Lyrical Emotion across Cultures

Milestone 1 Report; Team InSight; Team Member: Xu Weilun, Huang Xin, Wang Qi

1. Dataset

The WASABI [1] is a large-scale, multi-faceted dataset that contains metadata for over 2 million commercially released songs, along with information on approximately 200K albums and 77K artists. It integrates cultural metadata, lyrics annotations, and audio analysis results. The dataset is available here.

Key Features:

- **Lyrical Emotion**
  - Valence and arousal scores computed using NLP techniques.
  - Emotion tags collected on Last.fm (human data)
- **Language:** Categorical variable indicating the language of the lyrics.
- **Genre:** Musical genre classification.
- **Explicitness:** Indicator of whether a song's lyrics are explicit.

Data Quality and Preprocessing Considerations

>>> Strength: Diverse Metadata Sources

The dataset aggregates information from several external sources such as LyricsWikia, MusicBrainz, DBpedia, Deezer, Discogs, and more. This integration provides detailed data on song structure, lyrical topics, *emotions (human annotations)*, explicit content, chord analysis, and additional cultural context.

>>> Strength: Rich Annotations

It includes a wide range of annotations extracted using NLP and ML techniques. Examples are structural segmentation of lyrics, topic modeling via LDA, *emotion (valence-arousal coordinate) regression based on BERT*, and automatic chord extraction with confidence measures. These features allow for sophisticated cross-cultural and emotion analysis of lyrical content.

>>> Concern: Completeness

while the dataset is extensive, metadata was collected from different platforms, not all entries are complete or consistent, necessitating validation and possibly imputation.

>>> **Concern: Reliability of Emotion Annotations**

- BERT is typically pre-trained on large corpora that might be predominantly Western or otherwise not fully representative of all cultural expressions. This can lead to models that misinterpret or underrepresent emotions in lyrics that employ culturally specific idioms, references, or sentiment expressions.
- Emotional expression in music can vary significantly across cultures. What is considered a strong expression of sadness or joy in one cultural context might be expressed very differently in another. A BERT-based model may not capture these nuances if its training data lacks diversity, potentially resulting in inaccurate emotion predictions for non-Western lyrical content.
- Suitability: This dataset is well-suited for our project as it provides both textual lyrics for NLP analysis and structured metadata for trend visualization.

We may need to incorporate additional context-aware features that better capture the cultural variations in emotional expression.

2. Problematic

Main Goal: Analyze and visualize the evolution of lyrical emotions across different languages, cultures, and time periods. By leveraging NLP and data visualization, we seek to uncover patterns in how emotions are expressed in song lyrics worldwide across different cultural contexts.

Example Questions

- How have lyrical emotions evolved from 1960 to 2020?
- Do different languages or cultures exhibit distinct emotional trends in their lyrics?
- How do specific lyrical themes correlate with emotional expressions?

Main Axis for Development

1. **Emotional Dimensions Across Cultures:**

- **Valence and Arousal:** Visualize the distribution of valence (positivity/negativity) and arousal (intensity) scores across various cultural groups and languages.
- **Comparative Analysis:** Highlight differences and similarities between regions or language groups to uncover unique cultural expressions of emotion.

2. **Correlation with Cultural Metadata:**

- **Genre and Explicitness:** Integrate metadata such as genre classification and explicit content flags to examine how these factors intersect with the emotional content.
- **Temporal Trends:** Consider adding a time-based axis to explore changes in lyrical emotion over decades, reflecting evolving cultural contexts.

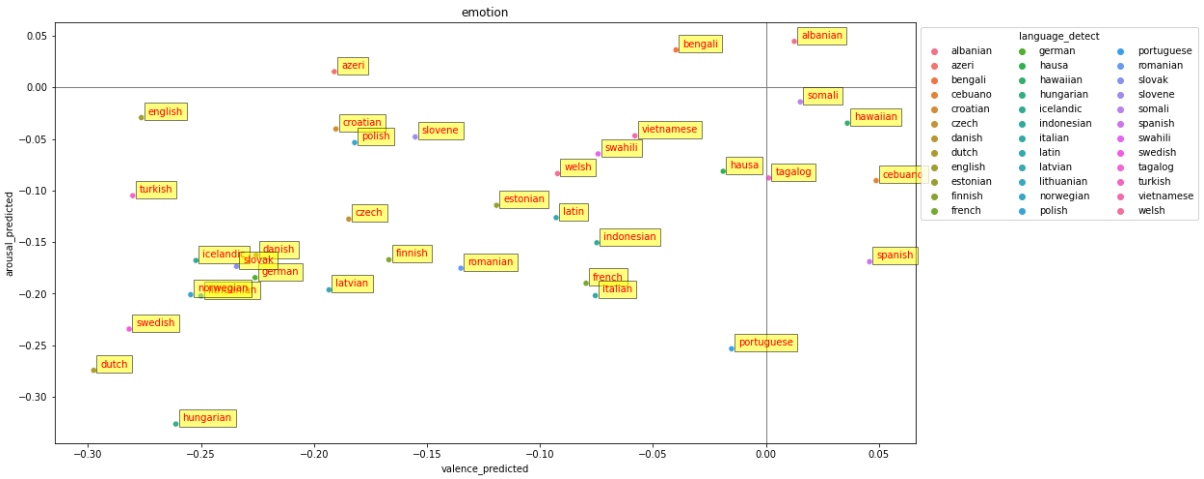
3. **Interplay Between Audio and Lyrical Features:** Although the focus is on lyrical emotion, include insights on how audio features (e.g., tempo, chords) might correlate with the emotional tone of the lyrics.

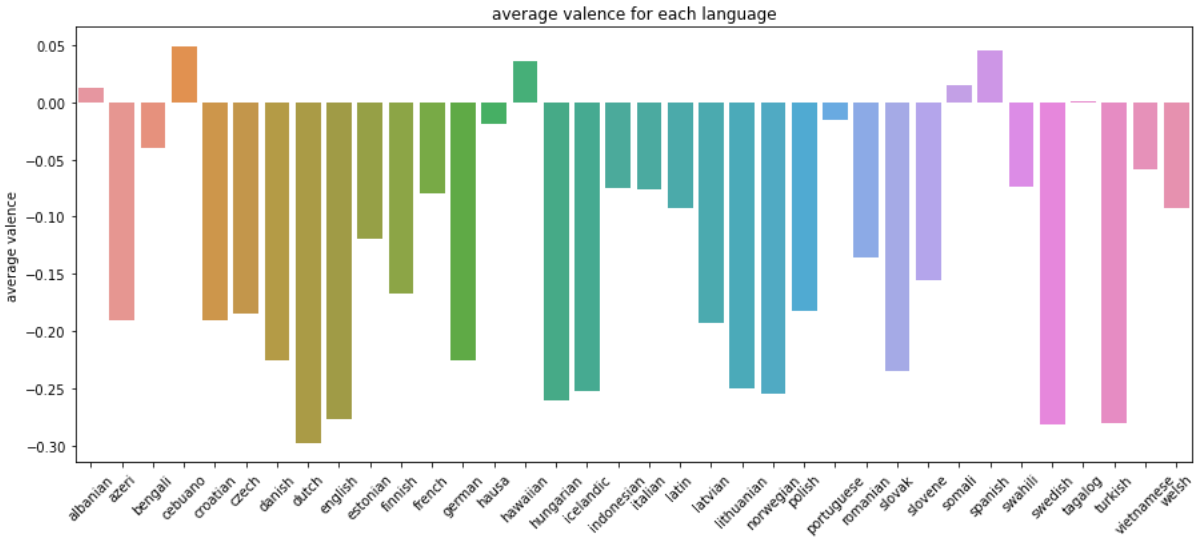
Target Audience

- **Academics:** Musicologists, cultural studies experts, and NLP researchers interested in the intersection of culture, emotion, and music.
- **Industry:** Data scientists, music streaming service analysts, and recommendation system developers seeking to refine their models with culturally nuanced insights.
- **General Public:** Enthusiasts and cultural commentators who are curious about how different communities experience and express musical emotions.

3. Exploratory Data Analysis

Emotion distribution across languages



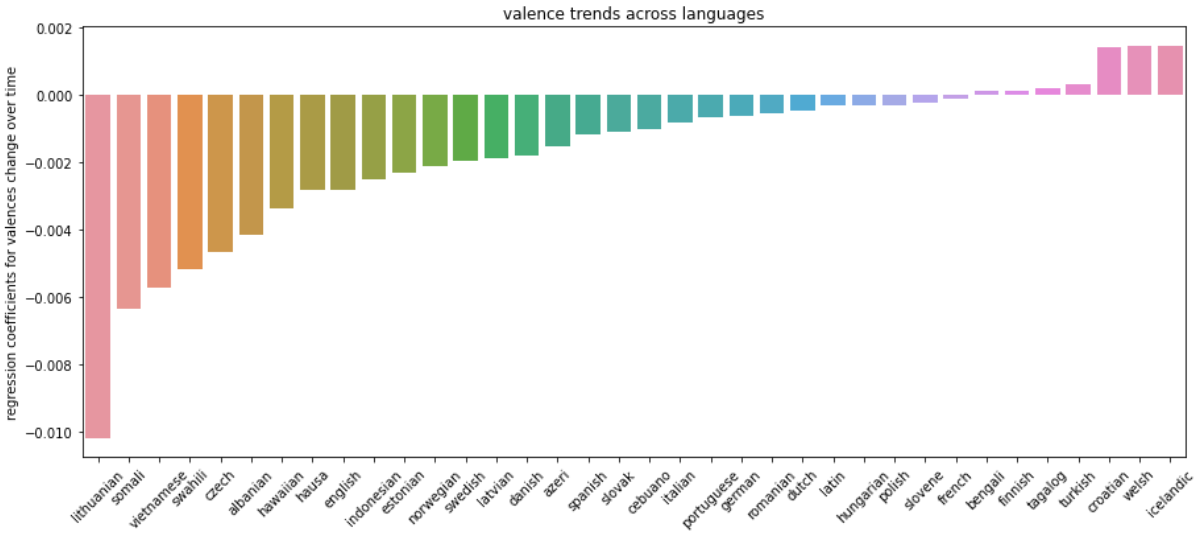


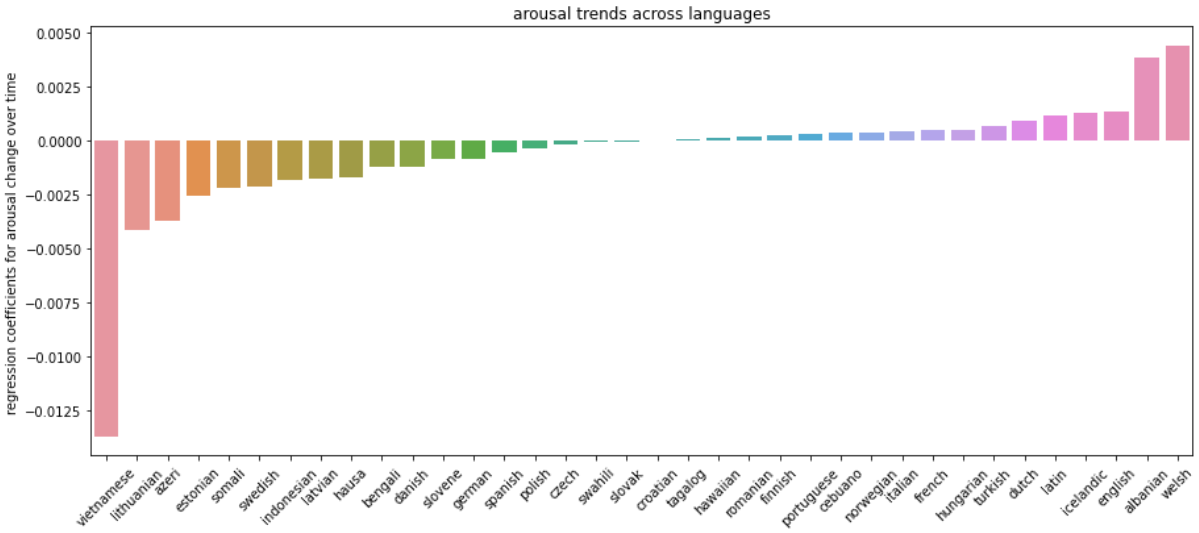
The scatterplot illustrates how various languages cluster in terms of their emotional attributes, with some (e.g., **English, Turkish, and Swedish**) tending toward moderate valence and arousal, while others (e.g., **Portuguese, Albanian, and Somali**) exhibit higher emotional intensity. Notably, **Bengali and Azeri** appear in the high-arousal quadrant, suggesting more emotionally charged lyrical content. The observed variation underscores cultural differences in musical expression, where some languages lean toward **calm and neutral tones**, while others exhibit **vibrant and expressive lyrical patterns**.

Notably:

- languages that generally show positivity: ['albanian', 'cebuano', 'hawaiian', 'somali', 'spanish', 'tagalog']
- languages that generally show high intensity: ['albanian', 'azeri', 'bengali']

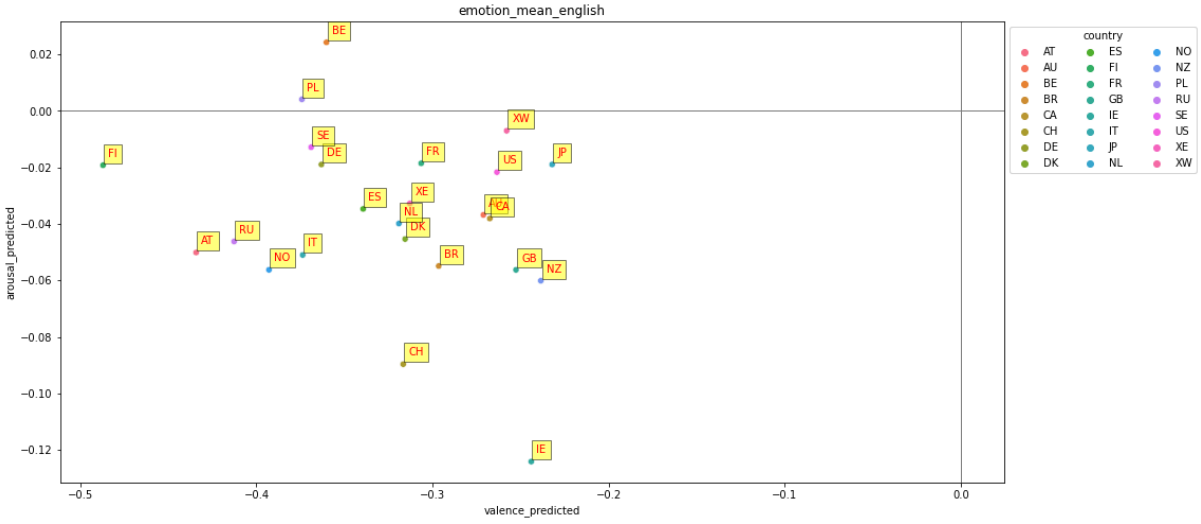
Trend Analysis





The analysis of **valence and arousal trends across languages** reveals significant shifts in lyrical emotion over time. While some languages exhibit a trend toward **more positive sentiment** in lyrics, others show a decline in valence, indicating a shift toward darker or more melancholic themes. Similarly, arousal trends highlight that certain languages have experienced a **drop in energetic intensity**, reflecting a move toward calmer, more subdued musical styles, while others display an increase in arousal, suggesting a preference for more emotionally intense lyrics.

Emotion distribution across countries



- **Language-based analysis** captures **the core emotional tendencies** of lyrics regardless of geographical location. For instance, **Spanish songs** may exhibit high positivity whether they originate from **Spain, Mexico, or Argentina**.

- **Country-based analysis** highlights **regional influences** and the impact of **multilingualism** within nations. For example, **Canada and Belgium** exhibit mixed valence-arousal scores due to their **multilingual populations**.
- Some trends align between both analyses, such as **Russian** (language) and **Russia** (country) tending toward lower valence, but **differences emerge** where global music trends influence national averages.
- **Industry & Media Effects**: Certain countries, like the **U.S. and U.K.**, dominate the global music scene and may exhibit more **blended emotional distributions** due to diverse influences.

Feature	Emotion Distribution by **Language**	Emotion Distribution by **Country**
Main Clustering Factors	Linguistic structure, cultural sentiment, and lyrical traditions.	Geopolitical regions, historical influences, and shared music industry trends.
Diversity of Points	Shows a broad range of emotional expressions across languages, with clear groupings.	More overlap, possibly due to multilingual influences within countries.
Valence (Positivity)	Some languages (e.g., **English, Spanish**) trend towards higher valence, while others (e.g., **Russian, Turkish**) show lower valence.	Certain countries (e.g., **Brazil, France, and the U.S.**) align with higher valence, whereas others (e.g., **Russia, Austria**) lean towards lower positivity.
Arousal (Energy Level)	Languages like **Bengali and Azeri** exhibit high arousal, while **Dutch and Swedish** are more subdued.	Countries like **Belgium and Poland** show higher arousal, while **Ireland and Switzerland** appear calmer.
Cultural Interpretation	Highlights **linguistic-driven emotion expression**, independent of geography.	Suggests that national identity and **global music trends** influence emotions beyond language alone.

- **Language-based analysis** reflects **how lyrics are emotionally expressed** based on **linguistic and cultural roots**.
- **Country-based analysis** captures **how geography, history, and industry trends** shape emotional trends in music.
- **Combining both perspectives** provides a deeper understanding of **how music emotions are shaped by both linguistic tradition and national culture**, offering valuable insights for cultural music studies and global trends.

4. Visualization Approach

1. **Static Analytical Visualizations**:
    - Time series plots tracking emotional trends across decades and cultures
    - Heatmaps displaying correlation between lyrical themes and emotions
    - Comparative bar charts for cross-cultural sentiment analysis
    - Network graphs showing relationships between languages, emotions, and themes
2. **Interactive Web Experience**:
    - Animated transitions showing the evolution of emotional content over time
    - Interactive filters allowing users to explore data by language, genre, time period, and emotion
    - Narrative-driven visualizations that guide users through key findings while allowing for open exploration

The visualization will employ a cohesive color scheme where different emotions are represented by distinct colors (e.g., pink for love, blue for sadness) across all visualization types, ensuring consistency in user experience and interpretation.

5. Related Work

Previous Work in Lyrical Emotion Analysis

This dataset has been leveraged in multiple academic investigations to explore NLP-based lyric analysis, multimodal music recommendation, cultural trend evaluations, and bias detection in song lyrics.

Several foundational studies establish WASABI's role as a structured knowledge graph and metadata-rich dataset. Fell et al. [**2**] examine the enrichment of the corpus through annotated lyrics, applying NLP techniques for sentiment analysis, explicit content classification, and structural analysis. Building on these developments, later research explores the dataset's functionalities for interactive navigation and data-driven tools [3].

Academic uses of WASABI extend beyond dataset structuring into applied research domains. Studies such as Betti et al. [4] employ the dataset to perform gender bias analysis in lyrics, while Monti et al. [5] integrate WASABI-derived lyric features into playlist recommendation systems. In another application, Fell et al. [6] leverage the dataset for lyrics segmentation tasks, aligning lyrics data with genre metadata to enhance text-music correlation studies.

Despite its extensive multimodal potential, challenges remain in fully realizing WASABI's capabilities in cross-modal AI applications and real-world implementations. Research has

predominantly focused on textual metadata and structural segmentation, while deeper multimodal feature fusion—combining lyrics, harmony, and sentiment analysis in predictive systems—remains underexplored. Additionally, issues such as metadata bias and ethical considerations warrant further study [1,4].

Beyond emotion annotation, the dataset has been employed for broader sociolinguistic examinations. One study utilized WASABI to investigate gender bias and sexism in song lyrics, using NLP techniques to assess linguistic patterns across five decades of popular music [3]. While providing insights into lyrical trends, this study did not engage with WASABI's emotion metadata or analyze cultural differences in lyrical emotion expression.

Despite the dataset's extensive annotations and its potential for large-scale analyses, no research has yet leveraged WASABI specifically for cross-cultural comparisons of lyrical emotion. This gap presents an opportunity for future studies to apply the dataset's emotion labels in cross-cultural contexts, validating linguistic and cultural differences in emotional expression through comparative frameworks.

How Our Approach is Original

While previous studies have analyzed lyrics, our approach is original in several key ways:

1. Multi-dimensional analysis combining emotional content, linguistic features, and cultural context in a single interactive visualization
2. Cross-linguistic comparison that goes beyond simple sentiment analysis to examine narrative structures across different languages and cultures
3. Temporal evolution visualization that allows users to see patterns across decades rather than isolated snapshots
4. Integration of pre-trained LDA topic models with emotion annotations to reveal connections between thematic content and emotional expression
5. Development of an interactive web experience that makes complex linguistic and emotional data accessible to both researchers and the general public through engaging, animation-based storytelling

Unlike most existing tools that focus only on English lyrics or basic sentiment classification, our visualization will provide insights into the nuanced ways emotional expression in music reflects cultural values and historical changes across the globe.

Sources of Inspiration

- Statistical visualisations from prior sentiment analysis projects on lyrics, such as FiveThirtyEight's work on song lyrics during 90s boy band.
- Lottie's "How Lyrics Have Changed Through the Decades" visualization approach.

References

[1] Buffa, M., Cabrio, E., Fell, M., Gandon, F., Giboin, A., Hennequin, R., ... & Winckler, M. (2021). The WASABI dataset: cultural, lyrics and audio analysis metadata about 2 million popular commercially released songs. In *The Semantic Web: 18th International Conference, ESWC 2021, Virtual Event, June 6–10, 2021, Proceedings 18* (pp. 515-531). Springer International Publishing

[2] Fell, M., Cabrio, E., Korfed, E., Buffa, M., & Gandon, F.L. (2019). Love Me, Love Me, Say (and Write!) that You Love Me: Enriching the WASABI Song Corpus with Lyrics Annotations. *ArXiv, abs/1912.02477*.

[3] Buffa, M., Lebrun, J., Pauwels, J., & Pellerin, G. (2019). A 2 Million Commercial Song Interactive Navigator.

[4] Betti, L., Abrate, C. & Kaltenbrunner, A. Large scale analysis of gender bias and sexism in song lyrics. *EPJ Data Sci.* **12**, 10 (2023). https://doi.org/10.1140/epjds/s13688-023-00384-8

[5] Monti, D., Palumbo, E., Rizzo, G., Lisena, P., Troncy, R., Fell, M., ... & Morisio, M. (2018). An ensemble approach of recurrent neural networks using pre-trained embeddings for playlist completion. In *Proceedings of the ACM Recommender Systems Challenge 2018* (pp. 1-6).

[6] Fell, M., Nechaev, Y., Cabrio, E., & Gandon, F.L. (2018). Lyrics Segmentation: Textual Macrostructure Detection using Convolutions. *International Conference on Computational Linguistics*.