Music Taste & Mental Health

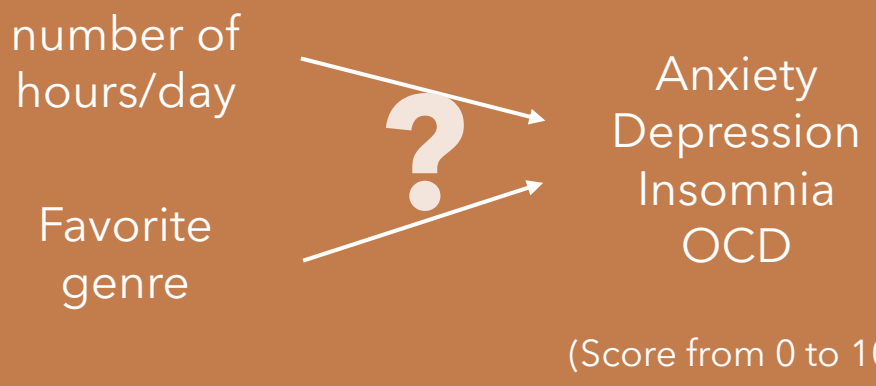
An analysis exploring the **Music & Mental Health (MxMH) dataset,**

from *Catherine Rasgaitis* **@Kaggle**

Ana Catarina Morais

Project Overview

Objective: Explore relationships between music habits & preferences and mental health.



Implementation Details

1st Cleaning!

What to keep and what to drop

Checking null values

- BPM variable (107/ 736 nulls)
- Other variables (1 to 8 nulls)

Making decisions

- Estimating average bpm of every favorite genre: *groupby("fav_genre")["bpm"].mean()*
- Replacing BPM null values with average bpm for corresponding favorite genre: *apply(lambda row:)*
- Droping rows (12) for other variables: *dropna()*

Removing crazy outliers for BPM

(999999999 is not a valid BPM for a song)

Checking for duplicates *(zero)*

Fixing data types

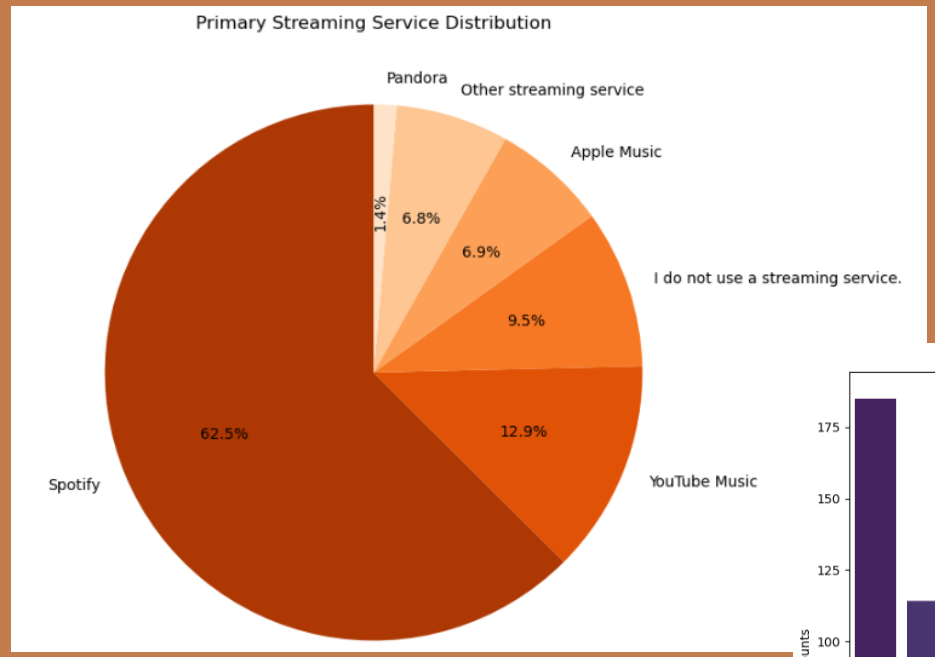
- Age
- Anxiety, Depression, Insomnia, OCD

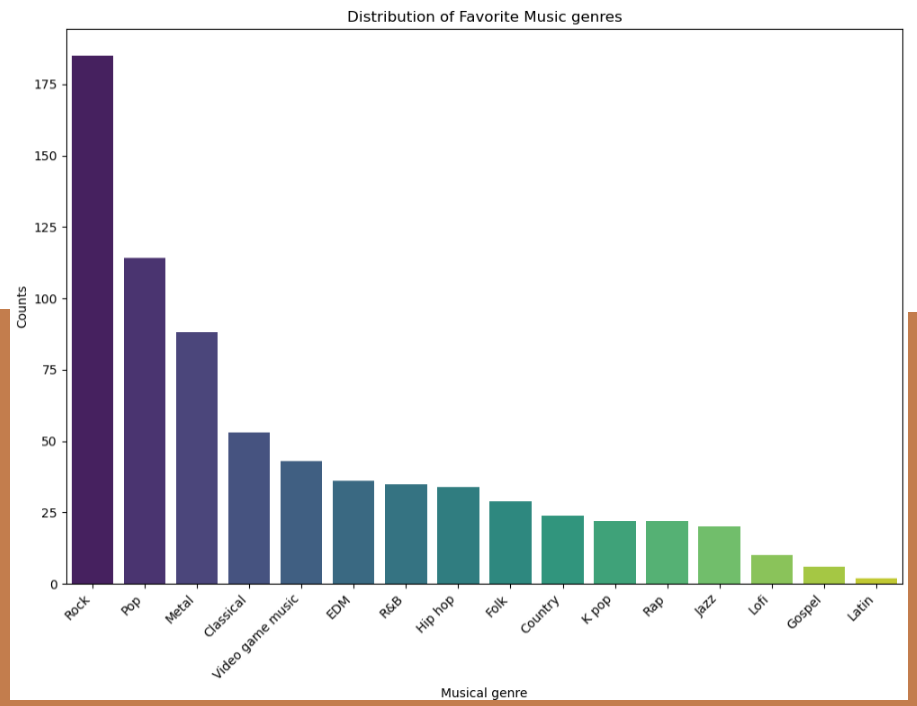
astype(int)

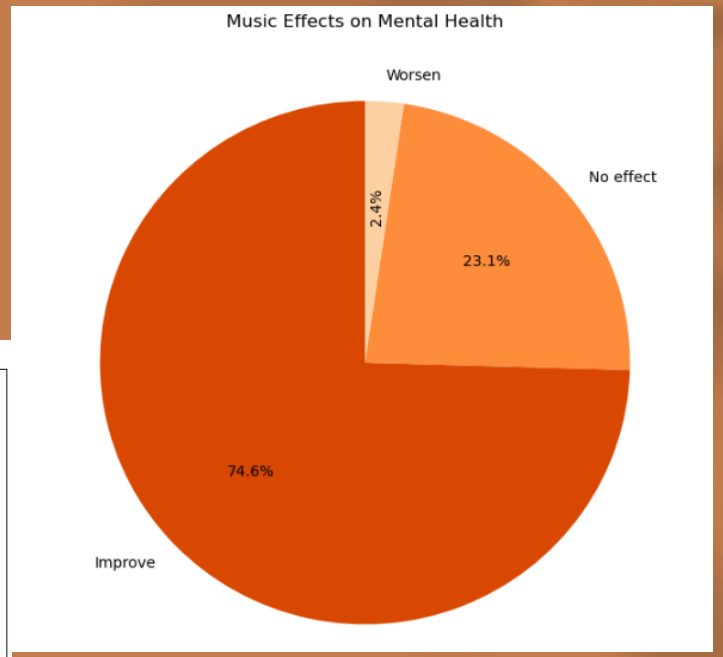
	age	primary_streaming_service	hours_per_day	while_working	fav_genre	bpm	anxiety	depression	insomnia	ocd	music_effects
2	18	Spotify	4.0	No	Video game music	132.0	7	7	10	2	No effect
3	61	YouTube Music	2.5	Yes	Jazz	84.0	9	7	3	3	Improve
4	18	Spotify	4.0	Yes	R&B	107.0	7	2	5	9	Improve

(723, 11)

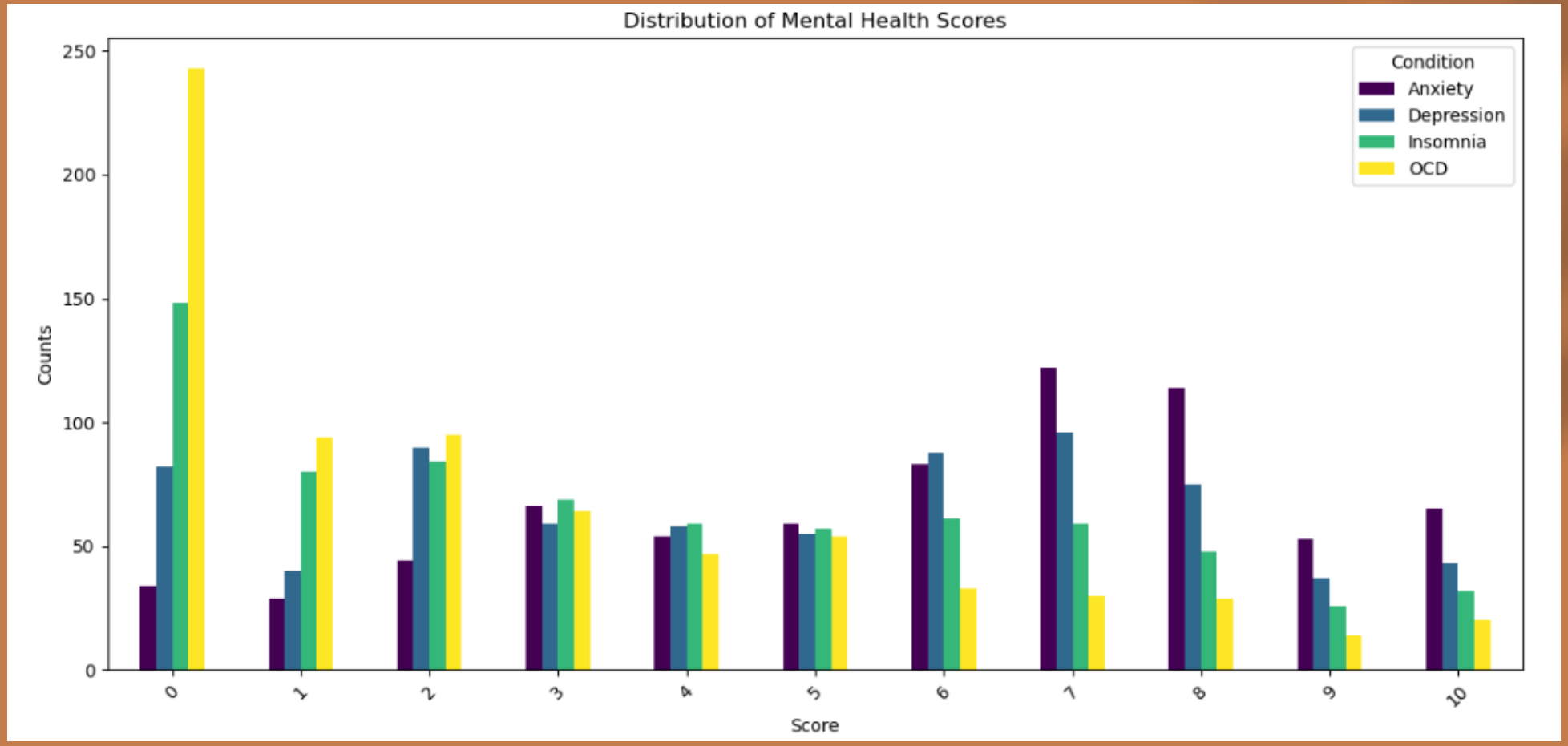
Data Analysis – Categorical Data



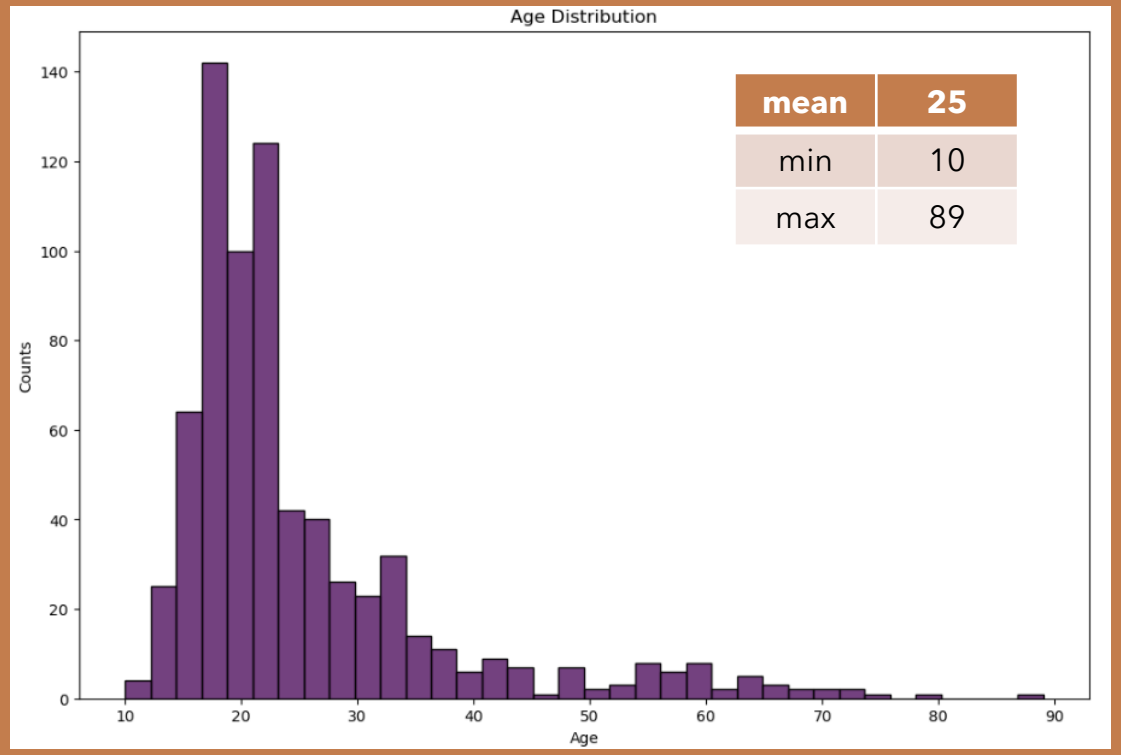




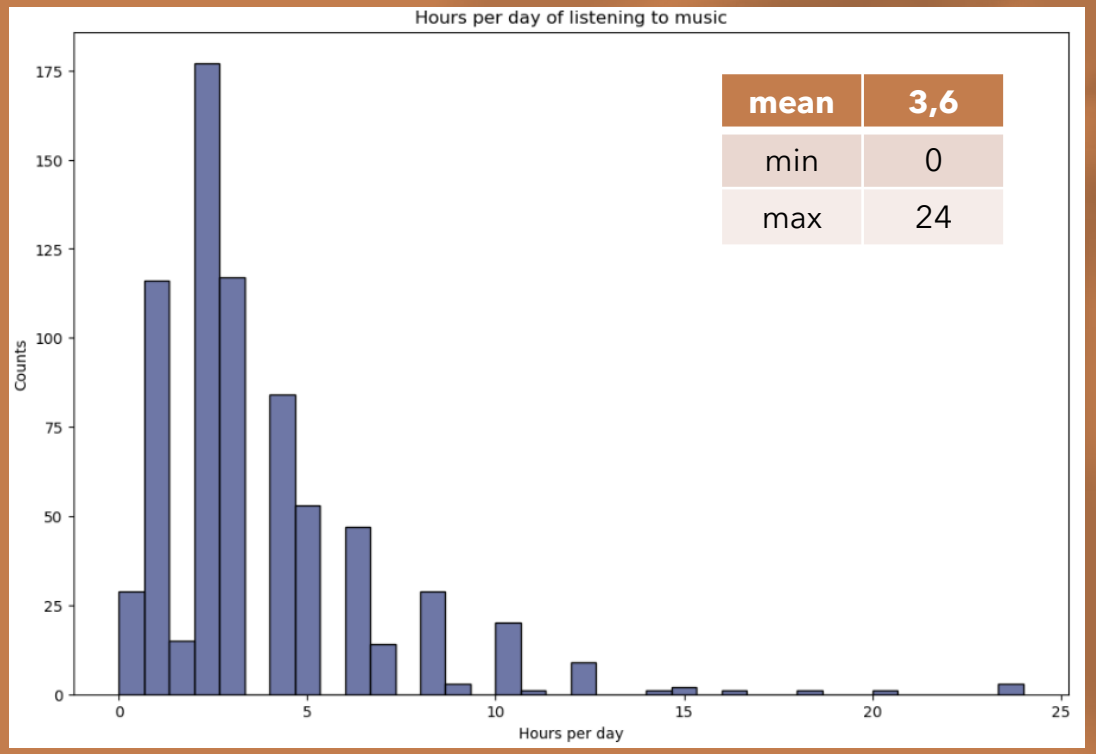
Data Analysis – Categorical Data



Data Analysis – Numerical Data

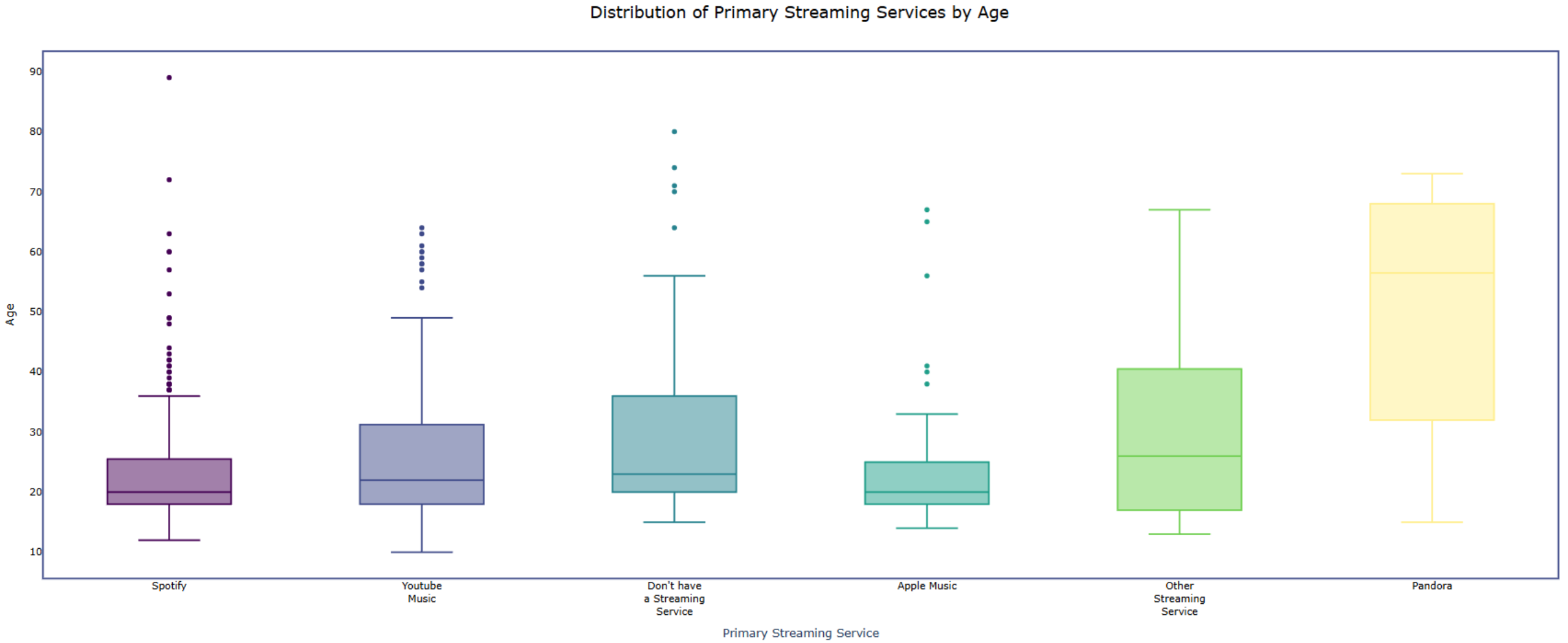


mean	25
min	10
max	89



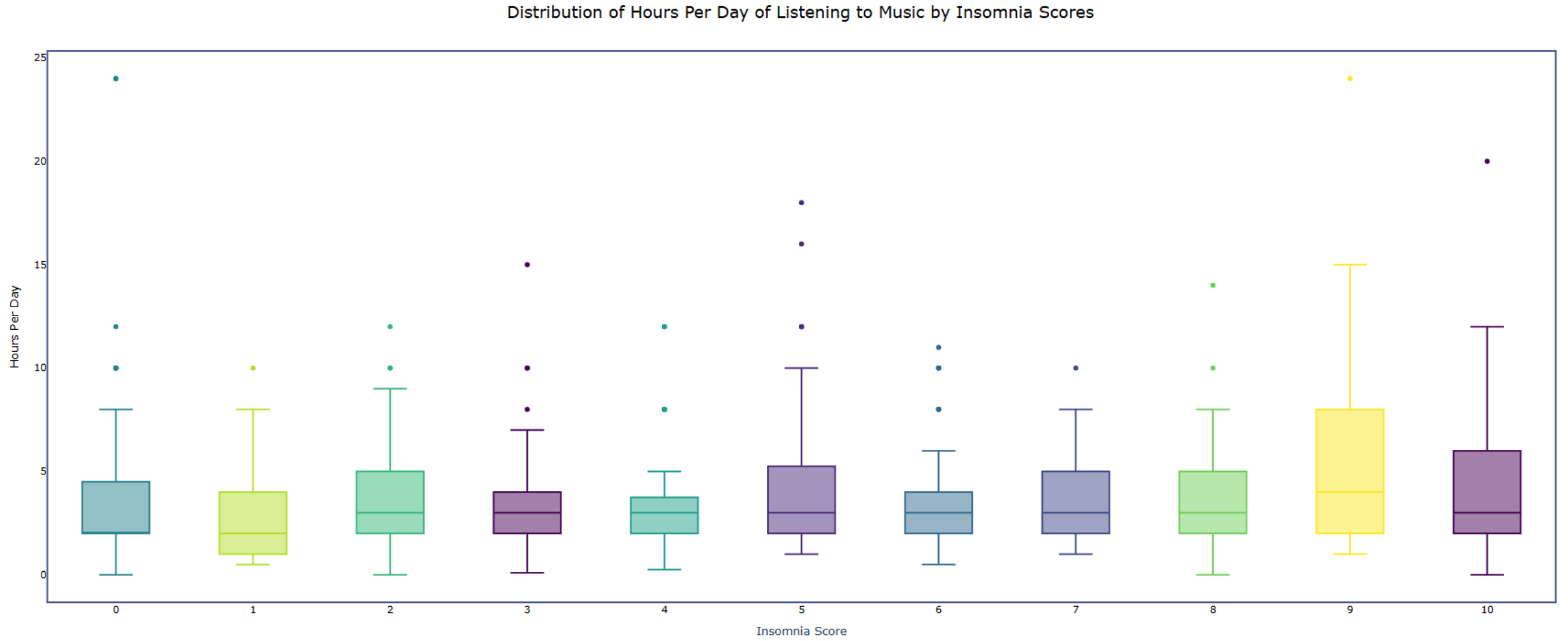
mean	3,6
min	0
max	24

Streaming Service Preferences



Listening time and mental health

Kruskal-Wallis test
P-value < 0.05



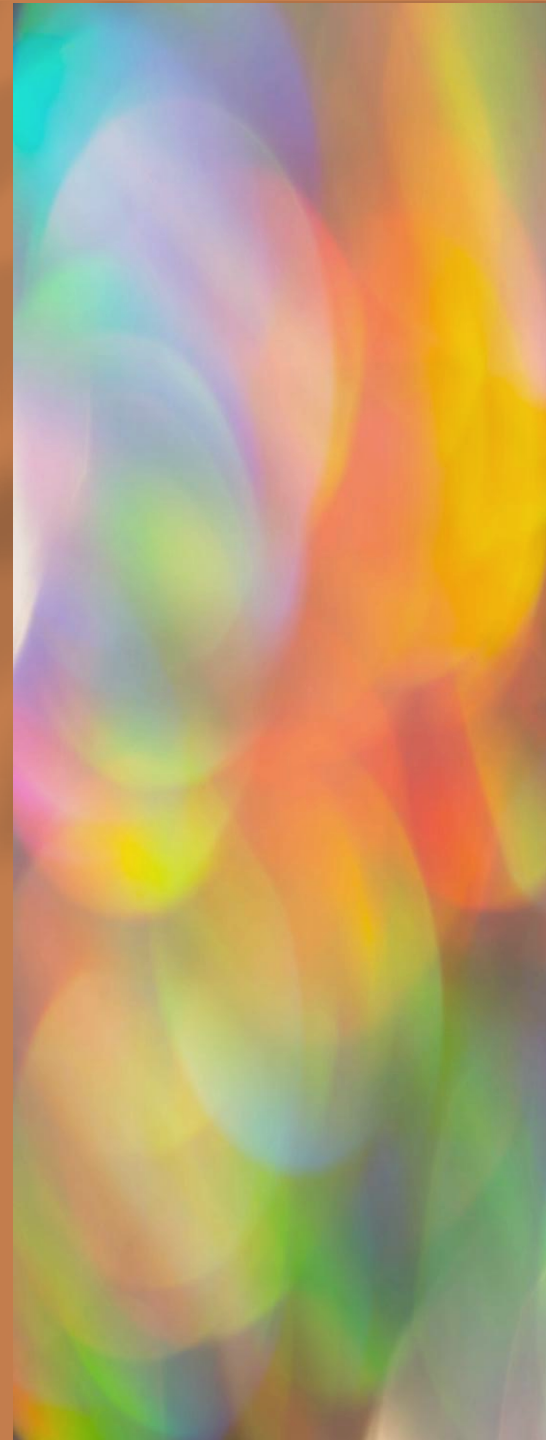
Listening time and mental health

Hours per day and reported mental health scores:

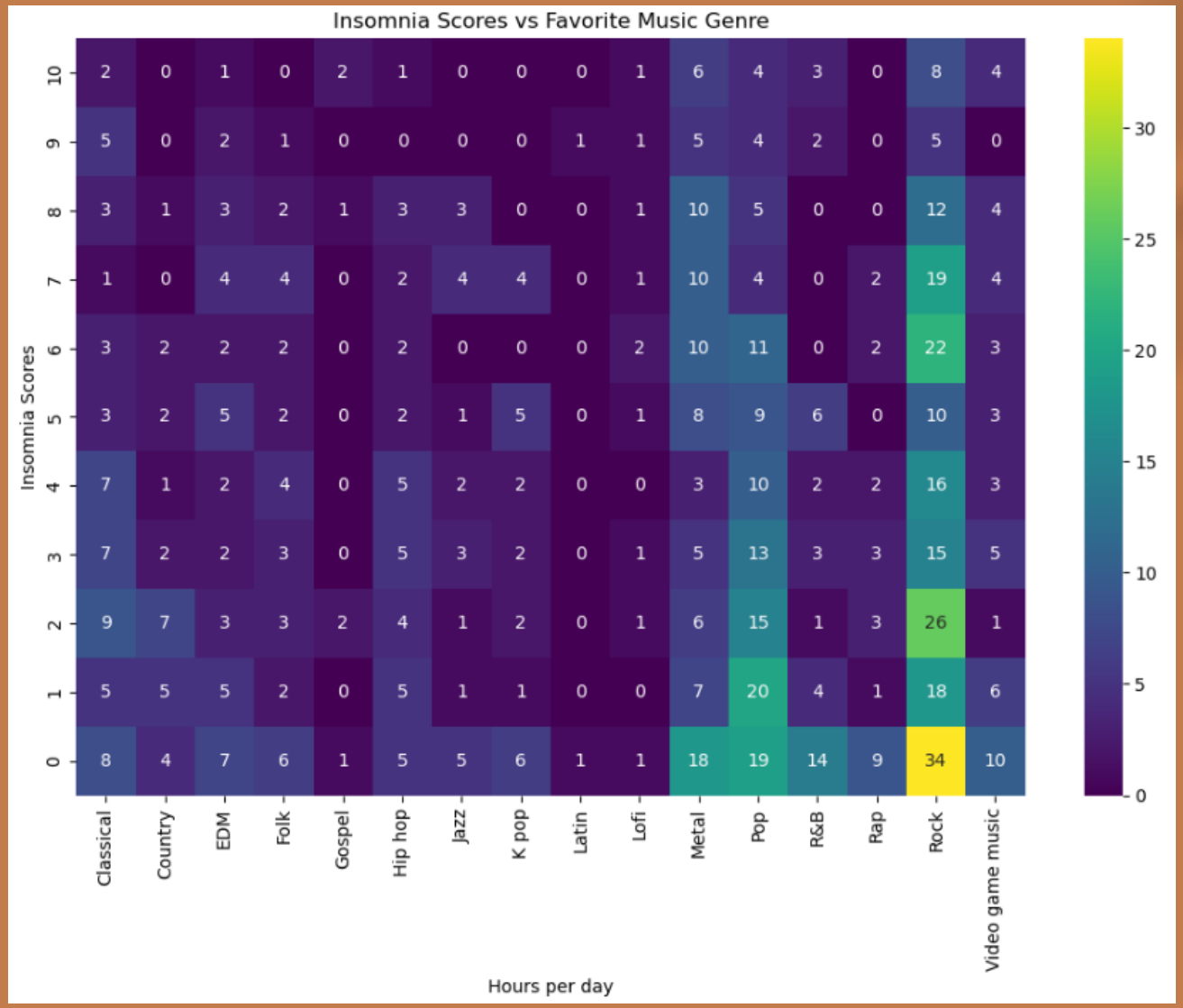
Higher scores for depression, insomnia and OCD correlated with longer hours spent listening to music per day,

While

Lower scores correlated with fewer (or zero) hours of listening to music.

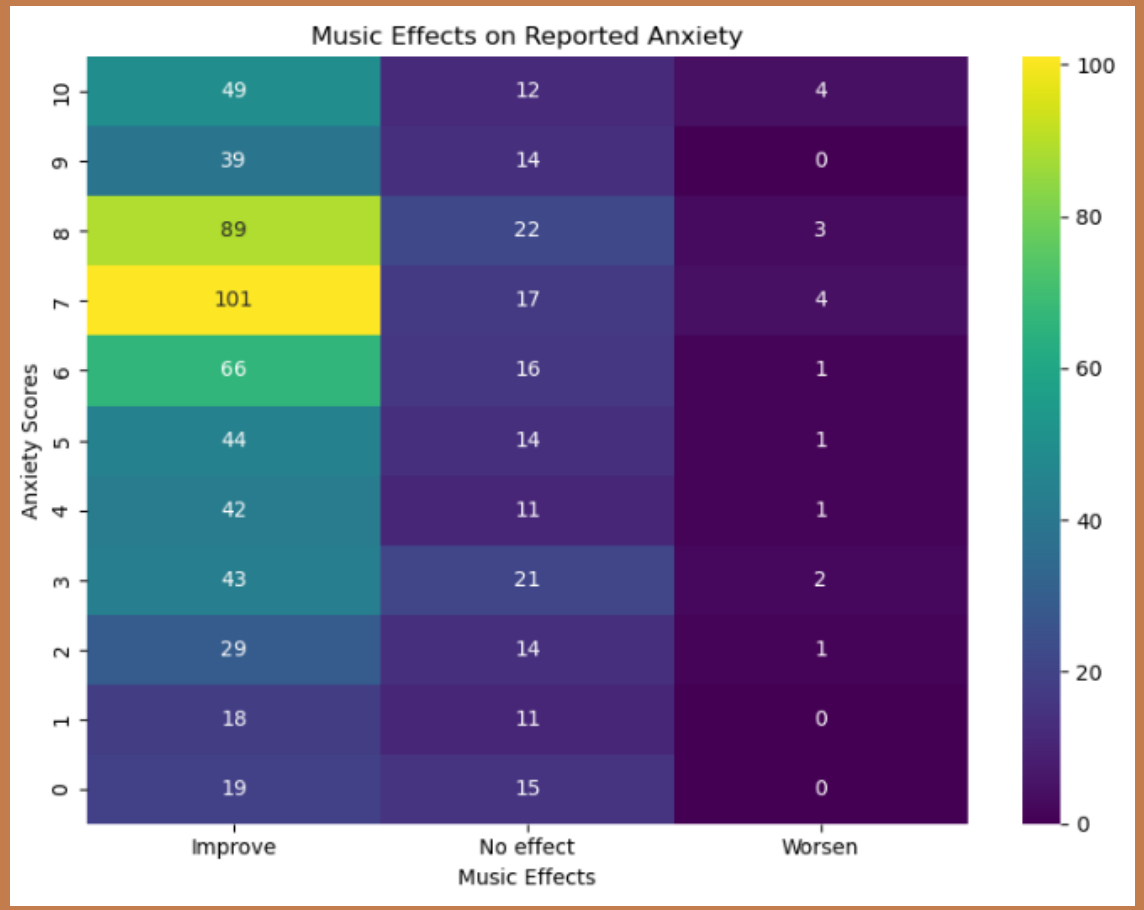


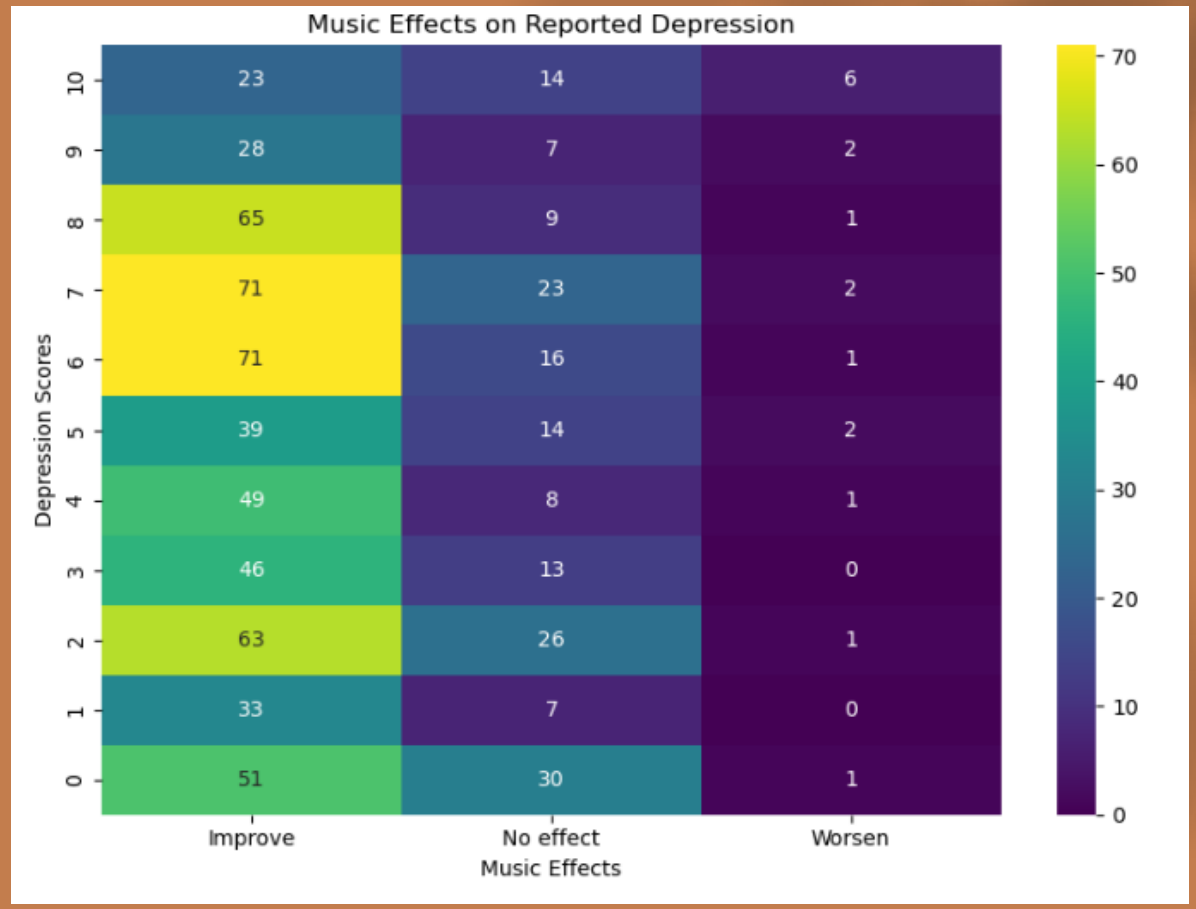
Favorite genre and mental health



Chi-squared test
P-value < 0.05

Music effects on mental health





Chi-squared test
P-value < 0.05

Main Insights

Listening Hours and Mental Health:

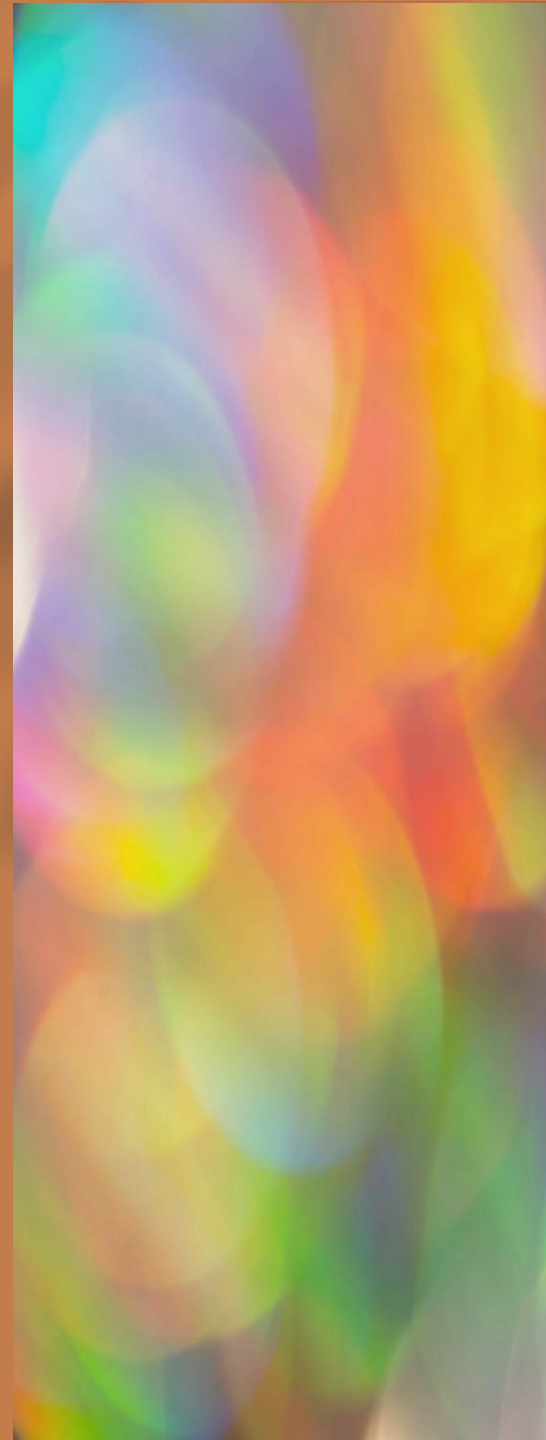
Significant differences in listening hours were observed across **depression**, **insomnia**, and **OCD** levels, but **not** across **anxiety** levels.

Higher scores for depression, insomnia, and OCD were associated with longer listening hours per day, possibly suggesting music as a coping mechanism.

Favorite Genre and Insomnia:

A significant association between **insomnia levels** and **favorite genre** was identified.

The lower the insomnia score, the higher the preference for **rock music**.



Main Insights and Next Steps

Biases in Population Data:

Favorite Genre Bias: Certain genres like **rock** and **pop** were overrepresented in the data, potentially skewing conclusions.

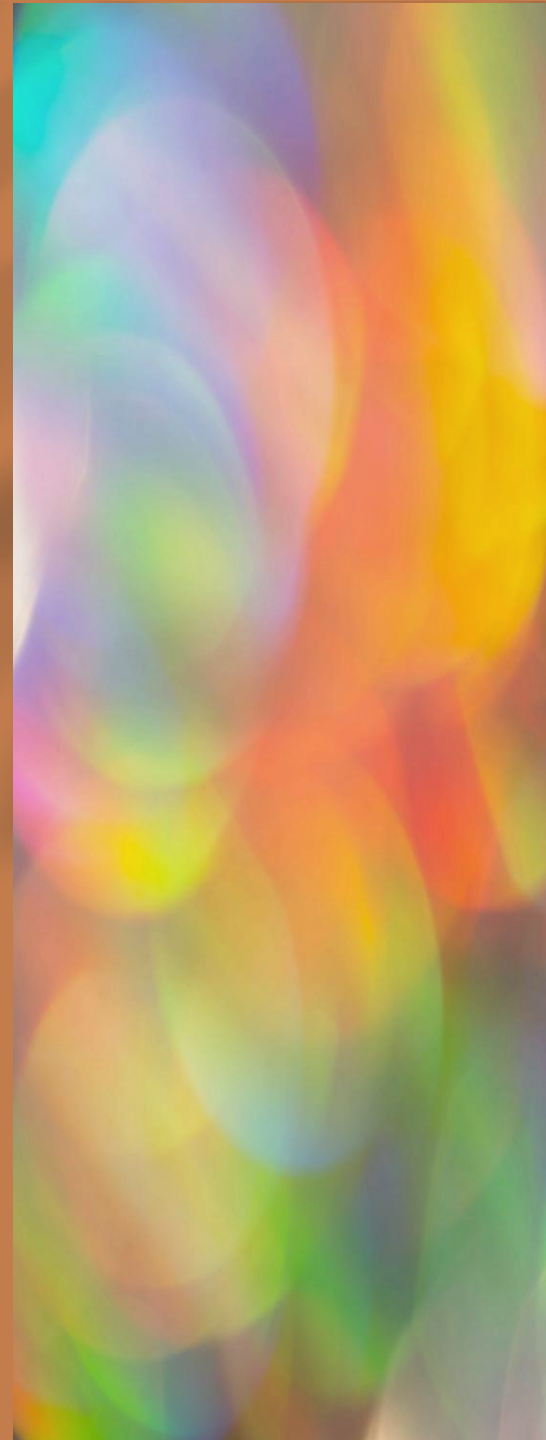
Age Distribution Bias: Some age groups appear to dominate the data.

What next?

Explore **interaction effects** between multiple variables (e.g., favorite gender, insomnia scores and listening hours) and statistical comparisons across **subgroups** (age, "while working").

Explore the effects of **other categories** that were not covered in this analysis.

Collect **additional data** (more behavioral and contextual variables and expand genre options) to reduce biases.



Thank You!

I will take your questions, now ;)