

Ficha Técnica: Business Intelligence

Título del Proyecto: Business Intelligence

Objetivo:

Proporcionar una base sólida para la toma de decisiones informada, permitiendo a los interesados tomar medidas estratégicas para mejorar la eficiencia operativa y la rentabilidad en el dinámico ecosistema de Airbnb.

Equipo:

Individual.

Herramientas y Tecnologías:

Bigquery, Power BI, Google Docs, Google Slices.

Insumos:

CSV Superstore.

CSV Competencia.

Diccionario de Datos:

Tabla rooms (Dimensión):

id: un identificador único para cada habitación.

name: el nombre del anuncio de Airbnb

neighbourhood: acrónimo del barrio en el que se encuentra el anuncio de Airbnb

neighbourhoodgroup: barrio en el que se encuentra el anuncio de Airbnb

latitude: la coordenada de latitud del anuncio de Airbnb

longitude: la coordenada de longitud del anuncio de Airbnb

roomtype: el tipo de habitación que ofrece el anuncio de Airbnb

minimum_nights: el número mínimo de noches necesarias para reservar el anuncio de Airbnb.

Tabla hosts (Dimensión):

hostid : un identificador único para cada host.

hostname: el nombre del anfitrión del anuncio de Airbnb.

Tabla reviews (Hechos):

id: un identificador único para cada habitación.

hostid : un identificador único para cada host.

price: el precio por noche del anuncio de Airbnb

numberofreviews: el número total de reseñas que ha recibido el anuncio de Airbnb

lastreview: la fecha de la última reseña que recibió el anuncio de Airbnb

reviewspermonth: El número promedio de reseñas que recibe el anuncio de Airbnb por mes

calculatedhostlistingscount: el número total de listados que tiene el anfitrión

availability365: la cantidad de días que el anuncio de Airbnb está disponible para reservar en un año.

Crear y preparar la Base de Datos:

- Conectar/importar datos a otras herramientas.

Dataset

hosts

reviews

rooms

Host: 37485 registros.

Reviews: 48875 registros.

Rooms: 48875 registros.

- Identificar y manejar valores nulos.

Hosts

Fila	NULL_host_id	NULL_host_name
1	0	18

Reviews

Fila	NULL_id	NULL_host_id	NULL_price	NULL_number_of_re	NULL_last_review	NULL_reviews_per_r	NULL_price_1	NULL_calculated_ho	NULL_availability_36
1	0	0	0	20	10039	10019	0	0	156

Rooms

NULL_id	NULL_name	NULL_neighbourho	NULL_neighbourho	NULL_latitude	NULL_longitude	NULL_room_type	NULL_minimum_nigt
0	0	0	0	0	0	0	0

- Identificar y manejar valores duplicados.

No se encontraron registros duplicados en las tablas.

- Identificar y manejar datos discrepantes en variables categóricas y variables numéricas.

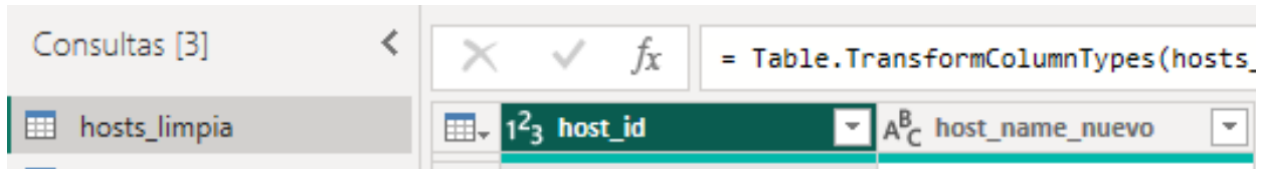
Se creó la tabla **hosts_limpia** con **36881** registros , luego de quitar nulos y datos discrepantes.

Se creó la tabla **reviews_limpia** con **38694** registros, luego de quitar nulos y datos discrepantes.

Se creó la tabla **rooms_limpia** con **48710** registros, luego de quitar nulos y datos discrepantes.

- Comprobar y cambiar tipo de dato

Hosts



Se cambió el `host_id` de la tabla **hosts_limpiar** a **número entero** en power bi.

Reviews_limpiar

hosts_limpiar	1 ² id	1 ² host_id	1 ² price	1 ² number_of_reviews	last_review
reviews_limpiar	1	34399221	6279234	148	1
					08-07-201

Se cambió el `id`, `host_id`, `price`, y `number_of_reviews` a **número entero** y el campo `last_review` a **fecha**, de la tabla **hosts_limpiar** en power bi.

Rooms_limpiar

hosts_limpiar	1 ² id	A ^B name	A ^B room_type_nuevo	A ^B neighbourhood_nuevo
reviews_limpiar	1	10000070	Spacious Large Private Room By Pr. Park w Aircon	private room
rooms_limpiar	2	100002	MANHATTAN Neat, Nice, Bright ROOM	private room
				washington heights

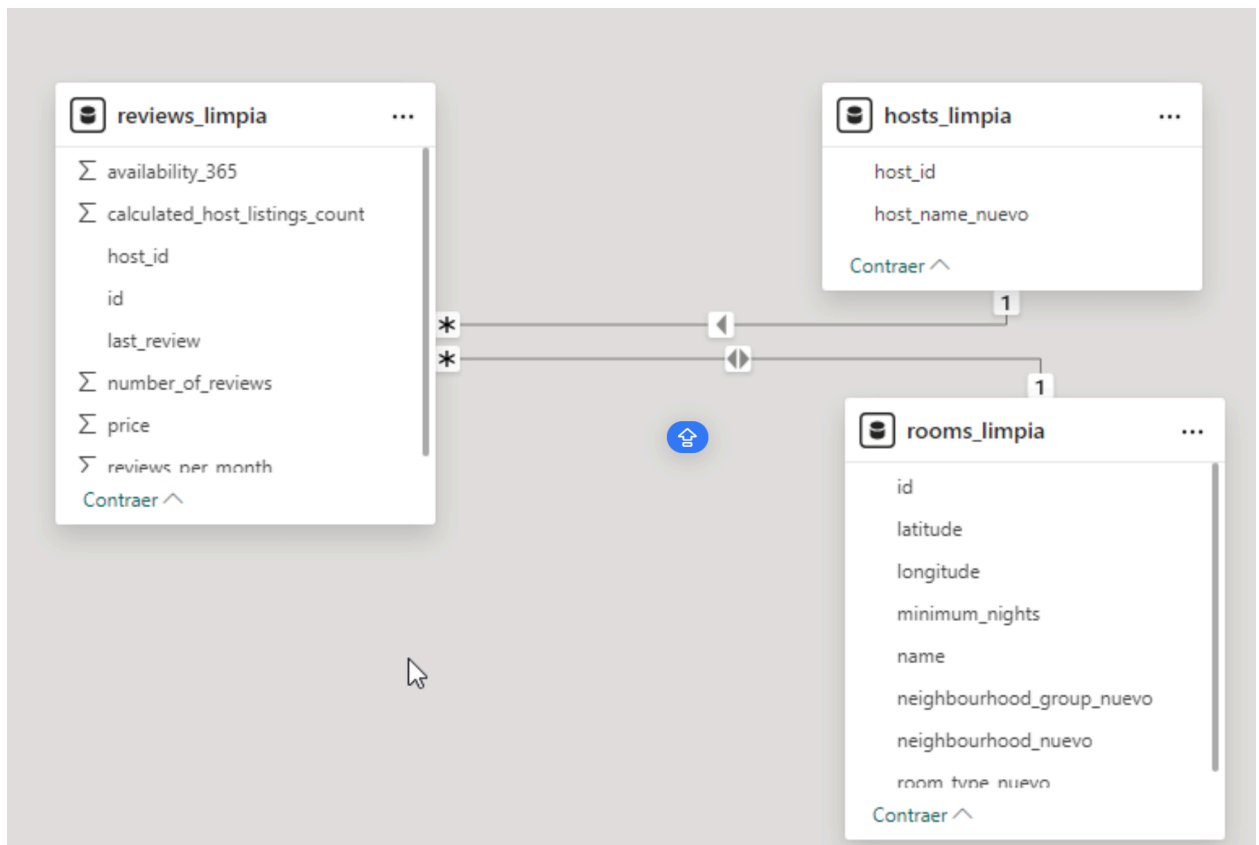
Se cambió el `id` a **número entero** de la tabla **rooms_limpiar** en power bi.

- Relacionar tablas

Se relacionaron las tablas:

host_limpiar y **reviews_limpiar** por el campo **host_id**.

rooms_limpiar y **reviews_limpiar** por el campo **id**.



- Crear nuevas variables

1 precio_x_hab = reviews_limpiar[price] * reviews_limpiar[availability_365]

1 id_review = COUNTROWS(FILTER('reviews_limpiar','reviews_limpiar'[id] <= EARLIER(reviews_limpiar[id])))

1 Año = YEAR('reviews_limpiar'[last_review])

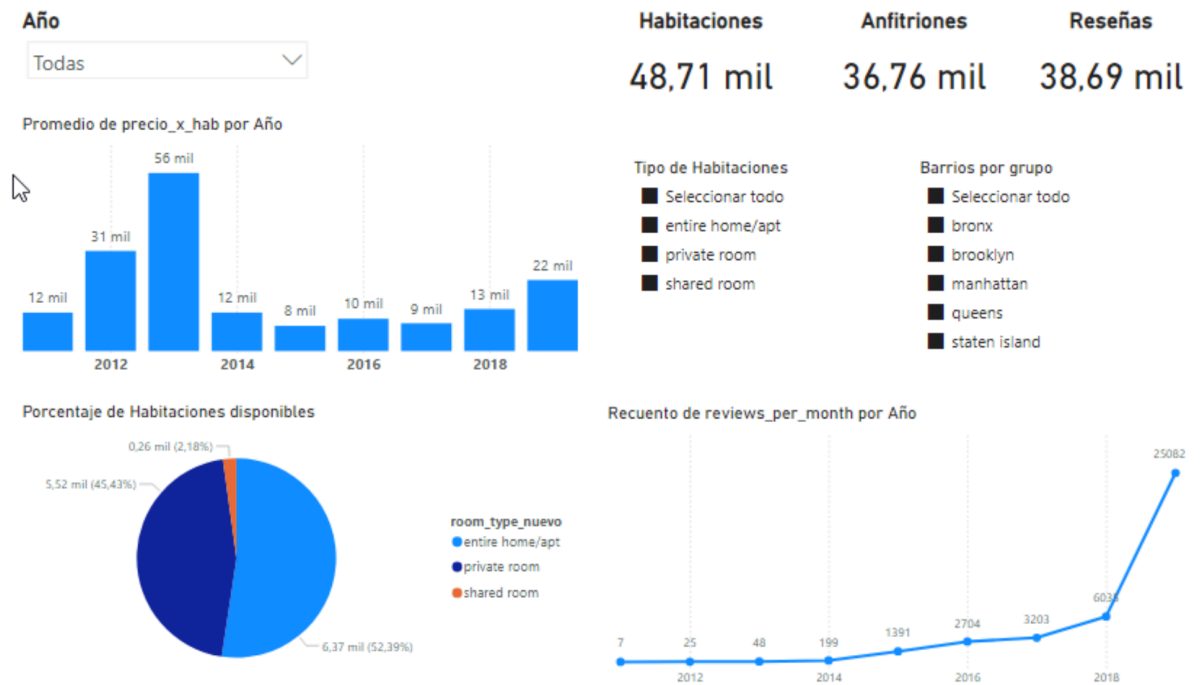
- Fórmulas DAX

1 porc_hab_disp = DIVIDE(SUM([availability_365]),COUNT([id])*365,0)

1 Total_huespedes_x_anio = 365/rooms_limpiar[minimum_nights]

Análisis exploratorio

- Agrupar datos según variables categóricas, visualizar las variables categóricas, visualizar el comportamiento de los datos a lo largo del tiempo.



- Aplicar medidas de tendencia central, visualizar distribución, aplicar medidas de dispersión.



Conclusiones:

- Barrios como Manhattan y Brooklyn presentan una demanda sostenida, reflejada en la alta cantidad de reseñas pero con baja disponibilidad. Esto indica una ocupación constante y rápida rotación de los alojamientos, lo que representa una oportunidad para los anfitriones actuales y potenciales.
- El análisis del precio promedio por tipo de habitación y barrio revela una diferencia significativa entre áreas. En barrios con alta demanda como Manhattan y Brooklyn, los precios son más elevados, lo que sugiere que los anfitriones podrían beneficiarse aún más ajustando las tarifas en función de la estacionalidad o la demanda.
- La disponibilidad de alojamientos ha crecido significativamente en varios barrios a lo largo de los

años. Este crecimiento sugiere que más anfitriones están entrando en el mercado, lo que podría influir en la competencia y, potencialmente, en los precios.

- Algunos barrios como Queens y Staten Island tienen menor cantidad de reseñas y baja disponibilidad. Esto podría ser una señal de menor demanda, pero también una oportunidad para atraer a más huéspedes a través de estrategias de marketing o precios competitivos.

Recomendaciones:

- Basado en la comparación de precios por barrio, se podrían ajustar los precios en temporadas altas o en barrios con alta demanda como Manhattan, aumentar las tarifas o implementar promociones específicas para captar más clientes maximizando la rentabilidad.
- Incentivar a los anfitriones a ofrecer más propiedades en áreas con alta demanda y poca disponibilidad, mejorando así el balance entre oferta y demanda.
- Los datos geográficos sugieren que hay barrios con menos saturación, lo que puede representar una oportunidad de expansión para captar más mercado.
- La estacionalidad es clara en los datos de disponibilidad a lo largo del año, por lo que una recomendación sería prepararse para la alta demanda en ciertas épocas del año y ajustar la oferta y precios de acuerdo con las tendencias.
- Expandir las operaciones a zonas con menos presencia de alojamientos pero con potencial de crecimiento de acuerdo con los datos.

Limitaciones:

Cantidad de información en en el Dataset.