The Impact of Music Listening Habits on Academic Performance and Stress Levels

Final Report

INFO 3237-090

Brielle Boling, Joe Flucus, Keon Herring, Jaden Smith, and Hannah Wood

April 22nd, 2025

Table of Contents **Pages**

Data Dictionary for Student Performance Dataset

Column Name	Data Type	Description
Student_ID	Object	Unique identifier for each student
First_Name	Object	Student's first name
Last_Name	Object	Student's last name
Email	Object	Contact email (can be anonymized)
Gender	Object	Student's gender (Male, female, other)
Age	Integer	Student's age
Department	Object	Student's academic department (e.g., CS, Engineering, Business)
Attendance (%)	Float	Attendance percentage (0-100%)
Midterm_Score	Float	Midterm exam score (out of 100)
Final_Score	Float	Final exam score (out of 100)
Assignments_Avg	Float	Average score of all assignments (out of 100)
Quizzes_Avg	Float	Average quiz scores (out of 100)
Participation_Score	Float	Score based on class participation (0-10)
Projects_Score	Float	Project evaluation score (out of 100)
Total_Score	Float	Weighted sum of all grades
Grade	Object	Final letter grade (A, B, C, D, F)
Study_Hours_per_Week	Float	Average number of study hours per week
Extracurricular_Activities	Object	Indicates if the student participates in extracurriculars (Yes/No)
Internet_Access_at_Home	Object	Does the student have internet access at home? (Yes/No)
Parent_Education_Level	Object	Highest education level of parents (None, High School, Bachelor's, Master's, PhD)
Family_Income_Level	String	Family income level (Low, Medium, High)
Stress_Level (1-10)	Integer	Self-reported stress level (1: Low, 10: High)
Sleep_Hours_per_Night	Integer	Average number of sleep hours per night.

Data Dictionary for Music and Mental Health Dataset

Column Name	Data Type	Description
Timestamp	Object	Date and time when the survey was submitted
Age	Float	Age of the respondent
Primary streaming service	Object	The respondent's primary music streaming service
Hours per day	Float	Number of hours the respondent listens to music daily
While working	Object	Indicates if the respondent listens to the music while studying/working (Yes/No)
Instrumentalist	Object	Indicates if the respondent plays an instrument (Yes/No)
Composer	Object	Indicates if the respondent composes music
Fav genre	Object	Respondents' favorite music genre.
Exploratory	Object	Indicates if the respondent actively explores new artists/genres (Yes/No)
Foreign languages	Object	Indicates if the respondent listens to music in languages they are not fluent in (Yes/No)
BPM	Float	Beats per minute of the respondent's favorite genre
Frequency [Classical]	Object	Frequency of listening to Classical music (never, rarely, sometimes, very frequently)
Frequency [Country]	Object	Frequency of listening to Country music (never, rarely, sometimes, very frequently)
Frequency [EDM]	Object	Frequency of listening to EDM music (never, rarely, sometimes, very frequently)
Frequency [Folk]	Object	Frequency of listening to Folk music (never, rarely, sometimes, very frequently)
Frequency [Gospel]	Object	Frequency of listening to Gospel music (never, rarely, sometimes, very frequently)
Frequency [Hip hop]	Object	Frequency of listening to Hip Hop music (never, rarely, sometimes, very frequently)
Frequency [Jazz]	Object	Frequency of listening to Jazz music (never, rarely, sometimes, very frequently)
Frequency [K-Pop]	Object	Frequency of listening to K-pop music (never, rarely, sometimes, very frequently
Frequency [Latin]	Object	Frequency of listening to Latin music (never, rarely, sometimes, very frequently)
Frequency [Lofi]	Object	Frequency of listening to Lofi music (never, rarely, sometimes, very frequently
Frequency [Metal]	Object	Frequency of listening to Metal music (never, rarely, sometimes, very frequently)
Frequency [Pop]	Object	Frequency of listening to Pop music (never, rarely, sometimes, very frequently)
Frequency [R&B]	Object	Frequency of listening to R&B music (never, rarely, sometimes, very frequently)
Frequency [Rap]	Object	Frequency of listening to Rap music (never, rarely, sometimes, very frequently)
Frequency [Rock]	Object	Frequency of listening to Rock music (never, rarely, sometimes, very frequently)
Frequency [Video Game]	Object	Frequency of listening to Video Game music (never, rarely, sometimes, very frequently)
Anxiety	Float	Self-reported Anxiety level (0-10 scale)
Depression	Float	Self-reported Depression level (0-10 scale)
Insomnia	Float	Self-reported Insomnia level (0-10 scale)
OCD	Float	Self-reported OCD level (0-10 scale)

Music effects	Object	Respondents' perception of music's impact on mental health (No Effect, Improves, Worsens)
Permission	Object	Indicates if the respondent gave permission to use their data (Yes/No)
BPM_Normalized	Float	A normalized measure of music tempo
Anxiety_Normalized	Float	Normalized score for anxiety

Project Summary

In modern academia, students face immense pressures from coursework, financial responsibilities, extracurricular commitments, and the complex process of transitioning into adulthood. As stress level concerns grow among student populations, it becomes critical to explore alternative avenues of support that extend beyond conventional methods. One such avenue is music. Music plays a central role in many students' daily routines, whether they are studying, commuting, exercising, or winding down after a long day. While many students believe that listening to music enhances their ability to concentrate or alleviates feelings of stress, the true impact of these habits on measurable outcomes such as academic performance and stress levels remains underexamined.

This project investigates how various dimensions of music listening behavior—such as favorite genres, listening frequency, beats per minute (BPM), and hours of daily listening—correlate with student's academic success (letter grades on 10-point scale) and self-reported stress levels ranked from 1 to 10. Our goal is to move beyond anecdotal claims and examine whether distinct music habits serve as indicators or even predictors of student stress levels and academic outcomes. Importantly, we focus our analysis through the lens of two major academic disciplines: STEM (Computer Science, Mathematics, and Engineering) and Business. These fields are known for their contrasting cognitive demands, and learning environments, all of which may influence how students in these domains use music to manage their academic and stressful lives.

The core objective of our analysis is to uncover meaningful patterns between music behaviors and academic or stress outcomes, particularly across academic disciplines. We seek to answer the following research question: **How do music habits impact the academic performance and stress levels of STEM and Business students?** To explore this, we utilized two different machine learning techniques, K-Means Clustering and Random Forests. Ultimately, the project aims to provide insights that may inform strategies for supporting student success and wellness through everyday behaviors like music listening.

Detailed Analyses of Data

Data Set Overview

This study utilizes two complementary datasets to explore the intersection of music habits, academic performance, and stress levels. The first dataset, *Students_Grading_Dataset* (5,000 records), includes real academic performance data from a private learning provider. Key attributes include age, attendance, major, total scores, and stress levels. While the dataset provides valuable insights into academic outcomes, limitations such as missing values and department imbalances must be considered. The second dataset is a *Music and Mental Health Survey* (736 responses), capturing respondents' preferred genres, music habits, study behaviors,

and self-reported mental health indicators (e.g., anxiety, depression, insomnia). While subjective in nature, it enables exploration of how perceived music effects relate to well-being and performance.

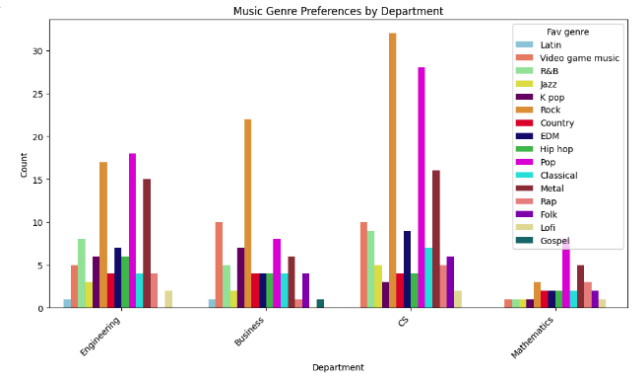
Data Cleaning and Preprocessing

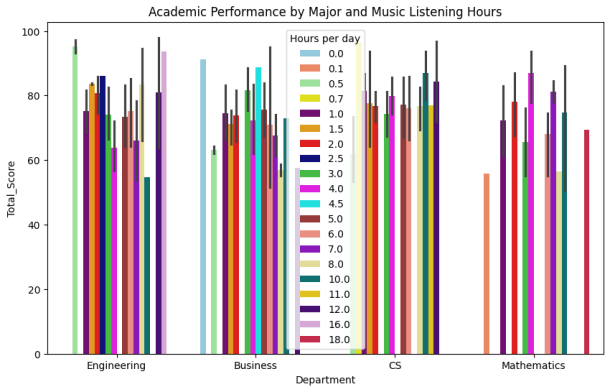
Before preprocessing, the datasets were merged to create a unified dataset for analysis. The original datasets, *studentperformance* and *musicmentalhealth*, did not share any common attributes or relational keys. To combine them, we used an inner join based on the index values of both datasets, ensuring that only rows with matching indices from both datasets were included. This approach allowed us to integrate the data effectively, aligning the performance and music habits of students with their mental health indicators for further analysis.

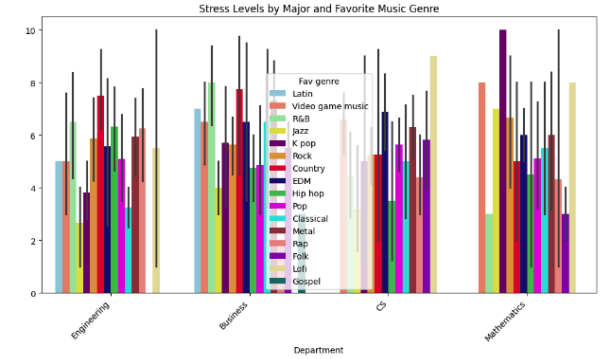
Once the datasets were merged, the data cleaning and preprocessing steps were carried out to ensure consistency and prepare the data for analysis. Missing values, which could have introduced biases or gaps in the analysis, were handled through imputation. For numerical features, the median value was used to replace any missing data, ensuring that outliers did not skew the imputation. For categorical variables, the mode was imputed to maintain consistency in the data. Additionally, to control the influence of extreme values, Winsorization was applied to continuous features like listening hours per day, limiting outliers to a predefined threshold. These preprocessing steps were essential for preparing the data for more accurate and reliable analysis.

Exploratory Data Analysis

Descriptive statistics and visual analysis were performed to explore distributions and trends in the data. Key patterns emerged in music genre preferences and their potential association with stress levels across departments. The most preferred genres varied notably by major—Rock and Pop were dominant in Computer Science, while Engineering showed a higher diversity in preferences, with notable counts for Rock, Pop, and Metal. Business students leaned toward Rock, whereas Mathematics students showed more modest genre preferences overall. Whilst still preferring Pop overall. The graph represents Academic performance by major and music listening preferences found that students who listen to music for a moderate amount of time—about 0.5 to 2.5 hours per day—tend to have higher scores, especially in Engineering and CS. CS students performed well across nearly all listening durations. However, students who listened for 10 or more hours per day had more mixed or lower scores, especially in Business and Math. Consequently, moderate music listening may help with academics, but too much could hurt—though some students still did well, showing study habits vary. Lastly, when analyzing stress levels, average scores varied by genre and department. Notably, students favoring Metal and EDM tended to report higher stress levels across most departments, while those preferring Classical or Jazz reported comparatively lower stress. These findings suggest possible links between music preference, academic discipline, and perceived stress.







Methodology

In this study, we applied several statistical and machine learning models to test our hypotheses and classify student outcomes and stress levels. The Random Forest Classifier was used to identify the most important predictors of academic performance and stress levels based on music habits. K-Means Clustering was also implemented to uncover natural groupings among students based on their listening patterns and mental wellness indicators. These methods were chosen for their strengths: Random Forests enhance accuracy and reduce overfitting, while K-Means clustering helps reveal hidden behavioral clusters, providing valuable exploratory insights into listening styles and the impact of music habits.

Results and Findings

K-Means Cluster Analysis

The Clustering analysis revealed four distinct student groups based on their music listening habits and associated stress and academic performance levels. Cluster 0 is characterized by high engagement across nearly all music genres, especially Classical, Jazz, and Lofi, along with slightly above average stress and academic performance scores. Cluster 1 shows a clear disengagement from most genres, particularly Metal, R&B, and Rap, and reports the lowest academic performance and stress levels. Cluster 2 exhibits low engagement with most genres but shows a strong preference for Metal and Rock; this group has moderate academic performance and slightly higher stress. In contrast, Cluster 3 demonstrates a strong preference for Hip Hop, K-Pop, R&B, and Rap, but lower engagement with Classical and Rock. This group has below-average academic scores and moderately low stress levels. These findings suggest potential correlations between genre preferences, stress and academic outcomes, highlighting the complexity of how music consumption may reflect or influence student well-being.

Random Forest 1: Music Habits and Academic Performance

In examining the relationship between students' music habits and academic performance, we created a Random Forest with an accuracy of 94.78%. Several patterns emerged suggesting that listening behavior may influence the letter grade a student receives. Key factors such as daily listening duration, BPM (beats per minute), and preferred genre were found to be significant indicators. Students who listened to music for shorter periods (typically ≤1.75 hours per day) and preferred softer or instrumental genres—such as Folk or Classical—tended to earn higher grades, falling within the A or B range. These students may be using music more strategically, balancing its benefits with academic focus.

Conversely, students who listened for extended periods (≥4.25 hours per day) or preferred high-BPM genres like EDM, K-pop, or Rap were more likely to fall into lower grade categories (C, D, or F). High tempo music and prolonged listening may contribute to distraction or reduced study efficiency. Overall, the results highlight that moderation and music type play an important role in academic success. Students who incorporate calming music for limited durations appear to perform better academically, while excessive or high-energy listening may negatively impact focus and outcomes.

Random Forest 2: Music Habits and Stress Levels

Utilizing an additional Random Forest, we explored the relationship between students' music habits and their self-reported stress levels (rated 1–10). The model achieved strong predictive performance with an accuracy of 91.1%. The first decision tree in this forest revealed that BPM (beats per minute) was the most significant predictor—lower BPM music consistently aligned with lower stress levels, while high BPM music was frequently associated with elevated stress. Listening time also emerged as an important variable. Students who listened to music for short to moderate durations (around 1 to 1.25 hours per day), especially to low-BPM tracks, tended to report the lowest stress levels (class 0.0–1.0). However, stress increased notably among those who listened for more than 3.5–4.5 hours per day, particularly when paired with faster music. This points to a potential overstimulation effect from prolonged high-energy listening.

The model also found that genre influences stress in meaningful ways. Genres like Classical, Jazz, Lofi, Country, and Gospel were consistently associated with lower stress, especially when paired with shorter listening durations and slower tempos. For instance, Gospel listeners with low BPM preferences ($\sim \leq 4.67$) almost exclusively appeared in the lowest stress category. The model identified "listening while working" as a notable stress predictor. Students who frequently listened to music while working—especially without matching music type to task—tended to experience elevated stress levels. This finding reinforces the importance of intentional listening routines, particularly during cognitively demanding activities.

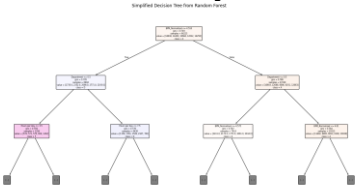
In contrast, genres such as Metal, EDM, Hip Hop, and Electronic were more common among students in higher stress classes (3.0–4.0), particularly when combined with long listening hours and high BPM. Pop showed mixed effects, reducing stress in some cases but increasing it in others depending on context. Overall, this Random Forest suggests that students' stress levels may be influenced not only by how much they listen to music, but also by what kind of music and how intense it is.

Random Forest 3: STEM vs. Business

To explore how music habits predict academic performance across STEM and Business majors, a Random Forest model was trained to classify students' letter grades using features like listening behavior, musical preferences, and department. After converting categorical variables such as department and favorite genre with label encoding and splitting the dataset, the Random Forest achieved a high accuracy of 94.91%, indicating strong predictive capability. The most influential feature in the model was BPM_Normalized, suggesting that the average tempo of music students listen to is significantly associated with academic outcomes. Following that, both Fav genre and Hours per day were tied in importance, indicating that what students listen to and how long they listen are nearly equally important. The final most important predictor was Department, though it contributed less to the model than musical characteristics suggesting that music behaviors may transcend academic major in predicting performance.

To gain deeper interpretability, rules from one of the decision trees in the forest were examined. For instance, if a student's favorite genre had a label index greater than 12.5 (corresponding to genres like Video Game Music or Metal), the model predicted a C average (3.0). However, when listening exceeded 7 hours per day but remained below 9 hours, the predicted grade increased to a B average (4.0), suggesting that moderate listening durations may not be detrimental. Notably, when listening time exceeded 9 hours and the favorite genre indexed below 12.0 (such as Hip Hop or Pop), students still achieved B-range grades (4.0), but if they favored higher-indexed genres (above 12.0), their predicted performance dropped to an F (0.0). This pattern suggests that genre choice may interact with listening time, amplifying negative outcomes when both BPM and exposure are high.

To visualize how these variables interact more intuitively, we extracted and plotted a simplified decision tree from the Random Forest. This tree placed BPM_Normalized at the root, reinforcing its dominant role in grade prediction. Students exposed to lower BPM ranged were generally predicted to receive higher grades (A or B), while higher BPM values often associated with genres like EDM or Metal were linked to lower outcomes (C to F). These findings suggest that excessive exposure to high-BPM music may impair cognitive performance, especially when combined with long listening durations or less cognitively engaging genres. Altogether, this analysis suggests that while major types play a role, music tempo, genre preference, and listening

habits have stronger predictive influence on academic performance. These insights can inform how students structure their study environments and make intentional choices about their music consumption. 

Recommendations

Based on our findings from cluster analysis and Random Forest models examining the relationship between music habits, academic performance, and stress levels, we offer several evidence-based recommendations for students. First, be intentional with your genre choices. Instrumental and low-BPM genres such as Classical, Lo-fi, Jazz, and Gospel are consistently associated with lower stress and higher academic performance. In contrast, high-BPM genres like EDM, Metal, and Hip Hop should be reserved for non-study times, as overuse of these styles is linked to elevated stress and lower grades.

Second, limit listening time throughout the day. Students who listened to music for one to three hours per day generally performed better academically and reported lower stress levels. In contrast, excessive listening—defined as more than four hours daily—was often correlated with decreased grades and heightened stress. Additionally, it's important to match your music to the task at hand. Calming or instrumental music works best for focused academic work, while more energetic tracks may be beneficial during breaks, workouts, or as a motivational boost. However, using high-energy music during complex study sessions may reduce concentration and hinder performance.

Finally, create structured listening routines. Establish boundaries around when and how you use music, such as curating specific playlists for study sessions or moments of stress relief. Students identified as "Focused Listeners" in our cluster analysis—those who used music selectively—demonstrated the best academic and mental health outcomes. In summary, music can be a powerful tool for enhancing academic life, but its benefits are highly dependent on genre, tempo, timing, and moderation. By making thoughtful and intentional choices, students can enjoy the advantages of music without compromising their performance or well-being.

Works Cited

Elhemaly, M. (n.d.). *Students grading dataset*. Kaggle. Retrieved March 23, 2025, from

https://www.kaggle.com/datasets/mahmoudelhemaly/students-grading-dataset

Rasgaitis, C. (n.d.). *MXMH survey results*. Kaggle. Retrieved March 23, 2025, from https://www.kaggle.com/datasets/catherinerasgaitis/mxmh-survey-results