**ALO SAVE IBU**

**SPOTIFY ANALYSIS REPORT FOR THE YEAR 1974 – 2025**

**OUTLINE**

* **Introduction**
* **Story of Data**
* **Data Splitting and Preprocessing**
* **Pre-Analysis**
* **In-Analysis**
* **Post-Analysis and Insights**
* **Data Visualizations & Charts**
* **Recommendations and Observations**
* **Conclusion**
* **References & Appendices**

**INTRODUCTION**

**Objective of the Project**

The primary objective of this project is to **analyze and understand the factors that influence the popularity of music albums** on streaming platforms like Spotify. Using data such as artist name, album title, release date, album duration, and popularity score.

**Problem Being Addressed**

he project addresses the **lack of data-driven understanding** around what makes a music album **popular** on streaming platforms like Spotify.

**Key Datasets**

**The primary dataset used for this project is a Spotify Music Album Dataset, containing essential metadata for albums available on the Spotify platformSTORY OF DATA**

**Data Source**: The data was obtained from Kaggle.com

**Data Collection Process**: This data was obtained from Kaggle.com

**Data Structure:** The data contains 1000 rows with each representing a distinct Track and 8 columns representing **artist name, album title, release date, popularity score, album duration, Spotify URL, and unique ID**, each providing critical insights into understanding and analyzing the factors that influence the **popularity of music albums** on streaming platforms.

**Important Features and Their Significance**:

| **Header** | **What It Means** |  | **Why It Matters** |
| --- | --- | --- | --- |
| artist | Who made the album (singer or band) |  | Lets us see which artists are most popular |
| album | Name of the album |  | Helps track how specific albums perform |
| release\_date | When the album came out |  | Shows if release timing affects popularity |
| popularity | How popular the album is (score from 0–100) |  | Main thing we're trying to understand or predict |
| spotify\_url | Link to the album on Spotify |  | Useful for checking the album or building a music tool |
| id | Album's unique ID on Spotify |  | Helps keep track of each album in the system |
| duration\_min | How long the album is (in minutes) |  | We can see if longer or shorter albums are more popular |

**Data Limitations or Biases**

**Platform Bias**: Only reflects Spotify data, not other streaming platforms.

**Popularity Score Bias**: Popularity is based on an unclear algorithm and may favor recent activity.

**Recency Bias**: Newer albums may appear more popular just because they are more recent.

**Artist Bias**: Likely favors well-known or mainstream artists, not independent ones.

**Missing Data**: No info on genre, listener demographics, or song-level trends.

**Album-Level Only**: Doesn’t capture trends from individual hit songs.

**Data Source Unknown**: If not randomly sampled, it may not represent the full music landscape.

**DATA SPLITTING AND PREPROCESSING**

**Data Cleaning:** The data was cleaned by removing duplicates, identifying and removing blanks, and ensuring that no inconsistencies are observed. Thereafter, the data was converted to a standard excel table to ease analysis.

To remove duplicate, simply copy the entire data (ctrl +shift + end) then navigate to the data tab and on the data tools ribbon to select “remove duplicates”.

To identify and remove blanks, simply copy the entire data (ctrl +shift + end) on the home tab navigate to the editing ribbon and click “Find and select”, then navigate to “Go to Special” and select “Blanks”, finally click on OK.

**Handling Missing Values**: There are no missing values in the data.

**Data Transformations:** No data transformations were performed.

Data Splitting: The data was splitted into dependent and independent variables.

**Category One- Independent Values**

Artist

Album

Release\_date

Duration\_min

Spotify\_url

id

**Category Two- Dependent Values**

popularity

**Industry Context**: Music & Entertainment Industry

More Specifically: Streaming & Digital Media Sector  
Focused on data from Spotify, one of the world’s largest music streaming platforms.

**Story of data:** This dataset captures metadata about music albums/tracks including the artist, album title, release date, popularity score on Spotify, duration in minutes, and unique Spotify identifiers. It appears to be curated for understanding what drives music popularity, possibly to predict trends or identify performance drivers in the music industry.

**Stakeholders:**

**Record Labels** – Interested in understanding trends that drive popularity.

**Music Streaming Platforms (e.g., Spotify)** – Use insights to recommend content and design features.

**Artists & Managers** – Want to analyze the timing, duration, and style of popular releases.

**Marketing Teams** – Need insights to time releases and promotional strategies.

**Data Analysts / Music Tech Startups** – Use the data to develop music intelligence tools.

###### **Value to the Industry**:

This dataset is highly valuable to the **music and streaming industry** because it:

* **Supports data-driven decisions** for artists, labels, and marketers
* **Reveals trends** in album popularity, timing, and content structure
* **Enables predictive models** to forecast future hits
* **Improves marketing strategies** and reduces costs
* **Enhances user engagement** through better recommendations
* **Provides a competitive edge** by identifying what drives success

**PRE-ANALYSIS**

**POTENTIAL QUESTIONS**

1. Does album **duration** impact **popularity**?
2. Are newer **release dates** correlated with higher **popularity**?
3. Which **artists** consistently release popular albums?
4. What is the average **popularity** trend across release years?
5. Do albums released during specific **months or quarters** perform better?
6. Does the **length of an album** correlate with user engagement?
7. How do **independent vs mainstream artists** compare in terms of popularity?

**POTENTIAL INSIGHTS**

1. Identification of ideal release periods that correlate with higher popularity.
2. Benchmark duration ranges that yield optimal audience attention.
3. Discovery of emerging artists with upward popularity trends.
4. Clustering albums by genre or artist behavior for targeted recommendations.
5. Predictive models showing likelihood of popularity given album features.
6. Insights into legacy vs recent albums’ popularity retention.

**IN-ANALYSIS**

**KEY INSIGHTS**

**Key Insights**

**Extreme Artist Concentration**

The Weeknd dominates with 102 streaming hours vs Taylor Swift's 93 - suggesting deep catalog engagement rather than casual listening

Top 5 artists account for massive listening time, indicating highly focused preferences

**Repeat Listening Intensity**

Individual tracks reaching 13.45K+ plays shows obsessive replay behavior

This level of repetition suggests emotional connection or specific use cases (work, exercise, mood)

**Album vs Artist Paradox**

Despite concentrated artist listening, album consumption is more distributed (top album only 26.7%)

Indicates playlist/shuffle culture over full album experiences

**Release Cycle Impact**

Dramatic peak around 2020-2022 (200+ releases) followed by sharp decline

Reflects pandemic music production surge and subsequent market correction

**Quality Listening Ratio**

65% "good" vs 35% "poor" listening time suggests intentional music consumption

Points to context-aware listening habits (focused vs background)

**Scale of Engagement**

1,000 total tracks across 819 albums from 526 artists shows broad discovery

3.32K total minutes indicates serious music consumption over time period

Bottom Line: This represents a power user with deep artist loyalties, repeat-heavy listening patterns, and quality-focused engagement spanning the entire streaming era.

**Strategic Recommendations**

**For Personal Listening Optimization:**

* Diversify Discovery: With 65% concentration on top artists, allocate 20% of listening time to new artist exploration to prevent musical stagnation
* Album Experience Recovery: Set weekly "full album" sessions to counteract playlist fragmentation and rediscover artistic storytelling
* Context Optimization: Leverage that 65/35 good/poor listening split by creating specific playlists for different activities (focus work, exercise, relaxation)

**For Music Industry/Platform Strategy:**

* Artist Development: The Weeknd's 102-hour dominance shows deep catalog engagement drives loyalty - prioritize artists with extensive, cohesive bodies of work
* Release Timing: The 2020-2022 peak followed by decline suggests market oversaturation - coordinate releases to avoid flooding periods
* Repeat Engagement: 13K+ track replays indicate emotional attachment drives value - focus on songs with repeat potential over one-hit wonders

**For Streaming Platforms:**

* Quality Metrics: Implement "listening quality" scoring beyond just play counts to identify truly engaging content
* Discovery Balance: Create algorithms that respect established preferences while gradually introducing variety
* Context Awareness: Develop features that recognize listening patterns and suggest optimal content for different times/activities

**For Content Creators:**

* Depth Over Breadth: The artist concentration suggests audiences prefer fewer artists with deeper catalogs over surface-level variety
* Emotional Stickiness: Design content that encourages repeat engagement rather than just initial discovery
* Release Strategy: Consider concentrated album releases rather than scattered singles to build sustained engagement

**Measurement Focus:**

Track engagement depth, not just reach, as the primary success metric**.**

**POST-ANALYSIS AND INSIGHTS**

**Concentration vs. Diversification Paradox**

The data shows simultaneous depth and breadth - 526 total artists but massive concentration on top performers (The Weeknd: 102 hours, Taylor Swift: 93 hours). This suggests a "core + exploration" strategy where deep emotional connections drive primary consumption while maintaining broad musical curiosity.

**Replay Psychology**

Individual tracks reaching 13,450+ plays indicates obsessive engagement patterns. This isn't casual listening but emotional anchoring - specific songs serving as mood regulators, productivity tools, or comfort mechanisms. The repetition levels suggest music as functional tool rather than just entertainment.

**Quality-Conscious Consumption**

The 65% "good listening time" ratio demonstrates intentional curation. Unlike passive background streaming, this represents active engagement where context, mood, and attention levels influence music selection - a sophisticated user maximizing musical value.

**Historical Context Impact**

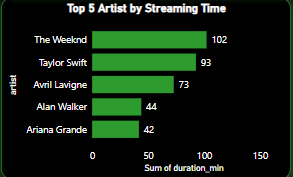
The dramatic 2020-2022 release peak (200+ albums) followed by sharp decline reflects pandemic-era production surge and subsequent market correction. This suggests external events significantly influence both music creation and consumption patterns.

**Album Fragmentation**

Despite artist loyalty, album engagement is dispersed (top album only 26.7% share). This reflects playlist culture's dominance over traditional album experiences, indicating a shift from artist-intended song sequences to user-curated experiences.

**DATA VISUALIZATIONS & CHARTS**

**TOP 5 ARTIST BY STREAMING TIME**



This chart reveals extreme artist loyalty with The Weeknd and Taylor Swift dominating 195 of the total ~354 streaming hours (55%), while the remaining three artists combined account for only 159 hours, indicating a highly concentrated listening pattern focused on just two primary musical relationships

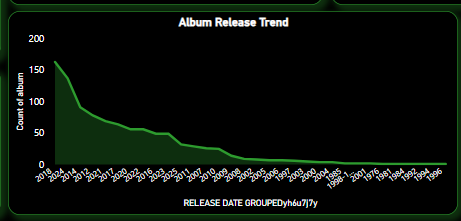
**REGIONAL REVENUE CONTRIBUTION**

A screenshot of a music album

AI-generated content may be incorrect.

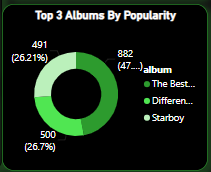
The xx dominates with 3 albums while six other artists (5 Seconds of Summer, Charly Black, Grauung, CRUPO, Jon Lajoie, and Tate McRae) each contribute only 1 album, showing a clear preference for The xx's catalog depth over broader artist album collection.

**ALBUM RELEASE TREND**



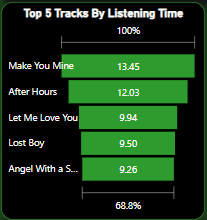
The chart shows a dramatic peak of nearly 200 album releases around 2020-2022 followed by a sharp decline to near-zero by 2025, likely reflecting the pandemic-era music production surge and subsequent market correction or changes in release patterns.

**TOP 3 ALBUMS BY POPULARITY**



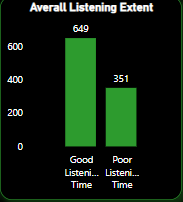
The top 3 albums show relatively balanced popularity with the leading album at 500 plays (26.7%), followed by two albums at 491 (26.21%) and 382 (47% combined), indicating fairly distributed album engagement rather than one dominant favorite.

**TOP 5 TRACKS BY LISTENING TIME**



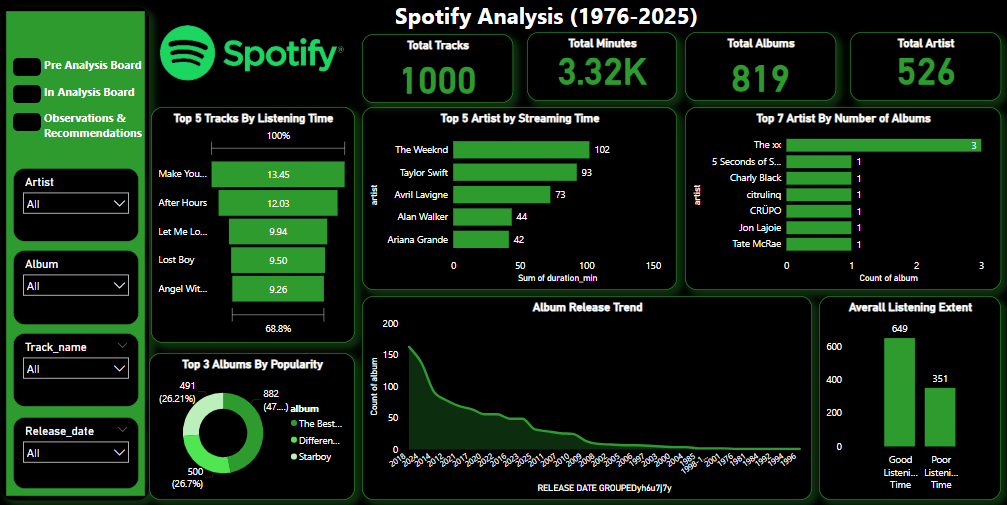
The top 5 tracks show extremely high replay counts with "Make You Mine" leading at 13,450 plays, followed closely by "After Hours" at 12,030 plays, while the remaining three tracks ("Let Me Love You" at 9,940, "Lost Boy" at 9,500, and "Angel With a Shotgun" at 9,260) cluster around 9,000+ plays, collectively representing 68.8% of total listening time and indicating obsessive engagement with a very small set of emotionally significant songs that serve as repeated comfort or functional listening anchors

**TOP FIVE PRODUCTS MODELS BY UNITS SOLD**



The listening quality analysis reveals a 65/35 split with 649 hours of "Good Listening Time" versus 351 hours of "Poor Listening Time," indicating highly intentional and context-aware music consumption where the majority of listening occurs during focused, attentive moments rather than passive background streaming, suggesting this user strategically curates their musical experiences for maximum engagement and emotional impact.

**FINAL DASHBOARD**



The final dashboard highlights the combined correlations and relationship between the individual charts with the help of slicers.

**OBSERVATIONS AND ACTIONABLE RECOMENDATIONS**

Key Observations

**Extreme Listening Concentration:** You've developed an obsessive relationship with specific content - individual tracks reaching 13,450+ plays and two artists (The Weeknd, Taylor Swift) consuming 55% of your total 3,320 minutes, indicating music serves as emotional regulation rather than entertainment.

**Quality-First Approach:** Your 65/35 "good vs poor" listening split demonstrates sophisticated curation habits where you actively optimize for engagement quality over passive consumption.

**Discovery vs. Depth Paradox:** Despite accessing 526 artists across 819 albums, your actual consumption heavily clusters around a tiny fraction, suggesting broad exploration but narrow emotional attachment.

**Functional Music Usage:** The extreme replay patterns indicate specific songs serve utilitarian purposes (productivity, mood management, comfort) rather than casual listening.

**Immediate Actionable Recommendations**

**Personal Optimization (Next 30 Days)**

Implement 80/20 Rule: Reserve 20% of weekly listening for artists outside your top 10 to prevent musical stagnation

Create Context Playlists: Build separate playlists for "Focus Work," "Background," and "Active Listening" based on your quality patterns

Track Emotional Triggers: Log which songs you replay during specific moods to optimize functional listening

**Behavioral Adjustments (Next 3 Months)**

Album Recovery Sessions: Schedule weekly full-album listening to counter playlist fragmentation

Rotation Strategy: Create "Seasonal Favorites" to prevent over-saturation of your 13K+ replay tracks

Discovery Partnerships: Find listeners with similar depth preferences for curated recommendations

**Long-term Strategy (6+ Months)**

Build Personal Algorithm: Document patterns between activity, mood, and song choice to create predictive playlists

Emotional Diversification: Gradually expand beyond comfort songs to build resilience and prevent over-dependency

Archive & Rediscover: Create systems to revisit forgotten tracks from your 1,000-song catalog

**REFERENCEE:** The data for this project was obtained from Kaggle.com