UNIVERSITY OF
EASTERN FINLAND

University of Eastern Finland

School of Computing

Master's Thesis

Web-Based Music Player for Music Performance Analysis

Fakrul M.A. Bhuiyan

30th September 2021

Abstract

Music and emotions have had an in-depth connection throughout history. People from different cultures perform and listen to music in unique ways to entertain, perform work, and relieve stress. The development of automatic music emotion recognition (MER) significantly improves the usability of online music streaming services such as Spotify, Apple Music, YouTube Music, and Amazon Music. Music emotion recognition aids in predicting music's affective content for listeners by applying machine learning and AI techniques. MER is efficient in understanding music, retrieval of music, and other music-related applications. An emotion-based music player detects the feelings of the listeners and helps to generate a playlist automatically. This thesis introduces EF Music, a web-based app developed with an Emotify dataset of 400 musical clips annotated with induced emotion from four genres: classical, rock, pop, and electronic. The implementation goal of EF music is to collect and preserve participant's feelings for each song to analyze the emotions similarities and connections. The investigation evaluates songs rated by Likert -scale to indicate how participant's feelings induce using emotions such as happy, sad, amusing, annoying, relaxing, dreamy, energizing, joyful, and neutral. For each sentiment, the influence on the participant's emotional experience is analyzed using a t-test and clustering to examine the correlation between the feelings.

Keywords: Music, Emotions, emotional annotation, Emotion Categorization, Music Performance.

Acknowledgments

I want to give my sincere thanks to the University of Eastern Finland (UEF) for accepting me to the IMPIT program. I would also like to thank all the professors who helped me gain knowledge in different areas of computer science. I have learned a lot during my studies at UEF.

I would also like to give my deepest gratitude to my supervisor, Professor Pasi Fränti, for his guidance during my study. His observation of my research and IT project with the thesis has allowed me to move forward with my skills and knowledge in computer science. I would also like to express my deepest gratitude to Dr. Radu Mariescu-Istodor, who has answered many technical questions during my master's study. His friendly support helps me to complete my IT project work.

I also give thanks to Abigail Wiafe, a Ph.D. student at the University of Eastern Finland. I also want to thank Oili Kohonen. She supported me a lot in completing my studies at UEF.

I would also like to express my deepest gratitude to my brother Tanvir Ahmed for his moral support and encouragement. I am grateful to my parents for everything. But my mother's dream was to get a master's degree from my childhood. I tried to make her dream come true. I am eternally grateful to my family and my life partner Farzana Haque.

List of Abbreviations

MIR   Music Information Retrieval

MER   Music Emotion Recognition

UEF   University of Eastern Finland

Table of Contents

List of Tables

List of Figures

1 Introduction

People are directly and indirectly dependent on music because music helps to boost emotions, release stress, helps sleep better, and serves as a powerful communication tool (Sharma, 2013). With the advancement of technology, people can access music using apps and over the internet in seconds.

Arjmand (2017) suggests that listening to pleasant music increases activities in the brain regions which is associated with emotion thus, music serves as a powerful strategy for eliciting emotions. (Meyer, 1956; Juslin & Sloboda, 2010), Emotions can be measured using physiological factors such as heart rate, transit time and amplitude, blood pressure, respiration, skin conductance, and skin temperature as their subjective indices of mood listen to in real-time. Krumhansl (1997), music generates a distinctive set of emotions in people based on the music style, environmental and physiological needs (Davis & Thaut, 1989). Scherer & Coutinho (2013) suggests that aesthetic emotions evoked by music result in novelty and complexity rather than direct importance to one's survival.

 Bakan, (2007) defines music as the humanly ordered sound that expresses feelings that includes and Sturm (2013) describes music composition features as performance parameters of music, including tempo dynamics, timbre. That it has a significant impact on the listener's perception (Clarke, 2002).

Over the last two decades, Music Information Retrieval (MIR) (Benetos et al., 2013) community has made significant breakthroughs in audio research, to significantly increase the volume of empirical data (Benetos et al., 2013) with most research analyzing performance recordings using a descriptive method which extracts musical features such as tempo curve (Palmer, 1989; Povel, 1977; Repp, 1990) or the loudness curve (Repp, 1998; Seashore, 1938), to gain general knowledge about performances or attributes that exist based on trends observed in the extracted data.

With the current quantities of musical databases, automatic music classification and similarity assessment methods are critical (Aljanaki et al. 2016). Emotion-based methods could be one of the most helpful access mechanisms for music collections. Implementing such procedures is not an easy process because of the limits of MER (music emotion recognition) and because the

emotional content of a musical work is inherently ambiguous. Within the realm of music-related emotions, there is a distinction between emotions expressed by music (but not necessarily felt by the listener) and emotions felt by the listener in reaction to music (which we refer to as induced emotions). There's no denying that music may elicit intense feelings in listeners (Krumhansl, 1997; Rickard, 2004). Many people use music for emotional self-regulation and music therapy (Gabrielsson, 2011), therefore developing algorithms that can automatically categorize and pick music based on these factors is critical. (Aljanaki, et al. 2016) contribute to the solution. In this study, we use (Aljanaki et al. 2016) dataset and implement a music player for observing the music emotion.

The link between expressed and produced emotion is not always straightforward. Expressed and induced emotion might link in four ways, according to Gabrielsson (2002): positive, negative, no systematic relation, or no relation; consequently, sentiment should not necessarily assume a positive relationship. Emotion induction is rare, even though a trained listener can almost always discern emotion presented in music. Juslin & Laukka (2004) says that Listeners experience significant emotions just around 55% of the time when listening to music, whereas 65 percent of musical episodes, music has an impact on how they feel (Juslin, Liljestrom, Vastfjall, Barradas, & Silva, 2008). Self-report, expressive behavior, and physiological reactions (heart rate, skin conductivity, blood pressure, and biochemical responses) can all be used to quantify emotional responses (Krumhansl, 1997; Rickard, 2004). In music, overt expressive action is not the norm. Self-reporting is perhaps the most extensively used and most valuable metric because it offers information on emotion's otherwise inaccessible cognitive component (Zentner & Eerola, 2011). Aljanaki et al. (2016) employ self-report to quantify induced emotional responses to music.

1.1 Problem Statement

This study finds numerous investigations that use composition features as performance parameters to extract musical performance. This study sees several works on affective music composition or music emotion prediction using machine learning algorithms, where the researchers focused on a predefined data set to predict. Xu et al. (2020) use a 60-excerpt dataset to predict emotion using machine learning. The research uses only three emotions (happy, sad, and relaxing), making it less appropriate to extract participants' feelings. This thesis surveys 181 participants with ten emotional

annotations, which helps to extract more accurate emotions. Another researcher, Aljanaki et al. (2016), collects emotional survey data from the participants through a music game. The music game has several questionnaires and rules, which makes the data collection process slow.

This thesis can say that this study needs a music player with a user-friendly interface to collect data fast for defining categorical emotional annotations. After collecting the data, this survey evaluates emotional annotations and ratings.

1.2 Motivation

The motivation to design and implement an app that aids in music listening and rating feelings is that most existing apps are complex, with songs lasting for more than a minute. Thus, listeners find it difficult to assess their feeling and sort music randomly. Therefore, the study adopts the Emotify dataset to design a simple and less complex music app capable of storing listener feelings after hearing a song.

1.3 Research Aims and Objectives

The research seeks to develop an automated web-based music player to analyze music performance - Listener's emotional responses are generated after this to a piece of music the particular piece/the whole song.

The objective of this research is to design and implement an automated web-based music player which meets the following requirement:

To generate the listener's emotional responses after hearing a song

To measure the effect of annotation button position.

To measure feelings relations.

To measure the song's association between ratings and emotional responses.

To calculate the time that takes a listener to provide emotional responses.

1.4 Contributions

This thesis introduces a user-friendly music performance extraction technique. The implemented system allows each user to listen to the music of their choice and select a category, such as melancholy, laughing, or any other emotion. They will be able to see how much time each person devotes to each song. If something goes wrong, users will be able to rate each music several times at any time, as well as modify their sensibility. By recording this data, this investigation hoped to build a solid system for music emotion identification. This program uses an emotional annotations dataset containing 400 song snippets (1 minute each) from four different genres (rock, classical, pop, and electronic). This thesis develops an EF music player to collect annotations for 400 musical excerpts. The annotations are publicly available. This thesis examines the survey data of the selected categorical emotional annotations to find the relation of different factors. The investigation evaluates the annotation button position to justify the data collection process is independent. This study also examines the annotations with participant's ratings.

1.5 Organization of the Thesis

The remainder of this paper is structured as follows. The following Section 2 presents the seven music components with examples. The explanation of music features is necessary because as this study is working with music, so music has a connection with different emotions. A discussion on the system architecture and implementation challenges appeared in Section 3. The subsequent sections focus on studies in music compositions and performance parameters for emotion extraction and relating them to either the music composition (Section 4) or the music emotion (Section 5). Section 6 describes the models for musical emotions where it explains several models for evaluating musical emotions. Section 7 deals with the assessment of the collected data and summarizes the findings from the analysis. The analysis extracts the correlation between music components and emotions. Section 7 concludes the investigation with the work to come.

2 Features of Music

Music is the art of organizing sounds to generate a composition using melody, harmony, rhythm, tempo, timbre. Those musical elements have influenced the emotions of listeners in a variety of ways. These have been defined as music performance metrics for music compositions by (Sturm, 2013). According to (Rink, 2002). Performance characteristics can have a significant impact on how a listener perceives music. Levitin (1999) claimed seven features: timbre, dynamic, form, rhythm, tempo, melody and harmony, essential for music composition. This study defines those features in the below description.

2.1 Melody

Melody is a musical phrase made up of a linear succession of musical notes or pitches that is agreeable to the ear. The melody's movement between successive notes or pitches creates a melodic contour that allows two melodies to be distinguished. Melody has several characteristics: pitch, interval, range, shape, phrase, cadence, and countermelody. Those characteristics are involved in emotion generating in a variety of ways. Melody in music emphasizes communication to listeners and aids in learning by speeding up memory and recall (Wallace, 1994). The form notes or pitches performed in a slow or rapid, jumpy or flowing manner can transmit happiness or sadness, and Schellenberg et al. (2000) investigate the impact of rhythms and pitches altered in melodies with the emotions observed.

Different forms of music and musicians use the melody in various patterns. The primary melody term of jazz musicians is "lead" or "head," and it is used as a beginning point for improvisation. Rock music and other types of popular and folk music tend to focus on one or two melodies. Because there are no chord changes in Indian classical music, it focuses on melody and rhythm rather than harmony. Composers of Western classical music frequently starts with a melody or subject and then add variations. Figure 1 shows the "Pop Goes the Weasel" melody.

Figure 1: The "Pop Goes the Weasel" Melody (Kliewer, 1975)

2.2 Rhythm

In music, rhythm is the manifestation of sound across time. Schulkind, (1999) defines rhythm as "the serial pattern of varied note durations in a song." Rhythm is a vital element in music because it can exist without a melody (as in drumbeat), but a melody cannot live without rhythm as the rhythm provides the pattern in the music. Listening to a rhythm in a piece of music sometimes results in the movement of the body. Rhythm features beat, accent, tempo, measure, meter, upbeat, downbeat, syncopation, nonmetric. According to (Balkwill & Thompson, 1999), rhythm helps extract happiness, anger, fear emotional responses and the appropriate instrument to implement the effect is the flute.

Holland (1994) defines the uses of harmony. One of the most significant parts of hip-hop music is the rhythmic delivery of the lyrics. The structure of the entire rhythmic pattern was constructed previously in Indian classical music. Western composers typically use odd meters and techniques like phasing and additive rhythm to create more rhythmically complex music. Figure 2 depicts a typical rhythmic pattern in western music.

Figure 2: A Proportional Representation of Rhythm as is Common in Western Music Notation (Honing, 2013)

2.3 Tempo

BBC Bitesize (2021) described tempo as a musical term that refers to the speed at which one should perform music. Tempo has many features such as. grave, largo, adagio, andante, allegro, vivace. According to (Juslin & Madison, 1999; Gagnon & Peretz, 2003), tempo musical features involve several emotional responses like happiness, sadness, anger.

In general, different countries have tempo-making formats. This study finds several articles where (Apel, 1969) provides examples of (German), France[1], and (Duke, 1989) explains English tempo-making formats. However, the common tempo patterns beat per minute (bpm) values are actual rough estimates for 4/4 time.

Figure **3** shows a speed meter to indicate tempo because tempo always represents the speed to perform a piece of music.

Figure 3: Tempo (Speed: How Fast or Slow)[2]

---

[1] https://theonlinemetronome.com/
[2] https://www.pngkey.com/detail/u2w7a9q8q8i1e6u2_speedometer-0shares-car-speed-meter-vector/

2.4 Harmony

Harmony is the sound of two or more notes played simultaneously in music. After hearing a piece of music, listeners can analyze the composition of individual sounds. Harmony's main characteristics are chord, scale, tonality, and texture (monophonic, heterophonic, homophonic, polyphonic). Harmony can extract happiness, sadness (nostalgia/longing), tenderness, and tension emotional responses using piano and strings claimed by (Lahdelma & Eerola, 2016).

Parallel perfect intervals are standard in Western religious music, and these intervals would preserve the clarity of the original plainsong. The English form has a sweeter tone and to be more suited to polyphony since it allows for more linear freedom in part-writing. Figure 4: Example of Implied Harmonies in J.S. Bach's Cello Suite no. 1 in G, BWV 1007, Bars 1–2 (Cole, 2019) depicts an example of inferred harmonies.

Figure 4: Example of Implied Harmonies in J.S. Bach's Cello Suite no. 1 in G, BWV 1007, Bars 1–2 (Cole, 2019)

2.5 Timbre

Timbre refers to the distinct aspects of sound produced by a musical instrument or the human voice when a piece of music is performed and can also refer to tonal color. Listeners can distinguish between two sounds with the same pitch, timing, and strength using timbre view as an attribute to auditory objects (Griffiths & Warren, 2004). Timbre can identify and analyze using harmonic content, vibrato, attack, decay, sustain and release (ADSR) envelopes. The harmonic content of a sound primarily determines the waveform of the sound signal changes over time. Timbre in the music domain is a significant factor that contributes to emotional expression in music (Eerola et

al., 2013), which impacts a listener. Liu et al. (2018) studied timbre features that stimulate emotions and concluded that the timbre of a simple and isolated musical instrument could convey emotions like emotional speech prosody. The emotional responses of happiness, sadness, anger, fear, tenderness produces by timbre claimed by (Hailstone et al. 2009; Liu et al. 2018)

Figure 5 shows a descending chromatic scale that encompasses a wide range of orchestral timbres. The woodwinds (flute and oboe) come first, followed by the massed sound of strings with the melody carried by the violins, and finally the brass (French horns).

Figure 5: Wagner Sleep Music from Act 3 of Die Walküre (Baileyshea, 2007)

2.6 Dynamic

Dynamic is the loudness or softness of a piece of music. By not making the sound dull, it demonstrates the shift in a bit of music. The importance of dynamics in music events and how listeners perceive it was described (Geringer & Breen, 1975; Krumhansl & Castellano, 1983). Dynamic features are crescendo, decrescendo, sforzando. Through dynamic musical features Happiness, Fear, Sadness, Anger emotional responses can extract described by (Juslin & Madison, 1999; Laukka & Gabrielsson, 2000)

Composers employ crescendo and diminuendo to shift the dynamics (also decrescendo) progressively. crescendo (cresc.): progressively increase the volume (<). diminuendo / decrescendo (dim. or decres.): play softer (>) over time. Below Figure 7 shows the changes in dynamics in music composition.

Figure 7: Dynamic Changes[4]

2.7 Form

Form refers to the framework of a musical composition or performance. Form features are repetition, contrast, and variation. The most common musical forms are Sectional, Binary, Rondo, Variational, Sonata-allegro.

The phrase "Binary Form" refers to a musical piece divided into two portions of roughly equal duration. AB or AABB are two ways to write Binary Form. Using Greensleeves as an example, the first system is nearly identical to the second system. The piece of music is in binary Form throughout AA'BB'. Figure 8 shows examples of binary form.

Figure 8: "Greensleeves" as an Example of Binary Form[5]

---

[3] https://sttps://stock.adobe.com/images/music-volume-sign-icon-vector-illustration/182558696
[4] https://courses.lumenlearning.com/musicappreciation_with_theory/chapter/dynamics-and-dynamics-changes/
[5] https://thereaderwiki.com/en/Form_(music)

3 Survey on Music Composition and Emotion

In this section, we discuss the previous literature of automatic music composition and models for music emotions.

3.1 Musical Features and Algorithmic Composition

Feelings and preferences are subjective when it comes to music. Every piece of music produced has its distinct rhythm and genres derived from its traditional background, way of life, and history. The most substantial reason for listening to music is emotional response (Panksepp,1995). This thesis finds numerous articles (Juslin & Madison, 1999; Gagnon & Peretz, 2003; Hunter et al., 2008; Laukka & Gabrielsson, 2000; Juslin & Laukka, 2000), exploring tempo musical features. The objective is to discover various emotional responses (Happy, sad, joyful, annoying, anxious, dreamy, energizing, relaxing, neutral, amusing). Juslin & Madison, (1999) investigation found that Piano, Electric Guitar, Electronic Drum, Flute are the best instruments for implementing the effect. (Juslin & Madison, 1999; Laukka & Gabrielsson, 2000) Also worked on dynamic musical features for extracting emotional responses. The way notes or pitches performed in a slow or rapid, jumpy or flowing manner can transmit happiness or sadness, and Schellenberg et al. (2000) investigate the impact of rhythms and pitches altered in melodies with the emotions observed.

Liu et al. (2018) investigated emotional timbre features and found that the timbre of simple, solitary musical instruments can express feelings like emotional speech prosody. This research also explored music composition algorithms which found several affective composition algorithms. Tokui & Iba (2000) used IEC to characterize the "CONGA" innovative approach to music creation, namely the composition of rhythms. The method's key feature is that it incorporates genetic algorithms (GA) and genetic programming (GP).

3.2 Models for Music Emotion

Music emotion recognition systems apply categorical and dimensional approaches as the two dominant models for labeling emotions in music and emotion research (Juslin & Sloboda, 2010; Zentner & Eerola, 2010). The categorical approach is characterized by discrete labels or words including happy, sad, fear and anger (Feng et al., 2003) that humans can easily perceive while the dimensional approach presents emotions in a dimensional space. The categorical model applies Ekman (1992) basic six emotions such as anger, fear, happy, sad, surprise, and disgust, or some domain (dependent) expressive terms are used to express emotional reaction or episode in the categorical model. Emotional states express are grouped into different categories owing to situational, cultural, or character differences. Thus, there is a challenge in grouping emotions into one classification. Besides, Sreeja & Mahalakshmi (2017) suggest that categorical models may not signify the various emotional states even though sets of emotional categories are defined. A factual representation of emotions is not displayed though subjects cannot select appropriate terms as part of the label.

The Dimensional approach applies quantifiable values on dimensions to describe an emotion. An example of the dimensional model is discussed by (Russell, 1980) and (Thayer, 1989). Russell, (1980) proposed the two-dimensional model: thus, valence (representing the positive and negative degree of emotions) and arousal (representing the intensity of the emotion) are represented on a two – dimensional plane with the x-axis representing valence and the y-axis representing arousal. It is well-known and frequently mentioned in cognitive science and psychology (Kuppens et al., 2017; Olson et al., 1989).

Thayer, (1989) proposed the Thayer's model adapted from the Russell circumplex model, implements two measures thus energy (referring to the volume or intensity of sound in music) and stress (representing the tonality and tempo of the music) corresponding to arousal and valence in the circumplex model. Based on the stress and energy, the mood in music is partitioned into four clusters thus: calm-energy, calm-tiredness, tense-energy, and tense-tiredness.

The two-dimensional models present an advantage of lessening uncertainty as a compared categorical model since a consistent measure is provided for measuring emotions in two well-defined dimensions.

4 System Design and Methodology

4.1 System Architecture

The system architecture of EF music, which is a music emotion tracking player, is discussed here. Figure 9 demonstrates the architecture of EF music player. The music player starts working with the command of the listener, where the listener gives some choice to listen to a song. At first, it creates its work by getting the live server. After getting a live server, the listener can send a request to the song list viewer to extract all possible songs by name/category/genre. The listener can make their choice from the 400 songs list. When the listeners fix their choice, the system sends the listener selection into the music auto player for playing the song. After finishing a song, the listener can provide their emotion-based feedback and rate the song. However, the listener must have to go through the registration process to enter the song review system. The music review system works for tracking the emotion of a which can change over time. The review analyzer works as an interpreter to analyze the reviews. The system has a database for storing the finalized reviews with the user data statistics. The problems in each of our components are discussed in detail below.
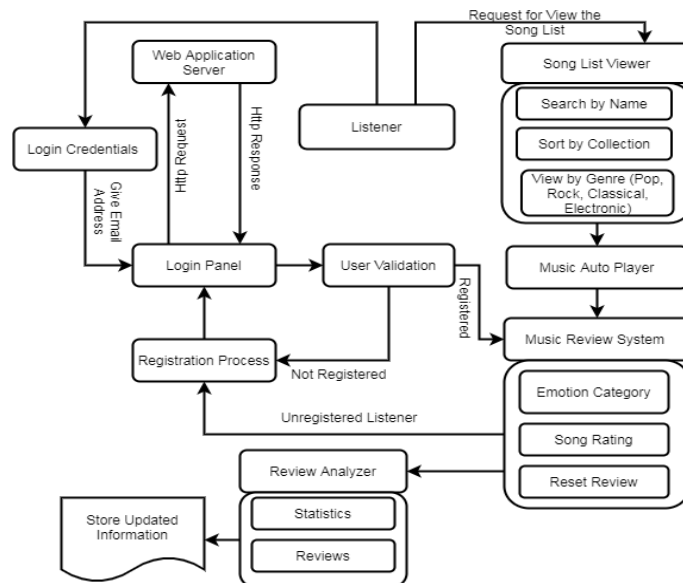
Figure 9: System Architecture of EF Music

13

4.1.1. User Validator

The above description of this research already discussed that anyone could enter the system as a general listener. The system ensures limited access to the general listener where the listener can only listen to a song. After listening to a song, the listener must register when they want to give their feedback about the music. The registration process is simple, where the listener has to provide their full name and mail address. The system has a verification process where the design matches the given mail address of the newly registered user. Suppose the mail address does not correspond with the provided email address shown in the registration process the system grants the user request. Then the system takes the user as a new user.

If the mail address matches the existing user, the system rejects the registration request. When a user completes their registration process, the new user has the login credentials to enter the system. The registered user gives their credentials to the login panel, where the login panel sends the information to the web server through an HTTP request. The web server checks the information validation. Then the server sends the validation information through an HTTP response. If the user validation ensures the user is valid, then the system allows entering the feedback option.

4.1.2. Song List Viewer

A user-friendly song list viewer panel is developed to aid the listener's sort the song list. A listener can find out the desired song by a simple search on the system and change the number of entities (songs) showing. The listener can fix the entity number as 10/50/100. The listeners can sort their playlist based on the music which is rated/played by them. The listeners also have the facility to sort the list based on the song that is not rated/ played. The song list viewer system has several song genres like pop, rock, classic, and electronic. The listener can sort their song by their favorite genre. The song list viewer also has other sorting systems like sort based on the collection (all, emotify), name, and song id number. The selected music for playing will be displayed in a grey hue on the music list so that the user can easily see that it is now playing. On the music list, rated music will be shown in a dark grayish cyan color so that the user can immediately see that they have previously rated the music.

4.1.3. Music Player

We implemented a music player where the system already has a song database. The user can easily select the song from there. Music streaming works like this: a streaming service sends tiny bits of data to a streamer, allowing the user to get pre-buffered music buffered a few minutes or even seconds before playing a song. If the user has a good internet connection, streaming technology delivers a continuous listening experience. The music player is straightforward to use. The player has the traditional features for playing a song. The player has an auto-play option that helps the user to play songs automatically one by one. The system has just play /pause/next option for controlling the song play, which is very simple to handle.

4.1.4. Music Review System

After hearing a song, a listener can provide their feelings through the music review system. The music review system has several emotion-based feeling options, i.e., happy, sad, amusing, annoying, anxious, relaxing, dreamy, energizing, joyful, neutral. A listener can select multiple options at a time to express their feelings. However, if the listener chooses neutral, then the other selections are removed immediately. The listener can also change their review later; the previous examination will be deducted from the review system. The listeners can also express the level of their feelings through the review system. They have to give the feelings class 1-5 by hovering over or clicking on the star rating. The developed system has a review reset option where the listeners can change their feelings. The procedure permanently stores the updated reviews.

4.1.5. Review Analyzer

The implemented system has a review analyzing feature divided into two parts review and user review statistics. Each music rating is described in detail in the review section. Everything about each song rating, including who rated it, their feelings, and how many stars (ratings) they gave it, as well as how long they listened to it, including song name and timestamp. The user statistics list displays all of the interaction's history of the listeners. How many songs did each person listen to, as well as how long they listened to them? Furthermore, this page contains a review count, which

means that a single person rating counter will be present so that anybody can see which impression has the most significant impact on the person.

4.2 Data Collection

According to Bhat (2021), data collection is the process of gathering, measuring, and analyzing precise insights for research using established approved procedures. Based on the facts gathered, we will evaluate the listener's feedback. Our research is survey-based quantitative research. Quantitative research, according to Bhandari (2021), is the process of collecting and interpreting numerical data. It can be used to look for patterns and averages, make predictions, test causal linkages, and extrapolate results to larger groups. This research uses a 400 songs dataset from Emotify to examine the listener's feelings after listening to a song. After listening to music, the registered users can express their feelings through a song survey panel. The system stores the provided information in a database where the listener can see their activity statistics. This further investigation analysis was held to define the feelings pattern of several listeners after listening to a song.

4.3 Data Preprocessing

Data pre-processing is a technique that helps to transform raw data into a valuable and efficient format. The significance of data pre-processing in creating a good and legitimate data set cannot be overstated. If the data is not vetted or inspected properly, the collected raw data is incomplete, inconsistent, and tempered. After acquiring the listener's reviews from the review system output, several approaches were used to validate the feelings present in a song. In the survey, this thesis recorded 3030 responses from 181 participants. For making data meaningful, this thesis applies several steps of information pre-processing. In the raw dataset, this thesis observes that some data repeats and some fields are missing, which have been fixed manually. This process is called data cleaning. The process helps to identify male and female song listening percentages for genres. The method also helps to justify emotions, rating growth, and time distributions. Then conducts data normalization where convert data 0.0 to 1.0. This process helps to perform data mining. Through the process, this thesis analyzes button position t-test, correlations, and associations.

16

4.3.1 Data

The following is a list of the information discovered in the file:

- Ten annotations by the participant
- Id of the music file
- Genre of the music file (whether emotion was strongly felt for this song or not). Experienced emotion indicates One indicates.
- The mood of the participant before playing music in the music player.

The section discusses the data collection methods applied in the study. The list of emotional annotations with adjective markers shows below in Table 1.

Table 1: Emotional Annotations with Adjective Markers as Used in EF Music Player.

Factors	Adjective Markers
Happy	radiant, elated, content
Sad	Sorrowful, depressed, sad
Amusing	admiring, fascinated, impressed, goose bumps, thrills
Annoying	Irritating, Madding
Anxious	nervous, revolted, tense, Dysphoria
Relaxing	Tranquility, Soothed, calm, in peace, meditative, serene
Dreamy	Tender Longing, Affectionate, softened up, melancholic, nostalgic, sentimental
Energizing	Activation, Disinhibited, excited, active, agitated, fiery
Joyful	Feels like dancing, bouncy feeling, animated
Neutral	No emotion, indifferent

4.3.2 Musical Database

This study uses the Emotify dataset (Aljanaki et al. 2016) in the experiment to establish ecological validity, which lasts for a minute. The music chosen was randomly selected from a larger music collection. The research put together a bundle of 400 musical pieces from Magna Tune (magnatune.com), 100 from each of four genres (classical, rock, pop, and electronic). The

recording firm assigned genres to the songs. The dataset includes music from 241 distinct albums by 140 different artists in the resultant dataset. The dataset chose the Magna melody for various reasons: it is high quality and relatively unknown (familiar music may preload generated feeling) (Schubert, 2007). The study also manually checked the music and eliminated some recordings (about 2%) due to poor quality.

4.3.3 Rating scale

This study uses the Likert scale for getting user opinions about a song. When responding to a question on a Likert scale, participants must choose from a list of star rating options which means " Extremely poor " to " Excellent " scale. The rating scales will be placed beneath each emotion annotation, and the music player collects the ratings for each emotional annotation. The scale's mean values are as follows: (1=extremely poor, 2=bad, 3=average, 4=good, 5=Excellent). This rating scale has collected the impacts on the participant's emotion after listening to a song.

5 Survey of Music Emotions

Emotions felt and perceived when listening to music are distinct from those felt and perceived in daily life. Music is a suitable medium for communication. This thesis develops an EF music player, which helps listeners provide their feelings after hearing a song. **Figure 10** provides the front-page information of the EF music player. This page contains the Music section, Music player section, Registration section, and others, including the about and review section. In the music section, it will be sort music by clicking every title, which helps users get their favorite genre quickly. The users can also search by specific name if they want, and the search is in real-time so that users don't need to wait a moment for searching.
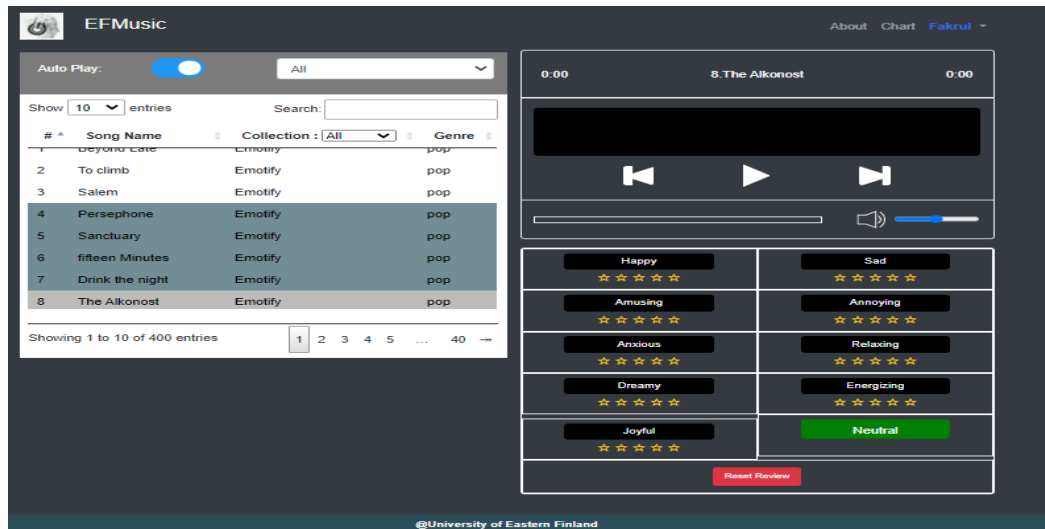
Figure 10: Landing Page of EF Music

This section contains details about each music rating. Figure 11 represents the EF music user review system. The review system shows the information about each song rating, like who rated the song, their feelings and how many ratings(star) they placed for the music, and how long they played that song, including song name and listened to timestamp.

Figure 11: Review Page of EF Music

The developed EF music can analyze the user emotions and other parameters through user review statistics. The user statistics list in Figure 12 shows every user interaction over the app. How many songs are listened to by any individual, as well as how long have they heard? Moreover, this page contains a review count, which means a single person rating counter will be here to know which impression affects the person most.

Figure 12: User Review Statistics of EF Music

20

6 Analysis of the Survey Data

The dataset comprises 400 music snippets (1 minute long) in 4 genres (rock, classical, pop, electronic). The survey data were collected using 10 emotional annotations. The participant can rate the song on a 1 to 5 rating scale. The participant could also select multiple feelings at a time. When a participant selects neutral, however, all other emotions are automatically deleted. The emotional categories with adjective markers describe below, as found in the music player. Participants were allowed to skip tracks and change genres, and they were encouraged to do so because an induced emotional reaction does not always occur when listening to music. As a result, less popular genres (among our sample of participants) received fewer annotations, as did less popular songs.

6.1. Participants and Annotations

In this study, 181 people participated in the survey, with 3030 responses collected throughout listening sessions. The following were the stylistic preferences: rock 24, classical is 15 percent, pop is 41 percent, and electronic is 20 percent (multiple genres were allowed). Table 2 shows emotion distribution among ten emotion categories.

Table 2: Frequency Analysis for the Existence of Emotions Among Ten Emotion Categories

Emotion	Frequency	Percentage
Happy	345	11%
Sad	263	9%
Amusing	374	12%
Annoying	281	9%
Anxious	305	10%
Relaxing	405	13%
Dreamy	292	10%
Energizing	320	10%
Joyful	253	8%
Neutral	218	7%

In Table 1 The frequency analysis for emotions in our sample (happy, sad, amusing, bothersome, anxious, relaxing, dreamy, energetic, joyful, and neutral) has shown. As shown in Table 2 below, Relaxing gets 13 percent of the review, while amusing and Happy get 12 percent and 11 percent, respectively. Only 7% of people chose neutral, which suggests they have no feelings. Dreamy, energizing, and nervous have a 10% frequency, while annoying, sad, and joyful have 9 percent, 9 percent, and 8% emotional frequency, respectively.

6.2 Participants, Professions and Country

The section discusses the gender, profession, and country distribution in terms of the rating provided by the participants. Figure 13 shows a domain-based analysis that focuses on gender, professions and countries. Through the study, this investigation reveals that most of the participants are from Ghana. The total number of Gahana participants is 158 where 125 males and 38 females. The investigation also shows that 136 participants are male, and 45 participants are female. This analysis reveals another point that participant are students by work. Of 181 participants 162 participants are students whereas 19 peoples are not student. For other country-based analysis, Finland has nine males and four females. This study also notices the participants of Bangladesh, which have two males and two females. The analysis of the profession clarifies that male participation is higher than female participation. For the student profession, male participants have 77% engagement whereas the female has 33% participation. The non-student domain shows male 12 and female 7 participants.
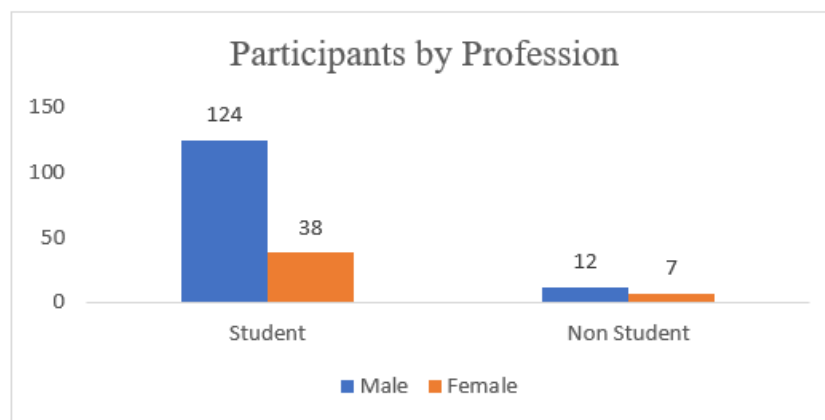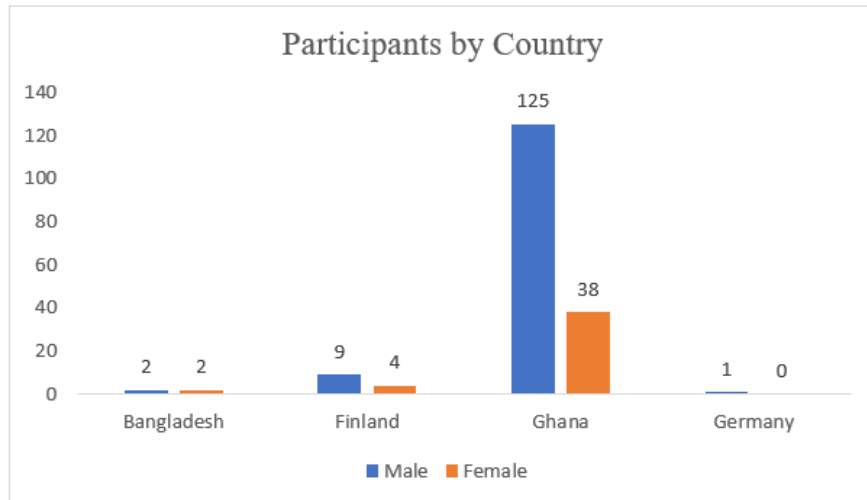
Figure 13: Participants, Profession and Country

6.3 Songs Listening Between Genders

Figure 14 shows the listening ratio of males and females for different genres. Through the analysis, this investigation find that most pop songs hear both male and female—the percentage of listening to a pop song is 38% male and 43% female. The least listening song is classic—the classic songs ratio for males is 15% and females 17%. However, if we focus the Figure 14, we can reveal that males prefer rock and electronic, whereas females choose pop and classic. The rock songs have a 25% male and 23% female listening ratio. This investigation finds a higher listening gap for electronic genre songs. Through the analysis, we discover that there is a 6% listening gap. The percentage of hearing electronic is 22% for males and 16% for females.
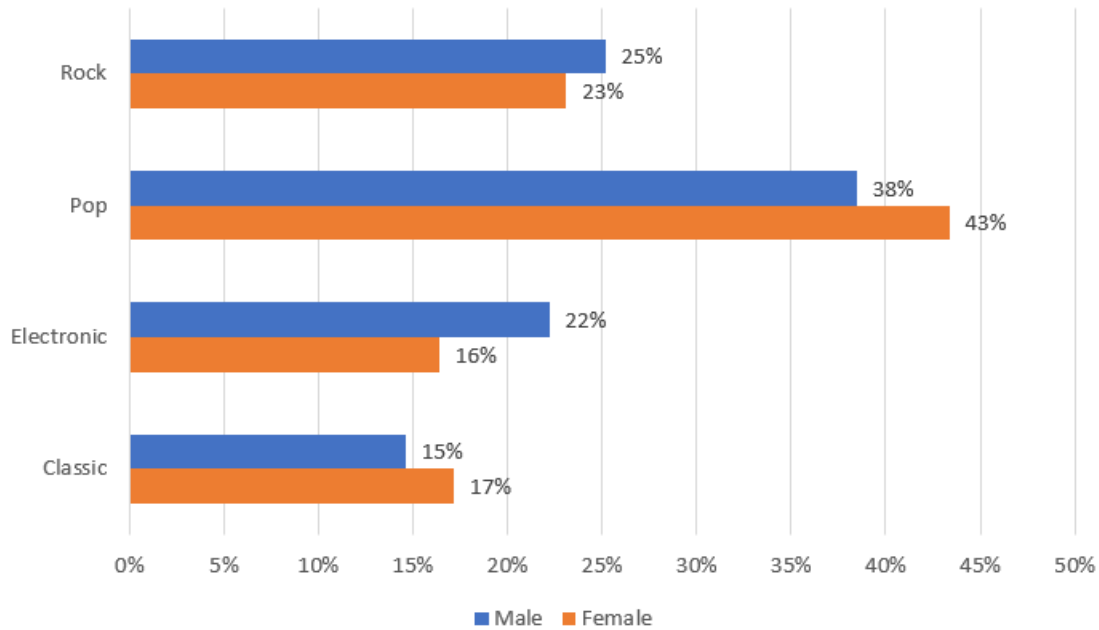
Figure 14: Song Listening Ratio between Genders

6.4 Emotion Distribution for Gender

In this investigation, we evaluate the emotion distribution between males and females. In another experiment, Figure 16 shows that for every genre relaxing selection percentage is higher. In this analysis, we also find that relaxing has a higher rate for both males and females whereas males 14% and 11%. The Figure 15 shows that female selects amusing annotations more compare to other annotations which are 12%. On the other side, males select relaxing annotations more than other annotations, which are 14%. So, we can say that males feel more relaxed after listening to a song, whereas females feel more amusing. Other annotations like dreamy, joyful, energizing, and sad have similar types of emotion distribution for both males and females. There are some differences in annoying, anxious, happy, and neutral annotation selection. This analysis finds it Annoying and anxious as similar distribution 9% male and 11% female whereas happy 12% male and 10% female and have neutral 8% male and 5% female. The analysis reveals that males are more neutral than females. The research also discovers female feels more annoying and anxious after hearing a song.
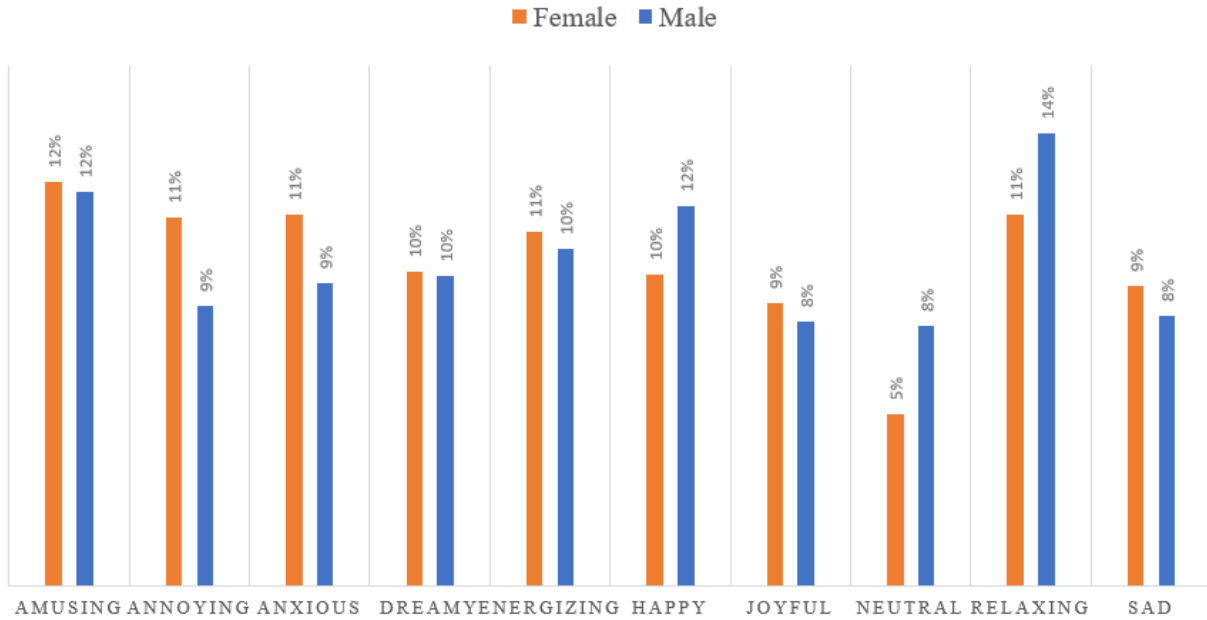
Figure 15: Emotion Distribution between Male and female

6.5 Influence of button order on the frequency of selection

Positions of the buttons (the Ten buttons with emotional labels on them) in the music player interface for each participant. Figure 10 depicts the location of the controls. We needed to see if the selected buttons were in specific areas more frequently than other buttons (regardless of the text on the button.). This investigation divides the emotional possibilities into three categories in this case: group 1 (happy, sad, amusing), group 2 (annoying, anxious, relaxing), and group 3 (dreamy, energizing, joyful). However, this study did not include neutral in any of the groups because neutral means no feelings. Table 3 provides an overview of the frequency selection.

Table 3: T-test for Button Position

Group 1 and Group 3		Group 2 and Group 3	
Group 1 Mean	326.67	Group 2 Mean	332
Group 3 Mean	288.33	Group 3 Mean	288.33
P- Value	0.381	P- Value	0.359

Participants choose from a variety of buttons during a single listening session. Buttons in the bottom were chosen less frequently in group 3 that in the first and second row.

The results of the paired Student's t-test have shown in Table 3. This analysis is seen, to have no discernible difference between the groups. This analysis did not find a less than p-value of 0.05 in this investigation. As a result, this investigation may conclude that button position has no bearing on the quality of annotation collections.

6.6 Influence of Genres on Emotional Annotations

This thesis uses the emotify dataset, which has four genres. Figure 16 shows an analysis of a relation between genres and emotional annotations. Through the experiment, we reveal an interesting fact that all genres have a standard connection with feelings. Happy, amusing, relaxing, and energizing have the most responses for pop, classical and electronic genres. In the case of pop song relaxing have 15% responses which is the highest percentage. The other happy, energizing, amusing and joyful annotations have 14%, 13%, 11%, and 11%, respectively. Classical and electronic happy have 22% and 16% responses, which is the highest percentage. However, this analysis finds an exception for the electronic genre. It contains 14% of neutral responses, which is higher among all genres for the tepid response. In the case of the rock genre, we discover a different scenario compared to the other three genres. For the rock genre, sad (14%), annoying (13%), anxious (15%), and relaxing (12%) have higher percentages. The other annotations like happy, amusing, energizing, and joyful have 8%, 12%, 7%, and 7%. Through the analysis, another general point discovers that relaxing emotional annotation always has a consistent percentage for all genres.
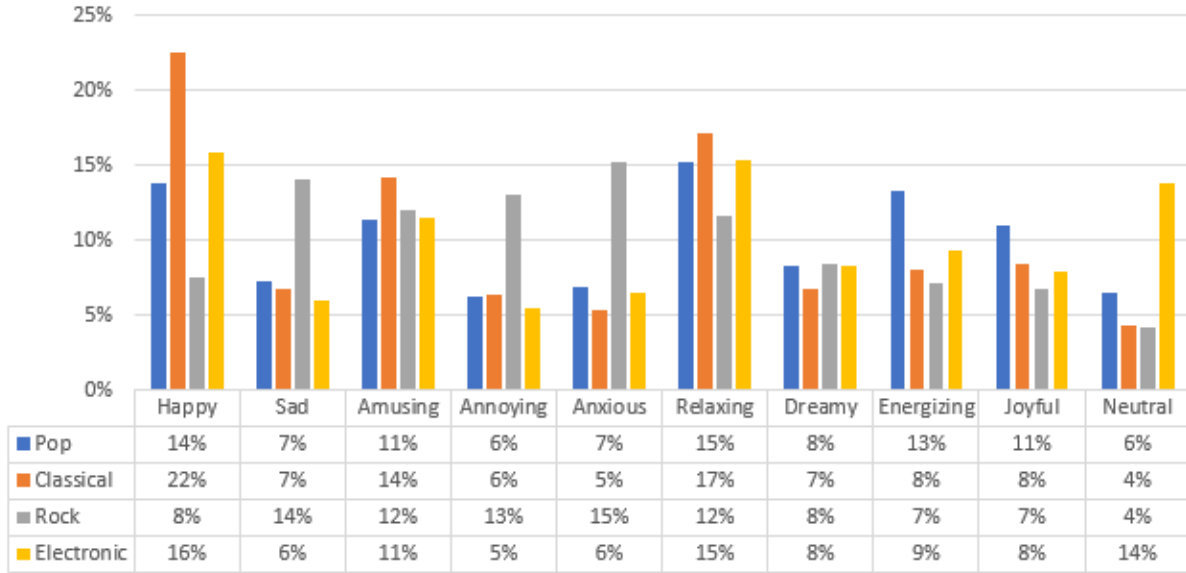
Figure 16: Ratio Analysis between Genre and Emotional Annotations

6.7 Correlation Analysis for Emotional Annotations

This thesis applied Pearson 2 tailed correlation (Benesty et al. 2009) to identify the relationship between emotional annotations. The strength of a relationship between two variables is measured using correlation coefficients. In statistics, the Pearson correlation is the most widely utilized. This metric assesses the strength and direction of a two-variable linear relationship. The values are always between -1 (strong negative relationship) and +1 (strong positive association) (strong positive relationship). A linear relationship is weak or non-existent if the values are at or near zero. The formula of Pearson correlation has given below.

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{1}$$

The formula descriptions are $\rho_{X,Y}$= Pearson product-moment correlation coefficient. $cov(x,y)$= covariance of variables *x and y*, σ_x= standard deviation of *x, and* σ_y= standard deviation of *y* used the formula for calculating correlations between emotional annotation categories (see Table 4). In the correlation analysis this thesis picks one answer for a single annotation and prepare a song annotation dataset. Strong positive correlations mean that the correlated categories were either often selected together (co-occurring emotions) or often chosen by different people for the same

music (confused and potentially redundant types). Table 4 shows correlation between emotional categories.

Table 4: Correlations Between Emotional Categories.

	Sad	Amusing	Annoying	Anxious	Relaxing	Dreamy	Energizing	Joyful	Neutral
Happy	-0.093	**0.445****	-0.030	0.052	**0.172****	**0.399****	**0.324****	**0.472****	**-0.150****
Sad		-0.014	**0.170****	**0.439****	**0.198****	0.048	-0.008	0.032	-0.095
Amusing			0.030	0.104*	0.109*	**0.381****	**0.312****	**0.482****	**-0.243****
Annoying				**0.171****	-0.064	-0.007	0.050	0.003	0.067
Anxious					0.092	0.086	0.055	0.065	-0.118*
Relaxing						**0.172****	0.052	**0.169****	**-0.158****
Dreamy							**0.334****	**0.408****	**-0.169****
Energizing								**0.291****	**-0.171****
Joyful									**-0.307****

Prominent examples are happy and amusing with r=0.445, happy and relaxing with r=0.172, happy and dreamy with r=0.324, happy and joyful with r=0.472, sad and annoying with r=0.170, sadness and anxious with r=0.439. There are so many other strong positive correlations such as amusing with dreamy, energizing, and joyful, relaxing with joyful, nostalgic with energizing, joyful, and energizing with joy. Strong correlation means (**) Correlation is significant at the 0.01 level (2-tailed), and just correlation means (*) Correlation is significant at the 0.05 level (2-tailed). Some solid negative correlations are happy and neutral with r=-0.150, relaxing and neutral with r=-0.243, dreamy and neutral with r=-0.169, energizing and neutral with r=-0.171, and joyful and neutral with r=-0.307.

6.8 Influence of personal factors on induced emotion

Personal and situational circumstances can significantly impact the emotion evoked by music (Dibben, 2004). In this section, we'll look at the extent to which various elements affect.

6.8.1 Influence of ratings

According to new research, people's perceptions of music change depending on their mood. In this thesis induced musical emotion study, expected to observe a rating effect. Chi-square test on category selection frequencies grouped by participants' ratings and discovered significant variations in the sadness, annoying, energizing, and neutral categories. The most apparent inclination is for unpleasant and invigorating. Table 5 shows that 5 ratings get a higher number of selections in every case. There is a definite trend that participants give more responses for higher ratings all of the time.

Table 5: Ratings and Frequency of Emotion Selection. Bold if *p*-value<0.05

Emotion	Participants Ratings					Chi Squire	*P*- value
	1	2	3	4	5		
Happy	14	28	67	88	148	7.271	0.201
Sad	**16**	**35**	**68**	**61**	**83**	**22.144**	**0.000**
Amusing	17	45	76	72	164	3.499	0.624
Annoying	**23**	**27**	**68**	**51**	**112**	**12.862**	**0.025**
Anxious	9	33	68	57	138	4.567	0.471
Relaxing	17	40	85	97	166	6.204	0.287
Dreamy	7	31	57	64	133	4.774	0.444
Energizing	**22**	**30**	**79**	**68**	**121**	**11.970**	**0.035**
Joyful	13	24	64	63	89	7.514	0.185
Neutral	**0**	**0**	**0**	**0**	**218**	**336.848**	**0.000**

6.8.2 Influence on emotions with the change of time

This study has a Song list where each song has a length of 1 min. In this case, there is a big question that is there any influence of time on emotions. After collecting data, this analysis divides the time into six parts, where every part contains ten seconds. Table 6 finds that almost half 48% of emotional responses come from the first twenty seconds through the experiment. However, the analysis sees an exception in neutral. More than half around 120 answers come from the last 10

seconds. The lowest response time is between 41-50 seconds, where only 10% of the participants give their responses.

Table 6: Time and Frequency of Emotion Selection.

Emotion	Time (Sec)					
	0-10	**11-20**	**21-30**	**31-40**	**41-50**	**51-60**
Happy	97	66	52	50	36	44
Sad	67	65	36	43	26	26
Amusing	104	81	52	55	43	39
Annoying	59	69	45	46	32	30
Anxious	77	70	41	45	32	40
Relaxing	105	98	69	50	39	44
Dreamy	86	63	39	38	32	34
Energizing	84	92	52	40	27	25
Joyful	70	63	44	34	23	19
Neutral	8	21	31	26	12	120
Total	**25%**	**23%**	**15%**	**14%**	**10%**	**14%**

6.9 Comparison of Rating Growth for Genres

This section analyzed the rating for four genres. This thesis finds the highest number of ratings for the pop genre. The total number of ratings for the pop genre is 1269. Through the experiment, this study finds similarities among classical, rock, and electronic. There is a curve observed in the Figure 17 between 2 and 4. Those three genres get most of the ratings for 3 and 5. In the case of rock songs 1=14, 2=85, 3=222, 4=136 and 5=283. On the other side classic and electronic get similar type ratings where 1= (20, 19), 2= (33, 54), 3= (102, 120), 4= (70,105) and 5= (219, 308). However, an exception comes to the front for a pop song. This thesis finds an exponential growth for pop song ratings where 1=98, 2=123, 3=200, 4=296, and 5=552. The rating growth of pop songs is different for the other three genres.
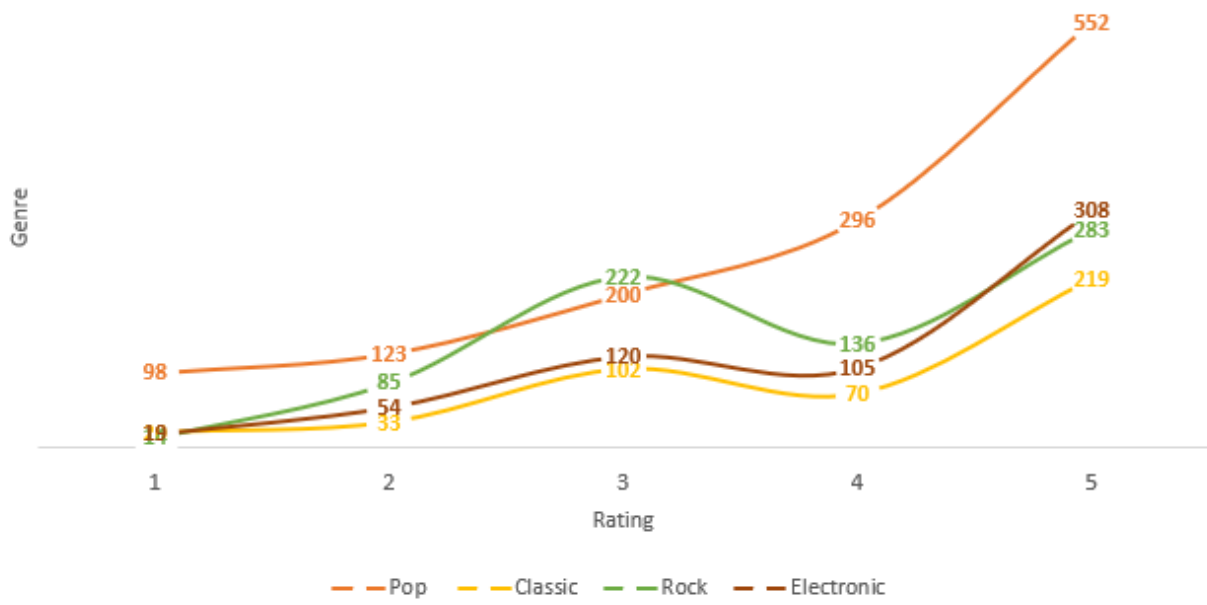
Figure 17: Rating Growth

6.11 Survey Data Findings

In our music player, a significant concern is about the button position. Because (Aljanaki et al. 2016) found a significant difference in emotional category selection for button position. For that, they used randomized button positions each session. However, in our t-test experiment, this study found a p-value>0.05. So, this thesis can conclude that there is not statistically significant effect from annotations collection.

This study finds an interesting fact in an analysis of the influence of genres on emotional annotations. All genres have a good ratio for relaxing annotation. Another factor we observed for pop, classical and electronic genres is that most given annotations are happy, amusing, relaxing, energizing, and joyful. On the other hand, rock songs obtained a higher percentage of sad, anxious, and annoying emotional annotation.

The analysis of participants, country, and profession this research found 163 participants from Ghana. The domain of the participants is primarily students. This study found 124 students. In another analysis, this study sees male feels more relaxing after hearing a song whereas female feels amusing. The percentage of comforting male emotion is 14%, and amusing female emotion is 12%.

From the survey data, this thesis observed note-able information through correlation coefficient analysis. The analysis shows the emotions that select together. Through the investigation, this thesis reveals an essential relation. For example, happiness does not correlate with sadness, annoying and anxiety. However happy have a strong correlation with amusement, joy, relaxation, energizing. This thesis found another similarity finds for the negative correlation coefficient. All the emotions have a negative correlation coefficient with neutral except annoying. Annoying does not correlate with neutral.

The Pearson's $\chi 2$ with p-value among the factors have displayed in Table 05. P-values with less than 0.01(<1%) among association factors represent a highly significant association, whereas p-value less than 0.05 (<5%) are donated as substantial. The investigation observe that four factors are significant where sadness (p=0.000, $\chi 2$=22.144) and neutral (p=0.000, $\chi 2$=336.848) is a highly effective association with ratings. The meaningful relationship between annoying – ratings (p=0.025 $\chi 2$=12.862) and energizing – ratings (p=0.035, $\chi 2$=11.970).

This study runs another investigation for checking the influence of time over emotional category. The survey data confirm that almost half of the responses from participants have come from the twenty seconds whereas 0-10 sec 25% and 11-20 sec 23%. However, this study also found that most of the neutral responses form the last ten seconds.

In the experiment of rating growth for each genre, this survey collects the total number of responses for each rating. The investigation ensures exponential growth for pop songs. However, the other genres like rock, classical and electronic have similar type growth patterns. Those genres have a carve between 2 and 4. They retrieved most of the responses for the 3 and 5 rating numbers.

7 Conclusions

This thesis analyzes the emotions induced by listening to the EF music app. This study chose ten annotations for this research because the study finds these annotations are adequate for evaluating ground truth. This study also finds those annotations are sufficient for an extensive dataset. This study conducts among 181 participants, where 90% of participants are students and from Ghana. Through the study, a significant point reveals that almost 48% of musical emotions conducts in the first 20 seconds. This analysis clarifies that people can connect their musical emotions very beginning. However, this thesis got an exceptional case for neutral most of the responses comes from the lost ten seconds. So, when the listeners cannot connect the emotions, they wait for a long time. An analysis for emotion among gender reveals that each genre relaxing have a higher ratio. Pop and electronic 15%, classical 17%, rock 12%. This analysis clarifies that listeners feel relaxing after a song. However, the analysis among males and females shows that males feel 14% relaxed whereas females feel 11% relaxed, which is lower than amusing. So, this study finds male feels more relaxing, and female feels amusing after enjoying a performance. In another investigation found an interesting correlation between emotional factors. This study sees happy and amusing, relaxing, dreamy, joyful have a strong correlation. The investigation also reveals sad and annoying, sadness and anxiety have strong correlations. This thesis also finds associations between ratings and emotional factors. Sadness, Annoying and energizing has significant associations with ratings. This study collects survey data, so the enormous amount of data may vary for the relations. This study hopes that those finding and implemented tool helps to develop more effective tools in future.

References

Aljanaki, A., Wiering, F., & Veltkamp, R. C. (2016). Studying emotion induced by music through a crowdsourcing game. *Information Processing & Management*, *52*(1), 115-128.

Arjmand, H. (2017). Emotional Responses to Music: Shifts in Frontal Brain Asymmetry Mark Periods of Musical Change. *Frontiers*

Apel, G. (1969). Protecting the health of children and adolescents in the German Democratic Republic. *Gigiena i Sanitariia*, *34*(7), 64-67.

Bakan, M. (2007). World music: Traditions and transformations. McGraw-Hill Higher Education.

Baileyshea, M. (2007). The struggle for orchestral control: power, dialogue, and the role of the orchestra in Wagner's Ring. *19th-Century Music*, *31*(1), 003-027.

Benetos, E., Dixon, S., Giannoulis, D., Kirchhoff, H., & Klapuri, A. (2013). Automatic music transcription: challenges and future directions. *Journal of Intelligent Information Systems*, *41*(3), 407-434.

Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing* (pp. 1-4). Springer, Berlin, Heidelberg.

Bhandari, P. (2021, February 15). An introduction to quantitative research. Scribbr.

Bhat, A. (2021, June 14). Data Collection: Definition, Methods, Example and Design. QuestionPro.

Burton, A. R., & Vladimirova, T. (1997). Applications of genetic techniques to musical composition.

Seashore, C. E. (1938). A musical ornament, the vibrato. *Proc. of Psychology of Music, 1938*, 33-52.

Clarke, E.F. (2002). Listening to Performance. *In John Rink, editor, Musical Performance — A Guide to Understanding*. Cambridge University Press, Cambridge.

Cole, W. (2019). Notation and the origins of Bach's Cello Suite in C minor (bwv1011). *Early music*, *47*(2), 241-254.

Coutinho, E., & Scherer, K. R. (2012). Towards a brief domain-specific self-report scale for the rapid assessment of musically induced emotions. In *Proceedings of the 12th International Conference of Music Perception and Cognition (ICMPC12)* (pp. 229-229).

Davis, W. B., & Thaut, M. H. (1989). The influence of preferred relaxing music on measures of state anxiety, relaxation, and physiological responses. *Journal of music therapy*, *26*(4), 168-187.

Duke, R. A. (1989). Effect of melodic rhythm on elementary students' and college undergraduates' perceptions of relative tempo. *Journal of Research in Music Education*, *37*(4), 246-257.

Dibben, N. (2004). The role of peripheral feedback in emotional experience with music. *Music Perception*, *22*(1), 79-115.

Eerola, T., Friberg, A., & Bresin, R. (2013). Emotional expression in music: contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, 4, 487.

Ekman, P. (1992). Are there basic emotions?. *Psychological Review*, 99(3), 550-553

Feng, B., Yao, P. M., Li, Y., Devlin, C. M., Zhang, D., Harding, H. P., ... & Tabas, I. (2003). The endoplasmic reticulum is the site of cholesterol-induced cytotoxicity in macrophages. *Nature cell biology*, *5*(9), 781-792.

Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae scientiae*, *5*(1_suppl), 123-147.

Gabrielsson, A. (2011). *Strong experiences with music: Music is much more than just music*. Oxford University Press.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition and Emotion*, 17(1), 25-40.

Geringer, J. M., & Breen, T. (1975). The role of dynamics in musical expression. *Journal of Music Therapy*, 12(1), 19-29.

Golberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. *Addion wesley*, 1989(102), 36.

Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object?. *Nature Reviews Neuroscience*, *5*(11), 887-892.

Hailstone, J. C., Omar, R., Henley, S. M., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *Quarterly Journal of Experimental Psychology*, 62(11), 2141-2155.

Hartmann, A. (1932). Untersuchungen über das metrische Verhalten in musikalischen Interpretationsvarianten [Investigation about the metric relationships in musical interpretation variants]. *Archiv für die gesamte Psychologie*, 84, 103-192.

Heller, J. J., & Campbell, W. C. (1971). Music performance analysis. *Bulletin of the Council for Research in Music Education*, 1-9.

Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *The psychology of music* (pp. 369–404). *Elsevier Academic Press*. https://doi.org/10.1016/B978-0-12-381460-9.00009-2

Holland, S. (1994). Learning about harmony with Harmony Space: an overview. *Music education: An artificial intelligence approach*, 24-40.

Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2008). Mixed affective responses to music with conflicting cues. *Cognition & Emotion*, 22(2), 327-352.

Jaimovich, J. (2013). Emotion recognition from physiological indicators for musical applications. PhD thesis, Queen's University Belfast, Belfast, United Kingdom.

Juslin, P. N., & Laukka, P. (2000). Improving emotional communication in music performance through cognitive feedback. *Musicae Scientiae*, *4*(2), 151-183.

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: listener, music, and situation. *Emotion*, *8*(5), 668.

Juslin, P. N., & Sloboda, J. (Eds.). (2011). *Handbook of music and emotion: Theory, research, applications*. Oxford University Press.

Juslin, P. N., & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception*, 17(2), 197-221.

Juslin, P. N., & Sloboda, J. A. (eds) (2010). *Handbook of Music and Emotion: Theory, Research and Applications.* New York, NY: Oxford University Press.

Kliewer, V. (1975). Melody: Linear aspects of twentieth-century music. *Aspects of twentieth-century music*, 270-321.

Koza, J. R., & Koza, J. R. (1992). Genetic programming: on the programming of computers by means of natural selection (Vol. 1). MIT press.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Can. J. Exp. Psychol*. 51, 336–352. doi: 10.1037/1196-1961.51.4.336

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition*, 11(4), 325-334.

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of personality*, *85*(4), 530-542.

Lahdelma, I., & Eerola, T. (2016). Single chords convey distinct emotional qualities to both naïve and expert listeners. *Psychology of Music*, 44(1), 37-54.

Laukka, P., & Gabrielsson, A. (2000). Emotional expression in drumming performance. *Psychology of Music*, 28(2), 181-189.

Levitin, D. J. (1999). The Psychology of Music, edited by Diana Deutsch. *Music Perception*, 16, 495-506.

Liu, X., Xu, Y., Alter, K., & Tuomainen, J. (2018). Emotional connotations of musical instrument timbre in comparison with emotional speech prosody: Evidence from acoustics and event-related potentials. *Frontiers in psychology*, 9, 737.

Meyer, L. B. (1956). "Emotion and meaning in music," *in Handbook of Music and Emotion: Theory, Research and Applications*, eds P. N. Juslin and J. A. Sloboda (Oxford: Oxford University Press), 279–312.

Olson, R. K., Wise, B. W., & Rack, J. P. (1989). Dyslexia: Deficits, genetic aetiology and computer-based remediation. *The Irish Journal of Psychology*, *10*(4), 494-508.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of experimental psychology: human perception and performance*, 15(2), 331.

Panksepp, J. (1995). The emotional sources of" chills" induced by music. *Music perception*, 13(2), 171-207.

Povel, D. J. (1977). Temporal structure of performed music: Some preliminary observations. *Acta Psychologica*, 41(4), 309-320.

Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *The Journal of the Acoustical Society of America*, 88(2), 622-641.

Repp, B. H. (1998). A microcosm of musical expression. I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *The Journal of the Acoustical Society of America*, 104(2), 1085-1100.

Rink, J. (Ed.). (2002). Musical performance: A guide to understanding. Cambridge University Press.

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of music*, *32*(4), 371-388.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science 340*, 216–219. doi: 10.1126/science.1231059.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them?. *Journal of new music research*, *33*(3), 239-251.

Scherer, K. R., & Coutinho, E. (2013). "How music creates emotion: a multifactorial process approach," in *The Emotional Power of Music*, eds T. Cochrane, B. Fantini, & K. R. Scherer (Oxford: Oxford University Press). doi: 10.1093/acprof:oso/9780199654888.003.0010.

Schubert, E. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, *35*(3), 499-515.

Schellenberg, E. G., Krysciak, A. M., & Campbell, R. J. (2000). Perceiving emotion in melody: Interactive effects of pitch and rhythm. *Music Perception*, 18(2), 155-171.

Schulkind, M. D. (1999). Long-term memory for temporal structure. *Memory & cognition*, 27(5), 896-906.

Sharma, P. B. (2013). Music: A tool for non verbal communication (with special reference to Indian music). *Int. J. Commun. Media Stud*, *3*(4), 1-6.

Sreeja, P. S., & Mahalakshmi, G. S. (2017, February). Concept identification from poems. In *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)* (pp. 211-216). IEEE.

Sturm, B. L. (2013). An introduction to audio content analysis: applications in signal processing and music informatics by alexander lerch. *Computer Music Journal*, 37(4), 90-91.

Tokui, N., & Iba, H. (2000, December). Music composition with interactive evolutionary computation. In *Proceedings of the third international conference on generative art* (Vol. 17, No. 2, pp. 215-226).

Torres-Eliard, K., Labbe, C., & Grandjean, D. (2011). Towards a dynamic approach to the study of emotions expressed by music. *In Proceedings 4th international ICST conference on intelligent technologies for interactive entertainment (pp. 252–259).*

Vuoskoski, J. K., & Eerola, T. (2010). Domain-specific or not? The applicability of different emotion models in the assessment of music-induced emotions. *In Proceedings of the 10th international conference on music perception and cognition (pp. 196–199).*

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1471.

Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian journal of psychology*, *13*(1), 241-257.

Xu, L., Wen, X., Shi, J., Li, S., Xiao, Y., Wan, Q., & Qian, X. (2020). Effects of individual factors on perceived emotion and felt emotion of music: Based on machine learning methods. *Psychology of Music*, 0305735620928422.

Zagrodski, M. (2013). Influence of musical context on the perception of emotional expression of music. *In Proceedings of the 3rd international conference on music & emotion*.

Zentner, M., & Eerola, T. (2011). *Self-report Measures and Models, din cadrul Handbook of Music and Emotion: Theory, Research, Applications*.