



Tecnológico de Monterrey

Unidad de formación:

Gestión de proyectos de plataformas tecnológicas (Gpo 201)

Actividad 2 (Regresión no lineal)

Profesor:

Alfredo García Suarez

Alumna:

Samantha Aja | A01738365

Fecha de entrega:

3 de octubre de 2025

Partí de la base de datos con información de listados de Airbnb de Ciudad de México. Mi objetivo fue evaluar la relación entre variables dadas por el ejercicio `host_response_rate`, `host_acceptance_rate`, `host_total_listings_count`, `accommodates`, `reviews_per_month` y `Price` y construir modelos de regresión no lineal para explicar mejor el comportamiento de cada objetivo. Con la base limpia calculé la matriz de correlación de Pearson sobre las variables numéricas. A partir de cada variable objetivo identifiqué las tres correlaciones absolutas más altas para priorizar predictores con señal más fuerte y estable. Eso me dio una “lista corta” para modelar.

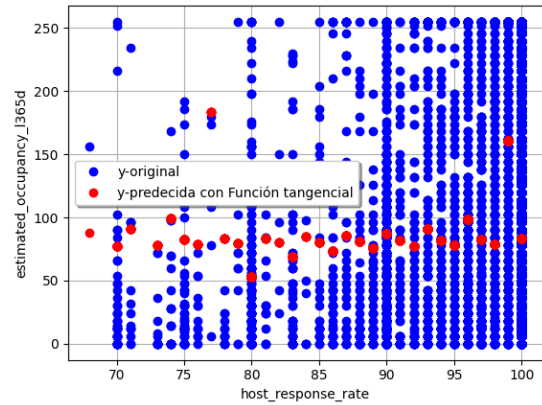
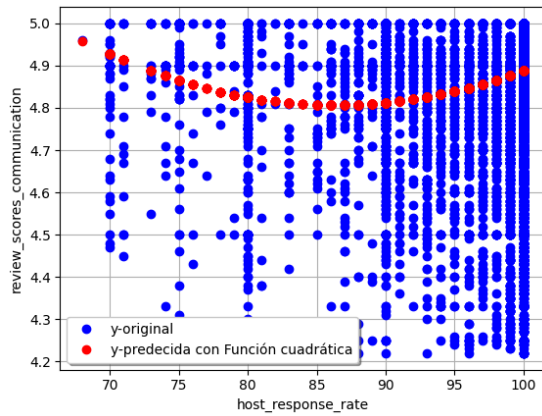
Dando el siguiente resultado:

Top 3 correlaciones para: <code>host_response_rate</code>		Top 3 correlaciones para: <code>accommodates</code>	
	Correlación		Correlación
<code>review_scores_communication</code>	0.093594	<code>bedrooms</code>	0.649348
<code>estimated_occupancy_1365d</code>	0.070185	<code>beds</code>	0.551342
<code>review_scores_cleanliness</code>	0.067495	<code>bathrooms</code>	0.383753
Top 3 correlaciones para: <code>host_acceptance_rate</code>		Top 3 correlaciones para: <code>reviews_per_month</code>	
	Correlación		Correlación
<code>minimum_nights</code>	-0.060021	<code>number_of_reviews_ltm</code>	0.775723
<code>minimum_minimum_nights</code>	-0.057223	<code>estimated_occupancy_1365d</code>	0.751280
<code>maximum_maximum_nights</code>	0.056526	<code>number_of_reviews_ly</code>	0.587285
Top 3 correlaciones para: <code>host_total_listings_count</code>		Top 3 correlaciones para: <code>price</code>	
	Correlación		Correlación
<code>calculated_host_listings_count</code>	0.621283	<code>accommodates</code>	0.361541
<code>calculated_host_listings_count_entire_homes</code>	0.353657	<code>bedrooms</code>	0.354766
<code>review_scores_rating</code>	-0.133247	<code>bathrooms</code>	0.296714

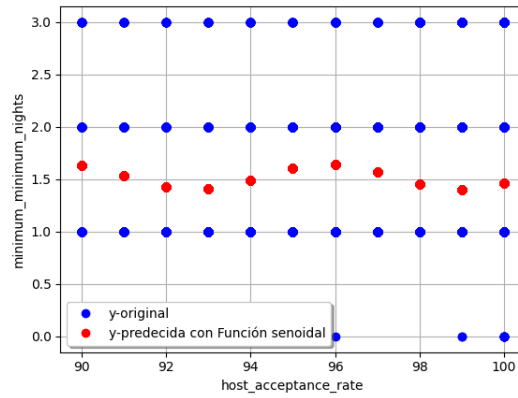
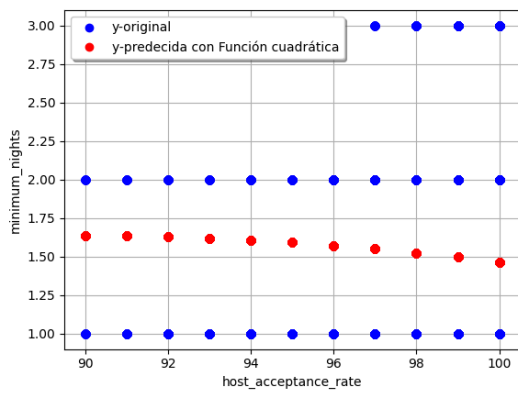
El criterio de aceptación que utilicé consistió en comparar el coeficiente de correlación obtenido de cada modelo de regresión no lineal (R) con el coeficiente de correlación lineal previamente calculado entre las mismas variables. Un modelo se consideró aceptado únicamente cuando su R fue mayor que el valor absoluto de la correlación lineal y al mismo tiempo utilizar dos funciones diferentes para cada variable dependiente como lo pide el ejercicio, ya que esto demuestra que el ajuste no lineal ofrece una mejor capacidad explicativa que la relación lineal simple. Este criterio es importante porque evita aceptar modelos que no aporten mejora real respecto a la correlación básica, garantizando así que los modelos seleccionados realmente capturan patrones adicionales en los datos y representan una ventaja al utilizar técnicas de regresión no lineal.

Una vez hechas las predicciones de cada modelo, se graficó la “y” original con las predicciones, así podemos ver como se grafican diferentes funciones las cuales se ajustarán de mejor manera de acuerdo a los datos de nuestra variable objetivo.

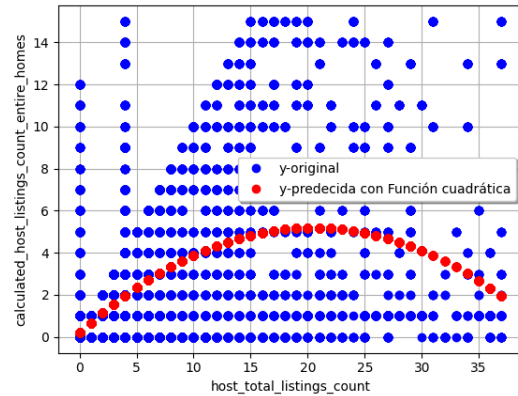
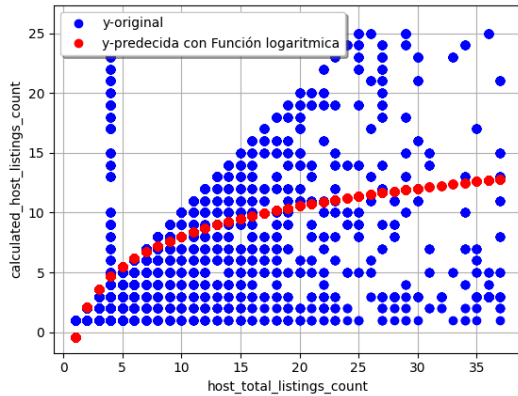
Host response rate:



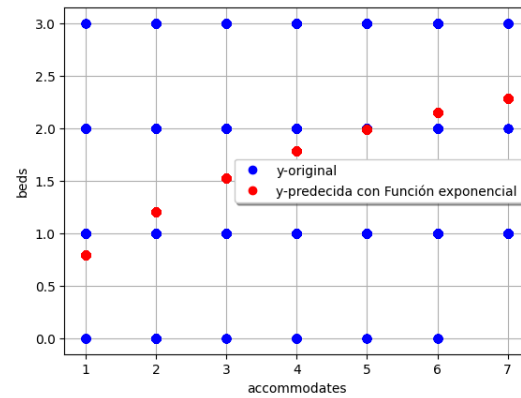
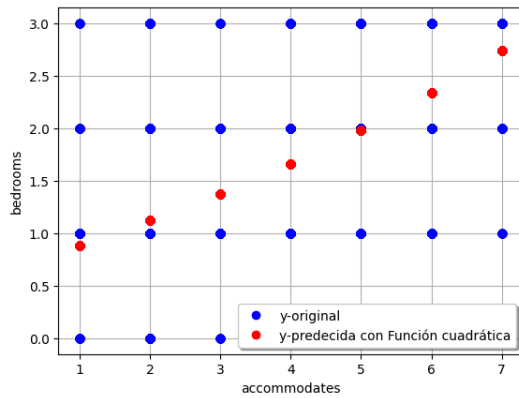
Host acceptance rate:



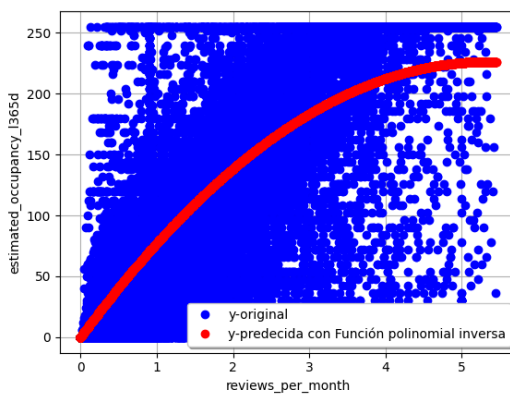
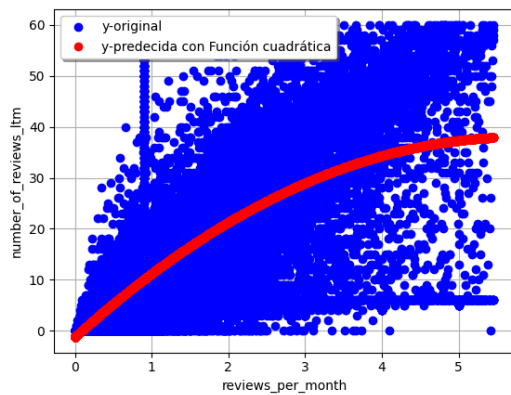
Host total listings count:



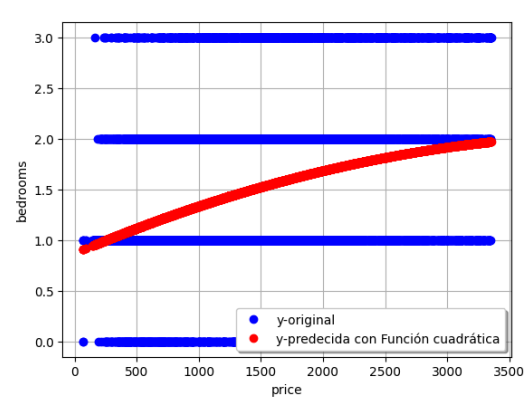
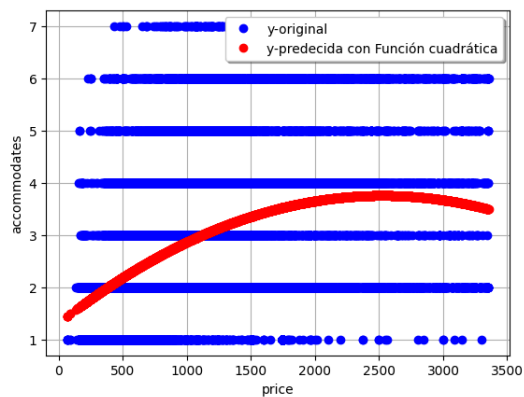
Accommodates:



Reviews per month:



Price:



Por último se hizo una última tabla para comparar los modelos y su correlación linear con la nueva R que se hizo con las regresiones no lineales para ver si es aceptado o no de acuerdo a los lineamientos antes fijados.

Modelo	Relación	Corr_lineal	R	R2	Aceptado
1	host_response_rate vs review_scores_communication	0.094	0.132	0.017	Sí
2	host_response_rate vs estimated_occupancy_l365d	0.070	0.133	0.018	Sí
3	host_acceptance_rate vs minimum_nights	-0.060	0.061	0.004	Sí
4	host_acceptance_rate vs minimum_minimum_nights	-0.057	0.074	0.005	Sí
5	host_total_listings_count vs calculated_host_l...	0.621	0.678	0.459	Sí
6	host_total_listings_count vs calculated_host_l...	0.354	0.454	0.206	Sí
7	accommodates vs bedrooms	0.649	0.653	0.426	Sí
8	accommodates vs beds	0.551	0.565	0.319	Sí
9	reviews_per_month vs number_of_reviews_ltm	0.776	0.789	0.623	Sí
10	reviews_per_month vs estimated_occupancy_l365d	0.751	0.767	0.588	Sí
11	price vs accommodates	0.362	0.391	0.153	Sí
12	price vs bedrooms	0.355	0.358	0.128	Sí

En conclusión, un modelo no lineal resulta superior a uno lineal cuando la relación entre las variables no sigue una tendencia recta, sino que presenta comportamientos curvos, exponenciales o de crecimiento variable. Mientras que los modelos lineales solo pueden representar relaciones proporcionales constantes, los modelos no lineales permiten capturar patrones más complejos y realistas del fenómeno analizado. Esto los hace más precisos para describir y predecir situaciones donde los cambios en una variable no generan efectos uniformes en la otra, logrando así un mejor ajuste a los datos y una explicación más completa de la variabilidad observada.