

# ■ Enhanced Data Analysis Report

## Dataset: Spotify History

Generated: 2025-09-14 22:00:00

Domain: \*\* \*\*Music Streaming\*\*

## ■ Executive Summary

This report presents a comprehensive analysis of the spotify history dataset. The dataset contains 148,675 records with 11 features, including 1 numeric and 8 categorical variables. The analysis identified this as a \*\* \*\*Music Streaming\*\* domain dataset.

## ■ Dataset Overview

- Total Records: 148,675
- Total Features: 11
- Numeric Features: 1
- Categorical Features: 8
- Missing Values: 0
- Duplicate Records: 0

## ■ Feature Analysis

name	dtype	num_missing	num_unique	mean	std
spotify_track_uri	object	0	16527	nan	nan
ts	object	0	140422	nan	nan
platform	object	0	6	nan	nan
ms_played	int64	0	43082	128316.63509275323	117840.06033150152
track_name	object	0	13839	nan	nan
artist_name	object	0	4113	nan	nan
album_name	object	0	7948	nan	nan
reason_start	object	143	13	nan	nan
reason_end	object	117	15	nan	nan
shuffle	bool	0	2	0.7445816095022021	0.4360975869593898

## ■ Generated Analysis Questions

The following questions were automatically generated to guide the analysis:

1. What is the distribution of `ms\_played` across different platforms?
2. Is there a correlation between the `track\_name` and the `platform`?
3. Does the `ms\_played` correlate with the `track\_name` or other features?
4. Are there any relationships between the `spotify\_track\_uri` and other columns like `platform` or `ms\_played`?
5. What are the most frequent values of `ms\_played` and how do they compare across platforms?
6. Are there any `spotify\_track\_uri` values that stand out significantly in terms of their `ms\_played` or other features?
7. Are there any patterns in the `platform` column that might suggest unusual activity or data entry errors?
8. Can we predict the `ms\_played` for a new track based on its `track\_name`, `platform`, and other features?

## ■ User Analysis Sessions

The following 2 questions were asked by the user with AI-generated insights:

### **Q1. [2025-09-14 21:58] Identify the Pattern of music listening on weekdays and weekends.**

Answer:

Comparison Analysis: ms\_played by ts • Highest average: 2021-03-04 00:27:48 (1561125.00) • Lowest average: 2017-08-27 00:00:23 (0.00) • Significant variation across ts groups Insight: This data shows that there's a significant difference in the amount of music played each day of the week. The highest amount of music was played on March 4th, 2021, while the lowest amount was played on August 27th, 2017. This suggests that there might be a pattern to music listening habits, potentially influenced by factors like time of day or day of the week.

Key Finding: ■ Comparison Analysis: ms\_played by ts • Highest average: 2021-03-04 00:27:48 (1561125.00) • Lowest ...

■ Included data visualization

### **Q2. [2025-09-14 21:59] Identify top 5 artists**

Answer (Used 1 previous insights):

■ AI Analysis: 1. \*\*The top 5 artists based on `ms\_played` are The Beatles (13530), The Killers (6748), John Mayer (4797), Lana Del Rey (285386), and Calvin Harris (61865).\*\* 2. This analysis reveals the most popular artists based on the total number of minutes played, with The Beatles and The Killers being the top two. 3. \*\*High average `ms\_played` suggests a strong preference for these artists, indicating potential for further analysis of their popularity and impact.\*\* Let me know if you'd like me to analyze other aspects of the data!

Key Finding: ■ AI Analysis: 1

■ Included data visualization

## ■ Analysis Insights

No analysis insights available yet. Run some analysis queries to generate insights.

## ■ Recent Visualizations

Found 2 recent visualizations:

- spotify\_history\_20250914\_215918.html (Created: 2025-09-14 21:59:18)
- spotify\_history\_20250914\_215806.html (Created: 2025-09-14 21:58:06)

## ■ Statistical Summary

*Key Statistics for Numeric Features:*

Feature	Mean	Std	Min	Max
ms_played	128060.68	117817.77	0.00	1561125.00

---

Report generated by AI Data Analysis Agent on 2025-09-14 22:00:00