



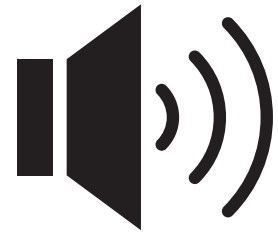
Exploratory Data Analysis

SPOTIFY

Lucas Zamora Vera

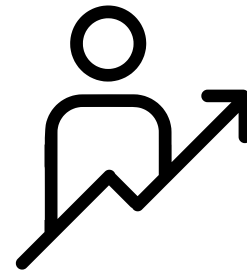
Datos Spotify

Canciones



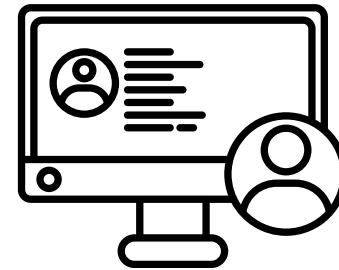
- Identificador único
- Artista
- Album
- Género

Popularidad



- Popularidad de la canción en base a los “likes” recibidos por los oyentes

Características



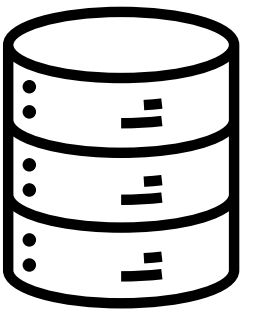
- Duración
- Tonalidad
- Clave
- Tempo
- Año de lanzamiento

Rasgos



- Explicidad
- Bailabilidad
- Energía
- Volumen
- Cantidad de palabra
- Instrumentación
- Nivel de directo
- Positividad

BBDD

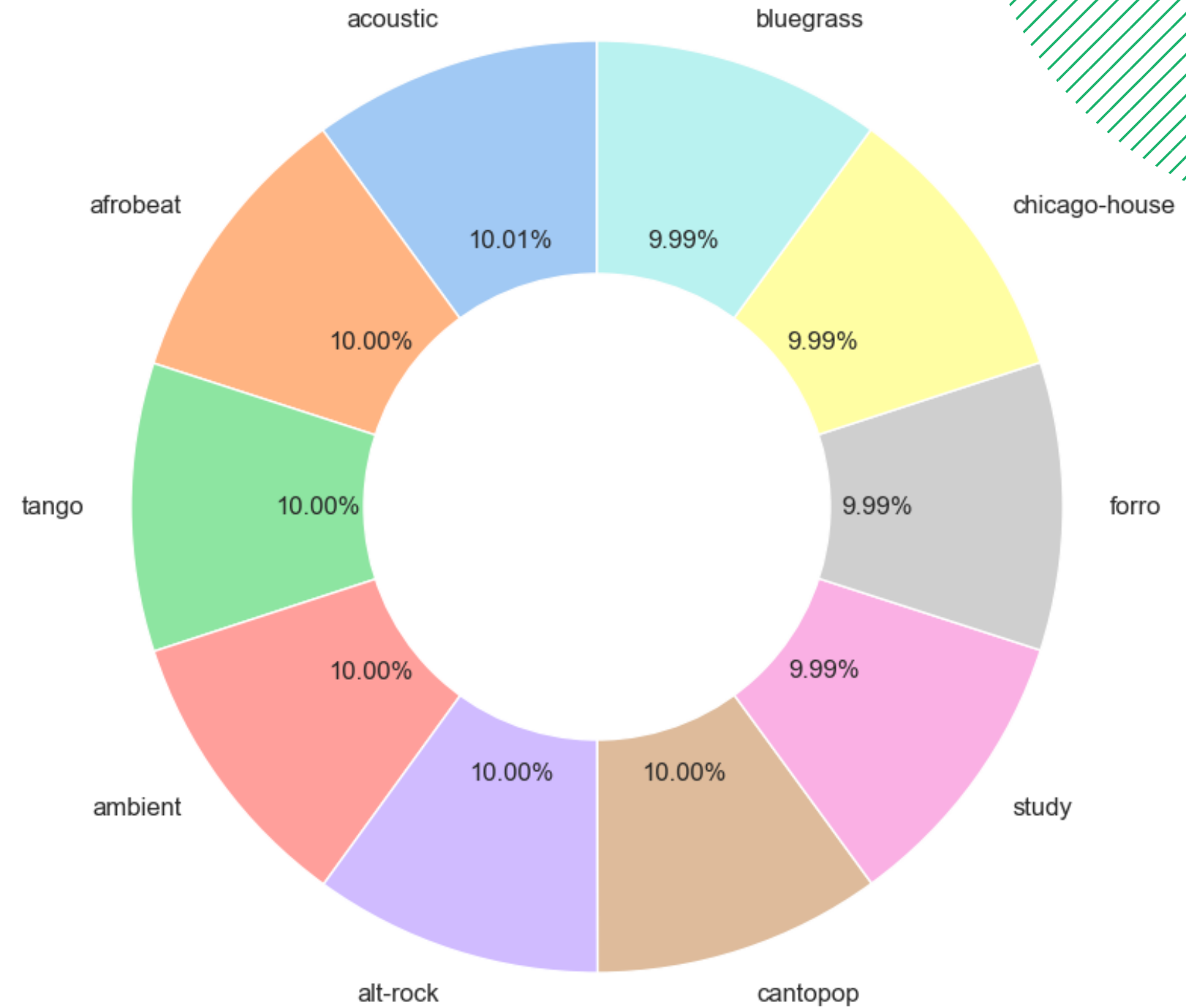


- 90.000 canciones
- Hasta 2021

Géneros mas presentes

- Los 10 géneros que cuentan con más canciones está distribuidos de forma muy similar

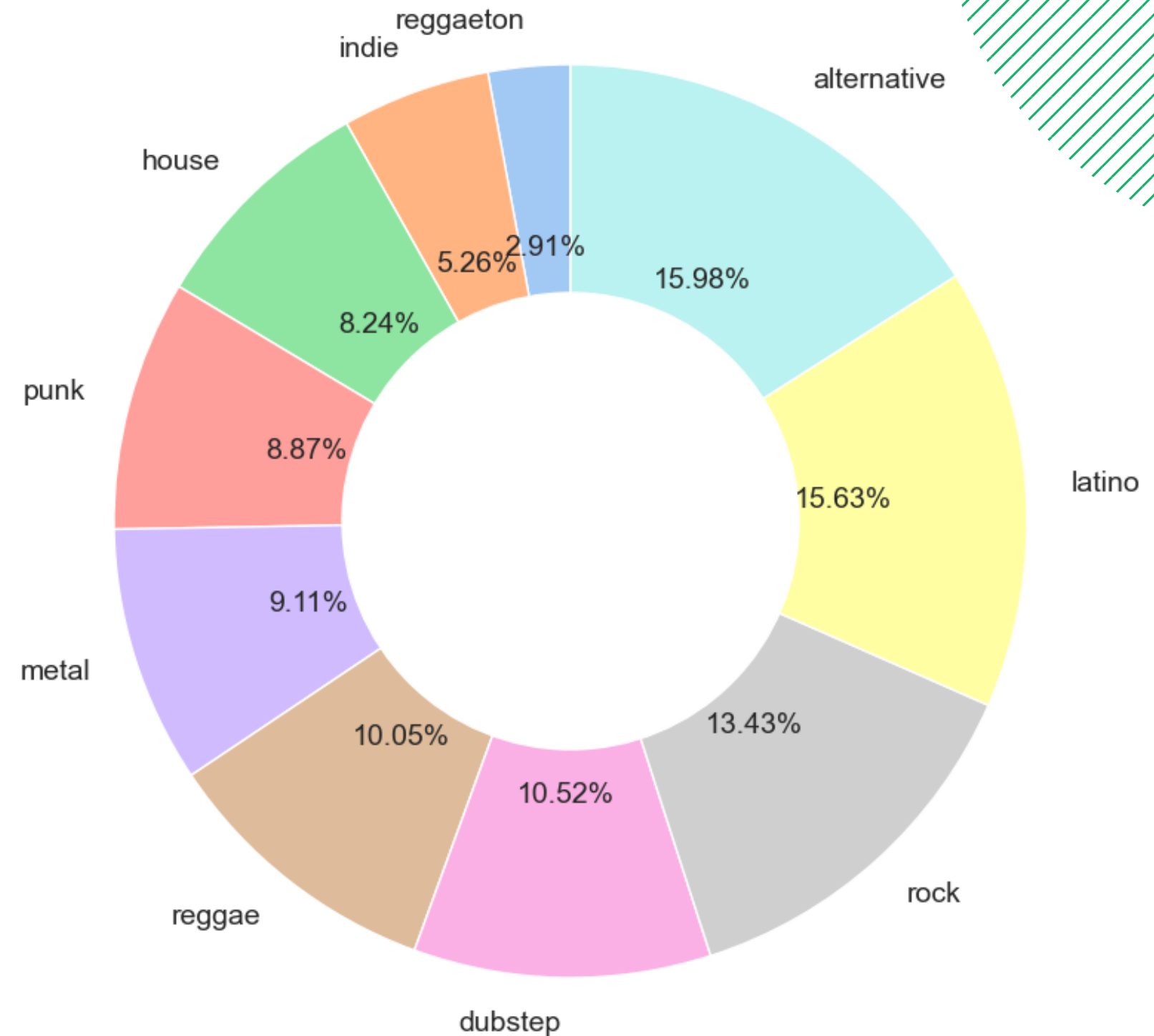
Distribution of Top 10 most present genres



Géneros menos presentes

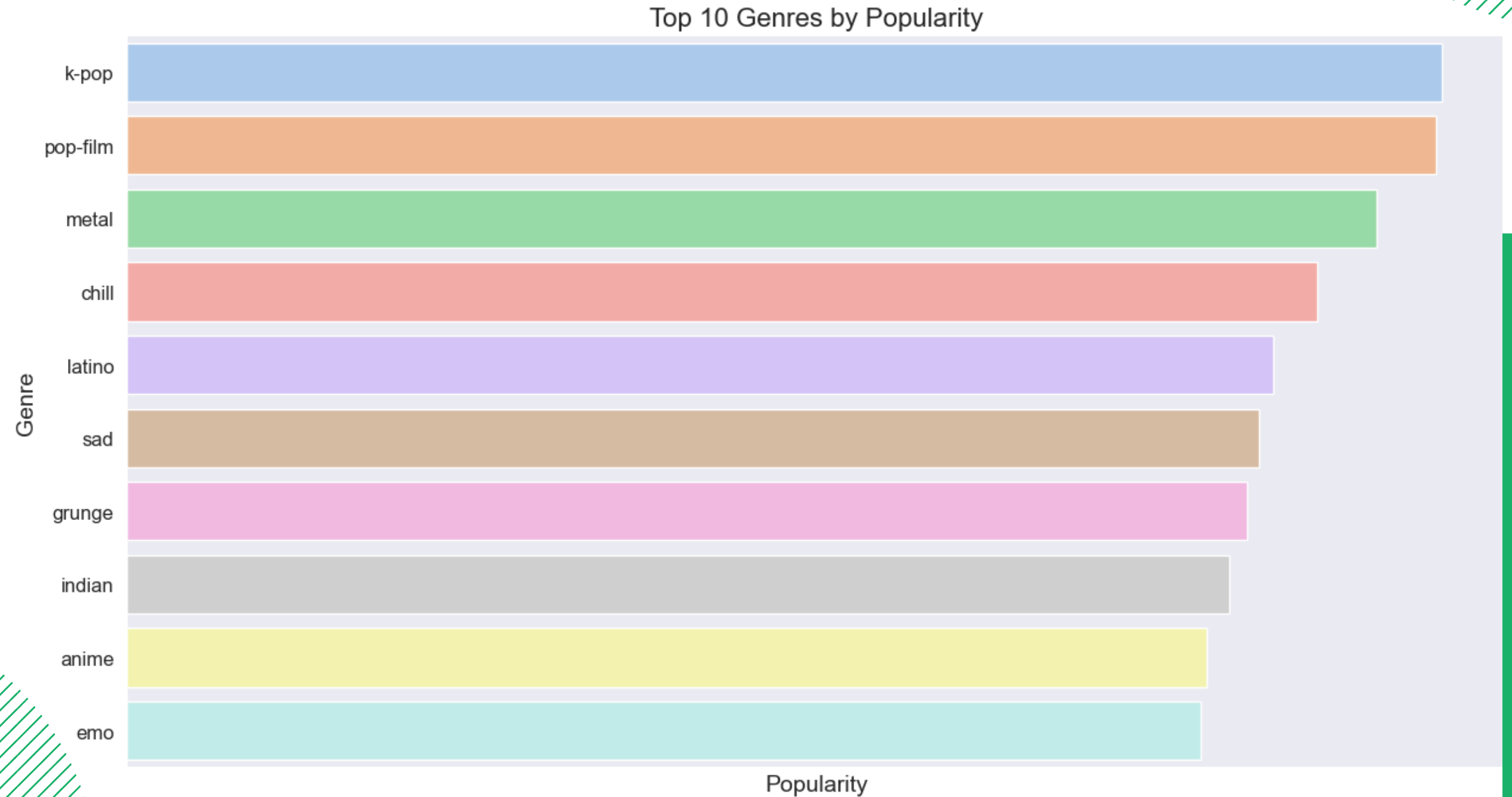
- Reggaeton, indie, rock, house entre los que menos representación tienen
- Géneros categorizados erróneamente

Distribution of Top 10 least present genres



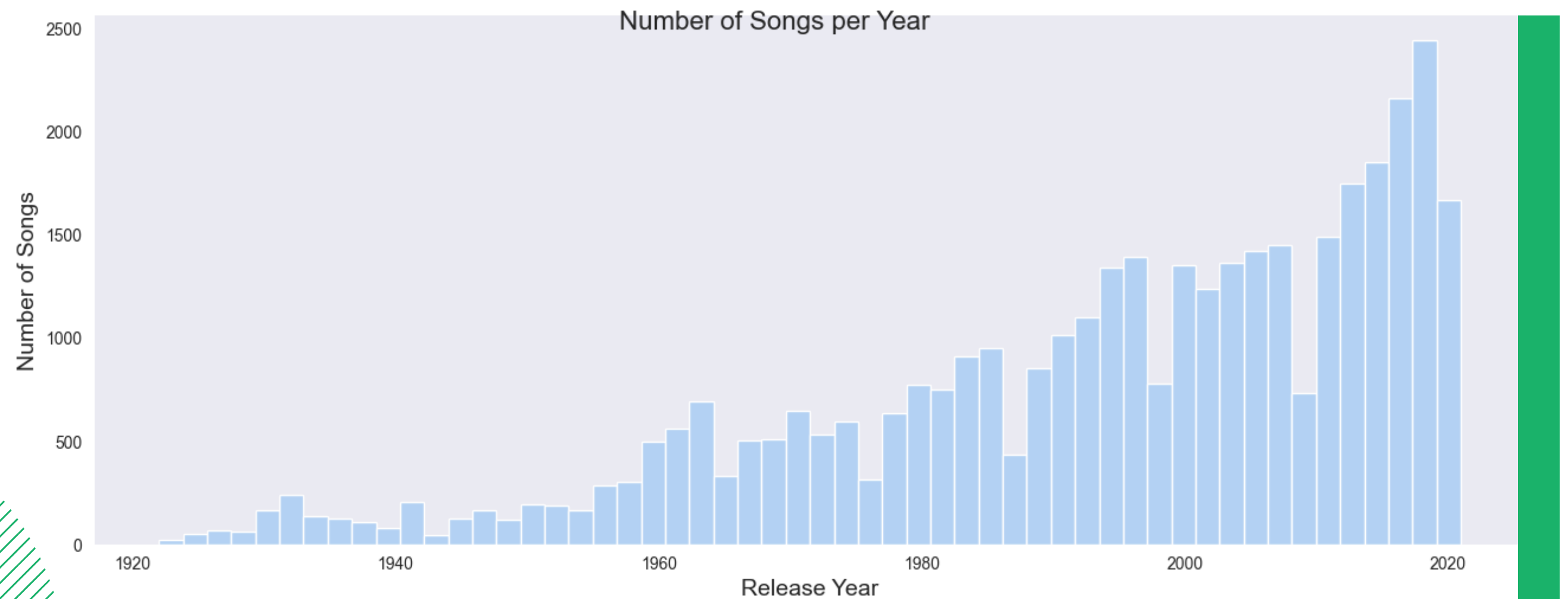
Géneros más populares

- K-Pop género más popular
- Diferentes subgéneros de Rock y Metal
- Música india entre los más populares



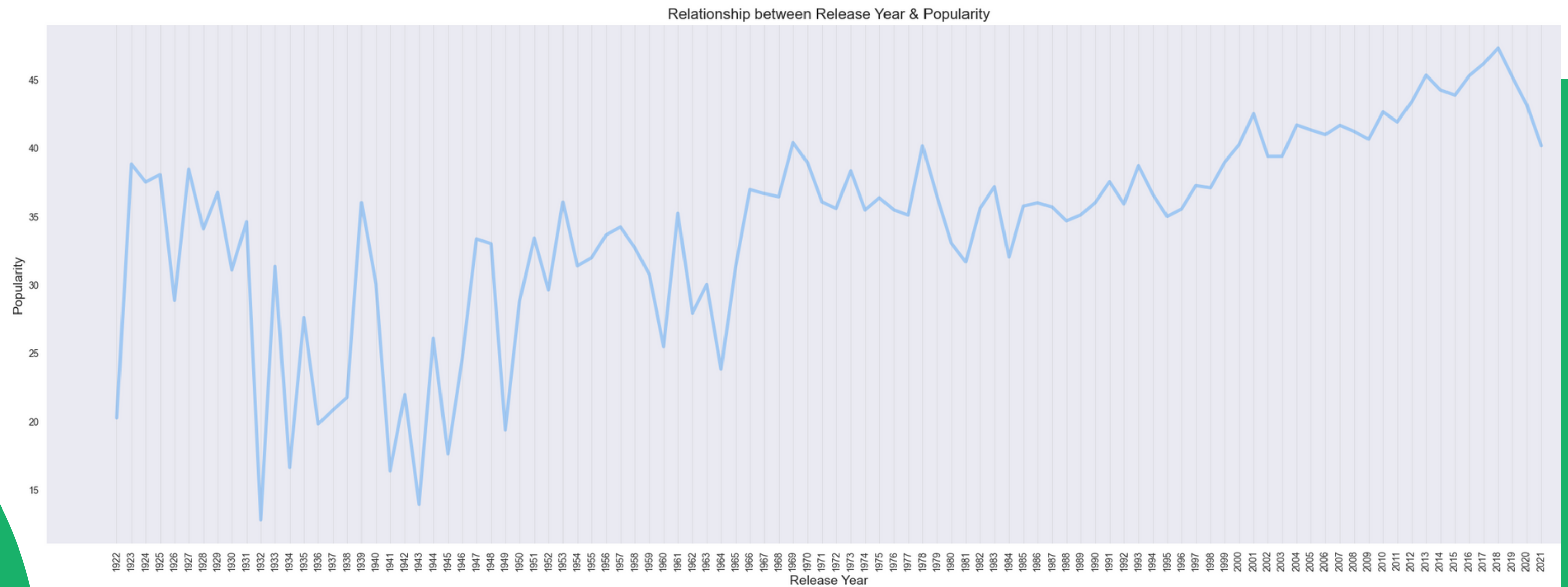
Canciones por año

- A partir de 1980 incrementa exponencialmente el número de canciones
- Nuevos medios de producción y grabación



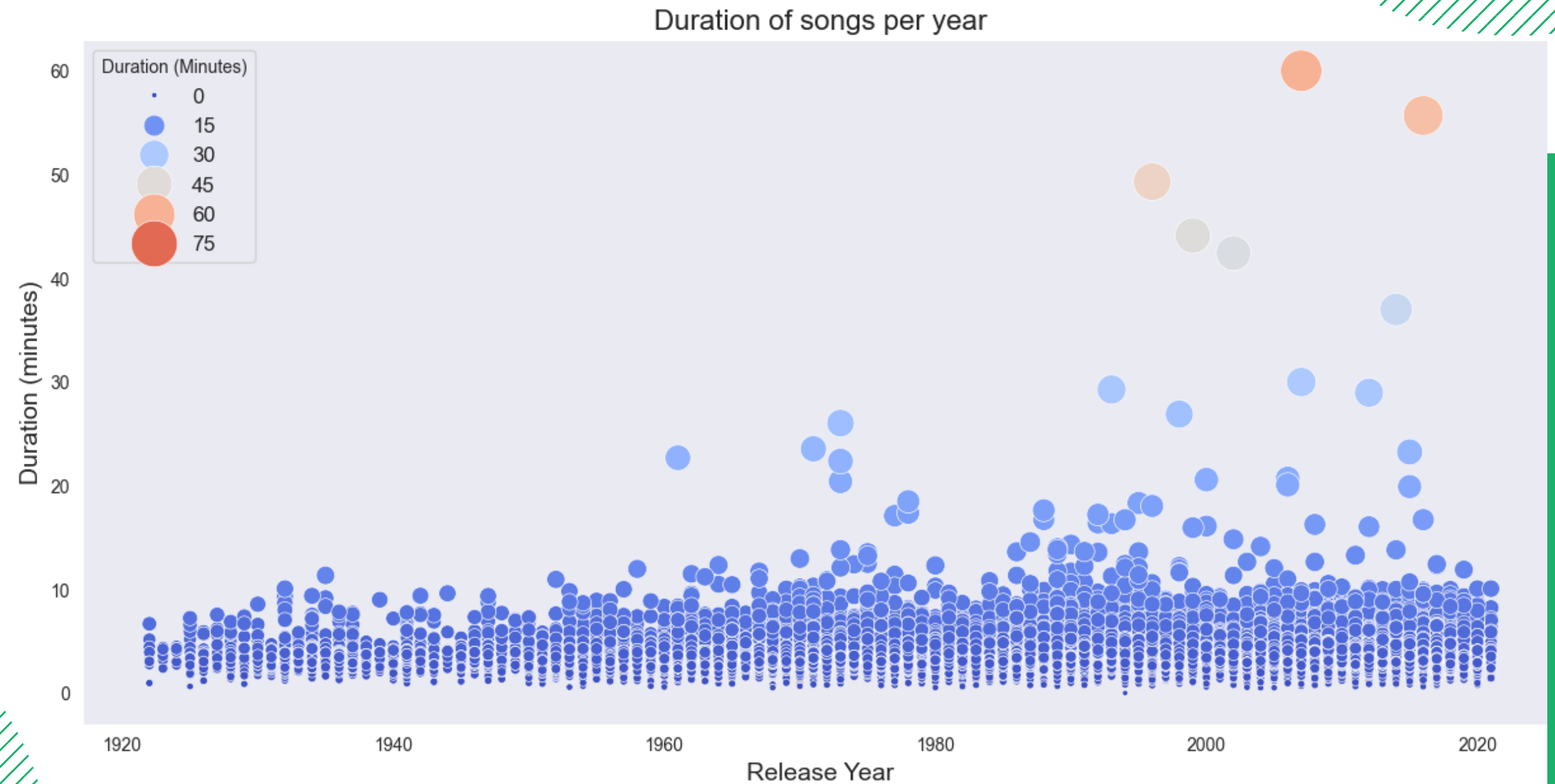
Popularidad por año

- A partir del '67 hay un incremento gradual de la popularidad de las canciones



Media de duración por año

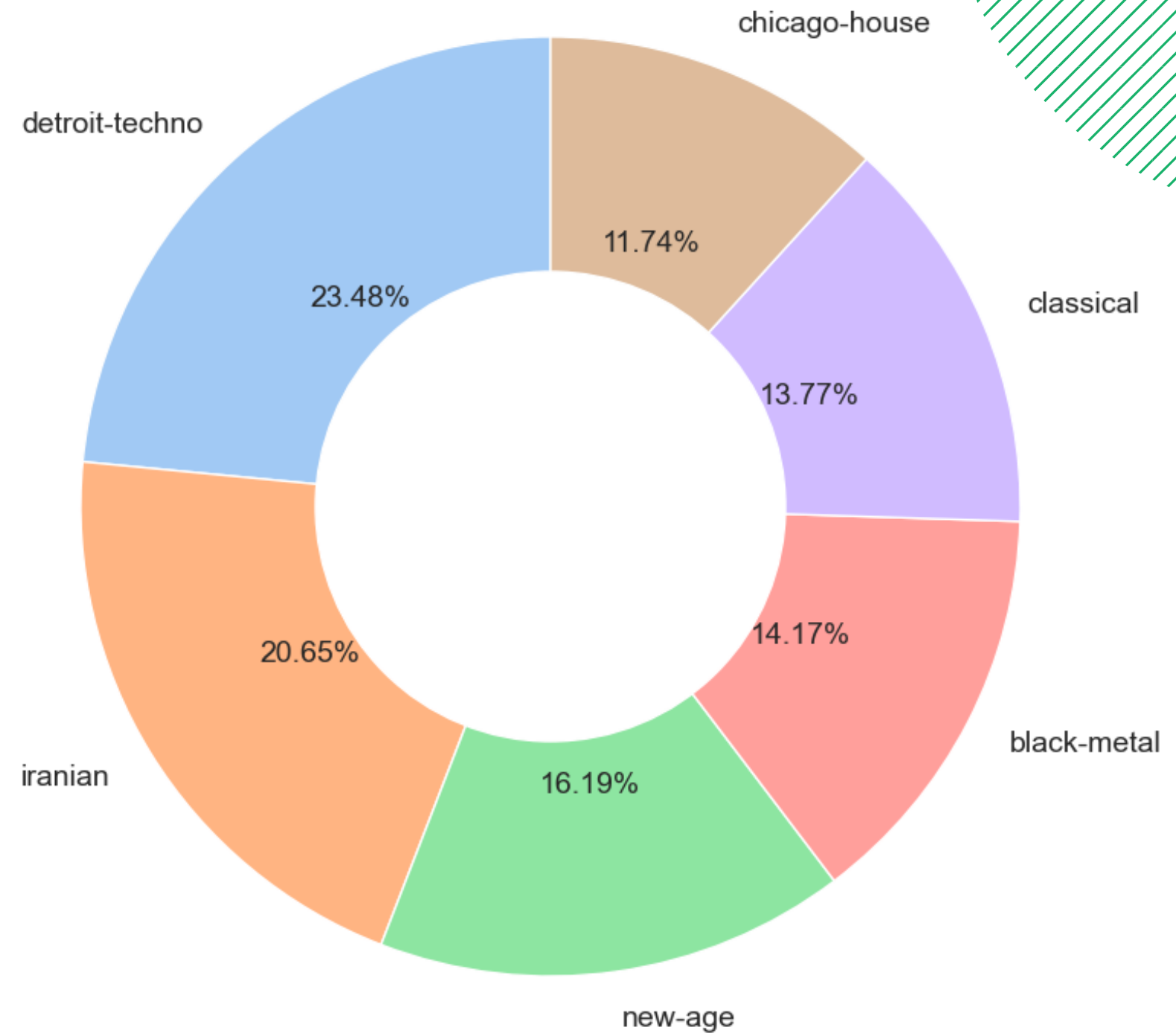
- A partir de 1960 empieza a haber un incremento en la duración de las canciones
- Nuevos soportes de grabación (cinta, cd, mp3)



Duración > 10 minutos

- Techno, House, experimental
- Música clásica
- Música Iraní
- 1960 aparecen nuevos géneros e instrumentos

Genres with duration > 10 minutes



Duración y popularidad

- Las canciones más populares tienen entre 2.5 y 4 minutos
- A partir de los 4 minutos la popularidad es muy baja



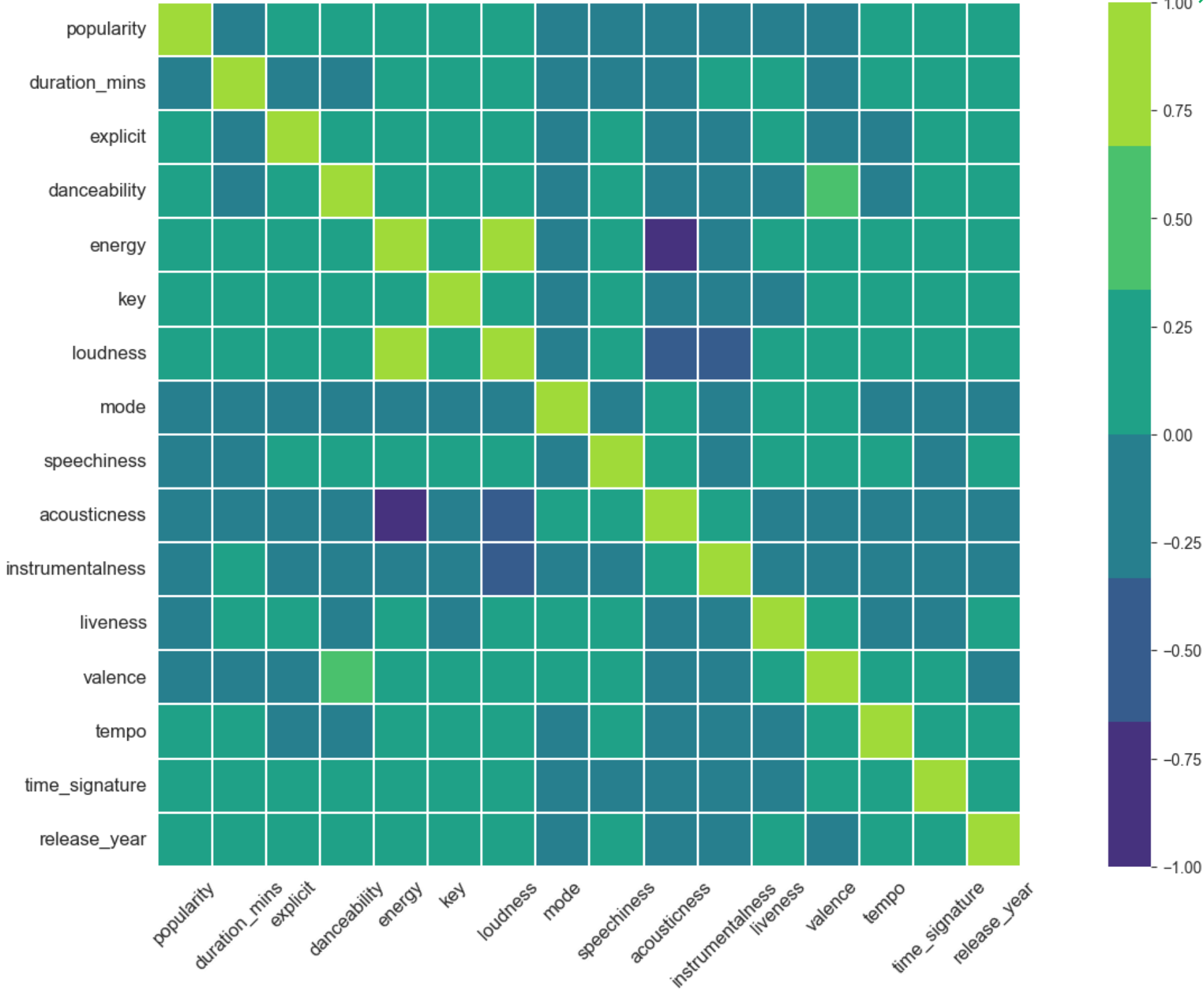
Correlación entre características



- Disponibilidad y estado anímico que genera

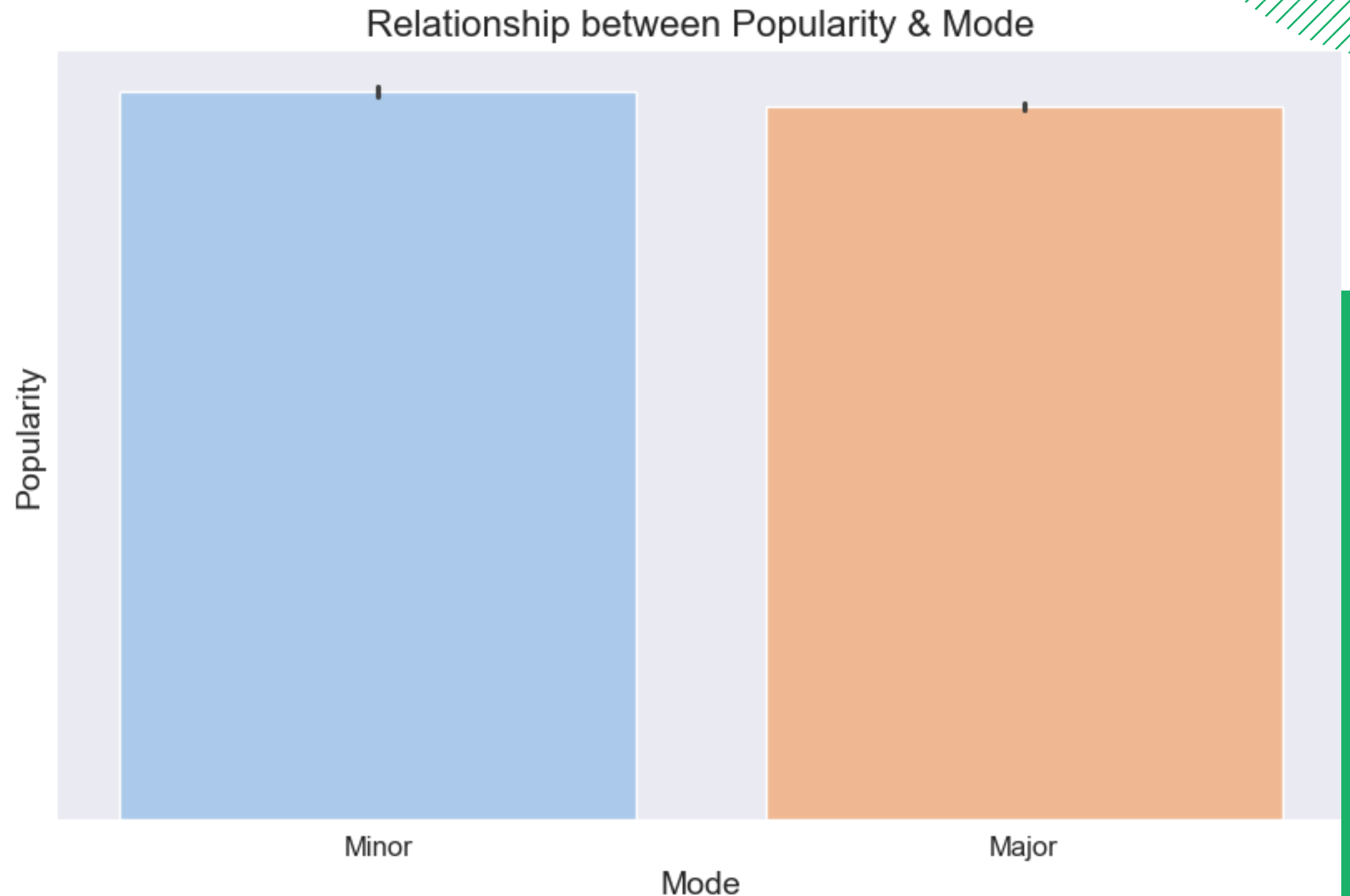


- Relación negativa entre energía y acusticidad



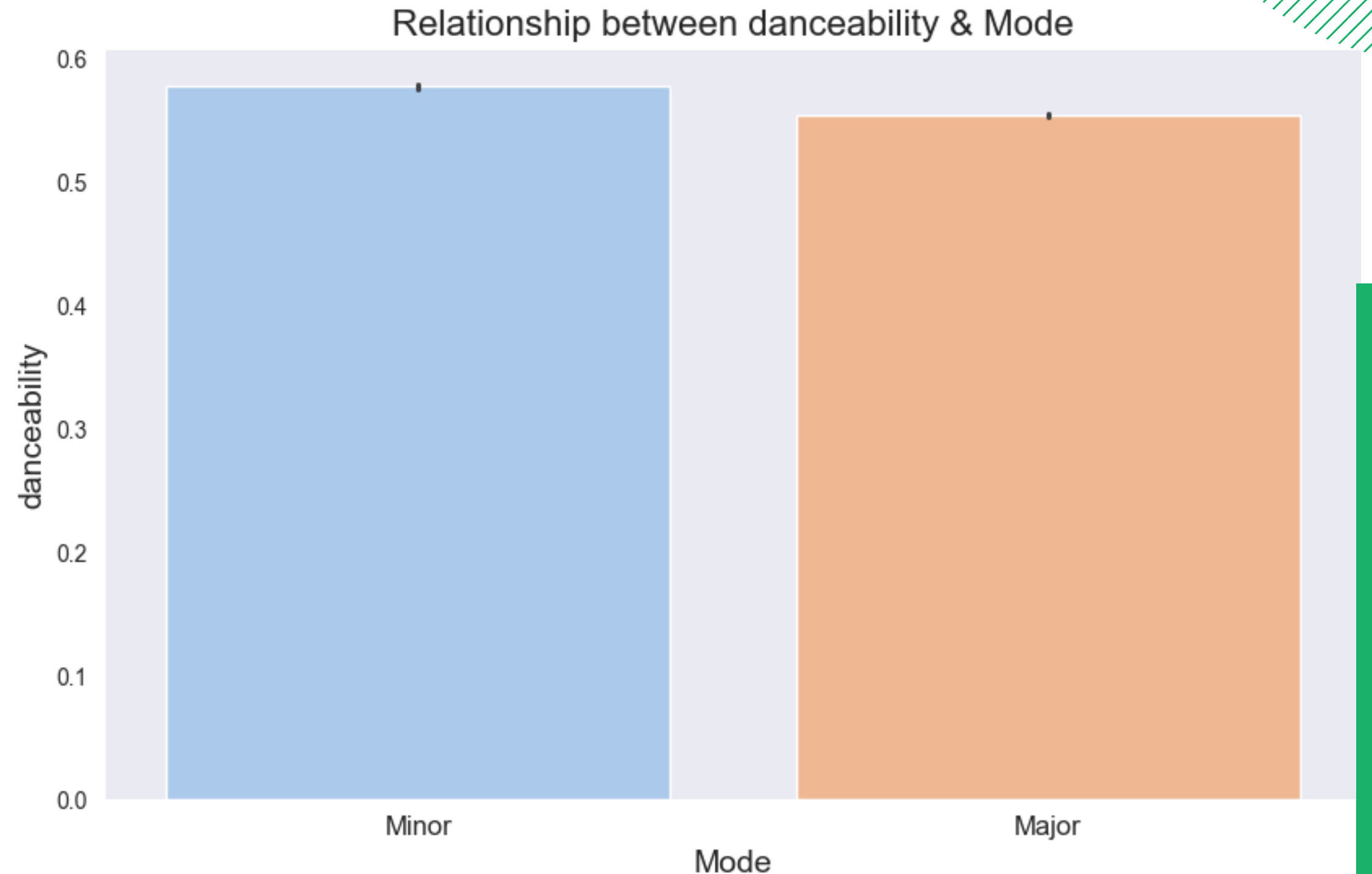
Tonalidad y popularidad

- No hay diferencia significativa entre tonalidades mayores y menores
- Ligera preferencia por tonalidades menores



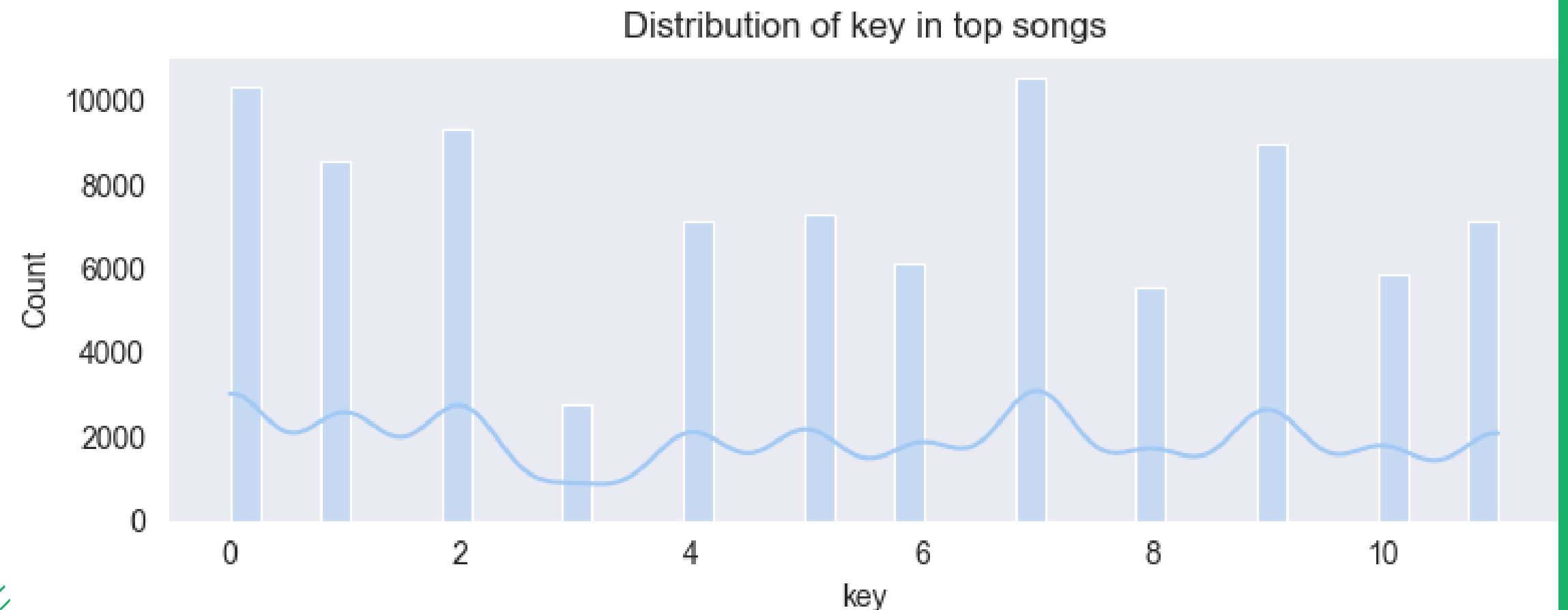
Tonalidad y bailabilidad

- Tampoco hay una gran diferencia entre la bailabilidad de los 2 modos tonales
- Las tonalidades menores cuentan con un grado más alto de bailabilidad



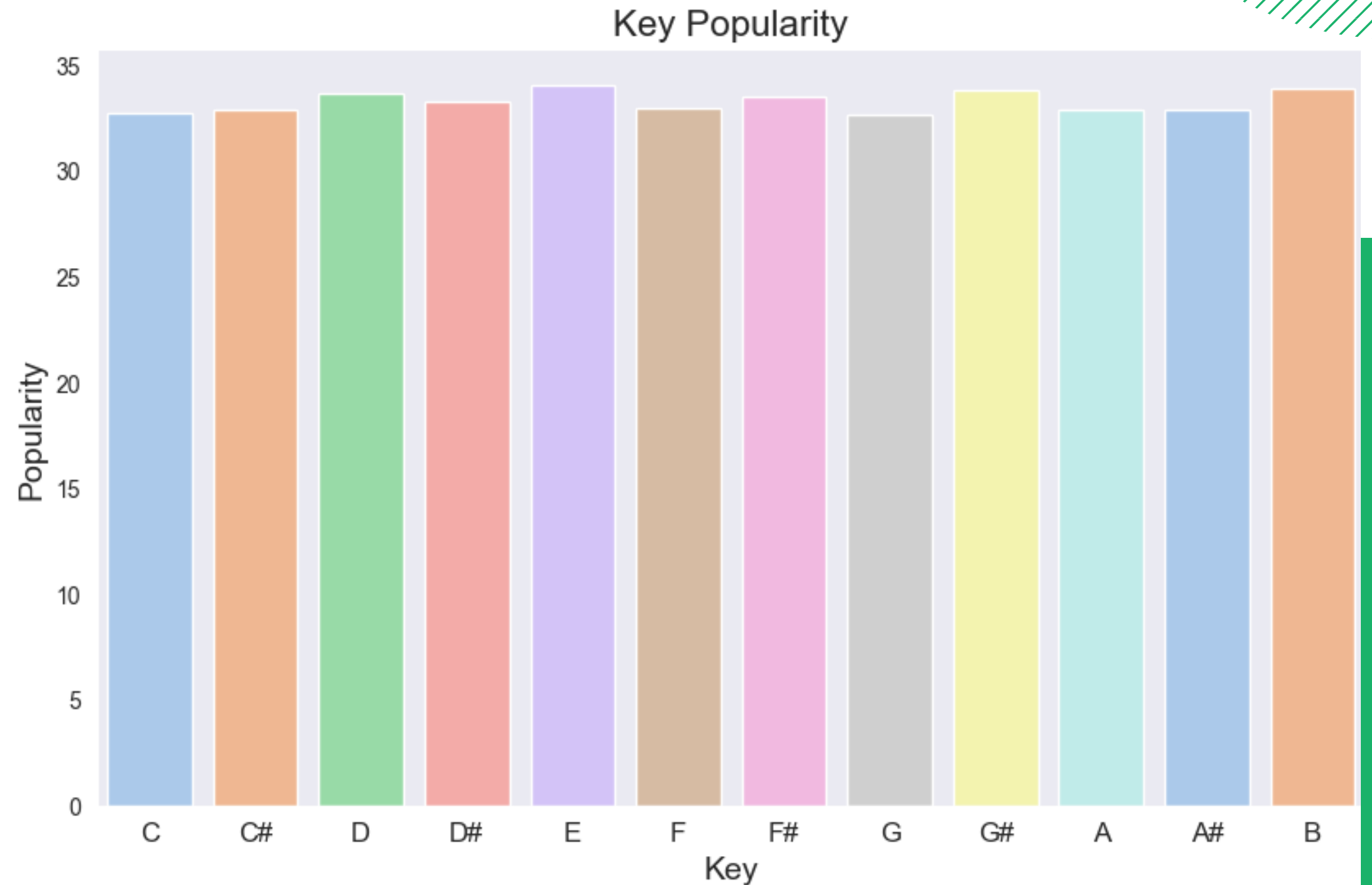
Tonalidad

- Hay mayor número de canciones en tonalidad de DO, RE y SOL
- RE# la tonalidad con menor número de canciones



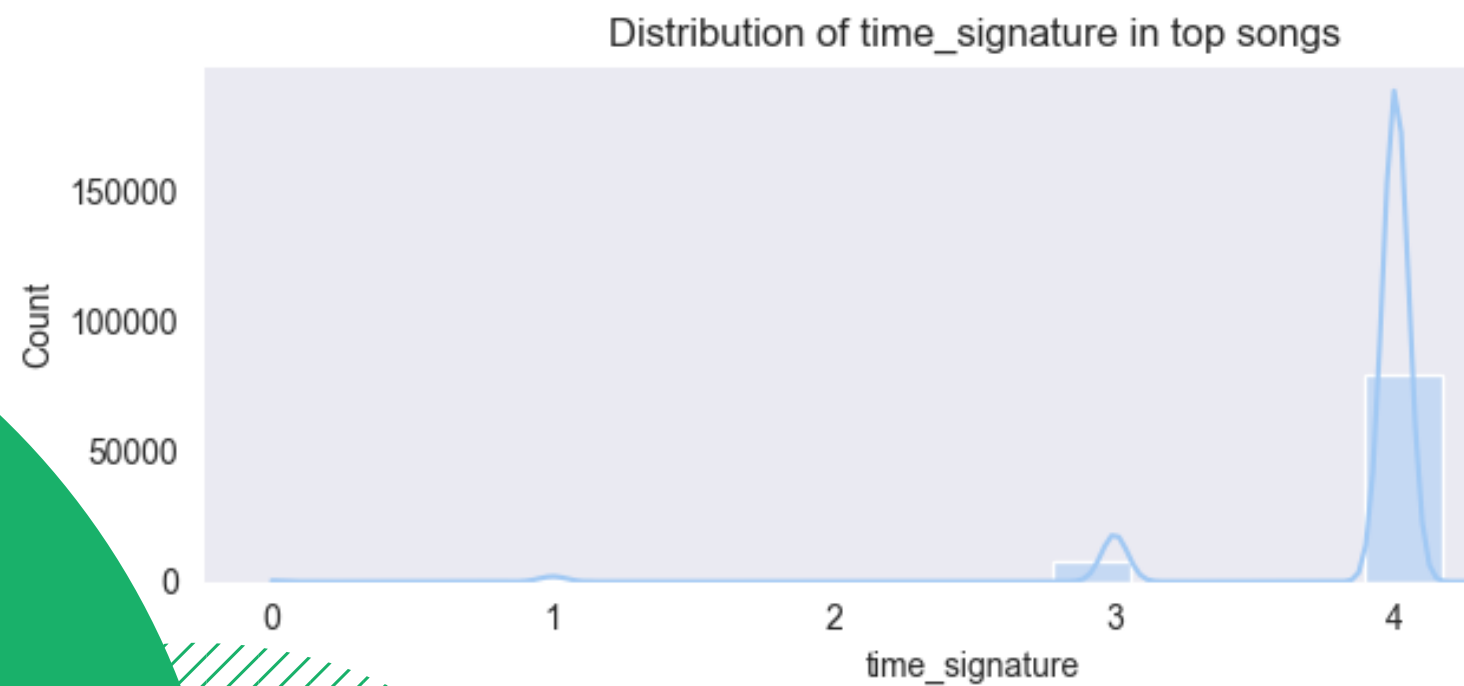
Tonalidad

- No hay diferencia significativa en la media de popularidad de las diferentes tonalidades

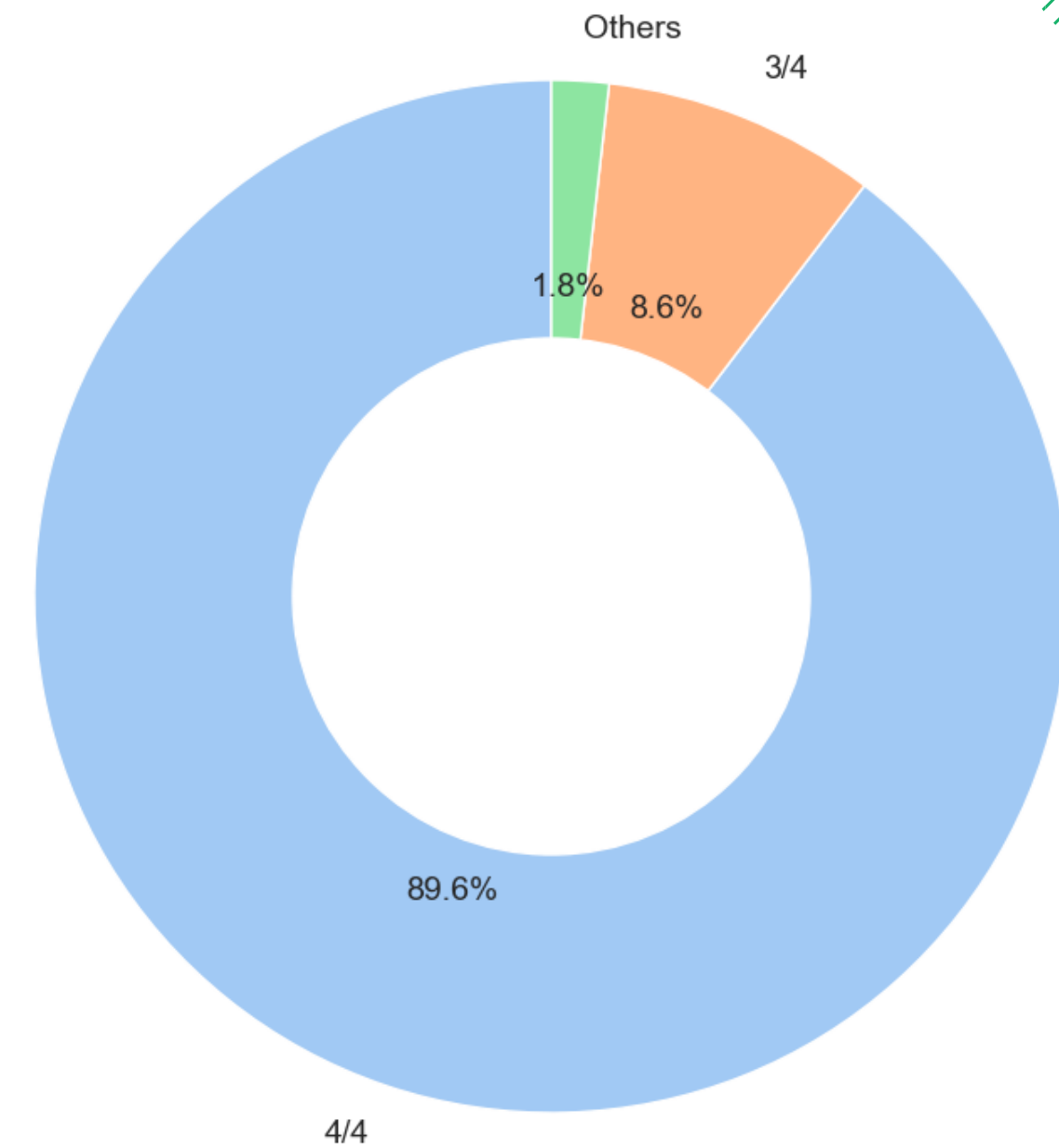


Compás

- La gran mayoría tienen un compás de 4/4
- Seguido de un 8% de canciones en 3/4
- Porcentaje muy pequeño de otros compases



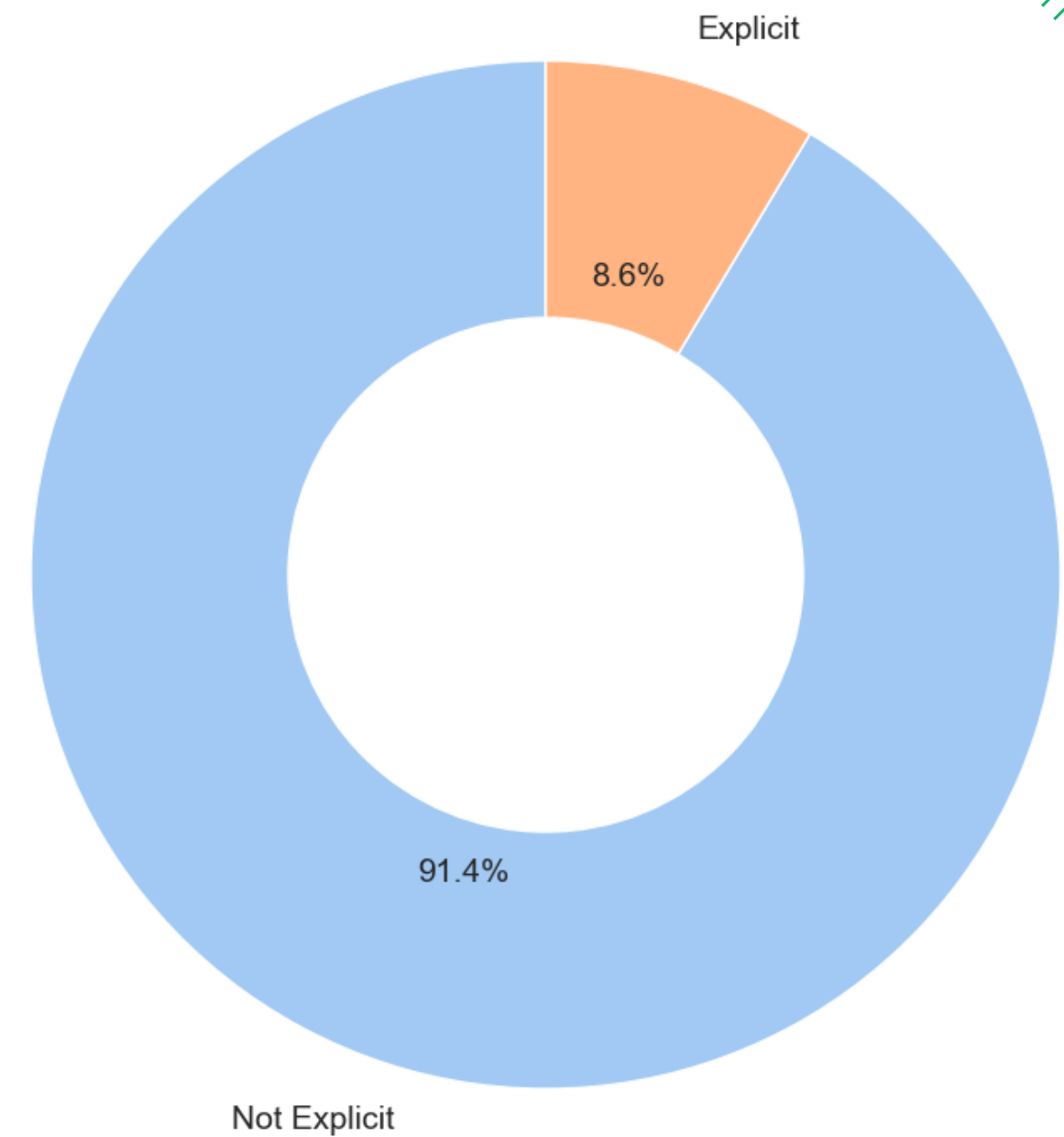
Time Signature distribution



Canciones explícitas

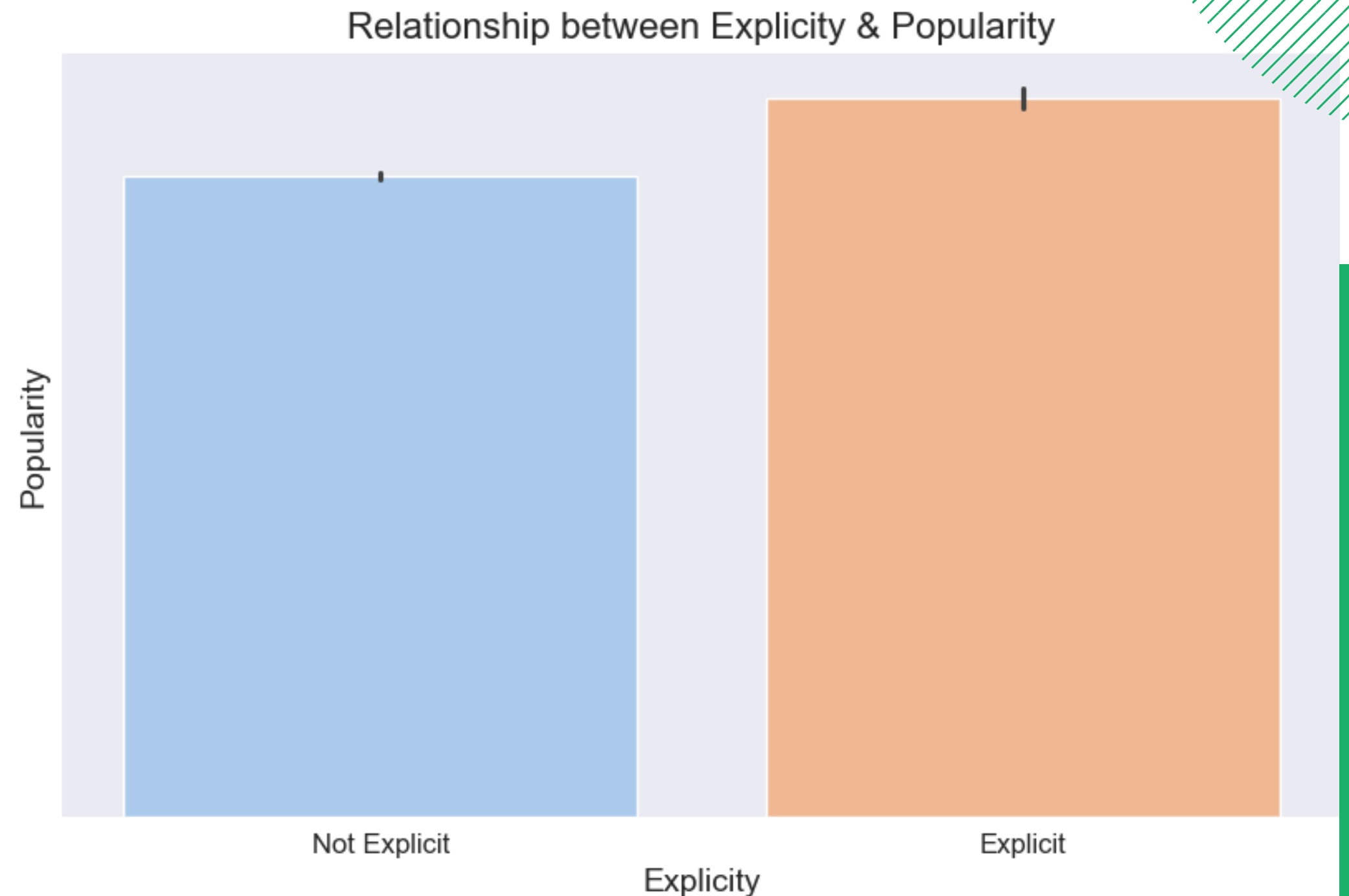
- Más del 90% no son explícitas
- Versión con y sin censura en plataformas de streaming
- Censura para emisión en radio y televisión

Distribution of explicit songs



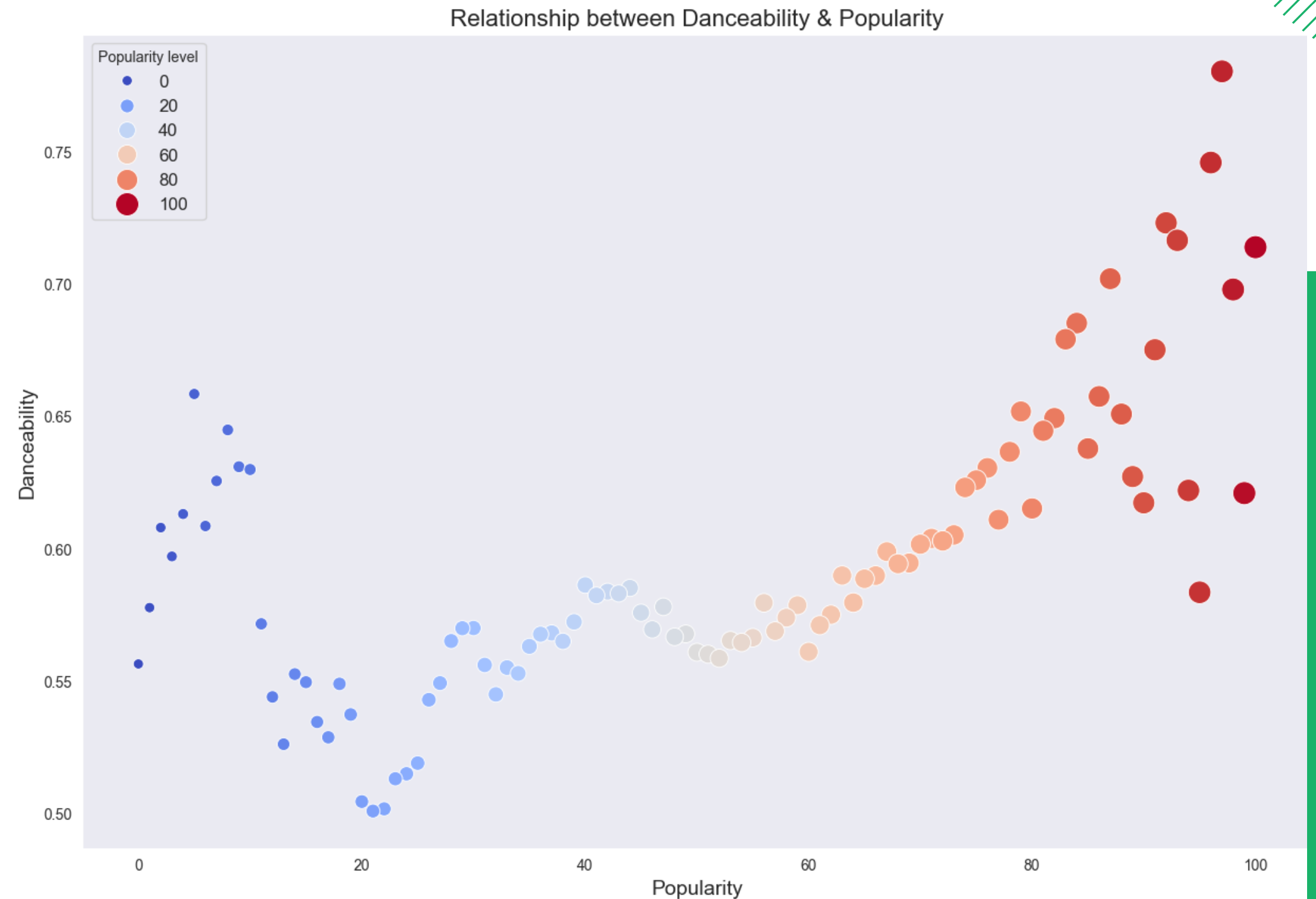
Canciones explícitas

- Ligera preferencia por parte de los usuarios por las canciones explícitas



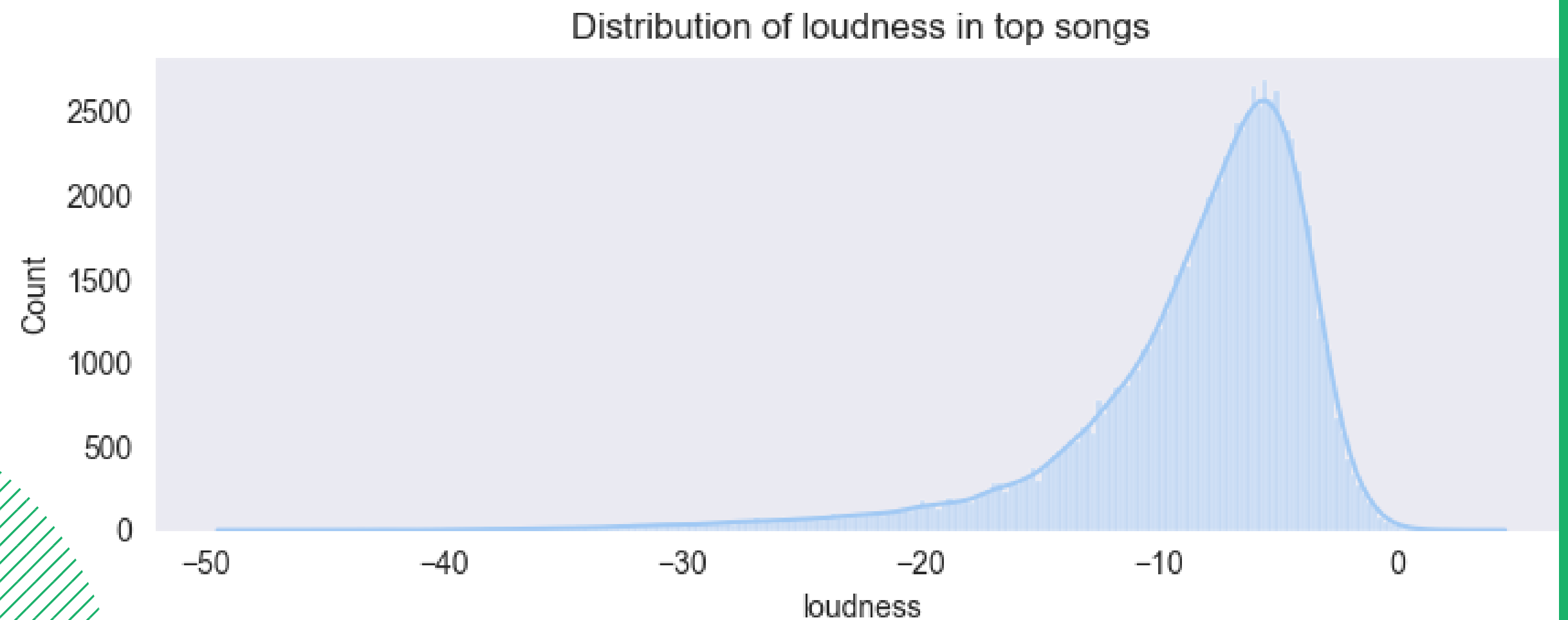
Bailabilidad y popularidad

- Las canciones más populares tienen una media de bailabilidad más alta



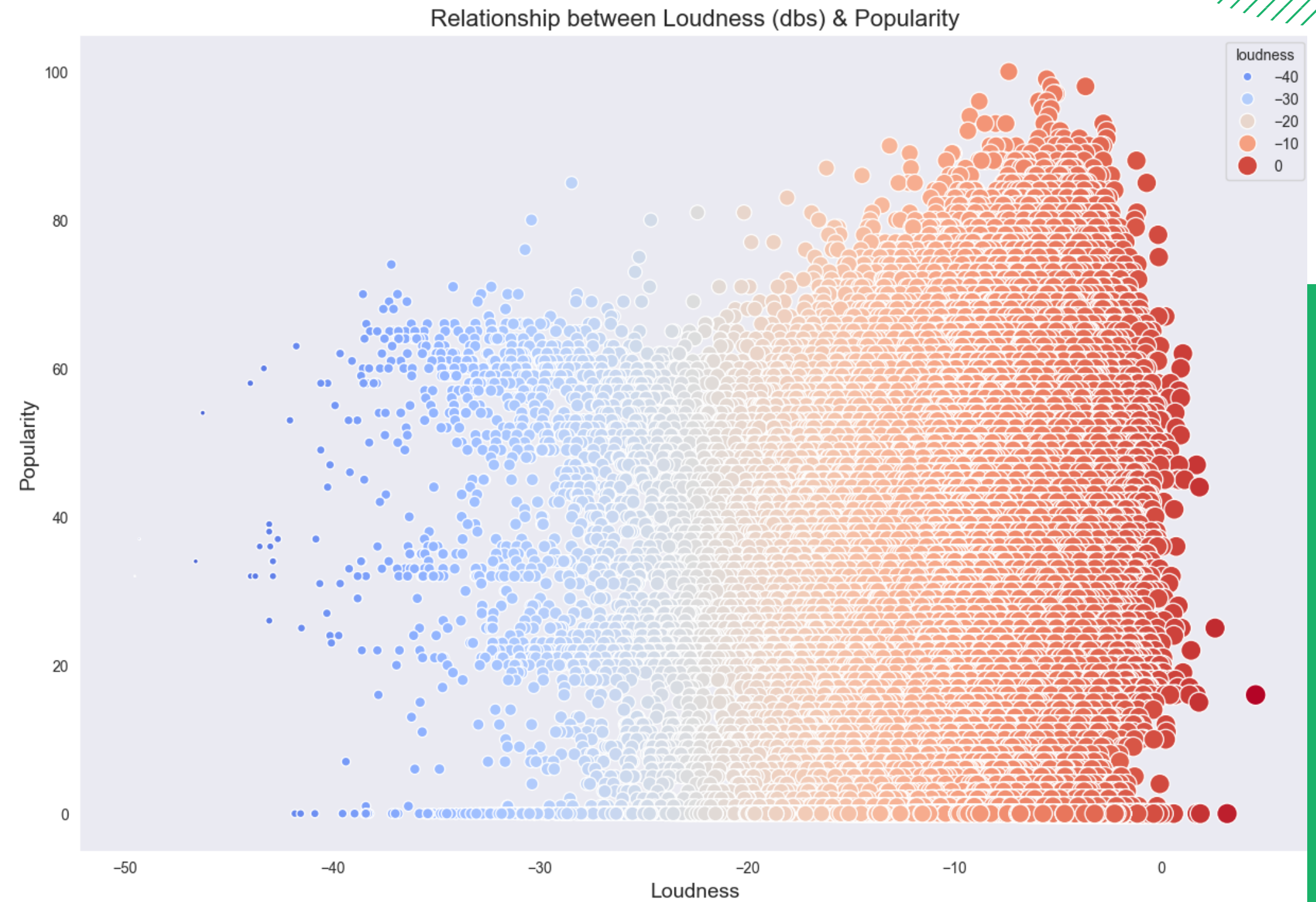
Volumen

- Estándar de volumen en servicios de streaming (-18, -14)db
- La gran mayoría de canciones alrededor de -6db
- Pérdida de calidad



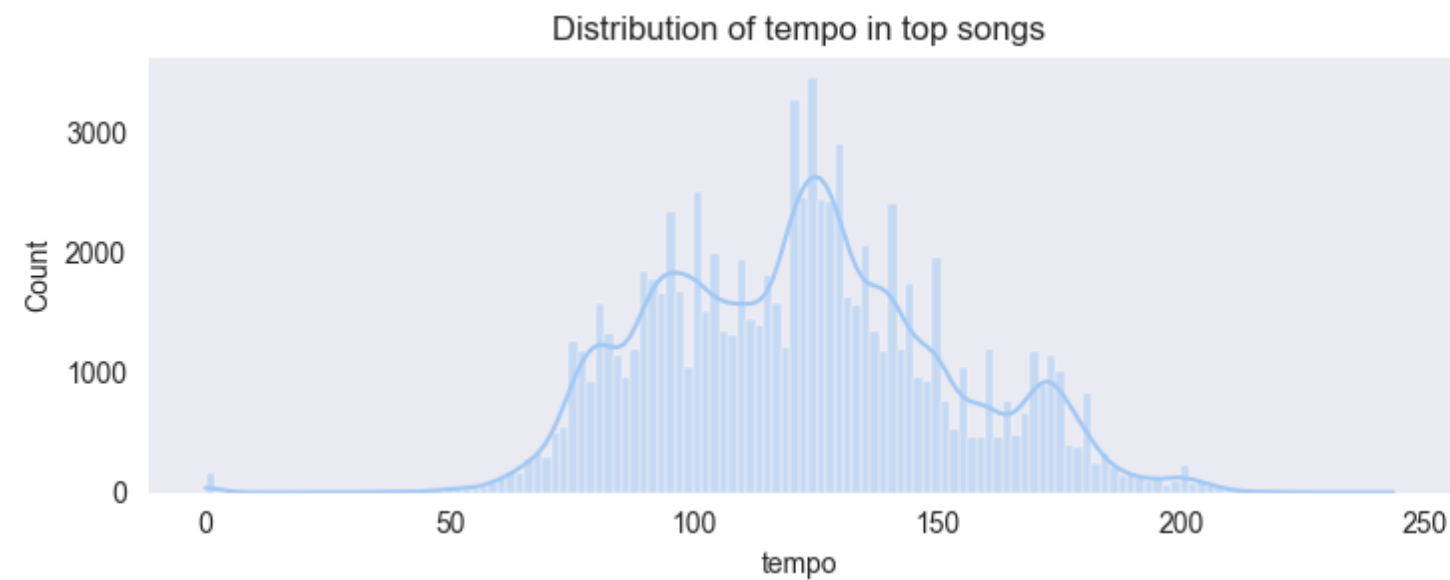
Volumen

- La gran mayoría de canciones populares están por encima de los -10 dbs



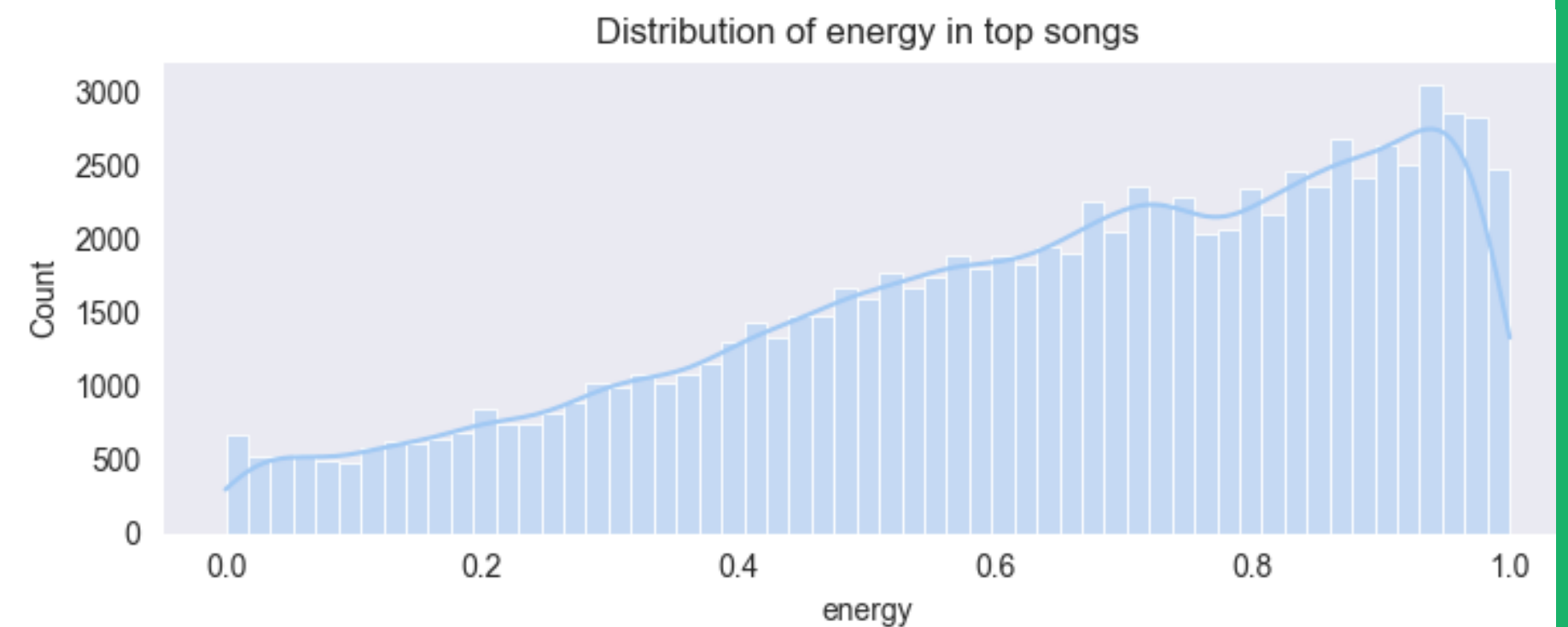
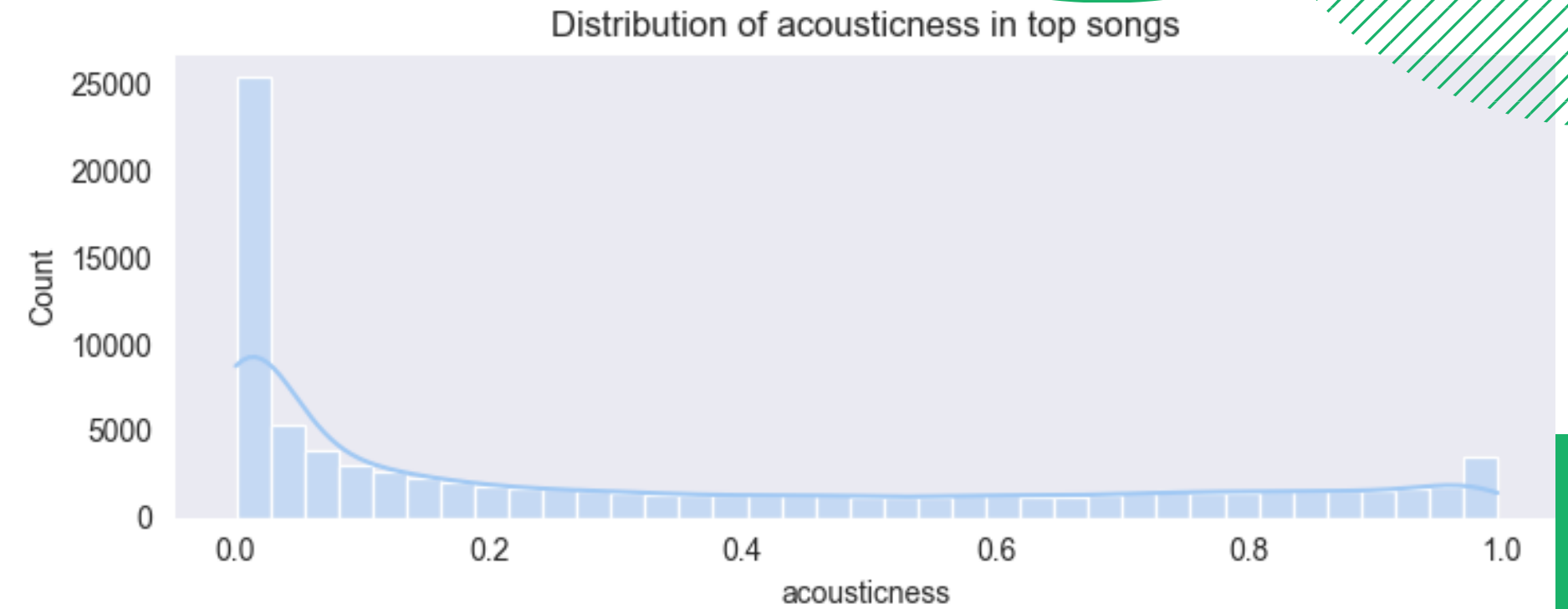
Tempo

- Las canciones más populares tienen un tempo de entre 115 y 130 BPM
- Anomalías:
- 100 BPM Reggaeton
- + 135 BPM “hardcores”, metal



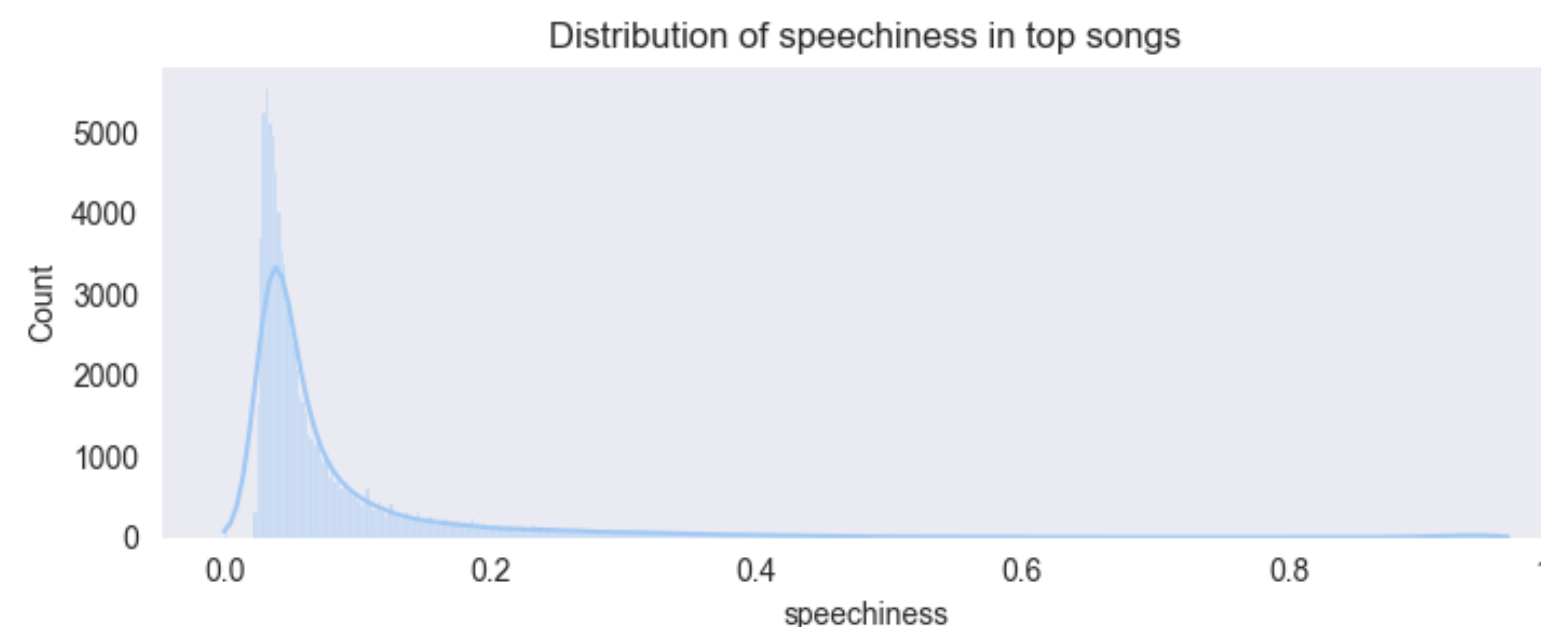
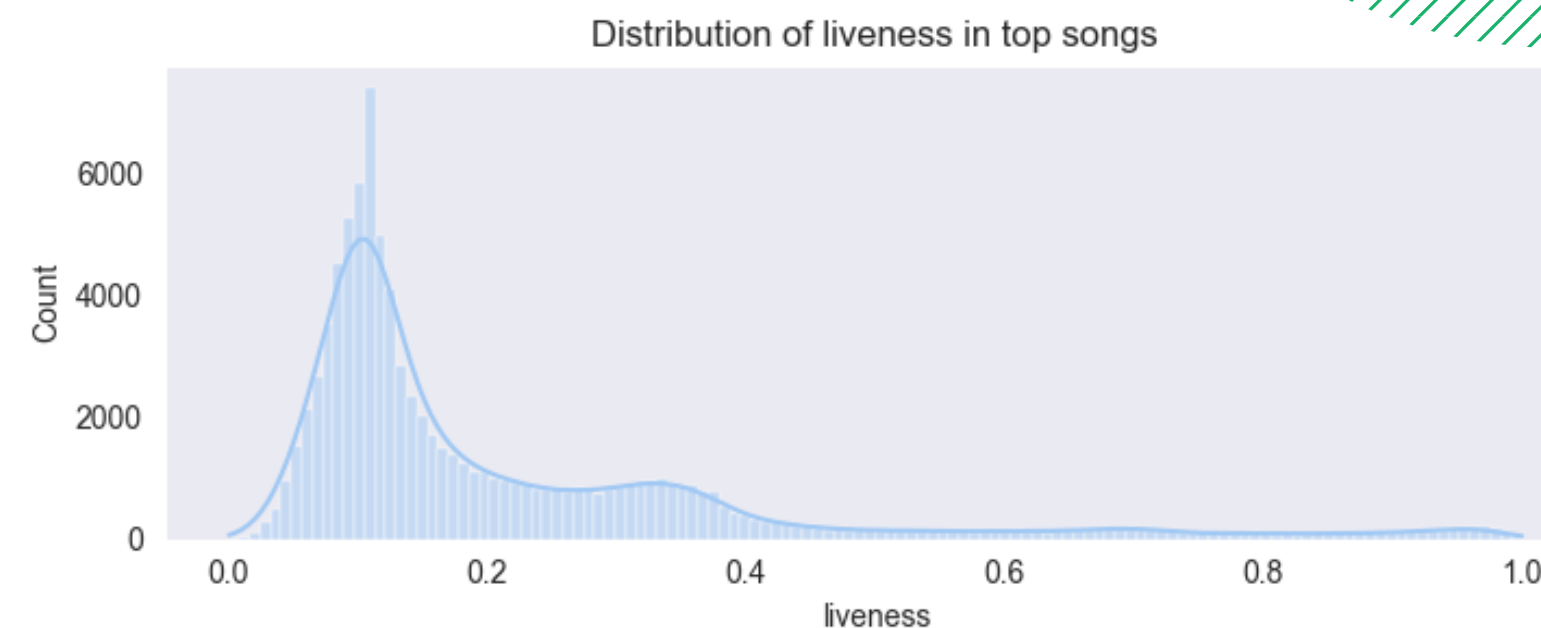
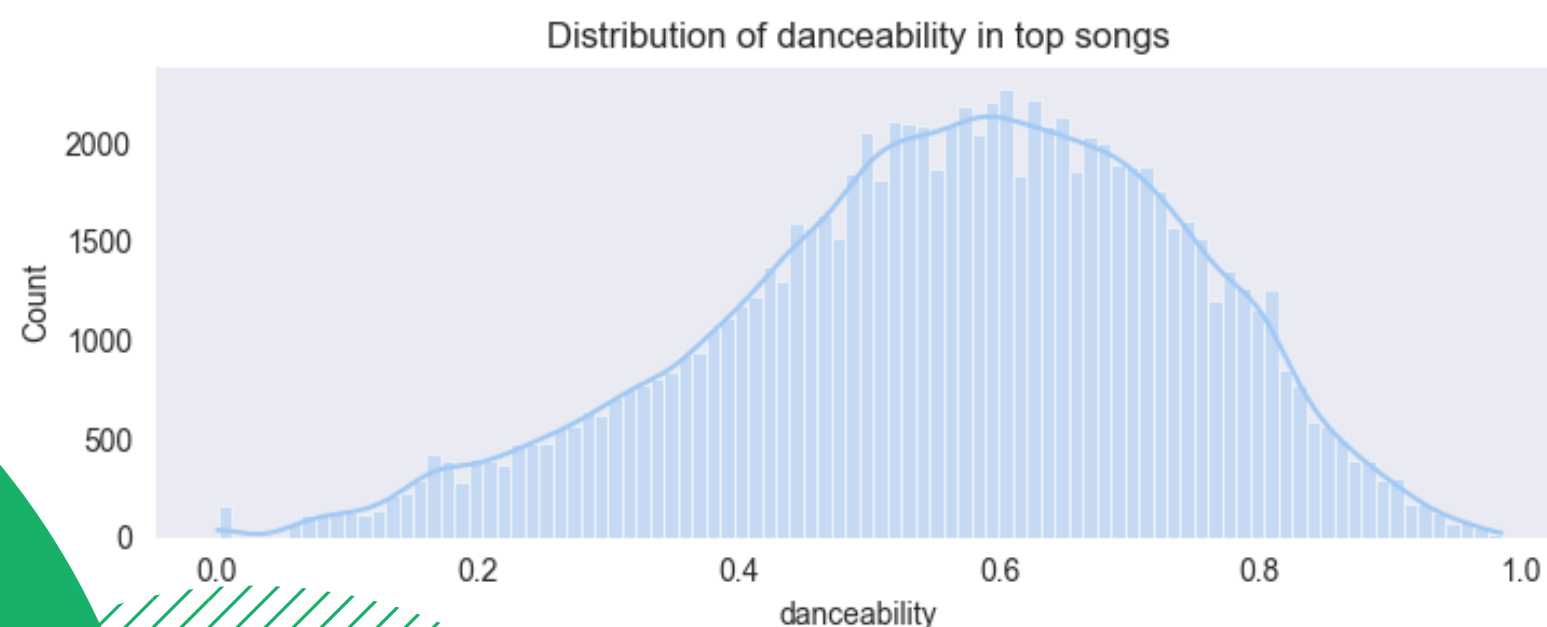
Energía y acusticidad

- La mayoría de canciones tienen un grado de acusticidad muy bajo
- Más canciones con mayor grado de energía



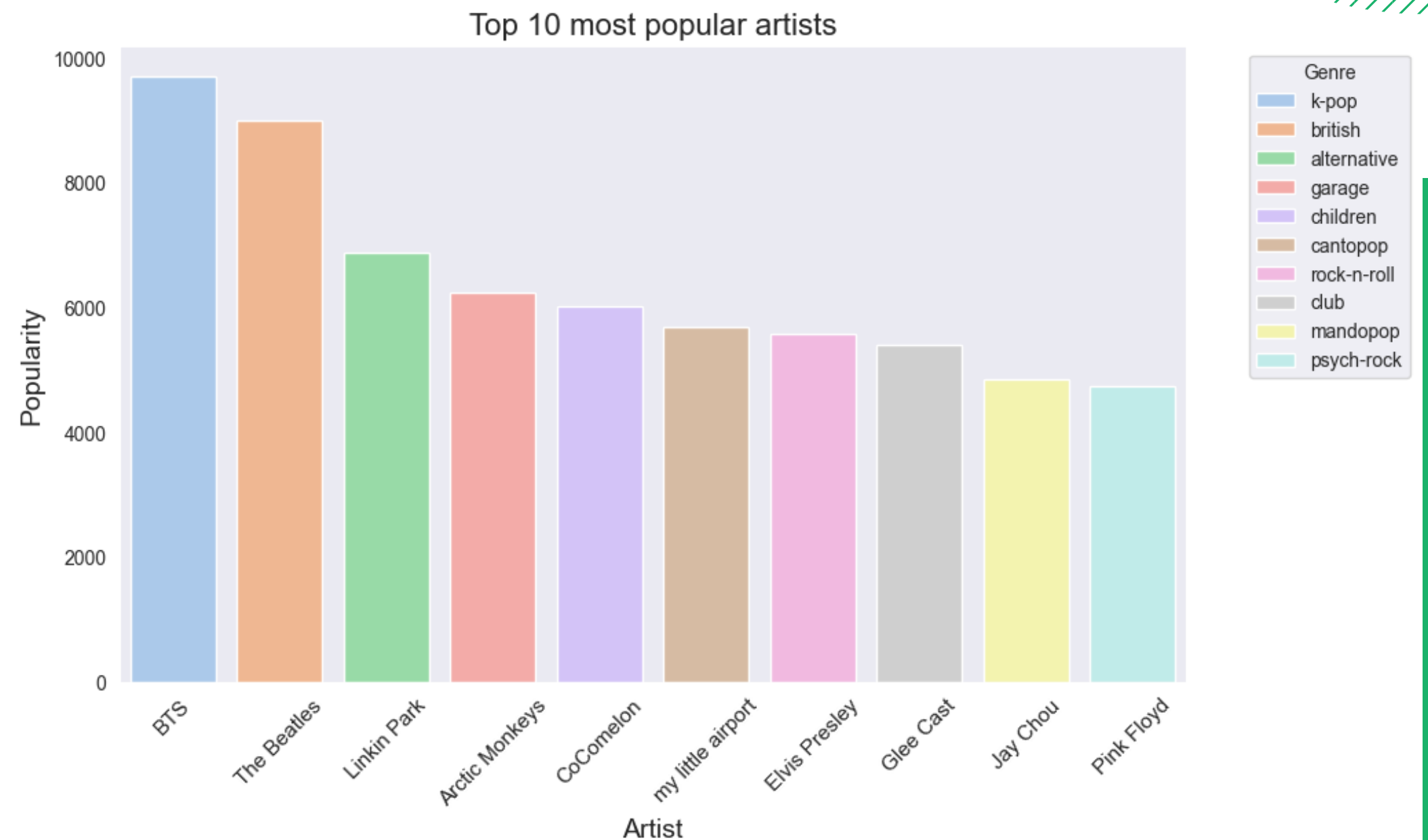
Bailabilidad, palabra y directo

- La mayoría de canciones tienen un grado bajo de directo
- La bailabilidad global está por encima de la mediana
- Hay poca cantidad de palabra hablada en las canciones



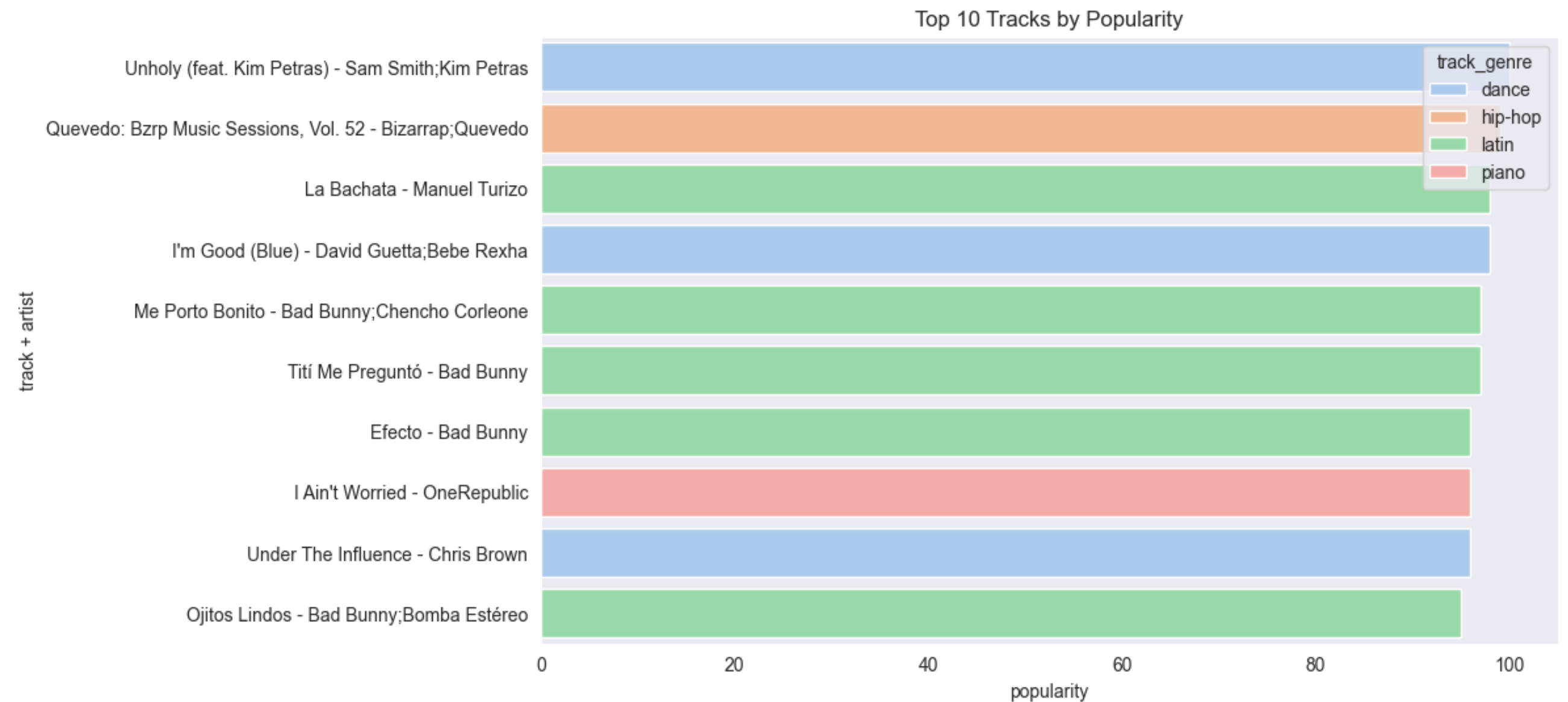
Artistas mas populares

- K-Pop y pop chino en el top-10
- Subgéneros del rock
- Música infantil



Canciones mas populares

- N° 1 pertenece al género dance
- Bad Bunny 4 posiciones en el top 10
- 6 de los 10 son música latina





Conclusión:

Aunque mi hipótesis principal era que los rasgos pueden influir en la popularidad, no puedo afirmar que esto sea realmente así.

Si que se puede ver que los usuarios pueden tener algo más de preferencia por diferentes claves, tempos o modos tonales pero ninguna de estas características es decisiva en la popularidad de una canción.

De ser así, el éxito de la música estaría reducido a una "formula mágica", Hago una canción a 128 BPM, en SOL Mayor, explícita y masterizada a -6db y tengo el éxito asegurado...

Pero obviamente esto no es así. Si que estas características pueden ayudar al usuario a que la canción le pueda "entrar" mejor de primeras, pero hay muchos factores externos que afectan a la popularidad de una canción, como la popularidad del artista, la aparición de la canción en diferentes medios, la estructura de esta, etc...