Boston University

Music and Depression:

DS110 Project

Nora Amer,

Kyosuke Asano,

Mark Castro, and

Srushti Karve

Professor Gold

DS110

8 December 2023

**Introduction**

According to Mental Health America, depression is one of the most common mental disorders with 8% (21 million) of American adults suffering from it annually (Mental Health America). The term “depression” emerged in the 19th century, and since then, extensive research and numerous treatments have been explored. Nevertheless, many people continue to struggle with depression. Is there anything effective to alleviate depression? In an effort to find a solution to this disorder, we found data on music and mental health. With the help of a dataset created by Catherine Rasgaitis, can we predict a person’s depression level based on their music listening habits? Let’s take this step to discover the impact of music on mental health.

**Prior Work**

Within the twenty-first century, many studies have focused on exploring the relationship between music and depression, other mental illnesses, and even general well-being as well. Many sources focus on recreational listening habits of individuals and trends within that domain. For example, in a conducted study by Sunkyung Yoon and Jonathan Rottenberg at the University of South Florida, 77 participants (39 with clinical depression, and 38 without) were assessed in their listening habits, examining the average tempo within their daily listening, as well as their reasons for listening to music whenever happy or sad. The study reported that on average, the favorite songs of the depressed group within the data had a slower tempo than those of the non-depressed group, and that they on average preferred sadder songs (Yoon). As this study identifies a certain trend, many other researchers have claimed that an individual’s intentions when listening to music have an important impact on their interaction. For example, In multiple studies conducted by Adam J. Lonsdale and Adrian C. North at the Heriot Watt University in Edinburgh, UK, research found that listeners typically listen to music for 6 reasons: Positive mood management, Diversion, Negative mood Management, Interpersonal Relationships, Personal Identity, and Surveillance. Within this phenomenon of mood regulation, research found that listeners were split between listening to music in order to create positive moods, but interestingly to also alleviate and create a space for negative feelings such as anxiety and stress to manifest, in order to cope (Lonsdale). This relation between intention and impact can also be seen in research conducted by Sai Charan Kanagala at University of Social Sciences and Humanities in Warsaw, which investigates the presence of maladaptive music engagement, and its relation to depression symptoms. Essentially, maladaptive strategies can include, “rumination, avoidant coping, and social isolation” as one is listening to music, and according to the study’s results, this maladaptive interaction resulted in an increase of odds in experiencing all of the depression symptoms in its participants [(](https://doi.org/10.1177/20592043211057217)Kanagala). Research has also addressed the much more scientific impact of music on an individual’s mental health. A study by Valorie N. Salimpoor at McGill University found that there was an actual internal impact on levels of dopamine induced in the striatal system of the brain when listening. The study also reported that when individuals were listening to pleasurable music and also neutral music, an increase in heart rate and respiration also occurred in creating an individual's emotional arousal when listening to music (Salimpoor). This increase in dopamine combats the low levels of dopamine that an individual’s depression can cause, and this application can be seen in the efficacy of music therapy. Research by Jin-Liang Wang at Southwest University in Chongqing, compared the alleviation of depression symptoms on a group of college students that went through music therapy, and a group who did not, showing that the ones who did go through music therapy were relieved of their symptoms much more than the group who did not (Wang). In essence, it can be concluded that prior studies known about the impact music has on depression levels are very multifaceted, considering they can vary in focus- such as in a medical sense and habitual sense. There is lots of research on the relationship between the two, but not too much insight on how the extent of one’s depression can vary based on the music listened to, which is the objective of this investigation.

**Methodology**

***Data Preparation***

The first step of data preparation was dropping columns that had irrelevant data, like columns where most of the answers were blank, or one that just gave permission for the study to use their data. Next, all the categorical data was modified so that we could apply machine learning techniques. For all the “Frequency [Genre]” columns, a response of “Never” was given a value of 0, “Rarely” a value of 1, and so on. For all the questions that were Yes/No questions, “Yes” was given a value of 1 and “No” a value of 0. Rows with missing values were deleted, and a parameter was set on the “Average BPM” column so only realistic values were considered. From there, the “Primary streaming service” and “Fav genre” column were one-hot encoded. Finally, all the columns were turned into integers.

***Statistics***

First, a Pearson’s correlation coefficient was conducted upon the variables “Hours per day” of music listened and “Depression” Levels in order to see if there was any specific correlation between the amount of hours listened to and the level of depression an individual would have. This would answer the question of whether listening to a high or low amount of music causes high or low depression levels. Therefore, its null hypothesis (H0) is that there is no impact of the amount of hours of music listened to on an individual’s depression levels. A one-way ANOVA test was also conducted with depression levels measured upon the frequencies of 15 different music genres listened to. This ANOVA test would hopefully determine if there were any statistical significance between the frequency of a genre listened and its respective depression level. Therefore, its null hypothesis (H0) is that there is no significant difference between the average depression levels of every genre and that they do not differ in any way.

***Machine Learning***

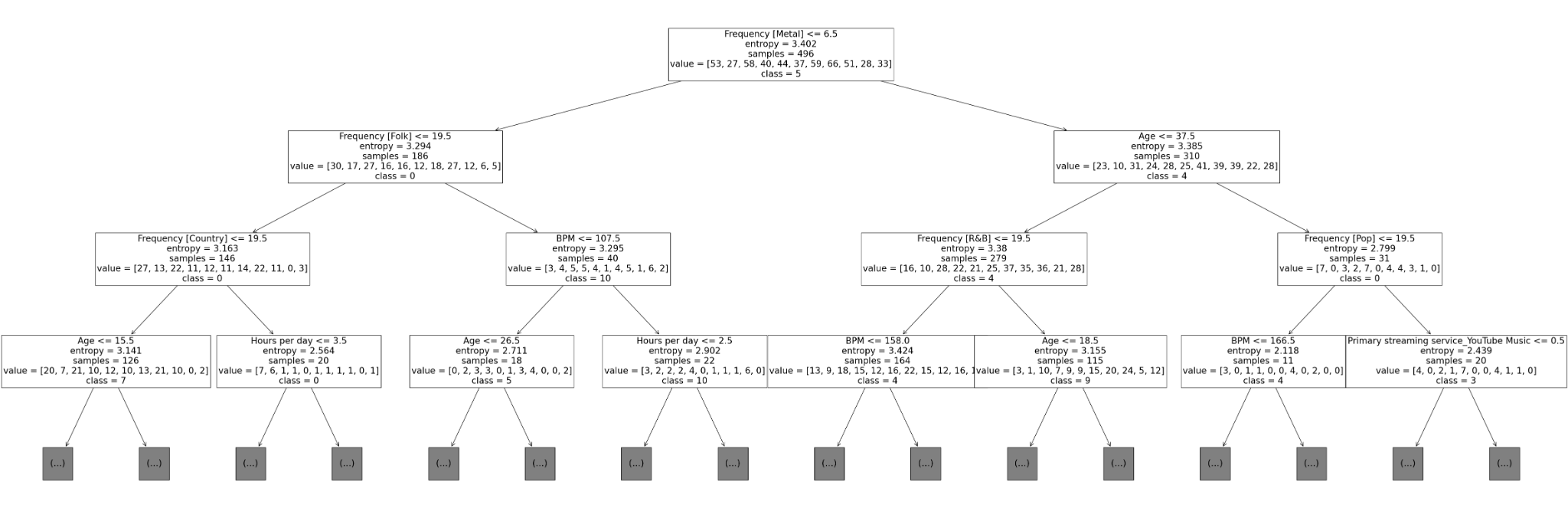
For both the random forest and decision tree, levels of depression are being predicted and almost every other column was used to predict depression. In order to optimize the tree to give the highest score, we iterated over many different combinations of max tree depth and max leaves. This function was run many times, and the most common values were max\_tree\_depth = 4 and max\_leaves = 40. These parameters were used for both methods.

**Results**

***Statistics***

When computing a Pearson's correlation coefficient with the variables “Hours per day” of music listened and “Depression” Levels, the correlation coefficient between the two returned to be only 0.11. For a Pearson’s correlation coefficient, the closer the value is to 1, the more correlated two variables are. As this value is very low, our data is not statistically significant and cannot reject the null hypothesis that the hours of music listened per day has no impact on one’s depression levels. Within the one-way ANOVA test, after evaluating on the average depression levels of the frequencies of 15 different music genres, the ANOVA test returned a p-value of 0.017, which for a ANOVA test, if its p-value < 0.05, the data is statistically significant. Since the data is statistically significant, the null hypothesis can be rejected that every genre has the same mean depression level, and that they do not differ in any way. As it was rejected, it can be said that no two means of the depression levels of the genres are the same, meaning that each genre has its own sort of impact (Hamel).

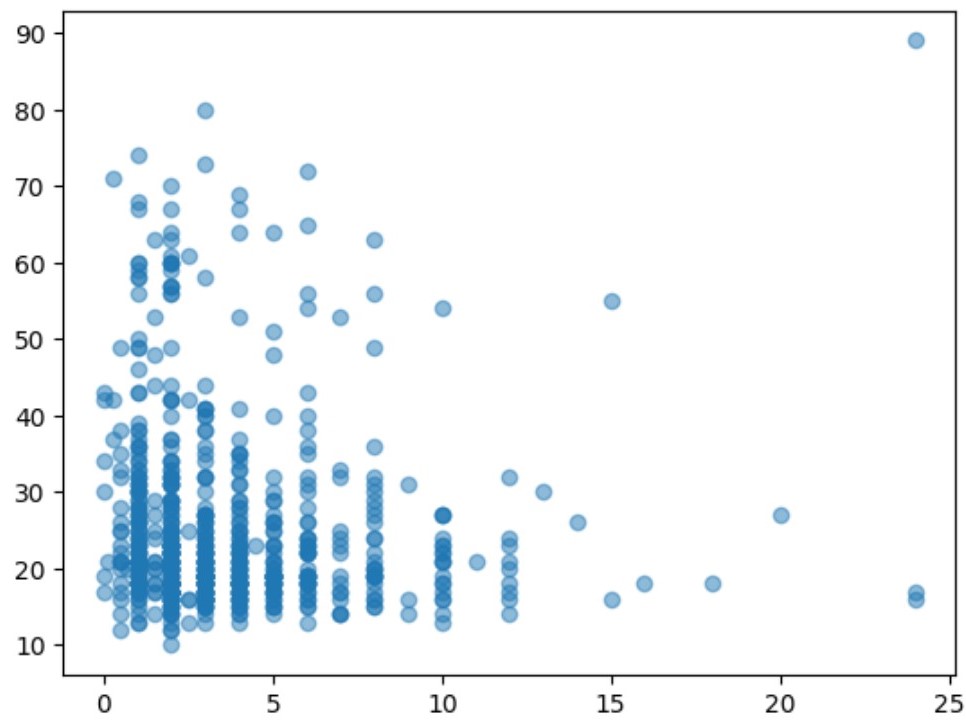
***Machine Learning***

In order to classify different respondents on a depression scale ranging from 0-10, a decision tree was utilized. 80% of our data was used to train the model, and 20% was used to test. When scoring the model on the training data, the score was a measly 0.25. The model is not very good at classifying the data. The score for the test data is 0.129, which is extremely low. The first three rows of the decision tree are represented below. The model is not able to classify the data, which means that it’s possible that music listening habits and depression levels are not closely related.

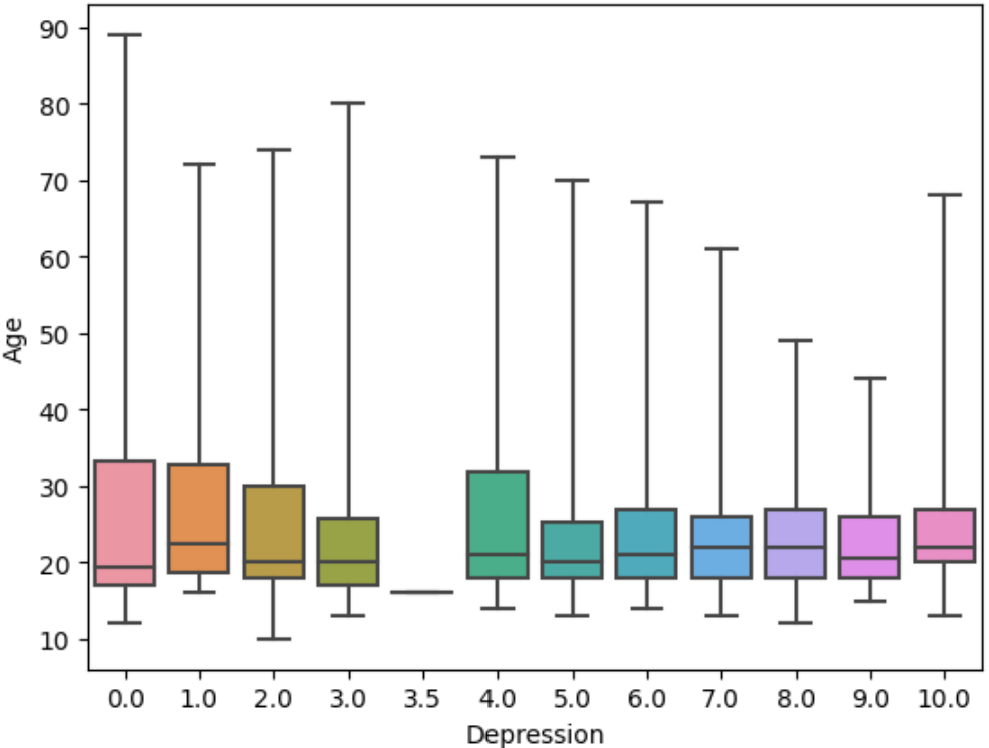
To minimize possible errors, a random forest was also utilized on the same data set, using the same method to optimize parameters. The score was a bit higher, at 0.137. Because of the overall low performance of these models, it is appropriate to say that music listening habits are not related to depression levels close enough to use machine learning.

***Visualizations***

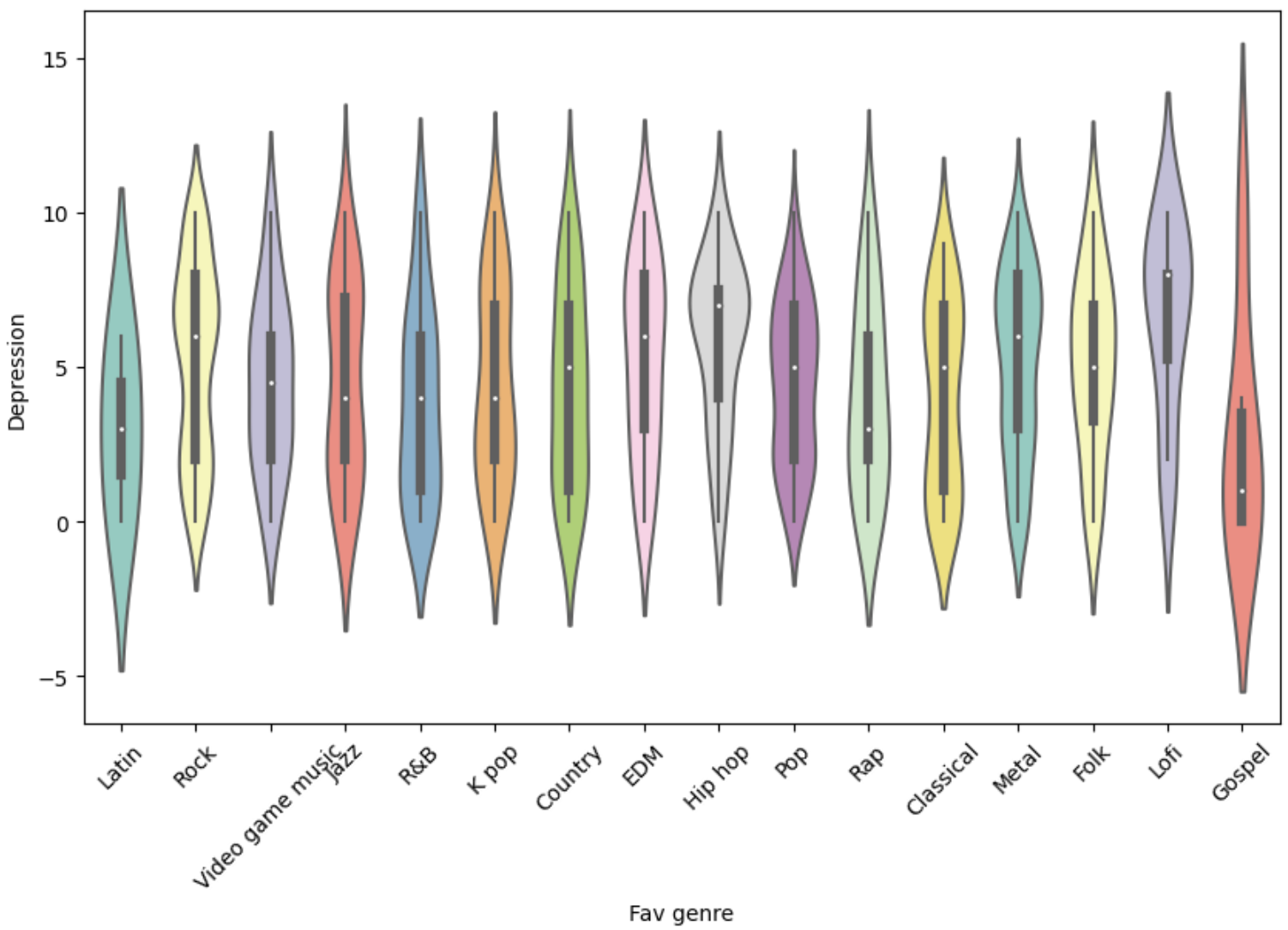
In order to understand our dataset better, we graphed a Scatter Plot using Matplotlib that represents the correlation between the ‘Hours per day’ that an individual listens to music for, and their ‘Age.’

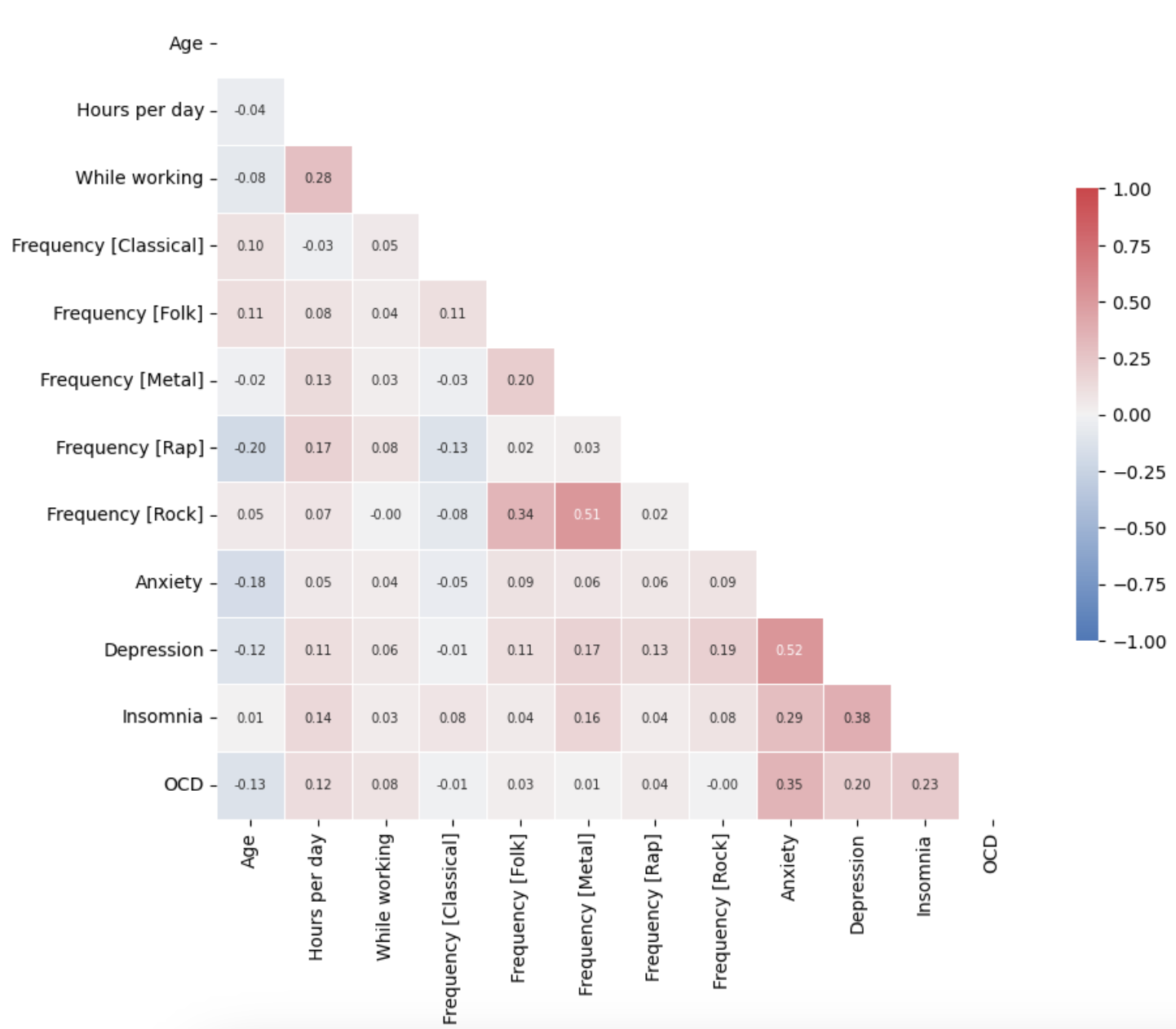


Additionally, we also used a Seaborn Box and Whisker Plot to visualize the relationship between the levels of ‘Depression’ in a person and their ‘Age.’

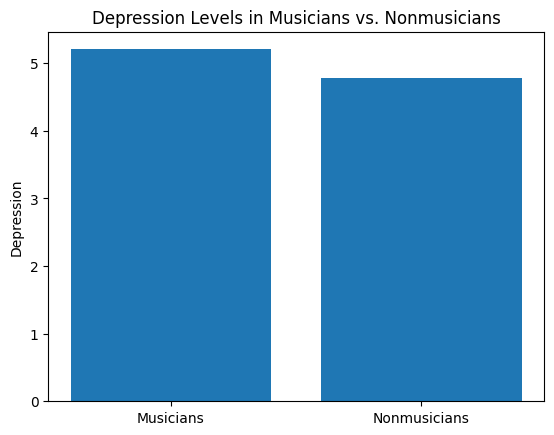


Moreover, we used a Seaborn Violinplot to see the relationship between the levels of ‘Depression’ in a person and their ‘Favorite genre.’



Then, we visualized the correlation coefficients between each column using Seaborn Heatmap, making it visually understandable. 

Finally, using a bar chart, we compared average depression rates between Musicians who answered yes to the questions “Are you an instrumentalist?” and “Are you a composer?” vs non-musicians to see if there was a significant difference between the two groups.



**Conclusion**

Unfortunately, we could not find a relationship between music listening habits and depression levels from our analysis using machine learning. However, it may be premature to conclude that there is no connection between them. According to the description on Kaggle, the data we used was collected from the results of Google Forms distributed on social media platforms like Discord. This raises the question of whether we are truly capturing a random sampling of individuals. Do people with depression use such platforms? Furthermore, do people with depression actively choose to listen to music? In order to do a detailed analysis of the data, it might be necessary to obtain data from a place where there are many people with depression, such as a mental hospital. For instance, if we had data that tracked the same depression patient over a long period, showing how music affects them over time, it would enable us to do a deep-dive into the impact of music. Through this research, we’ve learned the importance of considering how data is sourced and the characteristics of the subjects it represents.

Works Cited

“Depression.” *Mental Health America*, mhanational.org/conditions/depression#:~:text=Major%20depression%20is%20one%20of,of%20American%20adults%20each%20year. Accessed 8 Dec. 2023.

Hamel, Greg. “Python for Data 26: ANOVA.” *Kaggle*, Kaggle, 17 Dec. 2020, www.kaggle.com/code/hamelg/python-for-data-26-anova/notebook.

Kanagala, S. C., Schäfer, T., Greenberg, D. M., & Gabińska, A. (2021). Depression Symptoms are Linked to Music Use. Music & Science, 4. <https://doi.org/10.1177/20592043211057217>

Lonsdale, A.J. and North, A.C. (2011), Why do we listen to music? A uses and gratifications analysis. British Journal of Psychology, 102: 108-134. <https://doi.org/10.1348/000712610X506831>

“Plotting a Diagonal Correlation Matrix#.” *Plotting a Diagonal Correlation Matrix - Seaborn 0.13.0 Documentation*, seaborn.pydata.org/examples/many\_pairwise\_correlations.html. Accessed 8 Dec. 2023.

Rebecchini, Lavinia. “Music, mental health, and immunity.” Brain, behavior, & immunity - health vol. 18 100374. 21 Oct. 2021, doi:10.1016/j.bbih.2021.100374

Rasgaitis, Catherine. “Music & Mental Health Survey Results.” *Kaggle*, 21 Nov. 2022, www.kaggle.com/datasets/catherinerasgaitis/mxmh-survey-results/code.

Salimpoor, V., Benovoy, M., Larcher, K. *et al.* Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat Neurosci* 14, 257–262 (2011). <https://doi.org/10.1038/nn.2726>

“Violinplot from a Wide-Form Dataset#.” *Violinplot from a Wide-Form Dataset - Seaborn 0.13.0 Documentation*, seaborn.pydata.org/examples/wide\_form\_violinplot.html. Accessed 8 Dec. 2023.

Wang, Jin-Liang. (2011). Impact of group music therapy on the depression mood of college students. Health. 03. 151-155. 10.4236/health.2011.33028.

Yoon S, Rottenberg J. Listening to the blues: An ecological momentary assessment of music choice in depression. Emotion. 2021 Sep;21(6):1177-1187. doi: 10.1037/emo0000911. Epub 2020 Nov 30. PMID: 33252935.