



Final Presentation

Spotify 1.2M+ Songs • *Mazza Luca, Hoxha Roeld, Masciocchi Andrea*



Spotify 1.2M+ Songs

What does the dataset represent?

- **Songs** - Spotify's dataset of 1.2M+ entries (1'204'025)
- **Row** - Each row represents a song
- **Columns** - Each column is a characteristic of the song (24 columns)
- **Time Span:** 1908 - 2020 (Publication Date)
- **Worldwide** - Songs from all over the world

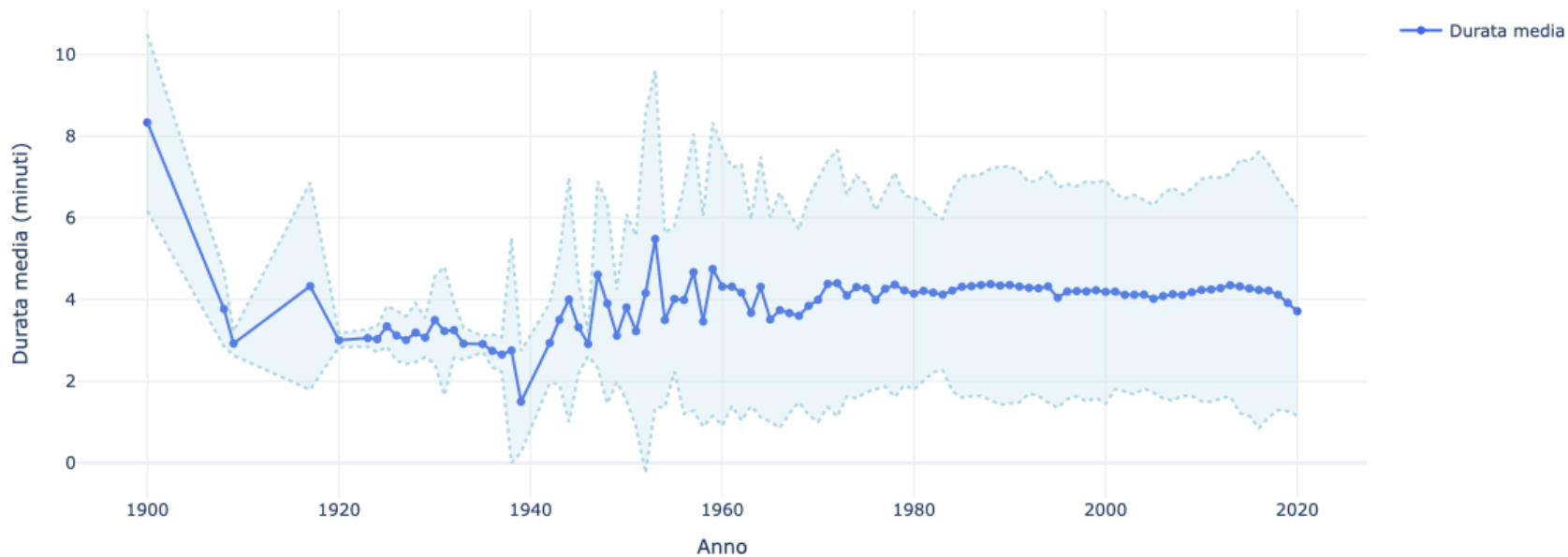
Why?

Why is this dataset interesting (to us)?

- Spotify is an every-day use for most of us
- Data Credibility
- Songs from all over the world
- Changes over time / history on music

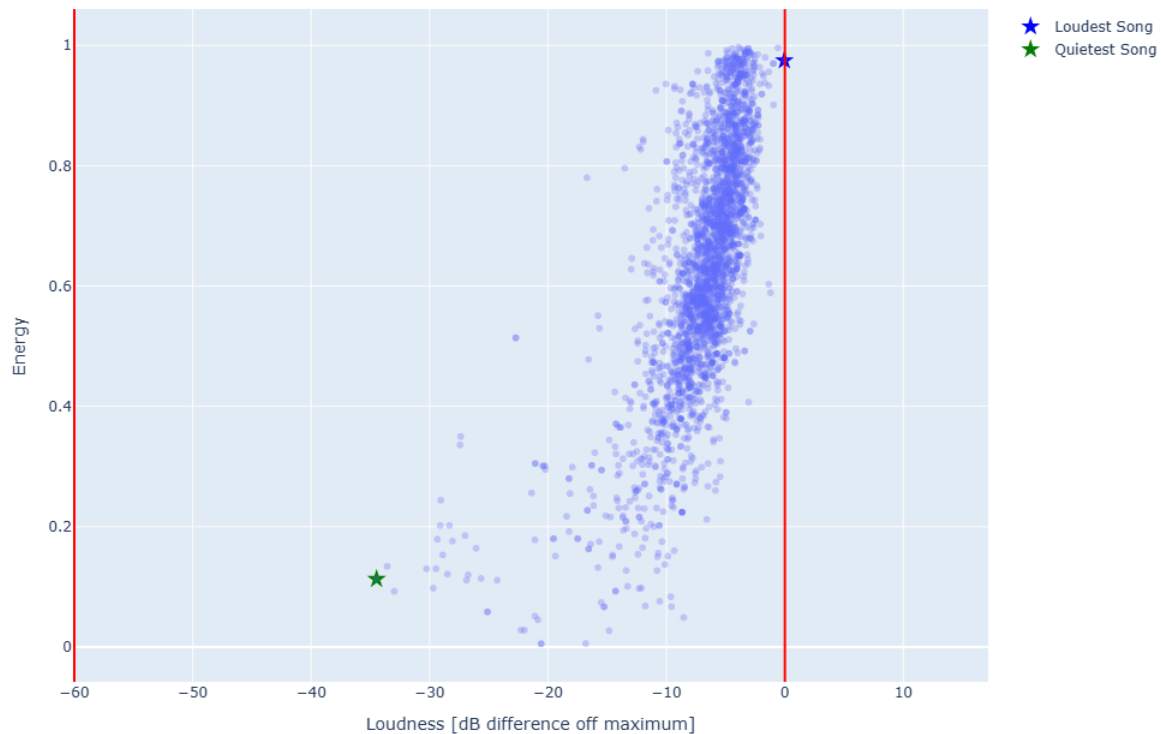
Average Song Duration

Average Song Duration from 1900 to present (with st. deviation)



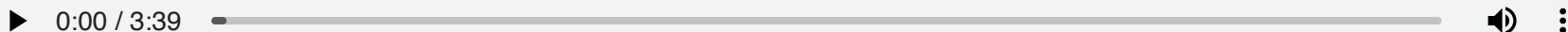
Loudness vs Energy

Loudness vs Energy (Top 500 Most Popular Artists)

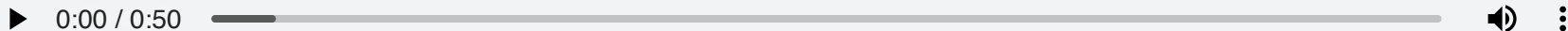


Songs Samples

💣 **Loudest Song:** Alarm - Marshmello Remix - *Anne-Marie, Marshmello*

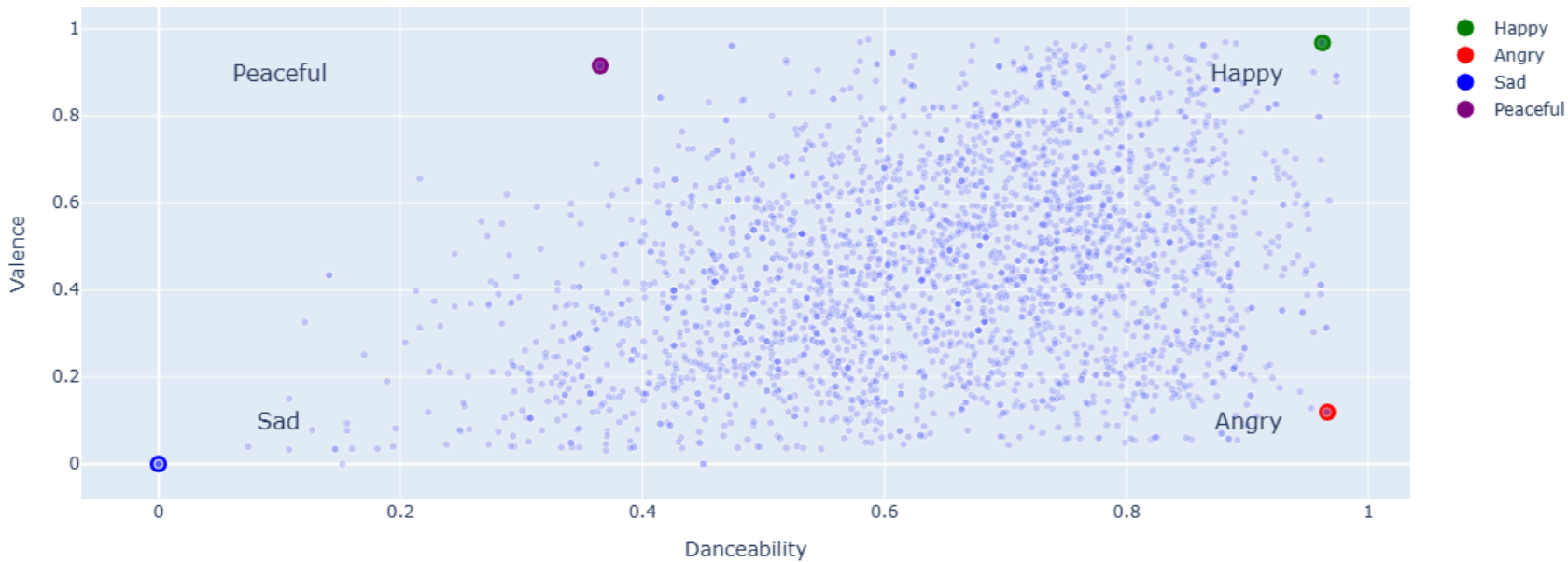


🤫 **Quietest Song:** The Explanation - *XXXTENTACION*



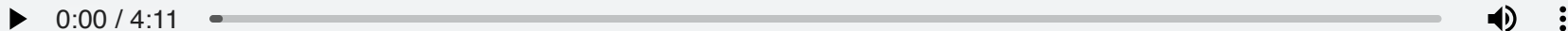
Danceability vs Valence

Danceability vs Valence (Top 500 Most Popular Artists)

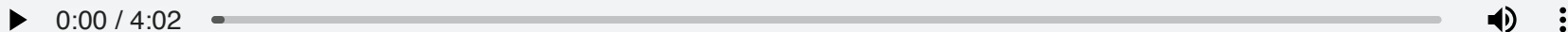


Songs Samples

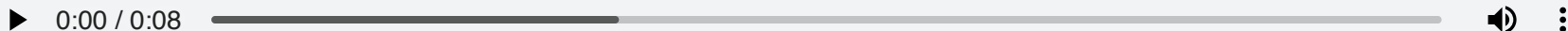
🎉 **Happy:** Merrymaking at My Place - *Calvin Harris*



😡 **Angry:** good day - 21 Savage



😭 **Sad:** Vault Character - Calvin Harris



😊 **Peaceful:** Leave Before the Lights Come On - *Arctic Monkeys*

Songs by Genre

Thank you!

Dataset · GitHub