**SQL AirBnb Problem Statement**

**Background**

AirBnB is a global vacation rental platform that connects property owners (hosts) with travellers seeking accommodations. The platform features millions of listings, from budget-friendly apartments to luxurious villas. To remain competitive, AirBnB continuously analyses its data to improve user experiences, optimise pricing, and enhance host and guest satisfaction.

The analytics team at AirBnB is tasked with delivering insights into key performance areas such as host activity, booking trends, and guest feedback. By leveraging the platform's dala, they aim to optimize revenue, improve guest satisfaction, and identity areas of growth or underperformance.

**Goal**

The goal is to extract actionable insights from AirBnB's datasets, focusing on

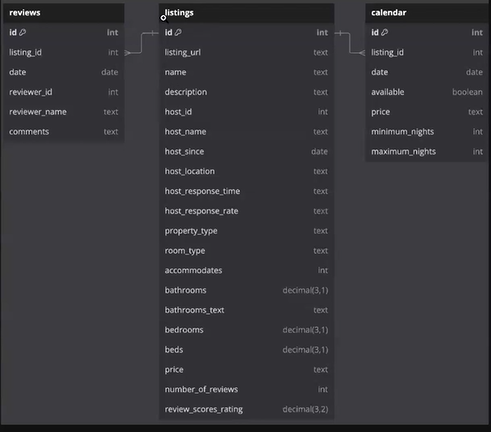
1. Booking trends and occupancy rates.

2. Host performance and listing popularity.

3. Guest feedback and property diversity.

**Dataset**

Screenshot:



**Reviews Table**

Tracks review comments on properties listed on AirBnB:

* **id:** Unique identifier for the reviews.
* **listing\_id:** Unique identifier for the listing.
* **date**: date on which the review was made by a reviewer.
* **reviewer\_id**: Unique identifier for reviewer writing comments.
* **reviewer\_name**: Name of the reviewer.
* **comments**: The review written by the reviewer on a property they have stayed at.

**Listings Table**

Tracks detailed information about properties listed on AirBnB:

* **id:** Unique identifier for the listing.
* **name:** Title of the listing.
* **description:** Detailed property description.
* **host\_id:** Identifier for the host managing the property.
* **host name:** Name of the host.
* **property type:** Type of property (e.g., house, apartment).
* **room\_type:** Type of room (e.g., entire home, private room).
* **accommodates:** Maximum number of guests.
* **review\_scores\_rating:** Average guest rating
* **number\_of\_reviews:** Total number of reviews

**Calendars Table**

Tracks review number of nights booked and price paid by guest for properties listed on AirBnB:

* **id:** Unique identifier for calendar table.
* **listing\_id:** Unique identifier for the listing.
* **date:** Date of the booking made by the guest on a property.
* **available:** Property availability for the date selected by the guest.
* **price:** Price of the listed property per night.
* **minimum\_nights:** Least number of nights required to book the listed property.
* **maximum\_nights:** At most number of nights that can be booked by the guest for a listed property.

**Question 1: Property Diversity**

With a wide range of property types available, from apartments and villas to unique stays like treehouses and boats, the management team wants to evaluate whether they are meeting customer needs.

This analysis will highlight the most common property types and uncover underserved categories

**Task**:

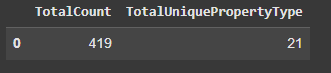
1. Calculate the total number of listings and unique property types.

2. Identify the top 5 most common property types and their respective counts.

**Answers:**

1. 

Output:

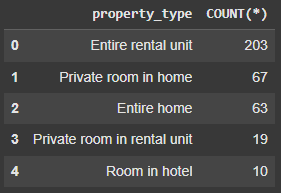


Explanation:

Count(Distinct) is used to count unique data types.

2. 

Output:



Explanation:

Group by will group every property as per property\_type

Order by clause is used to arrange columns in ascending or descending order.

**Question 2: Quest Ratings**

Guest satisfaction is critical to Airbnb's success, as high ratings lead to repeat bookings and positive word-of-mouth. Identifying top-rated listings can help Airbnb feature these properties in promotional campaigns and inspire other hosts to improve their offerings.

Conversely, understanding listings with poor ratings will allow Airbnb to address guest concerns through targeted interventions, such as host training or quality assurance checks.

**Task:**

1. Calculate the average review score across all listings.

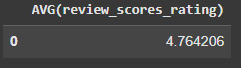
2. List the top 10 listings with the highest review scores.

3. Count the number of listings with review score rating below 4.0.

**Answers:**

1. 

Output:

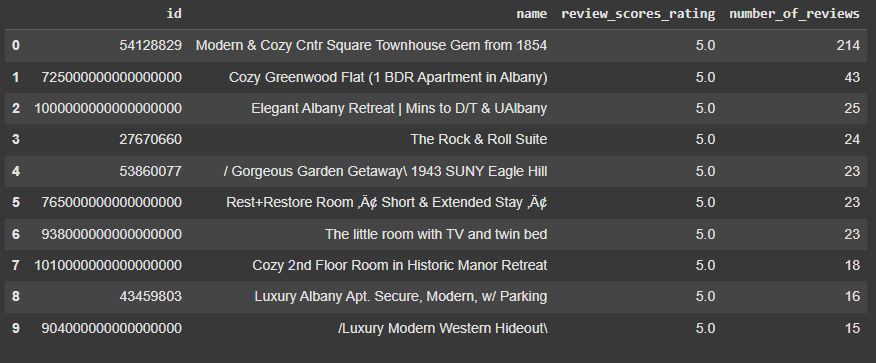


Explanation:

Average keyword is an aggregate function used to find the average value of a column.

2. 

Output:

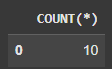


Explanation:

Limit keyword is used to display a certain number of rows as required and is positioned at the end of the query.

3. 

Output:



Explanation:

Where is the keyword used after table\_name to add condition.

**Question 3: Host Engagement**

Hosts are the backbone of Airbnb's platform, and their success directly impacts the company's growth. Hosts managing multiple listings contribute significantly to the platform's inventory and revenue. Airbnb wants to identify its most engaged hosts and understand their performance.

**Task:**

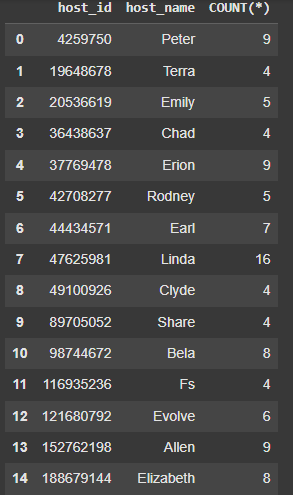
1. Identify hosts managing more than 3 listings.

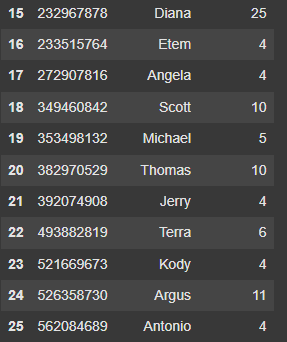
2. Calculate the average review score for each host across their listings.

**Answers:**



Output:



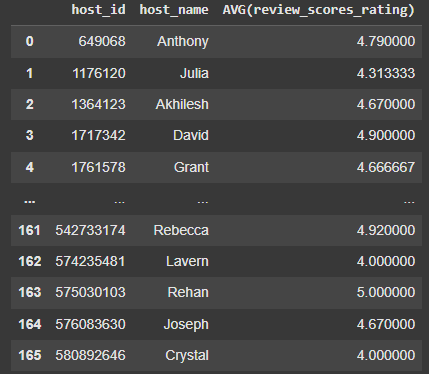


Explanation:

Having is always used after Group By to add condition. No other keyword after group by can be used else it will fetch errors.

2. 

Output:



Explanation:

“Is not null” is used to display the non-null values.

**Question 4: Pricing Patterns Across Property Types**

Airbnb's pricing strategy team wants to understand how property type influences nightly rates. Identifying undervalued property types can guide pricing adjustments, while understanding high-value categories can inform marketing priorities.

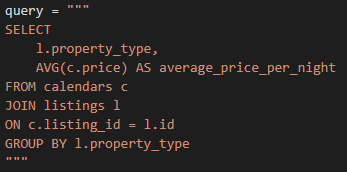
Analysing pricing patterns will also help identify trends in demand and inform strategic decisions about property offerings

**Task:**

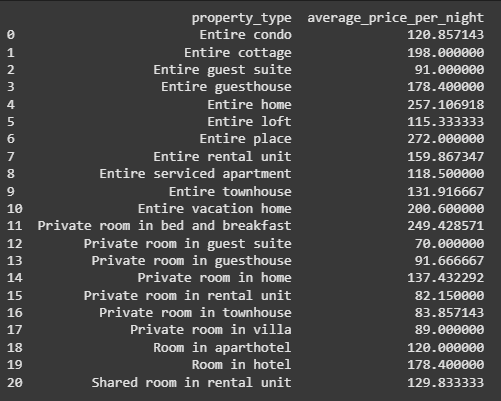
1. Calculate the average price per night for each property type.

2. Identify the top 5 listings with the highest average price per night and their property type.

**Answers:**

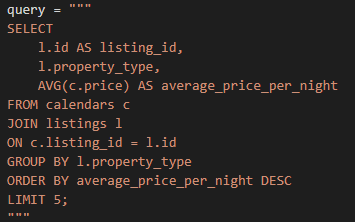
1. 

Output:

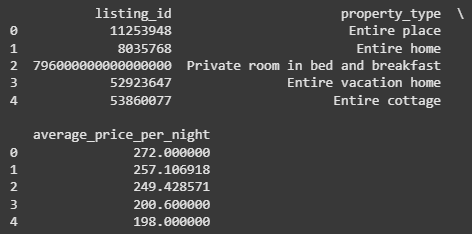


Explanation:

Join is used to join 2 tables using a primary key from first table which is a foreign key in second table.

2. 

Output:



Explanation:

A column can be renamed using the “as” keyword after the actual column name.

**Question 5: Seasonal Booking Trends**

Seasonal demand plays a critical role in the performance of Airbnb listings: Understanding how bookings vary across months can help optimize pricing strategies and promotional campaigns.

**Task:**

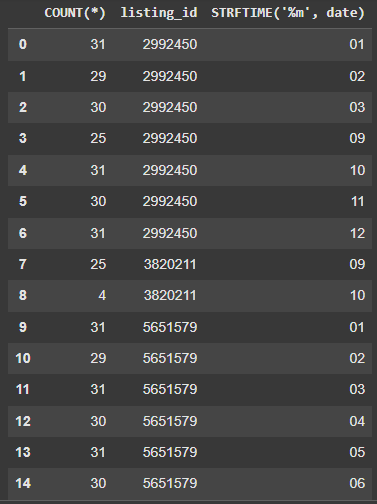
1. Calculate the total number of bookings (days when available f) for each listing, grouped by month.

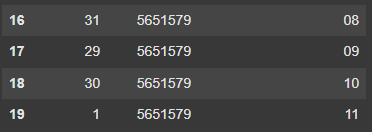
2. Identify the top 5 listings with the highest number of bookings during the peak month. (Peak Month is Nov).

**Answers:**

1. 

Output:



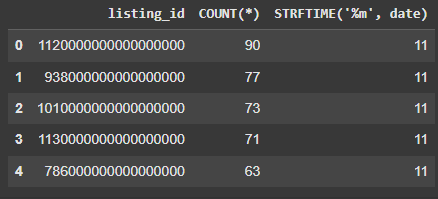


Explanation:

STRFTIME is the keyword used to fetch only date/ month/ year from a completely stated date.

2. 

Output:



Explanation:

“Like” keyword is used to fetch a certain letter in the data. It is used with % for an uncertain number of letters and \_ for a certain number of letters. Both can be used either before or after a letter as shown.