

Raymundo Díaz

Ciudad
País

Berlín
Alemania

index	Columna	Tipo	Importancia	Descripción
0	id	categorica	Clave primaria para unir con calendarios/reseñas y rastrear cambios.	Airbnb unique identifier for the listing.
1	listing_url	categorica	Permite auditorías puntuales y ver contexto/visual del anuncio.	URL to the listing’s page on Airbnb.
3	last_scraped	categorica	Define vigencia del dato; filtrar por actualidad.	UTC timestamp when this listing was last scraped.
4	source	categorica	Distinguir sesgos por método de captura (barrido vs histórico).	How the listing was found (neighbourhood search or previous scrape).
5	name	categorica	Útil para limpieza (detección de spam/comerciales) y NLP rápido.	Listing title provided by the host.
9	host_id	categorica	Llave para análisis de multilistado y profesionalización.	Airbnb unique identifier for the host.
10	host_url	categorica	Apoyo para validaciones puntuales fuera de modelo.	URL of the host’s profile page.
11	hots_name	text	El nombre de cada host, lo que permite insights más claros	Nombre de cada Host en la plataforma
12	host_since	categorica	Antigüedad del host; proxy de experiencia/confiabilidad.	Date the host profile was created.
13	host_location	categorica	Distancia host–listado; señales de gestión profesional.	Host’s self-reported location.
15	host_response_time	categorica	Velocidad de respuesta; impacto en conversión.	Typical response time to inquiries.
16	host_response_rate	numérica	Probabilidad de respuesta; afecta experiencia.	Percentage of messages the host responds to.
17	host_acceptance_rate	numérica	Tasa de aceptación; fricción de reserva.	Percentage of booking requests the host accepts.
18	host_is_superhost	boolean	Calidad/fiabilidad; efecto en precio y ocupación.	Whether the host has Superhost status (t/f).
21	host_neighbourhood	categorica	Comparar con barrio geocodificado; coherencia.	Neighbourhood specified by the host.
22	host_listings_count	numérica	Multilistado (según Airbnb); detectar operadores.	Number of listings the host has (per Airbnb).
23	host_total_listings_count	numérica	Portafolio total; segmentar profesional vs casual.	Total number of listings the host has (per Airbnb).
24	host_verifications	categorica	Confianza/seguridad del perfil.	Verification methods the host has completed.
25	host_has_profile_pic	categorica	Signal de legitimidad; filtrar cuentas vacías.	Whether the host has a profile picture (t/f).
26	host_identity_verified	categorica	Trust & safety; relacionar con pricing y reviews.	Whether the host’s identity is verified (t/f).
27	neighbourhood	categorica	Etiqueta libre; útil para QA y mappings.	Neighbourhood name (free text).
28	neighbourhood_cleansed	categorica	Barrio oficial; base para mapas y comparativos.	Neighbourhood label mapped via geocoded boundaries.
29	neighbourhood_group_cleansed	categorica	Agregación territorial; análisis macro por distrito.	Neighbourhood group label mapped via geocoded boundaries.
30	latitude	numérica	Geoespacial; proximidad a POIs/reglas de zona.	Listing latitude (WGS84).
31	longitude	numérica	Geoespacial; proximidad a POIs/reglas de zona.	Listing longitude (WGS84).
32	property_type	categorica	Tipo de activo; drivers de precio/ocupación.	Self-selected property type (e.g., apartment, house, hotel).
33	room_type	categorica	Segmento clave de mercado; restricciones regulatorias.	Room type (Entire home/apt, Private room, Shared room, Hotel).
34	accommodates	numérica	Capacidad; matching de demanda y revenue.	Maximum number of guests allowed.
35	bathrooms	numérica	Calidad/confort; contribuye a precio.	Number of bathrooms (numeric).
36	bathrooms_text	categorica	Granularidad (privadas/compartidas); limpieza de dato.	Bathroom info as text (e.g., '1.5 shared baths').
37	bedrooms	numérica	Capacidad/cómodo; correlaciona con precio.	Number of bedrooms.
38	beds	numérica	Capacidad efectiva; sensibilidad en grupos.	Number of beds.
39	amenities	categorica	Oferta de valor; features para pricing/segmentación.	JSON list of amenities included.
40	price	numérica	Precio base; insumo para ADR y comparativos.	Daily price in local currency (symbol is an export artifact).
41	minimum_nights	numérica	Restricciones de estancia; impacto regulatorio (Berlin).	Minimum nights required for a stay.
42	maximum_nights	numérica	Señal de estrategia (turismo vs mid-term).	Maximum nights allowed for a stay.
47	minimum_nights_avg_ntm	numérica	Promedio de reglas; modelado de ocupación.	Average minimum_nights over next 365 days.
48	maximum_nights_avg_ntm	numérica	Promedio de techos; stay pattern.	Average maximum_nights over next 365 days.
51	availability_30	numérica	Disponibilidad y comportamiento	Disponibilidad en 30 días
52	availability_60	numérica	Disponibilidad y comportamiento	Disponibilidad en 60 días

53	availability_90	numérica	Disponibilidad y comportamiento	Disponibilidad en 90 días
54	availability_365	numérica	Disponibilidad y comportamiento	Disponibilidad en 365 días
56	number_of_reviews	numérica	Conocer la popularidad	
61	estimated_occupancy_1365d	numérica	Estimar su ocupabilidad anual	Ocupación
62	estimated_revenue_1365d	numérica	Estimar las ganancias anuales	Ingresos
65	review_scores_rating	numérica	Conocer su posicionamiento	Calificaciones
67	review_scores_cleanliness	numérica	Conocer el estado de limpieza	Calificaciones
70	review_scores_location	numérica	Conocer la importancia de su ubicación	Calificaciones
71	review_scores_value	numérica	Conocer el puntaje	Calificaciones
78	reviews_per_month	numérica	Conocer las personas que se dieron el tiempo de dejar una reseña	Calificaciones