# Exploring the Impact of Music Engagement on Mental Health Outcomes

## Introduction

In recent years, the relationship between music consumption and mental health has been recognized. Often seen as a powerful tool for expressing and relieving stress, music may also play a complex role in mental health, revealing potential links between listening habits, music genres and mental health issues such as anxiety and depression. Some music listeners might use music to manage stress, but some might only see it as an art form. This raises an important question: does music, as a universally accessible art form, have a positive or negative impact on mental health based on different factors such as genre, age, listening habits, etc.?

Understanding the connection between music and mental health is valuable to researchers, mental health professionals, and the general public. With the increasing of mental health problems, especially among young people, a deeper understanding of how different music practices impact mental health outcomes could provide accessible strategies to improve physical and mental health. Additionally, this knowledge can enhance the resources and advice provided by therapists and counselors, making the approach to be more personalized. For music lovers, understanding how their music consumption patterns impact their mental health may lead to healthier music listening habits.

Prior research has explored various facets of this topic. Daniel E. Gustavson and colleagues explore the relationship between music engagement and mental health. They found that engaging with music by listening, singing, or playing is generally associated with improved mental health outcomes (Gustavson et al.). However, the findings also indicate that some individuals, particularly musicians, may be at a higher risk for mental health challenges. The study suggests more investigations. Therefore, we still need to study how genre, listening habits, and other aspects impact on mental health in diverse populations. This research aims to examining these aspects and how they relate to mental health problem.

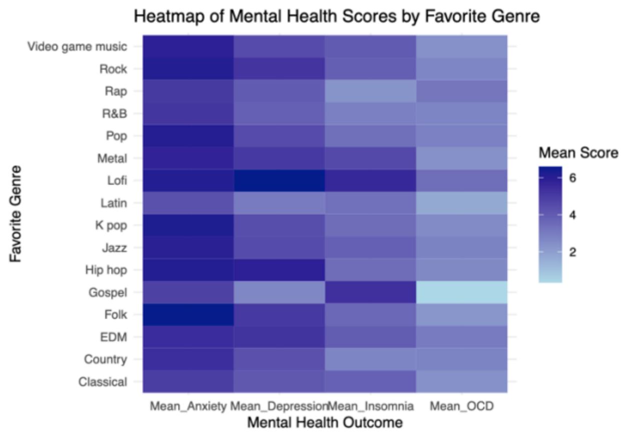
**Method**

This study utilizes the "Music & Mental Health Survey Results" dataset. It was originally collected and published by Catherine Rasgaitis on Kaggle. The dataset collected responses from individuals who participated in an online survey with total 736 observations. For this analysis, the primary response variable is the mental health status, assessed through self-reported level of anxiety, depression, insomnia, and OCD. The level is based on a scale from 0 to 10. I used each variable independently to test each mental health problem with music predictors. The predictor variable are: music genre preference: the primary genres participants listen to, such as classical, rock, pop, hip-hop, jazz, and electronic. Different genre might affect mental health differently. Music listening frequency: how often participants engage in music listening. Time spending might correlate with mental health conditions. Instrumentalist or composers: whether being related to music has direct influence on mental health. Music effect: participants perceive the impact on their mental health. Age: the age difference might also impact how music relates to the mental health.

The missing values were examined, particularly in key variables like mental health status and all the predictor variables. The analysis process includes statistical method which is multiple linear regression analysis will be performed to examine these relationships. During the analysis, I recoded the categories such the variable instrumentalist and composer, I gave yes to be 1 and no to be 0. I also separate the favorite genre to all different genres and the music effect to each category as “improve”, “no effect”, and “worsen” when putting them into regression models.

**Results**

After applying some models to my data, the results for my research questions are shown here with some models and tests. The first step was to visualize the relation between different music genre and mental health problem. Based on the heatmap below, it shows that for most the music genre, they tend to have a higher mean score with the anxiety as the blue is darker. For OCD, overall is lighter on the graph which means the mean score is smaller for most of the genre. So, I decided to explore more on which genre is significant with anxiety, and for other predictors which predictors are significant for different mental health problem.




Then, I applied multiple linear regression with anxiety as response variable and age, stream hours per day, and favorite genre as predictors. The anxiety is measured on a scale from 0 to 10 which is a continuous variable.

Table 1: Regression Estimates for Anxiety

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| (Intercept) | 5.9172564 | 0.4543744 | 13.0228639 | 0.0000000 |
| Fav.genreCountry | 0.4500753 | 0.6645997 | 0.6772126 | 0.4984895 |
| Fav.genreEDM | 0.3360083 | 0.5905146 | 0.5690093 | 0.5695280 |
| Fav.genreFolk | 1.6446750 | 0.6256885 | 2.6285842 | 0.0087577 |
| Fav.genreGospel | 1.2752233 | 1.2077704 | 1.0558491 | 0.2913928 |
| Fav.genreHip hop | 1.0580201 | 0.6031814 | 1.7540661 | 0.0798464 |
| Fav.genreJazz | 0.8619450 | 0.7237705 | 1.1909094 | 0.2340833 |
| Fav.genreK pop | 0.9466121 | 0.6601745 | 1.4338817 | 0.1520420 |
| Fav.genreLatin | -1.1068847 | 1.6316088 | -0.6784008 | 0.4977365 |
| Fav.genreLofi | 0.9696890 | 0.9457035 | 1.0253626 | 0.3055379 |
| Fav.genreMetal | 0.7924016 | 0.4767707 | 1.6620183 | 0.0969462 |
| Fav.genrePop | 1.0548025 | 0.4559532 | 2.3134009 | 0.0209826 |
| Fav.genreR&B | 0.2183809 | 0.5967306 | 0.3659622 | 0.7145012 |
| Fav.genreRap | -0.1052888 | 0.7003426 | -0.1503390 | 0.8805395 |
| Fav.genreRock | 1.2982717 | 0.4270702 | 3.0399489 | 0.0024523 |
| Fav.genreVideo game music | 0.7925300 | 0.5599101 | 1.4154594 | 0.1573680 |
| Age | -0.0443401 | 0.0088467 | -5.0120531 | 0.0000007 |
| Hours.per.day | 0.0459456 | 0.0342034 | 1.3433037 | 0.1795987 |

This table shows the regression estimates. The equation for my linear regression model here is:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

Where the intercept is $\beta_0$ and the rest estimates corresponds to each $\beta_i$, where i = 2…k. From the p value, we see that only Folk, Pop, and Rock music and age are less than the significant level of 0.05 which means they have significantly higher Anxiety scores. Other predictors are non-significant. To test if non-significant variables should be removed, I tested with a reduced model based on folk, pop, and rock genre and age as predictors. I compared the AIC value between the original model and reduced model, the difference is 0.152 which is small. So, those non-significant explanatory variables should be removed from the model. Then, I fitted multiple linear regression for each mental health problems one by one with the predictor variables besides favorite genre. Here is the table of regression estimates.

Table 6: Combined Regression Estimates for All Mental Health Models

| Model | term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|---|
| Anxiety | (Intercept) | 5.854 | 1.010 | 5.795 | 0.000 |
| Anxiety | Age | -0.038 | 0.008 | -4.452 | 0.000 |
| Anxiety | Hours.per.day | 0.038 | 0.034 | 1.133 | 0.257 |
| Anxiety | Instrumentalist | 0.035 | 0.239 | 0.146 | 0.884 |
| Anxiety | Composer | -0.150 | 0.296 | -0.506 | 0.613 |
| Anxiety | Music.effectsNo effect | 0.156 | 0.991 | 0.157 | 0.875 |
| Anxiety | Music.effectsImprove | 0.997 | 0.976 | 1.021 | 0.307 |
| Anxiety | Music.effectsWorsen | 1.669 | 1.174 | 1.421 | 0.156 |
| Depression | (Intercept) | 4.812 | 1.100 | 4.375 | 0.000 |
| Depression | Age | -0.028 | 0.009 | -3.006 | 0.003 |
| Depression | Hours.per.day | 0.102 | 0.037 | 2.777 | 0.006 |
| Depression | Instrumentalist | -0.157 | 0.261 | -0.603 | 0.547 |
| Depression | Composer | 0.324 | 0.322 | 1.006 | 0.315 |
| Depression | Music.effectsNo effect | -0.043 | 1.079 | -0.040 | 0.968 |
| Depression | Music.effectsImprove | 0.357 | 1.063 | 0.336 | 0.737 |
| Depression | Music.effectsWorsen | 2.709 | 1.279 | 2.118 | 0.034 |
| Insomnia | (Intercept) | 1.888 | 1.131 | 1.670 | 0.095 |
| Insomnia | Age | 0.005 | 0.009 | 0.520 | 0.604 |
| Insomnia | Hours.per.day | 0.134 | 0.038 | 3.550 | 0.000 |
| Insomnia | Instrumentalist | 0.007 | 0.268 | 0.028 | 0.978 |
| Insomnia | Composer | 0.632 | 0.331 | 1.909 | 0.057 |
| Insomnia | Music.effectsNo effect | 1.112 | 1.109 | 1.003 | 0.316 |
| Insomnia | Music.effectsImprove | 1.132 | 1.093 | 1.036 | 0.301 |
| Insomnia | Music.effectsWorsen | 2.082 | 1.315 | 1.583 | 0.114 |
| OCD | (Intercept) | 2.554 | 1.037 | 2.461 | 0.014 |
| OCD | Age | -0.028 | 0.009 | -3.256 | 0.001 |
| OCD | Hours.per.day | 0.107 | 0.035 | 3.088 | 0.002 |
| OCD | Instrumentalist | 0.026 | 0.246 | 0.106 | 0.916 |
| OCD | Composer | -0.084 | 0.304 | -0.275 | 0.783 |
| OCD | Music.effectsNo effect | 0.199 | 1.018 | 0.195 | 0.845 |
| OCD | Music.effectsImprove | 0.473 | 1.003 | 0.472 | 0.637 |
| OCD | Music.effectsWorsen | 0.913 | 1.206 | 0.757 | 0.449 |

Based on this regression result, for anxiety, only age has the p value that is smaller than the significant level 0.05. I applied a reduced model again with age and anxiety only, the result for AIC value for reduced model is 3.88 higher. This suggests that including the additional predictors provides a better fit, so the non-significant variables should not be removed. For depression, the age and hours.per.day are significant within the model, so I applied reduced model again, the AIC difference is 4.786 which the original model has a smaller value, so we prefer the original model. For insomnia, I applied reduced model with the significant variable hours.per.day and the AIC difference between is 4.72 and the reduced model is smaller, so we can remove the non-significant variable here. The last for OCD, I applied reduced model with significant variable age and hours.per.day and the AIC difference is 8.009 with reduced model smaller than the original model. So here we can remove the non-significant variables. I tested the normality of residuals based on the Shapiro-Wilk test. However, all the mental health problem are having a p-value that is lower than the significant level 0.05, which means that their residuals are not from a normal distribution. I applied a log transformation to try to fix the normality, but the result shows that the p values are still smaller than 0.05. To check if the residual variance is constant, I applied the Breusch-Pagan Test. The result shows that only the anxiety model with a p value bigger than 0.05 is homoscedasticity for the residual variance. Other three mental health variables all have non-constant variance.

**Discussion and Conclusion**

From the analysis, several insights into the relationship between music consumption and mental health were founded. The regression analysis indicated that preferences for Folk, Pop, and Rock music genres were significantly associated with higher self-reported anxiety scores. This finding might suggest that emotional music, which is often found in these genres, may evoke or resonate with anxious feelings. There is also a key factor which is age that is involved with different mental health problems. Age emerged as a significant predictor across four mental health conditions, suggesting how individuals perceive and are affected by music. In the results, the coefficient for age indicates a negative relationship, which means that as age increases, the levels of anxiety, depression, and OCD decrease. Depression was significantly linked to daily music listening hours, suggesting that excessive engagement with music might correlate with more depressive symptoms. Insomnia and OCD also demonstrated associations with hours spent listening, but their reduced models provided a better fit. While the regression models provided valuable insights, residual analysis gives a more detailed understanding on the model assumptions. Residuals for all mental health outcomes except anxiety were not normal, even after log transformations. Non-constant variance was observed in models for depression, insomnia, and OCD, indicating some issues with data variability. For music effect, participants' perceived effects of music on their mental health did not yield significant results in the models, suggesting that subjective experiences might not align with mental health outcomes.

For future studies, we might want to use more advanced statistical methods to address those issues such as the heteroscedasticity and residual normality. We could also explore more data on music attributes such as tempo, lyrics, and emotions. There are also limitations on this analysis. Larger and more diverse data could improve the generalizability of findings. Self-reported data contains bias, which could be over or underestimation of mental health symptoms or listening habits. This research shows the powerful connection between music and mental health, highlighting its potential for therapeutic applications and the importance of personalized approaches. Future studies can build on these findings by leveraging more diverse datasets and advanced analytical techniques, ultimately deepening our understanding of music as a tool for mental well-being.

**Reference**

Gustavson, Daniel E., et al. "Mental Health and Music Engagement: Review, Framework, and Guidelines for Future Studies." *Translational Psychiatry*, vol. 11, no. 370, 2021.