



BENEMÉRITA UNIVERSIDAD AUTÓNOMA
DE PUEBLA

FACULTAD DE CIENCIAS DE LA
COMPUTACIÓN

INGENIERÍA EN TECNOLOGÍAS DE LA
INFORMACIÓN

MATERIA: INTELIGENCIA DE NEGOCIOS

ENTENDIMIENTO DE LOS DATOS

FACILITADOR: PHD ALFREDO GARCÍA
SUÁREZ

ALUMNO: JORGE LEON VIVANCO

MATRICULA: 202129258

- 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	Descripción
1. id	Identificador único de la propiedad asignado por Airbnb.
2. listing_url	URL del anuncio en Airbnb.
3. scrape_id	Identificador del "scrape" en el que se recopiló este registro.
4. last_scraped	Fecha y hora en UTC en la que se recopiló la información del anuncio.
5. name	Nombre del anuncio.
6. description	Descripción detallada del anuncio.
7. neighborhood_overview	Descripción del vecindario proporcionada por el anfitrión.
8. picture_url	URL de la imagen principal del anuncio.
9. host_id	Identificador único del anfitrión.
10. host_url	URL del perfil del anfitrión en Airbnb.
11. host_name	Nombre del anfitrión (normalmente solo el nombre de pila).
12. host_since	Fecha en la que el anfitrión se unió a Airbnb.
13. host_location	Ubicación reportada por el anfitrión.
14. host_about	Información proporcionada por el anfitrión sobre sí mismo.
15. host_response_time	Tiempo promedio que tarda el anfitrión en responder.
16. host_response_rate	Porcentaje de respuesta del anfitrión a las solicitudes.
17. host_acceptance_rate	Porcentaje de solicitudes aceptadas por el anfitrión.
18. host_is_superhost	Indica si el anfitrión es un "Superhost" (t = verdadero, f = falso).
19. host_thumbnail_url	URL de la miniatura de la foto del anfitrión.
20. host_picture_url	URL de la foto del anfitrión.
21. host_neighbourhood	Vecindario donde reside el anfitrión.
22. host_listings_count	Número de anuncios del anfitrión.
23. host_total_listings_count	Número total de anuncios del anfitrión.
24. host_verifications	Métodos de verificación del anfitrión.
25. host_has_profile_pic	Indica si el anfitrión tiene una foto de perfil (t = verdadero, f = falso).
26. host_identity_verified	Indica si la identidad del anfitrión está verificada (t = verdadero, f = falso).
27. neighbourhood	Vecindario donde se encuentra la propiedad.
28. neighbourhood_cleansed	Vecindario según datos geocodificados.
29. neighbourhood_group_cleansed	Grupo de vecindarios según datos

	geocodificados.
30. latitude	Latitud de la ubicación de la propiedad.
31. longitude	Longitud de la ubicación de la propiedad.
32. property_type	Tipo de propiedad (ej., casa, departamento, etc.).
33. room_type	Tipo de habitación (ej., vivienda completa, habitación privada, compartida).
34. accommodates	Capacidad máxima de la propiedad.
35. bathrooms	Número de baños.
36. bathrooms_text	Descripción textual del número de baños.
37. bedrooms	Número de habitaciones.
38. beds	Número de camas.
39. amenities	Lista de servicios incluidos.
40. price	Precio por noche en moneda local.
41. minimum_nights	Número mínimo de noches por reserva.
42. maximum_nights	Número máximo de noches por reserva.
43. minimum_minimum_nights	Valor mínimo de noches permitido según el calendario (365 días futuros).
44. maximum_minimum_nights	Valor máximo de noches permitido según el calendario (365 días futuros).
45. minimum_maximum_nights	Valor mínimo del máximo de noches permitido según el calendario (365 días futuros).
46. maximum_maximum_nights	Valor máximo del máximo de noches permitido según el calendario (365 días futuros).
47. minimum_nights_avg_ntm	Promedio del mínimo de noches según el calendario (365 días futuros).
48. maximum_nights_avg_ntm	Promedio del máximo de noches según el calendario (365 días futuros).
49. calendar_updated	Fecha de última actualización del calendario.
50. has_availability	Indica si la propiedad está disponible (t = verdadero, f = falso).
51. availability_30	Días disponibles para reserva en los próximos 30 días.
52. availability_60	Días disponibles para reserva en los próximos 60 días.
53. availability_90	Días disponibles para reserva en los próximos 90 días.
54. availability_365	Días disponibles para reserva en los próximos 365 días.
55. calendar_last_scraped	Fecha en la que se recopiló por última vez el calendario del anuncio.
56. number_of_reviews	Número total de reseñas.
57. number_of_reviews_ltm	Número de reseñas en los últimos 12 meses.
58. number_of_reviews_l30d	Número de reseñas en los últimos 30 días.
59. first_review	Fecha de la primera reseña.

60. last_review	Fecha de la última reseña.
61. review_scores_rating	Calificación promedio del anuncio.
62. review_scores_accuracy	Calificación de precisión del anuncio.
63. review_scores_cleanliness	Calificación de limpieza del anuncio.
64. review_scores_checkin	Calificación de facilidad de check-in.
65. review_scores_communication	Calificación de comunicación con el anfitrión.
66. review_scores_location	Calificación de la ubicación del anuncio.
67. review_scores_value	Calificación de la relación calidad-precio.
68. license	Número de licencia o permiso del anuncio.
69. instant_bookable	Indica si la reserva es inmediata (t = verdadero, f = falso).
70. calculated_host_listings_count	Número de anuncios actuales del anfitrión en la ciudad o región.
71. calculated_host_listings_count_entire_homes	Número de anuncios de viviendas completas del anfitrión en la ciudad o región.
72. calculated_host_listings_count_private_rooms	Número de anuncios de habitaciones privadas del anfitrión en la ciudad o región.
73. calculated_host_listings_count_shared_rooms	Número de anuncios de habitaciones compartidas del anfitrión en la ciudad o región.
74. reviews_per_month	Promedio de reseñas por mes durante la vida del anuncio.

- 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)

Variable	Justificación
1. id	Permite identificar de manera única cada propiedad en el análisis.
2. listing_url	Es útil para acceder a información detallada sobre las propiedades y confirmar datos específicos.
3. last_scraped	Indica cuándo se recopiló la información, asegurando que los datos sean actuales y relevantes.
4. name	Permite identificar anuncios de forma específica y analizar cómo los nombres pueden impactar en la atracción de clientes.
5. description	Proporciona detalles clave sobre el anuncio, útiles para evaluar las características ofrecidas y su atractivo.
6. neighborhood_overview	Ayuda a comprender cómo los anfitriones describen los vecindarios, lo que puede influir en las preferencias de los clientes.
7. host_id	Identifica anfitriones individuales, crucial para analizar su desempeño o gestionar propiedades múltiples.
8. host_name	Permite analizar reputación o preferencias de los huéspedes hacia anfitriones específicos.
9. host_since	Indica la experiencia del anfitrión en la plataforma, lo que puede relacionarse con la calidad del servicio.
10. host_location	Ayuda a entender la ubicación geográfica de los anfitriones en relación con sus propiedades.
11. host_about	Proporciona información cualitativa sobre el anfitrión que puede influir en la percepción de los huéspedes.
12. host_response_time	Mide la rapidez de respuesta del anfitrión, un factor clave en la satisfacción del cliente.
13. host_response_rate	Indica la receptividad del anfitrión a las solicitudes, lo que afecta la experiencia del cliente.
14. host_acceptance_rate	Evalúa la frecuencia con la que los anfitriones aceptan reservas, relacionado con la disponibilidad efectiva.
15. host_is_superhost	Identifica anfitriones "Superhost", cuya reputación suele estar vinculada a un mejor

	servicio y experiencia.
16. neighbourhood	Proporciona información general sobre la ubicación, esencial para analizar demanda en distintas áreas.
17. neighbourhood_cleansed	Ofrece datos geográficos precisos para analizar ubicaciones específicas y su popularidad.
18. latitude	Coordena la ubicación geográfica exacta, útil para mapas y análisis espaciales.
19. longitude	Complementa la latitud en análisis geoespaciales, como mapas de calor o segmentación por área.
20. property_type	Clasifica los tipos de propiedades (e.g., casa, departamento), lo que ayuda a entender las preferencias del mercado.
21. room_type	Segmenta los espacios en función de su uso (e.g., vivienda completa, habitación privada), relevante para identificar tendencias.
22. accommodates	Determina la capacidad máxima de cada propiedad, útil para identificar perfiles de clientes (familias, viajeros solos, etc.).
23. bathrooms_text	Describe las instalaciones sanitarias, un factor clave en la percepción de calidad de los huéspedes.
24. bedrooms	Proporciona información sobre el tamaño de la propiedad, relacionado con la segmentación de mercado.
25. beds	Ayuda a evaluar la capacidad de alojamiento y comodidad de las propiedades.
26. amenities	Lista los servicios disponibles, esenciales para analizar preferencias de los clientes y comparaciones competitivas.
27. price	Variable fundamental para analizar estrategias de precios, demanda y competitividad en diferentes mercados.
28. minimum_nights	Ayuda a entender las restricciones de los anfitriones y su impacto en la disponibilidad.
29. maximum_nights	Evalúa la flexibilidad de las propiedades para diferentes tipos de clientes (e.g., estancias cortas o largas).
30. availability_30	Muestra la disponibilidad a corto plazo, crucial para evaluar tendencias inmediatas de ocupación.
31. availability_90	Indica la disponibilidad a mediano plazo, útil para planificaciones de mercado.
32. availability_365	Permite analizar la estacionalidad y disponibilidad a largo plazo.

33. number_of_reviews	Indica la popularidad general del anuncio, relevante para evaluar su atractivo en el mercado.
34. number_of_reviews_ltm	Reseñas en los últimos 12 meses, útil para analizar el rendimiento reciente del anuncio.
35. number_of_reviews_l30d	Reseñas en los últimos 30 días, reflejando la actividad actual del anuncio.
36. first_review	Fecha de la primera reseña, útil para medir la antigüedad y estabilidad del anuncio en la plataforma.
37. last_review	Fecha de la última reseña, relevante para evaluar la actividad reciente del anuncio.
38. review_scores_rating	Calificación general, un indicador clave de satisfacción del cliente.
39. review_scores_cleanliness	Evalúa la limpieza, un aspecto crucial para la experiencia del huésped.
40. review_scores_communication	Refleja la calidad de la interacción con el anfitrión, un factor importante en la experiencia del cliente.
41. review_scores_location	Mide la percepción de los huéspedes sobre la ubicación, relevante para el análisis de demanda geográfica.
42. review_scores_value	Evalúa la relación calidad-precio, un indicador importante de competitividad del anuncio.
43. license	Verifica el cumplimiento legal del anuncio, un aspecto crítico en ciertas jurisdicciones.
44. instant_bookable	Indica si el anuncio permite reservas instantáneas, lo que puede atraer clientes de última hora.
45. calculated_host_listings_count	Muestra el número total de propiedades gestionadas por un anfitrión, relevante para análisis de eficiencia.
46. calculated_host_listings_count_entire_homes	Proporciona datos específicos sobre viviendas completas, útiles para estrategias de mercado.
47. calculated_host_listings_count_private_rooms	Identifica habitaciones privadas, clave para segmentar preferencias de los clientes.
48. reviews_per_month	Muestra la actividad de las reseñas, reflejando la popularidad y estabilidad del anuncio en el tiempo.
49. host_thumbnail_url	Permite analizar visualmente cómo los anfitriones presentan su perfil, lo que podría influir en la confianza del cliente.
50. host_identity_verified	Indica si la identidad del anfitrión está verificada, un factor importante en la percepción de seguridad del cliente.