



BENEMERITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS DE LA COMPUTACIÓN

Inteligencia de Negocios

Actividad 1 (Entendimiento de los Datos)

Docente: Alfredo García Suárez

Alumno: Cesar Serafín Sampallo Amador

Entendimiento de los datos Australia – Tasmania

Traducción y descripción al español de las variables

Variable en ingles	Variable en español	Descripción en español
id	id	Identificador único del Airbnb del listado.
listing_url	url_de_la_publicación	Dirección de la publicación del Airbnb.
scrape_id	Id_del_scraping	Dentro del scrapeo esto era parte de.
last_scraped	ultimo_scrape	La fecha y hora en que la publicación fue scrapeada.
source	origen	Una de las “búsquedas del vecindario” o “scrapeo previo”. “búsquedas del vecindario” se refiere a que la publicación fue encontrada buscando la ciudad, mientras que “scrapeo previo” se refiere a que la publicación fue vista en otro scrapeo hecho en los últimos 65 días, y se confirmó que la publicación aún estaba disponible en el sitio de Airbnb.
name	nombre	Nombre de la publicación.
description	descripción	Descripción detallada de la publicación.
neighborhood_overview	descripción_del_vecindario	Descripción del vecindario del anfitrión
picture_url	url_de_la_imagen	Dirección de la imagen de tamaño regular alojada de Airbnb para la publicación.
host_id	Id_del_anfitrión	Identificador único de Airbnb para el anfitrión.

host_url	url_del_anfitrion	La página de Airbnb del anfitrión.
host_name	nombre_del_anfitrion	Nombre del anfitrión. Usualmente solo el nombre.
host_since	anfitrión_desde	La fecha de creación del anfitrión/usuario. Para anfitriones que son huéspedes de Airbnb, esta podría ser la fecha en que se registraron como invitados.
host_location	ubicación_del_anfitrion	La ubicación auto reportada del anfitrión.
host_about	acerca_del_anfitrion	Descripción acerca del anfitrión.
host_response_time	tiempo_de_respuesta_del_anfitrion	Tiempo en que responde el anfitrión.
host_response_rate	porcentaje_de_respuesta_del_anfitrion	Porcentaje del 0% al 100% que se le da al tiempo de respuesta de un anfitrión
host_acceptance_rate	porcentaje_de_aceptacion_del_anfitrion	Porcentaje en que un anfitrión acepta solicitudes de reserva.
host_is_superhost	anfitrión_es_superanfitrion	Anfitrión que ha alcanzado un nivel destacado de excelencia en la plataforma.
host_thumbnail_url	url_miniatura_del_anfitrión	Dirección de la miniatura del anfitrión.
host_picture_url	url_foto_del_anfitrion	Dirección de foto del anfitrión.
host_neighbourhood	vecindario_del_anfitrion	Vecindario del anfitrión.
host_listings_count	número_de_publicaciones_del_anfitrión	El número de publicaciones que tiene el anfitrión (por cálculos desconocidos de Airbnb).
host_total_listings_count	número_de_publicaciones_totales_del_anfitrión	El número de publicaciones totales que tiene el anfitrión (por cálculos desconocidos de Airbnb).

host_verifications	verificaciones_del_anfitrión	Total de fuentes de contacto (email, teléfono. etc) verificados del anfitrión.
host_has_profile_pic	foto_de_perfil_del_anfitrión	¿El anfitrión tiene foto de perfil?
host_identity_verified	identidad_del_anfitrión_verificada	¿El anfitrión ha verificado su identidad?
neighbourhood	vecindario_destacado	¿Pertenece a un vecindario destacado?
neighbourhood_cleansed	vecindario_corregido	El vecindario como geocodificado utilizando la latitud y longitud contra los vecindarios según lo definido por los archivos de forma digitales abiertos o públicos.
neighbourhood_group_cleansed	vecindario_de_grupo_corregido	El grupo vecinal como geocodificado utilizando la latitud y longitud contra los vecindarios según lo definido por los archivos de forma digitales abiertos o públicos.
latitude	latitud	Utiliza el Sistema Mundial Geodésico (WGS84) para la latitud y longitud.
longitude	longitud	Utiliza el Sistema Mundial Geodésico (WGS84) para la latitud y longitud.
property_type	tipo_de_propiedad	Tipo de propiedad autoelegido. Los hoteles y los alojamiento y desayuno son descritos como tales por sus anfitriones en este campo.
room_type	tipo_de_cuarto	[casas enteras, departamentos, cuarto privado, cuarto compartido, hotel]. Todas las casas están agrupadas dentro de los siguientes tres grupos de cuarto: Casa completa

		Habitaciones privadas Habitaciones compartidas
accommodates	capacidad_alojamiento	La capacidad máxima de alojamiento.
bathrooms	baños	El numero de baños en el alojamiento.
bathrooms_text	especificacion_de_baños	El número de baños en la publicación. En el sitio web de Airbnb, el campo de baños ha evolucionado de un número a una descripción textual. Para las versiones anteriores de los datos, se utiliza 'baños'.
bedrooms	habitaciones	Número de habitaciones.
beds	camas	Número de camas.
amenities	amenidades	Productos y servicios adicionales ofrecidos en el alojamiento.
price	precio	Precio diario en la moneda local.
minimum_nights	minimo_de_noches	Mínimo de noches de estadía en el alojamiento (el calendario de asignaciones puede cambiar).
maximum_nights	maximo_de_noches	Máximo de noches de estadía en el alojamiento (el calendario de asignaciones puede cambiar).
minimum_minimum_nights	noches_mínimas_requeridas	El valor mínimo de noches requeridas en el calendario.
maximum_minimum_nights	noches_mínimas_máximas	El valor más grande de noches mínimas requeridas en el calendario.
minimum_maximum_nights	número_mínimo_de_noches_máximas	El valor más pequeño de noches maximas requeridas en el calendario.

maximum_maximum_nights	número_máximo_de_noches_máximas	El valor más grande de máximo_noches requeridas en el calendario.
minimum_nights_avg_ntm	promedio_de_noches_minimas	El valor promedio de mínimo_noches del calendario.
maximum_nights_avg_ntm	promedio_de_noches_maximas	El valor promedio de máximo_noches del calendario.
calendar_updated	calendario_actualizado	
has_availability	tiene_disponibilidad	¿Tiene disponibilidad?
availability_30	disponibilidad_30	La disponibilidad del alojamiento x días en el futuro según lo establecido en el calendario.
availability_60	disponibilidad_60	La disponibilidad del alojamiento x días en el futuro según lo establecido en el calendario.
availability_90	disponibilidad_90	La disponibilidad del alojamiento x días en el futuro según lo establecido en el calendario.
availability_365	disponibilidad_365	La disponibilidad del alojamiento x días en el futuro según lo establecido en el calendario.
calendar_last_scraped	ultimo_scrapeo_en_el_calendario	Ultimo scrapeo en el calendario
number_of_reviews	numero_de_reseñas	El número de reseñas que tiene la publicación.
number_of_reviews_ltm	numero_de_reseñas_ltm	El número de reseñas que tiene la publicación en los últimos 12 meses.
number_of_reviews_l30d	numero_de_reseñas_l30d	El número de reseñas que tiene la publicación en los últimos 30 días.
first_review	primera_reseña	La fecha de la primera reseña.
last_review	ultima_reseña	La fecha de la última reseña más reciente.

review_scores_rating	puntuación_de_reseñas	Puntuación de las reseñas. Dando de 1 a 5 estrellas.
review_scores_accuracy	Precisión_de_la_puntuacion_del_anuncio	Puntuación dada a la descripción del anuncio. Dando de 1 a 5 estrellas.
review_scores_cleanliness	puntuación_de_limpieza	Puntuación dada a la limpieza del alojamiento.
review_scores_checkin	puntuación_de_check-in	Puntuación dada sobre el proceso de registro.
review_scores_communication	puntuación_de_comunicacion_anfitrión	Puntuación dada a la comunicación del anfitrión.
review_scores_location	puntuación_de_reseñas_ubicación	Puntuación dada a la ubicación de la propiedad.
review_scores_value	puntuación_de_reseñas_valor	Puntuación dada sobre la relación calidad-precio de la propiedad.
license	licencia	El número de licencia/permiso/registro
instant_bookable	reservable_inmediatamente	¿Se permite hacer reservas instantáneas?
calculated_host_listings_count	cantidad_calculada_de_publicaciones_del_anfitrión	El número de publicaciones que el anfitrión tiene en la geografía de la ciudad/región.
calculated_host_listings_count_entire_homes	cantidad_calculada_de_publicaciones_de_casas_completas_del_anfitrión	El número de publicaciones con casa completa que el anfitrión tiene en la geografía de la ciudad/región.
calculated_host_listings_count_private_rooms	cantidad_calculada_de_publicaciones_de_cuartos_privados_del_anfitrión	El número de publicaciones con cuartos privados que el anfitrión tiene en la geografía de la ciudad/región.
calculated_host_listings_count_shared_rooms	cantidad_calculada_de_publicaciones_de_cuartos_compartidos_del_anfitrión	El número de publicaciones con cuartos compartidos que el anfitrión tiene en la geografía de la ciudad/región.
reviews_per_month	reseñas_por_mes	El número promedio de reseñas por mes que tiene la

		publicación durante la vida útil de la publicación.
--	--	---

Selección de las 50 variables

1. **Id:** Sirve para identificar de forma única cada publicación. Es fundamental para el seguimiento y análisis de datos.
2. **descripción:** Es útil para evaluar la calidad y claridad de la información proporcionada, lo que influye en la percepción del cliente.
3. **ultimo_scrape:** Sirve para conocer la fecha y hora más reciente en que se obtuvo información de una publicación.
4. **descripción_del_vecindario:** Útil para evaluar la calidad del vecindario, se podría relacionar con la ubicación del Airbnb.
5. **Id_del_anfitrión:** Es esencial para vincular datos relacionados con el anfitrión y analizar patrones en su desempeño.
6. **anfitrión_desde:** Refleja la experiencia del anfitrión en la plataforma, lo cual influye a la decisión de un cliente al reservar con anfitriones con más experiencia.
7. **tiempo_de_respuesta_del_anfitrión:** Un tiempo de respuesta rápido puede aumentar la confianza del cliente y mejorar la experiencia.
8. **porcentaje_de_respuesta_del_anfitrión:** Se relaciona directamente con el tiempo de respuesta del anfitrión, además podría relacionarse con la decisión final del cliente.
9. **porcentaje_de_aceptacion_del_anfitrión:** Se relaciona con el porcentaje y tiempo de respuesta del anfitrión. Además, yo creo que a más porcentaje de aceptación mas probabilidad hay de aceptar un cliente.

10. **anfitrión_es_superanfitrión:** La etiqueta de "superanfitrión" es un factor de confianza y puede influir en la decisión final de un cliente.
11. **número_de_publicaciones_del_anfitrión:** Muestra la experiencia del anfitrión y podría correlacionarse con la calidad del servicio.
12. **foto_de_perfil_del_anfitrión:** Si el anfitrión tiene foto de perfil más confianza tendrá el cliente de reservar.
13. **vecindario_corregido:** La ubicación del vecindario es necesaria al tomar una decisión por parte del cliente.
14. **Latitud:** puede servir para analizar precios basados en ubicación y accesibilidad del Airbnb.
15. **Longitud:** pueden servir para analizar precios basados en ubicación y accesibilidad del Airbnb.
16. **precio:** Es una variable clave para la toma de decisiones del cliente, ya que cada cliente tiende a tener un presupuesto diferente.
17. **verificaciones_del_anfitrión:** Refleja el nivel de seguridad y confianza que puede ofrecer el anfitrión a los clientes.
18. **vecindario_destacado:** La ubicación es uno de los principales factores que los clientes consideran al elegir un alojamiento.
19. **tipo_de_propiedad:** Los clientes tienen diferentes preferencias según sus necesidades (casa, apartamento, etc.).
20. **tipo_de_cuarto:** Ya que define el nivel de privacidad que los clientes pueden esperar.
21. **capacidad_alojamiento:** Esencial para grupos de clientes que busquen una propiedad grande.
22. **baños:** Es importante para grupos grandes, y puede relacionarse con la capacidad de alojamiento y tipo de propiedad.
23. **habitaciones:** Influye en la comodidad y privacidad del cliente. Además, se relaciona con los baños, tipo de propiedad y capacidad de alojamiento.
24. **camas:** Influye en la decisión final del cliente. Además, se relaciona con el tipo de propiedad y capacidad de alojamiento.

25. **amenidades:** Es una variable clave que puede influir en la decisión del cliente al ofrecer comodidades adicionales.
26. **minimo_de_noches:** Afecta a la flexibilidad del cliente para reservar.
27. **maximo_de_noches:** Influye en las reservas a largo plazo.
28. **tiene_disponibilidad:** Influye en la toma de decisión, si un alojamiento no esta disponible en días que el cliente requiera, este será descartado.
29. **disponibilidad_30:** Influye en la toma de decisión de un cliente, ya que muestra cuántos días están disponibles en los próximos 30 días. De igual manera para detectar tendencias de ocupación en los próximos 30 días.
30. **disponibilidad_60:** Influye en la toma de decisión de un cliente, ya que muestra cuántos días están disponibles en los próximos 60 días. De igual manera para detectar tendencias de ocupación en los próximos 60 días.
31. **disponibilidad_90:** Influye en la toma de decisión de un cliente, ya que muestra cuántos días están disponibles en los próximos 90 días. De igual manera para detectar tendencias de ocupación en los próximos 90 días.
32. **disponibilidad_365:** Influye en la toma de decisión de un cliente, ya que muestra cuántos días están disponibles en los próximos 365 días. De igual manera para detectar tendencias de ocupación en los próximos 365 días.
33. **numero_de_reseñas:** Un alto número de reseñas indica que el anuncio es popular y puede generar más confianza en los clientes.
34. **Precisión_de_la_puntuacion_del_anuncio:** La precisión de la descripción del anuncio influye en las expectativas del cliente y en su satisfacción al llegar al alojamiento.
35. **puntuación _de_limpieza:** El nivel de limpieza influye en la decisión del cliente.
36. **puntuación_de _check-in:** Un proceso de registro fácil y rápido mejora la experiencia del cliente.

37. **puntuación_de_comunicacion_anfitrión:** Una buena comunicación ayuda a resolver dudas rápidamente, lo que aumenta la confianza y la satisfacción del cliente.
38. **puntuación_de_reseñas_ubicación:** La ubicación es uno de los factores más importantes para los clientes al seleccionar un alojamiento.
39. **reservable_inmediatamente:** Un cliente que obtenga una reserva inmediata por parte de una publicación tiende a elegirla.
40. **cantidad_calculada_de_publicaciones_del_anfitrión:** La cantidad de publicaciones de un anfitrión impacta en la decisión del cliente, ya que tiene más opciones a reservar.
41. **cantidad_calculada_de_publicaciones_de_casas_completas_del_anfitrión:** Permite analizar cuántas publicaciones del anfitrión corresponden a casas completas, lo cual es relevante para clientes que buscan privacidad total durante su estadía.
42. **cantidad_calculada_de_publicaciones_de_cuartos_privados_del_anfitrión:** Ayuda a identificar cuántos alojamientos ofrecen habitaciones privadas, podría ser importante para clientes que desean una opción intermedia entre privacidad y costo.
43. **cantidad_calculada_de_publicaciones_de_cuartos_compartidos_del_anfitrión:** Proporciona información sobre cuántas publicaciones del anfitrión son habitaciones compartidas, podría ser útil para analizar opciones más económicas o para clientes que buscan socializar.
44. **reseñas_por_mes:** Si tiene mas frecuencia de publicaciones, más probabilidad hay de que se le recomiende a un cliente.
45. **licencia:** La licencia está ligada al anfitrión, por ende, es una variable que influye en su selección por parte del cliente.
46. **numero_de_reseñas_ltm:** Servirá analizar la popularidad y demanda del alojamiento a lo largo del último año, identificando tendencias anuales.
47. **numero_de_reseñas_l30d:** Servirá analizar la popularidad y demanda del alojamiento en los últimos 30 días.

- 48. **primera_reseña:** Permite identificar cuánto tiempo lleva el alojamiento recibiendo reseñas, lo que ayuda a evaluar su antigüedad.
- 49. **ultima_reseña:** Indica la actualidad del alojamiento en términos de actividad reciente, ayudando a evaluar su popularidad reciente.
- 50. **ultimo_scrapeo_en_el_calendario:** Permite asegurar que la información analizada esté actualizada y refleje la realidad actual.