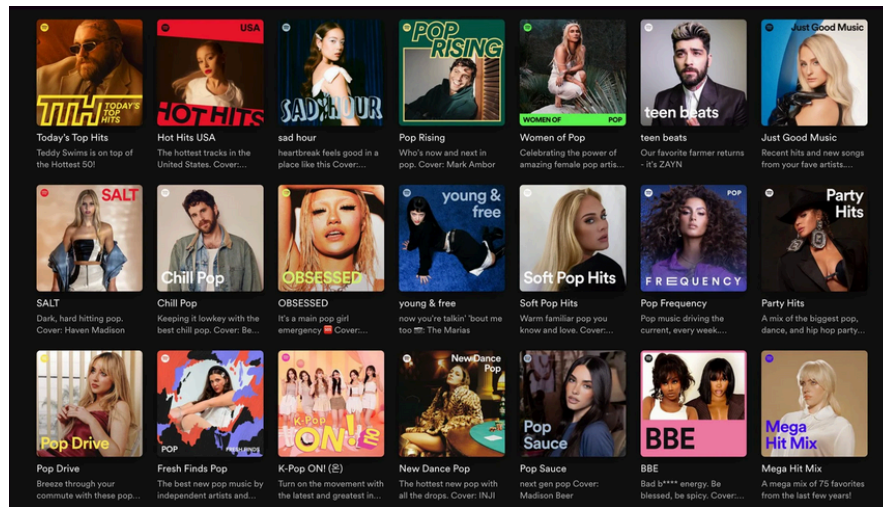# The Invisible Playlist Problem:

Why Spotify's Editorial Sound Doesn't Represent Real Curation

---

Introduction

Music streaming platforms have reshaped not only how we listen to music, but how music is organized and made visible. At the center of this ecosystem sits the playlist — the primary interface through which listeners encounter new artists, moods, and genres. For most people, playlists are personal objects shaped by memory, emotion, and routine. For a small fraction, however, playlists become public-facing products optimized for reach and discoverability.

For aspiring curators, this raises an uncomfortable question: if playlists are everywhere, why do only a handful gain visibility?
Most large-scale analyses focus almost exclusively on highly popular or editorial playlists — those already amplified by platforms. This creates a distorted picture of what "successful" curation looks like, while obscuring a basic reality: most playlists are expressive and never meant to scale.

This project asks a question that matters directly to anyone trying to break into the music curation scene:

> *Do the playlists that rise to the top reflect how people actually curate music — or are they selected for a narrow, platform-friendly template?*

By comparing playlist structure, title language, emotional signals, and audio features across popularity levels, this analysis shows that as playlists gain visibility, expressive diversity narrows — helping explain why distinctive playlists struggle to surface, and the tradeoffs curators face in today's editorial landscape.

How Does Playlist Structure Change as Playlists Gain Visibility?

Before looking at sound or language, it's useful to ask a basic question: do popular playlists even look like personal ones? We looked at how many songs appear in playlists at different levels of popularity, comparing personal playlists with those that attract large followings.
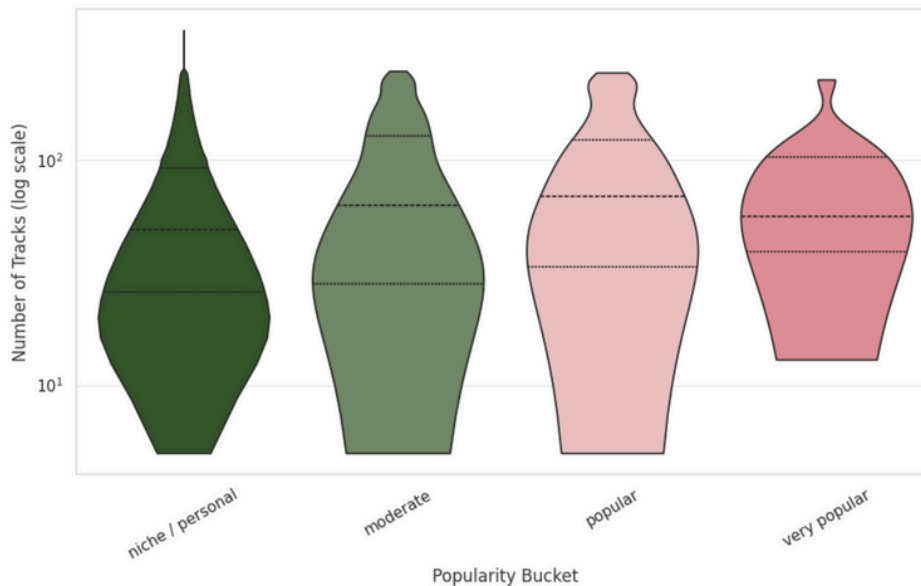

Figure 1: Playlist length distribution

Personal playlists tend to be much longer and uneven, often growing past 100 songs as people add tracks over time. Popular playlists, on the other hand, stick to a tighter range of about 20–50 songs. This shows that personal playlists act more like personal collections, while popular ones are trimmed down and shaped to look polished and easy to consume.

> *Once playlists are structurally constrained, how do curators change the way they present them?*

How Do Playlist Titles Shift from Personal Expression to Public Signaling?

2.1 Do Titles Become More Standardized as Playlists Grow?

Titles are one of the few explicitly human-authored elements of playlists. They reveal whether a playlist is meant for oneself or for discovery by others. We looked at playlist titles to see how their length and wording change as playlists move from personal use to wider popularity.

Personal playlists usually have short, casual titles, though some can be unusually long or playful. As playlists gain more followers, their titles become more uniform and predictable, sticking to a narrower range of lengths and using more generic terms.
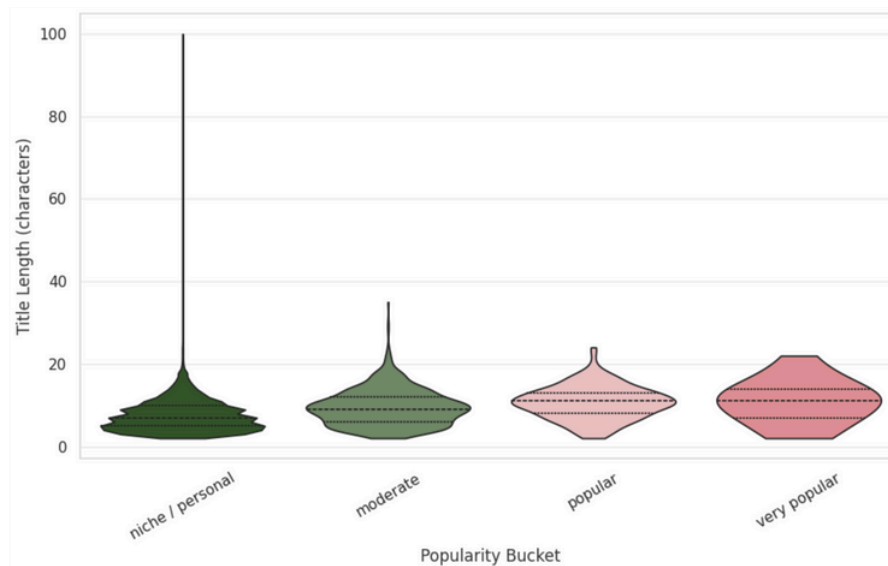
Figure 2: Title length by popularity

This shows that popular playlists are shaped less by personal expression and more by the need to be clear and searchable.

2.2 What Happens to Linguistic Uniqueness at Scale?

Length alone doesn't capture creativity. A long title can still be generic.
We measured lexical rarity — how uncommon title words are relative to the full dataset.
Personal playlists are the most creative — only about 6.6% use generic words, and you'll see quirky names like 4x4, Belters, or #ViNTAGE. Moderately popular ones stay fairly original too (8.4% generic). But once playlists get very popular, originality drops fast: around 40% of them use generic keywords. In other words, the bigger the audience, the more titles start sounding alike, even if they're longer.

> *If language becomes more neutral and standardized, does the music itself follow the same pattern?*

Figure 3: Spotify PR campaign emphasizing the individuality of personal playlist titles

2.3 Does Emotional Expression Disappear from Playlist Titles?

Emotion is central to how people use music privately. If it vanishes at scale, that signals a deeper transformation. We applied a transformer-based emotion classifier to playlist titles and aggregated results by popularity bucket.
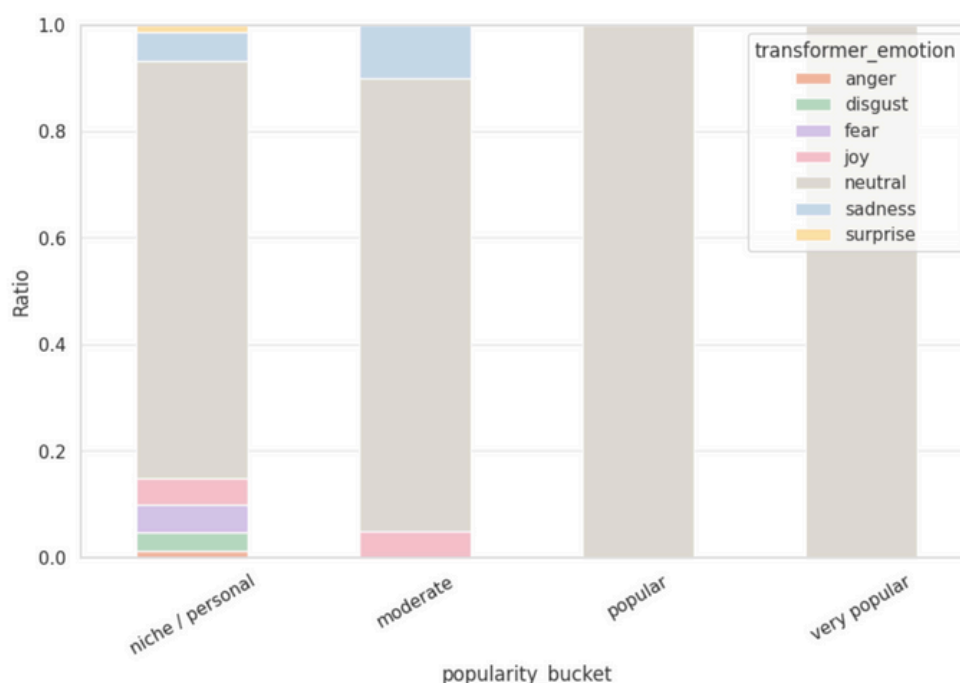

Figure 4: Transformer-Based Emotion Composition of Playlist Titles by Popularity

Personal and niche playlists show the widest emotional range — playful, sad, joyful, even quirky. But as playlists get more popular, those emotions fade fast. By the time you reach the very popular ones, almost all titles are neutral, with little emotional flavor. In short: the bigger the audience, the more titles shift from expressive to plain and functional.

How Does Playlist Sound Change as Popularity Increases?

Structure and language can be optimized intentionally. Sound is harder to fake. If convergence appears here too, it suggests deeper selection pressures.
We extracted audio features related to rhythm, energy, brightness, and timbre from track previews and compared their distributions across popularity levels.

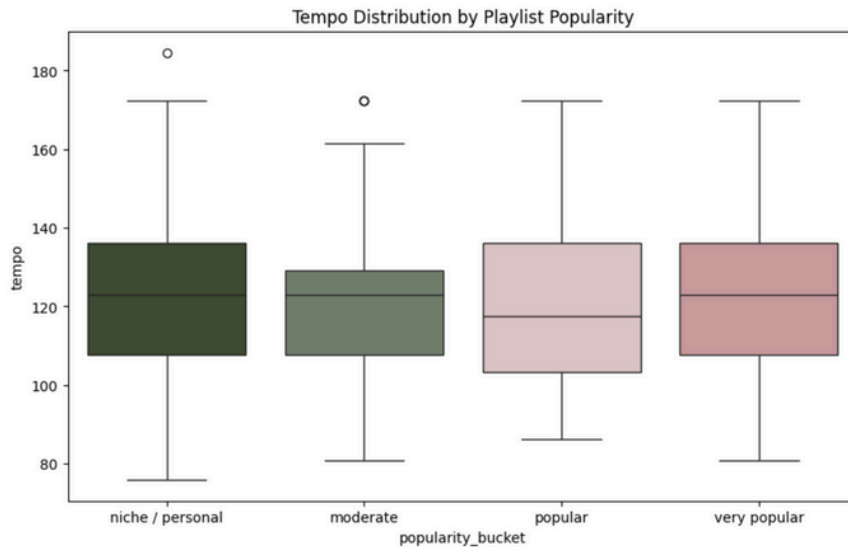## 3.1 Do Popular Playlists Avoid Rhythmic Extremes?

Figure 5. Tempo distribution by playlist popularity

Personal playlists span wide tempo ranges; popular playlists cluster narrowly around mid-tempo (100–120 BPM).

3.2 Does Energy and Brightness Converge with Popularity?

We measured RMS energy (how strong or loud a song feels) and spectral centroid (how bright it sounds, bass vs. treble). Personal playlists swing between soft and loud, dark and bright. Popular playlists avoid those extremes, keeping intensity and brightness steady for a smoother listen.
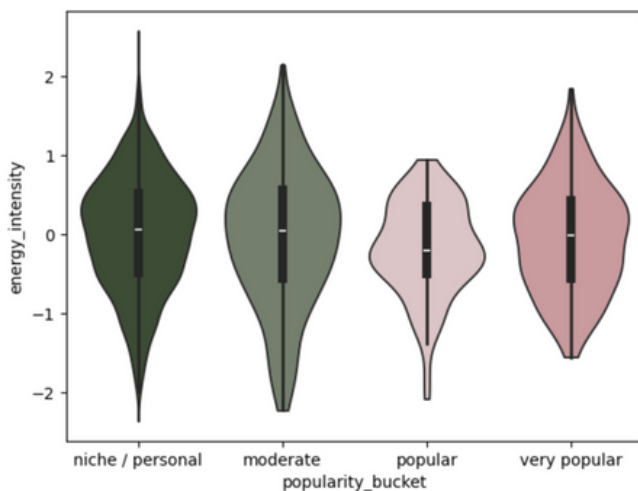

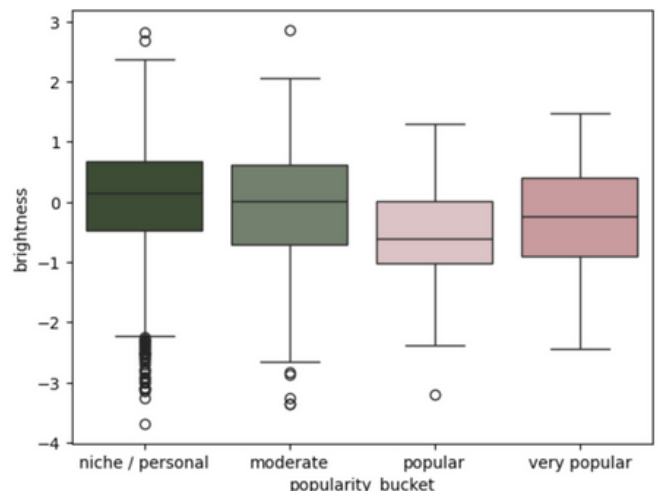**Figure 6(a).** RMS energy distribution by popularity bucket

Figure 6(b). Spectral centroid by playlist popularity

Is Variability the Hidden Factor That Separates Personal from Popular Playlists?

4.1 Do Popular Playlists Allow Less Sonic Variety Within a Playlist?

We measured how much variety exists within a playlist by looking at tempo (speed), RMS energy (loudness), spectral centroid (brightness, bass vs. treble), and timbre (tone color, captured with MFCCs). This told us whether playlists allow wide swings in sound or keep things tightly controlled.
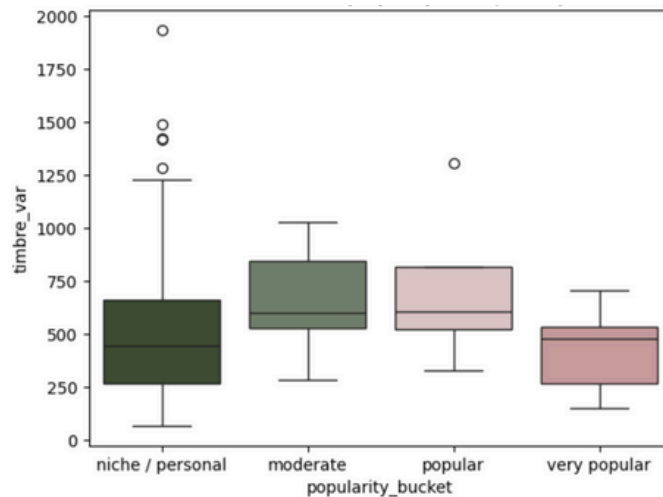

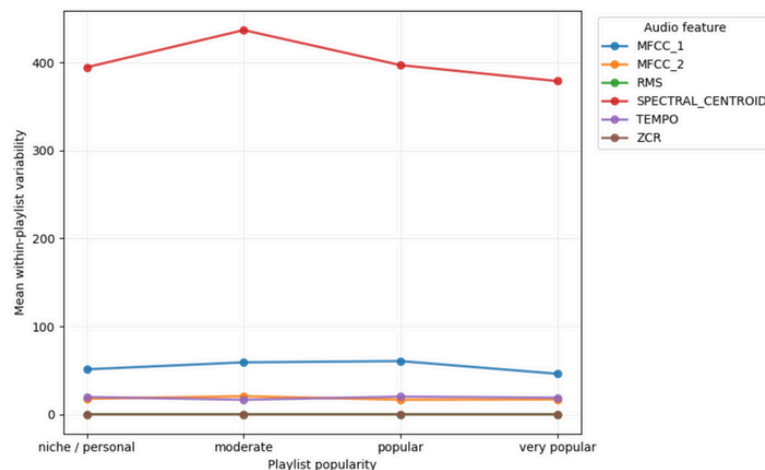**Figure 7(a).** MFCC variance by popularity bucket

Figure 7(b). Sonic Variability Decreases with Playlist Popularity figure.

Niche playlists show the most freedom, with big shifts in timbre and loudness. Very popular playlists are much more uniform, sticking to a consistent sound. Tempo is the one feature that stays more flexible even in mainstream playlists, while other qualities remain steady. In short, popularity tends to favour cohesion, but a little tempo variety helps keep listeners engaged.

4.2 Are Popular Playlists More Sonically Cohesive Overall?

We used Principal Component Analysis (PCA) to study how songs in a playlist cluster together based on features like rhythm, loudness (RMS energy), brightness (spectral centroid), and timbre (tone color). For each playlist, we measured the average distance of songs from the "center point" of that playlist in feature space — smaller distances mean tighter cohesion.
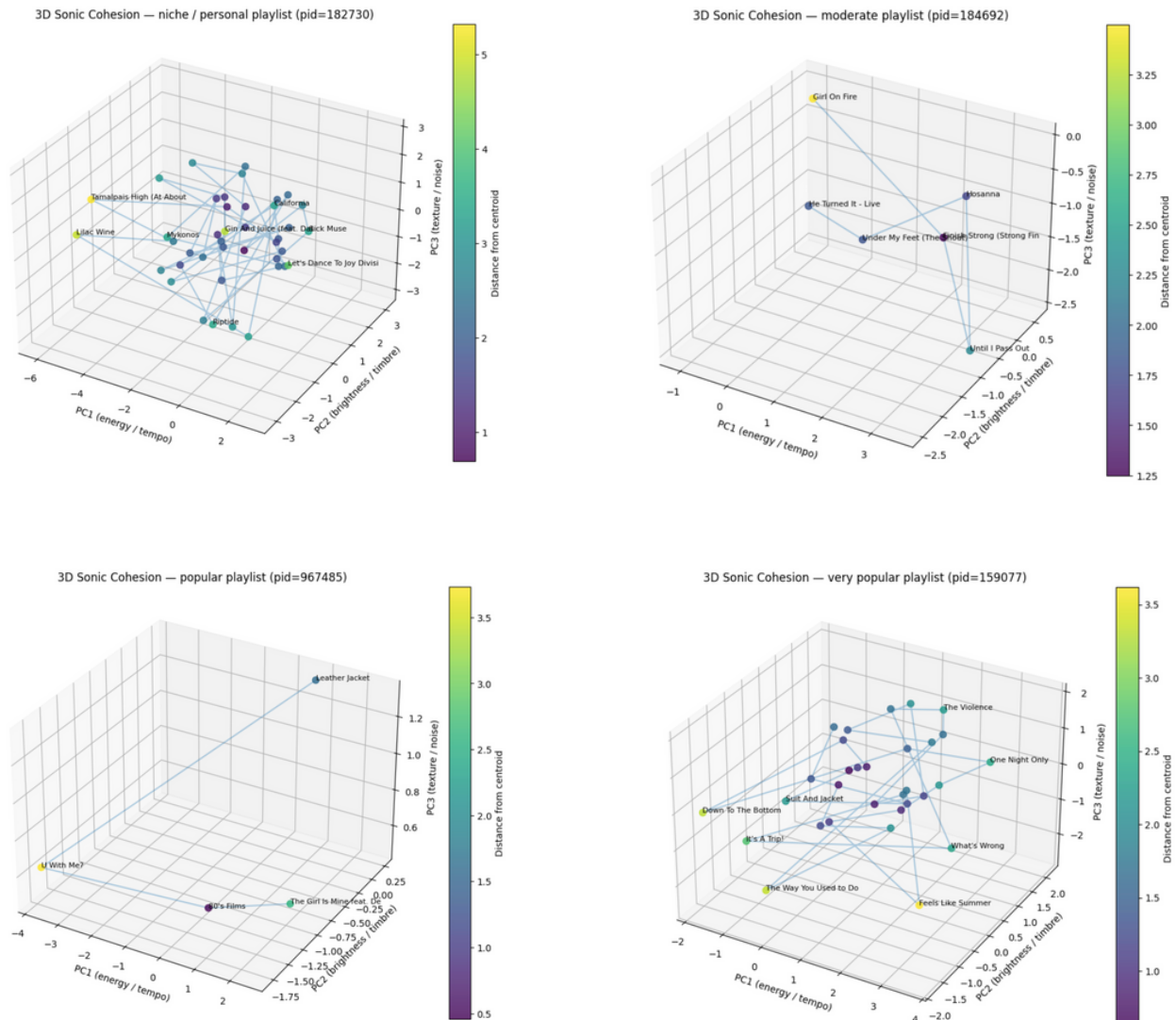

Figure 8. PCA projections of audio features for representative playlists across popularity buckets.

Very popular playlists had the smallest distances (around 1.76), showing they are tightly packed and sonically uniform. Personal playlists were much more spread out (around 1.82), reflecting looser, exploratory sequencing. This pattern held across multiple samples: popularity consistently narrows variety, creating compact clusters where rhythm, energy, brightness, and timbre all converge. In short, mainstream playlists sound more polished and cohesive, while personal ones allow greater diversity.

pid	popularity_bucket	mean_dist
159077	very popular	1.739744
184692	moderate	2.05772
967485	popular	2.087068
182730	niche / personal	2.478965

Table 1. Mean distance from centroid for playlists in Fig. 8 as a measure of playlist cohesion.

What Changes as Playlists Become Popular — and Why?

Personal playlists work like personal diaries — messy, playful, and full of sudden shifts in mood, tempo, and tone. Their titles are quirky, and the music spans wide ranges of loudness, brightness, and timbre, showing that variety is part of the charm, not a flaw.
As playlists attract more followers, this freedom shrinks. Popular ones lean toward smooth cohesion, neutral titles, and predictable emotional tone. Analyses like PCA confirm that very popular playlists cluster tightly in sound, with less spread across tempo, energy, brightness, and timbre. Emotional cues in titles also fade, replaced by plain labels.
This isn't accidental: platform algorithms and editorial norms reward playlists that minimize surprise and maximize clarity. Cohesion helps them travel farther, but it comes at the cost of expressive risk. What changes with popularity isn't just audience size — it's the range of creativity that remains visible.

What Does This Mean for Platforms, Curators, and Artists?

The data shows that today's playlist system leans toward predictability over variety, and that choice has real effects.
For **platforms**, adding signals that reward diversity — like variation in tempo or tone — could balance things out without making playlists harder to use.
For **curators**, popularity doesn't automatically mean quality. Playlists that feel smooth and cohesive spread more easily, but it's the quirky variation that makes curation unique. Knowing when to polish and when to keep things personal becomes a creative strategy.
For **artists**, the takeaway is clear: mainstream playlists favor consistent sound, while personal ones leave room for experimentation. Building music for both — polished tracks for scale and diverse ones for niche listeners — fits how visibility actually works.

What Gets Lost When Playlists Scale — and Who Decides?

Personal playlists are colorful and varied — they mix moods, tempos, and playful titles, showing music as a personal expression. Popular playlists, by contrast, are smoother, more neutral, and built for scale. As they move into public view, variety in structure, language, and sound steadily declines.

This doesn't mean expressive playlists lack worth. It means visibility is shaped by platform rules that favor a narrow template. Seeing this gap helps curators, artists, and listeners decide when to follow the mold, when to experiment, and what kind of musical culture they want playlists to reflect.

Data Sources and References

Dataset
- The analysis relied on the Spotify Million Playlist Dataset (MPD), but because Spotify's Web API audio-feature endpoints were deprecated, low-level descriptors like tempo and energy could no longer be retrieved directly. To address this, we turned to YouTube track previews as a substitute source.

Audio Feature Extraction
- McFee, B., Raffel, C., Liang, D., Ellis, D. P. W., McVicar, M., Battenberg, E., & Nieto, O. (2015). librosa: Audio and music signal analysis in Python. Proceedings of the 14th Python in Science Conference.

Emotion Classification
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., et al. (2020). Transformers: State-of-the-art natural language processing. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. Association for Computational Linguistics.

Methodological References
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065).

What More to Do With One More Week

- **Temporal Playlist Growth Analysis**
  - Use last_modified_date metadata in MPD to track how playlists evolve over time.
  - Compare whether playlists that eventually become popular show early signs of cohesion or neutrality in titles.
- **Sequence Modeling**
  - Apply models such as Markov chains or RNNs to playlist orderings.
  - Test if popular playlists show more predictable sequencing (e.g., tempo arcs, energy flow) compared to personal/niche ones.
- **Cross-Bucket Track Overlap**
  - Measure frequency of tracks from niche playlists appearing in popular ones.
  - Assess whether popular playlists recycle personal curation or bypass it entirely.