The Changing Sound of Music:
An Exploratory Corpus Study of Vocal Trends Over Time

Elena Georgieva[1], Pablo Ripollés[123], Brian McFee[14]

[1]NYU Music & Audio Research Lab (MARL), [2]NYU Dept. of Psychology, [3]NYU Center for Language Music & Emotion (CLAME), [4]NYU Center for Data Science

Abstract

- We conducted an exploratory study of **43,153 vocal tracks** of popular songs spanning from **1924 to 2010.**
- We used **source separation** to extract the vocal stems and **fundamental frequency (f0)** estimation to analyze pitch tracks.
- We extracted the **mean pitch, total variation,** and **pitch class entropy** of each song.
- We conducted statistical analysis of vocal pitch across years and genres, and report significant trends in our metrics over time and between genres.

Dataset

- We used the union of the **HSP-S and HSP-L Datasets** ("Hit Song Prediction- Small and Large," respectively)[1].
- Audio files **30-60 seconds in length** were taken from a private mp3 sample collection of the MSD [2] .
- Songs are generally widely listened-to, and the majority come from North America or Europe.
- We observe a strong bias towards more recent songs, especially after 1990.
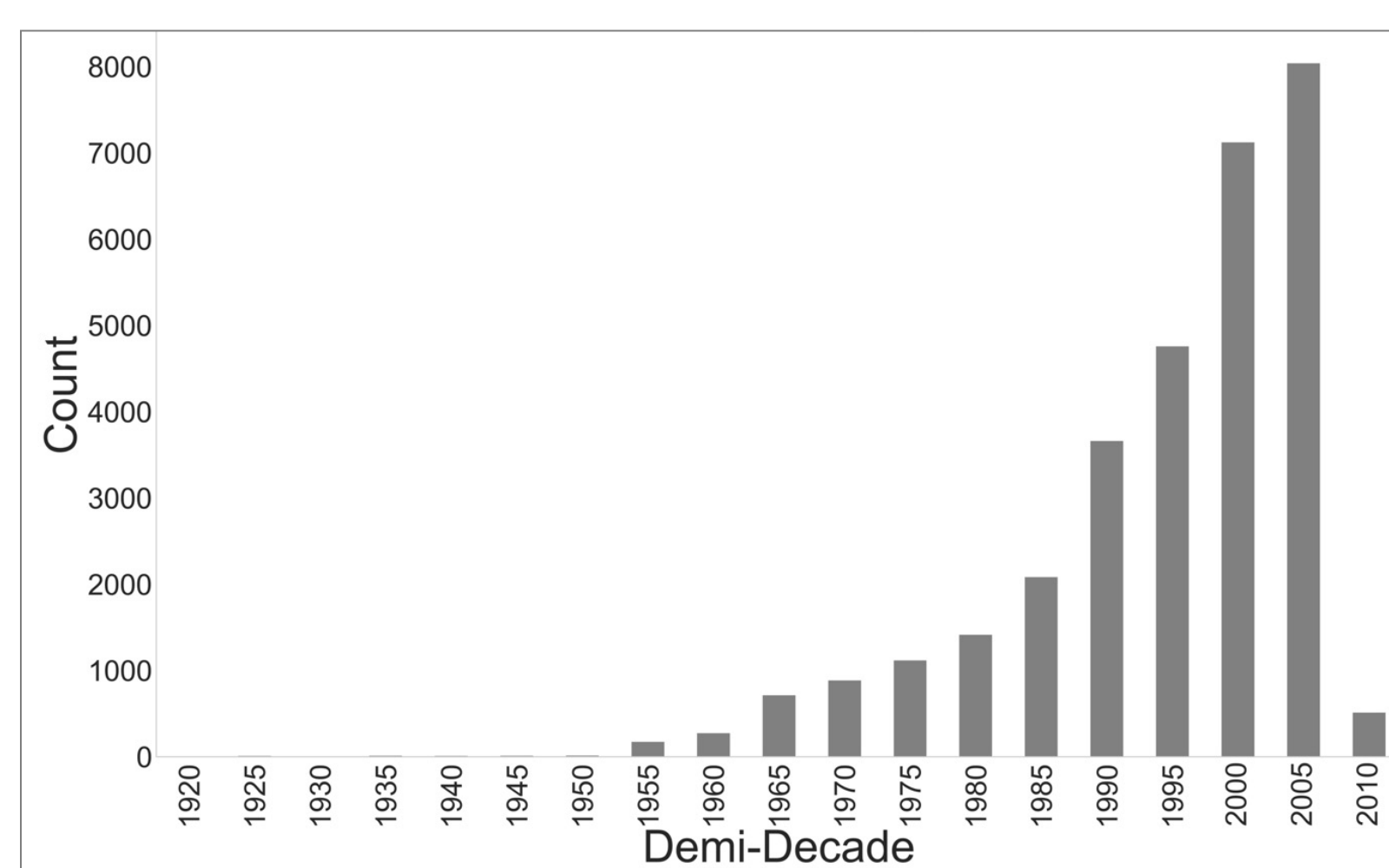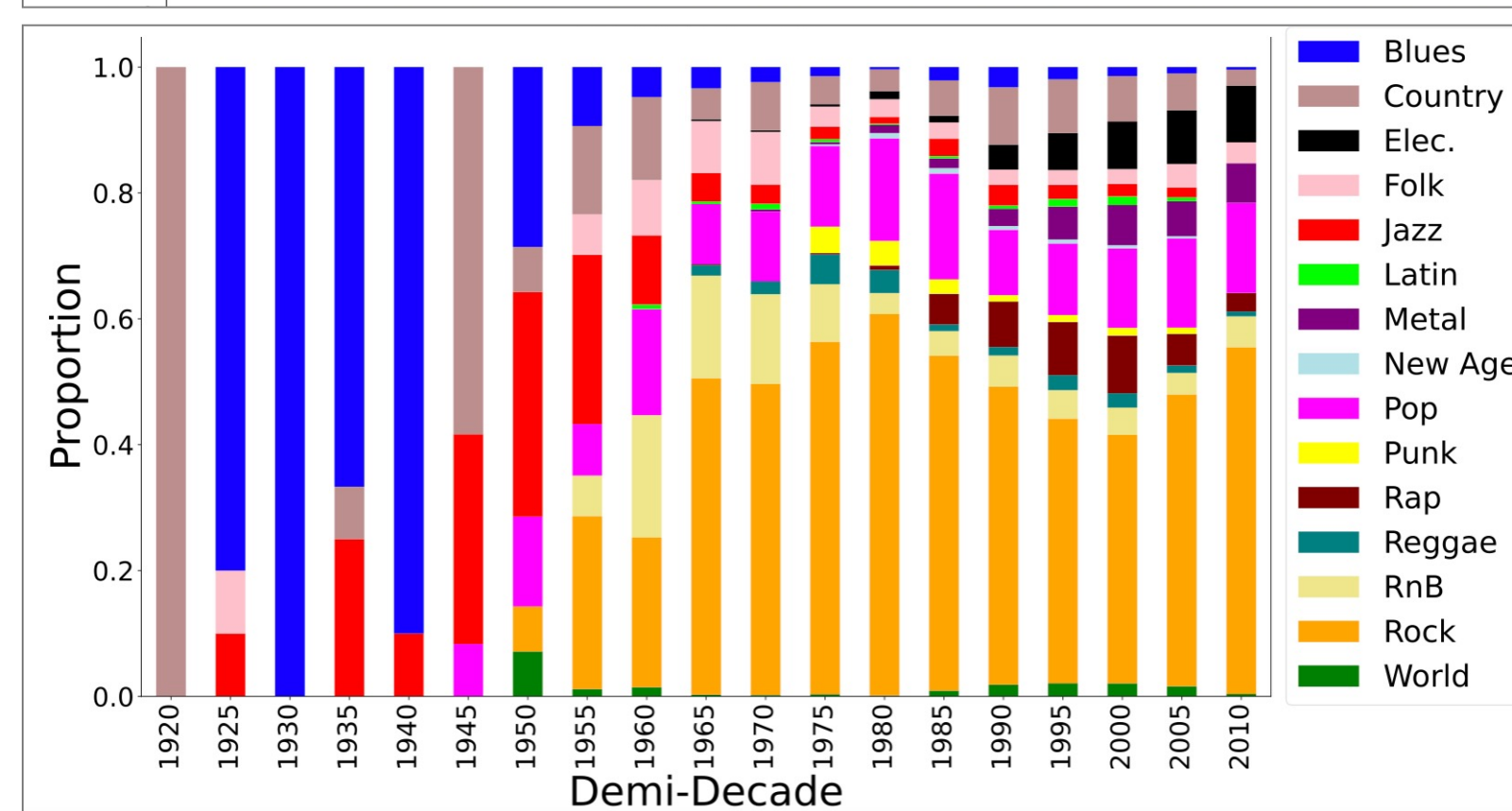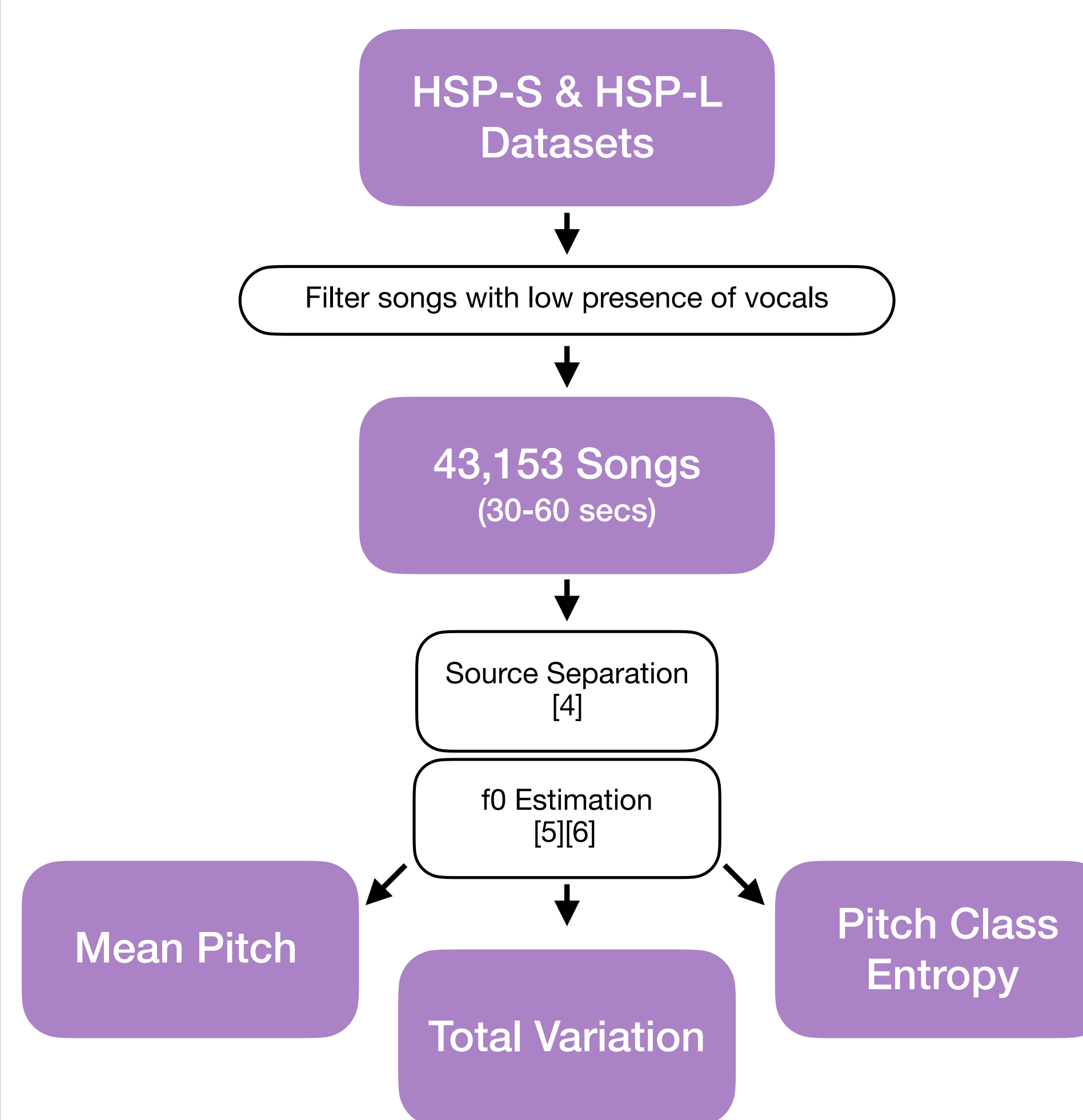
Figure 1. Chronological distribution of the dataset organized in 5-year demi-decade bins. Bottom: Relative distribution of genres in the dataset.

Methods

HSP-S & HSP-L Datasets

↓

Filter songs with low presence of vocals

↓

43,153 Songs (30-60 secs)

↓

Source Separation [4]

f0 Estimation [5][6]

→ Mean Pitch

→ Total Variation

→ Pitch Class Entropy

Mean Pitch

Mean pitch (in cents) calculated as the mean of each f0 array.

Total Variation (TV)

The rate of pitch change [7]

$$\text{TV}(x) = \frac{1}{N}\sum_{i=1}^{N-1}|x_{i+1} - x_i|$$

For a given f0 contour x = (x1,...,xN)

Pitch Class Entropy

The degree of unpredictability of the set of vocal pitches. Calculated over the probability of occurrence of each pitch class in the vocal.

Results

1. Linear regression analyses examining mean pitch & year for 8 musical genres.
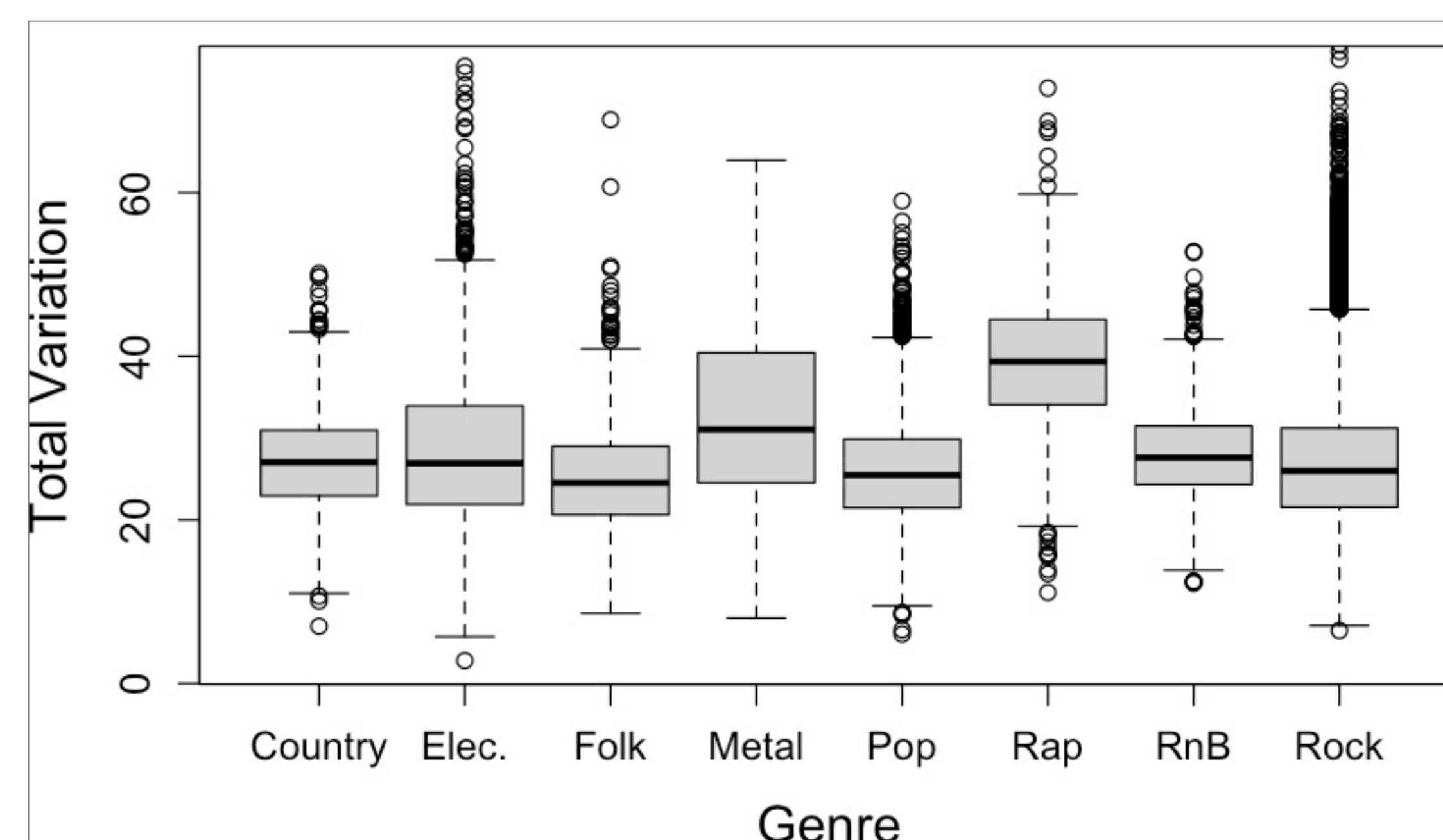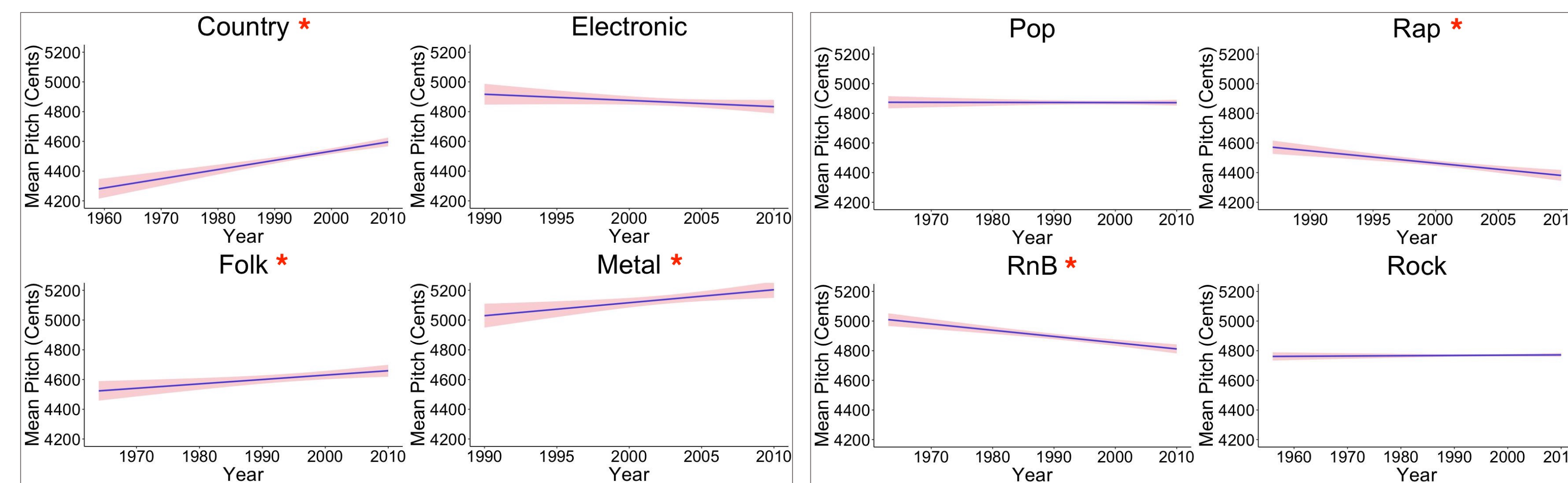
2. TV in each genre. Rap had a higher TV than the other genres.
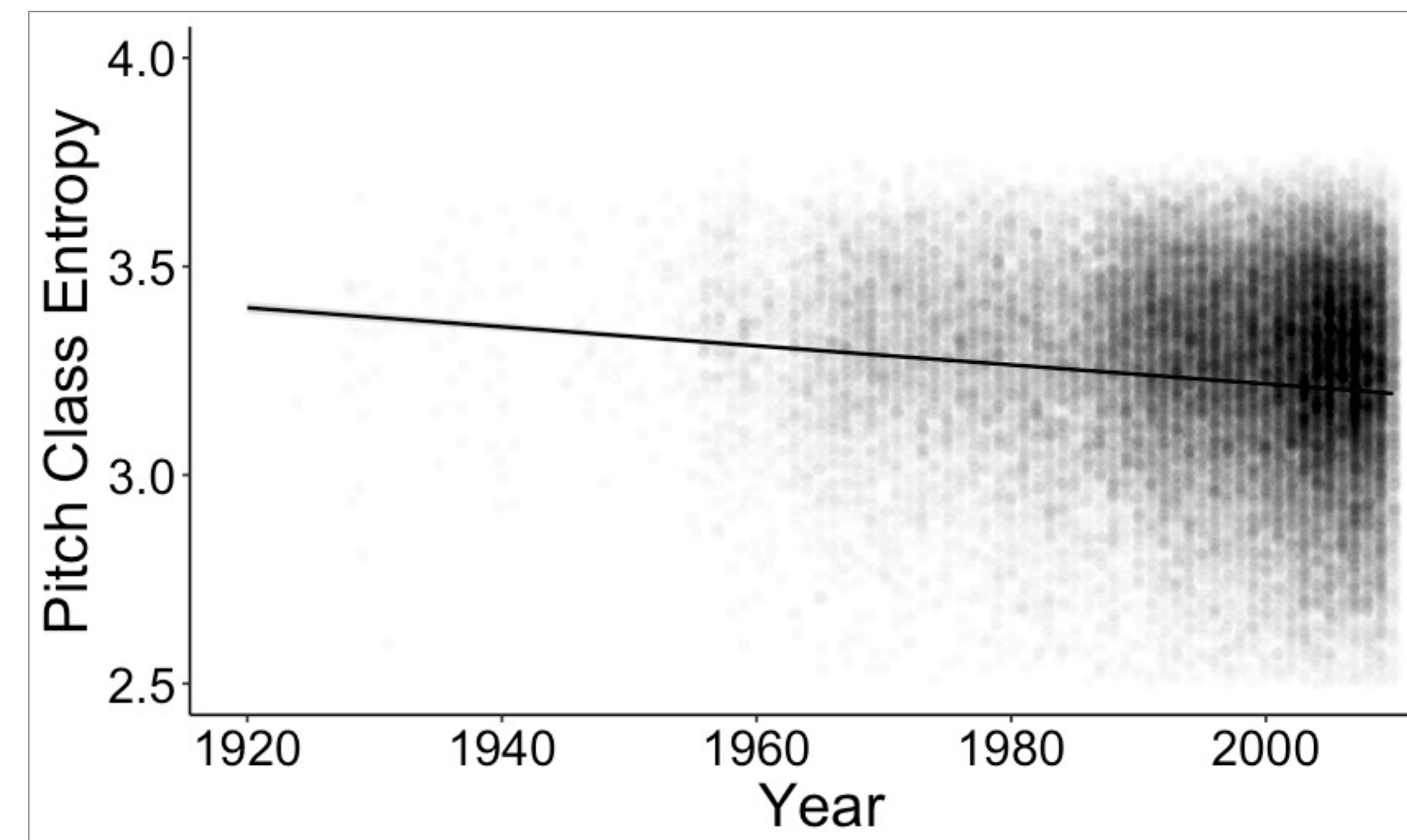
3. Pitch Class Entropy as a function of year. Each dot represents a song.

References

[1] M. Votter, M. Mayerl, G. Specht, and E. Zangerle, "Novel datasets for evaluating song popularity prediction tasks," IEEE International Symposium on Multimedia, 2021. [2] A. Schindler, R. Mayer, and A. Rauber, "Facilitating comprehensive benchmarking experiments on the million song dataset," ISMIR, 2012. [3] H. Schreiber, "Improving genre annotations for the million song dataset," ISMIR, 2015. [4] S. Rouard, F. Massa, and A. Défossez, "Hybrid transformers for music source separation," CoRR. [5] M. Mauch and S. Dixon, "Pyin: A fundamental frequency estimator using probabilistic threshold distributions," IEEE ICASSP, 2014. [6] B. McFee et al., "librosa/librosa: 0.8.1rc2," May 2021. [7] M. Panteli, R. Bittner, J. P. Bello, and S. Dixon, "Towards the characterization of singing styles in world music," EEE ICASSP, 2017.