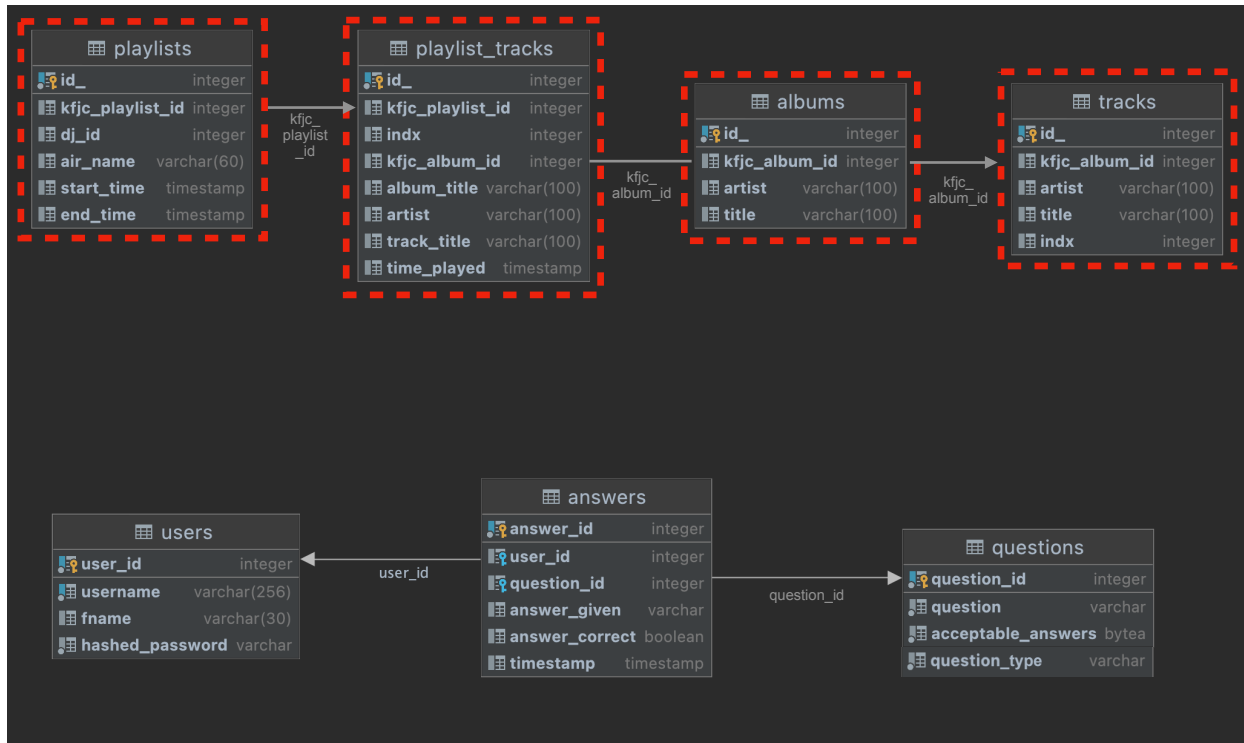


KFJC Trivia Robot Data Model



NOTES:

- The KFJC Station Data can be exported as csv and imported into PostgreSQL. I have no control or influence over the content or construction of these tables.
- The model is greatly simplified because I elected:
 - Not to import all KFJC Tables.
 - Not import all columns available to me.
 - Not import blank rows in playlist_tracks, usually indicative of a mic break.
 - The tracks table and collection_tracks tables have been collapsed into one table, tracks.
- In playlists, user_id has been renamed to dj_id, a more meaningful name for this project.

The condition of the data is: large and a bit dirty requiring a little cleanup on import.

The objective is "not to import a row I can't make a question out of."

So far:

- playlist_tracks is 2.5mil rows. Not importing blank rows cut that number by 10%. This number represents the 2.25 million songs that have been played on KFJC since September 1995. Wow.
- This required the "chunking" of data into 100,000 line chunks for importing in stages.
- KFJC's sound is edgy and eclectic. Let's just say a "Profanity Filter" has now been incorporated into the project.
- Yes, the artists field is duplicated in albums and tracks. Not the best practice but it allowed me to collapse the two tracks tables together, greatly simplifying the project for me.
- Indeed, album_title, artist and track_title are yet again duplicated in playlist_tracks but this is probably because DJs can bring albums from home or the internet and that record isn't a in the station library. This allows the DJ to keep an accurate log (fulfilling an FCC requirement) without modifying the station's albums table.

Users, Questions and Answers are completely controlled by me.

- The users table no longer needs the salt column for password salting because bcrypt can conveniently store them together.
- user_answers has been renamed answers.
- answer_correct has been added to answers. This will be a boolean with NULL allowed (indication a question skipped). The user response is graded immediately, the answer row is created and committed.
- A user should never be offered the same question twice. This can be accomplished by saving the question_id for each answer in the answers table.
- question_type was added to aid in displaying the answer correctly: tracks and indx for album questions, artist/album/track for track questions and dj rankings (by dates or number of shows).