



Spotify®

# AGENDA

1. Contexto y objetivos
2. Hipótesis y preguntas de interés
3. Datos
4. Análisis Exploratorio
5. Insights y recomendaciones



# CONTEXTO Y OBJETIVOS

---

**Spotify** es un servicio de streaming de **música en línea** que permite a los usuarios acceder a una amplia variedad de canciones, álbumes y artistas de todo el mundo. Los usuarios pueden crear sus propias listas de reproducción personalizadas, compartir música con amigos y seguir a sus artistas favoritos. Además, Spotify también ofrece contenido de podcast y otros formatos de audio. Fue fundada en Suecia en 2006 y actualmente es una de las plataformas de música en línea más populares a nivel mundial.

Con el objetivo de entender un poco más sobre la industria musical, presentamos un recorrido por los datos desde el año 1900 hasta el año 2021, con el fin de identificar patrones de comportamiento de los usuarios y de la música a lo largo de este siglo.



## Audiencia

Este análisis intenta contestar cualquier serie de preguntas que puedan surgir a partir del párrafo anterior, por lo cuál puede ser de utilidad para cualquier persona que esté interesado en saber un poco más sobre la industria musical, específicamente en spotify.

## Limitaciones

La limitación que nos presenta el siguiente análisis es que los datos no están actualizados hasta el presente año.



# HIPÓTESIS Y **PREGUNTAS DE INTERÉS**

---



A partir del contexto y los objetivos anteriores, se resaltan las siguientes preguntas:

- ¿Cuánto ha cambiado la producción de música a lo largo de los años?
- ¿Cuáles son las canciones y artistas más populares?
- Analizar las características de las canciones que son más o menos populares para comprender mejor lo que atrae a los usuarios, como el tempo, la energía, la instrumentalización, etc.

**DATOS**

---





Artistas /  
colaboraciones

**115062**



Cant. canciones

**586601**



Popularidad Media

**27,57**

Fechas:

**Mínima 1900-01-01**

**y**

**Máxima 2021-04-16**

El dataset (df) no  
contiene valores  
nulos ni duplicados

Eliminamos outliers inferiores,  
basándonos en la popularidad  
de la canciones. Por otro lado  
se mantienen los límites  
superiores ya que nos indican  
las canciones que realmente  
son o fueron populares.

## Columnas del data set:

Data columns (total 20 columns):

| #  | Column           | Non-Null Count  | Dtype          |
|----|------------------|-----------------|----------------|
| 0  | id               | 586601 non-null | object         |
| 1  | name             | 586601 non-null | object         |
| 2  | popularity       | 586601 non-null | int64          |
| 3  | duration_ms      | 586601 non-null | int64          |
| 4  | explicit         | 586601 non-null | int64          |
| 5  | artists          | 586601 non-null | object         |
| 6  | id_artists       | 586601 non-null | object         |
| 7  | release_date     | 586601 non-null | datetime64[ns] |
| 8  | danceability     | 586601 non-null | float64        |
| 9  | energy           | 586601 non-null | float64        |
| 10 | key              | 586601 non-null | int64          |
| 11 | loudness         | 586601 non-null | float64        |
| 12 | mode             | 586601 non-null | int64          |
| 13 | speechiness      | 586601 non-null | float64        |
| 14 | acousticness     | 586601 non-null | float64        |
| 15 | instrumentalness | 586601 non-null | float64        |
| 16 | liveness         | 586601 non-null | float64        |
| 17 | valence          | 586601 non-null | float64        |
| 18 | tempo            | 586601 non-null | float64        |
| 19 | year             | 586601 non-null | int64          |

dtypes: datetime64[ns](1), float64(9), int64(6), object(4)

## Informacion del dataset:

|                  | count    | mean          | std           | min    | 25%         | 50%           | 75%          | max         |
|------------------|----------|---------------|---------------|--------|-------------|---------------|--------------|-------------|
| popularity       | 586801.0 | 27.573212     | 18.369417     | 0.0    | 13.0000     | 27.000000     | 41.00000     | 100.000     |
| duration_ms      | 586801.0 | 230054.852626 | 126532.824981 | 3344.0 | 175083.0000 | 214907.000000 | 263887.00000 | 5621218.000 |
| explicit         | 586801.0 | 0.044091      | 0.205298      | 0.0    | 0.0000      | 0.000000      | 0.00000      | 1.000       |
| danceability     | 586801.0 | 0.563612      | 0.166101      | 0.0    | 0.4530      | 0.577000      | 0.68600      | 0.991       |
| energy           | 586801.0 | 0.542071      | 0.251910      | 0.0    | 0.3430      | 0.549000      | 0.74800      | 1.000       |
| key              | 586801.0 | 5.221594      | 3.519420      | 0.0    | 2.0000      | 5.000000      | 8.00000      | 11.000      |
| loudness         | 586801.0 | -10.205789    | 5.089422      | -60.0  | -12.8910    | -9.242000     | -6.48100     | 5.376       |
| mode             | 586801.0 | 0.658797      | 0.474114      | 0.0    | 0.0000      | 1.000000      | 1.00000      | 1.000       |
| speechiness      | 586801.0 | 0.104870      | 0.179902      | 0.0    | 0.0340      | 0.044300      | 0.07630      | 0.971       |
| acousticness     | 586801.0 | 0.449803      | 0.348812      | 0.0    | 0.0969      | 0.422000      | 0.78400      | 0.996       |
| instrumentalness | 586801.0 | 0.113425      | 0.266843      | 0.0    | 0.0000      | 0.000024      | 0.00955      | 1.000       |
| liveness         | 586801.0 | 0.213933      | 0.184328      | 0.0    | 0.0983      | 0.139000      | 0.27800      | 1.000       |
| valence          | 586801.0 | 0.552306      | 0.257673      | 0.0    | 0.3480      | 0.564000      | 0.76900      | 1.000       |
| tempo            | 586801.0 | 118.467930    | 29.762942     | 0.0    | 95.6060     | 117.387000    | 136.32400    | 246.381     |
| year             | 586801.0 | 1988.571729   | 22.826243     | 1900.0 | 1974.0000   | 1992.000000   | 2007.00000   | 2021.000    |

# ANÁLISIS **EXPLORATORIO**

---

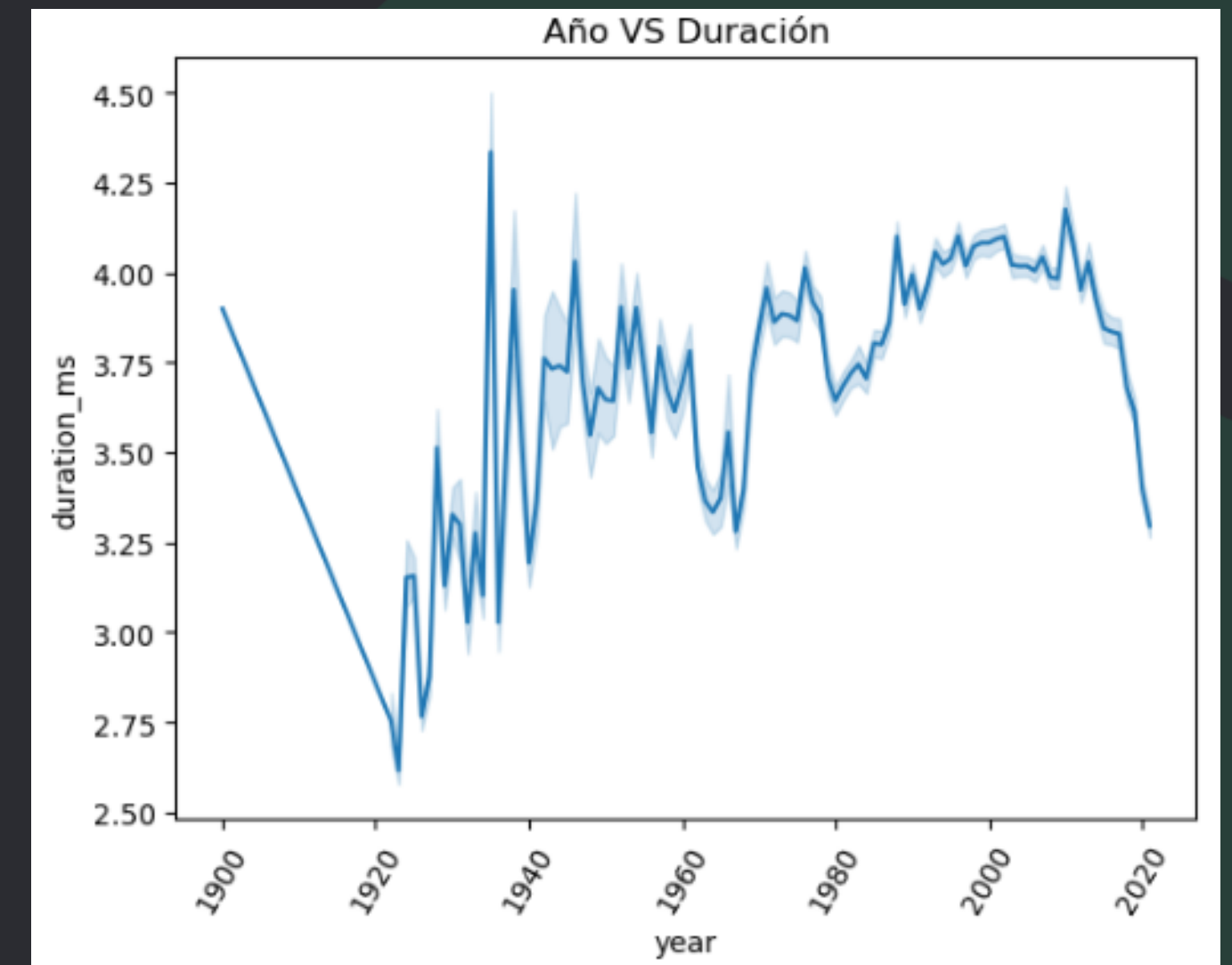
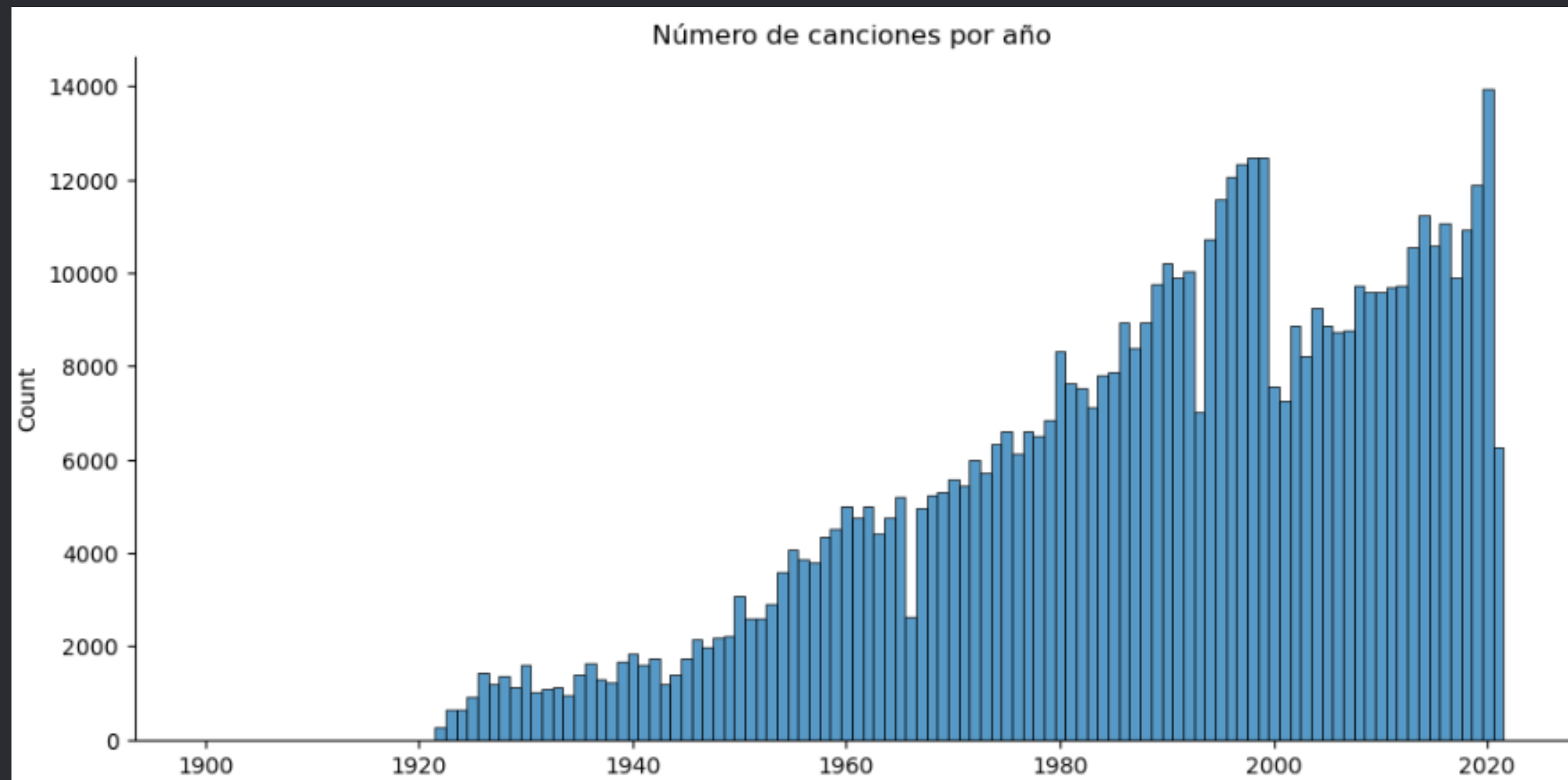
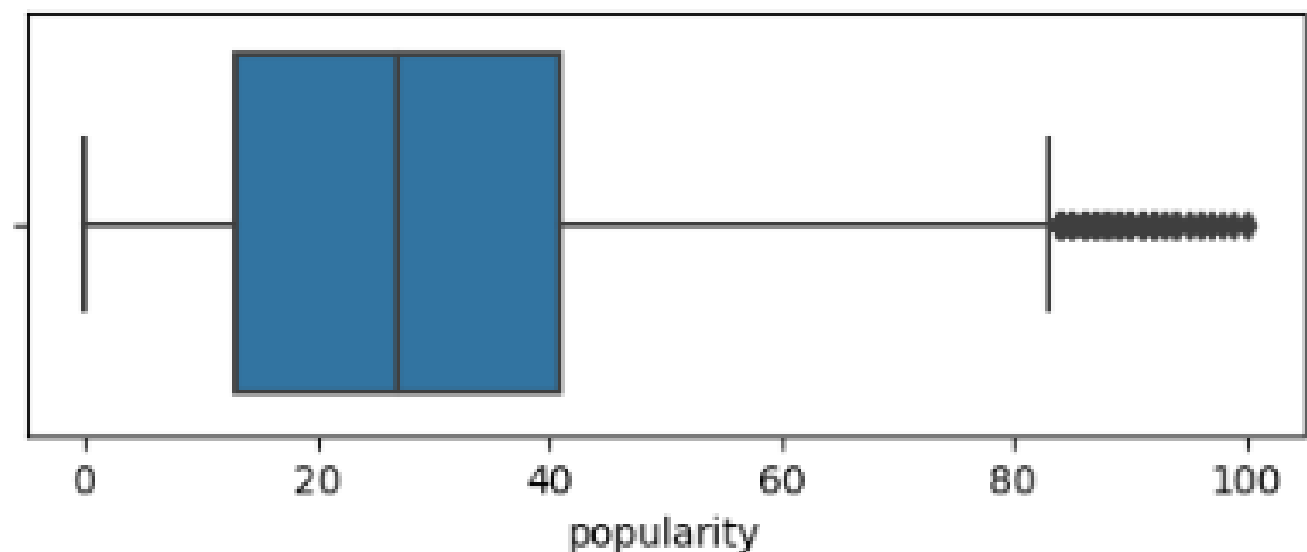
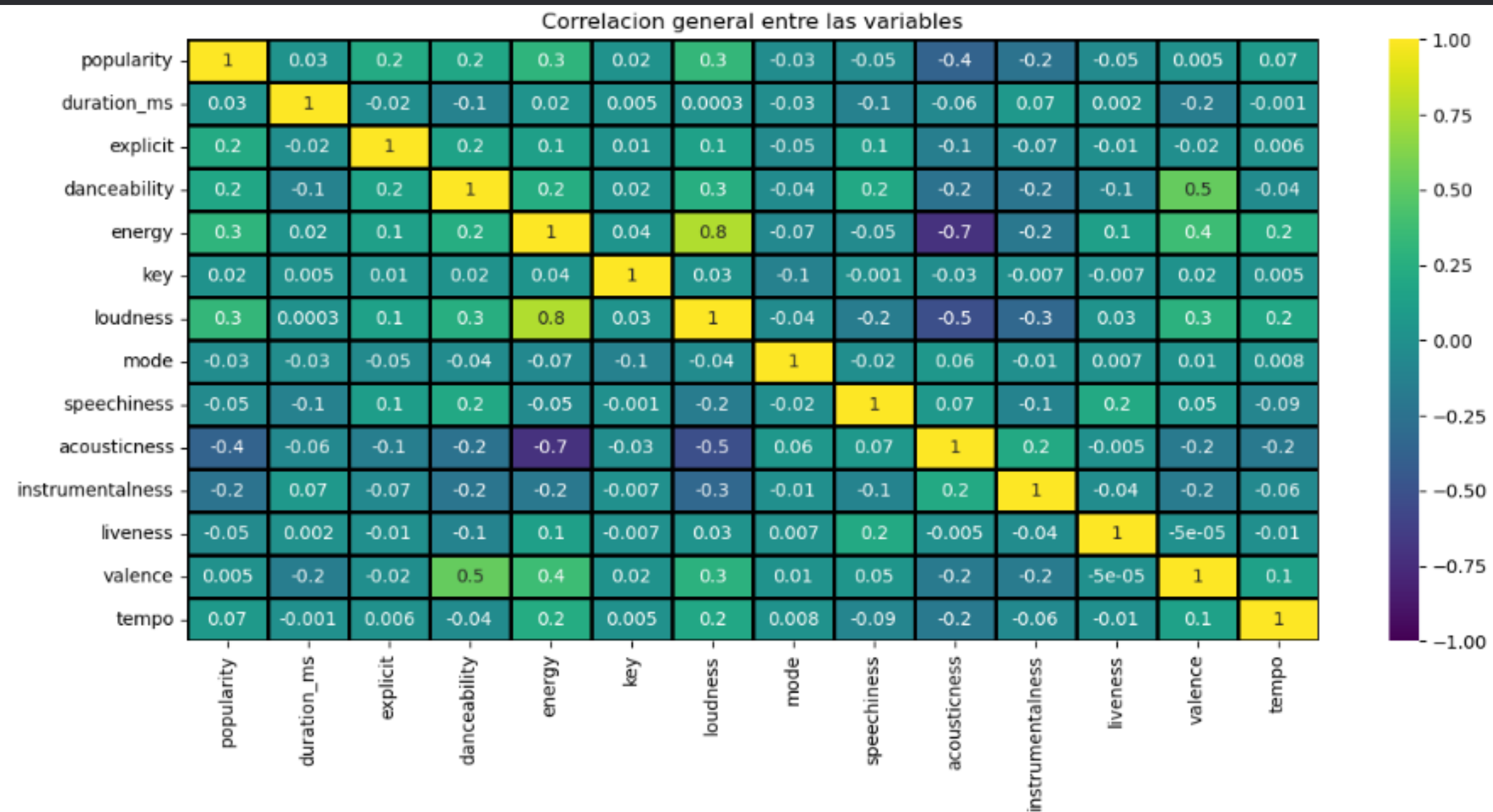


Gráfico general en base a la popularidad

<AxesSubplot:xlabel='popularity'>

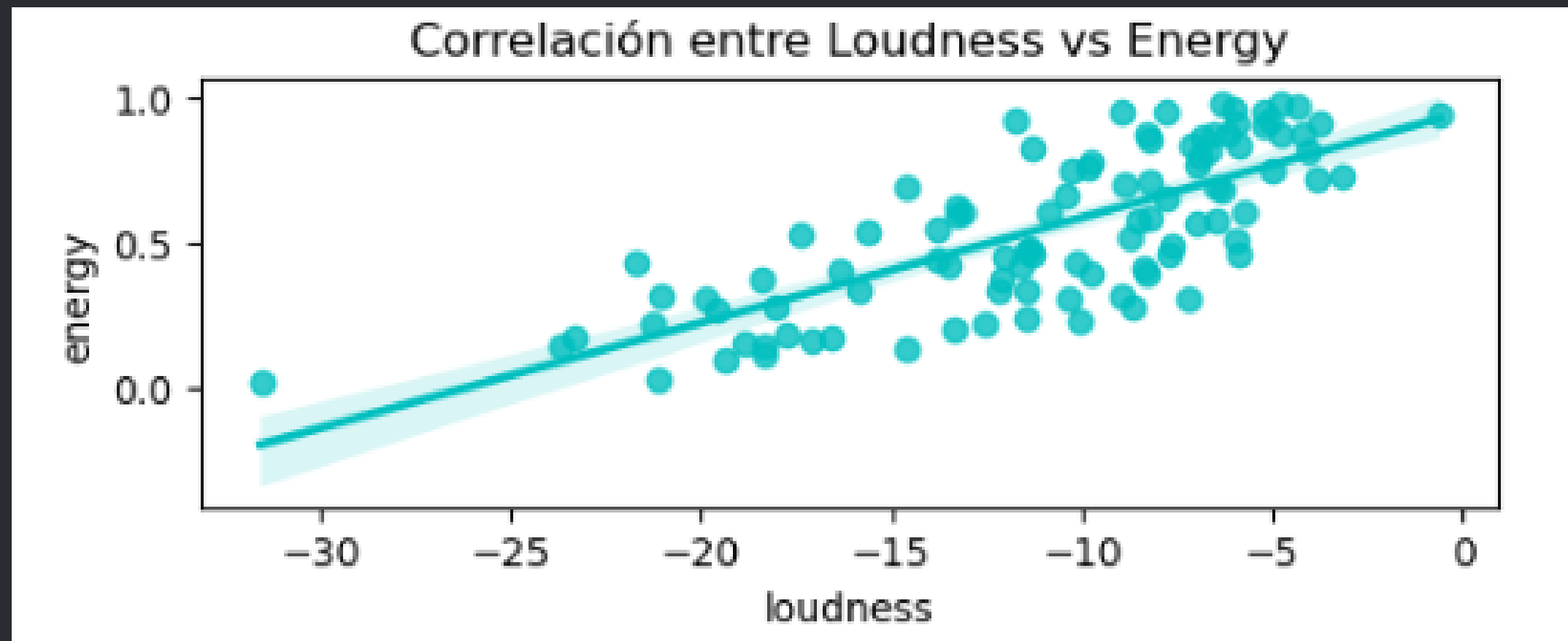


- En estos graficos podemos ver inicialmente un notable crecimiento en la cantidad de canciones desde 1900.
- 
- En siguiente gráfico a nivel general, podemos ver la variación de la duración de las canciones en minutos, desde los años de 1900, si bien podemos ver picos importantes, la tendencia es creciente y decae en el 2020 por la falta de datos.
- 
- para e último gráfico podemos ver que la media de la popularidad esta en 27.57 aproximadamente, por lo que gran parte de las canciones son poco populares.



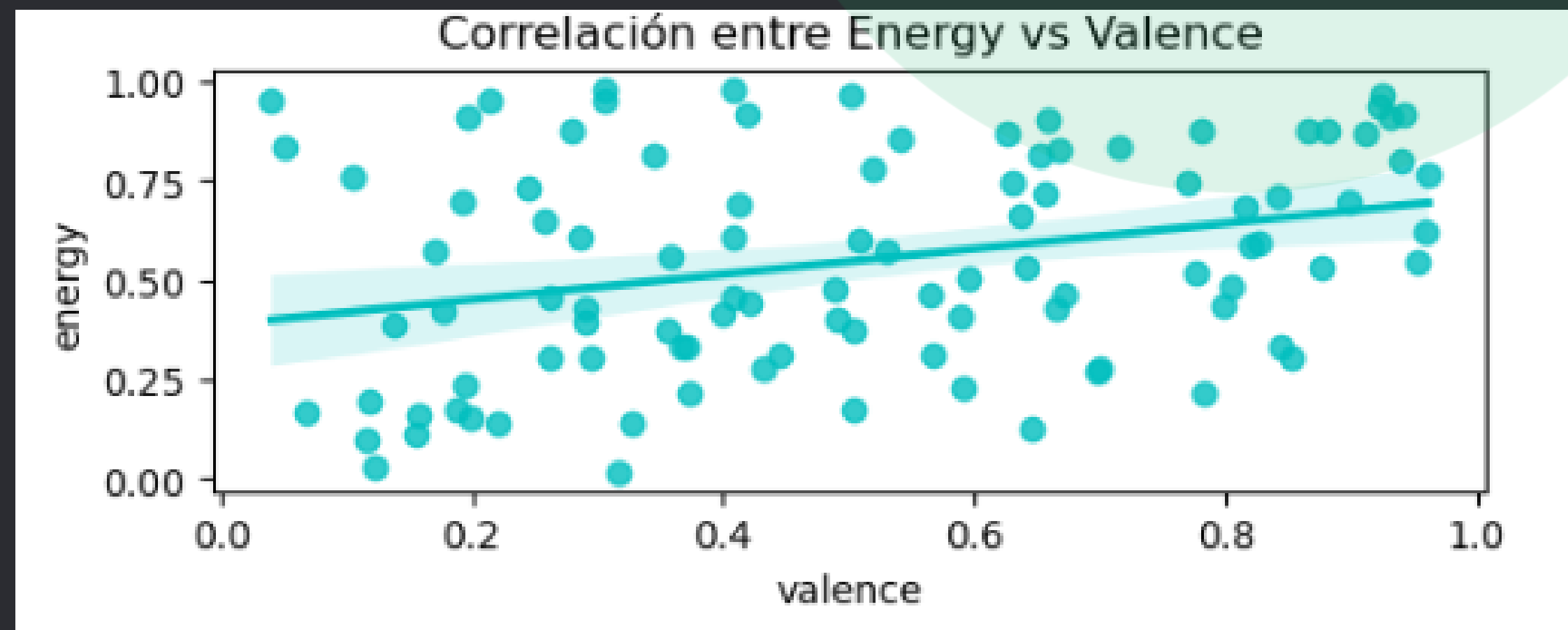
Podemos ver una mayor correlación entre las variables:

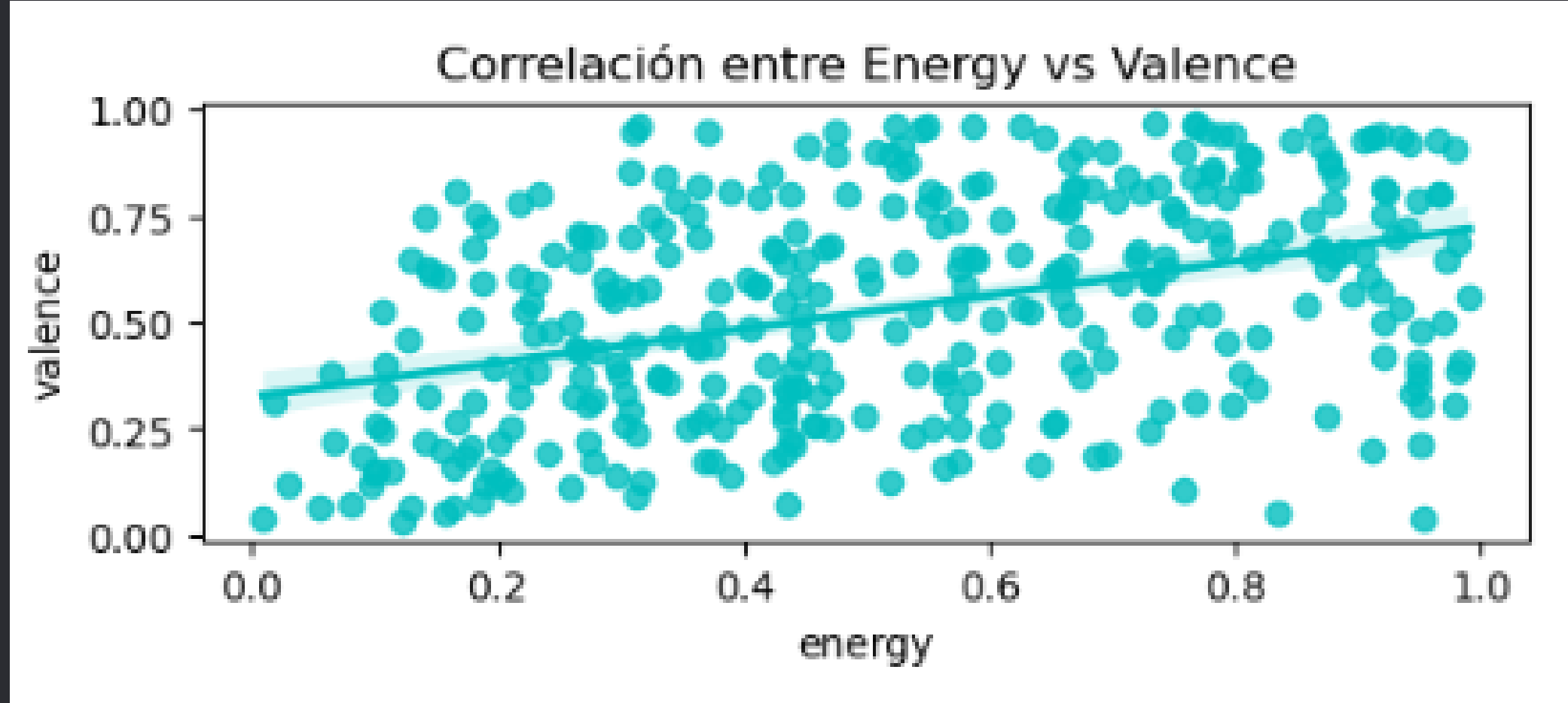
- Energy y loudness (siendo esta la más significativa)
- Danceability y valence
- Energy y valence



Se toma una muestra aleatoria de la base de datos principal, en este caso podemos ver que mientras más energía tiene la canción más ruidosa es esta, es decir, los niveles de sonoridad son más altos.

En este caso podemos ver que aunque los puntos estén dispersos, a nivel general hay una ligera relación entre "valence" que es el nivel de positividad o negatividad, siendo 1 el tope de positividad, y la energía de la canción. Lo que nos indica que no necesariamente las canciones energéticas transmiten positividad.



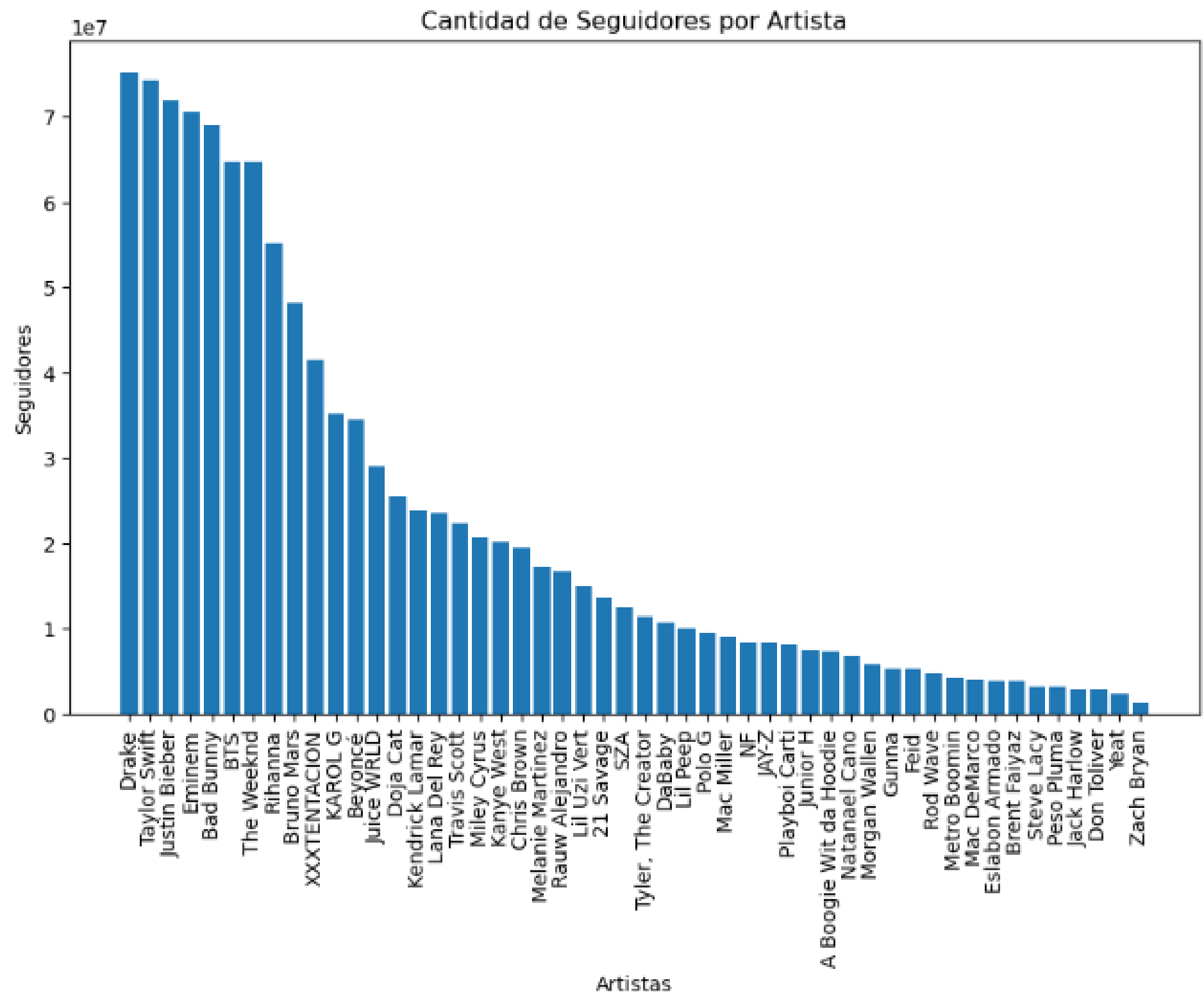


En este caso podemos ver que aunque los puntos estén dispersos, a nivel general hay una ligera relación entre "valence" que es el nivel de positividad o negatividad, siendo 1 el tope de positividad, y la energía de la canción. Lo que nos indica que no necesariamente las canciones energéticas transmiten positividad.





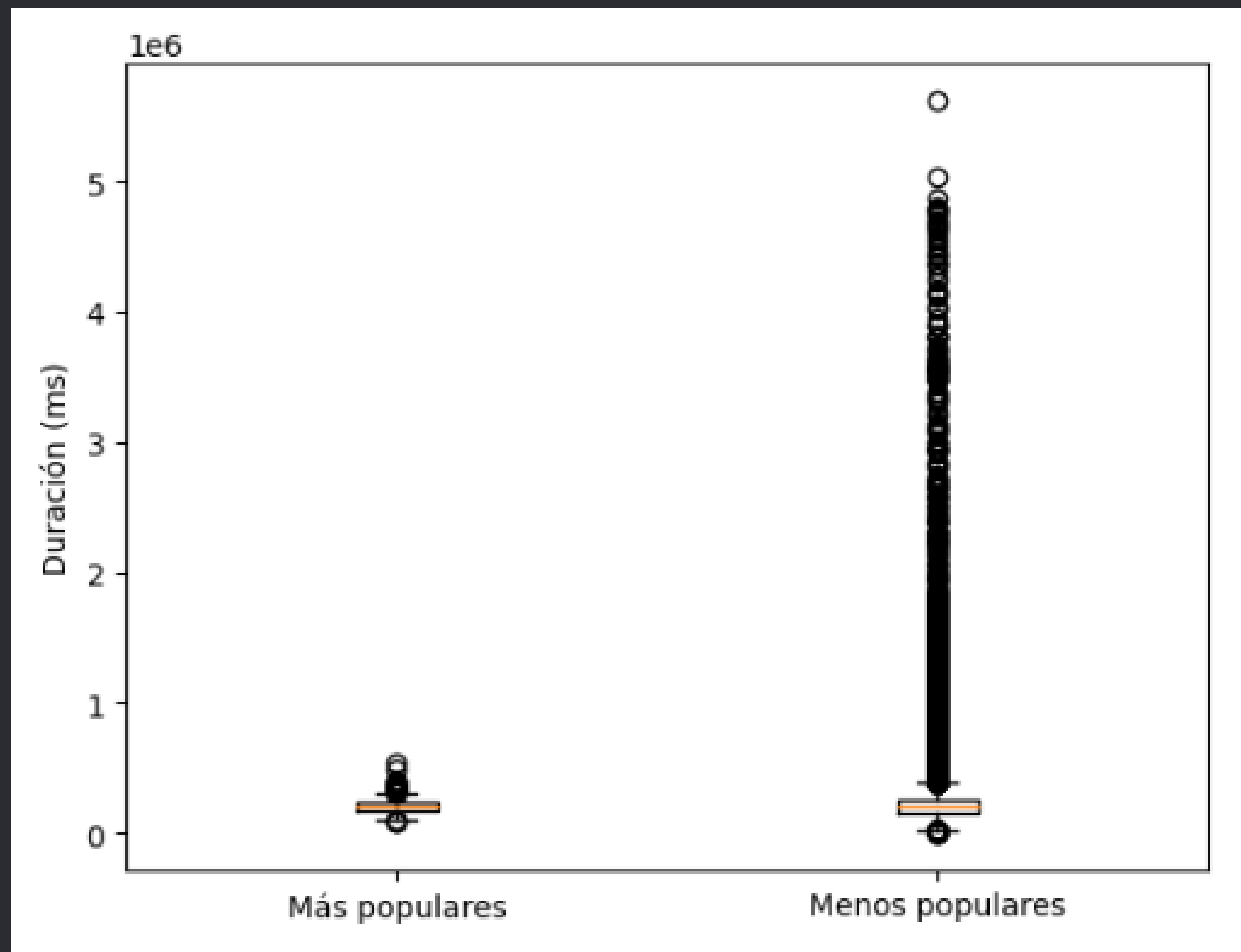
# Canciones y artistas más populares



Top 10 artistas 2021

|    | name          | popularity | genres   |
|----|---------------|------------|--|
| 0  | Taylor Swift  | 100        | [pop]  |
| 5  | Bad Bunny     | 97         | [reggaeton, trap latino, urbano latino]            |
| 3  | The Weeknd    | 96         | [canadian contemporary r&b, canadian pop, pop]     |
| 1  | Drake         | 96         | [canadian hip hop, canadian pop, hip hop, rap,...] |
| 8  | Peso Pluma    | 94         | [corrido, nueva musica mexicana, sad sierrano]     |
| 48 | Feid          | 92         | [colombian pop, pop reggaeton, reggaeton colomb... |
| 16 | Eminem        | 91         | [detroit hip hop, hip hop, rap]                    |
| 25 | KAROL G       | 91         | [reggaeton, reggaeton colombiano, trap latino,...] |
| 7  | SZA           | 91         | [pop, r&b, rap]                                    |
| 2  | Morgan Wallen | 91         | [contemporary country]                             |

Si bien en el primer gráfico podemos ver que Drake es artista con más seguidores, en el segundo gráfico vemos que Taylor Swift fue la artista más popular del 2021, o hasta la última fecha de recolección de datos. Lo que nos indica que no el mas popular tiene



En este gráfico podemos ver que las canciones mas populares tienen una duración bastante menor a las canciones con una popularidad menor al 80, teniendo en cuenta una escala de 0 a 100.

# INSIGHTS & **RECOMENDACIONES**

---

En línea general podemos ver que la energía de las canciones y la sonoridad o decibeles tienen una muy buena relación, por lo que se puede intuir que mayormente se busca hacer canciones con mucha energía y ruidos altos, lo podemos ver también que estos dos atributos tienen un número relevante en la popularidad de las canciones. Pero, por otra parte, podemos ver que mientras más energía tiene la canción menos acústica y menos ruidosas son estas, por lo tanto estos atributos, tanto la acústica y la instrumentación no son tan populares en el ranking, lo que nos llevaría a pensar que mayormente se buscan canciones con sonidos mas artificiales y electronicos.

Podemos ver también una relación bastante buena entre la bailabilidad de las canciones con el nivel de ánimo que transmiten estas.

También se hace notar la corta duracion de las canciones mas populares, lo que nos hace pensar en que mas larga la canción esta se vuelve menos popular.

