Phase 3 work

Finalization Phase

Abstract:

A database was created for Airbnb business. The business has two types of users. Hosts and guest. Each category can perform a number of actions and create different types of data that need to be stored in relation to the category. A table is used to store all users that sign up on Airbnb website. Two separate tables are used to categorize users as hosts and guests. For these two tables, there are a number of different tables created that relate back to these tables and some of these tables also have relationships with each other. For hosts, their information needs to be stored in tables such as: personal host information, rental information like type of rental, number of rooms, available services, pictures and location etcetera. Similarly, for guests, information like how many bookings they make, how many payments they make, in what currency, service charges, taxes and their social accounts etcetera. Then, some user generated data like message conversations between hosts and guests, their reviews and ratings, their profile ratings etcetera.

Initially, the design consisted of about 23 tables but the design implementation was inadequate. This was because cardinality between different entities was not matching up to the realistic requirement. For example, ratings of hosts and guests were to be stored in a single table but both hosts and guests can give multiple ratings so storing that in one table meant that there would be a use of primary keys multiple times. This would be error prone so a better approach was to store guests' and hosts' reviews separately with their own table and primary keys. And using a junction table we can link back and forth between who gave reviews to whom and how much. Similar approach was taken to properly define relationships between payments and taxes table pair and messages and languages table pair. This was a number of junction tables were introduced to correctly define the required cardinality.

The database stores everything in its internal structure in a relational format. Some of the tables have NULL values allowed for their attributes, even though from the perspective of the database this is not good, but in the realistic application this makes sense because, for instance, in the `booking` table the attributes `check-in` and `check-out` will not be filled until the guest actually checks in the stay site and then checks out on his final day, so not allowing NULL values in the database will not work. These attribute values will be later updated by the system and stored in the database. Similarly, for the table `listing` the foreign key attribute `listing_addressID` is allowed NULL because, depending on how the Airbnb website works for the host, the listing will be created first and then later the listing address will be stored for that listing, so if the `listing_addressID` attribute is not allowed to be NULL then the database would not store the listing at all, which would not be acceptable for the application. Hence, other tables have similar cases that are solved by setting attributes as NULL.

Some key features of the database are the segregation of tables into two main hierarchies in the ERD sense. This should aid the developer querying the database make sense of how the schema is logically built. As for the database sense, there is enough data to get answers for many different questions so very complex queries can be run. For example, a developer can perform simple sentiment analysis by querying results that show counts of specific sentimental words – like ‘Good’, ‘Great’ or ‘okay’, ‘bad’ etc – in reviews for all the rentals stored in the database. Which can be further segregated into sentiment by country or type of rental etcetera.

Metadata:

Table	Rows	Size
amenities	20	32.0 KiB
availability	20	32.0 KiB
booking	49	48.0 KiB
coupons	43	32.0 KiB
currency	179	16.0 KiB
guest	21	48.0 KiB
guest_rating_on_host	49	48.0 KiB
host	17	64.0 KiB
host_language	20	48.0 KiB
host_ratings_on_guest	255	48.0 KiB
images	70	80.0 KiB
language	393	48.0 KiB
listing	21	64.0 KiB
listing_address	20	16.0 KiB
listing_images	80	48.0 KiB
messages	795	1.9 MiB
messages_language	999	96.0 KiB
payments	49	96.0 KiB
payments_taxes	753	80.0 KiB
property_category	20	32.0 KiB
property_type	20	32.0 KiB
ratings_on_guest	255	48.0 KiB
ratings_on_host	49	48.0 KiB
reviews	43	48.0 KiB
room_type	20	32.0 KiB
service_charges	20	16.0 KiB
taxes	20	32.0 KiB
user	38	16.0 KiB
user_social_networks	141	32.0 KiB
wishlsit	79	48.0 KiB
Total		
30	4558	3.2 MiB