Tunes Beyond Borders:
Exploring Music Taste Similarities on Spotify Across Regions

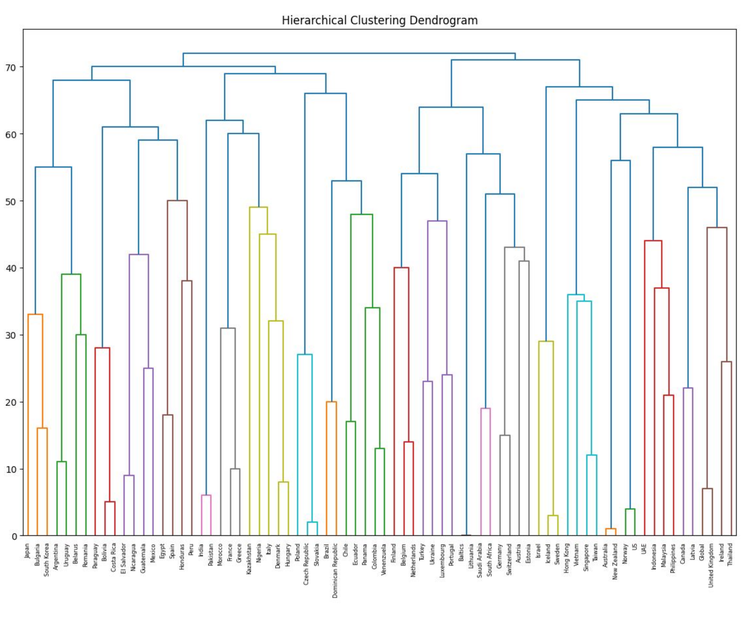
Nan-Hsin Lin, Yi-Chun Wang

Abstract

People from different regions often prefer to songs in their native languages, reflecting the cultural and emotional connections fostered by familiar lyrics and melodies. However, music also has the potential to cross language barriers, with global hits resonating in regions far from the language of the lyrics. This duality raises fascinating questions: What drives the global appeal of certain songs? How similar or distinct are the musical preferences of listeners across regions? The project aims to uncover the extent of overlap in music tastes across regions and identify opportunities for cross-cultural music recommendations.

Data

- Scope: 74 Spotify regional & global daily chart (Nov 2024)
- 12 Features: acousticness, danceability, energy, instrumentalness, key, liveness, loudness, mode, speechiness, tempo, valence, and duration (in milliseconds)
- A total of 4716 tracks

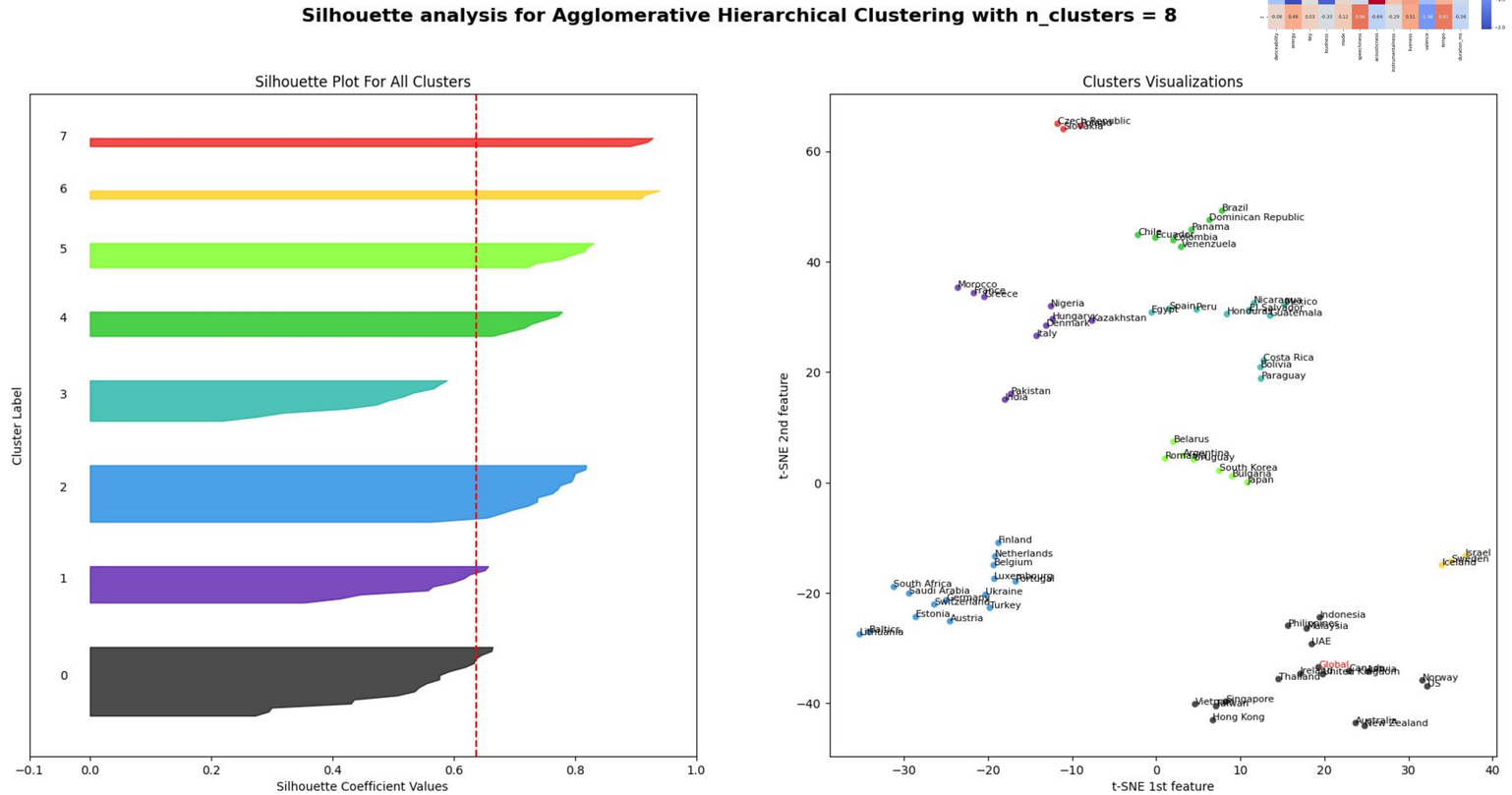


Methodology & Benchmark

- Dimensionality reduction techniques: Principal Component Analysis (PCA), Multidimensional Scaling (MDS), and t-SNE.
- Clustering algorithms: K-Means, Gaussian Mixture Model (GMM), Agglomerative Hierarchical Clustering, and DBSCAN.
- Benchmark: A silhouette score greater than 0.5 indicates good clustering performance.

Best Model: Agglomerative Hierarchical Clustering with t-SNE

The most effective clustering model turned out to be Agglomerative Hierarchical Clustering applied to t-SNE reduced features with n_components=2 and n_clusters=8, achieving a silhouette score of 0.64.



Findings

- Music taste indeed crosses national boundaries, with unexpected similarities between distant regions.
- Regional proximity still plays a role, as seen in the similarity between Australia and New Zealand.
- Cluster 6 (Iceland, Israel, and Sweden) is characterized by low danceability, energy, and loudness but high acousticness, reflecting a preference for acoustic, softer music.
- Cluster 7 (Czech Republic, Poland, and Slovakia) exhibits higher energy, faster tempo, and lower valence, suggesting a preference for energetic, fast-paced, and emotionally negative tracks.