

Proyecto de Análisis de Datos

# **Análisis, Visualización y Exploración de Datos de Airbnb Málaga**

Autor: Isidro Javier García Fernández

Agosto de 2023



# Índice

<b>1. Introducción</b>	<b>4</b>
1.1. Introducción	4
1.2. El Motor del Análisis de Datos	4
1.3. El Enfoque: Análisis de Airbnb en Málaga	4
1.4. El Alcance del Proyecto	5
1.5. Motivación	5
<b>2. Datos</b>	<b>6</b>
<b>3. Implementación en PostgreSQL</b>	<b>8</b>
3.1. Creación de Tablas	8
3.2. Índices	11
3.3. Creación de Vistas	12
3.4. Funciones	13
<b>4. Consultas</b>	<b>14</b>
4.1. Número total de alojamientos	14
4.2. Consulta de número de alojamientos por tipo, porcentaje del mismo , precio promedio y duración promedio de la estancia de los alojamientos por tipo	17
4.3. Consulta de número de alojamientos por zona, porcentaje, precio promedio de los alojamientos por zona y duración promedio de la estancia de los alojamientos por zona	19
4.4. Consulta de promedio de las calificaciones	22
4.5. Consulta de los alojamientos que tienen más comentarios y su valoración	22
4.6. Consulta del numero mínimo de noches	22
4.7. Consulta de anfitriones más activos y su relación con las revisiones	23
4.8. Consulta de porcentaje de alojamientos con algún comentario negativo	24
4.9. Consulta de alojamientos según la proximidad al centro de Málaga	24
4.10. Consulta de alojamientos según la proximidad al centro de Málaga ordenados de mejor a peor valoración en función de la distancia.	25
4.11. Consulta de valoración promedio del alojamiento en función del tipo de habitación.	25
4.12. Consulta de porcentaje de propietarios con más de una propiedad	26
4.13. Consulta de cuántos hosts se unen cada año	27
4.14. Consulta de porcentaje y nº camas por alojamiento	27
4.15. Consulta de porcentaje de precio	28
4.16. Porcentaje de mínimo numero de noches	30
4.17. Porcentaje de puntuaciones	30
4.18. Consulta de media de precios de alojamiento en función de su valoracion	31
<b>5. Vistas</b>	<b>34</b>
5.1. Vista de los alojamientos con mayor disponibilidad anual	34
5.2. Vista de la disponibilidad y el precio diario de los listados	34
5.3. Vista del precio medio según el número de habitaciones	35
5.4. Vista de las tendencias de precios medios a lo largo de los años	35
5.5. Vista de las tendencias de disponibilidad en los próximos meses	36
5.6. Vista de los alojamientos mejor valorados por categoría de propiedad	39
5.7. Vista de los alojamientos con comentarios negativos	40
5.8. Vista de los propietarios con más reseñas con comentarios negativos	40
<b>6. Análisis de Correlación de Atributos</b>	<b>42</b>
<b>7. Conclusiones</b>	<b>44</b>
<b>8. Bibliografía</b>	<b>46</b>

## Índice de cuadros

1.	Datos de alojamientos por tipo . . . . .	17
2.	Datos de alojamientos por zona . . . . .	19
3.	Alojamientos con más comentarios y su valoración . . . . .	22
4.	Anfitriones más activos y su relación con las revisiones en el año 2023 . . . . .	23
5.	Porcentaje de alojamientos con comentarios negativos . . . . .	24
6.	Alojamientos según proximidad al centro de Málaga . . . . .	24
7.	Alojamientos según proximidad al centro de Málaga (ordenados por valoración) . . . . .	25
8.	Valoración promedio del alojamiento según el tipo de habitación . . . . .	25
9.	Número de nuevos hosts que se unieron cada año en Airbnb en Málaga. . . . .	27
10.	Porcentaje y número de camas por alojamiento en Málaga. . . . .	28
11.	Porcentaje de precio de alojamientos en Málaga. . . . .	28
12.	Porcentaje de mínimo número de noches en alojamientos de Málaga. . . . .	30
13.	Porcentaje de puntuaciones en alojamientos de Málaga. . . . .	30
14.	Media de precios de alojamiento en función de su valoración en Málaga. . . . .	31
15.	Alojamientos con mayor disponibilidad anual en Málaga. . . . .	34
16.	Disponibilidad y precio diario de los alojamientos en Málaga . . . . .	34
17.	Precio medio según el número de habitaciones de los alojamientos en Málaga. . . . .	35
18.	Tendencias de precios medios a lo largo de los años en alojamientos de Málaga. . . . .	36
19.	Tendencias de disponibilidad en los próximos meses en alojamientos de Málaga. . . . .	37
20.	Alojamientos mejor valorados por categoría de propiedad. . . . .	39
21.	Alojamientos peor valorados por categoría de propiedad. . . . .	39

## Índice de figuras

1.	Esquema Entidad/Relación de la base de datos. . . . .	8
2.	Zonas/barrios de Málaga con límites de códigos postales. . . . .	14
3.	Distribución de alojamientos por zona y ubicación en coordenadas geográficas. . . . .	15
4.	Distribución de alojamientos por tipo y ubicación en coordenadas geográficas. . . . .	16
5.	Gráficas de área y densidad de alojamientos por zona/barrio en Málaga. . . . .	16
6.	Distribución de alojamientos por precio y tipo. . . . .	18
7.	Porcentaje de alojamientos por tipo de habitación. . . . .	18
8.	Distribución de alojamientos. . . . .	19
9.	Número de alojamientos por tipo de habitación. . . . .	20
10.	Proporción de Tipo de Habitación en cada Zona. . . . .	20
11.	Precio promedio por zona. . . . .	21
12.	Precio promedio por código postal. . . . .	21
13.	Hosts con más valoraciones junto con sus propiedades. . . . .	23
14.	Número de alojamientos por usuario. . . . .	26
15.	Comparación porcentaje de Hosts con +5 alojamientos. . . . .	26
16.	Distribución de Precios. . . . .	29
17.	Distribución de Precios (log). . . . .	29
18.	Comparación de número de reseñas y precios. . . . .	32
19.	Evolución del precio medio y disponibilidad. . . . .	37
20.	Reseñas a lo largo de los años. . . . .	38
21.	Promedio de reseñas por mes. . . . .	38
22.	Promedio de reseñas por mes por tipo de habitación en cada Zona. . . . .	39
23.	Mapa de correlación de atributos. . . . .	42

# 1. Introducción

## 1.1. Introducción

En la era contemporánea de la tecnología y la interconexión, Airbnb, Inc. destaca como una prominente empresa estadounidense que ha revolucionado la industria de alojamiento y turismo. Fundada en 2008, Airbnb ha forjado un mercado en línea que reconfigura la forma en que los viajeros experimentan nuevos destinos. Con su sede en San Francisco, California, la plataforma brinda acceso tanto a través de su sitio web como de su aplicación móvil, empoderando a millones de usuarios a explorar alojamientos únicos y experiencias personalizadas en todo el mundo.

Desde su inicio, Airbnb ha transformado el modo en que las personas planifican y experimentan sus viajes. Esta plataforma ha superado las barreras tradicionales, permitiendo que tanto huéspedes como anfitriones participen en una comunidad global de viajeros. Mediante la disposición de viviendas particulares para alquiler temporal y actividades turísticas, Airbnb ha redefinido la concepción del alojamiento, abriendo nuevas vías para que los viajeros se conecten con diversas culturas y entornos.

## 1.2. El Motor del Análisis de Datos

Estos millones de listados generan una gran cantidad de datos, datos que pueden analizarse y usarse para garantizar la seguridad de los usuarios, comprender los patrones de comportamiento, evaluar el rendimiento de los anfitriones y los huéspedes, orientar las iniciativas de marketing y engendrar servicios innovadores que continúan redefiniendo la experiencia de viaje.

El conjunto de datos en cuestión compone una combinación de valores categóricos y numéricos. A través de este análisis, se pretende desentrañar una serie de interrogantes: ¿Qué nos pueden revelar los datos sobre los distintos anfitriones y las diversas áreas? ¿Qué valiosas lecciones pueden extraerse de las predicciones, ya sea en términos de ubicación, precios o reseñas? ¿Quiénes son los anfitriones más activos y qué factores propician su éxito? Asimismo, se busca identificar diferencias notables en el tráfico entre áreas, y esclarecer las posibles razones subyacentes.

## 1.3. El Enfoque: Análisis de Airbnb en Málaga

Este proyecto de análisis de datos se enfoca en el estudio detallado del conjunto de datos de Airbnb en la ciudad de **Málaga**.

El propósito fundamental de este proyecto radica en llevar a cabo una evaluación minuciosa de los datos relacionados con Airbnb en Málaga. Para lograr este objetivo, se utilizarán varias herramientas fundamentales, entre las que se incluyen:

- **Python:** Un lenguaje de programación versátil y potente que se utilizará para la manipulación, limpieza y análisis inicial de los datos. La naturaleza flexible de Python permite la implementación de algoritmos y procesamientos específicos para extraer información valiosa de los conjuntos de datos.
- **PostgreSQL:** Como sistema de gestión de bases de datos relacional de código abierto y altamente escalable, PostgreSQL constituirá el núcleo de almacenamiento y análisis de este gran volumen de datos en este proyecto. Con su amplia gama de características, permitirá gestionar eficientemente la información y realizar consultas sofisticadas para obtener insights cruciales.
- **Tableau:** Una herramienta de visualización de datos que facilitará la creación de representaciones gráficas y visuales de los resultados obtenidos a partir del análisis. Tableau posibilita la comunicación efectiva de los hallazgos a través de gráficos interactivos y paneles intuitivos.

## 1.4. El Alcance del Proyecto

En este proyecto nos enfocamos en realizar análisis detallados de los datos de Airbnb en Málaga. A través de consultas y cálculos inteligentes, buscamos obtener información importante sobre alojamientos, disponibilidad, precios, reseñas y otros aspectos interesantes.

El resultado de estas operaciones es que conseguimos tener una visión completa y profunda del mercado de alquileres en Málaga. Al combinar estas herramientas con nuestra metodología, también podemos descubrir tendencias y patrones realmente interesantes y realizar predicciones. Estos descubrimientos nos dan una idea sólida de cómo funciona el mercado y nos ayudarán a tomar decisiones más informadas en este sector que está siempre en movimiento.

## 1.5. Motivación

Este proyecto se cimienta en la necesidad de arrojar luz sobre el mercado de alquileres de Airbnb en **Málaga** a través de un análisis exhaustivo de datos. Si bien hemos esbozado el enfoque y alcance, es crucial profundizar en por qué este análisis posee un valor inherente.

Este análisis va más allá de una simple recopilación de datos; representa un esfuerzo deliberado por desentrañar patrones, relaciones y tendencias que pueden pasar desapercibidos a simple vista. A medida que el mercado de alquileres continúa evolucionando, surge la necesidad de entender los matices detrás de las fluctuaciones de precios, las preferencias de ubicación y la satisfacción del cliente.

En este contexto, este proyecto adquiere relevancia. Los datos por sí solos son información cruda; el análisis los convierte en conocimiento accionable. Proporciona una hoja de ruta para tomar decisiones estratégicas fundamentadas en datos sólidos, tanto para los propietarios de propiedades de alquiler como para los huéspedes que buscan experiencias enriquecedoras.

## 2. Datos

Los datos empleados en este proyecto fueron obtenidos de [Inside Airbnb](#).<sup>[1]</sup>

Inside Airbnb es un proyecto independiente que recopila y analiza datos procedentes de diversas ciudades alrededor del mundo que forman parte de la plataforma Airbnb. Fue creado con el objetivo de brindar transparencia y comprensión sobre el impacto de Airbnb en los mercados de alquiler a corto plazo. Inside Airbnb recopila datos públicos de las listas de Airbnb, como información sobre las propiedades, los anfitriones y la disponibilidad de alojamiento. Estos datos se utilizan para crear conjuntos de datos estructurados que permiten realizar análisis y visualizaciones detalladas sobre la oferta de alojamiento en cada ciudad.

Los datos utilizados en este estudio están disponibles en su página web <sup>[2]</sup>. Adicionalmente, es posible explorar una representación interactiva de los datos en la sección “Explorar datos”, donde se presenta un mapa de Málaga <sup>[3]</sup>. Esta herramienta brinda una síntesis informativa sobre diversos aspectos de los datos relativos a Airbnb en la ciudad <sup>[4, 6, 7, 8, 9, 10, 11, 12, 13, 14]</sup>

- **Tipo de habitación:** Muestra el porcentaje de diferentes tipos de habitaciones disponibles en las listas de Airbnb en la ciudad. Por ejemplo, se muestra el porcentaje de viviendas/apartamentos completos, habitaciones privadas, habitaciones compartidas y habitaciones de hotel.
- **Actividad:** Muestra datos relacionados con la actividad de las propiedades de alquiler en Airbnb. Se incluyen estadísticas como el número promedio de noches reservadas, el precio promedio por noche y los ingresos promedio generados. Estos datos se calculan utilizando la estancia mínima, el precio y el número de reseñas de cada propiedad.
- **Alquileres a corto plazo:** Proporciona información sobre la proporción de alquileres a corto plazo en comparación con los alquileres a largo plazo. Esto puede ayudar a comprender si el mercado se inclina más hacia estancias cortas o largas, y si los alquileres a corto plazo están afectando la disponibilidad de viviendas a largo plazo.
- **Listados por anfitrión:** Muestra el porcentaje de anfitriones que tienen múltiples listados en Airbnb. Esto indica si hay anfitriones que operan varios alojamientos y sugiere la posibilidad de que estén administrando un negocio de alquileres a corto plazo.
- **Principales anfitriones:** Esta sección probablemente proporcionaría una lista de los principales anfitriones en la ciudad, basada en ciertos criterios como el número de reservas, los ingresos generados o la satisfacción de los huéspedes.





## 3. Implementación en PostgreSQL

### 3.1. Creación de Tablas

Para gestionar y analizar eficazmente los datos recopilados de Inside Airbnb, se ha optado por implementarlos en una base de datos PostgreSQL. En esta etapa, se han creado tres tablas fundamentales para organizar la información de manera estructurada y accesible. Estas tablas son:

- **listings:** Contiene información detallada sobre las propiedades de alquiler, como descripciones, ubicaciones, precios y calificaciones de los anfitriones.
- **reviews:** Contiene las reseñas de los usuarios sobre las propiedades.
- **calendar:** Proporciona información sobre la disponibilidad de las propiedades en diferentes fechas.



Figura 1: Esquema Entidad/Relación de la base de datos.

## Explicación de los datos: tabla 'listings'

- **id**: Identificador único del listado (BIGINT).
- **listing\_url**: URL del listado (TEXT, no nulo).
- **scrape\_id**: Identificador de la extracción (BIGINT).
- **last\_scraped**: Fecha en la que se realizó la última extracción (DATE).
- **source**: Fuente del listado (TEXT).
- **name**: Nombre del listado (TEXT).
- **description**: Descripción del listado (TEXT).
- **neighborhood\_overview**: Descripción del vecindario (TEXT).
- **picture\_url**: URL de la imagen del listado (TEXT).
- **host\_id**: Identificador único del anfitrión (BIGINT, no nulo).
- **host\_url**: URL del perfil del anfitrión (TEXT, no nulo).
- **host\_name**: Nombre del anfitrión (TEXT).
- **host\_since**: Fecha en la que el anfitrión se unió (DATE).
- **host\_location**: Ubicación del anfitrión (TEXT).
- **host\_about**: Descripción del anfitrión (TEXT).
- **host\_response\_time**: Tiempo de respuesta del anfitrión (TEXT).
- **host\_response\_rate**: Tasa de respuesta del anfitrión (TEXT).
- **host\_acceptance\_rate**: Tasa de aceptación del anfitrión (TEXT).
- **host\_is\_superhost**: Indica si el anfitrión es un superhost (TEXT).
- **host\_thumbnail\_url**: URL de la imagen en miniatura del anfitrión (TEXT).
- **host\_picture\_url**: URL de la imagen del anfitrión (TEXT).
- **host\_neighbourhood**: Vecindario del anfitrión (TEXT).
- **host\_listings\_count**: Número de listados del anfitrión (INTEGER).
- **host\_total\_listings\_count**: Número total de listados del anfitrión (INTEGER).
- **host\_verifications**: Métodos de verificación del anfitrión (TEXT).
- **host\_has\_profile\_pic**: Indica si el anfitrión tiene una foto de perfil (TEXT).
- **host\_identity\_verified**: Indica si la identidad del anfitrión ha sido verificada (TEXT).
- **neighbourhood**: Nombre del vecindario (TEXT).
- **neighbourhood\_cleansed**: Nombre del vecindario (limpiado) (TEXT).
- **neighbourhood\_group\_cleansed**: Nombre del grupo de vecindarios (limpiado) (TEXT).
- **latitude**: Latitud de la ubicación del listado (FLOAT).
- **longitude**: Longitud de la ubicación del listado (FLOAT).
- **property\_type**: Tipo de propiedad (TEXT).
- **room\_type**: Tipo de habitación (TEXT).
- **accommodates**: Capacidad de alojamiento (INTEGER).
- **bathrooms**: Número de baños (FLOAT).
- **bathrooms\_text**: Descripción de los baños (TEXT).
- **bedrooms**: Número de dormitorios (INTEGER).
- **beds**: Número de camas (INTEGER).
- **amenities**: Comodidades del listado (TEXT).
- **price**: Precio del listado (TEXT).
- **minimum\_nights**: Número mínimo de noches para reservar (INTEGER).
- **maximum\_nights**: Número máximo de noches para reservar (INTEGER).
- **minimum\_minimum\_nights**: Número mínimo de noches requeridas en la reserva más restrictiva (INTEGER).
- **maximum\_minimum\_nights**: Número máximo de noches requeridas en la reserva más restrictiva (INTEGER).
- **minimum\_maximum\_nights**: Número mínimo de noches permitidas en la reserva más permisiva (INTEGER).
- **maximum\_maximum\_nights**: Número máximo de noches permitidas en la reserva más permisiva (INTEGER).
- **minimum\_nights\_avg\_ntm**: Promedio mínimo de noches requeridas para reservar (FLOAT).
- **maximum\_nights\_avg\_ntm**: Promedio máximo de noches requeridas para reservar (FLOAT).

- **calendar\_updated**: Fecha de actualización del calendario (TEXT).
- **has\_availability**: Indica si hay disponibilidad (TEXT).
- **availability\_30**: Número de días de disponibilidad en los próximos 30 días (INTEGER).
- **availability\_60**: Número de días de disponibilidad en los próximos 60 días (INTEGER).
- **availability\_90**: Número de días de disponibilidad en los próximos 90 días (INTEGER).
- **availability\_365**: Número de días de disponibilidad en los próximos 365 días (INTEGER).
- **calendar\_last\_scraped**: Fecha en la que se realizó la última extracción del calendario (DATE).
- **number\_of\_reviews**: Número total de reseñas (INTEGER).
- **number\_of\_reviews\_ltm**: Número de reseñas en los últimos 12 meses (INTEGER).
- **number\_of\_reviews\_l30d**: Número de reseñas en los últimos 30 días (INTEGER).
- **first\_review**: Fecha de la primera reseña (DATE).
- **last\_review**: Fecha de la última reseña (DATE).
- **review\_scores\_rating**: Puntuación de las reseñas (FLOAT).
- **review\_scores\_accuracy**: Puntuación de precisión en las reseñas (FLOAT).
- **review\_scores\_cleanliness**: Puntuación de limpieza en las reseñas (FLOAT).
- **review\_scores\_checkin**: Puntuación de check-in en las reseñas (FLOAT).
- **review\_scores\_communication**: Puntuación de comunicación en las reseñas (FLOAT).
- **review\_scores\_location**: Puntuación de ubicación en las reseñas (FLOAT).
- **review\_scores\_value**: Puntuación de valor en las reseñas (FLOAT).
- **license**: Licencia relacionada con el listado (TEXT).
- **instant\_bookable**: Indica si el listado se puede reservar de forma instantánea (TEXT).
- **calculated\_host\_listings\_count**: Número de listados calculados para el anfitrión (INTEGER).
- **calculated\_host\_listings\_count\_entire\_homes**: Número de listados calculados para el anfitrión que son hogares completos (INTEGER).
- **calculated\_host\_listings\_count\_private\_rooms**: Número de listados calculados para el anfitrión que son habitaciones privadas (INTEGER).
- **calculated\_host\_listings\_count\_shared\_rooms**: Número de listados calculados para el anfitrión que son habitaciones compartidas (INTEGER).
- **reviews\_per\_month**: Número de reseñas promedio por mes (FLOAT).

### Explicación de los datos: tabla 'reviews'

- **listing\_id**: Identificador del listado (BIGINT).
- **id**: Identificador único de la reseña (BIGINT).
- **date**: Fecha de la reseña (DATE).
- **reviewer\_id**: Identificador del revisor (BIGINT).
- **reviewer\_name**: Nombre del revisor (TEXT).
- **comments**: Comentarios de la reseña (TEXT).

### Explicación de los datos: tabla 'calendar'

- **listing\_id**: Identificador del listado (BIGINT).
- **date**: Fecha del calendario (DATE).
- **available**: Indicador de disponibilidad (CHAR(1)).
- **price**: Precio (VARCHAR(30)).
- **adjusted\_price**: Precio ajustado (VARCHAR(30)).
- **minimum\_nights**: Número mínimo de noches (INTEGER).
- **maximum\_nights**: Número máximo de noches (INTEGER).

## Manejo de los datos

1. **Clave primaria:** La clave primaria identifica de manera única cada registro en una tabla.
  - Para la tabla "listings", se ha seleccionado el atributo `id` como clave primaria de tipo **BIGINT**. El atributo `id` proporciona un identificador único para cada registro de la tabla y permite la indexación eficiente y la recuperación rápida de datos.
  - En las tablas `reviews` y `calendar`, se ha utilizado una combinación de atributos ("*listing\_id*.*id*.*en* reviews" "*listing\_id* "date.*en* calendar") como clave primaria compuesta para garantizar la unicidad de los registros.
2. **Clave foránea:** Las claves foráneas aseguran la integridad referencial y permiten realizar consultas y operaciones que involucren datos relacionados en las diferentes tablas. Se ha establecido una relación entre las tablas `reviews` y `calendar` con la tabla "listings" mediante el uso de claves foráneas.
  - En la tabla `reviews`, el atributo "*listing\_id*" se utiliza como clave foránea que referencia el atributo `id` de la tabla "listings".
  - De manera similar, en la tabla `calendar`, el atributo "*listing\_id*" se utiliza como clave foránea que referencia el atributo `id` de la tabla "listings".
3. **Tipos de datos:** Se han elegido tipos de datos adecuados para cada atributo en función de la naturaleza de los datos que se almacenan. Por ejemplo:
  - Los atributos que almacenan texto largo, como descripciones y comentarios, se han definido como tipo **TEXT** para permitir una capacidad de almacenamiento más amplia.
  - Los identificadores y conteos, como `id`, "*host\_id*" y "*number\_of\_reviews*", se han seleccionado como tipo **BIGINT** para manejar valores enteros grandes.
  - Los datos de fecha se han definido como tipo **DATE** para almacenar información sobre fechas específicas.
  - Los precios se han definido como tipo **TEXT** o **VARCHAR(30)** para permitir la representación de diferentes formatos y monedas.

### 3.2. Índices

La creación de índices es una estrategia importante para mejorar la velocidad de búsqueda en bases de datos. A continuación, se justifica la creación de índices para cada uno de los siguientes atributos en las tablas correspondientes:

#### Tabla "listings"

- **Atributo `id`:** Se ha creado el índice `idx_listings_id` en el atributo `id` de la tabla "listings" para acelerar las búsquedas por este identificador único.
- **Atributo "`host_id`":** Se ha creado el índice `idx_listings_host_id` en el atributo "`host_id`" de la tabla "listings" para mejorar la velocidad de búsqueda de registros relacionados con un `host` específico.
- **Atributos "`host_id`" y "`property_type`":** El índice `idx_listings_host_id_property_type` en los atributos "`host_id`" y "`property_type`" permite realizar consultas más eficientes que involucren la combinación de estos dos atributos.
- **Atributos de ubicación:** Se ha creado el índice `idx_listings_location` utilizando el método de índice espacial `gist` en los atributos de latitud y longitud. Esto mejora la velocidad de búsqueda de registros basados en la ubicación geográfica.
- **Atributos de distancia:** El índice `idx_listings_distance` se ha creado utilizando una expresión de cálculo de distancia. Esto mejora la eficiencia en las búsquedas que involucren cálculos de distancia utilizando la fórmula proporcionada.

## Tabla “reviews”

- **Atributo “listing\_id”:** Se ha creado el índice `idx_reviews_listing_id` en el atributo “listing\_id” de la tabla reviews para acelerar las búsquedas por el identificador de las propiedades asociadas a las revisiones.
- **Atributo “comments”:** El índice `idx_listings_review` se ha creado utilizando el método `gin` para la creación de índices de texto completo en el atributo “comments”. Esto permite realizar búsquedas de texto en el campo de comentarios de manera más rápida y eficiente.

## Tabla “calendar”

- **Atributos “listing\_id” “date”:** El índice `idx_calendar_listing_id_date` en los atributos “listing\_id” “date” de la tabla calendar mejora la velocidad de búsqueda en consultas que involucren estos dos atributos de manera conjunta.
- **Atributo “.available”:** Se ha creado el índice `idx_calendar_availability` en el atributo “.available” para acelerar las búsquedas de disponibilidad en el calendario.
- **Atributo “listing\_id”:** El índice `idx_calendar_listing_id` en el atributo “listing\_id” mejora la eficiencia en búsquedas relacionadas con el identificador de las propiedades en el calendario.

## 3.3. Creación de Vistas

Se crean varias vistas en la base de datos para facilitar el análisis y la visualización de los datos. Estas vistas incluyen:

- `year_availability`: Muestra los alojamientos con mayor disponibilidad anual.
- `listing_availability_price`: Muestra la disponibilidad y el precio diario de los listados.
- `price_by_bedrooms`: Muestra el precio medio según el número de habitaciones.
- `price_trends`: Muestra las tendencias de precios medios a lo largo de los años.
- `availability_trends`: Muestra las tendencias de disponibilidad a lo largo de los años.
- `top_rated_listings`: Muestra los alojamientos mejor valorados por categoría de propiedad.
- `negative_reviews`: Muestra los alojamientos con comentarios negativos.
- `owner_negative_reviews`: Muestra los propietarios con más reseñas con comentarios negativos.

### 3.4. Funciones

El código también incluye la definición de funciones que realizan diferentes cálculos y facilitan el posterior análisis de los datos. Estas funciones incluyen:

- `calculate_average_rating()`: Calcula el promedio de las calificaciones de los alojamientos.

```
CREATE OR REPLACE FUNCTION calculate_average_rating()
    RETURNS FLOAT AS $$
DECLARE
    avg_rating FLOAT;
BEGIN
    SELECT AVG(review_scores_rating) INTO avg_rating
    FROM listings;
    RETURN avg_rating;
END;
$$ LANGUAGE plpgsql;
```

- `calculate_max_minimum_nights()`: Calcula el número mínimo de noches para alojarse en los listados.

```
CREATE OR REPLACE FUNCTION calculate_max_minimum_nights()
    RETURNS INTEGER AS $$
DECLARE
    max_min_nights INTEGER;
BEGIN
    SELECT MAX(minimum_nights) INTO max_min_nights
    FROM calendar
    GROUP BY listing_id
    ORDER BY max_min_nights DESC
    LIMIT 1;
    RETURN max_min_nights;
END;
$$ LANGUAGE plpgsql;
```

- `calculate_multiple_properties_percentage()`: Calcula el porcentaje de propietarios con más de una propiedad.

```
CREATE OR REPLACE FUNCTION calculate_multiple_properties_percentage()
    RETURNS FLOAT AS $$
DECLARE
    total_hosts INTEGER;
    hosts_with_multiple_properties INTEGER;
    percentage FLOAT;
BEGIN
    SELECT COUNT(DISTINCT host_id) INTO total_hosts
    FROM listings;

    SELECT COUNT(*) INTO hosts_with_multiple_properties
    FROM (
        SELECT host_id
        FROM listings
        GROUP BY host_id
        HAVING COUNT(*) > 1
    ) AS multiple_properties;

    percentage := (hosts_with_multiple_properties::FLOAT / total_hosts::FLOAT) * 100;
    RETURN ROUND(percentage);
END;
$$ LANGUAGE plpgsql;
```

## 4. Consultas

A continuación, se presentan algunos ejemplos de consultas utilizadas en el proyecto, junto con sus resultados y conclusiones.

### 4.1. Número total de alojamientos

Se llevó a cabo una consulta para obtener el recuento exacto de alojamientos que se analizarán en el estudio. El propósito de esta consulta es obtener una visión general del crecimiento de los alojamientos en Málaga. Esto es útil para comprender la competencia en el mercado y analizar la disponibilidad de alojamientos. A continuación se muestra el código utilizado y los resultados obtenidos:

```
SELECT COUNT(*) AS total_listings
FROM listings;
```

La ejecución de esta consulta reveló que en la ciudad de Málaga existe un total de 7,534 alojamientos disponibles. Este dato es fundamental para comprender la dimensión y disponibilidad del mercado de alquileres vacacionales en la ciudad de Málaga.

### Distribución geográfica de los alojamientos

Para visualizar la distribución geográfica de los alojamientos en Málaga, se creó un mapa utilizando el software Tableau. En la figura se muestra el mapa coloreado de zonas/barrios de Málaga, con los límites de los códigos postales.

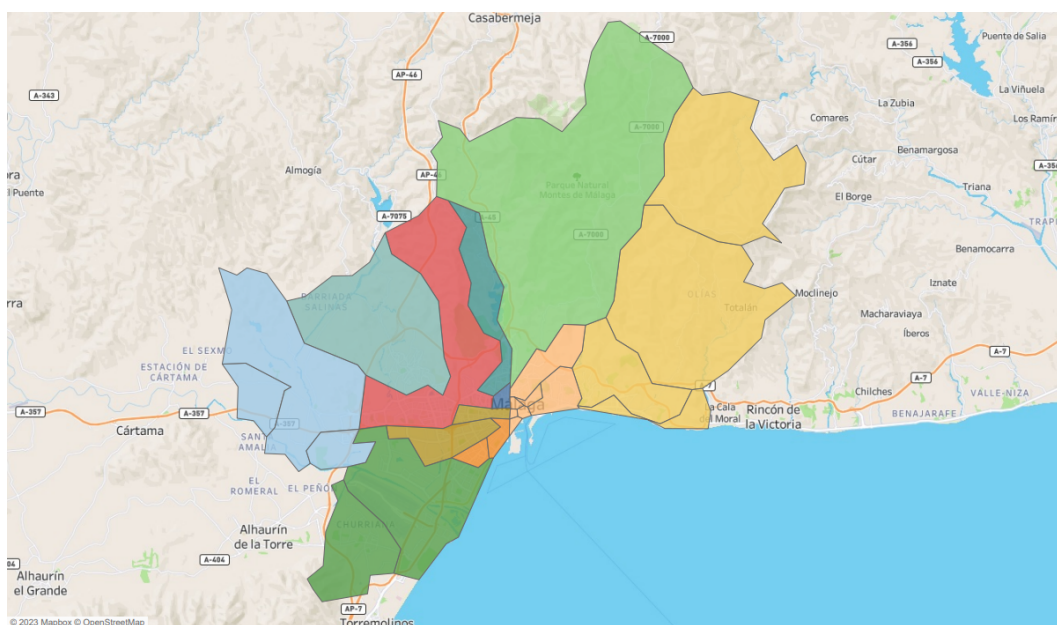


Figura 2: Zonas/barrios de Málaga con límites de códigos postales.

Neighbourhood
Bailen-Miraflores
Campanillas
Carretera de Cadiz
Centro
Churriana
Ciudad Jardin
Cruz De Humilladero
Este
Palma-Palmilla
Puerto de la Torre
Teatinos-Universidad

La visualización se ha segmentado en diferentes zonas o barrios de la ciudad, cada uno definido por la combinación de varios códigos postales con características similares.

Los límites de los códigos postales se muestran en el mapa, permitiendo una clara identificación de las áreas que abarca cada zona. Esta representación gráfica es esencial para identificar patrones de concentración y densidad de alojamientos en diferentes partes de la ciudad. La representación se basa en el modelo del proyecto [Inside Airbnb](#)

Los códigos postales asignados a cada zona son:

- **Bailén-Miraflores:** 29009
- **Campanillas:** 29196, 29590, 29591
- **Carretera de Cádiz:**
- **Centro:** 29001, 29005, 29008, 29012, 29013, 29015, 29016
- **Churriana:** 29004, 29140
- **Ciudad Jardín:** 29014
- **Cruz de Humilladero:** 29006, 29007
- **Este:** 29017, 29018, 29197, 29193
- **Palma-Palmilla:** 29011
- **Puerto de la Torre:** 29190
- **Teatinos-Universidad:** 29010

En esta representación, hemos utilizado coordenadas de latitud y longitud para mapear la ubicación de los alojamientos en el espacio geográfico. Cada alojamiento se representa mediante un punto de color en el mapa, donde el color del punto se corresponde con una zona o barrio específico. Esta codificación de colores es similar a la empleada en el mapa presentado previamente.

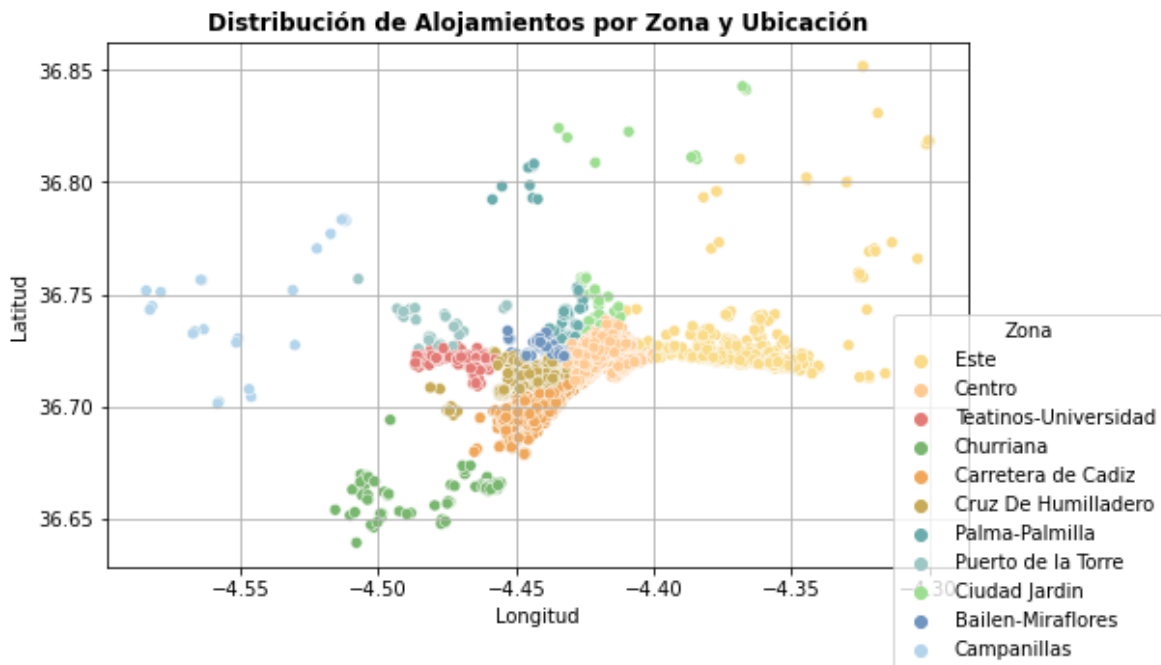


Figura 3: Distribución de alojamientos por zona y ubicación en coordenadas geográficas.

Esta visualización nos proporciona una visión de cómo los alojamientos se distribuyen en el entorno geográfico de Málaga. Al analizar la agrupación de puntos en diferentes zonas, podemos identificar áreas que pueden ser especialmente populares entre los huéspedes o donde los anfitriones pueden estar más activos. Asimismo, esta representación permite evaluar cómo los alojamientos se dispersan en las diversas partes de la ciudad y cómo se relaciona con los patrones observados en otras visualizaciones.

Observamos que existe una mayor concentración de alojamientos a medida que se aproxima al centro de la ciudad.



Consideramos, además de la distribución geográfica de los alojamientos en Málaga, los tipos de habitaciones ofrecidos por los anfitriones.

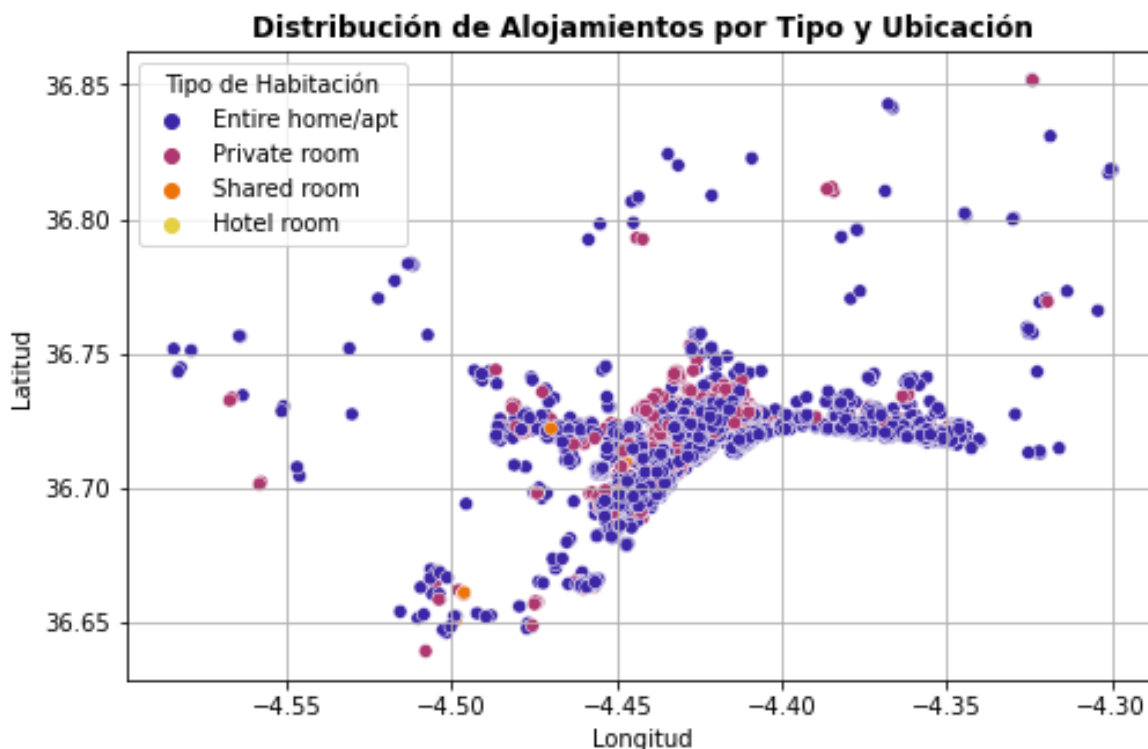


Figura 4: Distribución de alojamientos por tipo y ubicación en coordenadas geográficas.

Los colores de los puntos están codificados según los tipos de habitaciones, y esto nos ayuda a identificar patrones en la oferta de alojamientos, como la prevalencia de ciertos tipos de habitaciones en áreas específicas de la ciudad como las propiedades enteras o apartamentos, que son los que más abundan.

A continuación, se presenta una representación gráfica relacionada con la distribución de las zonas de estudio de Málaga. Observamos dos gráficas que proporcionan información sobre cómo los alojamientos se distribuyen en función del área de las zonas o barrios en Málaga y su densidad respectiva.

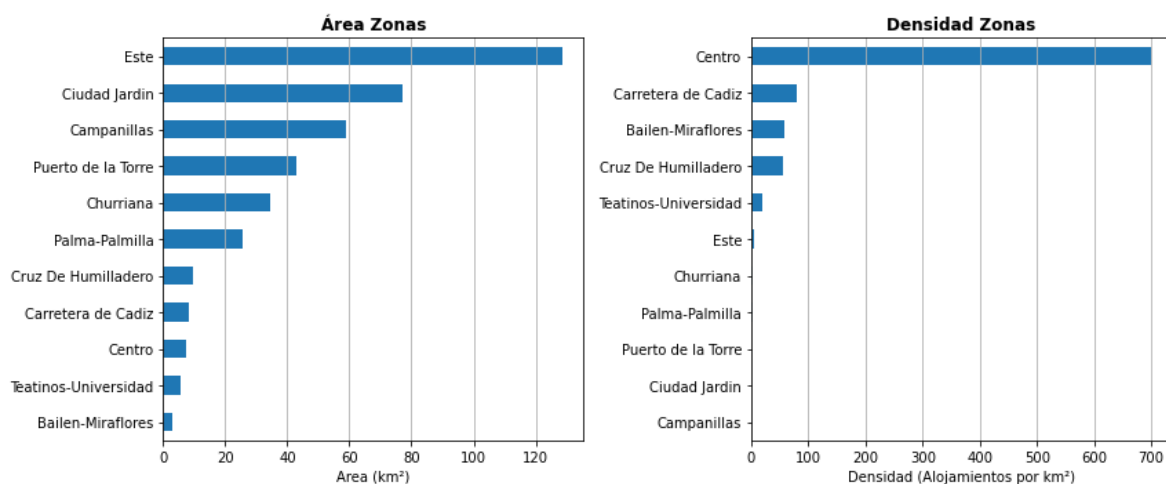


Figura 5: Gráficas de área y densidad de alojamientos por zona/barrio en Málaga.

En la primera gráfica se muestra el **área** de cada zona o barrio en Málaga, permitiéndonos identificar visualmente cuáles son las áreas más extensas y cuáles son más compactas en términos de espacio geográfico. Esta información contextual es fundamental para comprender la distribución de los alojamientos en la ciudad y cómo podría estar relacionada con la disponibilidad de espacio.

La segunda gráfica representa la **densidad** de alojamientos en cada zona o barrio. La densidad se refiere al número de alojamientos por metro cuadrado en cada zona. Esta gráfica nos proporciona una perspectiva única de cómo los alojamientos se agrupan o dispersan en relación con el espacio disponible.

La interacción entre ambas gráficas nos permite obtener conclusiones más precisas sobre cómo se distribuyen los alojamientos en Málaga y si hay alguna tendencia en la relación entre el área y la densidad de alojamientos. Observamos que el distrito *Este* es el más extenso pero no es que tiene una mayor densidad de alojamientos. Es el distrito *Centro* el que tiene una mayor concentración de alojamientos por kilómetro cuadrado.

## 4.2. Consulta de número de alojamientos por tipo, porcentaje del mismo , precio promedio y duración promedio de la estancia de los alojamientos por tipo

En esta consulta, se obtuvieron datos sobre el número de alojamientos, el precio promedio y la duración promedio de la estancia según el tipo de propiedad. A continuación se muestra el código utilizado y los resultados obtenidos:

```
SELECT property_type, COUNT(*) AS count,
       ROUND(AVG(CAST(REGEXP_REPLACE(listings.price,
       '[^\d.]', '', 'g') AS NUMERIC))) AS avg_price,
       ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM listings), 2)
       AS percentage,
       ROUND(AVG(minimum_nights)) AS avg_minimum_nights
FROM listings
GROUP BY property_type
ORDER BY count DESC;
```

Tipo	Nº de aloj.	Precio prom.	Porcentaje (%)	Dur. prom.
Entire rental unit	5151	219 €	68.37	3
Private room in rental unit	642	75 €	8.52	5
Entire condo	328	167 €	4.35	6
Entire loft	264	121 €	3.50	5
Entire home	207	232 €	2.75	9
Entire serviced apartment	193	175 €	2.56	2
Entire vacation home	126	190 €	1.67	2
Private room in home	107	55 €	1.42	3
Entire townhouse	51	1741 €	0.68	4
...	...	...	...	...

Cuadro 1: Datos de alojamientos por tipo

A través de un enfoque visual, es posible resumir de manera efectiva los resultados obtenidos con respecto al precio medio según el tipo de propiedad en Málaga. Esta gráfica, generada utilizando la herramienta **Python** de **'bokeh'**, ofrece una instantánea que destaca la variabilidad en los precios de los distintos tipos de alojamientos.

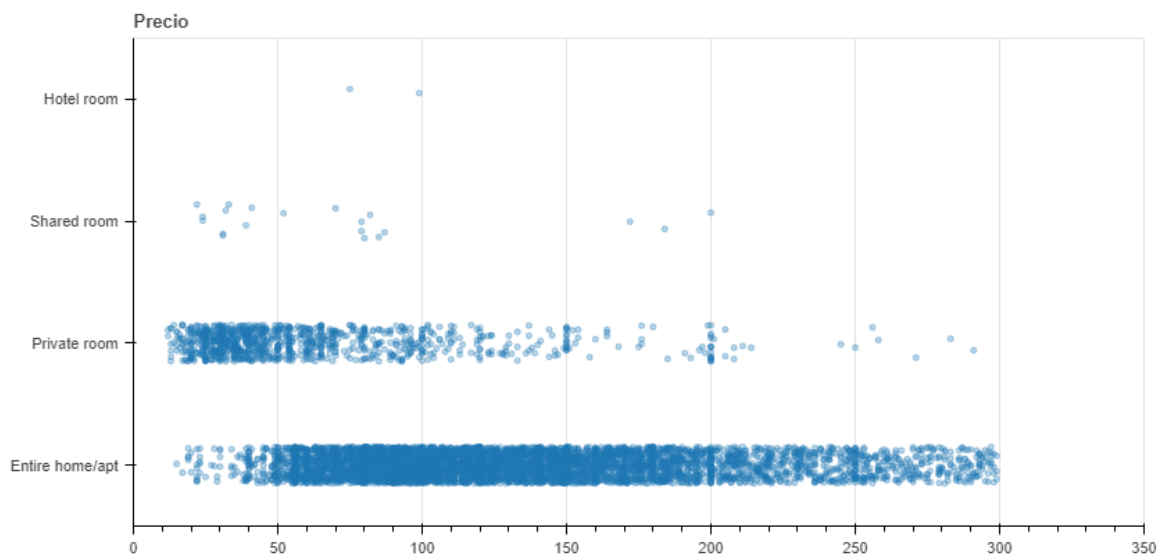


Figura 6: Distribución de alojamientos por precio y tipo.

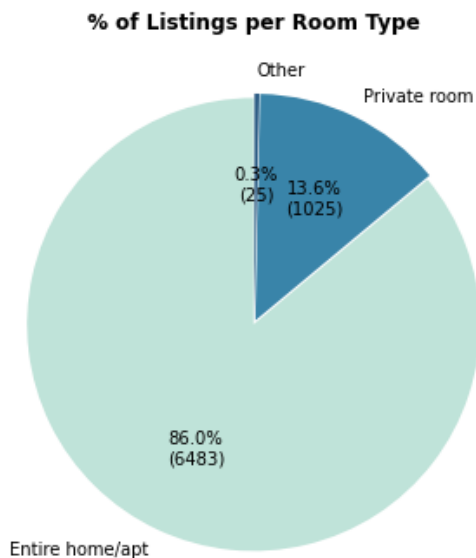


Figura 7: Porcentaje de alojamientos por tipo de habitación.

El resultado revela una amplia variedad de opciones disponibles para los huéspedes. Los alquileres completos son los más comunes, seguidos de habitaciones privadas en unidades de alquiler. Los precios promedio varían según el tipo de propiedad, siendo las casas completas las más costosas. La duración promedio de la estancia también varía, con una media de 3 a 5 noches.

#### 4.3. Consulta de número de alojamientos por zona, porcentaje, precio promedio de los alojamientos por zona y duración promedio de la estancia de los alojamientos por zona

En esta consulta, se obtuvieron datos relevantes sobre el número de alojamientos, el porcentaje, el precio promedio y la duración promedio de la estancia de los alojamientos por zona en Málaga. El propósito de esta consulta es analizar las características y tendencias de cada zona en relación con los alojamientos disponibles. A continuación se muestra el código utilizado y los resultados obtenidos:

```
SELECT neighbourhood_cleansed, COUNT(*) AS count,
       ROUND(AVG(CAST(REGEXP_REPLACE(price, '[^\d.]', '', 'g')
       AS NUMERIC))) AS avg_price,
       ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM listings), 2)
       AS percentage,
       ROUND(AVG(minimum_nights)) AS avg_minimum_nights
FROM listings
GROUP BY neighbourhood_cleansed
ORDER BY avg_price DESC;
```

Zona	Nº de aloj.	Precio prom.	Porcentaje (%)	Dur. prom.
Bailen-Miraflores	178	503 €	2.36	5
Este	738	298 €	9.80	4
Cruz De Humilladero	531	253 €	7.05	3
Carretera de Cadiz	658	237 €	8.73	4
Churriana	101	182 €	1.34	4
Centro	5054	174 €	67.08	3
Campanillas	26	168 €	0.35	3
Puerto de la Torre	31	156 €	0.41	3
Teatinos-Universidad	105	127 €	1.39	6
Ciudad Jardin	54	106 €	0.72	2
Palma-Palmilla	58	92 €	0.77	3

Cuadro 2: Datos de alojamientos por zona

Explorando más a fondo la distribución de alojamientos en Málaga, podemos utilizar una serie de gráficas para destacar aspectos clave. Podemos analizar mediante un gráfico circular la distribución de los alojamientos en Málaga.

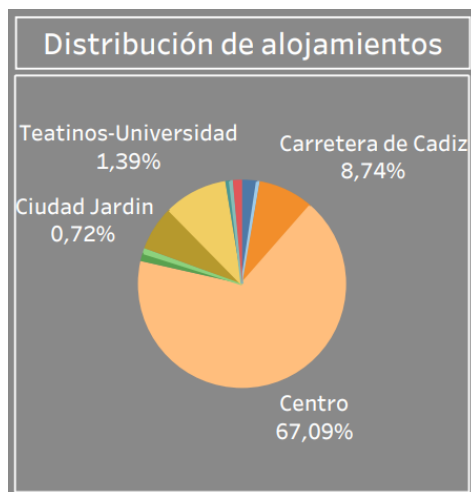


Figura 8: Distribución de alojamientos.

Comprobamos que efectivamente existe una concentración mayor de alojamientos en el distrito *Centro* con un 67% y en segundo lugar en *Carretera de Cádiz* con aproximadamente un 9%.

Continuamos analizando el número de alojamientos por tipo y por zona, utilizando un gráfico de barras. Es evidente que tanto las casas/apartamentos como las habitaciones privadas son las opciones más comunes en el distrito *Centro*, reforzando su popularidad entre los huéspedes.



Figura 9: Número de alojamientos por tipo de habitación.

En la siguiente gráfica se muestra la proporción de tipo de habitación en cada zona. Es esencial para conocer qué tipo de alojamiento se oferta en cada zona y permite identificar patrones.

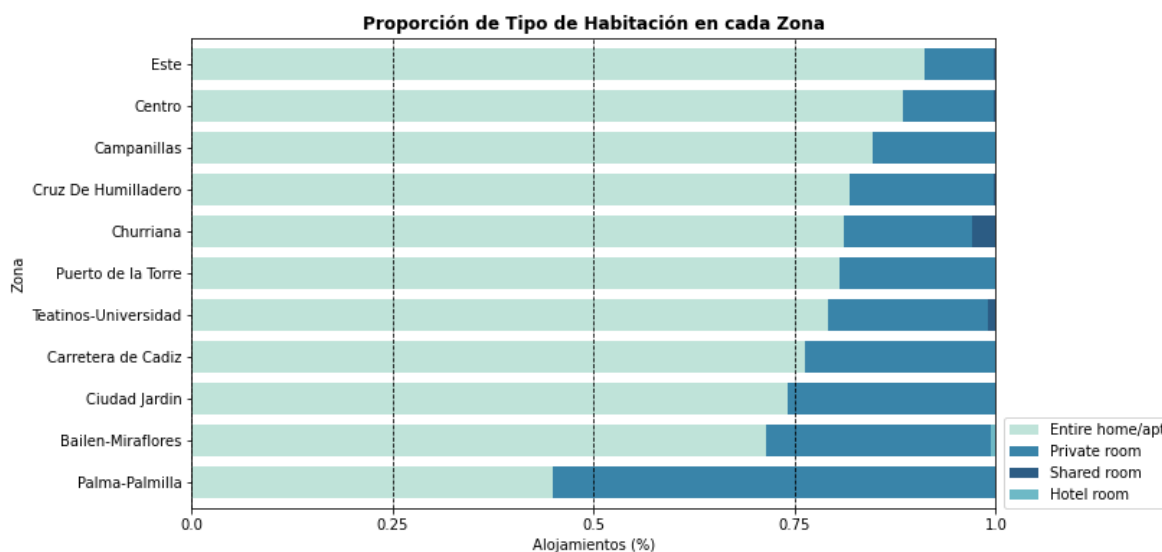


Figura 10: Proporción de Tipo de Habitación en cada Zona.

Observamos un dato curioso: el distrito de *Churriana* es el que mayor proporción de habitaciones compartidas tiene. Sin embargo, como hemos analizado anteriormente, este distrito no tiene una alta concentración de alojamientos públicos en Airbnb.

La relación entre el precio promedio y la ubicación geográfica es crucial. Esta gráfica arroja perspectivas significativas: las zonas *Bailen-Miraflores* y *Este* presentan los precios más altos, mientras que *Ciudad Jardín* y *Palma-Palmilla* tienen precios más bajos.

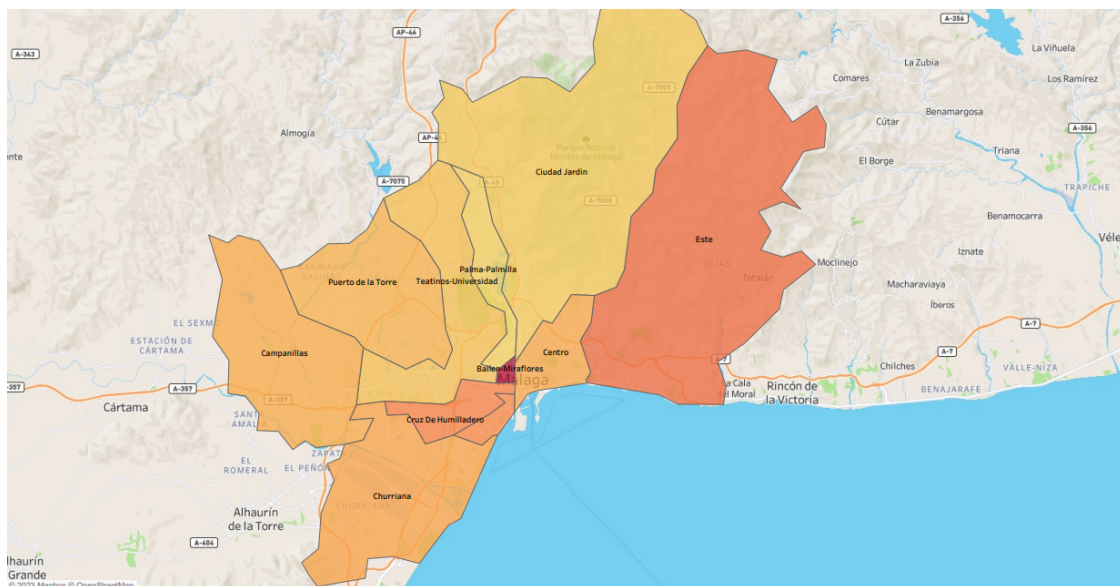


Figura 11: Precio promedio por zona.

Podemos llevar a cabo un análisis más detallado desglosando los datos no solo por barrio, sino también por código postal. Este enfoque nos permitirá descubrir qué zonas específicas de Málaga presentan una mayor demanda en términos de precio.

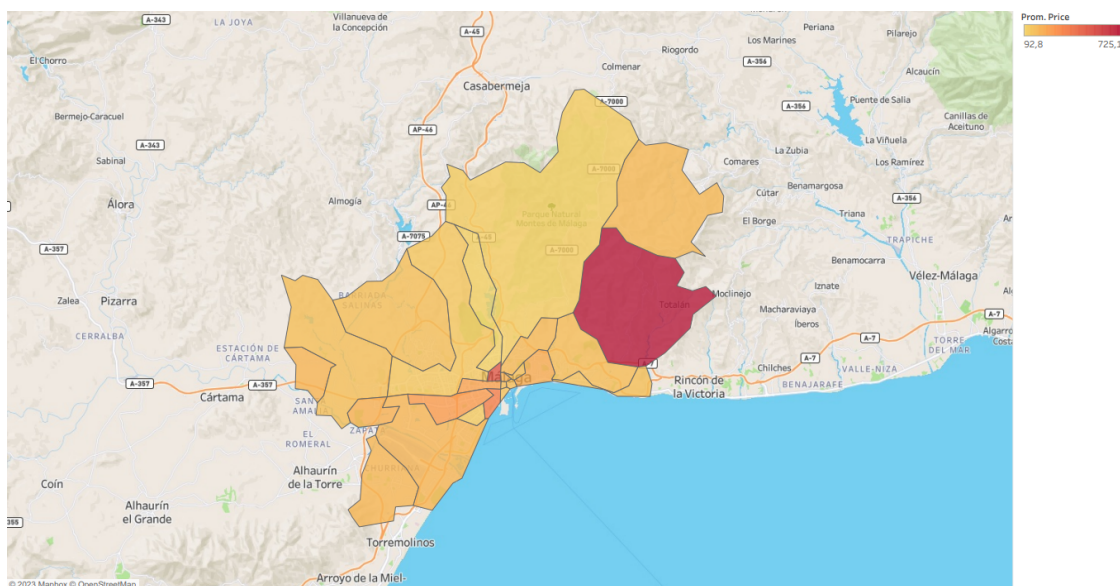


Figura 12: Precio promedio por código postal.

Destacan dos códigos postales en particular: el 29197, ubicado en el distrito *Este*, con un precio medio de 725,1€. Este valor sugiere una fuerte demanda en esta área específica. Otro código postal significativo es el 29009, correspondiente a *Bailen Miraflores*, con un precio medio de 503,4€.

Finalmente, evaluamos la duración promedio de la estancia por zona. Se observa una variación entre 2 y 6 noches, resaltando la diversidad de preferencias de los huéspedes en diferentes áreas.

Estos resultados proporcionan una visión de la distribución geográfica de los alojamientos, los precios promedio y la duración promedio de la estancia en Málaga, siendo información relevante para entender las preferencias de los huéspedes.

#### 4.4. Consulta de promedio de las calificaciones

En esta consulta, se obtiene la calificación promedio de todos los alojamientos en la ciudad de Málaga. El objetivo es determinar el nivel promedio de satisfacción de los huéspedes. A continuación se muestra el código utilizado y el resultado obtenido:

```
SELECT calculate_average_rating() AS average_rating;
```

El resultado revela que la puntuación media es de 4.61. Esto indica que, en promedio, los alojamientos en Málaga han recibido una calificación alta por parte de los huéspedes.

#### 4.5. Consulta de los alojamientos que tienen más comentarios y su valoración

En esta consulta, se identifican los alojamientos que han recibido la mayor cantidad de comentarios, junto con su valoración correspondiente. El objetivo es conocer los alojamientos más populares y bien valorados por los huéspedes. A continuación se muestra el código utilizado y los resultados obtenidos:

```
SELECT li.listing_url, li.name,  
       li.review_scores_rating, COUNT(*) AS comment_count  
FROM reviews  
JOIN listings li ON li.id = reviews.listing_id  
GROUP BY li.listing_url, li.name, li.review_scores_rating  
ORDER BY comment_count DESC;
```

Enlace del alojamiento	Nombre	Valoración	Nº de comentarios
<a href="https://www.airbnb.com/rooms/10802138">https://www.airbnb.com/rooms/10802138</a>	Room next to Airport A/C"	4.66	775
<a href="https://www.airbnb.com/rooms/2343063">https://www.airbnb.com/rooms/2343063</a>	"Nice Flat in the center VFT/MA/00202"	4.78	758
<a href="https://www.airbnb.com/rooms/30141259">https://www.airbnb.com/rooms/30141259</a>	^partamento Centro Málaga+Garaje Privado Gratis"	4.76	672
<a href="https://www.airbnb.com/rooms/2336145">https://www.airbnb.com/rooms/2336145</a>	"Lovely apartment in the center VFT/MA/00204"	4.77	671
<a href="https://www.airbnb.com/rooms/18369830">https://www.airbnb.com/rooms/18369830</a>	"Picasso Home"	4.85	526
<a href="https://www.airbnb.com/rooms/6955285">https://www.airbnb.com/rooms/6955285</a>	"1 Acogedora hab. cerca de la playa"	4.54	521
<a href="https://www.airbnb.com/rooms/13315721">https://www.airbnb.com/rooms/13315721</a>	Relax, culture and beach in the heart of the city."	4.95	510
<a href="https://www.airbnb.com/rooms/21466209">https://www.airbnb.com/rooms/21466209</a>	Çity Home"	4.83	508
...	...	...	...

Cuadro 3: Alojamientos con más comentarios y su valoración

Estos resultados destacan los alojamientos más populares y bien valorados en Málaga. Estos alojamientos han logrado atraer a una gran cantidad de huéspedes satisfechos y reciben una alta valoración por su calidad y servicio ofrecidos.

#### 4.6. Consulta del numero mínimo de noches

En esta consulta, se busca determinar el número mínimo de noches aceptado por los alojamientos en Málaga. El objetivo es comprender la duración mínima de estadía requerida por los anfitriones. A continuación se muestra el código utilizado y el resultado obtenido:

```
SELECT calculate_max_minimum_nights() AS max_minimum_nights;
```

El resultado revela que el número mínimo de noches aceptado por los alojamientos de Málaga es de 3 noches. Este dato es importante ya que indica la cantidad mínima de noches que los huéspedes deben reservar al planificar su estadía en la ciudad. Este resultado resalta la política común entre los anfitriones de exigir un mínimo de 3 noches de estancia. Con esta información, los huéspedes pueden planificar su viaje considerando esta restricción y ajustar su itinerario en consecuencia.

## 4.7. Consulta de anfitriones más activos y su relación con las revisiones

En esta consulta, se muestra la información de los anfitriones más activos en función del número de revisiones recibidas. El objetivo es identificar a los anfitriones con mayor participación y su impacto en las revisiones de los huéspedes. A continuación se muestra el código utilizado y el resultado obtenido:

```
SELECT host_name, host_url, host_listings_count, number_of_reviews
FROM (
    SELECT DISTINCT ON (host_name, host_url)
        host_name, host_url, host_listings_count,
        number_of_reviews
    FROM listings
    ORDER BY host_name, host_url, number_of_reviews DESC
) AS subquery
ORDER BY number_of_reviews DESC;
```

Nombre	Enlace de usuario	Nº aloj.	Nº revisiones
'Christina'	<a href="https://www.airbnb.com/users/show/47971136">https://www.airbnb.com/users/show/47971136</a>	16	775
'Frédéric Et Catherine'	<a href="https://www.airbnb.com/users/show/11931754">https://www.airbnb.com/users/show/11931754</a>	2	758
'Rafa&Pilu'	<a href="https://www.airbnb.com/users/show/226466430">https://www.airbnb.com/users/show/226466430</a>	3	672
'Mónica Y Juan'	<a href="https://www.airbnb.com/users/show/127195773">https://www.airbnb.com/users/show/127195773</a>	3	526
'Sandra'	<a href="https://www.airbnb.com/users/show/36462988">https://www.airbnb.com/users/show/36462988</a>	2	521
'Carmen Y Muni'	<a href="https://www.airbnb.com/users/show/15535210">https://www.airbnb.com/users/show/15535210</a>	2	510
'Ramon'	<a href="https://www.airbnb.com/users/show/32285652">https://www.airbnb.com/users/show/32285652</a>	1	495
'Francisco'	<a href="https://www.airbnb.com/users/show/44170578">https://www.airbnb.com/users/show/44170578</a>	5	493
'Encarni'	<a href="https://www.airbnb.com/users/show/47246887">https://www.airbnb.com/users/show/47246887</a>	1	491
'Christan'	<a href="https://www.airbnb.com/users/show/122669535">https://www.airbnb.com/users/show/122669535</a>	8	479
'Lola'	<a href="https://www.airbnb.com/users/show/268097476">https://www.airbnb.com/users/show/268097476</a>	3	472
'Daniel'	<a href="https://www.airbnb.com/users/show/39846040">https://www.airbnb.com/users/show/39846040</a>	2	459
'Nataliya'	<a href="https://www.airbnb.com/users/show/72851313">https://www.airbnb.com/users/show/72851313</a>	1	456
'Rachel & Hendrix'	<a href="https://www.airbnb.com/users/show/42226210">https://www.airbnb.com/users/show/42226210</a>	1	455

Cuadro 4: Anfitriones más activos y su relación con las revisiones en el año 2023

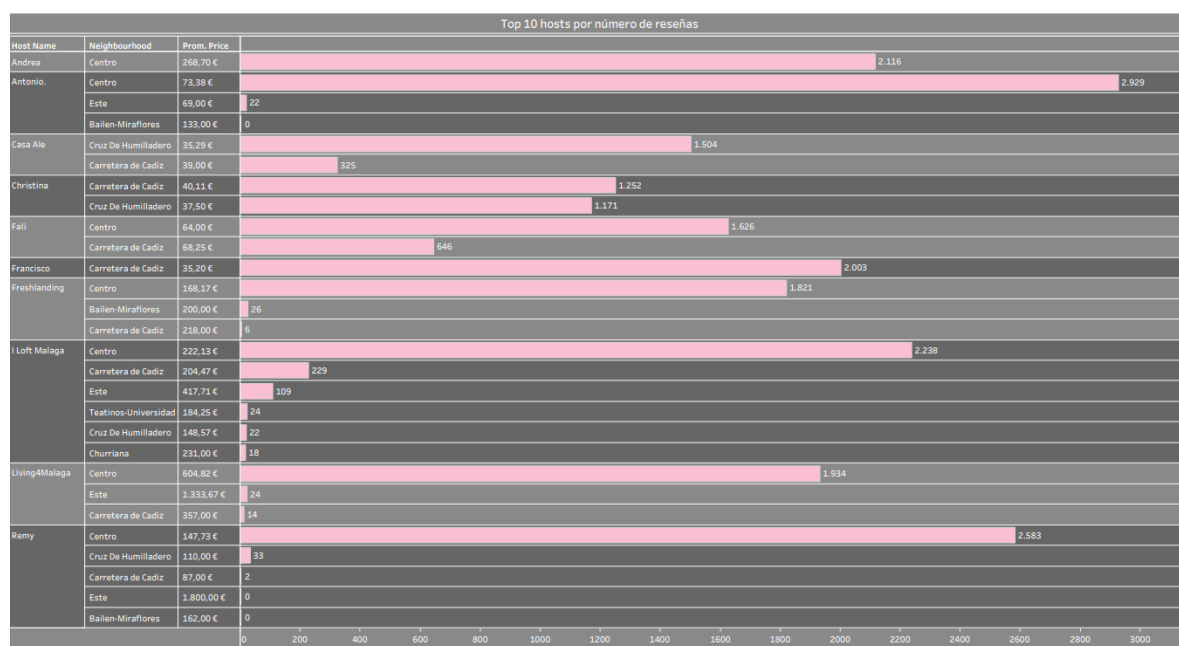


Figura 13: Hosts con más valoraciones junto con sus propiedades.



Estos resultados destacan a los anfitriones más activos y su influencia en las reseñas de los huéspedes. Por ejemplo, 'Christina' es el anfitrión más activo con 16 alojamientos y 775 reseñas en el año 2023. Sin embargo, mirando las reseñas totales vemos en el gráfico de barras que el usuario 'Antonio' tiene varios alojamientos en propiedad y sumando las reseñas totales consigue 2951 revisiones.

Estos anfitriones muestran un alto nivel de participación y una relación positiva con las revisiones de los huéspedes, lo que indica una experiencia satisfactoria para aquellos que se alojaron en sus propiedades. De hecho podemos ver los hosts con más valoraciones junto con sus propiedades

#### 4.8. Consulta de porcentaje de alojamientos con algún comentario negativo

En esta consulta, se analiza el porcentaje de alojamientos que tienen al menos un comentario negativo. Para considerar un comentario como negativo, se buscan palabras clave como 'sucio', 'desordenado' o 'ruidoso' en el campo de comentarios.

Total de alojamientos	Alojamientos con comentarios negativos	Porcentaje (%)
6331	998	15.76

Cuadro 5: Porcentaje de alojamientos con comentarios negativos

El resultado de la consulta muestra que de un total de 6,331 alojamientos analizados, se encontraron 998 alojamientos con comentarios negativos, lo que representa un 15.76 % del total. Esto indica que un porcentaje significativo de alojamientos ha recibido críticas negativas relacionadas con la limpieza, el orden o el ruido. Estos resultados pueden ser útiles para identificar áreas de mejora en los alojamientos y tomar medidas para abordar las preocupaciones de los huéspedes.

#### 4.9. Consulta de alojamientos según la proximidad al centro de Málaga

En esta consulta, se analizan los alojamientos más cercanos al centro de Málaga. El centro de Málaga se define tomando como referencia la [Plaza de la Marina](#), con coordenadas geográficas (36.718437, -4.419820). Para calcular la distancia entre cada alojamiento y el centro, se utiliza la norma euclídea que tiene en cuenta las coordenadas de latitud y longitud de cada alojamiento.

Enlace del alojamiento	Nombre	Proximidad
<a href="https://www.airbnb.com/rooms/42308266">https://www.airbnb.com/rooms/42308266</a>	Super-central double bedroom in Málaga	0.0176
<a href="https://www.airbnb.com/rooms/29068848">https://www.airbnb.com/rooms/29068848</a>	Location! Excellent studio in best location	0.0204
<a href="https://www.airbnb.com/rooms/23742313">https://www.airbnb.com/rooms/23742313</a>	Disfruta de la inmejorable ubicación de este sofisticado y acogedor apartamento	0.0310
<a href="https://www.airbnb.com/rooms/609893377699487005">https://www.airbnb.com/rooms/609893377699487005</a>	Precioso apartamento junto a la Catedral.	0.0381
<a href="https://www.airbnb.com/rooms/43994696">https://www.airbnb.com/rooms/43994696</a>	Calle Larios B	0.0414
<a href="https://www.airbnb.com/rooms/51134076">https://www.airbnb.com/rooms/51134076</a>	Malaga Center Flat Cathedral Balcony City Views	0.0427
<a href="https://www.airbnb.com/rooms/2431701">https://www.airbnb.com/rooms/2431701</a>	Apartamento centro de Málaga	0.0451
<a href="https://www.airbnb.com/rooms/51134069">https://www.airbnb.com/rooms/51134069</a>	Malaga Center Flat Cathedral Balcony Love II	0.0453
<a href="https://www.airbnb.com/rooms/35124090">https://www.airbnb.com/rooms/35124090</a>	iloftmalaga Larios Strachan II	0.0482
<a href="https://www.airbnb.com/rooms/27733098">https://www.airbnb.com/rooms/27733098</a>	Marble Penthouse/Prime location/High Ceilings	0.0486
<a href="https://www.airbnb.com/rooms/51900319">https://www.airbnb.com/rooms/51900319</a>	iloftmalaga Strachan 4	0.0559
<a href="https://www.airbnb.com/rooms/49554979">https://www.airbnb.com/rooms/49554979</a>	Obispo- Espectacular Apart. junto a la Catedral.	0.0611
<a href="https://www.airbnb.com/rooms/41928032">https://www.airbnb.com/rooms/41928032</a>	Malaga Center Flat Cathedral Balcony Relax	0.0655
...	...	...

Cuadro 6: Alojamientos según proximidad al centro de Málaga

Los primeros alojamientos en la lista son los más cercanos al centro. Es interesante observar que existe una gran variedad de alojamientos que están ubicados muy cerca del centro de Málaga. Estos alojamientos ofrecen una ubicación conveniente para los huéspedes que deseen explorar el centro de la ciudad y disfrutar de sus atracciones cercanas.

#### 4.10. Consulta de alojamientos según la proximidad al centro de Málaga ordenados de mejor a peor valoración en función de la distancia.

En esta consulta, se muestra una lista de alojamientos cerca del centro de Málaga, ordenados de mejor a peor valoración en función de su distancia relativa al centro. Al igual que en la consulta anterior la proximidad se calcula utilizando la fórmula de la norma euclidiana aplicada a las coordenadas de latitud y longitud de cada alojamiento.

Enlace del alojamiento	Nombre	Proximidad	Valoración
<a href="https://www.airbnb.com/rooms/42308266">https://www.airbnb.com/rooms/42308266</a>	Super-central double bedroom in Málaga	0.0176	4.85
<a href="https://www.airbnb.com/rooms/29068848">https://www.airbnb.com/rooms/29068848</a>	Location! Excellent studio in best location	0.0204	4.41
<a href="https://www.airbnb.com/rooms/23742313">https://www.airbnb.com/rooms/23742313</a>	Disfruta de la inmejorable ubicación de este sofisticado y acogedor apartamento	0.0310	4.96
<a href="https://www.airbnb.com/rooms/609893377699487005">https://www.airbnb.com/rooms/609893377699487005</a>	Precioso apartamento junto a la Catedral.	0.0381	4.86
<a href="https://www.airbnb.com/rooms/43994696">https://www.airbnb.com/rooms/43994696</a>	Calle Larios B	0.0414	2.5
<a href="https://www.airbnb.com/rooms/51134076">https://www.airbnb.com/rooms/51134076</a>	Malaga Center Flat Cathedral Balcony City Views	0.0427	4.33
<a href="https://www.airbnb.com/rooms/2431701">https://www.airbnb.com/rooms/2431701</a>	Apartamento centro de Málaga	0.0451	4.56
<a href="https://www.airbnb.com/rooms/51134069">https://www.airbnb.com/rooms/51134069</a>	Malaga Center Flat Cathedral Balcony Love II	0.0453	4.67
<a href="https://www.airbnb.com/rooms/35124090">https://www.airbnb.com/rooms/35124090</a>	iloftmalaga Larios Strachan II	0.0482	4.65
<a href="https://www.airbnb.com/rooms/27733098">https://www.airbnb.com/rooms/27733098</a>	Marble Penthouse/Prime location/High Ceilings	0.0486	4.9
<a href="https://www.airbnb.com/rooms/51900319">https://www.airbnb.com/rooms/51900319</a>	iloftmalaga Strachan 4	0.0559	4.96
<a href="https://www.airbnb.com/rooms/49554979">https://www.airbnb.com/rooms/49554979</a>	Obispo- Espectacular Apart. junto a la Catedral.	0.0611	4.5
<a href="https://www.airbnb.com/rooms/41928032">https://www.airbnb.com/rooms/41928032</a>	Malaga Center Flat Cathedral Balcony Relax	0.0655	4.74
<a href="https://www.airbnb.com/rooms/14961348">https://www.airbnb.com/rooms/14961348</a>	New Marina Square by Living4Malaga	0.0663	4.5

Cuadro 7: Alojamientos según proximidad al centro de Málaga (ordenados por valoración)

Estos resultados permiten a los viajeros elegir entre alojamientos bien valorados y convenientemente ubicados. La proximidad al centro puede ser un factor relevante para aquellos que deseen estar cerca de las atracciones principales de la ciudad. Sin embargo, es importante considerar tanto la ubicación como la valoración general del alojamiento al tomar una decisión.

#### 4.11. Consulta de valoración promedio del alojamiento en función del tipo de habitación.

La consulta realizada tiene como objetivo obtener la valoración promedio de los alojamientos según el tipo de habitación en función de los puntajes de las revisiones. Se agrupan los alojamientos por el tipo de habitación y se calcula el promedio redondeado de las valoraciones. Los alojamientos que no tienen una valoración registrada se excluyen de la consulta.

```
SELECT room_type, ROUND(AVG(review_scores_rating)) AS avg_rating
FROM listings
WHERE review_scores_rating IS NOT NULL
GROUP BY room_type
ORDER BY avg_rating DESC;
```

Tipo de habitación	Valoración promedio
'Entire home/apt'	5
'Private room'	5
'Hotel room'	5
'Shared room'	4

Cuadro 8: Valoración promedio del alojamiento según el tipo de habitación

Se observa que los tipos de habitación de casa completa/apartamento y las habitaciones de hotel tienen una valoración promedio de 5, mientras que las habitaciones compartidas ('Shared room') tienen una valoración promedio de 4. Esto indica que, en general, los huéspedes tienden a otorgar una valoración más alta a los alojamientos que ofrecen mayor privacidad y comodidad.

#### 4.12. Consulta de porcentaje de propietarios con más de una propiedad

La consulta tiene como objetivo calcular el porcentaje de propietarios que poseen más de una propiedad de alojamiento en la ciudad de Málaga.

```
SELECT calculate_multiple_properties_percentage() AS percentage;
```

El resultado de la consulta es 32. Esto significa que aproximadamente el 32 % de los propietarios en Málaga tienen más de una propiedad de alojamiento en Airbnb. En la siguiente gráfica circular podemos ver desglosado por porcentajes según cuántos propietarios tienen cierto número de alojamientos publicados en Airbnb.

**Número de Alojamientos por Usuario**

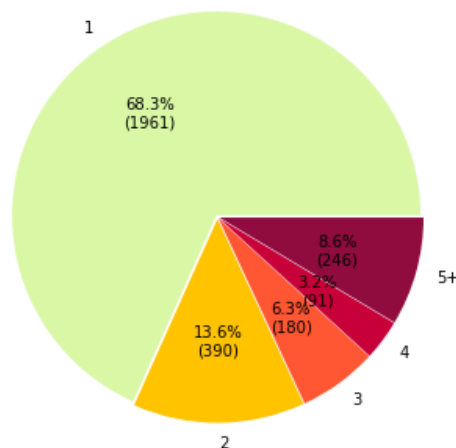


Figura 14: Número de alojamientos por usuario.

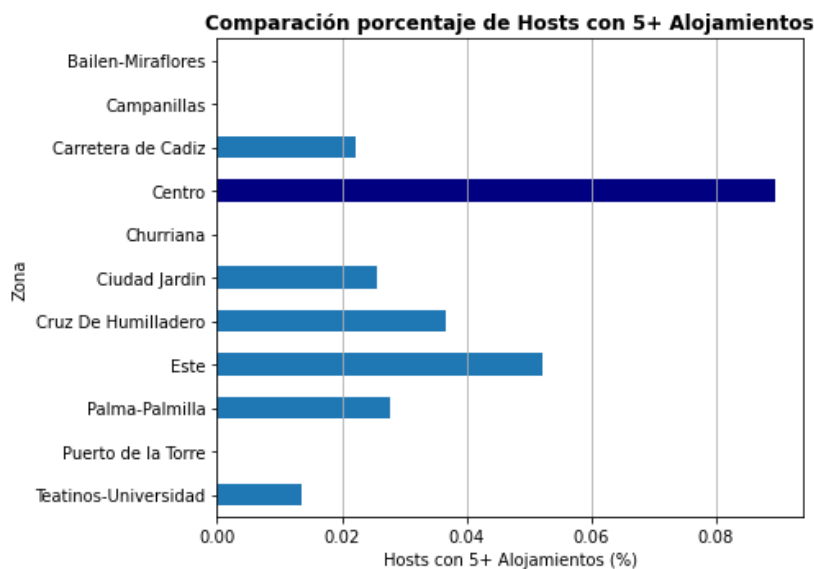


Figura 15: Comparación porcentaje de Hosts con +5 alojamientos.

Vemos que el distrito *Centro* un 10% de los propietarios tiene más de 5 alojamientos publicados en Airbnb. Es interesante notar que un porcentaje significativo de propietarios tiene múltiples propiedades, lo que puede sugerir una tendencia hacia el negocio de alquileres a corto plazo en la ciudad. Este dato puede ser relevante para el mercado inmobiliario y la industria turística, ya que muestra el nivel de participación de ciertos propietarios en el mercado de alquileres vacacionales.

#### 4.13. Consulta de cuántos hosts se unen cada año

La consulta tiene como objetivo mostrar cuántos hosts (propietarios) se unen cada año en Airbnb en la ciudad de Málaga. Estos datos son útiles para comprender la evolución del mercado de alquileres vacacionales en la ciudad y pueden ser útiles para las autoridades locales y profesionales del sector turístico para tomar decisiones y adaptar sus estrategias en función de las tendencias y cambios en la plataforma.

```
SELECT EXTRACT(YEAR FROM host_since) AS joined_year,  
       COUNT(DISTINCT host_id) AS new_hosts  
FROM listings  
GROUP BY joined_year  
ORDER BY joined_year;
```

Año de Unión	Nuevos Hosts
2009	1
2010	1
2011	26
2012	88
2013	140
2014	246
2015	357
2016	422
2017	360
2018	283
2019	280
2020	132
2021	184
2022	282
2023	70

Cuadro 9: Número de nuevos hosts que se unieron cada año en Airbnb en Málaga.

El resultado de la consulta nos muestra que ha habido un crecimiento notable a lo largo de los años fruto de la incipiente popularidad de esta plataforma. Sugiere además un crecimiento continuo de la oferta de alojamientos en la plataforma en la ciudad.

#### 4.14. Consulta de porcentaje y nº camas por alojamiento

Esta consulta muestra el porcentaje de alojamientos en Málaga según el número de camas que ofrecen, así como la cantidad absoluta de alojamientos en cada categoría. El objetivo es comprender la distribución de camas en la oferta de alojamientos en la ciudad. Los resultados obtenidos son los siguientes:

```
SELECT beds, COUNT(*) AS count,  
       ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM listings), 2) AS percentage  
FROM listings  
GROUP BY beds  
ORDER BY percentage DESC;
```

Número de Camas	Cantidad	Porcentaje (%)
1	2308	30.63
2	2248	29.84
3	1370	18.18
4	833	11.06
5	326	4.33
6	167	2.22
7	79	1.05
8	44	0.58
9	26	0.35
10	31	0.41
...	...	...

Cuadro 10: Porcentaje y número de camas por alojamiento en Málaga.

La tabla muestra que la mayoría de los alojamientos en Málaga ofrecen 1 o 2 camas, representando aproximadamente el 60.47% del total. Por otro lado, los alojamientos con más de 4 camas son menos comunes, y los que ofrecen 6 camas o más representan menos del 6% del total. Estos datos son útiles para entender la capacidad y distribución de alojamientos disponibles para los viajeros en Málaga.

#### 4.15. Consulta de porcentaje de precio

Esta consulta proporciona información sobre la distribución de los precios de alojamientos en Málaga y su frecuencia relativa. Los resultados obtenidos son los siguientes:

Precio	Cantidad	Porcentaje (%)
56.00 €	148	1.96
150.00 €	124	1.65
100.00 €	123	1.63
120.00 €	110	1.46
90.00 €	91	1.21
60.00 €	78	1.04
70.00 €	78	1.04
80.00 €	76	1.01
200.00 €	72	0.96
65.00 €	66	0.88
...	...	...

Cuadro 11: Porcentaje de precio de alojamientos en Málaga.

La tabla muestra la distribución de los precios de alojamientos en Málaga y su frecuencia relativa en la base de datos. El precio modal es 56.00 € con una frecuencia relativa del 1.96%. Es interesante observar cómo los precios se distribuyen a lo largo de diferentes rangos y cómo la mayoría de los precios tienen una frecuencia baja, lo que indica una diversidad de opciones de alojamiento para los visitantes.

Además, se presentan dos visualizaciones gráficas que profundizan en la distribución de los precios de los alojamientos:

Esta figura muestra la distribución de precios de alojamientos en Málaga. Es evidente que la concentración de datos en el rango más bajo y la dispersión en los rangos superiores hacen que la gráfica sea difícil de interpretar. Algunos valores atípicos extremos pueden estar afectando la visualización.

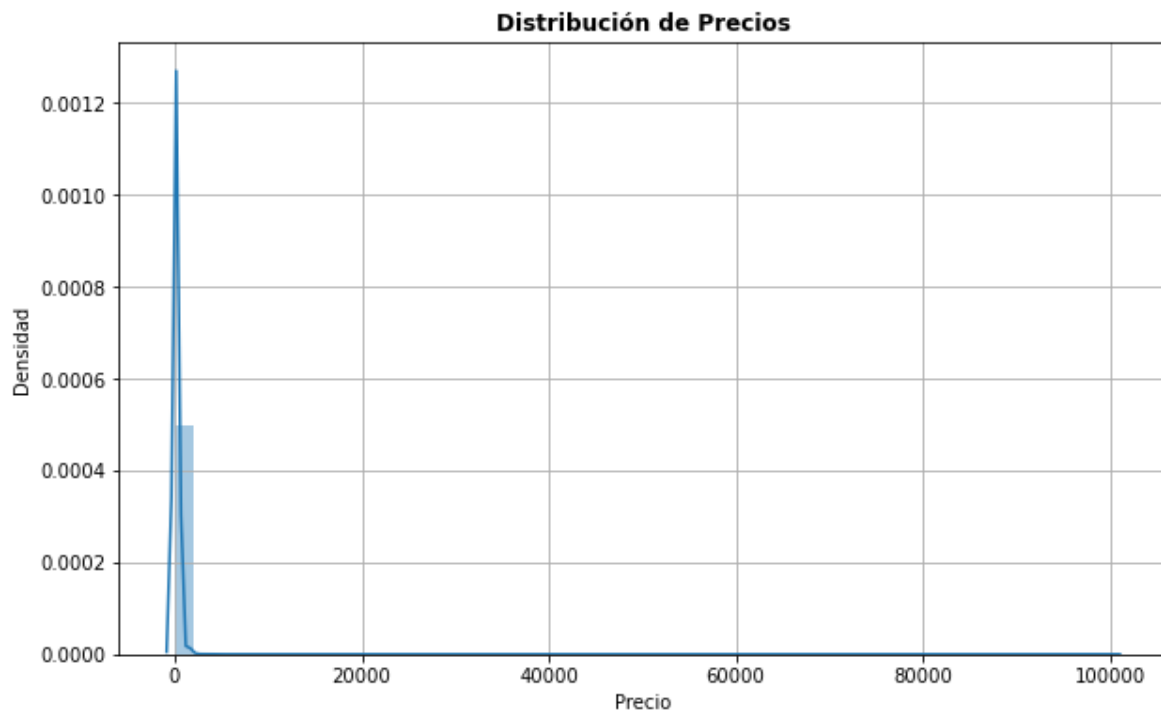


Figura 16: Distribución de Precios.

Para abordar la concentración en la parte inferior de la gráfica, se ha aplicado una transformación logarítmica a los precios. Esto ayuda a suavizar la visualización y hacer que los datos sean más comprensibles. La gráfica presenta una distribución más uniforme de los precios en comparación con la gráfica anterior.

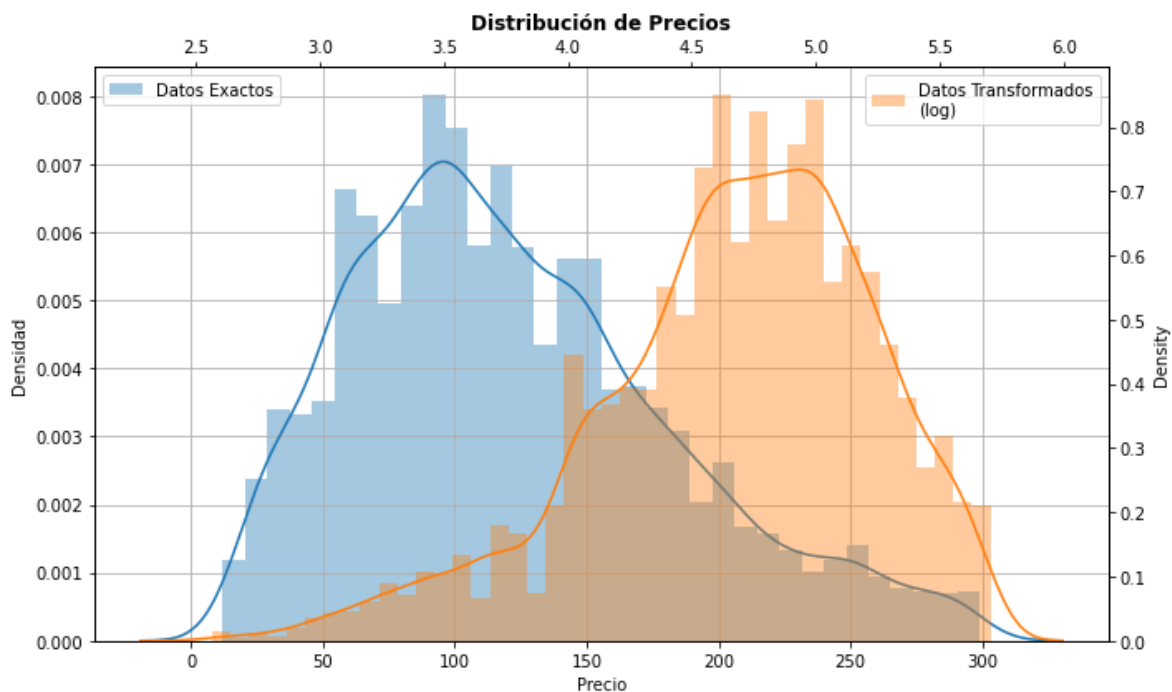


Figura 17: Distribución de Precios (log).

#### 4.16. Porcentaje de mínimo numero de noches

La siguiente consulta muestra el porcentaje de alojamientos en Málaga según el número mínimo de noches requerido para reservar: Los resultados obtenidos son los siguientes:

Mínimo de Noches	Cantidad	Porcentaje (%)
3	6350	82,29
4	457	6.07
5	271	3.60
7	153	2.03
30	70	0.93
6	70	0.93
60	51	0.68
10	19	0.25
28	14	0.19
15	13	0.17
32	9	0.12
61	8	0.11
...	...	...

Cuadro 12: Porcentaje de mínimo número de noches en alojamientos de Málaga.

Podemos observar que la mayoría de los alojamientos en Málaga tienen un mínimo de 3 noches requeridas para reservar, representando el 82,29 %. Los mínimos de 4, 5 y 6 noches también son comunes, mientras que los mínimos más largos (30, 60, etc.) tienen un porcentaje mucho menor en la oferta de alojamientos.

#### 4.17. Porcentaje de puntuaciones

La siguiente consulta muestra el porcentaje de puntuaciones en los alojamientos de Málaga:

```
SELECT review_scores_rating, COUNT(*) AS count,
       ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM listings), 2)
       AS percentage
FROM listings WHERE review_scores_rating IS NOT NULL
GROUP BY review_scores_rating
ORDER BY percentage DESC;
```

Puntuación	Cantidad	Porcentaje (%)
5.00	1061	14.08
4.50	248	3.29
4.00	239	3.17
4.67	208	2.76
4.75	148	1.96
4.83	134	1.78
4.80	118	1.57
4.88	108	1.43
...	...	...

Cuadro 13: Porcentaje de puntuaciones en alojamientos de Málaga.

Encontramos que la mayoría de los alojamientos en Málaga tienen una puntuación de 5.00 (la calificación más alta posible), representando el 14.08 % del total de alojamientos. También es notable la presencia de puntuaciones en torno a 4.0 y 4.5, aunque con porcentajes menores. La distribución de las puntuaciones muestra una tendencia hacia calificaciones altas, lo que sugiere que la mayoría de los alojamientos en Málaga son bien valorados por los huéspedes.

#### 4.18. Consulta de media de precios de alojamiento en función de su valoración

La siguiente consulta muestra la media de precios de los alojamientos en Málaga según su valoración. A continuación se muestra el código de la vista y un fragmento de los resultados:

```
SELECT rating_range,
       ROUND(AVG(CAST(REGEXP_REPLACE(price, '[^\d.]', '', 'g')
                        AS NUMERIC)), 2) AS average_price
FROM ( SELECT
      CASE
        WHEN review_scores_rating BETWEEN 0 AND 1 THEN '0-1'
        WHEN review_scores_rating > 1 AND review_scores_rating <= 2 THEN '1-2'
        WHEN review_scores_rating > 2 AND review_scores_rating <= 3 THEN '2-3'
        WHEN review_scores_rating > 3 AND review_scores_rating <= 4 THEN '3-4'
        WHEN review_scores_rating > 4 AND review_scores_rating <= 5 THEN '4-5' END
      AS rating_range,
      price
    FROM listings
    WHERE review_scores_rating IS NOT NULL
  ) AS subquery
GROUP BY rating_range
ORDER BY rating_range;
```

Los resultados obtenidos son los siguientes:

Valoración	Precio Medio
0-1	154.77 €
1-2	136.25 €
2-3	609.79 €
3-4	300.56 €
4-5	187.05 €

Cuadro 14: Media de precios de alojamiento en función de su valoración en Málaga.

Se observa que los alojamientos con una valoración de 2 a 3 tienen un precio medio excepcionalmente alto en comparación con otras categorías de valoración, mientras que aquellos con una valoración de 1 a 2 y 4 a 5 tienden a tener precios medios más bajos.

Estos resultados sugieren que los alojamientos mejor valorados y peor valorados en Málaga pueden ofrecer una mejor relación calidad-precio en términos de precios promedio.

A continuación, exploramos una gráfica que arroja luz sobre la relación entre el número de reseñas y los precios de los alojamientos en Málaga. Esta visualización proporciona información valiosa sobre cómo la popularidad y la valoración de un alojamiento pueden influir en su precio promedio. Muestra cómo esta relación podría tener implicaciones para los anfitriones y los huéspedes.



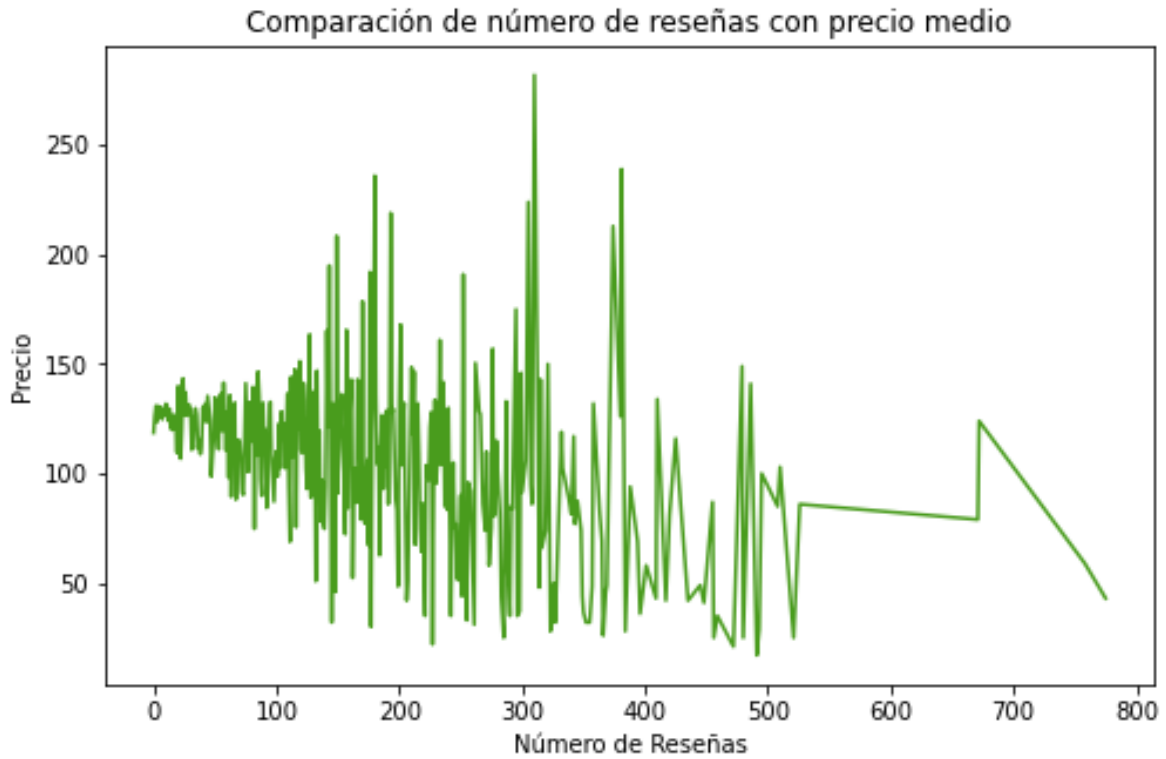


Figura 18: Comparación de número de reseñas y precios.

La fluctuación de precios en función del número de reseñas sugiere que existe una relación compleja entre la popularidad de un alojamiento y su precio promedio. Se observa que, en general, alojamientos con alrededor de 100 reseñas tienden a tener precios medios que oscilan entre 80 y 130 euros. A medida que el número de reseñas aumenta, los precios también fluctúan considerablemente. Por ejemplo, alojamientos con alrededor de 300 reseñas varían en precio desde 80 euros hasta más de 200 euros. Sin embargo, en el rango de 600-700 reseñas, la variabilidad es menor, con precios estabilizándose en torno a 100 euros o 130 euros. Esto puede deberse a que haya una menor muestra de alojamientos con tantas reseñas.



## 5. Vistas

A continuación, se presentan algunos ejemplos de vistas creadas en el proyecto, junto con sus resultados y conclusiones.

### 5.1. Vista de los alojamientos con mayor disponibilidad anual

En esta consulta podemos ver los alojamientos con mayor disponibilidad anual. La vista 'year\_availability' muestra los alojamientos con mayor disponibilidad durante todo el año en Málaga. Permite facilitar a los usuarios la identificación de opciones con una alta disponibilidad para reservar en cualquier época. A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW year_availability AS
SELECT name, availability_365
FROM listings
ORDER BY availability_365 DESC
```

Nombre del alojamiento	Enlace del alojamiento	Disponibilidad anual
EXCELENTE PISO CASCO HISTORICO	<a href="https://www.airbnb.com/rooms/678778297636694785">https://www.airbnb.com/rooms/678778297636694785</a>	365
Habitación acogedora cerca del centro de Málaga	<a href="https://www.airbnb.com/rooms/634843206096078765">https://www.airbnb.com/rooms/634843206096078765</a>	365
Modern and Sunny Loft Terrace closed to beach	<a href="https://www.airbnb.com/rooms/26564973">https://www.airbnb.com/rooms/26564973</a>	365
Charming apartment in Málaga near seabeach	<a href="https://www.airbnb.com/rooms/819266217210297893">https://www.airbnb.com/rooms/819266217210297893</a>	365
Welcoming apartment at the center of Málaga	<a href="https://www.airbnb.com/rooms/617319758965018121">https://www.airbnb.com/rooms/617319758965018121</a>	365
1 Bedrooms Apartment La Mundial	<a href="https://www.airbnb.com/rooms/44010462">https://www.airbnb.com/rooms/44010462</a>	365
Luxurious Apartment Gold Mile	<a href="https://www.airbnb.com/rooms/790376052631156910">https://www.airbnb.com/rooms/790376052631156910</a>	365
...	...	...

Cuadro 15: Alojamientos con mayor disponibilidad anual en Málaga.

La vista 'year\_availability' proporciona una lista de alojamientos que permanecen disponibles durante todo el año, lo que puede ser útil para viajeros que buscan opciones flexibles en cuanto a fechas de estancia. Existe una gran cantidad de alojamientos con plena disponibilidad durante todo un año. Si consultamos la vista desde la última fila observaremos los alojamientos con disponibilidad más corta, que puede ser de gran utilidad para viajeros con planes más concretos o aquellos que buscan estancias más cortas.

### 5.2. Vista de la disponibilidad y el precio diario de los listados

Se creó una vista llamada 'listing\_availability\_price' para mostrar la disponibilidad y el precio diario de los listados. Con esta vista, los usuarios pueden obtener información sobre las fechas disponibles para cada alojamiento y su precio diario correspondiente. A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW listing_availability_price AS
SELECT l.listing_url, l.name, MIN(c.date) AS start_date,
MAX(c.date) AS end_date,
ROUND(AVG(CAST(REGEXP_REPLACE(REPLACE(l.price, '$', ''),
'[\d.]', '', 'g') AS NUMERIC)), 2) AS price
FROM calendar c
JOIN listings l ON c.listing_id = l.id
WHERE c.available = 't'
GROUP BY l.id, l.name, l.price;
```

URL	Nombre del alojamiento	Fecha de inicio	Fecha de fin	Precio diario
<a href="https://www.airbnb.com/rooms/31039156">https://www.airbnb.com/rooms/31039156</a>	CENTRO HISTÓRICO 2 HABITACIONES LAZCANO	2023-04-09	2023-09-30	103.00 €
<a href="https://www.airbnb.com/rooms/85444784432131414">https://www.airbnb.com/rooms/85444784432131414</a>	Bluemarine room for tour	2023-04-03	2023-04-30	65.00 €
<a href="https://www.airbnb.com/rooms/574649101626526498">https://www.airbnb.com/rooms/574649101626526498</a>	Perfect Suite Train station	2023-04-09	2023-06-29	76.00 €
<a href="https://www.airbnb.com/rooms/749258261295533795">https://www.airbnb.com/rooms/749258261295533795</a>	Nuevo loft - Piscina, Playa & relax	2023-04-09	2024-03-29	45.00 €
<a href="https://www.airbnb.com/rooms/594786518117713438">https://www.airbnb.com/rooms/594786518117713438</a>	La Perla de Málaga	2023-04-02	2024-03-30	118.00 €
...	...	...	...	...

Cuadro 16: Disponibilidad y precio diario de los alojamientos en Málaga

### 5.3. Vista del precio medio según el número de habitaciones

Se creó una vista llamada 'price\_by\_bedrooms' para mostrar el precio medio según el número de habitaciones. Esto puede ser útil para los viajeros que deseen encontrar alojamientos que se ajusten a sus necesidades y presupuesto en función del tamaño del grupo con el que viajan.

A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW price_by_bedrooms AS
SELECT bedrooms,
       ROUND(AVG(CAST(REGEXP_REPLACE(price, '[^\d.]', '', 'g')
                          AS NUMERIC))) AS avg_price
FROM listings
GROUP BY bedrooms
ORDER BY avg_price DESC;
```

Número de Habitaciones	Precio Medio
25	1285 €
24	1285 €
12	1106 €
11	676 €
10	600 €
15	564 €
8	561 €
9	465 €
7	442 €
5	428 €
6	413 €
3	292 €
4	277 €
2	213 €
1	146 €

Cuadro 17: Precio medio según el número de habitaciones de los alojamientos en Málaga.

Se puede observar que los alojamientos con un número inusualmente alto de habitaciones, como 24 o 25 habitaciones, tienen precios medios excepcionalmente altos. Esto sugiere que los alojamientos de gran capacidad tienden a ser más costosos en promedio, lo cual es comprensible debido a su capacidad para acomodar a un gran número de personas.

### 5.4. Vista de las tendencias de precios medios a lo largo de los años

La vista 'price\_trends' se ha creado con el propósito de mostrar las tendencias de los precios medios de alojamientos a lo largo de los años. Esta información puede ser útil para los viajeros y propietarios de alojamientos, ya que proporciona una visión general de cómo han evolucionado los precios a lo largo del tiempo.

A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW price_trends AS
SELECT TO_CHAR(date, 'YYYY-MM') AS year_month,
       ROUND(AVG(CAST(REGEXP_REPLACE(price, '[^\d.]', '', 'g')
                          AS NUMERIC))) AS avg_price
FROM calendar
GROUP BY TO_CHAR(date, 'YYYY-MM')
ORDER BY year_month;
```

Año-Mes	Precio Medio
2023-03	196 €
2023-04	181 €
2023-05	174 €
2023-06	165 €
2023-07	184 €
2023-08	205 €
2023-09	195 €
2023-10	208 €
2023-11	197 €
2023-12	215 €
2024-01	215 €
2024-02	236 €
2024-03	265 €

Cuadro 18: Tendencias de precios medios a lo largo de los años en alojamientos de Málaga.

Los viajeros pueden utilizar esta información para planificar sus viajes en momentos que se ajusten a su presupuesto, mientras que los propietarios de alojamientos pueden tomar decisiones informadas sobre cómo ajustar sus tarifas según la demanda en diferentes períodos del año.

## 5.5. Vista de las tendencias de disponibilidad en los próximos meses

Se creó una vista llamada 'availability\_trends' para mostrar las tendencias de disponibilidad que se presentarán a lo largo de los próximos meses. Esta información es valiosa tanto para los viajeros como para los propietarios de alojamientos, ya que proporciona una visión general de la disponibilidad de alojamientos en diferentes períodos del tiempo.

A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW availability_trends
AS
SELECT TO_CHAR(date, 'YYYY-MM') AS year_month,
       sum(
         CASE
           WHEN calendar.available = 't'::bpchar THEN 1
           ELSE 0
         END) AS available_listings,
       sum(
         CASE
           WHEN calendar.available = 'f'::bpchar THEN 1
           ELSE 0
         END) AS unavailable_listings
FROM calendar
GROUP BY TO_CHAR(date, 'YYYY-MM')
ORDER BY year_month;
```

Para complementar la información proporcionada por la tabla, presentamos la siguiente gráfica que muestra la evolución del precio medio de los alojamientos en los mismos meses en comparación con su disponibilidad. Esta gráfica nos permitirá visualizar cómo se relaciona la disponibilidad con el precio promedio, y si existen correlaciones entre estos dos aspectos.

Año-Mes	Alojamientos Disponibles	Alojamientos No Disponibles
2023-03	364	1443
2023-04	71907	154113
2023-05	115274	118280
2023-06	131775	94245
2023-07	132310	101244
2023-08	128159	105395
2023-09	141768	84252
2023-10	132191	101363
2023-11	133738	92282
2023-12	135198	98356
2024-01	113267	120287
2024-02	99131	119355
2024-03	101726	122403

Cuadro 19: Tendencias de disponibilidad en los próximos meses en alojamientos de Málaga.

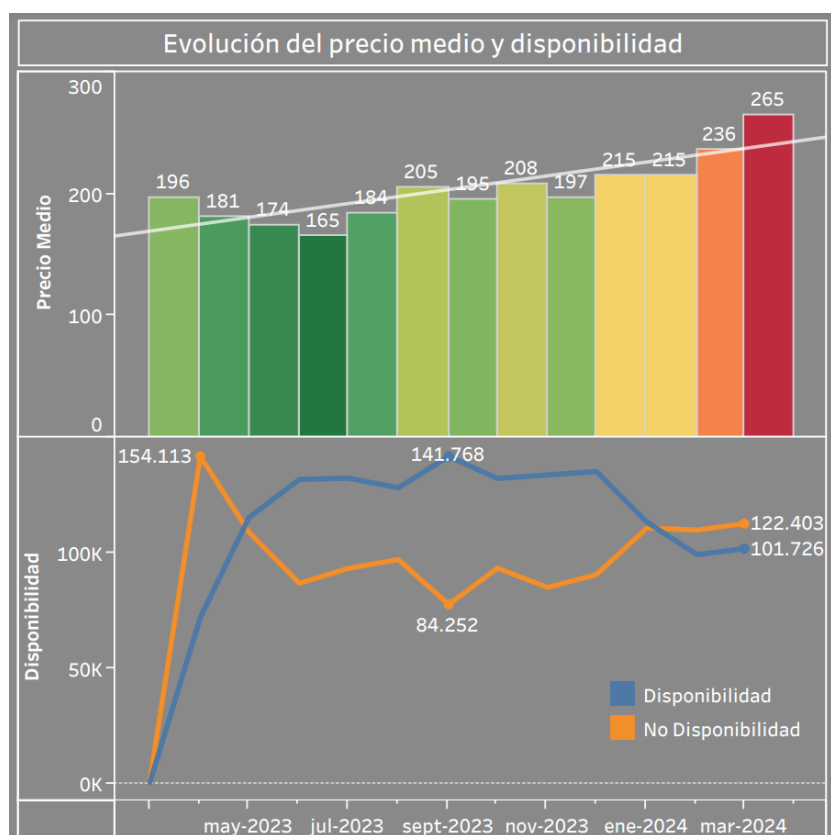


Figura 19: Evolución del precio medio y disponibilidad.

En algunos meses, como abril, mayo y junio, la disponibilidad es más alta, lo que puede indicar una mayor demanda durante las temporadas turísticas o vacacionales. Por otro lado, en meses como enero y febrero, la disponibilidad tiende a ser más baja, lo que podría relacionarse con la temporada baja después del invierno en la ciudad.

Vamos a analizar en detalle la relación entre las valoraciones y las reseñas en diferentes aspectos. A continuación, se presentan una serie de gráficas que nos proporcionan información valiosa sobre las tendencias en función de las valoraciones y reseñas:

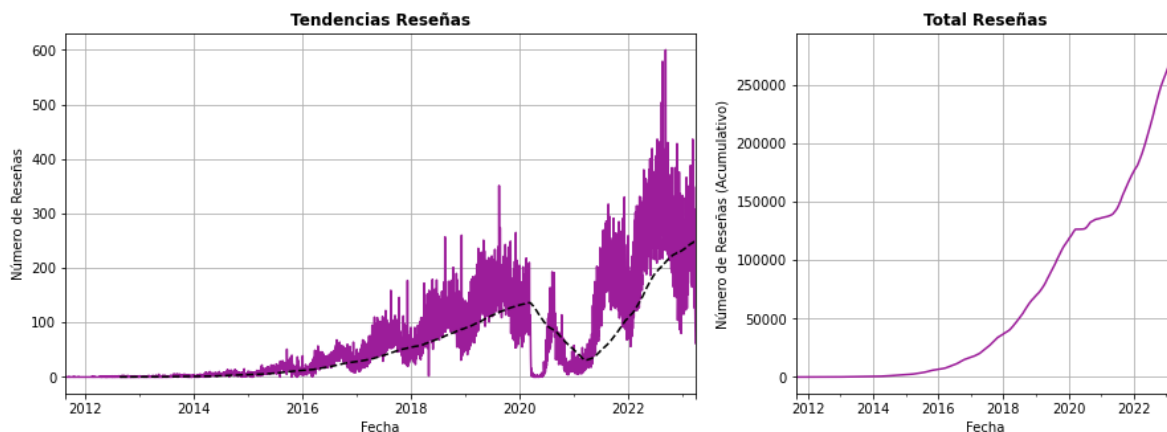


Figura 20: Reseñas a lo largo de los años.

Se observa un patrón de crecimiento constante, con un notable aumento a partir de 2018. Este aumento puede ser atribuido al aumento de la popularidad de plataformas como Airbnb. Sin embargo, en 2020, se observa una disminución marcada en las reseñas, posiblemente debido a la pandemia de COVID-19 que afectó los viajes y la hospitalidad. A partir de 2021, las reseñas vuelven a aumentar, lo que sugiere una recuperación en el sector.

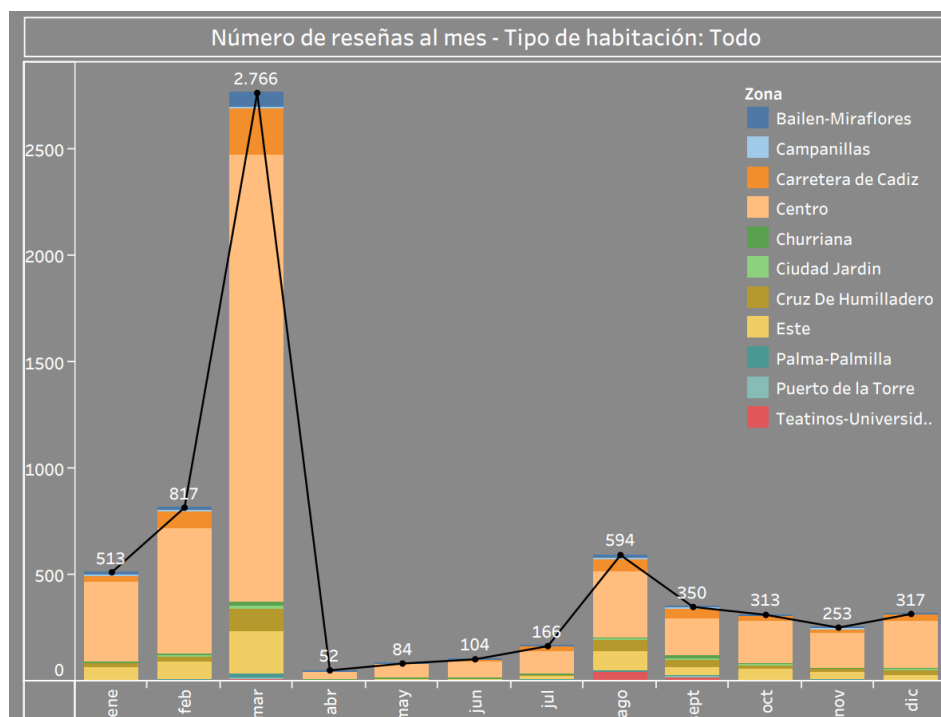


Figura 21: Promedio de reseñas por mes.

Esta gráfica muestra el promedio de reseñas por mes y por barrio. Resalta que en marzo, el número promedio de reseñas es más alto, alcanzando un pico de 2766 reseñas en ese mes. Además, se observa que el distrito *Centro* tiene la mayoría de las reseñas, pues es donde se sitúan la mayoría de alojamientos.

Promedio de reseñas al mes											
Room Type	Bailen-..	Campa..	Carretera de Ca..	Centro	Churria..	Ciudad Jardin	Cruz De Humilla..	Este	Palma..	Puerto de la To..	Teatino..
Casa/aptarta..	1,639	0,886	1,165	1,415	0,746	0,797	0,823	0,818	1,005	0,687	0,577
Habitación compartida				1,155	1,430			0,310			
Habitación de hotel	0,770			0,300							
Habitación privada	1,665	1,845	2,591	1,588	1,644	1,238	2,436	0,731	1,403	1,790	0,871

Figura 22: Promedio de reseñas por mes por tipo de habitación en cada Zona.

Se presenta un gráfico que muestra el promedio de reseñas por mes y por tipo de habitación en cada zona. Se puede notar que en el distrito *Carretera de Cádiz* y *Cruz de Humilladero*, las habitaciones privadas tienen una mayor cantidad de reseñas en comparación con otros tipos de habitaciones. Esto podría indicar que estas zonas son preferidas por los huéspedes que buscan habitaciones privadas en lugar de alquileres completos o compartidos.

## 5.6. Vista de los alojamientos mejor valorados por categoría de propiedad

Se creó una vista llamada 'top\_rated\_listings' para mostrar los alojamientos mejor valorados por categoría de propiedad. Esta información puede ser de gran interés para los viajeros que buscan opciones de alojamiento con altas valoraciones por parte de otros usuarios.

A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW top_rated_listings AS
SELECT listing_url, property_type, name, review_scores_rating
FROM listings
WHERE review_scores_rating IS NOT NULL
ORDER BY review_scores_rating DESC;
```

Enlace del Alojamiento	Tipo de Propiedad	Nombre	Valoración
<a href="https://www.airbnb.com/rooms/25233010">https://www.airbnb.com/rooms/25233010</a>	Entire loft	GUADALMAR SEA&GOLF	5
<a href="https://www.airbnb.com/rooms/33438511">https://www.airbnb.com/rooms/33438511</a>	Entire rental unit	MALAGUETA PLAYA AZUL MEDITERRANEO	5
<a href="https://www.airbnb.com/rooms/33422857">https://www.airbnb.com/rooms/33422857</a>	Entire rental unit	Apartamento Las Dalias, Málaga, Costa del sol.	5
<a href="https://www.airbnb.com/rooms/41525349">https://www.airbnb.com/rooms/41525349</a>	Room in boutique hotel	Hotel Del Pintor, Habitación económica	5
<a href="https://www.airbnb.com/rooms/33073492">https://www.airbnb.com/rooms/33073492</a>	Entire rental unit	VILLA CANDADO PRIVATE SWIMMING POOL	5
<a href="https://www.airbnb.com/rooms/34592596">https://www.airbnb.com/rooms/34592596</a>	Entire home	Playa 45, first line beach house in El Palo Málaga	5
<a href="https://www.airbnb.com/rooms/34515076">https://www.airbnb.com/rooms/34515076</a>	Entire rental unit	Studio Arts Malaga-Soho	5
<a href="https://www.airbnb.com/rooms/33728797">https://www.airbnb.com/rooms/33728797</a>	Entire rental unit	Apartamento Palma	5
<a href="https://www.airbnb.com/rooms/33108239">https://www.airbnb.com/rooms/33108239</a>	Entire rental unit	Apartamento El Marqués 6 San Juan 1	5
...	...	...	...

Cuadro 20: Alojamientos mejor valorados por categoría de propiedad.

Si consultamos la vista desde la última fila observaremos los alojamientos con peor valoración. Esto puede ser útil para los viajeros que desean evitar opciones con calificaciones bajas y asegurarse de tener una experiencia positiva durante su estancia.

Enlace	Tipo de Propiedad	Nombre	Valoración
<a href="https://www.airbnb.com/rooms/24762693">https://www.airbnb.com/rooms/24762693</a>	Entire home	Apartamento en centro de Málaga	0
<a href="https://www.airbnb.com/rooms/41375127">https://www.airbnb.com/rooms/41375127</a>	Entire loft	Loft Larios, único y exclusivo, repetirás!	0
<a href="https://www.airbnb.com/rooms/41902864">https://www.airbnb.com/rooms/41902864</a>	Entire rental unit	Camino Suárez 2	0
<a href="https://www.airbnb.com/rooms/50916471">https://www.airbnb.com/rooms/50916471</a>	Entire rental unit	Holidays2Malaga Reyna Manescau - Huelin Area just 450 mts from beach	1
<a href="https://www.airbnb.com/rooms/702141724072577174">https://www.airbnb.com/rooms/702141724072577174</a>	Entire rental unit	"Nuevo apartamento Catedral (centro Málaga)"	1
...	...	...	...

Cuadro 21: Alojamientos peor valorados por categoría de propiedad.



## 5.7. Vista de los alojamientos con comentarios negativos

Se creó una vista llamada 'negative\_reviews' para mostrar los alojamientos con comentarios negativos. Los comentarios son extraídos de la tabla de reseñas y se comparan con palabras clave que podrían indicar aspectos negativos de la estancia, como suciedad, desorden o ruido.

A continuación se muestra el código de la vista.

```
CREATE OR REPLACE VIEW negative_reviews AS
  SELECT li.host_name, li.listing_url, li.host_id, li.host_url,
         r.comments
  FROM reviews r
 JOIN listings li ON r.listing_id = li.id
 WHERE r.comments LIKE '%sucio%' OR
        r.comments LIKE '%desordenado%' OR
        r.comments LIKE '%ruidoso%'
 ORDER BY li.id;
```

Al revisar esta vista, los anfitriones y propietarios pueden identificar áreas de mejora y tomar medidas para brindar una estancia más satisfactoria a sus huéspedes en el futuro. Además, los viajeros pueden usar esta información para tomar decisiones informadas al elegir su alojamiento, evitando potenciales problemas y asegurando una experiencia positiva durante su estancia en Málaga.

## 5.8. Vista de los propietarios con más reseñas con comentarios negativos

Se creó una vista llamada 'owner\_negative\_reviews' para mostrar los propietarios con más reseñas con comentarios negativos. Esta vista se ha creado para identificar a los propietarios que han recibido más comentarios negativos en sus alojamientos, lo que puede ser útil para tomar decisiones informadas al seleccionar un lugar para hospedarse.

A continuación se muestra el código de la vista y un fragmento de los resultados:

```
CREATE OR REPLACE VIEW owner_negative_reviews AS
  SELECT li.host_name, li.host_url, COUNT(*)
         AS negative_review_count
  FROM reviews r
 JOIN listings li ON r.listing_id = li.id
 WHERE r.comments LIKE '%sucio%' OR
        r.comments LIKE '%desordenado%' OR
        r.comments LIKE '%ruidoso%'
 GROUP BY li.host_name, li.host_id, li.host_url
 ORDER BY negative_review_count DESC;
```

Propietario	Enlace del Propietario	Cantidad de Reseñas Negativas
Malaga	<a href="https://www.airbnb.com/users/show/252959314">https://www.airbnb.com/users/show/252959314</a>	18
Homeabout	<a href="https://www.airbnb.com/users/show/17499500">https://www.airbnb.com/users/show/17499500</a>	17
MimoRooms	<a href="https://www.airbnb.com/users/show/7082899">https://www.airbnb.com/users/show/7082899</a>	16
Urbe10	<a href="https://www.airbnb.com/users/show/151527355">https://www.airbnb.com/users/show/151527355</a>	14
Ana	<a href="https://www.airbnb.com/users/show/9744958">https://www.airbnb.com/users/show/9744958</a>	11
Alberto	<a href="https://www.airbnb.com/users/show/138864019">https://www.airbnb.com/users/show/138864019</a>	11
Welcome	<a href="https://www.airbnb.com/users/show/329266025">https://www.airbnb.com/users/show/329266025</a>	11
Antonio.	<a href="https://www.airbnb.com/users/show/89705945">https://www.airbnb.com/users/show/89705945</a>	10
Vintage Flat	<a href="https://www.airbnb.com/users/show/324236176">https://www.airbnb.com/users/show/324236176</a>	9
La Recepción	<a href="https://www.airbnb.com/users/show/5890675">https://www.airbnb.com/users/show/5890675</a>	9
Casa Ale	<a href="https://www.airbnb.com/users/show/82314525">https://www.airbnb.com/users/show/82314525</a>	9
Mariano	<a href="https://www.airbnb.com/users/show/3301281">https://www.airbnb.com/users/show/3301281</a>	9
Remy	<a href="https://www.airbnb.com/users/show/278821830">https://www.airbnb.com/users/show/278821830</a>	9
Lu&Cia	<a href="https://www.airbnb.com/users/show/126386697">https://www.airbnb.com/users/show/126386697</a>	9
..	...	...

Al analizar esta vista, podemos identificar a los propietarios que pueden tener problemas recurrentes en sus propiedades y evitar potenciales experiencias desagradables.



## 6. Análisis de Correlación de Atributos

En el contexto del mercado de alquileres vacacionales en Málaga, hemos utilizado diversas herramientas y visualizaciones para explorar la información contenida en los datos de Airbnb. Uno de los enfoques más poderosos es la correlación de atributos, que nos permite identificar relaciones y patrones entre variables, permitiendo analizar la interacción entre distintos aspectos del mercado.

En nuestra investigación, hemos utilizado la técnica de correlación de atributos para examinar la relación entre diferentes variables presentes en el conjunto de datos. Para visualizar estas relaciones, creamos un mapa de correlación, también conocido como *heatmap*, que nos permite identificar las magnitudes y direcciones de las correlaciones entre pares de atributos.

En el mapa de correlación presentado en la figura, cada celda representa la relación entre dos atributos. Los valores oscilan entre -1 y 1, donde 1 indica una correlación positiva perfecta, -1 una correlación negativa perfecta y 0 una falta de correlación.

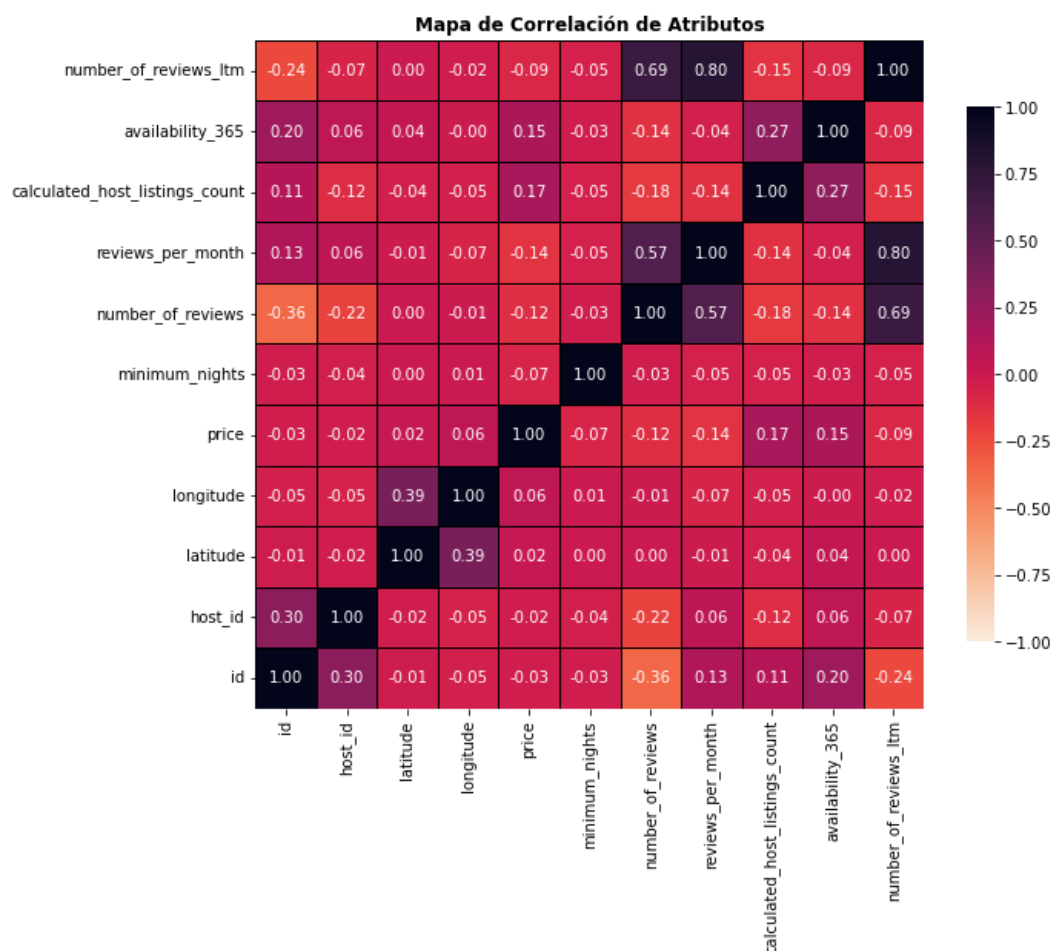


Figura 23: Mapa de correlación de atributos.

Al analizar el mapa de correlación, se destacan varias conclusiones clave. La alta correlación entre el ID del alojamiento y el `host_id` sugiere que ciertos anfitriones pueden tener múltiples propiedades en el mercado de alquileres. Además, la correlación entre `latitude` y `longitude` es coherente, ya que estas dos variables están intrínsecamente relacionadas con la ubicación geográfica de las propiedades. Es interesante observar la correlación entre el número total de reseñas (`number_of_reviews`) y la tasa de reseñas mensuales (`reviews_per_month`), lo que sugiere que las propiedades que reciben más reseñas en general también tienden a recibir más reseñas por mes.

Sin embargo, llama la atención la baja correlación entre el número total de reseñas y el ID del alojamiento, indicando que la popularidad de una propiedad medida por el número de reseñas no necesariamente está correlacionada con su posición en la secuencia de IDs.



## 7. Conclusiones

Este proyecto de análisis de datos de Airbnb en Málaga proporciona una visión detallada del mercado de alquileres en la ciudad. A través de consultas y cálculos realizados en PostgreSQL, análisis con el lenguaje Python y visualización con Tableau, se obtiene información valiosa sobre la disponibilidad, los precios, las valoraciones y otros aspectos relevantes de los alojamientos de Airbnb.

El estudio revela tendencias y patrones significativos en el mercado de alquileres, lo que puede ser útil para propietarios de alojamientos, huéspedes y profesionales del sector inmobiliario. La información obtenida puede guiar la toma de decisiones informadas en cuanto a la gestión de propiedades, fijación de precios y mejora de la calidad del servicio.

Es importante destacar que el análisis de datos es un proceso continuo y dinámico. A medida que se obtengan más datos y se realicen más análisis, se pueden descubrir nuevas ideas y oportunidades. Por lo tanto, este proyecto puede servir como punto de partida para investigaciones futuras y estudios más exhaustivos sobre el mercado de alquileres de Airbnb en Málaga.

### Dashboard en Tableau

Para una visualización detallada de los resultados obtenidos en este análisis, se ha creado un dashboard interactivo utilizando Tableau. Este dashboard proporciona una forma intuitiva de explorar los datos y comprender mejor las tendencias y patrones identificados en el informe.

Puedes acceder al dashboard en el siguiente enlace: [Dashboard de Tableau](#)

En el dashboard, podrás interactuar con gráficos interactivos, filtrar datos y obtener una visión más completa de los resultados de este análisis.



## 8. Bibliografía

### Referencias

- [1] Inside Airbnb. <http://insideairbnb.com/> Inside Airbnb, "Get the Data", Inside Airbnb. [En línea]. Disponible: <http://insideairbnb.com/get-the-data/>. (Fecha de acceso: 10 de Agosto de 2024).
- [2] Inside Airbnb - Obtener los datos. <http://insideairbnb.com/get-the-data/>
- [3] Inside Airbnb - Datos de Málaga. <http://insideairbnb.com/malaga>
- [4] Avery Smith - Data Career Jumpstart, 11 ago 2022, *Analysis of AirBnB Data Using SQL — Data Portfolio Project*, Acceso: 10 de agosto de 2023, [Vídeo en línea]. Disponible en <https://youtu.be/CHb-QvGcRjw>
- [5] AbsentData, 23 may 2021, *Data Analyst Portfolio Project #1 — Build a Database — Develop SQL — Create a Dashboard*, Acceso: 10 de agosto de 2023, [Vídeo en línea]. Disponible en <https://youtu.be/S2zBHmkRbhY>
- [6] R. Rusiano, "Housing Crisis and Tourism in Madrid", Kaggle. [En línea]. Disponible: <https://www.kaggle.com/code/rusiano/housing-crisis-and-tourism-in-madrid>. (Fecha de acceso: 10 de Agosto de 2024).
- [7] V. Viktoriaw, "Madrid Data Analysis & Price Prediction", Kaggle. [En línea]. Disponible: <https://www.kaggle.com/code/viktoriaw/madrid-data-analysis-price-prediction>. (Fecha de acceso: 10 de Agosto de 2024).
- [8] C. Teh, "Data Visualization with Bokeh", Kaggle. [En línea]. Disponible: <https://www.kaggle.com/code/cindyteht/data-visualization-with-bokeh>. (Fecha de acceso: 10 de Agosto de 2024).
- [9] Vaibhav Gupta, 31 ago 2022, *My First Ever Analytical Video — Airbnb Booking Analysis — EDA (Exploratory Data Analysis)*, Acceso: 10 de agosto de 2023, [Vídeo en línea]. Disponible en <https://youtu.be/9Iq4cM.bAGM>
- [10] Shashank Kalanithi, 8 feb 2021, *Day in the Life of a Data Analyst - SurveyMonkey Data Transformation*, Acceso: 10 de agosto de 2023, [Vídeo en línea]. Disponible en <https://youtu.be/pKvWD0f18Pc>
- [11] Gaurav Sahani, 23 feb 2022, *Tableau Dashboard for Airbnb Data — Dashboard from Scratch — Part 1*, Acceso: 10 de agosto de 2023, [Vídeo en línea]. Disponible en [https://youtu.be/bg1Gzo\\_YvAU](https://youtu.be/bg1Gzo_YvAU)
- [12] Data Science RoadMap, 10 mar 2023, *Create an Amazing Interactive Tableau Dashboard in 40 minutes — Airbnb NYC*, Acceso: 15 de agosto de 2023, [Vídeo en línea]. Disponible en [https://youtu.be/47\\_JlqPxKjY](https://youtu.be/47_JlqPxKjY)
- [13] Python Soldiers, 18 nov 2021, *I Did Analysis of Airbnb Bookings Data Using Python, Excel and Tableau*, Acceso: 17 de agosto de 2023, [Vídeo en línea]. Disponible en <https://youtu.be/Q2AFVafpRJA>
- [14] Ashi Ranka, 30 nov 2020, *Data Analysis, Visualization and Exploration on NYC AirBnb*, Acceso: 22 de agosto de 2023, [Vídeo en línea]. Disponible en <https://youtu.be/OFWHY0GJTIM>