

POPULARITY

*¿Qué hace a una
canción popular?*

Contexto y Audiencia

Hipótesis / Preguntas de interés

Metadata

Análisis Exploratorio

Insights y Recomendaciones

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

21

VARIABLES

ANÁLISIS

EXPLORATORIO

DISPERSIÓN DE POPULARIDAD

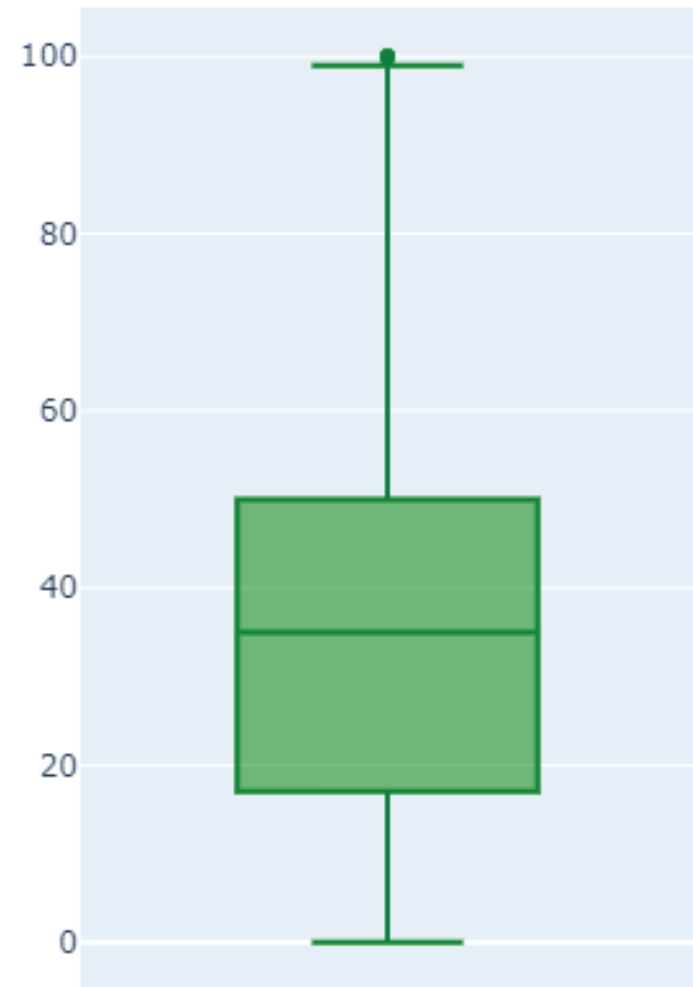
En la gráfica de caja y bigotes vemos que la media de la mediana de popularidad está por debajo de 40. Lo que nos podría llevar a pensar que el repertorio de música no está bien aceptado. O que se está incluyendo canciones que no se escuchan. Y vemos que la distribución tiene asimetría hacia la izquierda. 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).

35

POPULARIDAD
MEDIA

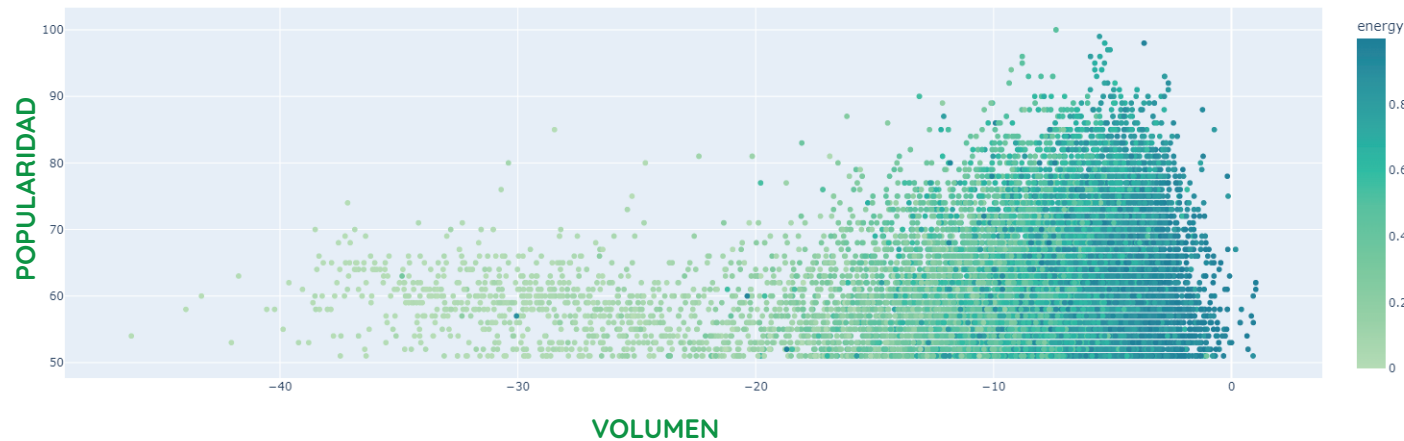
50 a 100

POPULARIDAD
CUARTIL SUPERIOR



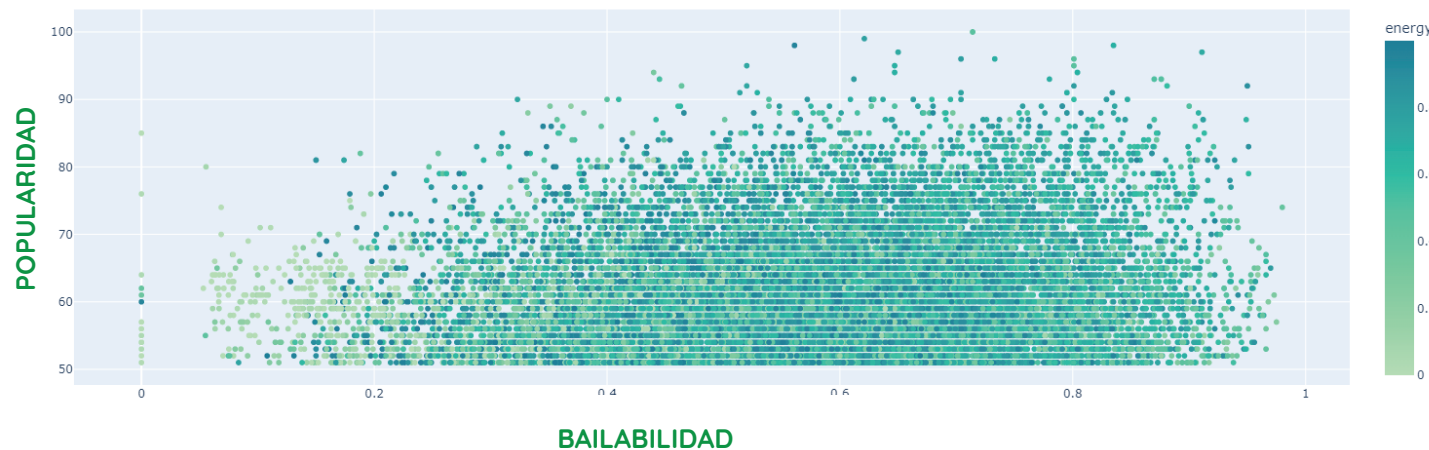
POPULARIDAD

CARACTERÍSTICAS FÍSICAS



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.



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

DIFERENCIAS

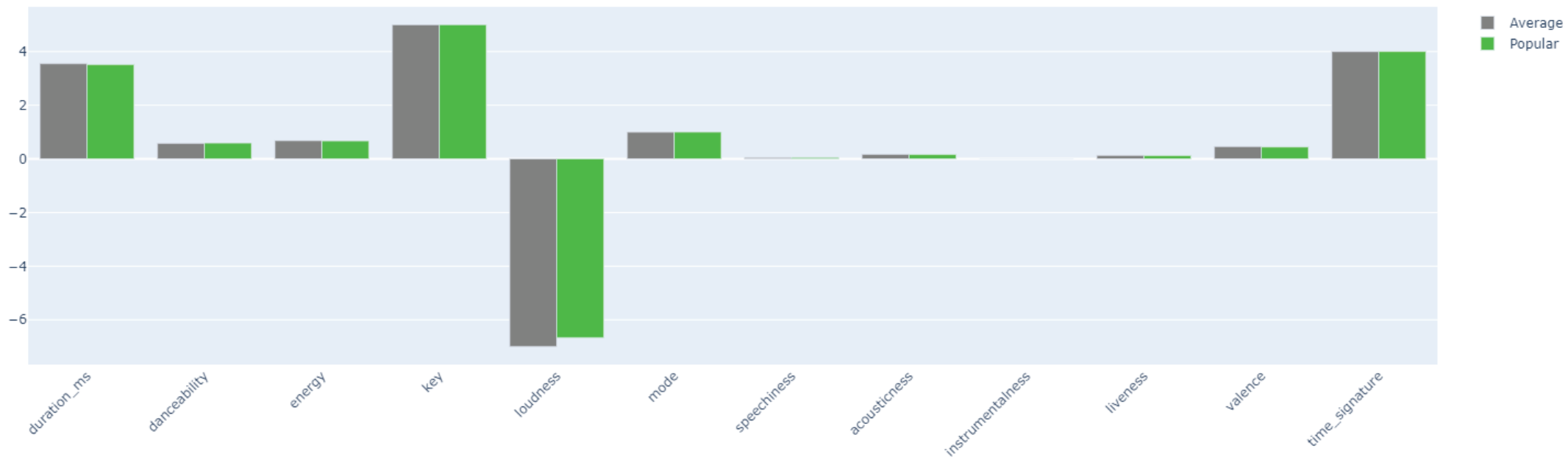
¿hay alguna diferencia que resalte entre las canciones más populares y el resto de canciones?

Comparando las medias de todos los campos de las canciones más populares y las medias del dataset general. Vemos que no existen grandes diferencias entre ellas.

Las diferencias más notorias son en "liveness", "acousticness", "valence" y "loudness" Siendo "liveness", "valence" y "acousticness" más bajas en las más populares. Pero menos de 10%

La diferencia es mínima. Por lo que podemos asumir que estos valores por si solos no son definitorios para la popularidad de la canción

No hay diferencia que resalte



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

Pero si tomamos el valor de popularidad mediana del cuartil superior de las canciones vemos que solo el genero "pop" está en ese rango.

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".

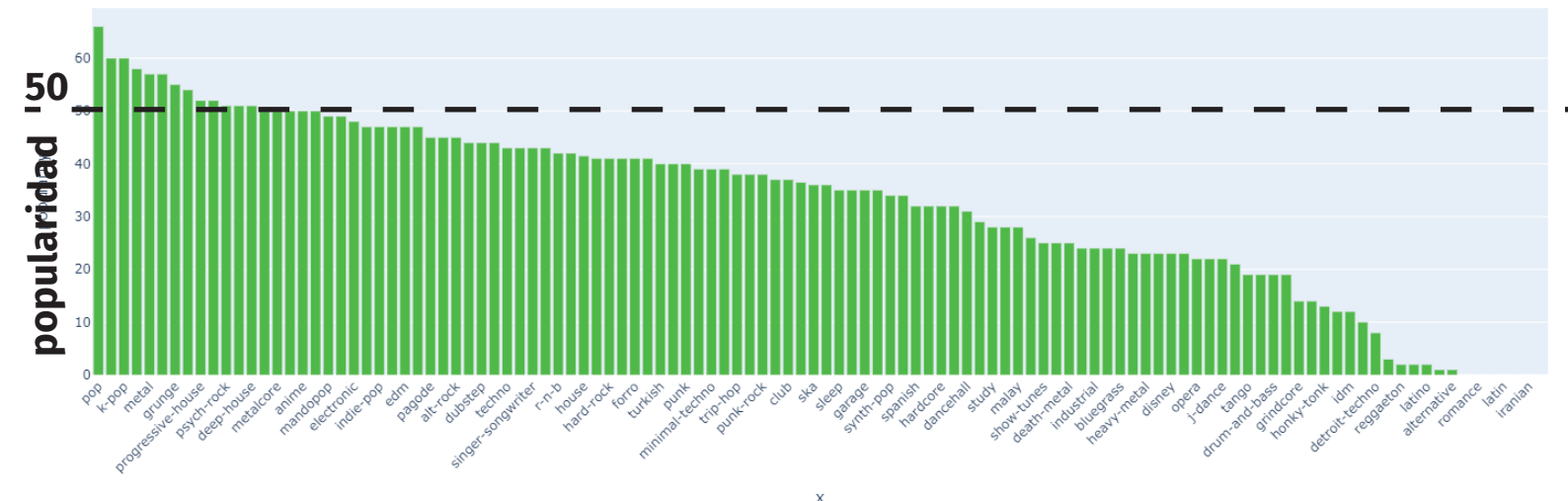
pop

k-pop

pop-film

hip-hop

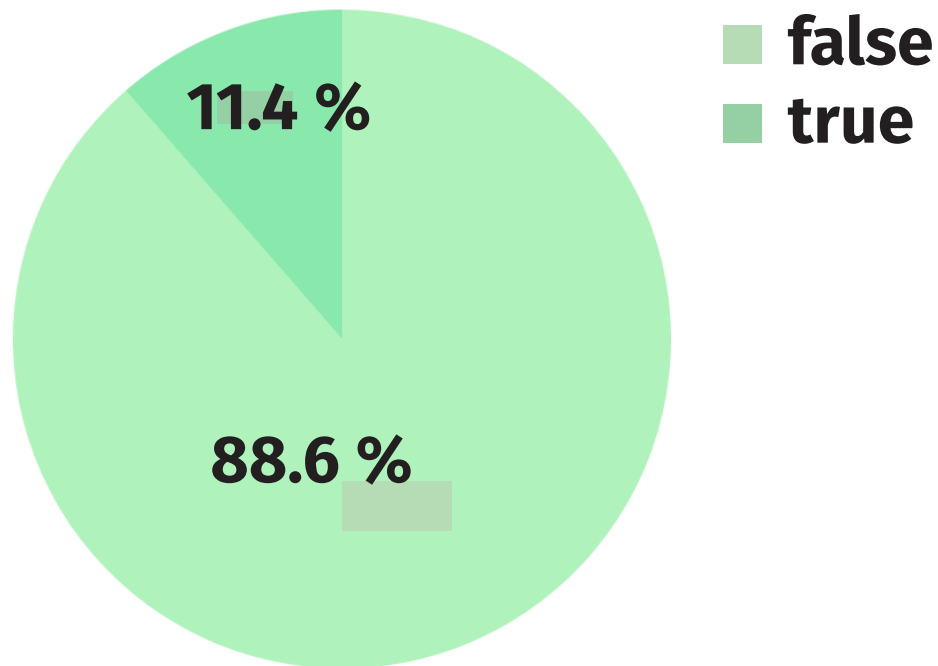
metal



ANÁLISIS POR CONTENIDO EXPLÍCITO

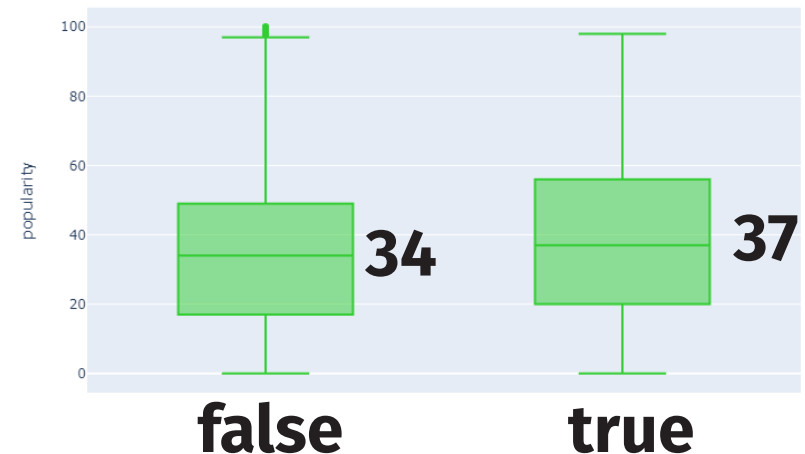
Vemos que más de un 85% de las canciones del cuartil superior de popularidad no tienen lenguaje explícito. Sin embargo, las canciones con lenguaje explícito parecen ser más populares.

Porcentaje de canciones con contenido explícito



Las canciones tienen mayor popularidad.

Distribución de popularidad por contenido explícito



INSIGHTS

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

Comparando las medias de todos los campos de las canciones más populares y las medias del dataset general. Vemos que no existen grandes diferencias entre ellas.

Se ve que la cantidad de géneros que tienen una media de popularidad mayor a 50 es únicamente el 11% de la cantidad total de géneros

RECOMENDACIONES

Encontramos demasiado riesgoso enfocarse en el género más popular. Por lo que recomendamos que se elija un género de los 4 siguientes.

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

Que la duración de la canción sea aproximadamente 3 minutos

Puede que generar una letra explícita, ayude en su popularidad. Pero no es seguro

si
Letra explícita

pop
k-pop
pop-film
hip-hop
metal
Géneros

< 3min
Duración