Predictors of Music's Impact on Mental Health

Present By Arshad, Hao Yu, Jae Hee, Qi Heng

Data Set & Objectives

Music & Mental Health Survey Results

- Objective: To develop a machine learning model that determines which factors best impact mental health.
- Context: Understanding the role of music in mental well-being, particularly for young adults.
- Potential Applications: Music therapy, personalised music recommendations for mental health improvement

List of Contents

1.INITIAL DATA PREPARATION

Removed unnecessary columns like “Primary streaming service” and “Permissions”

Rows with missing values are removed to ensure data integrity.

```
# Create a copy of the original DataFrame
df_clean = df.copy()

# Remove unwanted columns
df_clean.drop(['Primary streaming service', 'Permissions', 'Timestamp','Instrumentalist', 'Composer', 'While working',

# Remove rows with any missing values
df_clean.dropna(inplace=True)
```

Filtered the data to include only relevant data

```
#restrict data using age: 18-30
df_clean = df_clean[(df_clean['Age'] >= 18) & (df_clean['Age'] <= 30)]
df_clean
```

Combined categorical data for better data analysis

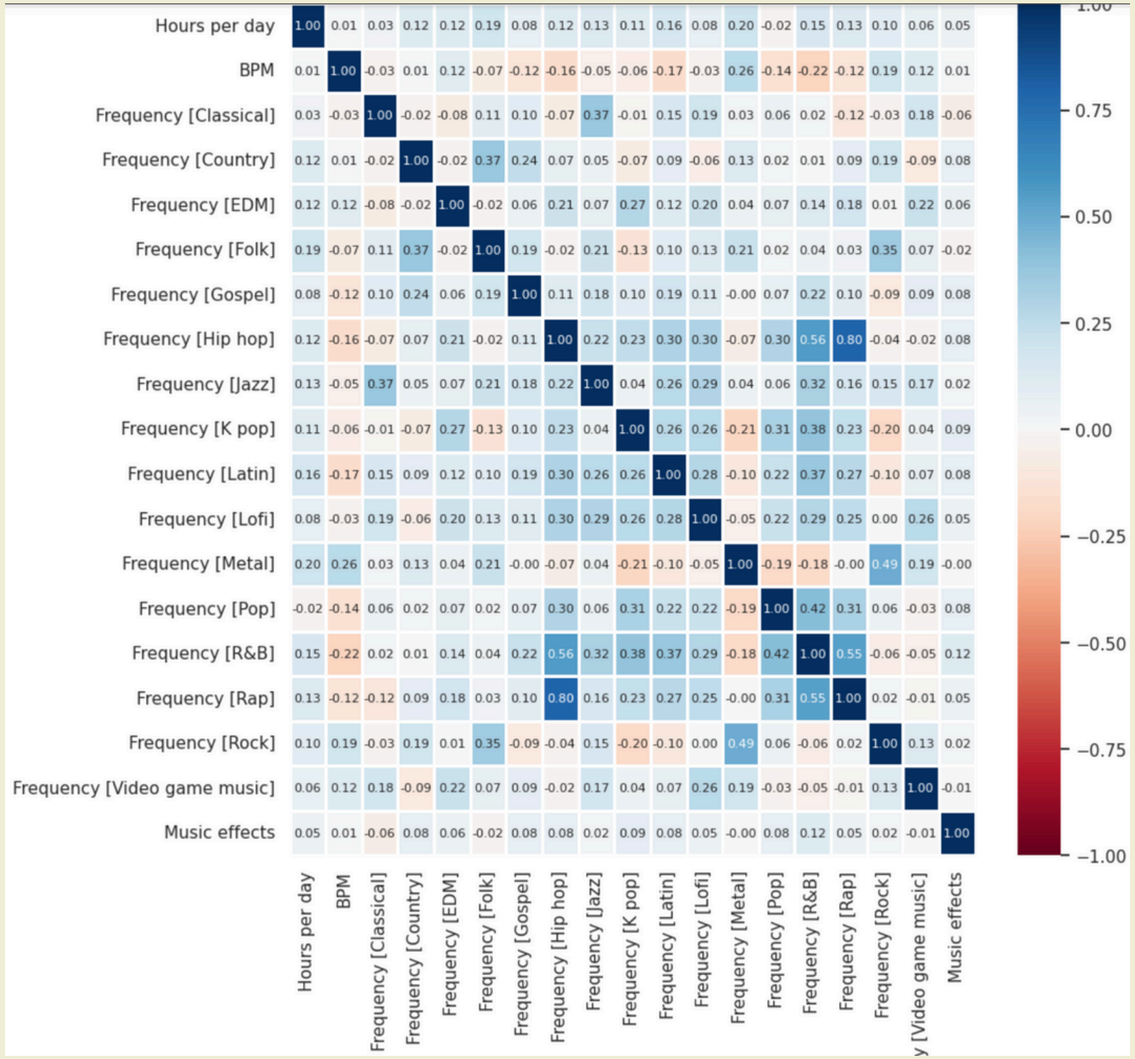
```
# Check Counts and Relative Frequency of Values for a feature
df_clean['Music effects'].value_counts(normalize=True)
print(df_unencoded['Music effects'].value_counts())
print("\n")

#combine no effect and worsen to worsen
df_clean['Music effects'] = df_clean['Music effects'].replace(['No effect', 'Worsen'], 'Worsen')
#assign 1 to improve and 0 to worsen
df_clean['Music effects'] = df_clean['Music effects'].replace({'Improve': 1, 'Worsen': 0})
df_clean
```

```
Improve      471
No effect    138
Worsen        15
```

EXPLORATORY DATA ANALYSIS AND OBSERVATION

EXPLORATORY DATA ANALYSIS AND OBSERVATION

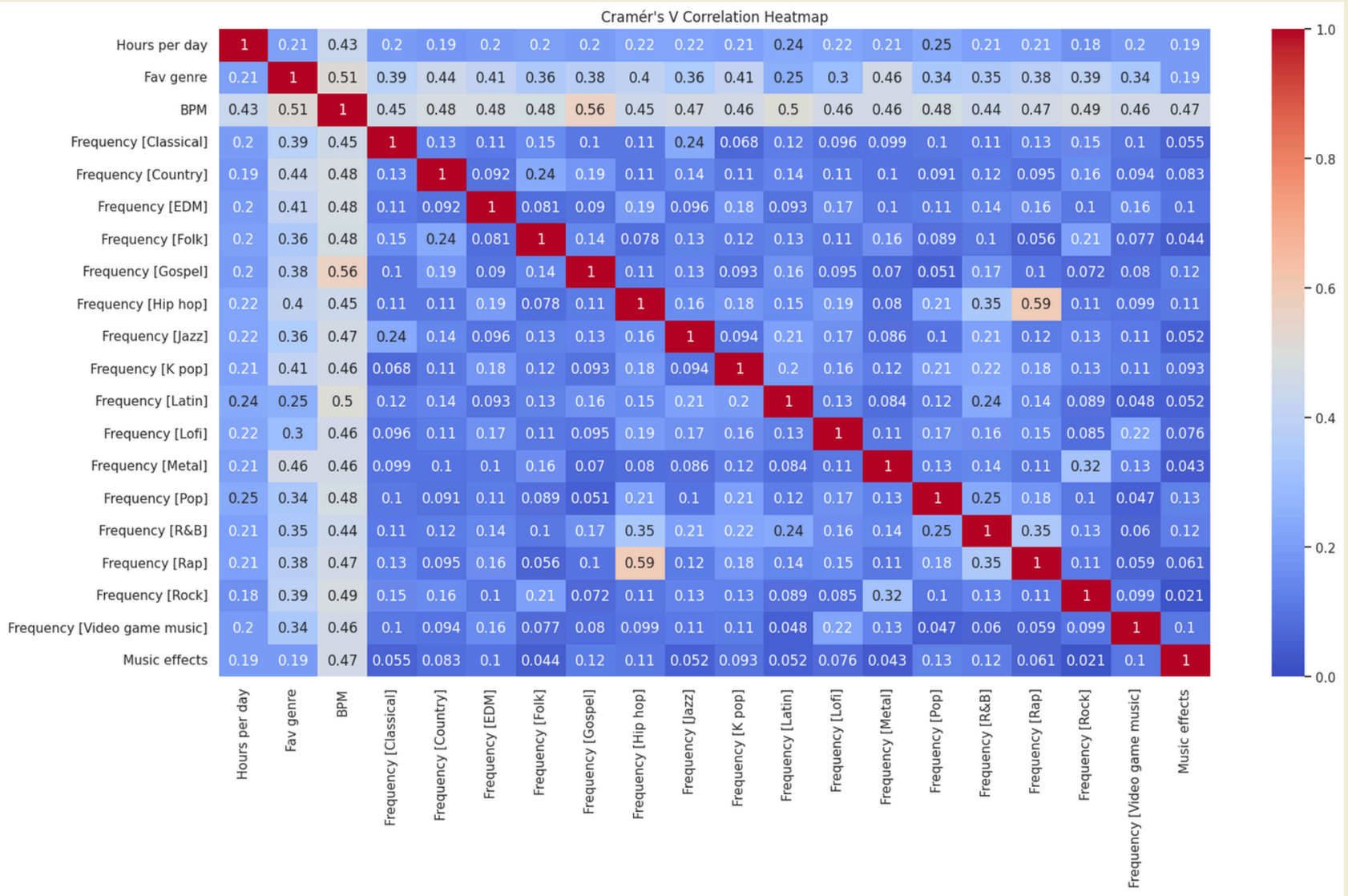
Correlation Matrix



Due to the lack of insightful data analysis from normal correlation, We'll try for different approach with Cramér's V

EXPLORATORY DATA ANALYSIS AND OBSERVATION

Cramér's V



We chose Cramér's V as it measures the strength of association between two nominal variables.

As seen from the plotted graph, We'll be focusing on BPM, Hours per day, Frequency [Gospel],Frequency [Hip hop], Frequency [Pop], Frequency [R&B] that has the highest association with music effects

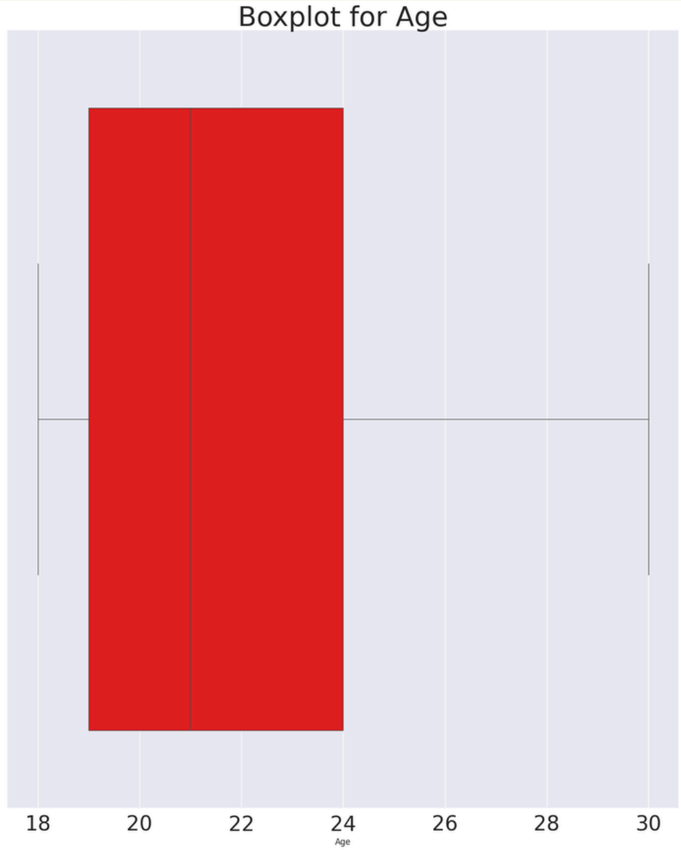
EXPLORATORY DATA ANALYSIS AND OBSERVATION

Statistical Distribution of numerical variables

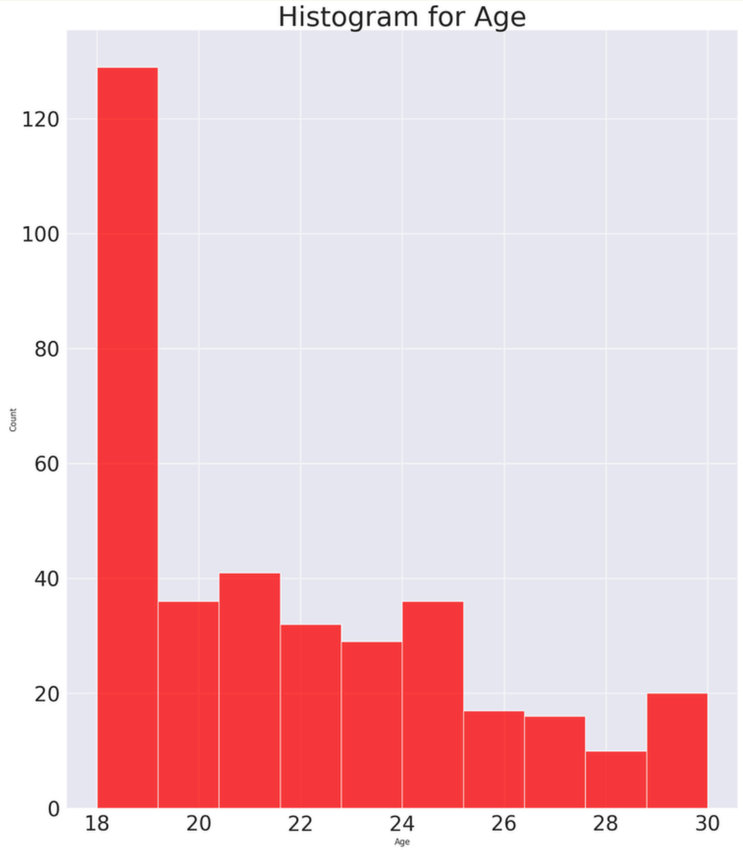
	Age float64	Hours per day flo...	BPM float64
count	366	366	366
mean	21.73497268	3.70273224	123.1584699
std	3.412825267	2.817653385	32.59728574
min	18	0.5	20
25%	19	2	99
50%	21	3	120
75%	24	5	145
max	30	18	220

EXPLORATORY DATA ANALYSIS AND OBSERVATION

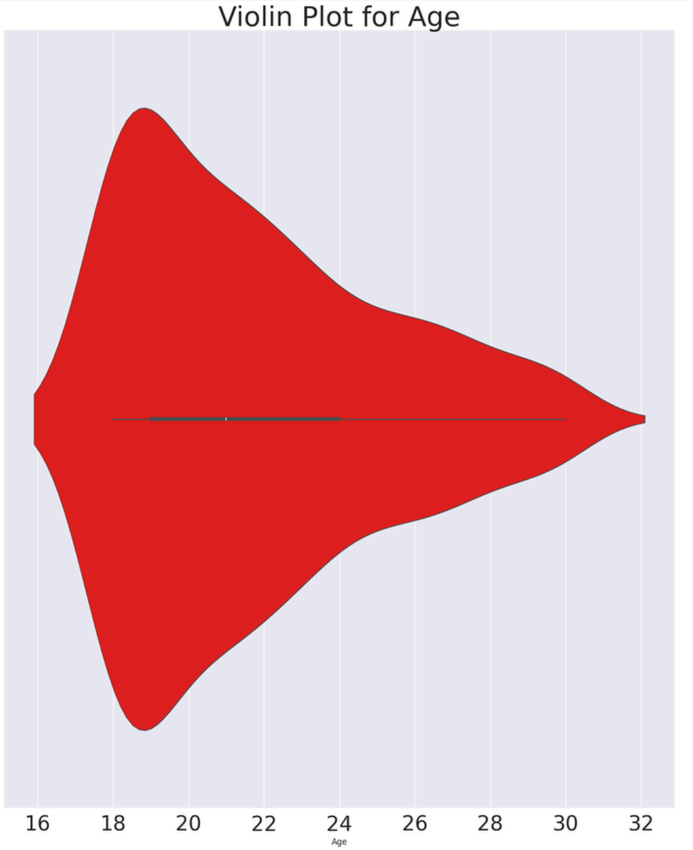
Age



Box Plot



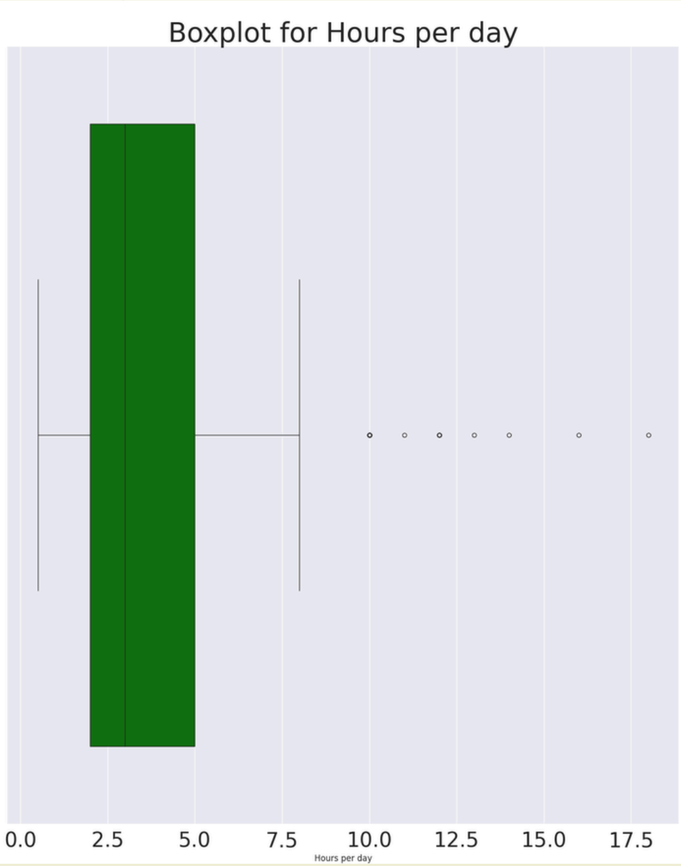
Histrogram



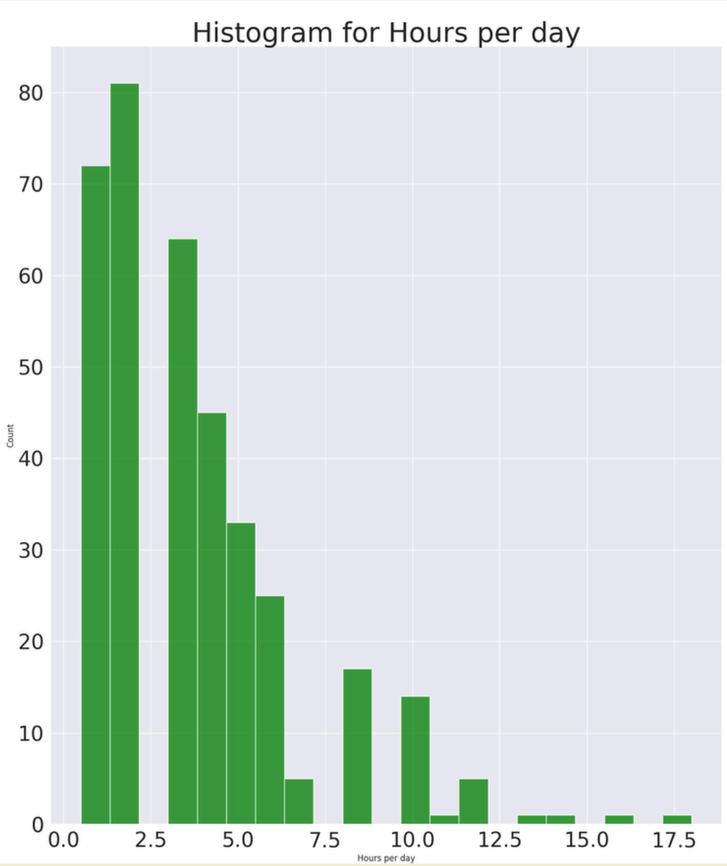
Violin Plot

EXPLORATORY DATA ANALYSIS AND OBSERVATION

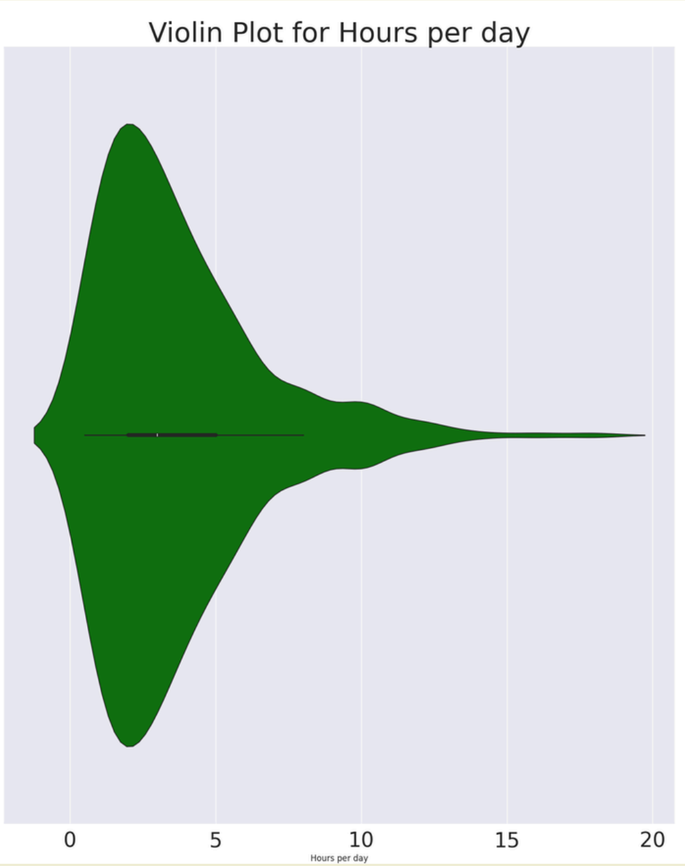
Hours per day



Box Plot



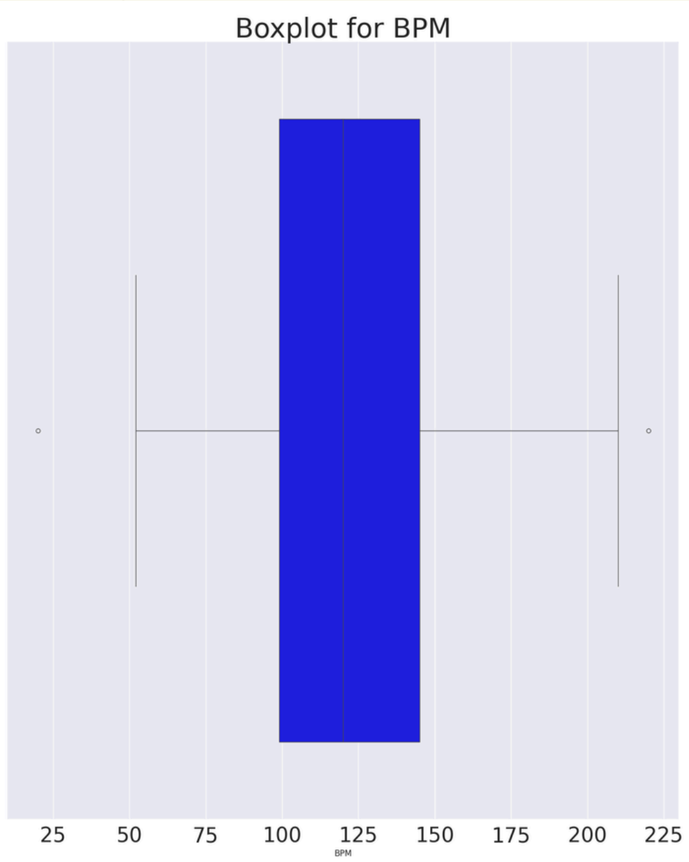
Histrogram



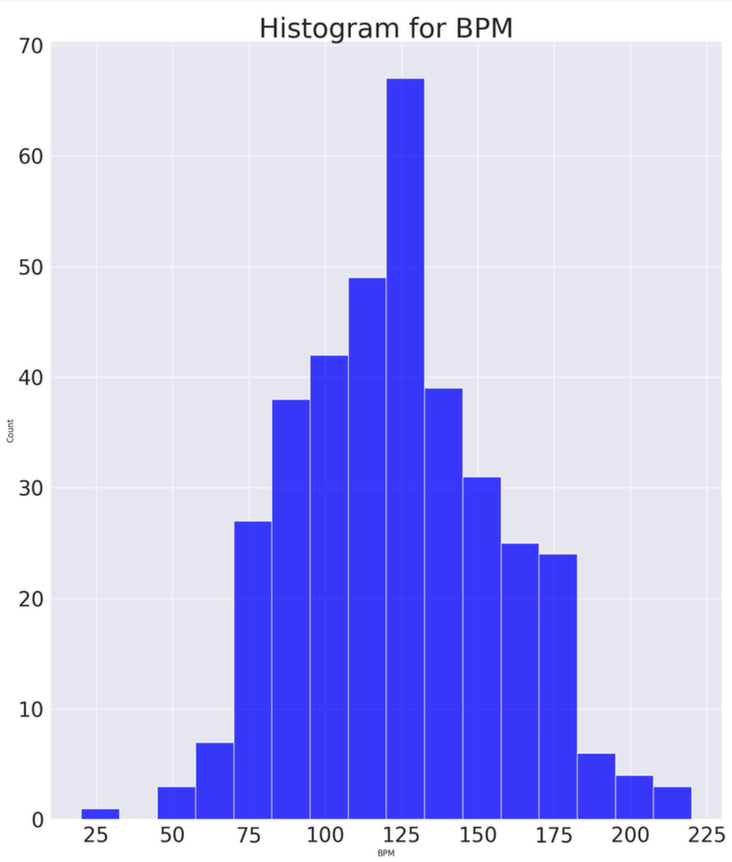
Violin Plot

EXPLORATORY DATA ANALYSIS AND OBSERVATION

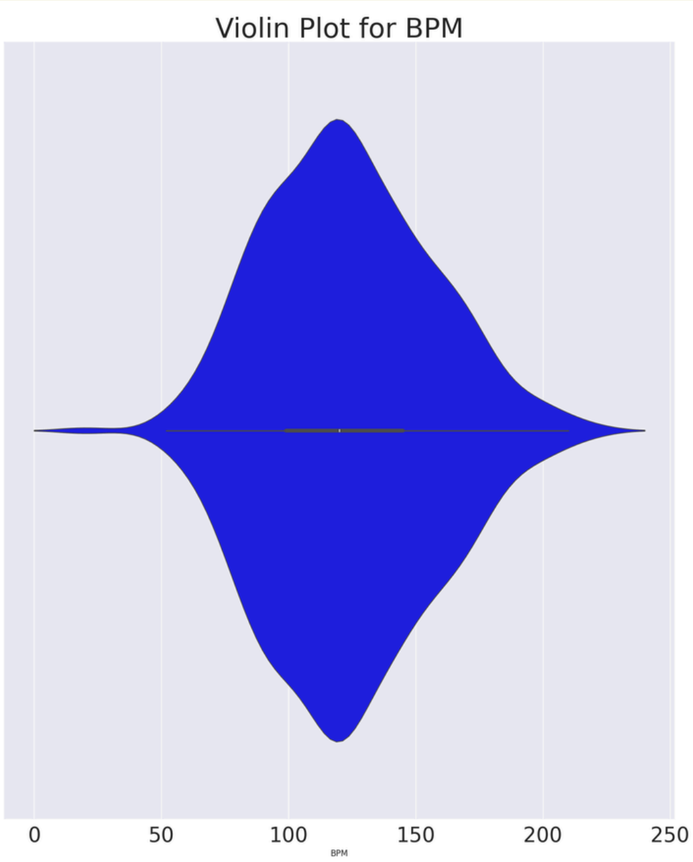
BPM



Box Plot



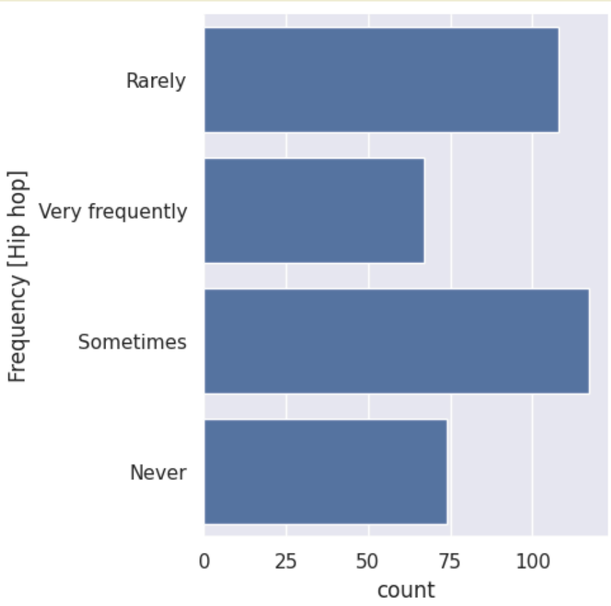
Histrogram

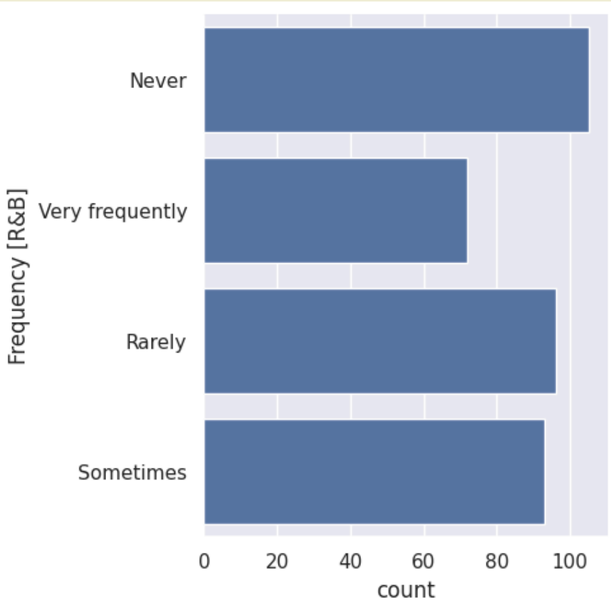


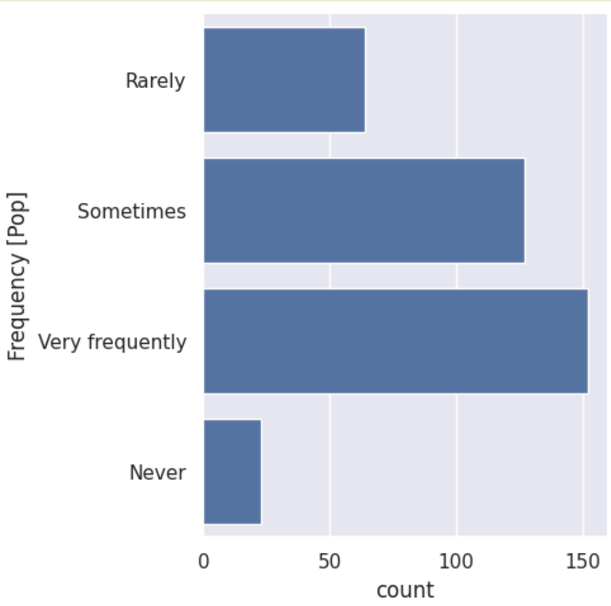
Violin Plot

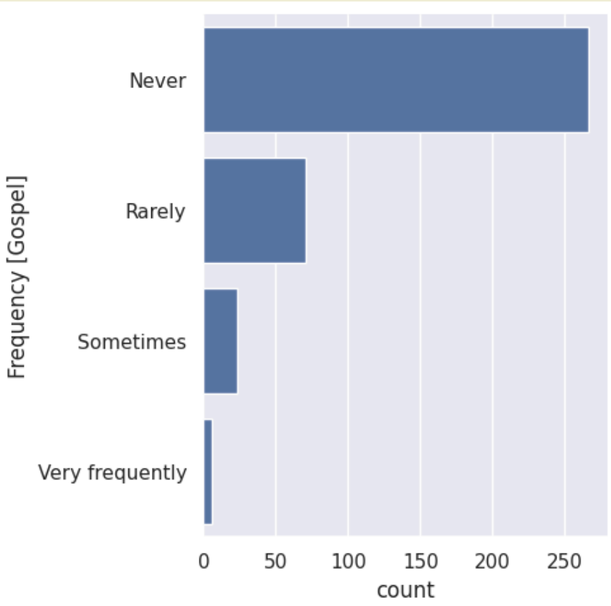
EXPLORATORY DATA ANALYSIS AND OBSERVATION

CATPPLOT



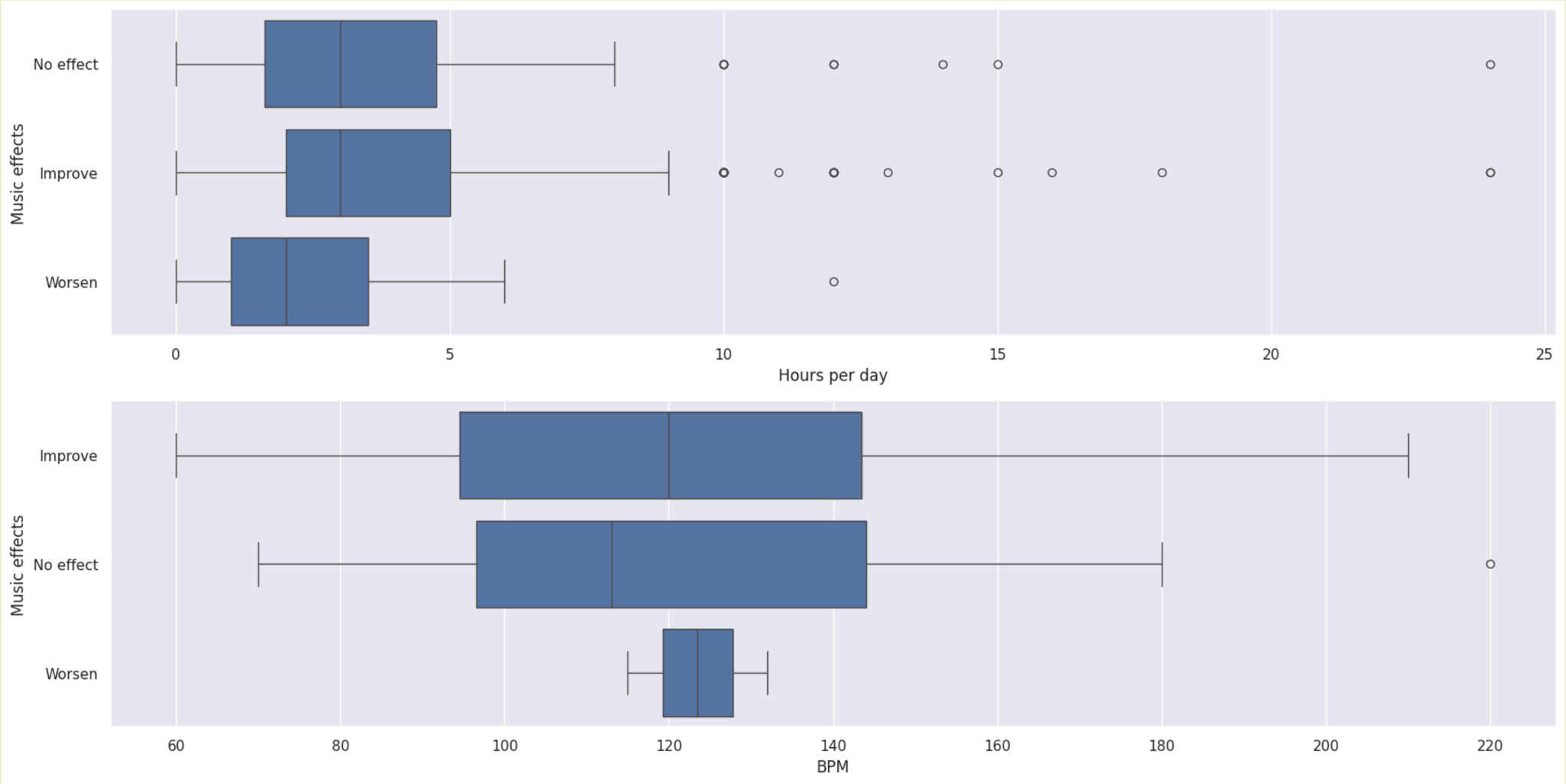






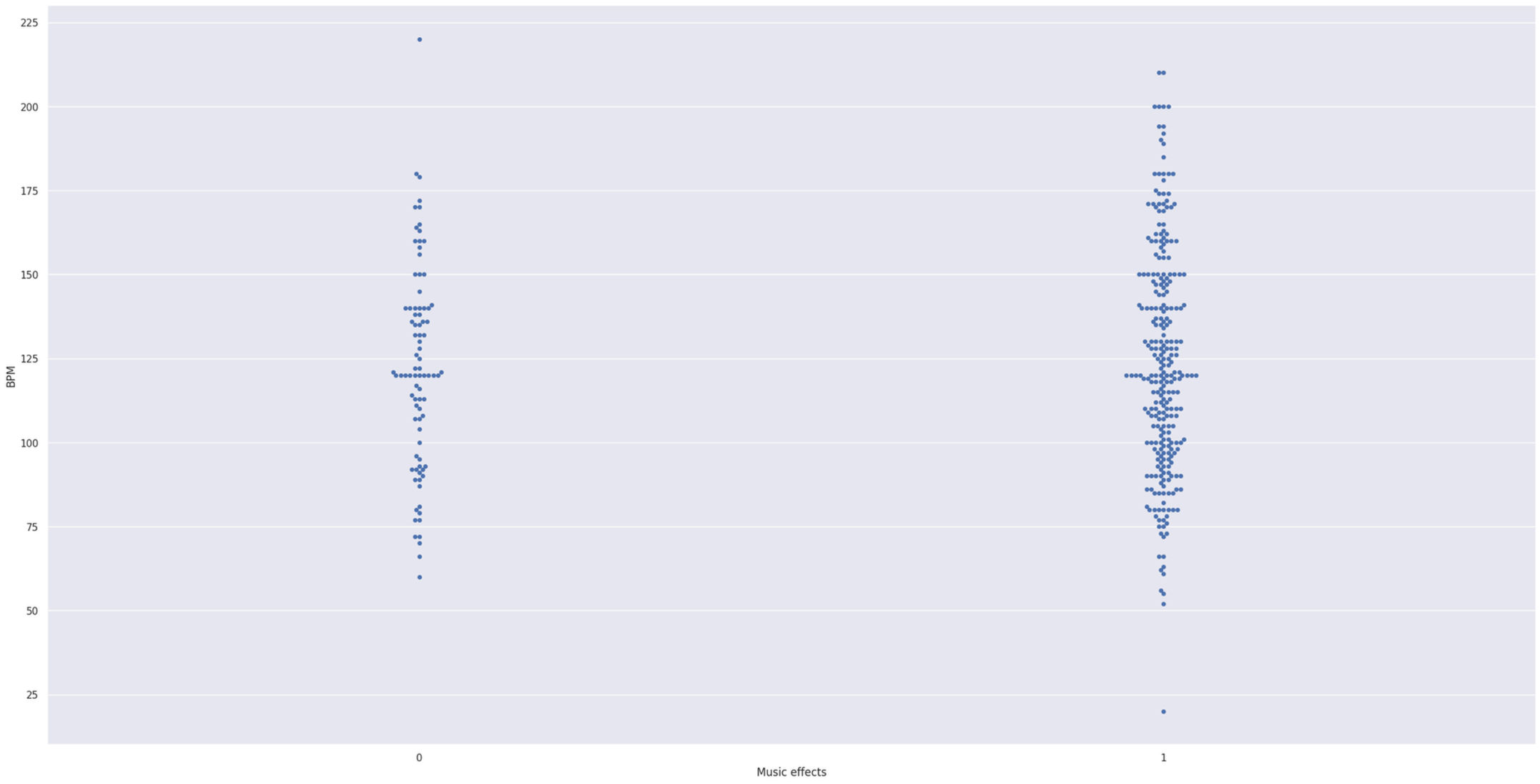
EXPLORATORY DATA ANALYSIS AND OBSERVATION

Boxplot on Music Effects



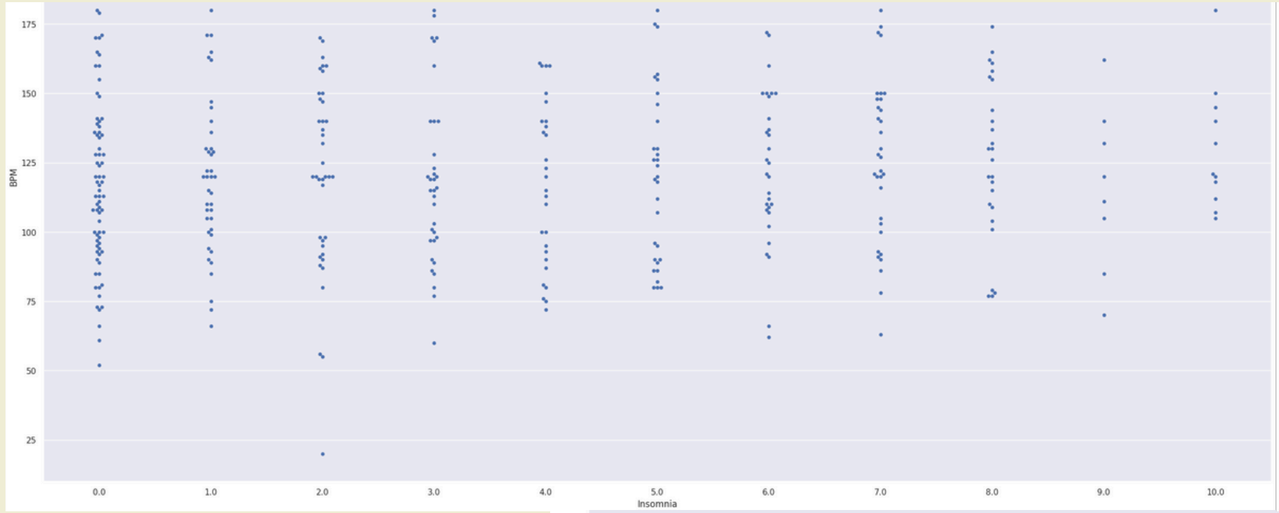
EXPLORATORY DATA ANALYSIS AND OBSERVATION

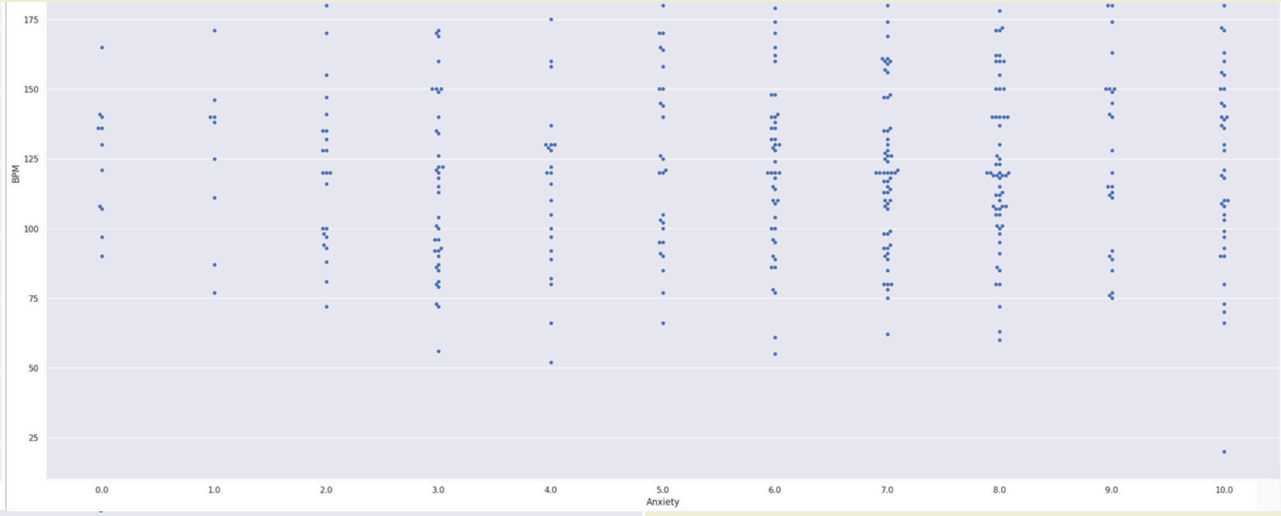
BPM & music effects Swarmplots

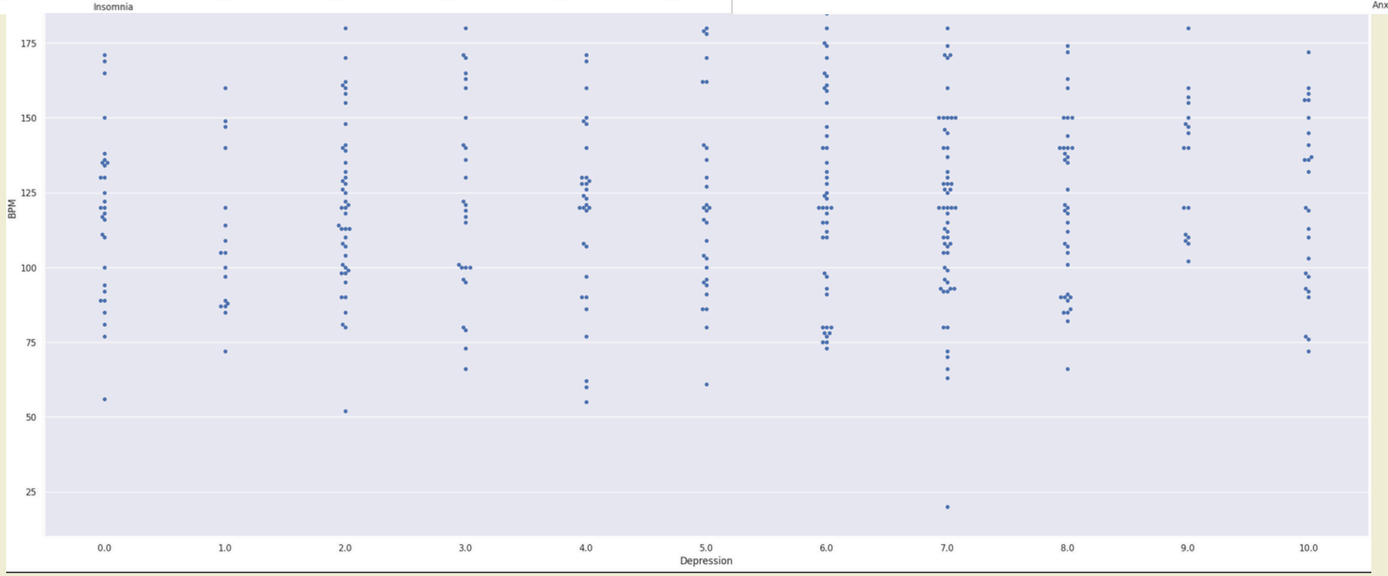


EXPLORATORY DATA ANALYSIS AND OBSERVATION

BPM Swarmplots on types of mental health

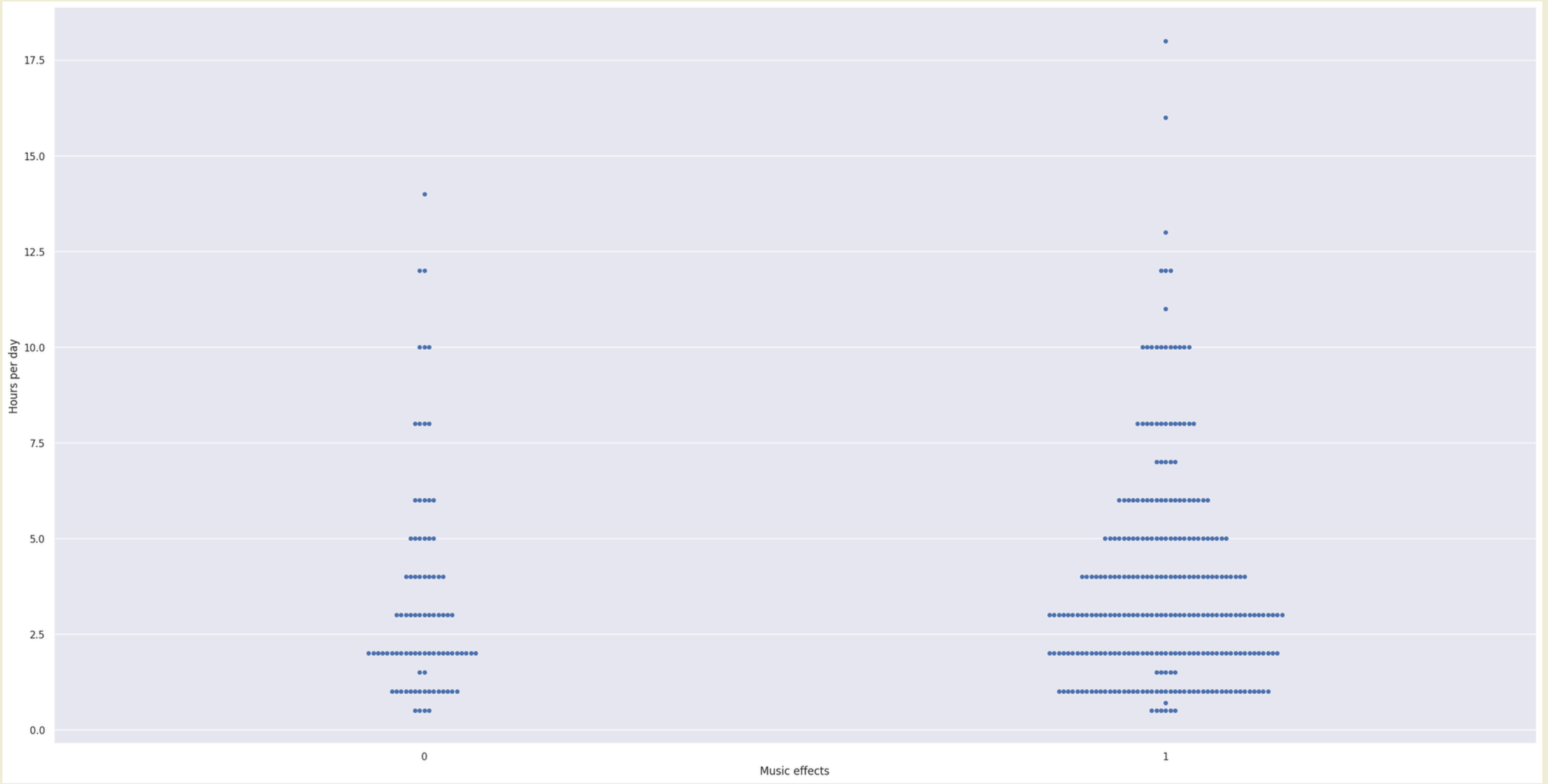






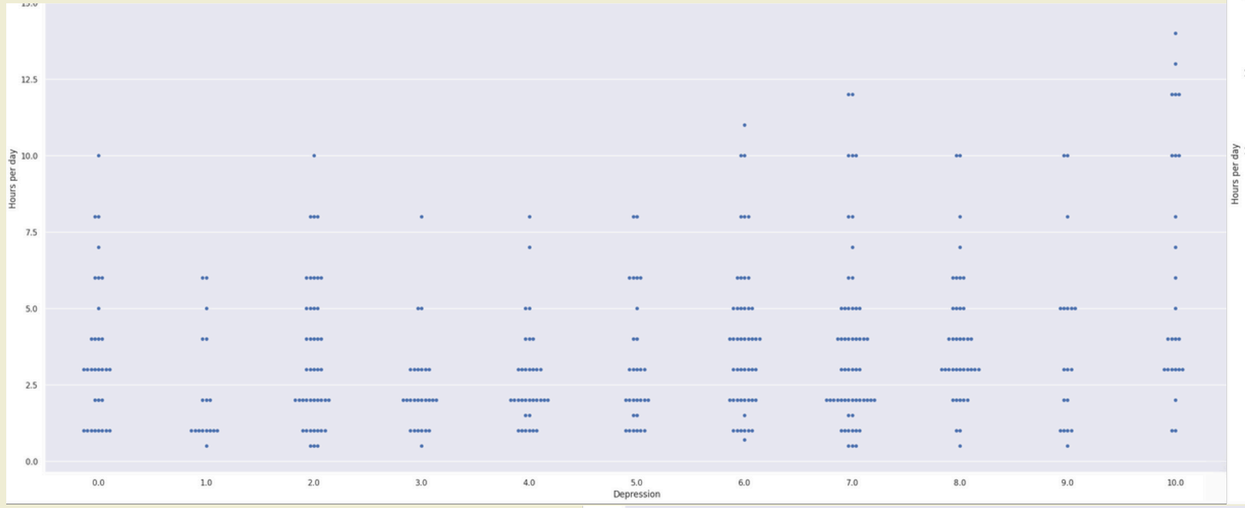
EXPLORATORY DATA ANALYSIS AND OBSERVATION

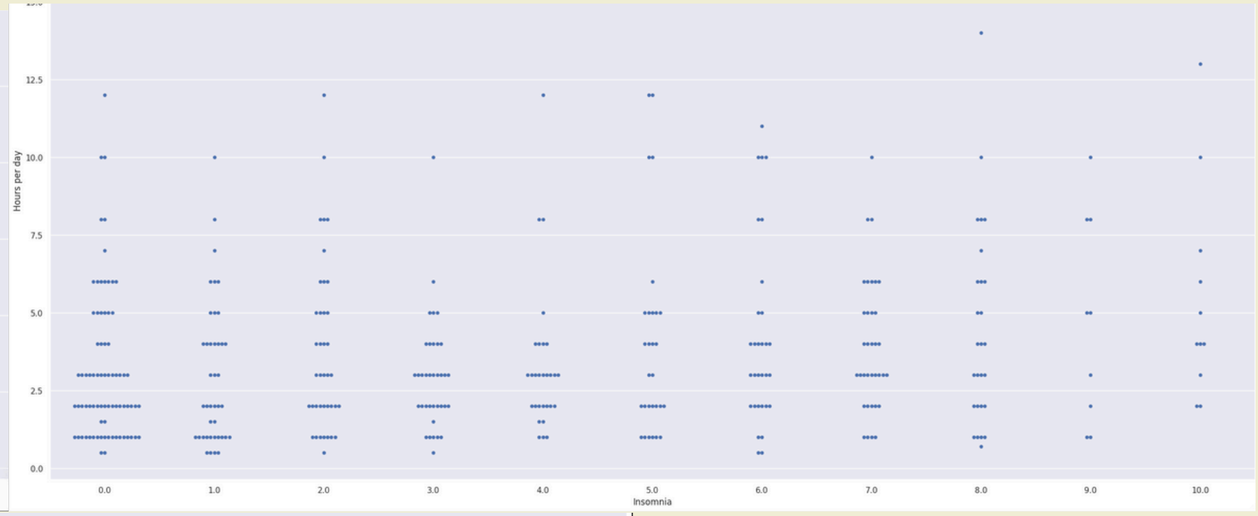
Hours per day & music effects Swarmplots

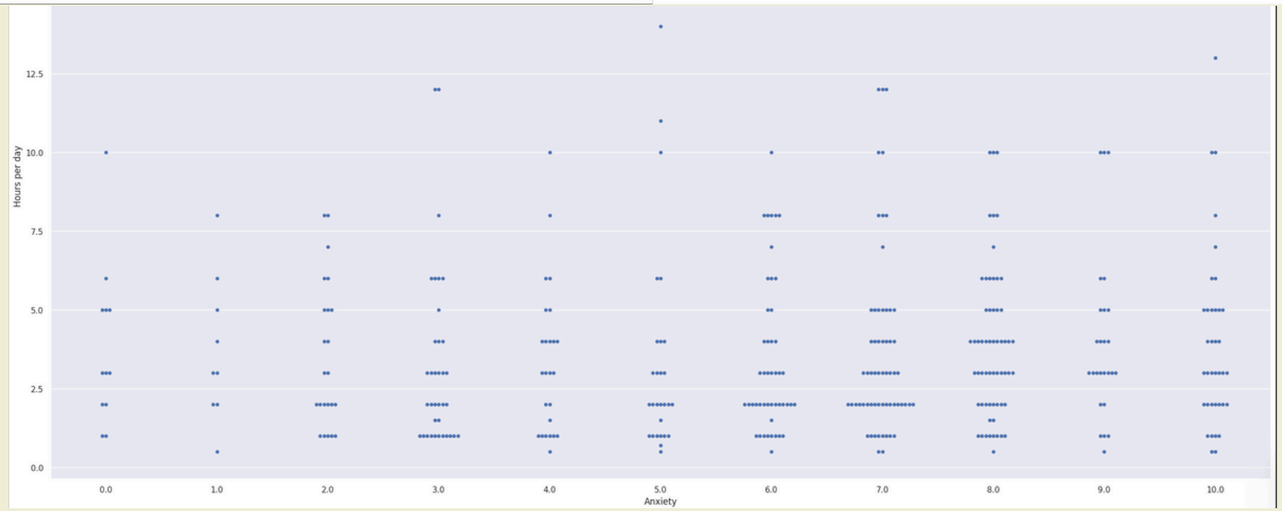


EXPLORATORY DATA ANALYSIS AND OBSERVATION

Hours per day Swarmplots on types of mental health







MACHINE LEARNING MODELS

Classification Tree and Confusion Matrix

Classification Tree with separate models discussing about whether factors pose a positive effect on the target group

Confusion Matrix for Train and Test

Train_Test-Split

Preparing train sets and test for
'BPM' and 'Music Effect'
& 'Hours per Day' and 'Music Effect'

```
Train Set : (257, 1) (257, 1)
Test Set  : (109, 1) (109, 1)
```

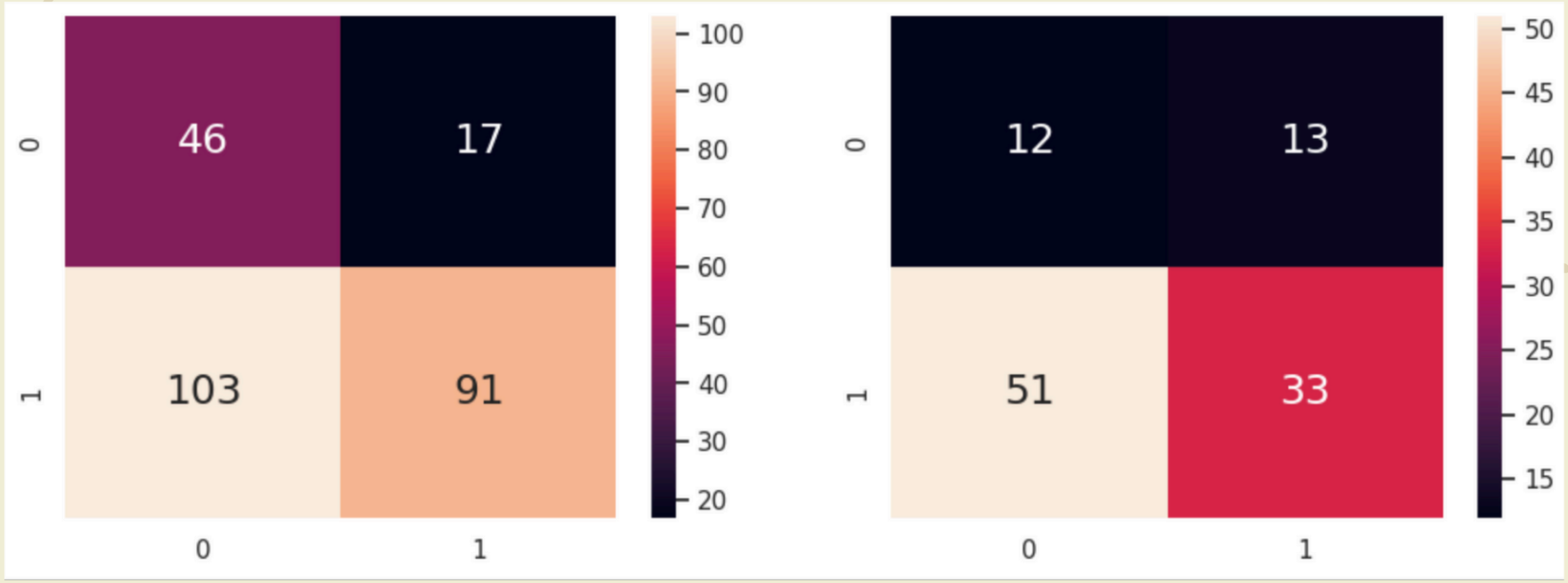
The train : test ratio to be 0.7 : 0.3

- **Adequate training data**
- **Reliable conclusions**
- **Aviod overfitting**

Classfication Tree

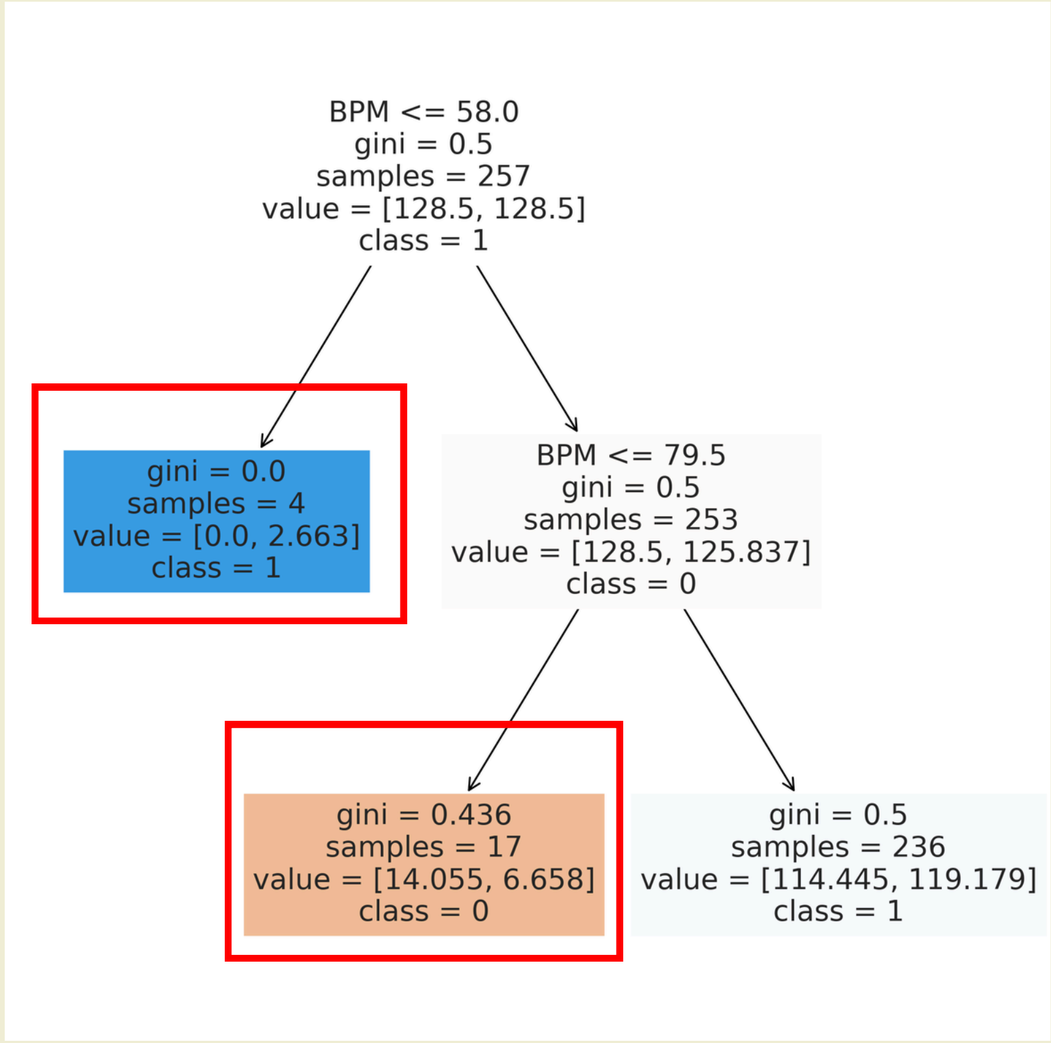
Classfication Accuracy and Confusion Matrix ('BPM' and 'Music Effect')

```
Goodness of Fit of Model        Train Dataset
Classification Accuracy         : 0.7392996108949417

Goodness of Fit of Model        Test Dataset
Classification Accuracy         : 0.7431192660550459
```



Classfication Tree

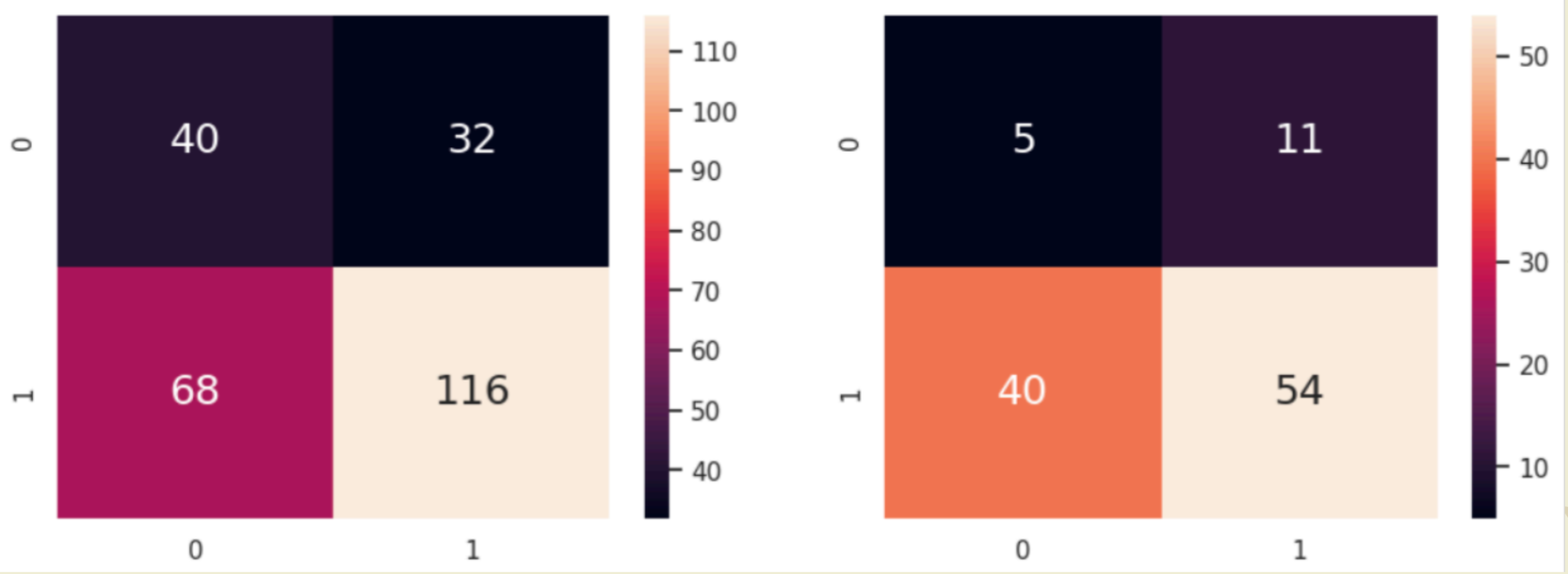
(‘BPM’ and ‘Music Effect’)



Classfication Tree

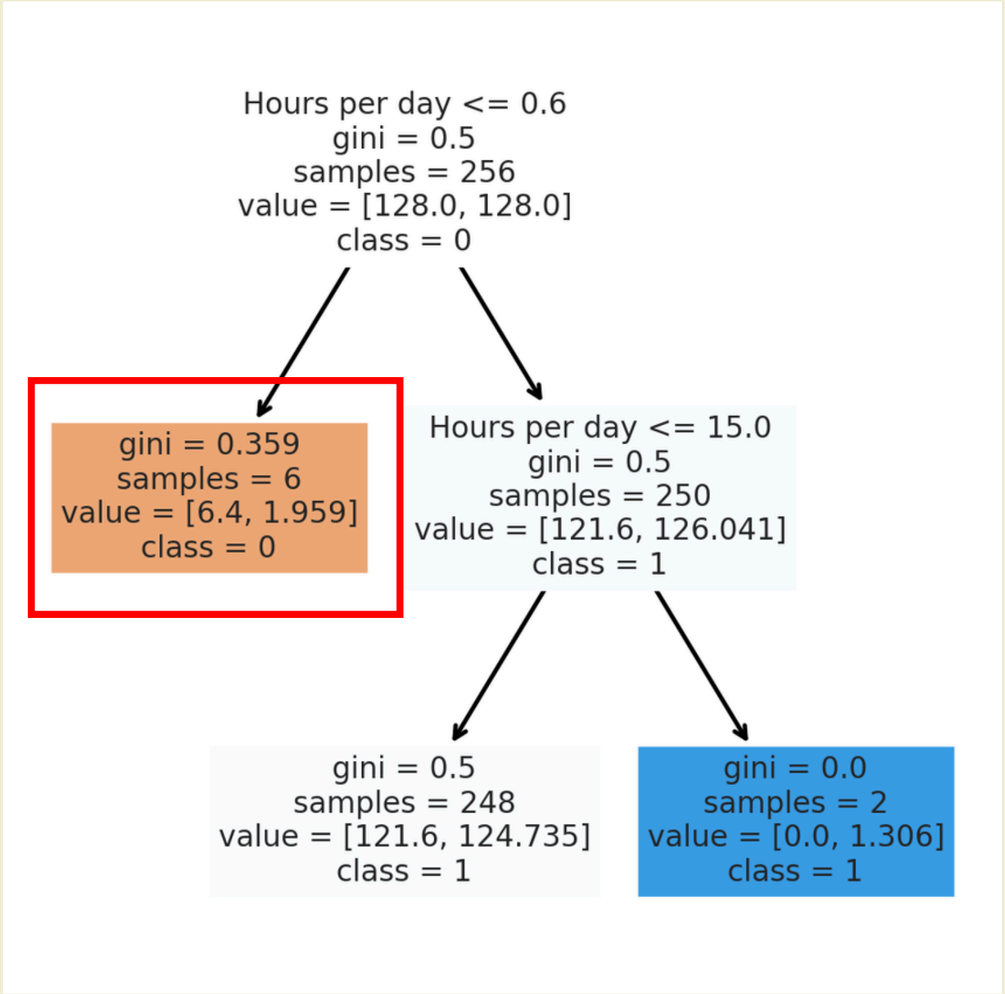
Classfication Accuracy and Confusion Matrix ('Hours per Day' and 'Music Effect')

```
Goodness of Fit of Model        Train Dataset
Classification Accuracy         : 0.765625

Goodness of Fit of Model        Test Dataset
Classification Accuracy         : 0.7272727272727273
```



Classfication Tree

('Hours per Day' and 'Music Effect')

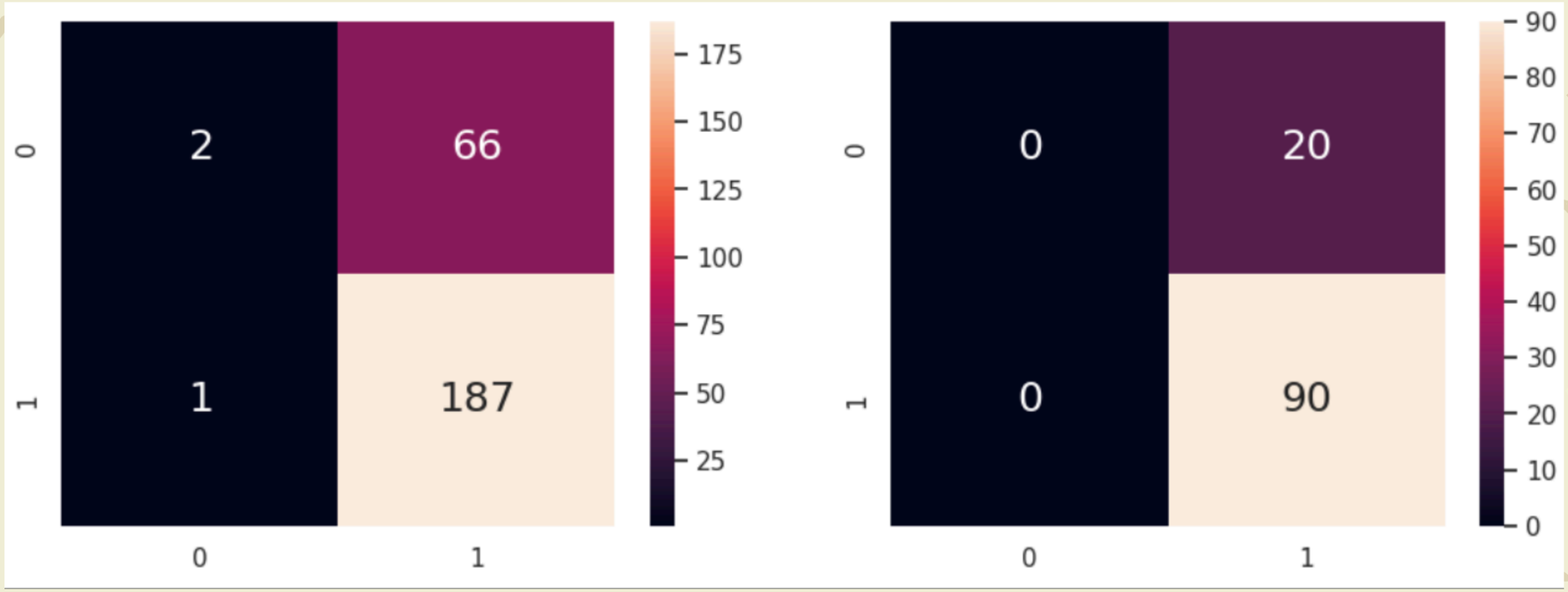


Multi variate Classfication Tree

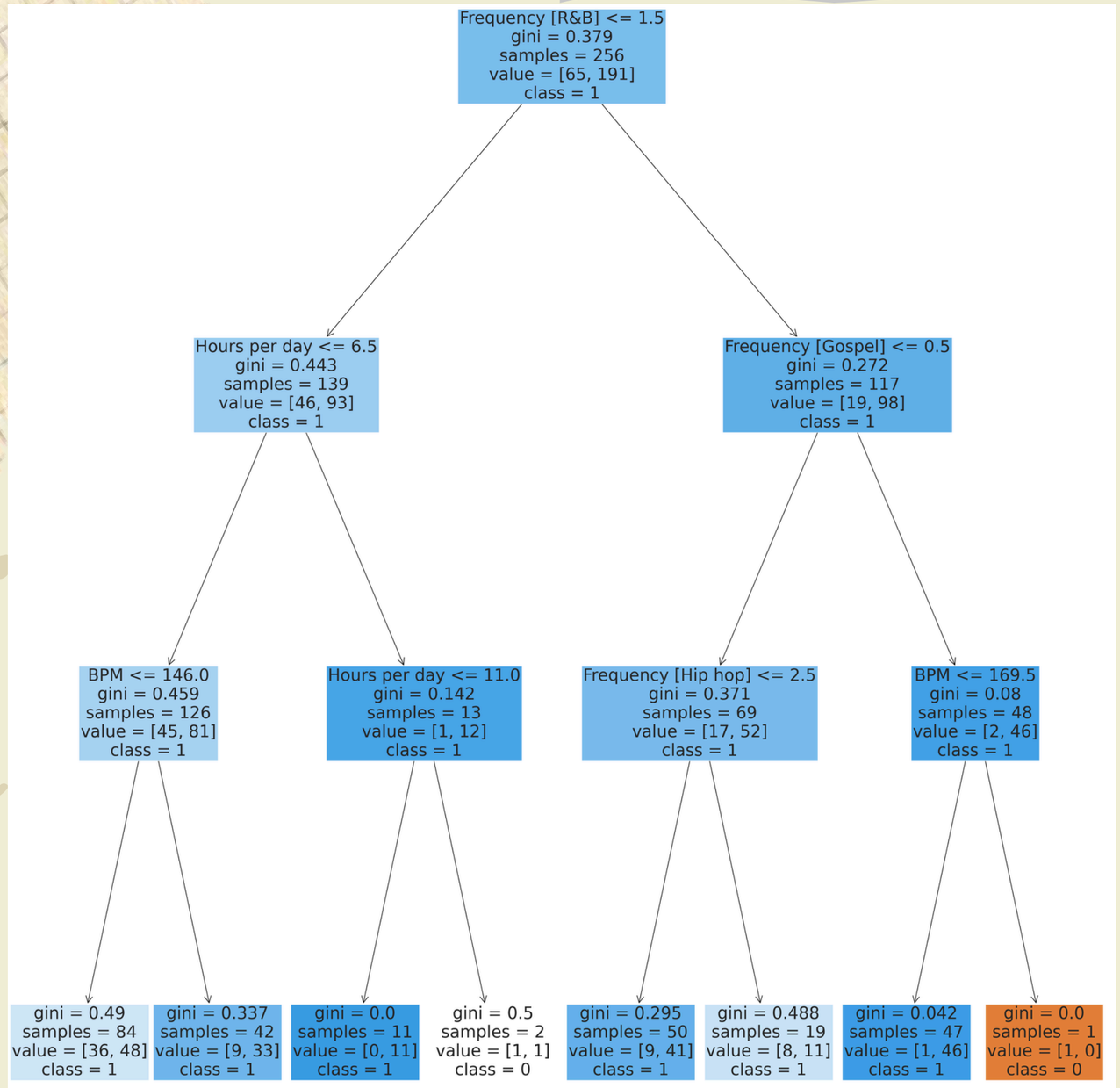
We decided to use multi variate decision tree based on most important factors: BPM, Hours per day, Frequency [Gospel], Frequency [Hip hop], Frequency [Pop], Frequency [R&B]

Multi variate Classfication Tree

```
Goodness of Fit of Model        Train Dataset
Classification Accuracy         : 0.73828125

Goodness of Fit of Model        Test Dataset
Classification Accuracy         : 0.8181818181818182
```



Multi variate decision tree



Similarly, due to the lack of insightful data analysis from the multi-variate decision tree. Therefore we'll try for different approach with Random Forest

Weight Balance

```
17  dectree = DecisionTreeClassifier(max_depth=2, class_weight='balanced')
18  dectree.fit(X_train, y_train)
```

What is the use?

Class weight balance is used to handle imbalanced datasets, which can an lead to a biased model that favors the majority class.
Class_weight = 'balanced'
-> assign higher weights to the samples of the minority classes
-> so we can get better accuracy in decision tree

Application of random forest algorithm

```
#Define the Random Forest model
#Bigger number of trees = more accurate
rf_model = RandomForestClassifier(
    n_estimators=200,              # Number of trees (Decision tree)
    max_depth=5,                   # Control tree depth (Decision tree)
    class_weight='balanced',       # Handles class imbalance
    random_state=42                # For reproducibility
)
```

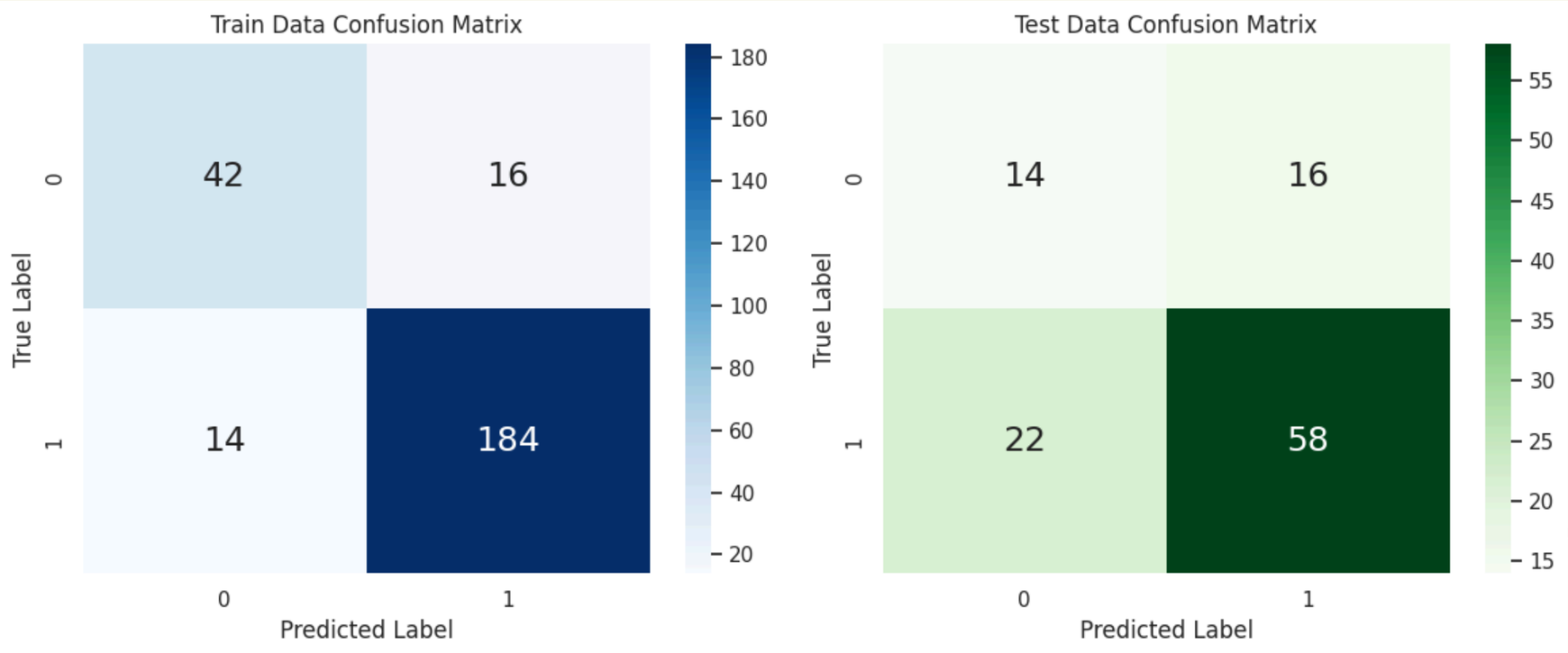
Objective: To predict the music effect using random forest classifier that handles class imbalance and prevents overfitting

n_estimators =200: initialize to 200 trees
max_depth=5: Restricted tree depth to 5 levels to prevent overfitting.
class_weight: Adjusted to ‘balanced’ for handling class imbalance
random_state= 42: controls the randomness in machine learning processes, ensuring consistency.

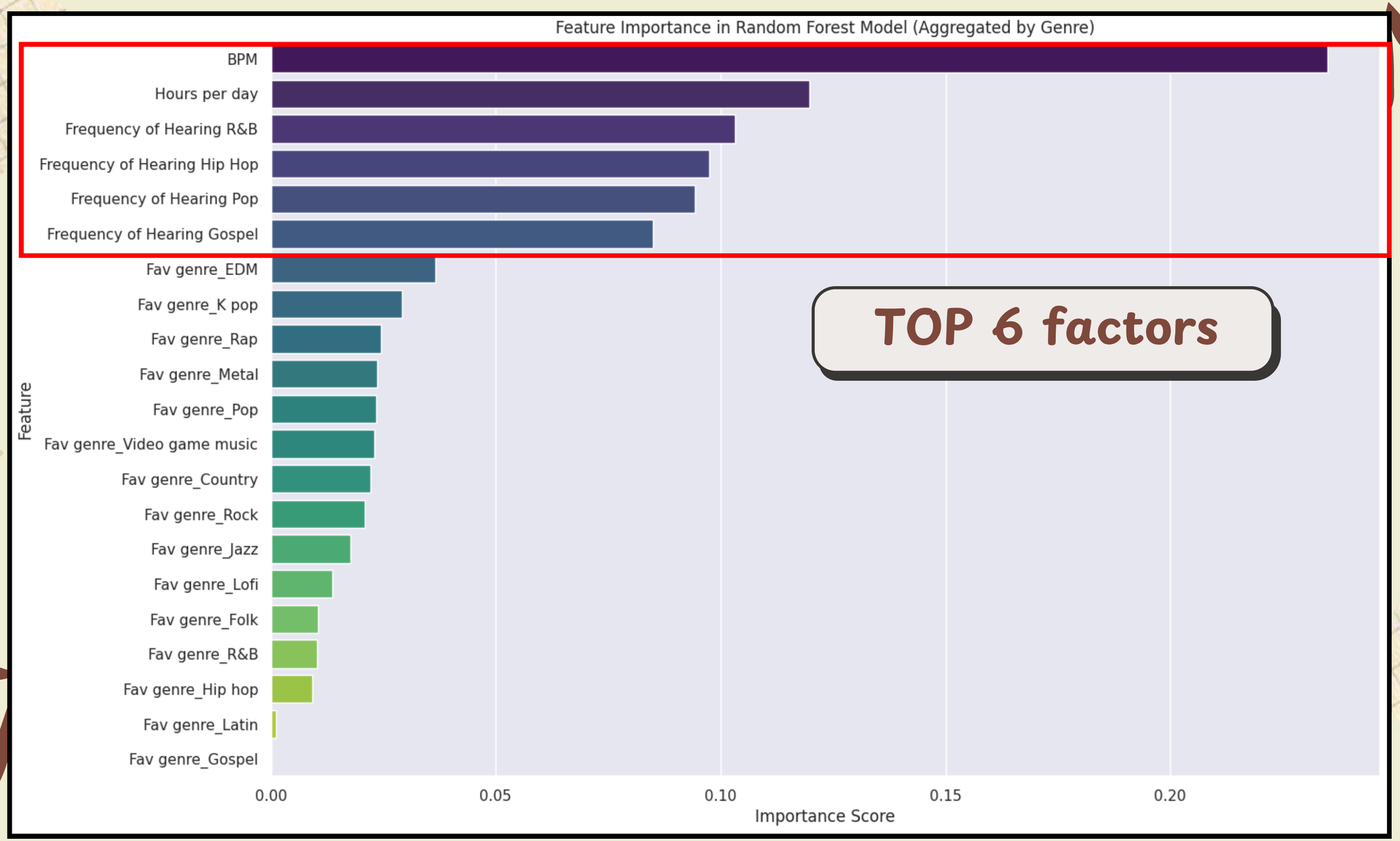
Random Forest

Accuracy and Confusion Matrix

```
Random Forest Model Accuracy
Train Accuracy: 0.8828125
Test Accuracy: 0.6545454545454545
```



Random forest plots



Outcome

In summary, we used supervised learning algorithms like decision tree and random forest, as well as unsupervised learning algorithms like clustering to evaluate the data.

We have built a relatively effective model of >70% accuracy

From our Machine Learning Models, we were able to identify variables are the most influential in amplifying the effect of music on our mental wellbeing

Conclusion

- Numerical variables like BPM and Hours per day are the most influential features in predicting music effect.
- Supported by both the Decision Tree and Random Forest, frequency of listening to R&B music is likely to be more influential than the rest of the genres
- Order of influence by genre in decreasing order: R&B, Hip Hop, Pop, Gospel

Conclusion

Our findings are actually supported by the Australian & New Zealand Mental Health Association

Mood

It can also inspire you to express your feelings. Music is a great reflection of your mood and your emotions. The genre of music you listen to can tell a lot about your mental health and also contribute to your mental health.

Sometimes just listening to your favourite jam can help you forget about all the stress in life and elevate your mood. The rhythmic throb and the thud in music help boost your mental health by giving you an adrenaline rush.

Listen to RnB or pop to distract yourself when you are feeling low or when you feel drained. You can also put on heavy metal or rock music when you're enraged.

https://anzmh.asn.au/blog/mental-health/music-for-your-mental-health

Conclusion

Therefore based on our analysis, we recommend incorporating music therapy that involve listening to music with lower BPM values (Below 80) and R&B music we enjoy for at least 40 minutes a day to potentially improve our mental well-being.

Thank You!