SOC-25 FINAL PROJECT

MUSIC RECOMMENDATION SYSTEM

Submitted by
Esha Yindukuri
24B2119

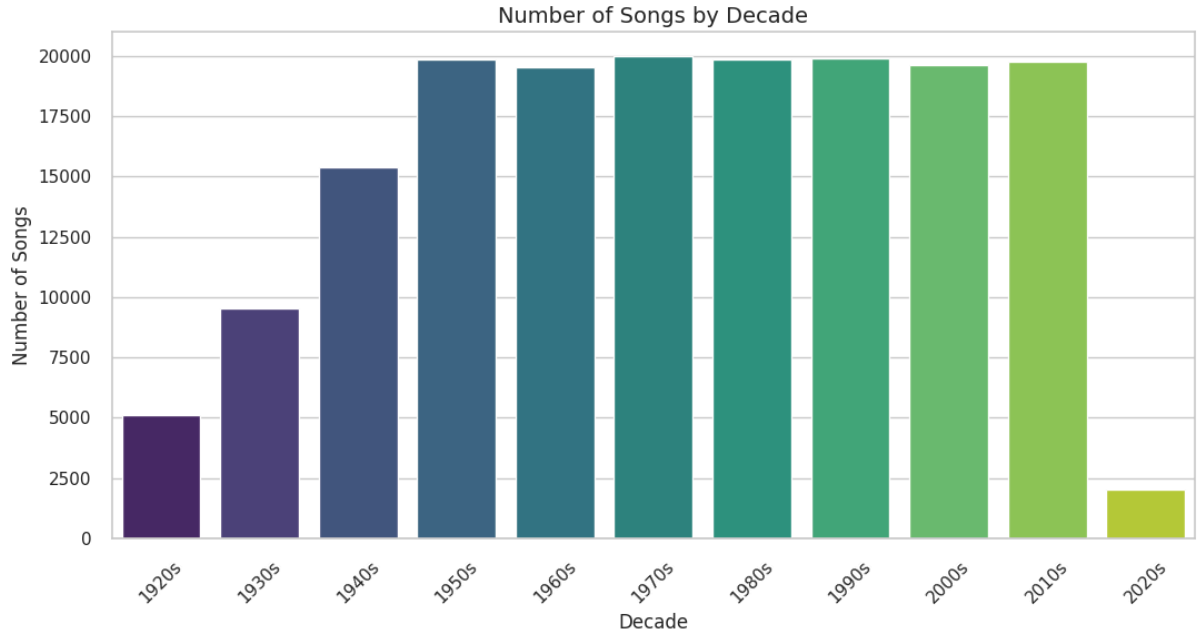
Objective

The goal of this project is to build a content-based music recommendation system using the Spotify Dataset, which contains audio features of thousands of songs across different genres and years. This will

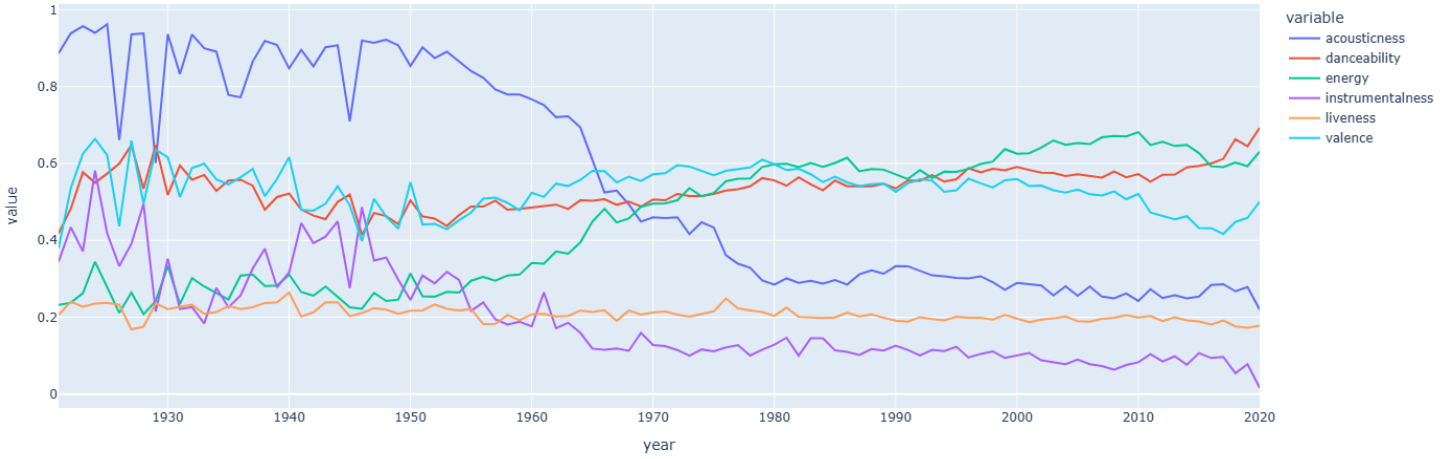
- Understand the characteristics and trends of music using exploratory data analysis (EDA)
- Group similar songs and genres using clustering algorithm
- Recommend songs that sound similar to a given input song using audio-based features and similarity metrics

Exploratory Data Analysis

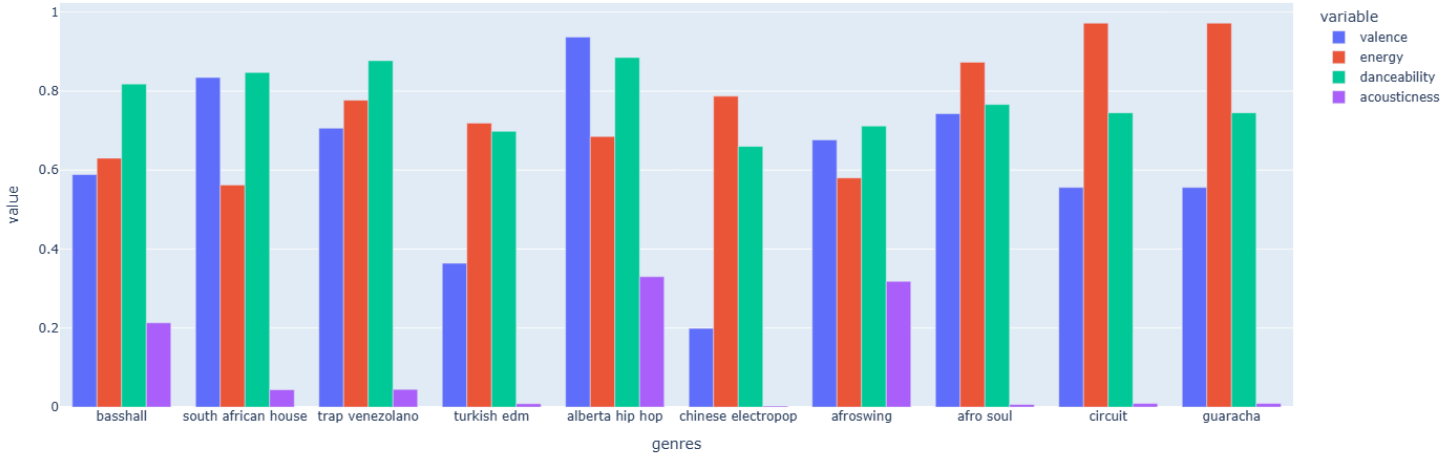
Songs Released Over Decades : Shows how music production has varied over time. Peaks indicate popular eras in music history.



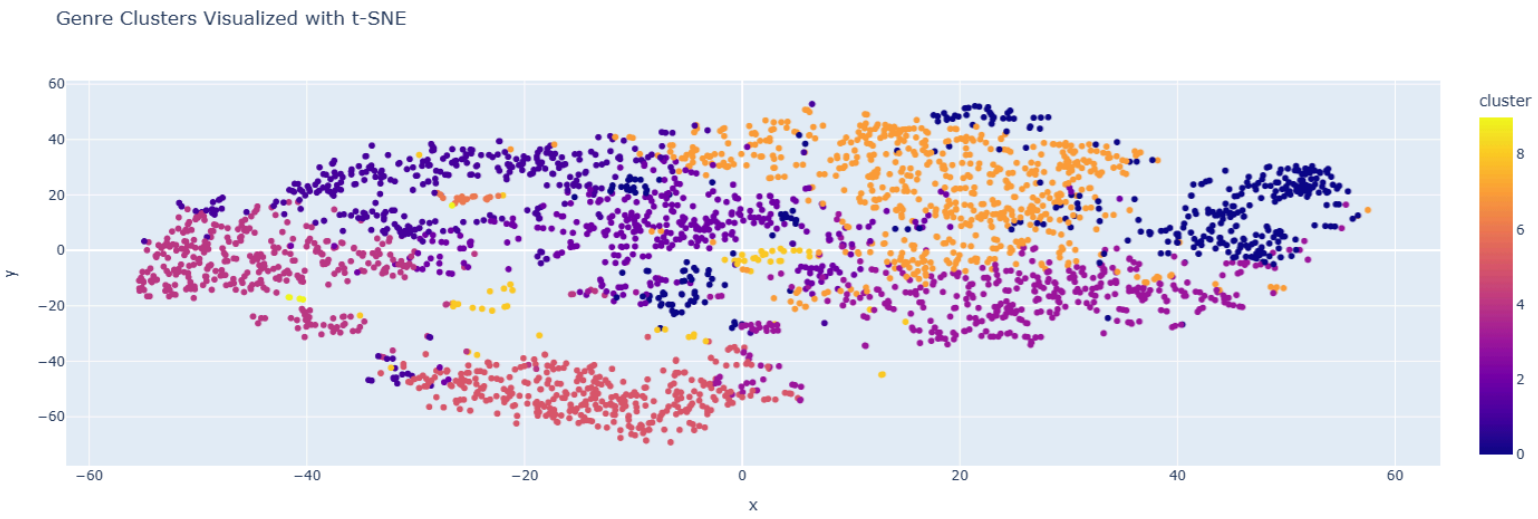
Sound Features Over Time (1921–2020): Tracks changes in audio features like energy, danceability, and valence over years.



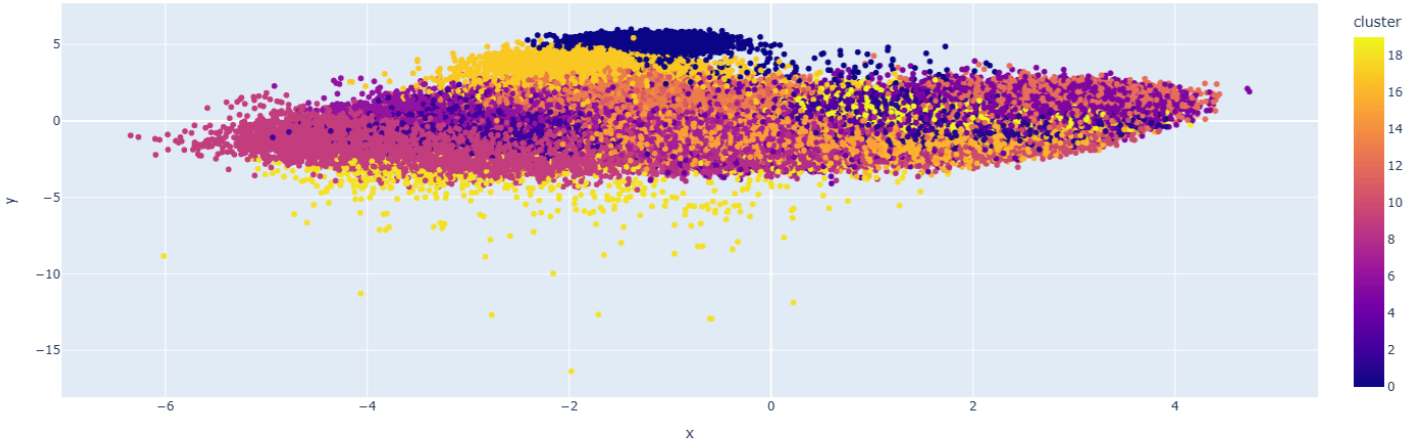
Top 10 Popular Genres – Feature Comparison : Compares key audio traits (valence, energy, etc.) for the most popular genres.



Genre Clusters with t-SNE



Song Clusters with PCA



Recommendation system using cosine similarity

```
Enter song name: Don't start now
Enter artist name: Dua Lipa

🎧 Recommended Songs:
```

Song	Artist	similarity
Don't Start Now	['Dua Lipa']	0.999847
Believer	['Imagine Dragons']	0.994600
New Rules	['Dua Lipa']	0.984614
Waka Waka (This Time for Africa) [The Official 2010 FIFA World Cup (TM) Song] (feat. Freshlyground)	['Shakira', 'Freshlyground']	0.983978
One Kiss (with Dua Lipa)	['Calvin Harris', 'Dua Lipa']	0.982699
Only Girl (In The World)	['Rihanna']	0.980289
Disturbia	['Rihanna']	0.979301
Piece Of Your Heart	['MEDUZA', 'Goodboys']	0.979151
Filter	['BTS']	0.976263
Can't Feel My Face	['The Weeknd']	0.975325