**Spotify Top 1000 Tracks Analysis Report**

**🌐 Project Overview**

This project analyzes the **Top 1000 Most Played Spotify Songs of All Time**, exploring trends in popularity, artist performance, sentiment of track titles, duration distribution, and release patterns. The goal is to uncover meaningful insights for playlist curators, record labels, and music marketers.

**🔗 Data Source**

* The dataset was obtained from Kaggle: [Top 1000 Most Played Spotify Songs of All Time](https://www.kaggle.com/datasets/kunalgp/top-1000-most-played-spotify-songs-of-all-time?resource=download).
* It includes fields such as track\_name, artist, album, release\_date, popularity, duration\_min, and spotify\_url.

**⚖️ Libraries & Tools Used**

* pandas – Data manipulation and cleaning
* matplotlib & seaborn – Visualization
* vaderSentiment – Sentiment analysis on track titles
* wordcloud – Word cloud generation for sentiment-based title analysis
* Python – Primary programming language

**🎓 Data Cleaning Summary**

* **Duplicates removed** based on unique id
* **Invalid release dates** (23 entries) handled with datetime conversion and exclusion from time-based charts
* **Outliers removed** from duration\_min using the 1st and 99th percentiles, leaving 980 valid entries

**📊 Key Insights**

**1. Data Cleaning & Structure**

* After removing duplicates and outliers, **980 valid tracks** remained for analysis.
* Some missing or invalid dates were excluded from time-based insights, but core features (popularity, artist, duration) were preserved.

**2. Popularity Trends**

* The most popular tracks (popularity score 90+) include artists like **Kendrick Lamar**, **The Weeknd**, and **Gracie Abrams**.
* **Popularity distribution** is right-skewed: most songs have a popularity score between **30 and 80**, with a few extreme hits.

**3. Artist Analysis**

* **The Weeknd**, **Drake**, and **Taylor Swift** dominate in the number of tracks on the list.
* **Childish Gambino** and **Taylor Swift** appear among artists with longer average track durations.
* The most **popular artists** (by average score) include those with fewer, high-performing songs.

**4. Time-Based Analysis**

* There's an upward trend in **track releases from 2015 onward**, with a peak around 2020–2022.
* **Average popularity over the years** shows minor fluctuations but suggests recent songs maintain consistent listener interest.
* Early 2025 tracks are predominantly released between **January and April**.

**5. Track Duration**

* Most tracks fall within **2.5 to 4 minutes**, aligning with mainstream radio standards.
* Longest tracks: *After Hours* (The Weeknd), *Enchanted* (Taylor Swift).
* Shortest tracks: *Walls Could Talk* (Halsey), *Hope* (XXXTENTACION).
* There's **no strong correlation** between popularity and track duration, meaning both short and long tracks can succeed.

**6. Album-Level Insights**

* Albums like **"Starboy"**, **"Dreamland"**, and **"Encore"** have multiple top tracks.
* Albums with high **average popularity** tend to feature fewer but highly successful songs (e.g., *Indigo*, *Planet Her*).

**7. Sentiment Analysis on Track Titles**

* Majority of track titles have **positive or neutral sentiment**.
* Positive words: *Love*, *Good*, *Dream*, percentage of 19.6
* Negative words: *Cry*, *Die*, *Alone*, percentage of 14.3
* Neutral was predominant with a percentage of 66.1
* **Positive titles** tend to have **slightly higher average popularity**.

**8. Sentiment Over Time & Artists**

* Sentiment trends show a **balanced mix** of emotional tones across years.
* Artists like **The Weeknd** and **Drake** show a tendency toward **more emotional or negative-toned titles**.
* Positive sentiment artists include **Dua Lipa** and **Shawn Mendes**.

**✅ Recommendations**

**🔁 For Curators & Playlist Designers:**

* Focus on tracks **2.5–4 minutes long**, which align with typical listener preferences.
* Feature **positive sentiment songs** and upbeat titles during promotional cycles.

**🎙️ For Artists & Labels:**

* Shorter songs don’t necessarily reduce popularity—**experiment with interludes** or short emotional tracks.
* Releasing multiple hits from a single album can enhance album loyalty. Examples like *Starboy* and *Encore* validate this strategy.

**📅 For Release Strategy:**

* Monthly trends (Jan–Apr 2025) show strong early-year releases. **Q1 drops** might benefit from less competition and higher visibility.

**🧠 For Data Teams & Marketing:**

* Use sentiment insights to **segment marketing content** (e.g., promote sadder tracks for late night or fall playlists).
* Combine release year + sentiment + artist to build **emotionally themed playlists** (e.g., "2023 Sad Vibes" or "2020 Feel-Good Hits").

**⚠️ Challenges & Resolutions**

* **Challenge:** Missing or invalid release dates disrupted time-based visuals
  + **Solution:** Converted dates using pd.to\_datetime and excluded invalid entries
* **Challenge:** heatmap initially displayed all months
  + **Solution:** Filtered release\_month dynamically to only display months present in the dataset
* **Challenge:** Seaborn palette warning due to improper usage
  + **Solution:** Switched to compatible built-in palettes like mako and crest

**Prepared by:** Nwaukwa Gracious

**Notebook:** SPORTIFY\_ANALYSIS.ipynb