

CS400-010 Fall 2021

X5 Team Project Proposal - X-Team 4:

- Adam Cook, cook29@wisc.edu
- Felix Lin, flin45@wisc.edu
- Matthias Schmitz, mjschmitz6@wisc.edu
- Tomas Perez, tperez7@wisc.edu

Title: Music Library Analytics

Problem: Music management software (e.g. iTunes, Rekordbox) does not typically provide the user with a way to report on their music library with summarizations and visualizations.

Primary Stakeholder: The primary users of this application will be people who own digital music collections, such as music enthusiasts and disc jockeys. Music collection owners will be able to use this application to summarize and analyze their collection. They'll get a breakdown of how much of certain types of music they have, and how their collection has grown over time.

Graphical User Interface: The interface of the application will provide the following options:

1. **Find similar tracks:**

A free text field will allow the user to search for a track by title. Depending on capabilities of the GUI, it will completion-match or search for track names that start with what the user entered to help them pick one that exists in the database. The program will output tracks done by artists who have collaborated with the artist(s) who made the selected track. It will have the following options:

- a. A checkbox to restrict suggestions to the same genre as the selected track (no genre restriction if de-selected)
- b. After searching one level deep, an option to search a 2nd level deep

2. **Genre statistics:**

A free text field will allow the user to enter a genre. Depending on capabilities of the GUI, it will completion-match or search for genres that start with what the user entered to help them pick one that exists in the database. The program will produce a graph or table report (depending on GUI capabilities) which shows how many tracks of this genre were added to the user's library over time. The summarization will compare the growth of this genre in the library to the library as a whole to see if it takes up a smaller or larger percentage of the library over time.

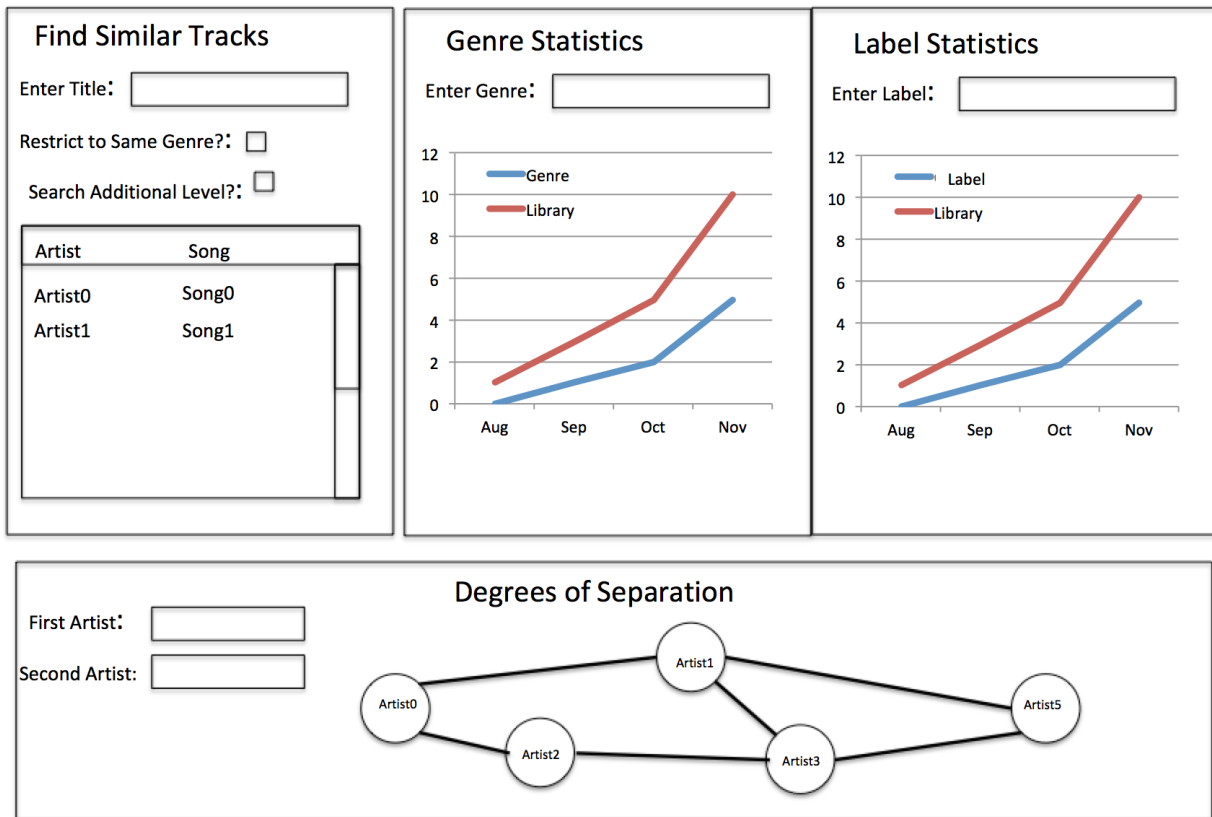
3. **Label statistics:**

This will function the same way as what is described above for genre statistics options a & b, except it will summarize based on the record label that released each track.

4. **Find degrees of separation between artists:**

A free text field will allow the user to search for 2 artist names. Depending on capabilities of the GUI, it will completion-match or search for artist names that start with what the user entered to help them pick one that exists in the database. Using a graph of artists who have collaborated with each other, the program will display both the shortest and longest paths of separation between the 2 artists. All ties will be displayed for both shortest and longest paths.

Music Library Analytics



Data: Users will be able to export their music collection to a .tsv file, which is an option in some music management software, like Rekordbox. Then they will import that file into the Music Library Analytics application. Our application will expect the .csv file to contain common music data for each file, such as Artist, Title, Genre, Date Added, and Label. The following structures will be used:

- All tracks will be stored in a hash table, where the hash key is generated based on a concatenated string of Artist, Track, Album, and Date Added.
- Each searchable field in the GUI (Title, Genre, Label) will have a B-tree which facilitates that searching. The tree will hold keys which can then be used to look up the objects for each track in the hash table.
- All artists will also be placed in a graph data structure, where the edges represent having collaborated on at least one track with the connected artist. Collaboration means either being credited as co-artists together on a track, or one of the artists remixing one of the other's tracks. This facilitates the similar track suggestions and the degrees of separation between artists feature.

Input Data File Format:

The input file will be a .tsv file with a row for each track and at least the following columns. Represented as a table below instead of as a .tsv for ease of reading.

#	Track Title	Artist	Genre	Date Added	Label	Remixer
1	The Melody (Original Mix)	Black Loops	Deep House	2020-04-04	Freerange Records	
2	Blue Pill (Original Mix)	Black Loops	Deep House	2020-04-04	Freerange Records	
3	Moving Light feat. Lisa Shaw (Deep Feels Dub)	Miguel Migs	Deep House	2020-04-04	Salted Music	
4	Vibin' (Saison Remix)	N-You-Up	Deep House	2020-04-04	No Fuss Records	Saison
5	What I Say (Extended Mix)	Dillon Nathaniel	Tech House	2020-04-04	SOLOTOKO	
6	Party People (Original Mix)	Caravaca	Tech House	2020-04-04	Selected.	
7	I'm Free To Love You (Extended Mix)	Monki, DJ Rae	Progressive House	2020-04-04	Toolroom	
8	I Like To Cha Cha (Original Mix)	GruuvElement's, Gianluca Rattalino	Tech House	2020-04-04	Safe Underground	
9	Strength (The Shapeshifters Club Remix)	Carolyn Harding, Freeform Five	Disco House	2020-04-04	Glitterbox Recordings	The Shapeshifters

Output Example:

An example of the output after requesting a graph of country songs added over time:

Search: Country

Title	Artist	Features	Album	Genre	Length	Date Added
Fancy Like	Walker Hayes		Country Stuff	Country	2:42	11/1/2021
Old Town Road - Remix	Lil Nas X	Billy Ray Cyrus	7 EP	Country	2:37	11/5/2020

Statistics:

Total Lil Nas X Songs: 2

Total Country Songs: 2



Related Artists:

[Billy Rae Cyrus](#)

[Doja Cat](#)

[Megan Thee Stallion](#)

[Miley Cyrus](#)

The degrees of separation will output as text like this:

Shortest path between artist A and B:

Artist A -> Artist B

Longest path between artist A and B:

Artist A -> Artist C -> Artist D -> Artist B