AIRBNB DATABASE PRESENTATION

PHASE 2 - IMPLEMENTATION

By Tom Schmaeling



TABLE OF CONTENTS

- I. Introduction and Presentation Structure
- 2. Changes compared to Phase I
- 3. General Database Structure
- 4. Examining Individual Table



INTRODUCTION

This presentation is a part of phase 2 and intends to provide extensive documentation of the database structure, it's tables, relationships and constraints, as well as explain the provided test cases, and their results.

We will first explain the changes to the concept, the overall structure and elements of the database before examining each of the 27* tables that make up the database.

- I encourage the reader to run the test commands in their own environment for better readability.

*One table has been removed; all changes can be found on the following slide.



CHANGES

- I have decided to remove the triphistory table. The idea behind the table was to improve the access to the bookings of a guest for easier access. There is however not a significant enough improvement to justify the redundancy.
- There have been some adjustments to the attribute distribution of the user, guest and host tables after reflecting the requirements/constraints of the individual attributes.
- The primary key's now have more descriptive names.
- Small additional changes include:
 - Addresses now have an 'address_type' attribute
 - User attribute "government_id" was changed to governmentid_image_id
 - Messages now have an 'author user id' attribute
 - Bookings no longer have the 'transaction id' attribute
 - PropertyListing now has the 'owning_host_id' attribute
 - PropertyReview attributes have changed to better align with the AirBnB app.
 - Currency attribute 'amount usd' was changed to 'amount'
 - BankInformation now has a 'name' attribute for the bank's name



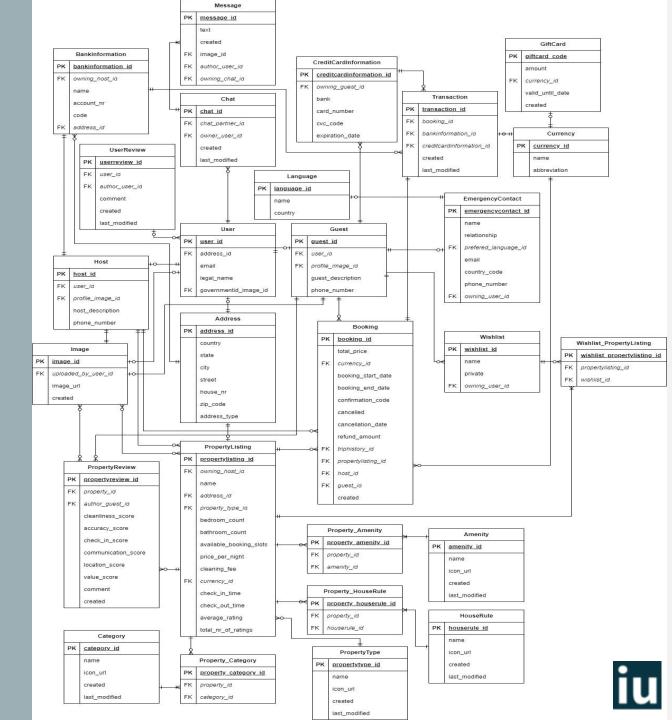
STRUCTURE

There are two main "blocks" of information the database needs to support: Users and Properties.

There are multiple tables facilitating each of these data sets, and their attributes.

The two categories of users, 'guest' and 'host' share a base 'user' class creating a joined subclass table strategy.

The 'PropertyListing' table includes relevant attributes for the properties offered on the page, multiple of which use N:M relations and therefore need to be normalized via additional tables.



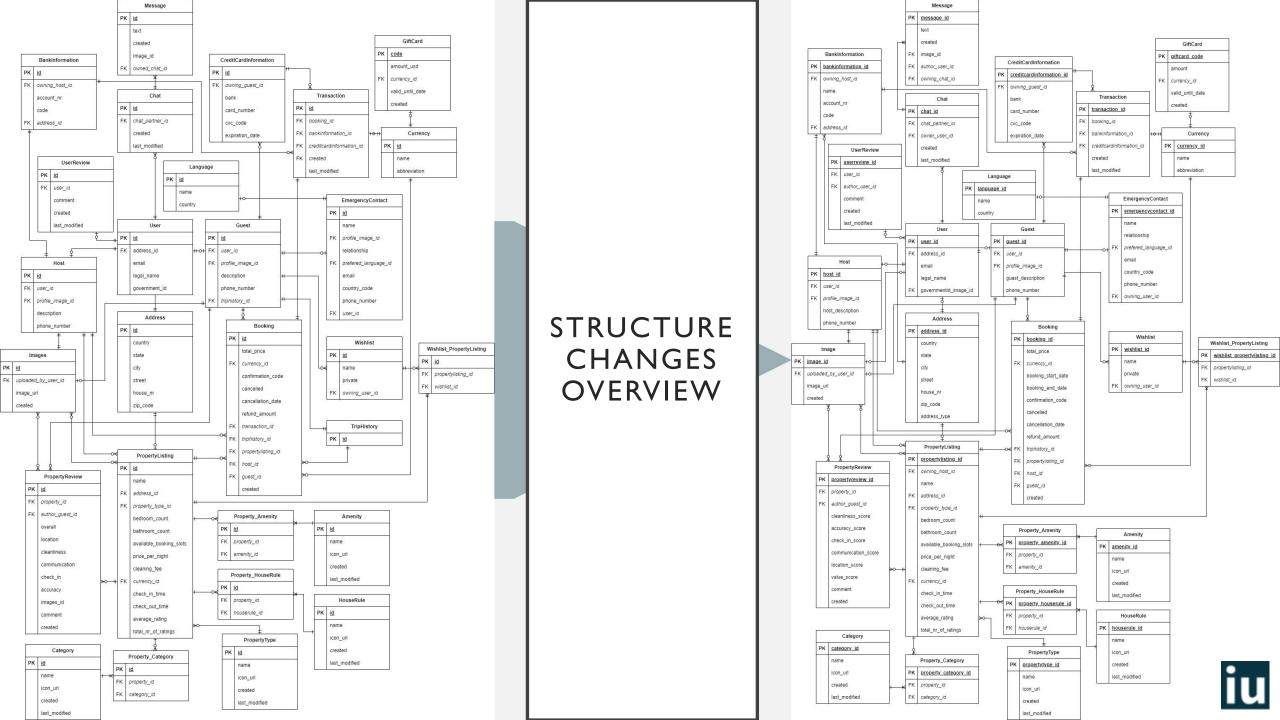


TABLE - USER

% N = 5	User
PK	user id
FK	address_id
	email
	legal_name
FK	governmentid_image_id

Tested via 'Guest' and 'Host' test cases

- The User is the base/super class for all users and holds attributes any user will have
- The 'address_id' is a foreign key that references the address table
- The 'governmentid_image_id' is a foreign key that references the image table.
- This table is tested via the host and guest tables which can be seen in the following two slides



TABLE - GUEST



LEFT JOIN EmergencyContact EC ON U.user_id = EC.owning_user_id

LEFT JOIN CreditCardInformation CCI ON G.guest_id = CCI.owning_guest_id;

- The 'user_id' attribute refers back to the base class
- 'profile_image_id' references the id of an image that can be loaded to display the users profile image
- The test case is meant to test the relationship between the 'Guest' table/class and its super class 'User'. As well as testing other relevant relationships.





TABLE - HOST

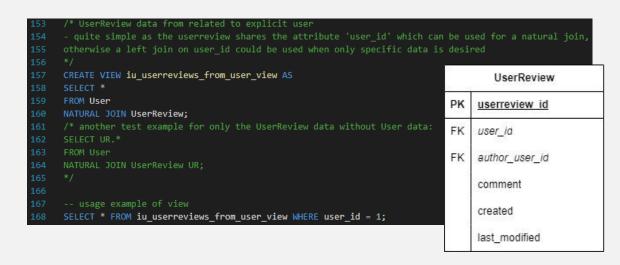
```
CREATE VIEW iu_userhost_view AS
   U.user id,
   U.legal_name,
                                                                                                 Host
   U.email,
   A.country,
   A.state,
                                                                                       host id
   A.city,
   A.street,
   A.house_nr,
                                                                                       user_ia
   A.zip_code,
   H.host_description,
   H.phone number.
                                                                                        profile image id
   IProfile.image_url AS profile_image_url,
   IGovernID.image_url AS governmentid_image_url,
   BI.account_nr AS bank_account_nr,
   BI.code AS bank code,
                                                                                        host description
   BI.address id AS bank address id,
   BIA.country AS bank_address_country
                                                                                       phone_number
   User U
JOIN Address A ON U.address id = A.address id
JOIN Host H ON U.user_id = H.user_id
JOIN Image IProfile ON H.profile_image_id = IProfile.image_id
JOIN Image IGovernID ON U.governmentid image id = IGovernID.image id
LEFT JOIN BankInformation BI ON H.host_id = BI.owning_host_id
JOIN Address BIA ON BI.address id = BIA.address id;
```

SELECT * FROM iu userhost view WHERE user id = 21;

- The 'user_id' attribute refers back to the base class
- 'profile_image_id' references the id of an image that can be loaded to display the users profile image
 - The reasoning for having this attribute in both subclasses is that it is only required for
- The 'triphistory_id' is used to link together
 - The test case is very similar to that of the previous guest class, adjusting where relevant relationships change compared to the previous slide.



TABLE - USERREVIEW



- The 'UserReview' table holds all the reviews given on users.
- These are comments left by hosts on the guest pages and are different from reviews left by guests on the property listing.
- The comment attribute holds the user created written text
- User and author_user ids and timestamps are also saved
- The test case tests the content and relationship of user reviews and users by returning all data entries of user reviews for a given user.

	_userreviews_from_user_view V							
			governmentid_image_id					last_modified
1 1 1	max.musterman@example.com max.musterman@example.com max.musterman@example.com	Max Musterman		1 2 3	21 23 24	They were a lovely guest, we hope to meet you again some time! There were no problems, and they left the property clean and in order. Thanks for your stay!	2024-01-08 10:03:57 2024-01-08 10:03:57 2024-01-08 10:03:57	2024-01-08 10:03:57 2024-01-08 10:03:57 2024-01-08 10:03:57



TABLE - ADDRESS

	Address	
PK	address id	
	country	
	state	
	city	
	street	
	house_nr	
	zip_code	
	address_type	

Tested via multiple other test cases

- The address attributes themselves are selfexplanatory in meaning.
- The reasoning for the selection of attributes is the categorization or search user flow on the website.
- An address can be used by either users, properties (listings) or banks.
- This table is tested is a part of many other tables and is therefore already tested more than enough. There is still a simple select all test to check for completeness of content in the test.sql file.



TABLE - IMAGE

W	Image
PK	image id
FK	uploaded_by_user_id
	image_url
	created

Tested via multiple other test cases

- This table assumes that the images themselves are stored by a cloud storage provider for example.
- This means that the attribute itself can be of type VARCHAR instead of having to save the image as a BLOB, which is inefficient.
- The Table also stores the id of the user that uploaded the image.
- The 'created' attribute is defaulted to the current timestamp, meaning it shows the time the image was uploaded.
- Like the 'Address' table, the 'Image' table is also part of many other tables, meaning that the relationships of this table are already thoroughly tested. There is, again, still a select all test to check for completeness of content.



TABLE - CURRENCY

Currency			
PK	currency id		
	name		
	abbreviation		

'Currency' is a simple table, see test case in test.sql

- The currency table represents all currencies available in the application.
- Name represents the name of the currency while country represents the country it is used in.
- Abbreviation is used for display purposes.
- Currency is not a table that will see frequent changes and is rather used as a reference.
- This table is very simple and does not require complicated testing, a simple select all test can be found in the test.sql file.



TABLE - LANGUAGE

	Language			
PK	language id			
	name			
	country			

'Language' is a simple table, see test case in test.sql

- The language table represents all languages available in the application.
- Name represents the name of the language while country represents the country it is used in.
- The reason for this is that Airbnb differentiates between, for example, American and British English.
- Language is not a table that will see frequent changes and is rather used as a reference.
- Like the 'Currency', this table is also very simple, a test for content should suffice for this table, which can be found in the test.sql file.



TABLE - CHAT

```
two examples, the first one will display general details of the chat, the second text will
return all messages of a given chat
CREATE VIEW iu_chat_details_view AS
   UO.legal name AS owning user name,
   UP.legal_name AS partner_guest_name,
   COUNT(M.message_id) AS message_count
JOIN User UO ON C.owner user id = UO.user id
                                                                                Chat
JOIN User UP ON C.chat partner id = UP.user id
LEFT JOIN Message M ON M.owning chat id = C.chat id
                                                                 PK chat id
GROUP BY C.chat id, UO.legal name, UP.legal name;
-- usage example of view
                                                                        chat partner id
SELECT * FROM iu_chat_details_view WHERE chat_id = 1;
CREATE VIEW iu chat messages view AS
                                                                        owner_user_id
   M.message_id,
                                                                        created
   M.image id,
   C.chat_id AS owning_chat_id_ref
                                                                        last_modified
FROM Message M
JOIN Chat C ON C.chat_id = M.owning_chat_id;
-- usage example of view
SELECT * FROM iu_chat_messages_view WHERE owning_chat_id_ref = 1;
```

- A chat is a collection of messages, in the next slide we will inspect the 'message' table.
- These messages are linked to a chat by sharing the same chat id.
- Other than the 'created' and 'last_modified' timestamps, the chats table also holds both chat participants.
- The test case for this table works in combination with the message table (next slide). The first test case checks the general content of the chat table.



TABLE - MESSAGE

```
two examples, the first one will display general details of the chat, the second text will
return all messages of a given chat
CREATE VIEW iu_chat_details_view AS
    UO.legal_name AS owning_user_name,
    UP.legal name AS partner guest name,
    COUNT(M.message_id) AS message_count
                                                                         Message
JOIN User UO ON C.owner user id = UO.user id
JOIN User UP ON C.chat partner id = UP.user id
                                                                    message id
LEFT JOIN Message M ON M.owning chat id = C.chat id
GROUP BY C.chat_id, UO.legal_name, UP.legal_name;
                                                                    text
-- usage example of view
SELECT * FROM iu chat details view WHERE chat id = 1;
                                                                    created
CREATE VIEW iu chat messages view AS
                                                                    image_id
    M.message id,
    M.text,
                                                                    author_user_id
    M.image id,
    C.chat_id AS owning_chat_id_ref
                                                                   owning_chat_id
FROM Message M
JOIN Chat C ON C.chat id = M.owning chat id;
 -- usage example of view
SELECT * FROM iu_chat_messages_view WHERE owning_chat_id_ref = 1;
         1 | Content for message 1 |
```

row in set (0.00 sec)

- The message is linked to the chat via the chat id, represented here as 'owning_chat_id'. The 'author_user_id' attribute holds the author; Both id's can be used to reconstruct a chat.
- As Airbnb enables users to share images in chats, a message may contain an image instead of text. This necessitates the optional 'image_id' attribute.
- The created timestamp is essentially the 'sent' timestamp of the message.
- The second test case in the screenshot test the relationship of 'Chat' and 'Message' tables by returning all data entries of messages for a given



TABLE - EMERGENCYCONTACT

	EmergencyContact
PK	emergencycontact id
	name
	relationship
FK	prefered_language_id
	email
	country_code
	phone_number
FK	owning_user_id

'Language' is a simple table, see test case in test.sql

- The emergency contact different to a user as it is only a collection of information relevant to the contact without any of the functionality of an actual account.
- All the information regarding the contact should be self-explanatory.
- The 'owning_user_id' refers to the user that owns it.
- This table is covered in the Guest user test cases, another simple select all statement is provided in the test.sql file to check for completeness of content.



TABLE - WISHLIST

```
- test case that shows the wishlist, user and property listing data to proves the proper
implementation of links between them. Goal is to see the relationship of a user owning
a wishlist, which in turn 'owns' (multiple) propertylistings.
                                                                                  Wishlist
CREATE VIEW iu wishlist details view AS
                                                                            wishlist id
   W.*,
   U.legal name AS owning user name,
                                                                            name
    PL.name AS property_listing_namne
                                                                            private
   Wishlist W
JOIN User U ON W.owning user id = U.user id
                                                                            owning user id
JOIN Wishlist PropertyListing WPL ON W.wishlist id = WPL.wishlist id
JOIN PropertyListing PL ON WPL.propertylisting id = PL.propertylisting id;
-- usage example of view
SELECT * FROM iu wishlist details view WHERE wishlist id = 1;
-- this view gets all data regarding the property listings that are in a given wishlist
CREATE VIEW iu wishlist propertylistings view AS
   W.wishlist id,
   PL.*
FROM Wishlist W
JOIN Wishlist PropertyListing WPL ON W.wishlist id = WPL.wishlist id
JOIN PropertyListing PL ON WPL.propertylisting_id = PL.propertylisting_id;
-- usage example of view
SELECT * FROM iu wishlist propertylistings view WHERE wishlist id = 1;
```

- Wishlists in the Airbnb application are a collection of property listings.
- As this is a M:N relation it needs to be normalized, we achieve this by using this 'wishlist' and a 'wishlist_propertylisting' table (see next slide).
- A user can have multiple wishlists, hence the need for a name.
- The 'wishlist_id' is used in the next table to normalize the M:N relation.
- The table works in close relation to the 'wishlist_propertylisting' table (next slide). This first test case checks for general data of the chat table and other related tables in conjunction. The second test checks the relation of a chat and its messages.

wishlist_id	name	private	owning_user_id	owning_user_name	property_listing_namne
	wishlist 1 wishlist 1 wishlist 1 wishlist 1 wishlist 1	1 1 1 1 1		Max Musterman Max Musterman Max Musterman Max Musterman Max Musterman	Cozy Studio Apartment Cozy Studio Apartment Spacious Loft in the City Luxurious Beachfront Villa Mountain Retreat Cabin



TABLE - WISHLIST_PROPERTYLISTING

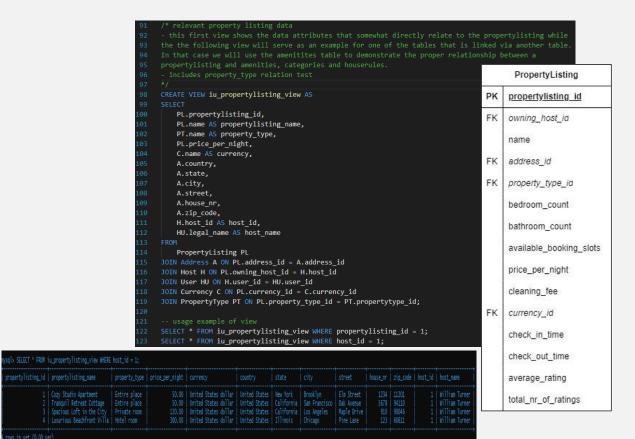
```
test case that shows the wishlist, user and property listing data to proves the proper
 implementation of links between them. Goal is to see the relationship of a user owning
a wishlist, which in turn 'owns' (multiple) propertylistings.
                                                                      Wishlist_PropertyListing
CREATE VIEW iu_wishlist_details_view AS
                                                                       wishlist propertylisting id
    U.legal name AS owning user name,
                                                                 FK
                                                                       propertylisting_id
    PL.name AS property_listing_namne
                                                                       wishlist_id
    Wishlist W
JOIN User U ON W.owning_user_id = U.user_id
JOIN Wishlist PropertyListing WPL ON W.wishlist id = WPL.wishlist id
JOIN PropertyListing PL ON WPL.propertylisting_id = PL.propertylisting_id;
-- usage example of view
SELECT * FROM iu wishlist details view WHERE wishlist id = 1;
-- this view gets all data regarding the property listings that are in a given wishlist
CREATE VIEW iu wishlist propertylistings view AS
    W.wishlist_id,
    PL.*
FROM Wishlist W
JOIN Wishlist_PropertyListing WPL ON W.wishlist_id = WPL.wishlist_id
JOIN PropertyListing PL ON WPL.propertylisting id = PL.propertylisting id;
SELECT * FROM iu_wishlist_propertylistings_view WHERE wishlist_id = 1;
```

- This table, as mentioned, is used to normalize the wishlist – propertylisting relation.
- The table matches Propertylistings to Wishlists using the two foreign key ids.
- The propertylisting table will be introduced in the following slides.
- As explained in the previous slide, this table stands in close relation to the wishlist table. The second test case demonstrates the relationship of a wishlist with the propertylisting table via the link of this table very well.

mysql> SELECT :			ew WHERE wishlist_id = 1;		 +	 +	+	+		+		
wishlist_id	propertylisting_id	owning_host_id		address_id		available_booking_slots						
1	1	NULL	Cozy Studio Apartment	61	1						0.0	0
1			Cozy Studio Apartment Spacious Loft in the City	62 63			50.00 120.00		15:00:00 16:00:00	11:00:00 10:00:00	0.0	
1		NULL NULL	Luxurious Beachfront Villa Mountain Retreat Cabin	64 65			300.00 80.00		14:00:00 12:00:00	12:00:00 10:00:00	0.0	
F rows in set	(0.00.505)											



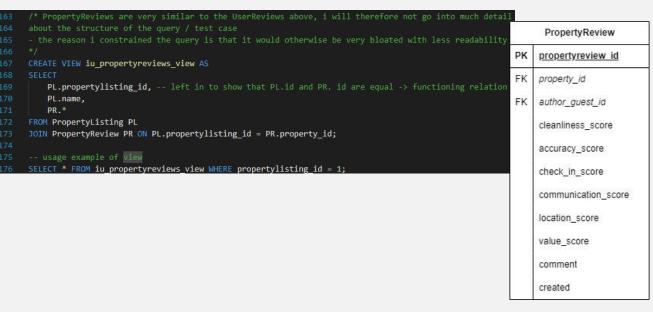
TABLE - PROPERTYLISTING



- The 'PropertyListing' table marks the second important 'block' of data mentioned in the introduction.
- This table holds all relevant information for the listings and has multiple M:N relations that are not shown in the table itself.
- These relations rely on the 'propertylisting_id' and are introduced in the following slides.
- Although many of the relations of this table have already been tested, this table of great importance to the overall system. It is therefore reasonable to create a test case that generally tests all relationships of this table. Said test case can be seen here, it focuses on returning relevant from all tables that have a relationship with the propertylisting table. (Either via property or host id)



TABLE - PROPERTYREVIEW



- Similar to a User Review this table holds the reviews given by guests to properties.
- It holds multiple ratings/scores which are displayed on the listing page. These are represented as integers with values from 0 to 5. (Star rating)
- Users can also add comments to the review.
- The test case simple checks the relationship between propertylistings and reviews, as well as the review content.





TABLE - PROPERTY_X

Property_Category

PK property_category id

PK property_amenity id

PK property_id

FK property_id

FK property_id

FK category_id

FK amenity_id

FK house

```
Property_HouseRule

PK property houserule id

FK property_id

FK houserule_id
```

```
/* this test case shows the selected amenities for a certain propertylisting
- this same structure will work for categories and houserules as well, which will not be included as
they function identically to this view (with different names, etc.)
- i have essentially decided to view amenities, categories and houserules as an equivalence class
and chose Amenity as the representative. this can reduce testing workload/effort
(this can/could be reasoned as trying to reduce testing costs for example)

*/

CREATE VIEW iu_propertylisting_amentities_view AS

SELECT

PL.propertylisting_id,
PL.name AS property_name,
A.name AS amenity_name

FROM Propertylisting PL

JOIN Property_Amenity PA ON PL.propertylisting_id = PA.property_id

JOIN Amenity A ON PA.amenity_id = A.amenity_id;

-- usage example of view

SELECT * FROM iu_propertylisting_amentities_view WHERE propertylisting_id = 1;
```

*X = Category/Amenity/HouseRule

- Each table is used to normalize a relation between the PropertyListing and their respective table. I see these tables as one equivalence class.
- The naming convention of these tables is meant to represent the link between 'PropertyListing' and 'Amenity', 'Category' or 'HouseRule' table.
- The individual tables they are related to will be introduced in the following slide.
- These table are used to link other tables and hence do not necessarily need an id attribute.

 (I am considering removing these for the finalization phase and would appreciate feedback regarding this)
- The test case checks the link of the propertylisting to the respective table via the link of these tables by returning data of each data entry.



TABLE - AMENITY/CATEGORY/HOUSERULE

	Category		Amenity	HouseRule		
PK	category id	PK	amenity id	PK	houserule id	
	name		name		name	
	icon_url		icon_url		icon_url	
	created		created		created	
	last_modified		last_modified		last_modified	

```
305  /* Amenity data */
306  SELECT * FROM Amenity;
307
308  /* Category data */
309  SELECT * FROM Category;
310
311  /* HouseRule data */
312  SELECT * FROM HouseRule;
```

Simple tables, which are tested in previous test case.

- These are the tables that are in a M:N relationship with the 'PropertyListing' table.
- Like the tables in the last slide, I also consider these tables one equivalence class.
- These table hold unexpectedly little information because they are only represented by their name and an icon in the application.
- Relevant for test cases are the relationships of these tables, these are already tested other test cases, resulting in these simple queries, which check for content.



TABLE - PROPERTYTYPE

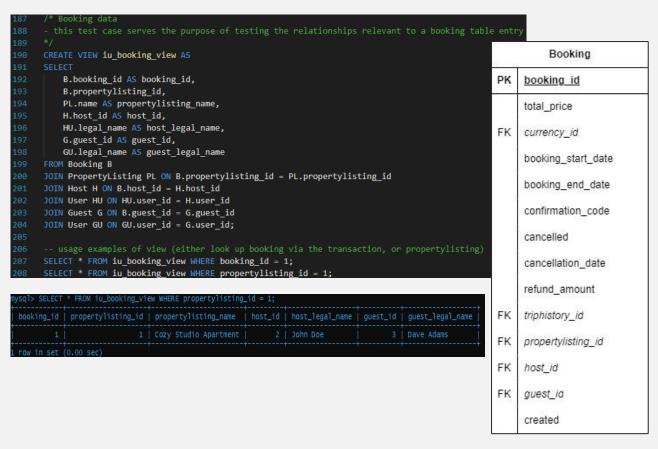
	PropertyType	
PK	propertytype id	
	name	
	icon_url	
	created	
	last_modified	

'PropertyType' is a simple table, see test case in test.sql

- Different to the previous three tables, the 'PrototypeType' is not in a M:N relation with the PropertyListing table.
- In the Airbnb application, each property is listed based on its type.
- These types are not frequently changed, and if they are, then they are changed by an admin.
- The relevant relationship of this table is already table is already tested in the propertylisting test case, the content itself is once again tested via a simple select all statement in the test.sql file.



TABLE - BOOKING



- Once a guest books a property, an entry in the 'Booking' table is made.
- Most of the information listed is obvious and inferred by the attribute name.
- The time frame is stated by the booking start and end dates.
- The optional 'cancelled' Boolean and the relevant 'refund_amount' is only used in case the booking is cancelled.
- The test cases displays that the relationships work as intended, and the related data is properly drawn into one coherent data entry.



TABLE - BANKINFORMATION

% S = 50	BankInformation				
PK	bankinformation id				
FK	owning_host_id				
	name				
	account_nr				
	code				
FK	address_id				

Tested via multiple other test cases

- As stated in the problem statement, Bank Information is only relevant for hosts.
- The 'BankInformation' table references its address as well its owning host user via ids.
- The rest of the information given in this table is example data relevant to Banks.
- This tables relationships are tested via other test cases, the content itself is tested once again via a select all query at the end of the test.sql file.



TABLE - CREDITCARDINFORMATION

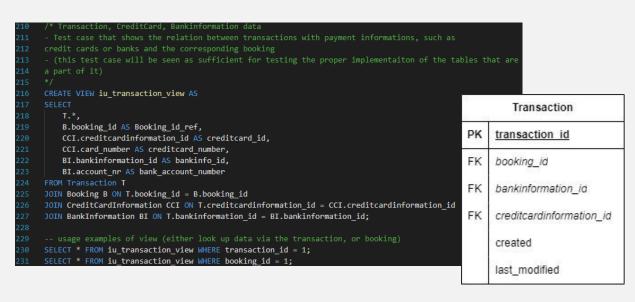
CreditCardInformation		
PK	creditcardinformation id	
FK	owning_guest_id	
	bank	
	card_number	
	cvc_code	
	expiration_date	

Tested via multiple other test cases

- As stated in the problem statement, Credit Card Information is only relevant for guests.
- The 'CreditCardInformation' table references its address as well its owning host user via ids.
- The rest of the information given in this table is example data relevant to Banks.
- Counterpart to the previous table, both are also tested in the test cases of the 'Transaction' table in the following slide.



TABLE - TRANSACTION



- The transaction holds relevant payment information for each booking.
- The table references 'BankInformation' and 'CreditCardInformation', as well as the booking it was made for.
- The test case can be used to check if the relationships of the transaction table and the respective payment information and 'booking_id' are plausible.

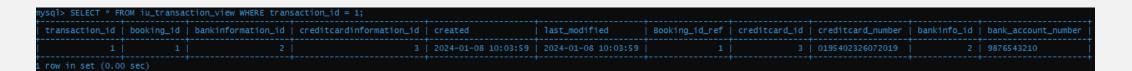




TABLE - GIFTCARD

GiftCard		
PK	giftcard code	
	amount	
FK	currency_id	
	valid_until_date	
	created	

'GiftCard' is a simple table, see test case in test.sql

- The 'GiftCard' table contains information relevant to gift cards.
- The 'amount' attribute saves the value of gift card in us dollars.
- The 'currency_id' attribute can then be used to convert it, if necessary
- The GiftCard table is a very simple, which is not as important as other tables to the functionality of the system and therefore does not warrant a complicated test case, content is checked again at the end of the test.sql file.



THANKS FOR READING

Additional information is given in the Pebble Pad submission as well as in code via comments.

