



ANALYZING GLOBAL MUSIC TRENDS USING SPOTIFY AUDIO FEATURES

A Data-Driven Analysis in R

Shristi

B.Tech CSE (Data Science), Indian Institute of Information Technology, Una (IIIT UNA)



useR! 2026

July 6-9, 2026
Warsaw, Poland

Dr. Prince Sharma

Faculty, School Of Computing, IIIT Una



114,000+
Tracks

Spotify Dataset

2.24% R^2

Audio features explain only 2.24%
of popularity variance.

Popularity is likely influenced by external factors beyond audio features.



8 Audio
Features

Predicting Popularity
(0-100)

1. RESEARCH QUESTION

Which Spotify audio features significantly influence song popularity, and how much variance do successive feature blocks explain?

2. DATASET

- 114,000+ Tracks
- 22 Variables
- 8 Audio Features
- 0 Missing Values
- Global Dataset

Outcome Variable

- Popularity (0-100)

3. METHODOLOGY



Spotify Dataset



Data Cleaning

(clean_names, drop_na, scaling)



Exploratory Data Analysis

(Distributions, Outliers)



Correlation Analysis

(Feature Relationship)



Hierarchical OLS Regression

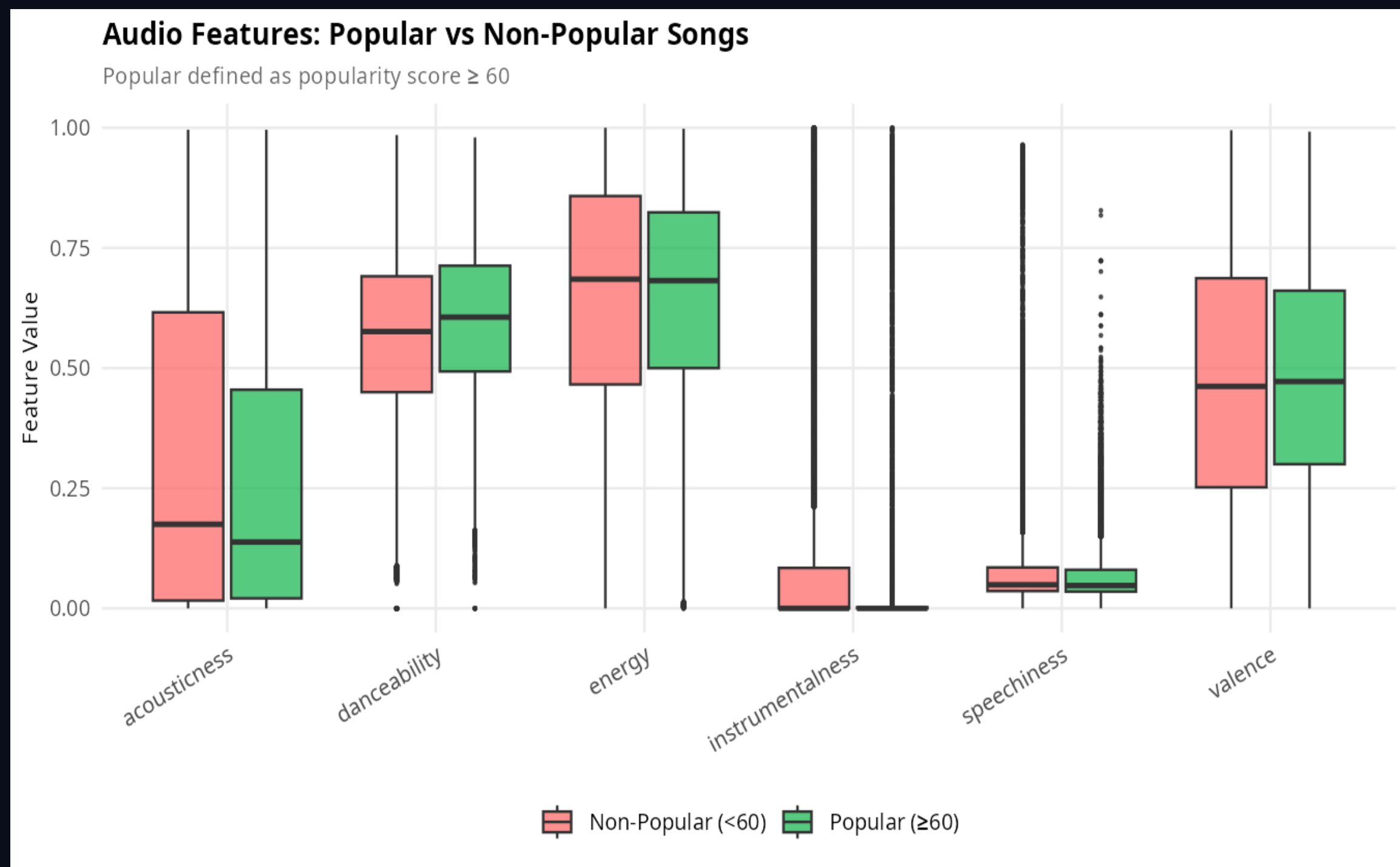
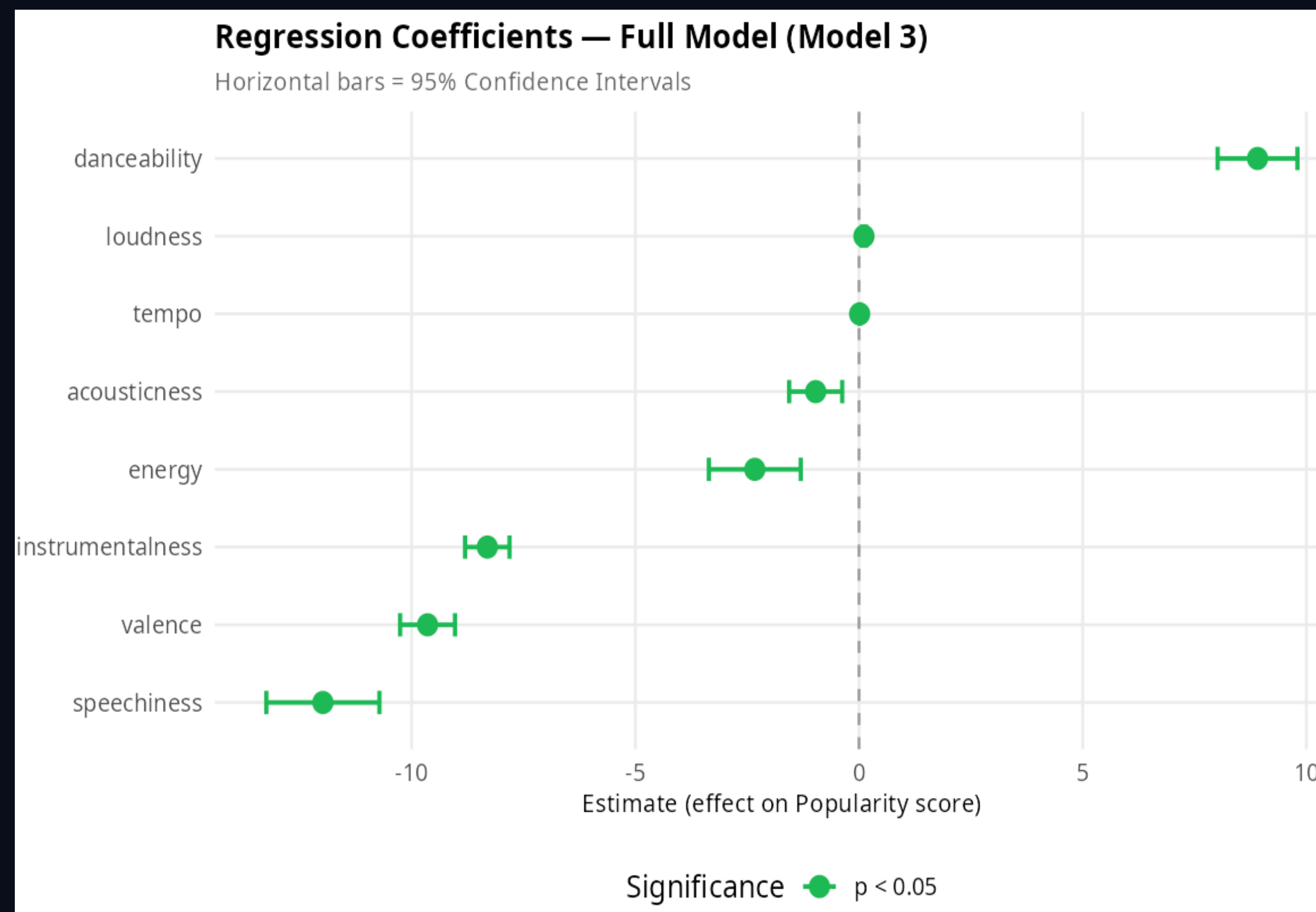
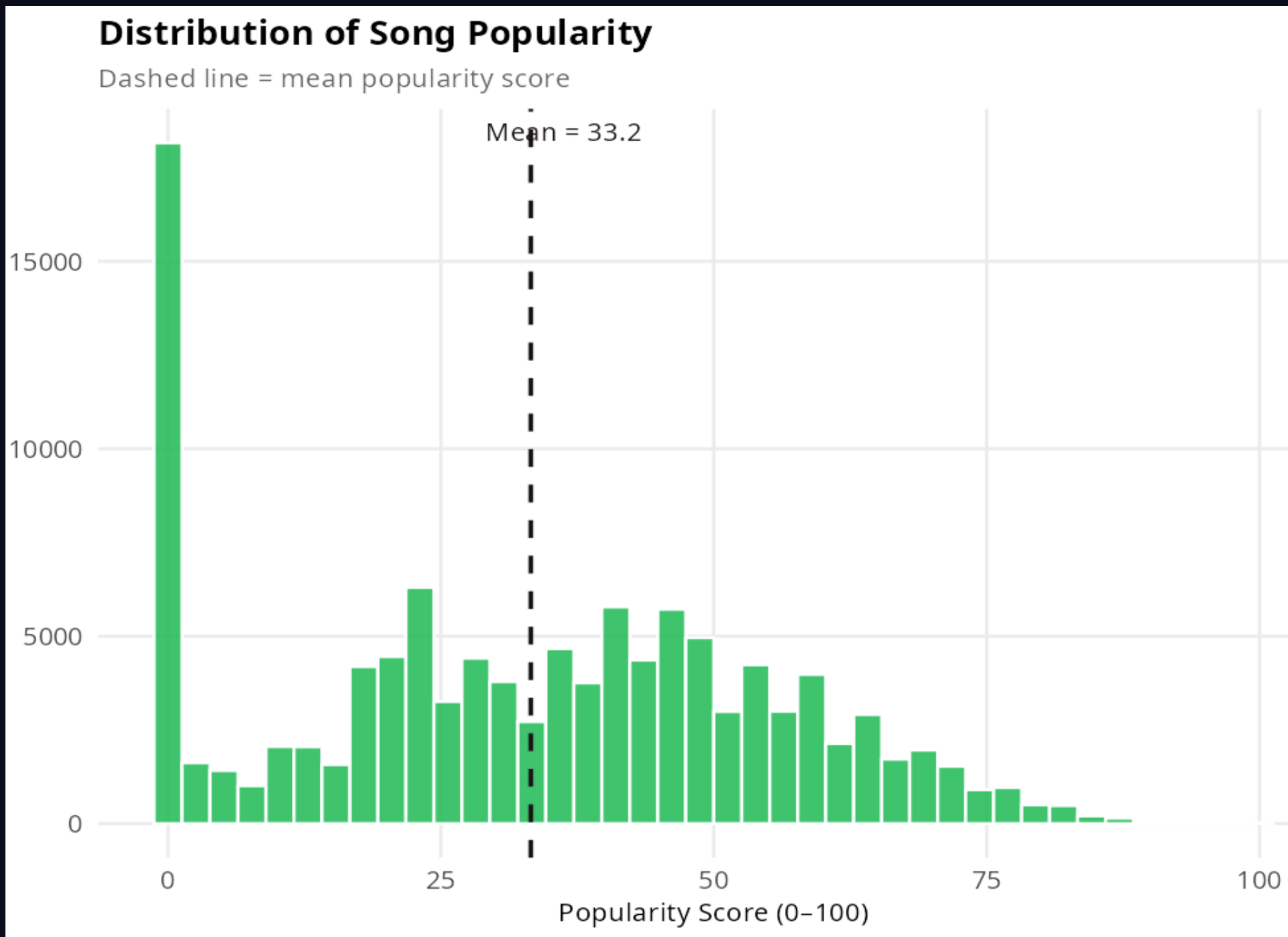
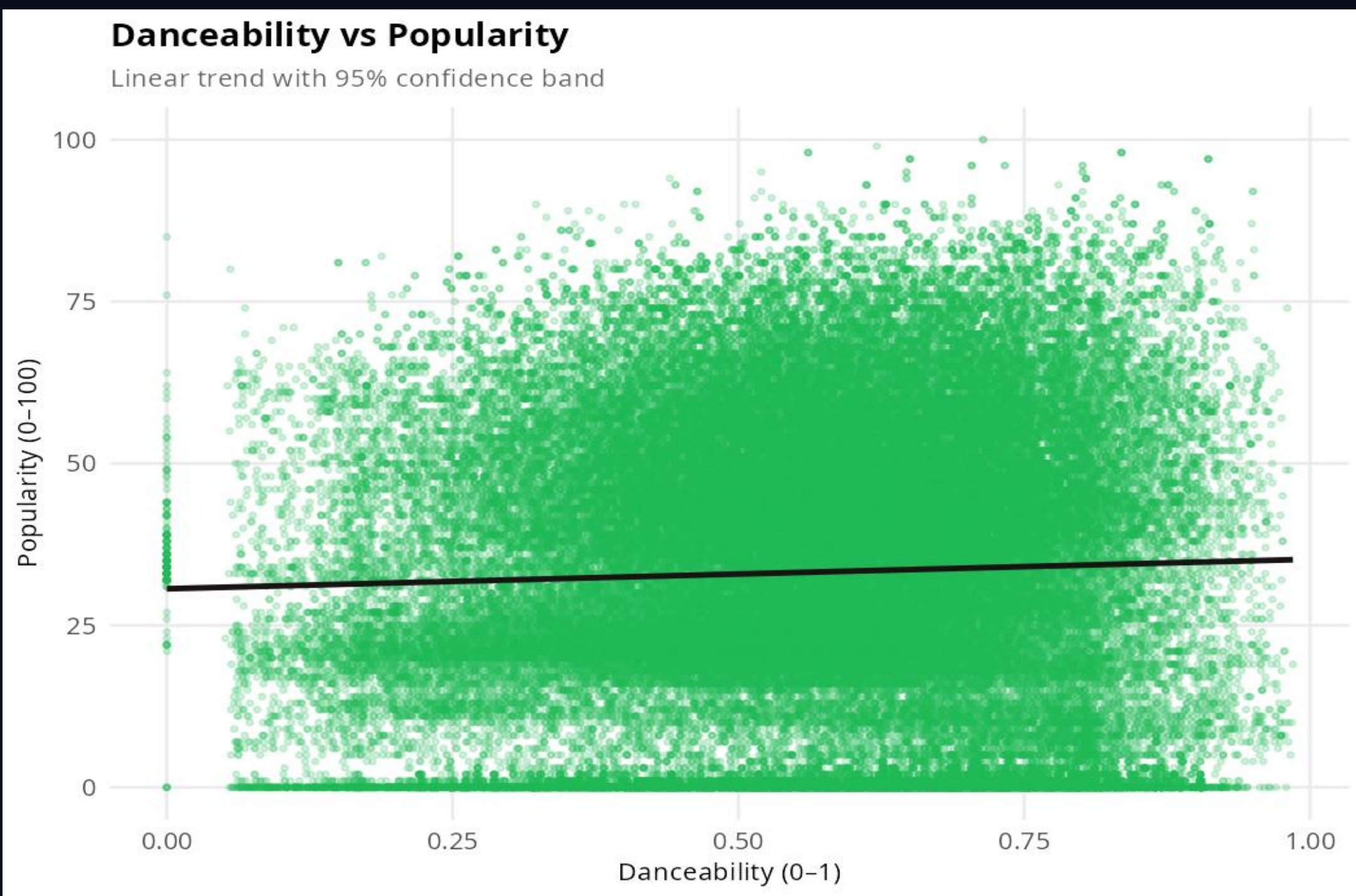
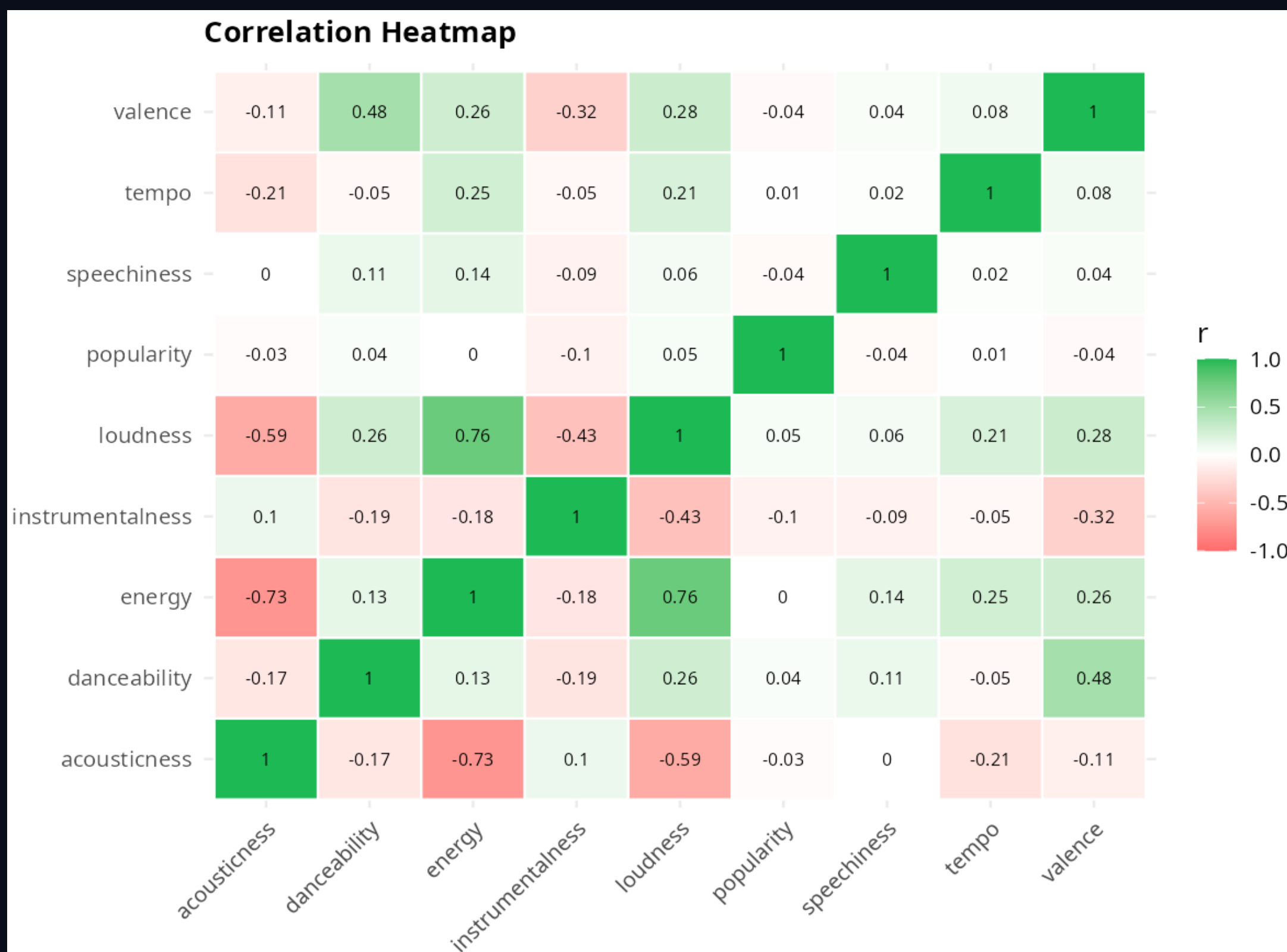
(Three Models)



Model Evaluation

(R^2 , Adj. R^2 , AIC, ANOVA)

4. EXPLORATORY DATA ANALYSIS (EDA)



7. FUTURE SCOPE

- Genre-wise popularity modelling
- ML-based prediction
- Temporal music trend analysis
- Artist metadata integration
- Playlist and social media impact analysis

8. CONCLUSIONS

- Danceability increases popularity, while speechiness and instrumentalness reduce.
- Audio features explain only 2.24% of popularity variance, highlighting the importance of external factors in song success.

5. MODEL PERFORMANCE (Hierarchical Regression)

Model	Predictor Block	R^2	Key Insight
Model 1	Danceability, Energy, Valence	0.0057	Baseline
Model 2	+ Acousticness, Instrumentalness, Speechiness	0.0218	Largest Gain
Model 3	+ Tempo, Loudness	0.0224	Small Improvement

6. KEY FINDINGS

- Danceability was the strongest positive predictor ($\beta = +8.91$)
- Speechiness was the strongest negative predictor ($\beta = -12.00$)
- Instrumental tracks showed lower popularity ($\beta = -8.31$)
- Loudness and tempo had small positive effects
- External factors (artist reputation, marketing, playlist placement) likely play a larger role in song success
- Audio features explained only 2.24% of popularity variance