

ENERO 2025

INTELIGENCIA DE NEGOCIOS

ACTIVIDA1 (ENTENDIMIENTO DE LOS DATOS)

CIUDAD:VENECIA

ROCÍO RAMÍREZ FABIÁN
202153429

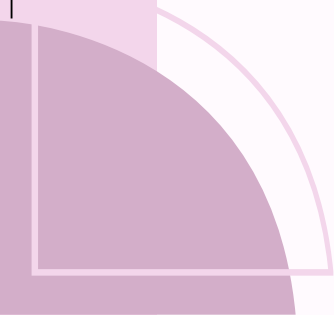


INTRO

Airbnb es una plataforma en línea que conecta a personas que buscan alojamiento con anfitriones que ofrecen sus propiedades. Los anfitriones pueden ser personas que tienen una casa, apartamento, habitación o cualquier otro espacio disponible.

Para que un espacio pueda ser anunciado en Airbnb, el anfitrión debe cumplir con ciertos requisitos, como: Respetar la normativa local, Obtener los permisos necesarios, Asegurar que la propiedad está en buenas condiciones, Describir con detalle el alojamiento, Incluir fotos atractivas.

Los viajeros pueden buscar alojamiento en Airbnb ingresando el lugar y las fechas de su viaje. Pueden utilizar filtros para encontrar el alojamiento que más se ajuste a sus necesidades, como el precio, el número de habitaciones o la ubicación.



INFORME DE VARIABLES DEL DATASET

id	listing_url	scrape_id	last_scraped	source	name	description
----	-------------	-----------	--------------	--------	------	-------------

id

6623

INT(tipo entero)

Identificador único de Airbnb para el anuncio

scrape_id

raspar id

BIGINT (tipo entero grande)

Dentro de Airbnb "Scrape" esto era parte de

last_scraped

ultimo raspado

DateTime(tipo fecha y hora)

La fecha y hora en que se "eliminó" este listado.

source

fuentes

Text(tipo texto)

Uno de "búsqueda de barrio" o "raspado previo". "búsqueda de 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 todavía estaba disponible en el sitio de Airbnb.

name

nombre

Text(tipo texto)

Nombre del listado

description

descripción

Text(tipo texto)

Descripción detallada del listado

neighborhood overview

descripción general del vecindario

Text(tipo texto)

Descripción del anfitrión del vecindario.

picture url

URL_imagen

Text(tipo texto)

URL de la imagen de tamaño normal alojada en Airbnb para el anuncio.

host id

identificación del host

Integer(entero)

Identificador único de Airbnb para el anfitrión/usuario

host url

url del host

Text(text)

La página de Airbnb para el anfitrión.

host name

nombre del host

Text(tipo de txt)

Nombre del anfitrión. Por lo general, solo el nombre(s).

host_since

anfitrión_desde

Date(tipo fecha)

La fecha en que se creó el host/usuario. Para los anfitriones que son invitados de Airbnb, esta podría ser la fecha en la que se registraron como invitados.

host_location

ubicación host

Text(tipo texto)

La ubicación autoinformada por el anfitrión.

host_about

anfitrión sobre

Text(tipo texto)

Descripción sobre el anfitrión.

picture_url

URL_imagen

Text(tipo texto)

URL de la imagen de tamaño normal alojada en Airbnb para el anuncio.

host_id

identificación del host

Integer(entero)

Identificador único de Airbnb para el anfitrión/usuario

host_response_time

tiempo de respuesta del host

host_acceptance_rate

tasa de aceptación del anfitrión

Esa tarifa a la que un anfitrión acepta solicitudes de reserva.

host_is_superhost

La anfitriona es superhost

Boolean(tipo verdadero o falso)

Nombre del anfitrión. Por lo general, solo el nombre(s).

host_thumbnail_url

URL en miniatura del host

Text(tipo texto)

host_picture_url

URL de la imagen del host

Text(tipo texto)

host_neighbourhood

barrio anfitrión

Text(tipo texto)

host_listings_count

recuento de listados de anfitriones

Text(tipo texto)

La cantidad de anuncios que tiene el anfitrión (según cálculos desconocidos de Airbnb)

host_total_listings_count

recuento total de listados de anfitriones

Text(tipo texto)

La cantidad de anuncios que tiene el anfitrión (según cálculos desconocidos de Airbnb)

host_verifications

verificaciones de host

host_has_profile_pic

La anfitriona tiene foto de perfil

host_identity_verified

identidad del host verificada

neighbourhood

vecindario

neighbourhood_cleansed

barrio limpio

Text(tipo texto)

El vecindario está geocodificado utilizando la latitud y longitud frente a los vecindarios definidos por archivos de forma digitales abiertos o públicos.

neighbourhood_group_cleansed

grupo de vecinos limpiado

Text(tipo texto)

El grupo de vecinos está geocodificado utilizando la latitud y longitud de los barrios definidos por archivos de forma digitales abiertos o públicos.

latitude

latitud

Numerico(tipo numerico)

Utiliza la proyección del Sistema Geodésico Mundial (WGS84) para latitud y longitud.

longitude

longitud

Numerico(tipo numerico)

Utiliza la proyección del Sistema Geodésico Mundial (WGS84) para latitud y longitud.

property_type

tipo de propiedad

Text(tipo texto)

Tipo de propiedad autoseleccionado. Los hoteles y Bed and Breakfast son descritos como tales por sus anfitriones en este campo.

room_type

tipo de habitacion

Text(tipo texto)

[Casa entera/apto|Habitación privada|Habitación compartida|Hotel]

Todas las viviendas se agrupan en los siguientes tres tipos de habitaciones:

Todo el lugar

habitación privada

Habitación compartida

Todo el lugar: Los lugares enteros son mejores si buscas un hogar lejos del hogar. Con un alojamiento completo, 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 proporcionar más detalles sobre el anuncio.

Habitaciones privadas: Las habitaciones privadas son ideales 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 podrás compartir algunos espacios con otras personas. Es posible que tengas que caminar por espacios interiores que otro anfitrión o huésped pueda ocupar para llegar a tu habitación.

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 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 estancias económicas.

`accommodates`

se adapta

Integer(entero)

La capacidad máxima del anuncio.

`bathrooms`

baños

Numerico(tipo numerico)

El número de baños en el listado.

`bathrooms_text`

texto de baños

String(tipo cadena)

El número de baños del listado.

En el sitio web de Airbnb, el campo de los baños ha evolucionado de un número a una descripción textual. Para raspaduras más antiguas, se utilizan baños.

`bedrooms`

dormitorios

Integer(entero)

El número de dormitorios

`beds`

camas

Integer(entero)

El número de camas

`amenities`

comodidades

JSON(formato de datos de texto)

`price`

precio

currency(cálculos monetarios y de punto fijo)

Precio diario en moneda local.

NOTA: el signo \$ es un artefacto técnico de la exportación, ignórelo

`minimum_nights`

noches mínimas

INTEGER(entero)

Número mínimo de noches de estadía para el alojamiento (las reglas del calendario pueden ser diferentes)

maximum_nights

noches máximas

INTEGER(entero)

Número máximo de noches de estadía para el alojamiento (las reglas del calendario pueden ser diferentes)

minimum_minimum_nights

noches mínimas mínimas

INTEGER(entero)

el valor mínimo de noche más pequeño del calendario (mirando 365 noches en el futuro).

maximum_minimum_nights

mínimo máximo noches

INTEGER(entero)

el valor máximo noche más pequeño del calendario (mirando 365 noches en el futuro)

maximum_maximum_nights
máximo máximo de noches

INTEGER(entero)

el valor máximo_noche más grande del calendario (mirando 365 noches en el futuro).

minimum_nights_avg_ntm

noches mínimas promedio ntm

NUMERIC(Tipo numerico)

el valor promedio de noche mínima del calendario (mirando 365 noches en el futuro)

maximum_nights_avg_ntm

Promedio máximo de noches ntm

NUMERIC(Tipo numerico)

el valor promedio de máxima_noche del calendario (mirando 365 noches en el futuro)

calendar_updated

calendario actualizado

DATE(tipo fecha)

has_availability

tiene disponibilidad
BOOLEAN(boleano)

availability_30

disponibilidad 30
INTEGER(entero)

disponibilidad_x. La disponibilidad del listado x días en el futuro según lo determine el calendario. Tenga en cuenta que es posible que un anuncio no esté disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.

availability_60

disponibilidad 60
INTEGER(entero)

disponibilidad_x. La disponibilidad del listado x días en el futuro según lo determine el calendario. Tenga en cuenta que es posible que un anuncio no esté disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.

availability_90

disponibilidad 90
INTEGER(entero)

disponibilidad_x. La disponibilidad del listado x días en el futuro según lo determine el calendario. Tenga en cuenta que es posible que un anuncio no esté disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.

availability_365

disponibilidad 365
INTEGER(entero)

disponibilidad_x. La disponibilidad del listado x días en el futuro según lo determine el calendario. Tenga en cuenta que es posible que un anuncio no esté disponible porque ha sido reservado por un huésped o bloqueado por el anfitrión.

calendar_last_scraped

calendario raspado por última vez
DATE(Fecha)

number_of_reviews

número de reseñas
INTEGER(entero)

El número de reseñas que tiene el anuncio.

`number_of_reviews_ltm`

número de reseñas ltm

INTEGER(entero)

El número de reseñas que tiene el anuncio (en los últimos 12 meses)

`number_of_reviews_l30d`

número de reseñas l30d

INTEGER(entero)

El número de reseñas que tiene el anuncio (en los últimos 30 días)

`first_review`

primera revisión

DATE(fecha)

La fecha de la primera/más antigua revisión.

`last_review`

última revisión

DATE(fecha)

La fecha de la última/más reciente revisión.

`review_scores_rating`

calificación de puntuaciones de revisión

`review_scores_accuracy`

revisar la precisión de las puntuaciones

`review_scores_cleanliness`

revisar puntuaciones limpieza

`review_scores_checkin`

revisar puntajes registrarse

`review_scores_communication`

revisar puntuaciones comunicación

`review_scores_location`

revisar la ubicación de las puntuaciones

`review_scores_value`

revisar el valor de las puntuaciones

`license`

licencia

TEXT(texto)

El número de licencia/permiso/registro

`instant_bookable`

reservable al instante

BOOLEAN(booleano)

[t=verdadero; f=falso]. Si el huésped puede reservar automáticamente el alojamiento sin que el anfitrión tenga que aceptar su solicitud de reserva. Un indicador de una cotización comercial.

`calculated_host_listings_count`

recuento calculado de listados de anfitriones

INTEGER(entero)

La cantidad de listados que el anfitrión tiene en el scrap actual, en la geografía de la ciudad/región.

`calculated_host_listings_count_entire_homes`

Los listados de anfitriones calculados cuentan casas enteras.

INTEGER(entero)

La cantidad de listados completos de casas/apto. que el anfitrión tiene en el scrape actual, en la geografía de la ciudad/región.

`calculated_host_listings_count_private_rooms`

Los listados de anfitriones calculados cuentan las habitaciones privadas.

INTEGER(entero)

La cantidad de listados de habitaciones privadas que el anfitrión tiene en el scrap actual, en la geografía de la ciudad/región.

`calculated_host_listings_count_shared_rooms`

Los listados de anfitriones calculados cuentan las habitaciones compartidas.

INTEGER(entero)

La cantidad de listados de habitaciones compartidas que el anfitrión tiene en el scrap actual, en la geografía de la ciudad/región.

`reviews_per_month`

revisiones por mes

NUMERIC(numerico)

El número promedio de reseñas por mes que tiene el listado durante su vida útil.

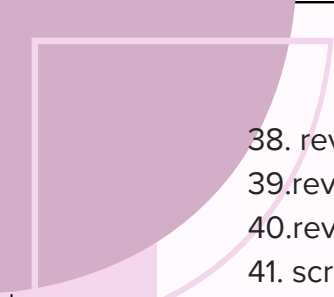

Seudocoe/~SQL:

SI $\text{fecha_scrape} - \text{primera_revisión} \leq 30$ ENTONCES
número_de_revisiones

ELSE $\text{número_de_revisiones} / ((\text{scrape_date} - \text{primera_revisión} + 1) / (365/12))$

4. 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)

- 1.id: Es el identificador único de cada propiedad. Es fundamental para cualquier análisis, ya que permite diferenciar y rastrear registros.
- 2.description: Proporciona detalles clave sobre la propiedad.
- 3.neighbourhood :Ubicación general del listado.
- 4.latitude: Coordenada para análisis geoespacial.
- 5.longitude: Complementa el análisis de ubicación.
- 6.property_type :Determina el tipo de alojamiento.
- 7.room_type: Influye en el precio y el mercado objetivo.
- 8.accommodates: Define la capacidad máxima.
- 9.host_id: Identificador único del anfitrión.
- 10.bathrooms - Afecta la comodidad percibida.
- 11.bedrooms - Influye en el precio y la demanda.
- 12.beds - Proporciona un indicador de capacidad.
- 13.amenities - Diferenciador clave entre propiedades.
- 14.host_id - Identificador único del anfitrión.
- 15.host_since - Indica experiencia en la plataforma.
- 16.host_is_superhost - Influye en la confianza de los huéspedes.
- 17.host_response_time - Mide la rapidez de respuesta.
- 18.host_response_rate - Porcentaje de mensajes respondidos.
- 19.host_acceptance_rate - Proporción de solicitudes aceptadas.
- 20.host_listings_count - Indica el número de propiedades gestionadas.
- 21.host_verifications - Factores que validan la identidad del anfitrión.
- 22.price - Es el eje central de muchos análisis.
- 23.minimum_nights - Define el público objetivo.
- 24.maximum_nights - Restricción máxima para reservas.
- 25.availability_365 - Número de días disponibles al año.
- 26.instant_bookable - Si permite reservas instantáneas.
- 27.cancellation_policy - Afecta la decisión de reserva.
- 28.security_deposit - Puede disuadir a algunos huéspedes.
- 29.cleaning_fee - Costos adicionales a considerar.
- 30.number_of_reviews - Indicador de popularidad.
- 31.first_review - Marca la antigüedad de las reseñas.
- 32.last_review - Última interacción con huéspedes.
- 33.review_scores_rating - Calificación promedio general.
- 34.review_scores_accuracy - Precisión de la descripción.
- 35.review_scores_cleanliness - Limpieza percibida.
- 36.review_scores_checkin - Experiencia al registrarse.
- 37.review_scores_communication - Calidad de comunicación.

- 
- 
- 38. review_scores_location - Opiniones sobre la ubicación.
 - 39. review_scores_value - Relación calidad-precio.
 - 40. reviews_per_month - Frecuencia de reseñas.
 - 41. scrape_id - Fecha y hora del último "scrape".
 - 42. last_scraped - Última vez que se actualizó el listado.
 - 43. source - Origen de la información.
 - 44. neighbourhood_cleansed - Ubicación específica, más detallada.
 - 45. zipcode - Para análisis a nivel de código postal.
 - 46. market - Mercado objetivo general.
 - 47. country - País del listado, para segmentación global.
 - 48. is_business_travel_ready - Si está optimizado para viajes de negocios.
 - 49. calculated_host_listings_count: Indica el número total de listados que un anfitrión tiene activos en un área específica. Es útil para identificar anfitriones profesionales o aquellos que gestionan múltiples propiedades, lo cual puede influir en la calidad y consistencia del servicio.
 - 50. extra_people: Representa el costo adicional por cada huésped que excede el límite de guests_included. Esta variable puede impactar las decisiones de reserva de grupos grandes y es clave para entender estrategias de precios.