
Proyecto Hipótesis: Análisis de datos “Plataformas de Música”

Betsabé Morales
Minerva Bobadilla
Latam
15-07-2024

Caso

De análisis

En este proyecto se centró en un análisis de datos para diversas plataformas de Música tales como “Spotify, Apple, Deezer y Shazam”.

En un mundo en el que la industria musical es extremadamente competitiva y está en permanente evolución, la capacidad de tomar decisiones basadas en datos se ha convertido en un activo invaluable.

En este contexto, una discográfica se enfrenta al emocionante desafío de lanzar un nuevo artista en el escenario musical global. Afortunadamente, cuenta con una herramienta poderosa en su arsenal: un extenso dataset de Spotify con información sobre las canciones más escuchadas en 2023.

Por lo anterior el objetivo principal es Validar (refutar o confirmar) estas hipótesis mediante el análisis de los datos, y proporcionar recomendaciones estratégicas basadas en los hallazgos. En última instancia, el objetivo es que la discográfica y el nuevo artista puedan tomar decisiones informadas que aumenten sus posibilidades de conseguir el “éxito”.

Periodo Analizado 2023.

Hipótesis: “Plataformas Musicales”

Para el estudio se cuenta con una base de datos de 950 tracks, la cual se procesó con técnica de Análisis de Datos y desarrollo la metodología de segmentación por cuartiles para segmentar por características técnicas.

950

total track

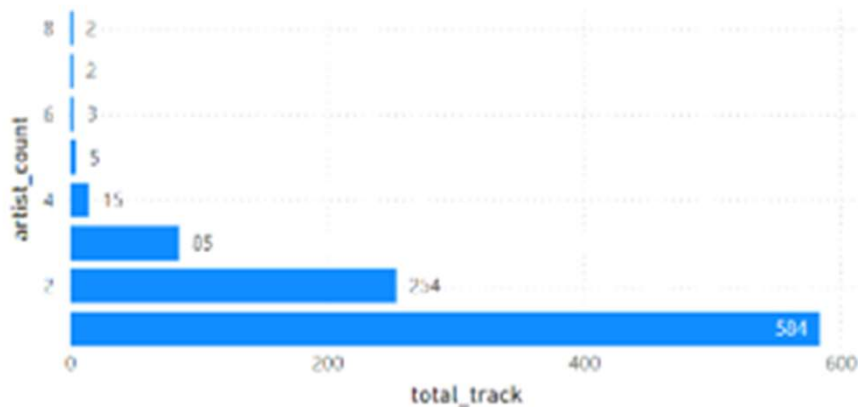
1480

Suma de artist_count

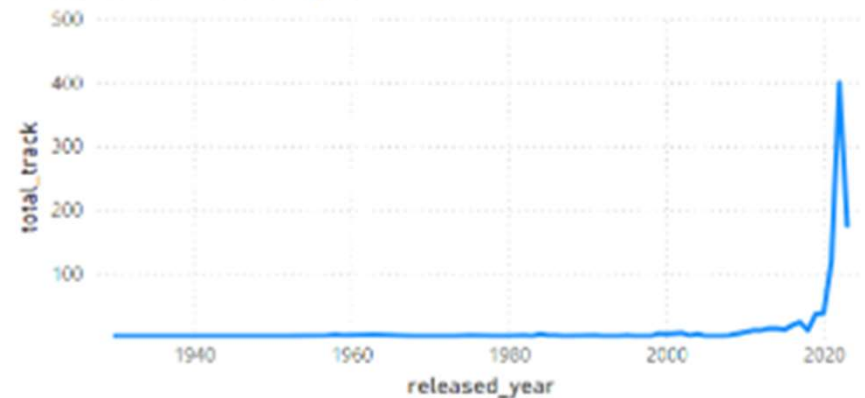
5 mill.

Suma de total playlists

total_track por artist_count



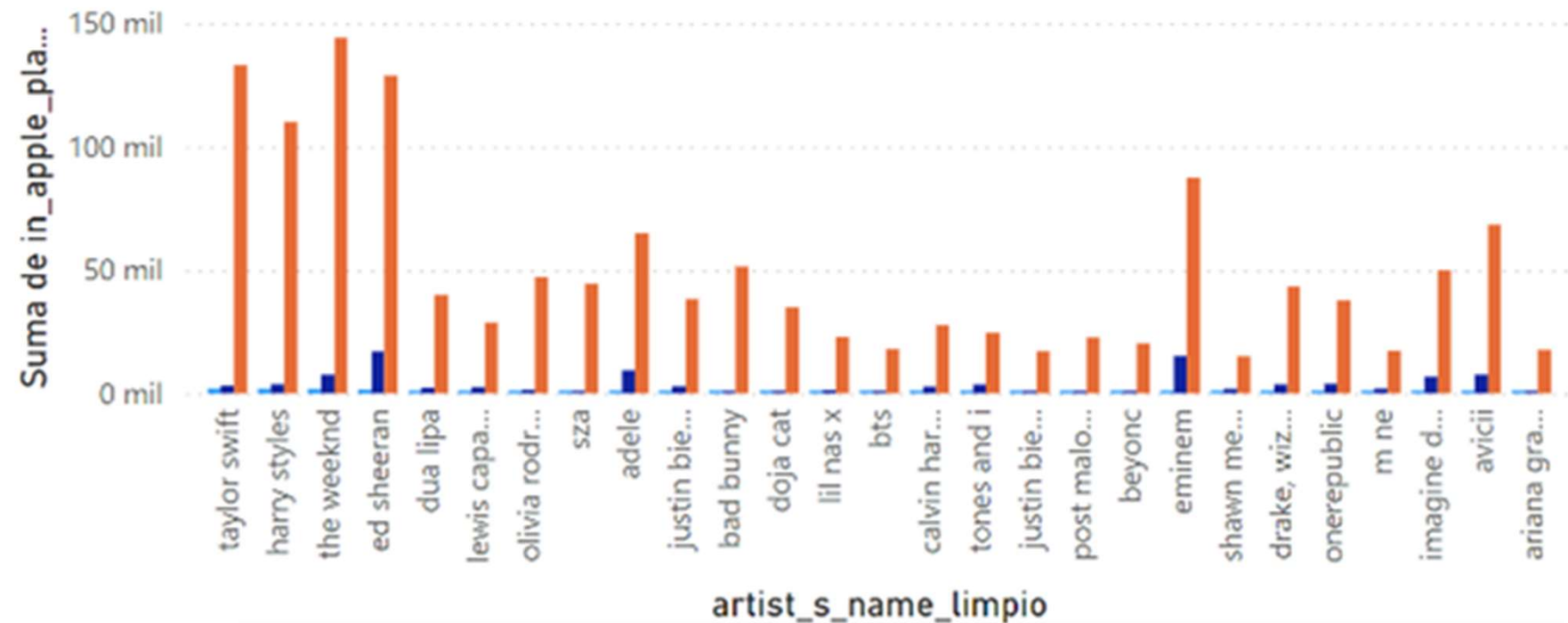
total_track por released_year



Hipótesis: “Plataformas Musicales”

Suma de in_apple_playlists, Suma de in_deezer_playlists y Suma de spotify_playlists por artist_s_name_limpio

● Suma de in_apple_playlists ● Suma de in_deezer_playlists ● Suma de spotify_playlists



Hipótesis: “Plataformas Musicales”

Las Hipótesis entregadas por la discográfica y que se deben validar, para saber qué hace que una canción sea más escuchada son las siguientes:

- I) Las canciones con un mayor BPM (Beats Por Minuto) tienen más éxito en términos de cantidad de streams en Spotify.
 - II) Las canciones más populares en el ranking de Spotify también tienen un comportamiento similar en otras plataformas como Deezer.
 - III) La presencia de una canción en un mayor número de playlists se relaciona con un mayor número de streams.
 - IV) Los artistas con un mayor número de canciones en Spotify tienen más streams.
 - V) Las características de la canción influyen en el éxito en términos de cantidad de streams en Spotify.
-

Hipótesis: “Plataformas Musicales”

I) Las canciones con un mayor BPM (Beats Por Minuto) tienen más éxito en términos de cantidad de streams en Spotify.

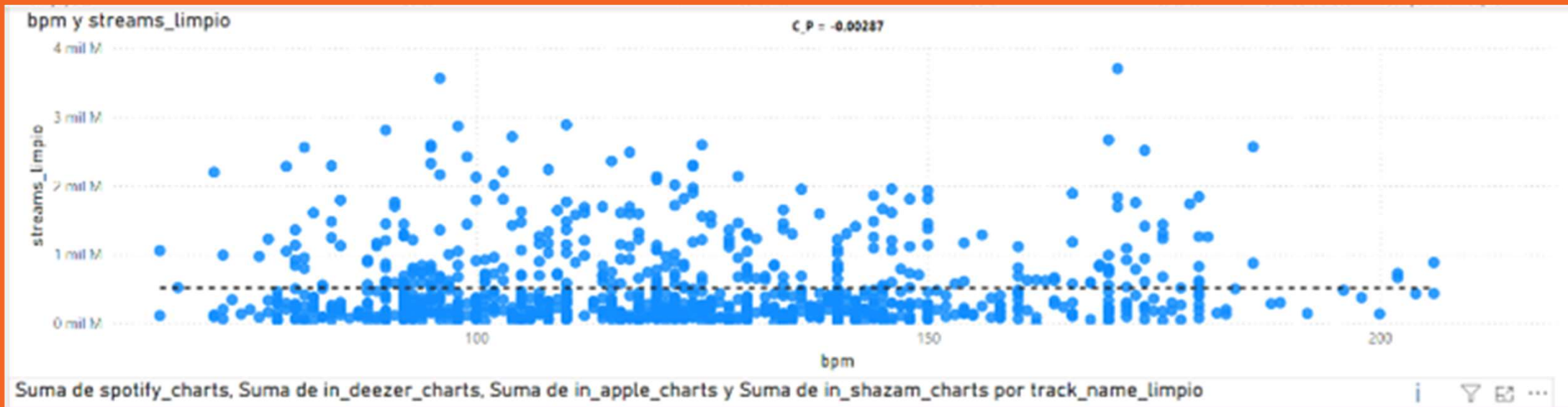


Figura N° 1: bpm vs streams_limpio

Hipótesis: “Plataformas Musicales”

II) Las canciones más populares en el ranking de Spotify también tienen un comportamiento similar en otras plataformas como Deezer.

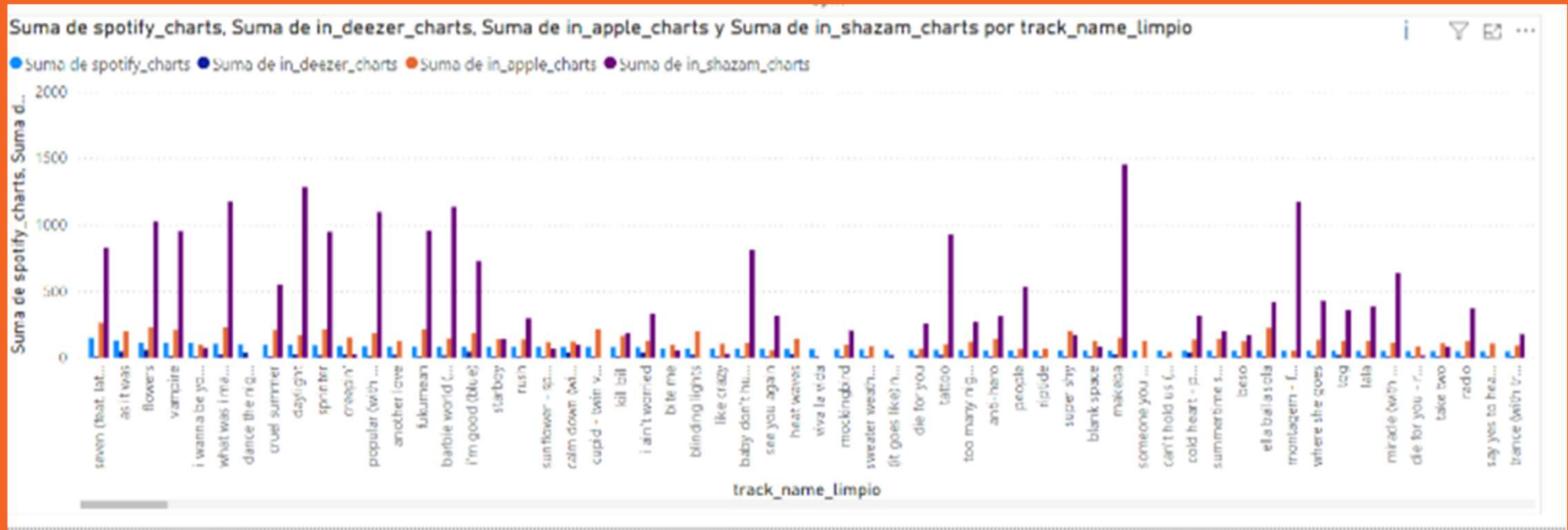


Figura N° 2: track_name_limpio vs suma_plataforma_charts

Hipótesis: “Plataformas Musicales”

III) La presencia de una canción en un mayor número de playlists se relaciona con un mayor número de streams.

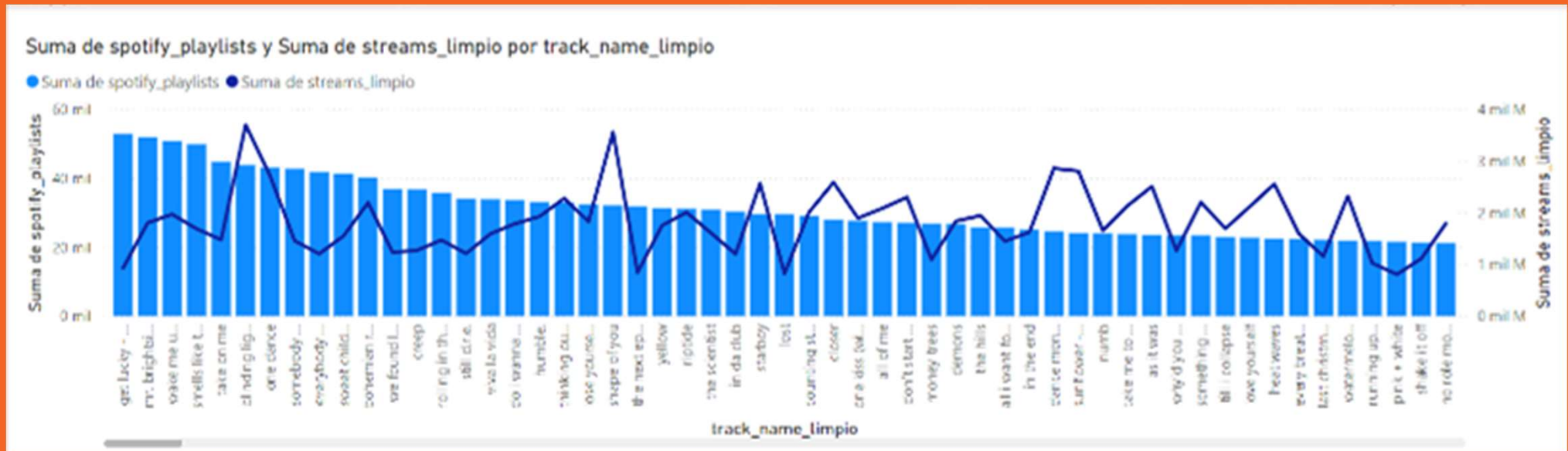


Figura N° 3: track_name_limpio vs suma_spotify_playlists, suma_streams_limpio.

Hipótesis: “Plataformas Musicales”

IV) Los artistas con un mayor número de canciones en Spotify tienen más streams.

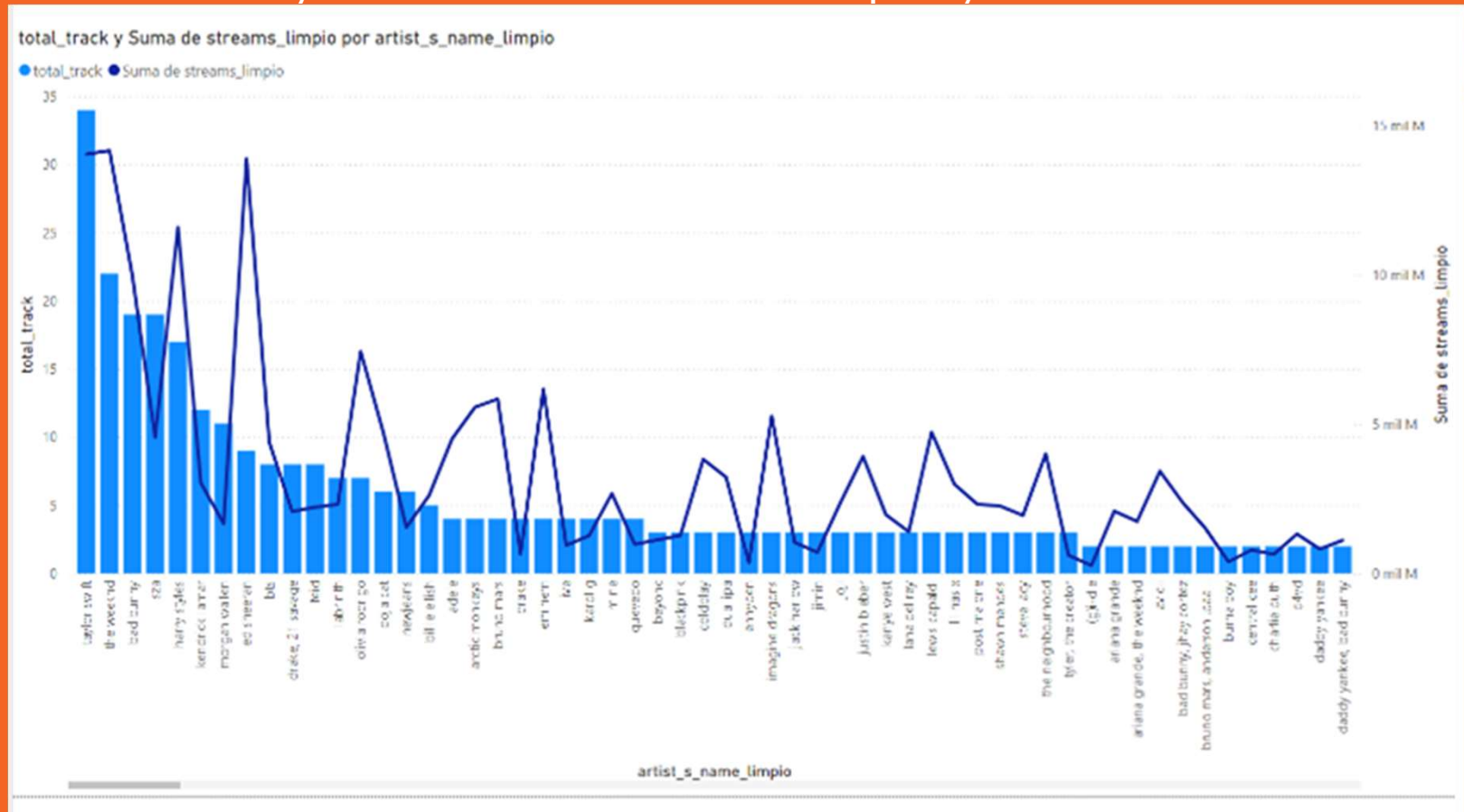


Figura N° 4: artist_name_limpio vs total_tracks, suma_streams_limpio.

Hipótesis: “Plataformas Musicales”

V) Las características de la canción influyen en el éxito en términos de cantidad de streams en Spotify.

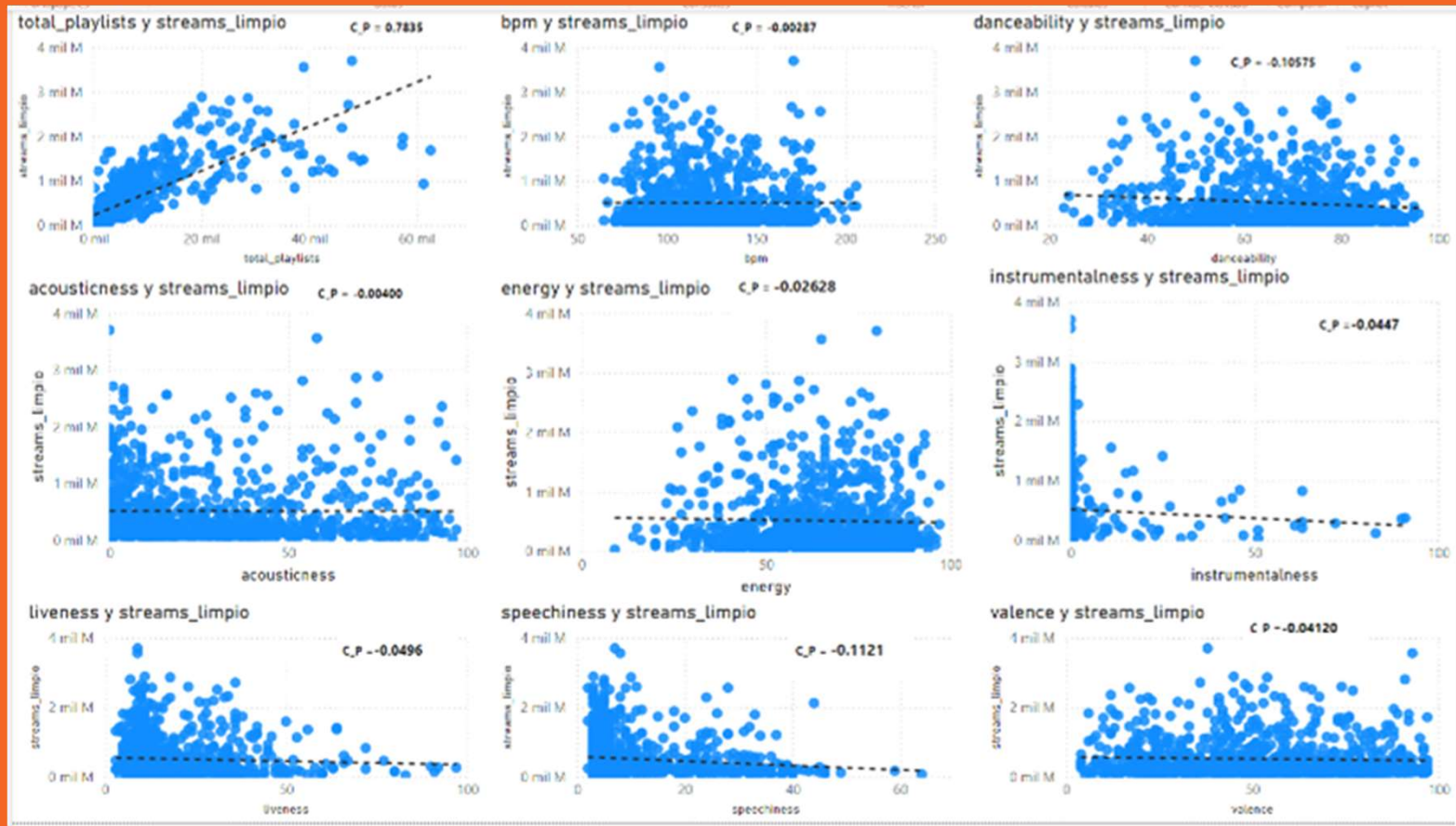


Figura N° 5: característica técnica vs streams _limpio; total_playlists vs streams_limpio.

Conclusiones

Durante el desarrollo del proyecto se utilizó el proceso de Análisis de Datos y la metodología de segmentación (Cuartiles). El Periodo analizado fue el año 2023, la discográfica planteó una serie de hipótesis, sobre qué hace que una canción sea más escuchada. Estas hipótesis se detallan a continuación y de esto se puede concluir lo siguiente:

I) Las canciones con un mayor BPM (Beats Por Minuto) tienen más éxito en términos de cantidad de streams en Spotify.

Se refuta esta hipótesis ya que a mayor BPM no hay mayor cantidad de streams, por lo tanto las canciones de mayor bpm no son las que tienen mas cantidad de streams. Se deduce de gráfica de dispersión Figura N° 1: bpm vs streams_limpio. Esto también se ve reflejado en el valor de la correlación de Pearson = -0,002 lo cual es un valor alejado del 1 para el ajuste perfecto.

Conclusiones

II) Las canciones más populares en el ranking de Spotify también tienen un comportamiento similar en otras plataformas como Deezer.

Se refuta esta hipótesis ya que, las canciones populares en ranking de Spotify no presentan el mismo comportamiento de éxito en las diferentes plataformas. Ver gráfica Figura N° 2: track_name_limpio vs suma_plataforma_charts.

III) La presencia de una canción en un mayor número de playlists se relaciona con un mayor número de streams.

Se refuta la hipótesis ya que no se puede comprobar dicho comportamiento, dado que el contar con un mayor número de playlists por canción no es una relación directamente proporcional al número de streams, esto se verifica en gráfico de Figura N° 3: track_name_limpio vs suma_spotify_playlists, suma_streams_limpio.

Conclusiones

IV) Los artistas con un mayor número de canciones en Spotify tienen más streams. Esta hipótesis se refuta con la gráfica de `artist_name_limpio` vs `total_tracks`, `artist_name_limpio` vs `suma_de_streams_limpio`, quien tiene mas canciones no son quienes presentan mayor cantidad de streams ver ejemplo de Ed Sheeran vs Sza. ver grafica Figura N° 4: `artist_name_limpio` vs `total_tracks`, `suma_streams_limpio`.

V) Las características de la canción influyen en el éxito en términos de cantidad de streams en Spotify.

Esta hipótesis es refutada ya que las características técnicas como `bpm`, `acousticness`, `danceability`, `energy`, `instrumentalness`, `liveness`, `speechiness` y `valence` todas presentan un bajo valor en su correlación de Pearson respectivamente, ver gráficos Figura N° 5: característica técnica vs `streams_limpio`; `total_playlists` vs `streams_limpio`, excepto la correlación entre `total_playlists` vs `streams_limpio` que es la más cercana a 1 con un valor de 0,7835, sin embargo no es una característica técnica.

Recomendaciones

Dado que todas las hipótesis fueron refutadas, sin dar una característica a sobre qué hace que una canción sea la más escuchada, entonces se solicita entregar más información: de género de la canción, datos en el tiempo (streams, datos técnicos, datos demográficos de oyentes, etc), para lograr un análisis que realmente pueda indicar el éxito de una canción y entregar una conclusión certera y adecuada.

Muchas gracias
