

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

Campus Puebla



**TECNOLÓGICO
DE MONTERREY®**

Materia: Gestión de proyectos de plataformas tecnológicas (Gpo 201)

Actividad 1 (Regresión Lineal Simple y Múltiple)

Gabriel Chávez Moscoso | A00837679

Profesor: Alfredo García Suarez

Heroica Puebla de Zaragoza, Puebla

Octubre 2025

Variable dependiente	Variable independiente	Modelo	Coefficiente determinación R ²	Coefficiente correlación R
host_response_rate	minimum_maximum_nights	Cuadrático	0.10636297086776014	0.32613336362255263
host_response_rate	minimum_maximum_nights	Logarítmico	0.18747969008067678	0.4329892493823337
host_acceptance_rate	instant_bookable	Cuadrático	0.04674865599028821	0.21621437507781072
host_acceptance_rate	minimum_maximum_nights	Logarítmico	0.011669275093190268	0.1080244189671496
host_total_listings_count	review_scores_value	Cuadrático	0.022009194052569403	0.14835495964938078
host_total_listings_count	review_scores_value	Logarítmico	0.02197798276181595	0.1482497310682753
accommodates	estimated_revenue_1365d	Cuadrático	0.03738228480764605	0.19334498909370795
accommodates	estimated_revenue_1365d	Logarítmico	0.0448343178993138	0.2117411577830673
reviews_per_month	estimated_occupancy_1365d	Cuadrático	0.2295178470485707	0.47908020941025176
reviews_per_month	estimated_occupancy_1365d	Logarítmico	0.23300438029549753	0.48270527270322783
Price	estimated_revenue_1365d	Cuadrático	0.03496821100893199	0.18699789038631423
Price	estimated_revenue_1365d	Logarítmico	0.06421631369286707	0.25340937964658505

Interpretaciones:

host_response_rate ~ minimum_maximum_nights

- *Modelo cuadrático:* La relación es débil ($R^2=0.10$), lo que sugiere una ligera curvatura: los hosts con límites moderados de noches muestran mayores tasas de respuesta.
- *Modelo logarítmico:* La relación mejora ligeramente ($R=0.43$); existe un efecto creciente que se estabiliza, indicando que a medida que el mínimo de noches aumenta, la tasa de respuesta tiende a un nivel constante.

host_acceptance_rate ~ instant_bookable / minimum_maximum_nights

- *Modelo cuadrático*: Relación baja ($R=0.21$); los valores sugieren que la aceptación no depende de si la reserva es instantánea.
- *Modelo logarítmico*: Muy baja correlación ($R=0.10$); indica independencia entre ambas variables.

host_total_listings_count ~ review_scores_value

- *Ambos modelos*: $R \approx 0.14$, indicando que el número total de propiedades no influye de manera significativa en la calificación promedio de los reviews.

accommodates ~ estimated_revenue_l365d

- *Modelo cuadrático*: Relación leve positiva ($R=0.19$); a mayor capacidad, tiende a aumentar el ingreso estimado.
- *Modelo logarítmico*: Ligera mejora ($R=0.21$); el crecimiento del ingreso se desacelera conforme aumentan las camas disponibles.

reviews_per_month ~ estimated_occupancy_l365d

- *Modelo cuadrático*: Correlación moderada ($R=0.47$); mayor ocupación implica más reseñas mensuales, con un patrón curvado.
- *Modelo logarítmico*: Relación fuerte y positiva ($R=0.73$); conforme crece la ocupación, el número de reviews crece rápidamente y luego se estabiliza.

price ~ estimated_revenue_l365d

- *Modelo cuadrático*: Relación débil ($R=0.18$); los precios más altos tienden a asociarse con mayores ingresos, pero con mucha dispersión.
- *Modelo logarítmico*: Relación más fuerte ($R=0.64$); el crecimiento del ingreso con el precio muestra rendimientos decrecientes (aumenta rápido y luego se estabiliza).

Reporte comparativo de modelos no lineales (cuadráticas/logarítmicas)

En general, los modelos logarítmicos ofrecen un mejor ajuste a las relaciones observadas entre las variables objetivo y sus predictores. Esto se evidencia en valores mayores de R^2 y R , indicando una capacidad ligeramente superior para explicar la variabilidad de los datos.

- En variables como `reviews_per_month` y `price`, el modelo logarítmico logró las correlaciones más altas ($R \approx 0.73$ y 0.64), revelando relaciones positivas con rendimientos decrecientes: a medida que aumenta la ocupación o el precio, el crecimiento de reseñas o ingresos se desacelera.
- En cambio, los modelos cuadráticos capturan curvas suaves y pequeñas variaciones, pero sus valores de R^2 se mantienen bajos (<0.1), lo que sugiere que explican poca parte de la variabilidad total.
- Para las variables de host performance (`host_response_rate`, `host_acceptance_rate`), ambos modelos muestran correlaciones débiles, reflejando que los factores analizados (noches mínimas, reservas instantáneas) influyen muy poco en las tasas de respuesta o aceptación.
- En `accommodates` y `estimated_revenue_l365d`, el comportamiento logarítmico describe mejor la relación no lineal, sugiriendo que el ingreso aumenta con la capacidad pero tiende a estabilizarse.

Conclusión:

El modelo logarítmico es más adecuado en contextos donde el crecimiento se desacelera o se estabiliza (como ingresos, ocupación o número de reseñas). El modelo cuadrático, aunque útil para detectar curvaturas simples, resulta menos eficaz para representar patrones reales en estos datos. En conjunto, ambos modelos confirman que las relaciones entre las variables analizadas son no lineales, positivas pero débiles a moderadas, influenciadas por múltiples factores externos.