

**Instituto Tecnológico y de Estudios
Superiores de Monterrey**



**Tecnológico
de Monterrey**

Analítica de Datos y Herramientas de Inteligencia Artificial 1

Actividad 1 (Entendimiento de los Datos)

ALUMNO

André Calmus González

A01733529

Grupo 101

Profesor: Alfredo García

16, Febrero 2025

Tabla con Traducción de las 75 Variables y Descripción

Nombre encabezado	Traducción al Español	Descripción Español
id	Identificador del anuncio	ID único del anuncio en Airbnb.
listing_url	URL del anuncio	URL de la página del anuncio.
scrape_id	Identificador del scrape	ID del proceso de recolección de datos ("scrape") de Inside Airbnb.
last_scraped	Fecha del último scrape	Fecha y hora en que se obtuvieron los datos del anuncio.
source	Fuente de información	Origen de los datos: búsqueda por vecindario o scrape anterior.
name	Nombre del anuncio	Nombre del anuncio.
description	Descripción del anuncio	Descripción detallada del anuncio.
neighborhood_overview	Descripción del vecindario	Descripción del vecindario por el anfitrión.
picture_url	URL de la imagen	URL de la imagen del anuncio.
host_id	Identificador del anfitrión	ID único del anfitrión en Airbnb.
host_url	URL del anfitrión	URL de la página del anfitrión.
host_name	Nombre del anfitrión	Nombre del anfitrión.
host_since	Fecha de alta del anfitrión	Fecha en que se creó el perfil del anfitrión.
host_location	Ubicación del anfitrión	Ubicación del anfitrión según su perfil.
host_about	Descripción del anfitrión	Descripción sobre el anfitrión.
host_response_time	Tiempo de respuesta	Tiempo que tarda el anfitrión en responder.
host_response_rate	Tasa de respuesta	Porcentaje de respuestas del anfitrión.
host_acceptance_rate	Tasa de aceptación	Porcentaje de solicitudes de reserva que acepta el anfitrión.
host_is_superhost	Es super anfitrión	Indica si el anfitrión es Superhost.
host_thumbnail_url	URL miniatura del anfitrión	URL de la miniatura de la foto del anfitrión.

host_picture_url	URL foto del anfitrión	URL de la foto del anfitrión.
host_neighbourhood	Vecindario del anfitrión	Vecindario del anfitrión.
host_listings_count	Cantidad de anuncios	Número de anuncios del anfitrión.
host_total_listings_count	Cantidad total de anuncios	Número total de anuncios del anfitrión.
host_verifications	Verificaciones del anfitrión	Métodos de verificación del anfitrión.
host_has_profile_pic	Anfitrión tiene foto	Indica si el anfitrión tiene foto de perfil.
host_identity_verified	Identidad verificada	Indica si la identidad del anfitrión está verificada.
neighbourhood	Vecindario	Vecindario del anuncio.
neighbourhood_cleansed	Vecindario normalizado	Vecindario geocodificado usando latitud y longitud.
neighbourhood_group_cleansed	Grupo de vecindario	Grupo de vecindario geocodificado usando latitud y longitud.
latitude	Latitud	Latitud del anuncio (WGS84).
longitude	Longitud	Longitud del anuncio (WGS84).
property_type	Tipo de propiedad	Tipo de propiedad (ej. casa, apartamento, hotel).
room_type	Tipo de dormitorio	tipo de dormitorio del alojamiento
accommodates	Capacidad	Capacidad máxima del anuncio.
bathrooms	Cantidad de baños	Número de baños en el anuncio.
bathrooms_text	Texto de baños	Descripción textual del número de baños.
bedrooms	Cantidad de dormitorios	Número de dormitorios en el anuncio.
beds	Cantidad de camas	Número de camas en el anuncio.
amenities	Comodidades	Lista de comodidades ofrecidas.
price	Precio diario	Precio por noche en la moneda local.
minimum_nights	Noches mínimas	Mínimo de noches de estancia.
maximum_nights	Noches máximas	Máximo de noches de estancia.
minimum_minimum_nights	Mínimo de noches mínimas	Mínimo de noches mínimas (calculado del calendario).

maximum_minimum_nights	Máximo de noches mínimas	Máximo de noches mínimas (calculado del calendario).
minimum_maximum_nights	Mínimo de noches máximas	Mínimo de noches máximas (calculado del calendario).
maximum_maximum_nights	Máximo de noches máximas	Máximo de noches máximas (calculado del calendario).
minimum_nights_avg_ntm	Promedio noches mínimas	Promedio de noches mínimas (calculado del calendario).
maximum_nights_avg_ntm	Promedio noches máximas	Promedio de noches máximas (calculado del calendario).
calendar_updated	Calendario actualizado	Fecha de la última actualización del calendario.
has_availability	Tiene disponibilidad	Indica si el anuncio está disponible.
availability_30	Disponibilidad en 30 días	Disponibilidad en los próximos 30 días.
availability_60	Disponibilidad en 60 días	Disponibilidad en los próximos 60 días.
availability_90	Disponibilidad en 90 días	Disponibilidad en los próximos 90 días.
availability_365	Disponibilidad en 365 días	Disponibilidad en los próximos 365 días.
calendar_last_scraped	Último scrape del calendario	Fecha del último "scrape" del calendario.
number_of_reviews	Cantidad de reviews	Número de reseñas del anuncio.
number_of_reviews_ltm	Reviews en los últimos 12 meses	Número de reseñas en los últimos 12 meses.
number_of_reviews_l30d	Reviews en los últimos 30 días	Número de reseñas en los últimos 30 días.
first_review	Fecha del primer review	Fecha de la primera reseña.
last_review	Fecha del último review	Fecha de la última reseña.
review_scores_rating	Puntaje de rating	Puntaje general de las reseñas.
review_scores_accuracy	Puntaje de precisión	Puntaje de precisión de las reseñas.
review_scores_cleanliness	Puntaje de limpieza	Puntaje de limpieza de las reseñas.
review_scores_checkin	Puntaje de checkin	Puntaje de facilidad de check-in de las reseñas.
review_scores_communication	Puntaje de comunicación	Puntaje de comunicación de las reseñas.
review_scores_location	Puntaje de ubicación	Puntaje de ubicación de las reseñas.

review_scores_value	Puntaje de valor	Puntaje de valor (calidad/precio) de las reseñas.
license	Licencia	Número de licencia o permiso.
instant_bookable	Reserva instantánea	Indica si se puede reservar al instante.
calculated_host_listings_count	Anuncios del anfitrión	Número de anuncios del anfitrión en la zona.
calculated_host_listings_count _entire_homes	Anuncios de casas/aptos	Número de casas/aptos del anfitrión en la zona.
calculated_host_listings_count _private_rooms	Anuncios de habitaciones	Número de habitaciones privadas del anfitrión en la zona.
calculated_host_listings_count _shared_rooms	Anuncios compartidos	Número de habitaciones compartidas del anfitrión en la zona.
reviews_per_month	Reviews por mes	Promedio de reseñas por mes.

Tabla con las 50 Variables Seleccionadas y su Justificación

Número de Variable	Variable	Justificación
0	Identificador del anuncio	Identificador único para cada anuncio, esencial para el análisis y cruce de datos.
2	Identificador del scrape	Permite identificar el lote de datos del cual proviene la información, útil para análisis de tendencias temporales.
4	Fuente de información	Permite identificar el origen de los datos, y si un anuncio fue encontrado por búsqueda de vecindario o un scrape anterior.
9	Identificador del anfitrión	Permite agrupar anuncios por anfitrión, útil para analizar el comportamiento de los anfitriones y su impacto en las propiedades.
15	Tiempo de respuesta	Importante saber cuanto tiempo demora el anfitrión en contestar.
16	Tasa de respuesta	Importante saber con qué frecuencia el anfitrión contesta.
17	Tasa de aceptación	Porcentaje de solicitudes de reserva aceptadas, indicativo de la calidad y demanda del anfitrión.
18	Es superanfitrión	Indica si el anfitrión está catalogado como un superanfitrión lo cual es importante para saber la calidad de su servicio.

22	Cantidad de anuncios	Número de anuncios que tiene el anfitrión, útil para analizar la escala de su negocio.
23	Cantidad total de anuncios	Número total de anuncios del anfitrión, incluyendo los inactivos o eliminados.
30	Latitud	Coordenada geográfica, fundamental para el análisis espacial y la ubicación de los anuncios.
31	Longitud	Coordenada geográfica, fundamental para el análisis espacial y la ubicación de los anuncios.
32	Tipo de propiedad	Tipo de propiedad, importante para entender las características de la oferta.
33	Tipo de dormitorio	Tipo de habitación, clave para segmentar la oferta y analizar las preferencias de los usuarios.
34	Capacidad	Número máximo de personas que pueden alojarse, influye en el precio y la demanda.
35	Cantidad de baños	Número de baños en el anuncio, afecta el confort y el precio.
36	Texto de baños	Descripción textual del número de baños.
37	Cantidad de dormitorios	Número de dormitorios en el anuncio, influye en el confort y el precio.
38	Cantidad de camas	Número de camas en el anuncio, afecta el confort y el precio.
39	Comodidades	Comunidades con las que cuenta el alojamiento, importante para conocer con lo que este cuenta.
40	Precio diario	Precio por noche, la variable más importante para entender el valor de la oferta.
41	Noches mínimas	Mínimo de noches de estancia, influye en la demanda y el tipo de cliente.
42	Noches máximas	Máximo de noches de estancia, influye en la demanda y el tipo de cliente.
43	Mínimo de noches mínimas	Mínimo de noches mínimas (calculado del calendario), da una idea de la flexibilidad del anfitrión.
44	Máximo de noches mínimas	Máximo de noches mínimas (calculado del calendario), da una idea de la flexibilidad del anfitrión.
45	Mínimo de noches máximas	Mínimo de noches máximas (calculado del calendario), da una idea de la flexibilidad del anfitrión.
46	Máximo de noches máximas	Máximo de noches máximas (calculado del calendario), da una idea de la flexibilidad del anfitrión.

47	Promedio noches mínimas	Promedio de noches mínimas (calculado del calendario), da una idea de la flexibilidad del anfitrión.
48	Promedio noches máximas	Promedio de noches máximas (calculado del calendario), da una idea de la flexibilidad del anfitrión.
50	Tiene disponibilidad	Indica si el anuncio está disponible, crucial para el análisis de la ocupación.
51	Disponibilidad en 30 días	Disponibilidad en los próximos 30 días, útil para predecir la ocupación a corto plazo.
52	Disponibilidad en 60 días	Disponibilidad en los próximos 60 días, útil para predecir la ocupación a medio plazo.
53	Disponibilidad en 90 días	Disponibilidad en los próximos 90 días, útil para predecir la ocupación a medio plazo.
54	Disponibilidad en 365 días	Disponibilidad en los próximos 365 días, útil para predecir la ocupación a largo plazo.
56	Cantidad de reviews	Número de reseñas, indicativo de la popularidad y antigüedad del anuncio.
57	Reviews en los últimos 12 meses	Número de reseñas en los últimos 12 meses, indicativo de la popularidad reciente.
58	Reviews en los últimos 30 días	Número de reseñas en los últimos 30 días, indicativo de la popularidad muy reciente.
61	Puntaje de rating	Puntaje general de las reseñas, clave para entender la calidad percibida.
62	Puntaje de precisión	Puntaje de precisión en las reseñas, indica si el anuncio es como se describe.
63	Puntaje de limpieza	Puntaje de limpieza en las reseñas, importante para la satisfacción del cliente.
64	Puntaje de checkin	Puntaje de facilidad de check-in, influye en la experiencia del usuario.
65	Puntaje de comunicación	Puntaje de comunicación en las reseñas, influye en la experiencia del usuario.
66	Puntaje de ubicación	Puntaje de ubicación en las reseñas, importante para el análisis espacial y la preferencia de los usuarios.
67	Puntaje de valor	Puntaje de valor (calidad/precio) en las reseñas, clave para entender la relación calidad-precio.
69	Reserva instantánea	Indica si se puede reservar al instante, influye en la demanda y el tipo de cliente.

70	Anuncios del anfitrión	Número de anuncios del anfitrión en la zona, útil para entender la competencia.
71	Anuncios de casas/aptos	Número de casas/aptos del anfitrión en la zona, útil para entender la competencia en ese segmento.
72	Anuncios de habitaciones	Número de habitaciones privadas del anfitrión en la zona, útil para entender la competencia en ese segmento.
73	Anuncios compartidos	Número de habitaciones compartidas del anfitrión en la zona, útil para entender la competencia en ese segmento.
74	Reviews por mes	Promedio de reseñas por mes, indicativo de la popularidad y la actividad reciente.