



# **airbnb**

**Automated Content  
Management System**



**Group Name: coconut**

Zhu Xiaoyu 23422041

Luo Jinjin 23467215

Luo Fan 23435968



**1**

## **Topic Introduction**

- Motivation Analysis
- Dataset
- Key Variables



**2**

## **Database Design**

- Table Creation
- Table Relationship
- Query



**3**

## **Template**

- Format
- Sample Output
- Generate content



# ***Contents***



# Topic Introduction



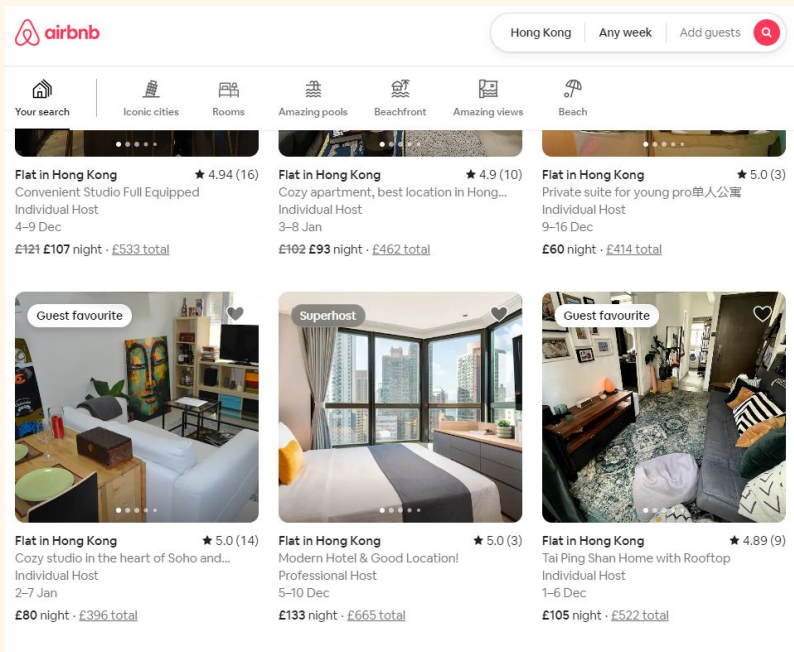
# About Project



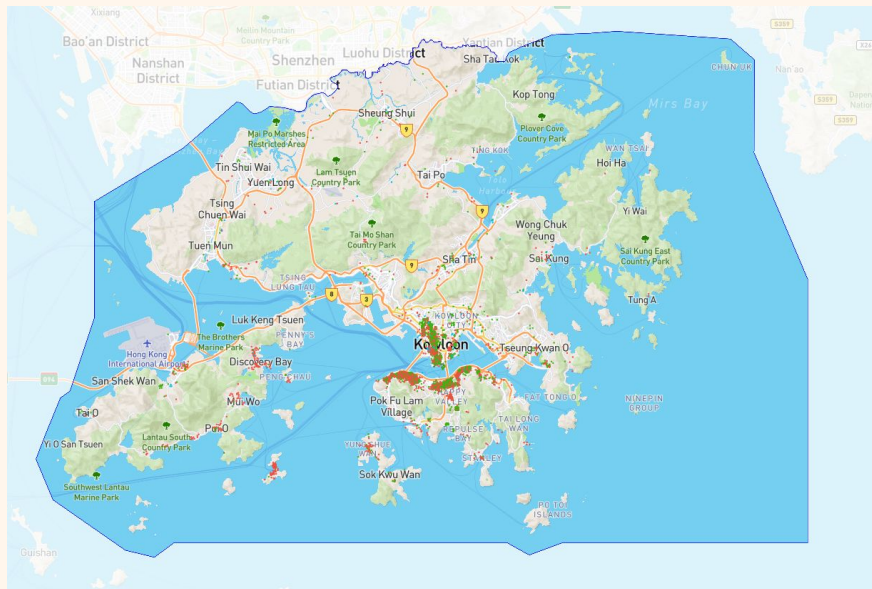
## Why this topic?

In recent years, Airbnb has revolutionized the way people explore new places, offering a more personal and immersive experience compared to traditional lodging options.

To capitalize on this, We used SQL database technology to develop a user-friendly query system specifically designed for analyzing Hong Kong Airbnb data. More than just a tool, the system helps users easily find the most popular room types and their locations, and explore the complex relationships between room types, guest reviews, booking length and review frequency.



# Motivation



## Analysis Problem

We hope to **address the following questions** by further processing and analysing the data:

1. Get the overall housing price trend and average price in HK;
2. Discover the most popular room types and the community they belong to;
3. Discover factors such as room type, minimum booking days, and reviews price impact;
4. Compare whether there are significant differences in flow between different regional locations;

# About The Dataset



The screenshot shows the 'Inside Airbnb' website interface. At the top, there's a navigation bar with links: 'Data', 'About', 'Support', 'Organise', and 'Donate'. Below this, the page is titled 'Hong Kong, Hong Kong, China' with a sub-header '17 September, 2023 (Explore)'. A table lists various data files available for download, including listings, calendar, reviews, and neighbourhoods data. The table has three columns: 'Country/City', 'File Name', and 'Description'. Below the table, there's a link to 'show archived data'.

Country/City	File Name	Description
Hong Kong	<a href="#">listings.csv.gz</a>	Detailed Listings data
Hong Kong	<a href="#">calendar.csv.gz</a>	Detailed Calendar Data
Hong Kong	<a href="#">reviews.csv.gz</a>	Detailed Review Data
Hong Kong	<a href="#">listings.csv</a>	Summary information and metrics for listings in Hong Kong (good for visualisations).
Hong Kong	<a href="#">reviews.csv</a>	Summary Review data and Listing ID (to facilitate time based analytics and visualisations linked to a listing).
Hong Kong	<a href="#">neighbourhoods.csv</a>	Neighbourhood list for geo filter. Sourced from city or open source GIS files.
Hong Kong	<a href="#">neighbourhoods.geojson</a>	GeoJSON file of neighbourhoods of the city.

The selected data set describes Airbnb's order activities and data indicators in HK in 2023

## *From Inside Airbnb*

Inside Airbnb is a mission driven project that provides data and advocacy about Airbnb's impact on residential communities.

Data Source: <http://insideairbnb.com/hong-kong>

# Key Variables



## Room Type

Airbnb hosts can list **entire homes/apartments, private, shared rooms**, and more recently **hotel room**

## Listing per Host

Some Airbnb hosts have multiple listings.

A host may list **separate rooms in the same apartment**, or multiple apartments or homes available in their entirety.

## Minimum Nights

By looking at the "minimum nights" setting for listings, we can see if the market has shifted to **longer-term stays.**

2

# Database Design





# Data cleaning



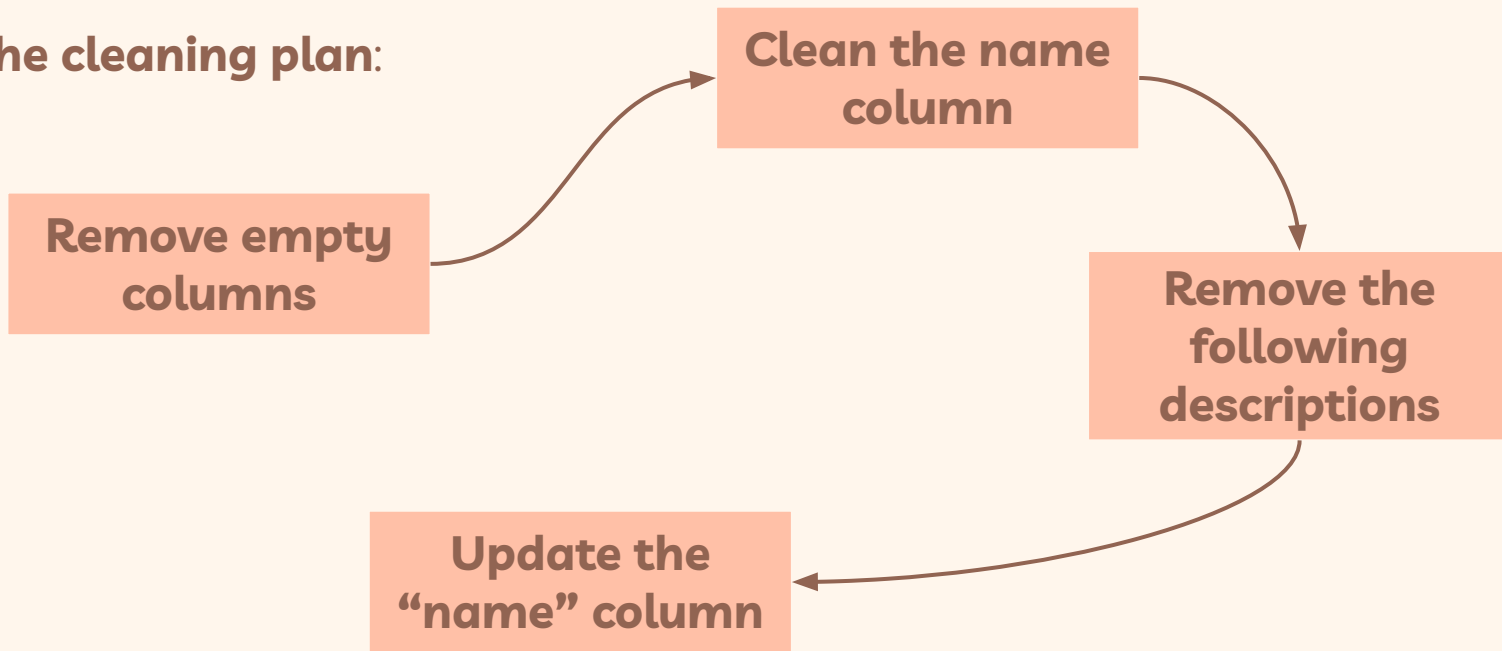
This data contains **empty columns**. For this data, the name of the Airbnb included the room descriptions, so the **same hotel may be displayed multiple times**, which is not good for our project.

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

# ***Data cleaning***



The cleaning plan:



# Table Creation



## Airbnb

id

name

Room type

price

Minimum night

## Host

id

Host id

Host name

## Location

id

neighbourhood

latitude

longitude

## Reviewed

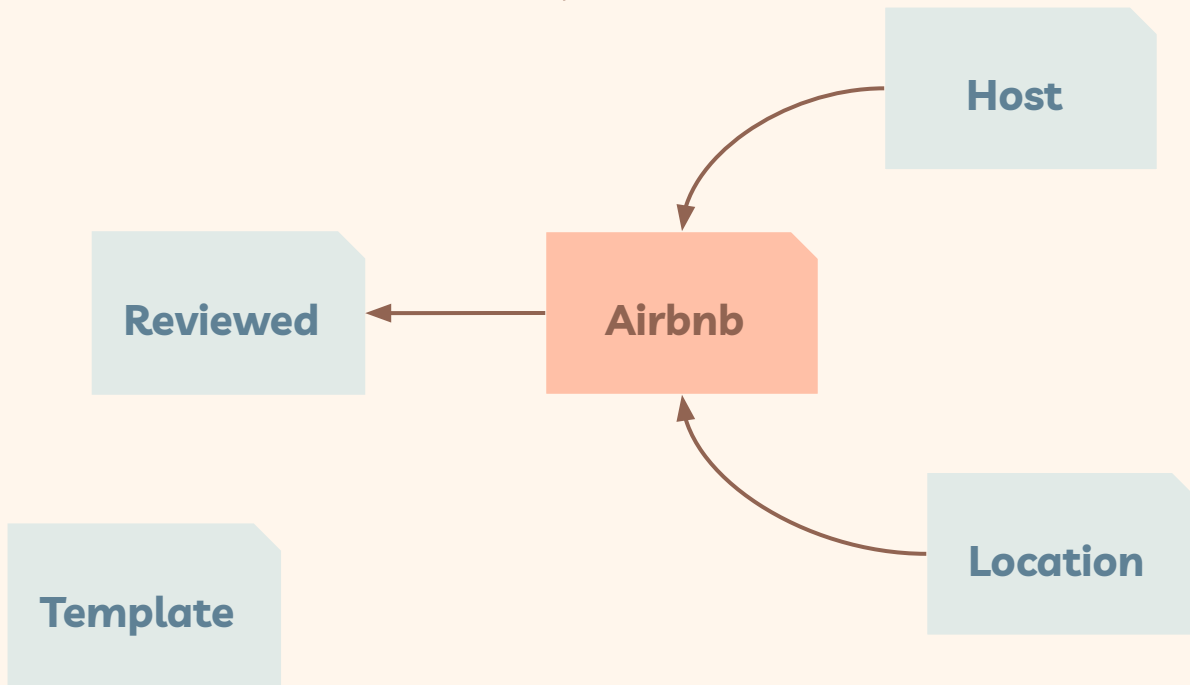
id

number\_of\_reviews

last\_review

reviews\_per\_month

# Table Relationship



# SQL Query



```
def printRoomInfo(id, conn):  
    """  
    Print out the infomation of a room given a id  
    """  
    with conn:  
        sql = "SELECT * FROM AIRBNB WHERE ID = ?"  
        cur = conn.cursor()  
        cur.execute(sql, [id])  
        result = cur.fetchone()  
        if result is None:  
            print("No such rooms")  
        else:  
            headers = ["id", "name", "room_type", "minimum_night", "price"]  
            for i in range(0, 5):  
                print(headers[i], ":", result[i])
```

# SQL Query



```
def listMostPopularRoomType(conn):  
  
    with conn:  
        sql = """  
            SELECT A.room_type, SUM(R.number_of_reviews) as total_reviews  
            FROM Airbnb A  
            JOIN Review R ON A.id = R.id  
            GROUP BY A.room_type  
            ORDER BY total_reviews DESC  
            LIMIT 1 """  
        cur = conn.cursor()  
        cur.execute(sql)  
        result_set = cur.fetchall()  
        if result_set:  
            for result in result_set:  
                room_type = result[0]  
                number_of_reviews = result[1]  
                print("The most popular room type is", room_type, "with total", number_of_reviews,"reviews")  
        else:  
            print ("NO such room type")
```

# SQL Query



```
def deleteRoom(conn):  
    """  
    Delete a flight from database by id  
    """  
    listALLRooms(conn)  
    line = input("Please input the id to delete: ")  
    line = line.strip()  
    print(line)  
    with conn:  
        sql = "DELETE * FROM AIRBNB WHERE ID = ?"  
        cur = conn.cursor()  
        cur.execute(sql, [line])  
        conn.commit()  
        print("succeed to delete room ", line)
```

3

***Template***





# Template content



## Examples:



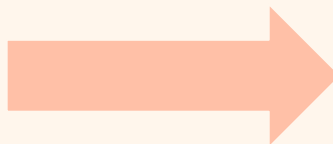
- |   |   |
|---|---|
| 4 | For this room type,its average minimum night that the customers need to book is {avg(minimum_night)}. |
| 5 | You can also take a look at the listings that sitting in the same price range like {name} and {name}. |
| 6 | The Most expensive room name is {name}, the district is nearby {neighbourhood}, the price is {price}. |

## content topics:

1. Introduction to certain accommodations
- 2.The most popular listing
- 3.The most expensive listing
- 4.Introduction of certain host and listing of her/him



**By executing different queries,get the template with metadata**



**Organize different templates to generate reports of certain topic**

# Sample output

## Airbnb

Listing Introduction and  
recommendation

### 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.

### 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.



## 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.



***Thanks for  
Watching***

