto económico en constante cambio como lo es la Argentina.



CONTENIDO EN LA PRESENTACIÓN

A continuación, se encontrará las tablas y el diagrama entidadrelación iniciales, que quiere decir esto, que para la nueva entrega se hicieron cambios, los cuales se verán detallados.

DIAGRAMA ENTIDAD-RELACION INICIAL





VISTA ORIGINAL DEL DATASET

	Α	В	С	D	Е	F	G	Н	I	J	K	L
1	money	brand	model	year	color	fuel_type	door	gear	motor	body_type	kilometres	currency
2	10350000	Toyota	Corolla Cross	2022	Plateado	Nafta	5.0	Automática		SUV	500	pesos
3	10850000	Jeep	Compass	2022	Blanco	Nafta	5.0	Automática	2.4	SUV	500	pesos
4	35500	Jeep	Compass	2022	Gris oscuro	Nafta	5.0	Automática	2.4	SUV	500	dólares
5	19000	Citroën	C4 Cactus	2022	Gris oscuro	Nafta	5.0	Automática		SUV	550	dólares
6	5800000	Toyota	Corolla	2019	Gris	Nafta	4.0	Manual	1.8	Sedán	9000	pesos
7	34500	Jeep	Compass	2022	Negro	Nafta	5.0	Automática	1.3	SUV	10500	dólares
8	25000	Kia	Sorento	2014	Negro	Diésel	5.0	Automática	2.2	SUV	156000	dólares
9	12300	Mercedes-Benz	Clase C	2008	Plateado	Nafta	4.0	Automática	1.8	Sedán	174000	dólares
10	35500	BMW	220	2016	Blanco	Nafta	2.0	Automática	TURBO	Coupé	66000	dólares
11	6800000	Volkswagen	T-Cross	2019	Naranja	Nafta	5.0	Automática	1.6 MSI Nafta	SUV	35000	pesos
12	38500	Mercedes-Benz	Clase A	2017	Plateado	Nafta	5.0	Automática	3	Hatchback	55394	dólares
13	20500	Mercedes-Benz	Clase A	2013	Blanco	Nafta	5.0	Manual	1.6	Hatchback	162080	dólares
14	3500000	Renault	Sandero	2022	Plateado	Nafta	5.0	Manual		Sedán	111111	pesos
15	5990000	Volkswagen	Polo	2022	Blanco	Nafta	5.0	Manual	1.6	Hatchback	111111	pesos
16	41900	BMW	Serie 1	2020	Negro	Nafta	5.0	Automática		Hatchback	34800	dólares
17	2100000	Ford	Ecosport	2011	Beige	Nafta	4.0	Manual	1.6	SUV	101000	pesos
18	38500	Chevrolet	Trailblazer	2019	Gris	Diésel	5.0	Automática		SUV	73000	dólares
19	2940000	Toyota	Etios	2015	Blanco	Nafta	5.0	Manual	1.5	Hatchback	83000	pesos
20	2050000		Clio	2014	Gris	Nafta	5.0	Manual	1.2	Hatchback	99000	pesos
21	4120000	Toyota	RAV4	2011	Gris	Nafta	5.0	Automática	2.4	SUV	180000	pesos
22	4100000	Toyota	Etios	2018	Gris	Nafta	5.0	Manual	1.5	Sedán	80000	pesos
23	15000		Cooper	2013	Rojo	Nafta	5.0	Manual	1,6	Hatchback		dólares
24	5099000		Renegade	2018		Nafta	5.0	Manual	1.8	SUV		pesos
25	8300000	Toyota	Hilux	2017	Rojo	Diésel	4.0	Manual	2.8	Pick-Up	170000	pesos
28	2800000		Bravo		Negro	Nafta	5.0	Manual		Hatchback		pesos
27	4500000	Ford	Ecosport	2019	Gris	Nafta	5.0	Manual	1.5	SUV	61000	pesos
28	2400000	Chevrolet	Classic	2014	Gris	Nafta	5.0	Manual	1.6	Hatchback		pesos
29	5200000	Renault	Master	2013	Blanco	Diésel	3.0	Manual	2.3	Hatchback	180000	pesos
30	7500000	Ford	Ranger	2021	Blanco	Diésel	4.0	Manual	2.2	Pick-Up		pesos
31	2450000	Peugeot	408	2013	Gris oscuro	Nafta	4.0	Manual	2	Sedán	95000	pesos
32	5200000	Honda	CR-V	2013	Blanco	Nafta	5.0	Automática	2	SUV	186700	pesos
33	10200000		Hilux		Negro	Diésel	4.0	Manual		Pick-Up		pesos
34	2450000		Journey		Blanco	Nafta	5.0	Automática	2.4	Crossover		pesos
35	2585000		Siena	2017		Nafta	5.0	Manual		Sedán		pesos
38	3150000		Fiesta	2016		Nafta	5.0	Manual		Hatchback		pesos
37	4300000	Hyundai	Tucson		Negro	Nafta	5.0	Manual	2	SUV		pesos
38	7589900		HR-V	2019	-	Nafta	5.0	Automática	1.8	SUV		pesos
39	1950000		Linea		Plateado	Nafta	4.0	Manual	1.8	Sedán	147000	
40	3050000		Onix	2016		Nafta	5.0	Manual		Hatchback		pesos
41	3150000		408		Blanco	Nafta	4.0	Manual		Sedán		pesos
42	9389900		Equinox		Plateado	Nafta	5.0	Automática		SUV		pesos
43	3290000		308		Blanco	Nafta	5.0	Automática		Hatchback		pesos
44	14900		A1	2011	Blanco	Nafta	3.0	Automática		Hatchback		dólares
45	3150000		Aircross		Plateado	Nafta/GNC	5.0	Automática		Hatchback		pesos
48	3100000		408	2015		Nafta	4.0	Manual	20	Sedán		nesos
70	STORAGE	e eadleth	400	2015	CHIN	- Common		Marina		- WAINIII	. anna	THE ALEX

El dataset fue adquirido luego de una búsqueda por la página de kaggle



VISTA FINAL DEL DATASET

	A	В	С	D	E	F		G	Н	1	J	К		L	М
1	ID_INFO =	Precio_USD =	Precio_ARS =	Marca =	Modelo [™]	ID_Color	₹	Año ₹	Kilometros =	ID_Combustible =	Puertas =	ID_Caja	マ	Motor =	ID_Carroceria =
2	422	\$6,774,69	\$1.200.000,00	Chevrolet	Corsa	#C0C0C0		2008	150000	1	3.0		2	1.6	4
3	289	\$8.000,00	\$1.417.040,00	Volkswagen	Gol Trend	#808080		2013	130000	1	5.0		2	1.6	4
4	109	\$8.750,64	\$1.550.000,00	Volkswagen	Gol	#000000		2012	102000	1	5.0		2	1.4	4
5	438	\$8.900,00	\$1.576.457,00	Volvo	C30	#000000		2008	173000	1	5.0		2	2.0	4
6	110	\$9.004,69	\$1.595.000,00	Renault	Clio	#FFFFFF		2014	85000	1	3.0		2	1.2	4
7	101	\$9.089,37	\$1.610.000,00	Renault	Sandero	#000000		2013	83000	1	5.0		2	1.6	4
8	144	\$9.315,19	\$1.650.000,00	Chevrolet	Aveo	#FFFFFF		2011	153000	1	4.0		2	1.6	2
9	105	\$9.315,19	\$1.650.000,00	Volkswagen	Upl	#FFFFFF		2018	79000	1	5.0		2	1.0	4
10	296	\$9.541,02	\$1,690,000,00	Chevrolet	Celta	#C0C0C0		2013	109000	1	3.0		2	1.4	4
11	444	\$9.586,18	\$1.698.000,00	Ford	Ka	#808080		2011	97000	1	3.0		2	1.6	4
12	59	\$9.597,47	\$1,700,000,00	Citroën	C3	#0000FF		2010	141000	1	5.0		2	1.6	4
13	187	\$9.879,75	\$1.750.000,00	Ford	Ka	#808080		2013	70000	1	3.0		2	1.0	4
14	208	\$9.879,75	\$1,750,000,00	Volkswagen	Gol Trend	#C0C0C0		2009	119000	1	5.0		2	1.6	4
15	279	\$9.936,21	\$1,760,000,00	Fiat	Uno	#FF0000		2010	128000	1	5.0		2	1.4	4
16	441	\$10.000,00	\$1.771.300,00	Volvo	960	#808080		1995	125000	1	5.0		2	3.0	8
17	148	\$10,444,31	\$1.850.000,00	Chevrolet	Celta	#FF0000		2012	85000	1	3.0		2	1.4	4
18	472	\$10.500,76	\$1.860.000,00	Renault	Logan	#000000		2014	86000	1	4.0		2	1.6	2
19	251	\$10.599,00	\$1.877.400,87	Mercedes-Benz	Clase C	#C0C0C0		2008	250000	1	4.0		1	2.2	2
20	131	\$10.670,13	\$1,890,000,00	Renault	Sandero	#FF0000		2014	107000	1	5.0		2	1.6	4
21	348	\$10.698,36	\$1.895.000,00	Suzuki	Swift	#808080		2009	190000	1	5.0		2	1.5	4
22	213	\$10,726,58	\$1,900,000,00	Chevrolet	Celta	#FF0000		2012	101000	1	3.0		2	1.4	4
23	219	\$10.726,58	\$1,900,000,00	Fiat	Palio	#808080		2011	190000	1	5.0		2	1.6	8
24	38	\$11.008,86	\$1,950,000,00	Fiat	Linea	#C0C0C0		2014	147000	1	4.0		2	1.8	2
25	468	\$11.008,86	\$1,950,000,00	Peugeot	504	#FF0000		1996	99000	2	4.0		2	2.3	2
26	368	\$11.008,86	\$1.950.000,00	Renault	Clio	#000000		2012	122000	1	5.0		2	1.2	4
27	102	\$11.065,32	\$1,960,000,00	Volkswagen	Gol Trend	#C0C0C0		2015	65500	1	5.0		2	1.6	4
28	129	\$11.150,00	\$1.975.000,00	Fiat	Uno	#FFFFFF		2016	87000	1	5.0		2	1.4	4
29	479	\$11.234.69	\$1,990,000,00	Citroën	Berlingo	#C0C0C0		2009	220000	1	5.0		2	1.4	9
30	138	\$11,293,96	\$2,000,500,00	Ford	Focus II	#808080		2009	160000	1	4.0		2	2.0	2
31	140	\$11.293.96	\$2,000.500.00	Peugeot	207	#808080		2012	110000	2	5.0		2	1.4	4
32	139	\$11,293,96	\$2,000,500,00	Renault	Clio	#808080		2014	88000	1	5.0		2	1.2	4
33	418	\$11.573.42	\$2.050.000.00	Chevrolet	Meriva	#808080		2012	98000	3	5.0		2	1.8	7
34	19	\$11,573,42	\$2,050,000,00	Renault	Clio	#808080		2014	99000	1	5.0		2	1.2	4
35	415	\$11.573.42	\$2.050.000.00	Volkswagen	Gol	#808080		2014	43000	1	5.0		2	1.4	4
38	380	\$11,799.24			Bora	#000000		2011	168200	1	4.0		2	2.0	2
37	16	\$11.855,70				#F5F5DC		2011	101000	1	4.0		2	1.6	1
38	104	\$11.855,70				#C0C0C0		2015	54500	1	4.0		2	2.0	2
39	90	\$11.858,52			Focus II	#FFD700		2009	133000	1	5.0		2	2.0	4
40	107	\$12,137,98				#C0C0C0		2014	57000	1	4.0		2	2.0	2
41	171	\$12,183,14				#C0C0C0		2014	63000	3	4.0		2	1.4	2
42	434	\$12,222,66	\$2,165,000,00			#000000		2014	69000	3	5.0		2	1.4	4
43	8	\$12,300.00		Mercedes-Benz		#C0C0C0		2008	174000	1	4.0		1	1.8	2
44	173	\$12.363.24				#FFFFFF		2014	79000	1	5.0		2	1.6	4
45	336	\$12,363,24				#808080		2016	89100	1	5.0		2	1.2	4
48	245	\$12,000,24	\$2,700,000,00			#000000		2010	137000	- 1	5.0		2	1.4	4
		Precio de Cocl				Combustibl	e =		ja ▼ Carro	ceria *					





TRANSFORMACIONES PARA LA SEGUNDA ENTREGA

Para la segunda entrega se rellenó espacios nulos buscando en internet el dato faltante utilizando la información que se tenia, también se movieron columnas y se modificó el formato de algunas, como por ejemplo en la columna precio se agregó el formato de moneda.





LISTADO DE COLUMNAS DE TABLAS INICIAL

PRECIO DE COCHES EN ARGENTINA

TIPO DE CLAVE	CAMPO	TIPO DE DATO
РК	ID_INFO	INT
-	Moneda	text(200)
-	Precio	varchar(200)
-	Marca	text(200)
-	Modelo	varchar(200)
FK	ID_Color	varchar(200)
-	Año	INT
-	Kilómetros	INT
FK	ID_Combustible	INT
-	Puertas	varchar(200)
FK	ID_Caja	INT
-	Motor	varchar(200)
FK	ID_Carrocería	INT



COLOR

TIPO DE CLAVE	САМРО	TIPO DE DATO
PK	ID_Color	varchar(200)
-	Color	text(200)

TIPO DE COMBUSTIBLE

TIPO DE CLAVE	САМРО	TIPO DE DATO
PK	ID_Combustible	INT
-	Combustible	varchar(200)



TIPO DE CAJA

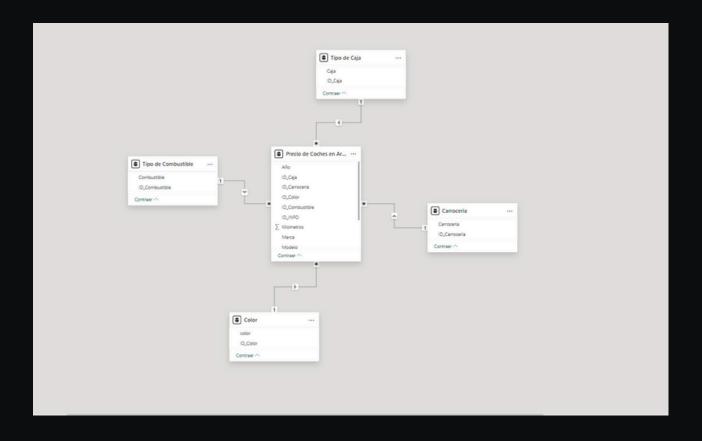
TIPO DE CLAVE	CAMPO	TIPO DE DATO
РК	ID_Caja	INT
-	Caja	text(200)

CARROCERÍA

TIPO DE CLAVE	САМРО	TIPO DE DATO
PK	ID_Carroceria	INT
-	Carroceria	text(200)



DIAGRAMA ENTIDAD-RELACION INICIAL



Para diagrama relacional generado automáticamente por Power BI, no se tuvo que realizar cambios, ni hacer tablas puente.



DIAGRAMA ENTIDAD-RELACION FINAL

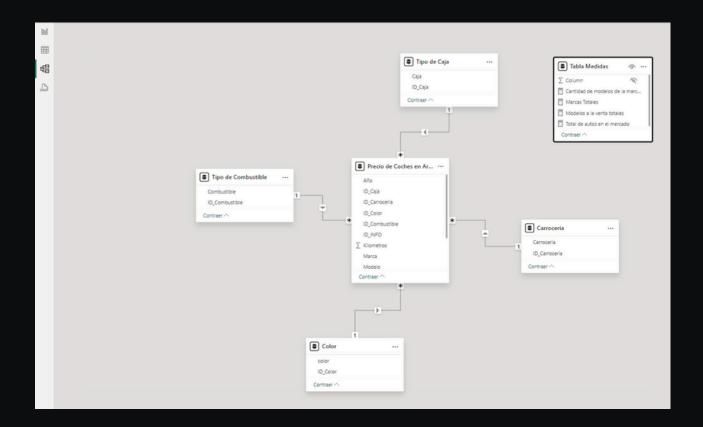
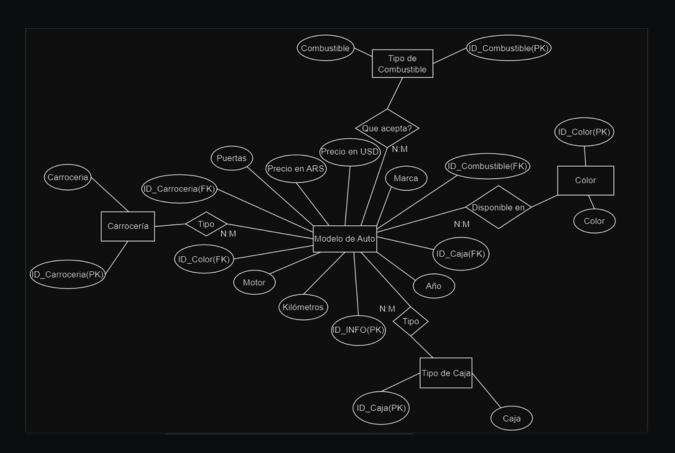








DIAGRAMA ENTIDAD-RELACION FINAL







LISTADO DE TABLAS FINAL

TIPO DE CLAVE	CAMPO	TIPO DE DATO
РК	ID_INFO	INT
-	Precio_USD	varchar(200)
-	Precio_ARS	varchar(200)
-	Marca	text(200)
-	Modelo	varchar(200)
FK	ID_Color	varchar(200)
-	Año	INT
-	Kilómetros	INT
FK	ID_Combustible	INT
-	Puertas	varchar(200)
FK	ID_Caja	INT
-	Motor	varchar(200)
FK	ID_Carrocería	INT



COLOR

TIPO DE CLAVE	CAMPO	TIPO DE DATO
РК	ID_Color	varchar(200)
-	Color	text(200)

TIPO DE COMBUSTIBLE

TIPO DE CLAVE	САМРО	TIPO DE DATO
PK	ID_Combustible	INT
-	Combustible	varchar(200)



TIPO DE CAJA

TIPO DE CLAVE	CAMPO	TIPO DE DATO
РК	ID_Caja	INT
-	Caja	text(200)

CARROCERÍA

TIPO DE CLAVE	САМРО	TIPO DE DATO
PK	ID_Carroceria	INT
-	Carroceria	text(200)



TRANSFORMACIONES EN POWER BI

Para el diseño, se utilizó el mismo que el de esta presentación, con la misma paleta de colores, excepto en las tablas de barras en las que se utilizó distintos colores para diferenciar los valores.

Se crea la tabla "Medidas", a fin de tener datos más concretos y fáciles a la vista

MEDIDAS CALCULADAS Y FORMULAS

Las medidas realizadas en la tabla "Medidas" son las siguientes: Cantidad de modelos de la marca a la venta, Marcas Totales, Modelos a la venta totales, Total de autos en el mercado, Total de Precio_ARS para Marca, Total de Precio_USD para Marca.

 El fin de la medida "Cantidad de modelos de la marca a la venta" es saber la cantidad total de modelos de autos, contando los repetidos y cuando se filtra la marca, y opcionalmente el modelo, saber cuantos hay en el mercado. La fórmula es:

Cantidad de modelos de la marca a la venta = COUNT('Precio de Coches en Argentina'[Modelo])



 El fin de la medida "Marcas Totales", es saber cantidad de marcas de autos listadas. La formula es la siguiente: Marcas Totales = DISTINCTCOUNT('Precio de Coches en

Marcas Totales = DISTINCTCOUNT('Precio de Coches en Argentina'[Marca])

• A diferencia del primero, la medida "Modelos a la venta totales", da a saber la cantidad de distintas marcas que hay. La formula es la siguiente:

Modelos a la venta totales = DISTINCTCOUNT('Precio de Coches en Argentina'[Modelo])

 La medida "Total de autos en el mercado" nos informa de la cantidad de vehículos listados en el mercado. La formula es la siguiente:

Total de autos en el mercado = COUNT('Precio de Coches en Argentina'[Marca])

 La medida "Total de Precio_ARS para Marca" suma el precio de los vehículos filtrados en pesos argentinos. La formula es la siguiente:

Total de Precio_ARS para Marca = CALCULATE(

SUM('Precio de Coches en Argentina'[Precio_ARS]),
ALLSELECTED('Precio de Coches en Argentina'[Marca]))

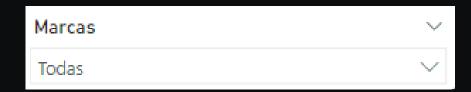


 La medida "Total de Precio_USD para Marca" es igual a la anterior, a diferencia que esta está en dólares estadounidenses. La formula es la siguiente:

Total de Precio_USD para Marca = CALCULATE(

SUM('Precio de Coches en Argentina'[Precio_USD]), ALLSELECTED('Precio de Coches en Argentina'[Marca]))

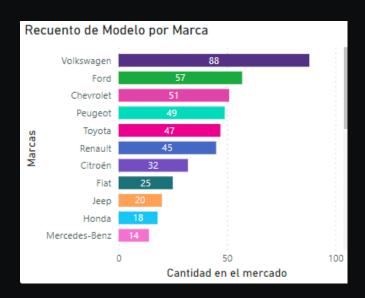
SEGMENTACIONES ELEGIDAS



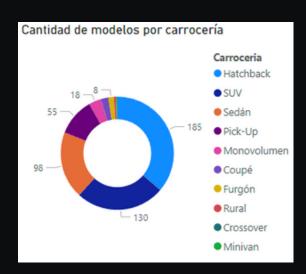
Se filtra por las marcas que se desee saber información.





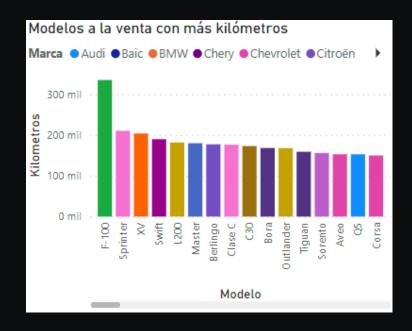


En esta tabla se podrá observar la cantidad de vehículos por marca hay listadas.



Acá podemos ver la cantidad de vehículos que hay por su tipo de carrocería.





Se puede observar los modelos con más kilómetros a la venta.



Estos son los indicadores de algunos de los indicadores en los que se colocaron las medidas realizadas.





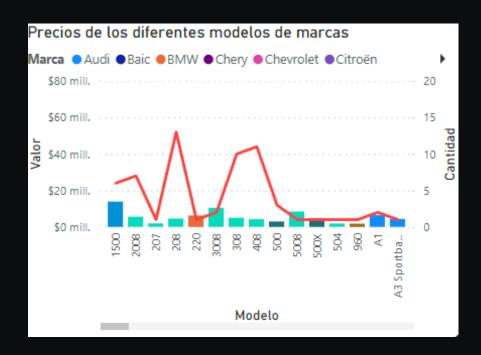
Modelo/s	~
Todas	~

Este filtro se utiliza en conjunto con el de marcas para buscar los modelos de las marcas filtradas.

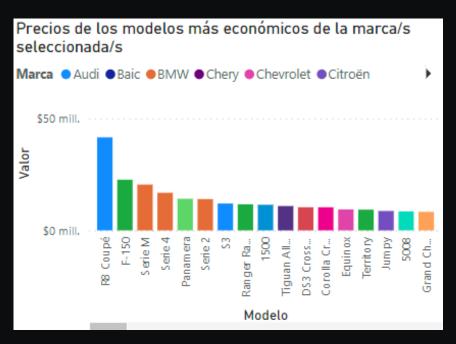
Precios		· ·
\$1.200.000	\$76.165.900	
\bigcirc		$\overline{}$
)		0
Kilometros		
500	335000	
\bigcirc		
)		O
Año		
1995	2022	
\bigcirc		
		\circ

Estos filtros van en conjunto con los anteriores, o sea se van modificando con respecto a los demás filtros. Y podemos filtrar entre valores según la preferencia de búsqueda que tengamos. (El filtro de precios esta en moneda Argentina)





Se puede observar los precios de los modelos más caros de cada serie y, adicionalmente, se puede ver la cantidad que hay de estos modelos.









A diferencia del anterior, este muestra el modelo más económico que hay entre series.

Para entender mejor ambos, por ejemplo si nosotros tenemos cinco modelos de Volkswagen Amarok la primer tabla nos mostrará el precio del modelo más caro y la segunda el modelo más barato, esto nos sirve para saber, en conjunto de cuantos modelos hay, entre que precios estarán los vehículos que deseemos averiguar.



510

Cantidad de modelos de la/s marca/s a la venta

Este indicador nos muestra la cantidad de modelos por marca e irá cambiando en relación con los filtros de "Marca" y "Modelo" que apliquemos.

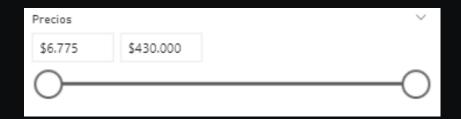
\$3 mil M
Total de Precio_ARS para Marca

Podremos ver la suma total de los precios (en pesos argentinos) entre las marcas y/o modelos seleccionados.

\$15 mill.
Total de Precio_USD para Marca

Este indicador es igual al anterior solo que suma los precios que están en dólares.



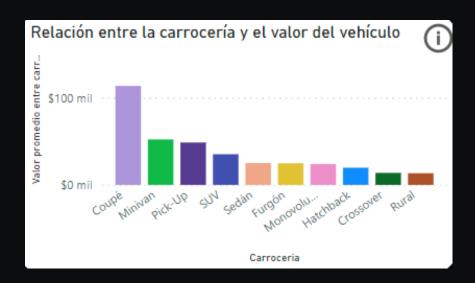


Este filtro entre valores es igual que el filtro nombrado anteriormente, a diferencia que este está en dólares

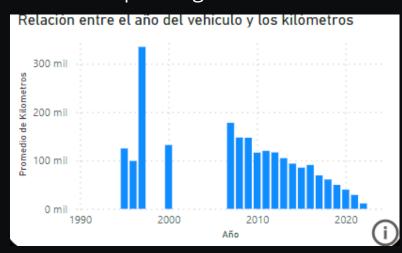


En esta tabla podremos observar que generalmente el valor del vehículo es mayor cuando es de caja automática y menor cuando es de caja manual.





En esta visualización se puede observar como aumenta o disminuye el valor del vehículo dependiendo del tipo de carrocería que tenga.



En esta visualización se puede observar que mientras más viejos sean los vehículos más cantidad de kilómetros recorridos tiene.





En esta visualización se puede observar como aumenta o disminuye el valor del vehículo dependiendo de los kilómetros recorridos del mismo.

ANEXO

Como vimos en esta presentación, hay dos filtros iguales que los diferencia el tipo de moneda, ¿Porqué se da esto? Esto es porque, por ejemplo el filtro que esta en dólares, se utiliza si el usuario está más interesado en averiguar los vehículos con su valor en dólares por si dispone de esta moneda o le interesa más para una búsqueda internacional.

