What Drives Music Popularity?

Behavioral Discovery & Underrated Track Analysis

ANALYTIC REPORT HERE

Lucy Roh
Behavioral Data Analytics · UX Research
click here : github profile

Research Question

We want to understand

Research Question

Why do some songs succeed while others fail?

1. Can we predict popularity using artist history?
2. When predictions fail, what explains the gap?

Dataset Overview

Overview & EDA

Dataset Overview

Dataset

Tracks enriched with:

- Spotify / Apple popularity
- artist popularity & followers
- release history (momentum)
- audio features (tempo, energy, brightness, etc.)

~6–7k tracks

Goal

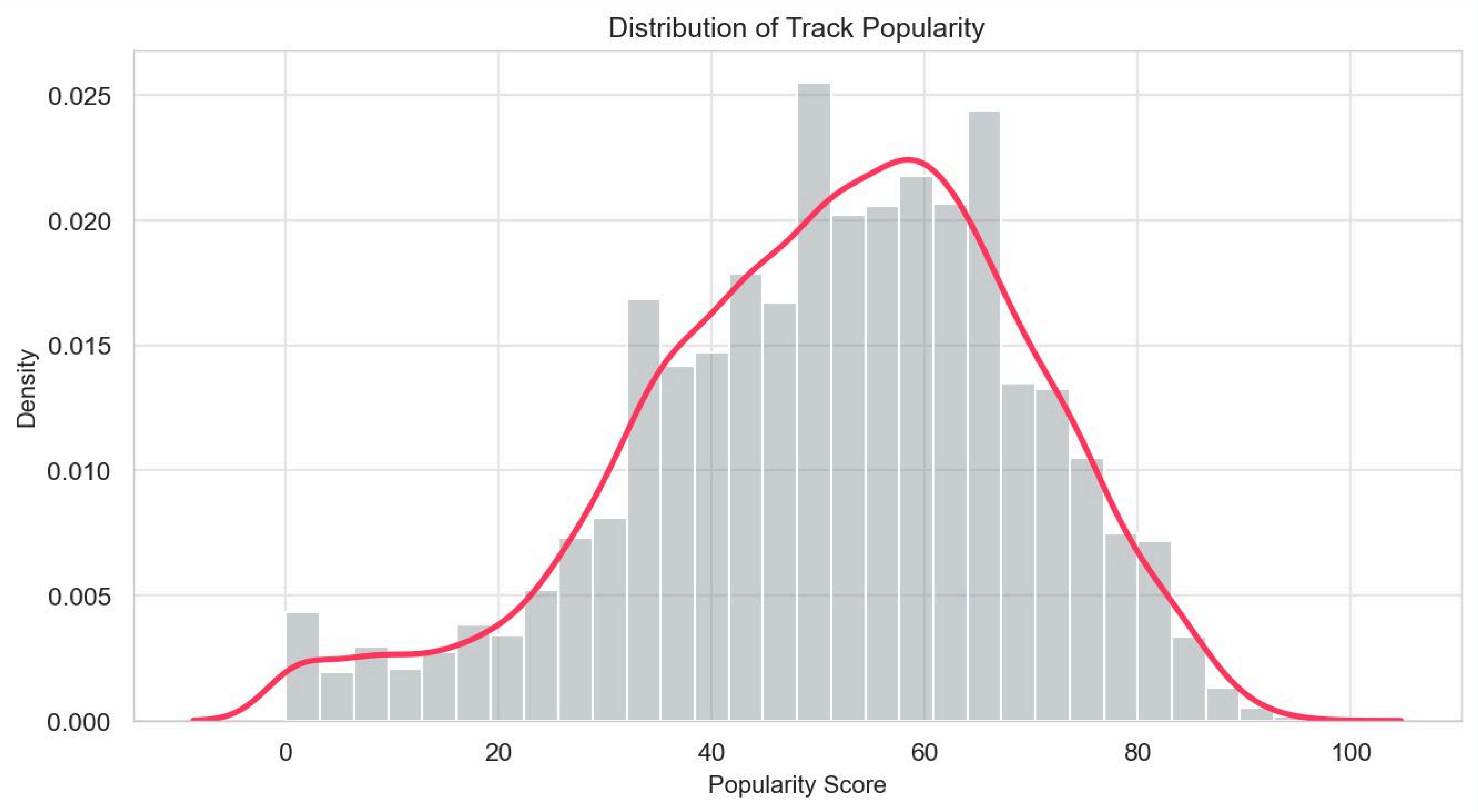
Model structural + behavioral signals together

Popularity Distribution (EDA)

Popularity Distribution

Observations:

- most songs cluster mid-range
- long tails of hits and failures
- high variance



Implications

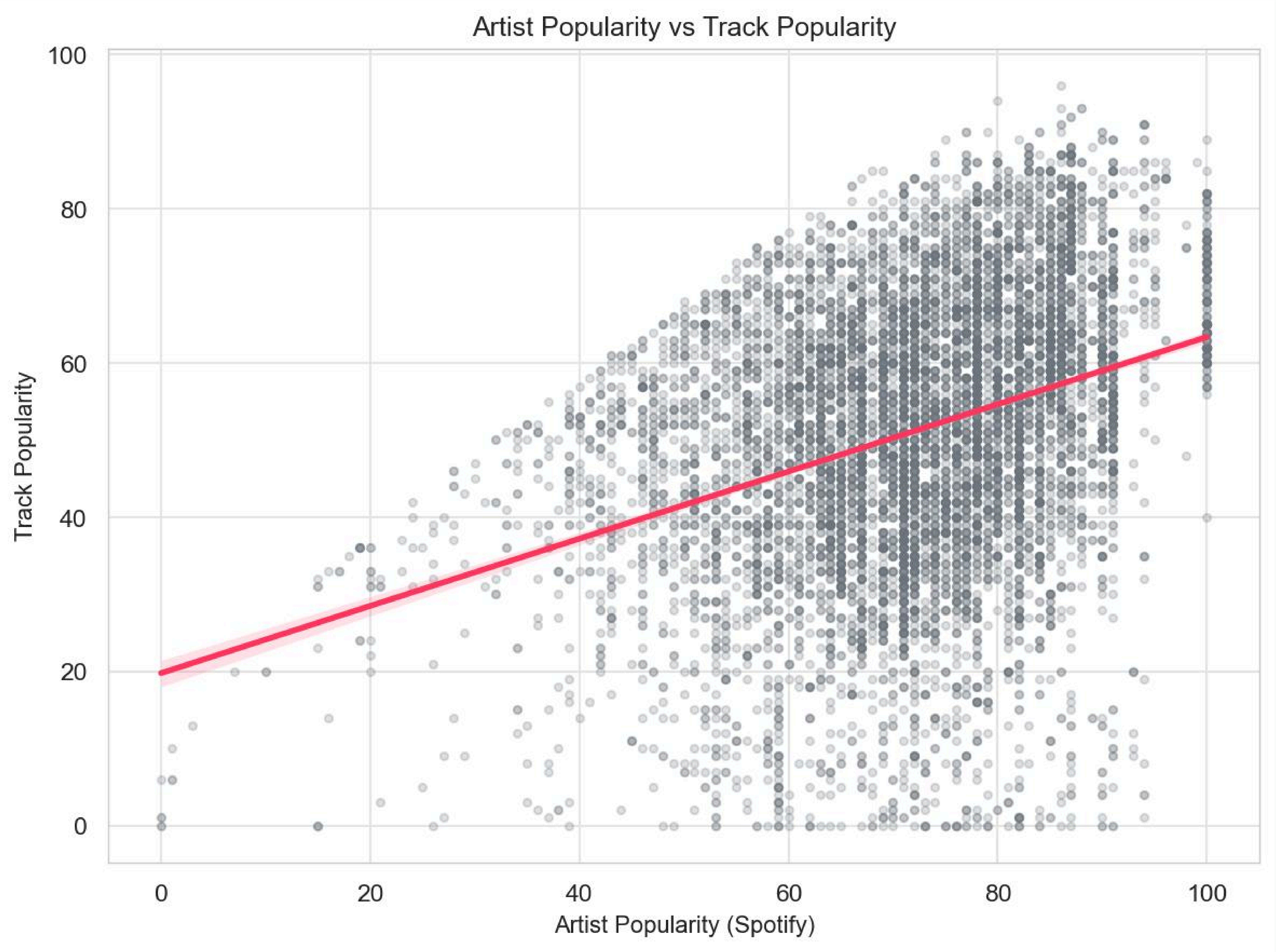
Success is noisy
It is **unlikely that there are pure rules.**

Artist Strength Matters (EDA)

Artist Strength Matters (EDA)

We observe:

- higher artist popularity → higher track success
- follower count correlates with visibility
- past success predicts future success



Interpretation

Cumulative advantage ("rich get richer") exists

Modeling

Baseline model

Model Training & Evaluation Setup

Goal:

Estimate expected success before release

Data split

• Train: 5,479 tracks

• Test: 1,370 tracks (future holdout)

Method

• Linear regression (OLS)

• Pre-release features only

• No leakage from future popularity

Evaluation

• Metrics reported on test set only

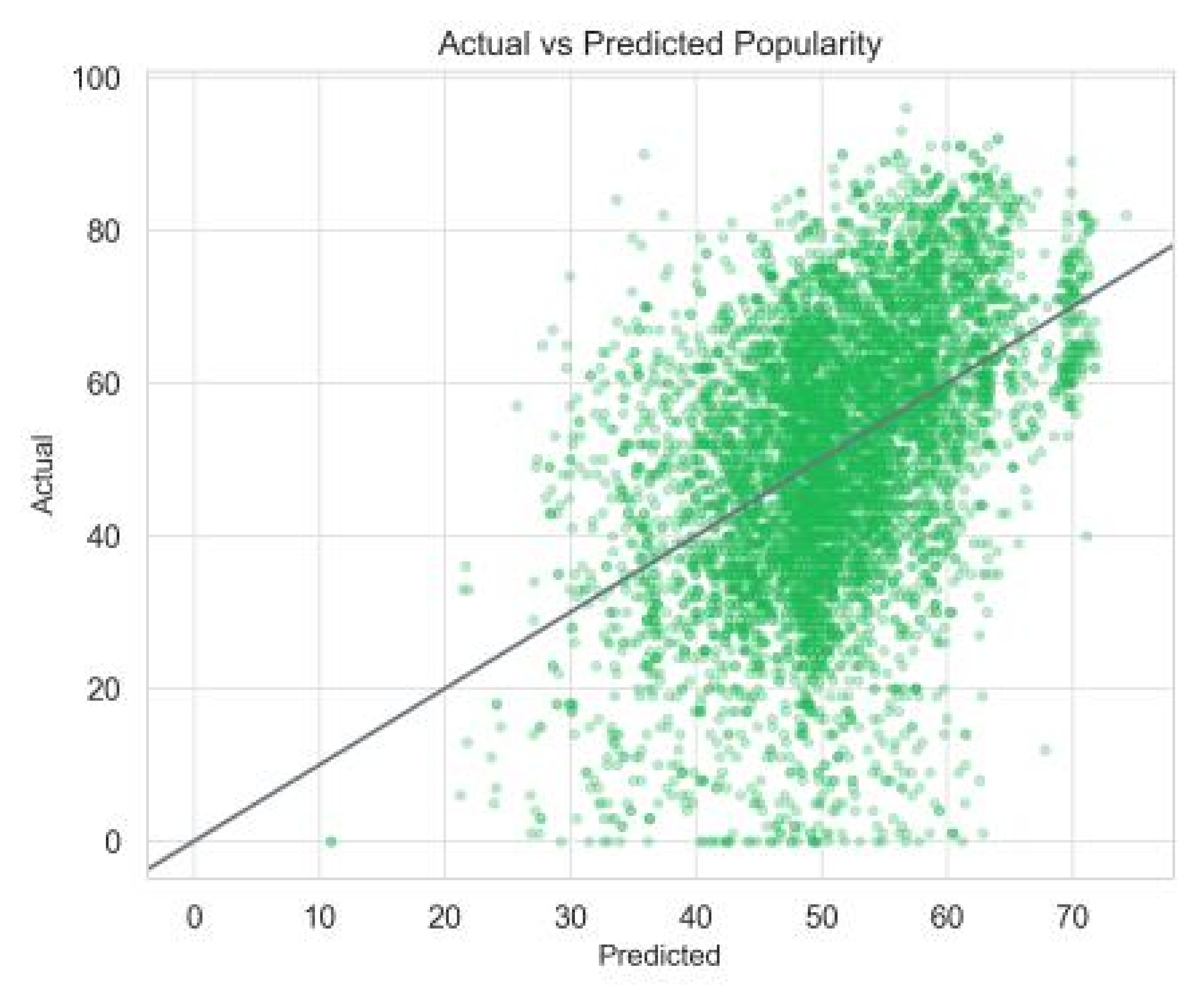
Baseline Model

Expected Popularity

We estimate **each track's expected popularity** using:

- artist reputation
- follower count
- previous release performance (momentum)
- historical averages
- track duration

Goal:
Estimate **expected success before release**

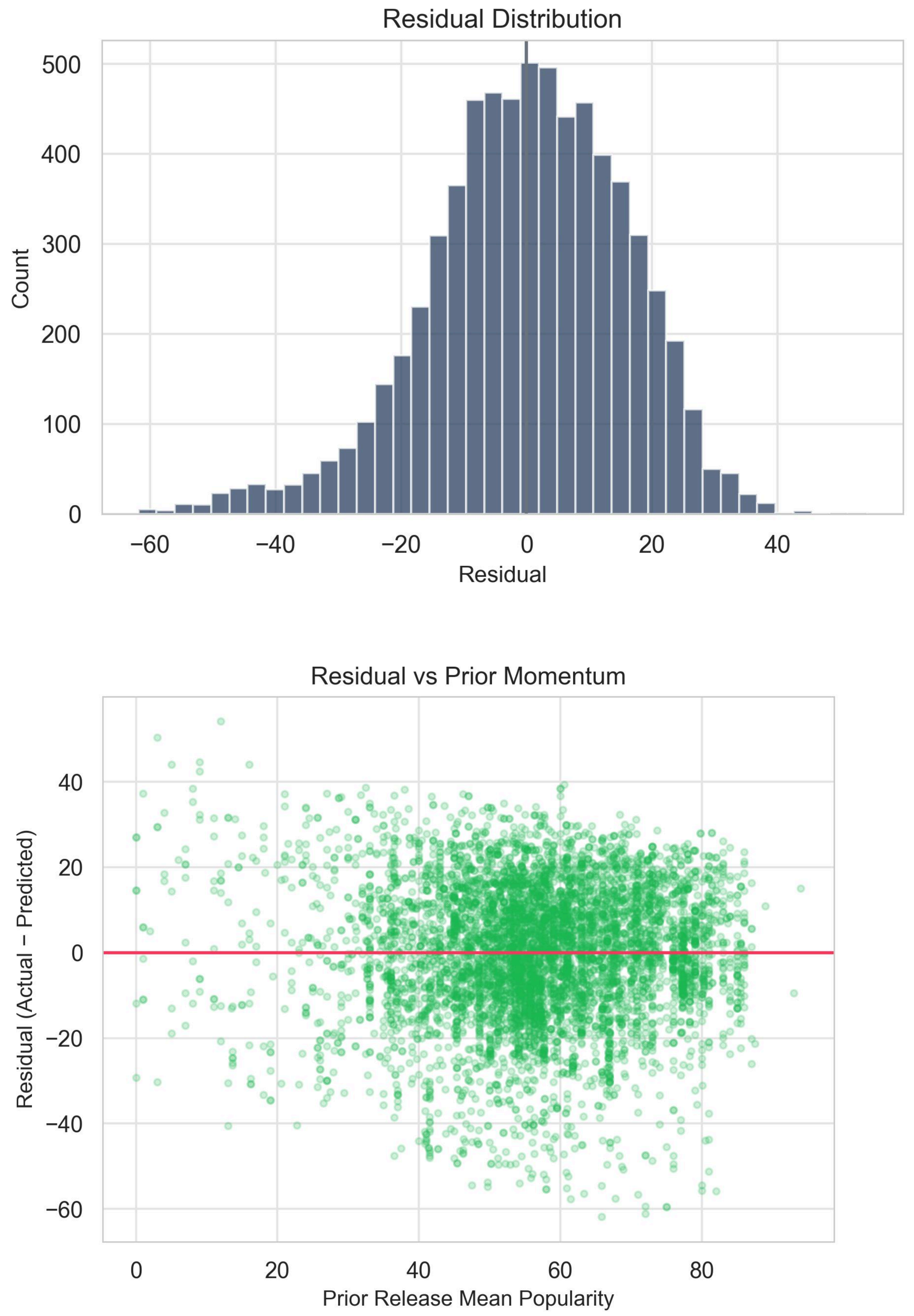


Model Performance

Model Captures Trends, Not Everything

Findings:

- captures general artist advantage
- explains baseline behavior
- but large unexplained variance remains



Key Observation

Many songs deviate strongly from predictions

Model Performance

How Much Can We Actually Predict?

Model	R^2	MAE	
Audio only	**0.003**	**14.17**	Audio has almost zero predictive power
Artist only	**0.104**	**13.34**	Artist reputation explains baseline popularity
Artist + prev song momentum	0.175	12.28	Short-term carryover adds strong improvement
Artist + prior mean momentum	0.197	12.07	Catalog reputation stronger than one-hit momentum
Artist + both momentum	**0.198**	**12.07**	Long-term dominates; short-term mostly absorbed

Test Performance (n = 1,370)

- Audio features ≈ no predictive power
- Artist reputation explains ~10%
- Even best model explains only ~20%
- 80% remains unexplained

Residual Analysis

Why do we still see large prediction errors?

New Question

If reputation explains baseline success...

Why do we still see large prediction errors?

What causes:

- surprise hits
- hidden gems
- unexpected failures

We investigate the gap :

Residual = actual – predicted

Residual Analysis

Behavior Gap Signal

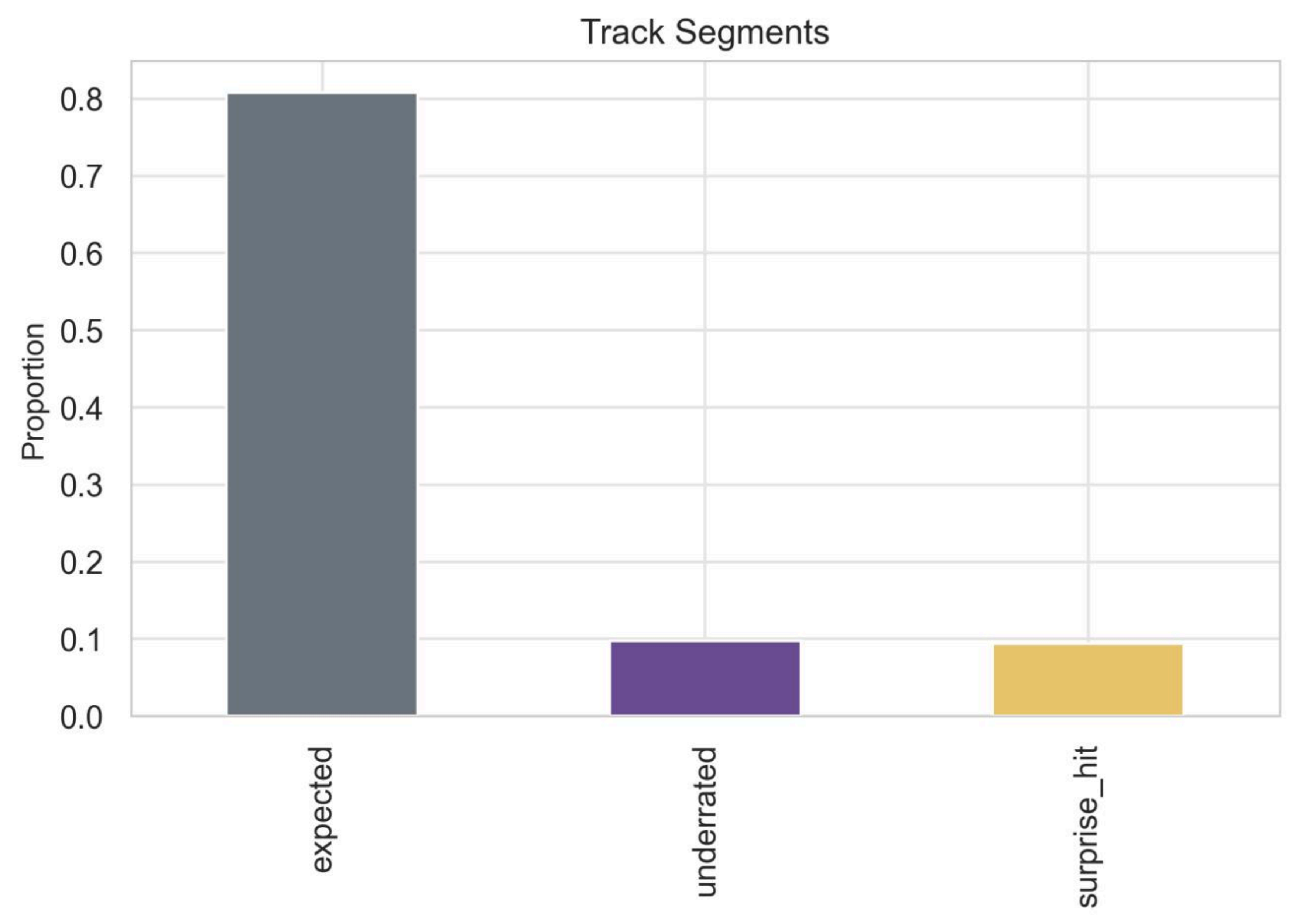
We classify tracks using percentiles:

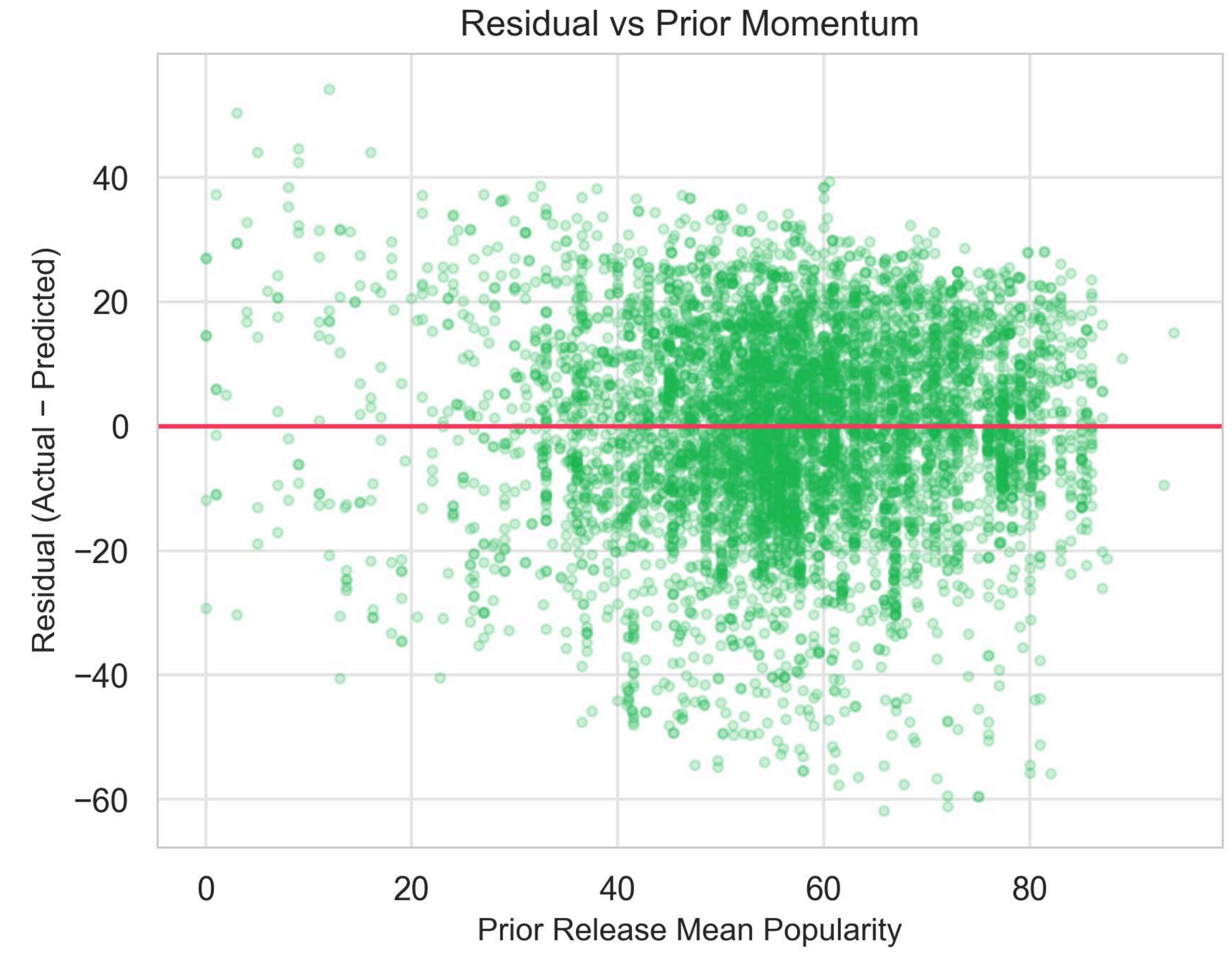
- Top 10% → Surprise Hits
- Middle 80% → Expected
- Bottom 10% → Underrated

→ **Approximately 20% of tracks behave unexpectedly.**

This represents meaningful opportunity:

- **overlooked content**
- **unexpected growth**
- **ranking inefficiencies**





Residual Analysis

We compute:

Residual = Actual – Predicted

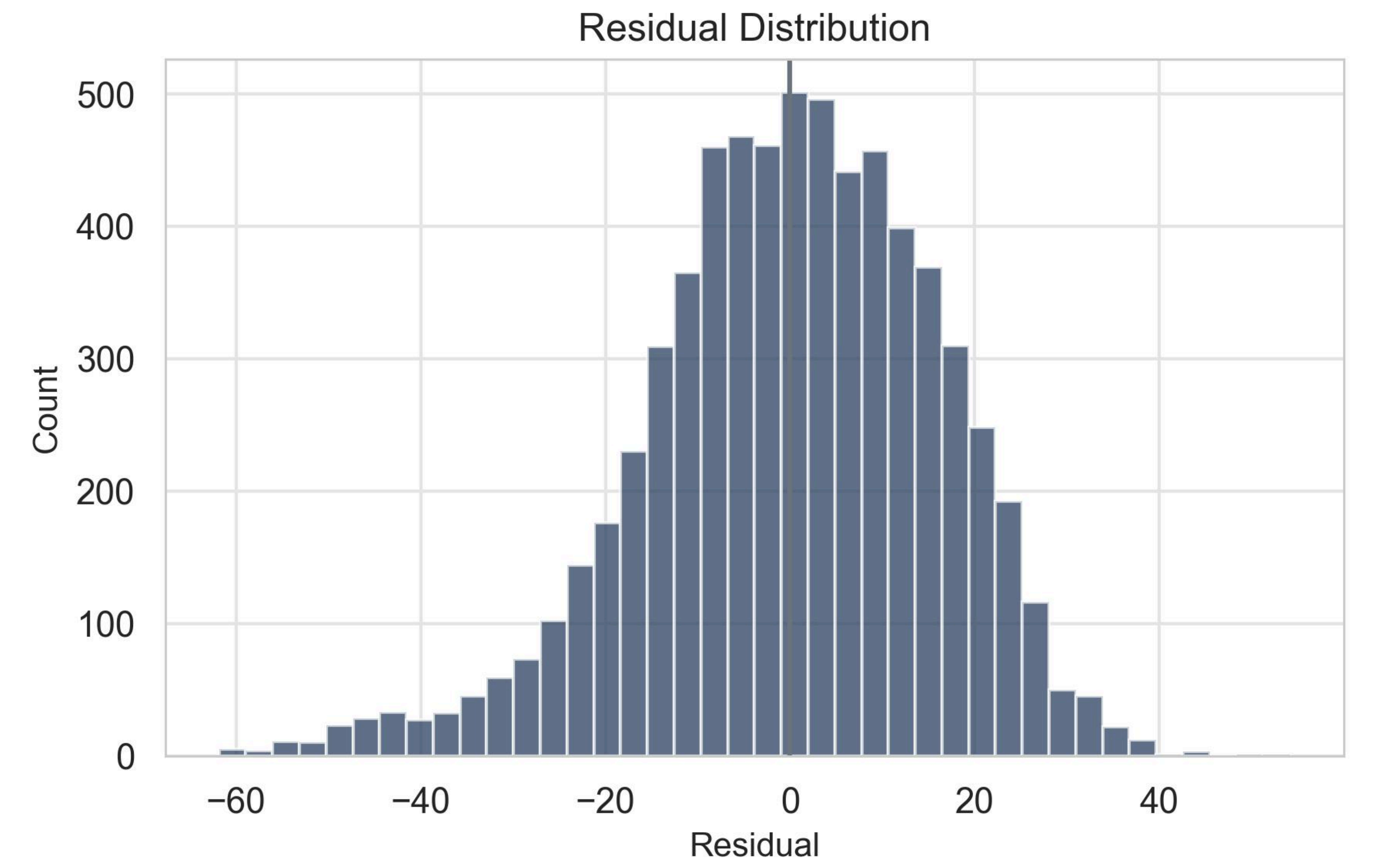
Interpretation:

- positive → surprise hit
- negative → underrated
- near zero → expected

Residuals isolate behavioral and exposure effects that modeling misses.

Observations:

Large deviations persist even for strong artists.
Success is not determined by reputation alone.



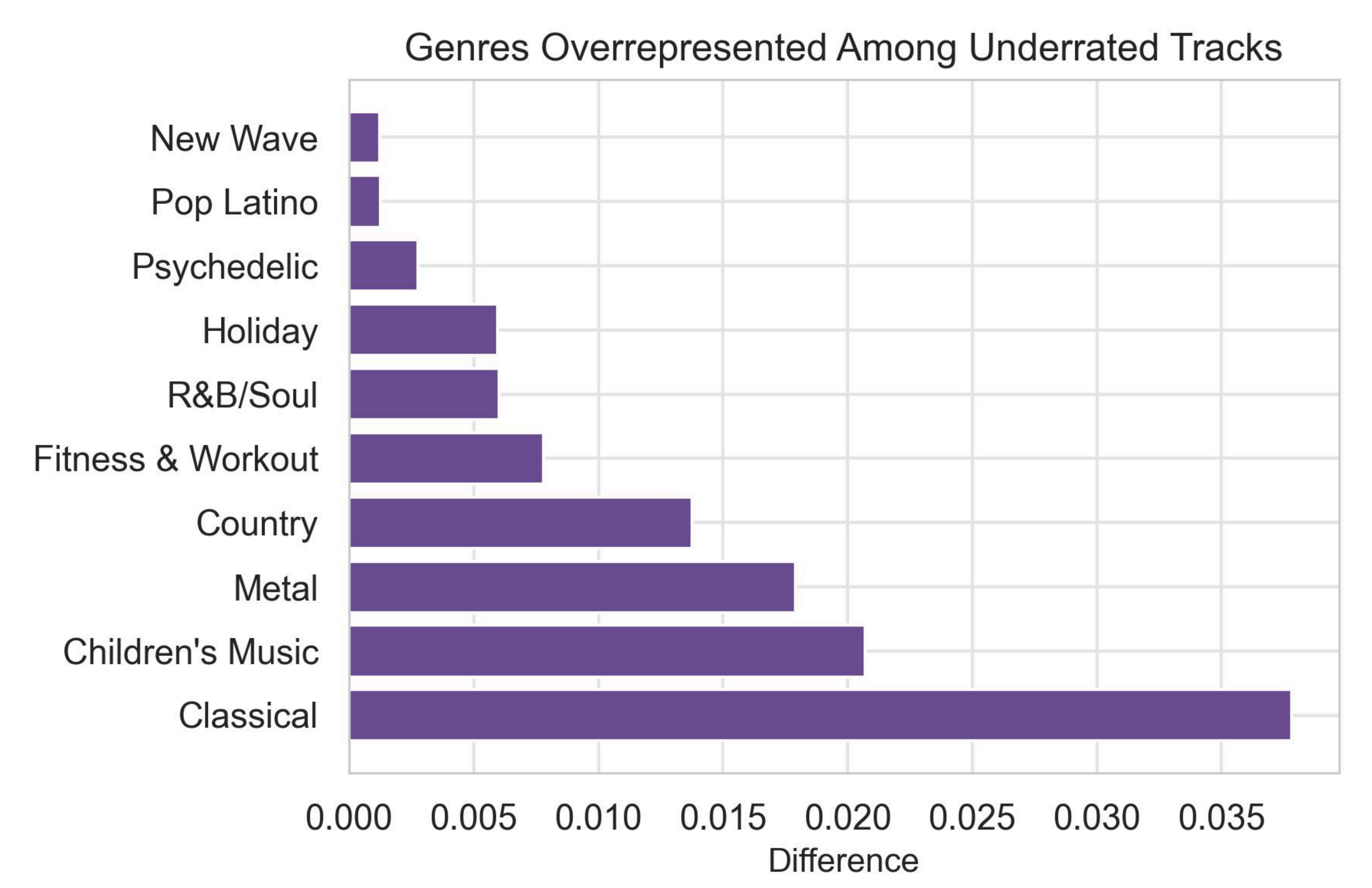
Underrated Tracks (Who Gets Overlooked?)

Underrated ≠ Weak Artists

Underrated tracks show similar:

- artist popularity
- follower counts
- historical momentum
- song duration

Yet they perform significantly worse than expected.



Key Difference

They are concentrated in niche genres:

- Classical, Children’s, Country, Metal

Interpretation

These tracks are likely underexposed, not lower quality;
indicating structural missed opportunities (songs predicted to fail but succeeded)

Surprise Hits (Who Gets Amplified?)

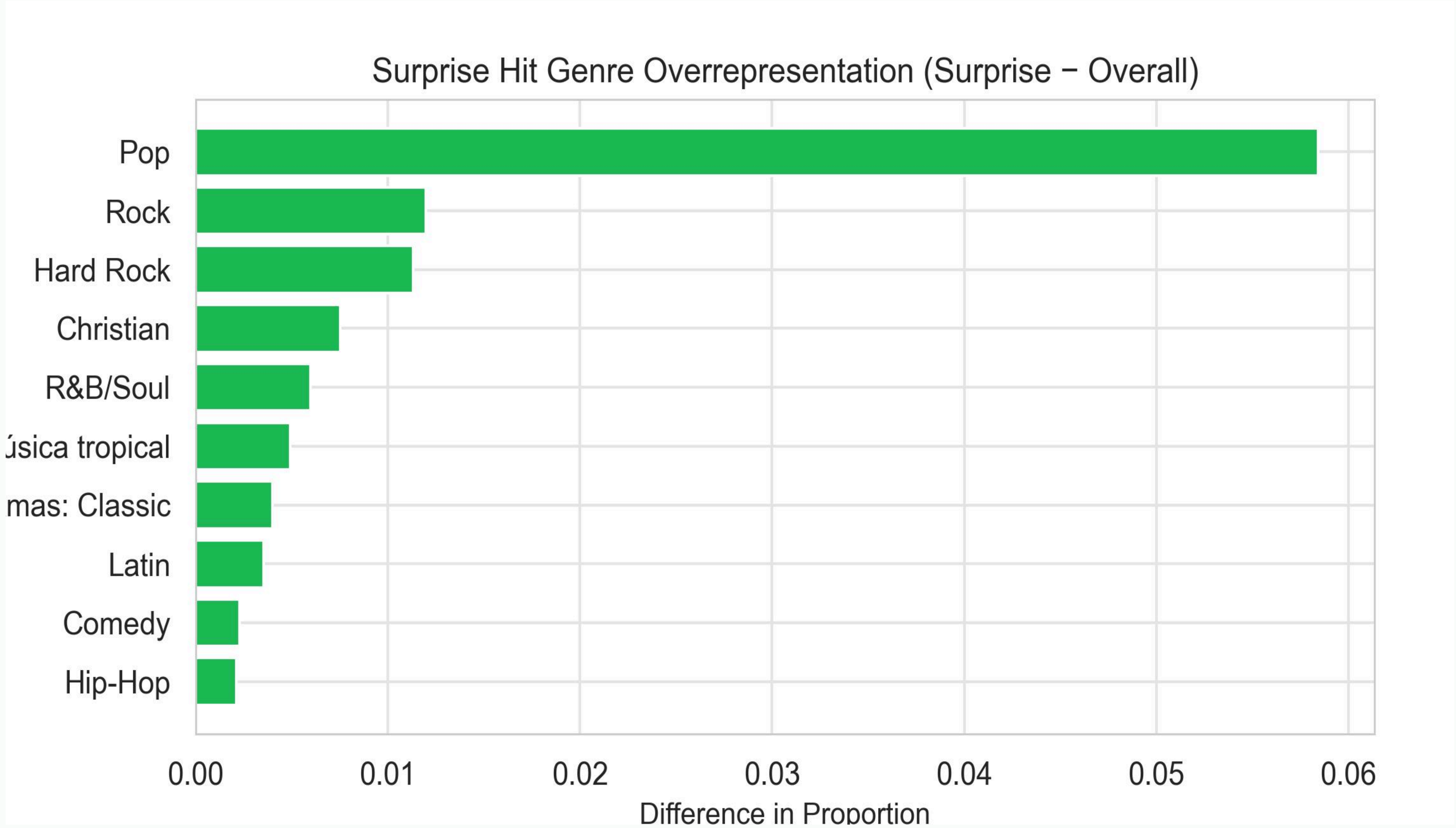
Surprise Hits ≠ Stronger Artists

Surprise hits often come from:

- average or modest momentum
- typical artist strength

But they disproportionately belong to:

- Pop
- Rock
- R&B



Interpretation

Mainstream genres receive broader exposure and amplification.

Platform exposure favors popular categories

Discovery systems leave measurable value untapped.

Insights

- 80% of success unexplained by reputation
- niche genres systematically underexposed
- mainstream genres disproportionately amplified

Opportunities

- boost hidden gems
- residual-aware ranking
- diversify recommendations
- improve long-tail engagement

Why This Matters

Current systems over-amplify mainstream tracks
→ wasted long-tail inventory

~20% of songs behave unexpectedly
→ missed discovery opportunities

Potential impact

- higher catalog utilization
- more diverse listening
- improved retention
- better fairness for smaller artists
- stronger recommendation quality