

# POPULARITY

*¿Qué hace a una  
canción popular?*

**Contexto y Audiencia**

**Hipótesis / Preguntas de interés**

**Metadata**

**Análisis Exploratorio**

**Features y conclusiones**

# Contexto, Audiencia y Limitación

## Contexto

### Industria musical

La producción musical es una industria impulsada por la creatividad, pero con los avances en la tecnología, ahora hay una oportunidad de incorporar ideas basadas en datos en el proceso creativo. Al aprovechar los algoritmos de aprendizaje automático, podemos analizar grandes conjuntos de datos de pistas de música populares y extraer características clave que son comunes entre las pistas exitosas. Estas características podrían incluir cosas como el tempo, las progresiones de acordes, las letras e incluso el estado de ánimo de la pista.

Al combinar la creatividad del productor con el poder analítico del aprendizaje automático, podemos llegar a un mejor resultado.

## Audiencia

### Productor musical latino

Estamos trabajando para un productor musical latino que está buscando su nuevo hit. Como el mercado es extremadamente competitivo necesita una guía sobre que tipo de canción será mejor recibida. Para así reducir el riesgo de la inversión.

## Limitación

### Fecha, Geografía y Demografía

El dataset con el que estamos trabajando no está actualizado y no tiene información sobre la demografía o ubicación geográfica en las que la canción es popular.

El dataset es de Octubre 2022

# **PREGUNTAS DE INTERÉS**

**¿Podemos predecir si una canción va a ser popular?**

**¿Cuál género es el más popular?**

**¿Hay alguna característica “física” que sea determinante?**

# RESUMEN METADATA

**114.000**

CANCIONES

**433.255**

MINUTOS DE MÚSICA

**114**

GÉNEROS

**31.437**

ARTISTAS

**20**

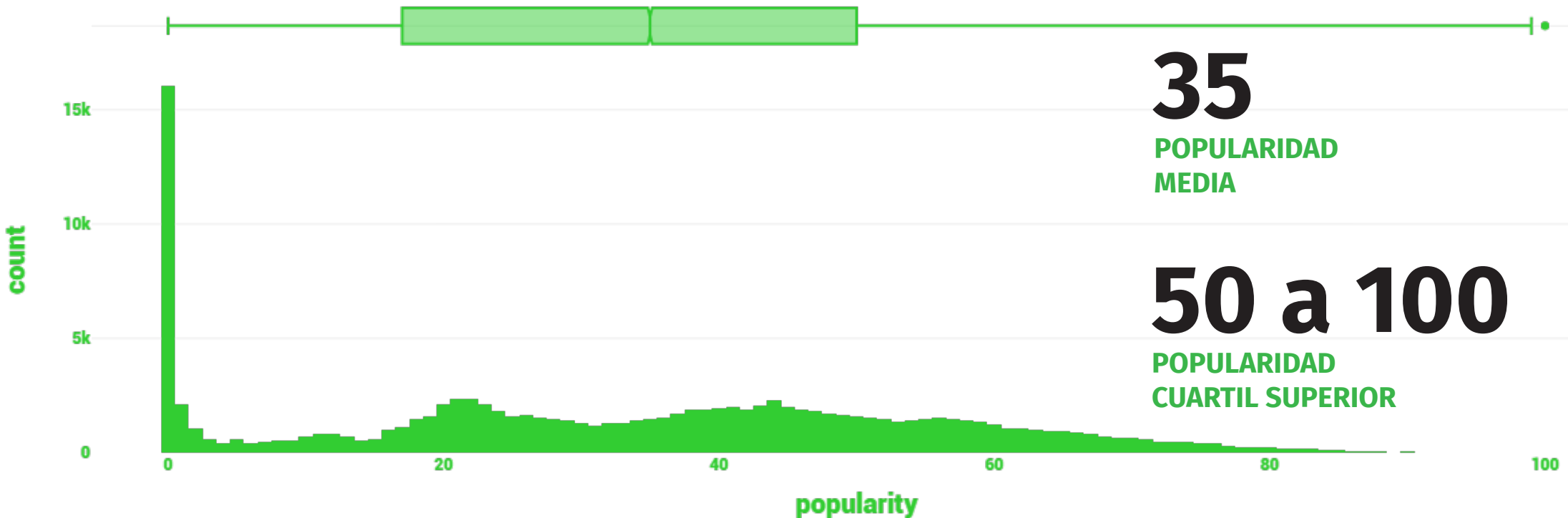
VARIABLES

# **ANÁLISIS EXPLORATORIO**

# DISPERSIÓN DE POPULARIDAD

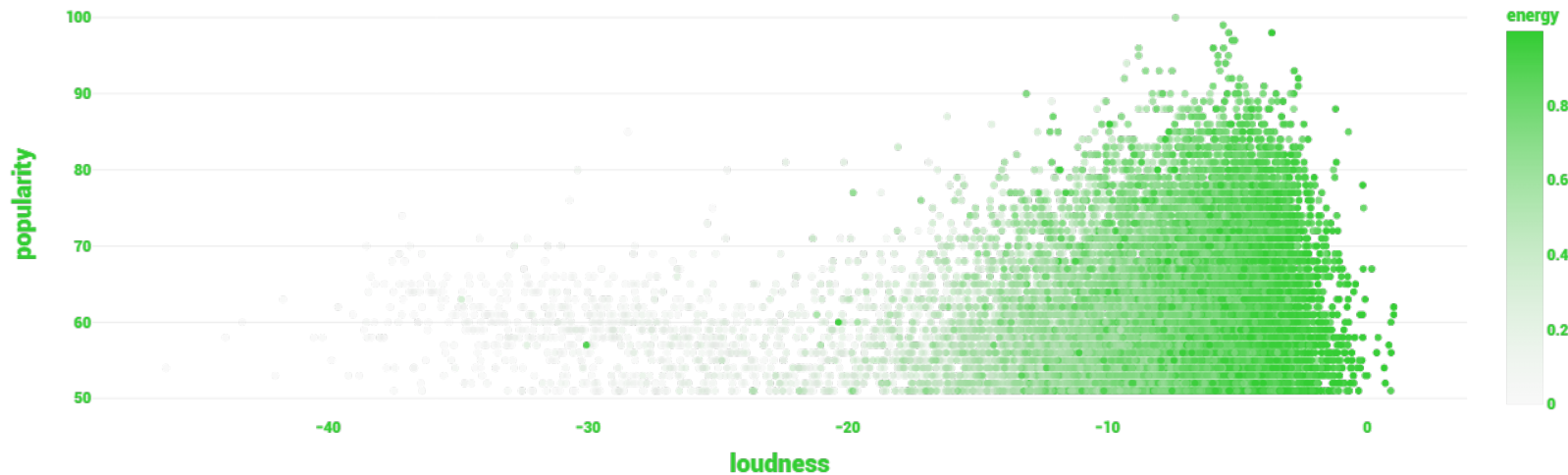
En la gráfica de caja y bigotes vemos que la media de la mediana de popularidad está por debajo de 40. Lo que nos podría llevar a pensar que el repertorio de música no está bien aceptado. O que se está incluyendo canciones que no se escuchan. Y vemos que la distribución tiene asimetría hacia la izquierda. También podemos notar que el punto de mayor densidad se encuentra alrededor del 0. (observación: hablamos del punto alrededor del 0 y no puntualmente en el 0 dado a que la probabilidad de un punto en una función continua es 0).

Boxplot popularidad



# CARACTERÍSTICAS FÍSICAS

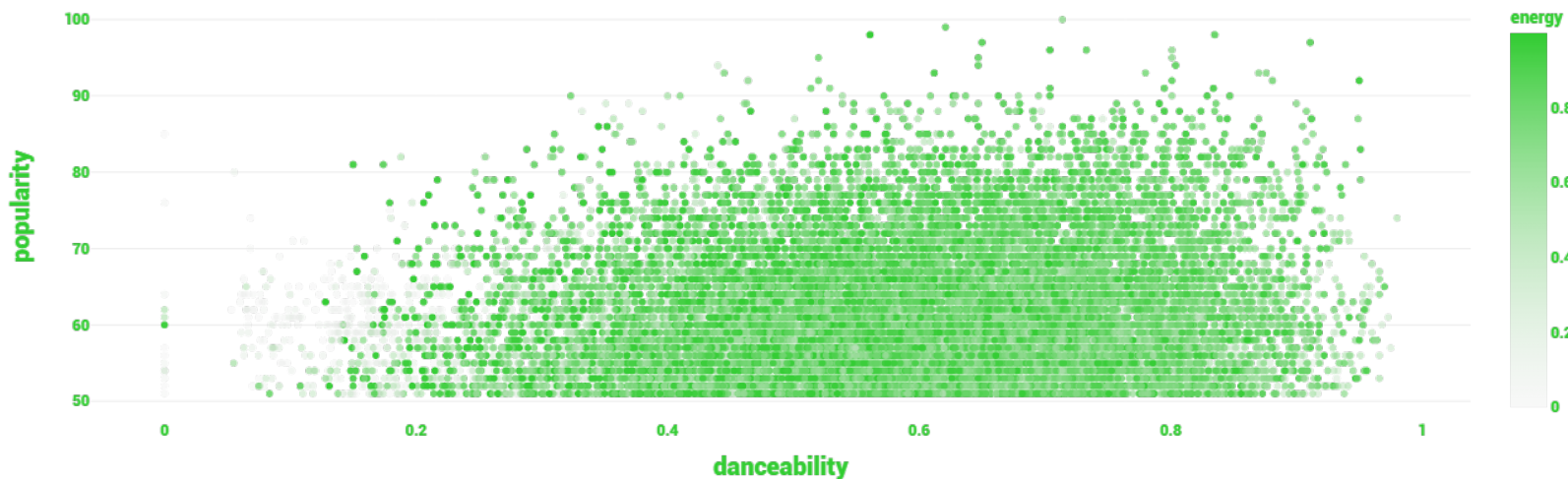
## Popularidad por volumen



Con las gráficas vemos que la bailabilidad no está directamente ligada a la energía. Y vemos casos de gran popularidad que la bailabilidad es muy alta pero la energía es muy baja.

Se podría deducir que las canciones más populares son ruidosas. Vemos que hay canciones que tiene muy buena aceptación que tiene un volumen bajo. Pero no hay canciones extremadamente populares en el extremo más ruidoso.

## Popularidad por bailabilidad



**Vemos que no hay correlaciones entre ninguna característica de la canción con el resultado comercial de la misma.**



# ANÁLISIS POR CONTENIDO EXPLÍCITO

**Las canciones son  
un poco mas populares**

Vemos que más de un 85% de las canciones del cuartil superior de popularidad no tienen lenguaje explícito. Sin embargo, las canciones con lenguaje explícito parecen ser más populares.

**Boxplot popularidad según contenido explícito**



# ANÁLISIS POR GÉNERO

Para que nuestra recomendación sea mas concreta. Podemos estudiar que géneros son los más populares. Para así generar una canción de ese género, asumiendo que va a tener una mediana de popularidad mayor al resto.

Como el valor mínimo de popularidad es 50. Vamos a ver que géneros tienen una media de popularidad de 50

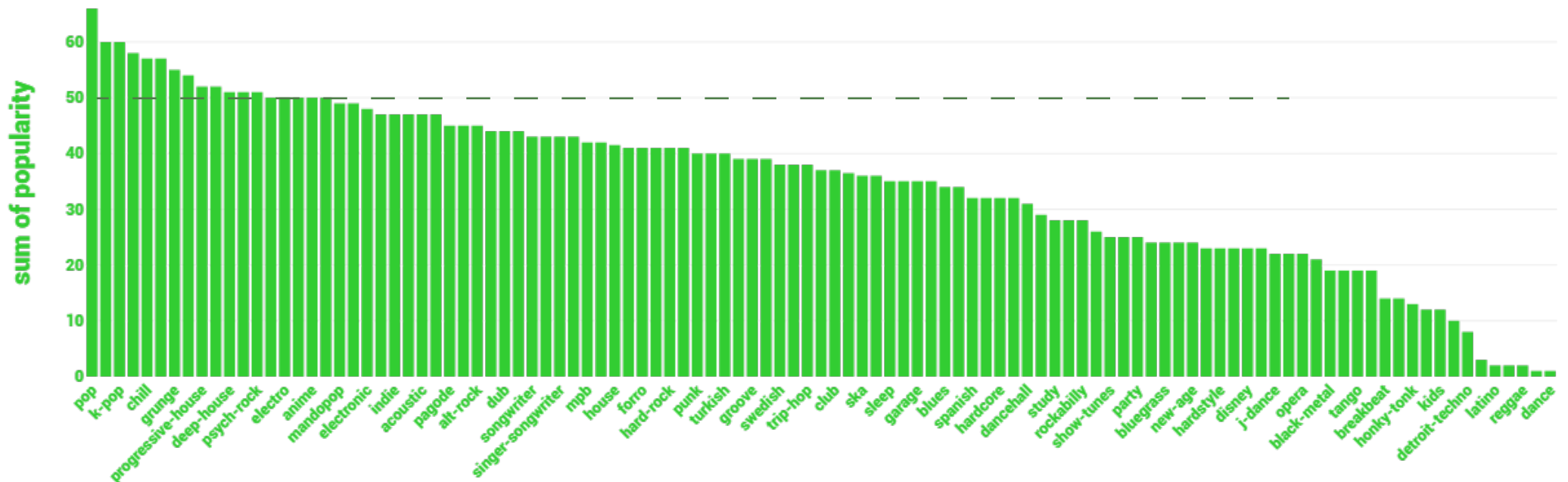
Se ve que la cantidad de géneros que tienen una media de popularidad mayor a 50% es unicamente el 11% de la cantidad total de géneros. Todos los generos tienen la misma cantidad de canciones (1000) por lo que podemos asumir que gran parte de las canciones en estos géneros son populares

Encontramos demasiado riesgo so enfocarse en el genero más popular. Por lo que recomendamos que se elija un genero de los 4 siguientes.

"k-pop", "pop-film", "hip-hop" y "chill". Teniendo en cuenta que estos son los cuatro generos más populares siguientes a "pop".

**pop**  
**k-pop**  
**pop-film**  
**hip-hop**  
**chill**

## Popularidad media por género



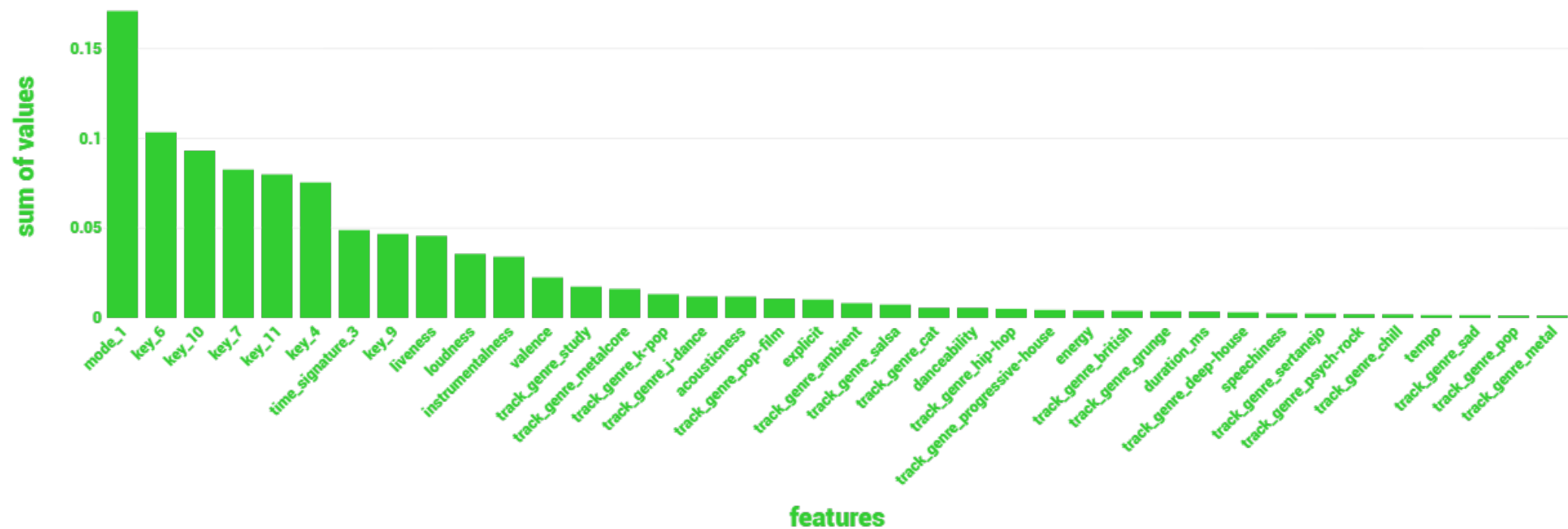
# FEATURES Y CONCLUSIONES

# IMPORTANCIA DE FEATURES

Mayor  
Modo

F#  
Clave

## ESTUDIO DE FEATURES



# CONCLUSIONES

pop  
k-pop

pop-film

hip-hop

chill  
Géneros

Mayor  
Modo

F#  
Clave

3:30 min  
Duración

LAS CARACTERÍSTICAS RECOMENDADAS EN ESTE TRABAJO NO ASEGURAN POPULARIDAD

Utilizar las características mencionadas puede que ayuden a generar tracción en el algoritmo de Spotify. Se podría utilizar algunas canciones que no sean la mejor expresión artística de nuestro cliente para generar tracción.

Y publicar las canciones pensadas para hit radial una vez se genere la suficiente tracción.

El uso de lenguaje explícito no es estrictamente necesario. Pero se nota una concentración de canciones con lenguaje explícito en el principio de la tabla

Siguiendo las recomendaciones anteriores vemos que las canciones que tienen estas características tienen una popularidad mayor a la mediana del total.

Algo que notamos es que no hay autores que tengan gran notoriedad. Por lo que concluimos que estas características no hacen a una canción popular en el sentido general.

Sino que hacen a la canción **POPULAR EN SPOTIFY**. Quizas el algoritmo de recomendación de spotify responde mejor a estas características

