



Tecnológico de Monterrey

Instituto Tecnológico y de Estudios Superiores de Monterrey
Campus Puebla

Analítica de datos y herramientas de inteligencia artificial I

Gpo 101

Profesor:

Alfredo García Suárez

Actividad 1: Entendimiento de los datos

Estudiante:

Jorge Alberto Cortes Sánchez | A01736236

19 de agosto del 2025

Oporto / Portugal

3) Justificar la selección de cada variable en una tabla

Variable original (Airbnb)	Traducción al español	Tipo	Justificación de selección
id	ID del anuncio	Numérica	Identificador único, clave para análisis y limpieza de datos.
name	Nombre del anuncio	Categórica	Puede dar contexto al tipo de propiedad, usado para análisis de texto.
host_id	ID del anfitrión	Numérica	Permite analizar comportamiento de anfitriones (cuántos anuncios maneja cada uno).
host_name	Nombre del anfitrión	Categórica	Identificación complementaria, útil para duplicados o reputación.
neighbourhood_group	Grupo de vecindario	Categórica	Segmentación geográfica clave para el análisis espacial.
neighbourhood	Vecindario	Categórica	Permite análisis más detallados de localización y demanda.
latitude	Latitud	Numérica	Importante para análisis geoespacial y mapas.
longitude	Longitud	Numérica	Permite geolocalizar propiedades.
room_type	Tipo de habitación	Categórica	Fundamental para diferenciar entre alojamiento

			completo, privado o compartido.
price	Precio por noche	Numérica	Variable objetivo principal en muchos análisis de negocio.
minimum_nights	Noches mínimas	Numérica	Impacta en accesibilidad y rentabilidad de anuncios.
number_of_reviews	Número de reseñas	Numérica	Métrica de popularidad y confianza del anuncio.
last_review	Última reseña	Categórica (fecha)	Sirve para medir actualidad/actividad del anuncio.
reviews_per_month	Reseñas por mes	Numérica	Mide consistencia de la demanda.
calculated_host_listings_count	Número de anuncios del anfitrión	Numérica	Identifica anfitriones individuales vs. empresas.
availability_365	Disponibilidad (días/año)	Numérica	Relacionado con la oferta real disponible.
accommodates	Capacidad de huéspedes	Numérica	Factor clave en fijación de precios y segmentación.
bathrooms	Número de baños	Numérica	Indicador de comodidad y calidad del anuncio.
bedrooms	Número de recámaras	Numérica	Relevante para diferenciar tamaño del alojamiento.
beds	Número de camas	Numérica	Complemento de capacidad y confort.

bed_type	Tipo de cama	Categórica	Afecta preferencia del cliente y experiencia.
amenities	Amenidades	Categórica (lista)	Información crítica sobre valor agregado y diferenciación.
square_feet	Metros cuadrados	Numérica	Variable de tamaño físico, útil para correlación con precio.
property_type	Tipo de propiedad	Categórica	Diferencia casas, departamentos, villas, etc.
cancellation_policy	Política de cancelación	Categórica	Impacta en reservas y confianza del cliente.
instant_bookable	Reserva instantánea	Categórica	Facilita la experiencia del usuario.
host_since	Fecha de registro del anfitrión	Categórica (fecha)	Mide experiencia del anfitrión en la plataforma.
host_response_time	Tiempo de respuesta del anfitrión	Categórica	Factor de calidad de servicio.
host_response_rate	Tasa de respuesta del anfitrión	Numérica	Influye en confiabilidad percibida.
host_acceptance_rate	Tasa de aceptación del anfitrión	Numérica	Afecta probabilidad de concretar una reserva.
host_is_superhost	Es superanfitrión	Categórica	Indicador de calidad y confianza.
host_listings_count	Número de anuncios del host	Numérica	Identifica hosts profesionales vs. particulares.

host_total_listings_count	Total de anuncios del host	Numérica	Relevante para medir concentración de oferta.
review_scores_rating	Calificación promedio	Numérica	Indicador global de calidad.
review_scores_accuracy	Calificación de precisión	Numérica	Evalúa qué tan real es la descripción del anuncio.
review_scores_cleanliness	Calificación de limpieza	Numérica	Factor determinante para clientes.
review_scores_checkin	Calificación del check-in	Numérica	Relacionado con experiencia de llegada.
review_scores_communication	Calificación de comunicación	Numérica	Afecta confianza huésped-anfitrión.
review_scores_location	Calificación de ubicación	Numérica	Determina la percepción del entorno.
review_scores_value	Calificación de relación valor/precio	Numérica	Evalúa satisfacción del cliente en costo-beneficio.
reviews_per_month	Reseñas por mes	Numérica	Indicador de dinamismo de reservas.
is_business_travel_ready	Apto para viajes de negocio	Categórica	Segmenta mercado corporativo.
license	Licencia de operación	Categórica	Importante para regulaciones locales.
requires_license	Requiere licencia	Categórica	Indicador de cumplimiento regulatorio.
jurisdiction_names	Jurisdicción	Categórica	Afecta regulaciones y análisis de mercado.

weekly_price	Precio semanal	Numérica	Útil para segmentación de estadías largas.
monthly_price	Precio mensual	Numérica	Relevante para estancias prolongadas.
security_deposit	Depósito de seguridad	Numérica	Impacta en accesibilidad y confianza.
cleaning_fee	Tarifa de limpieza	Numérica	Costos extra a considerar en análisis de precios.
extra_people	Cargo por persona extra	Numérica	Factor en análisis de ingresos adicionales.

4) Realizar la traducción a español y la descripción de cada una de las variables de la base de datos con ayuda del “Data Dictionary”

Variable en inglés	Variable en español	Descripción en español
id	id	Identificador único de la propiedad en Airbnb.
listing_url	URL del anuncio	Enlace a la página del anuncio en Airbnb.
name	nombre	Nombre del anuncio de la propiedad.
description	descripción	Descripción detallada del alojamiento.

neighborhood_overview	descripción_vecindario	Descripción que el anfitrión proporciona sobre el vecindario.
picture_url	URL_imagen	Enlace a la imagen principal del anuncio.
host_id	id_anfitrión	Identificador único del anfitrión en Airbnb.
host_name	nombre_anfitrión	Nombre del anfitrión (generalment e solo el primer nombre).
host_since	anfitrión_desde	Fecha en que el anfitrión se registró en Airbnb.
host_location	ubicación_anfitrión	Ubicación autodeclarada por el anfitrión.
host_response_time	tiempo_respuesta_anfitrión	Tiempo promedio en que el anfitrión responde a solicitudes.
host_response_rate	tasa_respuesta_anfitrión	Porcentaje de solicitudes respondidas por el anfitrión.

host_acceptance_rate	tasa_aceptación_anfitrión	Porcentaje de solicitudes aceptadas por el anfitrión.
host_is_superhost	es_superanfitrión	Indica si el anfitrión es considerado "Superhost" (t/f).
host_listings_count	cantidad_anuncios_anfitrión	Número de anuncios que el anfitrión tiene (según Airbnb).
host_total_listings_count	total_anuncios_anfitrión	Número total de anuncios del anfitrión.
host_identity_verified	identidad_verificada	Indica si la identidad del anfitrión ha sido verificada (t/f).
neighbourhood_cleansed	vecindario_limpio	Vecindario estandarizado o según coordenadas.
latitude	latitud	Latitud de la propiedad (WGS84).
longitude	longitud	Longitud de la propiedad (WGS84).
property_type	tipo_propiedad	Tipo de propiedad seleccionado por el anfitrión (ej.

		casa, hotel, B&B).
room_type	tipo_habitación	Tipo de habitación: vivienda entera, habitación privada o compartida.
accommodates	capacidad	Número máximo de huéspedes que puede alojar la propiedad.
bathrooms_text	baños_texto	Número de baños, expresado en texto.
bedrooms	dormitorios	Número de dormitorios en la propiedad.
beds	camas	Número total de camas en la propiedad.
amenities	amenidades	Lista de servicios disponibles (ej. wifi, cocina, TV).
price	precio	Precio diario en la moneda local.
minimum_nights	noches_mínimas	Número mínimo de noches para reservar.

maximum_nights	noches_máximas	Número máximo de noches permitido en la reserva.
availability_30	disponibilidad_30	Disponibilidad del anuncio en los próximos 30 días.
availability_60	disponibilidad_60	Disponibilidad del anuncio en los próximos 60 días.
availability_90	disponibilidad_90	Disponibilidad del anuncio en los próximos 90 días.
availability_365	disponibilidad_365	Disponibilidad del anuncio en los próximos 365 días.
number_of_reviews	número_reseñas	Número total de reseñas recibidas.
number_of_reviews_ltm	reseñas_últimos_12m	Número de reseñas en los últimos 12 meses.
number_of_reviews_l30d	reseñas_últimos_30d	Número de reseñas en los últimos 30 días.
first_review	primera_reseña	Fecha de la primera reseña recibida.

last_review	última reseña	Fecha de la última reseña recibida.
review_scores_rating	puntaje_calificación	Calificación general promedio del anuncio.
review_scores_accuracy	puntaje_precisión	Calificación de la precisión de la descripción del anuncio.
review_scores_cleanliness	puntaje_limpieza	Calificación de la limpieza del alojamiento.
review_scores_checkin	puntaje_checkin	Calificación del proceso de entrada (check-in).
review_scores_communication	puntaje_comunicación	Calificación de la comunicación con el anfitrión.
review_scores_location	puntaje_ubicación	Calificación de la ubicación del alojamiento.
review_scores_value	puntaje_valor	Calificación de la relación calidad-precio.
instant_bookable	reserva_inmediata	Indica si el anuncio puede reservarse de inmediato (t/f).

calculated_host_listings_count	anuncios_calculados	Número de anuncios activos del anfitrión en el área.
calculated_host_listings_count_entire_homes	anuncios_viviendas_enteras	Número de viviendas enteras activas del anfitrión.
calculated_host_listings_count_private_rooms	anuncios_habitaciones_privadas	Número de habitaciones privadas activas del anfitrión.
reviews_per_month	reseñas_por_mes	Promedio de reseñas recibidas por mes desde la publicación.