

Análisis de estadístico

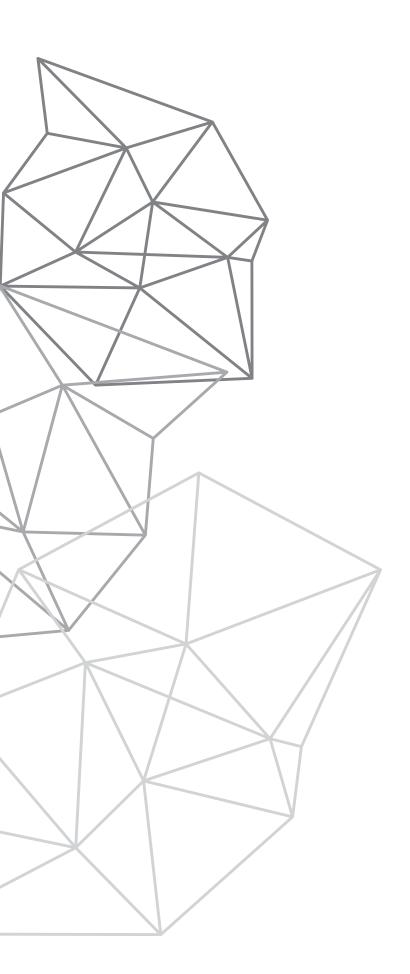
Los factores de impacto y su relación con el precio de venta de un automóvil

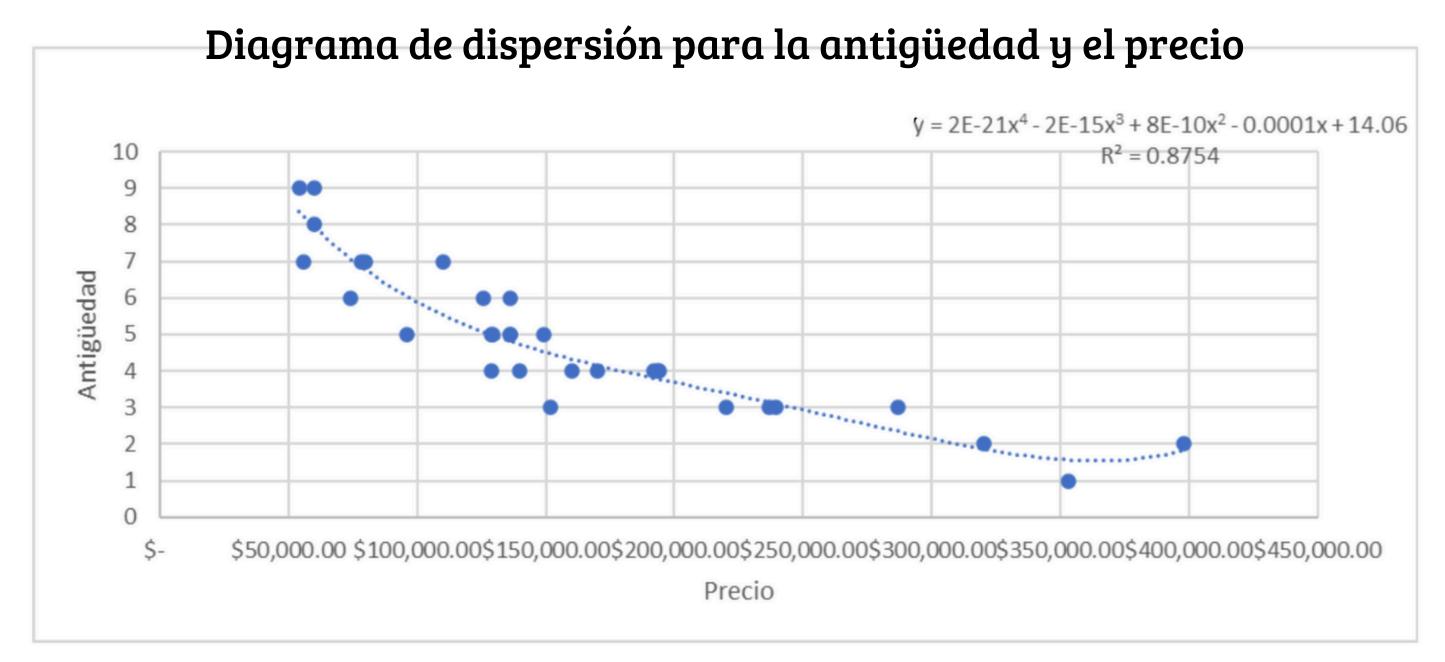
Jaime Nicolás González Lucio - A0127818 Omar Emilio Casillas Alday - A01712114 Ricardo Ruiz Cano - A01707233 Ángel David Candelario Rolon - A01712387 Jacob Espejel Guiza - A01712021



Una agencia de ventas de autos seminuevos desea determinar el precio de un automóvil de la marca Camaro.

Se recopilaron datos de 31 vehículos, antigüedad, kilometraje (en miles), condición (mala, promedio o excelente) y si el auto ha tenido uno o más dueños. Estos datos permitirán realizar un análisis preciso para estimar el precio de venta de los Camaros basándose en sus características específicas.

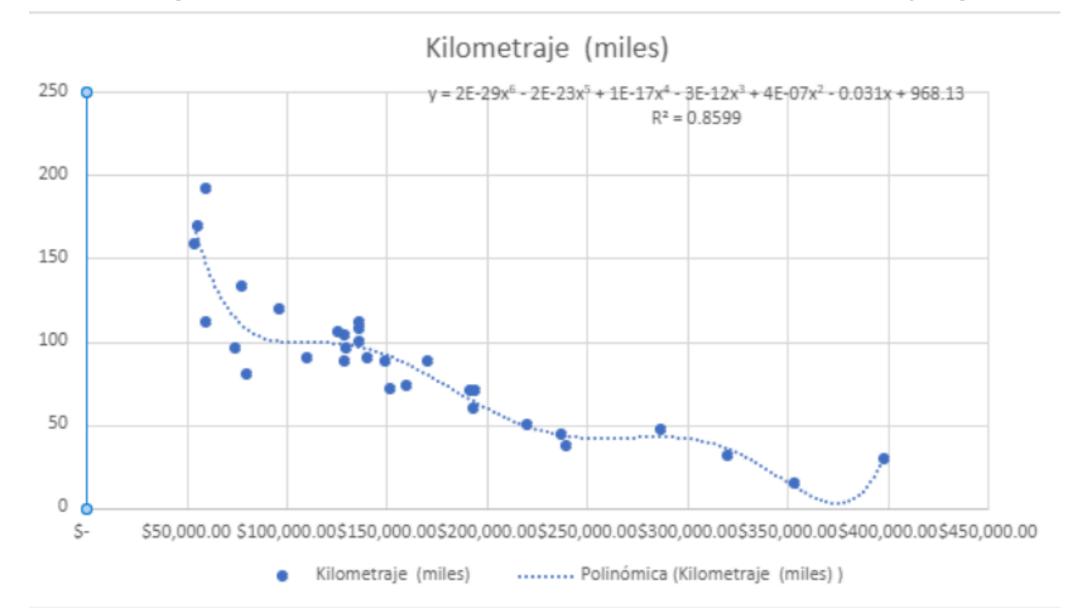




Optamos por una línea de tendencia polinomial de grado 4 debido a su alta correlación (R = -0.93562, $R^2 = 0.8754$) con la gráfica, indicando una fuerza de correlación muy alta. Aunque un modelo lineal también mostró una alta correlación (R = -0.8724, $R^2 = 0.7611$), la línea polinomial se ajusta mejor a los datos, con menos error.

Diagrama de dispersión para la Kilometraje y el precio





Optamos por una línea de tendencia polinomial de grado 4 debido a su alta correlación (R = -0.93562, $R^2 = 0.8599$) con la gráfica, indicando una fuerza de correlación muy alta. Aunque un modelo lineal también mostró una alta correlación (R = -0.8724, $R^2 = 0.7611$), la línea polinomial se ajusta mejor a los datos, con menos error.

Modelo de Regresión Lineal Múltiple para Estimar el Precio de Venta según Antigüedad y Kilometraje

y=350436.27-22743.71*Antiguedad-899.68*Kilometraje

e la regresión							
0.90066735							
0.81120168							
0.79771609							
39401.3822							
31							
'ARIANZA							
ados de liberto	ma de cuadrad	dio de los cuad	F	alor crítico de	F		
2	1.8677E+11	9.3386E+10	60.1532037	7.3109E-11			
28	4.3469E+10	1552468917					
30	2.3024E+11						
Coeficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	Superior 95.0%
350436.273	18850.6934	18.5900999			•	-	389050.168
-22743.7126	6710.73549	-3.38915349	0.00210044	-36490.0311	-8997.39408	-36490.0311	-8997.39408
-899.680781	330.212847	-2.72454809	0.01096916	-1576.09114	-223.270426	-1576.09114	-223.270426
	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de liberto 2 28 30 Coeficientes 350436.273 -22743.7126	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de libertana de cuadrad 2 1.8677E+11 28 4.3469E+10 30 2.3024E+11 Coeficientes Error típico 350436.273 18850.6934 -22743.7126 6710.73549	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de libertana de cuadraadio de los cuad 2 1.8677E+11 9.3386E+10 28 4.3469E+10 1552468917 30 2.3024E+11 Coeficientes Error típico Estadístico t 350436.273 18850.6934 18.5900999 -22743.7126 6710.73549 -3.38915349	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de liberta de cuadra dio de los cuac 2 1.8677E+11 9.3386E+10 60.1532037 28 4.3469E+10 1552468917 30 2.3024E+11 Coeficientes Error típico Estadístico t Probabilidad 350436.273 18850.6934 18.5900999 2.7345E-17 -22743.7126 6710.73549 -3.38915349 0.00210044	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de libertana de cuadraa dio de los cuac 2 1.8677E+11 9.3386E+10 60.1532037 7.3109E-11 28 4.3469E+10 1552468917 30 2.3024E+11 Coeficientes Error típico Estadístico t Probabilidad Inferior 95% 350436.273 18850.6934 18.5900999 2.7345E-17 311822.378 -22743.7126 6710.73549 -3.38915349 0.00210044 -36490.0311	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de libertana de cuadraa dio de los cuac	0.90066735 0.81120168 0.79771609 39401.3822 31 ARIANZA ados de libertana de cuadraddio de los cuad 2 1.8677E+11 9.3386E+10 60.1532037 7.3109E-11 28 4.3469E+10 1552468917 30 2.3024E+11 Coeficientes Error típico Estadístico t Probabilidad Inferior 95% Superior 95% Inferior 95.0% 350436.273 18850.6934 18.5900999 2.7345E-17 311822.378 389050.168 311822.378 -22743.7126 6710.73549 -3.38915349 0.00210044 -36490.0311 -8997.39408 -36490.0311

Para entender la relación entre precio de venta, antigüedad y kilometraje de vehículos: revisamos datos para errores y atípicos, usamos regresión lineal múltiple ajustando el modelo, solucionamos problemas de multicolinealidad y autocorrelación, ampliamos la muestra si posible y validamos con cruzada para asegurar fiabilidad

Análisis de Regresión Lineal Múltiple con Todas las Variables Disponibles

		ANÁLISIS DE VARIANZA								
		G	irados de liberto	ma de cuadrad	dio de los cuac	F	alor crítico de l	F		
Estadís Área de trazado si	ón	Regresión	5	2.0639E+11	4.1279E+10	43.275785	1.6162E-11			
	0.04670000	Residuos	25	2.3846E+10	953857459					
Coeficiente de correlación múltiple	0.94679903	Total	30	2.3024E+11						
Coeficiente de determinación R^2	0.89642841									
2000110101110 de determinación 111 2	0.05042041		Coeficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95.0%	uperior 95.0%
R^2 ajustado	0.87571409	Intercepción	297260.419	23696.9335	12.5442568	2.7593E-12	248455.671	346065.167	248455.671	346065.167
Error tímico	20004 5020	Aniguedad	-22663.8375	5489.99021	-4.12821091	0.00035622	-33970.684	-11356.991	-33970.684	-11356.991
Error típico	30884.5829	Kilometraje	-414.37698	294.460273	-1.4072424	0.17166741	-1020.82927	192.075305	-1020.82927	192.075305
Observaciones	31	Condicion	14455.9118	15933.7521	0.90725095	0.37293656	-18360.265	47272.0886	-18360.265	47272.0886
Observaciones	31	Excelente	65633.8673	22231.9657	2.95222961	0.00677051	19846.2769	111421.458	19846.2769	111421.458
		Unico Dueño	-15522.5204	18265.8568	-0.84981069	0.4034956	-53141.7566	22096.7158	-53141.7566	22096.7158

Al analizar los datos, empleamos regresión múltiple considerando el precio de venta (Y) y características como antigüedad, kilometraje, condición, y único dueño (X). El coeficiente de determinación R2R^2R2 de 0.8931 indica que el 89.31% de la variabilidad en el precio puede explicarse por estas variables, sugiriendo un modelo robusto.

	-1129 59	0001.04.11	
AÑO	TRIMESTRE 1	TRIMESTRE 2	TRIMESTRE 3
1	3	5	10
2	4	7	12
3	6	8	14
4	7.33333	9.66666	16

Pronostico

El análisis de los datos sugiere que el modelo lineal es más adecuado para el proyecto, mostrando un error promedio menor (menos de 0.03) en comparación con el método de promedio móvil. Debido a la cantidad limitada de datos, el ajuste lineal resulta más preciso que el promedio móvil, el cual es más efectivo en conjuntos de datos más grandes

AÑO	TRIMESTRE 1	TRIMESTRE 2	TRIME
1	3		
2	4		
3	6	3.5	6
4	?	5	6.25

AÑO	TRIMESTRE 1	TRIMESTRE 2	TRIME
1	3	2.8333	0.2777
2	4	4.3333	0.1111
3	6	5.8888	0.2777
4	?	7.3333	

AÑO	TRIMESTRE 2	PRONOSTICO PROMEDIO MOVIL	ERROR DE PRONOSTICO AL CUADRADO
1	5		
2	8		
3	8	6	8
4	?	7.5	4

AÑO	TRIMESTRE 3	PRONOSTICO PROMEDIO MOVIL	ERROR DE PRONOSTICO AL CUADRADO
1	10		
2	12		
3	14	11	14
4	?	13	9

PROMEDIO ERROR: 6

PROMEDIO ERROR: 11.5

AÑO	TRIMESTRE 2	PRONOSTICO PROMEDIO MOVIL	ERROR DE PRONOSTICO AL CUADRADO
1	5	5.1666	0.0277
2	7	6.6666	0.1111
3	8	8.1666	0.0277
4	?	9.6666	

AÑO	TRIMESTRE 3	PRONOSTICO PROMEDIO MOVIL	ERROR DE PRONOSTICO AL CUADRADO
1	10	10	0
2	12	12	0
3	14	14	0
4	?	16	

PROMEDIO ERROR: 0.02777

PROMEDIO ERROR: 0