Melody and Mind: Linking Musical Characteristics to Listener Well-Being

DSCI 510

Nuojinli Xu

December, 4, 2025

Introduction

This project investigates how musical characteristics influence listeners' emotional well-being by integrating data from multiple online platforms. By combining YouTube engagement metrics, Spotify audio features, and mental-health survey responses, the study aims to explore whether measurable components of music—such as tempo, energy, valence, and acoustic properties—correlate with users' emotional reactions and perceived benefits from music. Through data cleaning, feature integration, exploratory visualization, and preliminary modeling, this project seeks to provide insights into how different musical attributes may affect mood, emotion, and overall mental wellness.

Data Sources

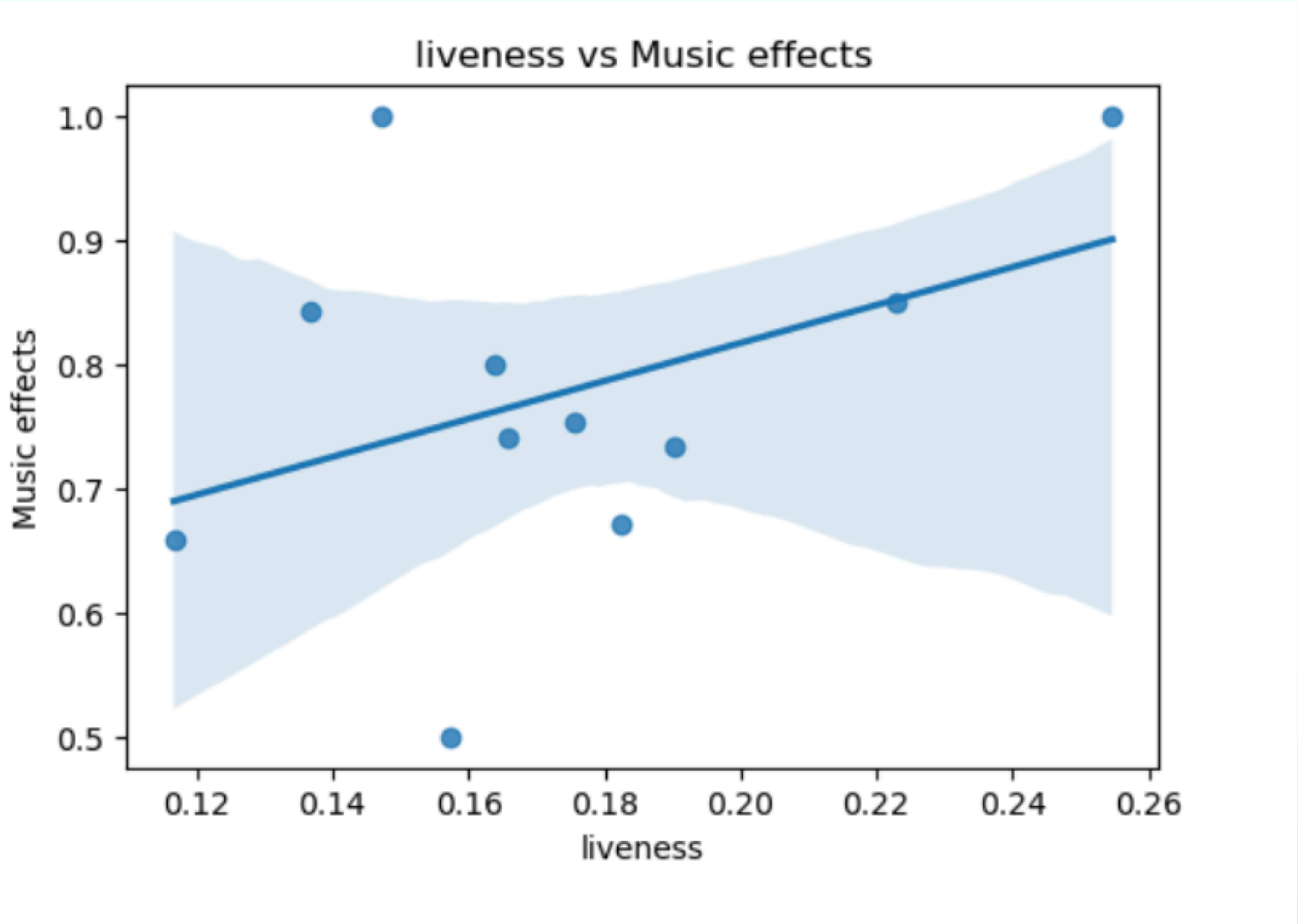
1 . Spotify Dataset
This dataset provides objective acoustic and musical attributes for each track, including tempo, energy, valence, danceability, instrumentalness, and more. These features are used to quantify musical characteristics and analyze how they relate to emotional responses.

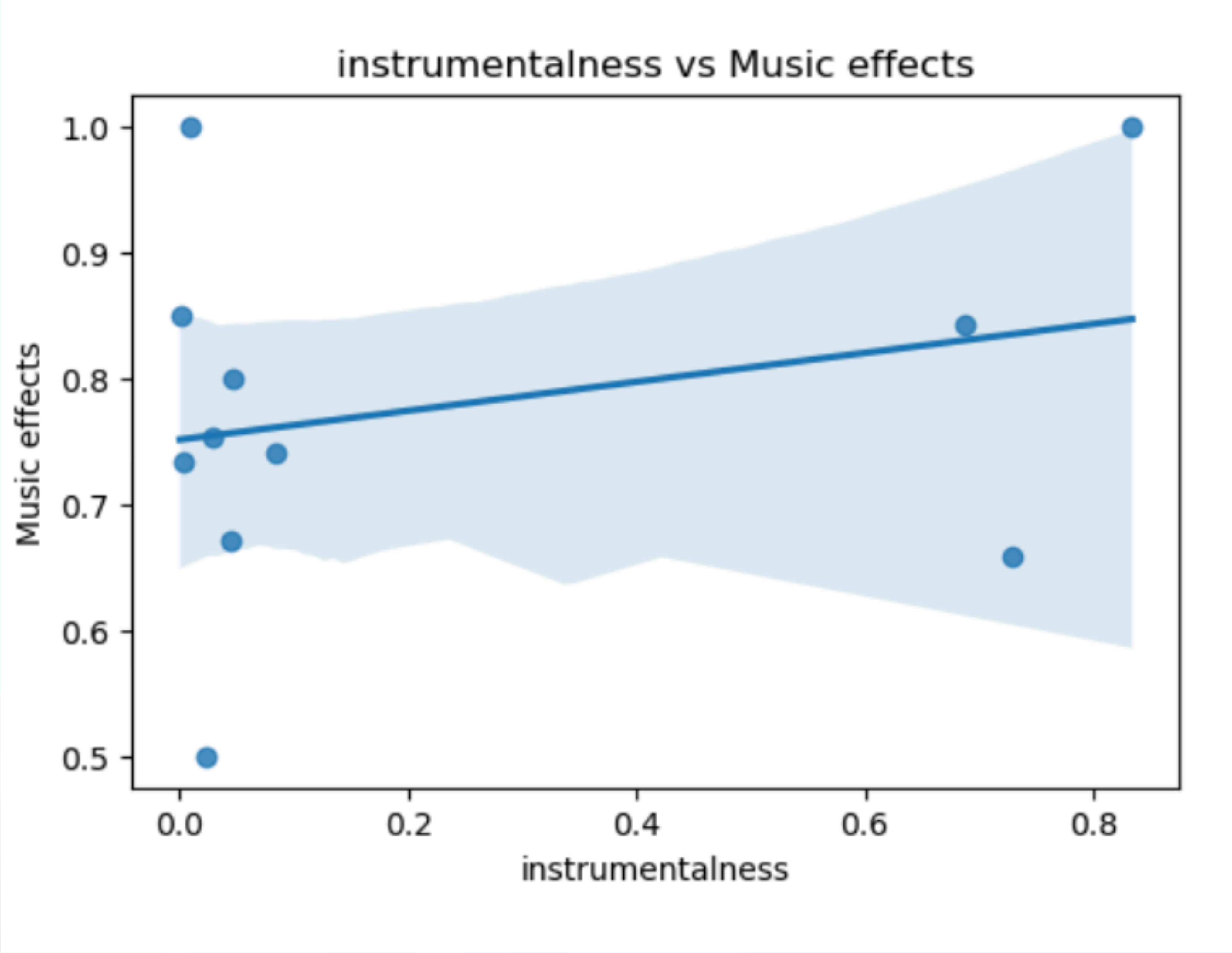
2 . Music & Mental-Health Dataset
This dataset includes emotional states, wellness indicators, and responses related to music-listening habits. It enables linking musical attributes and online interaction patterns to actual well-being outcomes, forming the basis for understanding music's emotional impact from the listener's perspective.

3 . YouTube Data (Views, Likes, Comments)
Collected using the YouTube Data API, this dataset contains engagement information for a curated set of music videos. It includes video views, likes, the top-level comments, and newly added comments from users. These data help capture audience reactions, popularity patterns, and qualitative sentiment from user-generated feedback.

Music VS Mental

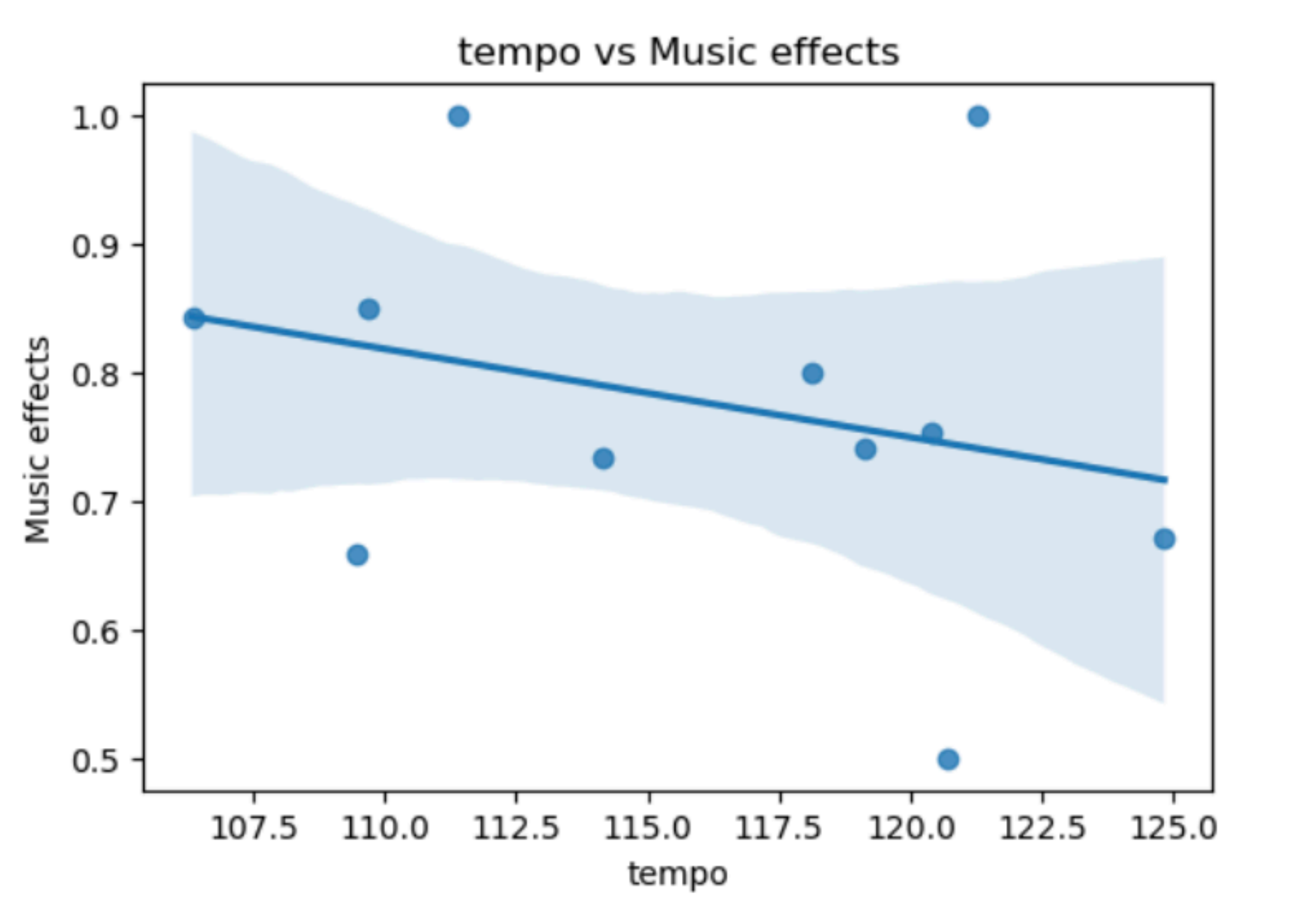
Worsen: -1 No Effect: 0 Improve: 1

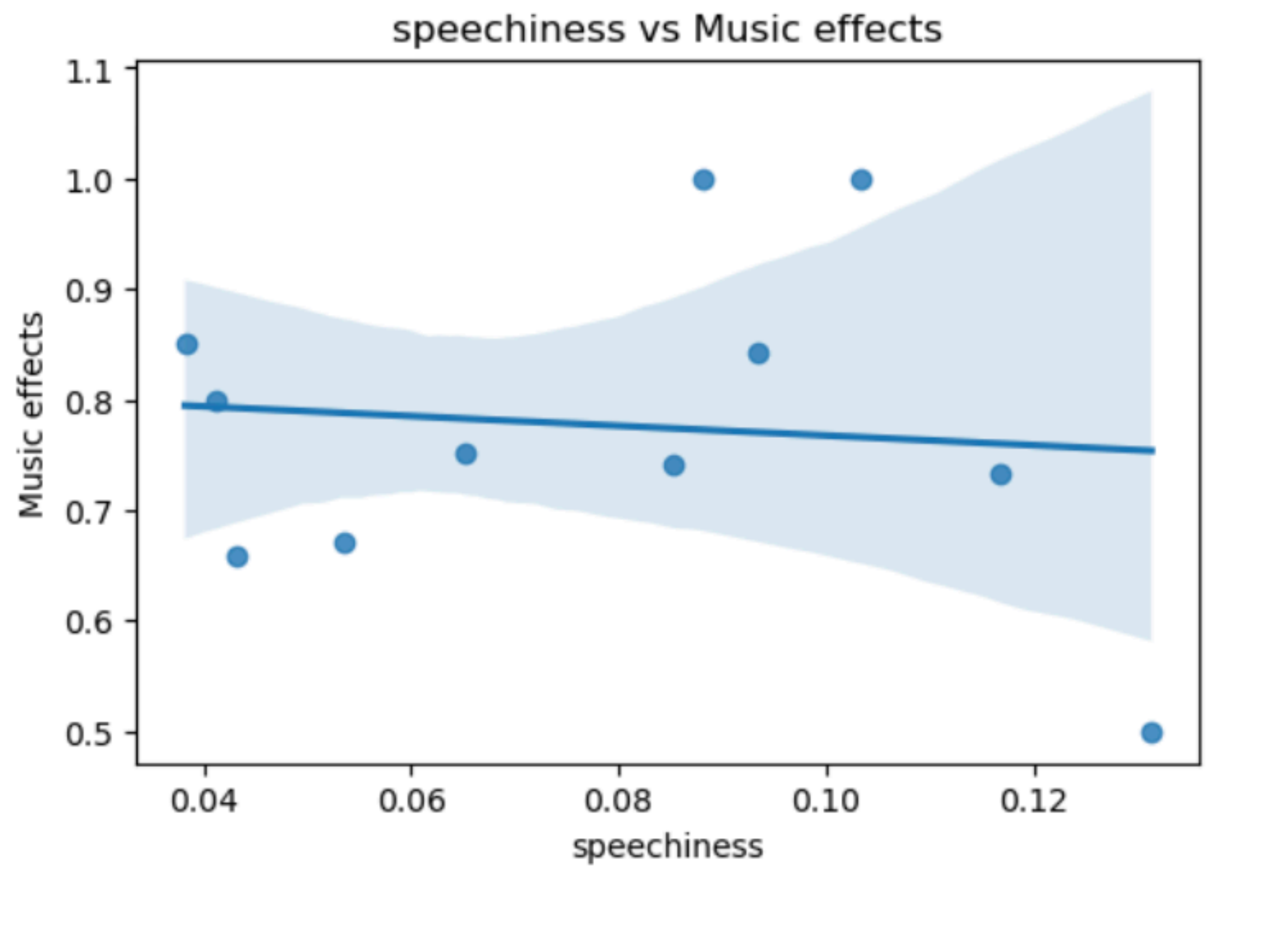




Music VS Mental

Worsen: -1 No Effect: 0 Improve: 1





Comment emotional analysis

Good Words & Bad Words & Intensifier & Negation

+1 & -1 & * 2 & * -1

"good", "great", "amazing", "awesome", "beautiful", "wonderful", "fantastic", "excellent", "love", "loved", "lovely", “perfect”...

"bad", "sad", "terrible", "awful", "hate", "hated", "angry", "mad", "worst", “disappointin g", "boring", “annoying”...

“very”, ”really”, “extremely”, “super", “so”, “absolutely", “totally", "incredibly ...

"not", "no", "never", "don't", "doesn't", "didn't", "isn't", "aren't", "won't", "can't", “cannot”...

Correlation Heatmap: Comment Score vs Music Features

Challenges

1. Noisy and Unstructured YouTube Comments
YouTube comments vary greatly in style, length, and language usage. Many include slang, emojis, or mixed emotions, making them difficult to interpret and requiring careful preprocessing before analysis.

2. Designing a Rule-Based Sentiment Scoring System
Creating a custom sentiment model using positive/negative word lists introduced challenges such as handling intensifiers, negations, and repeated modifiers. Ensuring consistent scoring across all comments required iterative refinement.

References

Ameh, S. (2023). Spotify Music Dataset [Data set]. Kaggle. https://www.kaggle.com/datasets/solomonameh/spotify-music-dataset

Rasgaitis, C. (2020). Music & Mental Health (MxMH) Survey Results [Data set]. Kaggle.https://www.kaggle.com/datasets/catherinerasgaitis/mxmh-survey-results/data

Google Developers. (2024). YouTube Data API v3 Documentation. https://developers.google.com/youtube/v3

Thank You!