



Benemérita Universidad Autónoma de Puebla

FCC

Alfredo García

Inteligencia de Negocios

Entendimiento de los datos

202154423

José Armando Ramos Ramos

Periodo: Primavera 2025

Punto número tres

Realizar la traducción al español y la descripción de cada una de las variables de la base de datos con ayuda del “**Data Dictionary**” proporcionado por la compañía Airbnb.

Variable o campo	Descripción
Identificador (Id)	Identificador único de Airbnb para el anuncio
URL del listado	Enlace directo a la página web asociada
Identificador de extracción (scrape_id)	Identificador único asignado a cada registro durante el proceso de extracción de datos
Última extracción	La fecha y hora en que se extrajo este listado.
Fuente	Una de las opciones de "búsqueda de barrio" o "búsqueda previa". "Búsqueda de barrio" significa que el anuncio se encontró al buscar en la ciudad, mientras que "búsqueda previa" significa que el anuncio se vio en otra búsqueda realizada en los últimos 65 días y se confirmó que el anuncio aún estaba disponible en el sitio de Airbnb.
Nombre	Nombre del anuncio
Descripción	Descripción detallada del anuncio.
descripción general del vecindario	Descripción del vecindario proporcionada por el anfitrión.
URL de la imagen (URL_imagen)	URL de la imagen principal del anuncio, alojada por Airbnb.
identificación del host (id_host)	Identificador único de Airbnb para el anfitrión/usuario.
URL del host (host_url)	URL de la página del anfitrión en Airbnb.
Nombre de host (nombre_host)	Nombre del anfitrión. Generalmente, solo incluye el/los primeros nombres.

Anfitrión desde	La fecha en que se creó el anfitrión/usuario. Para los anfitriones que son huéspedes de Airbnb, esta podría ser la fecha en que se registraron como huéspedes.
Ubicación del anfitrión (ubicación_anfitrión)	La ubicación informada por el anfitrión
Acerca del anfitrión	Descripción del anfitrión
Tiempo de respuesta del host	Tiempo promedio que tarda el anfitrión en responder.
Tasa de respuesta del host	Porcentaje de respuestas del anfitrión a las consultas recibidas.
Tasa de aceptación del host	Porcentaje de solicitudes aceptadas por el anfitrión.
El host es superhost	Indica si el anfitrión tiene estatus de "superhost" ("t" para verdadero, "f" para falso).
URL de la miniatura del host	URL de la imagen en miniatura del anfitrión.
URL de la imagen del host	URL de la imagen del anfitrión.
vecindario_anfitrión	Barrio donde reside el anfitrión, según lo informado por él.
recuento_de_anfitriones_en_listado	Número total de anuncios publicados por el anfitrión.
recuento_total_de_anfitriones	La cantidad de anuncios que tiene el anfitrión (según cálculos desconocidos de Airbnb).
verificaciones del host	Métodos de verificación usados por el anfitrión.
el host tiene una foto de perfil	Indica si el anfitrión tiene una foto de perfil ("t" para verdadero, "f" para falso).
identidad del host verificada	Indica si la identidad del anfitrión está verificada ("t" para verdadero, "f" para falso).
vecindario	Nombre del vecindario donde está ubicado el alojamiento.

Vecindario_estandarizado	El vecindario se geo codifica utilizando la latitud y la longitud en comparación con los vecindarios definidos por archivos de forma digitales abiertos o públicos.
agrupación de vecindario normalizado	El grupo de vecindarios se geo codifica utilizando la latitud y la longitud en comparación con los vecindarios definidos por archivos de forma digitales abiertos o públicos
Latitud	Utiliza la proyección del Sistema Geodésico Mundial (WGS84) para latitud y longitud
Longitud	Utiliza la proyección del Sistema Geodésico Mundial (WGS84) para latitud y longitud
Tipo_propiedad	Tipo de propiedad seleccionado por el usuario. Los hoteles y los Bed and Breakfast se describen como tales por sus anfitriones en este campo
Tipo_habitación	<p>Tipo de habitación ofrecida, como habitación privada, compartida o alojamiento entero.</p> <p><b>Casa entera</b></p> <p>Las casas enteras son las mejores si buscas un hogar lejos de casa. Con una casa entera, tendrás todo el espacio para ti. Esto suele incluir un dormitorio, un baño, una cocina y una entrada independiente y exclusiva. Los anfitriones deben indicar en la descripción si estarán en la propiedad o no (por ejemplo: "El anfitrión ocupa el primer piso de la casa") y</p>

	<p>proporcionar más detalles en el anuncio.</p> <p><b>Habitaciones privadas</b>  Las habitaciones privadas son ideales para cuando prefieres un poco de privacidad y aún valoras una conexión local. Cuando reservas una habitación privada, tendrás tu propia habitación privada para dormir y puedes compartir algunos espacios con otras personas. Es posible que debas caminar por espacios interiores que otro anfitrión o huésped puede ocupar para llegar a tu habitación.</p> <p><b>Habitaciones compartidas</b>  Las habitaciones compartidas son para cuando no te importa compartir un espacio con otras personas. Cuando reservas una habitación compartida, dormirás en un espacio compartido con otras personas y compartirás todo el espacio con otras personas. Las habitaciones compartidas son populares entre los viajeros flexibles que buscan nuevos amigos y estadías económicas</p>
Capacidad	Número máximo de huéspedes que el alojamiento puede recibir
Baños	El número de baños en el listado
Baños_mencionados	<p>"El número de baños que hay en el anuncio.</p> <p>En el sitio web de Airbnb, el campo de baños ha evolucionado de un número a una descripción textual.</p> <p>En los anuncios más</p>

	antiguos, se utiliza bathroom"
Dormitorios	El número de dormitorios
Camas	El número de cama (s)
Comodidades	Lista de amenidades o servicios proporcionados en el alojamiento
Precio	"Precio diario en moneda local. NOTA: el signo \$ es un artefacto técnico de la exportación, ignórelo"
Noches mínimas	Número mínimo de noches de estancia para el alojamiento (las reglas del calendario pueden ser diferentes)
Noches máximas	Número máximo de noches de estancia para el alojamiento (las reglas del calendario pueden ser diferentes)
Noches mínimas mínimas	El valor mínimo de noche más pequeño del calendario (mirando 365 noches en el futuro)
Máximo mínimo noches	El valor mínimo de noche más grande del calendario (mirando 365 noches en el futuro)
Mínimo máximo noches	El valor máximo de noche más pequeño del calendario (mirando 365 noches en el futuro)
Máximo máximo de noches	El valor máximo de noche más grande del calendario (mirando 365 noches en el futuro)
Noches minimas_promedio ntm	el valor mínimo promedio de la noche según el calendario (mirando 365 noches en el futuro)
Noches máximas promedio ntm	el valor máximo promedio de la noche del calendario (mirando 365 noches en el futuro)

Calendario actualizado	Frecuencia con la que se actualiza el calendario del alojamiento
Tiene disponibilidad	Indica si el alojamiento tiene disponibilidad para reservas ("t" para verdadero, "f" para falso)
Disponibilidad_30	Disponibilidad. La disponibilidad del alojamiento dentro de x días según lo determine el calendario. Tenga en cuenta que un alojamiento puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión
Disponibilidad_60	Disponibilidad. La disponibilidad del alojamiento x días en el futuro según lo determine el calendario. Tenga en cuenta que un alojamiento puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión
Disponibilidad_90	La disponibilidad del alojamiento x días en el futuro según lo determine el calendario. Tenga en cuenta que un alojamiento puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión
Disponibilidad_365	Número de noches disponibles para reserva en los próximos 365 días
Calendario raspado por última vez	Fecha y hora en formato UTC en la que se capturó la información del calendario
Número_de_reseñas	El número de reseñas que tiene el anuncio
Número_de_revisiones_ltm	La cantidad de reseñas que tiene el anuncio (en los últimos 12 meses)

Número_de_reseñas_l30d	El número de reseñas que tiene el anuncio (en los últimos 30 días)
Primera_revisión	La fecha de la primera/más antigua revisión
Última_revisión	La fecha de la última/más reciente revisión
Calificación de las puntuaciones de revisión	Puntaje promedio general basado en las reseñas de los huéspedes
Precisión de las puntuaciones de revisión	Puntaje promedio de precisión basado en las reseñas de los huéspedes
Reseñas_puntuaciones_limpieza	Puntaje promedio de limpieza basado en las reseñas de los huéspedes
Revisión_puntuaciones_verificación	Puntaje promedio del proceso de check-in basado en las reseñas de los huéspedes
Revisión_puntuaciones_comunicación	Puntaje promedio de comunicación basado en las reseñas de los huéspedes
Ubicación de las puntuaciones de revisión	Puntaje promedio de ubicación basado en las reseñas de los huéspedes
Valor de las puntuaciones de revisión	Puntaje promedio de relación calidad-precio basado en las reseñas de los huéspedes
Licencia	El número de licencia/permiso/registro
Reservable_instantáneamente	[t=true; f=false]. Si el huésped puede reservar automáticamente el alojamiento sin que el anfitrión tenga que aceptar su solicitud de reserva. Indicador de un alojamiento comercial
Recuento_de_listas_de_hosts_calculados	La cantidad de listados que tiene el anfitrión en el scrap actual, en la geografía de la ciudad/región
Los listados de anfitriones calculados cuentan casas enteras	La cantidad de listados completos de casas/apartamentos que el



	anfitrión tiene en el scrap actual, en la geografía de la ciudad/región
Recuento de listas de hosts calculados para habitaciones privadas	La cantidad de listados de habitaciones privadas que el anfitrión tiene en el scrap actual, en la geografía de la ciudad/región
Recuento de habitaciones compartidas en listas de hosts calculados	La cantidad de listados de habitaciones compartidas que el anfitrión tiene en el scrap actual, en la geografía de la ciudad/región
Reseñas por mes	La cantidad promedio de reseñas por mes que tiene el listado durante su vigencia

Punto número cuatro.

Seleccionar 50 variables que consideres más importantes para el análisis de datos respecto a los requerimientos del cliente **Airbnb. (Justificar la selección de cada variable).**

Para seleccionar las variables para que se puedan hacer un análisis de datos debemos centrarnos en aquellas que nos puedan decir algo o interpretar acerca de, en este caso, la aplicación de Airbnb. De forma general que nos digan el desempeño de los anfitriones, es decir, analizando su desempeño, la experiencia que este tiene con el cliente y viceversa, entre otros. Por ello seleccione los siguientes.

Id. El id es un identificador único para poder rastrear específicamente un alojamiento.

Listing\_url. Su utilidad recae en que podemos consultar directamente los detalles del anuncio en Airbnb.

Name. El nombre del anuncio puede ser de atracción inicial para el cliente.

Descripción. Nos otorga información acerca de las características del alojamiento, por ende, creo que es importante a la hora de tomar la decisión.

Neighborhood\_overview. Nos brinda detalles acerca del vecindario, lo cual es un factor importante para el cliente.

Host\_id. Este permite analizar patrones de comportamiento y rendimiento por el anfitrión.

Host\_name. Nos da información relevante para poder personalizar análisis o consultas.

Host\_since. Acerca de la experiencia del anfitrión, es decir, su antigüedad, puede correlacionarse con la calidad del servicio.

Host\_response\_time. Mide la capacidad del anfitrión para su tiempo de respuesta, lo cual es un factor crítico para la satisfacción del huésped.

Host\_response\_rate. Permite evaluar la eficiencia del anfitrión en la gestión de consultas.

Host\_acceptance\_rate. Nos indica la disposición que tiene el anfitrión para aceptar reservas.

Host\_is\_superhost. En caso de los super host suelen estar asociados con una mejora de experiencia hacia los clientes.

Host\_verification. Nos asegura que los anfitriones estén verificados, lo cual aumenta la confianza de los clientes.

Neighbourhood\_cleansed. Otorga información de forma estandarizada y útil para análisis geográficos.

Latitud y longitud. Nos da una ubicación precisa para su análisis de manera especial o su visualización geográfica.

Property\_type. Identifica el tipo de propiedad, el cual, es un factor clave a la hora de que el huésped haga su elección.

Room\_type. Determina las opciones de privacidad disponibles para que los huéspedes escogan.

Acommodates. El número de huéspedes que se pueden alojar es un elemento crucial para las expectativas del cliente.

Bedrooms. Un aspecto importante para definir la capacidad y sus características del alojamiento.

Beds. Identifica cuantas personas pueden dormir cómodamente, un aspecto indispensable a la hora de tomar una decisión.

Amenities. La calidad y cantidad de amenidades puede influir en las reservas.

Price. Un factor fundamental en la decisión de los huéspedes.

Minimum\_nights y maximum\_nights. Determina las restricciones en la duración de la estancia.

Availability\_365. Nos proporciona una perspectiva acerca de la disponibilidad anual del alojamiento.

Number\_of\_views. Declara la popularidad del anuncio.

Number\_of\_view\_ltm. Indica la popularidad reciente, el cual es útil para identificar tendencias.

Reviews\_scores\_rating. Es el puntaje general, crucial para poder evaluar la calidad del alojamiento.

Reviews\_scores\_cleanliness. La limpieza es uno de los factores más importantes en el criterio de los huéspedes.

Review\_scores\_communication. La comunicación afecta de forma directa con la satisfacción de los huéspedes.

Review\_scores\_location. Ayuda a evaluar la convivencia de la ubicación del alojamiento.

Reviews\_scores\_value. Nos indica la percepción de calidad-precio del huésped.

Instant\_bookable. La opción de reserva instantánea puede atraer más usuarios, debido a (en su mayoría) los viajes imprevistos.

Calculated\_host\_listings\_count. Mide el tamaño de la operación del anfitrión.

Reviews\_per\_month. Indica con qué frecuencia los huéspedes dejan reseñas, es decir, nos refleja la actividad.

Host\_has\_profile\_pic. Genera confianza en el cliente al ver una foto del anfitrión.

License. Nos confirma la legibilidad del alojamiento, el cual, es un aspecto importante para los huéspedes.

Calendar\_update. Refleja la atención por parte del anfitrión al mantener la información actualizada.

Bathrooms. Al mostrar la información, nos ofrece información cuantitativa que nos ayuda a complementar la descripción textual.

Maximum\_nights\_avg\_ntm. Ayuda a dar una compresión acerca de las restricciones promedio en reservas prolongadas.

First\_review. Identifica el historial de actividad de alojamiento.

Last\_view. Indica la relevancia reciente del alojamiento.

Availability\_30. Es útil para poder analizar la disponibilidad inmediata.

Availability\_60. Complementa los análisis de disponibilidad, pero a corto plazo.

Calculated\_host\_listings\_count\_entire\_homes. Nos ayuda a segmentar el análisis en tipos específicos acerca del alojamiento.

Calculated\_host\_listings\_count\_private\_rooms. Permite distinguir entre alojamientos en habitaciones privadas.

Calculated\_host\_listings\_count\_shared\_rooms. Util para poder analizar opciones de habitaciones compartidas.

Host\_identify\_verified. Proporciona seguridad adicional a los huéspedes y con ello más atracción.

Host\_neighbourhood. Ayuda a identificar el área donde el anfitrión reside y con ello más seguridad.

Scrape\_id. Identifica sesiones de recopilación de datos específicas, para poder dar un resultado o a tomar decisiones.

Dichas 50, las tome en sentido que podamos hacer análisis de datos, es decir, que sean influyentes que de cierta manera pueda decirme algo o a los huéspedes para poder decidirse por dicho servicio.