Instructions. Write up your answers to the following questions and submit them to your GitHub repository for HW-2 by the end of the day on the due date (more details on submitting to GitHub will be provided in Piazza). Your answers must be typed (not handwritten), organized, and clearly marked (i.e., with respect to which question and part is being answered). It is okay to draw the query trees in Question 4 by hand and include a picture of each in your submission. Additionally, include your first and last name and the assignment (HW-2) at the top of your write up. Your answers must be submitted as a single PDF file called hw-2.pdf.

1. [15pts] Define a relational schema for storing the following information about music albums, tracks, songs, music groups, group members (artists), music genres, and record labels. Your relations must support the following requirements and constraints. For each relation primary keys, candidate keys, and foreign keys as appropriate in your design. Your schema should not use any “surrogate” keys.

   - Each album has at a minimum a title, the year it was recorded, the group that recorded the album, the tracks on the album, and the album’s record label. While two albums can have the same title, it is okay to assume that a group only has one album with a given title.

   - Each music group can be associated with zero or more music genres. Each genre has a unique label, e.g., “jazz”, “rock”, “pop”, “rap”, and a longer description of the genre itself. For instance, the description of “indie pop” is “Combines guitar pop with DIY ethic in opposition to the style and tone of mainstream pop music.”

   - Each group can be influenced by zero or more other music groups. Each music group has a name, assumed to be unique across groups, and the year the group was formed.

   - Music artists can be members of zero or more music groups. An artist is a member of a group within a certain range of years (e.g., from 1991 to 2000). It could be the case that the artist is still a member of a group (e.g., from 2007 to the present). Each artist has a name, and a birth year. For simplicity, you can assume no two artists have the same name.

   - In cases where a record album is recorded under the name of an individual artist, assume that the album is recorded by a group whose name is the artists’ name. In this way, both the artist and the other musicians involved in recording the music for the album are captured.

   - A music track represents a recorded song. Each music track is associated with one or more artists that produced it and the song it is a recording of. Albums consist of multiple tracks, but a track can also be released as a single (i.e., not part of any particular album). A particular track can also be included on multiple albums. Each track has the year it was recorded along with a unique track name (as a numeric identifier).
• A song has a title and the year it was written. In addition, each song is associated with the artists that it was written by. For simplicity, you can assume that no two songs have the same title.

2. [8 pts] Using your answer in Question 1, create a schema diagram of your schema.

3. [2 pts] We typically think of a “well designed” schema as one with minimal redundancy (note that there are times when a bit of redundancy can help performance). For instance, representing the above information in a single table would result in considerable redundancy. Based on this, rate how “good” you think your schema design is and why. Be specific (e.g., use examples) when describing your rating and list where you think redundancy could exist if at all.

4. [10 pts] Assume the following relational schema for representing computer information.

\[
\text{product(maker, model, type)} \\
\text{pc(model, speed, ram, hd, price)} \\
\text{laptop(model, speed, ram, hd, screen, price)}
\]

Note that computer speed (measured in GHz) and screen size (in inches) are floating point values, and ram and hd (for hard-disk size) are integer values with the unit GiB, and price is an integer value. The maker attribute represents the name of the product manufacturer. For each part below, write a corresponding relational algebra query and its query tree.

(a). Find all of the makers of laptops with a hard disk of at least 256 GiB.

(b). Find the model number and price of all products (of any type) made by manufacturer B (assuming makers are named A, B, C, and so on).

(c). Find all of the makers of laptops but not PCs.

(d). Find all of the specific hard-disk sizes that occur in two or more PCs.

(e). Find all of the makers of PCs with the highest available speed.