



Tecnológico de Monterrey

Campus Puebla

José Manuel Flores Pérez | A01733838

Actividad 3.1: Reporte

Alfredo García Suárez

13 de abril de 2025

En la etapa inicial del análisis, se realizó una limpieza y transformación de los datos, las columnas `host_response_rate` y `host_acceptance_rate` fueron convertidas de formato porcentaje a valores numéricos, mientras que la columna `price` fue limpiada de símbolos monetarios para su conversión a tipo flotante. Asimismo, se eliminaron registros con valores nulos, con el fin de asegurar que el ajuste de las curvas se realizará sobre datos válidos y consistentes.

Posteriormente, se aplicaron dos tipos de funciones: una función cuadrática de la forma $y = ax^2 + bx + c$ y una función logarítmica de la forma $y = a \cdot \log(x) + b$. Estas funciones se ajustaron utilizando `curve_fit`, y la calidad del ajuste se evaluó con el coeficiente de determinación R^2 y su raíz cuadrada r , que representa la correlación entre los valores predichos y los reales.

Se analizaron seis combinaciones de variables, la relación entre `host_response_rate` y `host_acceptance_rate` fue modelada mediante una función cuadrática, mostrando un ajuste visualmente adecuado; por otro lado, la variable `host_total_listings_count` fue comparada con `price` utilizando una función logarítmica, mostrando un comportamiento en el cual la tasa de crecimiento disminuye conforme aumentan los precios; también se aplicaron funciones cuadráticas para estudiar el vínculo entre `accommodates` y `reviews_per_month`, así como entre `reviews_per_month` y `host_acceptance_rate`. Las relaciones logarítmicas fueron útiles para capturar comportamientos en los que el efecto de una variable se estabiliza, como en el caso de `price` frente a `host_response_rate` y `host_acceptance_rate` frente a `accommodates`.

Los valores de R^2 y r fueron calculados e impresos en cada caso dentro del notebook, junto con las gráficas que muestran tanto los datos reales como las curvas ajustadas; estas visualizaciones permitieron observar de forma clara la calidad del ajuste y la naturaleza de la relación entre las variables.

En conclusión, este análisis confirma que las regresiones no lineales permiten capturar relaciones más complejas entre variables que una regresión lineal no puede modelar eficientemente, las funciones cuadráticas resultaron útiles en casos donde existe una curvatura evidente en los datos, mientras que las funciones logarítmicas ayudaron a describir relaciones con crecimiento desacelerado, la combinación de estos modelos, junto con una visualización adecuada, proporciona una herramienta poderosa para el análisis exploratorio y predictivo en conjuntos de datos similares.