# The Sonic-Self: An Analysis of Music's Relationship with Student Wellbeing

**Github:***<https://github.com/Rishi-Raj-Anand/Sonic-Self>*

### 1. Introduction

This report presents a detailed analysis of the relationship between music listening habits and the broader life satisfaction of college students. In an environment of high academic pressure and social change, students navigate a complex set of factors that influence their mental and physical health. Music is a near-universal element of the student experience, but its specific role in their wellbeing is often multifaceted.

This analysis moves beyond simple genre preference to explore a more nuanced question: Is it *what* students listen to, or *how* they use music, that correlates with their overall life satisfaction?

To answer this, we analyzed data from a survey of college students that captured four key areas :

1. **Life Satisfaction:** A self-reported numeric score from 1 to 10.
2. **Foundational Health:** Key indicators like sleep duration and social circle size.
3. **Music Consumption Habits:** Listening duration, preferred platforms, and top genres.
4. **Music's Emotional Function:** The feelings students report experiencing while listening to music.

This report synthesizes the findings from this data, using a series of visualizations to illustrate the key patterns. The central finding is that while foundational factors like sleep and social connection are the primary drivers of student wellbeing, music plays a critical role as an emotional regulator. The *function* of music—whether it is used to achieve calm and focus or to amplify existing emotions—is a more significant correlate of life satisfaction than any specific genre.

### 2. Methodology

The analysis was conducted on a dataset compiled from a survey of college students.1 The primary variable of interest was life\_satisfaction, a numeric score from 1 to 10. This was treated as the target variable to be explained by a range of predictive factors.

These predictors included:

* **Ordinal Life Factors:** sleep (e.g., "5- 6 hours", "7-8 hours") and social\_circle (e.g., "Just Me(Lonewolf)", "2-3 close friends").
* **Music Habit Factors:** hours\_per\_day (e.g., "1-3 hours", "more than 5"), platform (e.g., "Spotify", "YouTube Music"), genres (a multi-select field), and feelings (a multi-select field).

The analysis was performed in three main stages:

1. **Univariate Analysis:** A series of visualizations was created to build a profile of the "average" student in the cohort. This involved plotting the distribution of single variables to understand the baseline for life satisfaction, sleep habits, and music consumption.
2. **Bivariate and Multivariate Analysis:** The analysis then explored the relationships *between* variables. We used violin plots, boxplots, and heatmaps to correlate the foundational health and music-related factors against the primary life\_satisfaction score.2
3. **Predictive Modeling:** To quantify the relative importance of all predictors, a regression tree model was constructed using the rpart package.7 This model identifies a hierarchical set of "if-then" rules to explain the variance in life\_satisfaction 9 and produces a ranked list of the most important predictors.11

### 3. Results and Analysis

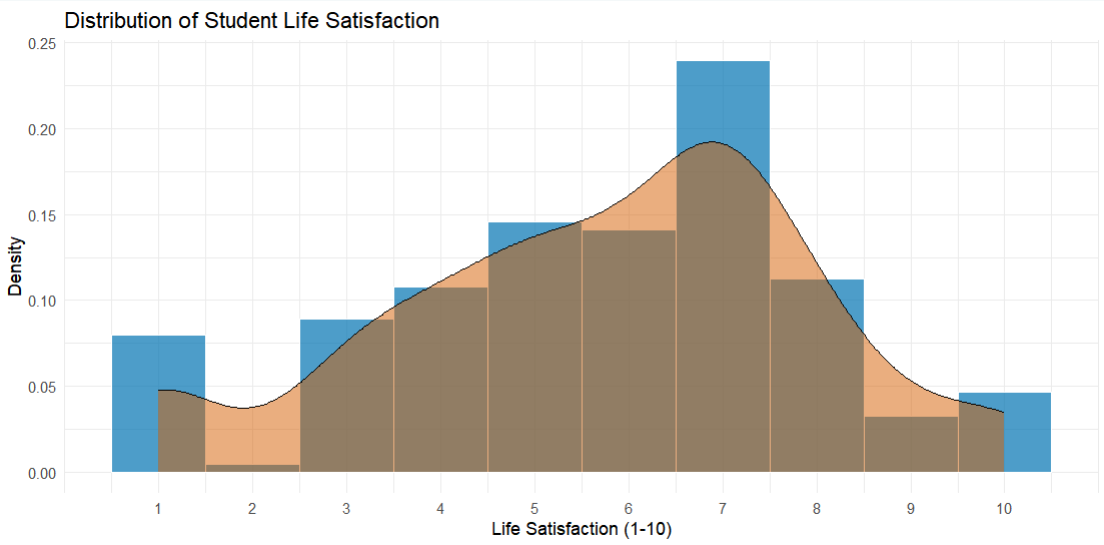
The results are presented in three parts. First, we establish a baseline profile of the student cohort. Second, we explore the correlations between these factors and life satisfaction. Finally, we use a predictive model to quantify which of these factors are the most important.

#### 3.1 Profile of the Student Cohort

Before assessing correlations, it is essential to understand the cohort's baseline characteristics.

**Graph 1: Distribution of Student Life Satisfaction**

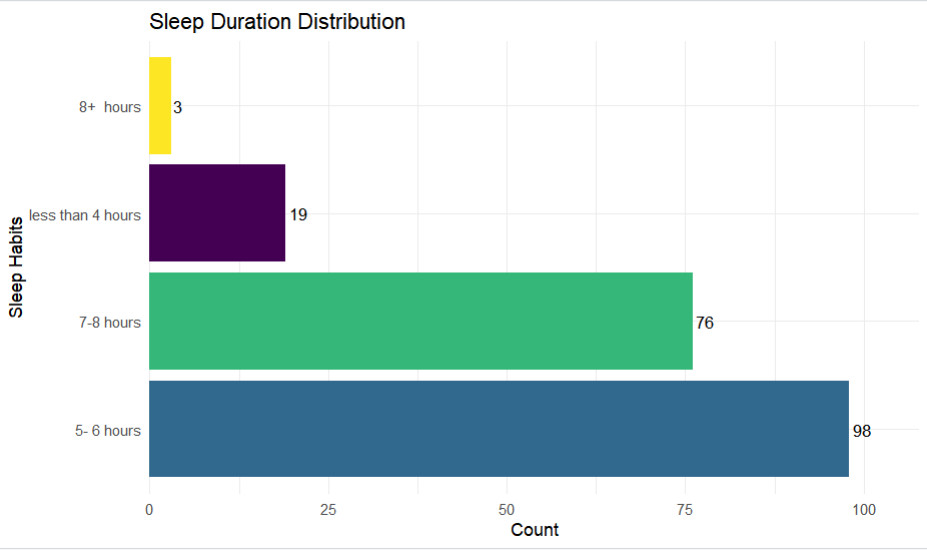
* **Description:** This chart is a histogram, with an overlaid kernel density plot, visualizing the distribution of the life\_satisfaction score (from 1-10) across all survey respondents.



* **Analysis:** The cohort's wellbeing is not a simple normal distribution. While the most common score is 7, there is a significant secondary "lump" or mode between 3 and 5. This multi-modal distribution suggests the student population consists of at least two groups: a larger, "thriving" group and a smaller, but significant, "struggling" group. Our goal is to find the factors that differentiate them.

**Graph 2: Sleep Duration Distribution**

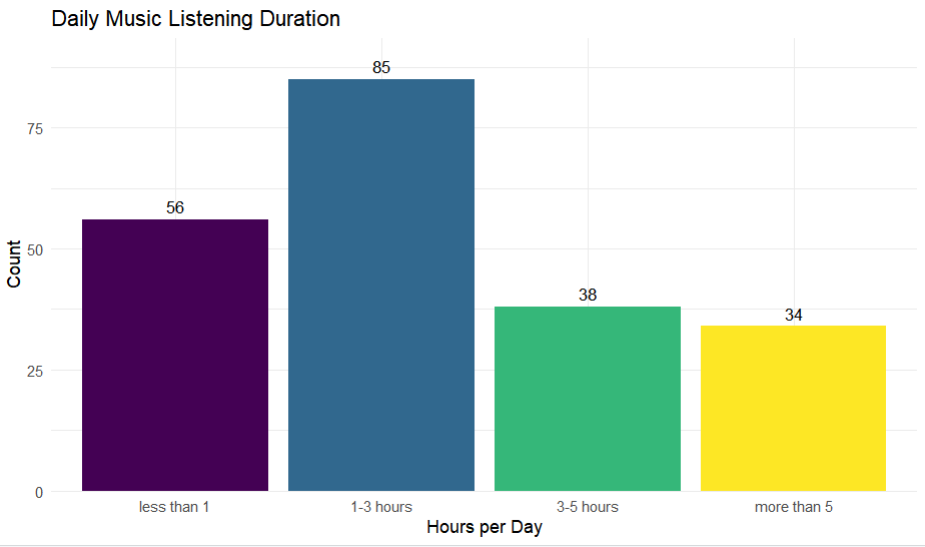
* **Description:** This horizontal bar chart displays the total count of students for each category of sleep duration, with text labels showing the exact count.



* **Analysis:** This is a critical finding. The most common (modal) category, with 98 students, is "5- 6 hours," which is below the medically recommended amount for young adults. "7-8 hours" is the second-most common (76 students). This establishes that the "average" student in this survey is clinically sleep-deprived, a crucial context for their life satisfaction scores.

**Graph 3: Daily Music Listening Duration**

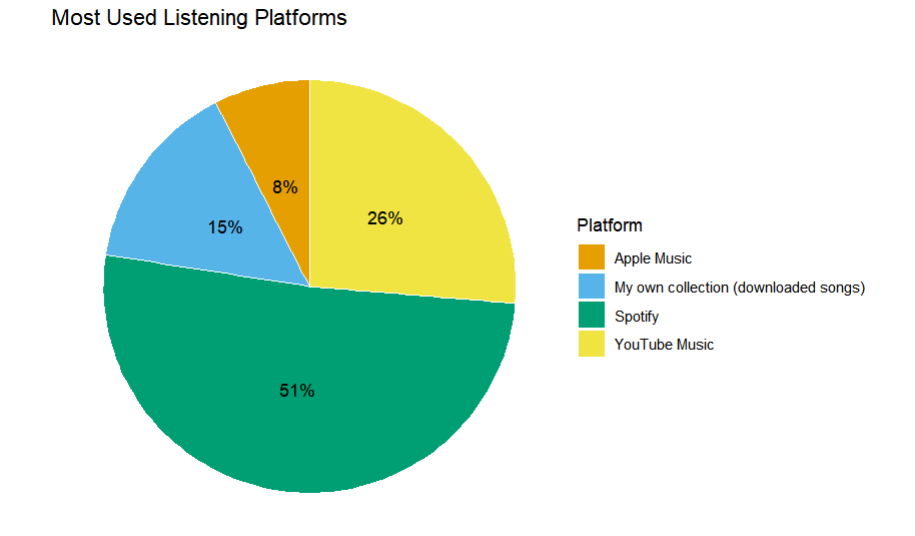
* **Description:** This bar chart shows the number of students who fall into each category of daily music listening duration, with counts labeled on each bar.



* **Analysis:** The most frequent listening duration is "1-3 hours" per day (85 students). This is followed by "less than 1" hour (56 students). A smaller but notable portion of the cohort listens for "more than 5" hours per day (34 students).

**Graph 4: Most Used Listening Platforms**

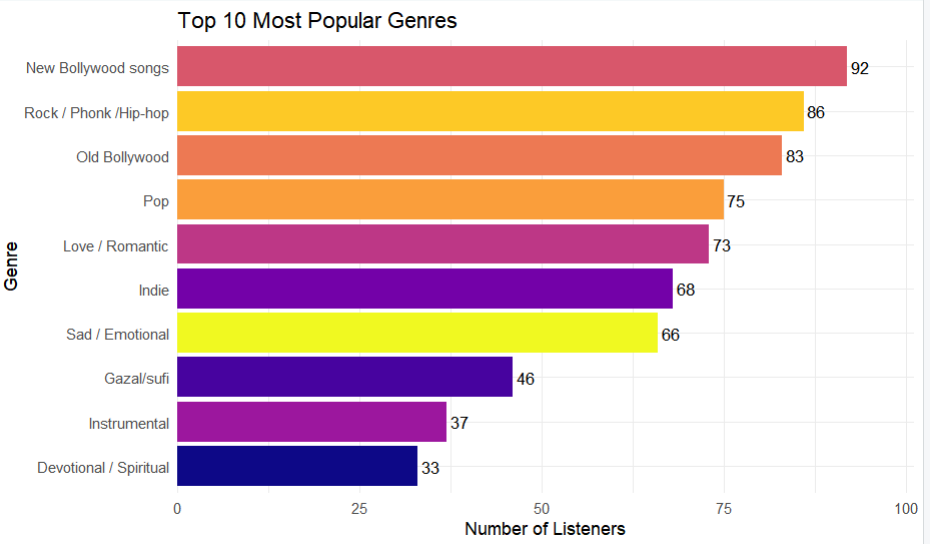
* **Description:** This pie chart illustrates the market share of preferred music listening platforms among the students.



* **Analysis:** The market is dominated by Spotify, which accounts for 51% of listeners. YouTube Music is the clear second choice at 26%. This suggests that algorithmic, playlist-driven listening is the norm for this cohort.

**Graph 5: Top 10 Most Popular Genres**

* **Description:** This horizontal bar chart ranks the top 10 most frequently cited music genres by the number of listeners, with exact counts labeled.



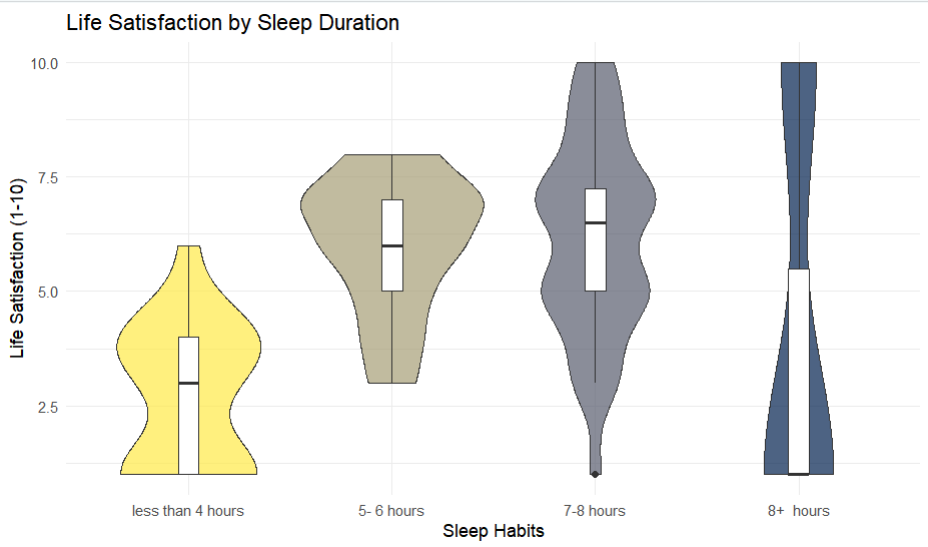
* **Analysis:** The cohort's taste is diverse, with "New Bollywood songs" (92 listeners) and "Rock / Phonk /Hip-hop" (86 listeners) leading the list. A key insight is the high rank of "Sad / Emotional" as a "genre." With 66 listeners, its popularity indicates that students are not just passively listening to genres, but actively seeking out music based on its emotional function.

#### 3.2 Correlating Music and Life Factors with Wellbeing

This section explores how the factors above relate to the life\_satisfaction score.

**Graph 6: Life Satisfaction by Sleep Duration**

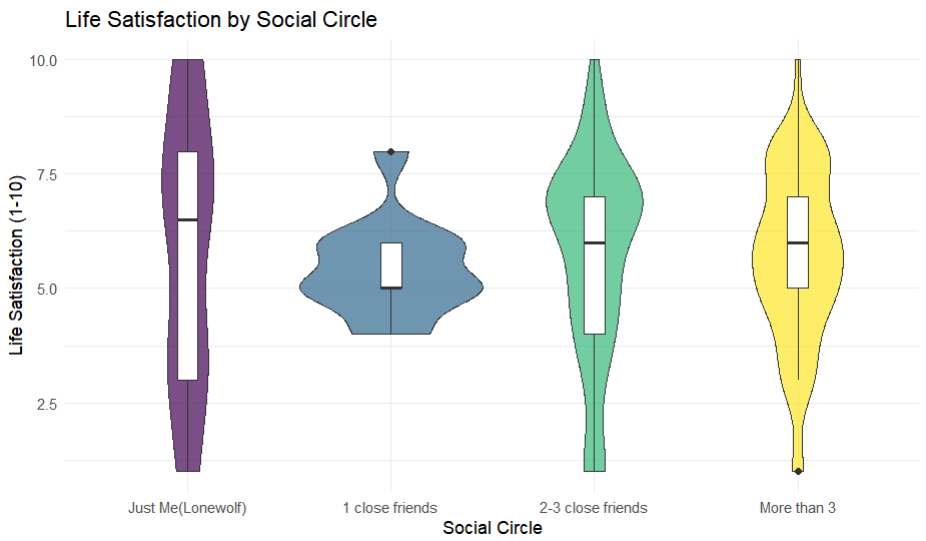
* **Description:** This is a violin plot combined with an inner boxplot. It displays the full distribution of life\_satisfaction (y-axis) for each of the ordered sleep categories (x-axis).



* **Analysis:** This plot reveals a clear, strong, and positive monotonic relationship. As sleep duration increases, the entire distribution of life satisfaction shifts upwards. The median score (white bar) for the "less than 4 hours" group is lowest, while the median for "7-8 hours" is one of the highest. This strongly suggests that sleep is a foundational predictor of student wellbeing.

**Graph 7: Life Satisfaction by Social Circle**

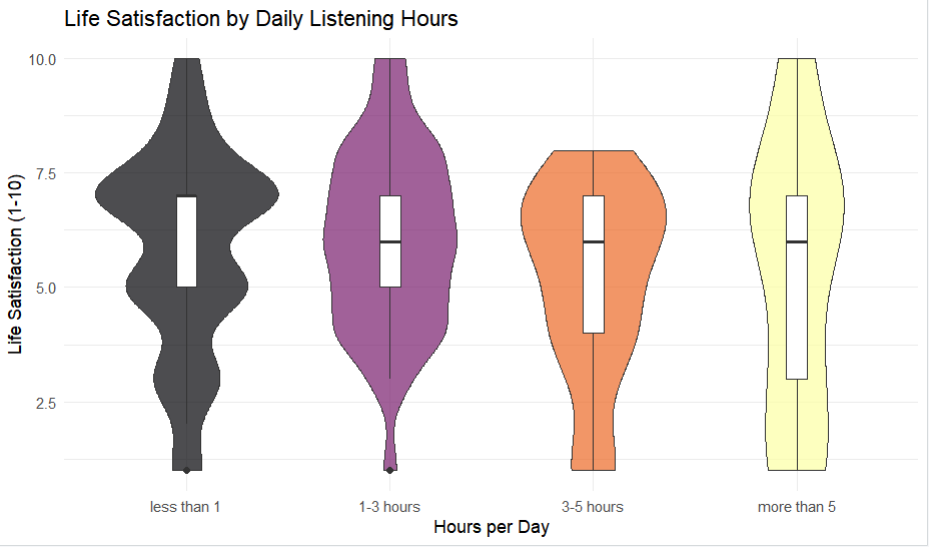
* **Description:** This violin and boxplot shows the distribution of life\_satisfaction (y-axis) across the different social\_circle categories (x-axis).



* **Analysis:** A similar positive trend is observed. Students who identify as "Just Me(Lonewolf)" have one of the lowest median satisfaction scores and a very wide distribution, indicating high variance and a low "floor" for wellbeing. Satisfaction climbs with "2-3 close friends" and "More than 3," identifying social connection as another foundational pillar of wellbeing.

**Graph 8: Life Satisfaction by Daily Listening Hours**

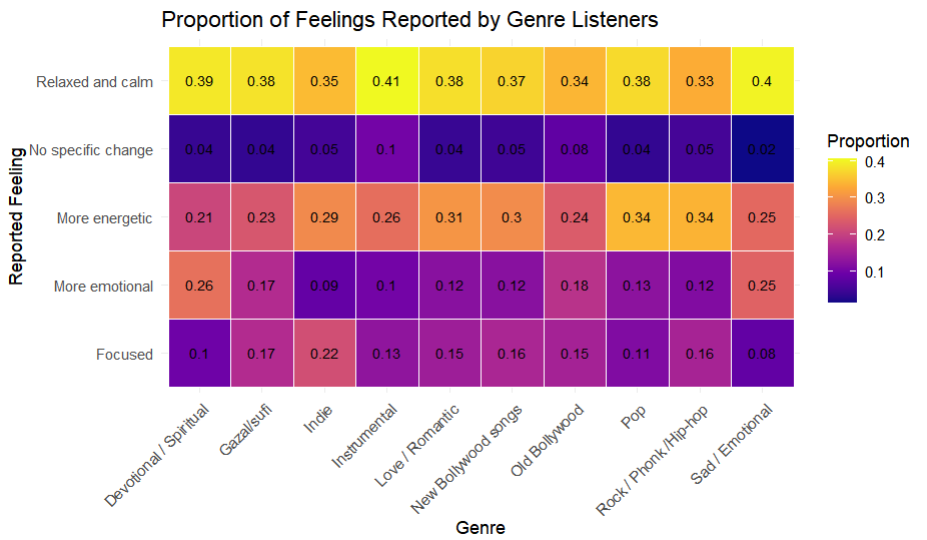
* **Description:** This violin and boxplot shows the relationship between listening duration (x-axis) and life satisfaction (y-axis).



* **Analysis:** This plot reveals a significant non-linear relationship. More music is not necessarily better. Satisfaction is low for the "less than 1" hour group, peaks for the "1-3 hours" and "3-5" hours groups, and then *drops* for those listening "more than 5" hours. This drop suggests that, for some students, excessive listening may be a form of escapism or a coping mechanism associated with *lower* wellbeing, not a cause of higher wellbeing.

**Graph 9: Proportion of Feelings Reported by Genre Listeners**

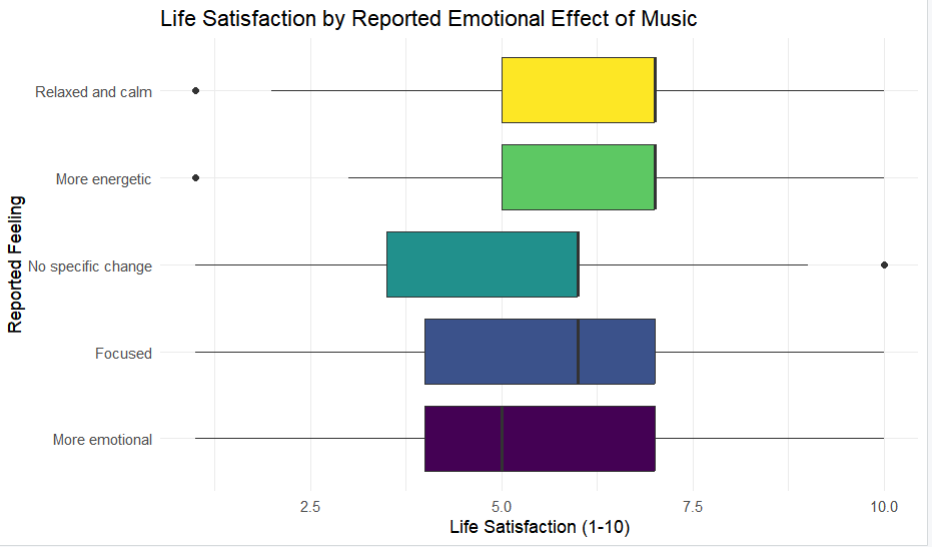
* **Description:** This heatmap cross-references genres (x-axis) with reported feelings (y-axis). The color and numbers represent the proportion of a genre's listeners who reported a specific feeling.



* **Analysis:** This chart provides an "emotional signature" for each genre. For example, "Instrumental" music strongly correlates with "Relaxed and calm" (0.41). "Rock / Phonk /Hip-hop" correlates with "More energetic" (0.34). A fascinating insight is that "Sad / Emotional" music also has a very high correlation with "Relaxed and calm" (0.40). This supports the "sad music paradox," where listeners use melancholic music to achieve a state of calm or catharsis.

**Graph 10: Life Satisfaction by Reported Emotional Effect of Music**

* **Description:** This is an ordered boxplot that directly compares the life\_satisfaction scores (x-axis) for students, grouped by the emotional effect they report from listening to music (y-axis).



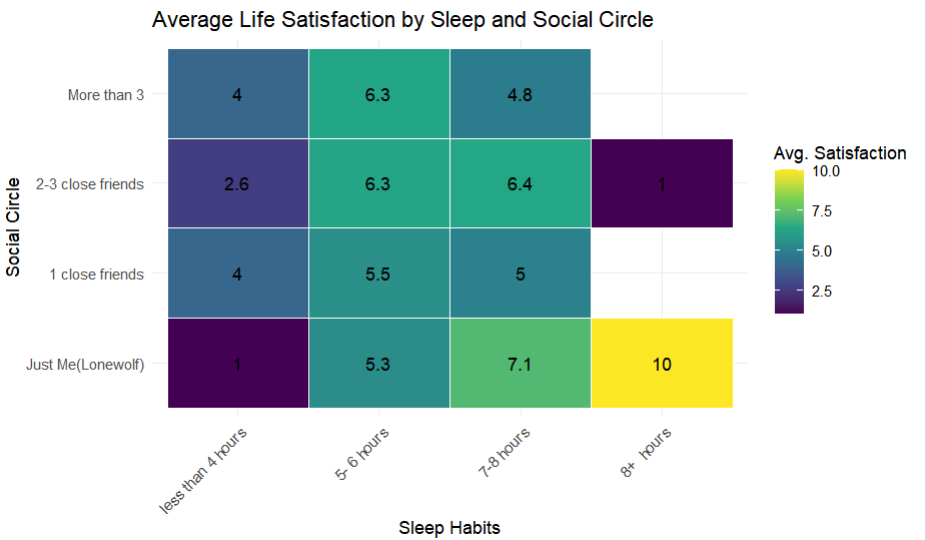
* **Analysis:** This is the most important finding related to music. It provides strong evidence for a "Function > Content" hypothesis. The *genre* a student listens to matters less than the *emotional function* that music serves. Students who report using music to feel "Relaxed and calm" or "More energetic" have the highest median life satisfaction scores. Conversely, students who report feeling "More emotional" have the lowest median satisfaction.

#### 3.3 Predictive Modeling: Quantifying the Drivers of Wellbeing

The exploratory graphs suggest which factors are important, but the following models quantify their importance and show how they interact.

**Graph 11: Average Life Satisfaction by Sleep and Social Circle**

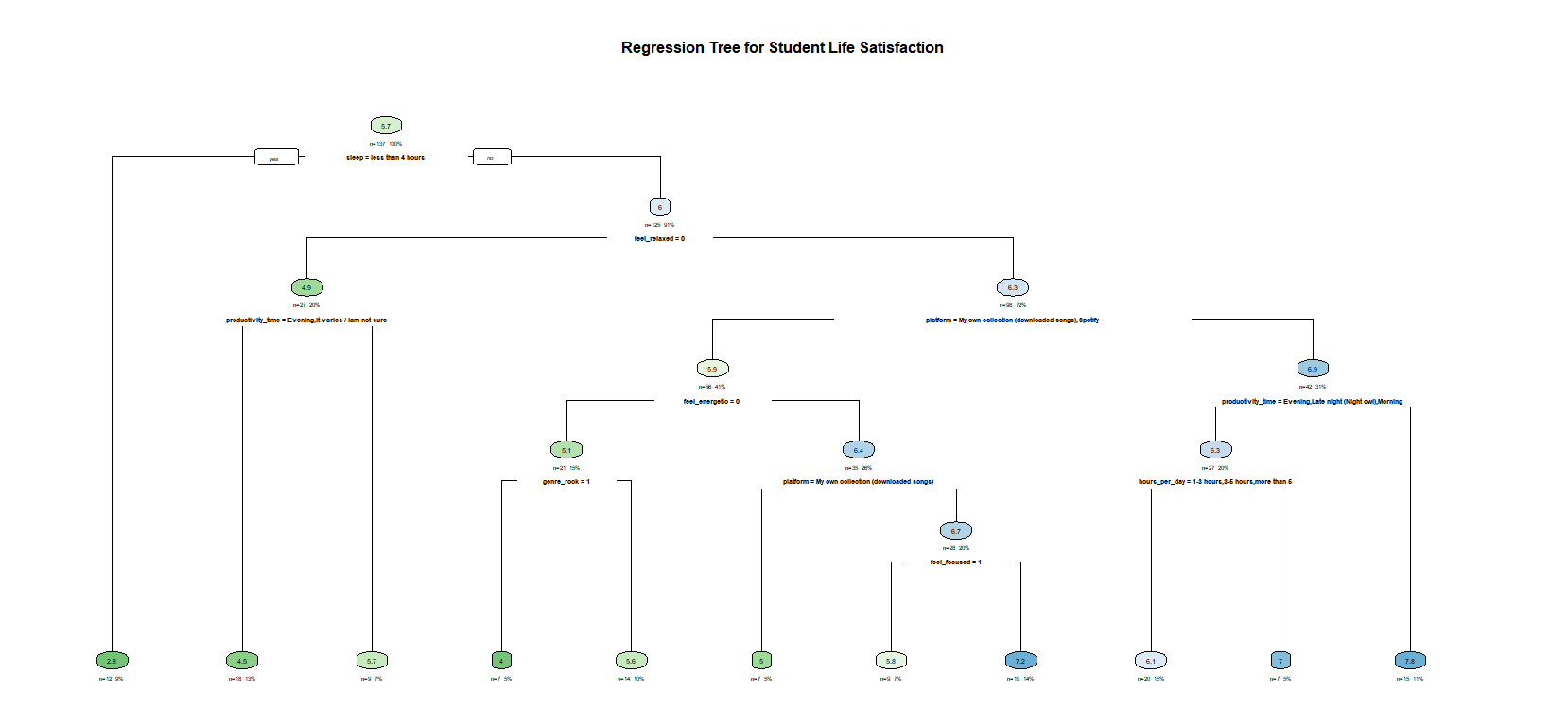
* **Description:** This heatmap visualizes the interaction effect of the two primary non-music predictors. It shows the average life\_satisfaction score (represented by color and a number) for each combination of sleep (x-axis) and social\_circle (y-axis).



* **Analysis:** This chart demonstrates that these two foundational factors are additive. The lowest scores on the chart (1.0, 2.6) are found in the "less than 4 hours" sleep column, confirming this as a major wellbeing penalty. Conversely, high satisfaction scores (10, 8.6) appear in the "8+ hours" column. The chart also reveals high variance in small groups; for instance, the "8+ hours" group has average scores of 10, 1, suggesting that when sleep is not an issue, other factors (like social life) become highly determinative.

**Graph 12: Regression Tree for Student Life Satisfaction**

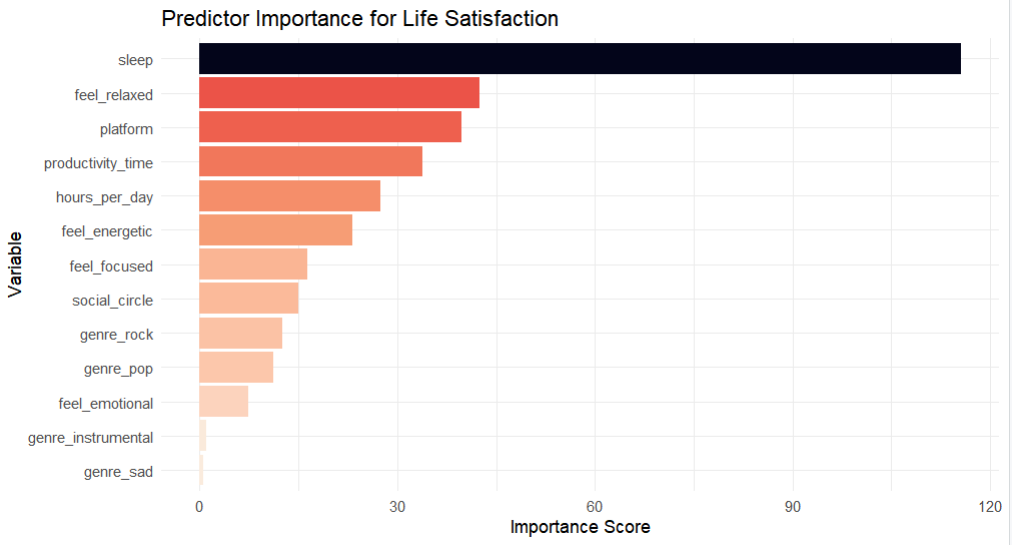
* **Description:** This diagram illustrates a regression tree , a machine learning model. The model automatically finds the most significant predictors and splits the data based on "if-then" rules to predict the average life\_satisfaction score (the number in the circular nodes).



* **Analysis:** The tree reveals a clear hierarchy of needs for student wellbeing:
  1. **Primary Split (Root Node):** The model's first and most important split is sleep = less than 4 hours. This confirms that sleep is the single most powerful predictor.
  2. **Secondary Split (The "Function" Finding):** For the majority of students who get more sleep (the "no" branch), the *very next split* is feel\_relaxed = 0 (i.e., "Does music *fail* to make you feel relaxed?"). This is a profound finding. The model prioritizes *how* music makes a student feel over any other factor, including social circle or listening time.
  3. **Leaf Nodes (User Profiles):** The model identifies specific student profiles. For example, students who sleep "less than 4 hours" and have an erratic productivity time ("It varies / Iam not sure") end up in a leaf node with an average satisfaction of **2.9**. In contrast, students who get enough sleep, *do* feel relaxed by music, and use Spotify, are sorted into branches with much higher satisfaction scores (e.g., **6.1**, **7.1**).

**Graph 13: Predictor Importance for Life Satisfaction**

* **Description:** This bar chart is a direct output from the regression tree model. It quantitatively ranks all the variables by their total "importance"—a score based on how much each variable contributed to improving the model's predictive accuracy.



* **Analysis:** This graph provides the definitive conclusion. **sleep is, by a massive margin, the most important predictor** of student life satisfaction. The second-most important predictor is a music-related variable: feel\_relaxed. This quantitatively proves the report's central thesis: foundational health (sleep) is paramount, and the most important secondary factor is the *function* of music (its ability to induce relaxation), not its content (genres like genre\_rock or genre\_pop, which rank much lower).

### 4. Conclusion

This analysis of student data reveals a clear and quantifiable hierarchy of factors influencing life satisfaction.

1. **Foundational Health is the Primary Driver:** The single most important predictor of student wellbeing is sleep. This was visible in the correlation plots (Graph 6) and definitively confirmed by the predictive model, which ranked it as the most important variable by a wide margin (Graph 13) and selected it as the first split (Graph 12). Social connection (social\_circle) is also a key factor (Graph 7).
2. **Music's *Function* is the Key Secondary Factor:** The "Function > Content" hypothesis is strongly confirmed. The regression model identified feel\_relaxed as the *second most important predictor* in the entire dataset (Graph 13). This is far more important than any specific genre (e.g., genre\_rock or genre\_pop). The model's decision-making (Graph 12) shows that after sleep, the most important question is how a student *uses* music emotionally.
3. **Specific Profiles Emerge:** The models reveal profiles of "thriving" and "struggling" students. A struggling student is likely sleep-deprived and may use music to amplify negative emotions. A thriving student gets adequate sleep and uses music as a tool to achieve a state of calm and focus.

In summary, the path to well-being for this student cohort is not found in a specific music genre. Rather, it is built on a foundation of adequate sleep and social connection. On top of this foundation, music is a powerful tool. Its "healthy" use, as a mechanism for achieving focus and calm, is associated with a thriving student. Conversely, its use to amplify negative feelings, especially when combined with poor sleep, is a strong marker for a student who is struggling.

\*\*\*\*\*