

15 de febrero del 2025



Tecnológico de Monterrey

Actividad 1
(Entendimiento de los Datos)

Alumno:

Sergio Giles Bracamontes

Matricula:

A01425434

Unidad de formación:

Analítica de datos y herramientas
de inteligencia artificial I

Docente:

Alfredo García Suárez

Reporte Informativo: Entendimiento de los Datos de Airbnb en Porto, Norte, Portugal

Introducción

El presente reporte tiene como objetivo proporcionar un entendimiento detallado de la base de datos de Airbnb correspondiente a la ciudad de **Porto, Norte, Portugal**. A través de la traducción y descripción de las variables, así como la selección de las 50 más relevantes, se busca facilitar el análisis de los datos para identificar patrones, tendencias y oportunidades en el mercado de alojamientos turísticos en esta región.

Descripción de la Base de Datos

La base de datos utilizada, “**listings.csv**”, contiene información detallada sobre los listados de Airbnb en Porto. Incluye variables relacionadas con las propiedades, los anfitriones, las reseñas, la disponibilidad y las comodidades ofrecidas. Cada variable ha sido traducida al español y descrita para facilitar su interpretación. Por ejemplo, las categorías principales:

- **Información del Listado:** Incluye datos como el identificador único del listado, el nombre, la descripción, el tipo de propiedad y el tipo de habitación.
- **Información del Anfitrión:** Contiene detalles sobre el anfitrión, como su identificador único, nombre, fecha de registro, tasa de respuesta y estatus de Superanfitrión.
- **Ubicación:** Proporciona datos geográficos, como la latitud, longitud y vecindario del listado.
- **Reseñas:** Incluye información sobre las calificaciones y comentarios de los huéspedes, como la puntuación promedio, la limpieza y la ubicación.
- **Disponibilidad:** Ofrece datos sobre la disponibilidad del listado en diferentes periodos (30, 60, 90 y 365 días).
- **Precios y Políticas:** Contiene información sobre el precio por noche, el número mínimo y máximo de noches de estancia, y la posibilidad de reserva instantánea.

Selección de las 50 Variables más Relevantes

Se han seleccionado 50 variables considerando su importancia para el análisis de datos, priorizando aquellas de tipo numérico y categórico. Estas variables permiten realizar un análisis integral de los listados, desde la perspectiva del huésped, el anfitrión y la propiedad. Categorías de las variables seleccionadas:

Variables Numéricas: Incluyen datos cuantitativos como el precio, la capacidad de alojamiento, el número de baños, dormitorios y camas, así como las puntuaciones de las reseñas y la disponibilidad en diferentes periodos.

Variables Categóricas: Abarcan información como el tipo de propiedad, el tipo de habitación, el vecindario y el estatus de Superanfitrión.

Variables Relacionadas con el Anfitrión: Contienen datos como la tasa de respuesta, la tasa de aceptación, la cantidad de listados y la verificación de identidad.

Variables de Reseñas: Incluyen el número total de reseñas, la puntuación promedio y las calificaciones específicas (limpieza, comunicación, ubicación, etc.).

La selección de estas variables se justifica por su capacidad para proporcionar insights sobre el mercado de alojamientos en Porto, permitiendo analizar aspectos como la calidad de los listados, el perfil de los anfitriones, las preferencias de los huéspedes y la dinámica de precios.

Análisis de las Variables Seleccionadas

Las variables seleccionadas permiten abordar diversas preguntas de investigación, tales como:

¿Cuál es el perfil típico de un anfitrión en Porto?

Variables como `host_since`, `host_response_rate`, `host_is_superhost` y `host_listings_count` proporcionan información sobre la experiencia, responsabilidad y escala de operaciones de los anfitriones.

¿Qué factores influyen en el precio de los listados?

Variables como `price`, `property_type`, `room_type`, `accommodates` y `amenities` permiten analizar cómo las características de la propiedad y las comodidades ofrecidas afectan el precio.

¿Cómo es la experiencia del huésped en los listados de Porto?

Variables como `review_scores_rating`, `review_scores_cleanliness`, `review_scores_communication` y `number_of_reviews` ofrecen insights sobre la satisfacción de los huéspedes y la calidad de los servicios.

¿Cuál es la disponibilidad de los listados a lo largo del año?

Variables como `availability_30`, `availability_60`, `availability_90` y `availability_365` permiten evaluar la ocupación y la demanda en diferentes periodos.

Conclusión

El análisis de las 50 variables seleccionadas proporciona un entendimiento eficiente sobre el mercado de alojamientos de Airbnb en Porto, Norte, Portugal. Estas variables permiten explorar aspectos clave como la calidad de los listados, el comportamiento de los anfitriones, las preferencias de los huéspedes y la dinámica de precios