



AIDM7360

Group Project Report

Semester 1, AY 2023/2024

23422041 | Zhu Xiaoyu

23435968 | Luo Fan

23467215 | Luo Jinjin

Topic & motivation

Hong Kong, as a world-renowned tourist destination, has naturally become the first choice for many tourists. To better serve this wave of the tourism boom, we used SQL database technology to create a user-friendly query system to analyse Airbnb-related data in Hong Kong. This system is designed to help users efficiently filter out the most popular room types and the neighbourhoods in which they are located while also exploring the correlation between specific room types and reviews, the number of days booked, and the number of reviews. This system will significantly facilitate the editor's conveniently generating reports, helping readers make accommodation choices faster and more accurately.

Analysis of the problem

Using automated content management systems to process and analyse Airbnb data through in-depth analysis of Airbnb Hong Kong listings, user reviews, host information and location data, we can have a clearer understanding of user preferences, price trends, service quality and other indicators. And make some templates to store this information in a database that will be used in standard reports.

The Solutions

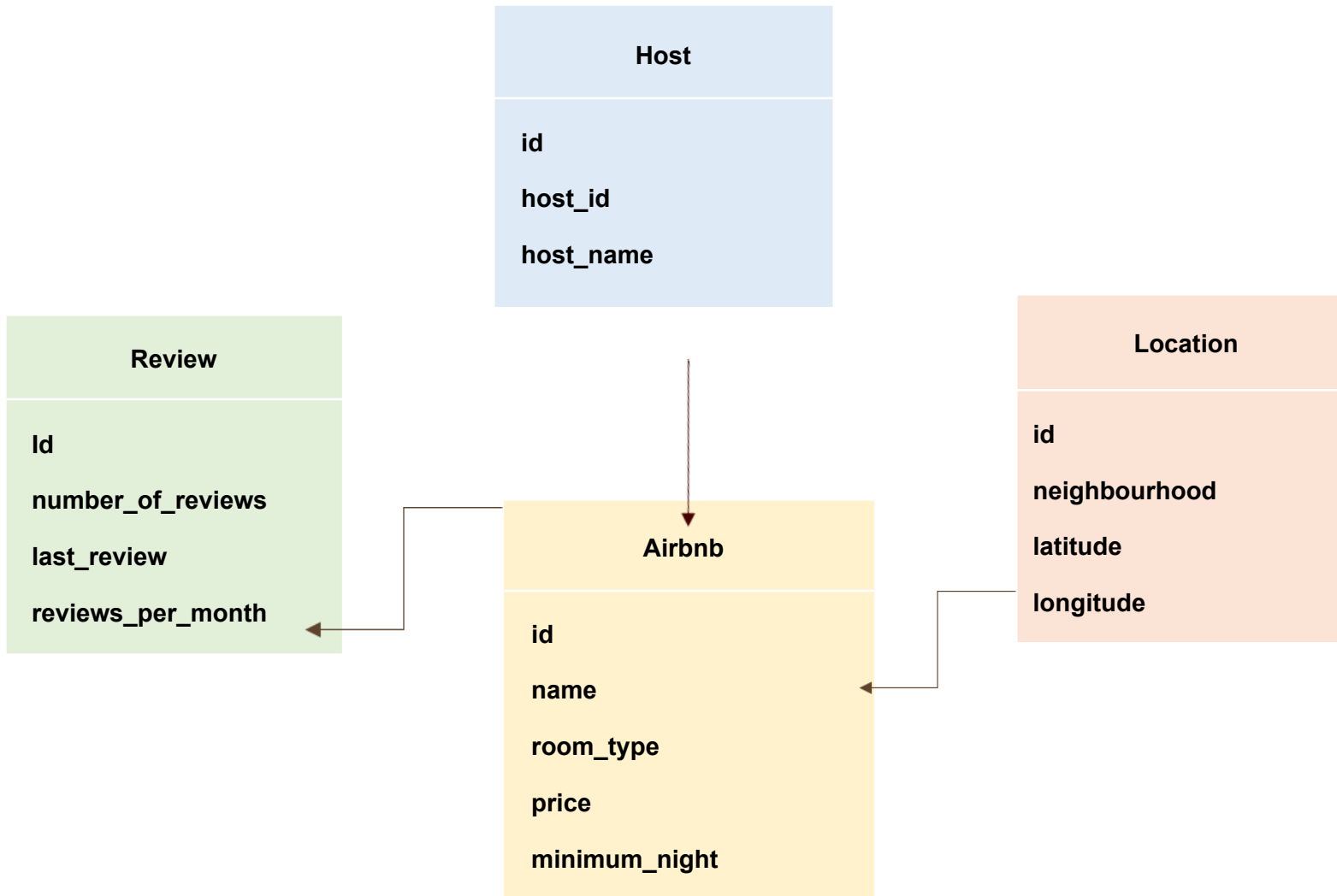
Data analysis & data cleaning

This data contains empty columns. For this data, the name of the Airbnb included the room descriptions, so the same hotel may be displayed multiple times, which is not good for our project.

id	name	host_id	host_name	neighbourhood_group	neighbourhood
17891	Rental unit in Hong Kong Island · ★4.76 · Studio · 1 bed · 1 bath	69063	Candace		Central & Western
72571	Rental unit in Sheung Wan · ★4.22 · Studio · 1 bed · 1 bath	304876	Brendan		Central & Western
103760	Rental unit in Central · ★4.46 · 2 bedrooms · 3 beds · 1 bath	304876	Brendan		Central & Western
104626	Rental unit in Hong Kong · ★4.38 · 1 bedroom · 1 bed · 1.5 baths	544166	Celine		Central & Western
132773	Rental unit in Hong Kong Island · ★4.50 · 2 bedrooms · 3 beds · 1 bath	304876	Brendan		Central & Western
163664	Rental unit in Sheung Wan · ★4.28 · 2 bedrooms · 3 beds · 1 bath	304876	Brendan		Central & Western
163742	Rental unit in Sheung Wan · ★4.33 · 2 bedrooms · 3 beds · 1 bath	304876	Brendan		Central & Western
248140	Rental unit in Sheung Wan · ★4.88 · 1 bedroom · 1 bed · 1 bath	1300549	Darren		Central & Western

The cleaning plan is first to remove the empty columns. Then remove the duplicate hotel (Using the latitude and longitude because they have unique IDs but have the same location). After that, clean the name column. Lastly, remove the stars and room descriptions, then store the new name in the “name” column.

Relational database design (4 tables)



Approach & methods

Template

Choose the most popular listing

Tid 1: Find the most popular room on Airbnb. Users choose option 2, the system will have an output of a specific room name and the total number of reviews this room has received.

Based on the survey, it can be found that the most popular Airbnb listing in Hong Kong in 2023 was the {name}, which received the most {number_of_reviews} items of reviews, and you can search it by id: {id}.

Select room type and price (by id)

Tid 2: Select the price of a specific room type. Users choose option 9, the system will have an output of a price and room type.

What is concerned the most is that the price of this listing is HK\$ {price}, with a {room_type} room type.

Select host name and neighbourhood (by id)

Tid 3: Find the landlord's name under the listing. Users choose option 10, the system will have an output of a specific room name and the total number of reviews this room has received.

This Airbnb listing is owned by {host_name}, which is located in community {neighbourhood}.

Select minimum night booking (by room type)

Tid 4: Find minimum night booking requirements for a room type. Users choose option 5, the system will have an output of the number of nights booked.

For this room type, its average minimum night that the customers need to book is {avg(minimum_night)} nights.

Select the listing host (by host name)

Tid5: Find the specific owners and prices of the most popular homes. Users choose option 6 and the system will have an output of information about it.

For this popular host, you can know that the listing he/she owns is {name}, with a {room_type} type and holds on HK\${price}.

List the most expensive room name

Tid 6: Find the most expensive room name. Users choose option 3, the system will have an output of a neighbourhood and price.

The Most expensive room name is {name}, the district is nearby {neighbourhood}, the price is HK\${price}.

List recently reviewed room names and the number of reviews per month

Tid 7: Find the names of rooms that have recently received reviews and the average number of reviews per month. Users choose option 8, the system will have an output of name and customer comments per month.

The listing {name} has just received the reviews, and it got the average {reviews_per_month} item of customer comments per month.

Results & findings

Sample Report

Based on the survey, it can be found that the most popular Airbnb listing in Hong Kong in 2023 was the Rental unit in Hong Kong, which received 501 items of reviews, and you can search it by id:850531. When you enter the id: 17891, what is concerned the most is that the price of this listing is 1400, with an Entire home/apt room type. Besides, When you want to book a private room, you can know in advance that you need to book this room type for at least 21 nights.

Sample Output



The most popular listing

Based on the survey, it can be found that the most popular Airbnb listing in Hong Kong in 2023 was the Rental unit in Hong Kong, which received the most 501 items of reviews, and you can search it by id:850531.



The most expensive listing

The Most expensive room name is Rental unit in Hong Kong, the district is nearby Central & Western, the price is HK\$150000.The listing Rental unit in Hong Kong has just received the reviews, and it got the average 0.45 item of customer comments per month.





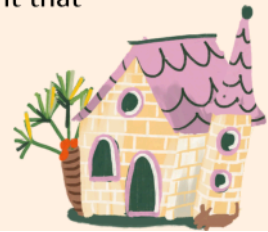
Introduction to certain accommodations

The listing Rental unit in Hong Kong Island has just received the reviews, and it got the average 1.25 item of customer comments per month. This airbnb listing is owned by Candace, which is located in community Central & Western. What is concerned the most is that the price of this listing is HK\$ 1400, with a Entire home/apt room type.



Introduction of certain host and listing of her/him

For this popular host, you can know that the listing he/she owned is Rental unit in Hong Kong, with a Entire home/apt type and hold on HK\$ 2336. For this room type, its average minimum night that the customers need to book is 21.0 nights.



Workload:

Members	Data cleaning	Tables	SQL query	Template	Report	Slide
ZHU Xiaoyu		√	√		√	√
LUO Fan			√	√	√	√
LUO Jinjin	√	√	√		√	√

