

AUTOR: Valeria Becerril Hernández | A01736860

FECHA DE EXTRACCIÓN : 27 MARZO, 2025

PAÍS: ITALIA
REGIÓN: LOMBARDÍA
CIUDAD: BÉRGAMO

INTRODUCCIÓN: Para comprender de manera integral el mercado de Airbnb en Bergamo, se seleccionaron 50 variables clave que permiten analizar en profundidad los anuncios, los anfitriones, la ubicación y la experiencia de los huéspedes. Estas variables fueron escogidas por su utilidad para distintos tipos de análisis, incluyendo descriptivo, comparativo, geoespacial, de tendencias y de comportamiento de usuarios y anfitriones. Entre ellas se incluyen características del anuncio, como tipo de propiedad, número de habitaciones y camas, baños, capacidad máxima y amenidades, así como información sobre precios, disponibilidad y restricciones de estancia, lo que permite evaluar la oferta disponible, segmentar propiedades y comprender la dinámica de la demanda en diferentes contextos.

Asimismo, se incorporaron variables relacionadas con el desempeño y la profesionalización de los anfitriones, como antigüedad en la plataforma, nivel de verificación, estatus de superhost y cantidad de anuncios gestionados, lo que permite analizar su confiabilidad y estrategias de gestión. También se consideran indicadores de satisfacción de los huéspedes, como el número y frecuencia de reseñas y las calificaciones en limpieza, ubicación, comunicación y relación calidad-precio. Las variables geoespaciales y de vecindario, incluyendo coordenadas y barrios geocodificados, facilitan estudios de proximidad, segmentación por zonas y evaluación del impacto de la ubicación sobre precios, ocupación y experiencia del huésped. En conjunto, estas 50 variables proporcionan una base sólida para realizar un análisis completo del ecosistema Airbnb en Bergamo, ofreciendo información clave para la toma de decisiones estratégicas en la gestión de propiedades, optimización de precios y desarrollo de negocios turísticos.

DICCIONARIO DE DATOS				
COLUMNA	TIPO	SELECCIÓN	DESCRIPCIÓN	JUSTIFICACIÓN
id	integer		Identificador único de Airbnb para el anuncio	
listing_url	text		Link del anuncio	
scrape_id	bigint		Dentro de Airbnb es parte de «Scrape»	
last_scraped	datetime		UTC. Fecha y hora en que se «extrajo» este anuncio.	
source	text	X	Una de las opciones es «búsqueda por vecindario» o «raspado anterior». «Búsqueda por vecindario» significa que el anuncio se encontró buscando en la ciudad, mientras que «raspado anterior» significa que el anuncio se vio en otro raspado realizado en los últimos 65 días y se confirmó que el anuncio seguía disponible en el sitio web de Airbnb.	Para conocer cómo se obtuvo el anuncio y evaluar posibles sesgos en los datos.
name	text	X	Nombre del anuncio	Puede usarse para identificar estilos, promociones o marketing del anuncio.
description	text		Descripción detallada del anuncio	
neighborhood_overview	text	X	Descripción del vecindario por parte del anfitrión	Describe el vecindario, útil para análisis de ubicación y atracción turística.
picture_url	text		URL de la imagen de tamaño normal alojada en Airbnb para el anuncio.	
host_id	integer	X	Identificador único de Airbnb para el anfitrión/usuario.	Identificador único del anfitrión, esencial para vincular propiedades y analizar comportamiento.
host_url	text		Página de Airbnb para el anfitrión	
host_name	text		Nombre del anfitrión. Normalmente solo el nombre o nombres de pila.	
host_since	date	X	Fecha en la que se creó el anfitrión/usuario. Para los anfitriones que son huéspedes de Airbnb, podría ser la fecha en la que se registraron como huéspedes.	Permite analizar experiencia del anfitrión y su relación con calificaciones o precios.
host_location	text		Ubicación indicada por el anfitrión.	
host_about	text		Descripción del anfitrión	
host_response_time		X	Tiempo promedio de respuesta	Indicador de servicio.
host_response_rate		X	Tasa de respuesta	Evalúa confiabilidad del anfitrión.
host_acceptance_rate			La tasa a la que un anfitrión acepta solicitudes de reserva.	
host_is_superhost	boolean [t=true; f=false]	X	Anfitrión destacado	Influye en precio, reservas y confiabilidad.
host_thumbnail_url	text		Link del la foto pequeña del anfitrión	
host_picture_url	text		Link del la foto mediana del anfitrión	
host_neighbourhood	text		Vecindario del anfitrión	
host_listings_count	text		El número de anuncios que tiene el anfitrión (según cálculos desconocidos de Airbnb).	
host_total_listings_count	text		El número de anuncios que tiene el anfitrión (según cálculos desconocidos de Airbnb).	
host_verifications		X	Métodos de verificación	Indicador de seguridad y confianza.
host_has_profile_pic	boolean [t=true; f=false]	X	Perfil con foto	Análisis de confianza y profesionalismo.
host_identity_verified	boolean [t=true; f=false]	X	Verificación de identidad	Indicador de seguridad.

neighbourhood	text		Vencindario	
neighbourhood_cleansed	text	X	El barrio geocodificado utilizando la latitud y la longitud en relación con los barrios definidos por archivos de forma digitales abiertos o públicos.	Barrio geocodificado; útil para análisis por zonas.
neighbourhood_group_cleansed	text	X	El grupo de barrios geocodificados utilizando la latitud y la longitud en comparación con los barrios definidos por archivos de forma digitales abiertos o públicos.	Agrupación de barrios; permite análisis regional más amplio.
latitude	numeric	X	Utiliza la proyección del Sistema Geodésico Mundial (WGS84) para la latitud y la longitud.	Coordenada geográfica para análisis espacial y proximidad a puntos de interés.
longitude	numeric	X	Utiliza la proyección del Sistema Geodésico Mundial (WGS84) para la latitud y la longitud.	Coordenada geográfica para mapeo y clusterización de propiedades
property_type	text	X	Tipo de propiedad seleccionado por el usuario. Los hoteles y los bed and breakfast son descritos como tales por sus propietarios en este campo.	Tipo de propiedad; permite análisis según apartamento, casa, B&B, etc.
room_type	text	X	<p>«[Casa/apartamento completo Habitación privada Habitación compartida Hotel]</p> <p>Todas las casas se agrupan en los siguientes tres tipos de habitaciones:</p> <p>Lugar completo Habitación privada Habitación compartida</p> <p>Lugar completo Los lugares completos son la mejor opción si buscas un hogar lejos de casa. Con un lugar completo, tendrás todo el espacio para ti. Por lo general, esto incluye 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 la primera planta de la casa») y proporcionar más detalles en el anuncio.</p> <p>Habitaciones privadas Las habitaciones privadas son ideales si prefieres un poco de privacidad, pero sigues valorando el contacto con los lugareños. Cuando reservas una habitación privada, tendrás tu propia habitación para dormir y es posible que compartas algunos espacios con otras personas. Es posible que tengas que atravesar espacios interiores que otro anfitrión u huésped pueda ocupar para llegar a tu habitación.</p> <p>Habitaciones compartidas Las habitaciones compartidas son para cuando no te importa compartir un espacio con otras personas. Cuando reservas una habitación compartida, dormirás en un espacio que compartes con otras personas y compartirás todo el espacio con ellas. Las habitaciones compartidas son populares entre los viajeros flexibles que buscan nuevos amigos y estancias económicas.</p>	Determina privacidad y estilo de la estancia, útil para segmentar clientes.
accommodates	integer	X	La capacidad máxima del anuncio	Capacidad máxima del anuncio; clave para dimensionar precios y demanda.
bathrooms	numeric	X	El número de baños del anuncio	Cantidad de baños, indicador de comodidad y lujo.
bathrooms_text	string		<p>«El número de baños del anuncio.</p> <p>En el sitio web de Airbnb, el campo de los baños ha pasado de ser un número a una descripción textual. Para los anuncios más antiguos, se utiliza el número de baños».</p>	
bedrooms	integer	X	El número de habitaciones	Número de habitaciones; fundamental para determinar capacidad y segmentar propiedades.
beds	integer	X	El número de camas	Número de camas; importante para estimar ocupación máxima y tipo de huéspedes.
amenities	json	X	Lista de servicios	Fundamental para segmentación y análisis de confort
price	currency	X	<p>«Precio diario en moneda local.</p> <p>NOTA: el signo \$ es un artefacto técnico de la exportación, por favor ignórelo».</p>	Precio diario; variable principal para análisis económico y comparativo.
minimum_nights	integer	X	Número mínimo de noches de estancia para el anuncio (las reglas del calendario pueden ser diferentes).	Define estancias mínimas y restricciones de reserva.
maximum_nights	integer	X	Número máximo de noches de estancia para el anuncio (las reglas del calendario pueden ser diferentes).	Permite evaluar flexibilidad del anfitrión y demanda de estancias
minimum_minimum_nights	integer		El valor mínimo de noches más pequeño del calendario (mirando 365 noches en el futuro).	
maximum_minimum_nights	integer		El valor mínimo de noches más grande del calendario (mirando 365 noches en el futuro).	
minimum_maximum_nights	integer		El valor máximo de noches más pequeño del calendario (mirando 365 noches en el futuro).	

maximum_maximum_nights	integer		El valor máximo_noche más alto del calendario (mirando 365 noches hacia el futuro)	
minimum_nights_avg_ntm	numeric		El valor mínimo_noche promedio del calendario (mirando 365 noches hacia el futuro)	
maximum_nights_avg_ntm	numeric		El valor máximo_noche promedio del calendario (mirando 365 noches hacia el futuro)	
calendar_updated	date		Última fecha actualizada	
has_availability	boolean	X	[t=verdadero; f=falso]	Indica si la propiedad puede reservarse actualmente.
availability_30	integer	X	disponibilidad_x. La disponibilidad del anuncio x días en el futuro según lo determinado por el calendario. Tenga en cuenta que un anuncio puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.	Disponibilidad a 30 días, indicador de ocupación futura.
availability_60	integer	X	disponibilidad_x. La disponibilidad del anuncio x días en el futuro según lo determinado por el calendario. Tenga en cuenta que un anuncio puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.	Disponibilidad a 60 días, útil para análisis de planificación a mediano plazo.
availability_90	integer	X	disponibilidad_x. La disponibilidad del anuncio x días en el futuro según lo determinado por el calendario. Tenga en cuenta que un anuncio puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.	Disponibilidad a 90 días, análisis de ocupación estacional.
availability_365	integer	X	disponibilidad_x. La disponibilidad del anuncio x días en el futuro según lo determinado por el calendario. Tenga en cuenta que un anuncio puede no estar disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.	Disponibilidad anual, importante para entender la utilización total de la propiedad.
calendar_last_scraped	date	X	Última fecha de actualización de calendario	Útil para validar vigencia de la información.
number_of_reviews	integer	X	El número de reseñas que tiene el anuncio	Cantidad total de reseñas; indicador de popularidad y actividad del anuncio.
number_of_reviews_ltm	integer	X	El número de reseñas que tiene el anuncio (en los últimos 12 meses)	Reseñas últimos 12 meses, útil para medir tendencias recientes.
number_of_reviews_l30d	integer	X	El número de reseñas que tiene el anuncio (en los últimos 30 días)	Reseñas últimos 30 días; indicador de actividad inmediata
first_review	date	X	La fecha de la primera reseña/la más antigua	Fecha de primera reseña; permite analizar historial y antigüedad del anuncio.
last_review	date	X	La fecha de la última reseña/la más reciente	Fecha de última reseña; útil para identificar actividad reciente.
review_scores_rating		X	Puntuación general del anuncio	Puntuación general, indicador de satisfacción del huésped.
review_scores_accuracy		X	Puntuación de exactitud	Precisión de la descripción; útil para confianza del huésped.
review_scores_cleanliness		X	Puntuación de limpieza	Limpieza, clave en percepción de calidad.
review_scores_checkin		X	Puntuación del check-in	Facilidad de check-in; indicador de experiencia de usuario.
review_scores_communication		X	Puntuación de comunicación	Comunicación anfitrión-huésped; indicador de servicio.
review_scores_location		X	Puntuación de ubicación	Valoración de la ubicación; útil para análisis geográfico.
review_scores_value		X	Puntuación calidad-precio	Valoración de la ubicación; útil para análisis geográfico.
license	text	X	El número de licencia/permiso/registro	cumplimiento normativo y legalidad.
instant_bookable	boolean	X	[t=verdadero; f=falso]. Si el huésped puede reservar automáticamente el anuncio sin que el anfitrión tenga que aceptar su solicitud de reserva. Un indicador de un anuncio comercial.	Si puede reservarse al instante; relevante para análisis de conveniencia.
calculated_host_listings_count	integer	X	El número de anuncios que tiene el anfitrión en el rastreo actual, en la ciudad/región geográfica.	Total de anuncios del anfitrión; permite evaluar profesionalización.
calculated_host_listings_count_entire_homes	integer	X	El número de anuncios de casas/apartamentos completos que tiene el anfitrión en el rastreo actual, en la ciudad/región geográfica.	Casas/apartamentos completos; útil para segmentación por tipo de alojamiento.
calculated_host_listings_count_private_rooms	integer	X	El número de anuncios de habitaciones privadas que tiene el anfitrión en el rastreo actual, en la ciudad/región geográfica.	Habitaciones privadas; analiza oferta específica.
calculated_host_listings_count_shared_rooms	integer	X	El número de anuncios de habitaciones compartidas que tiene el anfitrión en el rastreo actual, en la ciudad/región geográfica.	Habitaciones compartidas; segmentación y demanda específica.
reviews_per_month	numeric	X	El número promedio de reseñas por mes que tiene el anuncio a lo largo de su vida útil. Psuedocoe/~SQL: IF scrape_date - first_review <= 30 THEN number_of_reviews ELSE number_of_reviews / ((scrape_date - first_review + 1) / (365/12))	Promedio de reseñas por mes; indicador de popularidad y rotación de huéspedes.