



Analyzing Data Products in the Platform Economy: Spotify's Discover Weekly

Presented by:

- Sakib Hosen Himel
- Mahmuda Akter
- Zabihullah Azimy
- Prashant Kumar

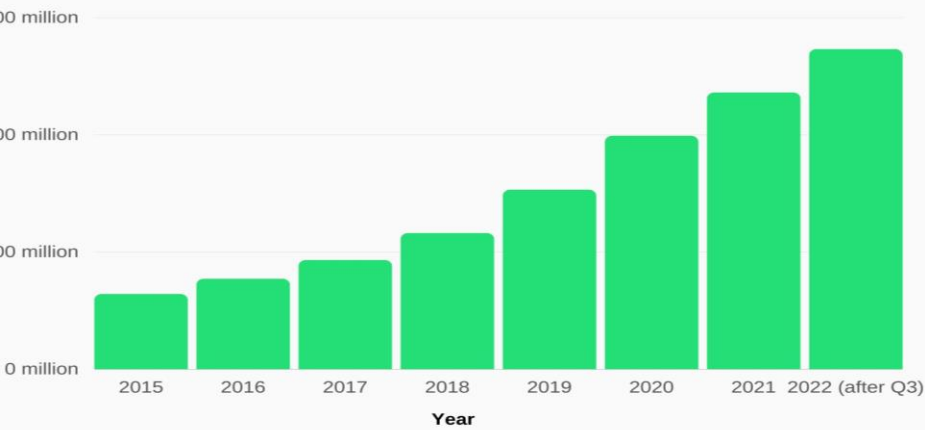
Master of Data Science
Business intelligence 2023-24

Company overview

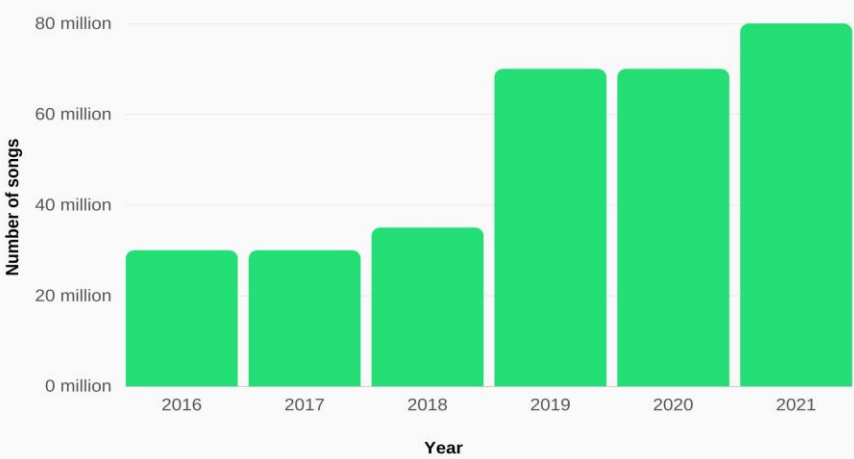
- Spotify Technology S.A. is a music streaming service.
- Developed by Swedish company Spotify Technology.
- Owned by its founders, Daniel Ek and Martin Lorentzon
- It is headquartered in Stockholm, Sweden. Spotify launched their service on 7 October 2008.
- As of 20 August 2018, it is available in 65 regions.



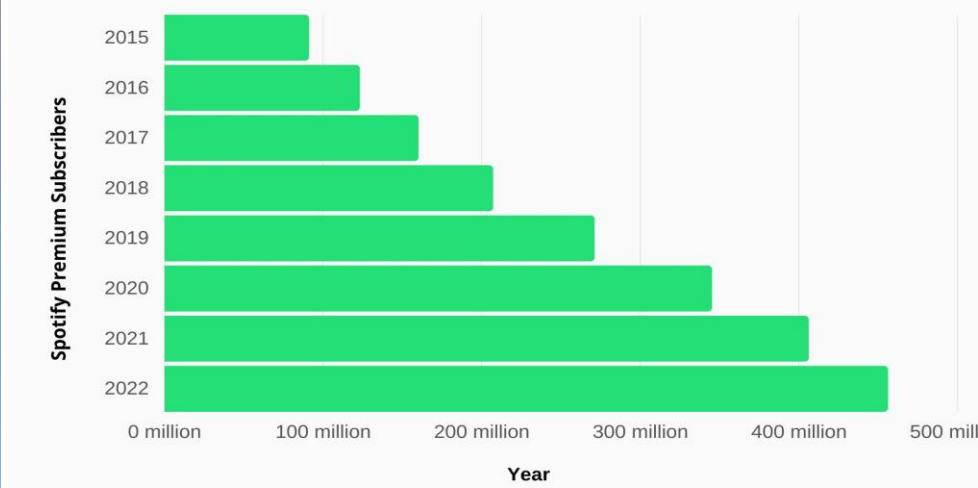
Spotify: Number of Ad-Support Users



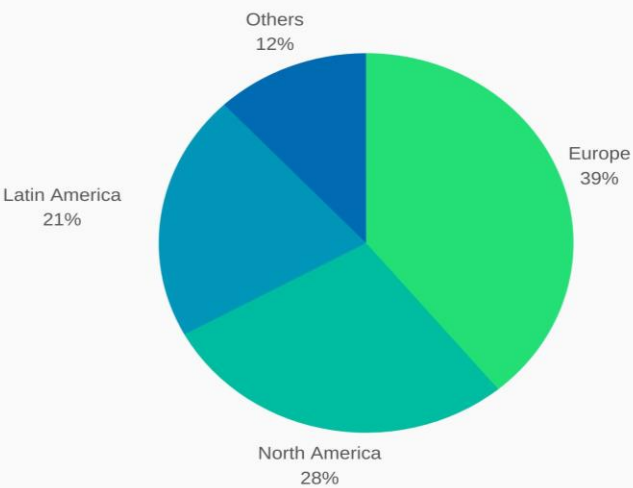
Spotify: Number of Songs



Spotify: Number of Premium Subscribers



Spotify: Percentage of Premium Subscribers Worldwide



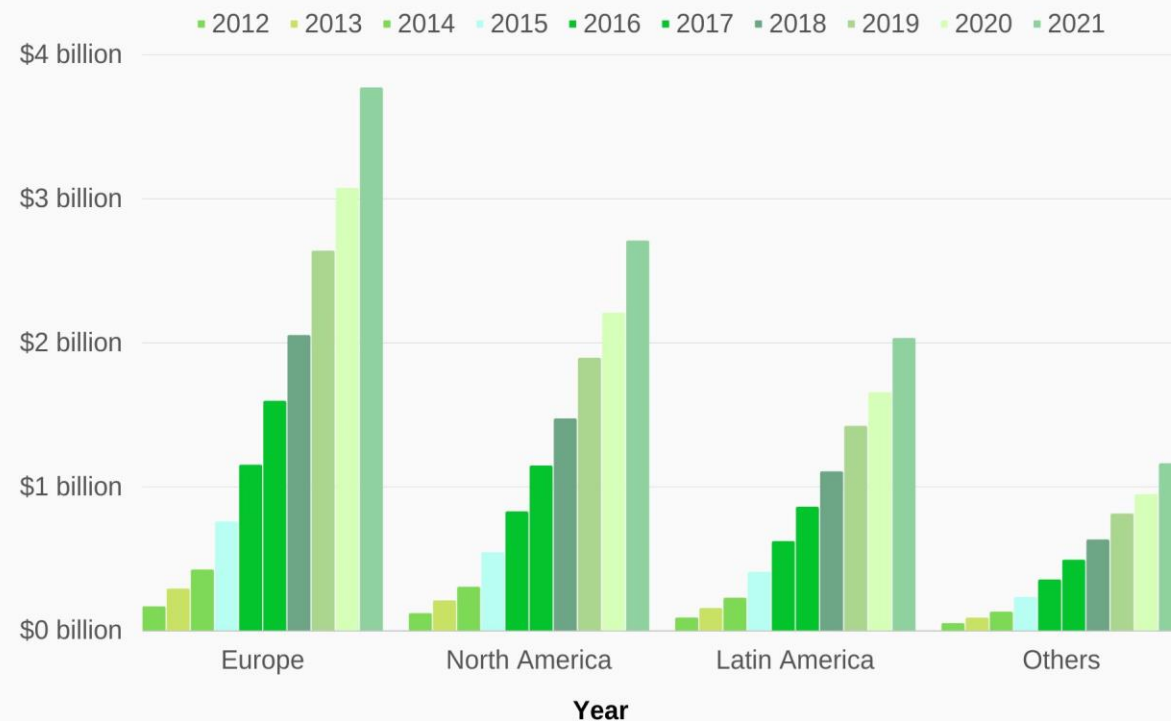
Key accomplishments:

- As of December 2023, Spotify has around **574m** active users (over **40%** of which are premium subscribers)
- **226 million** paying subscribers.
- In Europe **336m** users, around **157m** premium users (**58%**).
- Spotify has over **100 million** tracks in its library

Spotify Revenue by Segment



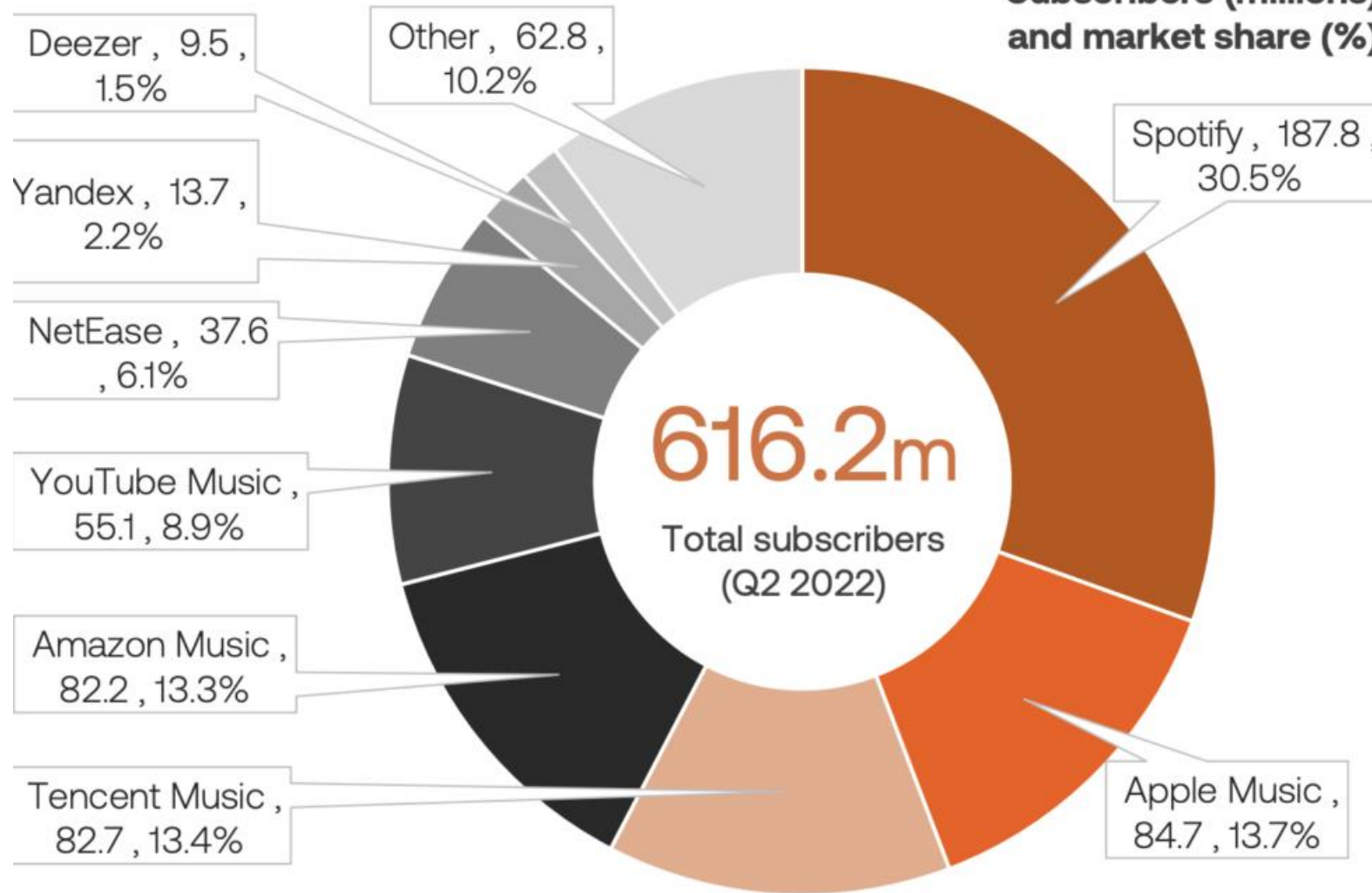
Spotify: Average Revenue Per User by region



Revenue

As of September 30, 2023, the revenue was \$13.647B, a 12.01% increase year-over-year.

**Subscribers (millions)
and market share (%)**



**Market
Dominance**

Products overview:

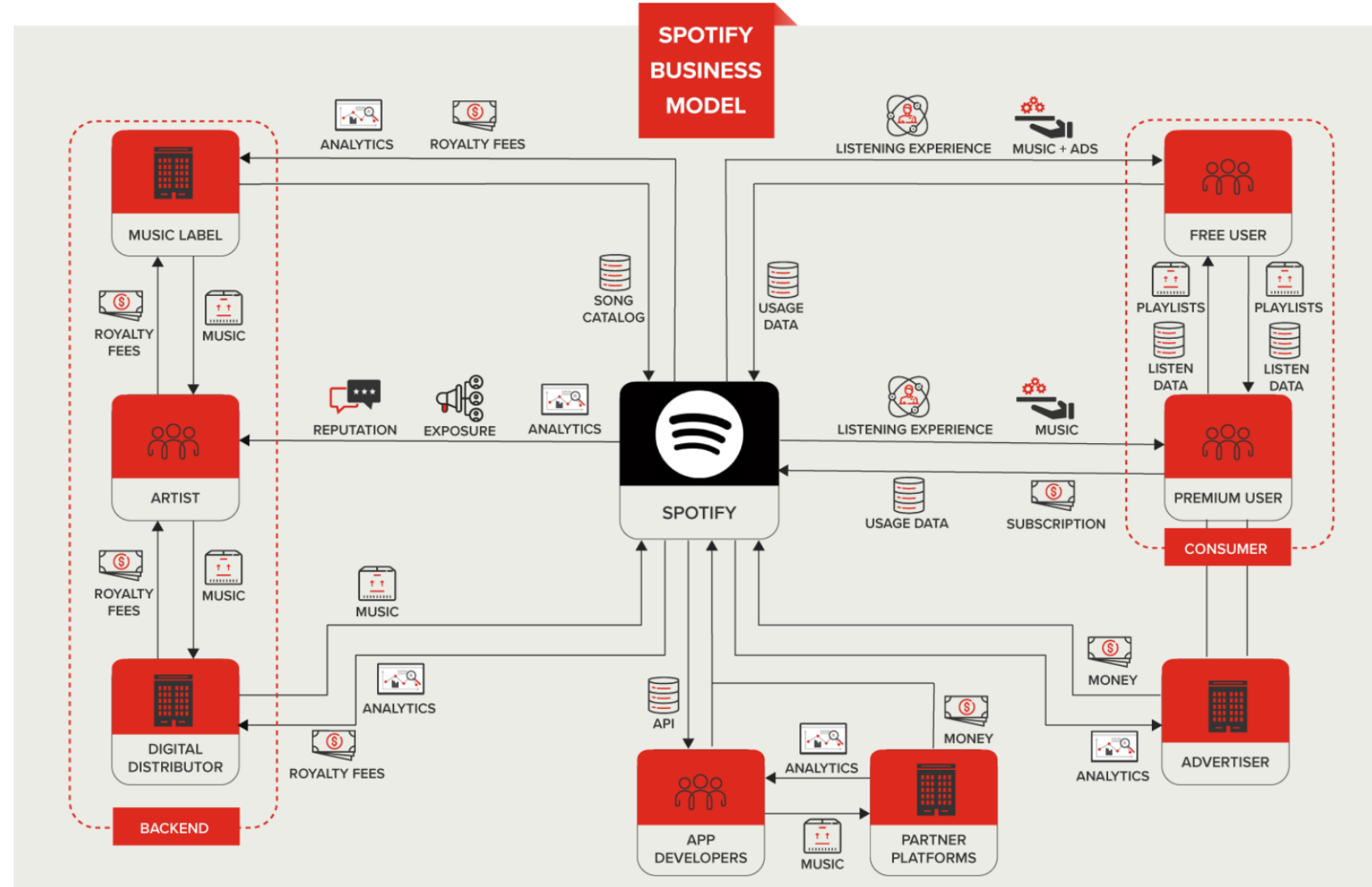
Spotify offers a range of products and features designed to enhance the music listening experience.
A comprehensive music and podcast streaming service with a variety of products and features to suit every need.

- Music Streaming Service (Free, premium)
- Podcast Platform(Original, Discovery)
- Audio Book, etc.

Business model patterns

- ✓ Spotify operates on a multi-sided platform business model.
- ✓ Offers both free and premium subscription options for users.
- ✓ A business-to-business-to-consumer (B2B2C) model for partnering with other companies.

1. Freemium model
2. Premium model



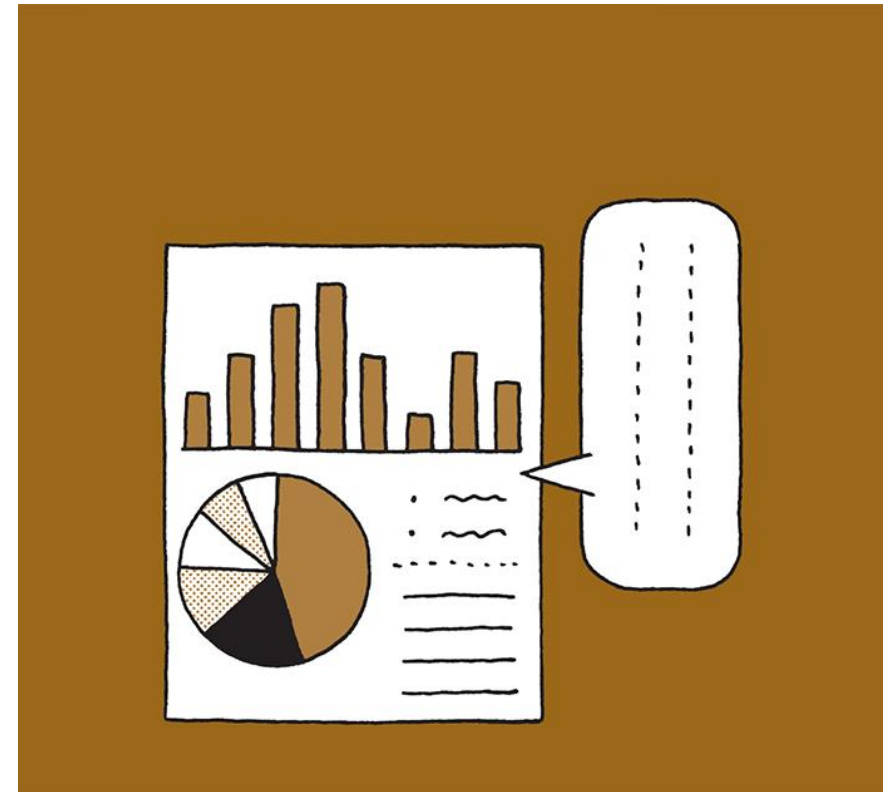
Data Product - Discover Weekly

Customer data in Spotify is used to give the listeners individualized recommendations of what other music might be interesting to them.

How Does it Work?

Spotify's algorithms analyse user data:

- listening history
- saved songs
- user-generated playlists



Data Product Canvas

PROBLEM

1. Limited music discovery
2. Narrow music horizons
3. Constrained musical landscape (Less Genre diversity)

DATA

- User listening history
- User-generated playlists
- Audio features of songs
- External metadata

SOLUTION

- Collaborative Filtering
- Content based filtering
- Audio Features

KPI'S

- Average time spent listening
- Listeners Demography
- Retention rate
- Daily Active Users
- Save Rate
- Play count and more

HYPOTHESIS

- Hypothesis 1: Goals for discovery
- Hypothesis 2: Context from individual habits
- Hypothesis 3: One great recommendation

ACTORS

- Active Spotify users
- Spotify data scientists and engineers
- Music industry partners

ACTIONS

- Engage with Discover Weekly playlists
- Monitor and improve Discover Weekly recommendations
- Provide relevant metadata and insights

VALUES

- Increased user engagement and satisfaction
- Improved retention rates
- Growth in user base and revenue

RISKS

- Over-reliance on algorithms
- New music releases
- Changes in user preferences

PERFORMANCE/IMPACT

- Introducing users to new music
- Attracting and retaining users
- Differentiating Spotify from its competitors
- Increased revenue
- Reduced marketing costs

Problems

- users often prefer repeatedly listening to a small selection of their favourites.
- Some people get stuck in a specific genre or a handful of artists.
- They never venture into the musical unknown.
- For these users, as well as for the musically adventurous among us:

Spotify delivers a compelling solution called **Discover weekly**



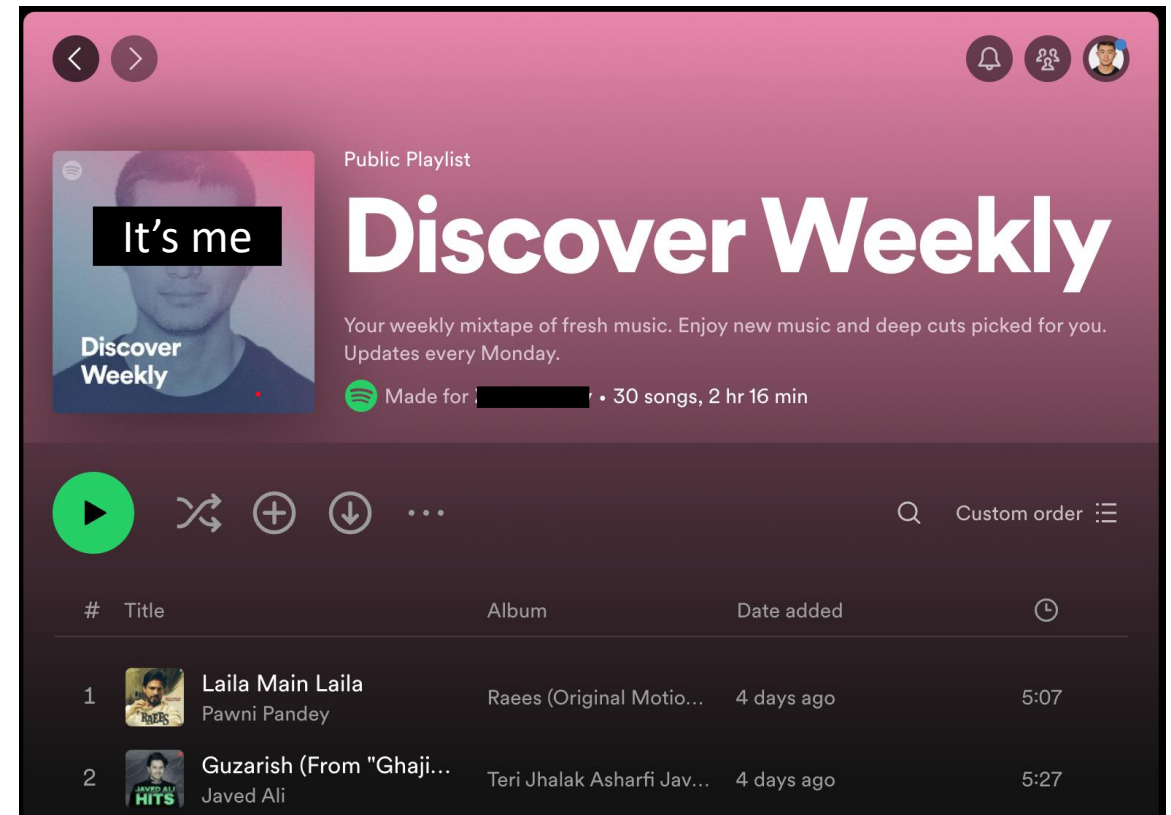
Recommendation systems

- A recommendation system is an algorithm which tries to predict the rating or preference a user would give to an item.
- The basic technology that powers recommendation systems is largely the same, regardless of the specific domain of application (like movies, books or films).



How Spotify's discover weekly works?

- Discover weekly recommendation model is a combination of a number of **effective recommendation** techniques previously used by other industry players.
- This has led to a **uniquely powerful** music recommendation engine.



Spotify's recommendation engine is based on three recommendation models

- **Collaborative filtering** consists of collecting and analyzing users' behaviors.
- **Content-based filtering** looks at the descriptions of songs and artists.
- **Audio features** are extracted from the raw audio through machine learning.



**Your
Discover
Weekly**

Collaborative filtering

- With collaborative filtering, recommendations are outsourced to the users.
- Listening behaviours are analysed and used as a way to predict users' preferences.
- People who listen to similar music likely have similar musical tastes.
- Conversely, if the same group of like-minded people listens to two different songs, they're probably similar.

Users

1	0	0	0	1	0	0	1
0	0	1	0	0	1	0	0
1	0	1	0	0	0	1	1
0	1	0	0	0	1	0	0
0	0	1	0	0	1	0	0
1	0	0	0	1	0	0	1

Songs

A matrix in action.

Content-based filtering

- Content-based filtering, compares descriptors of an item against a user profile
- The user profile is built from the same tags that describe the content the user consumes, like 'rock' or 'classical', i.e., semantic information.
- In 2014, Spotify acquired **Echo Nest**, a company that employed Natural Language Processing (NLP)
- To then extract semantic information from music-related text content.



Content-based filtering

- Spotify crawls the Internet to figure out what people think about artists and songs.
- News articles, blogs and online reviews are analysed to infer the adjectives and nouns.
- These NLP algorithms also spot connections between different songs by looking at the language related to different artists.

A circular diagram with 'NLP' in the center. Surrounding it are five icons: a book, a head with gears, a speech bubble, a telephone, and a Wi-Fi symbol, all connected by a circular line.

NLP

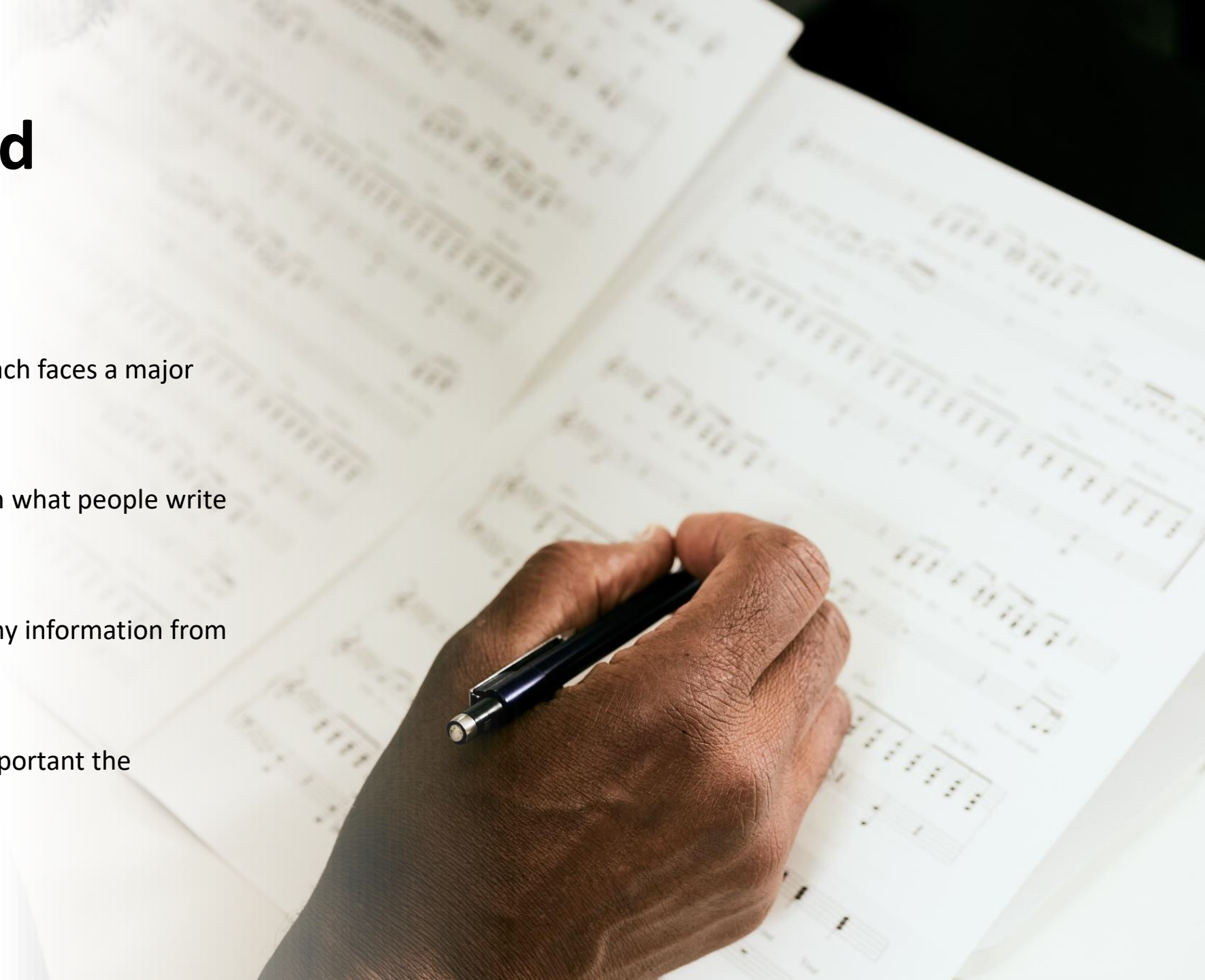
Content-based filtering

The content-based filtering approach faces a major issue:

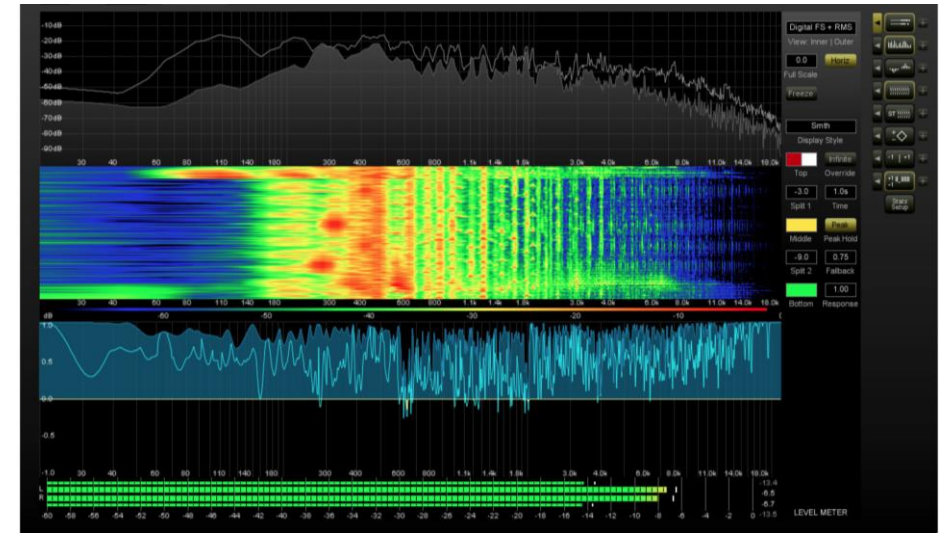
- All the information derives from what people write about the music and the artists.

Or

- The NLP algorithms don't get any information from the songs themselves.
- This **fails** to account for how important the actual *sound* of a track is.



What to do then?



It's better to leverage that information through competent audio analysis, in order to achieve more well-rounded recommendations.

Spotify

Your Discover Weekly

Introducing Your Discover Weekly
30 tracks you'll love,
delivered every Monday.

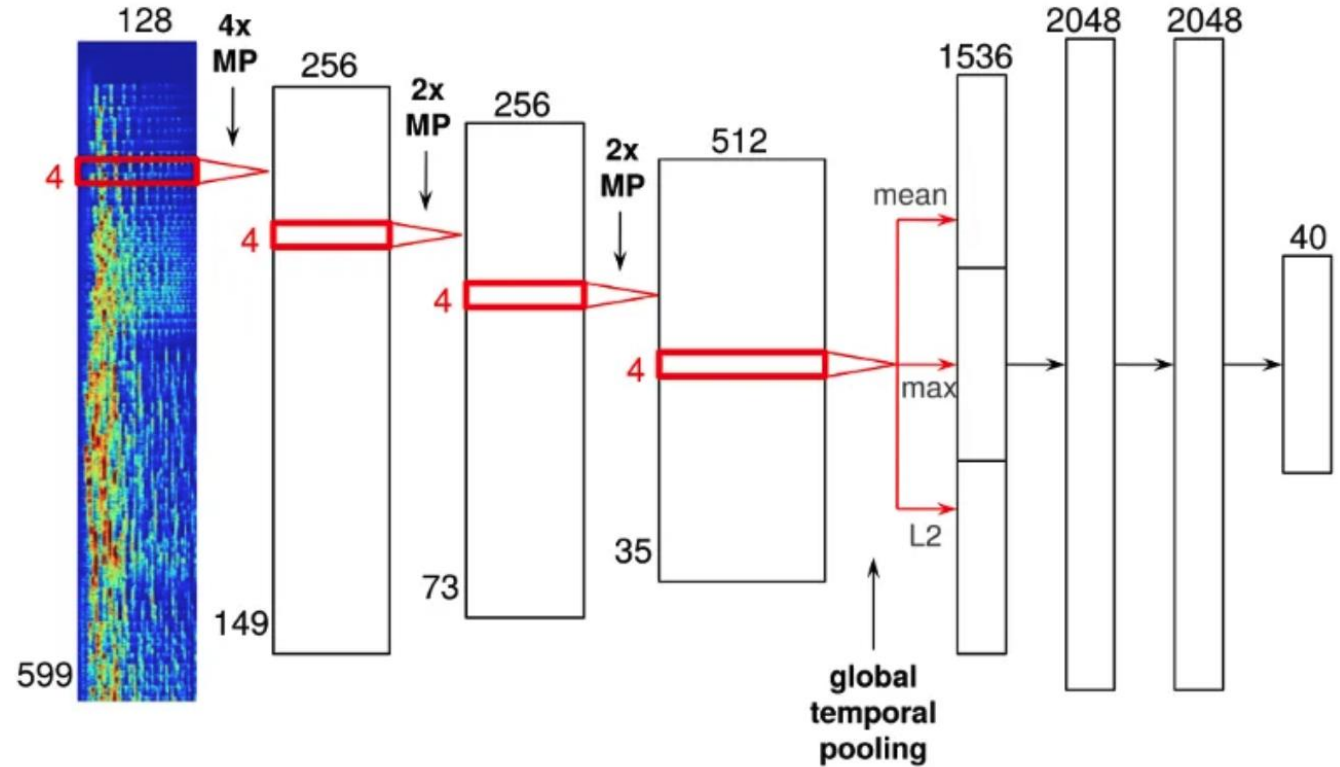
Audio features

- Spotify uses CNN to extract musical features directly from raw audio.
- Convolutional networks have mainly been used with visual data.
- Data scientists have successfully applied them to image detection.
- This is achieved by feeding a dataset of images to the network, pixel by pixel.
- Once trained, the algorithm is capable of classifying different objects that appear in images that are new to the network.



Audio features

In the case of Spotify, the network has been modified to accept audio data as the input instead of pixels.

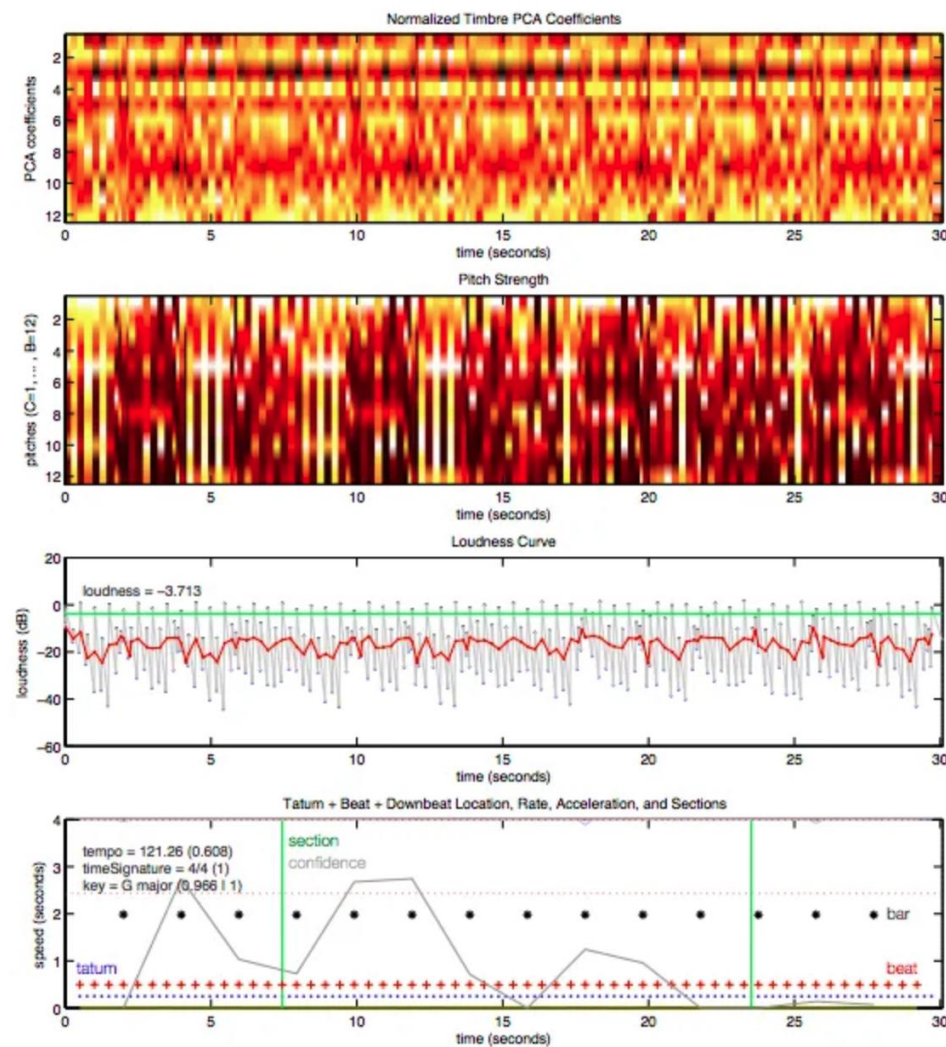


The architecture of the Spotify convolutional network.

Audio features

The output of the neural network consists of an understanding of the song, which includes audio features such as:

- **Key**
- **Mode**
- **Tempo**
- **Loudness**
- **Time signature**



Plot of the output of the network for 30 seconds of 'Around the World' by Daft Punk.

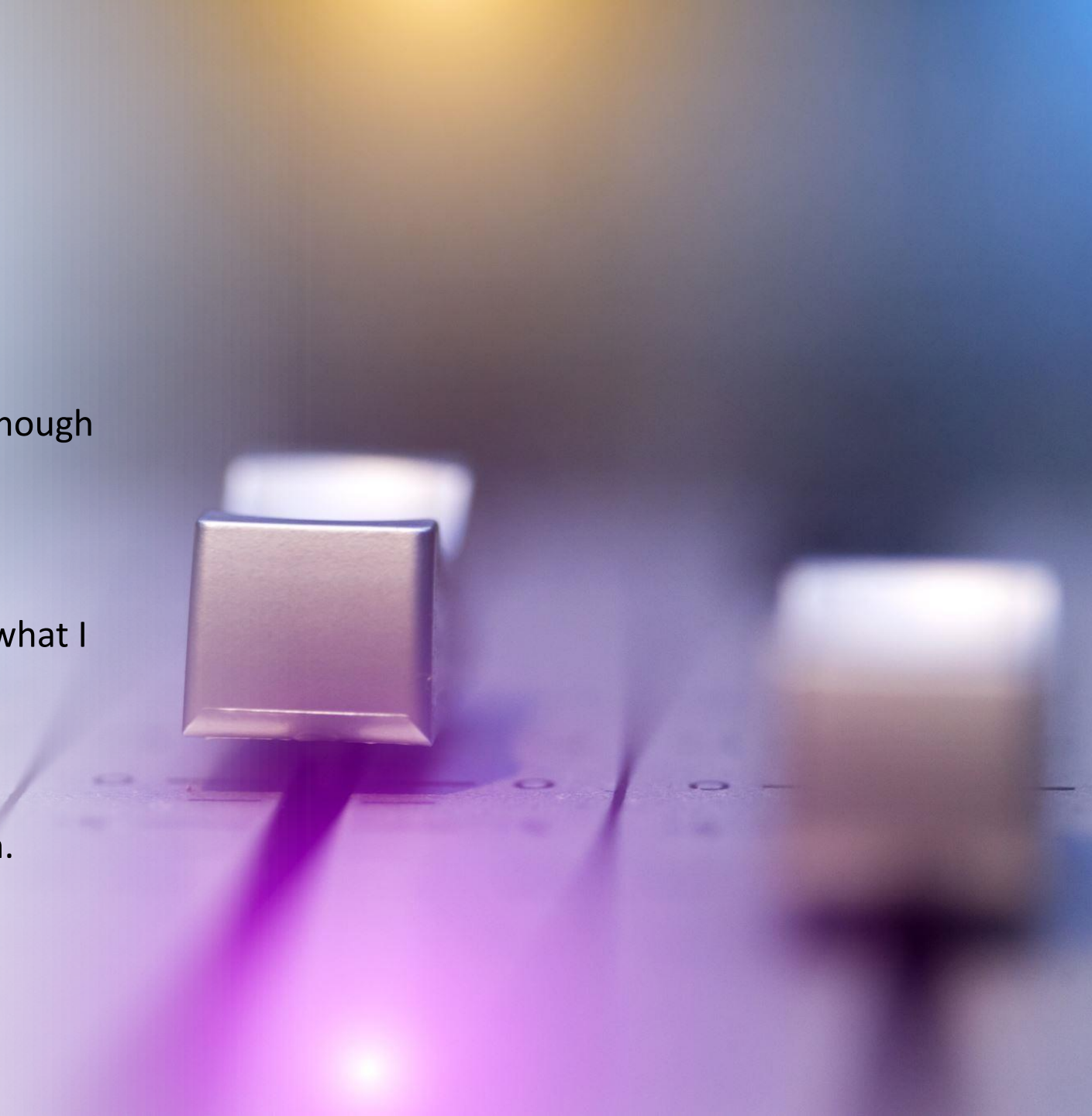
Similarity vs Diversity

- The thing is, people tend to like new music that, in degrees, is similar to what they're used to.
- If a recommendation system was to consider only similarity, the job would be relatively easy.
- The motto for such recommender system would be '**same old blues**'.
- But, after a while people get bored of listening to infinite variations on the same theme.



Meet Diversity!

- A recommendation engine should be bold enough to propose music that is outside our musical comfort zone.
- If the music suggested is too different from what I like, then I likely won't enjoy it!
- So, this is risky! When someone doesn't like listening, he/she might stop the subscription.



What might help?

Thinking of high similarity and high diversity as the two ends of a continuous spectrum.

With a variance of behaviours and expectations as large as the scope of the human experience, we should think of a solution!

Spotify could ask their users to customise the discovery engine in order to respond to their recommendation needs.



Hypotheses:

Hypothesis 1: Goals for discovery

On-platform behaviours provide clearer signals when they're viewed within the context of the user's goals.

Hypothesis 2: Context from individual habits

That metrics should be normalized relative to each user's typical behaviour, rather than assuming the same standards for all users.

Hypothesis 3: One great recommendation

That just one *great* recommendation could have a large positive effect on satisfaction.



Hypotheses Validations



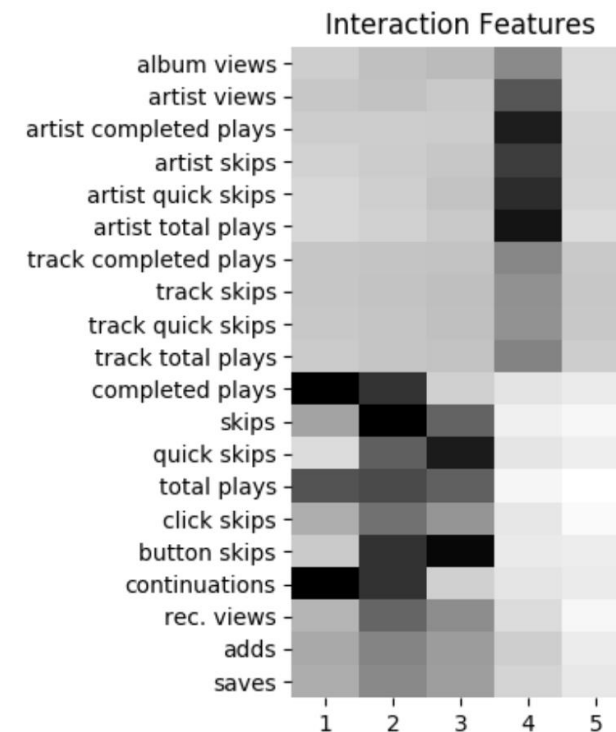
Let's see at hypotheses validations
in the next slides



These validations are done
quantitatively – using data

Validating our first hypotheses- Goal for Discovery

- **Cluster 1:** consists of completed plays and few skips, corresponding to goal 1 (play new music in the background).
- **Cluster 2:** consists of plays, skips, and saves, corresponding to goal 2 (listen to new music now and later).
- **Cluster 3:** Users mostly skip and save (goal 3: find new music for later);
- **Cluster 4:** Folks in cluster 4 listen downstream (goal 4: engage with new music).



Validating our second hypotheses- Context from Individual Habits

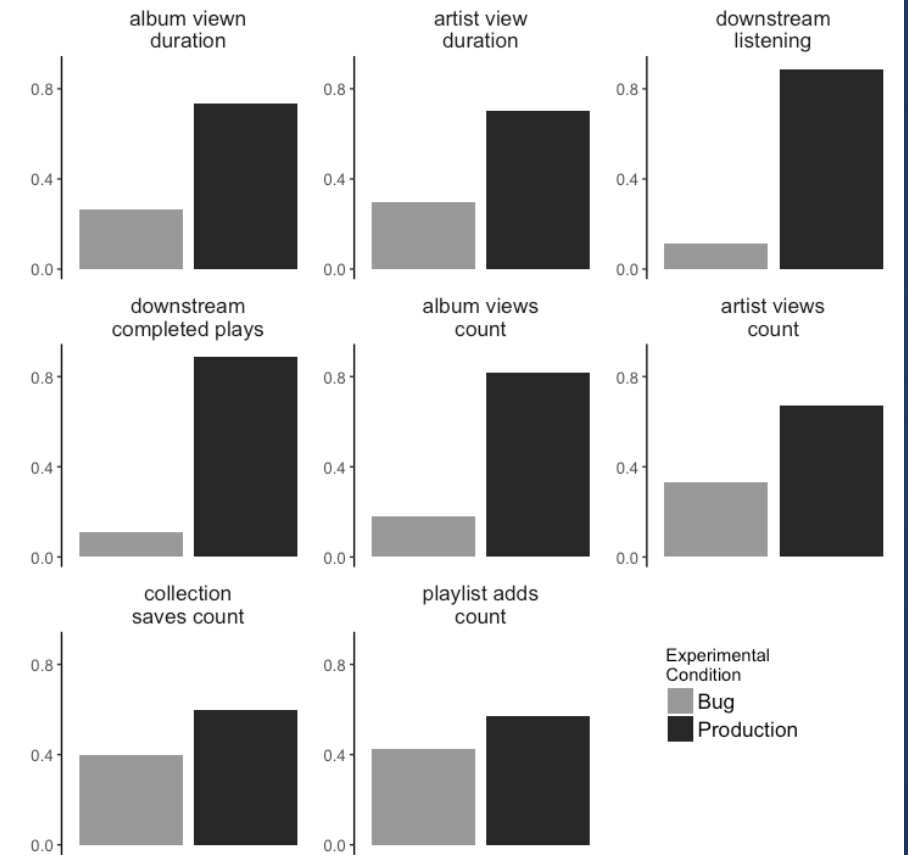
User Judged their playlist relative to previous week and level of enagement varied between users.

Although behaviours differ according to users' goals, certain behaviours were unambiguously positive are-

- Saving / adding a tract.
- Viewing an artist or album page.
- Listening downstream.

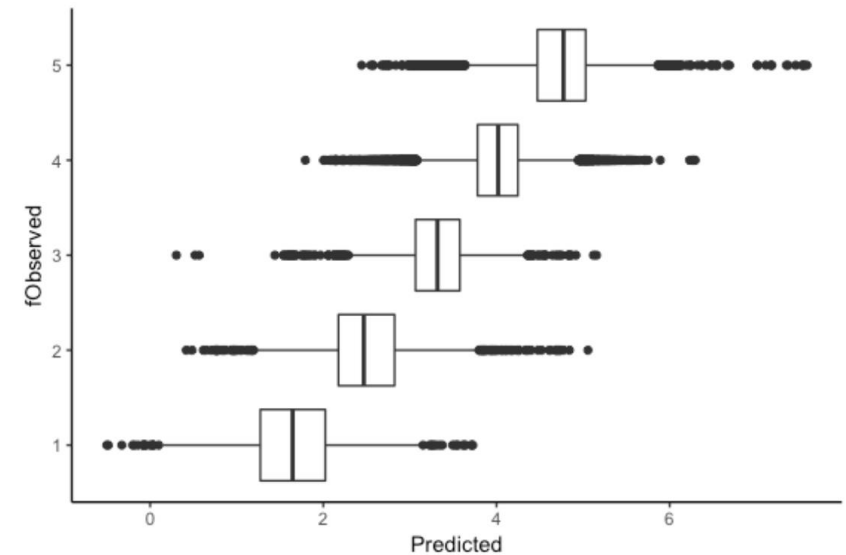
- Check out how these signals differ between the production system, and a “degraded” version of the algorithm containing a bug.
- We observed significant differences between the systems, especially for the normalized versions of the metrics.

Thus: metrics should be normalized to each user.

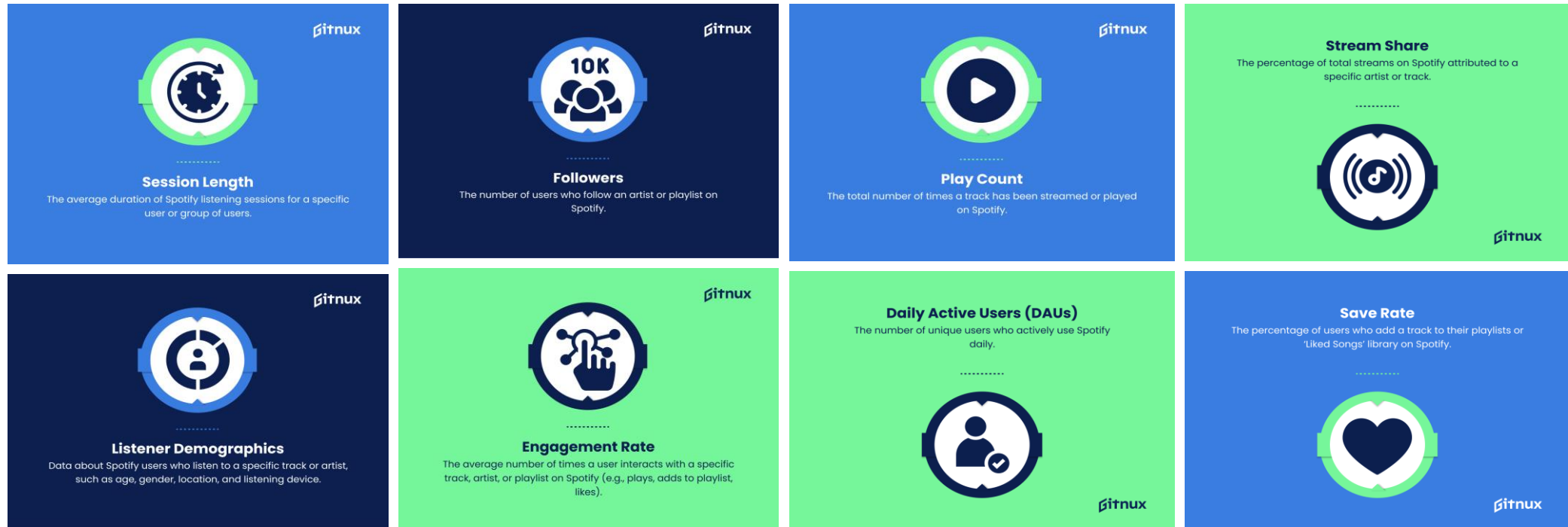


Validating our third hypotheses- One Great Recommendation

- Collected data were leble data, with overall and current with satisfaction from users by spotify.
- Model trained gradient boost decision tree.
- The x axis is the predicted values, and the y axis is the observed values; a perfect model would show a straight line going up and to the right.
- We see that the model is successfully differentiating between positive and negative experiences.



Key Performance Indicators



<https://gitnux.org/spotify-metrics/>

Actors



Active Spotify Users



**Spotify Data Scientists
and Engineers**



**Music Industry Partners
– Record companies**

Actions

Actors	Actions	Outcomes	Metrics
Active Users	Engage with Discover weekly playlist	Enhance user engagement and satisfaction, increase retention rates, and attract new users.	1. No of users engaging with Discover Weekly playlists. 2. Average time spent listening to Discover Weekly playlists. 3. No of new songs discovered through Discover Weekly
Data Scientist and Engineers	Monitor Discover weekly recommendation	Improve the quality of recommendations, increase user satisfaction, and expand user engagement.	1. Precision of recommendations. 2. User satisfaction with recommendations. 3. No of recommend songs that users add to their playlists.
Music Industry partners	Provide relevant metadata and insights	Strengthen relationships with music industry partners, enhance music discovery, and increase revenue.	1. Number of new music partners. 2. Increased revenue from music partnerships.

Values

- Increased User Engagement and Satisfaction
- Improved Retention Rates
- Growth in User Base and Revenue



Risks



Over-reliance on Algorithms

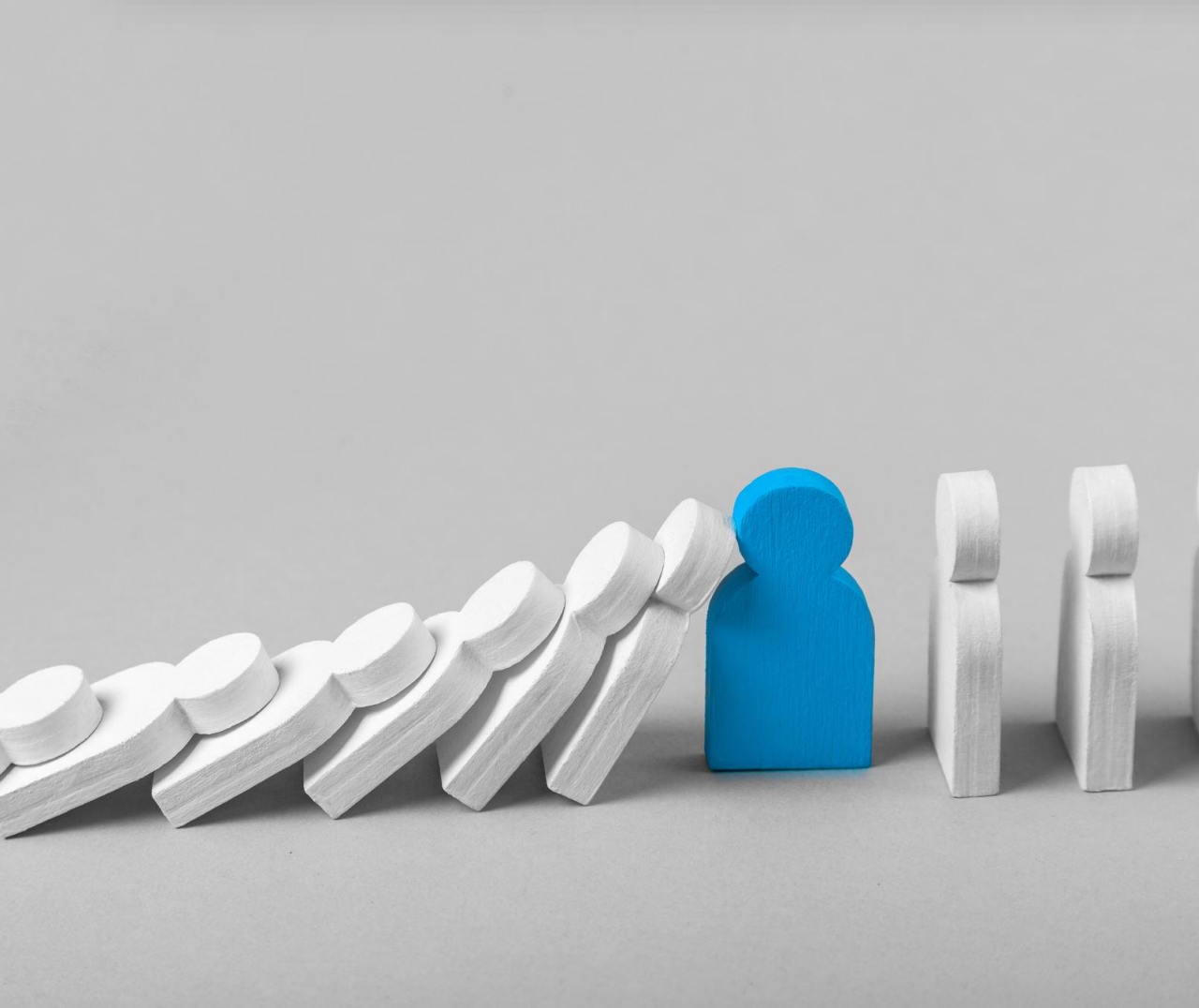


New Music Releases



Changes in User Preferences

Performance/Impact



- **New Music and Attracting/Retaining Users**
- **Differentiation from Competitors**

How Spotify Attract Consumers for Growth?

- Personalized & Easy-to-Use App Interface
- Attracting Users with Freemium Model
- Personalized Recommendations
- Regular Updates and Innovations
- Place: Expanding into the metaverse



References:

<https://research.atspotify.com/2018/07/understanding-and-evaluating-user-satisfaction-with-music-discovery/>

<https://medium.com/the-sound-of-ai/spotify-s-discover-weekly-explained-breaking-from-your-music-bubble-or-maybe-not-b506da144123>

<https://gitnux.org/spotify-metrics/>

<https://businessmodelnavigator.com/case-firm?id=94>