



AIRBNB LISTINGS

By Apala Mishra

TABLE OF CONTENTS

Sr. No.	Title	Slide No.
1	Project Description	3
2	Logical Data Model	4
3	Physical Data Model	5
4	Insights	6
5	Visualization Charts	11

PROJECT DESCRIPTION

Airbnb is a vacation rental company based out in North America. It operates as an online marketplace, comprising of two parties- guests and hosts, and focuses on short stays and experiences

The business model followed by them is peer to peer platform, they gain service fee from the customers who book the homestays and gain commission from the hosts. The service fee ranges from 5-15% for guests and 3% for hosts

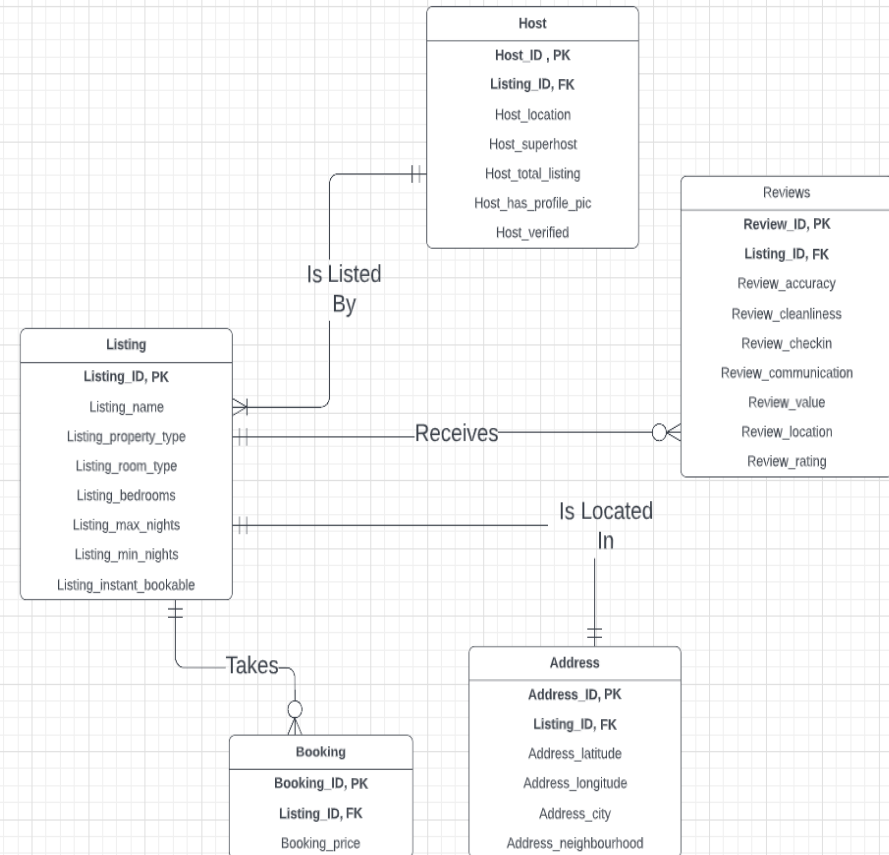
Their platform makes booking easier across the world and expands the industry making it viable to book accommodation all over the globe. The revenue Airbnb generated in 2021 is USD 5.99 Billion a 77% increase from 2020

This dataset from Airbnb sourced from Kaggle.com has 250K listings across 10 major cities and includes information on hosts, pricing of the stay, property details, and reviews on multiple criteria



LOGICAL DATA MODEL

- Our ERD model has 5 entities- the host, listing, reviews, address, and booking tables
- The booking entity includes the booking ID, listing ID, and booking price
- It is separate from the listing table as there may be multiple bookings of a single listing throughout the given time period, which can be identified through the booking ID
- All the tables in this ERD have only one primary key, which eliminates the possibility of any partial dependencies
- The address table has been separated from the listing table as well, due to the table not being in 3rd normal form. Once done, all our tables are in second normal form



PHYSICAL DATA MODEL

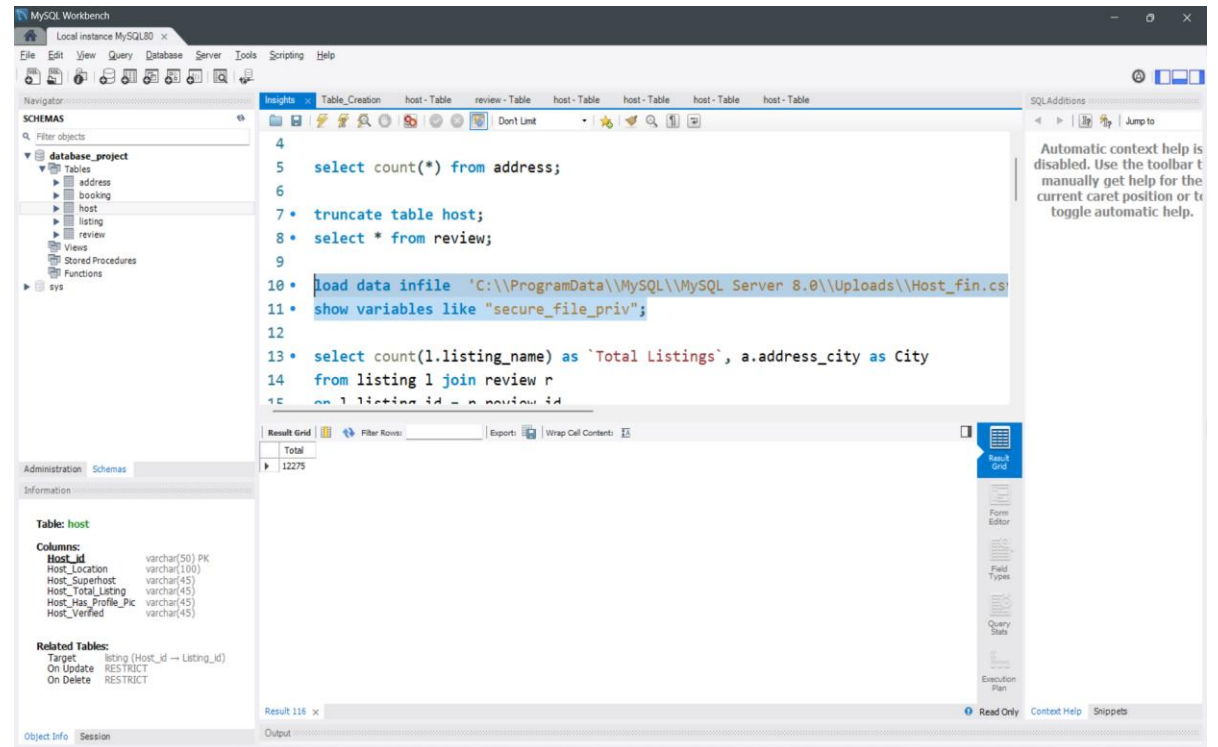


Tool used is MySQL Workbench



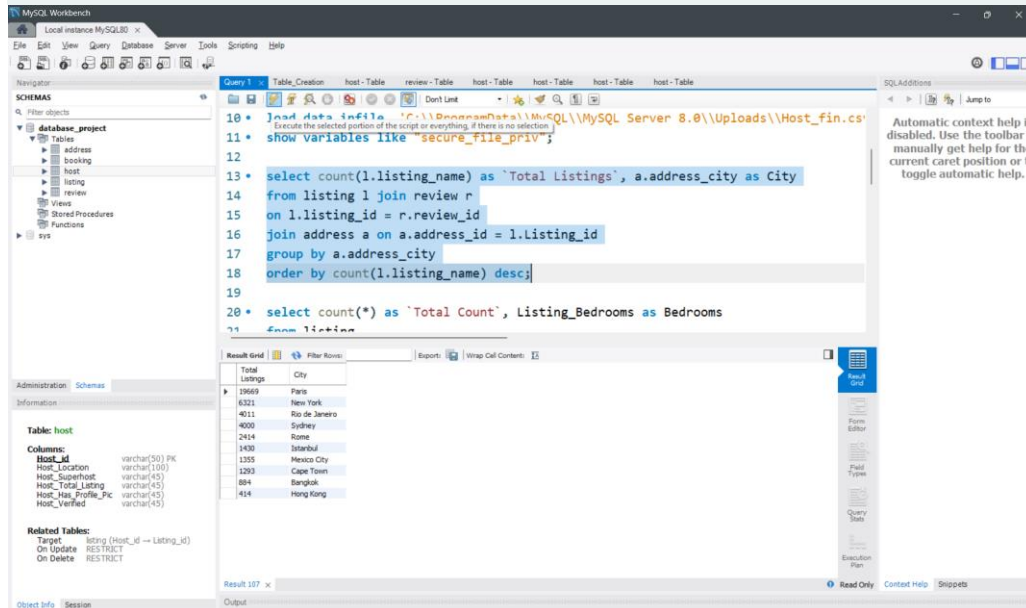
Data Loading Command:

```
load data
infile 'C:\\ProgramData\\MySQL
\\MySQL Server
8.0\\Uploads\\Host_fin.csv' into
table host fields terminated by ','
enclosed by '"' lines terminated
by '\\n' ignore 1 rows;
```



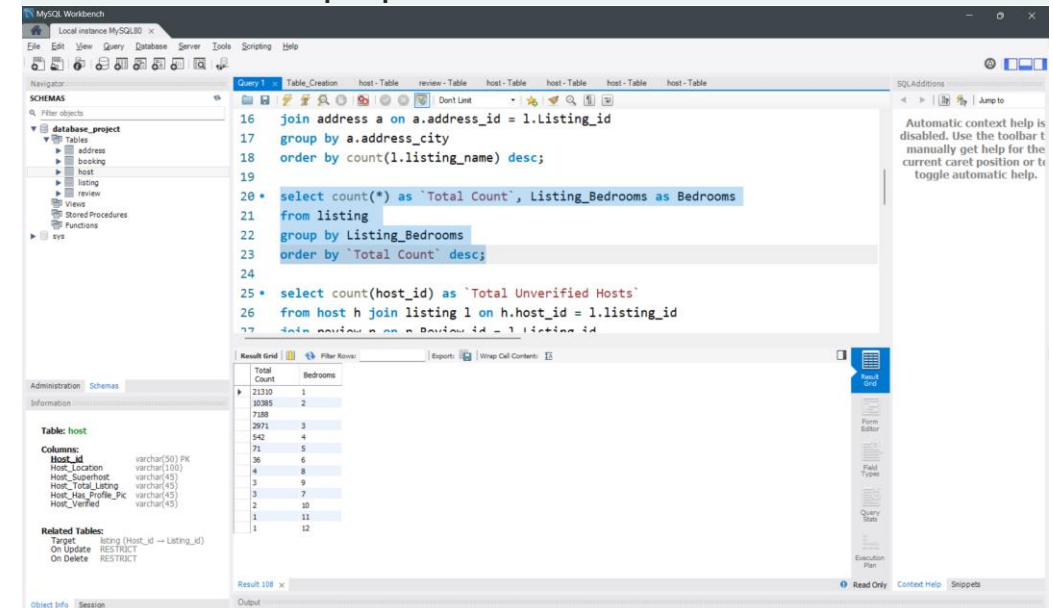
INSIGHTS

Query 1: How many listings are there in each city?



As seen in the output, the city with the highest number of listings is Paris, with more than 19k listings. This is followed by New York, Rio de Janeiro, Sydney, and more. With this insight, Airbnb can collaborate with brands or tourism destinations or airlines that frequently land in Paris, and make sure that the listings are booked throughout the year by giving collaborative deals. It can also do a lot in other cities to promote more hosts to offer their properties as listings.

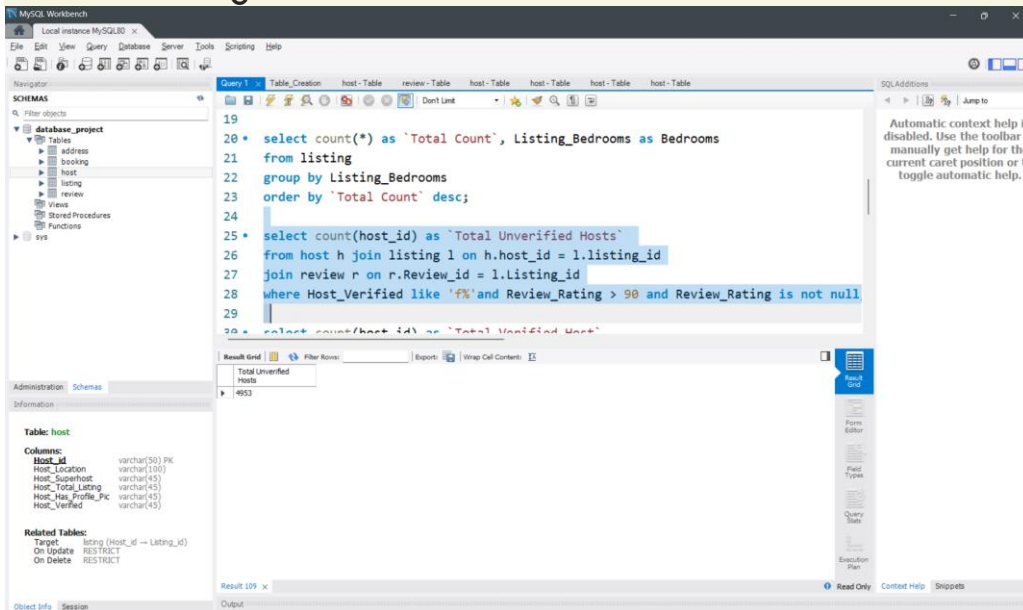
Query 2: What is the frequency table of the number of bedrooms in the properties?



As seen in the output, it is concluded that more than 21k of the properties have only 1 bedroom, followed by 10k properties having 2 bedrooms. This signifies that Airbnb has a lot of options for singles or couples travelling together, and can advertise to these parties, while also working towards expanding their options for accommodations with more rooms for larger parties.

INSIGHTS

Query 3: How many unverified hosts are there who have overall ratings more than 90%?



The screenshot shows the MySQL Workbench interface with Query 3. The SQL editor contains the following code:

```
19
20 * select count(*) as 'Total Count', Listing_Bedrooms as Bedrooms
21 from listing
22 group by Listing_Bedrooms
23 order by 'Total Count' desc;
24
25 * select count(host_id) as 'Total Unverified Hosts'
26 from host h join listing l on h.host_id = l.listing_id
27 join review r on r.Review_id = l.Listing_id
28 where Host_Verified like 'f%' and Review_Rating > 90 and Review_Rating is not null
29
30 * select count(host_id) as 'Total Verified Hosts'
```

The result grid shows a single row with the value 4953 for 'Total Unverified Hosts'.

Table: host

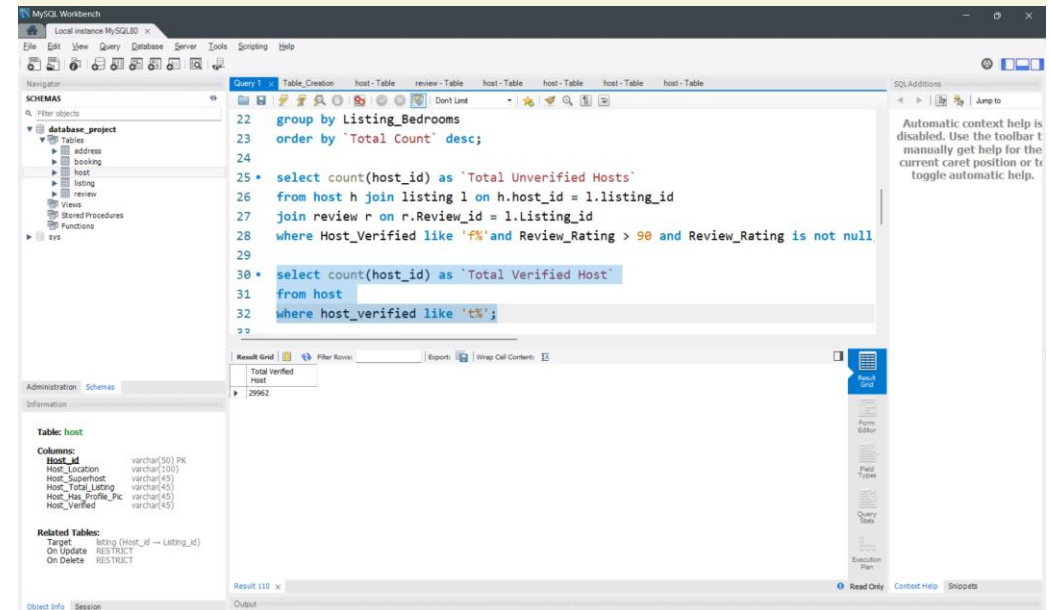
Column:	host_id	Host_Location	Host_Superhost	Host_Total_Listing	Host_Has_Profile_Pic	Host_Verified
host_id	varchar(50) PK	varchar(100)	varchar(45)	varchar(45)	varchar(45)	varchar(45)

Related Tables:

Target	listing (host_id → listing_id)
On Update	RESTRICT
On Delete	RESTRICT

There are almost 5k unverified hosts that have an overall rating of more than 90%. Airbnb should reach out to them and urge them to get verified, so their profile can be more attractive, and their bookings can increase, generating more commission for the company.

Query 4: How many verified hosts are there?



The screenshot shows the MySQL Workbench interface with Query 4. The SQL editor contains the following code:

```
22 group by Listing_Bedrooms
23 order by 'Total Count' desc;
24
25 * select count(host_id) as 'Total Unverified Hosts'
26 from host h join listing l on h.host_id = l.listing_id
27 join review r on r.Review_id = l.Listing_id
28 where Host_Verified like 'f%' and Review_Rating > 90 and Review_Rating is not null
29
30 * select count(host_id) as 'Total Verified Hosts'
31 from host
32 where host_verified like 't%';
33
```

The result grid shows a single row with the value 29962 for 'Total Verified Hosts'.

Table: host

Column:	host_id	Host_Location	Host_Superhost	Host_Total_Listing	Host_Has_Profile_Pic	Host_Verified
host_id	varchar(50) PK	varchar(100)	varchar(45)	varchar(45)	varchar(45)	varchar(45)

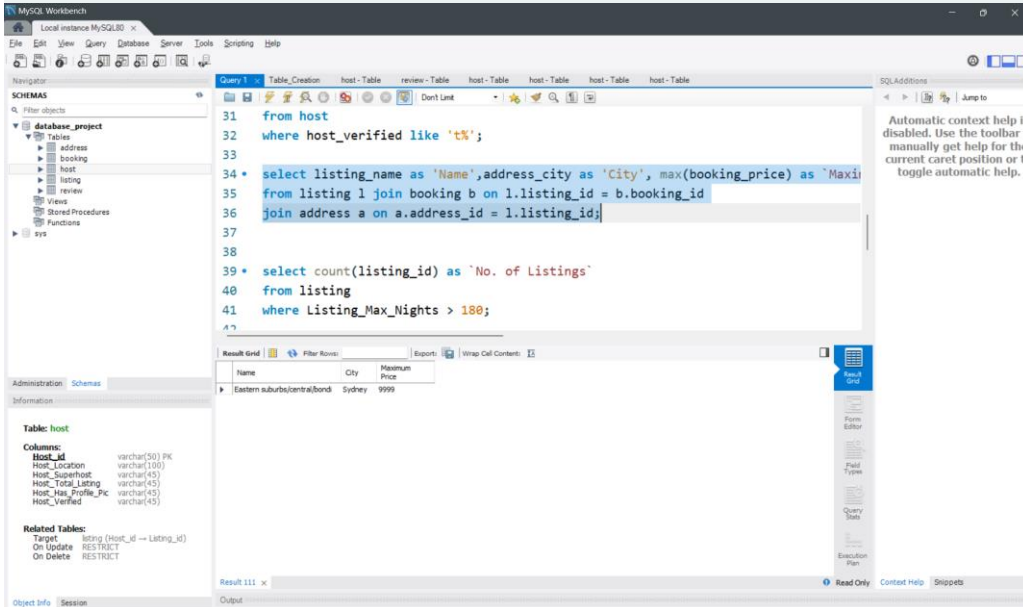
Related Tables:

Target	listing (host_id → listing_id)
On Update	RESTRICT
On Delete	RESTRICT

There are almost 30k verified hosts on Airbnb. To improve this number, Airbnb should give hosts an incentive to get verified, add a profile picture, and do other things that make their listings and profiles more appealing.

INSIGHTS

Query 5: What is the maximum booking price in the data?



The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
31 from host
32 where host_verified like 't%';
33
34 * select listing_name as 'Name', address_city as 'City', max(booking_price) as 'Maximum Price'
35 from listing l join booking b on l.listing_id = b.booking_id
36 join address a on a.address_id = l.listing_id;
37
38
39 * select count(listing_id) as 'No. of Listings'
40 from listing
41 where Listing_Max_Nights > 180;
```

The results grid shows the following data:

Name	City	Maximum Price
Eastern suburbs/central/bondi	Sydney	9999

The information panel on the left shows the schema for the 'host' table:

Table: host

Columns:

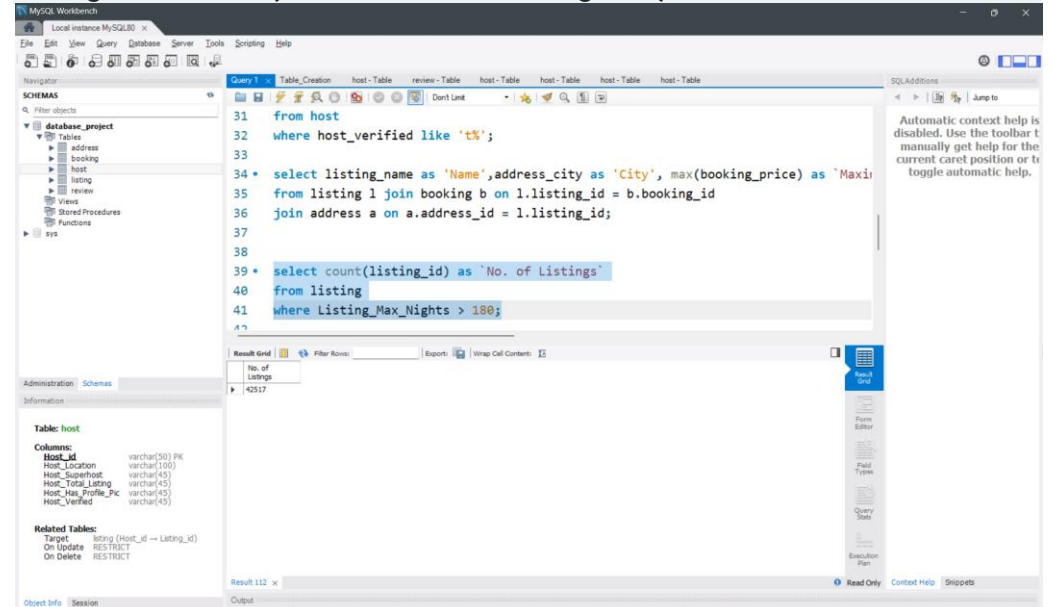
- Host_id: varchar(50) PK
- Host_Location: varchar(100)
- Host_Superhost: varchar(45)
- Host_Total_Listing: varchar(45)
- Host_Max_Profile_Pic: varchar(45)
- Host_Verified: varchar(45)

Related Tables:

- Target: listing (Host_id → Listing_id)
- On Update: RESTRICT
- On Delete: RESTRICT

The maximum booking price in the data is \$9999. This information can be used to study the location and guests of this booking price's listing and use it to increase the number of such listings and guests, respectively.

Query 6: How many listings allow guests to stay for a long duration (more than 180 nights?)



The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
31 from host
32 where host_verified like 't%';
33
34 * select listing_name as 'Name', address_city as 'City', max(booking_price) as 'Maximum Price'
35 from listing l join booking b on l.listing_id = b.booking_id
36 join address a on a.address_id = l.listing_id;
37
38
39 * select count(listing_id) as 'No. of Listings'
40 from listing
41 where Listing_Max_Nights > 180;
```

The results grid shows the following data:

No. of Listings
42517

The information panel on the left shows the schema for the 'host' table:

Table: host

Columns:

- Host_id: varchar(50) PK
- Host_Location: varchar(100)
- Host_Superhost: varchar(45)
- Host_Total_Listing: varchar(45)
- Host_Max_Profile_Pic: varchar(45)
- Host_Verified: varchar(45)

Related Tables:

- Target: listing (Host_id → Listing_id)
- On Update: RESTRICT
- On Delete: RESTRICT

Almost 43k listings allow guests to stay for longer than 6 months. Airbnb should see that hosts are open to longer stays, and hence incentive guests into booking for longer durations.

INSIGHTS

Query 7: How many listings in each city are instantly bookable?

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
37
38
39 * select count(listing_id) as 'No. of Listings'
40 from listing
41 where Listing_Max_Nights > 180;
42
43 * select count(listing_id) as 'No. of Listings', address_city as 'City'
44 from listing l join address a
45 on l.listing_id = a.address_id
46 where Listing_Instant_Bookable like 't%'
47 group by address_city;
```

The Results grid shows the following data:

No. of Listings	City
679	Istanbul
1299	Name
531	Cape Town
1461	New York
4543	Paris
1193	Sydney
1061	Rio de Janeiro
634	Mexico City
415	Bangkok
112	Hong Kong

The Table: host section shows the following columns:

- Host_id: varchar(50) PK
- Host_Location: varchar(100)
- Host_Superhost: varchar(45)
- Host_Tota_Listing: varchar(45)
- Host_Has_Profile_Pic: varchar(45)
- Host_Verified: varchar(45)

Related Tables:

- listing (Host_id → Listing_id)
- On Update: RESTRICT
- On Delete: RESTRICT

The output shows the number of listings that are instantly bookable in each city. We can see that this number is not very high for cities like Paris, New York etc. Airbnb should incentivize the hosts to provide this feature so that the booking process is quicker and more convenient.

Query 8: What is the range of the booking prices of listings in Paris?

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
45 on l.listing_id = a.address_id
46 where Listing_Instant_Bookable like 't%'
47 group by address_city;
48
49 * select min(booking_price) as 'Minimum Price', max(Booking_Price) as 'Maximum Price'
50 from booking b join listing l on b.booking_id = l.listing_id
51 join address a on l.listing_id = a.address_id
52 where address_city = 'Paris';
53
54 * select listing_Name as "Name", Listing_Property_Type as "Property Type", Listing
55 from listing
56 where Listing_Property_Type = 'Entire Apartment' and Listing_Property_Type = 'Entire Apartment';
```

The Results grid shows the following data:

Minimum Price	Maximum Price
10	999

The Table: host section shows the following columns:

- Host_id: varchar(50) PK
- Host_Location: varchar(100)
- Host_Superhost: varchar(45)
- Host_Tota_Listing: varchar(45)
- Host_Has_Profile_Pic: varchar(45)
- Host_Verified: varchar(45)

Related Tables:

- listing (Host_id → Listing_id)
- On Update: RESTRICT
- On Delete: RESTRICT

In Paris, a city with a very high number of listings, and a popular tourist destination- the booking price ranges from \$10-\$999. This shows that the city is open to all people within that range, and offers a wide variety of experiences to the guests. This should be encouraged at all popular travel destinations.

INSIGHTS

Query 9: How many listings are there where the entire place is available and has more than 3 bedrooms? How appealing is Airbnb to larger groups?

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
52 where address_city = 'Paris';
53
54 * select listing_Name as "Name", Listing_Property_Type as "Property Type", Listing
55 from listing
56 where Listing_Property_Type= 'Entire Apartment' and Listing_Bedrooms > 3
57 order by Listing_Bedrooms desc;
58
59 * select count(*) as 'Total'
60 from review
61 where Review_Accuracy <= 4;
62
```

The Results Grid shows the following data:

Name	Property Type	No. of Bedrooms
MASSIVE TRIPLEX 9 BR/6BATH W/ ELEVATOR	Entire apartment	9
BEST LOCATION Beachfront (Close to NYC)	Entire apartment	9
Amazing Large Group's Paradise 6 Bedr...	Entire apartment	9
Reagan's Apartment	Entire apartment	8
Beautiful sea view apartments close from the be...	Entire apartment	8
8 bedrooms apartment in front of metro station	Entire apartment	8
Travellers Exclusive 8 Bedrooms twin Apartme...	Entire apartment	8
Bonita casa para descansar o disfrutar festiv...	Entire apartment	7
Cosy & Spacious Duplex 4200 in Ridgewood	Entire apartment	7
Duplex Copacabana Partial Sea View	Entire apartment	7
Gorgeous Penthouse 3 floors Spanama beach	Entire apartment	6
Package 3 Apartments at PHOENIX	Entire apartment	6
Comfort, convenience and safety your stay	Entire apartment	6
Twins same building 464-466 6 bedrooms Baku...	Entire apartment	6
Roo009 - Luxury 6 bedroom penthouse in Copac...	Entire apartment	6
Apartamento per group Colosseum - Terrace	Entire apartment	6
5 star Penthouse with 1800A° ocean view Spa...	Entire apartment	6

The left sidebar shows the database schema for 'database_project' with tables: address, booking, host, listing, review, and sys. The 'listing' table is selected, showing columns: Host_id, Host_Location, Host_Superhost, Host_Total_Listing, Host_Has_Profile_Pic, and Host_Verified. The 'Related Tables' section shows 'listing' as the target table for 'Host_id' with a 'RESTRICT' relationship.

115 listings are there that are available with these features. In case there is a demand for larger properties, Airbnb should incentivize people with larger properties to become hosts, so it can attract bigger parties of people.

Query 10: How many listings have reviews with review accuracy being rated less than or equal to only 40%?

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
52 where address_city = 'Paris';
53
54 * select listing_Name as "Name", Listing_Property_Type as "Property Type", Listing
55 from listing
56 where Listing_Property_Type= 'Entire Apartment' and Listing_Bedrooms > 3
57 order by Listing_Bedrooms desc;
58
59 * select count(*) as 'Total'
60 from review
61 where Review_Accuracy <= 4;
62
```

The Results Grid shows the following data:

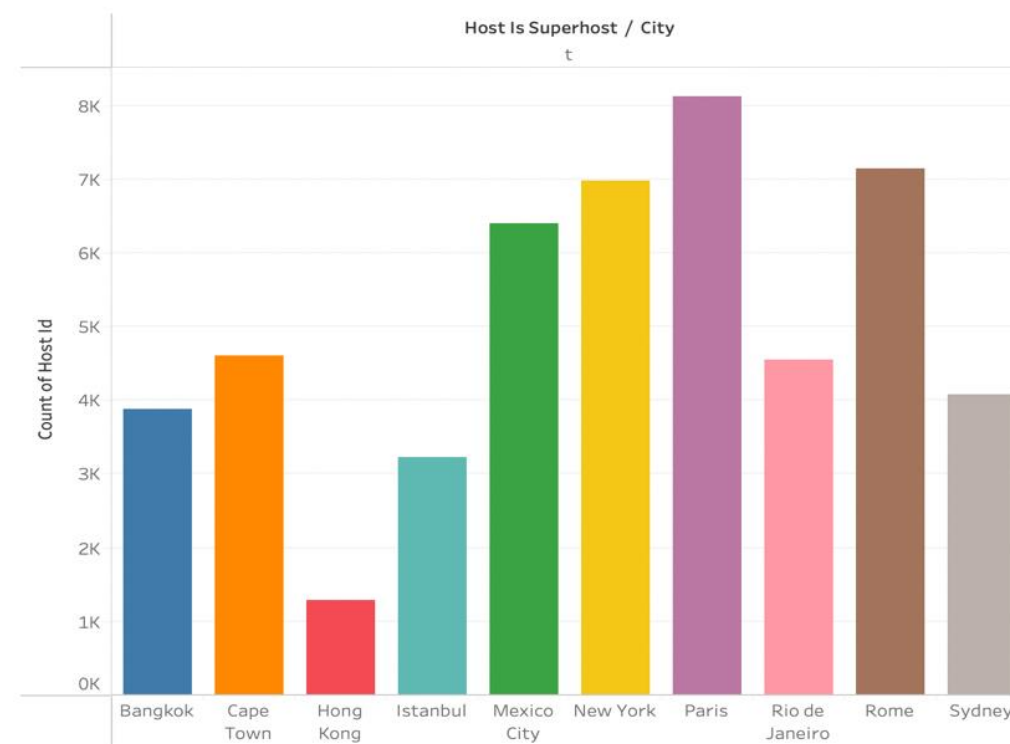
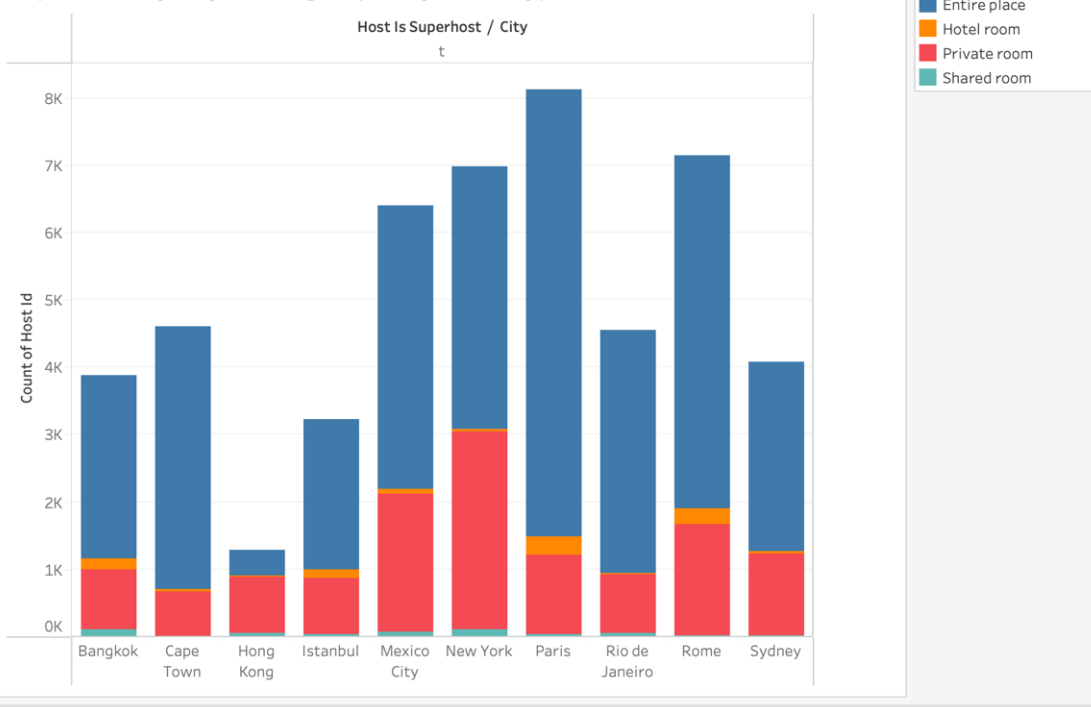
Total
12279

The left sidebar shows the database schema for 'database_project' with tables: address, booking, host, listing, review, and sys. The 'review' table is selected, showing columns: Host_id, Host_Location, Host_Superhost, Host_Total_Listing, Host_Has_Profile_Pic, and Host_Verified. The 'Related Tables' section shows 'review' as the target table for 'Host_id' with a 'RESTRICT' relationship.

There are around 12,200 listings where the review for accuracy is low. This is not a good look for Airbnb as customers will feel the descriptions are misleading. The company should look into these listings and ask the hosts to be transparent to avoid dissatisfaction, refunds, and cancellations.

VISUALIZATIONS

Super Host by City name, grouped by Room Types



A photograph of a bright, airy dining room. In the center, a wooden table is covered with a light blue patterned tablecloth and set with white dishes, including a large bowl and several smaller plates. To the right, a large, leafy green plant sits on a wooden surface. In the background, two open windows reveal a lush garden with green trees and white flowers. A small wooden table with more plants is positioned between the main table and the windows. Two pendant lights hang from the ceiling: a white one over the table and a black one over the plant. The overall atmosphere is warm and inviting.

THANK YOU