CSCI 2141 ASSIGNMENT PART 2

Section 0: Adjust Your Dataset (if Necessary)

According to the feedback, I have added information about the primary keys and foreign keys to the database schema.

1. SONG – This table is going to be a table containing data about the various songs available on Spotify.

The primary key is going to be song\_id and the foreign key is going to be album\_id which references the album\_id in ALBUM.

1. CHART- which is a table that contains data about the top songs of the year 2022 on Spotify.

The primary key is going to be chart\_id.

The foreign key is going to be artist\_id which references the primary key in ARTIST i.e., artist\_id.

1. ARTIST- this table is going to contain information about the various artists on Spotify.

The primary key is artist\_id.

The foreign key is song\_id which references the song\_id in SONG.

Each artist can have multiple song and each song can have multiple artists.

1. ALBUM- this table is going to contain information about albums released by various artists.

The primary key is album\_id.

Since, I have added foreign keys to the table, I have decided to remove some columns from my database to reduce redundancy, from 2 of my tables:

SONG & CHART:

These tables are not going to have duration\_ms & speechiness as attributes anymore. Furthermore, I am removing the attribute weeks\_on\_chart from CHART.

ARTIST:

This table is not going to have top\_song as an attribute, it is going to be replaced by the foreign key song\_id.

I have also decided to use ChatGPT to generate the data to increase the number of rows in my dataset and make the process of dataset generation more convenient for ARTIST & ALBUM.

Section 1: Design Your Database in 3NF

Design the 3NF for the database.

The database is already in 1 NF form since, there are no comma separated lists in the database.

Dependency diagram before normalization:

A blue rectangular object with black text

Description automatically generated

A blue rectangular object with black text

Description automatically generated

Dependency diagram after normalization:

Album:

A blue rectangular box with black text

Description automatically generated

Artist:

A blue rectangular sign with black text

Description automatically generated

Song:

A blue rectangular object with black text

Description automatically generated

Chart:

A blue rectangular sign with black text

Description automatically generated

Internal Schema before normalization:

A computer screen shot of a computer

Description automatically generated

Internal schema after normalization:

A computer screen shot of a computer

Description automatically generated

There are no partial dependencies in the database because there is only one primary in in every single table i.e., the id of each entity. Therefore, they are in 2NF.

However, there are some transitive dependencies in the tables SONG & CHART due to the addition of foreign key attributes to the table. Which can be removed by the process of normalization, but this would not increase the number of tables in the database because, they are foreign keys which means that a table defining these attributes already exists in the database.

Which implies that the dependent non-prime attributes only need to be removed from SONG & CHART.

SECTION 2:

Please refer the corresponding sql file.

SECTION 3:

Where clause:

A screenshot of a computer program

Description automatically generated

Inner join:

A screen shot of a computer

Description automatically generated

Group by:

A black screen with white text

Description automatically generated