Development of a Data Mart – Airbnb Clone

Abstract:

Database is a great way of storing and organizing data in an efficient manner. The aim of this project is to create a sample database about customers and services of a rental company for holiday destination, particularly Airbnb. In this submission, I used MySQL Workbench to draw ER diagram and establish relations among entities and their cardinalities. Meanwhile, I had to chance to explore the user interface of the application which I liked and found useful for my further studies.

The approach used to design the DB is Snowflake schema. It is a bottom-up model where fact tables, dimension tables, and sub-dimension tables are contained. For the server side, I again used MySQL server to further test the database from my local pc.

I am planning to build a data mart that can be used to browse data about guests, hosts, reservation details, listing details, reviews of both hosts and guests, deals, dispute management etc.

Since I used airbnb to rent rooms before, I already know about how the system works. Still, to start off, I wanted to make sure that I do not miss a detail. Thus, I checked airbnb's website to see if there's any improvement. After surfing airbnb website for important details, I designed an entity relationship model that will be used as a background for designing this database. Entity relationship diagrams are used to display data schemas in a database. ER diagrams contains keys, relationships, entities, and attributes. In this project, I preferred crow's foot notation. Its easily understood style makes the diagram grasped at a glance.

The tables used in the ER diagram can be described below as following :

1- Reservations table will serve as the main table . This table will contain data about address of the room, details of the guest, details of listing which contains further information about the host, amenities, properties and rules of that room. Naturally, this table will contain data about each reservation's check-in and check-out dates.
2- Listing table will act as the second most important table. It contains information of that particular room together with its host details, price and other important details.
3- Address table contains data about the address of each room which will can be further listed by their respective hosts.
4- Hosts table is one another important table as it contains data about the owner of room, how many times its listings have been booked and its average review score.
5- Guests table acts similar to hosts table as it contains similar information of guests as in hosts table.

6- Staff table is like hosts and guests table, but only keeps the data of airbnb's employees.
7- Deals table is a specific table which only contains data about special deals created by any particular staff in airbnb organization by cities, along with discount ratio and validity duration.
8- Dependents table is like a sub-dimension table which keeps the data of guests' dependents, i.e spouse, kids etc.
9- Properties table is like a sub-dimension table which keeps the detailed data of listing's properties such as area, pool, gym, number of rooms etc.
10- Amenities table is like a sub-dimension table which keeps the data of listing's amenities.
11- Rules table is like a sub-dimension table which keeps the further data of listing's special rules, i.e pets allowed or not.
12- Payment table contains data as to which guest paid to which listing with preferred payment method and time.
13- Tickets table keeps the data of disputes or complaints between any host and guest for a particular stay.
14- Photos table keeps the urls of any listing that is to be uploaded during the creation of a listing by a host.
15- Reviews_hosts table keeps the data of comments and starts rated by guests for any particular listing.
16- Reviews_guests table keeps the data of comments and stars rated by hosts for any particular stay.
17- Messaging table keeps the messages that are sent and received by guests and hosts prior to reservation and/or after check-out.
18- City table is designed just to keep data of any address created by hosts.
19- Country table is designed just the keep data countries which is related to cities.
20- User details table is a separate table which keeps the data of users (guest, host or staff) other miscellaneous data.

Business Flow and Requirement Analysis

Hosts and guests register to system seperately.

They provide all of their information to system as most of the fields are mandotary and cannot be left blank. They also need to choose the type of user during registration.

Eeach address has to have certain details like neighboorhood, street, postcode , city and country.

Hosts can create listings by using these addresses.

Every listing has to have a certain address, name, description, city, country along with its properties, rules and amenities.

Every listing has to have a price and description which their host can enter.

Then the guest will enter queries to search bar in the website to list these listings from a specific area.

They will be able to see all the specs, properties and photos of the room which are previously entered by hosts.

Guests have the option to send messages to hosts which they can see in the same listing page.

If they reach to an agreement, their data will be held in reservation table which is also linked to other tables. The booking number (reservation id) will also be held in the same table along with check-in and check-out dates.

During reservation, guests enter their dependents, their check-in and check-out date, their payment method.

Hosts and/or guests can issue tickets in case they have a dispute which needs a resolution. Staff will have the option to check the ticket as solved.

Deals can be created by staff for a specific area with a discount rate.

Tables and Descriptions

I used TINYINT instead of BOOLEAN, respective data fields will be stored as 1 or 0 where necessary.

ADDRESS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
host_id	INT	References to primary key of hosts table, ID
street	VARCHAR(20)	street of property
neighborhood	VARCHAR(20)	neighborhood of property
number	INT	number of property, i.e door number
postcode	VARCHAR(10)	postcode of property
city_id	INT	FK - References to primary key of city table, ID
country_id	INT	FK - References to primary key of country table, ID

AMENITIES TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
kitchen	TINYINT	property has kitchen or not, boolean
ac	TINYINT	property has AC or not, boolean
heater	TINYINT	property has heater or not, boolean
washer	TINYINT	property has washer or not, boolean
dryer	TINYINT	property has dryer or not, boolean
internet	TINYINT	property has internet or not, boolean
tv	TINYINT	property has tv or not, boolean
smart_home	TINYINT	property has smart_home or not, boolean

CITY TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
name	VARCHAR(45)	name of the city
state	VARCHAR(45)	name of the state
country_id	INT	FK - References to primary key of country table, ID

COUNTRY TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
name	VARCHAR(45)	name of the country
continent	VARCHAR(45)	name of the continent

DEALS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
staff_id	INT	FK - References to primary key of staff table, ID
city_id	INT	FK - References to primary key of city table, ID
percent_off	INT	amount of discount in percentage
valid_from	DATE	deal start date
valid_until	DATE	deal end date
created_at	TIMESTAMP	deal creation date

DEPENDENTS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
guest_id	INT	FK - References to primary key of guest table, ID
spouse	TINYINT	guest has spouse or not, boolean
children	INT	number of children that guest has
infants	INT	number of infants that guest has
pets	TINYINT	guest has pets or not, boolean

GUESTS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
user_type	VARCHAR(1)	user type : guest, host or staff. Can be G,H or S
first_name	VARCHAR(35)	first name of the guest
last_name	VARCHAR(35)	surname of the guest
username	VARCHAR(20)	username taken during registration
password	VARCHAR(25)	password taken during registration
e-mail	VARCHAR(50)	e-mail taken during registration
gender	TINYINT	gender of the guest, 1 for male, 0 for female
birthdate	DATE	birthdate of the guest
job	VARCHAR(25)	job of the guest
mobile_phone	VARCHAR(20)	mobile phone of the guest
booking_count	INT	number of bookings guest has done
review_score	INT	guest's average review score given by the hosts

HOSTS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
user_type	VARCHAR(1)	user type : guest, host or staff. Can be G,H or S
first_name	VARCHAR(35)	first name of the host
last_name	VARCHAR(35)	surname of the host
username	VARCHAR(20)	username taken during registration
password	VARCHAR(25)	password taken during registration
e-mail	VARCHAR(50)	e-mail taken during registration
gender	TINYINT	gender of the host, 1 for male, 0 for female
birthdate	DATE	birthdate of the host
job	VARCHAR(25)	job of the host
mobile_phone	VARCHAR(20)	mobile phone of the host
booking_count	INT	number of bookings host have rented out
review_score	INT	host's average review score given by the guests

LISTING TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
caption	VARCHAR(60)	title for the ticket to appear on top
desc	TEXT	detailed description of the listing
price	INT	price / night of stay
address_id	INT	FK - References to primary key of adress table, ID
host_id	INT	FK - References to primary key of host table, ID
amenities_id	INT	FK - References to primary key of amenities table, ID
properties_id	INT	FK - References to primary key of properties table, ID
rules_id	INT	FK - References to primary key of rules table, ID
created_at	TIMESTAMP	displays the time and date listing has been created

MESSAGING TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key of a particular message , Artificial ID
sender_id	INT	Message sender's id
sender_type	VARCHAR(1)	user type : guest, host or staff. Can be G,H or S
receiver_id	INT	Message receiver's id
receiver_type	INT	user type : guest, host or staff. Can be G,H or S
message	TEXT	message content

PAYMENT TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
reservation_id	INT	FK - References to primary key of reservations table, ID
listing_id	INT	FK - References to primary key of listing table, ID
guest_id	INT	FK - References to primary key of guest table, ID
paid_at	TIMESTAMP	displays the time and date payment has been made at
payment_method	VARCHAR(25)	displays the payment method

PHOTOS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
listing_id	INT	FK - References to primary key of listing table, ID
photo_url	VARCHAR(100)	Keeps the link of the photo for the listing
created_at	TIMESTAMP	displays the time and date photo has been uploaded at

PROPERTIES TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
sqft	INT	Area of the property
rooms	INT	number of rooms property has
security	TINYINT	property has security or not, boolean
elevator	TINYINT	property has elevator or not, boolean
gym	TINYINT	property has gym or not, boolean
free_parking	TINYINT	property is suitable free parking or not, boolean
for_handicapped	TINYINT	property is suitable for handicapped or not, boolean
pool	TINYINT	property has pool or not, boolean

RESERVATIONS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
guest_id	INT	FK - References to primary key of guest table, ID
dependents_id	INT	FK - References to primary key of dependents table, ID
listing_id	INT	FK - References to primary key of listing table, ID
rented_at	TIMESTAMP	Time and date reservation has been made at
check_in	DATE	Check-in Date
check_out	DATE	Check-out Date

GUEST REVIEWS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
guest_id	INT	FK - References to primary key of guest table, ID
comment	TEXT	Comment for a spesific stay that guest has received from host
stars	INT	Given stars for the stay that guest has been given by host

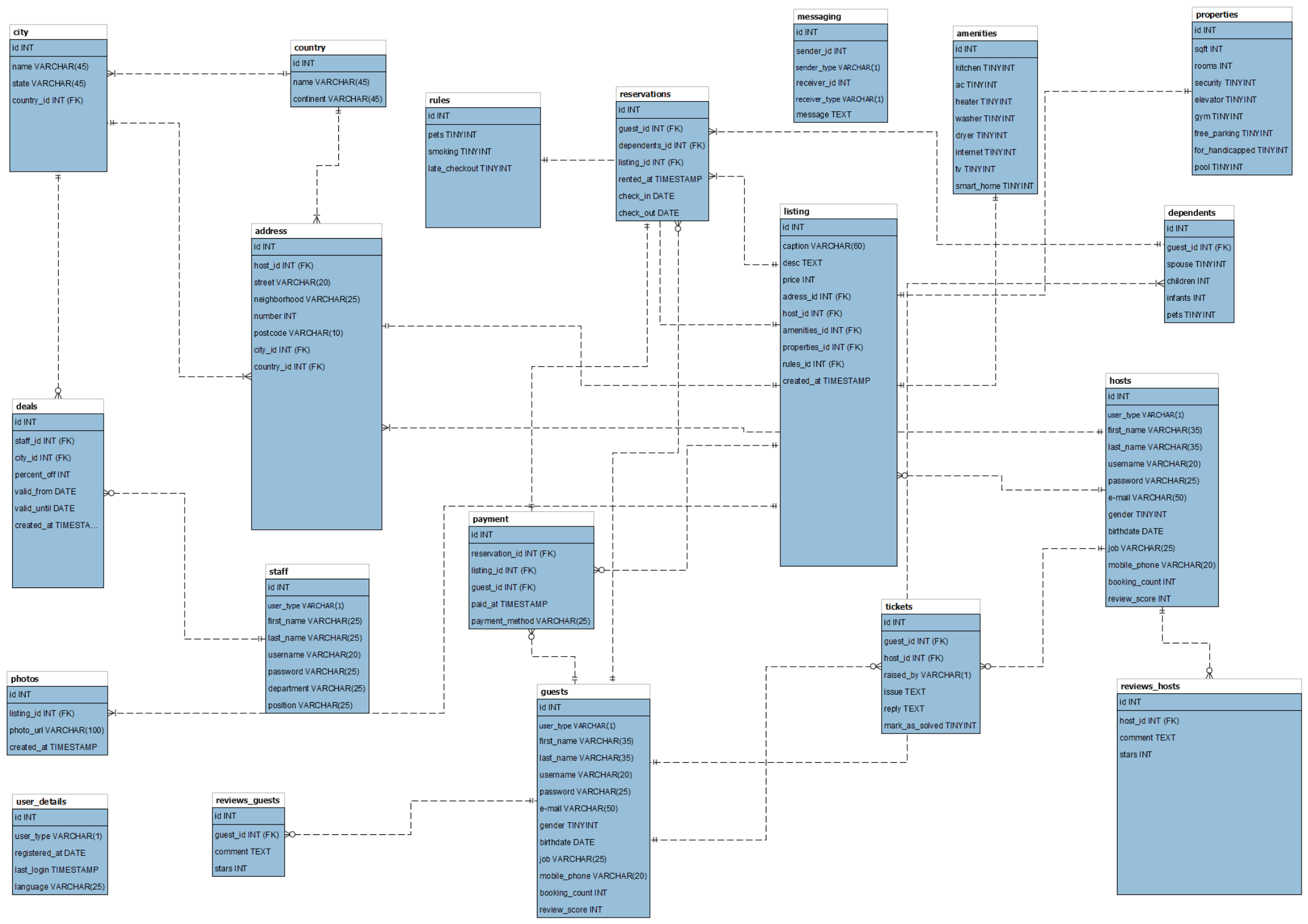
HOST REVIEWS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
host_id	INT	FK - References to primary key of host table, ID
comment	TEXT	Comment for a spesific stay that host has received from guest
stars	INT	Given stars for the stay that host has been given by guest

RULES TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
pets	TINYINT	pets allowed or not, boolean
smoking	TINYINT	smoking allowed or not, boolean
late_checkout	TINYINT	late checkout allowed or not, boolean

STAFF TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
first_name	VARCHAR(25)	name of staff
last_name	VARCHAR(25)	surname of staff
username	VARCHAR(20)	username of staff
password	VARCHAR(25)	password of staff
department	VARCHAR(25)	deparment staff works at
position	VARCHAR(25)	position of staff

TICKETS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
guest_id	INT	FK - References to primary key of guest table, ID
host_id	INT	FK - References to primary key of host table, ID
raised_by	VARCHAR(1)	displays the party who has raised the dispute, H or G
issue	TEXT	description of the issue
reply	TEXT	reply from the other party
mark_as_solved	TINYINT	staff sets this field as true or not, 1 or 0, boolean

USER DETAILS TABLE		
Column	**Data Type**	**Column Description**
id	INT	Primary key, Artificial ID
user_type	INT	user type : guest, host or staff. Can be G,H or S
registered_at	DATE	Registration date
last_login	TIMESTAMP	Last login details, as date and time
language	VARCHAR(25)	Preferred language during registration


city
id INT
name VARCHAR(45)
state VARCHAR(45)
country_id INT (FK)
country
id INT
name VARCHAR(45)
continent VARCHAR(45)
rules
id INT
pets TINYINT
smoking TINYINT
late_checkout TINYINT
reservations
id INT
guest_id INT (FK)
dependents_id INT (FK)
listing_id INT (FK)
rented_at TIMESTAMP
check_in DATE
check_out DATE
messaging
id INT
sender_id INT
sender_type VARCHAR(1)
receiver_id INT
receiver_type VARCHAR(1)
message TEXT
amenities
id INT
kitchen TINYINT
ac TINYINT
heater TINYINT
washer TINYINT
dryer TINYINT
internet TINYINT
tv TINYINT
smart_home TINYINT
properties
id INT
sqft INT
rooms INT
security TINYINT
elevator TINYINT
gym TINYINT
free_parking TINYINT
for_handicapped TINYINT
pool TINYINT
address
id INT
host_id INT (FK)
street VARCHAR(20)
neighborhood VARCHAR(25)
number INT
postcode VARCHAR(10)
city_id INT (FK)
country_id INT (FK)
listing
id INT
caption VARCHAR(60)
desc TEXT
price INT
adress_id INT (FK)
host_id INT (FK)
amenities_id INT (FK)
properties_id INT (FK)
rules_id INT (FK)
created_at TIMESTAMP
dependents
id INT
guest_id INT (FK)
spouse TINYINT
children INT
infants INT
pets TINYINT
deals
id INT
staff_id INT (FK)
city_id INT (FK)
percent_off INT
valid_from DATE
valid_until DATE
created_at TIMESTA...
hosts
id INT
user_type VARCHAR(1)
first_name VARCHAR(35)
last_name VARCHAR(35)
username VARCHAR(20)
password VARCHAR(25)
e-mail VARCHAR(50)
gender TINYINT
birthdate DATE
job VARCHAR(25)
mobile_phone VARCHAR(20)
booking_count INT
review_score INT
payment
id INT
reservation_id INT (FK)
listing_id INT (FK)
guest_id INT (FK)
paid_at TIMESTAMP
payment_method VARCHAR(25)
staff
id INT
user_type VARCHAR(1)
first_name VARCHAR(25)
last_name VARCHAR(25)
username VARCHAR(20)
password VARCHAR(25)
department VARCHAR(25)
position VARCHAR(25)
tickets
id INT
guest_id INT (FK)
host_id INT (FK)
raised_by VARCHAR(1)
issue TEXT
reply TEXT
mark_as_solved TINYINT
photos
id INT
listing_id INT (FK)
photo_url VARCHAR(100)
created_at TIMESTAMP
guests
id INT
user_type VARCHAR(1)
first_name VARCHAR(35)
last_name VARCHAR(35)
username VARCHAR(20)
password VARCHAR(25)
e-mail VARCHAR(50)
gender TINYINT
birthdate DATE
job VARCHAR(25)
mobile_phone VARCHAR(20)
booking_count INT
review_score INT
reviews_hosts
id INT
host_id INT (FK)
comment TEXT
stars INT
user_details
id INT
user_type VARCHAR(1)
registered_at DATE
last_login TIMESTAMP
language VARCHAR(25)
reviews_guests
id INT
guest_id INT (FK)
comment TEXT
stars INT