Music Therapy Survey Analytics Report

Executive Summary

Music therapy is a relatively new and understudied field of therapy. A self-reported study of 735 patients was analyzed using statistical testing and predictive modeling in order to identify predictors of improvement in music therapy patients. Three features we found to best predict improvement were whether or not a patient listened to music while they were working, their level of insomnia, and how frequently they listened to R&B. While our findings do provide valuable insights for clinicians and researchers, further research with more rigorous data collection is necessary to validate these before integrating them into the therapeutic process.

Introduction

Music therapy has emerged as a promising therapeutic approach, offering unique benefits to patients across multiple settings. Many academic institutions now have music therapy programs offering specialized training in this field. However, due to its novelty, limited research is available regarding the effectiveness of particular features, such as musical characteristics and listening habits, on patient improvement. Understanding the relationship between such factors and the outcomes of music therapy can help clinicians provide more effective treatment to patients thereby enhancing their outcomes.

To explore these relationships, an in-depth analysis was conducted on a music therapy survey organized by a computer science student at the University of Washington. The analysis followed a three step approach: (1) data cleaning to address the inconsistencies present in self-reported data, (2) exploration of feature interrelationships, and (3) predictive modeling to identify key features associated with treatment success.

Problem Statement

What features of a patient, their listening habits, or their music, effectively predict improvement following music therapy treatment?

Data Overview

Data Collection

The Music & Mental Health survey used for this analysis was distributed via Google Form where participants self-reported their symptoms, elaborated on listening habits, and included demographic details. 735 participants responded to the survey, which was posted in various

Reddit forums, Discord servers, and social media platforms. Posters and business cards were also used to advertise the form in libraries, parks, and other public locations.

In order to address missing values in the *BPM* column, additional data was collected from the Spotify API. The imported data included track features such as *danceability* and *acousticness* which were excluded as irrelevant for our analysis, therefore we only used the *tempo* (BPM) data.

Feature Overview

- **Timestamp** - Date and time when form was submitted
- **Age** - Respondent's age (integer)
- **Primary streaming service** - Respondent's primary streaming service (categorical)
- **Hours per day** - Number of hours the respondent listens to music per day (continuous)
- **While working** - Indicates if a respondent listens to music while working (binary: Yes/No).
- **Instrumentalist** - Indicates if a respondent plays an instrument regularly (binary: Yes/No).
- **Composer** - Indicates if a respondent composes music (binary: Yes/No).
- **Fav genre** - Respondent's favorite or top genre (categorical)
- **Exploratory** - Indicates if a respondent actively explores new artists or genres (binary: Yes/No).
- **Foreign languages** - Indicates if a respondent listens to music with lyrics in a language they are not fluent in (binary: Yes/No).
- **BPM** - Beats per minute of favorite genre (continuous).
- **Frequency (Genre)** - How frequently respondent listens to a particular genre (ordinal).
- **Anxiety** - Self reported anxiety on a scale of 1-10 (ordinal).
- **Depression** - Self reported depression on a scale of 1-10 (ordinal).
- **Insomnia** - Self reported insomnia on a scale of 1-10 (ordinal).
- **OCD** - Self reported OCD on a scale of 1-10 (ordinal).
- **Music effects** - Indicates effect of music on respondent’s condition (categorical: Improve/No effect/Worsen).
- **Permission** - Indicates permission to publicize data (binary: Yes/No).

Data Cleaning

In order to prepare the data for analysis and modeling, inconsistencies and missing values were appropriately addressed. The *Permissions*, *Timestamp*, and *Primary streaming service* columns were dropped because they provided no relevant information. *Age*, *Anxiety*, *Depression*, *OCD*, and *Insomnia* columns were converted to an integer data type for consistency. Participants who reported numbers outside of the requested scale for features such as *Anxiety* and *Depression* were dropped to maintain data integrity.

For columns besides *BPM*, missing values were imputed using either the mode, median, or mean. The mode was used for categorical features, and the mean for continuous features with a relatively normal distribution. The median was used for continuous features with significant skew, such as *Age* as seen in Figure 1.1, in order to keep the distribution of age consistent.

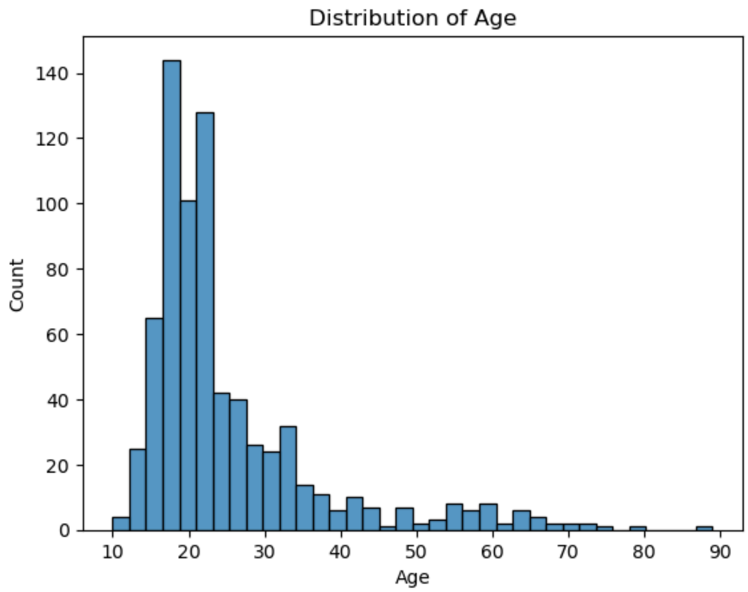

Figure 1.1 Distribution of Age in survey data

The *BPM* column was missing 107 values and contained a significant number of outliers. In order to maintain the volume of our data while preserving accurate BPM values, these were imputed using data extracted from Spotify. Specifically, genre-specific data, including BPM, was retrieved for the top 50 tracks of each genre using the Spotify API. Depending on the level of skew in the survey data, either the mean, median, or mode of the Spotify data was used to impute missing values.

Exploratory Data Analysis Findings

Differences between mean BPMs of various genres

To determine potential collinearity between BPM and Fav genre, we analyzed the distribution of BPM across the various genres and tested for significant differences in mean BPM. Figure 2.1 uses boxplots to illustrate the distribution of BPM across the various genres. Some genres exhibit significant skews, such as Jazz and Lofi, while others such as Latin and K pop exhibit a distribution which closely resembles a normal distribution.

An ANOVA test yielded an F-statistic of 4.9 and p-value of 3.52e-9, indicating that differences in mean BPMs across genres are statistically significant, and that mean BPM varied more between genres than within them. This allowed us to conclude that there are significant differences between mean BPMs between genres.

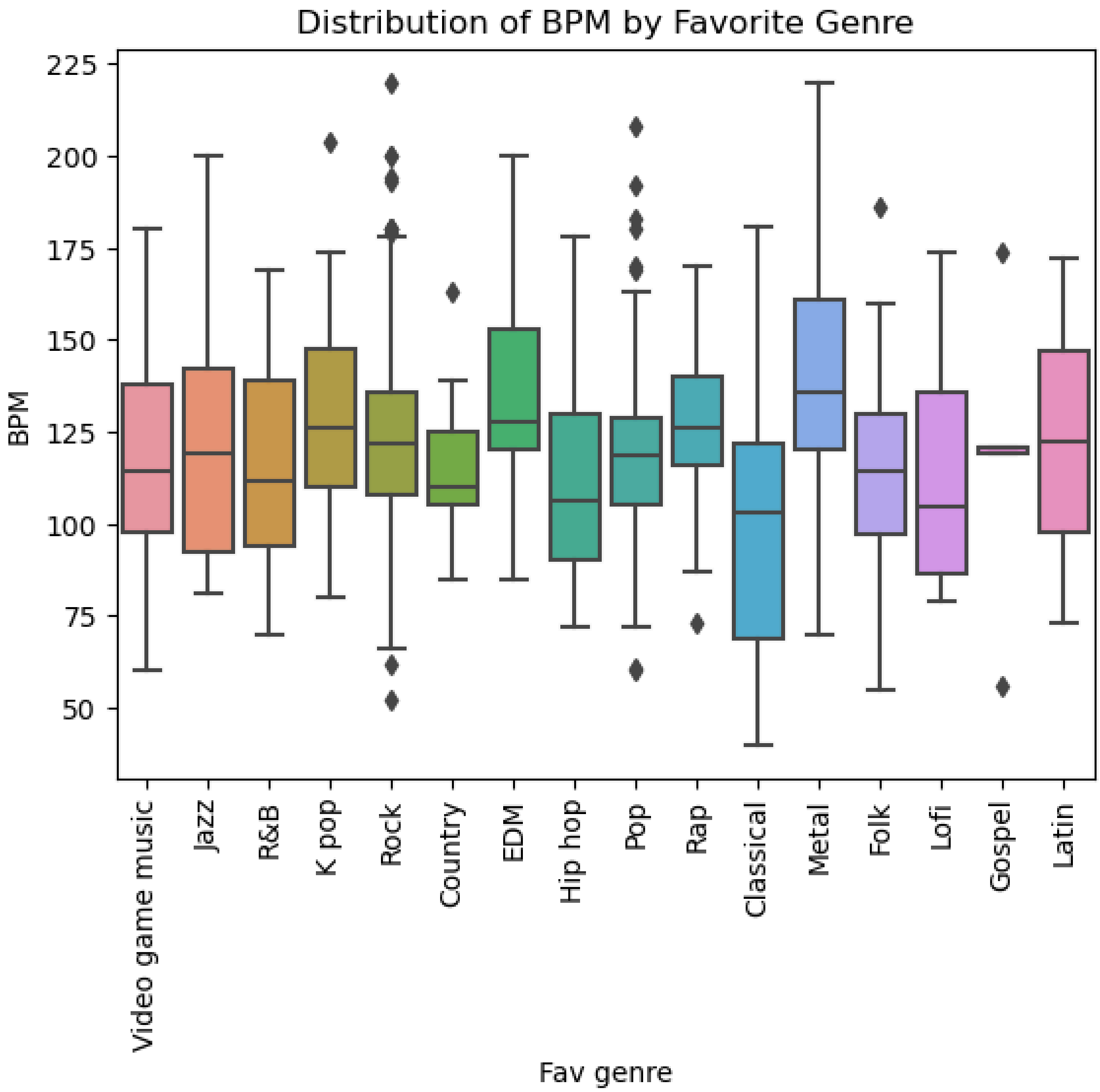

Figure 2.1: Distribution of BPM across genres

Relationship between being a musician and music therapy outcomes

Playing an instrument may influence how respondents respond to music therapy, as musicians often engage with music differently than nonmusicians Figure 2.2 demonstrates that while nonmusicians reported higher rates of improvement, a higher proportion of them worsened compared to musicians.

Due to a significant class imbalance (only 4 musicians worsened), bootstrapping was applied to balance the data. This resulted in the proportions present in figure 2.3. A chi-squared test with chi-squared equal to 50.35 and p-value equal to 1.15e-11 confirmed a statistically significant relationship between being a musician and music therapy outcomes.

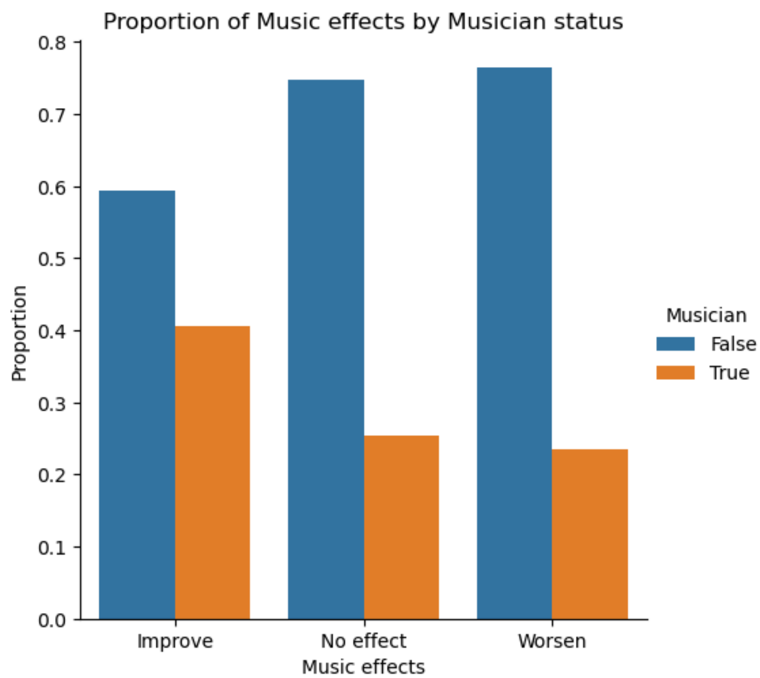

Figure 2.2 Effects of music therapy based on being a musician

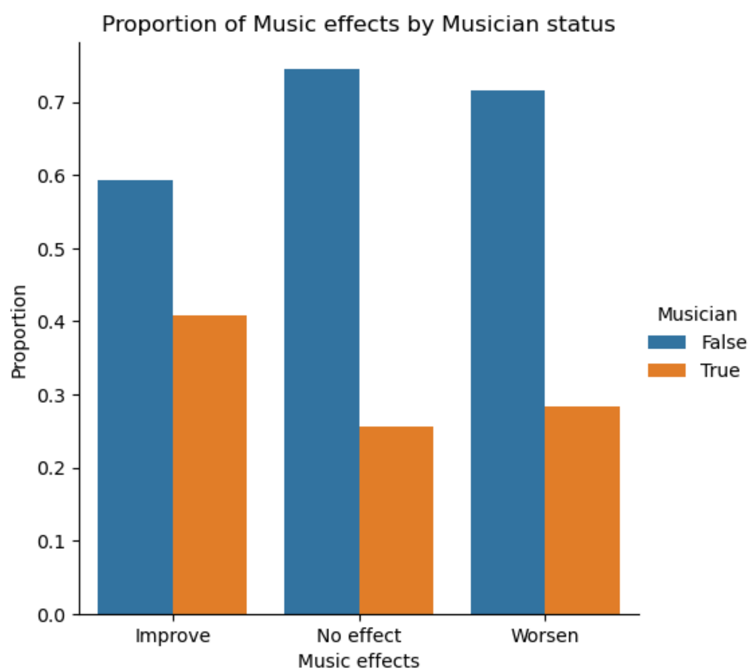

Figure 2.3: New chart representing therapy effects grouped by musician status using bootstrapped data.

Effects of mental disorders on therapy outcome

Chi-squared statistical tests between a patient's level of Depression (statistic: 3.33, p-value: 0.19), Insomnia (statistic: 1.44, p-value: 0.49), or OCD (statistic: 5.03, p-value: 0.08) and a positive outcome of music therapy did not yield a significant result. We were unable to reject the hypothesis that those suffering from OCD or Insomnia may worsen following treatment, suggesting that effects of music therapy are unclear on these individuals and further research is necessary. Participants suffering from Anxiety however, showed statistically significant improvement following therapy, suggesting music therapy could be effective for managing it. This remains true regardless of the increased levels of Anxiety as shown in figure 2.5.

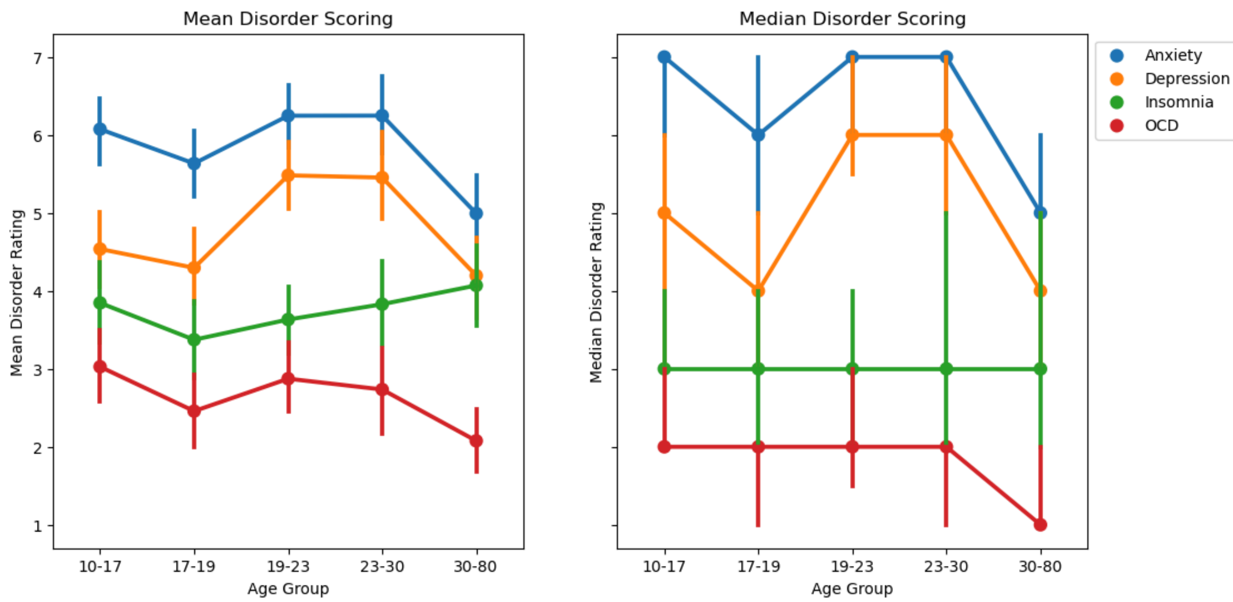

Figure 2.5: Mean and median disorder scoring grouped by age groups.

Modeling and Findings

Preprocessing included feature encoding and scaling. Binary variables were converted to dummy variables, ordinal variables were one-hot encoded, and continuous variables were standardized using StandardScaler. The *Music effects* feature was binarized into the target variable *Improve*. Data was split into training (70%) and test (30%) sets.

Given the categorical nature of the data, Logistic Regression, Random Forest, and Gradient Boosting Classifier models were tested. Due to class imbalance (no-information rate: 0.74), F1 scores were used to evaluate performance. Hyperparameter tuning (via grid/randomized search with 5-fold cross-validation) yielded models with mean accuracies of 85–86%. However, Gradient Boosting's test F1 score matched the no-information rate (0.74).

Although the models lacked strong predictive power, feature importance analysis was insightful. Gradient Boosting identified *While working* and *Anxiety level* as the top predictors of improvement, aligning with Logistic Regression's coefficients.

Recommendations

Investigate music therapy's effects on mental health disorders

Strengthen evidence by conducting targeted studies on OCD, Depression, Anxiety, and Insomnia. Through statistical testing and modeling our analysis suggests relationships between these mental disorders and varying therapy outcomes. Further research using rigorously collected and controlled datasets is necessary for making recommendations to practitioners.

Encourage music listening while working.

All predictive models identified listening to music as the strongest predictor of therapy success. Future studies should explore the mechanisms behind this effect and reevaluate its integration into therapy practices.

Enhance data collection methods.

To improve reliability, collaborate with practitioners globally to gather structured and consistent data. This will address the limitations of our self-reported data, including geographical biases and sample diversity, and allow for more robust recommendations.